T2R2 東京科学大学 リサーチリポジトリ
Science Tokyo Research Repository

論文 / 著書情報
Article / Book Information

題目(和文)	PDEに基づく画像処理を用いた低ビットレート映像符号化の前処理に関する研究
Title(English)	
著者(和文)	辻裕之
Author(English)	
出典(和文)	学位:博士（工学）, 学位授与機関:東京工業大学, 報告番号:甲第6566号, 授与年月日:2006年3月26日, 学位の種別:課程博士, 審査員:中嶋正之
Citation(English)	Degree:Doctor of Engineering, Conferring organization: Tokyo Institute of Technology, Report number:甲第6566号, Conferred date:2006/3/26, Degree Type:Course doctor, Examiner:
学位種別(和文)	博士論文
Type(English)	Doctoral Thesis

Powered by T2R2 (Science Tokyo Research Repository)

PDEに基づく画像処理を用いた低ビットレート映像符号化の前処理に関する研究

指導教官　中嶋 正之 教授

提出者　東京工業大学大学院
情報理工学研究科 計算工学専攻
博士課程
辻　裕 之

2006年2月28日

目 次

第 1 章　序論 **1**
1.1 背景 I：動画像符号化技術の現状 1
1.1.1 動画像符号化標準の現在 1
1.1.2 処理技術の観点から見た動画像符号化研究の方向性 2
1.2 背景 II：画像処理技術の現状 4
1.2.1 画像処理の抱える問題点 4
1.2.2 画像処理の新しい潮流 5
1.2.3 PDE に基づく手法と動画像符号化 5
1.3 目的 6
1.4 本論文の構成 8

第 2 章　PDE に基づく画像処理 **11**
2.1 はじめに 11
2.2 PDE に基づく画像処理の概要 11
2.2.1 スケールスペース画像表現 11
2.2.2 画像修復 14
2.2.3 セグメンテーション 16
2.2.4 オブジェクト・トラッキング 23
2.2.5 画像編集 25
2.3 非等方拡散法 28
2.3.1 Perona-Malik 拡散法 (PMD) 28
2.3.2 エッジ停止関数の正規化 31
2.3.3 PMD の正則化 33
2.4 幾何学的な動的輪郭法 34
2.4.1 測地線に基づく動的輪郭法 (GAC) 34
2.4.2 レベルセット法による定式化 36
2.5 PMD と GAC の数理的な関係 38
2.6 おわりに 40

第 3 章　ローレベル符号化における前処理技術 —プレフィルタを中心に— **41**
3.1 はじめに 41

3.2 ピクチャスキップ抑制のためのプレフィルタ制御法 42
3.2.1 背景と目的 42
3.2.2 フィルタ制御の原理 42
3.3 シミュレーションによる検証 46
3.3.1 シミュレーション結果 46
3.3.2 考察及び今後の課題 47
3.4 おわりに 48

第 4 章 非等方拡散法に基づく空間平滑化プレフィルタの構築 50
4.1 はじめに 50
4.2 非等方拡散法による平滑化の挙動とその問題点 51
4.3 PMD に基づくプレフィルタの構築 56
4.3.1 準備 56
4.3.2 理論的考察 58
4.3.3 エッジ停止関数の導出 59
4.4 シミュレーションによる検証 61
4.4.1 ノイズ除去性能の検証 61
4.4.2 エッジ保存性とエッジ強調効果の抑制 64
4.4.3 演算量の削減 66
4.4.4 画像情報量の削減 67
4.5 おわりに 68

第 5 章 時空間平滑化プレフィルタへの拡張 78
5.1 はじめに 78
5.2 時空間非等方拡散フィルタ 79
5.2.1 非等方拡散法 79
5.2.2 時空間フィルタへの拡張 80
5.2.3 時間方向の拡散処理に関する検討事項 82
5.3 従来の時空間平滑化フィルタとの関係 82
5.3.1 空間及び時間平滑化フィルタの定式化 82
5.3.2 空間方向の非等方拡散法に関する理論的考察 85
5.3.3 時間方向の非等方拡散に関する理論的考察 87
5.4 時間方向のエッジ停止関数に関する検討 89
5.4.1 候補となるエッジ停止関数 89
5.4.2 TDS 関数の挙動に関する考察 90
5.4.3 主観画質への影響 92
5.5 まとめ 92

第 6 章　ミッドレベル符号化における前処理技術 —VOP 生成手法を中心に—　94
6.1 はじめに 94
6.2 VOP 自動生成技術の概要 95
6.2.1 共通フレームワーク 95
6.2.2 従来研究 96
6.3 エッジベースの VOP 自動生成（Meier の方法） 96
6.3.1 アルゴリズムの概要 96
6.3.2 初期オブジェクトモデルの抽出 (Stage 1) 98
6.3.3 エッジ画像の生成 (Stage 2) 99
6.3.4 エッジ情報に基づくオブジェクト・トラッキング (Stage 3) 100
6.4 エッジ情報に基づくオブジェクト抽出 101
6.4.1 関連研究 102
6.4.2 Meier らのアプローチ 102
6.5 おわりに 105

第 7 章　時間発展方程式に基づく輪郭線抽出 (I) – 輝度情報を用いる方法　106
7.1 はじめに 106
7.2 提案手法 107
7.2.1 GAC に関する設定 107
7.2.2 アルゴリズムの概要 108
7.2.3 擬似コード 109
7.3 シミュレーションによる検証 113
7.3.1 オブジェクトの抽出過程 113
7.3.2 従来手法との比較 115
7.4 おわりに 116

第 8 章　時間発展方程式に基づく輪郭線抽出 (II) – 2 値エッジ情報を用いる方法　119
8.1 はじめに 119
8.2 輪郭線の存在可能性を表す場の生成 120
8.2.1 アルゴリズムの概要 120
8.2.2 接触円による補間 121
8.2.3 存在確率の算出 121
8.3 存在可能性の場を用いた閉輪郭線の時間発展処理 124
8.3.1 測地線による動的輪郭モデルの修正 124
8.3.2 レベルセット法による定式化 124
8.4 シミュレーションによる検証 (1) 125
8.5 接ベクトル導出手法の改善 126
8.5.1 従来法の問題点 126
8.5.2 改善手法の提案 127

8.6 シミュレーションによる検証 (2) 128
8.6.1 基本図形による実験 128
8.6.2 標準画像による実験 129
8.7 おわりに 130

第 9 章 結論 **132**
9.1 まとめ 132
9.2 今後の課題 — 動画像符号化への応用 134

参 考 文 献 **136**

付 録 A 命題の証明 **149**
A.1 PMD の安定性に関する命題 (第 2 章) 149

図 目 次

1.1 画像符号化標準の技術的分類 2
1.2 本論文で検討する前処理技術の位置づけ 7
1.3 本論文の構成 10

2.1 スケールスペース 13
2.2 Image Inpainting による画像修復の例 [8] 16
2.3 動的輪郭法によるオブジェクト抽出例 16
2.4 曲線 $\mathbf{c}(q)$ と法線ベクトル $\mathbf{N}$ 18
2.5 レベルセット法による曲線の時間発展 [77] 19
2.6 平均曲率流方程式による閉曲線の時間発展 20
2.7 勾配を用いない動的輪郭法の例 [15] 22
2.8 Geodesic Active Tracking によるトラッキングの例 [66] 24
2.9 指標ベクトル場に基づく画像補間 [69] 25
2.10 Seamless Cloning の適用例 [69] 26
2.11 Trimap[81] 27
2.12 Poisson Matting による画像の合成例 [81] 28
2.13 典型的なエッジ停止関数のプロット 29
2.14 空間 4-近傍 30
2.15 PMD による平滑化処理 ($T = 15$, $\epsilon = 0.1$, $iter = 20$) 31
2.16 エッジ停止関数の正規化 32
2.17 PMD によるノイズ除去効果: (左から) 原画像，ノイズ重畳画像 (salt&pepper noise, ノイズ密度 0.1)，PMD(g_1) による平滑化結果，正則化 PMD(g_1) による平滑化結果 (いずれも，$T = 15$, $\epsilon = 0.1$, $iter = 20$) 33
2.18 エッジ検出関数 g のプロット 35
2.19 単一輪郭から複数オブジェクトを抽出した例 39
2.20 複数輪郭を単一オブジェクトに統合した例 39

3.1 送信側 VBV バッファモデル 43
3.2 プレフィルタの制御フロー 45
3.3 プレフィルタパラメータの変動 47
3.4 バッファレベルの変動 48

4.1 テスト画像 51
4.2 ２項ガウシアンマスク 51
4.3 PMD (with $g^{BS}(x)$) による第 120 ラインの推移（上段から $T = 16, 32, 64(\epsilon = 0.1)$，各段の左から $iter = 5, 10, 20$) 53
4.4 PMD (with $g^{PM}(x)$) による第 120 ラインの推移（上段から $T = 16, 32, 64(\epsilon = 0.1)$，各段の左から $iter = 5, 10, 20$) 54
4.5 正則化 PMD (with $g^{PM}(x)$) による第 120 ラインの推移（上段から $T = 4, 8, 12$，各段の左から $iter = 5, 10, 20$) 55
4.6 導出したエッジ停止関数の比較 61
4.7 ノイズ除去に伴う SNR の変動 62
4.8 最適 SNR 時の画像 63
4.9 平滑化画像における第 120 ライン波形の比較（左図：正則化従来法，右図：提案法） 64
4.10 PMD (with $g^{Prop}(x)$) による第 120 ラインの推移（上段から $T = 16, 32, 48$, 各段の左から $iter = 5, 10, 20$) 65
4.11 処理時間の比較 66
4.12 I-VOP における平均信号電力（分散値）の低減効果 70
4.13 I-VOP におけるエントロピーの低減効果 71
4.14 平均信号電力（分散値），エントロピーの低減効果 72
4.15 エントロピー 1bit/pel 削減時の主観画像品質の比較 (Flower Garden) 73
4.16 エントロピー 1bit/pel 削減時の主観画像品質の比較 (Mobile & Calendar) . . 74
4.17 エントロピー 1bit/pel 削減時の主観画像品質の比較 (Susie) 75
4.18 エントロピー 1bit/pel 削減時の主観画像品質の比較 (Diva with noise) 76
4.19 エントロピー 1bit/pel 削減時の主観画像品質の比較 (Football) 77

5.1 時空間 6-近傍 81
5.2 従来の時空間平滑化フィルタの構成 83
5.3 従来法の TDS 関数 84
5.4 係数 $\alpha(x)$ のプロット 84
5.5 候補となるエッジ停止関数 90
5.6 各エッジ停止関数に対応する Ψ 関数 90
5.7 g^{tp} による時間拡散の TDS 関数 91
5.8 時間平滑化による主観画像品質の比較 93

6.1 VOP 自動生成の共通フレームワーク [34] 95
6.2 VOP 自動生成のブロック図 97
6.3 初期オブジェクトモデルの抽出 (Stage 1): (a) 差分検出によるマスク (31^{st} - 32^{nd} frames), (b) マスクの細線化, (c) 短エッジの除去 (しきい値 8 pixels). . 98
6.4 背景エッジのフィルタリング (Stage 2) 99

6.5 Hausdorff マッチングに基づくオブジェクトモデルの更新 (Stage 3) 100
6.6 2 値エッジ画像からのオブジェクト輪郭抽出の問題 101
6.7 単純な画素補填を用いたオブジェクト抽出例 102
6.8 境界線修正法による修正例 [49] 103
6.9 コストの設定 [49] 104

7.1 単純な画素補填を用いたオブジェクト抽出例 109
7.2 初期輪郭線 (a) とこれに対応する符号付き距離関数 (b) 110
7.3 提案法によるオブジェクト輪郭の抽出過程 114
7.4 オブジェクト輪郭線の比較. (a),(e),(i) オブジェクトモデル from “Hall Monitor” (32^{nd}, 50^{th}, 70^{th} frames). (b),(f),(j) 単純画素補填による輪郭抽出結果. (c),(g),(k) 境界線修正法による輪郭抽出結果. (図左: オリジナル出力, 図右: 枝除去後の最終結果). (d),(h),(l) GAC に基づく提案法による輪郭抽出結果. . 118

8.1 接触円による補間 121
8.2 補間により求めた接線方向の分布 122
8.3 欠けた円弧に対する「存在可能性を表す場」の導出 ($h_{init} = 3$, $N_{min} = 7$) . . 123
8.4 オブジェクトモデルに対する「存在可能性を表す場」の導出 ($h_{init} = 3$, $N_{min} = 5$) 126
8.5 修正 GAC によるオブジェクト輪郭の抽出過程 127
8.6 P における接ベクトルの導出 128
8.7 基本図形のエッジ画像 129
8.8 存在可能性の場の導出結果（上段：従来法，下段：提案法） 130
8.9 輪郭の推定結果（上段：従来法，下段：提案法） 131
8.10 人物の輪郭推定（左から，エッジ画像，存在可能性の場，輪郭推定結果） . . 131

表 目 次

3.1 テストシーケンス 46
3.2 シミュレーション諸元 46
3.3 ピクチャスキップ発生回数 (fps の実効値) 47

4.1 SNR ピーク値と繰り返し回数 62

6.1 記号の定義 (図 6.2) 96

第1章　序論

1.1　背景 I ：動画像符号化技術の現状

本節では，本論文の背景となる動画像符号化技術の現状について述べるとともに，使用される処理技術の観点から動画像符号化手法の分類を試み，本論文における研究の方向性について述べる．

1.1.1　動画像符号化標準の現在

動画像圧縮符号化は，1990 年に ISDN(Integrated Services Digital Network) における TV 電話・会議サービスをターゲットとした H.261 が標準化されて以来，主に通信分野をターゲットとする ITU-T(International Telecommunication Union Telecommunication standardization sector) と放送・蓄積分野をターゲットとする ISO/IEC(International Organization for Standardization/International Electrotechnical Commission) とが相互補完しながら標準化が進められ，これに呼応する形で活発な研究・開発が行われてきた．また，これらの標準化技術は実用面においても大きな成功を収めており，現在でも DVD やディジタル放送には H.262/MPEG-2 が，インターネットやモバイルにおける動画像の通信や配信には H.263/MPEG-4 が，それぞれ不可欠な画像圧縮技術として日常的に使用されている．これらの後継となる符号化標準として，現時点で最新の規格は，2003 年に ITU-T と ISO/IEC の合同で標準化された H.264/AVC である．各世代の符号化標準を年代順に並べたものを図 1.1 に示す.

H.264/AVC は，従来の H.263/MPEG-4 に比べおよそ 2 倍の圧縮率を実現するといわれ，符号化効率の点で大幅な向上を実現している [57]．その主な要因としては，(1) 符号化モードの選択肢を大幅に拡大し，従来よりも精度の高い予測画像を生成可能としたこと，(2) 高度な可変長符号化を導入し，符号化シンボルの圧縮効率を高めたこと等が挙げられる．これらは近年のコンピュータの高性能化に加えて，技術的なブレイクスルーによっても支えられている．特に重要な技術は，Wiegand[80] による画期的なレート－歪みの最適化方式である．この技術の登場により，膨大な符号化モードの組み合わせの中から，最も符号量を低減する最適な組み合わせを選択する方法論が確立され，現在の H.264/AVC における符号化効率の向上につながっている．

しかし一方で，圧縮符号化の原理という観点から見ると，核となる技術は依然として「動き補償＋ DCT」のハイブリッド符号化に可変長符号化を組み合わせたものであり，これは

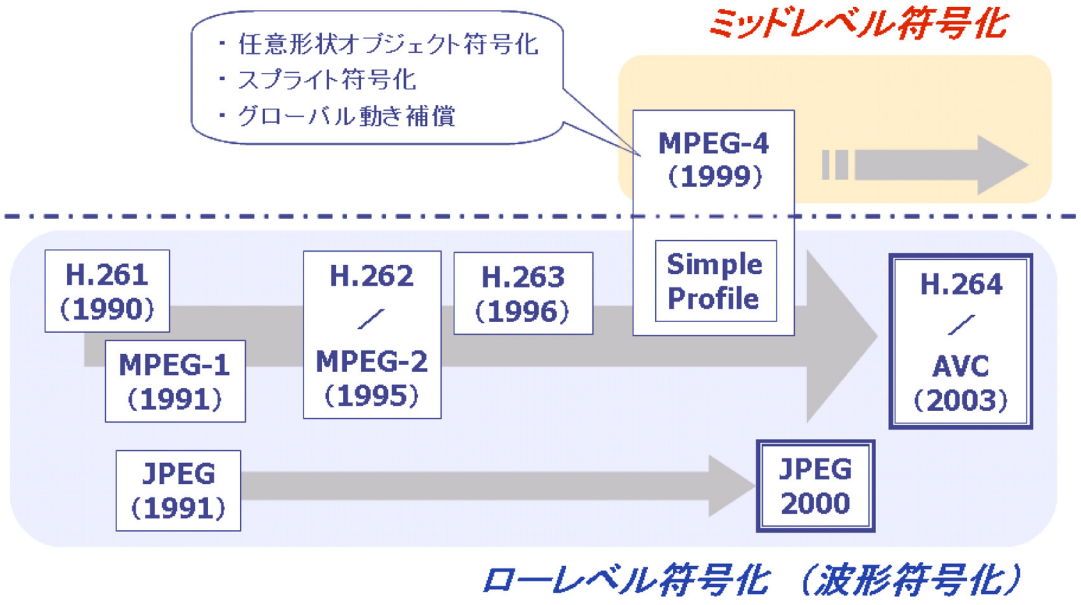


図 1.1: 画像符号化標準の技術的分類

H.261 から H.264/AVC まで一貫して変わっていない．各世代の符号化標準の差分は，主として多様なサービス要求に応えるために必要な符号化ツールを追加したものに過ぎないともいえる．その意味において，H.264/AVC における符号化効率の向上は，新しい圧縮符号化の原理に基づくものではなく，近年の高速・大容量なコンピュータを前提とする徹底した最適化技術のもとで，動き補償と離散コサイン変換 (DCT) を組み合わせた従来のハイブリッド符号化の性能を最大限に引き出すことによって実現したものである．すなわち，H.264/AVC は H.261 以降今日まで培われてきた「動き補償＋ DCT」ハイブリッド符号化のいわば最終形であり，その完成度は非常に高い．このため，今後の研究においては，異なる原理に基づく圧縮符号化を導入しない限り，これ以上の圧縮効率の改善は困難である．

1.1.2　処理技術の観点から見た動画像符号化研究の方向性

次に，これまで標準化されてきた符号化技術を処理技術の点から整理し，本論文における動画像符号化研究の方向性について考える．

動画像符号化手法は，使用されている処理技術の点から，ローレベル符号化，ミッドレベル符号化，ハイレベル符号化の 3 つに分類することができる [74]．ローレベル符号化とは，画像を 2 次元の信号波形と考え，その波形の様子を特徴付けるパラメータを抽出して符号化するディジタル信号処理的な手法を指す．したがって，予測符号化，直交変換符号化，サブバンド符号化，ウェーブレット変換符号化などのフィルタリングベースの波形符号化はすべてここに含まれる．次に，ミッドレベル符号化とは，画像を 3 次元被写体のディスプレイ平

面上への 2 次元投影像と考え，その投影像が有する属性の記述を画像特徴パラメータとして抽出し，符号化するものである．例えば，画像をディスプレイ面上で同一の動きを持った複数の領域に分割するレイヤ表現 [93] や 3 次元情報を扱えるようこれを拡張したセッティング表現 [89] を用いた動きセグメンテーションに基づく画像符号化手法はミッドレベル符号化に含まれる．最後に，ハイレベル符号化とは，被写体の 3 次元モデルをあらかじめ与えておき，その被写体を特徴付ける意味論的な画像構造パラメータを抽出して符号化する手法を指す．例えば人間の顔画像をその 3 次元モデルの特徴パラメータに基づき符号化するモデルベース符号化などはここに含まれる．

以上の分類に基づき，これまで標準化されてきた符号化技術を整理すると，図 1.1 に示すように，そのほとんどがローレベル符号化の範疇に含まれることがわかる．これは，標準化の性質上，汎用性と実用化のしやすさを優先しながら圧縮効率を高める工夫がなされてきたことによるもので，その結果として，H シリーズや MPEG の標準化技術が広く社会に受け入れられるようになったと評価できる．しかしながら，従来符号化の最終形ともいうべき H.264/AVC が標準化された現時点では，ローレベル符号化を志向する動画像符号化の研究における次なるブレイクスルーは決して容易ではない．このような状況における今後のローレベル符号化に関する研究として，低ビットレート動画像の主観画質の改善技術が挙げられる．従来符号化による圧縮効率が頭打ちとなり，にもかかわらずその画質がまだ不十分であるとすれば，低ビットレート動画像に符号化歪みは付き物という前提に立って，当面はそれをいかにきれいに見せるかという方向からの研究もまた重要である．これを実現するには，人間の視覚にとって重要な情報とそうでない情報を区別し，選択的に切り替えながら処理を行う高度な画像解析技術が必要となる．

一方，MPEG-4(Part2)[21, 22] で提示された別の方向性もまた，今後の動画像符号化の研究を考える上で興味深い．MPEG-4 のコア・プロファイルには任意形状オブジェクト符号化が採用され，VOP(Video Object Plane) と呼ばれる複数のレイヤに画像を分割して符号化するフレームワークが提供されている．また，メインプロファイルでは，動画像中に複数フレームにわたって共通して登場する背景を合成し，予めスプライトと呼ばれる平面オブジェクトを生成するスプライト符号化が提供されている．これらはいずれも画像中のオブジェクトをベースとし，従来の波形符号化とは一線を画す技術である．このように，動画像をオブジェクト単位で構造化しようというチャレンジングな試みは，上の分類でいえばミッドレベル符号化に相当する．しかしながら，これらのオブジェクト符号化ツールを含む上位プロファイルが実用の場において使用されることは稀である．例えば，インターネットやモバイルネットワークで，通常，MPEG-4 として使用されるのは，波形符号化ツールのみを用いたシンプル・プロファイルのエンコーダである．その主な理由の一つとして，自然画像から背景と前景オブジェクトを分離する等の高度な動画像解析技術がまだ未成熟である点が挙げられる．現時点では，いかなる入力画像に対しても安定して動作する汎用的な VOP の自動生成技術はまだ存在しておらず，このことが MPEG-4 の任意形状オブジェクト符号化の普及を阻む要因となっている．

このように，従来は主として画像処理やコンピュータビジョンの分野で検討されていた高

度な画像解析技術が，動画像符号化のような比較的実用的な分野においても求められつつあるというのが，近年における現状である．一般に，多様な自然画像に対してこのような処理を安定して行う汎用的な手法の確立は未だに困難な問題であり，今後，これを実現するための，新しいコンセプトに基づく画像処理技術の研究が望まれる．

1.2　背景 II：画像処理技術の現状

本節では，現在の画像処理の問題点について述べた後，本研究のもうひとつの背景となる，近年の画像処理技術に関する新しい潮流について述べる．

1.2.1　画像処理の抱える問題点

ミッドレベル符号化に要求される動画像解析を実現するためには，画像処理技術の高度化が不可欠である．画像処理は 1960 年代から続く長い歴史を持ち，フィルタ理論，スペクトル解析，確率・統計理論をベースとした有益な手法が数多く提案されてきた．例えば文献 [25, 36] などを見れば，これまでに提案されてきた定番の画像処理ツールの成果を知ることができる．現在の画像処理手法の多くは，このような蓄積の上で，先人たちの成果をどのようにうまく組み合わせ・アレンジすれば目的とする処理を効率的に実現できるか，という点に重点が置かれている．そして，これらの手法の多くはヒューリスティックスに基づく要素を含んでおり，ある特定の事例にはうまくいくが，あらゆる事例において安定に動作するかどうかは定かではない．

長尾は，このような現在の画像処理のあり方に対して次のような指摘をしている [53]：

(1) 処理目的・処理対象が限定されており，他の目的・対象にはほとんど役に立たないことが多く，汎用性が著しく乏しい．

(2) これまでに多数の処理アルゴリズムが提案されているものの，それらを総括的に論じる理論・方法論がない．

(3) これまでに実現されている画像処理は比較的容易なもので，人間が行うような知的な画像処理は未だ実現されていない．

これらの問題のうち特に第 1 項目を解決するために，“進化的画像処理”という興味深い方法論が提案されている [53]．進化的画像処理とは，画像処理に精通した研究者が既知の画像処理ツールの数々から有効なものを選び組み合わせて，目的の処理を構築するプロセスを最適化問題として捉えることにより，一般に多大な労力と時間を要する画像処理を自動構築する強力な方法論である．最適化問題を解く過程で，遺伝的アルゴリズムに代表される進化的手法が使用されるため，このように呼ばれる．自動構築したアルゴリズムが汎用性を持てば，究極的には第 3 項目を満たすようなシステムの構築につながっていく可能性がある．

進化的画像処理は工学的に魅力的なアプローチであるが，いくつかの問題点も指摘できる．例えば [53] で取り上げられているフィルタの自動構築の例でいえば，目的の処理を実現するために必要となる基本フィルタセットをどのように選ぶかという点が明らかでなく，結局，経験に伴うヒューリスティックスが必要となる．さらに，ある基本セットを用意した場合，目的とする処理がそのフィルタの組み合わせでどこまで正確に実現できるかについて，理論的な解析を行うことが困難である．また，最終的に得られたフィルタにどの程度の汎化能力があるかについても，理論的に明らかにすることはやはり困難である．

このように，多くの画像処理手法においては，その汎用性を理論的に保証することが困難であり，より多くの事例による実験によって信頼性を高めるしか手段がないのが現状である．

1.2.2　画像処理の新しい潮流

一方，近年では，20 世紀の現代数学を理論的な基盤とする，より洗練された画像処理の理論が提案されるようになった．Aubert らはこれまで提案されている 3 つの主要な方向として，確率場モデル，ウェーブレット解析，偏微分方程式（PDE: Partial Differential Equations）に基づく手法を挙げている [4].

確率場モデルは，1984 年に Geman 兄弟が提案したマルコフ確率場 (MRF: Markov Random Field) をベースとした画像理論であって，ノイズにより著しく劣化した画像の修復に関して著しい成果を挙げている [23]．ウェーブレット解析は，1980 年代後半に Meyer と Mallat によって多重解像度分解の理論が提案されて以来，信号処理や画像処理をはじめとする様々な応用分野で利用されている [28]．とりわけ，静止画像圧縮標準の最新規格である JPEG2000 にウェーブレット変換が取り入れられたことは記憶に新しい [58].

PDE に基づく画像処理はこれらよりやや新しく，1990 年に Perona らによって提案された非等方拡散法 [70] が注目を集めたことをきっかけに，活発な研究が行われるようになった．特に，数学分野でフィールズ賞を獲得した David Munford と Pierre-Louis Lions が相次いでコンピュータビジョンの分野に参入し，優れた成果を残しつつある現状を受け，最近では多くの応用数学者による高度な画像処理研究が見られるようになった．これにより，画像処理の理論を数理的に構築しようという新しい流れが，今後ますます活性化するものと予想される．

1.2.3　PDE に基づく手法と動画像符号化

本論文では，上で述べた画像処理の 3 つ潮流のなかでも，特に PDE に基づく方法 (PDE based method) に着目する．低ビットレート動画像符号化に求められる前処理技術との関連性から，以下にその理由を述べる．

ローレベル符号化と PDE

先に，ローレベル符号化の今後の研究課題として低ビットレート動画像の主観画質の改善技術を挙げた．動画像符号化で画質改善のためによく使われる前処理の一例として，平滑化プレフィルタがある．入力画像を予め平滑化することにより発生符号量を抑え，符号化歪みの発生を緩和することが目的である．低ビットレート環境では，それに見合った符号量を実現するために，かなり強い平滑化フィルタをかける場合がある．このようなとき，通常の線形フィルタではエッジなどの視覚的に重要な情報が失われてしまう．このため，動画像符号化の前処理で使用するフィルタは「視覚的に重要な画像特徴の保存性」を持っていることが望ましい．PDE に基づく手法は，エッジをはじめとする画像の特徴を保存しながら他の領域を平滑化するような非線形スケールスペースを容易に生成することができる．これはすなわち，PDE に基づく手法を利用すれば，様々な画像の特徴を保存する複雑な平滑化フィルタを，統一的な見地からより簡単に構築できることを意味している．

ミッドレベル符号化と PDE

先に，ミッドレベル符号化の今後の研究課題としてオブジェクトをベースとする符号化技術を挙げた．オブジェクトを単位とする動画像の構造化を実現する上で基礎となるのは，自然画像中のオブジェクト領域を人の手を介さず自動的に切り出すセグメンテーション技術である．コンピュータビジョンでは，セグメンテーションの問題をしばしば変分法による関数の最適化の問題として定式化される．このように変分法によって提示された問題は，Euler-Lagrange 方程式を通して PDE の問題に帰着することができる [20]．これは，セグメンテーションに限らず，画像修復や動き推定など，変分法によって提示される画像処理の問題は少なくない．また，必ずしも変分法を起源としない PDE に基づく画像処理を考慮に入れることにより，より広範な応用範囲への適用が期待できる．

以上述べたことは，ローレベル符号化とミッドレベル符号化の主要な問題が，PDE に基づく画像処理の枠組みに基づいて，統一的に扱える可能性を示唆している．特に PDE に基づく手法が，ミッドレベル符号化に必要なオブジェクト単位の画像処理へと展開できるポテンシャルを持っている点は，今後の動画像符号化の進展を考える上で，注目に値する．

1.3　目的

本論文では，以上の背景を踏まえ，近年の画像処理の成果である PDE に基づく手法を利用して，MPEG-4 をはじめとする低ビットレート動画像符号化の前処理技術の高度化について検討することを目的とする．本論文で扱う前処理技術とは，大きく分けて，ローレベル符号化に必要となるプレフィルタ技術とミッドレベル符号化に必要となる VOP 自動生成技術の 2 種類である．これらの前処理技術の動画像符号化における位置づけを図 1.2 に示す．

まず，ローレベル符号化に必要となる「エッジ保存平滑化プレフィルタ」について述べる．

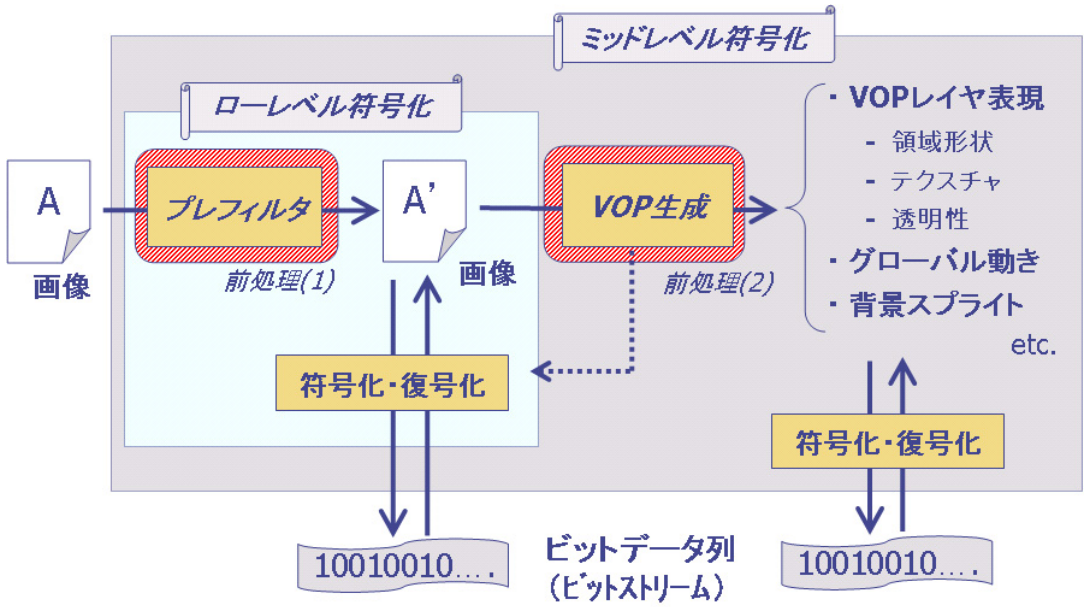


図 1.2: 本論文で検討する前処理技術の位置づけ

近年では，インターネットやモバイル・ネットワークを通して，H.263 や MPEG-4 シンプル・プロファイル相当のエンコーダで符号化された低ビットレートの映像が当たり前のように送受信されている．一般に，そのような帯域の保証されない低レート環境で映像を符号化する場合，ブロック歪み，モスキート，ピクチャスキップなど，ディジタル符号化特有の歪みが顕著に現れ，その品質は必ずしも十分とはいえない．平滑化プレフィルタには，このような環境において，符号化する動画像の情報量を予め削減し，復号化後の画像に現れる符号化歪みを緩和し，主観画像品質を高める効果があることが知られている．しかしながら，情報量を削減するために強い平滑化フィルタをかけてしまうと，オブジェクトに関わるエッジ等の視覚的に重要な情報が失われ，かえって主観画質を落としてしまうことにもなりかねない．このような状況を回避するには，エッジ等の視覚的に重要な特徴は保存しながら，それ以外の領域を優先的に平滑化することができる，高度な非線形フィルタを導入する必要がある．本論文のひとつの目的は，PDE に基づく画像処理手法を利用して，このようなエッジ保存平滑化プレフィルタを構築するとともに，その情報量削減効果について検証することである．なお，ここでは，Perona-Malik の非等方拡散法 [70] に基づいて得られる非線形スケールスペースを利用したフィルタ構築を行う．

