

Title	実時間マルチスケール空間フィルタ処理機能を有する 知能視覚センサの開発と応用
Author(s)	安川, 真輔
Citation	大阪大学, 2017, 博士論文
Version Type	VoR
URL	https://doi.org/10.18910/61748
rights	
Note	

Osaka University Knowledge Archive : OUKA

<https://ir.library.osaka-u.ac.jp/>

Osaka University

博士学位論文

実時間マルチスケール空間フィルタ処理機能を
有する知能視覚センサの開発と応用

安 川 真 輔

2016年 1月

大阪大学大学院工学研究科
電気電子情報工学専攻

目次

目次.....	i
内容梗概.....	iv
表記法.....	vi
第1章 序論.....	1
1.1 複数のスケールで視覚情報を解釈する生物の視覚系.....	1
1.2 マルチスケール解析の有用性.....	2
1.3 生体の視覚神経回路に学んだセンシング技術とその応用.....	4
1.4 研究目的.....	5
第2章 空間フィルタ処理とその高速化に関する先行研究.....	7
2.1 緒言.....	7
2.2 逐次型演算による空間フィルタ処理.....	7
2.2.1 空間フィルタ処理.....	7
2.2.2 逐次型演算器による空間フィルタ処理の問題点.....	9
2.3 低域通過型空間フィルタと帯域通過型空間フィルタ.....	11
2.3.1 ガウシアンフィルタの空間特性と周波数特性.....	11
2.3.2 DoG フィルタの空間特性と空間周波数特性.....	13
2.3.3 実空間と周波数空間.....	14
2.4 スケーラブルフィルタ.....	15
2.5 ソフトウェア・アルゴリズムの工夫による高速化.....	17
2.6 ハードウェアシステムによる高速化.....	17
2.7 生体の外網膜とアナログ抵抗回路網.....	18
2.8 シリコン網膜.....	22
2.8.1 内部構成と動作原理.....	22
2.8.2 FPGA によるシリコン網膜の制御.....	25
2.8.3 MOS 抵抗回路網と MOS 抵抗.....	27
2.9 結言.....	31
第3章 複数の帯域通過型空間フィルタ処理機能を有する知能視覚センサの開発.....	32
3.1 緒言.....	32
3.2 知能視覚センサの構成と処理の流れ.....	33
3.3 MOS 抵抗回路網の空間特性の測定.....	36
3.4 MOS 抵抗回路網の空間周波数応答の測定.....	38
3.5 画像に対するセンサの出力応答の測定.....	40

3.6 考察.....	40
3.6.1 提案するアーキテクチャの利点.....	40
3.6.2 単一フレーム内に適用できる空間フィルタの数について.....	42
3.6.3 MOS 抵抗の入力電圧レベルの依存性.....	42
3.7 結言.....	46
第4章 実時間マルチスケールエッジ検出機能を有する知能視覚センサの開発.....	47
4.1 緒言.....	47
4.2 ゼロ交差法によるエッジ検出.....	47
4.3 ゼロ交差検出回路.....	49
4.4 FPGA 上での差分画像とエッジ画像の並列処理.....	50
4.5 システムの出力応答の測定.....	51
4.5.1 単純な視覚対象に対する応答.....	51
4.5.2 複雑な視覚対象に対する応答.....	51
4.6 考察.....	54
4.7 結言.....	56
第5章 抵抗回路網を用いたスケール不変特徴量抽出の解析とシミュレーション.....	57
5.1 緒言.....	57
5.2 SIFT アルゴリズムとその関連研究.....	57
5.2.1 スケーラブルフィルタによるキーポイント検出.....	57
5.2.2 特徴量記述.....	59
5.2.3 キーポイントマッチング.....	60
5.2.4 関連研究.....	61
5.3 抵抗回路網フィルタとガウシアンフィルタの比較.....	61
5.4 DoG フィルタ処理画像と DoR フィルタ処理画像の比較.....	62
5.5 シミュレーション解析によるキーポイント検出の検討.....	66
5.6 シミュレーションによる特徴量記述の検討.....	68
5.7 考察.....	70
5.8 結言.....	71
第6章 実時間スケール不変特徴量抽出機能を有する知能視覚センサシステム.....	72
6.1 緒言.....	72
6.2 抵抗回路網による実時間スケラブルフィルタ処理.....	73
6.2.1 実時間マルチスケールフィルタ処理.....	73
6.2.2 MOS 抵抗回路網によるスケラブルフィルタ処理.....	73
6.3 FPGA によるキーポイント検出及び勾配ベクトル演算.....	75
6.3.1 FPGA による並列処理.....	75
6.3.2 SIFT キーポイント検出.....	77

6.3.3 勾配ベクトル画像の生成回路	79
6.4 ソフトウェアによる特徴ベクトル記述と対応点探索.....	81
6.5 評価実験.....	82
6.6 考察.....	85
6.6.1 アーキテクチャに対する考察.....	85
6.6.2 実環境への応用に対する考察.....	86
6.7 結言.....	87
第7章 結論.....	88
参考文献.....	90
謝辞.....	98
業績一覧.....	99

内容梗概

本論文は、筆者が主に大阪大学大学院工学研究科博士後期課程在学中に行った実時間マルチスケール空間フィルタ処理機能を有する智能視覚センサの開発と応用に関する研究をまとめたものであり、全7章から構成されている。

第1章 序論

本章では、本論文の目的と意義を述べた。視覚情報は様々なスケールの構造が組み合わされて構成されている。生体の視覚系はその並列階層的な構造により、視覚情報を効率的にマルチスケールで解析していると考えられている。従来の汎用イメージセンサと逐次演算型コンピュータとを組み合わせた画像センシングシステムでは、実時間で効率よく画像をマルチスケール解析することは難しい。従って、生体の視覚系に学び、アーキテクチャレベルからセンサシステム全体を見直すことが必要であることを論じた。

第2章 空間フィルタ処理とその高速化に関する先行研究

マルチスケール解析法の初段にはマルチスケール空間フィルタ処理がある。本章では、空間フィルタ処理と、その高速化の試みについての先行研究を述べた。まず逐次型演算による空間フィルタ処理とその問題点について述べた。次に、低域通過型及び帯域通過型空間フィルタの特性について述べた。また複数の空間フィルタ処理は多大な計算コストを要し、汎用コンピュータでは実時間処理が困難であることを述べた。これまでの空間フィルタ処理の高速化の手法について、先行研究を紹介した。最後に、本研究における高速化の鍵となる技術である生体網膜神経回路を模擬したアナログ抵抗回路網とそれを内蔵する智能視覚センサについて述べた。

第3章 複数の帯域通過型空間フィルタ処理機能を有する智能視覚センサの開発

本章では、複数の空間フィルタ処理の実行に適したハードウェア・アーキテクチャを提案し、試作した智能視覚センサの動作とその有用性について述べた。まず、本視覚センサが空間フィルタ処理を瞬時に実行するアナログ抵抗回路網と FPGA (Field-Programmable Gate Array) を有しており、このアナログ回路とデジタル回路の双方の利点を生かし、小型かつ低消費電力で実時間処理を実現していることを説明した。アナログ抵抗回路網は、その抵抗値によって空間特性が決定される。また本研究で用いる抵抗回路網の抵抗は MOS (Metal-Oxide-Semiconductor) 抵抗で構成されており、その抵抗値を外部信号により容易に制御できる。そこで、実機実験によって MOS 抵抗回路網の空間特性及び空間周波数特性が制御できることを示した。また FPGA からの制御信号により、取得画像に対して一フレーム取得

時間内に複数の空間フィルタを適用する手法を提案した。最後に試験画像を用いて、提案する知能視覚センサが実時間で複数の空間フィルタ処理を実行し、マルチスケール処理画像を取得できることを示した。

第4章 実時間マルチスケールエッジ検出機能を有する知能視覚センサの開発

本章では、第3章で述べた知能視覚センサの後段にデジタル演算回路を追加することで、マルチスケールの輪郭検出機能を有する知能視覚センサを構築したことについて述べた。まず帯域通過型空間フィルタ処理画像からエッジを検出する手法とその回路について述べた。次に、知能視覚センサからの画像をFPGA内のメモリに記録する際、画像読み出しと同時にFPGA上のデジタル回路で演算することで、フレームの遅れがなく、輪郭検出を実行する手法について述べた。最後に、複数の試験画像を用いて、提案する知能視覚センサが画像内の大まかな構造から細やかな構造まで抽出する、マルチスケール輪郭検出画像群が実時間で取得できることを示した。

第5章 抵抗回路網を用いたスケール不変特徴量抽出の解析とシミュレーション

本章では、マルチスケール解析で多く用いられているガウシアンフィルタの代わりに抵抗回路網フィルタを用いた、スケール不変特徴量の抽出についての解析とシミュレーションについて述べた。まず、従来のスケール不変特徴量抽出手法について簡潔に述べた。次に、スケール不変特徴量を抽出するために必要な空間フィルタのフィルタサイズの設定について述べた。また、ガウシアンフィルタと抵抗回路網フィルタの空間特性の違いが、スケール不変な特徴点の検出に与える影響をソフトウェア・シミュレーションで比較・検討した。最後に、異なる大きさのテスト画像群を用いて、二次元抵抗回路網によるスケール不変特徴量の抽出について検証した結果を述べた。

第6章 実時間スケール不変特徴量抽出機能を有する知能視覚センサシステム

本章では、画像からスケール不変特徴量を実時間で抽出する知能視覚センサシステムについて述べた。次いで、実時間スケール不変特徴量抽出を実現するためのシステム構成と、その動作について述べた。その中で特徴量抽出の高速化を図るために、視覚センサ内のFPGA上に新たに実装した並列演算モジュールについて述べた。実機においても、第5章で述べたようにスケール不変特徴量を抽出するために要求される抵抗回路網のフィルタサイズの設定ができることを示した。最後に、試験動画を視覚システムに提示し、本システムが実時間でスケール不変な特徴量を抽出できることを示した。

第7章 結論

本章では、本研究で得られた結果を総括した。

表記法

以下の表 1 に、本文中で各章を横断して用いられる主な記号をまとめる。

表 1 表記法リスト

表記	意味・説明
$E(\mathbf{X})$	抵抗回路網の入力端子に印加される光センサの画素電圧.
$V(\mathbf{X}, \lambda)$	抵抗回路網の出力端子電圧.
\mathbf{X}	画素位置 (x, y) を要素にもつベクトル. フレーム t 及び $t+1$ を更に要素として持つ時, \mathbf{X}_t 及び \mathbf{X}_{t+1} と表記する.
$f_{RN}(\lambda)$	スケールパラメータ λ をもつ RN (Resistive Network) フィルタ.
$f_{DR}(\lambda_1, \lambda_2)$	スケールパラメータ λ_1, λ_2 をもつ DoR (Difference of Resistive network) フィルタ. λ_1 と λ_2 の大きさの比が決まる時, $f_{DR}(\lambda_1)$ と略記.
λ	抵抗回路網フィルタのスケールパラメータ.
L	R_s と R_m の比の平方根.
R_s	抵抗回路網の側方抵抗.
R_m	抵抗回路網の縦方向抵抗.
D	抵抗回路網のノード間距離.
V_{rsbn}	R_s の nMOS 側バイアス電圧.
V_{rsbp}	R_s の pMOS 側バイアス電圧.
V_{rmbn}	R_m の nMOS 側バイアス電圧.
V_{rmbp}	R_m の pMOS 側バイアス電圧.
$S_{RN}(\mathbf{X}, \lambda)$	RN フィルタ処理画像.
$S_{DR}(\mathbf{X}, \lambda_1, \lambda_2)$	DoR フィルタ処理画像. λ_1 と λ_2 の大きさの比が決まる時, $S_{DR}(\mathbf{X}, \lambda_1)$ と略記.
$f_G(\sigma)$	スケールパラメータ σ をもつガウシアンフィルタ.
$f_{DG}(\sigma_1, \sigma_2)$	スケールパラメータ σ_1, σ_2 をもつ DoG フィルタ.
σ	ガウシアンフィルタのスケールパラメータ.

第1章 序論

1.1 複数のスケールで視覚情報を解釈する生物の視覚系

生体の網膜には外界の三次元構造を反映した二次元像が投影される。視覚とは、眼底に投影された二次元網膜像から目の前の三次元世界を解釈する情報処理過程である。図 1.1 に生体の視覚情報処理の流れの模式図を示す。視覚情報は網膜神経細胞内の初段の視細胞（図 1.1 内の Photoreceptor）により電気信号に変換され、出力段の神経節細胞（図 1.1 内の Ganglion cell）を介して、脳の視覚野（図 1.1 内の Visual cortex）へと送られる。この過程において並列階層的に配された様々な神経細胞群が、電気信号をやり取りすることで、視対象の大きさや動きや色、奥行きといった特徴の抽出から個体認識まで様々な視覚的な情報処理課題を解決している。網膜の段階から既に特徴抽出過程は始まっており、近年の研究では従来考えられていたよりも網膜神経回路は多くの機能を担っているという議論もある（総説として[1]）。この意味で、生体の視覚系は、外界の様子をきめ細かく、色再現性が高く切り出すことのできるデジタルカメラ等の撮像装置とは、目的が異なる情報処理装置であり、その設計原理も異なる。多くの視覚的課題は三次元世界から投影され、奥行き情報の欠落した二次元網膜像から推定することとなるため、解が一つに定まらない不良設定問題となる。生体の視覚系では（例えば、太陽のような光源が頭上にある等といった）物理的な仮説を有しており、これを拘束条件として、視覚的課題を解いていると考えられる[2]。

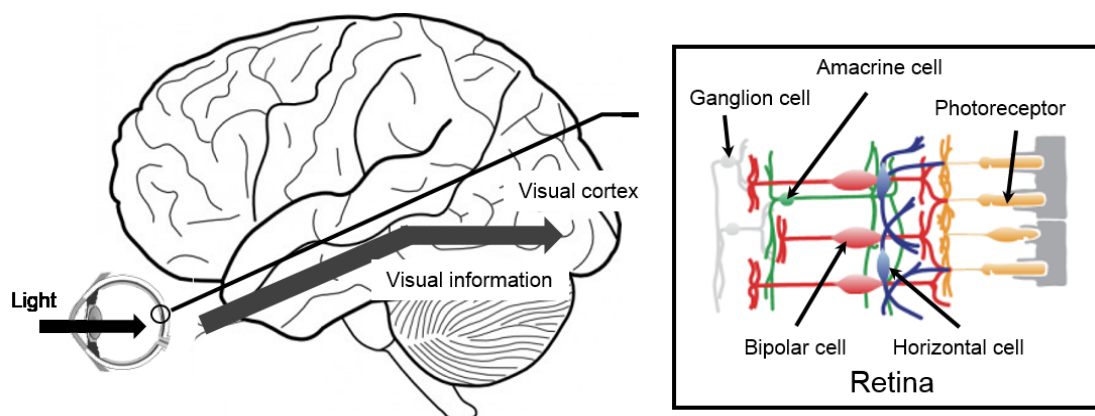


図 1.1 視覚の情報処理過程。まず外界からの光が網膜内の視細胞（Photoreceptor）で電気信号に変換される。その電気信号が網膜神経回路網でやり取りされることで、エッジや動き情報が抽出される。それらの視覚情報は神経節細胞（Ganglion cell）を介して視覚野に送られる。視覚野（Visual cortex）内において、更に高次の視覚的課題（形状認識等）が実行される。

三次元世界には様々な大きさの物体がある。従って、その様子が投影された二次元網膜像上には様々な大きさの幾何学的な特徴が表現されている。更に、眼と視対象との位置関係で対象の投影像が網膜を覆う領域の大きさ、すなわち対象の見えの大きさが変化する。生体の視覚神経回路では、このような様々な大きさの物体が表現されている像を解釈するために、視覚情報のあるスケールのみで処理せず、異なるスケールで並列に処理する仕組みを有していると考えられている[2]。ある神経細胞の応答に変化を生じさせる入力空間的な範囲を受容野と呼ぶが、その大きさは網膜神経細胞の段階から複数に分類されることが知られている[3]。視覚心理学の先行研究においても、生体の視覚系が視覚情報を空間周波数帯域ごとに分割し、その帯域ごとに、すなわちマルチスケールで処理することを示唆する実験結果が得られている [4]。更にマルチスケールで視野内の状況を解釈する生体の視覚系の計算原理に基づいた視覚情報処理モデルもまた提案されている[5, 6]。

上記の視覚情報のマルチスケール解析は有用なアプローチであるため、コンピュータビジョン分野のアルゴリズムにも広く利用されている[2, 7, 8]。この情報処理過程については、次節でその詳細を述べる。

1.2 マルチスケール解析の有用性

自然画像上には、物体の輪郭といった大局的な構造から、物体表面の模様といった局所的な構造まで、様々なスケールの様々な幾何学的特徴を持つ物体が表現されている[9]。そして、自然画像には様々な空間周波数成分が含まれている[10]。一般的に画像から情報を抽出する際、前処理として、着目した画像上の構造に応じたスケールに設定された空間フィルタを画像に適用することで、そのスケールの構造のみを強調する。この処理は、画像上のある空間周波数帯域の成分のみを強調することを意味する。マルチスケール画像解析では、複数の異なるスケールの空間フィルタを画像に適用することで、まず空間周波数帯域ごとに画像を分離する。これらの処理画像では、物体の輪郭や模様といった、それぞれ異なるスケールの構造が強調されている。マルチスケール画像解析ではそれらの画像群を個別に解析することで、入力画像上の局所的な構造から大局的な構造まで解析することができる。図 1.2(a) のシマウマの画像を用いてマルチスケールで二次元像を解析する有用性を論じる。マルチスケール解析では、その初段で異なる複数の大きさの空間フィルタを入力画像に適用する (図 1.2(b))。図 1.2(b) は (一次元) 空間フィルタの重み係数を表す。このマルチスケール空間フィルタ処理によって、細かい特徴を強調した画像から、粗い特徴を強調した画像まで複数の画像が生成される。この処理についての詳細は第 2 章で説明する。高周波帯域信号を通過させる小さなスケールの空間フィルタ (図 1.2(b) 上段) が適用された画像はシマウマの縞のような細かい構造が強調される (図 1.2(c) 上段)。一方で、低周波帯域信号を通過させる大きなスケールの空間フィルタ (図 1.2(b) 下段) が適用された画像は比較的細かい構造が除去され、シマウマの輪郭のみ強調される (図 1.2(c) 下段)。このようにマルチ

帯域通過型空間フィルタ

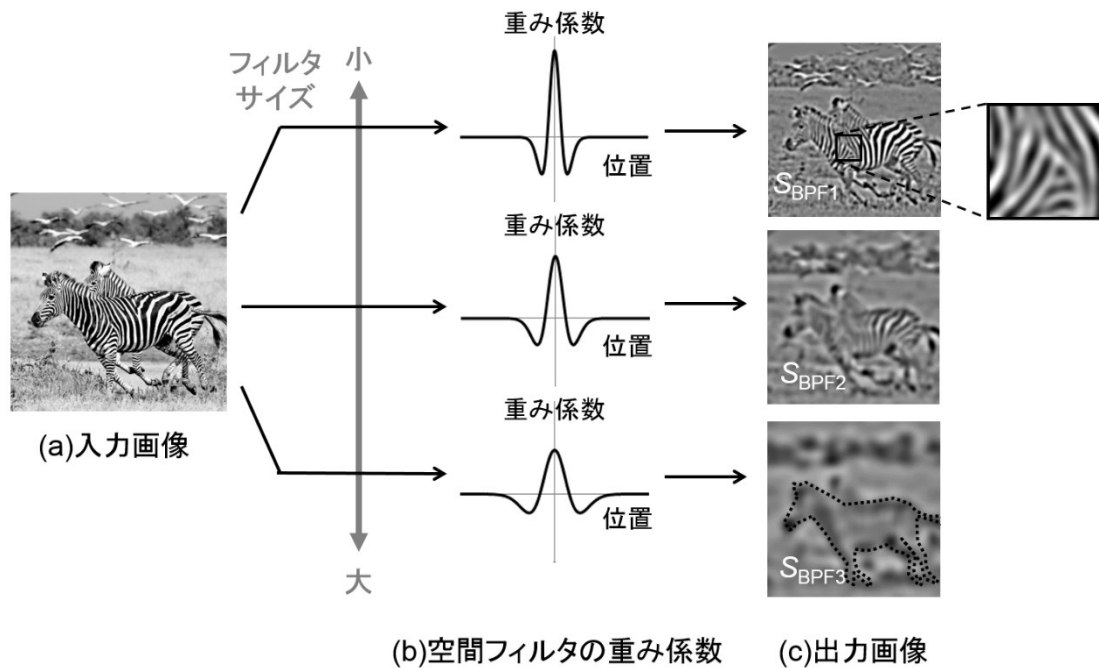


図 1.2 マルチスケール解析. (a) 入力画像. (b) 帯域通過特性を有する DoG (Difference of Gaussian) フィルタ群の重み係数 (一次元). 下のフィルタほどスケールが大きく, より低周波帯域の信号を通過させる. (c) 帯域通過型空間フィルタ処理画像. 入力画像にそれぞれ左に示される空間フィルタを適用することで得られる. 小さいスケールのフィルタでは細かい構造が, 大きいスケールのフィルタでは粗い構造がそれぞれ強調されている.

スケール画像解析は, 幅広い周波数帯域を含んだ画像情報から, 特定の周波数帯域のテクスチャやエッジのような画像特徴の抽出をする時に役立つとされている [2, 8]. マルチスケール画像解析は, コンピュータビジョン分野において, エッジ検出 [11], ステレオマッチング [12], スケール不変特徴記述 [13, 14] のような様々な応用 [15-21] で利用されており, その有用性が示されている.

画像センシングシステムでは, ここで例に挙げたシマウマのように動く対象を解析する場合, センサと対象物との相対的な距離が変化し, 像上での対象の見えの大きさも刻一刻と変化する. 従って, 画像センシングシステムには単一フレーム内に処理を完了させる実時間性が求められる. しかしながら, 複数の大きなサイズの空間フィルタを画像に適用する処理は, 計算コストが非常に高いため, 画像の実時間マルチスケール解析は難しい課題である. この点については 2.2.2 節でその詳細を述べる.

1.3 生体の視覚神経回路に学んだセンシング技術とその応用

視覚は外界の空間的な情報を一度に受容するため、生物が自己の置かれた状況を効率的に把握し、適切な行動を選択する上で重要な感覚情報である。外界の状況を効率よく解釈する画像センシングシステムを工学的に実現する上で、生体の視覚神経回路の計算原理や設計指針を手本とする研究がなされてきた。生物は自律行動という側面から、現状の工学的なシステムに比べて、非常に効率よく合理的な情報処理システムである。例えば国防高等研究計画局 (DARPA) によるロボットカーレース (DARPA グランド・チャレンジ) に参加した自律型ロボットカーは約 1kW のエネルギーが計算にかかること [22] に比べて、ハチは数 $10\mu\text{W}$ のエネルギーで視覚ナビゲーションを実現していると言われている [23]。

従来の画像センシングシステムの多くが、画像センサとマイクロコントローラのような汎用デジタルコンピュータから構成されている。近年の半導体プロセス技術や回路技術の発展に伴い、CMOS (Complementary Metal–Oxide–Semiconductor) イメージセンサに代表される画像センサの解像度や感度の向上が目覚ましい [24]。従来から用いられてきた CCD (Charge Coupled Device) イメージセンサと比較して、CMOS イメージセンサは安価であることや、低消費電力であること、また受光部の他に CMOS 回路が LSI に内蔵できることを理由に、現在では主流の画像センサとなり、携帯電話のカメラ等に広く採用されている。画像センシングシステムでは、この画像センサで取得された画像を汎用コンピュータ上のメモリに転送し、その中でデジタル画像処理が実行される。デジタル画像処理は一画素一画素、逐次的に情報を処理するために、処理が完了するまで次のフレームの画像を処理することができない。実時間処理が要求される応用の場合、一枚の画像に対する全ての処理を通常ビデオレート (33ms) 程度の時間内に完了させる必要がある。処理の高速化のために、CPU のクロックを高速にすると、それに伴う消費電力が上昇し、冷却用装置や大容量バッテリー等のハードウェアの規模の拡大を余儀なくさせる。

一方で、生体の神経回路の形態や計算原理の合理性に学び、それを電気電子回路で模倣し、システムを設計するニューロモフィック・エンジニアリング (Neuromorphic Engineering) と呼ばれる分野が 1980 年頃に創成された [25]。初期の研究では、MOS トランジスタ回路によって、ワンチップ上に神経回路の電氣的接続と演算機能を再現し、物理現象をそのまま演算に利用することで、最小のエネルギーで処理を実行するアナログ集積回路チップの開発が行われた。そこでは視覚や聴覚などの感覚系の神経回路や無脊椎動物が有する単純な神経回路を模倣したアナログ集積回路チップが複数の研究室で開発された (総説として [25-28]) が、中でも大きな成功をおさめた研究の一つとして、シリコン網膜 [29-36] の開発が挙げられる。シリコン網膜は、生体の網膜の合理的な構造に学んで、非常に効率よく空間フィルタ処理画像を取得できる画像センサである。ここで、本論文では情報処理機能を内蔵する画像センサを視覚センサと呼ぶこととする。この成功は、それまでに積み上げられてきた生体の網膜の構造やその電氣的特性に関する生理学的研究 [37-42] の知見に基づいて達成

された。シリコン網膜についての詳細はその基礎となった生理学的知見を 2.6 節で、技術的な内容を 2.7 節で述べる。

本グループで開発されたシリコン網膜はその後、後段に追加のアナログ回路、制御用・処理用デジタル回路を配することで、視細胞の受光特性を模した順応型視覚センサ[43]や、後段の視覚野の神経回路を模した方位選択性チップ[44]や色恒常機能の実現[45]、そして二つのカメラを用いた両眼視差検出システム[46]へと発展した。一方で、スイス連邦工科大学やスタンフォード大学を中心とした欧米のグループでは従来のデジタル画像の中には空間的・時間的に冗長な情報が含まれているとし、その変化分だけをアドレス・イベント表現 (Address-Event Representation ; AER) と呼ばれる手法で受け手側の計算機システムに送信するベースの視覚センサを開発してきた[28, 47]。

生体模倣型視覚センサを用いた応用もまた提案されており、例えばロボットの両眼システムによるナビゲーション[48]、行動モニタリング[49]、転倒検出[50]、ペン立て制御[51]、ゴールキーパーロボット[52]、神経補綴のための実時間シミュレータ[53]などのセンサシステムが研究・開発されてきた。

1.4 研究目的

本論文では、内蔵された専用の画像処理回路によって決められた簡易な前処理のみを実行する視覚センサに対して、より高度な画像処理機能を有するセンサである知能視覚センサを提案する。図 1.3 に知能視覚センサのコンセプト図を示す。本センサでは前段にレンズから入ってきた視覚情報を、受光と同時に画像処理する視覚センサを配し、後段に視覚センサ内への制御信号を生成するモジュールと、画像特徴を抽出する後処理モジュールを配する。知能視覚センサではセンサ内の画像処理回路の特性を動的に制御し、かつ後処理モジュールで更に情報処理を実行することで、より高度な画像処理機能を実現できる。

生物が外界の状況を把握するために、マルチスケール解析がその視覚系の神経回路内でなされていると考えられている[2]。本研究の目的は、生体の視覚系に学ぶことで、単一フ

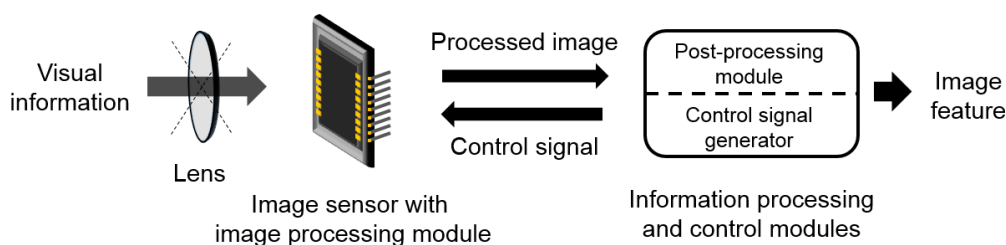


図 1.3 知能視覚センサのコンセプト図。知能視覚センサは画像処理機能を有する視覚センサと後処理及び制御機能を有するモジュールで構成される。

レーム内に効率よくマルチスケール空間フィルタ処理を実行できる視覚センサのアーキテクチャを提案し、それに基づく知能視覚センサを開発することである。提案する知能視覚センサは情報処理を実行するアナログ集積回路チップの他に、その特性を制御する信号を生成するデジタル回路を有している。上記の構成をとることで、本知能視覚センサはコンパクトでありながら、マルチスケール空間フィルタ処理のような高度な処理を実行する。更に、本論文では、この知能視覚センサの応用として実時間マルチスケール輪郭検出機能を有するセンサや、実時間でスケール不変特徴量抽出機能を有するセンサシステムを開発し、その有用性を示す。

第2章 空間フィルタ処理とその高速化に関する先行研究

2.1 緒言

空間フィルタ処理は、画像の認識過程における前処理の一つである。この処理は、演算を空間的に広範囲にわたって実行するため、逐次型演算器が不得意とする処理である。更にマルチスケール空間フィルタ処理では、空間フィルタ処理を複数回実行するため、逐次型演算器にとって非常に計算負荷が高い。

本章では、空間フィルタ処理について、ガウシアンフィルタを代表例として用いて説明する。また逐次型演算器における空間フィルタの計算コストについて述べる。次いで、空間フィルタ処理を高速化するためのソフトウェア・アルゴリズムやハードウェアの工夫についての先行研究をまとめる。最後に、本研究で鍵となる生体の網膜にある抵抗回路網と呼ばれる神経回路の構造と、網膜の構造を模したシリコン網膜と呼ばれるアナログ VLSI についての詳細を述べる。

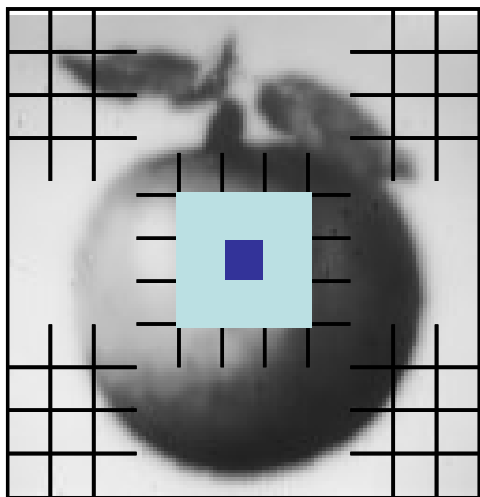
2.2 逐次型演算による空間フィルタ処理

2.2.1 空間フィルタ処理

空間フィルタ処理は、画像内のノイズ除去や、ある大きさの特徴を選択的に強調する機能を持つ画像処理技術の一種である。デジタルコンピュータ上での空間フィルタ処理を、図 2.1, 図 2.2 を用いて説明する。この処理では、図 2.1 の画像内にある格子状で示されるマスクと呼ばれる重み係数配列（フィルタカーネルともいう）を全画面にわたって、左上から右下まで一画素ずつ走査していく。画像配列とマスクが重なった部分（注目領域）において、要素ごとに重みを掛け合わせ、それらの合計を取ることで、マスク中心の画素の出力値が計算される。以上の計算は入力画像に対してフィルタの重み係数の畳み込み演算を実行したことを意味する。例えば図 2.2 のような重み係数配列（カーネル）と画像配列の注目領域を考える。このときの注目画素（中央の b_{22} ）の位置における畳み込み演算の結果 b'_{22} は次のようになる。

$$b'_{22} = \sum_{j=1}^3 \sum_{i=1}^3 a_{ij} b_{ij} \quad (2.1)$$

フィルタ処理前



フィルタ処理後

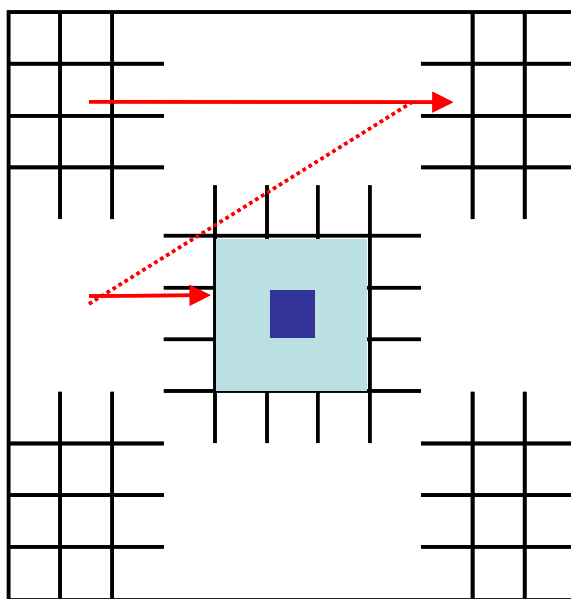
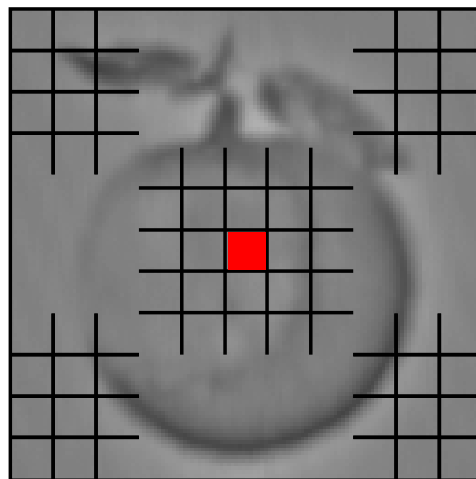


図 2.1 逐次型演算による空間フィルタ処理の例. 空間フィルタでは下の図内の矢印のように右上の画素から左下の画素まで重み係数配列を一画素ずつ走査していき、重み付き加算平均を実行していく.

図 2.1, 図 2.2 の例では 3×3 画素のフィルタカーネルを用いて説明したが, このフィルタカーネルの大きさ (以下, フィルタサイズ) は, 強調したい特徴の大きさに応じて設定される. 例えば, 画像内における大局的な変化を強調したい場合は, 大きなフィルタサイズの空間フィルタを入力画像に適用する必要がある.

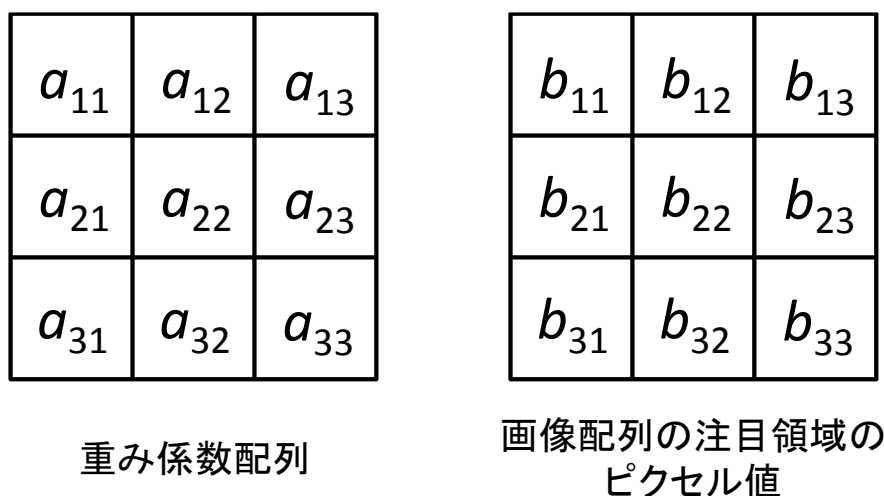


図 2.2 重み係数配列と注目領域.

2.2.2 逐次型演算器による空間フィルタ処理の問題点

図 2.3 を用いて、従来の画像センシングシステムによる空間フィルタ処理の問題点を説明する。このシステムは汎用の CCD/CMOS イメージセンサを内蔵したカメラと、メモリと CPU を内蔵した汎用デジタルコンピュータで構成されている。CCD/CMOS イメージセンサは各画素のフォトセンサによって外界からの光を受け、画像を生成する。その画像データは一画素ずつ走査的に読み出され、A/D (Analog-to-Digital) 変換器によってデジタル信号に変換後、画像保存用のメモリに転送される。その後、CPU はそのメモリにアクセスし、逐次的に処理を実行する。

空間フィルタ処理は、画像と空間フィルタの畳み込み演算である。従来のシステムによって、この処理を実行する場合、最も左上の画素から右下の画素まで順に注目画素と周辺の画素を読み出し、これらの重み付け加算を逐次的に実行することとなる。この空間フィルタ処理の積和演算 (Multiply-Accumulation; MAC) 数はそのフィルタサイズが大きくなるに従い、莫大な数に膨れ上がる。空間フィルタ処理において、積和演算の数 N_{ps} はフィルタ幅の二乗に比例する。従って、複数回空間フィルタ処理の計算量 N_{ps} は次の式で表現される。

$$N_{ps} = N_i \sum_{n=1}^{N_f} w_n^2. \quad (2.2)$$

ここで、 N_i は入力画像の画素数を表す。 w_n ($1 < n < N_f$) は空間フィルタの幅を、 N_f は適用するフィルタの数をそれぞれ表している。この式から、二次元空間フィルタの計算コストは、画像サイズに比例して増加するのに加えて、フィルタ幅の二乗に比例して増加することが分かる。

従来のシステムのアーキテクチャが、空間フィルタ処理の実行に不向きな理由は次の二つである。一つはメモリアクセス、もう一つは汎用 CPU に内蔵されている演算器である。

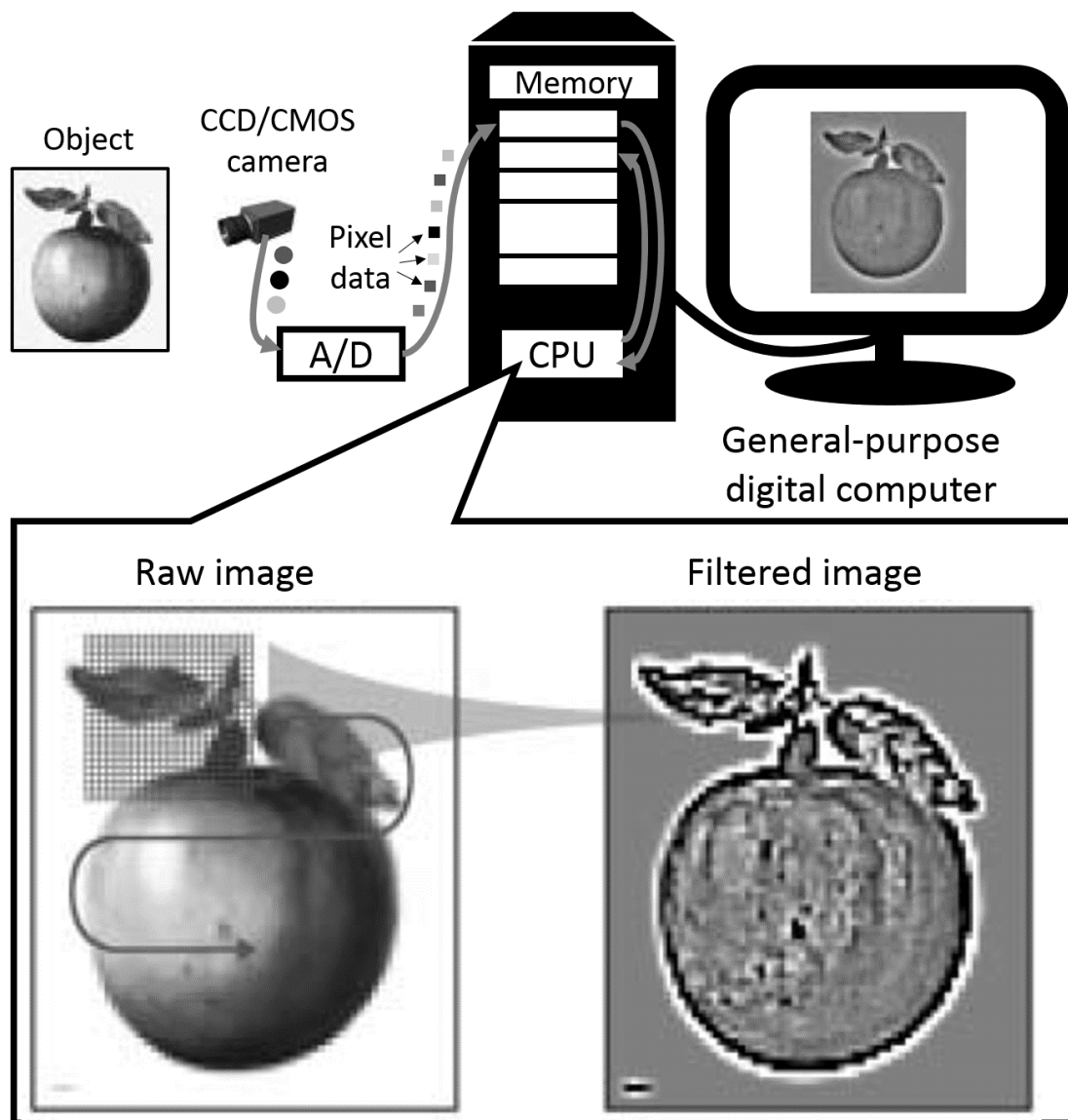


図 2.3 従来の逐次型演算を実行する視覚センシングシステムにおける空間フィルタ処理. CCD イメージセンサに代表される汎用イメージセンサとデジタルコンピュータで構成される. カメラで撮像された画像が一画素ずつメモリに保存される. CPU がメモリ内の画像に一画素ずつアクセスし, 処理を実行する.

まずメモリへのアクセスの問題について述べる. 逐次演算器によって画像処理をする場合, 外部メモリから CPU 内の一時メモリに画素情報, フィルタカーネル情報を読み込んで演算することとなる. この時, CPU がメモリにアドレスを指定して, メモリがそのアドレスからデータを出力するまでに一定の時間がかかる. このメモリへのアクセス時間は扱うデータが多いとその分だけ長くかかることになる. この問題はノイマンボトルネックといい, 逐次直列的に演算する方式では避けられない問題である.