次に，ミッドレベル符号化に必要となる「エッジ情報に基づく VOP 自動生成手法」について述べる．MPEG-4 の任意形状オブジェクト符号化が，実用上，ほとんど使用されていないことは既に説明した通りである．これを実現するには，任意の動画像から背景と前景オブジェクトを分離し，それぞれ個別の VOP としてレイヤ表現するという高度なハードルをクリアしなくてはならない．ここで必要となるのは，背景とは別に動いているオブジェクトを検出・追跡するオブジェクト追跡技術と，その形状を VOP として正確に抽出するオブ

ジェクト抽出技術である．本論文では，前者に対しては，Meier ら [48] によって提案されたエッジベースのオブジェクト追跡手法を前提として用いることとし，主としてそれに続くオブジェクト抽出手法について検討を行う．具体的には，Meier の方法によって得られた，2 値エッジ画像による不完全なオブジェクトモデルをもとに，正確なオブジェクト輪郭を抽出する手法を提案することが，本論文のもうひとつの目的である．なお，ここでは，PDE に従って閉輪郭線を時間発展させる，Geodesic Active Contours と呼ばれる動的輪郭モデル [12] を利用して，オブジェクト輪郭の抽出を実現する．

1.4　本論文の構成

本論文は 9 章から成り立っており，大きく 3 つのパートに分かれる．PDE に基づく画像処理の関連研究についてまとめた第 2 章，ローレベル符号化の前処理技術としてプレフィルタの検討を行った第 3 章〜第 5 章，ミッドレベル符号化の前処理技術として VOP 自動生成法の検討を行った第 6 章〜第 8 章である．図 1.3 に本論文の章構成を示す．以下，本章に続く各章の概要について述べる．

第 2 章では，PDE に基づく画像処理の研究動向について述べる．まず，PDE を用いた画像処理がどのような分野で成果を挙げているかを示すために，画像修復，セグメンテーション，オブジェクト・トラッキング，画像編集などの応用事例を取り上げ，その概要を説明する．次に，PDE を動画像符号化の前処理へ適用するにあたり次章以降で必要となる，非等方拡散法 (PMD: Perona-Malik Diffusion) と測地線による動的輪郭法 (GAC: Geodesic Active Contours) についてはやや詳しく説明する．最後に，PMD と GAC の数理的な関係性について述べる．

第 3 章では，ローレベル符号化として分類される MPEG-4 シンプル・プロファイルや H.263 などの低ビットレート符号化におけるプレフィルタが，低ビットレート動画像の画質を改善する上で非常に重要であることを示すために，著者が提案する VBV バッファに基づくプレフィルタ制御法について説明する．その結果を受け，入力画像に含まれるエッジなどの重要な特徴を保存して主観的な画質劣化を極力抑えながら，従来より多くの情報量を削減可能とする非線形な平滑化プレフィルタの必要性ついて問題提起を行う．

続く第 4 章〜第 5 章では，第 3 章で提起された問題を解決するため，視覚的に重要なエッジ情報をなるべく保存しながら画像を平滑化する非線形平滑化フィルタを，PDE を用いて構築する方法について述べる．本提案手法は，Perona-Malik による非等方拡散法を技術的な基盤としている．このようなエッジ保存形のフィルタは，特に低ビットレート映像符号化のプレフィルタとして有効と考えられる．まず，第 4 章では，非等方拡散法の平滑化効果を決定する上で重要となるエッジ停止関数の導出方法に関する理論的な考察を行う．Perona-Malik らの提案したエッジ停止関数について，空間平滑化フィルタを構築する上で問題となる点を述べ，これらを改善する新たなエッジ停止関数を理論的に導出し，その改善効果をシミュレーションにより確認する．第 5 章では，4 章で構築した空間平滑化フィルタを，時空間平滑化フィルタとして拡張する方法について述べる．一般に動画像符号化の前処理においては，

空間方向に加えて時間方向の平滑化を行うことにより，時間方向の冗長度をさらに削減できる効果がある．ここでは，第 4 章で提案した非等方拡散フィルタに時間拡散項を追加し，時空間 3 次元に広がる画像列に対して拡散処理を行うことにより時空間平滑化フィルタを実現する．さらに，提案法である時空間非等方拡散フィルタが，画像符号化において従来から使用されている非線形平滑化フィルタの自然な拡張として理解できることを，理論的な考察から明らかにする．

第 6 章では，MPEG-4 の任意形状オブジェクト符号化の前処理として必要となる VOP 自動生成手法ついて述べる．特にエッジ情報に基づく Meier ら [48] らのアルゴリズムを取り上げ，その概要を示す．さらに，2 値エッジ画像で表現されたオブジェクトモデルを基に正確なオブジェクト輪郭を抽出する問題を提示するとともに，その関連研究について述べる．

続く第 7 章〜第 8 章では，6 章で提起された問題を解決するため，2 値エッジ画像として与えられるオブジェクトモデルから，閉じたオブジェクト輪郭線を抽出する方法について検討している．基本的には，2 章で述べた GAC の方法に従い，初期輪郭線を PDE に基づいて時間発展させることによって目的のオブジェクト輪郭へと収束させるという共通の手法を採るが，エッジディテクタ関数の導出の仕方に応じて，各章でそれぞれ異なる 2 通りの手法を提示する．第 7 章では，輝度情報に基づくエッジディテクタを用いており，オリジナルの GAC をそのまま適用する方法を検討している．これに対し，8 章では，輝度情報を補助として用いることなしに，2 値エッジ画像のみからオブジェクト輪郭線の存在可能性を表す場を導出し，これをエッジディテクタの代わりに使用する方法について検討している．

最後に，第 9 章では，本論文のまとめと今後の課題について述べる．

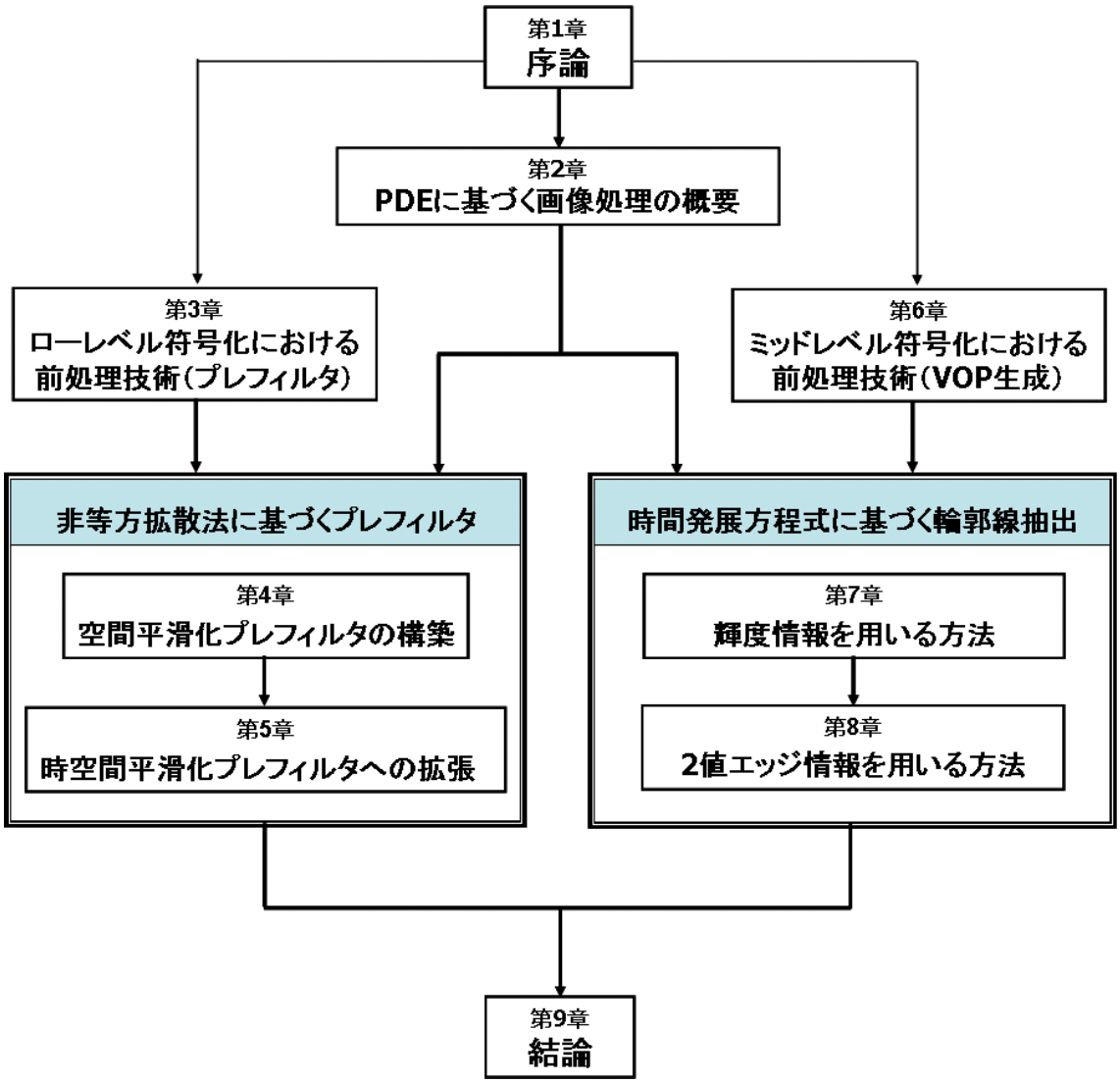

第1章
序論
第2章
PDEに基づく画像処理の概要
第3章
ローレベル符号化における
前処理技術（プレフィルタ）
第6章
ミッドレベル符号化における
前処理技術（VOP生成）
非等方拡散法に基づくプレフィルタ
第4章
空間平滑化プレフィルタの構築
第5章
時空間平滑化プレフィルタへの拡張
時間発展方程式に基づく輪郭線抽出
第7章
輝度情報を用いる方法
第8章
2値エッジ情報を用いる方法
第9章
結論

図 1.3: 本論文の構成

第2章　PDEに基づく画像処理

2.1　はじめに

本章では，従来研究の事例を通して，PDE に基づく画像処理の現状について述べる．また，本論文において，PDE に基づく手法を動画像符号化の前処理へ適用するにあたり，以降で特に重要な役割を果たす Perona-Malik の非等方拡散法（Anisotropic Diffusion）[70] と Caselles-Kimmel-Sapro の測地線による動的輪郭法（Geodesic Ative Contours）[12] については，技術的な内容を含めて詳細に説明する．

2.2　PDEに基づく画像処理の概要

本節では，PDE に基づく画像処理の従来研究として，スケールスペース画像表現，画像修復，セグメンテーション，オブジェクト・トラッキング，画像編集などへの応用事例について述べる．

2.2.1　スケールスペース画像表現

以下に，線形及び非線形のスケールスペース画像表現と PDE との数理的な関係を明らかにした従来研究の流れについて述べる．

線形スケールスペース [95], [44]

矩形領域 $\mathbf{\Omega} \subset \mathbf{R}^2$ で定義されたグレイスケール画像を I_0 とする．すなわち，I_0 は Ω から $\mathbf{R}$ への関数 $I_0 : \mathbf{x} = (x, y)^\top \in \Omega \mapsto I_0(x, y) \in \mathbf{R}$ である．分散 σ_τ を時間 $\tau \geq 0$ とともに単調に増加するパラメータとする．このとき，原画像 I_0 をガウシアンカーネル G_{σ_τ}

$$G_{\sigma_\tau}(x, y) = e^{-\frac{x^2+y^2}{2{\sigma_\tau}^2}} \tag{2.1}$$

によって平滑化した画像 I_{σ_τ} は，次の式で得られる．

$$I_{\sigma_\tau}(x, y) = G_{\sigma_\tau}(x, y) * I_0(x, y) \tag{2.2}$$

ここで，$G_{\sigma_\tau} * I_0$ は，関数 G_{σ_τ} と I の畳み込み演算を表す．平滑化画像 I_{σ_τ} を $\tau \in \mathbf{R}^+$ の関数と考えるとき，時間とともに増加する分散値 σ_τ に対し，各時刻の平滑化画像 I_{σ_τ} を時間

軸 τ 方向に連続につなげて得られる画像列 $I:(\tau,\mathbf{x})\mapsto I_{\sigma_\tau}(\mathbf{x})\in\mathbf{R}$ を線形スケールスペースと呼ぶ [95].

特に σ と τ の関係を $\sigma_\tau=\sqrt{2\tau}$ として定めるとき，式 (2.1) のガウシアンカーネルは，

$$G_\tau(x,y)=e^{-\frac{x^2+y^2}{4\tau}} \tag{2.3}$$

となり，これに基づく平滑化画像を I_τ と記述するとき，スケールスペース $I:(\tau,x,y)\mapsto I_\tau(x,y)\in\mathbf{R}$ の時系列変化の様子は，次の熱拡散方程式に従うことが知られている [44].

$$\begin{aligned}\frac{\partial}{\partial\tau}I(\tau,x,y) &= div\Big(\nabla I(\tau,x,y)\Big)\\ &= \Big(\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}\Big)I(\tau,x,y).\end{aligned} \tag{2.4}$$

このことは，原画像 I_0 を初期条件として熱拡散方程式 (2.4) を解くことにより，ガウシアンフィルタと同等な平滑化効果が得られることを示している．

非線形スケールスペース [70]

Perona-Malik は，熱拡散方程式とガウシアン平滑化フィルタとの同値性に着想を得て，式 (2.4) の PDE に，勾配の大きくなる方向への平滑化を抑制するエッジ停止関数に基づく重み係数を導入することにより，エッジ保存性を持った非線形スケールスペースを実現した．

$$\frac{\partial}{\partial\tau}I(\tau,x,y)=div\Big(g(|\nabla I(\tau,x,y)|)\cdot\nabla I(\tau,x,y)\Big) \tag{2.5}$$

図 2.1 に，ガウシアンによる線形スケールスペースと Perona-Malik による非線形スケールスペースの例を示す．線形スケールスペースにおける拡散は等方的であり，細かなテクスチャだけでなくオブジェクトのエッジに対しても分け隔てなく平滑化が行われるため，画像全体がぼやけた印象になってしまう．これに対し，Perona-Malik による非線形スケールスペースでは，画像の勾配が大きくなる方向への平滑化を抑制する非等方的な拡散が実現されるため，細かなテクスチャは平滑化されている一方で，オブジェクトのエッジは時間が経っても保存され，クリアな印象の平滑化画像が得られるというメリットがある．

Perona-Malik によるこの方法は，非等方拡散法（Anisotropic Diffusion）と呼ばれ，本論文においては，エッジ保存平滑化プレフィルタの構築にあたって重要な役割を果たすため，2.3 節でその技術的な詳細を述べることとする．

Alvarez-Guichard-Lions-Morel のスケールスペース理論 [3]

上で述べた線形／非線形のスケールスペースは，共通する性質として，時間 τ が増大するにつれて次第に精細さが失われ，画像が単純化されていくという「情報損失性」を持っている．Alvarez-Guichard-Lions-Morel は，このようなスケールスペースが満たすべき本質的な性質を，以下に示す 4 つの公理と 2 つの不変性によって特徴づけた．

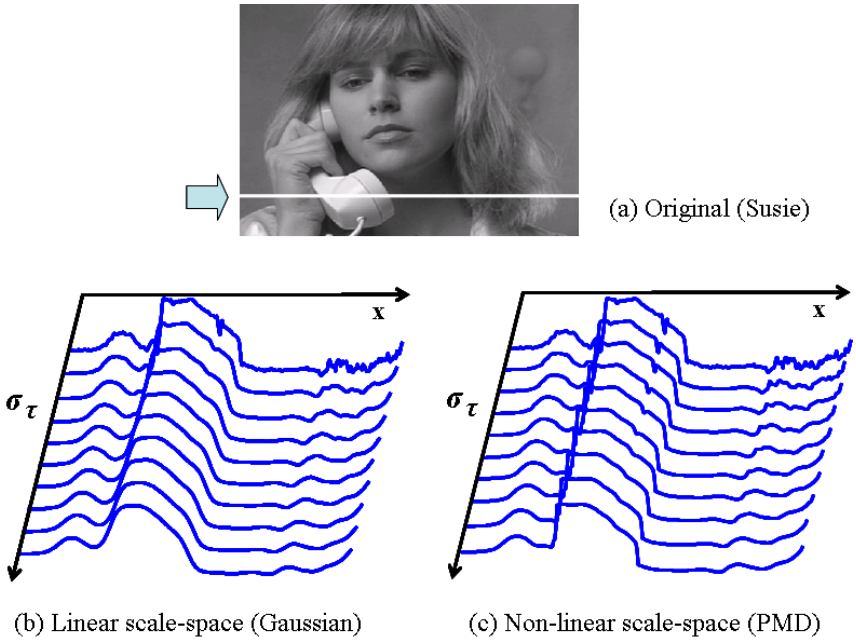


図 2.1: スケールスペース

スケールスペースの公理 [4]　画像信号 u に作用する作用素の族 $\{T_\tau\}_{\tau\geq 0}$ を用いて，$I(\tau,\mathbf{x}) = T_\tau(u_0)$ としてスケールスペースを定義する．ここで，u_0 は $\tau = 0$ における初期画像である．このとき，以下の性質が成り立つ．

(A1) Recursivity:　任意の画像 u に対して，

$$T_0(u) = u, \quad T_\sigma \circ T_\tau(u) = T_{\sigma+\tau}(u) \ \text{ for } \forall \sigma, \tau > 0.$$

(A2) Regularity:　任意の画像 u,v に対して，ある定数 C が存在して，

$$\|T_\tau(u + hv) - (T_\tau(u) + hv)\|_{L^\infty} \leq C \cdot ht \ \text{ for } \forall h, t \in [0, 1].$$

(A3) Locality:　$\nabla^\alpha u(\mathbf{x}) = \nabla^\alpha v(\mathbf{x})$ (for $\forall |\alpha| \geq 0$, $\forall \mathbf{x}$) を満たす任意の画像 u,v に対して[1]，

$$(\nabla^\alpha u - \nabla^\alpha v)(\mathbf{x}) = o(\tau), \quad \tau \to 0^+.$$

(A4) Comparison Principle:　R^2 上で，$u \leq v$ を満たす任意の画像 u,v に対して，

$$T_\tau(u) \leq T_\tau(v) \quad \text{on } R^2, \quad \text{for } \forall \tau \geq 0.$$

(I1) Gray-level shift invariance:　任意の画像 u と任意の定数 C に対して，

$$T_\tau(0) = 0, \quad T_\tau(u + C) = T_\tau(u) + C.$$

[1] $\nabla^\alpha u$ は，オーダー α の導関数 (derivative) を表す．

(I2) Translation invariance:　任意の画像 u に対して，

$$T_\tau(u(\mathbf{x}+\mathbf{h})) = (T_\tau(u))(\mathbf{x}+\mathbf{h}) \quad \text{for } \forall \mathbf{h} \in R^2 \text{ and } \tau \geq 0.$$

Alvarez-Guichard-Lions-Morel の功績は，これらの性質を公理として定義されるすべてのスケールスペース表現が，必ず，原画像を初期条件とする，ある偏微分方程式（PDE）によって記述できることを数学的に明らかにした点である．これにより，先に述べた線形・非線形のスケールスペースをはじめとして，従来から知られていたすべてのスケールスペース画像表現が，PDE という新しい視点から統一的に定式化されることとなった．その後も，このスケールスペース理論を踏まえ，様々な画像の特徴を保存する新しいタイプのスケールスペース画像表現 [94] や，Meyer-Maragos によって提案されるモルフォロジカルスケールスペース [50] などにおいて，PDE に基づく定式化が取り入れられている．

2.2.2　画像修復

以下に，PDE に基づく手法を画像修復に関する処理に利用した従来研究について述べる．

TV フィルタ [14]

TV フィルタは，Rudin-Osher-Fatemi[71, 72] によって提案された total variation(TV) に基づく非等方拡散モデルを用いて構成されたディジタルフィルタであり，ノイズ除去などの画像修復に効果を発揮する．以下にその原理を示す.

画像 I に関する TV を以下のように定義する．

$$TV(I) = \int_\Omega |\nabla I| d\mathbf{x} \tag{2.6}$$

ここで，$\Omega \subset \mathbf{R}^2$ は，画像の定義された矩形領域を表す．ノイズが重畳した画像を I^0 とするとき，制約条件

$$\frac{1}{|\Omega|}\int_\Omega (I - I^0)^2 d\mathbf{x} = \sigma^2 \tag{2.7}$$

を満たしながら，式 (2.6) を最小化することにより，ノイズが重畳した画像 I^0 を修復する．これは，Lagrange の未定乗数 λ を導入し，

$$J(I) = \int_\Omega |\nabla I| d\mathbf{x} + \frac{\lambda}{2}\int_\Omega (I - I^0)^2 d\mathbf{x} \tag{2.8}$$

を最小化することによって得られる．この Euler-Lagrange 方程式を考えることにより，その解へ収束するフローとして，

$$\frac{\partial I}{\partial \tau} = \nabla \cdot \left(\frac{\nabla I}{|\nabla I|} \right) + \lambda (I^0 - I) \tag{2.9}$$

が得られる．すなわち，この PDE を数値的に解くことにより，ノイズの除去が可能となる．

Shock フィルタ [60]

Osher-Rudin により提案された shock フィルタ [60] は，エッジ強調効果を持ったフローを提供することから，ボケ画像の修復に使用することができる．2 次元の画像に対する shock フィルタは，次の PDE により実現される．

$$\frac{\partial I}{\partial \tau} = -|\nabla I| \, F(\mathcal{L}(I)) \tag{2.10}$$

ここで，F は $F(0) = 0$ かつ $sign(s)F(s) > 0 \ \ (s \neq 0)$ を満たすリプシッツ連続な関数であり，$\mathcal{L}$ は非線形な楕円型偏微分作用素である．作用素 $\mathcal{L}$ としては，

$$\mathcal{L}(I) = I_{xx} + I_{yy} \tag{2.11}$$

$$\mathcal{L}(I) = \frac{1}{|\nabla I|^2}(I_x^2 I_{xx} + 2 I_x I_y I_{xy} + I_y^2 I_{yy}) \tag{2.12}$$

などが使用される．

Image Inpainting[8]

Inpainting とは，画像の一部が破損したり失われたりした箇所に対して適切な画素を補填して，自然な形に修復することをいう．Bertalmio-Sapiro らによって提案された Image Inpainting のアルゴリズム [8] は，修復したい領域 Ω の画素に対して，次の反復処理を行うことによって与えられる．

$$I^{n+1}(x,y) = I^n(x,y) + \Delta t \, I_t^n(x,y), \quad \forall (x,y) \in \Omega \tag{2.13}$$

ここで，添え字の n は式 (2.13) の反復回数を，(x,y) は修正する画素の座標位置を，I_t^n は画像 I^n への更新値をそれぞれ表している．以上の反復処理は，$I^{n+1}(x,y) = I^n(x,y) \ \ \forall (x,y) \in \Omega$ が成立した場合に停止する．

画像の修復結果の良し悪しは，更新値 I_t^n の与え方によって左右される．Bertalmio-Sapiro らは，更新値 I_t^n の定義として次の式を与えている．

$$I_t^n(x,y) = \nabla L^n(x,y) \cdot \mathbf{N}^n(x,y) \tag{2.14}$$

ここで，L^n と $\mathbf{N}^n$ は次式で定義される．

$$L^n(x,y) = I_{xx}^n(x,y) + I_{yy}^n(x,y), \quad \mathbf{N}^n(x,y) = \nabla I^n(x,y)^{\perp} \tag{2.15}$$

以上のアルゴリズムに基づく画像修復の例を図 2.2 に示す.

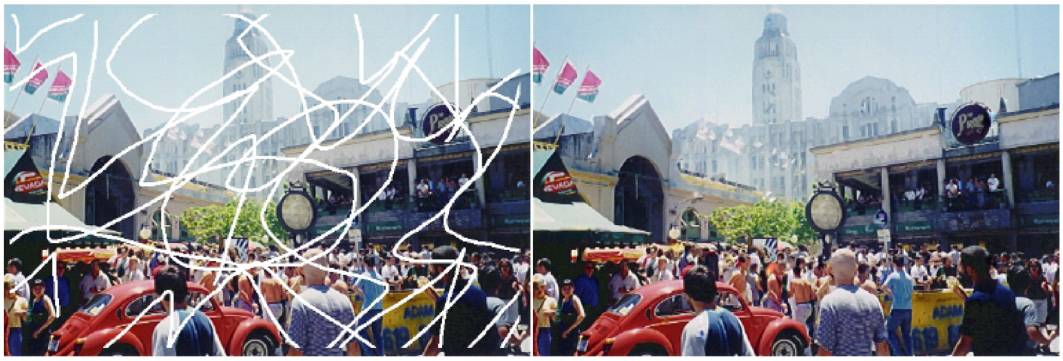


図 2.2: Image Inpainting による画像修復の例 [8]

2.2.3　セグメンテーション

以下では，特に動的輪郭法に基づくセグメンテーションに関する従来研究の流れについて述べる．動的輪郭法（あるいは snakes）とは，画面中の対象物体を取り囲むように配置した曲線を，画像のエッジ情報や曲線の滑らかさなどの情報を基に時間的に発展させ，対象物体の輪郭へ収束させるオブジェクト領域抽出法である．動的輪郭法によるオブジェクト抽出の典型的な例を図 2.3 に示す．

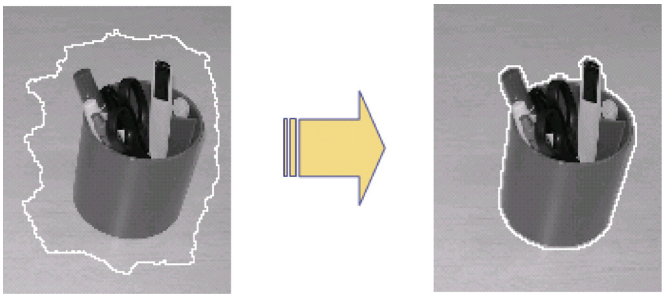


図 2.3: 動的輪郭法によるオブジェクト抽出例

古典的な動的輪郭モデル [38, 79]

Snakes の古典的な手法としては，1987 年に Kass-Witkin-Terzopoulos により提案されたモデルが有名である．このモデルにおいて，時間発展する輪郭の曲線はパラメトリックな曲線で表される．今, 画像平面 (x, y) 上で媒介変数表現された平曲線 $\mathbf{c}(q) = (x(q), y(q))(0 \leq q \leq 1)$ として snakes が表現されるものとする．曲線の滑らかさを表す内部エネルギーを $E_{int}(\mathbf{c}(q))$, 輝度勾配等の画像の特徴量から決まる外部エネルギーを $E_{ext}(\mathbf{c}(q))$ とすると，式 (2.16) に示すエネルギー関数の総和が最小化するように曲線 $\mathbf{c}(q)$ を変形することによって輪郭抽出

を行う．

$$E_{snakes} = \int_0^1 \{E_{int}(\mathbf{c}(q)) + E_{ext}(\mathbf{c}(q))\} dq. \tag{2.16}$$

ここで，典型的な内部エネルギー $E_{int}(\mathbf{c}(q))$ としては，曲線が連続かつ滑らかに保たれるように，$\mathbf{c}(q)$ の 1 階微分，2 階微分を用いて以下のように定義される．

$$E_{int}(\mathbf{c}(q)) = \alpha(q) \cdot \left| \frac{d\mathbf{c}(q)}{dq} \right|^2 + \beta(q) \cdot \left| \frac{d^2\mathbf{c}(q)}{dq^2} \right|^2. \tag{2.17}$$

また，典型的な外部エネルギー $E_{ext}(\mathbf{c}(q))$ としては，曲線がオブジェクトのエッジに引きつけられるように，画像の勾配を用いて以下のように定義される．

$$E_{ext}(\mathbf{c}(q)) = -\gamma(q) \cdot |\nabla I|. \tag{2.18}$$

以上述べた古典的な snakes には，以下に示すような欠点があることが知られている．本手法を数値的に実装するためには，曲線 $\mathbf{c}(q)$ 上のサンプル点をとって離散化しなくてはならないが，曲線が変形していくにつれ，サンプル点同士の位置関係や距離などの整合性を取る必要が生じ，数値的な扱いが難しくなる．このため，初期輪郭を対象物体の近くに配置しなくてはならない．さらに，snakes の振る舞いがエネルギーの各項の重みを決める係数 α, β, γ のとり方に敏感であり，目的の輪郭を検出する際のパラメータの値を決定するのが困難である．また，特に greedy algorithm に基づく snakes の実装 [83] を行った場合は，初期輪郭が少し変わっただけで収束する輪郭線の形状が異なるなど，結果の安定性に問題が生じる場合がある．

曲線の時間発展方程式 [75, 4]

一般に，次のような画像平面 $\Omega(\subset \mathbf{R}^2)$ 上の曲線 $\mathbf{c}(t, s)$ に関する時間発展方程式を考える．

$$\begin{cases} \dfrac{\partial \mathbf{c}}{\partial t} = F \cdot \mathbf{N} \\ \mathbf{c}(0, q) = \mathbf{c}_0(q) \end{cases} \tag{2.19}$$

ここで，$\mathbf{N}$ は曲線 $\mathbf{c}$ 上の各点における法線ベクトルであり，図 2.4 に示すように，曲線の接線方向に対して左向きの単位ベクトルとして与えられる．また，F は Ω 上で定義される速度関数である．このとき，方程式 (2.19) は曲線 $\mathbf{c}$ が時間とともにどのように変形していくかを表す曲線の運動方程式であり，F が法線方向への移動速度を表していることがわかる．曲線の運動を表すにあたり，法線方向に限らない任意のベクトルを用いることもできるが，一般に接線方向の成分は曲線の形状変化に影響しないことが知られている [40] ため，式 (2.19) のように，曲線の法線方向のベクトルのみを用いて曲線の運動を記述することが一般的である．

近年の動的輪郭モデルにおいては，オブジェクト輪郭の形状変化を記述する際に，上述のような曲線の時間発展方程式を積極的に利用する．通常，オブジェクトの輪郭線は閉曲線として与えられるが，特に閉曲線の運動を扱う場合には，次に述べるレベルセット法による定式化が有効である．

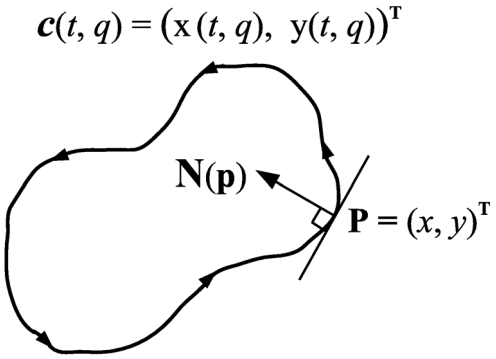


図 2.4: 曲線 $\mathbf{c}(q)$ と法線ベクトル $\mathbf{N}$

レベルセット法 [4, 78, 75, 61]

レベルセット法とは，x-y 平面上の輪郭曲線 $\mathbf{c}(t,q)$ を x-y 平面上で定義された超曲面 $u(t,x,y)$ のゼロ交差（レベルセット）として定義し，その超平面 u（レベル曲面）全体を時間発展させることにより，ゼロ交差である輪郭曲線 $\mathbf{c}$ の運動を表現する数値的な技法である．すなわち，曲面 $u : \mathbf{R}^+ \times \mathbf{\Omega} \to \mathbf{R}$ として，

$$u(t, \mathbf{c}(t,q)) = 0 \quad \text{for } \forall q, \forall t \geq 0 \tag{2.20}$$

を満たす関数を考え，曲線 $\mathbf{c}(t,q)$ そのものを時間発展させる代わりに曲面 $u(t,x,y)$ を時間発展させ，各時刻における曲線をそのレベルセットとして抽出する．式 (2.19) に相当する曲線 $\mathbf{c}(t,q)$ の運動は，レベル曲面 $u(t,x,y)$ に関する次の時間発展方程式により実現される[2]．

$$\frac{\partial u}{\partial t}(t,x,y) = F \cdot |\nabla u(t,x,y)|. \tag{2.21}$$

レベルセット法を実装する上で，ゼロ交差として $\mathbf{c}(q)$ を表現するレベル曲面 $u(x,y)$ は無数に存在し，性質の良いものであれば何を使ってもよいが，曲線 $\mathbf{c}(q)$ によって決まる符号付き距離関数（signed distance function）を使用するのが最も典型的である．符号付き距離関数の定義は次の通りである [4]:

$$u(\mathbf{x}) = \begin{cases} +d(\mathbf{x}, \mathbf{c}) & \text{if } \mathbf{x} \text{ is outside } \mathbf{c} \\ -d(\mathbf{x}, \mathbf{c}) & \text{if } \mathbf{x} \text{ is inside } \mathbf{c}, \end{cases} \tag{2.22}$$

ここで，$\mathbf{x} = (x,y) \in \Omega \subset R^2$ であり，$d(\mathbf{x}, \mathbf{c})$ は $\mathbf{x}$ から曲線 $\mathbf{c}$ へのユークリッド距離を表す．

レベルセット法において，曲線の運動を表すために，曲線そのものを直接扱わず，あえてこれをゼロ交差にもつレベル曲面を扱う主な理由は，数値計算上の利点による．特に，曲線

[2]証明は 2.4.2 に後述する．なお，式 (2.21) は Hamilton-Jacobi 方程式 $u_t + H(\nabla u) = 0$ において，特に $H(p) = -F \cdot |p|$ として得られるケースであると考えられる．

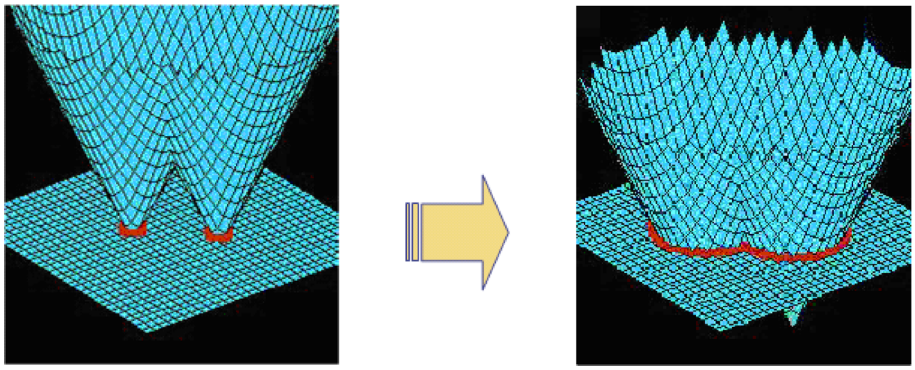


図 2.5: レベルセット法による曲線の時間発展 [77]

を直接離散化した場合にはトポロジー変化をうまく表現することが困難であるのに対して，曲面を使えばそれが容易に実現できる．ここで，トポロジー変化とは，単一の曲線が複数の曲線に分離しながら発展したり，逆に複数の曲線が融合して一つの曲線に統合しながら発展したりすることを指している．図 2.5 に符号付き距離関数に基づくレベルセット法の動作例を示す．図に示すように，二つの円錐を合体したような超曲面を考えれば，一つの円が二つに分離したり，逆に二つの円が一つに統合したりするトポロジー変化を表現することが可能である．

平均曲率流方程式 (mean curvature motion)[59, 24]

動的輪郭法において特に重要な役割を果たす時間発展方程式として，Osher と Sethian によって導入された平均曲率流方程式

$$\frac{\partial \mathbf{c}}{\partial t} = \kappa \cdot \mathbf{N} \tag{2.23}$$

が挙げられる．このフローでは，曲率 κ の大きくなる点ほど速い速度で移動することから，どのような複雑な形の曲線も最終的には単純な円に縮退し，消滅することとなる．平均曲率流に従って曲線が縮退していく様子を図 2.6 に示す．

一般に x-y 平面上の曲面 $u(x, y)$ の等高線に対する曲率 κ が

$$\kappa = div\left(\frac{\nabla u}{|\nabla u|}\right)$$

と書けることから，平均曲率流方程式 (2.23) は，レベルセット法を用いて，次のように表現できる．

$$\frac{\partial u}{\partial t} = |\nabla u| \cdot div\left(\frac{\nabla u}{|\nabla u|}\right) \tag{2.24}$$

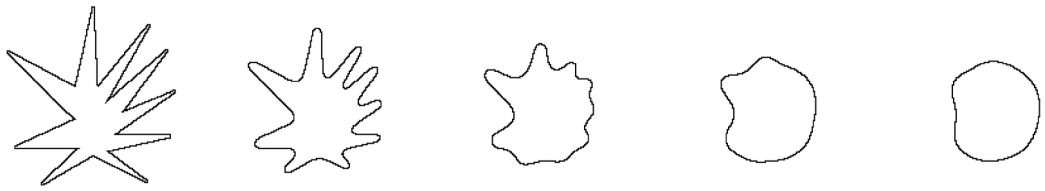


図 2.6: 平均曲率流方程式による閉曲線の時間発展

幾何学的な動的輪郭モデル [11, 45]

先に述べた Osher-Sethian の平均曲率流方程式の成果を受け，これベースとする幾何学的な動的輪郭モデルが Caselles らによって提案されている [11]．レベルセットを用いた PDE による定式化は次の通りである．

$$\frac{\partial u}{\partial t} = g(|\nabla I|) \cdot |\nabla u| \cdot \left(div\left(\frac{\nabla u}{|\nabla u|} \right) + \nu \right) \tag{2.25}$$

式 (2.25) において本質的なのは，平均曲率流方程式 (2.24) にエッジ検出関数による重み係数 $g(|\nabla I|)$ を掛けて得られる第 1 項である．この項によって得られる，曲線の時間発展の様子は次のように説明できる．まず，曲線が平坦な領域上にある場合には $g(|\nabla I|) \approx 1$ となり，曲線は平均曲率流に従って平滑化されながら縮退してゆく．その過程において曲線がエッジ点に遭遇すると，$g(|\nabla I|) \approx 0$ となることから，そこで曲線の時間発展が停止する．このようなプロセスを経て，初期輪郭をオブジェクト輪郭上に収束させることが，この幾何学的フローのねらいである．一方，式 (2.25) の第 2 項に見られる ν は，曲線の運動の速度を促進するため追加されたパラメータである．

Caselles らとほぼ同時期に，平均曲率流に基づくこれと同等な動的輪郭モデルが，Malladi らによっても提案されている [45]．同論文において，Malladi らはさらに次のような動的輪郭モデルについても言及している．

$$\frac{\partial u}{\partial t} = |\nabla u| \cdot \left(-\nu + \frac{\nu}{M_1 - M_2} \Big(|\nabla G_\sigma * I| - M_2 \Big) \right) \tag{2.26}$$

次に述べる「測地線に基づく動的輪郭法」は，ここで述べた幾何学的な動的輪郭モデルをさらに発展させた手法として位置づけられる．

測地線に基づく動的輪郭法 [12]

古典的な動的輪郭モデルでは，曲線 $\mathbf{c}(q)$ のパラメータ q を別のパラメータに変換すると，最小化すべきエネルギーの値もまた変化してしまう．曲線 $\mathbf{c}(q)$ のパラメータを変換しても曲線の形状自体は変わらないことを考えると，本質的に最小化すべきエネルギーとしては，曲線のパラメータ変換に関して不変な幾何学的な量であることが望ましい．

1997 年に Caselles-Kimmel-Sapiro によって提案された測地線に基づく動的輪郭法 (GAC) は，このような観点から導入された幾何学的な snakes である．GAC においては，画像平面 Ω 上の曲線 $\mathbf{c}$ の長さ $L(\mathbf{c})$ が，通常のユークリッド距離とは異なる尺度（リーマン計量）に基づいて次のように定義される．

$$L(\mathbf{c}) = 2\sqrt{\lambda}\int_a^b g\Big(|\nabla I(\mathbf{c}(q))|\Big)|\mathbf{c}'(q)|dq. \tag{2.27}$$

この距離尺度の下では，重み付け係数 $g(|\nabla I(\mathbf{c}(q))|)$ の影響から，勾配の大きなエッジ点における接ベクトルの長さがユークリッド距離に比べて過小評価されることになる．このため，長さ $L(\mathbf{c})$ が極小となるように初期輪郭線を時間発展させることにより，最終的にそのほとんどがエッジ点から構成されるオブジェクト輪郭曲線に収束することが期待される．このように，定められた 2 点を端点とする曲線のうち，その長さを極小化する曲線を特に測地線 (geodesics) と呼び [91]，これが本手法の名前の由来となっている．

GAC において極小化するエネルギーとは，画像平面 Ω にユークリッド距離とは異なる計量を入れた一般化曲面上の曲線の長さに他ならない．数学ではこのような一般化曲面をリーマン多様体 (Riemannian manifold) と呼ぶ．したがって，GAC とはリーマン多様体上の測地線を求める問題として理解することができる．リーマン多様体上の曲線の長さは幾何学的な量であるため，$\mathbf{c}(q)$ のパラメータを変換してもその値は変化しないという利点を持っている．また，GAC のもう一つの利点として，レベルセット法と組み合わせて使用できる点が挙げられる．これにより，古典モデルの欠点である数値的な実装に関する多くの問題を回避することができる．なお，測地線に基づく Caselles らの動的輪郭法は，適切な仮定の下で，Kass らの古典的な手法と同値であることが数学的に証明されている [4].

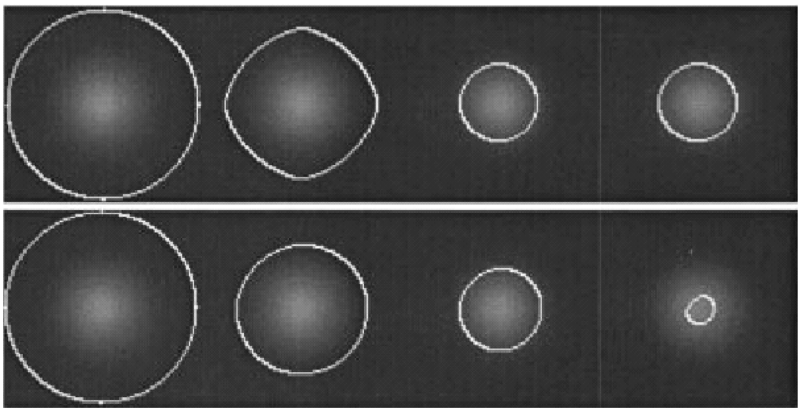
GAC は，本論文における VOP 自動生成アルゴリズムの構築にあたって重要な役割を果たすため，2.4 でその技術的な詳細を述べることとする．

勾配を用いない動的輪郭法 [15]

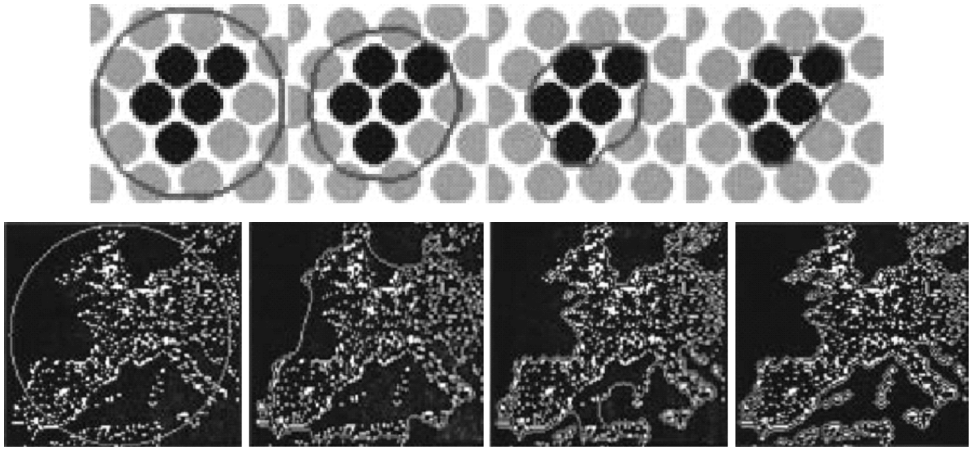
これまで述べた動的輪郭モデルは，輪郭曲線の時間発展を停止するための条件として，どれも画像勾配 $|\nabla I|$ を用いていた．このため，オブジェクトの輪郭が画像勾配に基づくエッジとして明確に検出されない場合には，エッジ検出関数による評価値 $g(|\nabla I|)$ が十分ゼロに近づかないことから曲線の時間発展を停止できず，正しい輪郭線を得られないという問題があった．Chan-Vese は，このような従来法が抱える問題を解決するため，画像勾配を用いなくてもオブジェクト輪郭を抽出できる “without edges” と呼ばれる画期的な動的輪郭モデルを提案した．このモデルは，次の最小化問題によって定式化される．

$$\begin{aligned}\inf_{c_1,c_2,\mathbf{c}} F(c_1,c_2,\mathbf{c}) = \mu\cdot Length(\mathbf{c}) \quad &+ \quad \lambda_1\int_{inside(\mathbf{c})}|I(x,y)-c_1|^2dxdy \\ &+ \quad \lambda_2\int_{outside(\mathbf{c})}|I(x,y)-c_2|^2dxdy\end{aligned} \tag{2.28}$$

ここで，$\mathbf{c}$ は画像平面状の閉曲線，c_1, c_2 は実数変数を表す．また，μ, λ_1, λ_2 はすべて正の固定値を表している．この最小化問題をレベルセット法によって定式化すると，レベル曲面



(a) 曖昧輪郭線の検出（上段: without edges，下段：GAC）



(b) 主観輪郭線の検出

図 2.7: 勾配を用いない動的輪郭法の例 [15]

$u(x,y)$ が満たすべき時間発展方程式は次の式で与えられる．

$$\frac{\partial u}{\partial t} = \delta_\epsilon(u)\left(\mu \cdot div\left(\frac{\nabla u}{|\nabla u|}\right) - |I - c_1|^2 + |I - c_2|^2\right) \tag{2.29}$$

$$c_1 = \frac{\int_\Omega I \cdot H(u)dx}{\int_\Omega H(u)dx},\ \ c2 = \frac{\int_\Omega I \cdot (1 - H(u))dx}{\int_\Omega (1 - H(u))dx}. \tag{2.30}$$

ここで，H はヘヴィサイド関数であり，$\delta_\epsilon = H'_\epsilon$ である．

本手法に基づくセグメンテーションの結果を図 2.7 に示す．図 2.7(a) には，本手法と GAC によるセグメンテーションの比較結果を提示した．GAC ではぼんやりとした曖昧な輪郭を持つオブジェクトを検出できないのに対して，“without edges”はこれを容易に実現することができる．また，図 2.7(b) の例のように，明確なエッジを持たないオブジェクトに対する主観的な輪郭線を捉えることも可能である．

その他の関連研究

動的輪郭法に関するその他の関連研究として，Xu らの Gradient Vector Flow(GVF) と Paragios らの Geodesic Active Regions がある．

Gradient Vector Flow[96] は，古典的な動的輪郭法の外部エネルギー項を改善し，オブジェクト輪郭への収束性を高めるために導入された．GVF に基づく外部エネルギーを導入することにより，ターゲットから離れた位置に置かれた初期輪郭もオブジェクト輪郭に収束するようになる．また，凹領域オブジェクトに対して，凹部分の輪郭への収束性が高まるなどの利点が得られる．後に，GVF をレベルセット PDE に導入するための手法が，Paragios らにより提案されている [67].

Geodesic Active Regions[64] は，エッジ情報に加え，領域情報を利用しながら動的輪郭をオブジェクトに収束させる，GAC の拡張手法である．統計量の測定とテクスチャ解析による一様な領域の探索とエッジディテクタによるエッジ点の探索による双方の結果に基づき，符号付き距離関数 u を時間発展させることにより，オブジェクト輪郭を抽出する．このため，GAC に比べて，よりロバストな輪郭線抽出が可能となる．また，GAC では膨張または収縮のどちらか一方向にしか時間発展できないため，曲線全体が完全にオブジェクトの外部もしくは内部に含まれるように初期輪郭を配置する必要があるのに対し，Geodesic Active Regions ではそのような制約は必要なく，自由な位置に初期輪郭を配置できるというメリットがある．

2.2.4　オブジェクト・トラッキング

以下に，PDE に基づくオブジェクト・トラッキングの従来研究について述べる．

Geodesic Active Tracking[62, 66]

先に述べた測地線に基づく動的輪郭法 (GAC) をオブジェクト・トラッキングに適用した最初の事例として，Paragios-Deriche による Geodesic Active Tracking が挙げられる．本手法は，固定カメラによる背景の動かない画像を前提としており，動き検出部とトラッキング部の 2 つのステップから成る．

動き検出部は，時間的に連続する 2 フレーム間の差分信号をもとに，動きのあるオブジェクト領域を切り出すステップである．まず，各画素 $\mathbf{x}$ がオブジェクト領域と背景領域の境界点である確率の場を考え，これを値域 $[0, 1]$ の画像とみなして得られる動き検出境界画像（motion detection boundary image）I_D を生成する．この画像 I_D を用いて，次の最小化問題を解くことにより，2 フレームに跨るオブジェクトの移動領域を求める．

$$E[\mathbf{c}(q)] = \int_b^a g(I_D(\mathbf{c}(q)))|\dot{\mathbf{c}}(q)|. \tag{2.31}$$

ここで，g はガウシアン関数

$$g(x,\sigma) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x^2}{2\sigma^2}} \tag{2.32}$$

を表す．このエネルギー汎関数を最小とする曲線は GAC と同様のレベルセット PDE による曲線の時間発展処理によって求めることができる．

また，トラッキング部は，動き検出部で求めたオブジェクトの移動領域の輪郭を起点として，その内部に存在する現画像の移動オブジェクトを通常の GAC に基づいて抽出するステップである．

これら処理を繰り返すことにより，画像シーケンスからオブジェクト輪郭を追跡することが可能となる．本手法に基づくオブジェクト・トラッキングの例を図 2.8 に示す．以上の 2 ステップ処理は，次のレベルセット PDE により定式化することができる．

$$\begin{aligned} \frac{\partial u}{\partial t} &= \Big\{\gamma\Big(g(I_D,\sigma_D)\kappa + \nabla g(I_D,\sigma_D)\cdot\frac{\nabla u}{|\nabla u|}\Big) \\ &\quad +(1-\gamma)\Big(g(|\nabla I|,\sigma_T)\kappa + \nabla g(|\nabla I|,\sigma_T)\cdot\frac{\nabla u}{|\nabla u|}\Big)\Big\}|\nabla u|. \end{aligned} \tag{2.33}$$

ここで，$\gamma = 1$ のときは動き検出部の処理を，$\gamma = 0$ のときはトラッキング部の処理を表している．なお，トラッキング部については，エッジ情報に加え，領域内部や動きに関する特徴量を利用して，オブジェクト輪郭抽出のロバスト性を改善する拡張手法についても提案が行われている [63, 65].

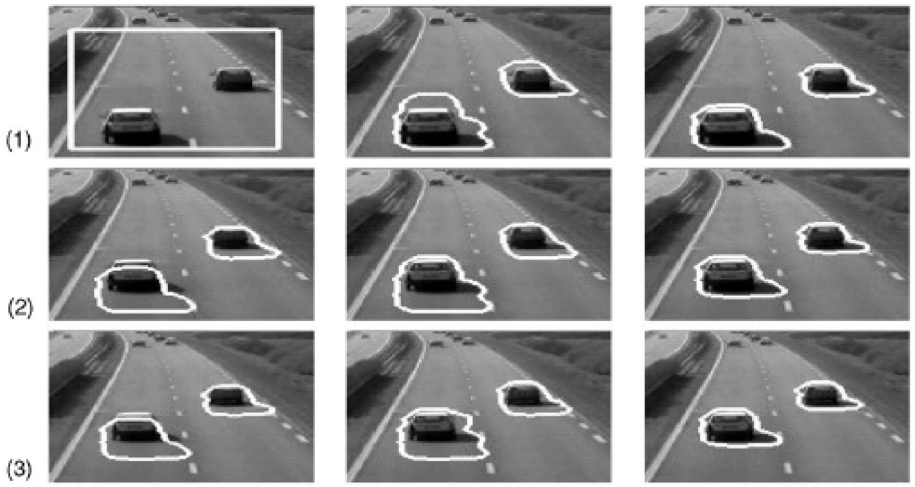


図 2.8: Geodesic Active Tracking によるトラッキングの例 [66]

その他の関連研究

レベルセット PDE を用いたオブジェクト・トラッキングの関連研究として，Bertalmio-Sapiro らの Morphing Active Contours[7] と Mansouri らの Region Tracking[46] がある．

Morphing Active Contours は，時間的に連続するフレーム間におけるオブジェクトの動き検出処理とオブジェクトの輪郭を抽出するトラッキング処理から構成されるという点において，基本的に Geodesic Active Tracking と同様の手法あるが，これらの処理を連立偏微分方程式 (coupled PDE) によって実現している点に新規性がある．

Region Tracking は，MDL 原理に基づくトラッキング手法である．本手法には，事前に動きベクトル場の計算が必要となること，トラッキングの結果が動きベクトルの精度に大きく影響されることなどの問題点があった．これらを踏まえた上で，その後，ベイズ推定に基づく画期的な Region Tracking 手法が，Mansouri によって新たに提案されている [47]．この新手法は，従来の Region Tracking が前提としていた動きベクトル情報を必要とせず，さらに背景が動いている場合にも対応できるなど，汎用手法としての優れた特長を持っている．

2.2.5　画像編集

PDE を用いたオブジェクト単位の画像編集の自動化に関する研究として，ポアソン方程式を基盤とする 2 つの手法について述べる．

Poisson Image Editing[69]

背景となる画像に別の画像から切り抜いたオブジェクトを貼り付けて合成画像を作るための手法である．自然な合成画像を作るには，オブジェクトを背景に対してシームレスに埋め込むことが重要である．Poisson Image Editing では，Dirichlet の境界条件を持つポアソン方程式を解くことにより，人手を介すことなく，これを実現することができる．

本手法の原理は次の通りである．$S(\subset R^2)$ 上の画像がその閉部分集合 Ω の内部において未定義であるとする．Ω の内部以外で定義された既知の画像を f^*，Ω 内部で定義される未知の補間画像を f として，f が Ω 外部の既知画像 f^* とシームレスにつながるようにしたい．

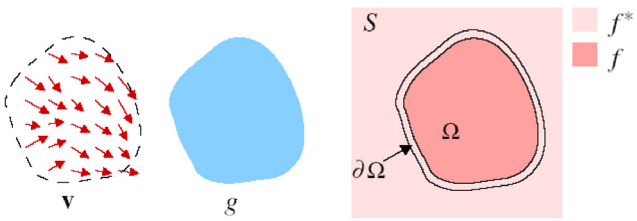


図 2.9: 指標ベクトル場に基づく画像補間 [69]

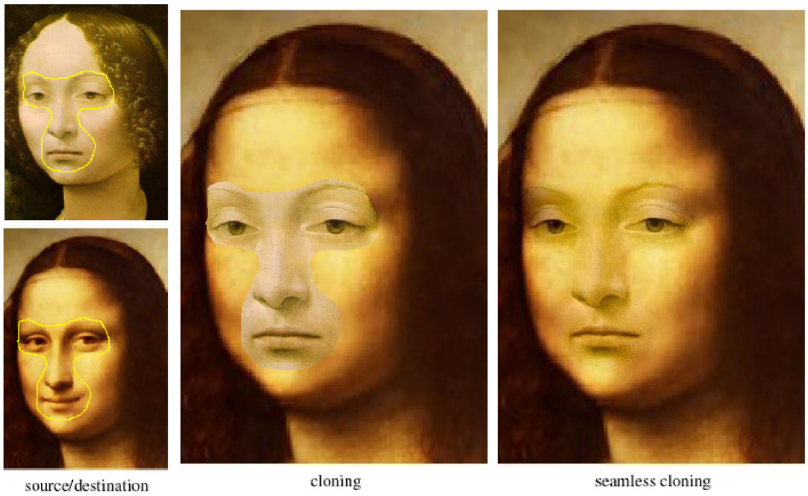


図 2.10: Seamless Cloning の適用例 [69]

最も簡単な補間の方法は，境界条件 $f|_{\partial\Omega} = f^*|_{\partial\Omega}$ の下でラプラス方程式を解くことであるが，この場合，補完した画像はディテールが失われてしまう．そこで，貼り付けるオブジェクトの画像 g を基に指標となるベクトル場 (guidance vector field) $\mathbf{v}$ を生成し，次の変分問題を解くことによって，ディテールを失わないシームレスな画像合成を実現する．

$$\min_f \int\int_\Omega |\nabla f - \boldsymbol{v}|^2 \quad \text{subject to} \quad f|_{\partial\Omega} = f^*|_{\partial\Omega}. \tag{2.34}$$

ここで，変分問題 (2.34) の極小解 f を求めることは，次の Dirichlet の境界条件を持つポアソン方程式 (2.35) を解くことと同値である．

$$\triangle f = div\ \boldsymbol{v} \quad \text{with} \quad f|_{\partial\Omega} = f^*|_{\partial\Omega}. \tag{2.35}$$

これらを有限差分法で離散化し，SOR 法や V-Cycle Multigrid 法 [27] で解いた場合，6 万画素レベルの画像領域の編集に要する処理時間は，Pentium4 マシンを使って 0.4 秒程度との報告がある [69]．Poisson Image Editing によるシームレスな画像合成の例を図 2.10 に示す．この例では，指標ベクトル場として，オブジェクトの画像 g に関する勾配ベクトル場 ∇g を使用している．

Poisson Matting[81]

自然画像を用いたオブジェクトのマット合成を実現するための手法である．マット合成とは，前景として抜き出す領域をマウスやタブレットなどを使って人手で指定する方法で，ブルーバックのように特殊な背景を持たないオブジェクトに対しても，マットキーを用いて合成画像を作成できるメリットがある [32]．Poisson Matting は，予めユーザが与える “trimap”

と呼ばれる領域分割（前景領域 Ω_F，背景領域 Ω_B，未知領域 Ω への分割）を前提とし，未知領域 Ω に対して Dirichlet の境界条件を持つポアソン方程式を解くことにより，高精度なマットキーを自動生成する点に特徴がある．

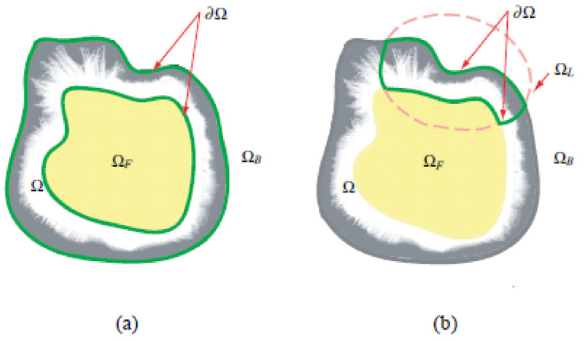


図 2.11: Trimap[81]

本手法の原理は次の通りである．合成画像を $I(x,y)$，背景画像を $B(x,y)$，前景画像を $F(x,y)$ とし，これらの合成に必要となるマットキーを $\alpha(x,y)$ とすると，マット合成の式は，

$$I = \alpha F + (1-\alpha)B \tag{2.36}$$

と書くことができる．ここで，マットキーは $0 \leq \alpha(x,y) \leq 1$ を満たし，前景領域では $\alpha|_{\Omega_F} = 1$，背景領域では $\alpha|_{\Omega_B} = 0$ である．未知領域 Ω に対する $\alpha|_{\Omega}$ を決めるために，まず α の勾配ベクトル $\nabla\alpha$ を求める．

$$\nabla\alpha = \frac{1}{F-B}(\nabla I - \mathbf{D}), \quad \text{where} \quad \mathbf{D} = \alpha\nabla F + (1-\alpha)\nabla B. \tag{2.37}$$

ここで，通常，$\mathbf{D}$ は $(F-B)\nabla\alpha$ に比べて小さな値をとることから，

$$\nabla\alpha \approx \frac{1}{F-B}\nabla I \tag{2.38}$$

を $\nabla\alpha$ の近似値として採用することができる．得られた $\nabla\alpha$ の近似値を指標ベクトル場とすれば，マットキー $\alpha|_{\Omega}$ は，次のポアソン方程式を解いて求めることができる．

$$\triangle\alpha = div\Big(\frac{1}{F-B}\nabla I\Big) \tag{2.39}$$

$$\text{with} \quad \alpha|_{\partial\Omega}(x) = \begin{cases} 1 & x \in \Omega_F \\ 0 & x \in \Omega_B \end{cases} \tag{2.40}$$

式 (2.40) によって生成したマットキーと，これに基づく画像合成の例を図 2.12 に示す．

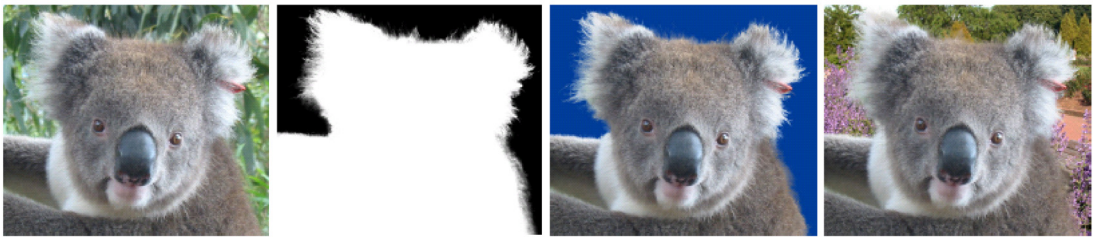


図 2.12: Poisson Matting による画像の合成例 [81]

2.3　非等方拡散法

本論文で提案する平滑化プレフィルタの基礎となる非等方拡散法 (Anisotropic Diffusion) の概要とその性質について述べる．ここでは，Perona-Malik ら [70] によって提案された最も基本的な実装法である Perona-Malik 拡散法 (PMD:Perona-Malik Diffusion) について説明する．

2.3.1　Perona-Malik 拡散法 (PMD)

非等方拡散法の定義

ガウシアンフィルタによる線形スケールスペース $I : (\tau, \mathbf{x}) \in \mathbf{R}^+ \times \Omega(\subset \mathbf{R}^2) \mapsto I(\tau, \mathbf{x}) \in \mathbf{R}$ の時系列変化の様子が熱拡散方程式に従うことは既に述べた．Perona-Malik はこの事実を非線形に拡張し，エッジを保存する非等方的な平滑化が次の PDE で記述できることを示した．

$$\frac{\partial I_\tau(\mathbf{x})}{\partial \tau} = div[g(\|\nabla I_\tau\|)\nabla I_\tau], \quad \mathbf{x} = (x, y)^\top. \tag{2.41}$$

ここで，$g(\|\nabla I\|)$ は重みづけ係数として機能し，エッジを跨る平滑化を抑制する役割を果たしている．このため，関数 g はエッジ停止関数 (edge-stopping function) と呼ばれる．しばしば利用されるエッジ停止関数としては，Perona-Malik[70] によって提示された 2 つのエッジ停止関数 g_1, g_2 や Black-Sapiro ら [9] によって提示されたエッジ停止関数 g_3 などが挙げられる．

$$g_1(x) = (1 + (x/K)^2)^{-1} \tag{2.42}$$

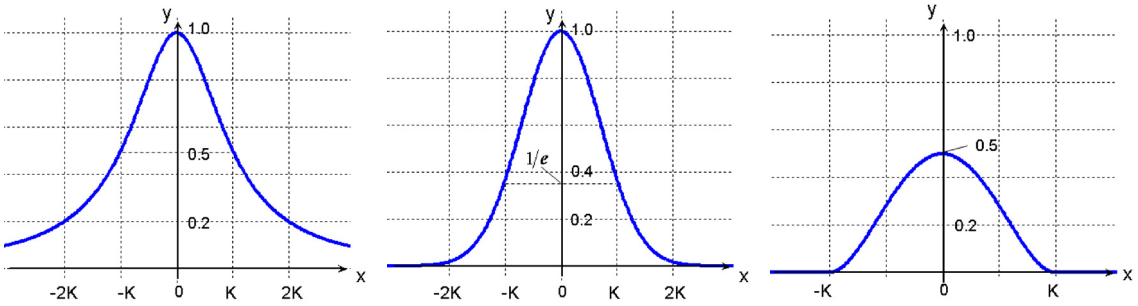
$$g_2(x) = e^{-\frac{x^2}{K^2}} \tag{2.43}$$

$$g_3(x) = \begin{cases} \frac{1}{2}(1 - (x/K)^2)^2 & \text{if } x \le K \\ 0 & \text{otherwise} \end{cases} \tag{2.44}$$