また汎用 CPU には積和演算器が少ないため、膨大な数の積和演算処理は不得意である。ハードウェア的な工夫についての先行研究は 2.5 節でその詳細を扱う。

上記の理由から、大きなサイズの空間フィルタを複数回、画像に適用するとすると、計算コストが深刻な問題となる。

2.3 低域通過型空間フィルタと帯域通過型空間フィルタ

2.3.1 ガウシアンフィルタの空間特性と周波数特性

ガウシアンフィルタ (Gaussian filter) は、中央値フィルタや平均化フィルタ等のフィルタと同様に、画像の平滑化やノイズ除去を目的として画像に適用される空間フィルタの一つである。ガウシアンフィルタという名称は重み付けの分布をガウス分布に則って生成している所に由来する。一次元ガウス関数 $f_G(x, \sigma)$ を式(2.3)に示す。

$$f_G(x, \sigma) = A_G \exp\left(-\frac{x^2}{2\sigma^2}\right). \quad (2.3)$$

ここで、 x は位置、 $A_G = 1/\sqrt{2\pi\sigma^2}$ は正規化係数、 σ はガウス関数の形状を決めるパラメータである標準偏差である。図 2.4 に一次元ガウス関数を示す。図中 σ はガウス関数の標準偏差であり、標準偏差が大きいほどガウス分布の広がりが大きくなる。そのため σ はスケールパラメータとも呼ばれる。画像は二次元であるため、画像処理に通常用いられるガウシアンフィルタのカーネルは式(2.4)で示される二次元ガウス関数である。

$$f_{2G}(x, y, \sigma) = A_{2G} \exp\left(-\frac{x^2+y^2}{2\sigma^2}\right). \quad (2.4)$$

ここで、 x, y は位置、 $A_{2G} = 1/2\pi\sigma^2$ は正規化係数である。

ガウシアンフィルタの処理結果を図 2.5 に示す。図 2.5(a) は入力画像であり、図 2.5(b)-(f) がガウシアンフィルタ処理画像である。図 2.5(a) に比べ、その他のガウシアンフィルタ処理画像は平滑され、円の輪郭部がぼけていることがわかる。また σ の大きなフィルタの方が、 σ の小さいフィルタよりも空間的に広い範囲で平滑化がなされていることがわかる。ガウシアンフィルタは注目画素の周囲の平均を計算する平均化フィルタに比べ、自然な平滑化がなされる。これは中央付近に大きな重みをもち、周辺に向かうにつれてなだらかに重みが小さくなるガウス関数の形状に起因している。ここで図 2.5 の出力画像の端はフィルタが処理できないため、黒く塗りつぶして表現している。

次にガウシアンフィルタの空間周波数特性について述べる。ガウシアンフィルタの空間周波数特性を図 2.6 に示す。図から平滑化フィルタは高周波数成分のみを遮断する低域通過

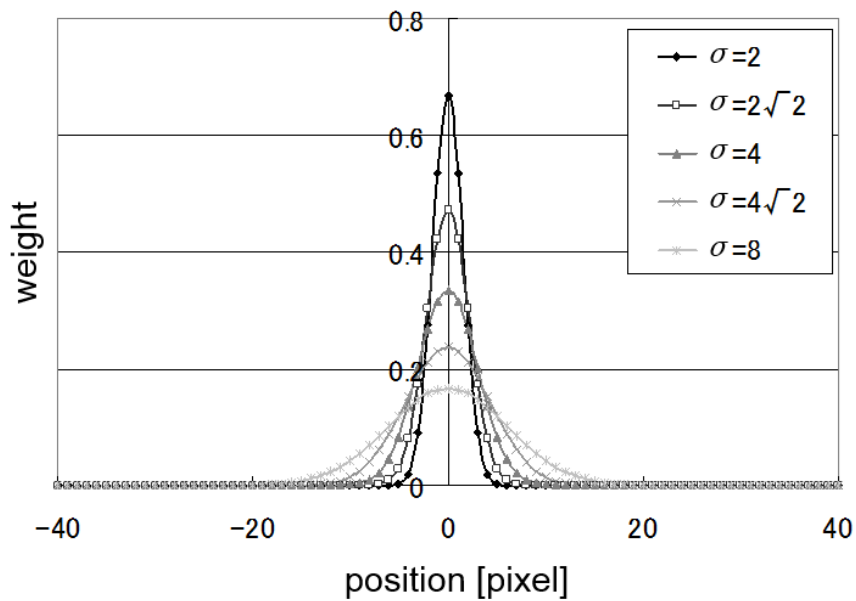


図 2.4 ガウシアンフィルタのフィルタカーネル（一次元）. σ が小さいほど鋭い形状をし、 σ が大きいほど広範囲に広がる.

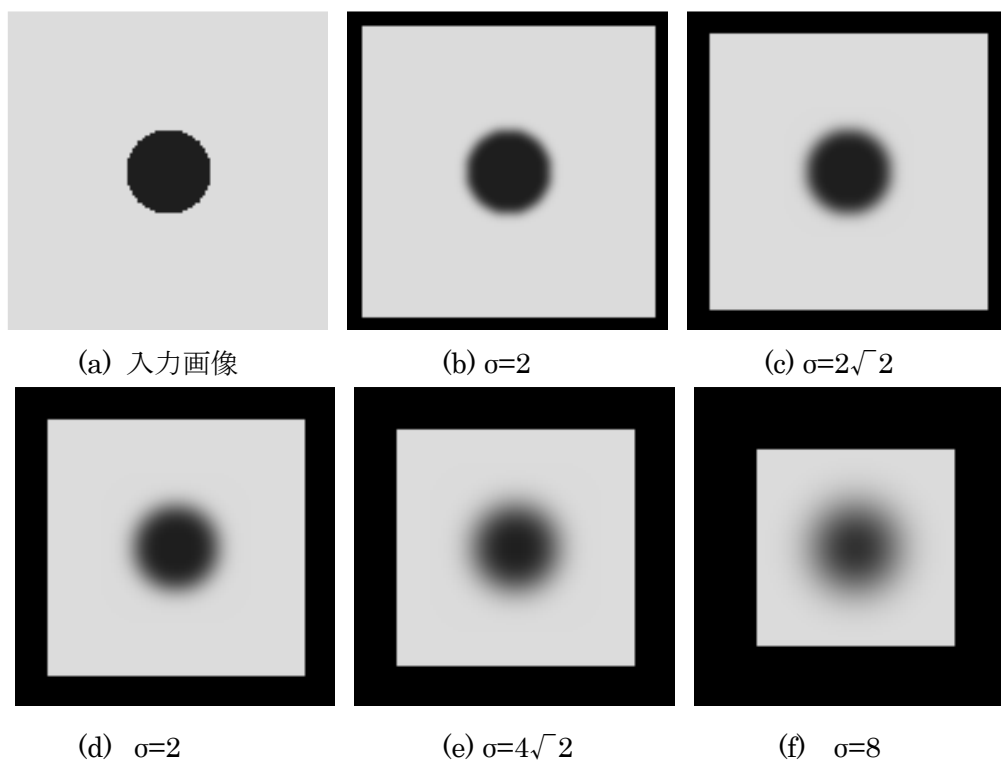


図 2.5 ガウシアンフィルタ処理結果. σ が大きいほど、平滑化度合いが強い. 画像の境界領域の黒色領域は入力画素がなく、演算できない領域を表す.

特性を有していることがわかる。実空間において空間フィルタのフィルタサイズが小さい時、すなわち σ が小さい時ほど、画像上で表現されている大まかな構造を残し、細やかな構造のみを除去する。すなわちカットオフ周波数が高域側にあり、高周波成分を多く含む細かい構造のみが除去される。一方、フィルタサイズと σ が大きい時、図 2.6 の矢印で示されるように、空間フィルタの周波数成分のカットオフ周波数はより低域側になり、細やかな構造は除去され、より大きく粗い構造のみが抽出される。

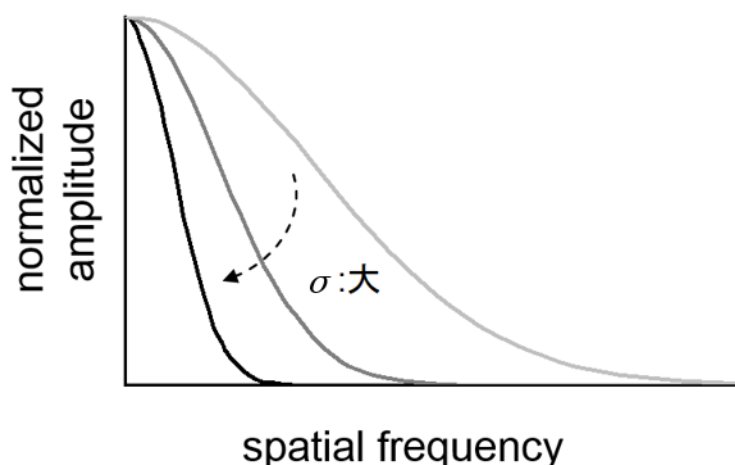


図 2.6 ガウシアンフィルタの空間周波数特性. σ が大きいほど、カットオフ周波数が低くなる.

2.3.2 DoG フィルタの空間特性と空間周波数特性

DoG (Difference of Gaussian) フィルタは異なる標準偏差の二つのガウス関数の差分のことであり、LoG (Laplacian of Gaussian) フィルタの近似として用いられることがある [13]. 一次元 DoG 関数を式 (2.5) に示す.

$$f_{DG}(x, \sigma_1, \sigma_2) = A_{G1} \exp\left(-\frac{1}{2} \frac{x^2}{\sigma_1^2}\right) - A_{G2} \exp\left(-\frac{1}{2} \frac{x^2}{\sigma_2^2}\right). \quad (2.5)$$

ここで σ_1 及び σ_2 はそれぞれのガウス関数の標準偏差であり、 $A_{G1} = 1/\sqrt{2\pi\sigma_1^2}$ 、 $A_{G2} = 1/\sqrt{2\pi\sigma_2^2}$ は正規化係数である。次に様々なスケールの一次元 DoG フィルタのフィルタカーネルを図 2.7 に示す。ここで、二つの標準偏差の関係は $\sigma_2 = \sqrt{2}\sigma_1$ として σ_1 を 2 から $4\sqrt{2}$ まで変更した。 σ を大きくするとフィルタカーネルの中心の重みが低くなり、周辺に重みが広がるのがわかる。ガウシアンフィルタと同様に、画像処理に通常用いられる DoG フィルタのカーネルは式(2.6)で示される二次元 DoG 関数である。

$$f_{2DG}(x, y, \sigma_1, \sigma_2) = A_{2G1} \exp\left(-\frac{1}{2} \frac{x^2+y^2}{\sigma_1^2}\right) - A_{2G2} \exp\left(-\frac{1}{2} \frac{x^2+y^2}{\sigma_2^2}\right). \quad (2.6)$$

ここで、 $A_{2G1} = 1/2\pi\sigma_1^2$ 、 $A_{2G2} = 1/2\pi\sigma_2^2$ は正規化係数である。

DoG フィルタは輪郭強調や特定の大きさの特徴を選択的に強調することを目的として画像に適用される。先ほどと同様の入力画像 (図 2.5(a)) に DoG フィルタを適用した場合の出力画像を図 2.8(a)-(d)に示す。 σ が小さいほど輪郭で鋭敏に応答しており、逆に σ が大きいほど輪郭で粗く応答していることが確認できる。

次に DoG フィルタの空間周波数特性の模式図を図 2.9 に示す。DoG フィルタの空間周波数特性はある特性の周波数帯域を通過させる特性をもつ。また DoG フィルタは σ が大きくなるとより低い帯域を通過させる特性となる。上記のように低域通過型空間フィルタであるガウシアンフィルタの差分は帯域通過型空間フィルタと数学的に等価となり、そのフィルタサイズを変更することで、異なる帯域通過特性が実現できる。

2.3.3 実空間と周波数空間

2.3.1 節と 2.3.2 節でガウシアンフィルタ及び DoG フィルタについて、それぞれ実空間における空間フィルタのフィルタカーネルの形状及び機能について説明した。実空間における画像と空間フィルタの畳み込み演算は、空間周波数領域では、画像の振幅スペクトルと

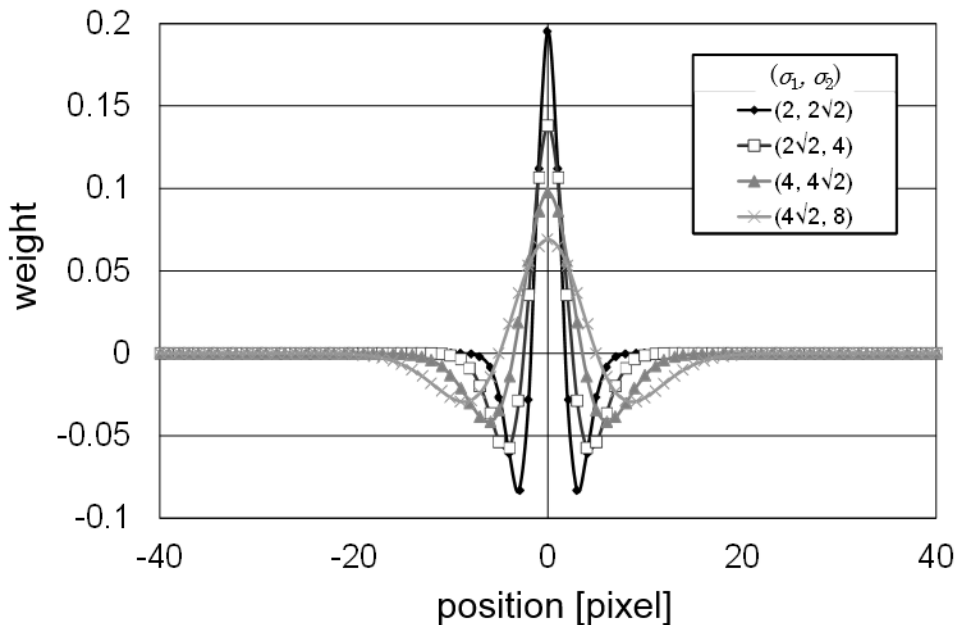


図 2.7 DoG フィルタのカーネル (一次元).

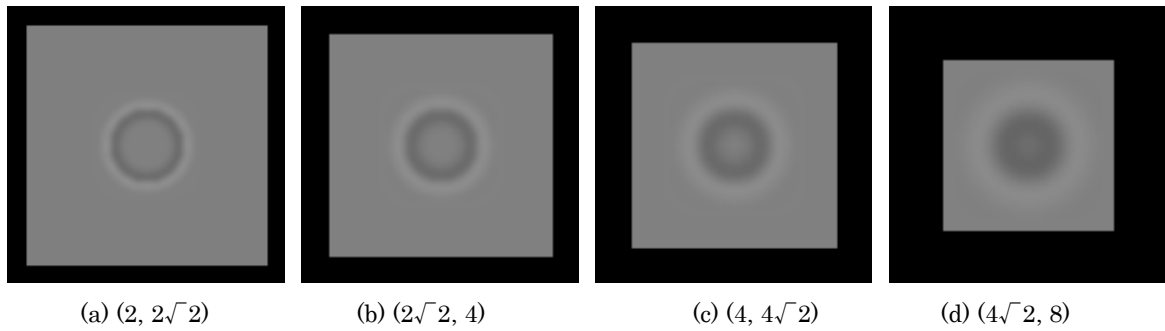


図 2.8 DoG フィルタ処理. (σ_1, σ_2) は DoG のスケールを表す.

空間フィルタの空間周波数特性の掛け算となる. 図 2.10 を用いてフーリエ変換を用いたフィルタ処理について説明する. 入力画像 (図 2.10(a)) をフーリエ変換し, 周波数空間で表現した結果が図 2.10(b) (振幅特性)である. また DoG フィルタ (図 2.10(c)) をフーリエ変換し, 周波数空間で表現した結果が図 2.10(d)である. これらを周波数空間上で掛け算し, 逆フーリエ変換することで図 2.10(f)に示されるように, 空間フィルタ処理を実行することが出来る. しかし, フーリエ変換は計算コストが高いため, 今回はこの方法を採用しない.

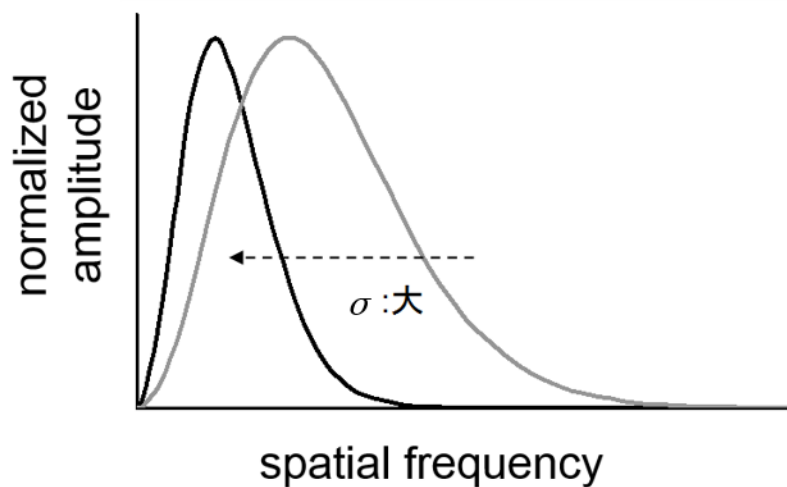


図 2.9 DOG フィルタの帯域通過特性. σ が大きいと, 通過帯域が低域側に移る.

2.4 スケーラブルフィルタ

マルチスケール解析では処理の初段で, 異なるフィルタサイズの複数の空間フィルタを画像に適用する. マルチスケール解析の技術の一つに本論文の 5 章, 6 章で対象となるスケール不変特徴量[13, 14]の抽出がある. この特徴量抽出のためには, 初段で適用する帯域通過型空間フィルタ群がスケラブルと呼ばれる特性を有していることが要求される. こ

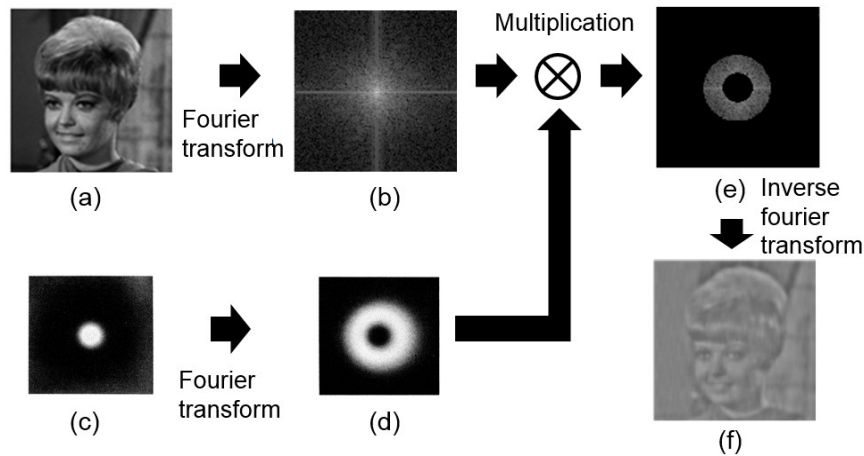


図 2.10 周波数空間でのフィルタ処理. (a) 入力画像を (b) フーリエ変換した結果と, (c) 二次元帯域通過型空間フィルタを (d) フーリエ変換した結果の掛け算した結果が (e) である. この (e) を (f) 逆フーリエ変換することで, 空間フィルタ処理が実行できる.

で, フィルタの重み関数の横方向の拡がり度合いが, あるパラメータで決定される時, そのフィルタはスケラブル性を備えているといい, このパラメータをフィルタのスケールもしくはスケールパラメータという. すなわち, スケラブル性を備えたフィルタ (以降, スケラブルフィルタ) とは, 重み関数であるフィルタカーネルが, 中心位置からの距離 d_s とスケールパラメータ a_s の比の関数 $f_s(d_s/a_s)$ であり, かつ正規化されたフィルタと定義される. スケール不変特徴量を抽出する手法の代表である SIFT (Scale-Invariant Feature Transform) アルゴリズム[13, 14]では, スケラブル性を備えるように標準偏差が設定された DoG (difference-of-Gaussian) 関数 $f_{bGn}(\sigma_n)$ を採用している. そこで, DoG フィルタがスケラブル性を備える条件を述べる.

問題を簡略化するために次元で考える. マルチスケール解析における n 番目のガウシアンフィルタのカーネルは式 (2.3) からガウス関数は次元において, 次の式で表される.

$$f_{Gn}(x/\sigma_n) = A_G \exp\left(-\frac{1}{2}\left(\frac{x}{\sigma_n}\right)^2\right). \quad (2.7)$$

式 (2.7) からガウシアンフィルタは x/σ_n の関数であるため, スケラブルフィルタである. 次に二つのガウシアンフィルタの小さいスケール, 大きいスケールの標準偏差をそれぞれ σ_n と σ_{n+1} である DoG フィルタを考える. この時, 標準偏差が式 (2.8) の関係をもつとする.

$$\sigma_{n+1} = K\sigma_n \quad (K: \text{正の整数}). \quad (2.8)$$

n 番目の一次元 DoG 関数は式 (2.5) と式 (2.8) から次のように表される.

$$f_{DGn}(x/\sigma_n) = A_{G1} \exp\left(-\frac{1}{2}\left(\frac{x}{\sigma_n}\right)^2\right) - A_{G2} \exp\left(-\frac{1}{2 \cdot K^2}\left(\frac{x}{\sigma_n}\right)^2\right). \quad (2.9)$$

式 (2.9) から, DoG 関数 f_{DGn} は x/σ_n の関数であるため, スケーラブルである.

2.5 ソフトウェア・アルゴリズムの工夫による高速化

2.2.2 節で述べたように, 空間フィルタ処理は従来の逐次型演算では非常に計算コストが高く, 複数回の空間フィルタ処理を画像に実行すると, その計算量は膨大なものとなる. この処理を高速化するために, ソフトウェア・アルゴリズムや専用ハードウェアが開発されてきた. 本節ではまずソフトウェア・アルゴリズムの先行研究をまとめる.

空間フィルタ処理を高速化するソフトウェア手法として, 主に処理の範囲を限定することで高速化する手法と, 空間フィルタ処理自体の処理速度を向上させる手法とに分類できる. 前者は空間フィルタに限らず一般的に画像処理を高速化する手法としてよく知られている. この手法ではコンピュータが事前知識をもっていない, 最初のフレームのみ画像内の全ての画素を対象に演算を実行し, 注目領域を抽出する. 次のフレームから, コンピュータはその領域のみ空間フィルタ処理を実行する. しかしながら, この手法は注目領域が移動する場合, 一度見失うと事前知識となる注目領域抽出処理からやり直すこととなる. また, 解析対象が画像全体でとらえている場合など, 画像内を大局的に解析したい場合, この手法は使うことができない. 後者は定時間空間フィルタ処理と呼ばれるソフトウェア・アルゴリズムである. 2.2.2 節で説明した通り, 畳み込み演算はフィルタ幅が大きくなればなるほど, 長い計算時間が要求される. 一方で, ここで紹介する手法はフィルタ幅に依存せずに, ガウシアンフィルタ処理を定時間で適用できる. これらの定時間空間フィルタ処理は Sugimoto らの研究[54]に詳しくまとめられているため, ここでは簡略化して述べる. 定時間空間フィルタ処理はフィルタの重ね掛け[55, 56]と再帰型フィルタ[57, 58], そして拡張積分画像によるフィルタ[59-61]の三つタイプに大別される. しかしながら入力画像のサイズが特に多いとき, これらのアルゴリズムであっても, マルチスケール画像を生成することは非常に計算時間がかかる.

2.6 ハードウェアシステムによる高速化

汎用 CPU には積和演算器が少ないため, 膨大な積和演算処理は不得意である. そのため,

近年、汎用 GPU (General-purpose graphics processing unit) [62], DSP (Digital signal processor), FPGA (Field-programmable gate array) (例えば [63]), そして ASIC (application specific integrated circuit) [64-67]といった専用プロセッサで画像処理を実行する研究がされており、それらが比較されている[68, 69]. 以下に、各ハードウェアでの空間フィルタ処理についてまとめる. 汎用 GPU を採用する場合, 高速に演算は実行されるが, 消費電力とシステムの規模が問題となる. またコンピュータ上のメモリから GPU 側のメモリへの画像転送速度もまた一定のボトルネックとなる. DSP を採用したとしても, DSP 内に用意されている積和演算命令に対応する演算器の数に比べて, 非常にフィルタサイズが大きい空間フィルタを画像に適用する場合, 複数の空間フィルタ処理を高速に実行することは困難である. FPGA は組み込み画像処理システムへの搭載が進んでおり, その開発の進展は目覚ましいものがあるが, 複数回空間フィルタ処理を数十 ms オーダー内で計算するには, FPGA に多くの内部ロジックを必要とする. ASIC は非常に小型・低消費電力ではあるが, 特定用途に特化しており, 例えばマルチスケール画像を解析することで, 注目領域の抽出と, そのテクスチャの解析を同時に抽出するなど様々な応用に利用していくための拡張性を備えることは難しい.

これらのデジタルベースのハードウェア高速化手法は処理時間と消費電力を, CPU ベースの画像処理システムに比べて削減できるが, それらとシステム規模は画像の入力サイズとフィルタ幅に強く依存する. 従って, 画像処理システムのフロントエンドにおいて, 空間フィルタ処理を行う演算器として, 最善の選択とはいえない.

2.7 生体の外網膜とアナログ抵抗回路網

生体の視覚神経回路で実行されている計算原理を学ぶことは, 情報処理機能を有する知能視覚センサのアーキテクチャを設計する上で有用である. なぜならば, 生体の神経回路内の計算は, 並列処理を基本としており, 空間的に広がりをもった情報である視覚情報を処理するためには並列処理が本来向いているためである. 一方, 現在多くの画像処理システムが逐次的に演算を実行するコンピュータで計算しているため, 効率化の向上にはアーキテクチャ的な限界がある.

これまでの研究で, 生体の視覚システムが多くの画像処理タスクをロバストに行うことが示されてきた. 生体の視覚系は周波数帯域ごとに画像を分離した後で, 各周波数に関して画像を処理する[4, 70]. 生物学的に合理的な視覚処理モデルがいくつか提案されており[5, 6], これらの前段の処理としても帯域分離処理が採用されている. 本研究では, 画像情報の帯域分離を目的に画像上の全画素並列に広域の空間フィルタを適用するにあたって, 生体の網膜の構造を模倣したアナログ抵抗回路網[37, 39]に着目する. 抵抗回路網による空間フィルタ処理は消費電力を最小化するために最適な解決法であると考えられている[71, 72].

図 2.10 に二次元抵抗回路網の回路図を示す. 抵抗回路網の各ノード (x, y) は画像上の画

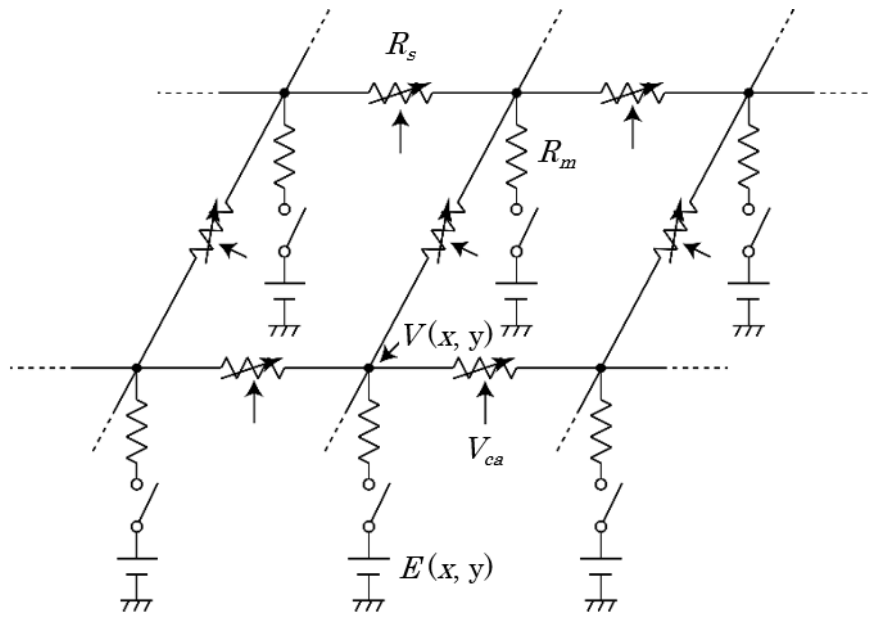


図 2.11 アナログ抵抗回路網. 抵抗回路網は入力ノードと出力ノード間を接続する R_m とノード間を接続する R_s とで構成される. 入力ノードには画素電圧 $E(x, y)$ が印加される. 画素信号は R_s を介して, 側方向に拡がり, 出力ノードには $V(x, y)$ が生成される. R_s の抵抗値の大きさで信号の拡がる度合が変わる. V_{ca} によって R_s の抵抗値の大きさを変更できる.

素に対応するように正方格子状に接続されている. 抵抗回路網は入力ノードと出力ノードを接続する抵抗 R_m と, 隣接するノードを接続する R_s で構成されている. 入力ノードには画素値に対応する電圧源 $E(x, y)$ が接続されており, 画素信号が側方に接続された抵抗 R_s を介して, 隣接するノードに拡がる. その結果, 出力ノードの画素電圧 $V(x, y)$ が生成される. R_s の抵抗値の大きさによって画素信号の拡がる度合が変わる. ここでは V_{ca} によって R_s の抵抗値の大きさを変更できるとする (後述).

図 2.11 のノード (x, y) においてキルヒホッフの電流則から, 次の式が得られる.

$$I_m(x, y) + \sum I_{adj}(x, y) = 0. \quad (2.10)$$

ここで, 電流 $I_m(x, y)$ はノード (x, y) における R_m を通る電流であり,

$$I_m(x, y) = (E(x, y) - V(x, y)) / R_m. \quad (2.11)$$

と表現できる. また $I_{adj}(x, y)$ は R_s を流れる電流であり, 次のように表現される.

$$I_{adj}(x, y) = (E(x+i, y+j) - V(x, y))/R_s, \quad (2.12)$$

となる．ここで， $(i, j) = \{(1, 0), (-1, 0), (0, 1), (0, -1)\}$ である．

問題を簡略化するために，一次元抵抗回路網を考える．一次元抵抗回路網のノード上で $x=0$ でインパルス入力を与えた時，インパルス応答は式 (2.13) のように記述される [39]．

$$V_{ir}(n\lambda) = A \exp\left(-D \frac{\ln|n|}{\lambda}\right) \quad (2.13)$$

$$2 \cosh\left(\frac{D}{\lambda} - 1\right) = \sqrt{\frac{R_s}{R_m}} = L \quad (2.14)$$

D は二つの隣接ノードの間の距離を表現し，ここで A はインパルス入力の振幅に比例する応答をとる．図 2.12 に λ が 1, 2, 4 の場合の応答の分布を示す．ここで λ の値が増加する時により多くの距離までの信号が広がる．抵抗回路網を構成する抵抗 R_s と R_m が理想的であるとき，抵抗回路網は線形である．従って，出力ノードには入力ノードの電圧に対して，インパルス応答 V_{ir} を畳み込んだ出力値の電圧分布が生成される．すなわち抵抗回路網は式 (2.13) のカーネルをもつ空間フィルタ処理を実行する．この空間フィルタを抵抗回路網フ

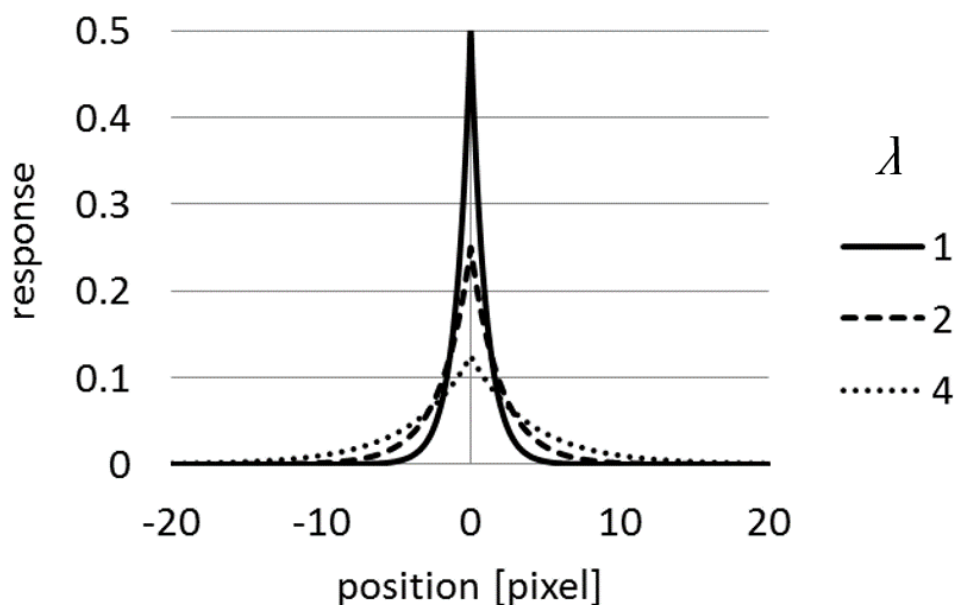


図 2.12 一次元抵抗回路網の空間インパルス応答． 応答はインパルスの入力位置から離れるほど指数的に減衰する． λ によって減衰の度合いが変わり，大きいほど側方向に広く信号が流れる．

フィルタ (RN フィルタ) と呼び, λ はそのスケールパラメータである. また λ は R_s と R_m を用いて式(2.14)のようにして表される.

これらの結果から, 抵抗回路網のフィルタサイズは抵抗値 R_s と R_m の比を変更することで制御できることがわかる. 抵抗回路網のインパルス応答, すなわち RN フィルタの重み係数は空間的に等方的であり, そして中心からの距離が増加すると, その重みが減衰する. このことから RN フィルタはガウシアンフィルタと同様に低域通過型空間フィルタとして機能することがわかる. また RN フィルタは中心からのノード番号 n とスケールパラメータ λ との比の関数となっている. 従って, RN フィルタはスケーラブル性を備えている. 抵抗回路網による空間フィルタ処理は抵抗回路網の回路の時定数に基づいた時間内 (数マイクロ秒 (μs) オーダー) に完了する[36].

抵抗回路網は外網膜の視細胞層や水平細胞層を模擬した構造である. この細胞層は視覚入力に対して平滑化フィルタ, すなわち低周波通過型空間フィルタとして働く. ここで, 同じく外網膜内にある双極細胞層は視覚入力に対して帯域通過型空間フィルタとして働くことが知られている[38]. この帯域通過型空間フィルタ処理を視覚入力に適用する神経回路のアーキテクチャは二つの抵抗回路網の出力の間の差としてモデル化されている[42]. 本論文では, この帯域通過型空間フィルタを DoR (Difference of Resistive network) フィルタと呼ぶ.

次に, DoR フィルタのスケーラブル性を考える. DoR フィルタ処理画像は二つの RN フィルタ処理画像の差によって生成される. DoR フィルタでは差をとる二つの RN フィルタのスケールパラメータを DoG フィルタの場合と同様に適切に設定することでスケーラブル性を備えた帯域通過型空間フィルタ群を用意できる. すなわち, 一つ目の RN フィルタのスケールパラメータ λ_1 と二つ目の RN フィルタのスケールパラメータ λ_2 の関係を $\lambda_2 = K\lambda_1$ とする. スケーラブル性を有する一次元 DoR フィルタの重み関数は次のように表現される.

$$f_{DR}(n/\lambda) = A_{R1} \exp\left(-D \frac{|n|}{\lambda}\right) - A_{R2} \exp\left(-D \frac{|n|}{K\lambda}\right), \quad (2.15)$$

ここで, A_{R1} と A_{R2} は正規化定数である. パラメータ λ と $K\lambda$ はそれぞれ抵抗回路網フィルタの比較的小さいスケールと比較的大きいスケールをもつ抵抗回路網フィルタの空間定数である. パラメータ n は中心からのノード番号である. この式(2.15)で示される DoR フィルタのカーネルは中心からのノード番号とスケールパラメータの比となっており, スケーラブル性を有している. このフィルタを活かすことで, デジタルマスクによる DoG フィルタ処理を, アナログ抵抗回路網の物理特性を活かした高速演算に置き換えることができ, 非常に低消費電力に瞬時にマルチスケール画像を生成できると考えられる.

2.8 シリコン網膜

生体の外網膜神経回路網の構造を模倣したアナログ抵抗回路網[37, 39]は、全ての画素に対して瞬時に低消費電力で大きなサイズの低域通過型空間フィルタ処理を入力画像に適用できる。特筆する性質として、抵抗回路網によるフィルタ処理の時間はフィルタの大きさと入力画像の大きさのどちらにも依存しない。Mead と Mahowald は小型なハードウェア構成で画像取得・フィルタ処理を同時に達成するためにアナログ集積回路技術によって、シリコン網膜と呼ばれる抵抗回路網を含む視覚センサを設計・製作した[25, 29]。しかしながら、従来のチップは、アナログ回路特有の素子特性のばらつきに起因する計算精度の低さにより、実際の応用に適用することは困難であった。

Mead らの先駆的な仕事に続いて、抵抗回路網を含んだ様々な小型・エネルギー効率の良い空間フィルタ処理機能を有する視覚センサが設計・製作されてきた[29-36]。これらのセンサのアーキテクチャは更に MOS トランジスタで構成された二層の抵抗回路網を含むように発展した[30, 31, 34, 35]。これらの二つの層の出力の差をとることで帯域通過型空間フィルタ処理画像を生成する。この画像の通過帯域は二つの層の空間周波数特性によって決まる。ガウシアンフィルタに近似的なカーネルの空間フィルタを適用できるスイットドキャパシタで構成された抵抗回路網[73]や、スイッチドレジスタ型の抵抗回路網[74]もまた開発されてきた。MOS 抵抗で構成された抵抗回路網の特筆すべき利点として、 μs オーダーでフィルタサイズが設定可能であることが挙げられる。

Kameda らによって開発されたシリコン網膜は自然照明下において比較的高い出力精度を持つことが確認されており、実用に近い性能が実現されている[35]。この集積回路はその後改良され、ニューラルイメージ社によって製品化された(“ASPECTUS-U”)。本研究ではマルチスケール空間フィルタ処理機能を有する知能視覚センサの一部として、このシリコン網膜チップを用いた。

2.8.1 内部構成と動作原理

本節ではチップの内部構成と動作原理を説明する。シリコン網膜チップの回路構成を図 2.13(a)に示す。受光部(フォトダイオード: 図中 PD) と処理回路からなる画素回路が、水平および垂直方向に 128×128 画素分並べられている。各画素は、受光面が正方格子に配置されるように並べられている(Kameda らのチップでは 40×46 画素が六方格子に配置されていた[35])。チップ出力は、垂直方向、水平方向それぞれに用意されたシフトレジスタからのデータ信号により、一画素ずつ出力線(図中 out) に接続され、順に読み出される。シリコン網膜の一画素分の回路構造を図 2.13(b) に示す。

光センサ部は、フォトダイオード(PD)の寄生容量成分を利用する電荷蓄積型の光センサを採用している。この光センサ部の動作のタイミングチャートを、図 2.14 に示す。まず、nMOSFET スイッチ(制御信号を s_{wp} と表記)を ON することで、PD に逆バイアスを印

加し、その寄生容量を充電する。これを光センサ部のリセット動作 (Reset) と呼ぶ。その後、 s_{wp} によりスイッチを OFF し、PD に光が当たると、その強度に応じて寄生容量に蓄えられた電荷が放電し、PD の電位が降下する。この電位降下の大きさが光センサ出力 (Photosensor output) となる。リセット動作を終えてから、光センサ出力値が確定するまでの時間を蓄積時間 (Accumulation time) と呼ぶ。この蓄積時間を長くすることで、光強度が弱い場合でも、出力ゲインを大きくすることができる。出力ゲインが高いこの方式を採用することにより、日常の自然照明下におけるような、照度の低い環境においても十分な感度が得られる。光センサの出力電圧は、二つの pMOSFET で構成されたソースフォロア回路を介して処理回路へと出力される。このような画素構造をもつ光センサを APS (Active pixel sensor) と呼ぶため、以降の光センサ部は APS と記述する。

本センサには二つの抵抗回路網が内蔵されている。一層目の抵抗回路網 (R_{m1}, R_{s1})、二層目の抵抗回路網 (R_{m2}, R_{s2}) は、それぞれ近傍 4 画素と接続され、光センサの出力を空間的に平滑化する。本研究では一層目の抵抗回路網は用いないため、各画素間の接続をスイッチでオフしており、二層目の抵抗回路網は用いるため、接続スイッチをオンしている。まず画素情報は固定パターンノイズ [24] をキャンセルする機能をもったノイズ補償バッファ (Nbuf) [75] に保存される。その出力が二層目の抵抗回路網 (R_{m2}, R_{s2}) で平滑化される。出力段のノイズ補償バッファは、一層目、二層目それぞれの抵抗回路網のノードの値がスイッチで切り替えられて入力、差分するように設計されているが、本研究では出力段の切り替えスイッチ ((制御信号を s_{ws} と表記)) は常に二層目側に接続し、二層目の電圧応答をそのまま出力する。抵抗回路網は、MOS 抵抗器により構成されており、図 2.13(b) の抵抗回路網に外部から印加するバイアス電圧 (V_{rmb2}, V_{rsb2}) によって、その抵抗値が変更できる。

ノイズ補償バッファの動作について、簡単に説明を行う (動作の詳細は [35] 参照)。ノイズ補償バッファでは、図 2.15(a) の二つのスイッチ (sw_{sh1}, sw_{sh2}) を、図 2.15(b) に示すように制御する事により、リセット時刻 (τ_{res}) における入力電圧 ($v_{in}(\tau_{res})$) と、リセット後の任意のデータ取得時刻 (τ_{acq}) における入力電圧 ($v_{in}(\tau_{act})$) の差 Δv_{in} に応じた電圧が出力される。出力電圧 v_{out} は、以下の式で表される。

$$v_{out} = -\frac{C_1}{C_2} \left(v_{in}(\tau_{acq}) - v_{in}(\tau_{res}) \right) + V_{ref} \quad (2.16)$$

このノイズ補償バッファの動作では、トランジスタのミスマッチによるオフセット電圧がキャンセルされる。このため、出力電圧における固定パターンノイズは 5 mV 以下となる [39]。また、光ショットノイズを始めとしたランダムノイズ [24] の影響も 5 mV 以下である。通常の入力条件では、数 100 mV の出力電圧範囲で使用できるため、ノイズは数%のオーダーとなり、大きな問題とはならない。

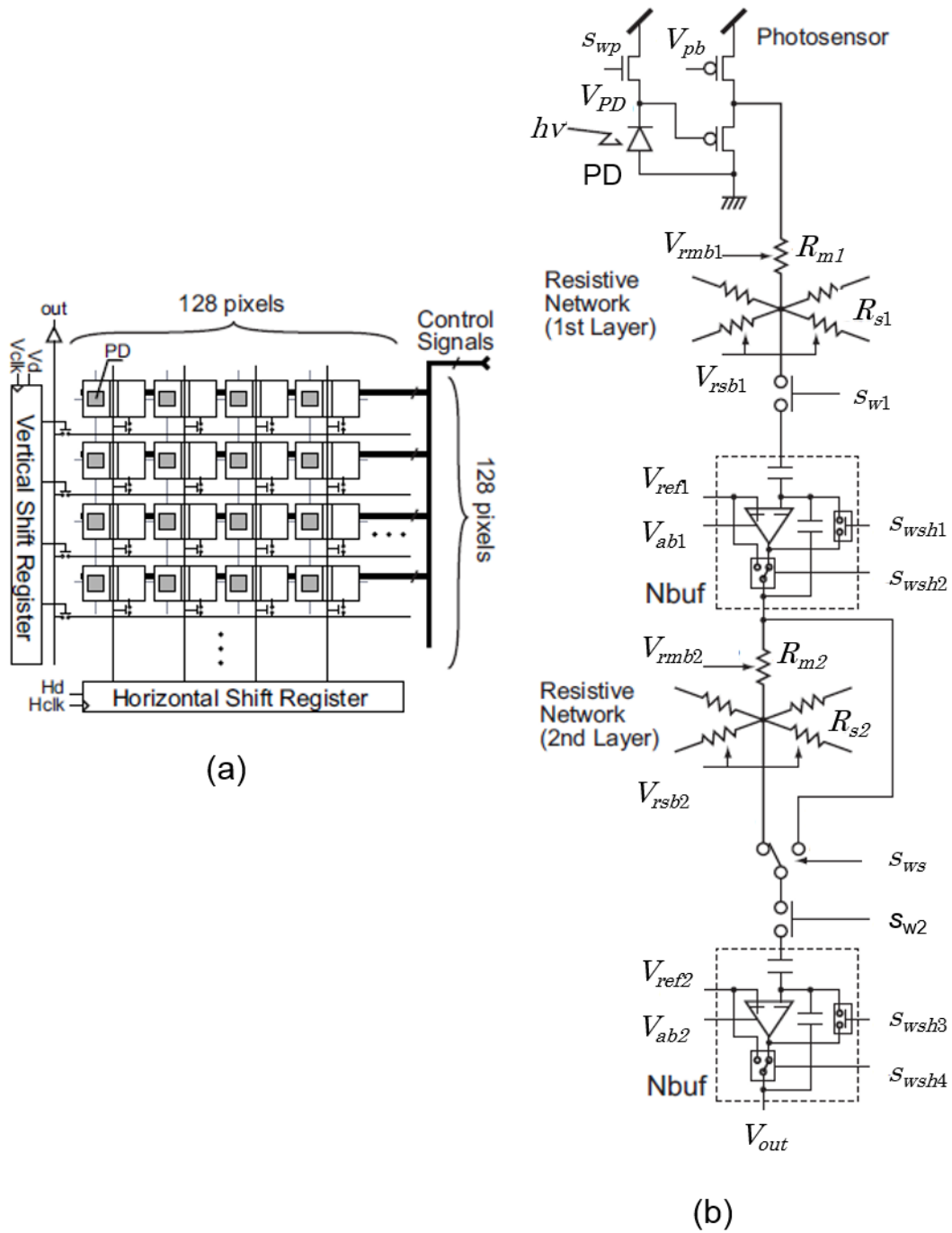


図 2.13 シリコン網膜の模式図. (a) チップ全体の回路構成.
(b) 一画素分の回路構成.