これらの関数のプロットを図 2.13 に示す．エッジ停止関数のプロットは，いずれも y 軸に対して対称で，$x = 0$ において最大値を取るとともに，$|x|$ が大きくなるにつれて 0 に収

束することがわかる．一般に，エッジ停止関数としては，$x \geq 0$ において単調減少で，かつ $\lim_{x\to\infty} g(x) = 0$ となる性質を持つ関数が選択される．

なお，$g \equiv 1$ とした場合には，式 (2.41) は式 (2.4) の熱拡散方程式に他ならず，画像の平滑化はエッジの有無に関わらず等方的に行われる．この場合の平滑化効果はガウシアンフィルタによる平滑化と同等である．



(a) g_1: Perona-Malik　(b) g_2: Perona-Malik　(c) g_3: Black-Sapiro

図 2.13: 典型的なエッジ停止関数のプロット

Perona-Malik による定式化 [70]

非等方拡散方程式 (2.41) を数値的に解くためには，有限差分法による離散化を行う必要がある．Perona-Malik は次のような定式化を行った．

$$
\begin{aligned}
I_{\mathbf{x}}^{(n+1)} &= I_{\mathbf{x}}^{(n)} + \frac{\lambda}{|\eta_{\mathbf{x}}|} \sum_{\mathbf{p}\in\eta_{\mathbf{x}}} g(\nabla I_{\mathbf{x},\mathbf{p}}^{(n)}) \cdot \nabla I_{\mathbf{x},\mathbf{p}}^{(n)}, \qquad (2.45)\\
\nabla I_{\mathbf{x},\mathbf{p}}^{(n)} &= I_{\mathbf{p}}^{(n)} - I_{\mathbf{x}}^{(n)}.
\end{aligned}
$$

ここで，$I_{\mathbf{x}}^{(n)}$ は n 回の反復演算後の画素 $\mathbf{x} = (x, y)^\top$ における画素値を表す．また，λ は数値計算を安定に収束させるためのパラメータである．$\eta_{\mathbf{x}}$ は画素 x の近傍領域で，ここでは図 2.14 に示すフレーム内の空間 4-近傍を指す．本稿では，原画像を初期画像として式 (2.45) による繰り返し処理によって実現される非等方拡散法を特に Perona-Malik 拡散法 (以下，PMD と記す) と呼ぶ．

非等方拡散法の特長として，エッジ停止関数の効果によってエッジを跨る平滑化が抑制され，フィルタ処理後もエッジが保存される点が挙げられる．このことは，Perona-Malik の定式化のもとで次のように説明される．式 (2.45) によれば，画像内のある画素 $\mathbf{x}$ における近傍点 $\mathbf{p}$ 方向への平滑化の度合いは，その勾配値をエッジ停止関数で評価した値 $g(\nabla I_{\mathbf{x},\mathbf{p}})$ によって決定される．すなわち，図 2.14 に示すように，勾配値が大きい場合は，エッジ停止関数による重み係数が 0 に近い値となり，その方向への平滑化が抑制される．逆に勾配値が小さい場合は，エッジ停止関数による重み係数が 1 に近い値となり，ガウシアンフィルタ

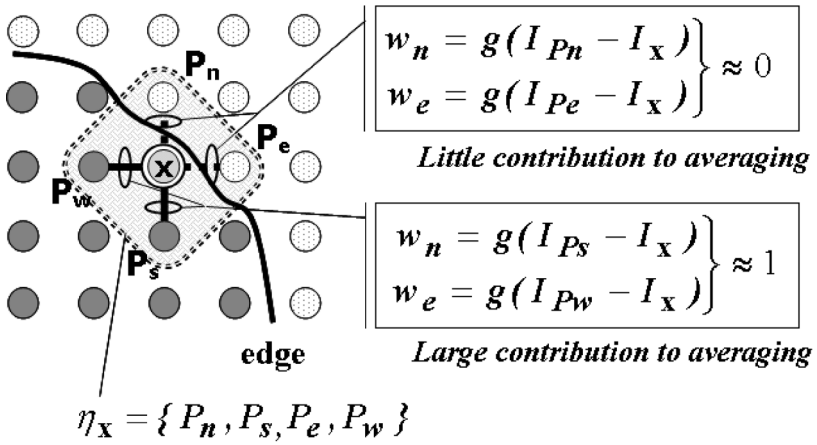


図 2.14: 空間 4-近傍

に近い等方的な平滑化が行われる．その結果として，輝度勾配の大きなエッジは保存され，そうでないテクスチャは選択的に平滑化されることになる．非等方拡散法が持つこのようなエッジ保存性は，低ビットレート映像符号化において，主観画質を保ちながらも，より強力な符号量削減を実現するプレフィルタを構築するにあたって重要であると考えられる．

図 2.15 に PMD による平滑化画像の例を示す．ここでは，$g(x) \equiv 1$ として得られるガウシアンフィルタと，Perona-Malik および Black-Sapiro のエッジ停止関数 g_1, g_3 を用いて得られる PMD において，反復回数 20 回で得られた平滑化画像を比較している．図 2.15(b) のガウシアン平滑化では画像全体が一様にぼやけてしまっているのに対して，同図 (c),(d) の PMD では髪の毛や背景などのテクスチャは生活化されているものの顔の各パーツのエッジが保存されており，主観的には非常に鮮明な印象を与える．また，２つのエッジ停止関数による結果を比較すると，g_3 の方がエッジ強調と平坦部の平滑化の程度が強く，やや人工的／絵画的な印象の画像になっていることがわかる．

PMD の安定性について

差分法によって得られた PMD の離散式 (2.45) を数値的に安定とするためには，安定性パラメータ λ を適切に設定する必要がある．これに関して，Perona-Malik の原論文 [70] では $0 \leq \lambda/|\eta_s| \leq 1/4$ という条件が記されている．本論文では 4-近傍を使用していることから $|\eta_s| = 4$ であり，安定性の条件は $0 \leq \lambda \leq 1$ となる[3]．

この条件の下で式 (2.45) が安定となる根拠について，原論文には特に述べられていないため，付録に証明を記載しておく．これは，PMD の離散式 (2.45) を時間微分に関する前進オイラー法とみなした場合の安定性条件として導くことができる．以下，本論文で扱う PMD では，特に断らない限り，$\lambda = 1$ に固定して議論を進めることとする．

[3]原論文では $\lambda/|\eta_s|$ を λ と表記しているため，安定性の条件は $0 \leq \lambda \leq 1/4$ として述べられている

(a) Original　(b) Gaussian ($g \equiv 1$)　(c) Perona-Malik (g_1)　(d) Black-Sapiro (g_3)

図 2.15: PMD による平滑化処理 ($T = 15$, $\epsilon = 0.1$, $iter = 20$)

2.3.2　エッジ停止関数の正規化

PMD においては，エッジ停止関数の選び方が，その平滑化効果に大きな影響を与える．ここでは，複数のエッジ停止関数を公平に比較する上で必要となるエッジ停止関数の正規化について述べる．

エッジ停止関数の正規化 [87]

PMD における平滑化に影響するパラメータとして，反復回数 $iter$ とエッジ停止関数の定数 K が挙げられる．

$Iter$ は，PMD の漸化式 (2.45) を何回反復した時点で拡散を止めるかを示すパラメータである．式 (2.45) の反復回数は，数値計算上，単位時間ステップ Δt を何回累積したかを示している．このため，その値が大きいほど拡散を開始してからの経過時間が長いことを意味し，その結果，より平滑化が進行することになる．

一方，K はエッジ停止関数 (2.42)〜(2.44) に使用されていた定数であり，関数の水平方向のスケールを決めることから，後に示す選択的テクスチャ平滑化の制御において重要な役割を果たす．しかしながら，式 (2.42)〜(2.44) のプロット（図 2.16）に見られるように，同じ K の値を用いていてもその水平方向のスケールは関数によってまちまちであり，このままでは異なるエッジ停止関数の平滑化効果を公平に比較・評価することが困難である．そこで以下では，定数 K に代わって，水平方向のスケールを正規化するための定数 T を導入し，エッジ停止関数を正規化する．

まず，エッジ停止関数の値域が $[0, 1]$ となるように予め y 軸方向に正規化しておき，得られた関数を g とする．このとき正規化定数 T を次のように定義する:

「十分小さな値 $\epsilon > 0$ に対して，$g(x) = \epsilon$ を満たす正の解 x を T とする」

この定義の下では，T 以上の勾配を持つ方向に対しては重み付け係数が ϵ 以下となって平滑

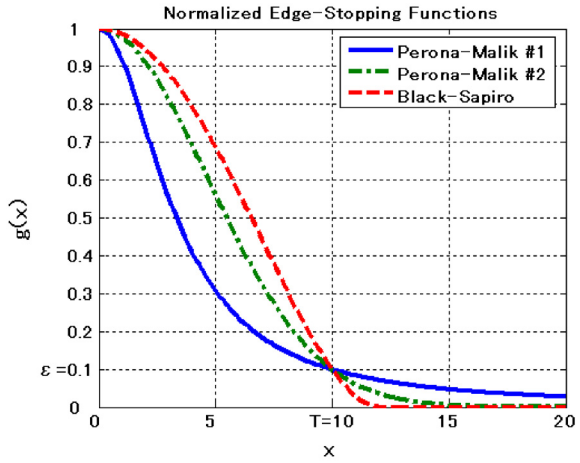


図 2.16: エッジ停止関数の正規化

化の影響が非常に小さくなることから，輝度勾配を基準として勾配 T 以下のエッジを選択的に平滑化したい場合の設定の目安として，本パラメータを使用することができる．例えば，より輝度勾配の大きなエッジまで平滑化したい場合には T を大きく取り，逆にできるだけ細かなエッジを残したい場合には T を小さく取ればよい．

式 (2.42)〜(2.44) で与えられるエッジ停止関数を，$T = 10$, $\epsilon = 0.1$ で正規化したプロットを図 2.16 に示す．一般に，正規化定数 T 及び ϵ が与えられたとき，これらのエッジ停止関数の定数 K は以下の式で与えられる．

$$g_1(x) \text{ の場合：} \quad K = T\sqrt{\frac{\epsilon}{1-\epsilon}} \tag{2.46}$$

$$g_2(x) \text{ の場合：} \quad K = \sqrt{\frac{T^2}{log(1/\epsilon)}} \tag{2.47}$$

$$g_3(x) \text{ の場合：} \quad K = \frac{T}{\sqrt{1-\sqrt{\epsilon}}} \tag{2.48}$$

本論文では，以下，Perona-Malik による g_1 と Black-Sapiro による g_3 を正規化したエッジ停止関数を主に扱い，これらをそれぞれ g^{PM}，g^{BS} と呼ぶ．すなわち，g^{PM} 及び g^{BS} は次のように定義される．

$$g^{PM} = (1 + (x/K)^2)^{-1}, \quad \text{where} \quad K = T\sqrt{\frac{\epsilon}{1-\epsilon}} \tag{2.49}$$

$$g^{BS} = \begin{cases} (1-(x/K)^2)^2 & \text{if } x \le K \\ 0 & \text{otherwise} \end{cases}, \quad \text{where} \quad K = \frac{T}{\sqrt{1-\sqrt{\epsilon}}} \tag{2.50}$$

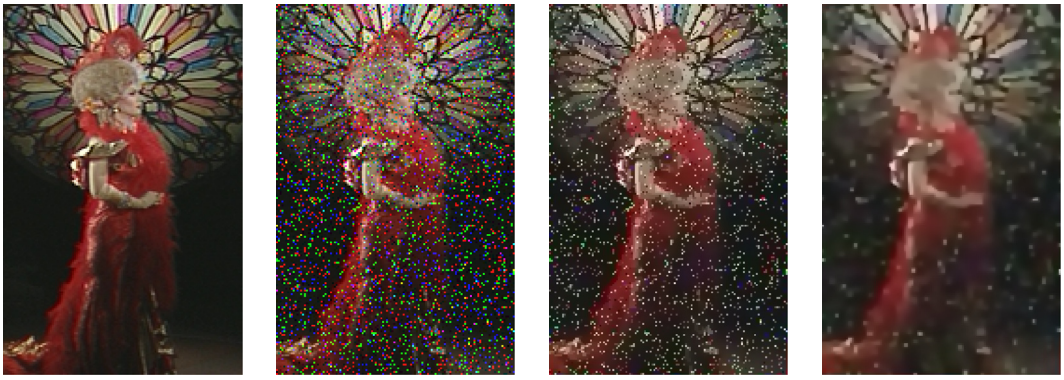


図 2.17: PMD によるノイズ除去効果 : (左から) 原画像，ノイズ重畳画像 (salt&pepper noise, ノイズ密度 0.1)，PMD(g_1) による平滑化結果，正則化 PMD(g_1) による平滑化結果 (いずれも，$T = 15$, $\epsilon = 0.1$, $iter = 20$)

2.3.3　PMD の正則化

PMD の問題点として，エッジ停止関数の影響により，特に salt & pepper ノイズに代表される輝度勾配の大きなノイズに対して安定した除去効果が得られないという点が知られている [70]．これに対し，Catte, Alvarez ら [13, 2] は，PMD に簡単な修正項を加えることにより，ノイズ除去を安定的に解決する方法を提案している．PMD におけるエッジ停止関数 g の引数が原画像の勾配値 ∇I であるのに対して，彼らの提案法では，これをガウシアン平滑化後の画像の勾配値 $\nabla(G * I)$ に置き換えることにより，ノイズ除去の性能を高めている．Catte, Alvarez らによるこの修正は PMD の正則化と呼ばれ，これにより式 (2.41) は次のように変更される．

$$\frac{\partial I(\tau, \mathbf{x})}{\partial \tau} = div\Big(g(|\nabla G * I|) \cdot \nabla I\Big), \quad \mathbf{x} = (x, y) \tag{2.51}$$

本論文では，ガウシアン修正項を導入することによりノイズ除去性能を改善した PMD を，正則化 PMD と呼ぶこととする．

ガウシアン修正項による正則化は単純であるが，非常に効果的である．ノイズを重畳した画像を PMD 及び正則化 PMD で平滑化した処理例を図 2.17 に示す．PMD による処理結果からわかるように，オリジナルの PMD では振幅の小さなテクスチャが平滑化される一方で，振幅の大きな salt & pepper ノイズの多くは平滑化されずに残っていることがわかる．これは，エッジ停止関数による重み付けが各画素の勾配値のみから決まることによる弊害である．この場合は，ノイズ重畳画素周辺における勾配値が大きくなった結果として，重み付け係数がゼロに近づき，平滑化が抑えられたことが原因となっている．これに対して，正則化 PMD による処理結果では，オブジェクトのエッジが保存されるだけでなく，ノイズもまた良好に除去できていることがわかる．このように，正則化 PMD は安定した雑音除去特性を持ち，PMD よりも平滑化フィルタに適していると考えられる．

2.4　幾何学的な動的輪郭法

本節では，近年，Caselles-Kimmel-Sapiro[12] によって提案された Geodesic Active Contours (GAC) と呼ばれる幾何学的な動的輪郭法について説明する．GAC は本論文で提案する VOP の自動生成アルゴリズムにおいて重要な役割を果たすため，技術的内容を含めて詳しく説明する．

2.4.1　測地線に基づく動的輪郭法 (GAC)

変分法による定式化

画像平面 Ω 上に与えられた 2 点 A,B を結ぶ曲線 $\mathbf{c}(q)(a \leq q \leq b)$ を考える．ただし，$\mathbf{c}(a) = A$, $\mathbf{c}(b) = B$ とする．このとき，曲線 $\mathbf{c}(q)$ の長さ L を次のように定義する．

$$L(\mathbf{c}) = \int_a^b g(|\nabla I(\mathbf{c}(q))|)|\mathbf{c}'(q)|dq. \tag{2.52}$$

ここで，g は非等方拡散法で使用したエッジ停止関数と同様の性質を持つ関数であり，GAC においては特にエッジ検出関数 (edge detector) と呼ばれる．特に $g(x) \equiv 1$ とすれば，式 (2.52) はユークリッド平面における通常の曲線の長さに他ならない．このように定義した曲線 $\mathbf{c}$ の長さ $L(\mathbf{c})$ は，パラメータの取り方によらない幾何学的な量になっている．

命題 2.1 式 *(2.52)* で定義される曲線の長さ $L(\mathbf{c})$ は，曲線のパラメータの変換に対して不変である．

[証明]　関数 $f : [a_1, b_1] \to [a, b]$，及びその逆関数 $f^{-1} : [a, b] \to [a_1, b_1]$ がともに滑らかで，$f(a_1) = a$, $f(b_1) = b$ となるパラメータ変換 $q = f(p)$ を考える．このとき，曲線 $\mathbf{c_1}(p) = \mathbf{c}(f(p))$ の長さ $L(\mathbf{c_1})$ は，仮定より $dq/dp > 0$ となることを利用すれば，

$$\begin{aligned}
L(\mathbf{c_1}) &= \int_{a_1}^{b_1} g\Big(|\nabla I(\mathbf{c}(f(p)))|\Big)\left|\frac{d\mathbf{c}(f(p))}{dp}\right|dp \\
&= \int_a^b g\Big(|\nabla I(\mathbf{c}(q))|\Big)\left|\frac{d\mathbf{c}(q)}{dq}\right| \cdot \left|\frac{dq}{dp}\right|dp \\
&= \int_a^b g\Big(|\nabla I(\mathbf{c}(q))|\Big) \cdot \left|\frac{d\mathbf{c}(q)}{dq}\right| \cdot \frac{dq}{dp}dp \\
&= \int_a^b g\Big(|\nabla I(\mathbf{c}(q))|\Big)\left|\frac{d\mathbf{c}(q)}{dq}\right|dq \;=\; L(\mathbf{c})
\end{aligned}$$

となり，$L(\mathbf{c})$ に等しい．∎

2 点 A,B を結ぶ曲線 $\mathbf{c}$ で，その長さ $L(\mathbf{c})$ が極小となるものを測地線 (geodesic) と呼ぶ．多くの場合，測地線は 2 点 A,B を結ぶ最短経路と考えてよい．測地線に基づく動的輪郭法 (GAC) とは，2 点 A,B を結ぶ任意の初期曲線 $\mathbf{c}(q)(a \leq q \leq b)$ を，式 (2.52) で定義された長

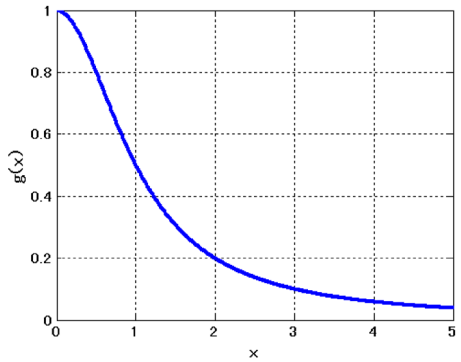


図 2.18: エッジ検出関数 g のプロット

さ L が減少するように時間発展させ，最終的にその極小点である測地線に収束させる手法である．このとき，$\mathbf{c}(a) = \mathbf{c}(b)$ という制約条件を加えることにより閉曲線を扱うことができる．

エッジ検出関数

エッジ検出関数としては，R^+ 上で単調減少となる次のような関数を使用するのが典型的である．

$$g(x) = (1 + (x/K)^2)^{-1}. \tag{2.53}$$

図 2.18 のプロットからわかるように，エッジ検出関数 $g(x)$ は画像の勾配値 (x) が大きくなるにつれて 0 に近づく性質を持ち，水平方向のスケールが定数 K によって決まるようになっている．

このように g を決めるとき，式 (2.52) に基づいて計算される曲線の長さ $L(\mathbf{c})$ は，重み係数 $g(|\nabla I(\mathbf{c}(q))|)$ の影響で，もはやよく知られたユークリッド距離による長さではないことに注意が必要である．エッジ検出関数 g の性質から，オブジェクトのエッジに沿うように走る曲線は，そうでない曲線に比べてその長さが短くなるように設定されている．このため，式 (2.52) を極小化する測地線を探索することにより，オブジェクトの輪郭線を得ることができる．

時間発展方程式の導出

ここでは，式 (2.52) を極小化する測地線を求める変分問題を解くことにより，GAC の時間発展方程式を導出する．

命題 2.2 κ を曲線 $\mathbf{c}(q)$ 上の曲率，N を曲線 $\mathbf{c}(q)$ 上の法線ベクトルとする．式 *(2.52)* で与え

られる長さ L のもとで，初期輪郭曲線を測地線へ収束させる時間発展方程式は，

$$\frac{\partial c}{\partial t} = (\kappa g - \langle \nabla g, \mathbf{N} \rangle)\mathbf{N}, \tag{2.54}$$

で与えられる．なお，ここでは，$\Omega(\subset \mathbf{R}^2)$ 上の関数 $g(|\nabla I|)$ を簡単のため g と表記している．

[証明]　まず，

$$L(t) = \int_a^b g\Big(|\nabla I(\mathbf{c}(t,q))|\Big)\left|\frac{\partial \mathbf{c}}{\partial q}(t,q)\right| dq. \tag{2.55}$$

を t に関して微分する．

$$\begin{aligned}
L'(t) &= \int_a^b g \cdot \Big\langle \frac{\frac{\partial \mathbf{c}}{\partial q}}{\left|\frac{\partial \mathbf{c}}{\partial q}\right|}, \frac{\partial^2 \mathbf{c}}{\partial t \partial q} \Big\rangle dq + \int_a^b \left|\frac{\partial \mathbf{c}}{\partial q}\right| \Big\langle \nabla g, \frac{\partial \mathbf{c}}{\partial t} \Big\rangle dq. && (2.56)\\
&= -\int_a^b \left[\Big\langle g \cdot \frac{\partial}{\partial q}\left(\frac{\frac{\partial \mathbf{c}}{\partial q}}{\left|\frac{\partial \mathbf{c}}{\partial q}\right|} \right) + \frac{\frac{\partial \mathbf{c}}{\partial q}}{\left|\frac{\partial \mathbf{c}}{\partial q}\right|} \Big\langle \nabla g, \frac{\partial \mathbf{c}}{\partial q} \Big\rangle, \frac{\partial \mathbf{c}}{\partial t} \Big\rangle \right] dq + \int_a^b \left|\frac{\partial \mathbf{c}}{\partial q}\right| \Big\langle \nabla g, \frac{\partial \mathbf{c}}{\partial t} \Big\rangle dq. && (2.57)\\
&= \int_a^b \left|\frac{\partial \mathbf{c}}{\partial q}\right| \left\langle \frac{\partial \mathbf{c}}{\partial t}, \nabla g - \frac{1}{\left|\frac{\partial \mathbf{c}}{\partial q}\right|} \frac{\partial}{\partial q}\left(\frac{\frac{\partial \mathbf{c}}{\partial q}}{\left|\frac{\partial \mathbf{c}}{\partial q}\right|} \right) g - \frac{\frac{\partial \mathbf{c}}{\partial q}}{\left|\frac{\partial \mathbf{c}}{\partial q}\right|} \Big\langle \nabla g, \frac{\frac{\partial \mathbf{c}}{\partial q}}{\left|\frac{\partial \mathbf{c}}{\partial q}\right|} \Big\rangle \right\rangle dq && (2.58)\\
&= \int_a^b \left|\frac{\partial \mathbf{c}}{\partial q}\right| \Big\langle \frac{\partial \mathbf{c}}{\partial q}, \nabla g - \kappa g \mathbf{N} - \langle \mathbf{T}, \nabla g \rangle \mathbf{T} \Big\rangle dq && (2.59)\\
&= \int_a^b \left|\frac{\partial \mathbf{c}}{\partial q}\right| \Big\langle \frac{\partial \mathbf{c}}{\partial q}, \langle \nabla g, \mathbf{N} \rangle \mathbf{N} - \kappa g \mathbf{N} \Big\rangle dq. && (2.60)
\end{aligned}$$

したがって，$L(t)$ が最も急速に減少する方向へ時間発展する PDE は，

$$\frac{\partial \mathbf{c}}{\partial t} = (\kappa g - \langle \nabla g, \mathbf{N} \rangle)\mathbf{N} \tag{2.61}$$

で与えられる．∎

2.4.2　レベルセット法による定式化

曲線の時間発展方程式のレベルセット表現

画像平面 $\Omega(\subset \mathbf{R}^2)$ 上の曲線 $\mathbf{c}(t,s)$ に関する時間発展方程式

$$\begin{cases} \dfrac{\partial \mathbf{c}}{\partial t} = F \cdot \mathbf{N} \\ \mathbf{c}(0,q) = \mathbf{c}_0(q) \end{cases} \tag{2.62}$$

のレベルセット法による定式化について考察する．ここで，$\mathbf{N}$ は曲線 $\mathbf{c}$ 上の各点における法線ベクトル，F は Ω 上の各点における速度関数である．レベルセット法では，

$$u(t, \mathbf{c}(t,q)) = 0 \quad \text{for } \forall q, \forall t \geq 0 \tag{2.63}$$

を満たす x-y 平面上の曲面 u を時間発展させることにより，そのゼロレベルセットとして曲線 $\mathbf{c}$ の運動を表現する．このとき，式 (2.62) に相当するレベル曲面 $u(t,x,y)$ の時間発展方程式に関して以下の命題が成り立つ．

命題 2.3 $u(t,x,y)$ が t 及び (x,y) に関して滑らかであるとき，曲面 u に関する時間発展方程式

$$\frac{\partial u}{\partial t}(t,x,y) = F \cdot |\nabla u(t,x,y)|. \tag{2.64}$$

は，そのレベルセットにおいて，式 *(2.62)* と同等である．

[証明]　式 (2.20) の両辺を t に関して微分すると，

$$\frac{\partial u}{\partial t} + \Big\langle \nabla u, \frac{\partial \mathbf{c}}{\partial t} \Big\rangle = 0 \tag{2.65}$$

となる．ここで，式 (2.19) より，

$$\frac{\partial u}{\partial t} + \Big\langle \nabla u, F\mathbf{N} \Big\rangle = 0 \tag{2.66}$$

となるが，曲面 u を閉曲面の内側で負，外側で正となるように選んでおけば，$\mathbf{N} = -\frac{\nabla u}{|\nabla u|}$ であることから，

$$\begin{aligned}
\frac{\partial u}{\partial t} - \Big\langle \nabla u, F \cdot \frac{\nabla u}{|\nabla u|} \Big\rangle = 0 \quad &\Leftrightarrow \quad \frac{\partial u}{\partial t} - \frac{F}{|\nabla u|}\Big\langle \nabla u, \nabla u \Big\rangle = 0 \\
&\Leftrightarrow \quad \frac{\partial u}{\partial t} - \frac{F}{|\nabla u|} \cdot |\nabla u|^2 = 0 \\
&\Leftrightarrow \quad \frac{\partial u}{\partial t} = F \cdot |\nabla u|
\end{aligned}$$

となる．すなわち，

$$\frac{\partial u}{\partial t}(t, \mathbf{c}(t,q)) = F \cdot |\nabla u(t, \mathbf{c}(t,q))| \tag{2.67}$$

が成立する．∎

レベルセット法による GAC の定式化

先に述べた GAC の時間発展方程式 (2.54) は，曲線の時間発展方程式 (2.62) と同様の形をしていることから，レベルセット法を用いたレベル曲面に関する PDE への拡張は容易である．レベルセット法による GAC 定式化として，次の命題が成り立つ．

命題 2.4 レベル曲面 u に関する *GAC* の時間発展方程式は以下の式で与えられる．

$$\frac{\partial u}{\partial t} = g\,|\nabla u|\; div\Big(\frac{\nabla u}{|\nabla u|}\Big) + \langle \nabla g, \nabla u \rangle. \tag{2.68}$$

[証明]　GAC の時間発展方程式 (2.54) は，式 (2.19) において，$F = \kappa g - \langle \nabla g, \mathbf{N} \rangle$ と置いたものに等しいことから，命題 2.4 より，レベル曲面 u に関する時間発展方程式は，

$$\frac{\partial u}{\partial t} = \Big(\kappa g - \langle \nabla g, \mathbf{N} \rangle \Big) \cdot |\nabla u| \tag{2.69}$$

と書ける．ここで，$\kappa = div\Big(\frac{\nabla u}{|\nabla u|}\Big)$，$\mathbf{N} = -\frac{\nabla u}{|\nabla u|}$ であることを利用すれば，

$$\begin{aligned} \text{式 (2.69)} \quad &\Leftrightarrow \quad \frac{\partial u}{\partial t} = \Big(g \cdot div\Big(\frac{\nabla u}{|\nabla u|}\Big) + \langle \nabla g, \frac{\nabla u}{|\nabla u|} \rangle \Big) \cdot |\nabla u| \\ &\Leftrightarrow \quad \frac{\partial u}{\partial t} = g \cdot |\nabla u| \cdot div\Big(\frac{\nabla u}{|\nabla u|}\Big) + \langle \nabla g, \nabla u \rangle \end{aligned}$$

が得られる．∎

式 (2.68) の右辺第 1 項 $g\,|\nabla u|\,\mathrm{div}\Big(\frac{\nabla u}{|\nabla u|}\Big)$ において $g \equiv 1$ とすると，よく知られた平均曲率流 (mean cuevature motion) の方程式 [4, 24] が得られる．すなわち，第 1 項は曲線を平滑化しながら領域の内側に向けて縮退する効果を持っていることがわかる．このときエッジ点に遭遇すると，重み付け係数 g の効果により曲線の時間発展が抑制され，結果的にオブジェクト輪郭線に収束することになる．一方，式 (2.68) の右辺第 2 項 $\langle \nabla g, \nabla u \rangle$ は，凹形に窪んだオブジェクトのエッジに対して，曲線を引き付ける役割を果たすことが知られている [4].

実用上は，snakes の収束を早めるため，しばしば次のような修正項が導入される [4].

$$\frac{\partial u}{\partial t} = g\,|\nabla u|\,\mathrm{div}\Big(\frac{\nabla u}{|\nabla u|}\Big) + \alpha\, g\,|\nabla u| + \langle \nabla g, \nabla u \rangle, \tag{2.70}$$

ここで，α は領域の内側に向かう時間発展を促進させるための正の定数である．α を大きく取るほど，輪郭線の時間発展のスピードが速くなる反面，輪郭線の滑らかさに対する制約が弱くなる．

GAC をレベルセット法で定式化したことにより，式 (2.68) で表される snakes は，これをゼロ交差として含むレベル曲面の時間発展によって実現され，その結果としてトポロジーの変化にも対応できるようになる．このため，初期輪郭として単一閉領域を与えながら複数のオブジェクト領域を切り出すことや，逆に複数の輪郭領域を単一の閉領域に統合することも可能となる．レベルセットに基づく GAC を用いて，単一輪郭から複数オブジェクトを抽出した例を図 2.19 に，複数輪郭を単一オブジェクトに統合した例を図 2.20 にそれぞれ示す．

なお，式 (2.68) の GAC は，有限差分法を用いて数値的に実現することができる．本論文における GAC の実装方法の詳細については，第 7 章のオブジェクト抽出アルゴリズムにあわせて説明する．

2.5　PMD と GAC の数理的な関係

これまで，PDE に基づく画像処理の代表的な事例として，Perona-Malik 拡散法 (PMD) と測地線に基づく動的輪郭法 (GAC) について述べてきた．PMD は画像内のテクスチャを

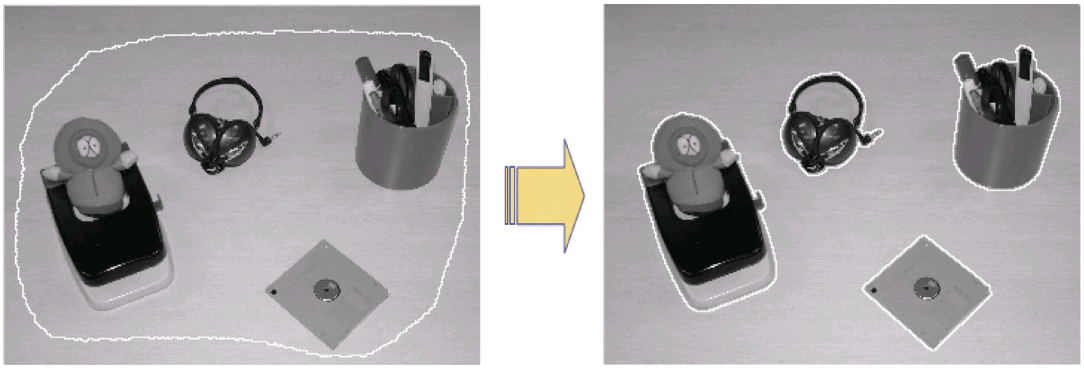


図 2.19: 単一輪郭から複数オブジェクトを抽出した例

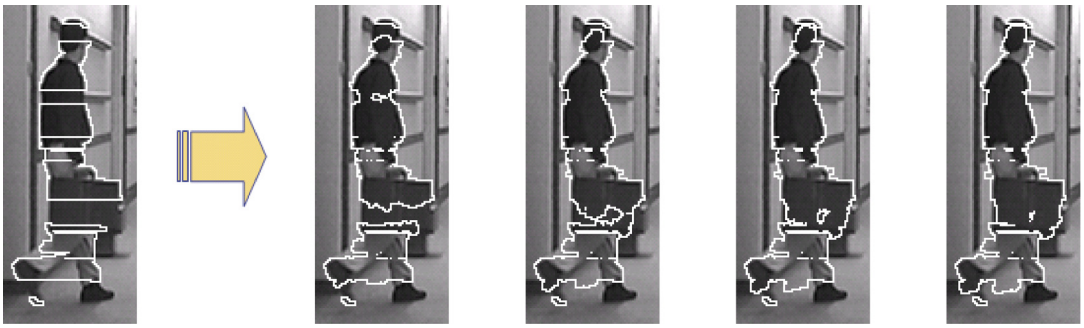


図 2.20: 複数輪郭を単一オブジェクトに統合した例

平滑化しながらエッジを先鋭化する画像の強調処理 (image enhancement) として，GAC は画像内に存在するオブジェクトの輪郭を切り出すオブジェクト検出処理 (object detection) として，それぞれ異なる目的のために導入された手法である．しかしながら，これらは単に PDE に基づく手法であるという形式上の共通点だけでなく，その原理において数学的に密接な関係があることが知られている [76]．以下に，その数理的な関係について補足的に説明しておく．

まず，レベルセット法による GAC の時間発展方程式 (2.68) は，

$$\frac{\partial u}{\partial t} = |\nabla u| \cdot div\left(g \cdot \frac{\nabla u}{|\nabla u|}\right) \tag{2.71}$$

と変形することができる．このことは，f をスカラー場，$\mathbf{v}$ をベクトル場とするとき，

$$div(f \cdot \mathbf{v}) = \langle \nabla f, \mathbf{v} \rangle + f \cdot div(\mathbf{v}) \tag{2.72}$$

が成り立つことを利用すれば，容易に確認することができる．さて，式 (2.71) のレベル曲面

$u(x,y)$ として，特に画像曲面 $I(x,y)$ を採用すると，

$$\frac{\partial I}{\partial t} = |\nabla I| \cdot div\left(g \cdot \frac{\nabla I}{|\nabla I|}\right) \tag{2.73}$$

が得られる．これは，画像曲面 $I(x,y)$ の等高線に関する GAC フローであり，すべての等高線が画像内のエッジ輪郭線へ収束していく様子を表すことから “self-snakes” と呼ばれる．

さて，ここで得られた式 (2.73) を，PMD の時間発展方程式

$$\frac{\partial I}{\partial t} = div\Big(\hat{g} \cdot \nabla I\Big) \tag{2.74}$$

と比較してみる．すると，エッジ停止関数 $\hat{g}$ とエッジ検出関数 g との間に

$$\hat{g} = \frac{g}{|\nabla I|}$$

という関係を与えれば，両者の違いは $|\nabla I|$ による係数の違いのみであることがわかる．すなわち，この考察から，PMD に基づく非等方拡散フローとは，画像曲面におけるあらゆる等高線をその画像内のエッジ輪郭線に向かって時間発展させる GAC フローに他ならない，という数理的に興味深い結論が得られる．

2.6　おわりに

本章では，PDE に基づく画像処理の従来研究について述べた．

まず 2.2 では，PDE に基づく手法の現状を理解するため，画像のスケールスペース表現，画像修復，セグメンテーション，オブジェクト・トラッキング，画像編集への適用事例として，それぞれ従来研究の中から重要なものを選び，その概要を説明した．

続く 2.3, 2.4 では，これらの中から特に Perona-Malik の非等方拡散法（PMD）と測地線に基づく動的輪郭法（GAC）を取り上げた．PMD は第 4 章，第 5 章でローレベル符号化向けのエッジ保存平滑化フィルタを構築する際に，また GAC は第 7 章，第 8 章でミッドレベル符号化向けの VOP 自動生成手法に求められるオブジェクト抽出技術を実現する際に，それぞれ重要な役割を果たす．このため，以降の章で必要となる事項については，基礎的な内容に踏み込んだ詳細な説明を行った．

最後に 2.5 では，一見したところ独立した技術のように見える PMD と GAC が，本質的には同じ数学的基盤の上に成り立っていることを指摘して，本章の総括とした．

第3章　ローレベル符号化における前処理技術 ―プレフィルタを中心に―

3.1　はじめに

インターネットやモバイルネットワークの普及を背景に，低ビットレート環境で映像の送受信を実現する映像圧縮技術の重要性が高まっている．近年では，H.263 や MPEG-4 シンプル・プロファイルをベースとしたソフトウェアエンコーダの登場により，インターネットやモバイルによる映像通信・配信サービスがごく当り前に見られるようになった．しかし，これらのサービスで流通している映像には，ブロック歪み，モスキートノイズ，不自然なフレームスキップなどのディジタル符号化特有の歪みが見られることが多く，その品質は必ずしも十分とはいえない．特に 1-pass の固定ビットレート（CBR）符号化を行った場合にその傾向は顕著である．

映像符号化の前処理としてしばしば使用されるプレフィルタは，このような歪みを緩和し主観的な画像品質を高める効果があることが知られている [37]．プレフィルタの役割は，本来，原画像に重畳したノイズを除去し符号化効率を向上することであったが，数百 Kbps 程度の低ビットレート符号化の場合は，さらに原画像自身の情報量を予め削減することに重点が置かれ，そのような観点に基づくプレフィルタの検討が行われている [84, 85, 90]．

本章では，MPEG-4 や H.263 などの低ビットレート符号化におけるプレフィルタが，低ビットレート動画像の画質を改善する上で効果的であることを示す研究事例として，著者が提案する VBV バッファに基づくプレフィルタ制御法について説明する．本提案法はピクチャスキップを改善し，主観画質を高めることをねらいとしている．ただし，ここで使用するプレフィルタは，パラメータにより平滑化効果を制御できる従来の線形フィルタ群によって構成されるものである．

本章の結びとして，上の提案法によるシミュレーション結果を踏まえ，画像特徴を保存して主観的な画質劣化を極力抑えながら従来より多くの情報量を削減可能とする非線形な平滑化プレフィルタの必要性ついて問題提起を行う．

3.2　ピクチャスキップ抑制のためのプレフィルタ制御法

3.2.1　背景と目的

MPEG-4 は非常に高い圧縮率の符号化方式であるが，帯域の保証されない低レート環境で映像を CBR 符号化する場合，ブロック歪み，モスキート，ピクチャスキップなど，ディジタル符号化特有の歪みが顕著に現れ，その品質は必ずしも十分とはいえない場合が多い．VBV(Video Buffering Verifier) バッファの破綻が原因となって発生するピクチャスキップもまた，視聴者にとって著しい視覚妨害となりうるディジタル歪みのひとつである．

MPEG-4 では通常，量子化スケール Qp を $1 \leq Qp \leq 31$ の範囲でマクロブロック毎に変化させることで発生符号量の制御を行うが，発生符号量の多いシーンが続くと符号量を抑えるため量子化スケールが上限値付近に張りついてしまい，顕著なブロック歪みが観測されるようになる．さらに，量子化スケールによる制御で符号量を抑制しきれなくなると，送信側 VBV バッファのオーバフローを回避するため，ピクチャスキップが発生する．特に発生符号量の多い I ピクチャではその傾向が強く，スポーツ等のスピード感が重視される映像コンテンツの場合，視聴者には著しい視覚妨害として認知されてしまうという問題がある．

上に述べた問題を回避し画像品質の向上を実現するためには，ブロック歪みが目立たない程度に量子化スケールを細かく保ちながら，一方で VBV バッファの破綻を回避する符号化制御方法を確立することが重要である．そこで本章では，この相反する要求条件を満足させるため，パラメータにより強度を可変に設定できる平滑化プレフィルタ群を導入し，VBV バッファ量と画像の複雑さに応じてその強弱を適応的に制御することにより，MPEG エンコーダの符号量制御を支援する方法を提案する．すなわち本提案は，各フレームへのビット割り当てが厳しい条件下での主観画質の向上，特にピクチャスキップの軽減を主な目標とするものである．

同様の図式に基づくプレフィルタ制御法の一例として [97] があるが，これは画像の特徴や符号化モードの選択状況に応じて時空間フィルタを切り替えるものであった．これに対し，本章のプレフィルタ制御は，VBV バッファの状態に応じて決まる「Qp 下限値」に基づきプレフィルタの強度を制御する点が特徴的である．ただし今回は空間フィルタのみを扱い，その制御は I ピクチャ符号化時にのみ行われるものとする．

3.2.2　フィルタ制御の原理

ここではプレフィルタの制御法について説明する．まず，VBV バッファの破綻を回避するための Qp 下限値を導入し，ある VBV バッファ残量と画像の複雑さの下での Qp 下限値の計算方法について述べる．本方式では，I ピクチャ毎に Qp 下限値を更新するが，この際 Qp 下限値が適正な値になるようにプレフィルタの強さを決定する．次に，簡単なパラメータにより平滑化の強弱をコントロールできるフィルタ群を導入する．最後に，Qp 下限値に基づくフィルタの制御フローについて述べる．

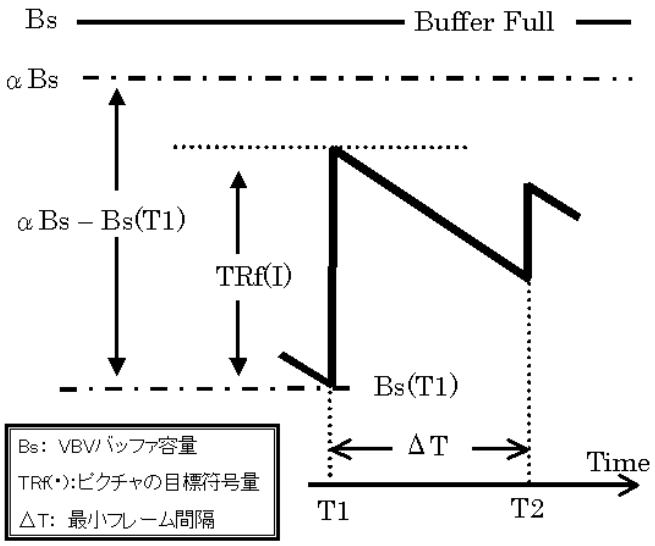


図 3.1: 送信側 VBV バッファモデル

Qp 下限値の算出法

容量 Bs の送信側 VBV バッファを持つ MPEG-4 符号化器を考える．ある I ピクチャが時刻 $T1$ に符号化されるとして，その直前のバッファ占有量を $B(T1)$ と表す．この送信側 VBV バッファの様子を図 3.1 に示す．

このとき，I ピクチャに対する目標符号量を $TRf(I)$，発生符号量を $Rf(I)$ と書くことにすれば，ピクチャスキップを回避するための必要条件は，

$$Rf(I) < min\{TRf(I), \alpha Bs - B(T1)\} \tag{3.1}$$

となる．ここで，α はピクチャスキップを行うかどうかのしきい値を決める比率を表し，$0.5 < \alpha \leq 1$ である．

次に，量子化パラメータ Qp と発生符号量 R との関係を示すレート－歪みモデルを利用する．ここでは MPEG-2 のテストモデル [33] に習い，以下の簡単なモデルを採用する．

$$X = Qp \cdot Rf \tag{3.2}$$

ここで，X は符号化すべき I ピクチャの複雑さ指標 [99] である．

さてこのとき，レート－歪みモデル (3.2) のもとで，条件 (3.1) を満足する最小の Qp のことを Qp 下限値として定義する．式 (3.1)，(3.2) より，Qp 下限値 Qp_{min} の算出式は以下で表される．

$$Qp_{min} = ceil(X/min\{TRf(I), \alpha Bs - B(T1)\}) \tag{3.3}$$

プレフィルタ制御の基本方針

I ピクチャ符号化における Qp の平均値が前節で導入した Qp 下限値以下であれば，理想的にはピクチャスキップは発生しないと期待できる．しかし，与えられた VBV バッファ残量が少ない場合や I ピクチャの複雑度が高い場合には，Qp_{min} の計算値が大きくなり，非現実的な結果が得られることがある．このような場合は，割り当てられた符号量に対して符号化すべき画像が複雑すぎることを示唆しているため，プレフィルタにより画像を平滑化して情報量を削減することで，バッファオーバフローに伴うピクチャスキップを回避する．

以上の目的を実現するには，パラメータにより平滑化の度合いを制御できるフィルタ群をプレフィルタとして採用すればよい．フィルタ群としては様々な候補が考えられるが，本章では，以下の 3×3 畳み込みマスクで与えられる 2 次元フィルタ群を使用することとした；

$$LSF(b) = \begin{pmatrix} 1 & 9-b & 1 \\ 9-b & (9-b)^2 & 9-b \\ 1 & 9-b & 1 \end{pmatrix} / (11-b)^2 \tag{3.4}$$

本フィルタはパラメータとして b を有し，$1 \leq b \leq 8$ の範囲でパラメータを変動させることで，異なる強さの線形平滑化フィルタが得られる．ここでは，b が増加するにつれて，平滑化が強くなるようにパラメータを設定した．

プレフィルタ制御フロー

ここでは図 3.2 に示す Qp 下限値に基づくフィルタの制御フローについて述べる．

まず GOV(Group of VOP) 先頭の I ピクチャを符号化するにあたり，プレフィルタ強度 b の初期値を決定する．符号化する I ピクチャがシーケンスの先頭またはカット点であれば，デフォルト値として指定される b_{start} が初期値として使われる．それ以外の場合は，直前の GOV で使用された b_{prev} が初期値として使用される．

こうして決まった初期 b 値により，原画像 Im をフィルタ $LSF(b)$ で平滑化し，得られた画像を Im' に格納する．すなわち，$Im' = LSF(b) * Im$ とする．そして，格納された平滑化画像 Im' に対して，複雑さ指標 X を推定する．本章では，Qp=15 固定で符号化して得られる符号量からその推定値 $\hat{X}$ を算出することとした．

次に上述した方法で，複雑さ指標の推定値 $\hat{X}$ 及び VBV バッファ残量から $\widehat{Qp}_{min}$ 値を算出する．ここで，算出した $\widehat{Qp}_{min}$ が大きすぎる場合には，バッファを破綻することなく符号化することが困難であるため，プレフィルタを強くして，画像の複雑さをさらに低減することで対処する．具体的には，$\widehat{Qp}_{min}$ の取り得る適正な値として，その上限値 TH_{high} と下限値 TH_{low} を設定し，$TH_{low} \leq \widehat{Qp}_{min} \leq TH_{high}$ を満たさない場合には，パラメータ b を調整してプレフィルタを再設定する．本章では，$\widehat{Qp}_{min}$ が上限値 TH_{high} を超える場合には $b = b + 1$ として平滑化フィルタを 1 段階強い方向へ制御し，下限値 TH_{low} に満たない場合にはプレフィルタをかけすぎている恐れがあるため，$b = b - 1$ として強度を 1 段階弱くすることとした．

以上のプレフィルタの調整は，条件 $TH_{low} \leq \widehat{Qp}_{min} \leq TH_{high}$ を満たすようになるまで複数回ループをまわすことによって実現される．ただし今回は，フィルタ制御のパラメータ b が上限値または下限値に達した時点で，$\widehat{Qp}_{min}$ が適正な範囲に入っていなくてもループを抜けることとした．

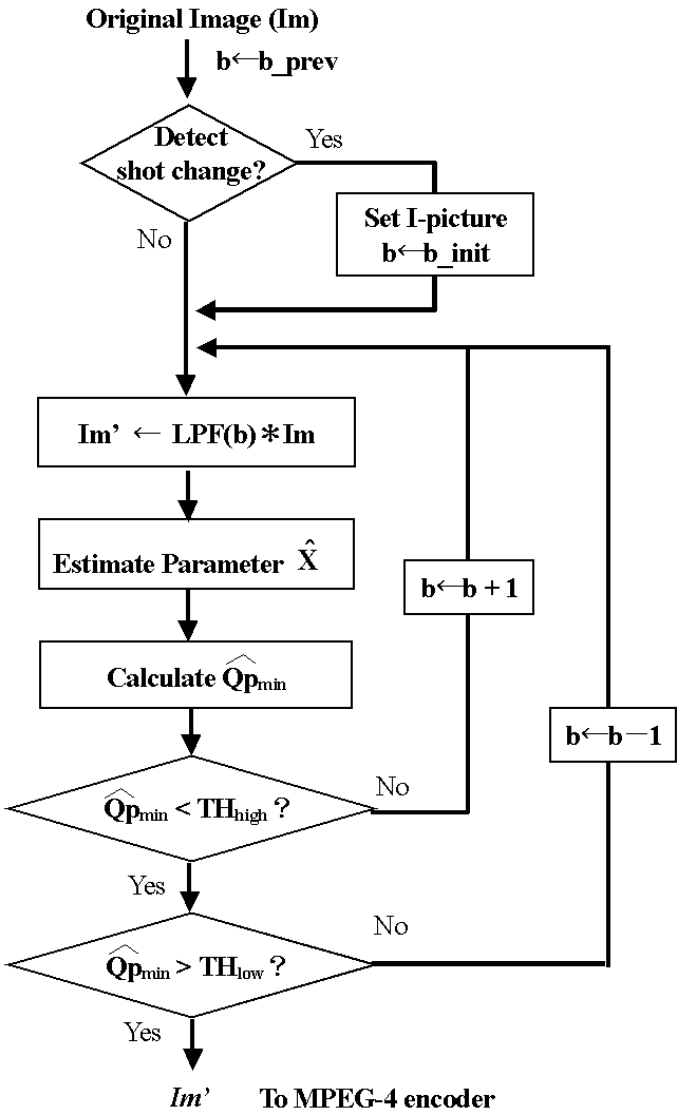


図 3.2: プレフィルタの制御フロー

3.3　シミュレーションによる検証

本節では，著者らの提案するプレフィルタ制御法を用いて MPEG-4 符号化シミュレーションを実施し，その効果を確認する．実験に使用したテスト画像シーケンスは表 3.1 の通りである．また，符号化においては MPEG-4 VM17 に基づくソフトウェアエンコーダ [35] を使用し，I/P ピクチャのみを使用する”Simple Profile”に準拠した．本実験における符号化パラメータ等の諸条件を表 3.2 に示す.

表 3.1: テストシーケンス

入力画像	SIF (352x240), 4:2:0, 890frames
コンテンツ	5 種類の MPEG 標準動画を連結したシーケンス（約 30 秒） 001〜150: Akiyo (akiyo) 151〜300: Football (ftbl) 301〜450: Susie (ssie) 451〜600: Flower garden (flow) 601〜890: Stefan (stefan)

表 3.2: シミュレーション諸元

アルゴリズム	MPEG-4 Simple Profile
符号化	180kbps, 30fps I ピクチャ間隔: N = 30 動き補償: 4MV モード ON
レート制御	マクロブロック・レート制御 バッファサイズ: 2 sec ピクチャスキップ閾値: $\alpha = 0.8$
フィルタ制御	フィルタ係数: $1 \leq b \leq 8$ Qp_{min} の変動域: [TH_low, TH_high] = [10, 20]

3.3.1　シミュレーション結果

本実験の目的は，複雑さ指標の高い画像（ftbl, flow, stefan）と低い画像（akiyo, ssie）を交互に並べた画像シーケンスにおいて，フィルタ制御が適切に行われていること，及びその結果としてピクチャスキップが低減しているかどうかを確認することである．本実験では，

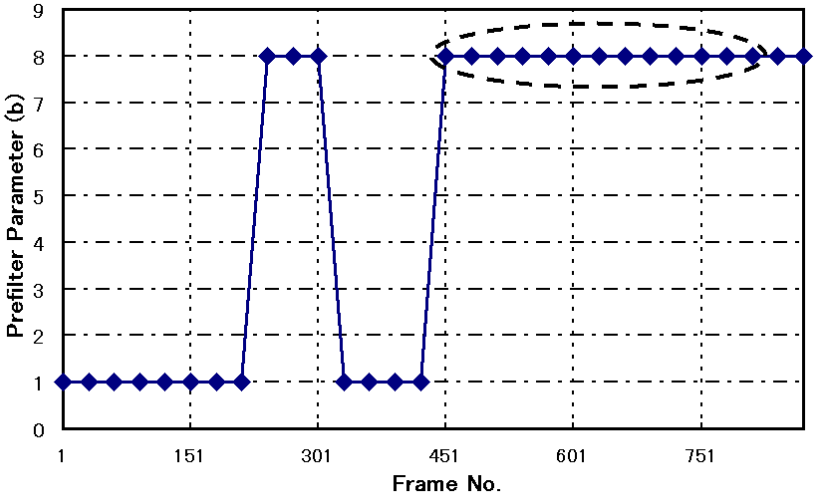


図 3.3: プレフィルタパラメータの変動

フィルタ制御を行わない VM17 による符号化と著者らの提案するフィルタ制御と VM17 を組み合わせた符号化の結果を比較した．

提案方式によるプレフィルタ強度 b の変動の様子を図 3.3 に示す．プレフィルタ強度は，複雑さの低い akiyo と ssie では弱く，複雑さの高い ftbl の後半や flow, stefan では強くなっており，適切なフィルタ設定がなされていることがわかる．また，表 3.3 はピクチャスキップの発生回数をまとめたものである．提案方式は VM17 に比べてピクチャスキップの発生を抑えられることがわかる．

表 3.3: ピクチャスキップ発生回数 (fps の実効値)

	画像シーケンス				
	akiyo	ftbl	ssie	flow	stefan
VM17	0 (30fps)	55 (19fps)	3 (29.4fps)	101 (9.8fps)	161 (13.3fps)
VM17 ＋提案法	0 (30fps)	42 (21.6fps)	4 (29.2fps)	0 (30fps)	23 (27.6fps)

3.3.2　考察及び今後の課題

図 3.4 に両方式による VBV バッファレベルの変動の様子を示す．提案法は VM17 に比べてバッファレベルにゆとりがあり，特に flow においてその効果が顕著である．しかし，空

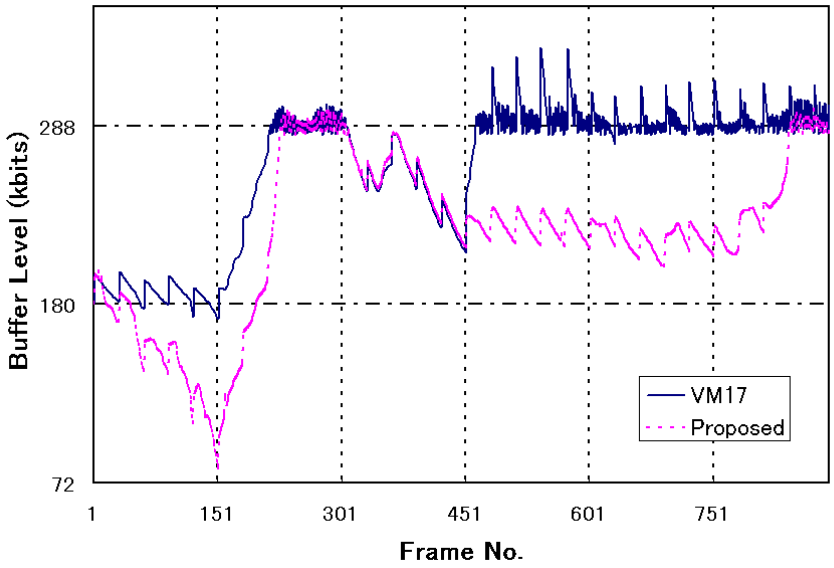


図 3.4: バッファレベルの変動

間フィルタのみを使用していることから，ftbl,stefan の後半部のように動きの激しいシーンではフレーム間予測誤差の低減が不十分であり，結果的にピクチャスキップの発生が見られた．このような変化の激しいシーンにおけるピクチャスキップをさらに改善するためには，フレーム間予測誤差を低減する時間フィルタの導入やフィルタとは別の原理に基づくフレームスキップのコントロール手法を検討する必要がある．

また，図 3.3 を見ると，フィルタの値として $b = 1$ と $b = 8$ を 2 値的に変化していることから，その間をあまり細かい間隔で設定することにはあまり意味がない結果となった．また，図 3.3 の点線で囲われた I ピクチャでは，バッファ残量が十分であったにもかかわらず，フィルタ強度を最大の $b = 8$ に設定しても条件 $TH_{low} \leq \widehat{Qp}_{min} \leq TH_{high}$ を満たすことができなかった．このことから，更に強い平滑化を実現するフィルタを導入する必要がある．

3.4　おわりに

結びとして，本章で取り上げた研究事例のまとめを述べ，それに基づく問題提起を行う．

本研究事例のまとめ

本章では，I ピクチャ符号化時に発生するピクチャスキップを抑制するため，VBV バッファ残量と入力画像の複雑さ指標から算出される Qp 下限値に基づき，プレフィルタの強度を適応的に制御する方法を提案した．シミュレーション実験を通して提案法を評価したところ，バッファの破綻に伴うピクチャスキップ数が減少し，主観画像品質を高める効果が得ら

れることを確認した．

本研究事例に基づく問題提起

本研究事例で使用した線形平滑化フィルタでは，最も平滑化の強い $b = 8$ に設定した場合，エンコーダに入力する段階で画像全体がかなりぼやけてしまうことが確認できる．このため，今回の実験では，これ以上平滑化を強くするとかえって主観画像品質を損なう可能性があるとの判断から，$b = 8$ を上限値と設定している．

しかしながら，本実験結果において，651 フレーム以降では常に b が上限値にはりついてしまい，最終的にはフレームスキップが発生していることを考えると，現実には上限値 $b = 8$ よりも情報量削減効果を高めたいケースが十分に起こり得ることがわかる．このことから，本実験で用いた線形平滑化フィルタほどには主観画質を損なうことなく，同時により多くの情報量を削減することができるプレフィルタの導入が望まれる．

次章では，以上の問題提起に基づき，視覚的に重要な画像特徴であるエッジを保存しながら，その他のテクスチャについては選択的に平滑化する非線形フィルタに関する検討を行う．すなわち，本実験の線形フィルタ群のように，パラメータで平滑化の強弱をコントロールできるエッジ保存平滑化フィルタ群を，非等方拡散法 (PMD) を用いて構築する方法について検討する．

第4章　非等方拡散法に基づく空間平滑化プレフィルタの構築

4.1　はじめに

低ビットレート映像符号化において使用されるプレフィルタは，原画像の情報量を予め削減することにより，符号化歪みを緩和し主観的な画像品質を高める効果がある．情報量を効果的に削減するには強い平滑化効果を持つフィルタの導入が不可欠であるが，この際生じる問題として，オブジェクトエッジのような視覚的に重要な特徴までが平滑化されてしまう点が挙げられる．このため，従来の映像符号化システムにおいては，メディアンフィルタ，モルフォロジカルフィルタなどのエッジ保存性を持つ非線形なフィルタが使用されていた [39]．また，Sobel Operator 等による簡易なエッジ検出に基づき，フィルタの平滑化の強弱を適応的に切り替えるなどの試みも行われていた [41, 42, 43]．

一方，コンピュータビジョンやパターン認識の分野においては，エッジ保存平滑化に関する研究が数多く報告されている．初期の研究で特に重要なものとして，非等方拡散法 [70]，Adaptive Smoothing [73] がある．近年では，同様の効果を実現する手法として Bilateral Filter [82] や Mean Shift [16] が提案され注目を集めている．これらの手法はいずれも数学的に密接な関係があることが知られており [5, 6]，共通した長所として，従来のプレフィルタに比べて非常に強い平滑化が実現できることが挙げられる．

本章では，低ビットレート映像符号化への適用を想定したエッジ保存型の強力な空間平滑化プレフィルタを提案することを目的として，特に非等方拡散法をベースとしたフィルタの構築方法について検討を行う．今回，非等方拡散法を採用するのは，各ステップの処理がシンプルで，対象画素の 4-近傍の値のみから計算可能であること，後述する 2 種類のパラメータ $iter, T$ により，エッジ保存平滑化の段階的な制御が容易であること等が主な理由である．画像圧縮の観点から見ると，Perona-Malik らの手法 [70] はノイズ除去の性能やエッジの強調効果などの点でまだ改善の余地がある．そこで本章では，非等方拡散法が空間平滑化プレフィルタとして望ましい性質を持つための条件を明らかにし，実際にその条件を満たす理想的な非等方拡散法を構築する．

本章の構成は次の通りである．まず 4.2 節では，PMD 及び正則化 PMD による平滑化の挙動を調べ，プレフィルタとして使用する際の問題点を明らかにする．続く 4.3 節では，提示した問題点を改善するためにエッジ停止関数が満たすべき必要条件を明らかにし，その条件に基づいた拡散法の導出を行う．4.4 節では，4.3 節で提案した拡散法の性能を計算機シミュレーションを通して検証し，最後に 4.5 節で本章の総括を行う．

4.2　非等方拡散法による平滑化の挙動とその問題点

PMD 及び正則化 PMD の平滑化の挙動を実験により確認すると共に，プレフィルタとしてこれを適用する際の問題点を提示する．本実験で使用するテスト画像を図 4.1 に示す．画面中央に見られる 3 つの垂直帯領域のレベルはそれぞれ $64, 128, 192$ であり，各領域にはそれぞれ輝度の高さが $\delta = 48, 32, 16$ となる白線が 8 画素間隔で重畳されている．本実験の目的は，これらの白線をテクスチャとみなし，PMD 及び正則化 PMD を用いてこのテクスチャを選択的に平滑化できることを示すことである．

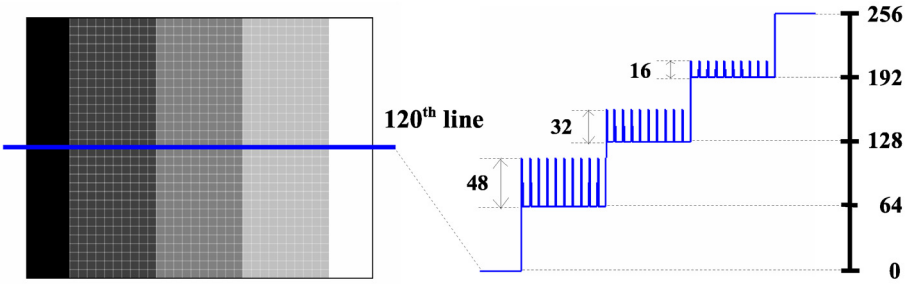


図 4.1: テスト画像

なお，以下の実験では，PMD を正則化する際に必要となるガウシアンフィルタの実装において，3 タップのフィルタ $[1, 2, 1]/4$ を水平・垂直方向に連続して施して得られる図 4.2 に示す 2 項ガウシアンマスク (Binomial Gaussian Mask) を使用する．

1	2	1
2	4	2
1	2	1

/ 16

図 4.2: 2 項ガウシアンマスク

PMD の場合

エッジ停止関数として g^{BS} 及び g^{PM} を用いて，PMD による平滑化の挙動を調べる．テスト画像において，輝度値 δ のライン上の点における勾配は $|\nabla I| = \delta$ であることから，$\epsilon = 0.1$ の下で $T = \delta$ $(\delta = 48, 32, 16)$ と設定する．各 T に対して平滑化処理を行い，各帯領域上の白線が反復回数 $iter$ の増加につれてどの様に平滑化されていくかを観察する．