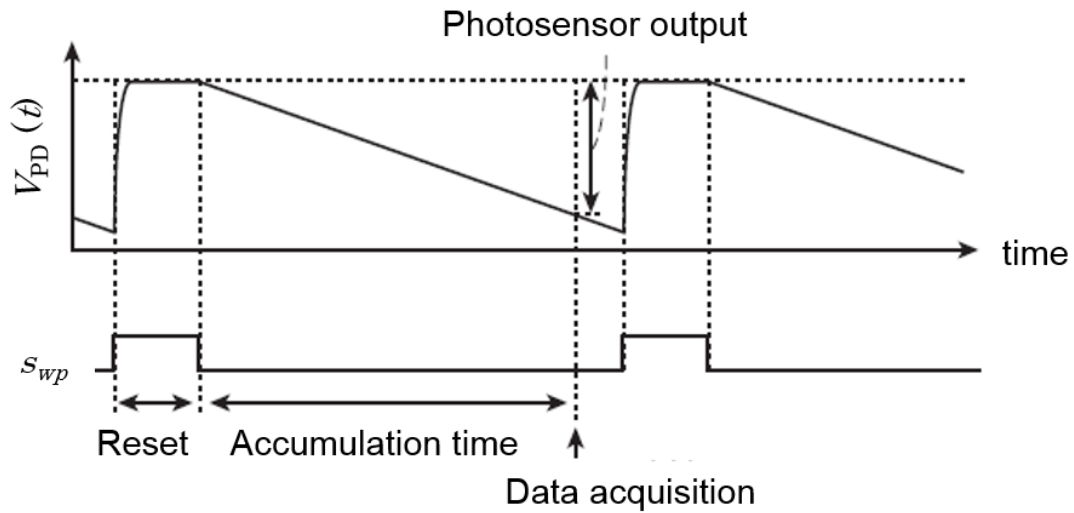


図 2.14 光センサ部の動作のタイミングチャート.

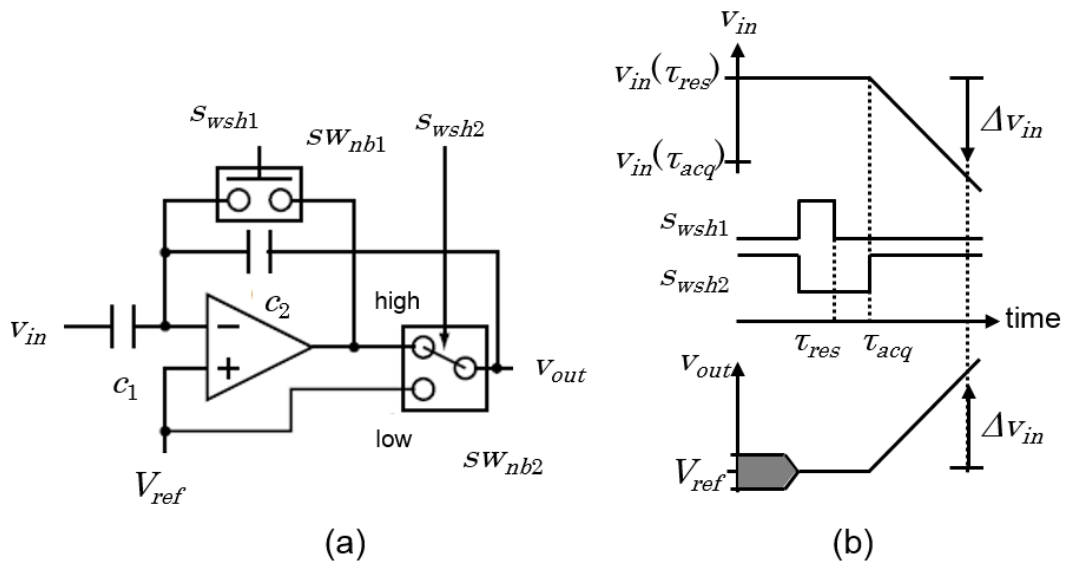


図 2.15 ノイズ補償バッファ. (a) 回路図. (b) 制御信号のタイミングチャート.

2.8.2 FPGA によるシリコン網膜の制御

シリコン網膜を用いて画像取得・空間フィルタ処理出力を得るために用いた制御信号を 図 2.16 に示す. チャート左に記された信号名は, シリコン網膜の回路図 (図 2.13) に記された信号名と対応する. 最上段はフォトダイオードの出力電圧を示し, 次段から八段目までのタイミングチャートはノイズ補償バッファのスイッチを制御する信号を示している. APS とノイズ補償バッファの動作については図 2.14, 図 2.15 で詳しく述べた. ノイズ補償

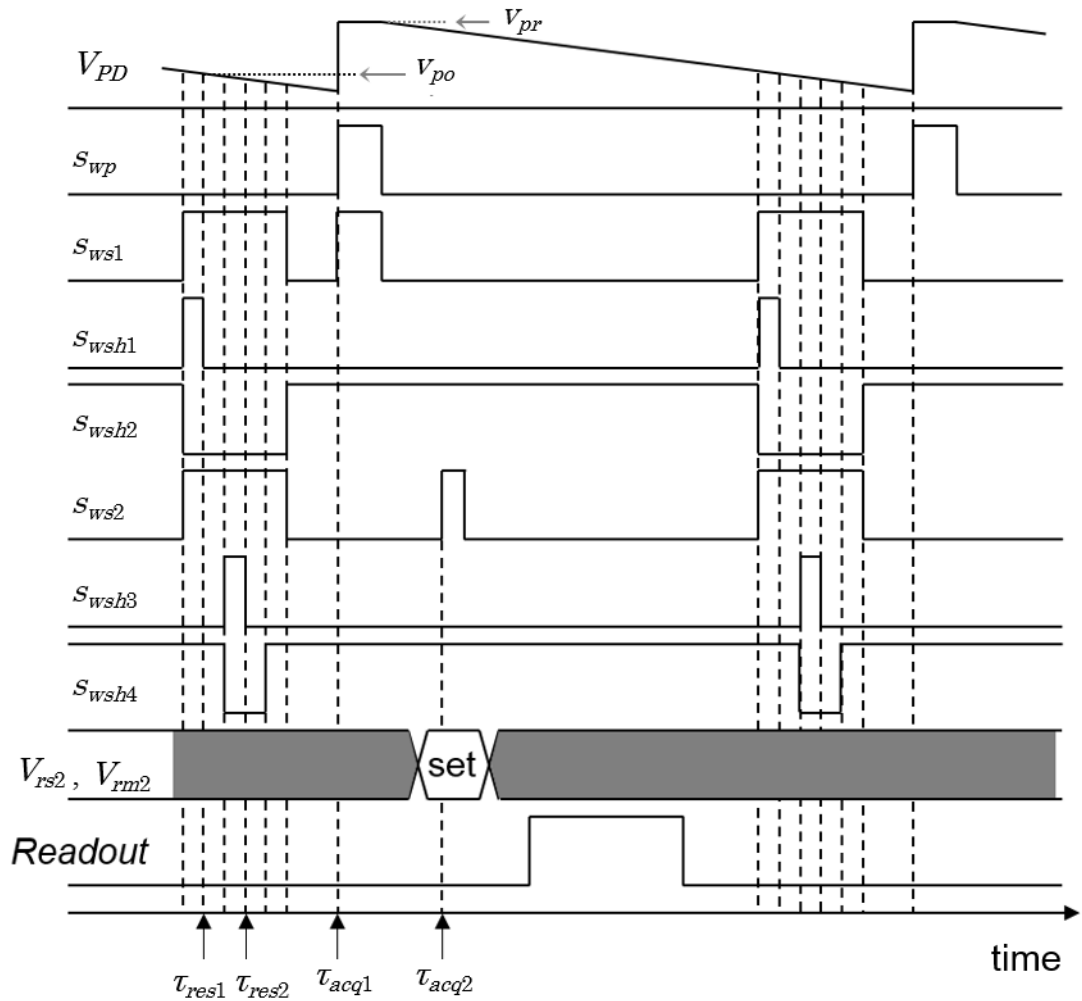


図 2.16 シリコン網膜の制御信号のタイミングチャート. 最上段はフォトダイオードの出力電圧 V_{PD} を, 次段から八段目まではノイズ補償バッファのスイッチを制御する信号のタイミングチャートを示す.

バッファはサンプルアンドホールド (S/H) 回路としての機能もまた果たしている. ここでは, 本研究で用いたシリコン網膜の持つ二つの S/H 回路 (S/H1, S/H2) を用いて APS の蓄積動作を途切れなく実行し, 二層目の抵抗回路網の出力画像を取得する制御手法を述べる.

二層目の抵抗回路網出力を取得する手法として, 前節と述べた APS 及び S/H 回路による画像取得と異なる部分は APS の蓄積動作終了時に S/H1, S/H2 のリセット時間 τ_{res1} , τ_{res2} を設定することである. τ_{res1} 時の S/H1 の出力 v_{out1} は式 (2.16) に従い,

$$v_{out1} = -\frac{C_1}{C_2} \left(v_{in1}(\tau_{acq1}) - v_{in1}(\tau_{res1}) \right) + V_{ref1}. \quad (2.17)$$

となる．この時刻では， $v_{in1}(\tau_{acq1})=v_{po}$ ， $v_{in1}(\tau_{res1})=v_{po}$ が等しいため， $v_{out1}=V_{ref1}$ となる．従って， τ_{res2} 時の S/H2 のリセット時の電圧 $v_{in2}(\tau_{res2})$ は，以下のようになる．

$$v_{in2}(\tau_{res2}) = V_{ref1}. \quad (2.18)$$

次に， τ_{acq1} 時の S/H1 の出力は，

$$v_{out1} = -\frac{C_1}{C_2} (v_{pr} - v_{po}) + V_{ref1}, \quad (2.19)$$

となり，最終的に τ_{acq2} 時の S/H2 の出力は，次式で表わされる．

$$v_{out2} = \frac{C_3}{C_4} \times G \left(\frac{C_1}{C_2} (v_{pr} - v_{po}) \right) + V_{ref2}. \quad (2.20)$$

ここで， $G(\cdot)$ は二層目の抵抗回路網による変換（平滑化处理）を表現している．式 (2.20) から， v_{pr} より v_{po} の方が電圧の値が大きいため， V_{ref2} を基準としてマイナス方向に信号応答が出力されるため，画素応答として正の値の方向でより明るい画素を表現するならば，A/D 変換後，外部メモリに保存する際，正負の反転処理が必要となる．また式 (2.19) は二層目の MOS 抵抗回路網への入力レベルと関係するため，次節で述べられる MOS 抵抗の動作域と関係する．

2.8.3 MOS 抵抗回路網と MOS 抵抗

シリコン網膜に内蔵する抵抗回路網の抵抗器は MOS 抵抗により実現されている[34]．MOS 抵抗は抵抗値を高く，消費電力を低く設定でき，回路構成も単純なので高集積化に適している．また外部から与えるバイアス電圧によりその抵抗値を変化させることが可能であり，抵抗回路網の空間フィルタ特性の変更の実現に重要な役割を果たす．まず最も単純化した MOS 抵抗について説明した後 ASPECTUS-U で用いられている MOS 抵抗回路の説明を行う．

MOS 抵抗回路の最小の構成単位である MOSFET (nMOS で説明) を図 2.17(a)に，その V_{in} - R_{ds} 特性を図 2.17(b)に示す[76]．MOS 抵抗は両端の電圧 V_{ds} と V_{gs} によって，抵抗値が定まる．MOS の構造は対称的であるため，nMOS の場合，ドレイン，ソースの高い電圧が印加した方がドレイン，もう一方がソースとなる．説明のため左側をドレインとし，入力電

圧を加え、右側をソースとする。ゲート電圧 V_g を一定 V_{dd} とし、 V_{in} を徐々に大きくする。 V_{in} が小さい時、 V_g の方が V_s より V_{th} 以上高くなり、MOS は非飽和領域で動作する。さらに V_{in} を大きくし、 $V_{dd} - V_{th}$ に近づくと V_{gs} が小さくなり、図のようにオン抵抗は急激に大きくなる。ここで抵抗の線形性が保たれた動作範囲として、 V_{ds} が $V_{dd} - V_{th}$ より十分小さいという条件が成り立つ場合、抵抗値は下式のようになる。

$$R_{on} = - \frac{1}{\mu_n C_{ox} \frac{W}{L} (V_g - V_{in} - V_{th})}. \quad (2.21)$$

例として R_{on} の変化を 1/4 に制限するときは入力レベルの最大値は下式で与えられる。この数式内の W はチャネル幅、 L はチャネル長を示す。

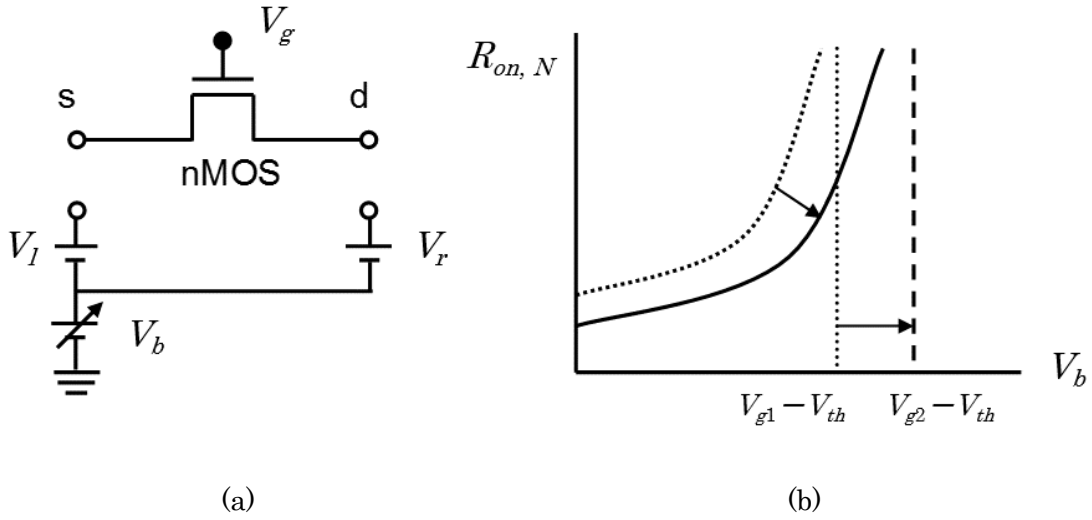
$$V_{in,max} = - \frac{3}{4} (V_g - V_{th}). \quad (2.22)$$

この値は $V_{dd}/2$ 程度まで劣化するため、入力振幅が極端に制限されることになる。またトランジスタのしきい値電圧は直接的に入力振幅を制限する。

次に線形性を高め、入力振幅範囲を広げるため、図 2.18(a) に示すような CMOS 構成による MOS 抵抗について考える。NMOS と PMOS のオン抵抗の増加特性が入力電圧レベルに対し、反対の特性を有することを利用する。この MOS 抵抗はグラウンドから V_{dd} までのレールツーレール動作が可能であるが、相補的な制御電圧 (V_{gn}, V_{gp}) が必要となる。図 2.18(b) に一般的な R_{on} の挙動を示す。オン抵抗の変動は各々単独の場合より非常に小さく、入力依存性が小さくなっていることがわかる。

本研究で用いた抵抗器の特性は抵抗回路網の空間周波数特性に影響を与えるため、画像処理として利用できるように、抵抗値を制御できなければならない。単独の MOSFET では、実現できる抵抗値に限界があるため、通常 MOS 抵抗は、複数の MOSFET により構成されている[34]。ここで Kameda らの提案したシリコン網膜に使用されている MOS 抵抗が nMOS ベースで構成されているのに対して[35]、その線形性を上げ、動作範囲を広げるため、本研究で採用する ASPECTUS-U の抵抗回路網では CMOS 構成に改善されている[36]。以下でその特性を説明する。

本研究では、二次元抵抗回路網の R_s の値を変更することによって、その空間特性を制御する。その MOS 抵抗回路の回路図、対応する回路記号とその特性の模式図をそれぞれ図 2.19(a)-(c) に示す。MOS 抵抗回路は抵抗部とバイアス電圧生成回路部の二部により構成される。各 MOS 抵抗はノード N_L と N_R で隣り合う抵抗に接続される。MOS 抵抗の抵抗値はソース-ゲート間の電圧に強く依存する。図 2.19(c) に示すように抵抗部は PMOS トランジスタと NMOS トランジスタで構成されており、その相補性から入力電圧依存性を抑える構

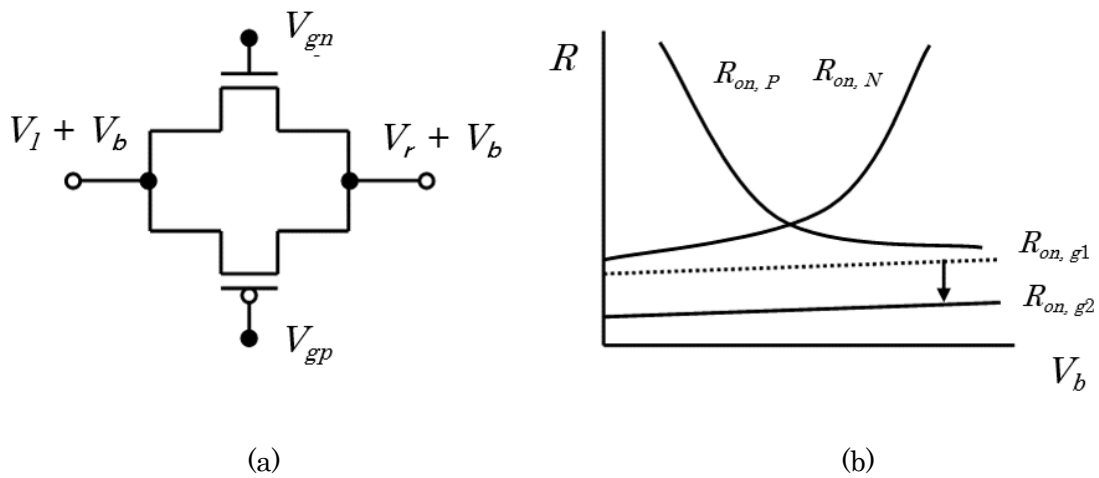


(a) (b)

図 2.17 MOS 抵抗. (a) nMOSFET. (b) 可変抵抗特性.

成をとっている.

バイアス回路部は V_{rsbn} と V_{rsbp} の電圧によって決定される電圧を PMOS トランジスタと NMOS トランジスタにそれぞれ適用する. バイアス電圧 V_{rsbn} を高くすることで抵抗部の PMOS の抵抗値を下げる事ができ, V_{rsbp} を低くすることでも, NMOS の抵抗値を下げる事ができる. 電流の通り道となる PMOS と NMOS の両方の抵抗値が増加, もしくは減少するように V_{rsbn} と V_{rsbp} を同時に制御することによって, R_s の抵抗値を制御することができる. バイアス電圧によって制御される MOS 抵抗回路の抵抗値を変化させるためにかかる時間は非常に短く, 本研究で用いた抵抗回路網の場合, μs オーダーである. また研究段階ではあるが, 更に抵抗値の精度の良い, 抵抗回路網もまた提案されている [74].



(a) (b)

図 2.18 CMOS 構成の MOS 抵抗. (a) CMOS 回路. (b) 可変抵抗特性.

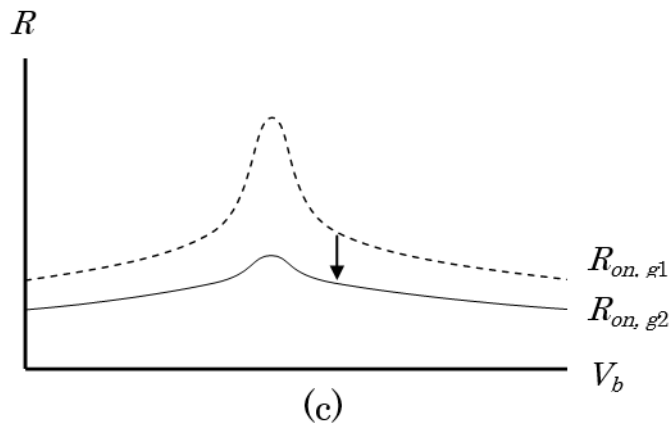
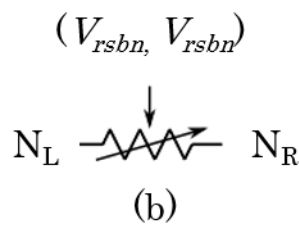
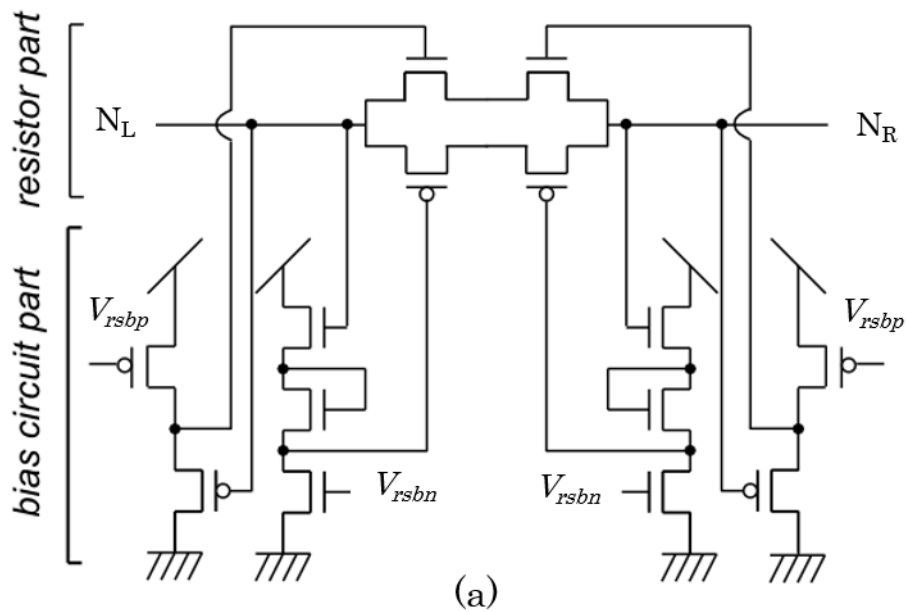


図 2.19 (a) MOS 抵抗回路 R_s の回路図[40]. (b) MOS 抵抗回路に対応する回路記号. 各 MOS 抵抗はノード N_L と N_R で隣り合う抵抗と接続されている. MOS 抵抗 R_s の抵抗値は外部制御電圧の組み合わせ (V_{rsbn} , V_{rsbp}) で決定される. (c) バイアス電圧に対する抵抗値の変化.

2.9 結言

本章では、逐次型演算による空間フィルタ処理について述べた後に、マルチスケール空間フィルタ処理の計算コストに関する問題点を述べた。空間フィルタ処理の高速化についてソフトウェアとハードウェアのそれぞれの観点から、これまで行なわれてきた高速化手法について紹介した。

本研究では、従来のデジタル演算器ではなく、生体網膜の神経回路の構造に学んだアナログ抵抗回路網に着目する。この抵抗回路網では全画素並列に効率的に空間フィルタ処理を実行できる。また構成している抵抗の値を変更することで、適用する空間フィルタの空間特性を制御できる。これまでの研究では、光センサと二層の抵抗回路網で構成されるシリコン網膜と呼ばれる知能視覚センサが提案されており、これは効率よく、ある一つの周波数帯域を通過させる帯域通過型空間フィルタ画像を取得できるものであった。

上記で述べたシリコン網膜は、CMOS集積回路技術で実現されており、その抵抗はMOSトランジスタで構成されている。MOS抵抗は外部からの制御電圧により、抵抗値を制御できる。次章からは、このMOS抵抗の抵抗値を制御することで、様々な空間特性の空間フィルタ処理を実時間で実行する知能視覚センサの開発について述べる。

第3章 複数の帯域通過型空間フィルタ処理 機能を有する知能視覚センサの開発

3.1 緒言

自然画像上には、様々なスケールの様々な幾何学的特徴を持つ物体が表現されている。そのような画像には様々な空間周波数成分が含まれている[10]。画像から情報を抽出するためには、複数の異なる帯域通過特性を持つ空間フィルタを用いて、空間周波数帯域ごとに画像を分割し、様々なスケールで特徴を解析することが有効である[2, 7, 8]。マルチスケール解析や周波数分離画像処理には多くの利点があることが先行研究によって示され、コンピュータビジョン分野の研究者に注目されてきた[13-21]。例えば、マルチスケール解析はステレオマッチング[12, 15]、オプティカルフロー推定[17]など幅広く利用されている。しかしながら、デジタル画像処理システム上での逐次計算によるアルゴリズムを採用した場合、積和演算の多い空間フィルタ処理を複数回画像に適用しなければならないため、この処理は非常に高い計算量となる。そのため実時間で画像を空間周波数帯域ごとに分離することは難しい。

第2章で述べたように、ソフトウェア[54-61]及びハードウェア[62-67]の工夫による画像センシングシステムの高速度の研究が行われてきた。本研究では網膜の構造を模した集積回路である、シリコン網膜と呼ばれる視覚センサに着目する。この視覚センサは全面素同時に瞬時に空間フィルタが実行できる抵抗回路網を有する。過去の研究では、異なる抵抗値の二層の抵抗回路網を用いて、固定の帯域通過特性を有する空間フィルタ、すなわちシングルスケールの空間フィルタを画像に適用するシリコン網膜が提案されている[35]。本研究では、単一フレーム内で異なる通過帯域をもつ複数の空間フィルタを画像に適用することにより、空間周波数帯域ごとに画像を分離することのできる知能視覚センサのアーキテクチャを提案する。更にそのアーキテクチャに基づいて新しい知能視覚センサを構築し、そのセンサに対してテスト画像を提示し、センサの出力画像を評価した。実験の結果から、本システムの有用性を示すことができた[77, 78]。本章では、その知能視覚センサの開発に関する研究内容をまとめる。

更に次節では、提案する知能視覚センサのアーキテクチャと処理の流れについて述べる。次いで、提案システムの空間特性と空間周波数特性を測定し、評価した結果について述べる。その後、センサの有用性を実証するため、開発した知能視覚センサにテスト画像を提示し、その出力を評価した結果を述べる。最後に本センサで単一フレーム内に適用できる空間フィルタ処理の数と、抵抗回路網の入力依存性について考察し、本章をまとめる。

3.2 知能視覚センサの構成と処理の流れ

図 3.1 に本研究で提案された知能視覚センサの構成を示す。本センサはシリコン網膜 (128 × 128 画素), FPGA (XC3S400, Xilinx), A/D 変換器, そして USB インターフェースで構成されている。このシリコン網膜は APS, 二つのサンプルアンドホールド回路 (図中

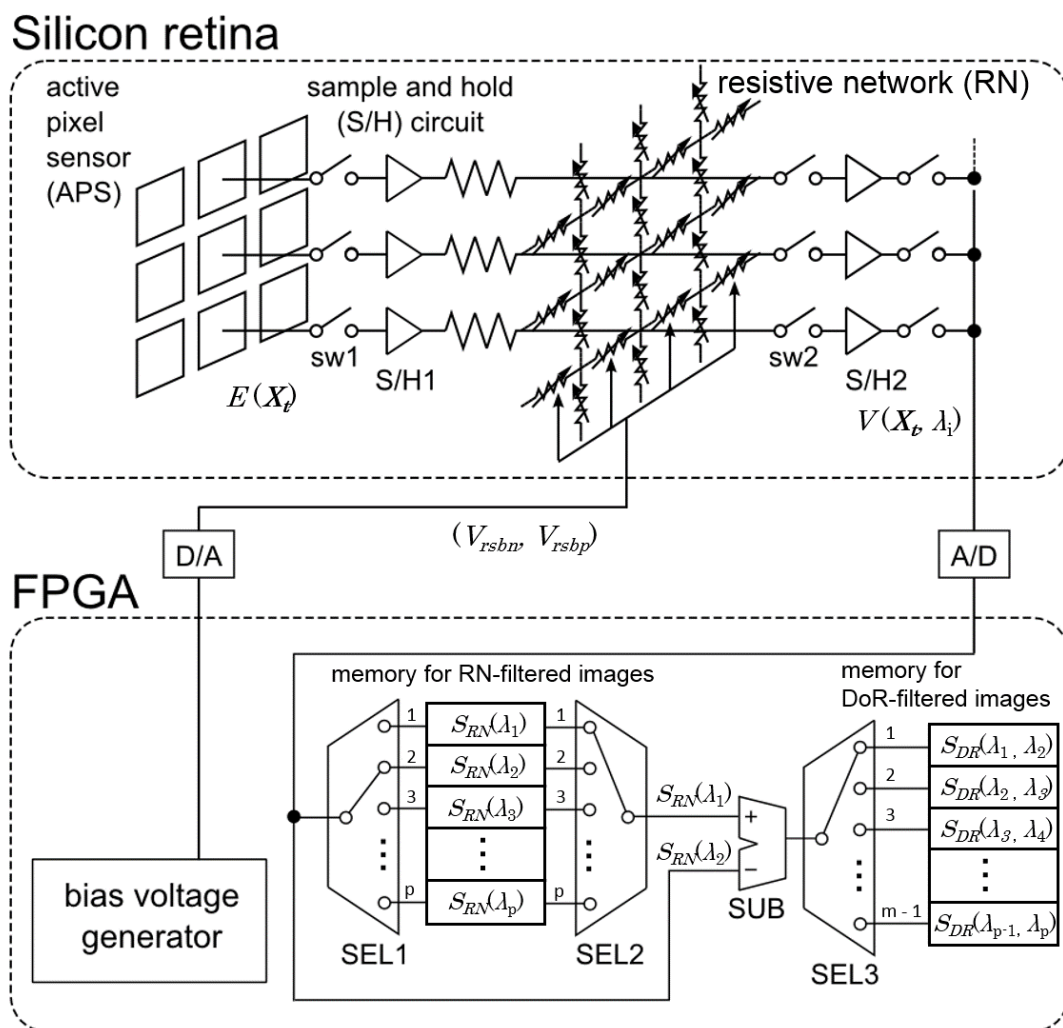


図 3.1 提案した知能視覚センサの構成。本センサはシステム網膜と FPGA で構成される。このシリコン網膜は APS と抵抗回路網で構成されている。抵抗回路網の抵抗値はバイアス電圧 V_{rsbw} と V_{rsbp} によって制御される。バイアス電圧は FPGA から D/A 変換器を介して抵抗回路網に与えられる。FPGA 内の $S_{RN}(\lambda_n)$ と $S_{DR}(\lambda_n)$ のブロックはそれぞれ RN フィルタ処理画像と DoR フィルタ処理画像を記録するためのメモリを表現している。SEL1, SEL2, そして SEL3 はセレクタを表現している。図 3.2 にこれらのセレクタによって選択された接続に関するタイミングチャートを示す。

S/H1, S/H2), 抵抗回路網 (図中 RN), そして読み出し回路で構成されている. FPGA 内には APS, S/H 回路と読み出し回路のための制御信号を生成するデジタル回路, 抵抗回路網へのバイアス電圧を指定するためのコマンド信号を生成するデジタル回路, そして画像処理のためのデジタル回路が実装されている. 図 3.2 に画像取得, RN フィルタの適用, FPGA 上での演算に関するタイミングチャートを示す. ここでは t フレーム目に実行される処理の流れが示されている.

まず図 3.2 内の時刻 τ_s に一つ目のスイッチ (図 3.1 内の sw1) が閉じ, x - y 平面に配置された t フレーム目の APS の出力 $E(\mathbf{X}_t)$ が S/H1 に保持される. ここで, $\mathbf{X}_t = (x, y, t)$ であり, $x=1, 2, \dots, 128$ と $y=1, 2, \dots, 128$ はそれぞれ画像の横方向と縦方向の画素の位置である. また $t=1, 2, \dots$ はフレーム数である. S/H1 への保持動作が終了後, sw1 を開き, APS が次のフレームの蓄積動作を始める. この動作により, 蓄積動作と後段の処理を切り離すことができ, 途切れなく画像取得・画像処理ができる.

S/H1 に保存された $E(\mathbf{X}_t)$ は, 抵抗回路網によって瞬時にフィルタ処理され, その出力応答 $V(\mathbf{X}_t, \lambda)$ が S/H2 に保持される. 2.6 節で述べたように, RN フィルタのスケールパラメータ λ の値は抵抗回路網を構成する抵抗器の抵抗値の比によって決定される. FPGA から送られた外部バイアス電圧 V_{rsbn} , V_{rsbp} の組み合わせによって, まず最もスケールパラメータが小さくなるように, すなわちカットオフ周波数が最も高くなるように, 抵抗回路網の側方抵抗 R_s の抵抗値が最も高い値に設定される. この時の抵抗回路網のスケールパラメータ λ_1 とし, 空間フィルタを $f_{RN}(\lambda_1)$ と表現する. sw2 が閉じ, S/H2 には $f_{RN}(\lambda_1)$ でフィルタ処理された画像 $S_{RN}(\lambda_1)$ が保存される. sw2 が開かれた後, フィルタ処理画像 $S_{RN}(\lambda_1)$ が A/D 変換器によって, 8bit に変換後, FPGA に送られ, $S_{RN}(\lambda_1)$ 用のメモリ領域に保存される.

次に, ある一定の値分 R_s の抵抗値を減少することで, スケールをより低いカットオフ周波数を持つ RN フィルタ $f_{RN}(\lambda_2)$ に設定する. 図 3.2 内の時刻 τ_1 のタイミングで, 空間フィルタ $f_{RN}(\lambda_2)$ を $E(\mathbf{X}_t)$ に適用する. その後 sw2 に閉じ, フィルタ処理画像 $S_{RN}(\lambda_2)$ を S/H2 に保存する. $S_{RN}(\lambda_1)$ の場合と同様に $S_{RN}(\lambda_2)$ は FPGA に転送され, $S_{RN}(\lambda_2)$ 用のメモリ領域に保存される.

次のサンプル時間 (図 3.2 内 τ_e) まで, この手順の繰り返すことで, 様々なカットオフ周波数の RN フィルタで処理された低域通過処理画像群が生成される. t フレーム目と同様に, $t+1$ 番目のフレームの APS 出力 $E(\mathbf{X}_{t+1})$ が, 時刻 τ_e で sw1 が再度閉じ, S/H1 に保存される. 抵抗回路網の処理時間は画像の読み出し時間と比較して十分短いため, 一フレームの中で適用できるフィルタの最大の数 m は読み出し時間によって制限される. このフィルタ処理画像の数 m については考察で議論する. 本研究では m の値は 4 とした.

取得された RN フィルタ処理画像群を FPGA 上の回路で差を取ることで帯域ごとに画像を分離する処理が実行される. 異なる低域通過特性を有する RN フィルタ処理画像間の差を取る処理は生画像に帯域通過型空間フィルタを適用することと等価である. この帯域通過型空間フィルタ処理を DoR フィルタ処理と呼び, 処理画像を DoR フィルタ処理画像と

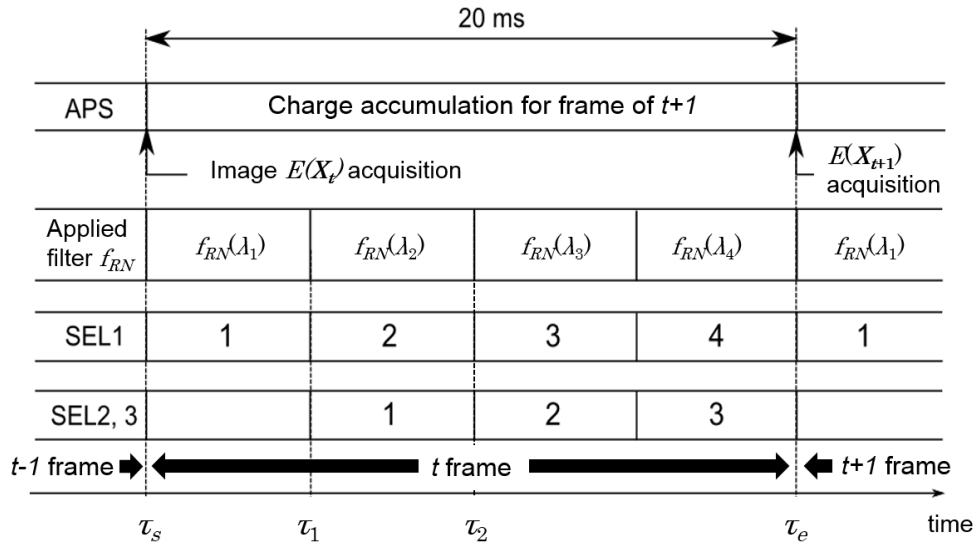


図 3.2 画像取得と抵抗回路網によるフィルタ適用についてのタイミングチャート。最上段の行は画像取得と APS の蓄積開始の時刻を示す。二番目の行は様々なカットオフ周波数で空間フィルタの適用の時刻。三番目と最下段の行は、個々の空間フィルタ適用とフィルタ処理画像の取得のためのセレクタ SEL1, SEL2, そして SEL3 の時刻をそれぞれ示す。

呼ぶ。 n 番目の DoR フィルタ処理画像 $S_{DR}(\lambda_n)$ は、シリコン網膜部から $S_{RN}(\lambda_{n+1})$ が読み出される区間に、並列に生成される。この処理について図 3.1 の記号を用い、図 3.2 内の τ_1 - τ_2 の区間を例に説明する。時刻 τ_1 において既に $S_{RN}(\lambda_1)$ は RN フィルタ処理画像を保存するメモリ領域 (図 3.1 中の memory for RN-filtered images ブロック。以下、RN 用メモリ) に保存されており、次の τ_2 までの区間で $S_{RN}(\lambda_2)$ が一画素ずつ読み出され、 $S_{RN}(\lambda_1)$ とは異なるメモリ番地に保存される。メモリの書き込み番地はセレクタ SEL1 によって指定され、 τ_1 - τ_2 区間では SEL1 は端子 2 に接続されている。 $S_{RN}(\lambda_2)$ は RN 用メモリへ保存されると同時に、差分器 SUB に入力される。SEL2 は RN 用メモリの読み出し番地を指定するためのセレクタである。一つ前の区間に保存された画像 $S_{RN}(\lambda_1)$ を読み出すために、SEL2 は端子 1 に接続される。読み出された $S_{RN}(\lambda_1)$ は SUB のもう一端に入力される。SUB で $S_{RN}(\lambda_1)$ と $S_{RN}(\lambda_2)$ の差が取られ、得られた $S_{DR}(\lambda_1)$ の画素値がセレクタ SEL3 を介して DoR フィルタ処理画像を保存するメモリ領域 (図 3.1 中の memory for DoR-filtered images ブロック。以下、DoR 用メモリ) に保存される。 $S_{DR}(\lambda_n)$ の λ_n の添え字 n が小さいほど、帯域がより高い周波数であることを表しており、より詳細な構造を表現している画像となる。最終的に、これらのフィルタ処理画像は、USB インターフェースを介して、デジタルコンピュータのような後処理用の画像処理装置に転送される。以上の仕組みにより、この智能視覚センサは単一フレーム (20ms) の間に複数の空間フィルタ処理とフィルタ処理画像の読み出しを完了できる。

3.3 MOS 抵抗回路網の空間特性の測定

提案した知能視覚センサを用いて、内蔵されている MOS 抵抗回路網の空間特性の測定を行った。図 3.3 に示すように、50cm の距離でセンサ前方に置かれた液晶ディスプレイ (Liquid Crystal Display ; LCD) 上に視覚刺激を提示することによって評価を行った。視覚刺激にスリット光のパターンを用いることで、抵抗回路網の線拡がり関数 (line spread function) を測定した。光パターンは搭載されたレンズ (最大 f 値 : 4) を介してセンサ受光面に投射される。抵抗回路網の空間特性は MOS 抵抗 R_s への電圧 V_{rsbn} と V_{rsbp} の両方を変更することで制御できる (2.6 節, 2.7.3 節を参照)。以下、バイアス電圧は (V_{rsbn} , V_{rsbp}) と表現する。まず四つの異なる低域通過特性をもつ RN フィルタ $f_{RN}(\lambda_1)$, $f_{RN}(\lambda_2)$, $f_{RN}(\lambda_3)$, $f_{RN}(\lambda_4)$ の空間特性を測定した。これらのフィルタのスケールパラメータ λ_1 , λ_2 , λ_3 , λ_4 に対応するバイアス電圧の組み合わせは、それぞれ (0.50V, 2.75V), (0.60V, 2.40V), (0.73 V, 2.20 V), そして (1.25V, 1.43V) である。またこの時の各空間フィルタの R_s の関係を以下に示す。

$$R_s \text{ of } f_{RN}(\lambda_1) > R_s \text{ of } f_{RN}(\lambda_2) > R_s \text{ of } f_{RN}(\lambda_3) > R_s \text{ of } f_{RN}(\lambda_4). \quad (3.1)$$

図 3.4(a) にスリット光を視覚刺激とした時の抵抗回路網による低域通過型空間フィルタ処理画像 $S_{RN}(\lambda_n)$ ($n=1, 2, 3, 4$) の水平方向の応答振幅の分布を示す。これはフィルタ処理画像の 44 番目から 84 番目までの水平方向 40 行の応答の平均である。この図が示すように、信号は R_s がより低い値の時ほど、側方向により長く拡がる。これは、2 章で示したシミュレーションと同様の結果である (図 2.12 参照)。

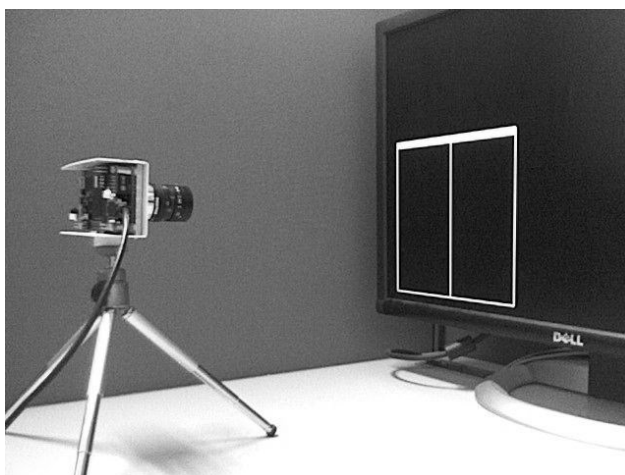


図 3.3 実験環境。知能視覚センサは LCD に描画された視覚刺激 (ここでは縦のスリット光) を受け取り、USB インターフェースを介して PC に帯域通過型空間フィルタ処理画像を転送する。

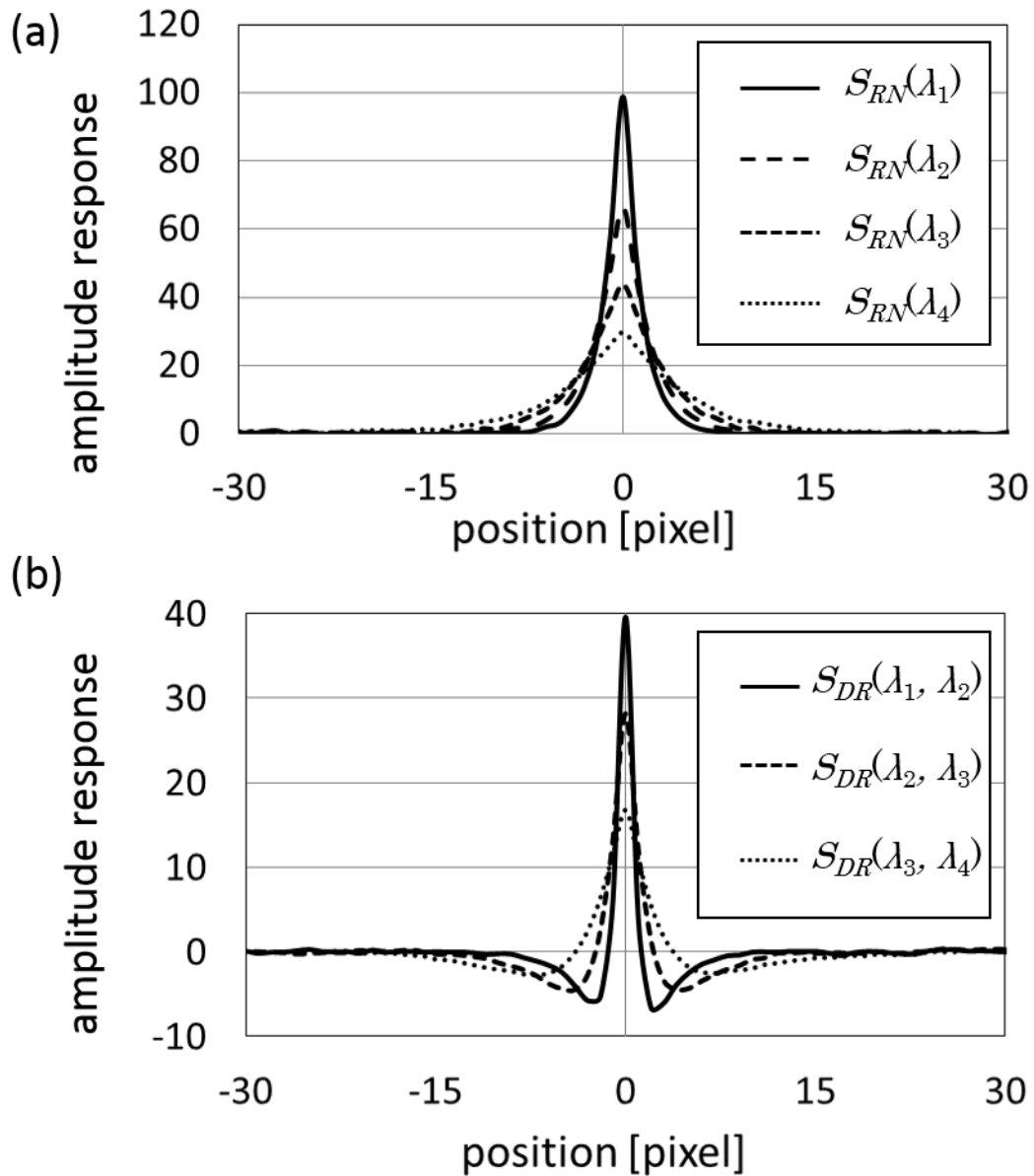


図 3.4 提案した知能視覚センサの空間特性. (a) 抵抗回路網による低域通過型空間フィルタの線拡がり関数. 凡例の λ の添え字 n は抵抗回路網フィルタの各スケールパラメータの番号を示す. この番号は図 3.2 と対応している. n がより大きいとき, よりスケールパラメータ λ も大きくなり, 抵抗回路網フィルタがより低いカットオフ周波数を持つことを表している. λ の値はバイアス電圧の組み合わせ (V_{rsbn} , V_{rsbp}) V で設定する. 四つの低域通過型空間フィルタ $f_{RN}(\lambda_1)$, $f_{RN}(\lambda_2)$, $f_{RN}(\lambda_3)$, $f_{RN}(\lambda_4)$ に与えられたバイアス電圧はそれぞれ (0.50, 2.75) V, (0.60, 2.40) V, (0.73, 2.20) V, (1.25, 1.43) V である. (b) 知能視覚センサによる帯域通過型空間フィルタの線拡がり関数. 三つのサイズの帯域通過型空間フィルタが実現できている.

次に知能視覚センサによって適用できる帯域通過特性を持つ DoR フィルタ f_{DR} の空間特性を同様の手法で測定した. 図 3.4(b) にスリット光を視覚刺激として提示した時に得られた三つの DoR フィルタ処理画像 $S_{DR}(\lambda_1, \lambda_2)$, $S_{DR}(\lambda_2, \lambda_3)$, $S_{DR}(\lambda_3, \lambda_4)$ の水平方向の応答振幅の分布を示す. RN フィルタ処理画像の場合と同様に, この応答振幅は 44 番目から 88 番目までの水平方向の応答振幅の平均である. $S_{DR}(\lambda_n, \lambda_{n+1})$ は図 3.4(a) に示される異なるカットオフ周波数で低域通過処理をされた RN フィルタ処理画像 $S_{RN}(\lambda_n)$ と $S_{RN}(\lambda_{n+1})$ の間の差を取ることで得られた. 各応答は, 他の帯域通過型空間フィルタの線拡がり関数と同様に, 中心で正に, 周辺で負に応答していることがわかる. この結果から, 開発した知能視覚センサにより三つのサイズの帯域通過型空間フィルタが適用できることを実証できた.

3.4 MOS 抵抗回路網の空間周波数応答の測定

LCD に様々な空間周波数の正弦波パターン (図 3.5(a)) を提示することによって, MOS 抵抗回路網の空間周波数特性を測定した. 前節で述べた四つの RN フィルタ $f_{RN}(\lambda_1)$, $f_{RN}(\lambda_2)$, $f_{RN}(\lambda_3)$, $f_{RN}(\lambda_4)$ の特性を測定対象とし, 抵抗回路網を構成する MOS 抵抗 R_s へのバイアス電圧は前節と同様とした. 空間周波数応答は各空間周波数の正弦波パターンに対する抵抗回路網のゲインである. 抵抗回路網のゲインは A_{out}/A_{in} の式で計算された. ここで A_{out} は RN フィルタ処理画像 $S_{RN}(\lambda_n)$ の正弦波 (図 3.5(a) 右図の灰色の実線) の振幅応答である. A_{in} はスケールパラメータ λ を最小に設定した場合の RN フィルタ処理画像 $S_{RN}(\lambda_0)$ の正弦波 (図 3.5(a) 右図の黒色の実線) の振幅応答である. ここで R_s へのバイアス電圧を制御し, その抵抗値を最大にした時のスケールパラメータを λ_0 とする.

図 3.5(b) にこうして求めた RN フィルタ $f_{RN}(\lambda_1)$, $f_{RN}(\lambda_2)$, $f_{RN}(\lambda_3)$, $f_{RN}(\lambda_4)$ の空間周波数応答を示す. この結果は抵抗回路網が低域通過型空間フィルタとして動作し, R_s の抵抗を変更することで, そのカットオフ周波数が制御できることを示している. 四つの抵抗回路網フィルタ $f_{RN}(\lambda_1)$, $f_{RN}(\lambda_2)$, $f_{RN}(\lambda_3)$, $f_{RN}(\lambda_4)$ の R_s の抵抗値は式(3.1)で述べた関係を持つ. 従って, 図から λ_n の添え字 n が小さいほど, すなわち R_s の抵抗値が増加するほどカットオフ周波数が増加していることがわかる.

次に, 知能視覚センサによって適用できる DoR フィルタ $f_{DR}(\lambda_1, \lambda_2)$, $f_{DR}(\lambda_2, \lambda_3)$, $f_{DR}(\lambda_3, \lambda_4)$ の空間周波数応答を測定した. 測定手法は DoR フィルタ処理画像 $S_{DR}(\lambda_n, \lambda_{n+1})$ の出力振幅を A_{out} とした点以外は RN フィルタを測定した場合と同様である. 測定結果を図 3.5(c) に示す. この結果は $f_{DR}(\lambda_1, \lambda_2)$, $f_{DR}(\lambda_2, \lambda_3)$, $f_{DR}(\lambda_3, \lambda_4)$ が, それぞれ画像中で 6.4 画素, 10.9 画素, 28.2 画素程度の大きさの周期構造を主に強調できるフィルタであると示している. 上記より, 開発した知能視覚センサが R_s を制御することで, 異なる通過帯域を有する空間フィルタを適用できることを実証できた.

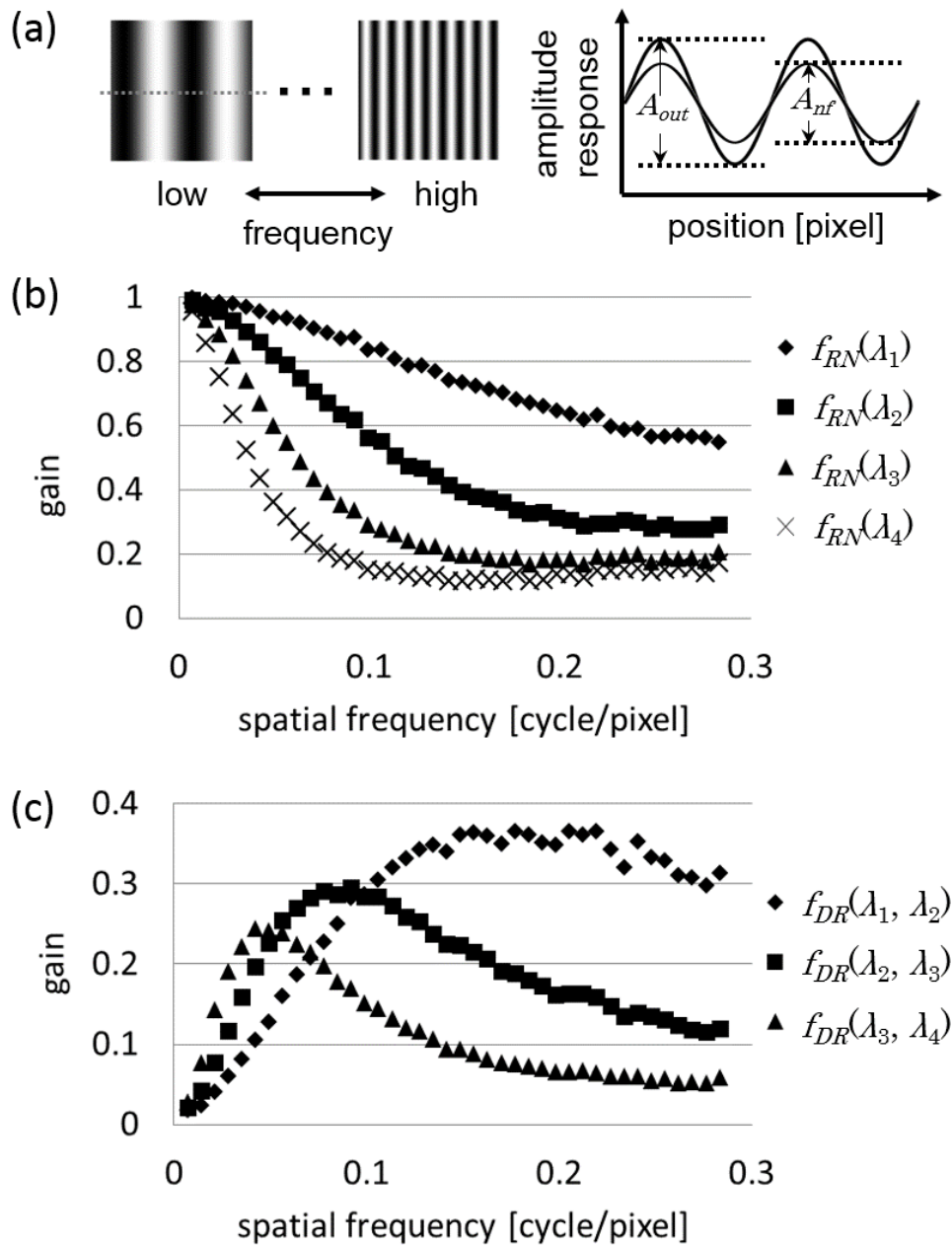


図 3.5 視覚センサの空間周波数応答. (a) 視覚刺激として用いられた正弦波グレーティングパターン (左). 正弦波グレーティングの灰色点線上の水平振幅応答 (右). A_{out} は RN フィルタ処理信号の振幅応答を, A_{in} は λ を最小に設定した場合 (λ_0) のフィルタ処理信号の正弦波の振幅応答をそれぞれ示す. (b) 低域通過特性を持つ RN フィルタ $f_{RN}(\lambda_1)$, $f_{RN}(\lambda_2)$, $f_{RN}(\lambda_3)$, $f_{RN}(\lambda_4)$ の空間周波数応答. バイアス電圧 (V_{rsbn} , V_{rsbp}) は図 3.4 に示した値と同じである. (c) 帯域通過特性を持つ DoR フィルタ $f_{DR}(\lambda_1, \lambda_2)$, $f_{DR}(\lambda_2, \lambda_3)$, $f_{DR}(\lambda_3, \lambda_4)$ の空間周波数応答. 帯域通過処理画像は二つの RN フィルタ処理画像の差をとることで得られた.

3.5 画像に対するセンサの出力応答の測定

最後に、LCD に画像を表示し、それを知能視覚センサに提示することで、センサの応答を評価する。視覚刺激として、図 3.6(a)に示す大きさの異なる複数の四角形を組み合わせた画像を用いた。図 3.6(b)に RN フィルタ $f_{RN}(\lambda_1)$, $f_{RN}(\lambda_2)$, $f_{RN}(\lambda_3)$, $f_{RN}(\lambda_4)$ によって、フィルタ処理された画像 $S_{RN}(\lambda_1)$, $S_{RN}(\lambda_2)$, $S_{RN}(\lambda_3)$, $S_{RN}(\lambda_4)$ を示す。図 3.6(c)に DoR フィルタ $f_{DR}(\lambda_1, \lambda_2)$, $f_{DR}(\lambda_2, \lambda_3)$, $f_{DR}(\lambda_3, \lambda_4)$ によってフィルタ処理された画像 $S_{DR}(\lambda_1)$, $S_{DR}(\lambda_2)$, $S_{DR}(\lambda_3)$ を示す。これらのフィルタの空間周波数特性は図 3.5 で示したものと同様である。図 3.6(d)に、図 3.6(c)の点線上の行の応答振幅の空間分布を示す。DoR フィルタ $f_{DR}(\lambda_1, \lambda_2)$ を用いて、フィルタ処理された画像の信号(図 3.6(d)内の黒色の実線)は他のフィルタ処理信号に比べ、より入力信号の細やかなテクスチャ構造を強調して表現していることがわかる。これは DoR フィルタ $f_{DR}(\lambda_1, \lambda_2)$ の通過帯域が適用した空間フィルタの間で最も高いためである。一方で、DoR フィルタ $f_{DR}(\lambda_3, \lambda_4)$ によってフィルタ処理された信号は低周波帯域を通過させるため、細やかなテクスチャ構造が抑制され、四角形のグループ全体の輪郭、すなわち大まかな構造を強調して表現していることがわかる(図 3.6(d)の灰色の実線)。これらの帯域通過処理画像は 20ms 内にフレーム落ちなく取得できた。以上の結果から、提案した知能視覚センサによって、異なる帯域通過特性を有する複数の空間フィルタ処理画像が実時間で取得できることが確認できた。

3.6 考察

3.6.1 提案するアーキテクチャの利点

我々は単一フレーム内に複数の空間フィルタを効率よく実行できる、新しい知能視覚センサのアーキテクチャを提案した。アーキテクチャに基づいて構築した試作機は複数回の空間フィルタ処理を実時間で実行できた。このアーキテクチャは高速で画像取得・フィルタ処理を実行するシリコン網膜とそれを制御するデジタル回路で構成されている。本アーキテクチャを採用した知能視覚センサは、抵抗回路網の空間特性を制御信号で動的に変更することで、コンパクトな構成でかつ、低消費電力で、実時間処理が実現できた。本知能視覚センサと、逐次型演算を実行する従来の画像センシングシステムとの主な違いは、センサのフロントエンドにある、全画素同時並列演算を実行するアナログ抵抗回路網と、次段にある画像情報の並列演算回路が実装できる FPGA で構成されている点である。本章では、FPGA は抵抗回路網の制御信号生成と差分演算の目的のみで利用したが、次章から詳細を述べるように本アーキテクチャは FPGA 内にミッションに応じた画像処理モジュール群を並列に配置することで、様々な画像特徴量を抽出できる拡張性を備えている。

提案したアーキテクチャに基づいて構築された試作機は USB バスパワー (2.5 ワット) で動作する。これは現在の一般的なデスクトップ・コンピュータが演算のために数十ワット

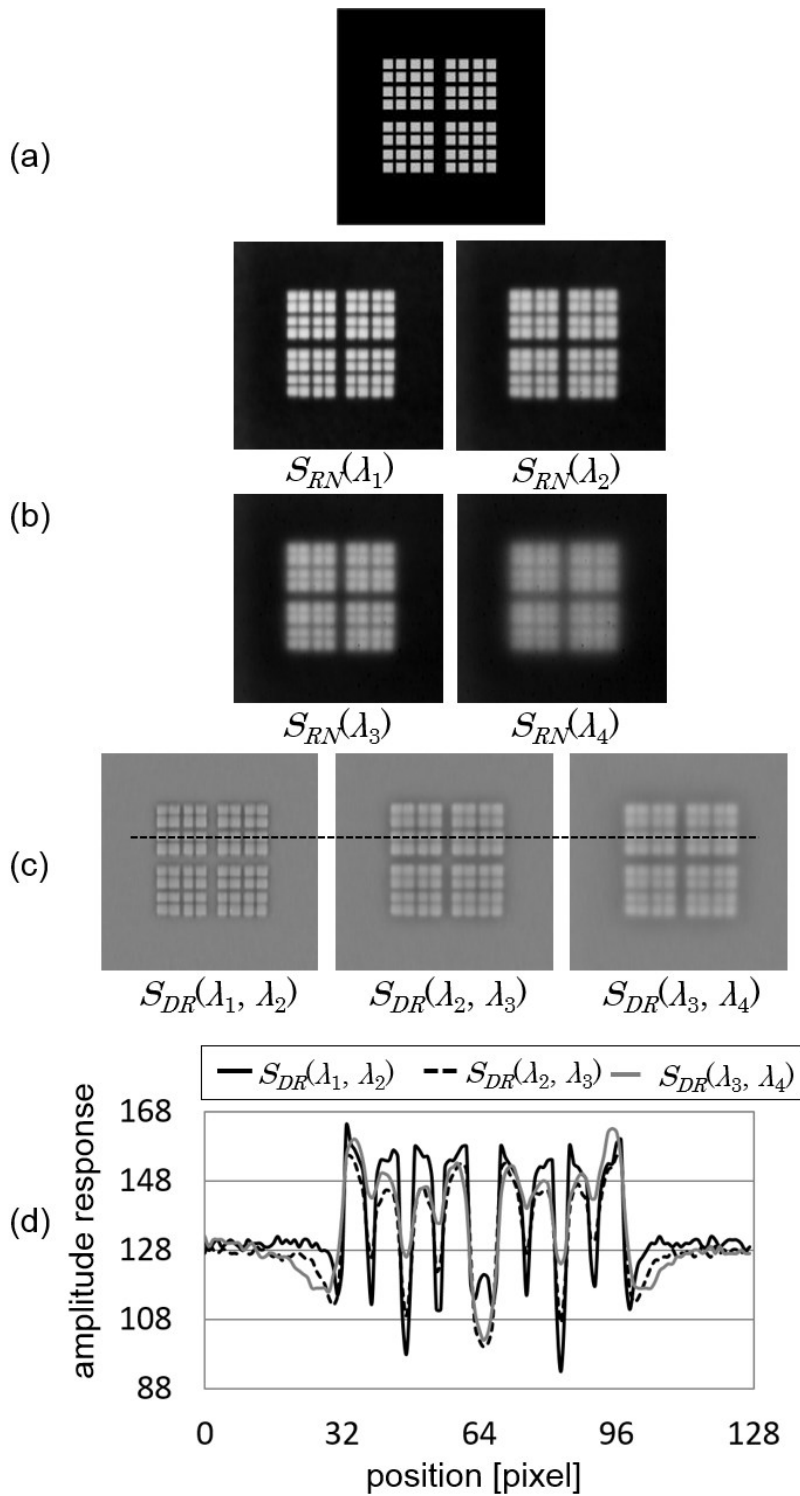


図 3.6 四角形のグループの画像への応答. (a) 入力画像. (b) RN フィルタ $f_{RN}(\lambda_1)$, $f_{RN}(\lambda_2)$, $f_{RN}(\lambda_3)$, $f_{RN}(\lambda_4)$ の出力画像. バイアス電圧は図 3.4 と同様である. (c) 帯域通過型空間フィルタ $f_{DR}(\lambda_1, \lambda_2)$, $f_{DR}(\lambda_2, \lambda_3)$, $f_{DR}(\lambda_3, \lambda_4)$ の出力画像. (d) カラム(c)の点線によって示された行の応答プロファイル.

かかることに比べて、非常に低消費電力である。

汎用デジタルコンピュータ上で複数回の空間フィルタ処理を実行すると、それにかかる計算時間は画像サイズや空間フィルタのサイズに、式(2.2)に示すように非常に強く依存するのに対して、本アーキテクチャを採用したセンサではほぼ影響しない。これは汎用デジタルコンピュータ上では、空間フィルタ処理のための積和演算を一画素一画素、逐次的に実行しているためである。一方で、抵抗回路網による空間フィルタ処理の計算コストは画像サイズにも空間フィルタのサイズにもほぼ依存せず、空間フィルタ処理は μs オーダーで実行できる。

3.6.2 単一フレーム内に適用できる空間フィルタの数について

本研究で提案した戦略の場合、空間フィルタ処理を複数回適用するために、単一フレーム内に適用できる帯域通過フィルタの最大の数は抵抗回路網の処理時間、画像の読み出し時間、蓄積時間の三つの要素で決定される。本研究で構築した試作機では画素数が 128×128 画素であり、その一画素を読み出すために約 200ns かかる。従って、試作したセンサでは全画像データを逐次的に読み出すため、画像の読み出しに約 3.3ms ほど必要である。抵抗回路網による処理と空間特性の変更に必要な時間は双方ともアナログ回路の整定時間に達するまでに要する時間となる。これらの期間は本研究で用いられた抵抗回路網では μs オーダーである。抵抗回路網による空間フィルタ処理に要する時間は画像の読み出しに比べて非常に短いため、蓄積時間内に帯域通過型空間フィルタの適用できる数は、読み出し時間によって決定される。従って、単一フレーム内に空間フィルタを適用できる数を増やすことを考えた場合、読み出し速度が現段階におけるボトルネックとなる。この問題の解決方法の一つとして列並列型の A/D 変換器の採用が挙げられる[79]。列並列型 A/D 変換は高速度での画像読み出しを達成できる。この技術を採用することで、提案したセンサは非常に高速に帯域通過型空間フィルタを適用でき、数百のフィルタ処理画像を実時間で得る見込みがある。

3.6.3 MOS 抵抗の入力電圧レベルの依存性

提案した知能視覚センサで採用した抵抗回路網内の MOS 抵抗は、2.7.3 節で述べたように、入力電圧レベルの依存性を抑えるために CMOS 抵抗の構成をとっている。しかしながら、CMOS 抵抗の構成をとっていてもなお、抵抗回路網の空間特性は入力電圧レベル、すなわち入力光の強度に依存する。本節では、その入力電圧レベルに対する依存性について実験し、結果を考察したことをまとめる。入力電圧レベル依存性は、屋外から屋内にカメラが移動する場合など、視覚環境が時々刻々と変化する応用において問題となる。この問題は、シリコン網膜がレベルシフト機能を有するために、抵抗回路網への入力信号のレベルを適応的にシフトすることによって解決できる見込みがある。以下にその詳細を述べる。

MOS 抵抗回路網の入力光依存性を検証した。まず 3.3 節で述べたように、MOS 抵抗回路網の線拡がり関数を測定した。ここでは小さなスケールパラメータ ΔL と大きなスケール

パラメータ λ_H に設定した二つの抵抗回路網フィルタ $f_{RN}(\lambda_L)$ と $f_{RN}(\lambda_H)$ を用いた。フィルタ特性を $f_{RN}(\lambda_L)$ と $f_{RN}(\lambda_H)$ に設定するために用いたバイアス電圧の組み合わせはそれぞれ、(0.36V, 2.69V) と (0.45V, 2.40V) とした。LCD からの光刺激はスリット光の強度を一定とし、背景光を明るくした場合 (平均画素値 : 約 168) と暗くした場合 (平均画素値 : 約 68) の二パターンを用意した。図 3.7 に実験結果を示す。 $f_{RN}(\lambda_H)$ に比べ $f_{RN}(\lambda_L)$ に顕著であるが、双方とも背景光が明るい場合の方が暗い場合に比べて応答が側方向に広がる傾向があることがわかる。ここで、3.3 節に比べ、どちらの線拡がり関数の振幅応答もノイズの影響が強い理由は水平方向の振幅応答の平均をとっていないためである。

次に $f_{RN}(\lambda_L)$ と $f_{RN}(\lambda_H)$ が入力光のレベルが変化した場合、空間周波数特性がどのように変化するか検討した。測定は 3.3 節と同様の手法で行った。正弦波の振幅は一定 (平均画素値 : 約 128) とし、背景光を基準の明るさに対して、明るい場合 (平均画素値 : 約 158) と暗い場合 (平均画素値 : 約 98) の結果を比較した。 $f_{RN}(\lambda_L)$ と $f_{RN}(\lambda_H)$ の結果をそれぞれ図 3.8 と図 3.9 に示す。この結果から抵抗比に関わらず入力光が明るくなるとカットオフ周波数が低くなることわかる。

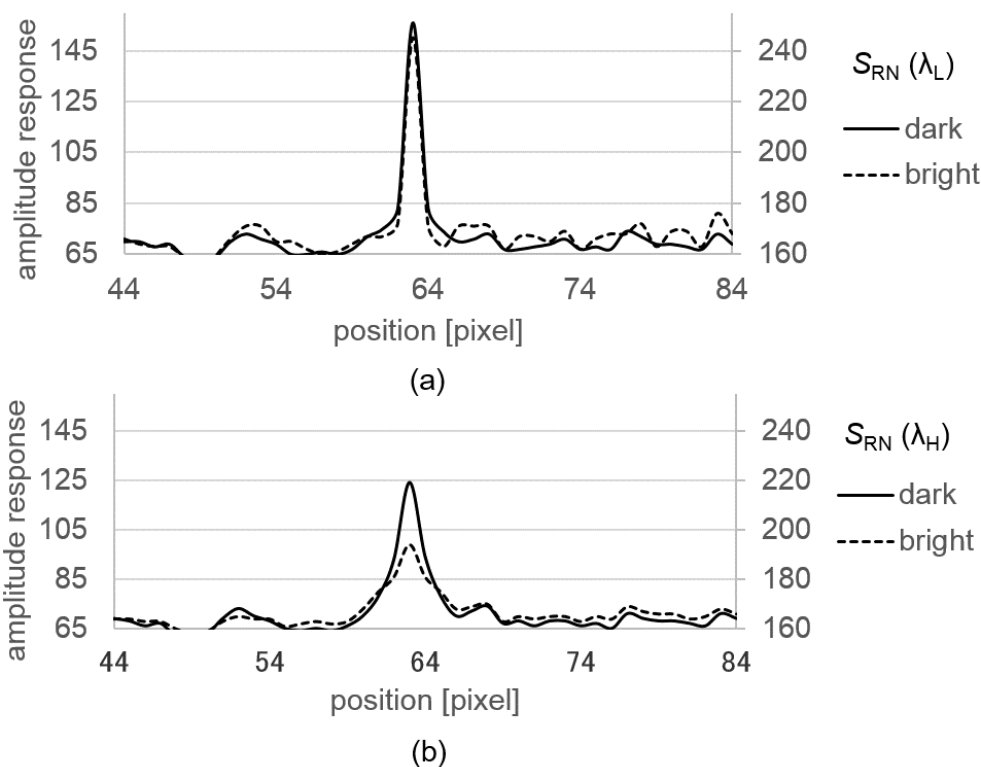


図 3.7 背景光を変化させた時の線拡がり関数. (a) $S_{RN}(\lambda_L)$ の場合. (b) $S_{RN}(\lambda_H)$ の場合. 左は背景光が暗い場合の応答のための目盛であり、右の目盛は明るい場合の応答のための目盛である。

次に入力光の振幅を変化させて同様の実験を行った。背景光は一定として、正弦波信号を基準の振幅に対して、大きい時と小さい時の結果を比較した。背景光の時と同様に抵抗比 λ が低い場合の空間フィルタ $f_{RM}(\lambda_L)$ と、高い場合の空間フィルタ $f_{RM}(\lambda_H)$ において調べた。図 3.10, 図 3.11 にそれぞれの結果を示す。この結果から振幅依存性は図 3.8, 図 3.9 に示した入力光レベルの依存性に比べ、比較的低いということがわかる。

本センサが刻一刻と環境光のレベルが変化する場面で用いられる場合、この入力光レベルの依存性は改善しなければならない問題である。抵抗回路網への入力光レベルが影響する理由は 2.7 節で述べた通りである。入力光レベルが低い場合は図 2.18(c)内の V_b が中間辺りの動作域で MOS 抵抗が動作しており、各 R_s の抵抗値が低くなり、スケールパラメータが小さくなる。一方で、入力光レベルが高い場合は R_s の抵抗値が先ほどより小さくなる領域で MOS 抵抗が動作しており、その結果スケールパラメータが比較的大きくなる。この説明は実験結果と一致する。さて本研究で用いたシリコン網膜では抵抗回路網への画素入力 は式 (2.19) で表される。問題は背景が明るいと、各画素での信号値 $v_{po} - v_{pr}$ が高くなり、各画素入力 v_{in} の値が高くなるということになる。従って、この問題は $v_{po} - v_{pr}$ の平均的な値に応じて、その DC 成分を打ち消すように S/H1 のリファレンス電圧 V_{ref} の値を制御し、 v_{in} を一定の範囲内に収めることによって緩和されると考えられる。

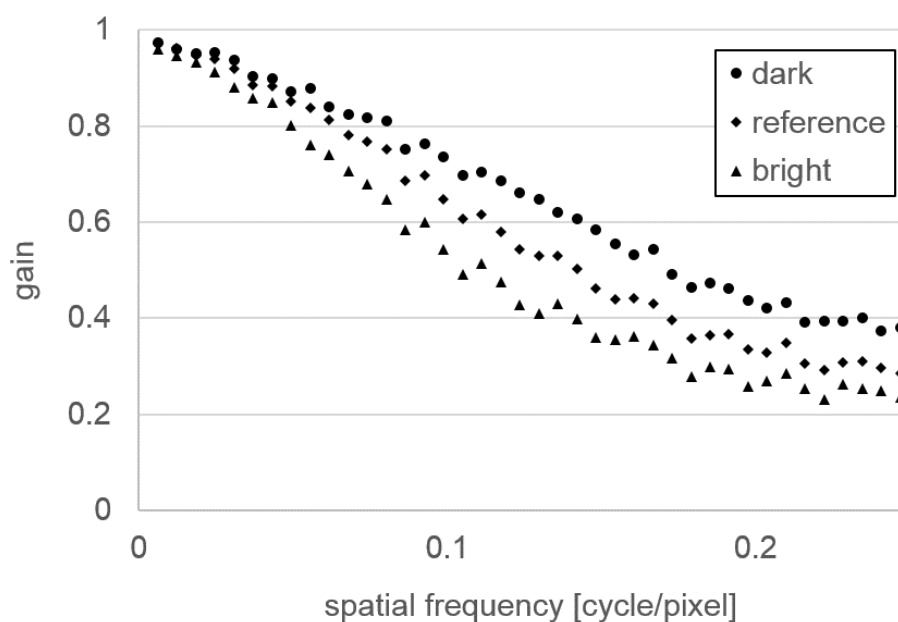


図 3.8 入力光レベルを変化させた場合の $f_{RM}(\lambda_D)$ の空間周波数特性への影響。暗く入力電圧レベルが低い場合に比べ、明るく入力電圧レベルが高い場合の方が抵抗回路網のカットオフ周波数が低くなる。

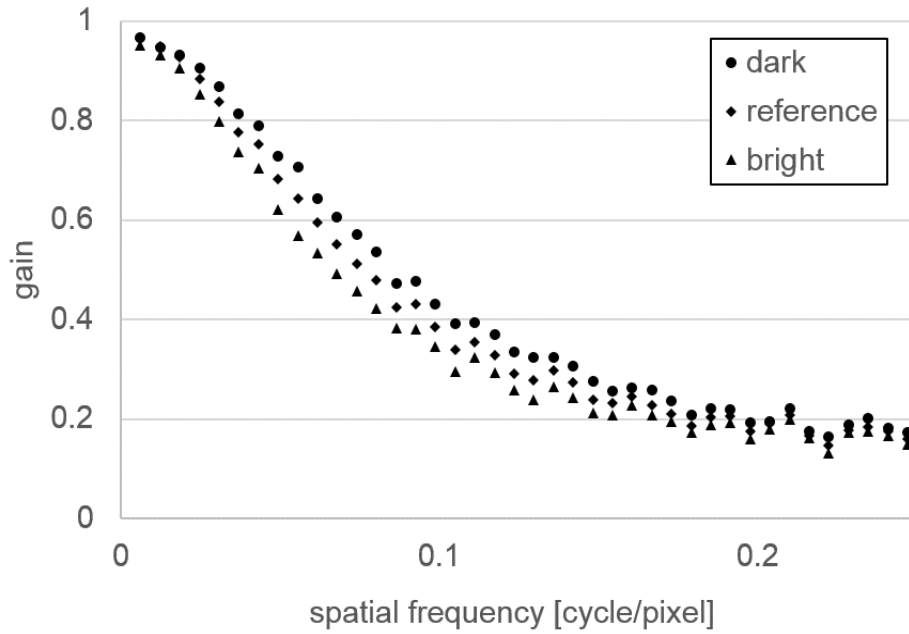


図 3.9 入力光レベルを変化させた場合の $f_{RN}(\lambda_H)$ の空間周波数特性への影響. $f_{RN}(\lambda_L)$ の場合と同様に, より明るく入力電圧レベルが高い場合の方が, 抵抗回路網のカットオフ周波数が低くなる.

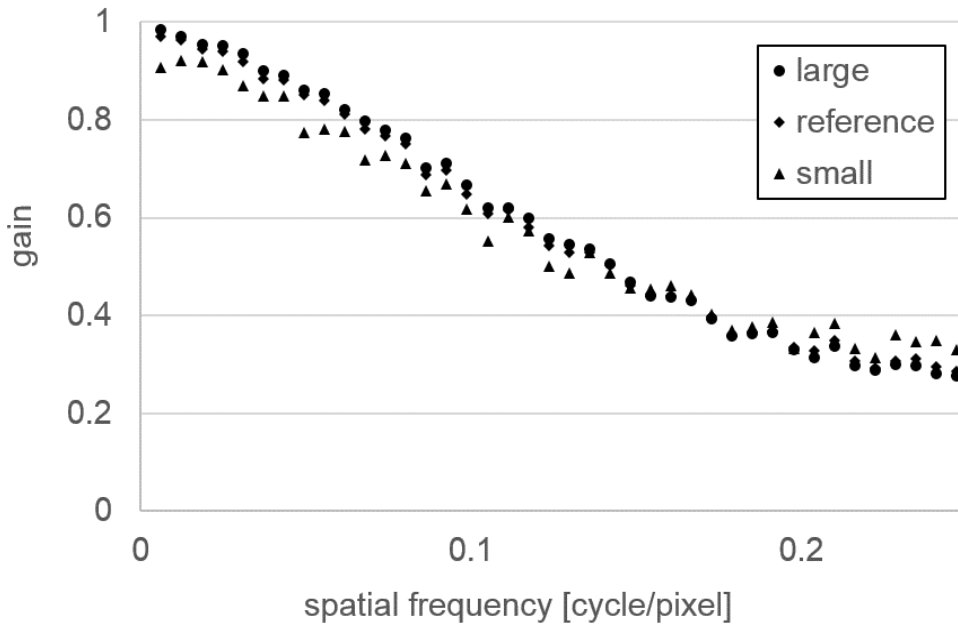


図 3.10 正弦波パターンの振幅強度を変化させた場合の $f_{RN}(\lambda_L)$ の空間周波数特性への影響. 振幅強度が大きい場合と小さい場合で, 抵抗回路網の空間周波数特性はほぼ変化がない.