エッジ停止関数 g^{BS}, g^{PM} を用いて $iter = 20$ まで繰り返し演算を行い，テスト画像の第 120 ラインが平滑化される様子を図 4.3，図 4.4 に示す．図の上段，中段，下段がそれぞれ $T = 16, 32, 48$ による平滑化結果を表す．また，各行のプロットは左から反復演算数 $iter = 5, 10, 20$ における結果を表している．本図から共通して言えることとして，いずれの場合も各垂直帯領域の境界線は平滑化されず，エッジが保存されていることが挙げられる．また，エッジは保存されるだけでなく，その波形から若干のエッジ強調効果があることが見て取れる．

これに対して各帯領域上の白線については，T の値に応じてある高さ以下の白線がテクスチャとして優先的に平滑化されている．図 4.3 のエッジ停止関数 g_3 を用いた平滑化では，例えば $T = 32$ の場合には，$\delta > 32$ の高さを持つ左の帯領域上の白線のみが平滑化されず，元の高さのまま保存されていることがわかる．このことから，平滑化したいテクスチャの振幅を示すパラメータとして，正規化定数 T を使用できることがわかる．

一方，図 4.4 のエッジ停止関数 g^{PM} を用いた平滑化においても同様の傾向が見られるが，この場合，テクスチャの振幅に基づく選択性は g^{BS} の場合ほど厳密ではない．例えば，$T = 32$ にもかかわらず，$\delta = 48$ の白線までがかなり平滑化の影響を受けていることがわかる．これは，エッジ停止関数 g^{PM} の値が，g^{BS} とは異なり $x > \delta$ の場合にも決してゼロにならないため，反復回数を重ねると，振幅が δ 以上の信号に対しても平滑化の影響を完全には避けられないことが原因である．このような場合は ϵ の値をさらに小さくとることによって，より厳密な選択的テクスチャ平滑化を実現できるようになると考えられる．

以上の考察から，PMD においては，パラメータ $T(\epsilon)$ をうまく調整することによって，エッジ保存性と輝度勾配に基づく選択的テクスチャ平滑化が同時に実現できることがわかった．パラメータ設定の指針としては，より平滑化を進行させるためには $iter$ を大きく取り，より強いエッジまで平滑化したい場合には T を大きく取ればよいことになる．この性質は，PMD を低ビットレート映像符号化のプレフィルタとして応用する際に有効である．例えば，エンコーダのバッファ残量に応じてパラメータ T,$iter$ を変えて平滑化の強弱をコントロールするなど，レート制御の要請に対応可能なフィルタ群を構築することが可能となる．

正則化 PMD の場合

次に，PMD よりも優れた雑音除去特性を持つ正則化 PMD による平滑化の挙動を調べる．本実験では，エッジ停止関数として g^{PM} を用いる．振幅 δ の白線上の点における勾配は $\|\nabla I\| = \delta$ であるが，正則化 PMD でエッジ停止関数の評価を行うのは $\|\nabla I\|$ ではなく，$\|\nabla G * I\|$ であることを考慮する必要がある．本章では，正則化のためのガウシアンフィルタの実装において，3 タップのフィルタ $[1, 2, 1]/4$ を水平・垂直方向に連続して施して得られる図 4.2 の 2 項ガウシアンマスク (Binomial Gaussian Mask) を使用している．このため，本テスト画像の場合，$\|\nabla G * I\| = \delta/4$ と評価されることになる．以上をふまえ，$T = \delta/4$ $(\delta = 48, 32, 16)$ に対して平滑化処理を行い，各帯領域上の白線の平滑化状況を観察することとした．

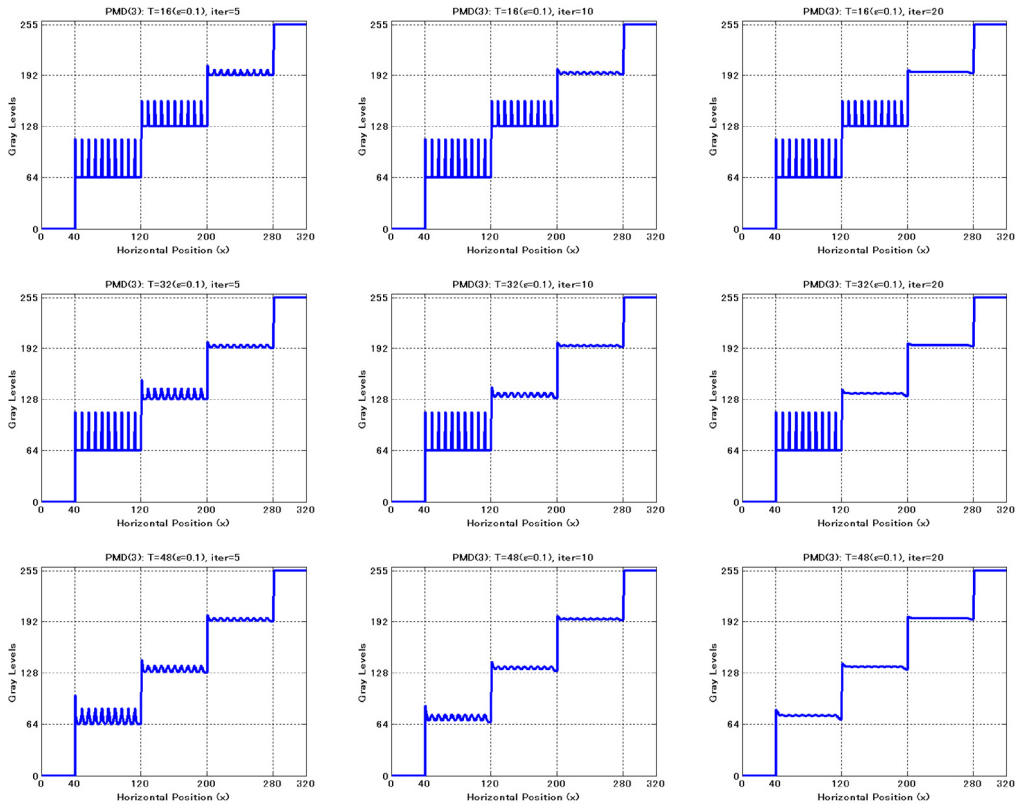


図 4.3: PMD (with $g^{BS}(x)$) による第 120 ラインの推移（上段から $T = 16, 32, 64(\epsilon = 0.1)$, 各段の左から $iter = 5, 10, 20$)

上記の各 T に対して，$iter = 20$ まで繰り返し演算を行った第 120 ラインの結果を図 4.5 に示す．図の各行（上段，中段，下段）がそれぞれ $T = 4, 8, 12$ による平滑化結果を表し，各行の画像は右に行くほど反復演算数 $iter$ が大きくなり，左から $iter = 5, 10, 20$ に相当する．まず本図から分かることとして，いずれの場合にも各垂直帯領域の境界線は平滑化されず，エッジが保存されていることが挙げられる．これに対して，各帯領域上の白線は T の値に応じてある程度選択的に平滑化されていることがわかる．例えば $T = 8$ の場合には，$\delta > 32$ の高さを持つ左の帯領域上の平滑化の進行が他の 2 つの帯領域と比較して遅くなっており，20 回の反復後には左の帯領域上にのみ白線がまだ残った状態になっている．

問題点

以上の実験により，PMD，正則化 PMD のいずれを用いても，エッジ保存性とテクスチャの振幅に基づく選択的テクスチャ平滑化が同時に実現できることがわかる．正則化 PMD の方が雑音除去特性に優れることを考慮すると，プレフィルタとしては正則化 PMD を選択す

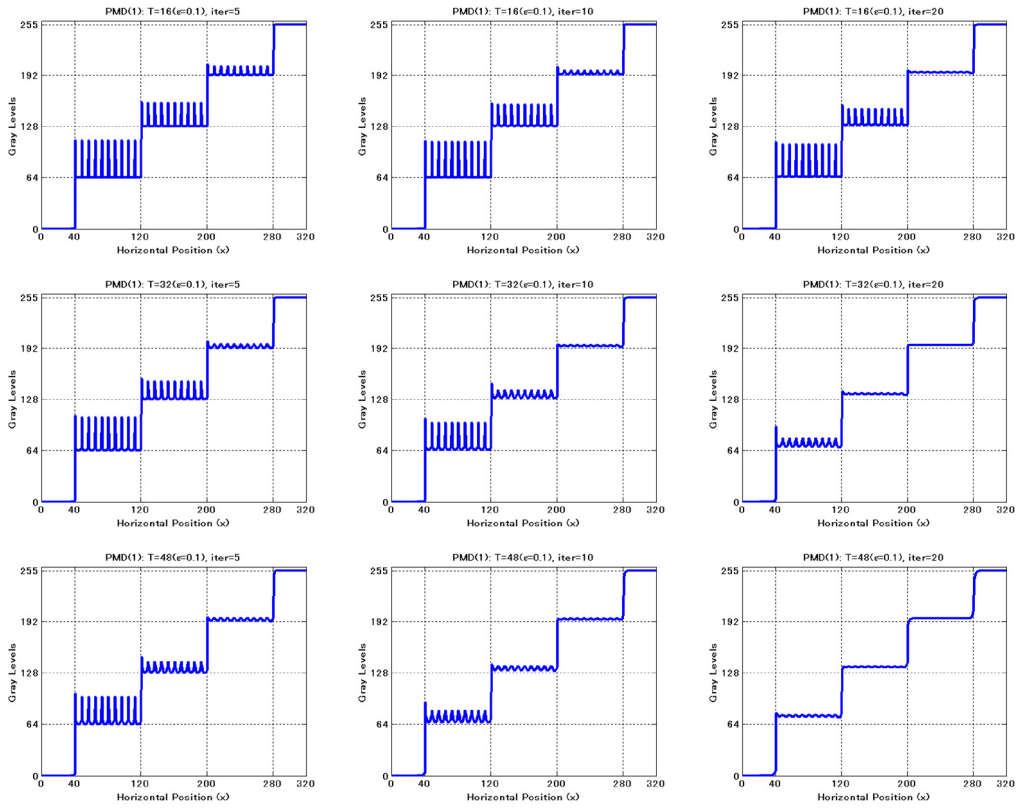


図 4.4: PMD (with $g^{PM}(x)$) による第 120 ラインの推移（上段から $T = 16, 32, 64(\epsilon = 0.1)$, 各段の左から $iter = 5, 10, 20$)

る方が妥当である．

しかし一方で，正則化 PMD を低ビットレート映像符号化のプレフィルタへの適用する場合の問題点として，エッジ強調効果による影響を指摘することができる．図 4.3，図 4.4 におけるステップエッジの波形を見ると，PMD による拡散には若干のエッジ強調効果が伴うことがわかる．正則化 PMD の場合にはこの傾向がさらに著しく，図 4.5 の波形では，繰り返し演算を行うにつれてステップエッジ部分でピークが発生してしまうことが確認できる．これは，画像符号化の観点から言えば，符号量の増加，符号化歪みの発生につながることとなり，主観画質への影響を考慮しても好ましくない現象であるといえる．

Perona-Malik らの PMD は，もともとエッジ抽出の前処理手法として，テクスチャの平滑化とエッジの強調を同時に行うことを目的としていた [70]．このことを考慮すると，彼らが異なる目的の下に選択したエッジ停止関数 g^{PM} をそのまま使用していることに問題があったと結論できる．以下では，このような問題点を踏まえ，非等方拡散法を用いて我々の目的に見合った平滑化効果を実現するために必要となるエッジ停止関数を一から構築するための検討を行う．

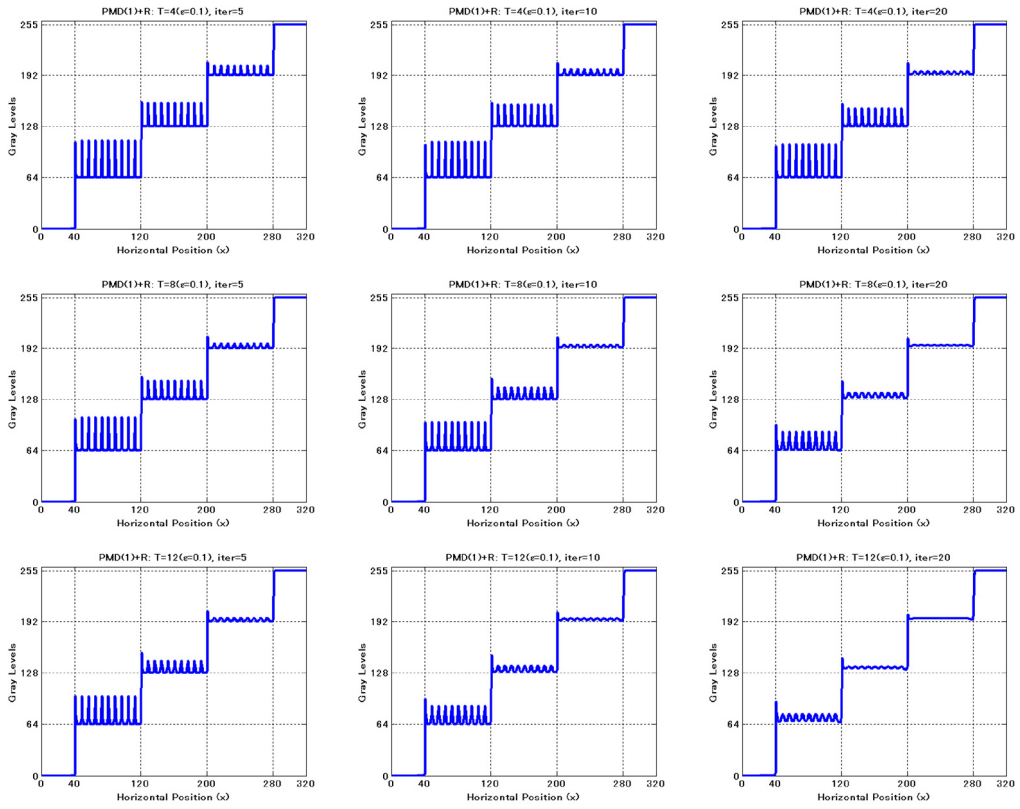


図 4.5: 正則化 PMD (with $g^{PM}(x)$) による第 120 ラインの推移（上段から $T = 4, 8, 12$，各段の左から $iter = 5, 10, 20$)

4.3　PMD に基づくプレフィルタの構築

本節では，適切なエッジ停止関数を導入することにより，次に挙げる 2 つの項目を同時に満足する PMD を構築する．

(1) 振幅の大きい雑音（salt & pepper ノイズ等）を除去できること

(2) エッジ部におけるエッジ強調効果が起こらないようにすること

まず，1 つめの項目を満たすエッジ停止関数を選定することにより，正則化技法を導入することなく，雑音除去特性の高いフィルタを構築することが可能となる．また，このことは PMD における反復ステップの演算量削減にもつながる．2 つめの項目は，符号量や符号化歪みを低減するために必要であるが，主観画質の観点からもその重要性を指摘することができる．一般に，平坦部の平滑化とエッジの強調が過度に行われた画像は，絵画的な印象の不自然な画像になってしまうことが多い．このため，より自然な印象の画像を得るためには，エッジを強調しすぎないことが重要である．

以下では，目的とするエッジ停止関数をどのように決定すればよいかを理論的に考察し，その構築手順の提案を行う．

4.3.1　準備

後の議論に必要となるいくつかの命題について述べる．まず，PMD とエネルギー最小化問題との関係について以下の命題が成り立つ．

命題 4.1 矩形領域 $\Omega \subset \mathbf{R}^2$ で定義された画像 $I(x,y)$ に関するエネルギー最小化問題

$$\boldsymbol{E}(I) := \min_{I} \int_{\Omega} f(\|\nabla I\|) d\Omega \tag{4.1}$$

の解を求める勾配降下フローは非等方拡散であり，そのエッジ停止関数は

$$g(x) := f'(x)/x \tag{4.2}$$

と書ける．ここで，関数 $f \in C^2[0,\infty)$ は誤差ノルムで，$f(x) > 0$, $f'(x) > 0$ $(\forall x > 0)$, および $f'(0) = 0$ を満たす単調増加関数である．

[証明]　画像 $I(x,y)$ の x 方向の微分を I_x，y 方向の微分を I_y と表す．このとき，$F(I_x, I_y) = f(\|\nabla I\|) = f(\sqrt{I_x^2 + I_y^2})$ とおけば，式 (4.1) の解が満たす Euler-Lagrange 方程式は，

$$\mathbf{E}'(I) = -\frac{\partial}{\partial x}\Big(\frac{\partial F}{\partial I_x}\Big) - \frac{\partial}{\partial y}\Big(\frac{\partial F}{\partial I_y}\Big) = 0$$

となる．この解を数値的に求めるには，以下の勾配降下フローを考えればよい．

$$\frac{\partial I(x,y;t)}{\partial t} = \frac{\partial}{\partial x}\Big(\frac{\partial F}{\partial I_x}\Big) + \frac{\partial}{\partial y}\Big(\frac{\partial F}{\partial I_y}\Big) \tag{4.3}$$

ここで，

$$\frac{\partial F}{\partial I_x} = f'(\|\nabla I\|)\frac{\partial \|\nabla I\|}{\partial I_x} = f'(\|\nabla I\|)\frac{I_x}{\|\nabla I\|}$$

であり，y についても同様の式が成り立つことから，式 (4.3) は

$$\begin{aligned}\frac{\partial I}{\partial t} &= \frac{\partial}{\partial x}\Big(\frac{f'(\|\nabla I\|)}{\|\nabla I\|}I_x\Big) + \frac{\partial}{\partial y}\Big(\frac{f'(\|\nabla I\|)}{\|\nabla I\|}I_y\Big) \\ &= div\Big(\frac{f'(\|\nabla I\|)}{\|\nabla I\|}\nabla I\Big) \end{aligned} \tag{4.4}$$

と書ける．ゆえに，PMD の定義式 (2.41) より，式 (4.4) はエッジ停止関数を $g(x) = f'(x)/x$ とする非等方拡散方程式に他ならない． ∎

最適化問題 (4.1) の大域解は明らかに定数画像 $I(x, y) = const$ である．しかし，エネルギー最小化問題を数値的に解く際に使用される勾配降下フローは，対応するエネルギー曲面の形状によっては局所解に収束してしまい，必ずしも大域解に辿りつけるとは限らない．このように，出発点としての初期画像が変わると様々な局所解に捕えられてしまう拡散法は “ill-posed”，逆にどのような初期画像から出発しても，正しく大域解に収束する拡散法は “well-posed” と呼ばれる．

エッジ停止関数 (4.2) に対応する PMD が well-posed となるための必要条件として，以下の命題が知られている．

命題 4.2 エッジ停止関数 *(4.2)* により実現される *PMD*

$$\frac{\partial I(\boldsymbol{s}, t)}{\partial t} = div\Big(\frac{f'(\|\nabla I\|)}{\|\nabla I\|}\nabla I\Big) \tag{4.5}$$

が *well-posed* となるための必要条件は，誤差ノルム f が $f'(\infty) \neq 0$ を満たすことである．

[証明]　証明は文献 [100] を参照． ∎

次に，PMD の幾何学的な意味を理解するには，拡散をエッジの接線方向と勾配方向に分解して考えるのが有効である．これに関して次の命題が成り立つ．

命題 4.3 式 *(4.5)* の非等方拡散は，エッジの勾配方向 η とこれに直交する方向 ξ への方向微分により，以下のように直交分解される．

$$\frac{\partial I(\boldsymbol{s}, t)}{\partial t} = g(\|\nabla I\|)I_{\xi\xi} + f''(\|\nabla I\|)I_{\eta\eta} \tag{4.6}$$

ここで，$g(x)$ は式 *(4.2)* で与えられるエッジ停止関数である．

[証明]　まず，エッジの勾配方向の単位ベクトルが $\eta = \nabla I/\|\nabla I\|$ と書けることから，$I_{\eta\eta}$ を計算すると，

$$\begin{aligned} I_\eta &= \nabla I \cdot (\nabla I/\|\nabla I\|) = \|\nabla I\|, \\ I_{\eta\eta} &= \nabla\|\nabla I\| \cdot \eta = \frac{1}{\|\nabla I\|}\nabla\|\nabla I\| \cdot \nabla I \end{aligned} \tag{4.7}$$

が得られる．次に，エッジ停止関数 $g(x) = f'(x)/x$ に基づく PMD の式を展開すると，

$$\begin{aligned} & div(g(\|\nabla I\|)\nabla I) \\ = \;& \nabla g(\|\nabla I\|) \cdot \nabla I + g(\|\nabla I\|) div(\nabla I) \\ = \;& g'(\|\nabla I\|)(\nabla \|\nabla I\| \cdot \nabla I) + g(\|\nabla I\|) div(\nabla I) \end{aligned}$$

となるが，ここで式 (4.7) より，$\nabla\|\nabla I\| \cdot \nabla I = \|\nabla I\| I_{\eta\eta}$ であること，及び，ラプラス作用素の等方性から，$div(\nabla I) = I_{xx} + I_{yy} = I_{\xi\xi} + I_{\eta\eta}$ となることを利用すれば，

$$\begin{aligned} = \;& g'(\|\nabla I\|)\|\nabla I\| I_{\eta\eta} + g(\|\nabla I\|)(I_{\xi\xi} + I_{\eta\eta}) \\ = \;& (g(\|\nabla I\|) + g'(\|\nabla I\|)\|\nabla I\|) I_{\eta\eta} + g(\|\nabla I\|) I_{\xi\xi} \\ = \;& f''(\|\nabla I\|) I_{\eta\eta} + g(\|\nabla I\|) I_{\xi\xi} \end{aligned}$$

となり，式 (4.6) の直交分解が得られる．ただし，最後の式変形にあたっては，命題 4.1 の式 (4.2) より $f'(x) = xg(x)$ となることから，$f''(x) = g(x) + xg'(x)$ が成立することを利用した．∎

4.3.2　理論的考察

命題 4.1〜4.3 に基づき，PMD において輝度勾配の大きなノイズの除去およびエッジの非強調を実現するための方針について述べる．

まず，ノイズ除去の方針は次の通りである．PMD においてノイズが除去しきれず残ってしまう問題は，その勾配降下フローが ill-posed であり，その過程において局所解に捕まってしまうことが原因と考えられる．そこで，命題 4.2 の必要条件に基づき，勾配降下フローが well-posed となるような誤差ノルム f を選択する．なぜなら，well-posed なフローでは大域解としての定数画像 $I(x, y) = const$ に収束することが保証され，結果として平滑化画像にノイズが残留するのを防ぐことができるからである．一旦，誤差ノルム f が決まれば，次は命題 4.1 で与えられた関係式 (4.2) を用いてエッジ停止関数 g を求めることが可能となる．

次に，エッジの非強調を実現するには命題 4.3 に示す直交分解を使用する．式 (4.6) によれば，PMD による拡散にはエッジの接線方向への拡散 $I_{\xi\xi}$ とエッジに垂直な勾配方向への拡散 $I_{\eta\eta}$ が同時に働いていることがわかる．接線方向への成分 $g(\|\nabla I\|)I_{\xi\xi}$ は $g(\|\nabla I\|) \geq 0$ であることから，常にエッジの接線方向への平滑化となり，目標とするエッジ保存型の拡散であることがわかる．一方，勾配方向への成分 $f''(\|\nabla I\|)I_{\eta\eta}$ は，$f''(\|\nabla I\|) > 0$ であればエッジを跨る平滑化，$f''(\|\nabla I\|) < 0$ であればエッジの強調となり，いずれにしても望ましくない効果が発生する．したがって，$\|\nabla I\| \geq K$ となるエッジを保存し，かつエッジの強調を回避したい場合には，誤差ノルム f が $x \geq K$ において $f''(x) = 0$ を満たすように調整し，エッジ点における勾配方向の拡散 $I_{\eta\eta}$ の作用を常にゼロにすればよい．

4.3.3　エッジ停止関数の導出

以上の考察をもとに，求めるエッジ停止関数 $g(x)$ を導出する手順をまとめると以下のようになる：

【エッジ停止関数 $g(x)$ の構成手順】

i)　以下の条件を満たす関数 ϕ を決定する．

$$\phi(x) \begin{cases} > 0 & \text{if } x < K \\ = 0 & \text{otherwise} \end{cases}$$

ii)　$\psi'(x) = \phi(x)$ を満たす $\psi'(x)$ を求める．　特に $\phi(x)$ が陽に与えられた場合は，

$$\psi(x) = \int \phi(x) dx.$$

iii)　$\psi(x)$ が以下の条件を満たすか確認する．満たさない場合は手順 (i) へ．

$$\psi(x) > 0 \text{ for } \forall x \in (0, \infty),\ \psi(0) = 0$$

iv)　求めるエッジ停止関数を次式で決定．

$$g(x) = \psi(x)/x.$$

以上の手順においては，手順 i) の条件を満たす関数 ϕ をどのように選ぶかという点が問題となる．本章で著者らが提案する基本的な指針は，原点の近傍 $|x| < \epsilon$ において $\phi(x) \approx g(x)$ となる関数を ϕ の候補として選択することである．命題 4.3 によれば，このような ϕ を選ぶことにより，$\|\nabla I\| < \epsilon$ となる画素では等方拡散が実現され，細かいテクスチャ領域などにおいては，ガウシアンフィルタと同等の効果的な平滑化効果が期待できるというメリットがある．以下では上記の指針に則り，具体的な 2 つのエッジ停止関数の導出例を示すこととする．

導出例 1　まず，最も理想的な例として，$0 \leq \forall x < K$ に対して，$\phi(x) \equiv g(x)$ を満たすエッジ停止関数 g を導出する．

$|x| < K$ の場合，$\psi(x) = xg(x)$ であるから，$\psi(x)$ が満たすべき式は $\psi(x) = x\phi(x)$，すなわち，$\psi(x) = x\psi'(x)$ である．これを解くと，$\psi(x) = \alpha x$（α は定数）が得られる．一方，$|x| \geq K$ の場合は，$\psi'(x) \equiv 0$ であるから，$\psi(x)$ が $x = K$ で連続となるためには $\psi(x) \equiv \alpha K$ となる必要がある．したがって，$\alpha = 1.0$ に正規化した場合，手順 *iv)* の関係式から，

$$g^{TV}(x) = \begin{cases} 1 \ \textit{(const)} & \textit{if } |x| < K \\ K/x & \textit{otherwise} \end{cases}$$

が得られる．∎

本導出例で得られたエッジ停止関数を観察すると，$0 \leq \|\nabla I\| < K$ を満たす画素においては，ガウシアンフィルタと等価な等方拡散が実現され，一方，$\|\nabla I\| \geq K$ を満たす画素

では，ノイズ重畳画像の復元に利用される Osher-Rudin の TV フィルタ（Total Variation Filter[14]）と等価な非等方拡散が実現されることがわかる[1]．上で得られたエッジ停止関数を g^{TV} と表記したのはこのためである．

残念ながら本導出例では，$g^{TV}(x)$ から得られる $\phi(x)$（$f''(x)$ に相当）を連続関数として構築できないという問題がある．この場合，エネルギー関数 f が $C^2[0,\infty)$ の元であるという仮定に反するという理論上の不都合が生じるだけでなく，著者らの検証においても，実際に g^{TV} を PMD のエッジ停止関数として使用すると過度のエッジ強調が発生し，安定した平滑化を実現できないことが確認されている．

次に，以下の導出例では上述した問題点の改善を試みる．

導出例 2 $\forall x < K$ に対し，

$$\phi(x) = \begin{cases} \alpha(1-(x/K)^2)^2 & \textit{if } x \le K \\ 0 & \textit{otherwise} \end{cases} \tag{4.8}$$

とする．ここで関数 ϕ は，*Black, Sapiro* らによって提案されたエッジ停止関数 *[9]* に基づく．$\phi'(0) = \phi'(K) = 0$ であるから，$\phi(x)$ は原点付近では等方拡散を実現し，$x = K$ で x 軸に滑らかにつながる性質を持っている．ここで，$\alpha = 1.0$ に正規化した関数を基に，手順 *(i)*〜*(iv)* に従ってエッジ停止関数を計算すると，以下の式が得られる．

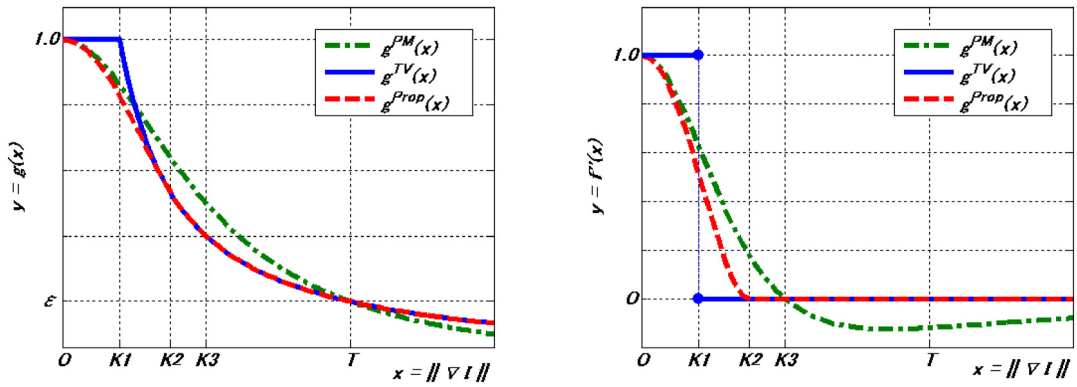
$$g^{Prop}(x) = \begin{cases} 1-\frac{2}{3}(\frac{x}{K})^2+\frac{1}{5}(\frac{x}{K})^4 & if\ |x| < K \\ \frac{8}{15}(\frac{x}{K})^{-1} & otherwise \end{cases}$$

なお，本関数では $T = \frac{8}{15}\frac{K}{\epsilon}$ が成立する．∎

上で導出した 2 つエッジ停止関数 $g^{TV}(x), g^{Prop}(x)$ に対応する PMD の特性を比較するため，エッジの接線方向への拡散 $I_{\xi\xi}$ の係数 $g(x)$ のプロットを図 4.6(a) に，勾配方向への拡散 $I_{\eta\eta}$ の係数 $f''(x)$ のプロットを図 4.6(b) に示す．比較のため，従来の $g^{PM}(x)$ に対するプロットも併記した．

$g(x)$ のプロットはいずれも右肩下がりの単調に減少する関数であるのに対し，$f''(x)$ のプロットにはかなりの違いが見られる．ここで，g^{TV}, g^{Prop}, g^{PM} が 0 以下となる境界値をそれぞれ $x = K_1, K_2, K_3$ で表記している．$\alpha = 1.0$，$\epsilon = 0.2$ とした場合には，$K_1 = 0.2T$, $K_2 = 0.375T$, $K_3 = 0.5T$ である．$g^{PM}(x)$ では，$x = K_3$ を境に $f''(x)$ が負の値を取り，エッジの強調が行われるのに対し，$g^{TV}(x), g^{Prop}(x)$ では，K_1, K_2 を境に $f''(x) = 0$ となることから，勾配方向にはエッジの平滑化も強調も行われず，エッジが保存されることがわかる．ただし $g^{TV}(x)$ の場合，$f''(x)$ のプロットが $x = K_1$ において連続性を失ってしまうことから，本章では $g^{Prop}(x)$ を提案手法として採用することとした．なお，g^{Prop} のサフィックスは “Proposed” の意味である．

[1] 正確には，TV フィルタから制約項を除いたものと等価となる．



(a) $g(x)$ のプロット　　(b) $f''(x)$ のプロット

図 4.6: 導出したエッジ停止関数の比較

4.4　シミュレーションによる検証

エッジ停止関数 $g^{Prop}(x)$ による PMD の性能を確認するため，シミュレーション実験を行った．以下の検証において，g^{PM} を用いた PMD を従来法（PMD），g^{Prop} を用いた PMD を提案法（Proposed），また従来法，提案法に対して正則化を適用したものをそれぞれ正則化従来法（PMD+R），正則化提案法 (Proposed+R) と呼ぶことにする．なお，正則化に伴う演算量の削減を考慮し，たたみ込み演算 $G * I$ には 3 タップのフィルタ $[1, 2, 1]/4$ を縦横に施して得られる二項ガウシアンマスクを使用している．

4.4.1　ノイズ除去性能の検証

まず，提案法によるノイズ除去性能の改善を確認するため，標準画像 Susie の第 100 フレームにノイズを重畳し，その平滑化の様子を確認するための数値実験を行った．重畳したノイズは，輝度勾配が高く通常の PMD では除去が困難な salt & pepper ノイズである．平滑化手法としては，従来法，正則化従来法，提案法，正則化提案法の 4 通り，および参考のため，ガウシアンと等価な等方拡散として $g \equiv 1$ を使用した．また本実験では，エッジ停止関数 g^{PM}, g^{Prop} において，いずれも $\epsilon = 0.2$，$T = 15$ と設定してシミュレーションを行った．