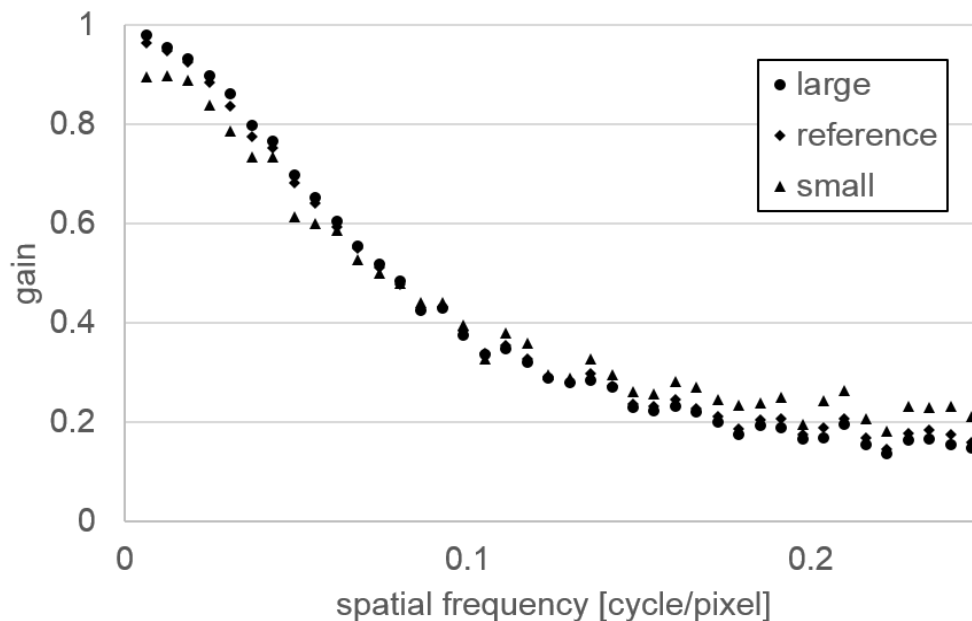


図 3.11 正弦波パターンの振幅強度を変化させた場合の $f_{RM}(\lambda_H)$ の空間周波数特性への影響. $f_{RM}(\lambda_L)$ の場合と同様に, 振幅強度が大きい場合と小さい場合で, 抵抗回路網の空間周波数特性はほぼ変化がない.

3.7 結言

本章では, 抵抗回路網を活用することで, 実時間で複数の空間フィルタ処理を効率よく適用できる知能視覚センサについて述べた. 我々の知能視覚センサは, コンパクトでかつ低電力で実時間演算を実現するために, アナログ/デジタル混合型のアーキテクチャ構成を採用した. 知能視覚センサは APS, S/H 回路, アナログ抵抗回路網, そして FPGA で構成される. 我々は, 抵抗回路網による低域通過型空間フィルタ及び知能視覚センサによる帯域通過型空間フィルタの空間特性と空間周波数特性を測定し, それらが外部制御信号によって制御できることを確認した. 次に, テスト画像をセンサに提示することで, 画像中にある複数の異なるスケールの構造が強調されている出力画像群をセンサが抽出できていることを示した.

この研究は低消費電力でかつ実時間で帯域分離処理を実行するシステムの実現に向けた第一歩である. 続く四章から六章までは帯域分離処理の例として実時間マルチスケールエッジ抽出と, スケール不変特徴量抽出の機能を有する知能視覚システムを構築したことについて述べる.

第4章 実時間マルチスケールエッジ検出機能を有する知能視覚センサの開発

4.1 緒言

前章ではマルチスケール空間フィルタ処理機能を有する知能視覚センサの開発について述べた。マルチスケール画像は、様々な特徴抽出の前段の処理として多く利用されている[13-21]。しかしながら、取得された出力画像群は非常にデータ量が多く、従来の逐次処理型デジタル画像処理システムで画像を処理する場合、多くの計算時間を要する。そこで、本章では前章で論じた知能視覚センサの後段に並列デジタル演算回路を配し、センサからの画素読み出しと同時に画像を処理することで、フレーム落ちを生じず、後段の処理を実行できるアーキテクチャを構築した[78]。まず特徴抽出機能として、画像処理の様々なタスクで利用されているエッジ検出を対象とした。本章では、提案したアーキテクチャに基づき、実時間マルチスケールエッジ検出機能を有する知能視覚センサのアルゴリズムを開発し、その出力を評価した。

4.2 ゼロ交差法によるエッジ検出

エッジ検出は画像理解において基本となる処理の一つであり、物体認識や領域分割などのタスクの前に実行される。エッジ検出のための基本的な戦略として、Marrらのゼロ交差法を採用した[2, 11]。この手法では、光学的な特性から、エッジがある位置では画素値の強度が空間的に急激に変化していると仮定し、その応答変化の位置を検出することでエッジを検出する。MarrとHildrithは強度変化を効率よく検出するために、ラプラシアンガウシアンフィルタを演算子として採用した[80]。応答強度の急激な変化は二次の導関数においてはゼロ交差を引き起こす。ラプラシアン演算子は二次の空間微分演算子として採用されており、ガウシアンフィルタはそのスケールパラメータによって、対象となる構造の大きさに同調させるために採用されている。例えば大きなスケールのフィルタを採用すれば、大まかな構造やぼやけた影のエッジを検出でき、小さなスケールのフィルタを採用すれば、画像中の鋭く焦点が合った微細な部分を検出できる[2]。ラプラシアンガウシアンフィルタがゼロ交差法で用いることのできる唯一の演算子ではなく、同じく帯域通過特性を有するDoGフィルタやDoRフィルタもまた演算子として用いることができる。

図 4.1 にゼロ交差法によるエッジ検出の概念図を示す。図 4.1(a)に示す明暗パターンの点

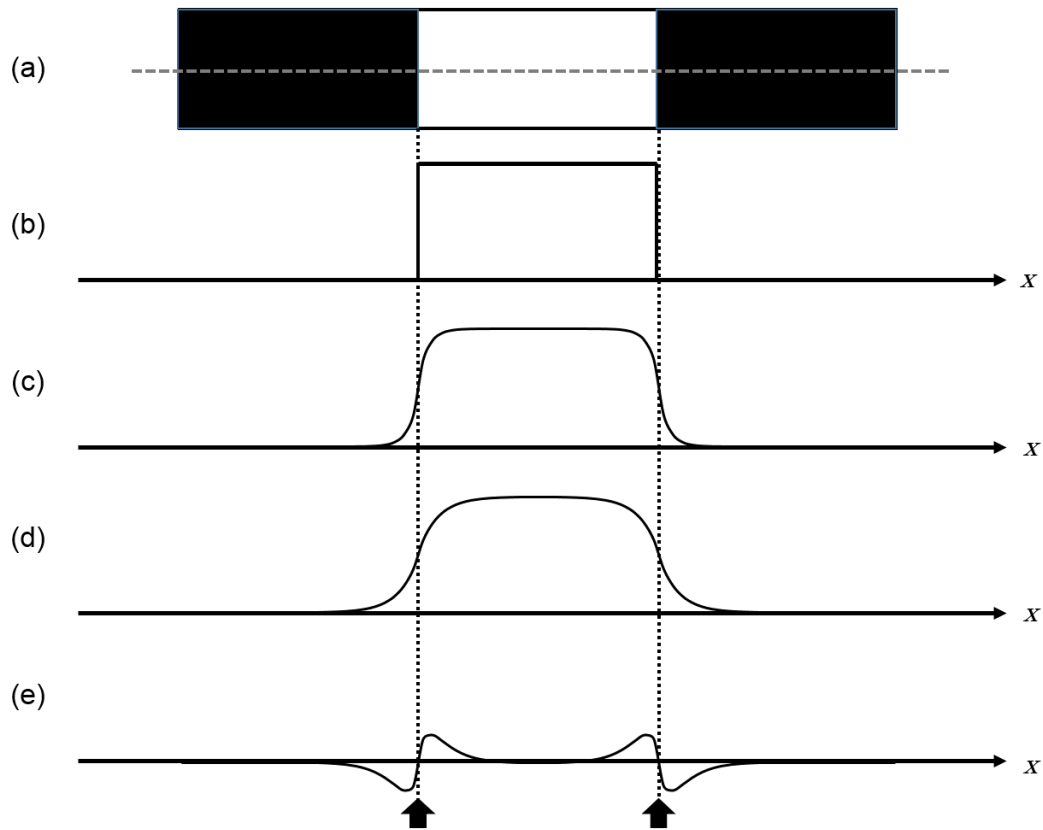


図 4.1 ゼロ交差法によるエッジ検出. (a) 明暗パターン. (b) (a)の点線上のラインプロファイル. (c) 小さなスケールの RN フィルタによる処理結果. (d) (c)に比べて大きなスケールの RN フィルタによる処理結果. (e) 二つの RN フィルタ処理信号の差を取ることによって得られた DoR フィルタ処理信号.

線上のラインプロファイル (図 4.1(b)) を入力信号とし, DoR フィルタをゼロ交差法に用いる演算子として用いた場合のエッジ検出の処理過程について述べる. 図 4.1(c)(d)は入力信号に RN フィルタを適用して得られたフィルタ処理信号である. 図 4.1(c)は大きなスケールの RN フィルタを, 図 4.1(d)は大きなスケールの RN フィルタを適用した結果である. 図 4.1(c)に比べて図 4.1(d)はより信号が側方向に広がっていることがわかる. 図 4.1(e)はそれらの差を取ることで得られた DoR フィルタ処理信号である. 図中矢印で示すように入力信号で応答強度の変化が強い付近で強く応答し, そのエッジ部分でゼロレベルを横切っていることがわかる.

4.3 ゼロ交差検出回路

本節では、4.2節で述べたゼロ交差法を実行するデジタル回路について述べる。FPGA上にゼロ交差検出回路を配することで、知能視覚センサの応用として実時間マルチスケールエッジ検出機能の実装を試みた。ゼロ交差検出回路は井上ら[81]の研究を参考に開発した。

一次元ゼロ交差検出回路の構成を図4.2に示す。図4.2の例は画像内の x 方向のゼロ交差を検出する回路である。本回路は主にレジスタ(図中REG)、比較器(図中COMP)、および論理ゲート(図中XOR, AND)によって構成される。まずDoRフィルタ処理画像の画素 $S_{DR}(x-1, y)$ がレジスタに保存される。画像内の横方向に隣り合う画素がゼロレベルを交差しているかを比較器COMP1、COMP2とXOR回路により、次式の演算を実行することで検出する。ここで、 S_{base} はゼロレベルを表す。

$$(S_{DR}(x, y) > S_{base}) \text{ xor } (S_{DR}(x-1, y) > S_{base}). \quad (4.1)$$

次に、エッジ由来のゼロ交差点とノイズによるゼロ交差点を分離するために、隣り合う画素間で大きく応答が変化している位置を検出する勾配検出器(図中点線内の回路)を用いた。この勾配検出器の出力を以下に示す。ここで、 S_{th} は応答変化の閾値を表す。

$$|S_{DR}(x-1, y) - S_{DR}(x, y)| > S_{th}. \quad (4.2)$$

最後にXOR回路と勾配検出器の出力のANDを取ることで、ゼロ交差点を検出する。

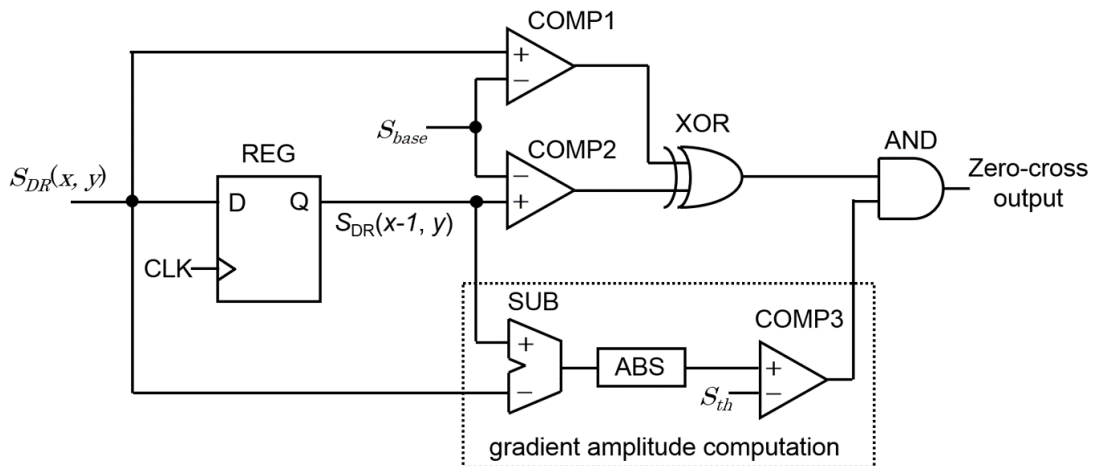


図4.2 一次元ゼロ交差検出回路。上部のゼロ交差検出器と下部の勾配検出器で構成される。勾配検出器の出力は信号とノイズとを区別するために用いられる。

提案した知能視覚センサには上記で述べた回路を二次元に拡張した回路を実装した。画像の x 方向だけでなく、 y 方向の回路を実装し、それらの OR を取った出力をエッジ検出回路の出力とした。

4.4 FPGA 上での差分画像とエッジ画像の並列処理

図 4.3 に示すタイミングチャートを用いて、実時間マルチスケールエッジ検出の流れを説明する。第 3 章で述べたように、4 つの異なるスケールの RN フィルタ処理画像 $S_{RN}(\mathbf{X}_t, \lambda_1)$, $S_{RN}(\mathbf{X}_t, \lambda_2)$, $S_{RN}(\mathbf{X}_t, \lambda_3)$, $S_{RN}(\mathbf{X}_t, \lambda_4)$ が A/D 変換器を介して FPGA 内のメモリに保存される。ここで、 $\mathbf{X}_t=(x, y, t)$ は画像の位置と時刻を表現するベクトルである。この RN フィルタ処理画像の読み出しと同時に並行的に DoR フィルタ処理画像 $S_{DR}(\mathbf{X}, \lambda_1, \lambda_2)$, $S_{DR}(\mathbf{X}, \lambda_2, \lambda_3)$, $S_{DR}(\mathbf{X}, \lambda_3, \lambda_4)$ とエッジ画像 $S_E(\mathbf{X}, \lambda_1, \lambda_2)$, $S_E(\mathbf{X}, \lambda_2, \lambda_3)$, $S_E(\mathbf{X}, \lambda_3, \lambda_4)$ が生成される。ここで、エッジ要素がある画素位置に 1 の値をもつ画像をエッジ画像とする。

図 4.3 内の時刻 t_1 における FPGA 上での処理を考える。 t_1 において、すでに RN フィルタ処理画像 $S_{RN}(\mathbf{X}_t, \lambda_1)$ がメモリに保存されている。 t_1 から t_2 までの区間で $S_{RN}(\mathbf{X}_t, \lambda_2)$ が一画素ずつ読み出される。すでに 3.2 節で説明をしたように $S_{RN}(\mathbf{X}_t, \lambda_2)$ の画素読み出しと同時にメモリから読み出された $S_{RN}(\mathbf{X}_t, \lambda_1)$ が一画素ずつ順に読み出され、共に差分器に入力されることで差分画像、すなわち DoR フィルタ処理画像 $S_{DR}(\mathbf{X}_t, \lambda_1, \lambda_2)$ が生成される。 $S_{DR}(\mathbf{X}_t, \lambda_1, \lambda_2)$ はメモリに保存される一方で、前節で述べたゼロ交差検出回路に入力される。このようにしてエッジ画像 $S_E(\mathbf{X}_t, \lambda_1, \lambda_2)$ の画素出力も FPGA 内で同時並行的に演算され、エッジ画像用メモリに順次に保存されていく。この処理を各 RN フィルタ処理画像の読み出し時に実行することで、三つの異なるスケールのエッジ画像を取得できる。

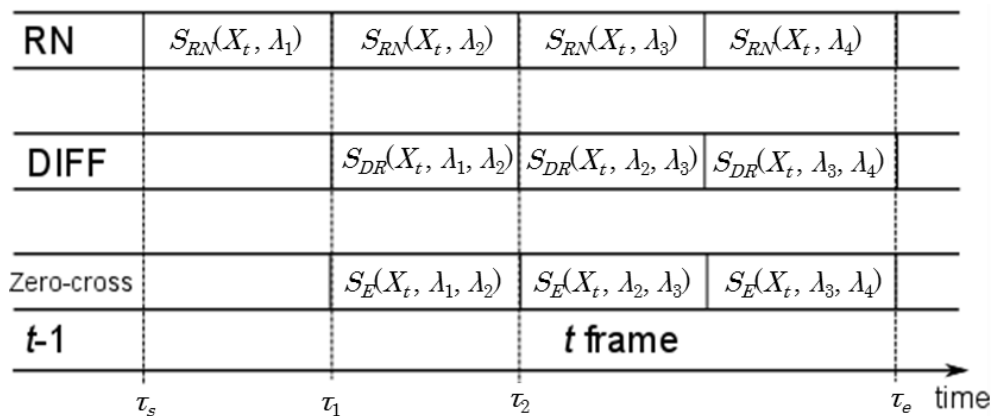


図 4.3 FPGA 上でのエッジ画像の演算についてのタイミングチャート。本知能視覚センサでは、FPGA 上で RN フィルタ処理画像 S_{RN} (一段目)と DoR フィルタ処理画像 S_{DR} (二段目)、そしてエッジ画像(三段目)を同時並行的に取得する。

4.5 システムの出力応答の測定

4.5.1 単純な視覚対象に対する応答

我々は図 3.6 と同様に図 4.4(a)に示す四角形のグループの画像を智能視覚センサに提示し、その出力画像を検討した。四角形の間幅は図 4.4(b)に示されるようにそれぞれ異なるように配置している。図 4.4(c)-(e)に t フレーム目の三つの DoR フィルタ処理画像 $S_{DR}(\mathbf{X}_t, \lambda_1, \lambda_2)$, $S_{DR}(\mathbf{X}_t, \lambda_2, \lambda_3)$, $S_{DR}(\mathbf{X}_t, \lambda_3, \lambda_4)$ を示す。本実験で用いた空間フィルタ群は図 3.5(c)に示された空間周波数特性を持つ。右側のグレースケールの階調を示すバーが帯域通過信号の応答振幅を表現する。DoR フィルタによって帯域通過処理された三つの画像の全てにおいて、四角形のない周辺領域の応答はゼロレベルを示す均一な灰色で表現されている。一方、四角形のグループの周辺と内部は正もしくは負の値をとる。図 4.4(f)は図 4.4(c)-(e)の黒色の点線によって指定された線上の応答を示している。

最も通過帯域が高い、 $f_{DR}(\mathbf{X}_t, \lambda_1, \lambda_2)$ によってフィルタ処理された信号(図 4.4(f)中の黒色の実線)は小さい四角形のそれぞれのエッジで鋭く変化している(x_1, x_2, x_3, x_4)。一方で最も通過帯域が低い、 $f_{DR}(\mathbf{X}_t, \lambda_3, \lambda_4)$ によってフィルタ処理された信号(図 4.4(f)中の灰色の実線)は四角形間の隙間の距離が狭い点(x_1, x_2, x_3)で大きい変化を示さず、 x_4 でのみ大きく変化を示した。

図 4.4(g)-(i)に画像(c)-(e)からそれぞれ抽出されたエッジ画像 $S_E(\mathbf{X}_t, \lambda_1, \lambda_2)$, $S_E(\mathbf{X}_t, \lambda_2, \lambda_3)$, $S_E(\mathbf{X}_t, \lambda_3, \lambda_4)$ を示す。これらの図の白色の点群は信号がゼロに交差する、ゼロ交差点を表現する。図 4.4(g)に示すように、 $S_E(\mathbf{X}_t, \lambda_1, \lambda_2)$ ではゼロ交差点が小さな四角形のそれぞれのエッジ(x_1, x_2, x_3, x_4)で抽出されていることがわかる。一方で、 $S_{DR}(\mathbf{X}_t, \lambda_3, \lambda_4)$ の場合は、四角形間の距離が広い x_4 の位置でのみエッジが現れる。従って、図 4.4(g)の白色の点群が小さな四角形の輪郭を表現する。一方で、図 4.4(i)のそれは四角形の四つのグループのエッジが表現する。

4.5.2 複雑な視覚対象に対する応答

複雑な視覚対象として二つの自然画像を智能視覚センサに提示し、出力画像群とその時系列データを検討した。

まず智能視覚センサによって図 4.5(a)に示す植物の写真に対する出力画像群を取得した。図 4.5(b)-(d)に t フレーム目の出力画像群である三つの DoR フィルタ処理画像 $S_{DR}(\mathbf{X}_t, \lambda_1, \lambda_2)$, $S_{DR}(\mathbf{X}_t, \lambda_2, \lambda_3)$, $S_{DR}(\mathbf{X}_t, \lambda_3, \lambda_4)$ を示す。 $S_{DR}(\mathbf{X}_t, \lambda_1, \lambda_2)$ では葉の形状のような細かいテクスチャパターンを表現されている(図 4.5(b))。一方で、 $S_{DR}(\mathbf{X}_t, \lambda_3, \lambda_4)$ ではこれらのパターンが平滑化されて除去されている(図 4.5(d))。図 4.5(e)-(g)は図 4.5(b)-(d)からそれぞれ抽出されたエッジ画像を示している。これらの図の白色の点群はエッジを表現する。図 4.5(e)において、葉の形や背景の壁の水平の線が表現されている。他方、植物の外郭や葉の固まりと

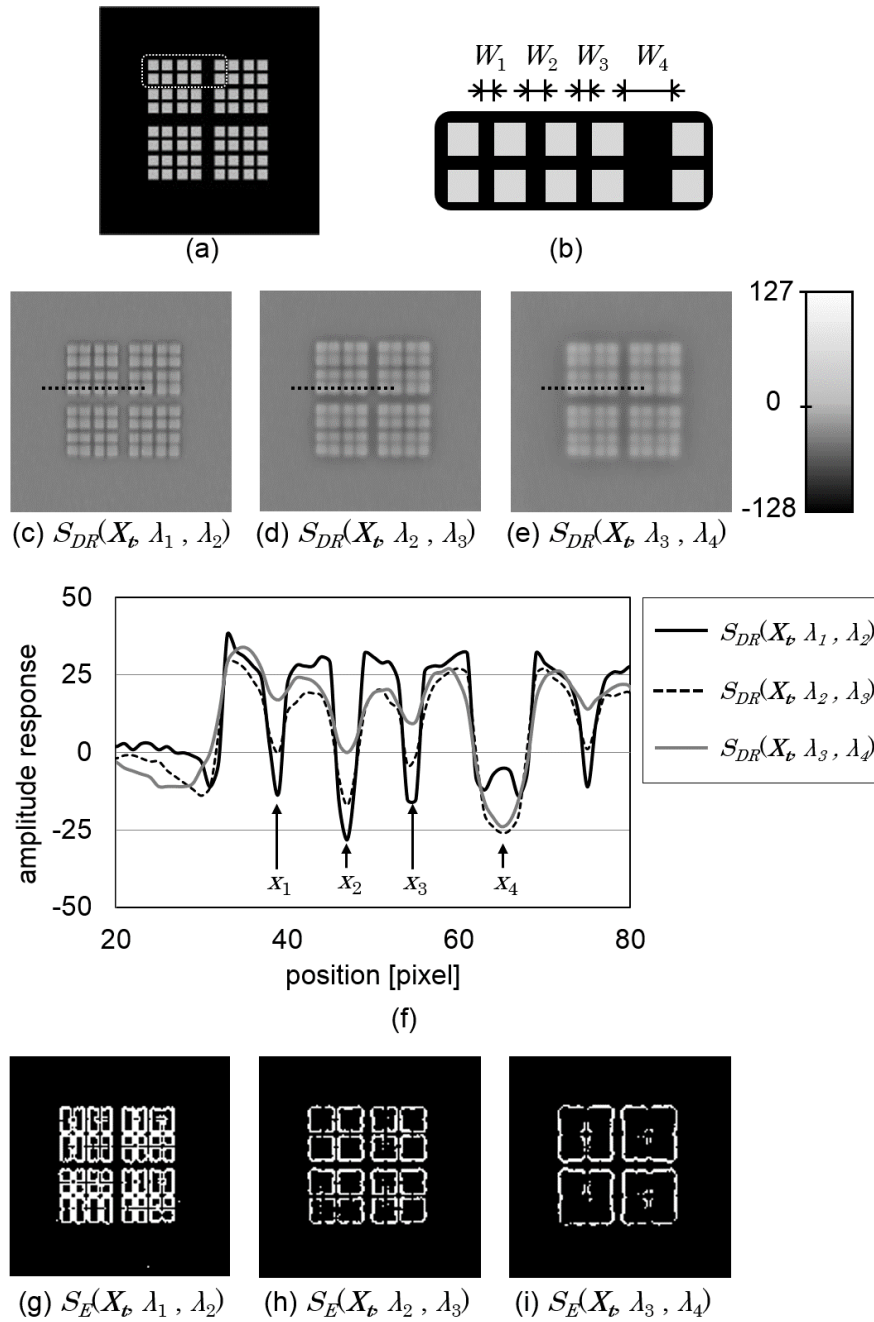


図 4.4 四角形のグループの画像に対する応答. (a) 入力画像. (b) 図 4.3(a)の白色の点線によって囲まれた領域の拡大図. 四角形の間距離は次の関係を持つ: $W_1 = W_3 < W_2 < W_4$. (c) (d) (e) それぞれ $S_{DR}(\mathbf{X}_b, \lambda_1, \lambda_2)$, $S_{DR}(\mathbf{X}_b, \lambda_2, \lambda_3)$, $S_{DR}(\mathbf{X}_b, \lambda_3, \lambda_4)$ を示す. 右側のグレースケールの階調を持つバーは帯域通過処理信号の応答のレベルを示す. (f) 図 4.3(c)-(e)の黒色の点線上の応答. (g) (h) (i) 図 4.3(c)-(e)から抽出されたエッジ画像 $S_E(\mathbf{X}_b, \lambda_1, \lambda_2)$, $S_E(\mathbf{X}_b, \lambda_2, \lambda_3)$, $S_E(\mathbf{X}_b, \lambda_3, \lambda_4)$. これらの図の白色の点群はゼロ交差点を表現している.

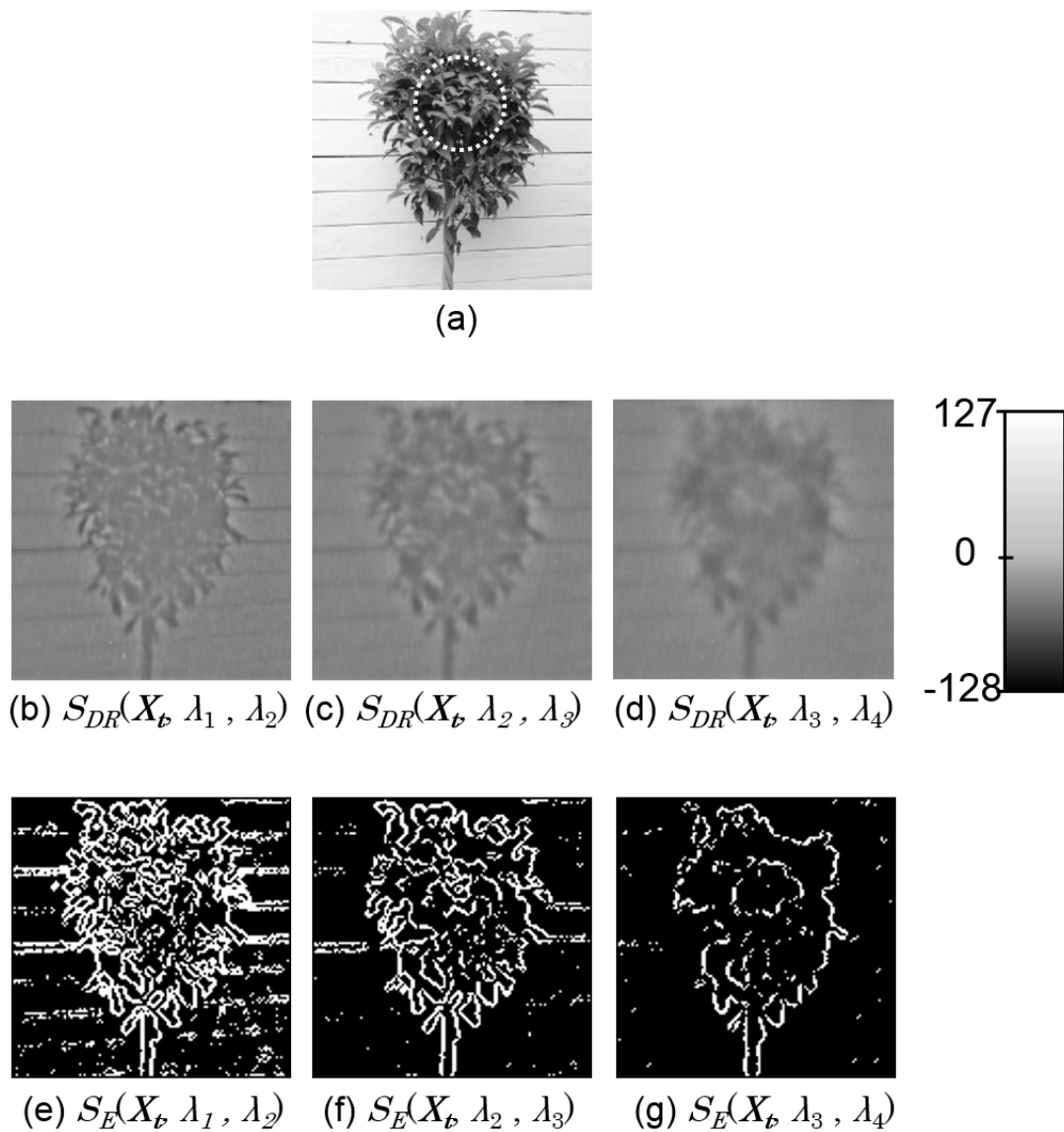


図 4.5 自然画像（植物の写真）に対する応答. (a) 入力画像. (b) (c) (d) DoR フィルタ処理 $S_{DR}(\mathbf{X}, \lambda_1, \lambda_2)$, $S_{DR}(\mathbf{X}, \lambda_2, \lambda_3)$, $S_{DR}(\mathbf{X}, \lambda_3, \lambda_4)$. (e) (f) (g) 画像(b)(c)(d)からそれぞれ抽出されたエッジ画像 $S_E(\mathbf{X}, \lambda_1, \lambda_2)$, $S_E(\mathbf{X}, \lambda_2, \lambda_3)$, $S_E(\mathbf{X}, \lambda_3, \lambda_4)$.

いった粗い構造が、図 4.5(a)で点線の円によって指定された図 4.5(g)が表現された。

次いで我々はシマウマが歩いているシーンの動画を提示し、知能視覚センサの出力画像の時系列データを検証した。図 4.6 は出力画像の四つの連続するフレームを示す。図 4.6(b)(c) に帯域通過処理画像 $S_{DR}(\mathbf{X}, \lambda_1, \lambda_2)$ と $S_{DR}(\mathbf{X}, \lambda_3, \lambda_4)$ から抽出されたエッジ画像 $S_E(\mathbf{X}, \lambda_1, \lambda_2)$ と $S_E(\mathbf{X}, \lambda_3, \lambda_4)$ を示す。図 4.6(b)において、細かいテクスチャパターンの輪郭とシマウマや他の物体の外郭の両方が表現されている。一方で、図 4.6(c)において、シマウマの縞のよう

な細かいテクスチャパターンが除去され、シマウマと他の物体の外郭線などの大まかな構造のみが表現されている。図 4.6(b)(c)の画像は全フレームで視覚センサが複数の大きさで輪郭を抽出できていることを示している。例えば、シマウマの前足に着目すると、その動きが時系列データとして適切に表現されている。これらの結果から本研究で提案された知能視覚センサの有用性が示された。

4.6 考察

我々は実時間マルチスケールエッジ検出機能を有する知能視覚センサを開発した。本知能視覚センサはマルチスケール空間フィルタ処理機能を有する視覚センサと後段の FPGA 上のゼロ交差検出回路とを組み合わせることで、効率的に粗い構造から細かい構造までの輪郭情報を抽出できた。高周波数成分から構成された帯域通過型空間フィルタ処理画像から抽出された輪郭抽出画像は、物体のテクスチャや輪郭のパターンを両方表現しているため、主に物体認識において利用できると考えられる。他方、低周波成分から構成された帯域通過型空間フィルタ処理画像から抽出された輪郭抽出画像は、細かい構造は除去され、物体の輪郭が主に表現されているために、物体の領域分離に利用できると考えられる。

シリコン網膜から FPGA 上のメモリへの画素読み出しには一定の時間を要する。一般的に、FPGA 上において、マルチスケール解析をする際も同様に逐次的に画像が保存されているメモリにアクセスし、画素情報を読み出し、処理を実行後、メモリ上の異なるアドレスに画素情報を書き込む。3.7.1 節で議論したように、この読み出し時間は本視覚センサのフレームレートのボトルネックとなるが、画素の読み出しとメモリへの書き込み時間は必ず必要な工程であるため、時間短縮には限界がある。本知能視覚センサのアーキテクチャでは FPGA 上の読み出し信号線の途中に画像処理回路を配することで、計算時間増大の原因となるメモリへのアクセスの回数を減らしている。本知能視覚センサは、すでにフィルタ処理済みの画像を読み出す上、画素を読み出し時に、FPGA 上の画像処理回路でパイプライン処理を実行するため、処理のための回路面積・効率の面で非常に優れた構成であると考えられる。

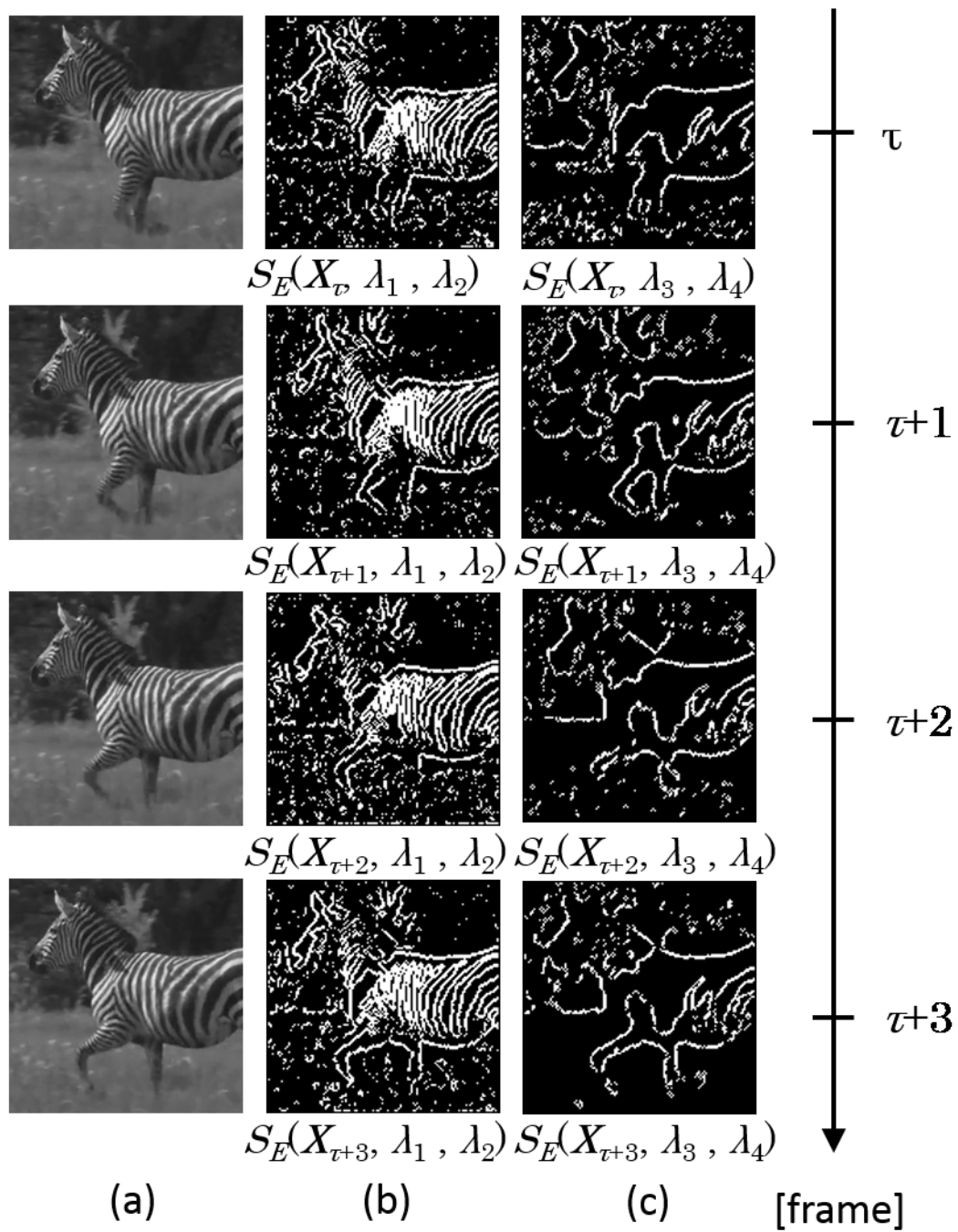


図 4.6 シマウマが歩いている動画に対する出力画像の時系列データ. (a) 入力画像の時刻 τ から $\tau+3$ までの連続する四フレーム分の入力画像. (b)(c) 帯域通過フィルタ $f_{DR}(\lambda_1, \lambda_2)$ と $f_{DR}(\lambda_3, \lambda_4)$ による帯域通過処理画像から抽出されたエッジ画像. 時刻 τ のエッジ画像を $S_E(X_{\tau}, \lambda_1, \lambda_2)$, $S_E(X_{\tau}, \lambda_3, \lambda_4)$ とする. (b)には細かい構造の動きが表現されており, (c)には粗い構造の動きが表現されている.

4.7 結言

本章では、実時間マルチスケールエッジ検出機能を有する知能視覚システムの開発と構築について述べた。開発した知能視覚センサでは、フィルタ処理画像の読み出しと同時に FPGA による並列処理を実行することで、実時間処理を実現した。実時間マルチスケールエッジ画像は粗い構造から細かい構造までの同時形状解析に利用できる他、これまでコンピュータビジョンの分野で研究されてきたマルチスケール画像を用いたエッジ検出法[11]や輪郭特徴ベースのステレオビジョンの粗密法[12]にも発展できる。

次章で構築されたスケール不変特徴量抽出機能を有する知能視覚センサシステムもまた、第三章で述べた複数回フィルタ処理機能を有する知能視覚センサと、本章で述べた FPGA による並列処理を活用することで実現されている。

第5章 抵抗回路網を用いたスケール不変特徴量抽出の解析とシミュレーション

5.1 緒言

画像情報による物体の認識は、画像処理の中心的課題の一つである。近年の画像認識システムの多くは探索画像から特徴点を抽出し、予め抽出しておいたテンプレート特徴点との特徴量同士との対応付けを用いる。この物体認識で用いられる特徴点が備えているべき重要な特性は、その特徴量が画像上での物体の見えの変化に対して不変性を有していることである。

本研究では、開発した知能視覚センサの一応用として、スケール不変特徴量の一つである SIFT 特徴量を実時間で抽出するハードウェアの構築を試みた。SIFT 特徴量とは、画像が抽出されるスケール変化や回転変化に不変な局所特徴量の種類である[13, 14]。しかしながら SIFT アルゴリズムは、前段にマルチスケール画像を生成する処理が必要であり、非常に計算コストが高く、実時間で実行することが難しい。

本研究ではマルチスケール画像を高速に生成するために、抵抗回路網 (図 2.11 参照) を用いる[82]。一方で、従来の SIFT アルゴリズムでは、マルチスケール画像生成のためにガウシアンフィルタが採用されている。そこで本章では、抵抗回路網 (RN) フィルタが SIFT アルゴリズムに用いることができるかを検討する。まずコンピュータ・シミュレーションにより、ガウシアンフィルタとの空間特性の違いが特徴点検出にどのように影響するか検討した。次に、様々なサイズのテストパターンを入力として、画像サイズによらず同じ場所で特徴点が検出できること、また画像サイズに応じたスケールで SIFT キーポイントと呼ばれる特徴点が検出できることを示す。

5.2 SIFT アルゴリズムとその関連研究

5.2.1 スケーラブルフィルタによるキーポイント検出

本研究では、入力画像から局所特徴量を抽出するために SIFT アルゴリズムを採用する。SIFT アルゴリズムについての詳細は Lowe の原著論文に記述されているため[14]、本節ではその概要を簡単に説明する。

図 5.1 に SIFT アルゴリズムの処理の流れを示す。SIFT アルゴリズムは主に、マルチスケール (図 5.1(a))、SIFT キーポイント検出 (図 5.1(b))、そして特徴量記述 (図 5.1(c)) の三つの処理で構成されている。

SIFT キーポイントとは画像の並進移動，スケール変化，回転変化に対して不変であり，かつ小さなノイズや画像の歪みに対しても影響を受け難い画像上の位置のことを指す[13]. SIFT キーポイントは次のように検出される.

図 5.1(a)に示すように，最初に，マルチスケールの帯域通過型空間フィルタ群が入力画像に適用される. この帯域通過型空間フィルタ群はスケラブル性を備えている必要がある. Lowe の提案した SIFT アルゴリズムでは，式(2.8)を満たすスケラブル性を備えた DoG 関数 (式 (2.9)) が帯域通過型空間フィルタとして採用されている.

次に，隣り合う三つのスケール $S_{DG}(\sigma_{k-1})$, $S_{DG}(\sigma_k)$, $S_{DG}(\sigma_{k+1})$ の DoG フィルタ処理画像間の画素値を比較することによって SIFT キーポイント候補点が探索される. ここで， σ_k はキーポイント探索スケールとする. 注目画像に隣り合う三つのスケールの画像上で，注目画素 (図 5.1(b)の黒色のセル) の周辺 3×3 画素の領域中の画素間 ((図 5.1(b)の灰色の 26 個のセル)) で，注目画素の応答が最大値か最小値であるならば，この画素は SIFT キーポイント候補点として選択される. 各候補点群に対して安定性についての検査をし，それを通過した点を最終的な SIFT キーポイントとする. この検査では，低コントラストである点と輪郭周辺の候補点が除去される. ここで，スケール σ_k の帯域通過型処理画像上でキーポイントが検出された時，この σ_k をキーポイントのスケールと定義する.

最後に，SIFT キーポイント周辺の特徴ベクトルが記述される. この特徴ベクトルはスケール及び方位の両方に不変である. SIFT キーポイントの特徴ベクトルはガウシアンフィルタ処理画像 $S_G(\sigma_k)$ から演算される. 特徴ベクトル記述についての詳細は次の節で述べる.

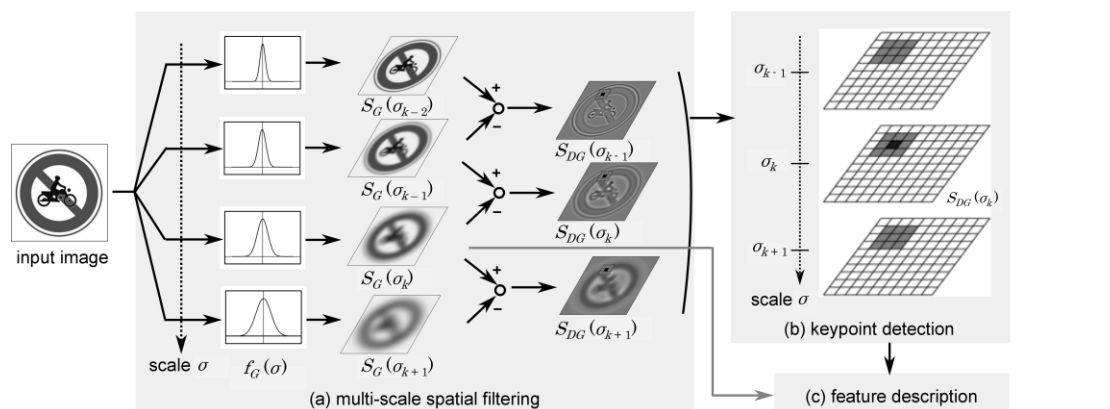


図 5.1 SIFT アルゴリズムにおける処理の流れ. SIFT アルゴリズムは (a) マルチスケール空間フィルタ，(b) SIFT キーポイント検出，そして (c) 特徴抽出で構成されている. この画像で， $S_G(\sigma_k)$ は入力画像にガウシアンフィルタ $f_G(\sigma_k)$ を適用することで得られたガウシアンフィルタ処理画像を示す. 更に， $S_{DG}(\sigma_k)$ は DoG フィルタ処理画像を表現し， σ_k はキーポイントが検出されるフィルタのスケールパラメータを表現する.

5.2.2 特徴量記述

各 SIFT キーポイント位置の周辺で、キーポイントのスケールをもつ勾配ベクトル画像から 128 次元の特徴量 (特徴ベクトル) $\mathbf{F} = (F_1, F_2, \dots, F_{128})$ が記述される. ここで勾配ベクトル画像 $\mathbf{S}_{GR} = (S_{GRM}, S_{GRO})$ は勾配強度 S_{GRM} と勾配方向 S_{GRO} から構成される. 勾配強度 S_{GRM} と勾配方向 S_{GRO} はそれぞれ次の式に従って計算できる.

$$S_{GRM}(\mathbf{X}, \sigma_k) = \{S_{Gx}(\mathbf{X}, \sigma_k)^2 + S_{Gy}(\mathbf{X}, \sigma_k)^2\}^{1/2}. \quad (5.1)$$

$$S_{GRO}(\mathbf{X}, \sigma_k) = \tan^{-1}\{S_{Gy}(\mathbf{X}, \sigma_k)/S_{Gx}(\mathbf{X}, \sigma_k)\}. \quad (5.2)$$

図 5.2 を用いながら更に特徴量の記述についての詳細を説明する. キーポイントの特徴ベクトルは勾配ベクトル画像 \mathbf{S}_{GR} から計算される. 式(5.2)によって求められた勾配方向 S_{GRO} は 0° から 360° までの連続的な値をもつが, その後の処理で離散化される (本研究では図 5.2(a)に示すように 8 方位に離散化する). キーポイントの特徴ベクトルを記述するために, まずキーポイント位置の周辺のベクトル \mathbf{S}_{GR} がガウシアンフィルタで重みづけられる. そのスケールは図 5.1 (a)内の灰色の円に示すように, $S_G(\sigma_k)$ のスケール σ_k の三倍大きくなるように設定される.

特徴ベクトルは SIFT キーポイント周辺の勾配ベクトルから形成される. まず勾配ベクトルから方位ヒストグラム H_{pq} を計算する. 図 5.2(a)に示すようにキーポイント周辺を 4×4 のサブ領域に分割する. 方位ヒストグラムはサブ領域内の全画素の勾配ベクトルのベクトル和をとることで求める. 勾配ベクトルは 8 方位に分類されるために, サブ領域でそれぞれ 8 方位のヒストグラム H_{pq} が得られる. ここで, 添え字 p ($1 \leq p \leq 4$) と q ($1 \leq q \leq 4$) はサブ領域の位置を意味する. これらの 16 個のサブ領域の中の 8 方位のビンから図 5.2(b)に示される特徴ベクトル \mathbf{F} が形成される. \mathbf{F} と H_{pq} はそれぞれ次式のように表現される.

$$\begin{aligned} \mathbf{F} &= (F_1, F_2, \dots, F_{128}), \\ &= (H_{11}, H_{12}, \dots, H_{44}), \end{aligned} \quad (5.3)$$

$$H_{11} = (F_1, F_2, \dots, F_8), \quad (5.4)$$

$$H_{12} = (F_9, F_{10}, \dots, F_{16}), \quad (5.5)$$

...

$$H_{44} = (F_{121}, F_{122}, \dots, F_{128}). \quad (5.6)$$

特徴記述の領域の広さは検出スケールに従って変更する. 図 5.2(a)において, 8 つの矢印は画素の勾配方位を指し, サブ領域は 2×2 画素のサイズである. 図 5.2(b)の垂直方向のバーの長さは特徴ベクトルの各要素の値の大きさを示す.

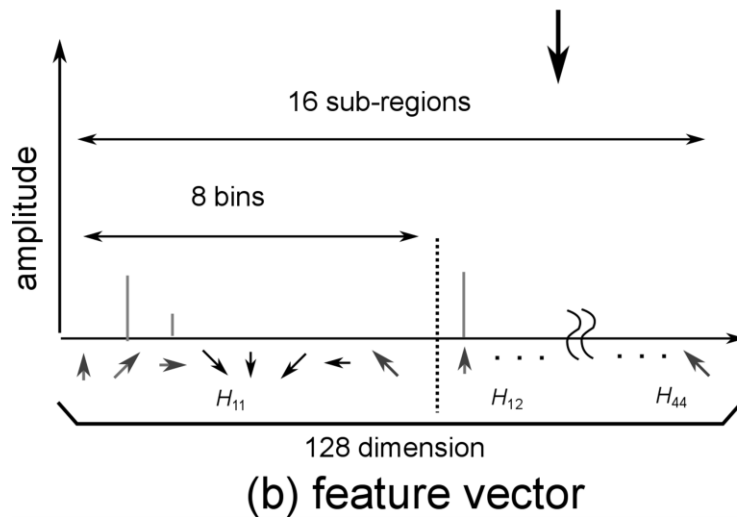
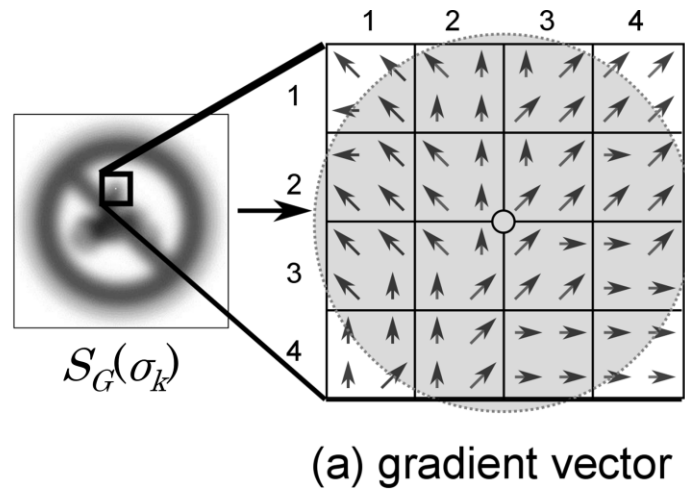


図 5.2 ベクトルの記述. (a) 現在のスケール σ_k の三倍のスケールを有するガウス関数で重みづけられた, $S_G(\sigma_k)$ のキーポイント周辺での勾配ベクトル. 点線の円によって示された領域がそのガウス関数を示す. キーポイントの周りの領域は特徴ベクトルを記述するために 4×4 のサブ領域に分割される. (b) 128 次元の特徴ベクトル. サブ領域内の勾配ベクトルから 8 方位のヒストグラム H が生成される. 各方位は 45 度ごとに離散化されている. バーの長さはサブ領域内の各方位の勾配ベクトルを加算することで求められた特徴ベクトルの各次元の値を示す.

5.2.3 キーポイントマッチング

物体認識のためには入力画像から抽出したキーポイントの特徴ベクトル $\mathbf{F}=(F_1, F_2, \dots, F_{128})$ と予め用意したテンプレートキーポイントの特徴ベクトル $\mathbf{T}=(T_1, T_2, \dots, T_{128})$ とを比較することでキーポイントマッチング (対応点探索) を実行する必要がある. 本研究ではキーポイントの類似度は式 (5.3) に示すユークリッド距離 E を用いて評価した.

$$E = \sqrt{\sum_{l=1}^{128} (F_l - T_l)^2}. \quad (5.7)$$

もし、入力画像から抽出されたキーポイントとテンプレートキーポイントとの間の類似度が閾値 E_{th} より小さく、かつすべてのキーポイントの間で最小となる時、キーポイントがテンプレートとの対応点として選択される。

5.2.4 関連研究

物体認識[13, 83]や分類[84], 歩行者トラッキング[85], ロボットの位置推定[86], そしてパノラマ画像生成[87]など, SIFT 特徴量を用いた様々な応用やその改良アルゴリズムが研究されてきた[88-97]. 局所特徴量の高速抽出が組み込み画像処理システムで実現されると, 更に様々な応用範囲が広がると考えられるため, SIFT アルゴリズムの効率化のためにソフトウェア及びハードウェア的工夫が提案されてきた。

ソフトウェアベースのアプローチとしては, 例えば Fast approximated SIFT [88], SURF (Speeded Up Robust Features) [89], BRISK (Binary Robust Invariant Scalable Keypoints) [90], そして ORB (Oriented FAST and Rotated BRIEF) [91]といった高速に抽出できる特徴点に関するアルゴリズムが挙げられる。これらの方法は抽出された特徴の質を犠牲にすることによって, 計算時間を減少することを狙いとしている。2.4 節, 2.5 節で述べたように, 先行研究としてガウシアンフィルタ処理の高速化技術もまた開発されている。これらの手法は SIFT の計算コストを減少することに成功しているが, 特に画像入力が高解像度である場合, 依然として計算コストは高くなる。

SIFT アルゴリズムを高速化するための専用ハードウェアもまた設計されてきた。これらのシステムは GPU, FPGA, そして ASIC を用いることによって高速化を試みている。GPU ベースのシステムは SIFT アルゴリズムを用いて三次元物体認識に適用されている[62]。GPU はその並列性から高い処理速度を提供するが, GPU ベースのシステムでは, 非常に消費電力が高いため, 携帯性の高い応用には不向きである等, 用途に限りがある。複数の計算機コアによって, 関心領域を抽出し, その領域のみ選択的に SIFT アルゴリズムを適用する物体認識を実行する ASIC もまた開発されてきた。その他の先行研究として FPGA (例えば [63]) や ASIC[98, 99] 上で実装されてきた。これらの専用ハードウェアでは計算速度は向上するものの回路規模やメモリ容量を増加させるなど, 空間フィルタの本質的な計算コストが依然として問題となる。

5.3 抵抗回路網フィルタとガウシアンフィルタの比較

Lowe の提案した SIFT アルゴリズムでは DoG フィルタにスケラブル性を持たせるために, フィルタサイズを決定する σ を等比的に増加させたガウシアンフィルタ群が要求さ

れる．本研究では RN フィルタによってガウシアンフィルタを置き換えるために，これらの空間特性を比較する必要がある．そこで，我々は MATLAB 上で二次元抵抗回路網の空間インパルス応答をシミュレーション解析した．図 5.3 は様々な λ の場合の二次元の正方格子抵抗回路網の単一インパルス応答を示す．各プロットは入力を受け取った画素を通る水平方向の振幅応答を示す．水平軸のゼロは入力位置を示し，図の凡例は λ の値を示している．信号は λ の大きい値においてより広く広がっている．抵抗回路網フィルタの中央近くの応答での減衰はガウシアンフィルタに比べて急激である一方で，抵抗回路網の中心から離れた位置での応答はガウシアンフィルタに比べてより大きい．図 5.4 は図 5.3 の抵抗回路網の λ とフィルタサイズの間関係を示す．ここで，抵抗回路網のフィルタサイズは 1/10 値幅 (FWTM : full width at one tenth of the maximum)，すなわち，抵抗回路網のノード上の応答振幅が最大値の 1/10 になる値の間のピクセル数として定義した．点線は $\text{FWTM} = 2.03 \lambda + 1.74$ を描画したものであり，図 5.4 はフィルタサイズが λ に比例して増加することを示している．一次元抵抗回路網の場合，式 (2.7) から振幅応答が x と λ の比の関数となっているため，フィルタサイズが λ に比例することは明らかである．図 5.4 の結果は解析的に解くことが難しい離散二次元抵抗回路網においても，スケールパラメータ λ とフィルタサイズがほぼ比例することを示している．この結果から，DoR フィルタを入力画像に適用するためには， λ を次のように等比的に設定すればよいことがわかる．

$$\lambda_s = K^{s-1} \lambda_0 (s=1,2,3,\dots), \quad (5.8)$$

ここで λ_0 は λ の最小値であり， s はフィルタの番号である．次の節からの画像処理シミュレーションでは，上記のインパルス応答を抵抗回路網フィルタのカーネルとして採用した．

5.4 DoG フィルタ処理画像と DoR フィルタ処理画像の比較

我々は 2.3 節で述べた DoG フィルタ処理画像と DoR フィルタ処理画像をシミュレーションにより比較した．図 5.5(a)(b) に DoG フィルタと DoR フィルタの単位インパルス応答をそれぞれ示す．DoG フィルタと DoR フィルタの両方が中心周辺拮抗型の空間特性を有している．それらの応答は σ や λ の値が大きいほど，広く広がる．次に我々はフィルタサイズがどのようにフィルタ信号に影響を与えるかを検討した．試験画像として，背景が黒く，中央部に 8 画素幅の白い四角形がある画像 (256×256 画素) を用いた (図 5.6(a)左)．この入力に対して DoG フィルタ $f_{DG}(\sigma_1, \sigma_2) = f_{DG}(1.41, 2)$ を適用した画像を図 5.6(a)右に示す．このフィルタ処理画像の白色の点線上の水平方向の応答振幅に着目する．図 5.6(b)(c) に様々なスケールにおける DoG フィルタ処理画像と DoR フィルタ処理画像の水平方向の応答振幅を示す．横軸は画素位置であり，縦軸の値の組はフィルタのスケールパラメータの組であ

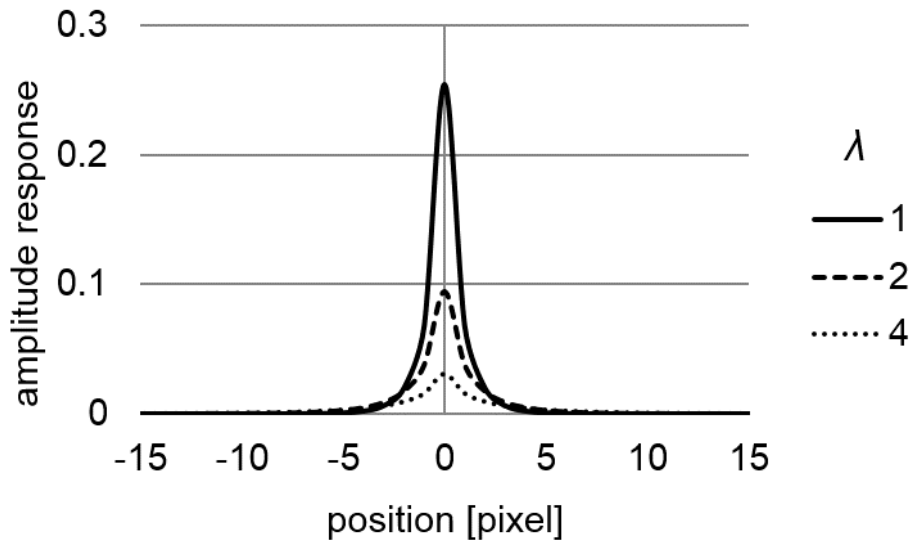


図 5.3 2次元正方格子抵抗回路網の空間インパルス応答. インパルス入力を受け取った画素を通る水平方向の振幅応答を示す.

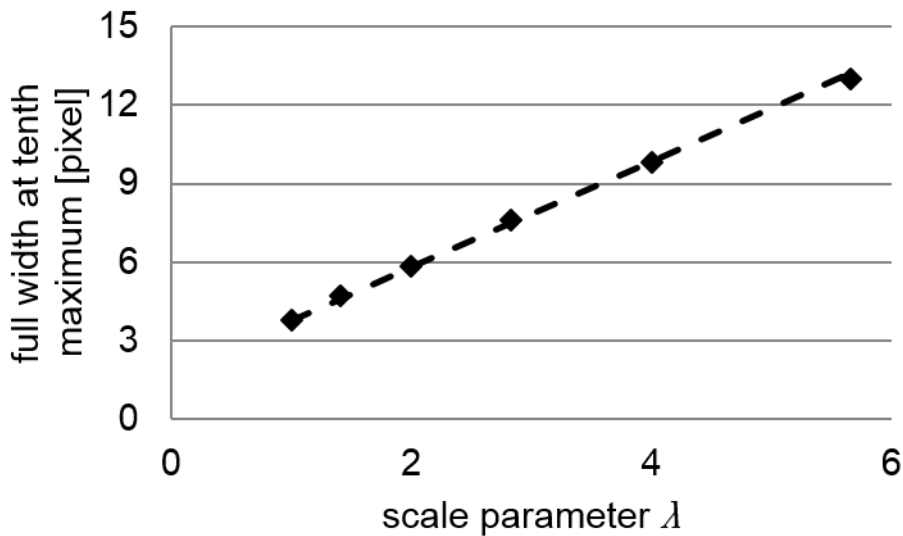


図 5.4 RN フィルタのフィルタサイズと λ との関係. ここでは 1/10 値幅をフィルタサイズとしている. 離散二次元抵抗回路網においてもフィルタサイズは λ に比例して増加していることがわかる.

る. DoG フィルタのスケールパラメータの組 (σ_1, σ_2) の関係は $\sigma_2 = K\sigma_1$ であり, DoR フィルタのスケールパラメータの組 (λ_1, λ_2) の関係は $\lambda_2 = K\lambda_1$ である. ここでシミュレーションのパラメータとして, K は $\sqrt{2}$ で固定し, σ_1 と λ_1 の値は 1.0 から 0.1 ずつ 4.0 まで 30 ステップほど変更した.

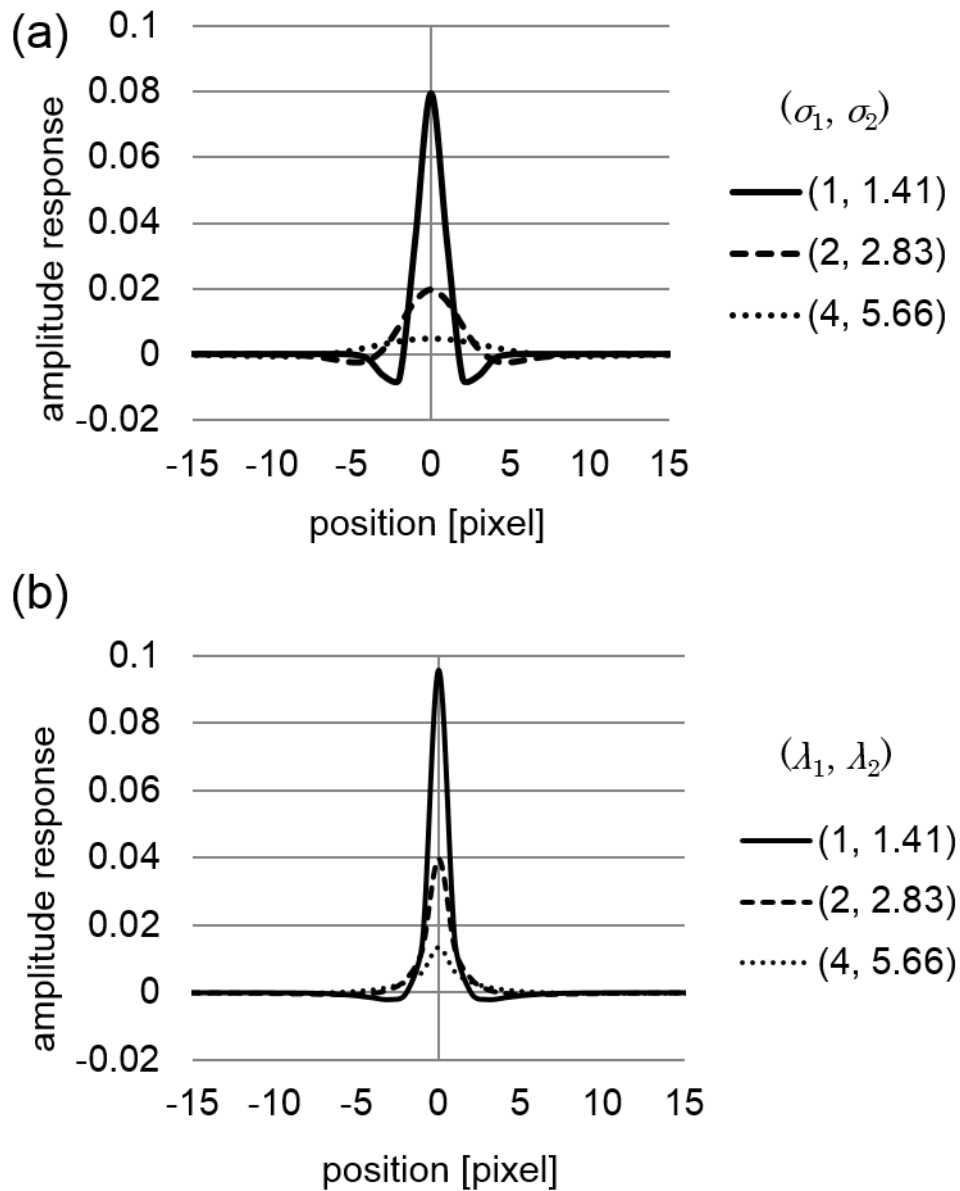


図 5.5 様々なスケールにおける (a) DoG フィルタと (b) DoR フィルタの空間インパルス応答. シミュレーションによって二次元空間インパルス応答を求めたが, ここではインパルス入力が入った画素を通る位置の水平方向の振幅応答を示す.

DoG フィルタと DoR フィルタは双方, 中心周辺拮抗型の空間フィルタを有するため, それらの出力画像が近い空間特性を持つ. 四角形の幅に比べて, それぞれのフィルタのフィルタサイズが小さい時 (図 5.6 (b)(c) の点線 A), 四角形の輪郭が大きく強調され, 中心の応答が輪郭部周辺より小さくなる. フィルタサイズが四角形の幅に合っている時 (図 5.6(b)(c) の点線 B), 四角形の中央が極値をとり, そしてこの画素がキーポイントとして選択される.

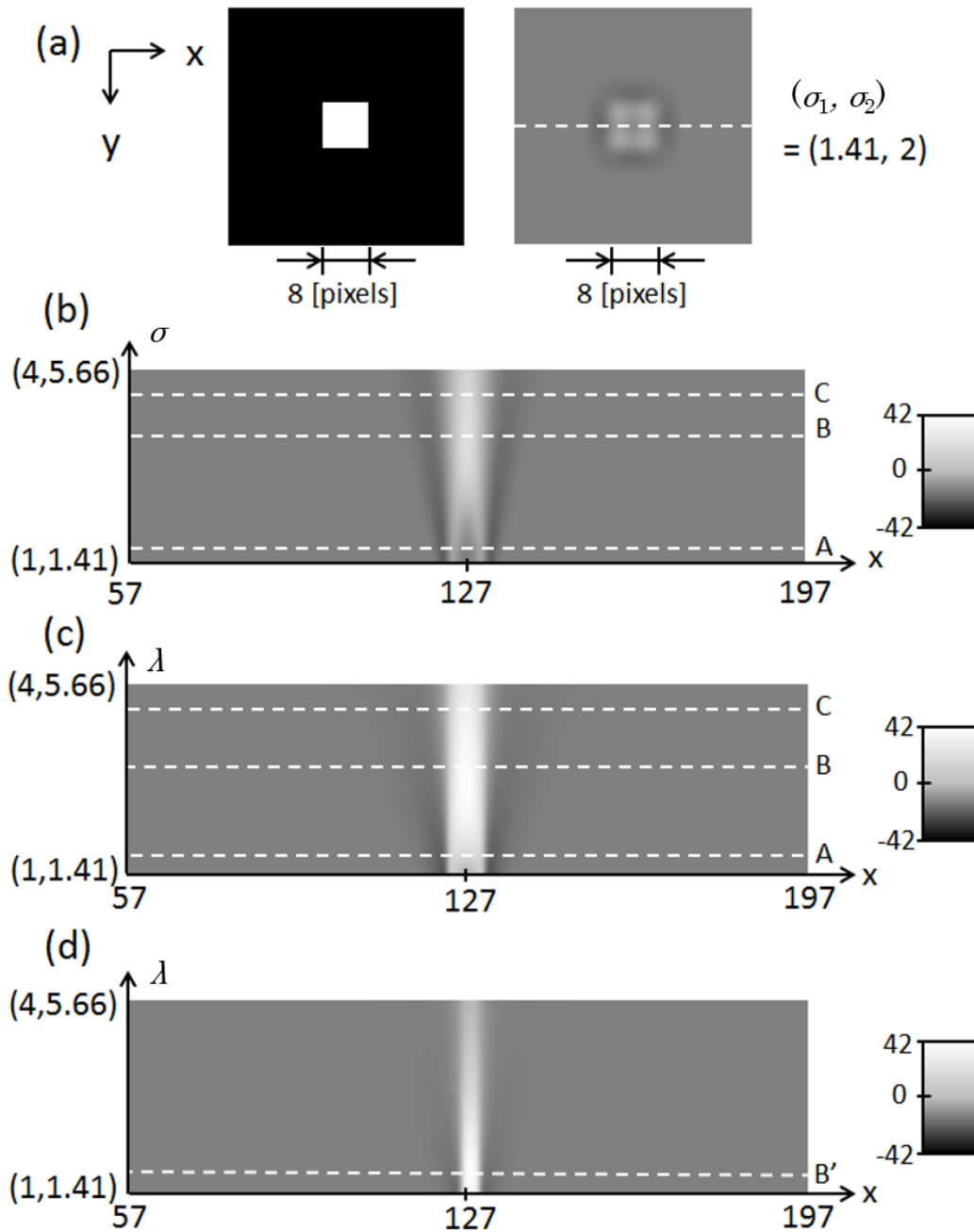


図 5.6 (a) DoG フィルタと DoR フィルタのサイズの変化に対する出力応答への影響を検討するために用いた 256 画素幅の入力画像 (左). 画像の中央には 8 画素の幅の四角形がある. DoG フィルタ $f_{DG}(\sigma_1, \sigma_2) = f_{DG}(1.41, 2)$ による処理画像 (右). (b) 様々なスケールにおける DoG フィルタ処理信号. (c) 様々なスケールにおける DoR フィルタ処理信号. (b)と(c)の各行は(a)の点線上のフィルタ処理信号を示す. また縦軸はスケールパラメータを示し, 横軸は画素位置を示す. (d) 4 画素幅の四角形を入力とした時の DoR フィルタ処理信号.

フィルタサイズは四角形を覆うほど十分に大きい場合 (図 5.6 (b)(c)の点線 C), 四角形の中心の応答は減衰する.

次に DoR フィルタを用いたキーポイント検出においても, 入力パターンのサイズの変更に応じて, SIFT キーポイントの検出スケールが適切に変化するか確認をした. 図 5.6(d)は中央部に 4 画素幅の四角形がある入力画像に対する DoR フィルタ処理信号を示す. 結果として, 上記の場合と比較して半分のスケールで極値が検出された (図 5.8(d)の点線 B').

図 5.6(b)(c)に示される二つのフィルタ処理信号の差は二つのフィルタの空間特性の差に起因する. キーポイント検出された周辺のフィルタ処理信号の応答に着目すると, 空間方向に対する DoR フィルタ信号の勾配は DoG フィルタ処理信号の勾配よりも大きい. 一方で, スケール方向に対する DoR フィルタの勾配は DoG フィルタ処理信号の勾配よりも小さい. 図 5.6 から, DoR フィルタの空間特性とスケール特性との間のバランスは DoG フィルタのそれとは異なるが, キーポイント検出は DoR フィルタを用いても可能であることを示すことができた.

5.5 シミュレーション解析によるキーポイント検出の検討

DoR フィルタを用いて視対象のサイズによらず同じ位置で SIFT キーポイントが検出されるかどうかを検討した. 図 5.7 にシミュレーション解析で用いたテスト画像のパターン (標識) を示す. 本実験では, 三つの異なるサイズの画像 (128×128 , 181×181 , 256×256 画素) を入力画像として用いた. SIFT キーポイントが検出された位置を表 5.1 に示す. 画像上でのキーポイントの相対位置 (x_r, y_r) は $(x_k/X, y_k/Y)$ として表現した. (x_k, y_k) はキーポ

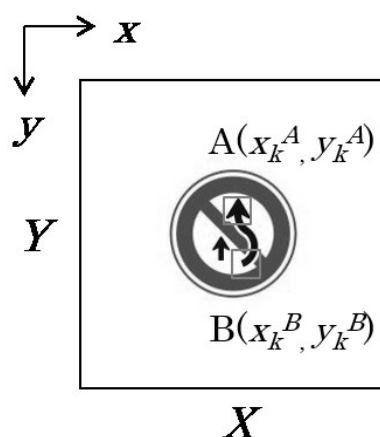


図 5.7 抵抗回路網フィルタによる SIFT アルゴリズムを評価するために使われた入力画像. 灰色の四角形は, 画像から検出された多くのキーポイントから抽出された二つのキーポイント $(A(x_k^A, y_k^A), B(x_k^B, y_k^B))$ の位置を示す. 三つの画像のサイズは 128×128 , 181×181 , そして 256×256 画素である.

表 5.1 キーポイント A の相対的な位置 T

Image size [pixels]	Location (x_k, y_k) [pixels]	Relative position (x_r, y_r)
128 × 128	(66, 53)	(0.516, 0.414)
181 × 181	(93, 76)	(0.514, 0.420)
256 × 256	(132, 107)	(0.514, 0.418)

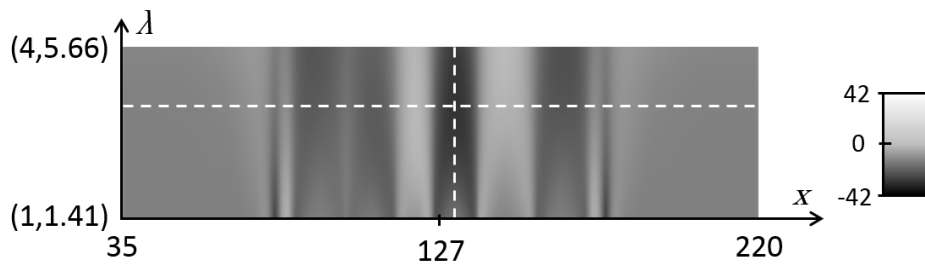


図 5.8 256 画素幅の画像を入力として用いた時の様々なスケールにおける DoR フィルタのフィルタ処理信号. 水平及び垂直な点線はキーポイント A の位置及びスケールをそれぞれ示す.

イントの位置を表現し, X, Y は画像の幅と高さを表現する. 三つのキーポイント相対位置 (x_r, y_r) の平均値は (0.5153, 0.4173) であり, 三つのキーポイントのどれもがこの位置から 1% 以下のずれの範囲に収まる. 従って, DoR フィルタを用いて画像サイズによらず, ほぼ同じ位置でキーポイントが検出されたといえる.

図 5.8 に 256 画素幅の画像の場合における様々なスケールでの DoR フィルタ処理信号を示す. 図 5.8 の行は図 5.7 のキーポイント A 上の水平ラインのフィルタ処理信号を示す. 垂直方向の白い点線はキーポイント A が検出された位置を示しており, $x_k=132$ 番目の画素で負の極値をとっていることがわかる. 次に, 水平方向の白い点線はキーポイント A が検出されたスケールを表しており, $x_k=132$ 番目の画素位置でスケール方向の変化に着目すると, $\lambda_1=2.83$ で負の極値をとっていることがわかる.

図 5.9 はキーポイント A におけるスケールパラメータ λ_1 と画素値の関係を示す. 図 5.9 に示すように, キーポイント A の検出された位置の画素値は画像が小さいサイズでは小さなスケールで最小値となり, 画像が大きいサイズでは大きいスケールで最小値となる. すなわち, 画像サイズに応じて適切なスケールで極値をとり, これがキーポイントのスケールとして選出された. 図 5.10 は画像サイズとキーポイントの検出スケールとの間の関係を示す. 画像サイズの大きさに応じて同じ位置で検出されるキーポイントのスケールが線形的に変化することがわかる. この結果から DoR フィルタを用いてキーポイントが適切なスケールで検出できることが示された.

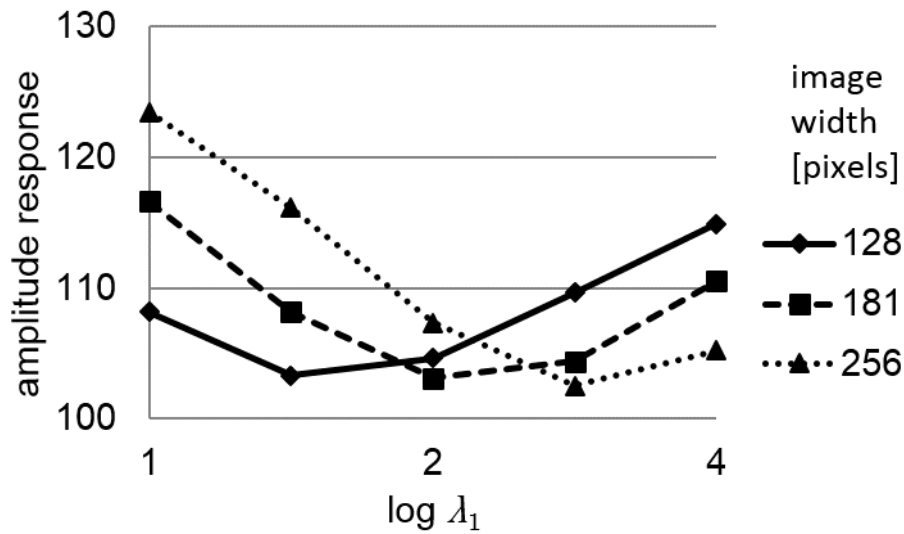


図 5.9 キーポイント A でスケールパラメータ λ_1 と画素値との間の関係.

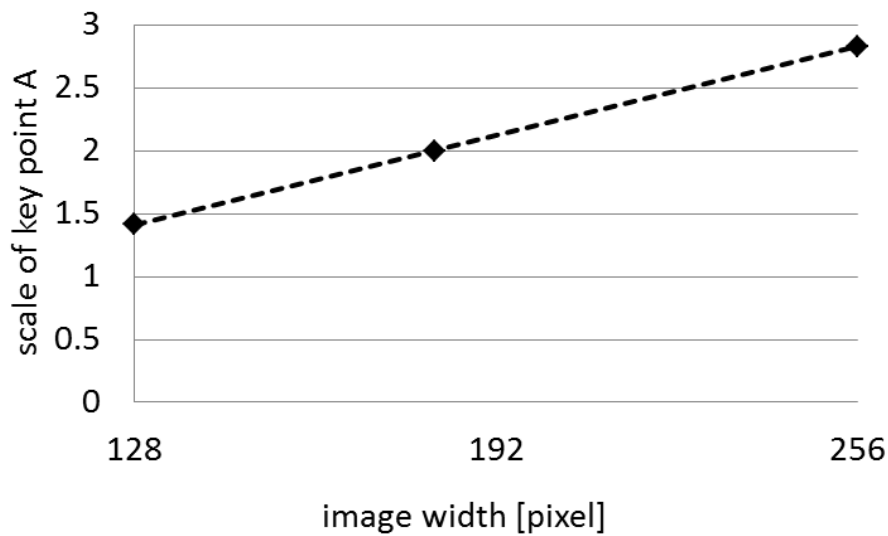


図 5.10 キーポイント A のスケールと画像の幅との間の関係. スケールは画像幅で線形的に増加する.

5.6 シミュレーションによる特徴量記述の検討

5.5 節において, DoR フィルタに用いてスケール不変な特徴点の検出が可能であることが示された. 次に 5.2.2 節で説明した特徴量記述の手法に基づいて, スケール不変特徴量のシミュレーションを行った. 従来の SIFT アルゴリズムとの違いはガウシアンフィルタの代わりに RN フィルタ処理画像の勾配画像から特徴ベクトルを求める点である. 入力画像には

三つの画像サイズの図 5.7 に示すパターンを用いた.

図 5.11 に, 図 5.7 のキーポイント A の特徴ベクトルを示す. 図 5.12 は 256 画素幅の画像上で検出されたキーポイント B で記述された特徴ベクトルを示す. 128 次元の特徴ベクトルはキーポイント周辺の画像勾配から計算された.

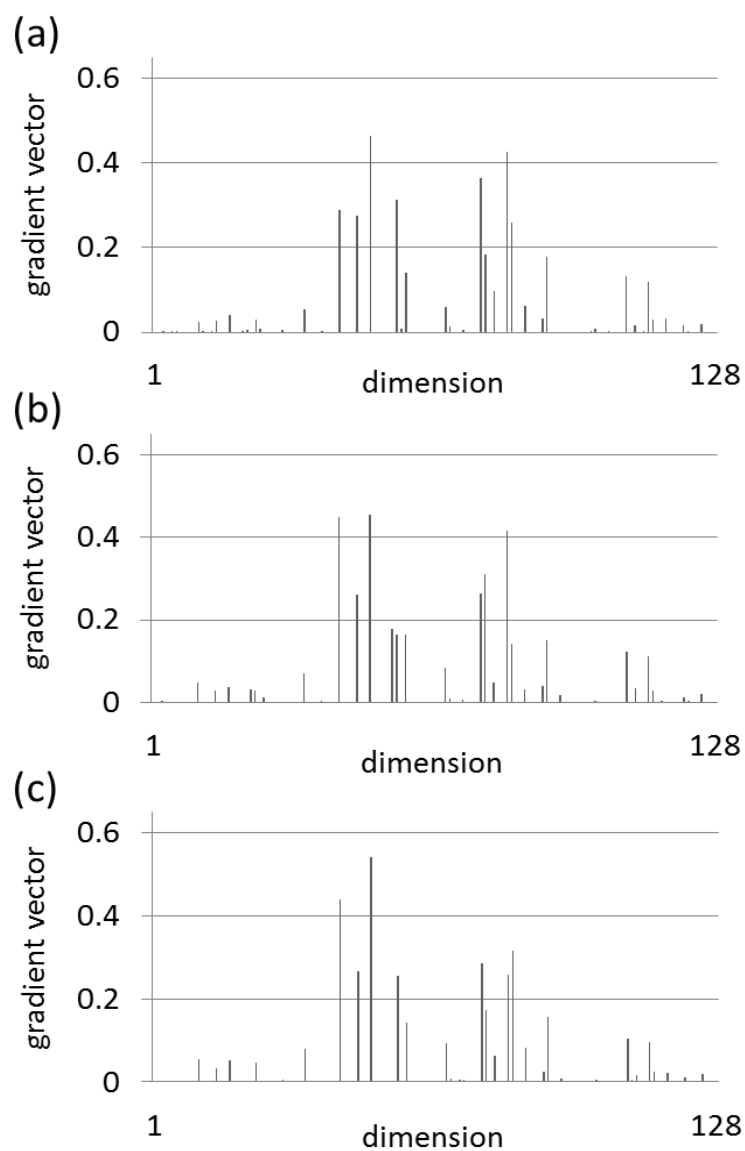


図 5.11 異なる画像サイズで検出されたキーポイント A の特徴ベクトル. (a) 128 × 128 画素. (b) 181 × 181 画素. (c) 256 × 256 画素.

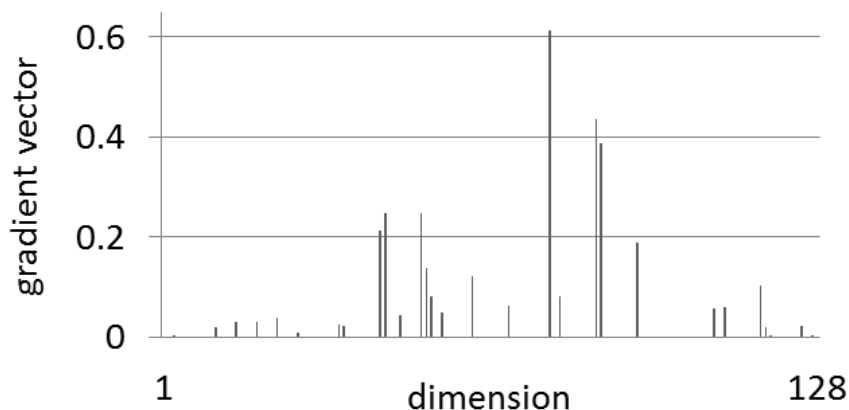


図 5.12 256 × 256 画素の画像から検出されたキーポイント B の特徴ベクトル.

表 5.2 キーポイントの間のユークリッド距離.