PMD の繰り返し演算回数に対する SNR の変動をプロットした結果を図 4.7 に示す．また，各方式の SNR のピーク値を表 4.7 にまとめた．

従来法では 100 回の繰り返し演算後もノイズが残留し，他方式に比べて SNR が明らかに低い結果となった．一方，提案法のプロットはノイズ除去の効果でいったん SNR が上昇してピークを迎え，その後テクスチャ平滑化の影響で SNR が落ち始める形を取っている．この点で従来法が抱えているノイズ除去の問題は改善されているが，ピークを迎えるまでの時間がやや長く，40 回以上の繰り返し演算が必要であった．これに対し，正則化を行うと SNR のピークを迎えるまでの繰り返し回数が大幅に減少し，ピークの SNR も等方拡散に比べて

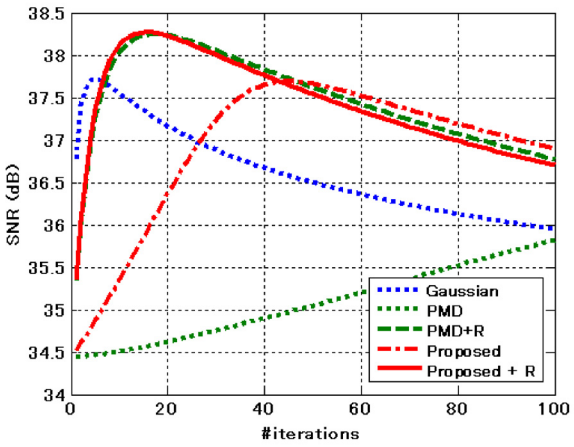


図 4.7: ノイズ除去に伴う SNR の変動

表 4.1: SNR ピーク値と繰り返し回数

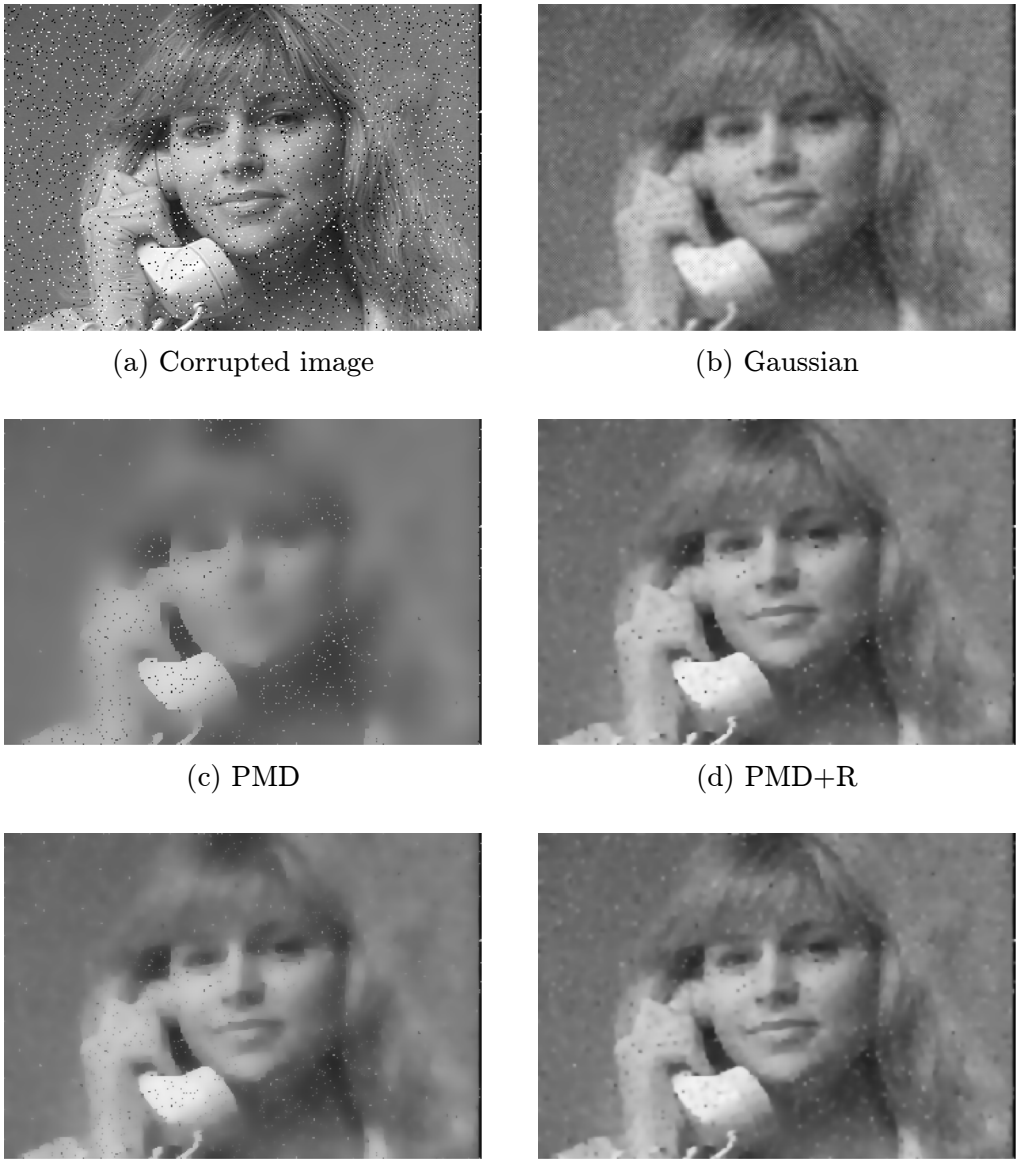
	#iters	SNR (dB)
Gaussian	5	37.71
PMD	183	36.32
PMD+R	17	38.25
Proposed	45	37.69
Proposed+R	16	38.27

0.6dB 程度向上することがわかる．正則化を行った両手法では，正則化提案法がやや SNR が高く，ピークを迎える反復回数もやや少ない結果となった．

各方式のピーク SNR 時の平滑化画像を図 4.8 に示す．提案法では，ピーク SNR を得るまでの繰り返し回数が多くなる影響からノイズはよく除去できている一方で，テクスチャ平滑化が必要以上に進行することにより，やや絵画的な不自然な画像になってしまっている．これに対し，正則化を行った両方式では，等方拡散の画像に比べてエッジがクリアであり，同時にノイズもよく平滑化されている．

著者らのシミュレーションでは，提案手法における 1 回の演算ループに要する処理時間が，正則化の導入に伴い，約 2 倍に増加した．しかし一方で，表 4.7 に見るように，SNR のピークを実現するまでに要する繰り返し回数が正則化によって半分以下に短縮されるため，正則化による演算量の増加は目標の平滑化画像を得るまでのトータルの処理時間としては特にネックにはならないことがわかる．

以上の検証により，ノイズ除去の観点から見ると，g^{Prop} による提案法は従来法よりも優れているといえる．また，主観画質を向上するためには，PMD の正則化が有効であることがわかる．



(a) Corrupted image　(b) Gaussian

(c) PMD　(d) PMD+R

(e) Proposed　(f) Proposed+R

図 4.8: 最適 SNR 時の画像

4.4.2　エッジ保存性とエッジ強調効果の抑制

提案法によるエッジ保存及びエッジ強調の抑制効果を確認するため，図 4.1 のテスト画像を用いた平滑化シミュレーションを行った．テスト画像の中央に見られる 3 つの垂直帯領域の輝度レベルはそれぞれ 64,128,192 であり，各領域にはそれぞれ輝度値 48,32,16 の白線が 8 画素間隔で重畳されている．比較のため用いた手法は正則化従来法と提案法の 2 種類であり，$\epsilon = 0.1$ の下，それぞれ $T_1 = 4$，$T_2 = 16$ に設定して，テスト画像の平滑化の様子を観察した．なお，ここでは正則化の影響を考慮し，同程度の勾配を持つエッジを保存対象とするよう $T_1 = T_2/4$ としてある．

図 4.9 に，両方式で平滑化された画像の 120 ラインの波形を示す．ここでは，主観的に同程度の平滑化が得られた時点での波形として，正則化従来法では $iter = 20$，提案法では $iter = 18$ におけるプロットを比較している．

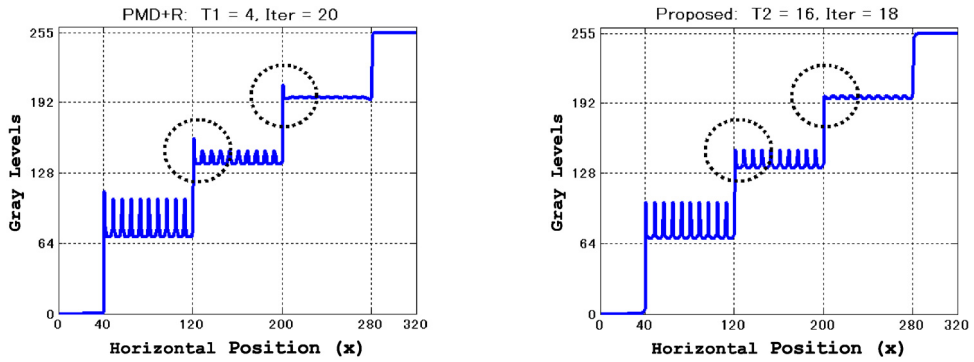


図 4.9: 平滑化画像における第 120 ライン波形の比較（左図：正則化従来法，右図：提案法）

図 4.9 の波形を見ると，図左の正則化従来法では，エッジ強調効果に伴い，帯領域境界部分に重畳したテクスチャが残留することにより，ステップエッジの波形にピークが発生していることが確認できる．このことから，たとえ正則化を導入しても，従来法では十分なエッジ強調の抑制効果が得られないことがわかる．これに対して，図右の提案法では従来法のようなピークの発生は起きておらず，エッジ停止関数の選定が適切であったことが確認できる．

次に提案法による PMD の選択的テクスチャ平滑化の様子を確認するために，エッジ停止関数 g^{prop} を用いて4.2 節と同様の平滑化シミュレーションを行った．ここでは，$T = 16, 32, 48$，$\epsilon = 0.1$ に対してそれぞれ $iter = 20$ まで平滑化の反復処理を行い，テスト画像の第 120 ラインが平滑化される様子を図 4.10 に示す．なお，プロットの提示の仕方は 4.2 節で行った実験のものと同様である．

図 4.10 各行のプロットに着目すると，振幅 $\delta < T$ のテクスチャが優先的に平滑化されており，ある程度の選択的テクスチャ平滑化は実現されていることがわかる．目安としては，T として積極的に平滑化したいテクスチャ振幅の上限値を採用すればよい．しかしながら，振幅 δ に基づくテクスチャの選択性は，エッジ停止関数として g^{BS} を使用した場合ほどには

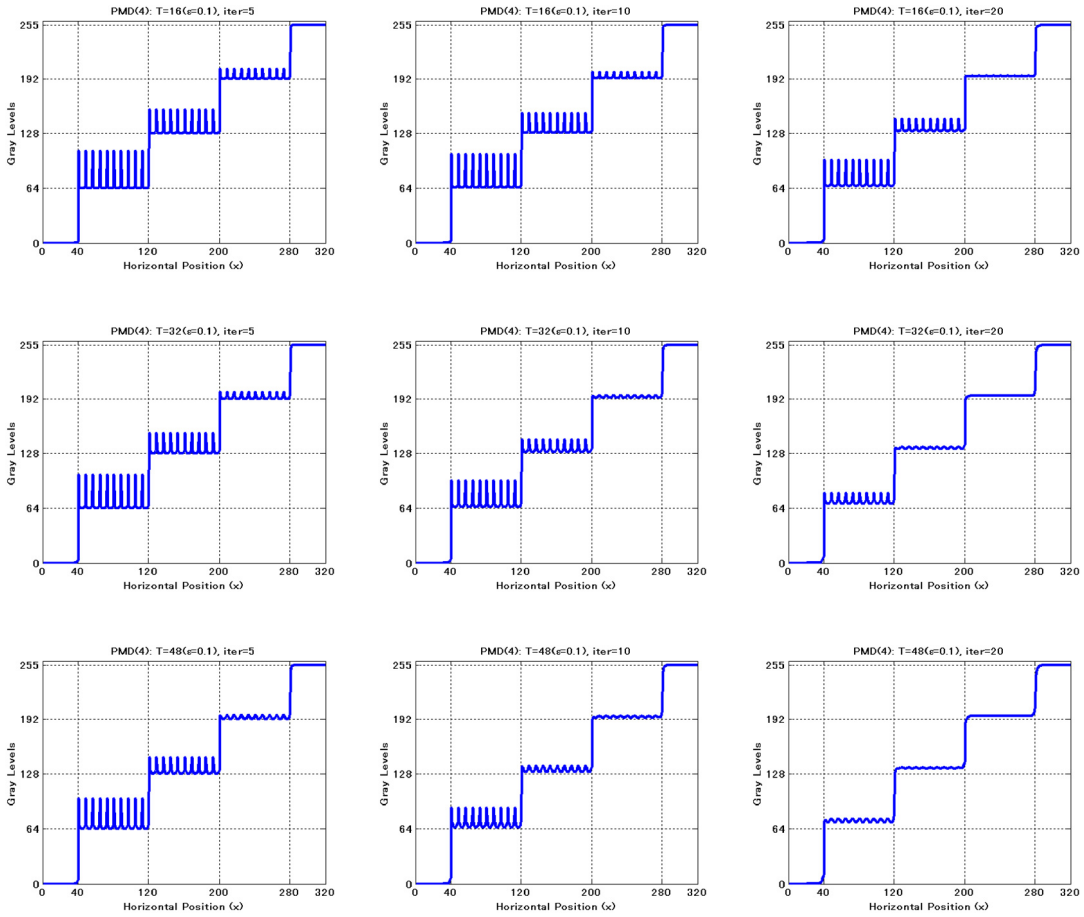


図 4.10: PMD (with $g^{Prop}(x)$) による第 120 ラインの推移（上段から $T = 16, 32, 48$，各段の左から $iter = 5, 10, 20$)

厳密ではない．例えば $T = 32$ の場合，$iter = 20$ の波形をみると，保存されるはずの $\delta = 48$ のテクスチャの平滑化がかなり進行していることがわかる．より厳密なエッジ保存平滑化を実現するには，ϵ をさらに小さな値に取る必要がある．

また，ステップエッジ部分に着目すると，特に正規化従来法において見られたエッジ強調に伴う波形のピークは，いずれのプロットにおいても全く発生していないことが確認できる．提案法は well-posed な拡散法であるため，繰り返し演算を継続することにより最終的には帯領域間のステップエッジもまた平滑化されてしまうはずである．しかしながら，本実験の結果を見る限り，20 回程度の繰り返し演算ではまだ十分良好にエッジが保存されていることがわかる．

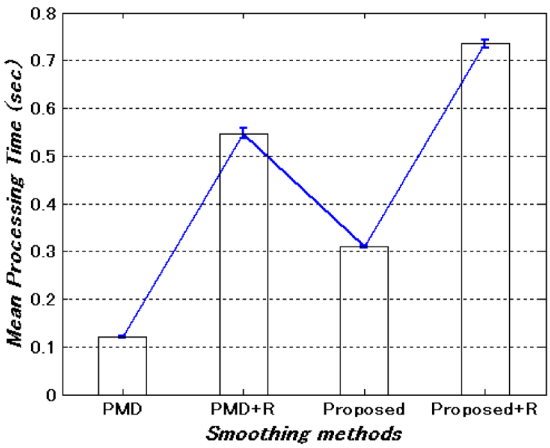


図 4.11: 処理時間の比較

4.4.3　演算量の削減

4 種類の平滑化手法（従来法，提案法，及びこれらを正規化したもの）に対して，繰り返しループに要する処理時間を測定して比較を行った．処理時間を棒グラフにまとめた結果を図 4.11 に示す．測定環境は Pentium4 2.2GHz の CPU と 512MB のメモリを搭載した Windows 2000 ワークステーションである．プログラムの開発環境には MATLAB（Ver.7）を使用し，M-file で作成したスクリプトプログラムの実行時間を測定した．図 4.11 に示された処理時間は，式 (2.45) あるいは式 (2.51) の反復演算を 100 回行い，1 回あたりの処理時間の平均を求めた値である．グラフには 95%信頼区間を示すバーが表示されているが，信頼区間が非常に短いため，グラフ上はやや確認しづらくなっている．

実験結果によると，従来法に対して提案法は 3 倍近くの処理時間がかかっている．これは，使用するエッジ停止関数が 2 つの関数を滑らかにつなげて作られているため，条件分岐による場合わけが必要になることや，$x < T$ において 4 次の多項式の計算が必要になることなどが原因で発生しており，ルックアップテーブルを使うなどの実装上の工夫によって吸収することが可能である．一方，正則化を行った場合は，いずれの手法においても約 0.4 秒のオーバヘッドが必要となることがわかるが，これは正則化処理に伴う本質的な差分である．提案法においては，正則化を行った結果，1 ループあたりの処理時間が 2 倍以上に増加するが，ノイズ除去の実験でも見たように，正則化によって半分以下の繰り返し処理で同等の平滑化効果が得られるようになることを考えれば，正則化を採用するメリットは十分考慮に値する．

4.4.4　画像情報量の削減

低ビットレートで使用される MPEG-4 Simple Profile では，フレーム内予測を行う I-VOP とフレーム間予測を行う P-VOP の 2 種類の画像タイプが使用される．ここでは，PMD に基づく非等方拡散フィルタによる，各画像タイプの情報量削減効果について検討する．

I-VOP の情報量削減効果

提案法による I-VOP の情報量削減効果について検証を行った．イントラフレームは，通常，8 × 8 ブロック単位で DCT 変換され，その DCT 係数が符号化される．その際，DC 成分は周囲のブロックから予測符号化されるため，ブロックの情報量を測る量としては AC 成分の信号電力が適切である．そこで本実験では，画素領域で AC 成分の信号電力に相当する各ブロックの分散を測定し，平滑化の繰り返し回数 *iter* の増加に伴い，分散値がどのように変化するかを確認した．さらに，*iter* の増加に伴うフレーム内予測誤差信号のエントロピーの変化についてもあわせて確認した．ここで，予測誤差信号のエントロピーは，画像内のすべての画素にわたりブロック内平均からの偏差の度数分布を求め，その分布から得られる各偏差レベルの出現確率をもとに算出した．

本実験で使用した画像は，標準画像 Flower Garden, Mobile & Calendar, Susie の第 1 フレーム，Diva with noise の第 112 フレーム，及び Football の第 50 フレームの計 5 種類である．上述の実験と同様，従来法，提案法，及びこれらを正規化した計 4 種類の平滑化手法で比較を行い，エッジ停止関数のパラメータは $\epsilon = 0.2$，$T = 20$ とした．

フィルタによる情報量削減効果に関するプロットを図 4.12, 図 4.13 に示す．図 4.12 が平均信号電力（分散値），図 4.13 がエントロピーの推移を示している．本実験の結果，5 種類のいずれの画像に対しても反復演算回数に対して単調に減少する傾向のプロットが得られた．電力のプロットは反比例に近い急激な落ち方を示すが，エントロピーのプロットは，これに比べればかなり緩やかな減少傾向を見せている．各手法のプロットを比較すると，複雑度の高い Flower Garden, Mobile & Calendar, Diva with noise では，提案法は従来法に比べて情報量削減効果が大幅に向上している．また，正則化を導入するとさらに情報量削減効果が上がり，その場合は，従来法よりも提案法の方がやや情報量削減効果が高くなることがわかる．一方，上の 3 つの画像に比べてもともと複雑度が低く符号量の少ない Susie と Football については，ガウシアンを除くどの平滑化手法を使用しても情報量削減効果にあまり差が見られなかった．

P-VOP の情報量削減効果

低ビットレートにおける動画像符号化では，特に P-VOP の情報量削減が重要となることから，イントラフレームの場合と同様に，提案法によるフレーム間差分信号の電力とエントロピーの低減効果についても検証を行った．Mobile & Calendar を用いた検証結果を図 4.14(a),(b) に示す．これらのプロットは，16×16 のマクロブロックをベースとした動き補

償フレーム間差分信号に対する分散およびエントロピーの低減効果を示している．本実験では，動きベクトルの算出に平滑化処理前の原画像を使用した．また，低ビットレートにおけるフレームレートとして 10fps 程度を想定しているため，フレーム間隔は 3 フレームとし，第 1 フレームと第 4 フレームの間で動き補償を行った．動きベクトルの探索範囲は 32 画素 × 32 画素とし，1 画素精度の全探索を行った結果を使用した．

図 4.14(a),(b) のプロットから，提案法によるフレーム内平滑化の結果として，フレーム間差分の電力とエントロピーもまた大幅に減少することがわかる．その低減効果は，フレーム内の場合と同様，正則化提案法，正則化 PMD，提案法，PMD の順に優れているが，特にエントロピーのプロットからは，PMD による結果が他の手法に比べて顕著に劣っていること，正則化提案法と正則化 PMD とではその効果にほとんど差がないこと，繰り返し回数が増えるにつれてこれら 2 つと提案法との低減効果の差がほとんどなくなること等が読み取れる．

主観画像品質の改善

フィルタをかけていない 5 種類の原画像 (Flower Garden, Mobile & Calendar, Susie, Diva with noise, Football) に対して各手法による平滑化を行い，エントロピーが 1 bit/pel 減少した時点での画像品質を主観的に比較した．これは，Mobile & Calendar の例でいえば，図 4.13(b) のプロットにおいてエントロピーが 5.9 bits/pel となるラインの平滑化画像に相当する．

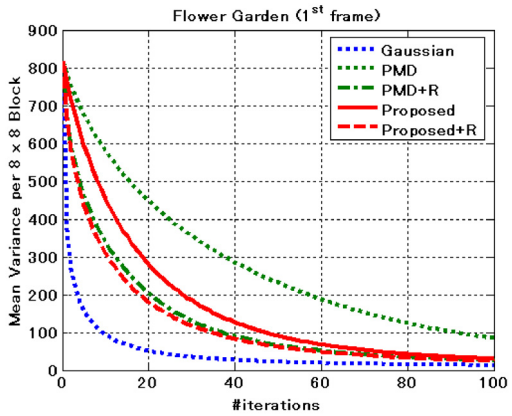
各手法によるエントロピー 1bit/pel 低減時の画像を図 4.15〜図 4.19 に示す．従来法では，1 bit/pel の情報削減のために多くの繰り返し演算が必要となることから，画像のディテイルが失われるとともに，エッジが過度に強調された不自然な平滑化画像になってしまう．これに対し，提案法を使用すれば，目標のエントロピーを実現するまでの繰り返し演算数を大幅に減少し，自然な平滑化画像を得ることができる．主観画質的にも，例えば Mobile & Calendar の玩具の汽車に着目すれば，ガウシアンによる等方的な拡散に比べてエッジが保存されよりクリアな印象の画像が得られることが確認できる．また，正則化を導入することにより，提案法とほぼ同等の平滑化画像をさらに少ない繰り返し演算数で実現できるようになることがわかる．本実験において，ガウシアンを除き，目標のエントロピーを最も少ない繰り返し回数で実現できたのは，複雑度の高い Flower Garden, Mobile & Calendar, Diva with noise に対しては正則化提案法，比較的複雑度の低い Susie, Football に対しては正則化従来法であった．

4.5　おわりに

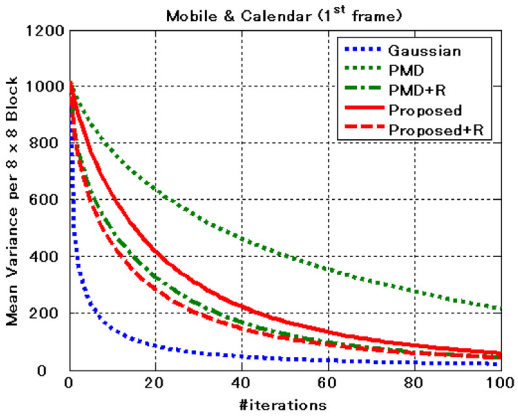
非等方拡散法に基づく空間平滑化フィルタの構築に関する基礎検討を行った．Perona-Malik らによって提案された PMD を空間平滑化フィルタとして使用する際には，輝度勾配の大きくなるノイズの除去とエッジ強調効果が問題となる．本章では，これらの問題を解決するた

めにエッジ停止関数が満たすべき条件を明らかにし，それに基づいて具体的なエッジ停止関数の導出を行った．さらにシミュレーション実験を通して，今回導出したエッジ停止関数によって実現される非等方拡散法において，上述の問題が解決されていることを確認した．また，低ビットレート映像符号化のプレフィルタとしての応用を想定して，本フィルタによる情報量削減効果についても実験を行い，提案手法の優位性を確認した．さらに本実験を通して，平滑化画像の主観画像品質を高める上で，ガウシアンによる正則化を導入することが効果的であることが確認された．

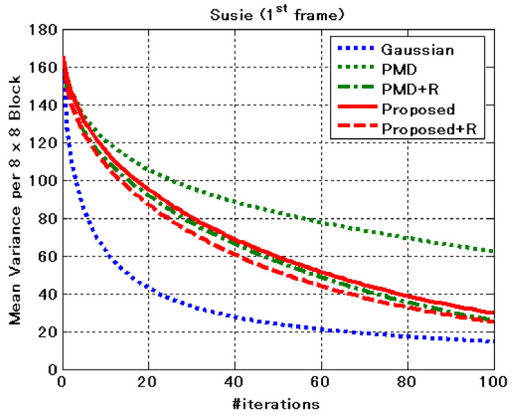
次章では，本章で構築した空間平滑化フィルタを踏まえ，これを時空間 3 次元フィルタへ拡張する方法について検討を行う．



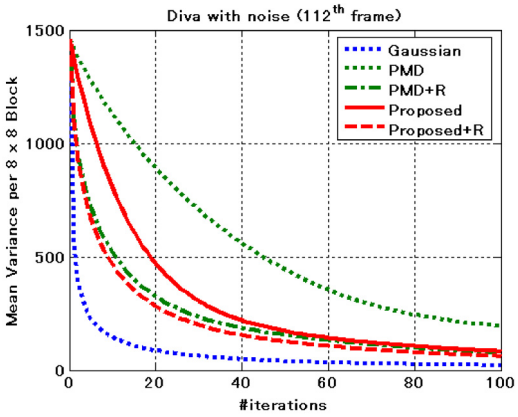
(a) Flower Garden



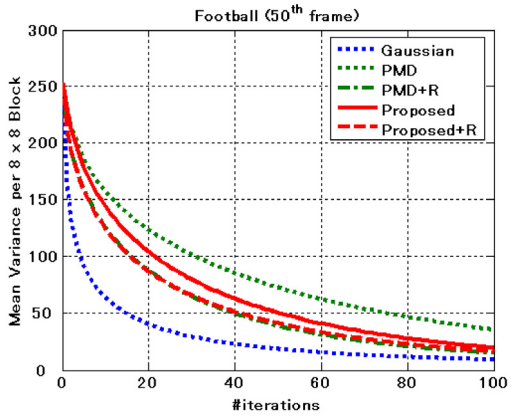
(b) Mobile & Calendar



(c) Susie

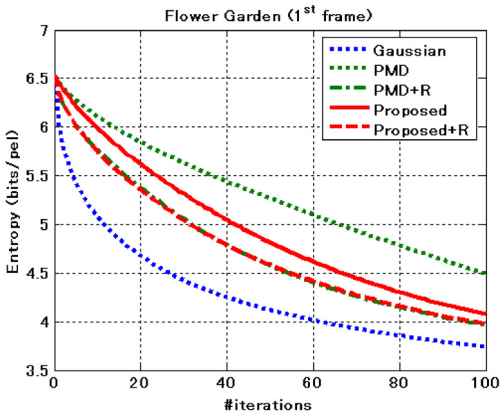


(d) Diva with noise

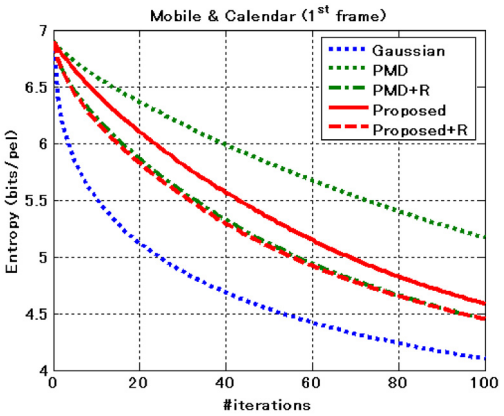


(e) Football

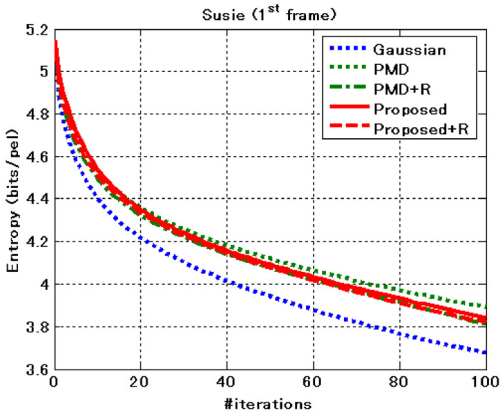
図 4.12: I-VOP における平均信号電力（分散値）の低減効果



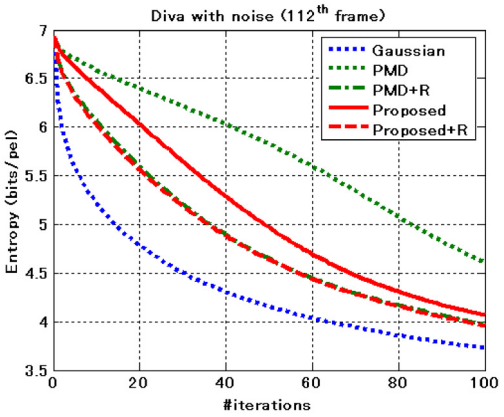
(a) Flower Garden



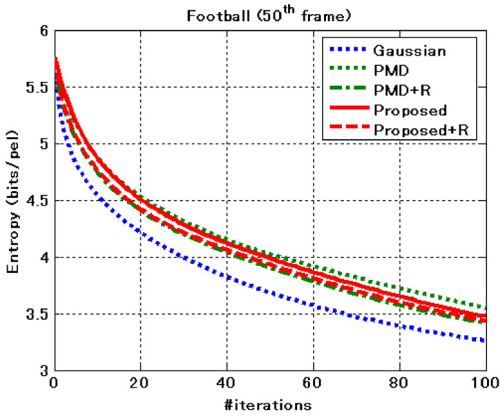
(b) Mobile & Calendar



(c) Susie

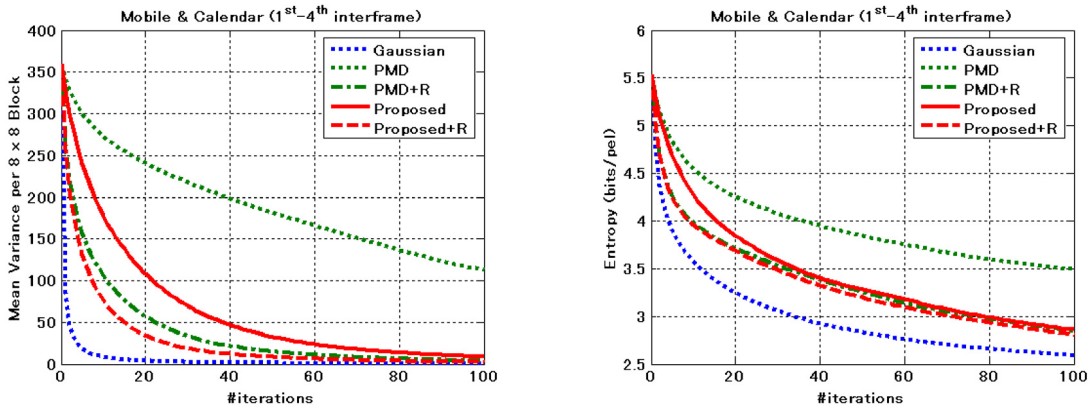


(d) Diva with noise



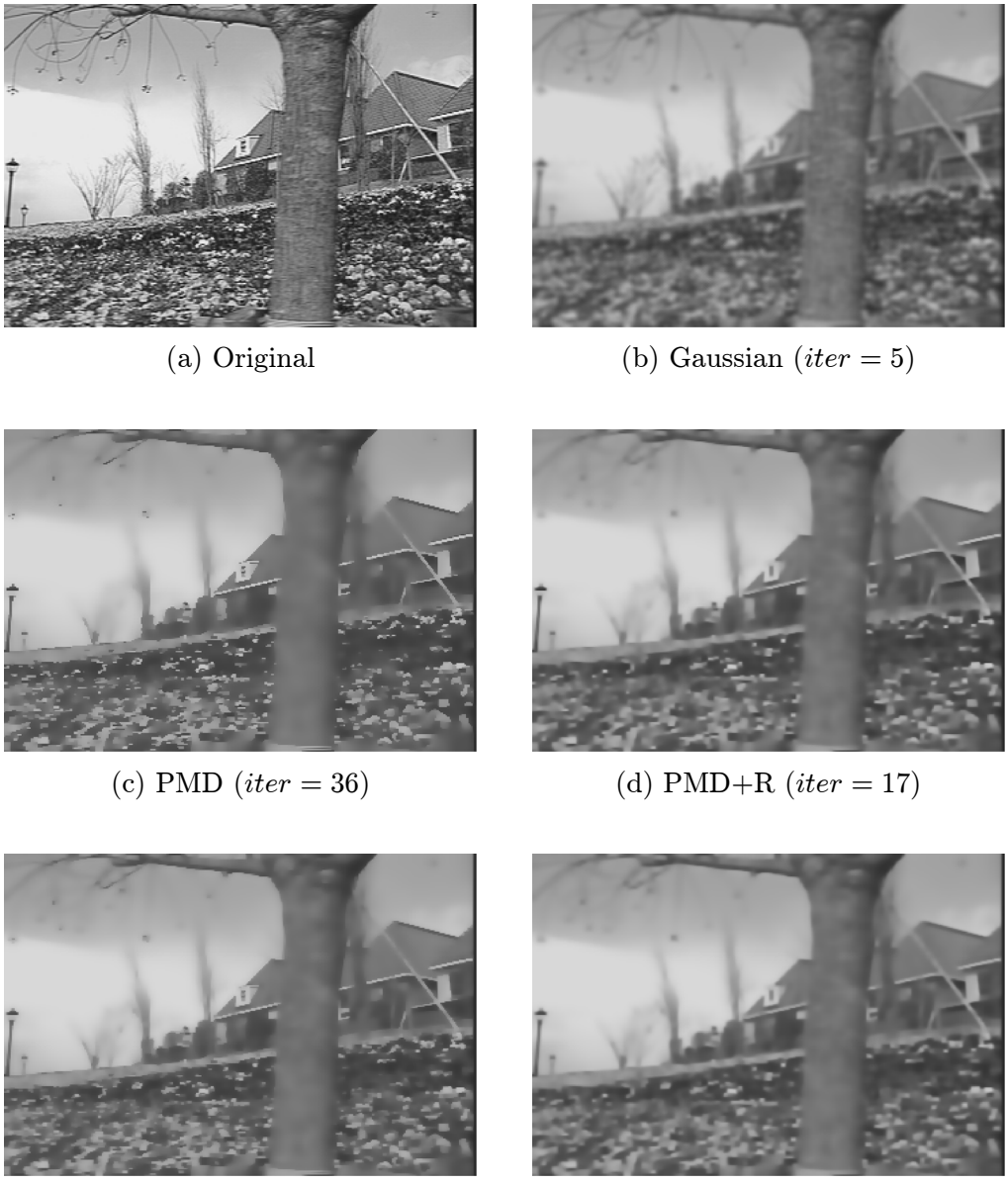
(e) Football

図 4.13: I-VOP におけるエントロピーの低減効果



(a) Variance reduction (P-VOP)　(b) Entropy reduction (P-VOP)