Key point location	Image size [pixels]	Euclidean distance
A	128 × 128	0.077
A	181 × 181	0.128
B	256 × 256	1.464

画像サイズ 256 × 256 画素におけるキーポイント A の特徴ベクトルと

他のキーポイントの特徴ベクトルとの間のユークリッド距離

同じ位置から抽出された特徴点がスケール不変な特徴量を有しているか検討するために、キーポイント間の類似性をユークリッド距離によって評価した。ユークリッド距離は式 (5.7) を用いて計算した。表 5.2 に 256 画素幅の画像で検出されたキーポイント A の特徴ベクトルと他のキーポイントの特徴ベクトルとの間のユークリッド距離を示す。異なる画像サイズで検出されたキーポイント A の特徴ベクトル間のユークリッド距離に対して、同じ画像サイズから検出されたキーポイント A と B の間のユークリッド距離が 10 倍以上、大きいことがわかる。

5.7 考察

前節までのシミュレーション結果から、正方格子の二次元抵抗回路網によって、スケール不変特徴量を抽出することが可能であると考えられる。DoR フィルタ処理信号と DoG フィルタ処理信号との違いは空間方向の勾配の強さとスケール方向の勾配の強さのバランスのみであり、これは特徴点検出において大きな問題にはならない。

非常にスケールパラメータが小さい場合、離散抵抗回路網のフィルタカーネルの形状がスケラブル性を備えなくなるという問題が考えられる。しかしながら、画像内で数画素の領域で小さく写っている対象はそもそも入力段階で、そのパターンがうまく表現できていない可能性が高く、実用上問題がないと考えられる。一方、スケールパラメータが大きい場合はフィルタの形状に関して問題がなく、むしろ計算コストの面で抵抗回路網を採用した利点が発揮される。

5.8 結言

スケール不変な特徴量は様々な画像処理応用に用いられている技術である。その特徴を抽出する手法の代表例として **SIFT** アルゴリズムがあるが、その中でのマルチスケール空間フィルタ処理の計算コストが高く、高速化は困難であった。

本研究では抵抗回路網を採用することが、マルチスケール空間フィルタ処理の高速化を実現する上で、非常に効果的な手法であると提案した。このマルチスケール空間フィルタ処理に用いられる低域通過型空間フィルタと帯域通過型空間フィルタにはスケラブル性が要求されていることを述べた。まず、一次元抵抗回路網モデルにおいて、スケールパラメータの値が等比的になるように設定することで、**RN** フィルタと **DoR** フィルタによってスケラブル性をもたせることができることを示した。また、それは二次元正方形格子抵抗回路網においてもシミュレーション実験を行うことで検討された。

様々なスケールにおける **DoG** フィルタ処理画像と **DoR** フィルタ処理画像の比較により、これらの出力が中心周辺拮抗型の空間特性を持ち、キーポイントの検出スケールが両方のフィルタにおいて同じように対象のサイズに応じて変化することが示された。**DoR** フィルタを用いた **SIFT** アルゴリズムのシミュレーションから、画像サイズに依らず、同じ場所でキーポイントが検出されること、そしてキーポイントのスケールが適切に検出されることが示された。更に表 5.2 は抵抗回路網フィルタから記述された特徴がスケール不変性を有していることを示唆させる結果であった。

本章で述べた研究は次章に議論する実時間・低消費電力で **SIFT** 特徴量を記述することができる知能視覚センサシステムに向けた第一歩となった。次章では、これを発展させ、第 3 章で提案した抵抗回路網の時分割制御と第 4 章で提案した **FPGA** による並列処理機構のアイデアに基づき、実時間スケール不変特徴量抽出機構を有する知能視覚センサの開発したことを論じる。

第6章 実時間スケール不変特徴量抽出機能を有する知能視覚センサシステム

6.1 緒言

実時間物体追跡は定点カメラによる人流監視から移動カメラによるロボットの制御まで様々な応用に期待される画像処理技術の一つである。視覚情報に基づく物体追跡では対象物の画像特徴量に基づいて、各画像フレーム内の対象物を探索・追従する。これらのタスクには特徴点検出器や特徴量抽出器（記述子）を使うことが効果的であることが知られている（総説として[96, 97]）。

特徴点ベースの物体追跡手法ではまず画像から特徴点を検出し、各特徴点において特徴ベクトルを記述する。例えば移動ロボットなど絶え間なく対象と搭載されたカメラとの距離が変化する場、画像上での見えの大きさが時々刻々と変化する。従って、物体追跡に適した特徴点にはスケール不変性が要求される。このような特徴点を高速に抽出する必要があることに加えて、移動ロボットに使われる画像センシングシステムはコンパクトでかつ低消費電力であることが要求される。

第5章において、SIFT キーポイントを高速に抽出するためにガウシアンフィルタの代わりに RN フィルタを採用する手法を提案した。コンピュータ・シミュレーションの結果から、RN フィルタ群を用いてスケール不変特徴量が抽出できることが明らかになった[82]。しかしながら、複数のスケールパラメータで RN フィルタを画像に適用するために、アナログ集積回路技術によって多段構造の抵抗回路網を実装することや、SIFT アルゴリズム内の処理のすべてをアナログ回路で実装することは現実的ではない。スケール不変特徴量をカメラの視野から実時間で抽出するためには、アナログ抵抗回路網とデジタル計算機の利点を組み合わせた新しい処理アーキテクチャが要求される。そこで我々はシリコン網膜と FPGA と汎用コンピュータを組み合わせることで、高速に SIFT 演算を実行する知能視覚センサシステムを提案する。このシステムでは SIFT 演算の前段のマルチスケール空間フィルタ処理には第3章で提案した抵抗回路網による複数回空間フィルタ処理が採用されており、中段のキーポイント検出と勾配ベクトル画像の計算には第4章で述べた FPGA 上での各画像の並列演算が採用されている。後段の汎用コンピュータでは、特徴量記述とキーポイントマッチングが実行される。

本章では、実時間物体追跡の実現のための知能視覚センサシステムについて論じる[100]。まず知能視覚センサシステムを構築し、SIFT キーポイントが画像から抽出できるか検討する。次いで、システムによる実時間 SIFT 演算及び特徴点追跡の実験を行い、その結果を考

察することで、提案するアーキテクチャの有用性を明らかにする。

6.2 抵抗回路網による実時間スケーラブルフィルタ処理

6.2.1 実時間マルチスケールフィルタ処理

SIFT アルゴリズムの計算コストを引き上げている要因の一つは、前段のマルチスケール空間フィルタ処理である。第 5 章で述べたように、二次元抵抗回路網は、スケーラブル性を備えた帯域通過型空間フィルタ群として適用できる。ただし、マルチスケール空間フィルタ処理のために、異なる抵抗値の抵抗群で構成された多段構造のアナログ抵抗回路網を実装することは、回路規模の増大を招く。そこで、可変抵抗器として用いることのできる CMOS 抵抗回路の利点を活かし、第 3 章で述べた単一画像フレームへの時分割法により、マルチスケール DoR フィルタ処理を高速に実行することを試みる。前章までと同様に、この時分割法を実装するために、シリコン網膜と FPGA で構成された視覚センサ”ASPECTUS-U”を用いた。3.2 節で述べたこととの違いは、DoR フィルタにスケーラブル性を備えさせるために、一フレームを時分割した各区間でスケールパラメータを等比的に $\lambda_2 = K\lambda_0, \lambda_1 = K^2\lambda_0, \dots$ と増加させていく点である。この点については、次の 6.2.2 節にまとめた。スケールパラメータは FPGA からの制御信号 (MOS 抵抗へのバイアス電圧) によって設定される。時分割された各区間において、等比的なスケールに設定された RN フィルタによる処理画像 $V(\mathbf{X}, K\lambda_0), V(\mathbf{X}, K^2\lambda_0), \dots$ が、抵抗回路網によって瞬間的に計算され、順次 FPGA に転送される。ここでは、6 枚の RN フィルタ処理画像を単一フレーム時間 33ms 内に取得できるように FPGA 内の制御回路を構築した。FPGA 内の回路についての詳細は、6.3 節で述べる。

6.2.2 MOS 抵抗回路網によるスケーラブルフィルタ処理

SIFT アルゴリズムではスケーラブル性を備えた帯域通過型空間フィルタ群を用いる。ここでは、DoR フィルタにスケーラブル性をもたせるように RN フィルタの空間特性を設定しなければならない。そこで、様々な空間定数における MOS 抵抗回路網の線拡がり関数を測定することで、適切に空間フィルタ群の空間特性を設定できるかを検討した。図 6.1 に異なるスケールに設定された MOS 抵抗回路網の線拡がり関数を示す。これらの線拡がり関数はシリコン網膜に対して、LCD を用いて垂直なスリット光 (図 6.1(a)) を提示することによって測定された。

図 6.1(b) に二つの異なるスケールをもつ RN フィルタの線拡がり関数を示す。凡例内の λ の添え字 s は各 RN フィルタのスケール λ の番号を指す。高い番号であればあるほど、大きなスケールを表現している。RN フィルタの線拡がり関数は $x=0$ の時、最も高い値が得られ、中心からの距離に従って徐々に減衰する。6 番目の RN フィルタ $f_{RN}(\lambda_6)$ のスケール λ_6 が $f_{RN}(\lambda_1)$ のスケール λ_1 より大きいため、 $f_{RN}(\lambda_6)$ によってフィルタ処理された信号は $f_{RN}(\lambda_1)$ によってフィルタ処理された信号よりも長い距離にわたって広がっている。

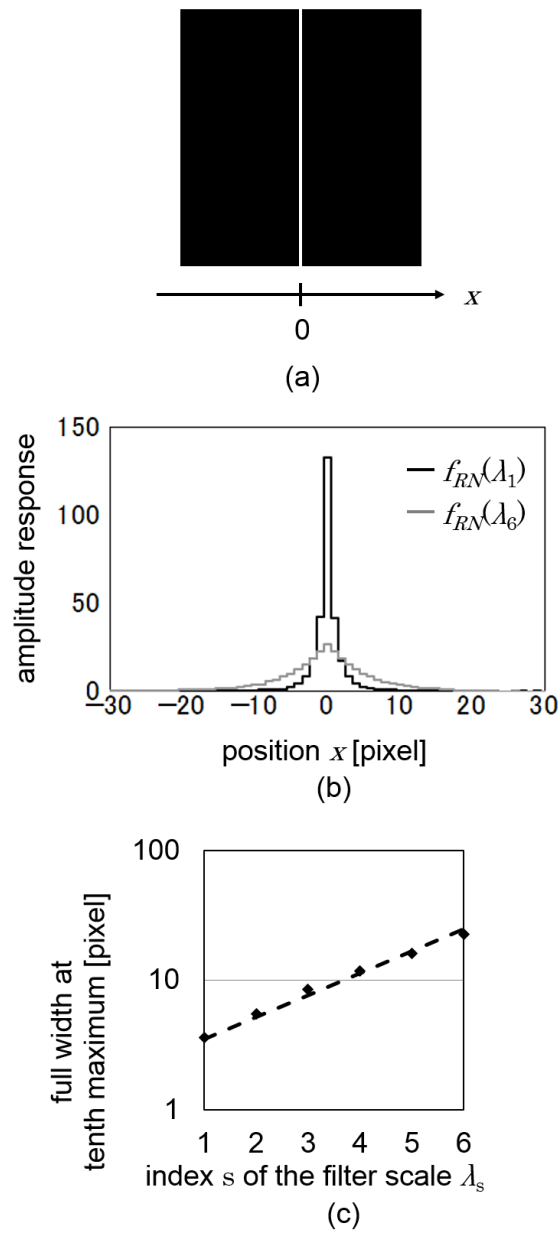


図 6.1 抵抗回路網の空間特性と測定に用いた入力画像. (a) 入力画像. (b) RN フィルタ $f_{RN}(\lambda_s)$ の線拡がり関数. 凡例の $f_{RN}(\lambda_s)$ の添え字 s は RN フィルタのスケールの番号を示す. そこで, 添え字 s の値がより大きい場合, λ の値が大きいことを表している. 信号は λ の値が高いほど, より長い距離にわたって伝搬する. (c) 応答の 1/10 値となる幅である FWTM をフィルタサイズと定義し, そのフィルタサイズと RN フィルタの番号との関係を示す. 各 RN フィルタの λ はフィルタサイズが等比的になるように制御信号によって設定された.

DoR フィルタにスケーラブル性を備えさせるためには、式(2.15)を満たすように、抵抗回路網のスケールパラメータ λ を等比的に増加するように、MOS 抵抗へのバイアス電圧を制御する必要がある。図 6.1(c)に 6 つの抵抗回路網フィルタのサイズを示す。ここで抵抗回路網のサイズは線拡がり関数の 1/10 幅である FWTM で表現した。この FWTM は抵抗回路網の空間定数 (RN フィルタのスケール) に比例する。FWTM は線形補間によって画素間の応答が近似された線拡がり関数から計算された。例えばフィルタ処理値が最大

(13.03) の 1/10 まで減衰する点 $x=1.80, -1.78$ であるため、 $f_{RN}(\lambda_1)$ の FWTM は 3.58 となる。図 6.1(c)の点線は $\text{FWTM} = 3.5 + 1.48 \times s$ を表現しており、この図はフィルタの大きさがほぼ等比的に増加するように MOS 抵抗回路網の空間特性が設定できているを示している。次節から述べられる RN フィルタはここで用いられたフィルタと同様のスケールに設定されている。

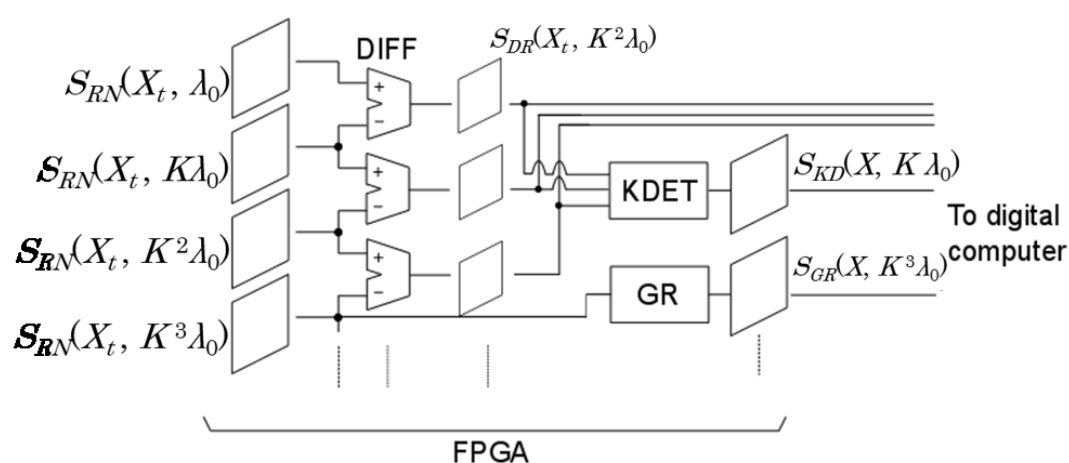
6.3 FPGA によるキーポイント検出及び勾配ベクトル演算

6.3.1 FPGA による並列処理

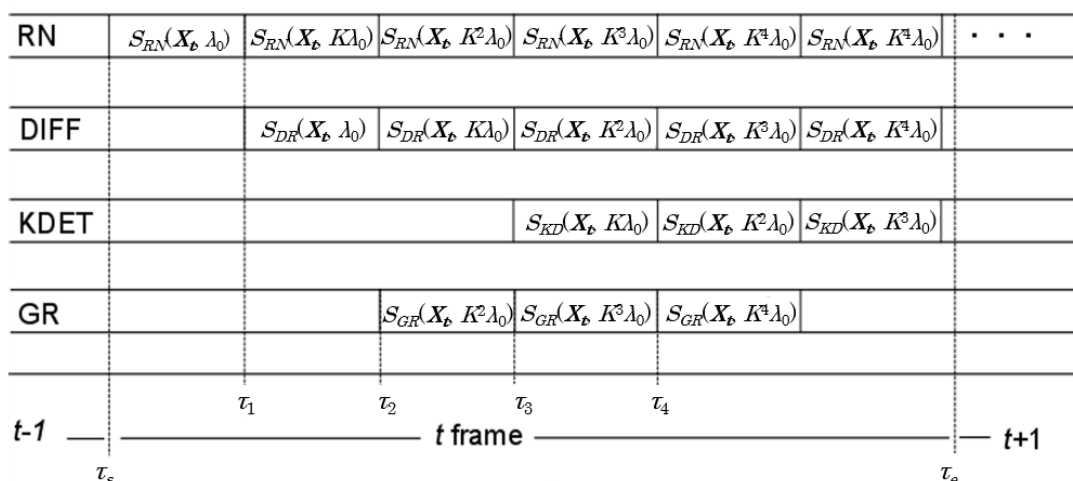
時分割法を用いれば抵抗回路網によるマルチスケールフィルタ処理と同時並行的にキーポイント検出が FPGA 上のデジタル回路で実行することができる。図 6.2(a)にスケール不変特徴量を抽出する FPGA 中のデジタル回路を示す。前節で説明したように、等比的にスケールを設定した RN フィルタにより処理された画像が、A/D 変換の後に FPGA 内のメモリ (図 6.2(a) $S_{RN}(\mathbf{X}_s, K^{s-1}\lambda_0)$ ($s=1,2,3,4$)) に保存される。ここで、 $\mathbf{X}_s=(x, y, t)$ は画像の位置と時刻 (フレーム) を表現するベクトルである。s 番目の抵抗回路網フィルタ処理画像が $S_{RN}(\mathbf{X}_s, K^{s-1}\lambda_0)$ 用メモリに保存される。キーポイント検出は FPGA に並列に配置された専用処理回路によって実行される。それらの回路は図 6.2(a)上で、ラベル付きのブロックとして表現されている。図中の DIFF, KDET, そして GR はそれぞれ差分器、キーポイント検出器、そして勾配ベクトル演算器を意味する。DIFF 回路は、隣接するスケールを有する二つの抵抗回路網フィルタ処理画像の差を計算する。この回路はマルチスケール DoR フィルタ処理画像を生成する。KDET 回路は DoR フィルタ処理画像内から、SIFT キーポイントを探査する。GR 回路は、抵抗回路網フィルタ処理画像の勾配ベクトルを計算する。これらの回路は独立に、かつ並列に動作する。従って、三種類の画像、DoR フィルタ処理画像、抵抗回路網の勾配分布、そしてキーポイント候補点分布は単一フレーム内に並列的に得られる。

図 6.2(b)に三つの専用処理回路のタイミングチャートを示す。ここでは、一フレームを各 RN フィルタ画像の読み出し時間に合わせて分割し、各区間で異なるスケールの画像に対して処理が実行される。これらの専用処理回路による処理の流れを、全ての回路が同時並列的に動作している t_3-t_4 の区間を例に説明する。まずシリコン網膜から出力される RN フィルタ処理画像 $S_{RN}(\mathbf{X}_s, K^3\lambda_0)$ が、 t_3-t_4 の区間をかけて一画素ずつ A/D 変換されて FPGA 内の

メモリに保存される。それと同時に RN フィルタの画素は DIFF 回路に入力される。同じタイミングで FPGA 内のメモリから一つ前のスケールの RN フィルタ画像 $S_{RN}(X_t, K^2\lambda_0)$ の画素が読み出され、DIFF 回路に入力される。ここで二つの画像の差が取られ、DoR フィルタ処理画像 $S_{DR}(X_t, K^2\lambda_0)$ が生成される。 $S_{DR}(X_t, K^2\lambda_0)$ はメモリに保存されると同時に KDET 回路に入力される。キーポイント検出では、三つのスケールの帯域通過型空間フィルタ処理画像を用いるため、KDER 回路には予め保存されていた $S_{DR}(X_t, \lambda_0)$, $S_{DR}(X_t, K\lambda_0)$ が同時に入力される。ここでキーポイント画像 $S_{KD}(X_t, K\lambda)$ が生成される。キーポイント画像とはキーポイントが検出された位置の画素値を 1 とするバイナリ画像である。GR 回路は



(a)



(b)

図 6.2 (a) シリコン網膜と FPGA 内のデジタル回路による並列処理の動作の流れ。

(b) FPGA 内のデジタル回路上での並列演算のためのタイミングチャート。

キーポイントの特徴記述を実行するために必要な勾配ベクトル画像 $S_{GR}(X_t, K^3\lambda)$ を計算する。この画像は RN フィルタ処理画像の画素読み出しタイミングと同時に生成される。図 6.2(b)に示すように、これらのデジタル回路が並列に動作するために、各工程は単一フレーム内に完了する。

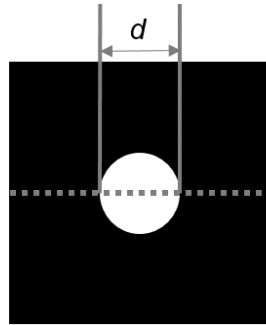
時刻 t_e で、APS は $t+1$ 番目のフレームの画像蓄積を完了し、フレーム t の出力画像 (S_{DoR} , S_{KD} , S_{GR}) が USB インターフェースを介してデジタルコンピュータに転送される。デジタルコンピュータ上では三つの出力画像を用いたキーポイントの特徴量の記述を実行する。特徴量の記述についての詳細は 6.4 節に述べる。このように抵抗回路網フィルタと FPGA の並列演算器を活用した戦略を採用することで画像サイズに影響をほぼ受けずに一フレーム内に処理を完了できる。

6.3.2 SIFT キーポイント検出

SIFT キーポイントは三つの異なるスケールの DoR フィルタ処理画像をまたぐ、注目画素周辺の 3×3 の領域に含まれる 27 画素の中で、最大値か最小値をとる画素として定義される。キーポイントのスケールはキーポイントが検出されたフィルタのスケールとして定義されている。図 6.3(a)に示す黒色の背景に白色の円で表現される画像を試験画像とし、その円のサイズとフィルタ処理画像の応答との間の関係を検討する。この出力は LCD に直径 d の円 (図 6.3(a)) を提示することによって取得された。図 6.3(b)(c)は円の中心にある灰色の点線で示される水平ライン上での DoR フィルタ処理画像群のラインプロファイルを示す。この時、それぞれの RN フィルタは図 6.2 で示される六つのスケール $\lambda_1 = \lambda_0$, $\lambda_2 = K\lambda_0$, $\lambda_3 = K^2\lambda_0$, ... に設定されている。

5 画素の直径の円が提示された時、中間のスケールのフィルタ処理画像 $f_{DR}(\lambda_2)$ におけるラインプロファイルは、 $x = 0$ の位置で最大値となる。この値はより小さいスケールの $f_{DR}(\lambda_1)$ や、より大きなスケールの $f_{DR}(\lambda_3)$ の場合における同画素の画素値より大きくなる。 $x = 0$ での $f_{DR}(\lambda_2)$ の画素は他の 26 個の近傍の画素よりも大きくなるため、この画素はスケール λ_2 のキーポイント候補点として選出される。7 画素の直径のより大きな円が提示された時、 $f_{DR}(\lambda_3)$ のラインプロファイルは $x = 0$ の画素で空間的に最大の値を持ち、またこの値は $f_{DR}(\lambda_2)$ や $f_{DR}(\lambda_4)$ の周辺の画素値より大きい。従って、この画素が上記と同じ理由でスケール λ_3 のキーポイントとして選出される。

図 6.4 にフィルタのスケールと、円の中心でのフィルタ処理画像の画素値との間の関係を示す。上で述べたように、直径 5 画素の円の画素値 (図 6.4 のダイヤ) と直径 7 画素の円の画素値 (図 6.4 でのクロス) が、それぞれスケール λ_2 と λ_3 で最大値をとる。直径 9 画素の円に対する応答はスケール λ_4 で最大値をとる (図 6.4 の円)。図 6.4 に示すように、大きな画像特徴では、フィルタ処理信号が極値となるスケールもまた高くなる。これらの結果はキーポイント候補が適切なスケールで検出されたことを示している。



(a)

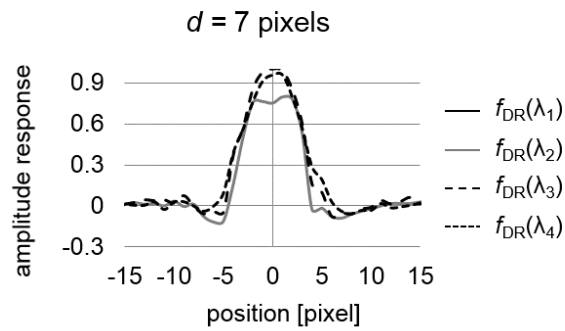
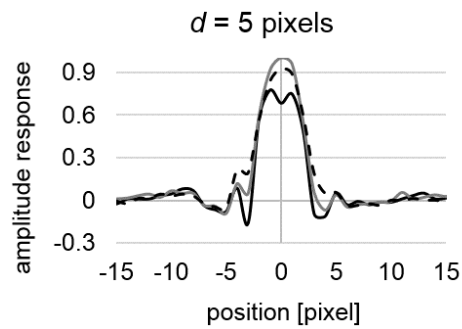


図 6.3 黒色の背景上にある白色の円に対する様々なスケールの DoR フィルタ処理信号. (a) 入力画像. (b) 5 画素 (c) 7 画素の直径の円の中心を通る水平方向の((a)の灰色の点線)の DoR フィルタ処理信号. 円の中心の位置を $(x, y) = (0, 0)$ と定義する.

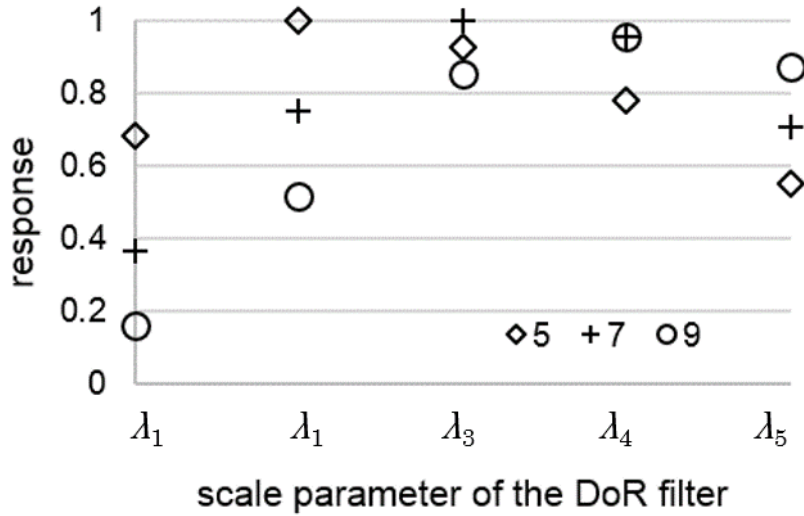


図 6.4 キーポイント位置におけるスケールパラメータと画素値の関係。

6.3.3 勾配ベクトル画像の生成回路

FPGA 内のデジタル演算回路 GR は抵抗回路網フィルタ処理画像 $S_{RM}(\mathbf{X}, K^{s-1}\lambda_0)$ から勾配ベクトル画像 $\mathbf{S}_{GR}=(S_{GRM}, S_{GRO})$ を計算する。 \mathbf{S}_{GR} は勾配強度 $S_{GRM}(\mathbf{X}, K^{s-1}\lambda_0)$ と勾配方位 $S_{GRO}(\mathbf{X}, K^{s-1}\lambda_0)$ から構成される。この勾配ベクトル画像はキーポイントの特徴ベクトルを記述するために用いられる。

勾配ベクトル画像を計算するデジタル演算回路 GR の、レジスタ転送レベル (Register Transfer Level ; RTL) の回路図を図 6.5 に示す。本回路は主に 256 画素分保存できるシフトレジスタ (図中 SHIFT), 加算器 (図中 ADD), 差分器 (図中 DIFF), 比較器 (図中 COMP), 二乗演算器 (図中 SQU), ビット深度変換器 (図中 BDC) で構成されている。本回路はシフトレジスタによる画素保存部, 勾配強度演算部 (図中 GRM), 勾配方位演算部 (図中 GRO) の三つに分けられる。まず $S_{RN}(\mathbf{X}, K^{s-1}\lambda_0)$ がシリコン網膜から一画素ずつ読み出され, FIFO (First in first out) 方式でシフトレジスタ (図中 SHIFT) に 257 画素分の画素値が一時記憶される。シフトレジスタ内の 128 番目のレジスタに保存された画素を注目画素とし, その位置 $\mathbf{X}=(x, y)$ における勾配ベクトルが計算され, 回路から出力される。ここで, 勾配強度演算と勾配方位演算に用いる周辺 3×3 画素を関心領域 (図中 Region of interest) とする。注目画素の上下左右の画素が演算に利用されるため, 対応するレジスタ (0 番目, 127 番目, 129 番目, 256 番目) の出力が勾配強度演算回路と勾配方位演算回路に入力される。

勾配強度演算回路 GRM では, 式(6.1)から式(6.3)の演算が実行される。

$$S_{GRM}(\mathbf{X}, K^{s-1}\lambda) = \left\{ S_{RN_x}(\mathbf{X}, K^{s-1}\lambda)^2 + S_{RN_y}(\mathbf{X}, K^{s-1}\lambda)^2 \right\}, \quad (6.1)$$

$$S_{RN_x}(\mathbf{X}, K^{s-1}\lambda) = S_{RN}(\mathbf{X}_{x-1}, K^{s-1}\lambda) - S_{RN}(\mathbf{X}_{x+1}, K^{s-1}\lambda), \quad (6.2)$$

$$S_{RNy}(\mathbf{X}, K^{s-1}\lambda) = S_{RN}(\mathbf{X}_{y-1}, K^{s-1}\lambda) - S_{RN}(\mathbf{X}_{y+1}, K^{s-1}\lambda). \quad (6.3)$$

ここで、 $\mathbf{X}_{x+1}=(x+1, y, \vartheta)$, $\mathbf{X}_{y-1}=(x, y-1, \vartheta)$, $\mathbf{X}_{x-1}=(x-1, y, \vartheta)$, $\mathbf{X}_{y+1}=(x, y+1, \vartheta)$ である。式(6.2)と式(6.3)は前段の差分器によって行われ、それらが二乗演算器にそれぞれ入力される。二つの二乗演算器の出力の和がとられることで、式(6.1)で表される演算が実行される。二乗演算器の出力の和は 16bit で表現されるが、本システムでは FPGA 内のメモリ容量を節約するために、勾配強度のビット深度を 16bit から 4bit に下げる。

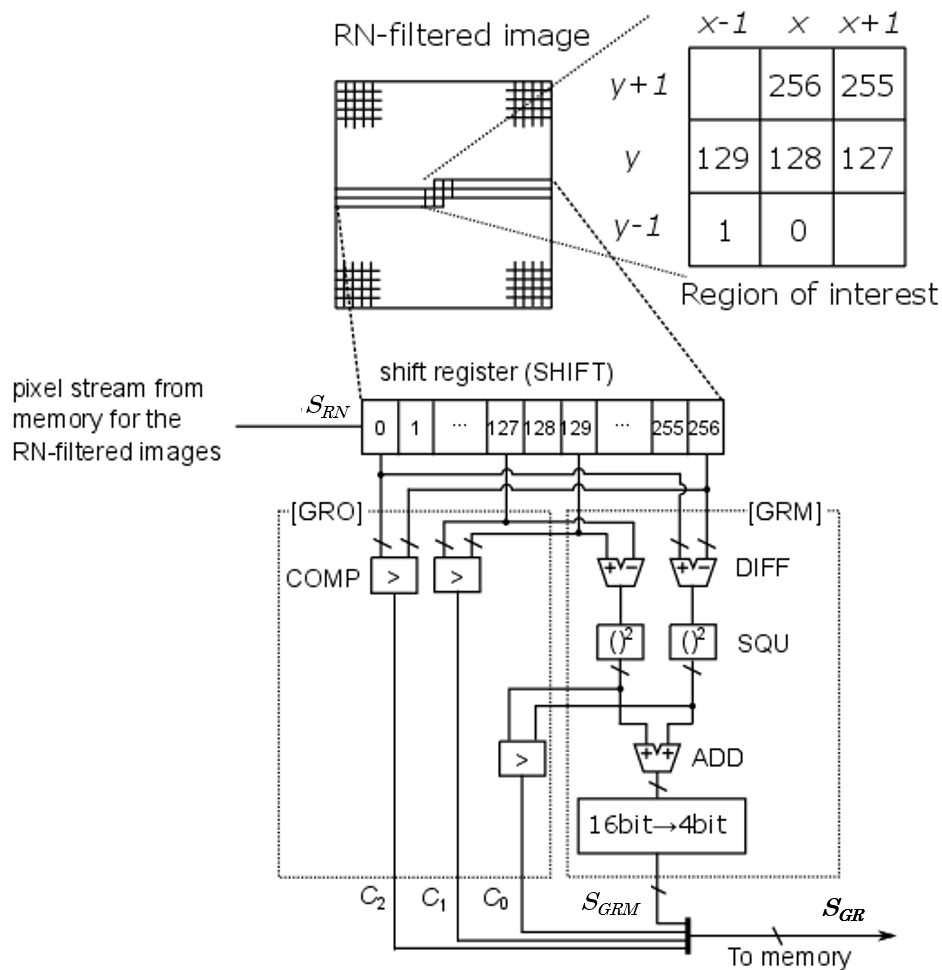


図 6.5 デジタル演算回路 GR の回路図。本回路は RN フィルタ処理画像 S_{RN} から勾配ベクトル S_{GR} を演算する。本回路は主に 257 画素分を保存するシフトレジスタと勾配強度演算部 GRM と勾配方位演算部 GRO で構成される。更に GRM は差分器 DIFF, 二乗演算器 SQU, 加算器 ADD とビット深度変換器で構成され、GRO は三つの比較演算器で構成されている。動作の詳細は本文中に述べる。

次に勾配方位演算回路 GRO の動作を述べる．勾配方位は次の条件式から得られる C_2 , C_1 , C_0 の 3bit を用いて 8 方位 ($45 \times i$ 度, ($0 \leq i \leq 7$)) の中に分類される．

$$C_2(\mathbf{X}, K^{s-1}\lambda) = \begin{cases} 1 (S_{RNy}(\mathbf{X}, K^{s-1}\lambda) \geq 0) \\ 0 (S_{RNy}(\mathbf{X}, K^{s-1}\lambda) < 0), \end{cases} \quad (6.4)$$

$$C_1(\mathbf{X}, K^{s-1}\lambda) = \begin{cases} 1 (S_{RNx}(\mathbf{X}, K^{s-1}\lambda) \geq 0) \\ 0 (S_{RNx}(\mathbf{X}, K^{s-1}\lambda) < 0), \end{cases} \quad (6.5)$$

$$C_0(\mathbf{X}, K^{s-1}\lambda) = \begin{cases} 1 ((S_{RNx}(\mathbf{X}, K^{s-1}\lambda))^2 \geq (S_{RNy}(\mathbf{X}, K^{s-1}\lambda))^2) \\ 0 ((S_{RNx}(\mathbf{X}, K^{s-1}\lambda))^2 < (S_{RNy}(\mathbf{X}, K^{s-1}\lambda))^2). \end{cases} \quad (6.6)$$

例えば, (C_2, C_1, C_0) が (0, 1, 0) である時, 特徴ベクトルの勾配方位の画素値は 0 と定義する．条件式(6.4), (6.5), (6.6)の比較演算は, それぞれコンパレータ COMP1, COMP2, COMP3 が実行する．

最終的に勾配強度演算回路の出力と勾配方位演算回路の出力が一つになり, 7bit の画素値の勾配ベクトルが回路から出力され, 順次 FPGA 内のメモリへと転送される．

6.4 ソフトウェアによる特徴ベクトル記述と対応点探索

汎用デジタルコンピュータ内のソフトウェアでは, キーポイント候補点からキーポイントの選出, 特徴ベクトルの記述, そして特徴点マッチングが実行される．

デジタルコンピュータ上で実行されるソフトウェアは, FPGA から送信されてきた DoR フィルタ処理画像, キーポイント画像, 勾配画像を使って特徴を記述する．SIFT アルゴリズムでの大部分の計算は, 本研究で開発されたハードウェアで実行されているため, デジタルコンピュータ上での計算負荷は大幅に削減されている．従って, 視野内の特徴点のトラッキングが実時間で実行できる．ソフトウェア・アルゴリズムを以下に述べる．

初めに, 安定性の低いキーポイント候補を次の二つの基準で除去する．キーポイント検出位置において, キーポイントスケールの DoG フィルタ処理信号の応答がある閾値 G_{th} より小さい場合, すなわちコントラストが低い場合, その位置のキーポイント候補点を除去する．加えて, 輪郭上のキーポイント候補点も窓問題を避けるために除去する[14]．この輪郭は DoR フィルタ処理画像にゼロ交差法を適用することで得られる．次に特徴ベクトル $\mathbf{F}=(F_1, F_2, \dots, F_{128})$ を勾配画像から計算する．この特徴ベクトルの記述については 5.2.2 節で述べた手法に従って実行される．デジタルコンピュータは入力画像から抽出したキーポイントの特徴ベクトル \mathbf{F} と予め用意したテンプレートキーポイントの特徴ベクトル \mathbf{T} とを比較することで対応点探索を実行する．キーポイントの類似度は式 (5.7) に示すユークリッド距離 E で評価した．

6.5 評価実験

本節では、提案した視覚システムの実時間キーポイント検出とトラッキングの性能の評価実験について述べる。実験では、図 6.6(a)に示すように、システム前方に配置した LCD によって知能視覚センサシステムへテスト動画を提示した。図 6.6(b)に示すように、テスト動画としてサイズが一定の速度で徐々に大きくなりながら右に動く交通標識の動画（全 300 フレーム）を用いた。撮像画像は水平方向に反転されるため、処理画像の交通標識は画像上で左方向へ動くこととなる。提案システムによってキーポイントの特徴ベクトルが実時間で抽出される。このキーポイント群は予め準備されたテンプレートのベクトルと比較され、その対応点を探索する。図 6.6(c)は 175 番目のフレームのキーポイント検出の結果を示す。座標 x と y は図 6.6(c)の左上に示す方向に定義された。灰色の四角形は画像で検出されたキーポイントの位置を示す。四角形の大きさがキーポイントのスケールの大きさと対応しており、例えばより大きな四角形はより大きなスケールでキーポイントが検出されたことを示す。

トラッキングの性能は、図 6.6(c)に示される自転車乗りの頭を中心付近で検出されたキーポイント A を使って評価された。図 6.6(d)はテンプレートキーポイントの 128 次元の特徴ベクトルを示す。キーポイントの類似度は上記で述べたユークリッド距離 E によって評価された。

自転車乗りの頭の位置に出るキーポイントをトラッキングした結果を図 6.7 に示す。0 番目のフレームにおけるキーポイント A の座標は(83, 73)であり、300 番目のフレームの最後の座標は(44, 51)であった。また頭の直径は 0 番目のフレームでの約 5 画素から 300 番目のフレームで約 9 画素に拡大した。図 6.7(a)-(e)は 50 番目、120 番目、150 番目、そして 300 番目のフレームでのキーポイント検出と特徴量マッチングの結果を示す。図 6.7(b)-(e)に示す白色の四角形が検出されたキーポイントを示す。図 6.7(a)-(e)の線は対応点間を接続している。図 6.7(b), (c), そして(e)で示すように、多くのフレームではキーポイント A が正確に検出され、正しく対応がとれている一方で、図 6.7(d)に示すように、まれに誤対応が生じた。

図 6.7(f)に全時刻での点 A と判断されたキーポイントの位置の変化を示す。点 A の位置はほとんどのフレームにおいて、正確にトラッキングされた。しかしながら、図 6.7(d)に示すように誤対応がいくつかのフレームにおいて生じた。図 6.7(g)に点 A と判断されたキーポイントのスケールの変化を示す。誤対応がいくつかのフレームにおいて生じているが、ほとんどのフレームにおいてスケールがパターンの拡大に応じて増加した。

誤対応の問題を軽減させるために、テンプレートとの対応が取れたキーポイント位置の時系列データに対してメディアンフィルタ処理を加えた。誤対応はまれに生じ、点の大半は正確な位置で対応がとれていたため、検出された点の中央値は対象が高速度で移動していない場合に限り、対象の近くであると期待される。フィルタへの入力ユークリッド距離を

用いて判断されたキーポイントの座標である。もしフレームに対応点がなければ、前のフレームからの結果が入力として用いられる。本実験で用いたメディアンフィルタは 7 フレームの中央値を出力とする。図 6.8 に示す結果から、メディアンフィルタにより誤対応を除去できたことが示された。

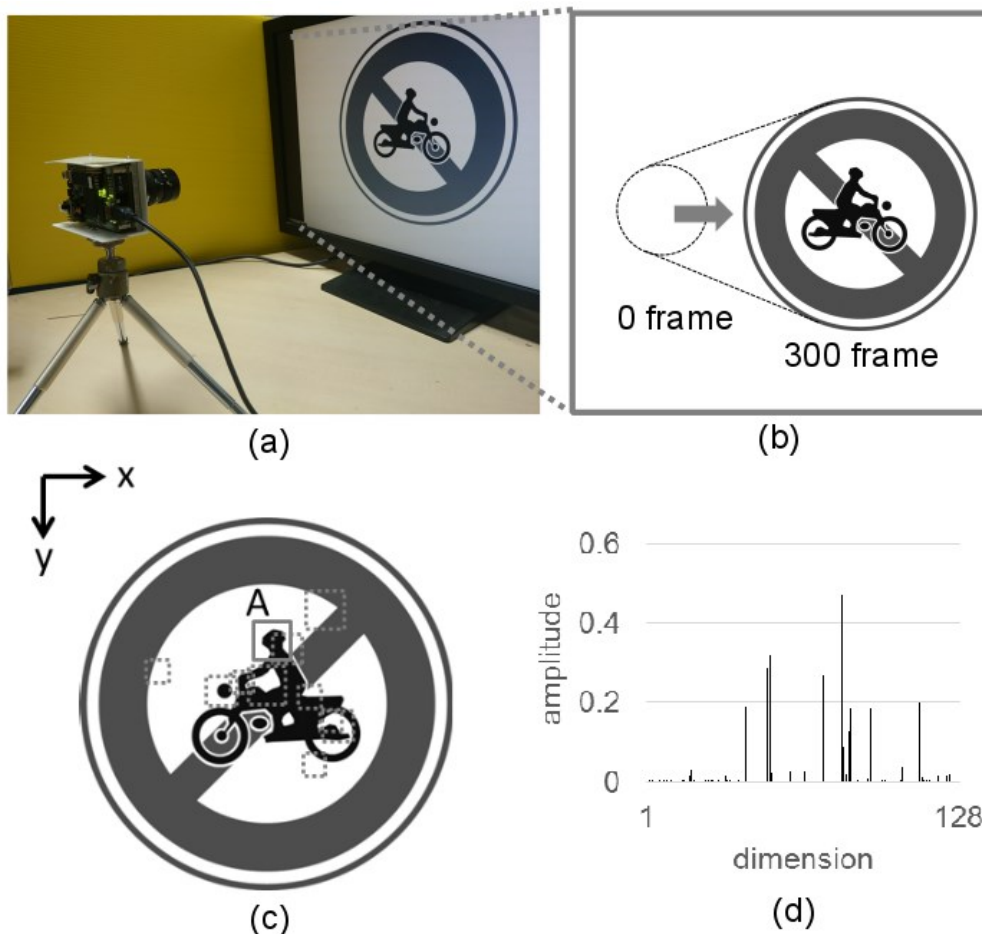


図 6.6 (a) 実験環境。(b) 提案した知能視覚システムの性能評を評価するために使用された入力動画のパターン。点線の円は測定を開始した 0 番目のフレームにおけるテストパターンの位置を表現する。(c) 175 番目のフレームのキーポイント検出の結果。灰色の四角形は画像内で検出されたキーポイントの位置を示す。より大きな四角形ほど大きなスケールの DoR フィルタ処理画像内で検出されえたキーポイントであることを示している。(d) キーポイント A のテンプレート特徴量。

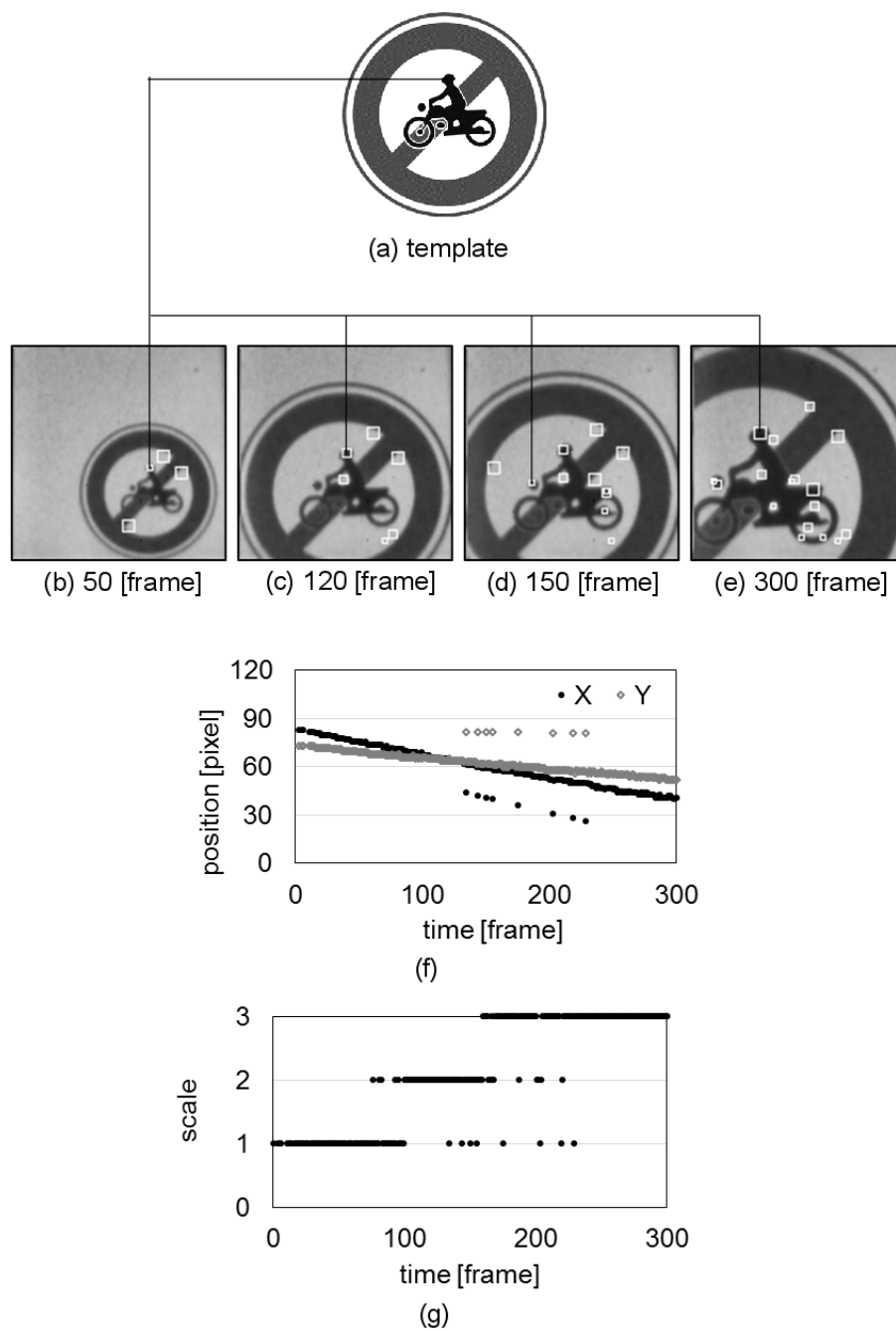


図 6.7 (a)-(e) 特徴点マッチングの結果. 白色の四角は画像内で検出されたキーポイントの位置を表現する. (a)と(b)-(e)を接続する線が対応点を指す. キーポイントのテンプレート特徴と対応が取れたキーポイントの(f) 位置と(g) スケール.

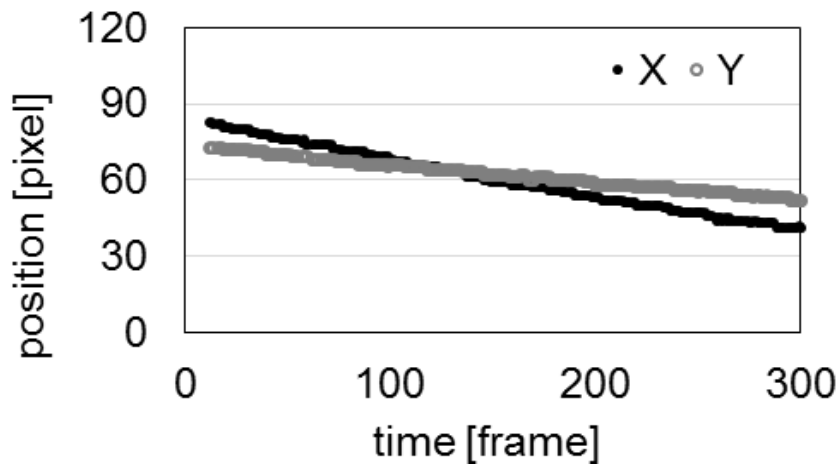


図 6.8 メディアンフィルタを適用後の対応がとれたキーポイントの位置の変化.

6.6 考察

6.6.1 アーキテクチャに対する考察

本章では画像からスケール不変特徴量を抽出し、実時間でこれらの特徴を連続的にトラッキングする智能視覚システム的设计と構築について述べた。提案した智能視覚システムはコンパクトでかつ低消費電力と実時間計算を達成するために、アナログ/デジタル混在型のアーキテクチャを有している。本智能視覚センサシステムは APS, S/H 回路, アナログ抵抗回路網, そして FPGA で構成されている。フルデジタルのアーキテクチャをもつ従来のシステムに対する本システムの利点は、前段のアナログ抵抗回路網による全画素同時演算と、次段の FPGA による特定用途に特化したデジタル並列演算である。本実験ではデスクトップ・コンピュータ (Intel(R) Core(TM) i7-4720HQ CPU, 2.60GHz) が FPGA で実行した以外のデジタル演算を実行するために用いられた。この処理にかかる計算量は近年のシングルボードコンピュータでも演算が実行できる範囲である。従って、全システムのサイズは更に小さく、消費電力は更に削減できる。

従来の逐次演算型の画像センシングシステムでは、実時間処理の実現のために 5.2.4 節で述べたようにソフトウェア・アルゴリズムの工夫が行われているが、基本的にはこれらの手法では抽出された特徴量の質と計算速度の間の関係はトレードオフである。一方で、我々の提案するアーキテクチャでは、計算コストの高い前段のマルチスケール空間フィルタ処理をアナログ専用回路に任せ、特徴記述を柔軟なデジタル演算器に担当させることによって、SIFT 特徴量のような高次元の特徴量の高速抽出を実現できる。

提案するアーキテクチャと、従来のアーキテクチャの計算時間を比較するために、我々はデスクトップ・コンピュータ (Intel(R) Core(TM) i7-4720HQ CPU, 2.60GHz) 上で動作する MATLAB を用いて、SIFT アルゴリズムにかかる計算時間を測定した。VGA (640 × 480

画素) サイズの入力画像から三つのスケールのキーポイントを検出するとして比較した. このコンピュータでは上記の条件で SIFT 演算の一部 (複数回フィルタ処理と勾配演算) のみで 42ms 要した. この時, 用いた空間フィルタの幅は, 図 6.1 で示す FWTM に抵抗回路網フィルタの幅に揃えた. この測定結果から, 現在のコンピュータでは, VGA 画像において本研究で用いたフレーム時間 (33 ms) 内に演算を完了することができないことが示唆された. 一方で, 提案するセンサシステムで採用されたアーキテクチャでは, 原理的に 128×128 画素でも VGA サイズでも要求される時間はほぼ変わらない. 単一フレーム内に検出できるキーポイントのスケールの最大数は画素読み出し速度に主に依存するが, この問題の解決方法は後述するように高速読み出し手法を採用することで解決できる.

次に示すシステムの限界を決める要素から, キーポイントのスケールの数は本研究では三つに限定されている. 一つ目の要素は抵抗回路網である. そこでは R_s と R_m の両方の抵抗の比を設定することでフィルタのスケールが決定されるが, 本研究ではシリコン網膜内の MOS 抵抗回路網の空間定数を, R_s の抵抗だけで制御している. このため制御できるスケールの範囲に限られる. R_m もまたスケールパラメータを決定する要素であるため, R_m の抵抗値を変更することで, 利用できるスケールが拡張できると考えられる. 二つ目の要素はシリコン網膜の読み出し時間である. μs オーダーの抵抗回路網の処理時間はミリ秒オーダーの画像の読み出し時間に比べて非常に短いため, 単一フレーム内でフィルタの最大数は読み出し時間によって主に決まる. カラム並列 A/D 変換器を採用することが, 高速画像読み出しについての実用的かつ効率的な手法である[79]. この技術により利用できるフィルタ処理画像の数を非常に増加させることができる.

抵抗回路網フィルタとガウシアンフィルタの両方が, 周辺で減衰していく同心円上の重み関数を有している. 従って, これらのフィルタは, ボックスフィルタを採用するような単純化した SIFT アルゴリズム[88]に比べ, より良いカットオフ特性をもつ. この差はキーポイント位置と特徴ベクトルの両方の正確さに影響する. 抵抗回路網を用いると, これらが単純化されたフィルタに比べ, より正確に計算される. 図 6.7 で示されるように, キーポイントの追従をする際, 誤対応が時折生じる. この問題への主な原因として, 固定パターンノイズのような回路ノイズが挙げられる. 本研究において用いられたシリコン網膜は, 十年程度前の古い半導体プロセスに基づいて開発されたため, シグナル・ノイズ比が最新版の技術のものより悪い. 最新技術でセンサを製造することはキーポイントの正確な位置決めと特徴ベクトルの生成への提案システムの能力が改善すると考えられる.

6.6.2 実環境への応用に対する考察

本章で行われた評価実験は接近する物体のシミュレーションである. 図 6.7 に示された実験の成功の結果は提案システムが, 例えば車の搭乗アシストシステムのような, 実環境への応用において実現可能性があることを示している.

提案したトラッキング能力は物体認識と組み合わせることで, その精度を更に向上でき

ると考えられる。SIFT 特徴量は物体認識において効果的ではあるが、それらは背景によって影響を受ける。この問題は物体認識に動物体のトラッキング結果を利用することで軽減されることが考えられる。これはトラッキングされている物体の背景が変化し続け、背景からの影響が平均化されるためである。

6.7 結言

本章では、画像からスケール不変特徴量を高速に抽出し、それらを実時間でトラッキングする知能視覚システムを設計し、構築したことについて述べた。我々のシステムはアナログ抵抗回路網と FPGA によるデジタル処理回路による、二つの並列処理手法の利点を活かした構成をとる。提案したシステムはスケール数の制限やアルゴリズムの簡略化による SIFT 特徴量の質の低下といった、いくつかの欠点を持つが、これらは基本的には考察したように、原理的には解決できると考えられる。

提案するアーキテクチャは CPU や GPU を用いた従来のシステムの場合に比べて、消費電力や処理時間において大きな利点がある。抵抗回路網フィルタと並列演算処理を実行するデジタル回路の組み合わせは実時間での視覚対象の探索や視覚対象トラッキングを実行するための有望なアーキテクチャである。

第7章 結論

視覚情報は外界の情報を多く含んでおり、これを効率よく計測・解析できる画像センシングシステムについて研究することは、これからの計測自動制御技術の発展において欠かせないことである。視覚情報は様々なスケールの構造が組み合わさって構成されており、生体の視覚系はその並列階層的な構造により、その視覚情報をマルチスケールで効率的に解析していると考えられている。現在、多くの画像センシングシステムで採用されている逐次演算型のコンピュータでは、画像をマルチスケールで解析する際に、初段で実行される複数回の空間フィルタ処理が不得意であり、その計算コストから実時間処理は困難である。本研究では、生体の視覚系のアーキテクチャに学ぶことで、効率よくマルチスケール空間フィルタ処理を実行する機能を有する知能視覚センサを開発した。この研究は、シングルスケールの空間フィルタ処理実行する機能を有する視覚センサであるシリコン網膜[35]の考え方を発展させることで遂行された。

本研究では、生体網膜の構造に学んだアナログ抵抗回路網による画素並列処理と、FPGAによる空間フィルタ処理の時分割制御法を活用することで、コンパクトな回路構成でありながら、実時間で効率的にマルチスケール空間フィルタ処理を実行できる知能視覚センサを開発した。本センサは小型・低消費電力で動作し、動作速度 50fps で三つの異なる帯域通過型空間フィルタ処理を画像に適用できる。実験結果から、粗い構造から細かい構造まで、それぞれ強調されたマルチスケール画像の取得ができることを示した。

本知能視覚センサの後段に FPGA による並列演算回路を追加することで、マルチスケール輪郭抽出機能を有する視覚システムを構築した。本知能視覚システムでは、センサからのマルチスケール画像の読み出しと FPGA 上での輪郭検出回路の演算が同時に実行されるため、処理に遅れやフレーム落ちが生じず、実時間で出力が得られる。

最後に、画像からスケール不変特徴量を高速に抽出し、それらを実時間でトラッキングをする知能視覚システムを開発した。本手法では従来の SIFT アルゴリズムで用いられているガウシアンフィルタの代わりに、瞬時に空間フィルタを適用できる抵抗回路網を採用した。この抵抗回路網フィルタによって、スケール不変特徴量が抽出できることをシミュレーションで確認した。その結果に基づき、アナログ抵抗回路網とデジタル回路の二つの並列処理手法の利点を活かしたアーキテクチャを考案し、ハードウェアを構築した。アーキテクチャレベルから見直した本知能視覚システムは、従来の画像センシングシステムの場合に比べて、消費電力や処理時間において大きな利点があることを示した。

本アーキテクチャ構成は、後段のデジタル演算部に拡張性を備えている。従って、本研究の開発した知能視覚システムは、本論文で提案したマルチスケール輪郭検出やスケール不変特徴量抽出といった応用に留まらず、FPGA 内に演算回路を追加することで、ステレオ

ビジョンの粗密法や、マルチスケール時空間フィルタ処理、顕著性マップ生成などの様々なマルチスケール解析機能を、単一フレーム内に実行できるように拡張でき、様々な応用に用いることができると期待される。

近年、労働の効率化、新しい資源や災害現場の探査や、より安全・快適な暮らしのために、ロボット工学分野や、IoT (Internet of Things) 分野における組み込みシステムに関する技術に期待が高まっている。その要素技術の中でも、外界の情報を多く抽出できる画像センシング技術の向上は、様々な分野の発展に大きな影響を与えられとされる。従来の画像センシングシステムは、そのシステム規模やエネルギー消費の点から、その多くが工場などの非常に限定された環境下で据え置き型として利用されてきた。一方で、本研究で提案した知能視覚センサは小型でかつ低消費電力で動作し、実時間性の要求が高い応用にも利用していくことができる。本研究での成果は、限定された環境に留まらず、第一次産業の現場から生活空間といった至る所まで画像センシング技術を行き届かせ、更に便利な高度情報化社会の発展に向けた第一歩であると考えている。

参考文献

- [1] Gollisch, T. and Meister, M. "Eye smarter than scientists believed: neural computations in circuits of the retina," *Neuron*, vol. 65, No. 2, pp. 150-164, 2010.
- [2] Marr, D. *Vision*, Freeman and Company, 1982.
- [3] Masland, R. H. "The fundamental plan of the retina," *Nature neuroscience*, Vol. 4, No. 9, pp. 877-886, 2001.
- [4] Wilson, H. R., and Bergen, J. R., "A four mechanism model for threshold spatial vision," *Vision research*, vol. 19, no. 1, pp. 19-32, 1979.
- [5] Itti, L., and Koch, C., "Computational modelling of visual attention," *Nature reviews neuroscience*, vol. 2, no. 3, pp. 194-203, 2001.
- [6] Riesenhuber, M., and Poggio, T., "Hierarchical models of object recognition in cortex," *Nature neuroscience*, vol. 2, no. 11, pp. 1019-1025, 1999.
- [7] Witkin, A. P., "Scale-space filtering: A new approach to multi-scale description," In *Acoustics, Speech, and Signal Processing, IEEE International Conference on ICASSP'84*, vol. 9, pp. 150-153, 1984.
- [8] Lindeberg, T., "Scale-space theory: A basic tool for analyzing structures at different scales," *Journal of applied statistics*, vol. 21 no. 1-2, pp. 225-270, 1994.
- [9] Koenderink, J. J., "The structure of images," *Biological cybernetics*, vol. 50, no. 5, pp. 363-370, 1984.
- [10] Field, D. J., "Relations between the statistics of natural images and the Response properties of cortical cells," *Journal of the Optical Society of America. A, Optics and image science*, vol. 4, no. 12, pp. 2379-2394, 1987.
- [11] Marr, D., and Hildreth, E., "Theory of edge detection," *Proceedings of the Royal Society of London. Series B. Biological Sciences*, vol. 207, no. 1167, pp. 187-217, 1980.
- [12] Marr, D., and Poggio, T., "A computational theory of human stereo vision," *Proceedings of the Royal Society of London B: Biological Sciences*, vol. 204, no. 1156, pp. 301-328, 1979.
- [13] Lowe, D. G., "Object recognition from local scale-invariant features," In *Computer vision, 1999, The proceedings of the seventh IEEE international conference on*, vol. 2, pp. 1150-1157, 1999.
- [14] Lowe, D. G., "Distinctive image features from scale-invariant keypoints," *International journal of computer vision*, vol. 60, no. 2, pp. 91-110, 2004.

- [15] Jones, D. G., and Malik, J., “A computational framework for determining stereo correspondence from a set of linear spatial filters,” In *Computer Vision—ECCV'92*, Springer Berlin Heidelberg, pp. 395-410, 1992.
- [16] Lindeberg, T., “Feature detection with automatic scale selection,” *International journal of computer vision*, vol. 30, no. 2, pp. 79-116, 1998.
- [17] Amiaz, T., Lubetzky, E., and Kiryati, N., “Coarse to over-fine optical flow estimation,” *Pattern recognition*, vol. 40, no. 9, pp. 2496-2503, 2007.
- [18] Li, R., and Sclaroff, S., “Multi-scale 3D scene flow from binocular stereo sequences,” *Computer vision and image understanding*, vol. 110, no. 1, pp. 75-90, 2008.
- [19] Duanggate, C., Uyyanonvara, B., Makhanov, S. S., Barman, S., and Williamson, T., “Object detection with feature stability over scale space,” *Journal of Visual Communication and Image Representation*, vol. 22, no. 4, 345-352, 2011.
- [20] Mégret, R., and Jolion, J. M., “Tracking scale-space blobs for video description,” *IEEE MultiMedia*, vol. 9, no. 2, pp. 34-43, 2002.
- [21] Mokhtarian, F., and Suomela, R., “Robust image corner detection through curvature scale space,” *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 20, no. 12, pp. 1376-1381, 1998.
- [22] Thrun, S., Montemerlo, M., Dahlkamp, H., Stavens, D., Aron, A., Diebel, J. Fong, P, Gale J, Halpenny, M. and Hoffmann, G, “Stanley: The robot that won the DARPA Grand Challenge. *Journal of field Robotics*,” vol. 23, no. 9, 661-692, 2006.
- [23] Liu, S. C., and Delbruck, T., “Neuromorphic sensory systems,” *Current opinion in neurobiology*, vol. 20, no. 3, pp. 288-295, 2010.
- [24] 米本和也, 「CCD/CMOS イメージ・センサの基礎と応用: 原理, 構造, 動作方式, 諸特性からシステム概要まで」, CQ 出版, 2003.
- [25] Mead, C, *Analog VLSI and Neural Systems*, Addison-Wesley, 1989.
- [26] Moini, A, *Vision chips*, Kluwer Academic Publishers, 2000.
- [27] Liu, S. C., Delbruck, T., Kramer, J., Indiveri, G., and Douglas, R, *Analog VLSI: Circuits and Principles*, MIT press, 2002.
- [28] Liu, S. C., Delbruck, T., Indiveri, G., Whatley, A., and Douglas, R. *Event-Based Neuromorphic Systems*, John Wiley & Sons, 2015.
- [29] Mead, C. A., and Mahowald, M. A., “A silicon model of early visual processing,” *Neural networks*, vol. 1, no. 1, pp. 91–97, 1988.
- [30] Boahen, K. A., and Andreou, A. G., “A contrast sensitive silicon retina with reciprocal synapses,” *Advances in Neural Information Processing Systems (NIPS)*, vol. 4, pp. 764-772, 1992.

- [31] Matsumoto, T., Shimmi, T., Kobayashi, H., Abidi, A., Yagi, T., and Sawaji, T., "A second order regularization vision chip for smoothing-contrast enhancement," Proc IJCNN Beijing, pp. 188-197, 1992.
- [32] Koch, C., and Mathur, B., "Neuromorphic vision chips," *IEEE Spectrum*, vol. 33. No. 5, pp. 38-46, 1996.
- [33] Wu, C. Y., and Chiu, C. F., "A new structure of the 2-D silicon retina," *Solid-State Circuits, IEEE Journal of*, vol. 30, no. 8, pp. 890-897, 1995.
- [34] Park, D. S., Kim, J. H., Kim, H. S., Park, J. H., Shin, J. K., and Lee, M., "A foveated-structure CMOS retina chip for edge detection with local light adaptation," *Sensors and Actuators A: Physical*, vol. 108, no. 1, pp. 75-80, 2003.
- [35] Kameda, S., and Yagi, T., "An analog VLSI chip emulating sustained and transient response channels of the vertebrate retina," *Neural Networks, IEEE Transactions on*, vol. 14, no. 5, pp. 1405-1412, 2003.
- [36] Shimonomura, K., Kameda, S., Iwata, A., and Yagi, T., "Wide-dynamic-range APS-based silicon retina with brightness constancy," *Neural Networks, IEEE Transactions on*, vol. 22, no. 9, pp. 1482-1493, 2011.
- [37] Naka, K. I., and Rushton, W. A. H., "The generation and spread of S - potentials in fish (Cyprinidae)," *The Journal of physiology*, vol. 192, no. 2, pp. 437-461, 1967.
- [38] Kaneko, A., "Receptive field organization of bipolar and amacrine cells in the goldfish retina," *The Journal of physiology*, vol. 235, no. 1, pp. 133-153, 1973.
- [39] Lamb, T. D., "Spatial properties of horizontal cell responses in the turtle retina", *The Journal of physiology*, vol. 263, no. 2, pp. 239-255, 1976.
- [40] Lamb, T. D., and Simon, E. J., "The relation between intercellular coupling and electrical noise in turtle photoreceptors," *The Journal of physiology*, vol. 263, no. 2, pp. 257-286, 1976.
- [41] Yagi, T., Arika, Y. and Funahashi, Y. "Dynamic model of dual layer neural network for vertebrate retina," In *Neural Networks, 1989. IJCNN.*, International Joint Conference on, pp. 787-789, 1989.
- [42] Yagi, T., Ohshima, S., and Funahashi, Y., "The role of retinal bipolar cell in early vision: an implication with analogue networks and regularization theory," *Biological cybernetics*, vol. 77, no. 3, pp. 163-171, 1997.
- [43] Okuno, H., and Yagi, T., "Image sensor system with bio-inspired efficient coding and adaptation," *IEEE transactions on biomedical circuits and systems*, vol. 6, no. 4, pp. 375-384, 2012.
- [44] Shimonomura, K., and Yagi, T. "Neuromorphic VLSI vision system for real-time

- texture segregation,” *Neural Networks*, vol. 21, no. 8, pp. 1197-1204, 2008.
- [45] Shimonomura, K., Kushima, T., and Yagi, T., “Binocular robot vision emulating disparity computation in the primary visual cortex,” *Neural Networks*, vol. 21, no. 2, pp. 331-340, 2008.
- [46] Shimonomura, K., and Yagi, T., “Vision sensor with brightness constancy: towards neuromorphic color vision,” *The Neuromorphic Engineer*, 2008.
- [47] Lichtsteiner, P., Posch, C., and Delbruck, T., “A 128×128 120 dB 15 μ s latency asynchronous temporal contrast vision sensor,” *IEEE journal of solid-state circuits*, vol. 43, no. 2, pp. 566-576, 2008.
- [48] 中司雅之, 下ノ村和弘, 八木哲也, “視覚モジュールを階層的に統合した脳型両眼ロボットビジョン,” 電子情報通信学会技術研究報告, NC, ニューロコンピューティング, vol. 107, no. 542, pp. 463-466, 2008.
- [49] 松村隆史, 武田真季, 奥野弘嗣, 山田憲嗣, 大野ゆう子, 八木哲也, “網膜チップセンサーを用いた病院内人物モニタシステムの開発,” 電子情報通信学会技術研究報告, MBE, ME とバイオサイバネティックス, vol. 110, no. 460, pp. 141-145, 2011.
- [50] Fu, Z., Culurciello, E., Lichtsteiner, P., and Delbruck, T., “Fall detection using an address-event temporal contrast vision sensor,” In 2008 IEEE International Symposium on Circuits and Systems, pp. 424-427, 2008.
- [51] Conradt, J., Cook, M., Berner, R., Lichtsteiner, P., Douglas, R. J., and Delbruck, T., “A pencil balancing robot using a pair of AER dynamic vision sensors,” In 2009 IEEE International Symposium on Circuits and Systems, pp. 781-784, 2009.
- [52] Delbruck, T., and Lang, M., “Robotic goalie with 3 ms reaction time at 4% CPU load using event-based dynamic vision sensor,” *Frontiers in Neuroscience*, vol. 7, no. 223, 2013.
- [53] Fehervari, T., Matsuoka, M., Okuno, H., and Yagi, T., “Real-time simulation of phosphene images evoked by electrical stimulation of the visual cortex,” In International Conference on Neural Information Processing, Springer Berlin Heidelberg, pp. 171-178, 2010.
- [54] Sugimoto, K., and Kamata, S. I., “Efficient Constant-time Gaussian Filtering with Sliding DCT/DST-5 and Dual-domain Error Minimization,” *ITE Transactions on Media Technology and Applications*, vol. 3, no. 1, pp.12-21, 2015.
- [55] Wells, W. M., “Efficient synthesis of Gaussian filters by cascaded uniform filters,” *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 8, no. 2, pp. 234-239, 1986.
- [56] Robinson, J., “Efficient gaussian filtering using cascaded prefix sums,” In Image Processing (ICIP), 2012 19th IEEE International Conference on, pp. 117-120, 2012.

- [57] Deriche, R., "Fast algorithms for low-level vision," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 12, no. 1, pp. 78-87, 1990.
- [58] Farnebäck, G., and Westin, C. F., "Improving Deriche-style recursive Gaussian filters," *Journal of Mathematical Imaging and Vision*, vol. 26, no. 3, pp. 293-299, 2006.
- [59] Unser, M., Aldroubi, A., and Eden, M., "B-spline signal processing. II. Efficiency design and applications," *Signal Processing, IEEE Transactions on*, vol. 41, no. 2, pp. 834-848, 1993.
- [60] Elboher, E., and Werman, M., "Cosine integral images for fast spatial and range filtering," In *Image Processing (ICIP), 2011 18th IEEE International Conference on*, pp. 89-92, 2011.
- [61] Sugimoto, K., and Kamata, S. I., "Fast Gaussian filter with second-order shift property of DCT-5," In *Image Processing (ICIP), 2013 20th IEEE International Conference on*, pp. 514-518, 2013.
- [62] Sinha, S. N., Frahm, J. M., Pollefeys, M., and Genc, Y., "GPU-based video feature tracking and matching," In *EDGE, Workshop on Edge Computing Using New Commodity Architectures*, vol. 278, 2006.
- [63] Bonato, V., Marques, E., and Constantinides, G., "A parallel hardware architecture for scale and rotation invariant feature detection," *Circuits and Systems for Video Technology, IEEE Transactions on*, vol. 18, no. 12, pp. 1703-1712, 2008.
- [64] van der Wal, G. S., and Sinniger, J. O., "Real time pyramid transform architecture", In *1985 Cambridge Symposium, International Society for Optics and Photonics*, pp. 300-305, 1985.
- [65] Lee, S., Oh, J., Park, J., Kwon, J., Kim, M., and Yoo, H.-J., "635 A 345 mw heterogeneous many-core processor with an intelligent inference engine for robust object recognition," *Solid-State Circuits, IEEE Journal of*, vol. 46, no. 1, pp. 42-51, 2011.
- [66] Kim, Y., Park, J., and Yoo, H. J., "A 34.1 fps scale-space processor with two-dimensional cache for real-time object recognition" In *Circuits and Systems (ISCAS), 2013 IEEE International Symposium on*, pp. 689-692, 2013.
- [67] Zeinolabedin, S. M. A., Karimi, N., and Samavi, S., "Low computational complexity hardware implementation of Laplacian Pyramid", In *Electrical Engineering (ICEE), 2010 18th Iranian Conference on*, pp. 465-470, 2010.
- [68] Pauwels, K., Tomasi, M., Diaz Alonso, J., Ros, E., and Van Hulle, M. M., "A comparison of FPGA and GPU for real-time phase-based optical flow, stereo, and local image features," *Computers, IEEE Transactions on*, vol. 61, no. 7, pp. 999-1012, 2012.
- [69] Cope, B., Cheung, P. Y., Luk, W., and Howes, L., "Performance comparison of

- graphics processors to reconfigurable logic: A case study,” *Computers, IEEE Transactions on*, vol. 59, no. 4, pp. 433-448, 2010.
- [70] Jones, J. P., and Palmer, L. A., “An evaluation of the two-dimensional Gabor filter model of simple receptive fields in cat striate cortex,” *Journal of neurophysiology*, vol. 58, no. 6, pp. 1233-1258, 1987.
- [71] Poggio, T., and Koch, C., “Ill-posed problems in early vision: from computational theory to analogue networks,” *Proceedings of the Royal Society of London B: Biological Sciences*, vol. 226, no. 1244, pp. 303-323, 1985.
- [72] Mead, C., “Neuromorphic electronic systems,” *Proceedings of the IEEE*, vol. 78, no. 10, pp. 1629-1636, 1990.
- [73] Fernandez-Berni, J., Carmona-Galan, R., and Rodríguez-Vázquez, Á., “FLIP-Q: A QCIF resolution focal-plane array for low-power image processing,” In *Low-Power Smart Imagers for Vision-Enabled Sensor Networks*, Springer New York, pp. 67-109, 2012.
- [74] 森本 昌介, 亀田 成司, 岩田 穆, “スイッチトレジスタ回路を内蔵したビジョンチップの開発,” 信学技法, IEICE Technical Report, NC, 2007.
- [75] Shibano, T., Lizuka, K., Miyamoto, M., Osaka, M., Miyama, R., and Kito, A., “Matched filter for DS-CDMA of up to 50 MChip/s based on sampled analog signal processing,” In *Solid-State Circuits Conference, 1997. Digest of Technical Papers. 43rd ISSCC., 1997 IEEE International*, pp. 100-101, 1997.
- [76] B.Razavi., “Design of Analog CMOS Integrated Circuits,” McGraw-Hill Higher Education, 2003.
- [77] Yasukawa, S., Okuno, H., Kameda, S., and Yagi, T., “High-speed multiple spatial band-pass filtering using a resistive network,” In *Mechatronics and Automation (ICMA), 2013 IEEE International Conference on*, pp. 1299-1304, 2013.
- [78] Yasukawa, S., Okuno, H., Kameda, S., and Yagi, T., “A vision sensor system with a real-time multi-scale filtering function,” *International Journal of Mechatronics and Automation*, vol. 4, no. 4, pp. 248-258, 2014.
- [79] Zhou, Z., Pain, B., and Fossum, E. R., “CMOS active pixel sensor with on-chip successive approximation analog-to-digital converter,” *Electron Devices, IEEE Transactions on*, vol. 44, no. 10, pp. 1759-1763, 1997.
- [80] Marr, David, and Ellen Hildreth. “Theory of edge detection.” *Proceedings of the Royal Society of London, series B, Biological Sciences*, vol 207, no. 1167, pp. 187-217, 1980.
- [81] 井上恵介, 亀田成司, 八木哲也., “シリコン網膜と FPGA を用いた実時間並列画像処理.” *The Journal of The Institute of Image Information and Television Engineers*,

- vol. 61, no. 3, pp. 316-324, 2007.
- [82] Yasukawa, S., Okuno, H., and Yagi, T., "Detection of scale-invariant key points employing a resistive network," In System Integration (SII), 2012 IEEE/SICE International Symposium on, pp. 877-882, 2012.
- [83] Ihara, A., Fujiyoshi, H., Takagi, M., Kumon, H., and Tamatsu, Y., "Improved matching accuracy in traffic sign recognition by using different feature subspaces," In *MVA*, pp. 130–133, 2009.
- [84] Csurka, G., Dance, C., Fan, L., Willamowski, J., and Bray, C., "Visual categorization with bags of keypoints," In Workshop on statistical learning in computer vision, ECCV , vol. 1, pp. 1-2, 2004.
- [85] Tsuduki, Y., and Fujiyoshi, H., "A method for visualizing pedestrian traffic flow using sift feature point tracking," In Advances in Image and Video Technology, Springer Berlin Heidelberg, pp. 25-36, 2009.
- [86] Se, S., Lowe, D., and Little, J., "Global localization using distinctive visual features," In Intelligent Robots and Systems, 2002. IEEE/RSJ International Conference on, vol. 1, pp. 226-231, 2002.
- [87] Brown, M., and Lowe, D. G., "Automatic panoramic image stitching using invariant features," *International journal of computer vision*, vol. 74, no. 1, pp. 59-73, 2007.
- [88] Grabner, M., Grabner, H., and Bischof, H., "Fast approximated SIFT," In Computer Vision—ACCV 2006, Springer Berlin Heidelberg, pp. 918-927, 2006.
- [89] Bay, H., Ess, A., Tuytelaars, T., and Van Gool, L. "Speeded-up robust features (SURF)," *Computer vision and image understanding*, vol. 110, no. 3, pp. 346-359, 2008.
- [90] Leutenegger, S., Chli, M., and Siegwart, R. Y., "BRISK: Binary robust invariant scalable keypoints," In Computer Vision (ICCV), 2011 IEEE International Conference on, pp. 2548-2555, 2011.
- [91] Rublee, E., Rabaud, V., Konolige, K., and Bradski, G., "ORB: an efficient alternative to SIFT or SURF," In Computer Vision (ICCV), 2011 IEEE International Conference on, pp. 2564-2571, 2011.
- [92] Schmid, C., and Mohr, R., "Local grayvalue invariants for image retrieval," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 19, no. 5, 530-534, 1997.
- [93] Brown, M., and Lowe, D. (2002). "Invariant Features from Interest Point Groups," In BMVC 2002: 13th British Machine Vision Conference, pp. 253-262, 2002.
- [94] Ke, Y., and Sukthankar, R., "PCA-SIFT: A more distinctive representation for

- local image descriptors,” In Computer Vision and Pattern Recognition, 2004. CVPR 2004. Proceedings of the 2004 IEEE Computer Society Conference on, vol. 2, pp. 506-513, 2004.
- [95] Mikolajczyk, K., and Schmid, C., “An affine invariant interest point detector,” In Computer Vision—ECCV 2002, Springer Berlin Heidelberg, pp. 128-142, 2002.
- [96] Mikolajczyk, K., and Schmid, C., “A performance evaluation of local descriptors,” *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 27, no. 10, pp. 1615-1630, 2005.
- [97] Tuytelaars, T., and Mikolajczyk, K., “Local invariant feature detectors: a survey. *Foundations and Trends® in Computer Graphics and Vision*, vol. 3, no. 3, pp. 177-280, 2008.
- [98] Huang, F. C., Huang, S. Y., Ker, J. W., and Chen, Y. C., “High-performance SIFT hardware accelerator for real-time image feature extraction,” *Circuits and Systems for Video Technology, IEEE Transactions on*, vol. 22, no. 3, pp. 340-351, 2012.
- [99] Su, Y. C., Huang, K. Y., Chen, T. W., Tsai, Y. M., Chien, S. Y., and Chen, L. G., “A 52 mW full HD 160-degree object viewpoint recognition SoC with visual vocabulary processor for wearable vision applications,” *Solid-State Circuits, IEEE Journal of*, vol. 47, no. 4, pp. 797-809, 2012.
- [100] Yasukawa, S., Okuno, H., Ishii, K., and Yagi, T., “Real-time object tracking based on scale-invariant features employing bio-inspired hardware,” *Neural Networks*, vol. 81, pp. 29-38, 2016.

謝辞

本研究は大阪大学大学院工学研究科電気電子情報工学専攻 八木哲也教授の御指導の下に行われたものであり、本研究を遂行するにあたり、終始懇切なる御指導と御鞭撻を賜りましたことに厚く御礼申し上げます。

本論文の作成にあたり、懇切なる御助言と御指導を頂きました大阪大学大学院工学研究科 森伸也教授、金子真教授に厚く御礼申し上げます。

本研究を遂行するにあたり、日々懇切なる御助言とご意見を頂くと共に、様々な面で御協力を頂きました奥野弘嗣助教（現・大阪工業大学講師）に深く感謝致します。

本論文の作成にあたり、貴重な御意見、御教示を頂きました大阪大学大学院工学研究科伊藤利道教授、森勇介教授、片山光浩教授、尾崎雅則教授、片山竜二教授、栖原敏明教授、近藤正彦教授に厚く御礼申し上げます。

本研究を遂行するにあたり、幾多の御指導と御意見を頂きました林田祐樹准教授、亀田成司特任准教授、石井和男教授（九州工業大学）、下ノ村和弘准教授（立命館大学）、小山内実准教授（東北大学）に厚く御礼申し上げます。

本研究を遂行するにあたり、幾多の御協力を頂くと共に、有意義な議論をして頂きました井上恵介氏、長谷川潤氏、故岡崎祐香氏をはじめとする八木研究室の研究員、大学院生、学部生、卒業生の皆様に厚く御礼申し上げます。特に、共に研究室内の回路班で研究を行った早川隆久氏と眞田忠氏には、感謝の意を表すとともに、今までの研究成果が論文という形になった喜びを伝えたいと思います。また、事務面で多大なご支援を頂きました小寺昌子女史、本多理乃女史、服部訓子女史に深く感謝致します。

最後に、学問に励む機会、環境をはじめ多くを支えて頂いた父と母、そして工学の道に進む切っ掛けを与えて頂いた兄に心より感謝致します。

業績一覧

学術雑誌投稿論文

- (1) S. Yasukawa, H. Okuno, K. Ishii, and T. Yagi,
“Real-time Object Tracking Based on Scale-invariant Features Employing Bio-inspired Hardware,”
Neural Networks, vol. 81, pp. 29-38, 2016.
- (2) S. Yasukawa, H. Okuno, S. Kameda, and T. Yagi,
“A Vision Sensor System with a Real-time Multi-scale Filtering Function,”
International Journal of Mechatronics and Automation, vol. 4, No. 4, pp. 248-258, 2014.

国際学会発表論文（査読付）

- (1) S. Yasukawa, H. Okuno, and T. Yagi,
“Detection of Scale-Invariant Key Points Employing a Resistive Network,”
in Proceedings of System Integration (SII), 2012 IEEE/SICE International Symposium on,
pp.877-882, 2012.
- (2) S. Yasukawa, H. Okuno, S. Kameda, and T. Yagi,
“High-Speed Multiple Spatial Band-Pass Filtering using a Resistive Network,”
in Proceedings of Mechatronics and Automation (ICMA), 2013 IEEE International Conference
on, pp.1299-1304, 2013.

国際会議発表（著者登壇）

- (1) S. Yasukawa, H. Okuno, and T. Yagi,
“Detection of Scale-Invariant Key Points Employing a Resistive Network,”
System Integration (SII), 2012 IEEE/SICE International Symposium on,
Fukuoka, Japan, 16-18, Dec., 2012.

- (2) S. Yasukawa, H. Okuno, S. Kameda, and T. Yagi,
“High-Speed Multiple Spatial Band-Pass Filtering using a Resistive Network,”
Mechatronics and Automation (ICMA), 2013 IEEE International Conference on,
Kagawa, Japan, 4-7, Aug., 2013.

国内学会発表（著者登壇）

- (1) 安川真輔, 奥野弘嗣, 八木哲也,
“輪郭強調フィルタの順応機能を有する智能視覚センサの開発”,
日本機械学会ロボティクス・メカトロニクス講演会 2010,
旭川 , 2010年6月13-16日.
- (2) 安川真輔, 奥野弘嗣, 八木哲也,
“複数の帯域通過型空間フィルタを有する視覚センサ”,
ニューロコンピューティング研究会 2011,
町田, 2011年5月7-9日.