図 4.14: 平均信号電力（分散値），エントロピーの低減効果



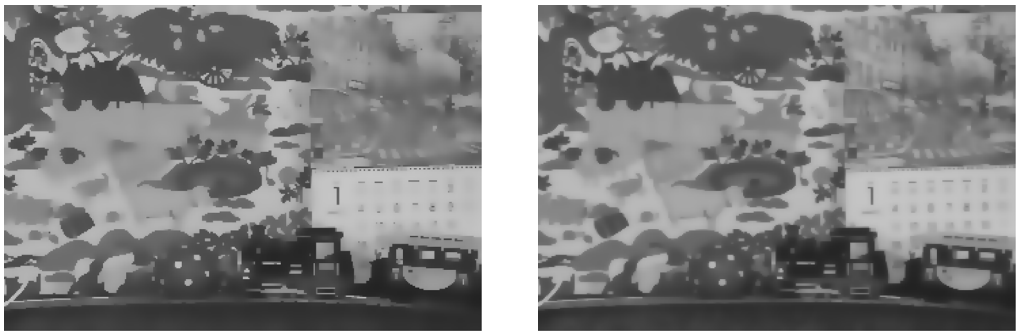
(a) Original　(b) Gaussian ($iter = 5$)

(c) PMD ($iter = 36$)　(d) PMD+R ($iter = 17$)

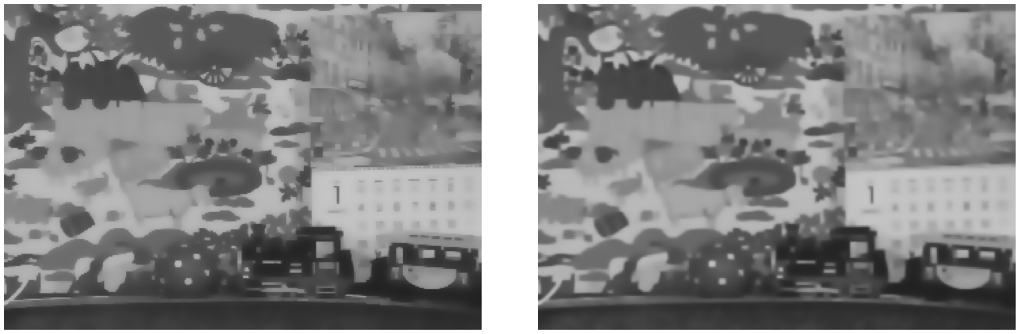
(e) Proposed ($iter = 24$)　(f) Proposed+R ($iter = 17$)

図 4.15: エントロピー 1bit/pel 削減時の主観画像品質の比較 (Flower Garden)

(a) Original　(b) Gaussian ($iter = 6$)



(c) PMD ($iter = 47$)　(d) PMD+R ($iter = 20$)



(e) Proposed ($iter = 28$)　(f) Proposed+R ($iter = 18$)

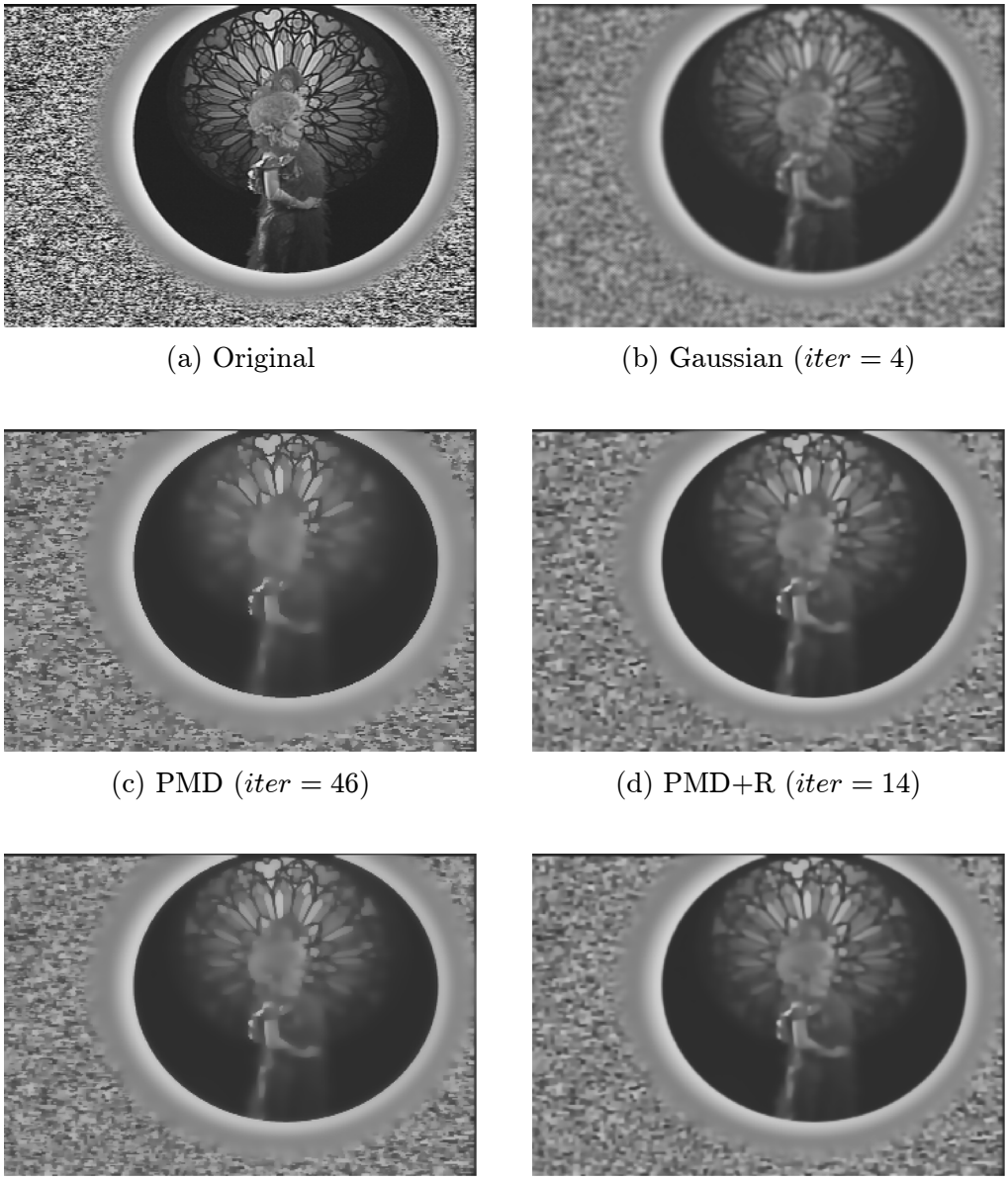
図 4.16: エントロピー 1bit/pel 削減時の主観画像品質の比較 (Mobile & Calendar)

(a) Original　(b) Gaussian ($iter = 27$)

(c) PMD ($iter = 47$)　(d) PMD+R ($iter = 41$)

(e) Proposed ($iter = 43$)　(f) Proposed+R ($iter = 42$)

図 4.17: エントロピー 1bit/pel 削減時の主観画像品質の比較 (Susie)

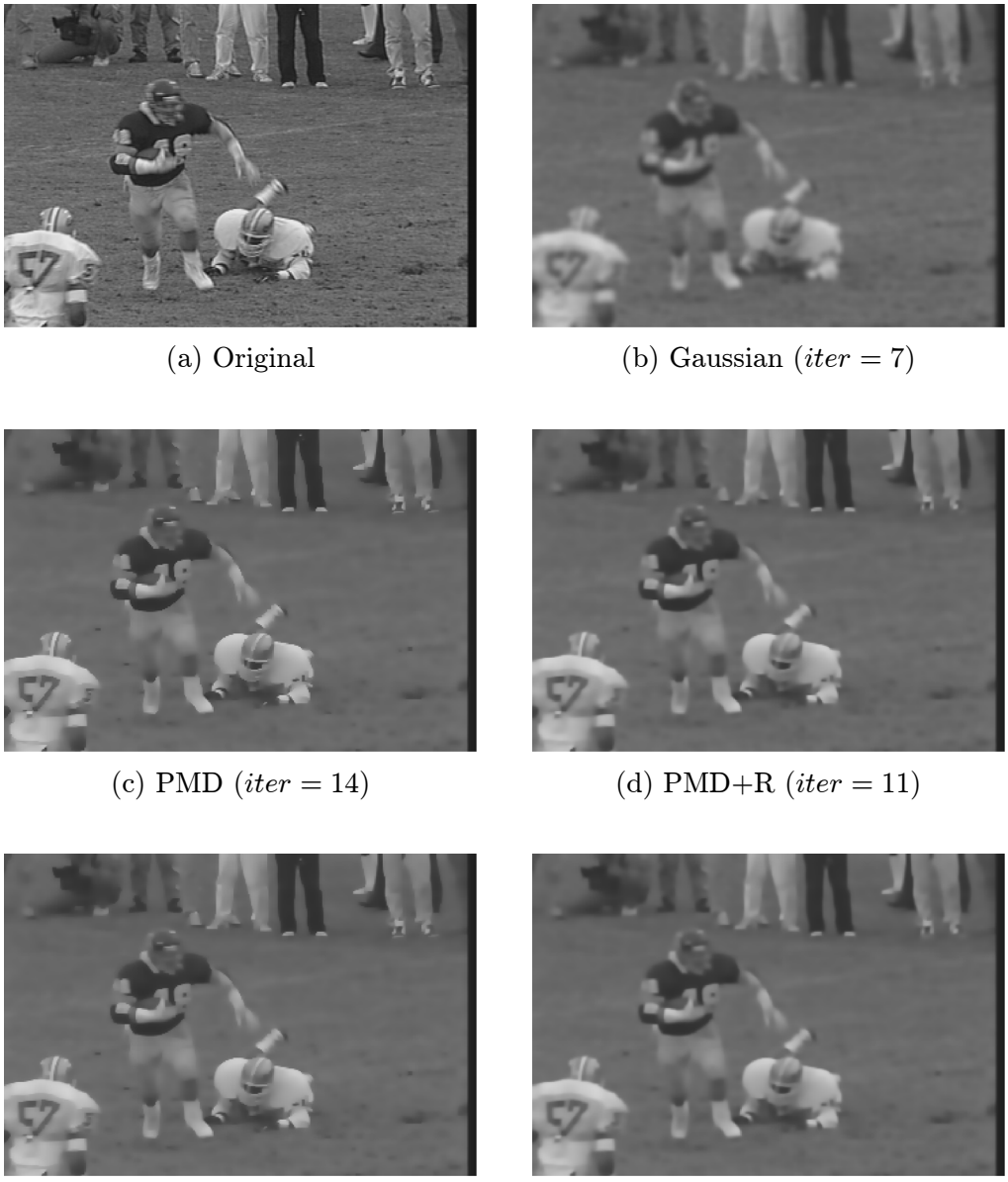


(a) Original　(b) Gaussian ($iter = 4$)

(c) PMD ($iter = 46$)　(d) PMD+R ($iter = 14$)

(e) Proposed ($iter = 24$)　(f) Proposed+R ($iter = 13$)

図 4.18: エントロピー 1bit/pel 削減時の主観画像品質の比較 (Diva with noise)



(a) Original　(b) Gaussian ($iter = 7$)

(c) PMD ($iter = 14$)　(d) PMD+R ($iter = 11$)

(e) Proposed ($iter = 14$)　(f) Proposed+R ($iter = 12$)

図 4.19: エントロピー 1bit/pel 削減時の主観画像品質の比較 (Football)

第5章　時空間平滑化プレフィルタへの拡張

5.1　はじめに

H.263 や MPEG-4 をベースとする低ビットレートの動画像符号化では，通常，符号化歪みに伴う画質の劣化を完全に回避することは困難である．このため，ある程度の歪みを許容しながらも，これを緩和し主観的な画像品質を高めることを目的として，しばしば平滑化プレフィルタによる前処理が導入される．この場合，プレフィルタの主たる目的は入力画像の符号量を削減することであるため，画像に重畳したノイズの平滑化に加えて，背景部の細かいテクスチャ等の平滑化までを視野に入れた，平滑化効果の高いフィルタを採用する必要がある．また，一方では，オブジェクトエッジに代表される視覚的に重要な情報をなるべく保存し，必要以上に主観画質を損なわないように留意することが望ましい．これらを両立するため，従来から，動画像符号化のプレフィルタとしてエッジ保存平滑化フィルタを用いる提案がなされてきた．本論文の第 4 章においては，非等方拡散法 [70] を用いたエッジ保存空間平滑化フィルタの構築方法について検討を行っている．非等方拡散法は繰り返し処理を伴う非線形フィルタであるが，従来法でしばしば必要となるエッジ抽出処理を明示的に行う必要がなく，与えられた方程式に従って数値的に入力画像を時間発展させるだけで，エッジ保存平滑化を実現できるという利点を持っている．

MPEG 等の動画像符号化においては，多くのフレームがフレーム間予測を行う P ピクチャや B ピクチャとして符号化される．一般に，フレーム間予測誤差は空間平滑化フィルタによっても低減するが，時間方向の平滑化を積極的に取り入れれば，さらに予測誤差を低減することができる．このため，動画像符号化の前処理においては，フレーム内の空間近傍だけでなくフレーム間に対しても平滑化を行う時空間フィルタがしばしば用いられてきた．初期の研究としては，TV 会議をターゲットにしたテレビジョン信号の符号化において，フレーム間差分信号を抑制するために導入された時間フィルタ [52, 18] が挙げられる．さらに，通常の SDTV(Standard Definition TeleVision) 信号に比べてカメラ入力にランダムノイズが多く含まれる傾向がある HDTV(High Definition TeleVision) 信号の符号化においても，フレーム内・フレーム間双方の予測誤差信号を抑制して符号化効率を改善するため，N.Mukawa ら [52] の手法を拡張した時空間平滑化フィルタが使用されていた [98]．また近年では，MPEG 等の動画像符号化をターゲットに，符号化歪みの発生を緩和して，主観画質を高める時空間プレフィルタの構築に関する試みが行われている [97, 92, 41, 42, 43]．

上述した HDTV 符号化のように，カメラ入力に伴うノイズの除去を主眼とするプレフィルタでは，平滑化の影響を視覚的に認知できるレベル以下に抑えるため，あまり強い平滑化効果を持ったフィルタは使用されない．これに対し，入力画像の符号量削減を主眼とするプ

レフィルタは，主観画質を維持しながらも，できるだけ強い平滑化を実現する必要があることから，従来法よりも高度に洗練されたエッジ保存平滑化プレフィルタを導入することが望ましい．著者らは，このような認識のもと，先に述べた非等方拡散法に基づく空間平滑化フィルタを導入したが，フレーム間にまたがる平滑化についてはまだ十分な検討を行っていなかった．しかしながら，今回ターゲットとする低ビットレート符号化においては，大部分のフレームが P ピクチャとして符号化され，I ピクチャは数秒に一度挿入されるだけというケースが多いため，フレーム間予測誤差の低減がトータルの符号量の削減に果たす役割は特に重要であるといえる．そこで本章では，Perona-Malik の非等方拡散法を時空間 3 次元の画像列に拡張して適用することで，エッジ保存形の時空間平滑化フィルタを実現する方法を提案する．ここでは，非等方拡散法の時間方向への拡張は著者がかつて提案した [84] の方法をベースとし，特に時間方向のエッジ停止関数の選定に関わる検討を中心に行う．

本章のもうひとつの目的は，非等方拡散法を前処理フィルタとして使用する正当性を主張することである．[52, 98, 18] で採用された従来のフィルタ手法はいずれも非線形フィルタであり，空間／時間方向の予測誤差を抑制する差分抑制関数を用いた統一的な定式化が可能である．本章で導入する非等方拡散法に基づく時空間平滑化フィルタは，これらの差分抑制関数に基づく時空間フィルタと数理的に密接な関係があり，これらの自然な拡張として理解できることを明らかにする．

以下，本章の構成は次の通りである．5.2 節では非等方拡散フィルタの時空間 3 次元への拡張方法について説明する．5.3 節では非等方拡散フィルタと従来の時空間平滑化プレフィルタの数理的な関係について考察を行い，理論的な立場から非等方拡散フィルタの正当性を主張する．続く 5.4 節では，非等方拡散法を時空間 3 次元に拡張する際に必要となる時間方向のエッジ停止関数について，理論及び実験の両面からその選定を行う．最後に 5.5 節でまとめと今後の課題を述べる．

5.2　時空間非等方拡散フィルタ

5.2.1　非等方拡散法

$I : \mathbf{x} = (x, y)^\top \mapsto I(x, y) \in \mathbf{R}$ を，矩形領域 $\mathbf{\Omega} \subset \mathbf{R}^2$ で定義されたグレイスケール画像とする．非等方拡散法とは，次の PDE (Partial Differential Equation) にしたがって初期画像 $I_0(\mathbf{x}) = I(\mathbf{x})$ を時間発展させ，任意時刻 τ[1]における画像 $I_\tau(\mathbf{x})$ を求めることで，エッジをまたがない平滑化を実現する方法である [70].

$$\frac{\partial I_\tau(\mathbf{x})}{\partial \tau} = div[g(\|\nabla I_\tau\|)\nabla I_\tau], \ \mathbf{x} = (x, y)^\top \tag{5.1}$$

ここで，関数 $g : \mathbf{R} \to \mathbf{R}$ はエッジ停止関数である．一般に，エッジ停止関数 g は $x \to 0$ のとき $g(x) \to 1$，$x \to \infty$ のとき $g(x) \to 0$ となるように設計される．このようにすれば，勾

[1]後に扱う画像列 $I(t, \mathbf{x})$ の時間パラメータ t と区別するため，ここでは時間発展のパラメータを τ と表記している．

配の大きな画素では重み係数 $g(\|\nabla I\|)$ が 0 に近づき，エッジ付近での平滑化を抑えることができるためである．特に $g \equiv 1$ としたとき，式 (5.1) は熱の拡散方程式であり，ガウシアンフィルタと同等な平滑化効果が得られることが知られている [44]．また，式 (5.1) が拡散方程式であることを考えれば，時刻 τ が大きくなるほど画像の平滑化が進行することが理解できる．

5.2.2　時空間フィルタへの拡張

Perona らによって提示された非等方拡散法はグレイスケールの静止画像に対する平滑化であったが，式 (5.1) は一般の n 次元に拡張しても成立する．そこで本論文では，これを時空間 3 次元で定義されるグレイスケールの時系列画像列（以下，画像列と呼ぶ）に対して適用する．

$I : (t, \mathbf{x}) \mapsto I(t, \mathbf{x}) \in \mathbf{R}$ を，時空間領域 $\mathbf{R}^+ \times \mathbf{\Omega}$ で定義された画像列とする．ここで，$\mathbf{\Omega}$ は画像を表す $\mathbf{R}^2$ 内の矩形領域である．任意時刻 τ における平滑化画像列 $I_\tau(t, \mathbf{x})$ を求めるには，静止画像の場合と同様に，初期画像を $I_0(t, \mathbf{x}) = I(t, \mathbf{x})$ とし，式 (5.1) にしたがって時間発展させればよい．ただし，画像列 $I(t, \mathbf{x})$ に対して式 (5.1) の非等方拡散法を適用する際には，式中のナブラ記号 ∇ は，$\nabla = (\frac{\partial}{\partial t}, \frac{\partial}{\partial x}, \frac{\partial}{\partial y})$ の意味で解釈する必要がある．

次に，式 (5.1) を時空間 3 次元に関する非等方拡散とみなして離散化する．非等方拡散法の時空間 3 次元への拡張は，従来の空間方向の近傍画素に時間方向の近傍画素を加えた時空間近傍を定義することによって実現される．本章では，フレーム内の空間 4-近傍に，時間方向の近傍画素として前後のフレームからそれぞれ 1 画素ずつを加えた時空間 6-近傍を用いて時空間非等方拡散フィルタを構築する．

具体的な式を記述する前に，画像列における画素，画素値，近傍などに関する表記法について述べる．以下，特に誤解が生じない限り，画像列内の画素を $\hat{\mathbf{x}} = (t, \mathbf{x})$ のようにフレーム時刻 t を省略して表記する．すなわち，画像列の画素 $\hat{\mathbf{x}}$ と言った場合には，これを含むフレーム内の 2 次元座標が $\mathbf{x}$ で与えられることを示しており，特にそのフレームの時間座標が必要な場合には，前後の文脈で $\hat{\mathbf{x}} = (t, \mathbf{x})$ のように明記する．また，簡単のため，画像列 I の画素 $\hat{\mathbf{x}}$ における画素値を $I_{\hat{\mathbf{x}}}$ と表記する．一方，画像列 I における時刻 t のフレームに主たる関心がある場合には，空間座標 $\mathbf{x}$ を省略して I_t と表記する．すなわち，この表記法の下では，$\hat{\mathbf{x}} = (t, \mathbf{x})$ のとき $I_{\hat{\mathbf{x}}} = I(t, \mathbf{x}) = I_t(\mathbf{x})$ が成り立つ．

さらに，画素 $\hat{\mathbf{x}}$ における時空間 6-近傍を $\eta_{\hat{\mathbf{x}}}$ と表記する．$\eta_{\hat{\mathbf{x}}}$ は $\hat{\mathbf{x}}$ を含むフレーム内の空間 4-近傍 $\eta^{sp}_{\hat{\mathbf{x}}}$

$$\begin{aligned} \eta^{sp}_{\hat{\mathbf{x}}} &\equiv \{\hat{\mathbf{x}}_N, \hat{\mathbf{x}}_S, \hat{\mathbf{x}}_E, \hat{\mathbf{x}}_W\} \\ &= \{(t, \mathbf{x}_N), (t, \mathbf{x}_S), (t, \mathbf{x}_E), (t, \mathbf{x}_W)\} \end{aligned} \tag{5.2}$$

と $\hat{\mathbf{x}}$ の前後のフレームにまたがる時間 2-近傍 $\eta^{tp}_{\hat{\mathbf{x}}}$

$$\begin{aligned} \eta^{tp}_{\hat{\mathbf{x}}} &\equiv \{\hat{\mathbf{x}}_F, \hat{\mathbf{x}}_B\} \\ &= \{(t+1, \mathbf{x}_F), (t-1, \mathbf{x}_B)\} \end{aligned} \tag{5.3}$$

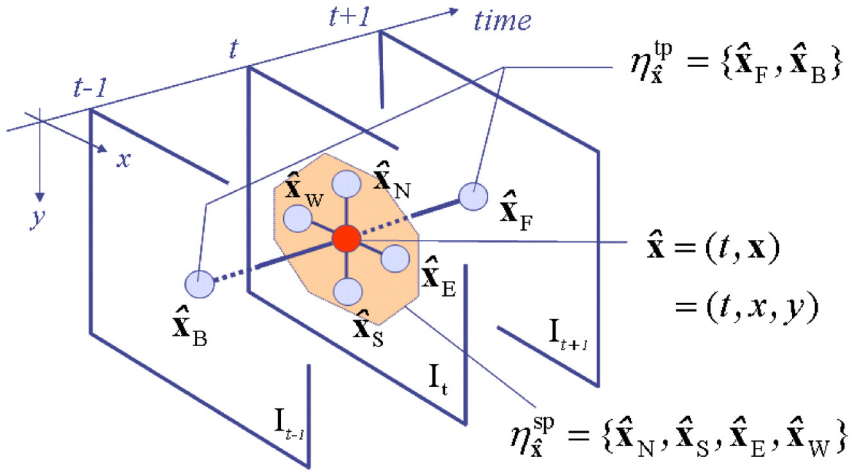


図 5.1: 時空間 6-近傍

の和集合 $\eta_{\hat{\mathbf{x}}} = \eta^{sp}_{\hat{\mathbf{x}}} \cup \eta^{tp}_{\hat{\mathbf{x}}}$ として定義する．画像列 I の画素 $\hat{\mathbf{x}} = (t, \mathbf{x})$ における時空間 6-近傍を図 5.1 に示す．

さて，Perona-Malik[70] の方法に習って，式 (5.1) を時空間 3 次元におけるエッジ保存形の拡散とみなして離散化すると次の式が得られる．

$$\begin{aligned} I^{(n+1)}_{\hat{\mathbf{x}}} &= I^{(n)}_{\hat{\mathbf{x}}} + \frac{\lambda_1}{5} \sum_{\hat{\xi} \in \eta^{sp}_{\hat{\mathbf{x}}}} g^{sp}(\nabla_{\hat{\xi}} I^{(n)}_{\hat{\mathbf{x}}}) \cdot \nabla_{\hat{\xi}} I^{(n)}_{\hat{\mathbf{x}}} \\ &\quad + \frac{\lambda_2}{2} \sum_{\hat{\xi} \in \eta^{tp}_{\hat{\mathbf{x}}}} g^{tp}(\nabla_{\hat{\xi}} I^{(n)}_{\hat{\mathbf{x}}}) \cdot \nabla_{\hat{\xi}} I^{(n)}_{\hat{\mathbf{x}}}, \\ &\quad \text{where} \quad \nabla_{\hat{\xi}} I^{(n)}_{\hat{\mathbf{x}}} \equiv I^{(n)}_{\hat{\xi}} - I^{(n)}_{\hat{\mathbf{x}}}. \end{aligned} \tag{5.4}$$

ここで，$I^{(n)}$ は n 回の繰り返し処理を行った平滑化画像列を表している．この意味において，$I^{(0)}$ は原画像列 I を表すものとする．λ_1, λ_2 は数値計算を安定に収束させるための係数である．また，一般に，画像列の空間方向と時間方向の信号特性はまったく異なるため，空間方向と時間方向では異なるエッジ停止関数 g^{sp}，g^{tp} を用いている．

エッジ停止関数の選定は，符号量の低減と主観画質の向上の点から重要と考えられる．第 4 章では，空間方向のエッジ停止関数の満たすべき性質について検討し，g^{Prop} を導出した．本章ではこれを g^{TN} と呼び，空間方向のエッジ停止関数として採用する．すなわち，g^{TN} の定義式は

$$g^{TN}(x) = \begin{cases} 1 - \frac{2}{3}(\frac{x}{K})^2 + \frac{1}{5}(\frac{x}{K})^4 & \text{if } x < K \\ \frac{8}{15}(\frac{x}{K})^{-1} & \text{otherwise} \end{cases} \tag{5.5}$$

となる．次に，時間方向のエッジ停止関数について検討を行う．

5.2.3　時間方向の拡散処理に関する検討事項

フレーム内を移動するオブジェクトがある場合やカメラに起因するフレーム全体のグローバルな動きがある場合には，時間方向の平滑化を行うにあたって，主観画質を損なわないよう十分な注意を払う必要がある．例えば，連続したフレーム間でオブジェクトが移動することによって現れるオクルージョン領域では，フレーム内の同じ位置にある画素同士で強い平滑化を行うと両者の画素がブレンドしてしまい，オブジェクトの残像が背景に現れたり，オブジェクトの輪郭がぼやける等の弊害が起きてしまうことがある．

このような弊害を回避する方法としては，２つの方向性が考えられる：(1) 輝度値の変動が大きく，時間的に連続しない画素間の時間平滑化を強力なエッジ停止関数で抑制する方法と，(2) フレーム内の各画素に対して動き補償を行い，時間的に連続した正しい対応画素同士で平滑化を行う方法である．本章では，特に (1) を実現する上で重要となる時間方向のエッジ停止関数の選定について理論的な考察を行うこととする．

5.3　従来の時空間平滑化フィルタとの関係

ここでは，動画像符号化において従来から用いられてきた時空間平滑化フィルタと非等方拡散フィルタとの数理的な関係について述べる．はじめに，従来の時空間フィルタを差分抑制関数に基づき定式化する．続いて，非等方拡散フィルタがある特定の条件の下でこの定式化に基づく従来のフィルタを拡張したものと理解できることを示す.

5.3.1　空間及び時間平滑化フィルタの定式化

画像符号化のプレフィルタとして従来から使用されている空間平滑化フィルタ及び時間平滑化フィルタの構成を図 5.2 に示す．いずれのフィルタも，原画像とある予測信号との間の誤差信号を差分抑制関数によって適応的に抑制し，その結果を元の予測信号に加えることによって平滑化を実現するという共通した手法がとられる．

従来研究 [98] では，空間フィルタの予測画像として平均値フィルタによる平滑化画像が，時間フィルタの予測画像として直前のフレーム I_{t-1} がそれぞれ使用されていた．このとき，時空間フィルタは，これらを空間フィルタ→時間フィルタの順に直列に接続することにより実現される．一方，[18] では動き補償を用いた時間平滑化を扱っているが，これは，動きベクトルに基づいて直前のフレーム I_{t-1} から生成した動き補償画像を，時間フィルタの予測画像として使用した事例と理解することができる．

以下，空間／時間フィルタをそれぞれ数式により定式化しておく．

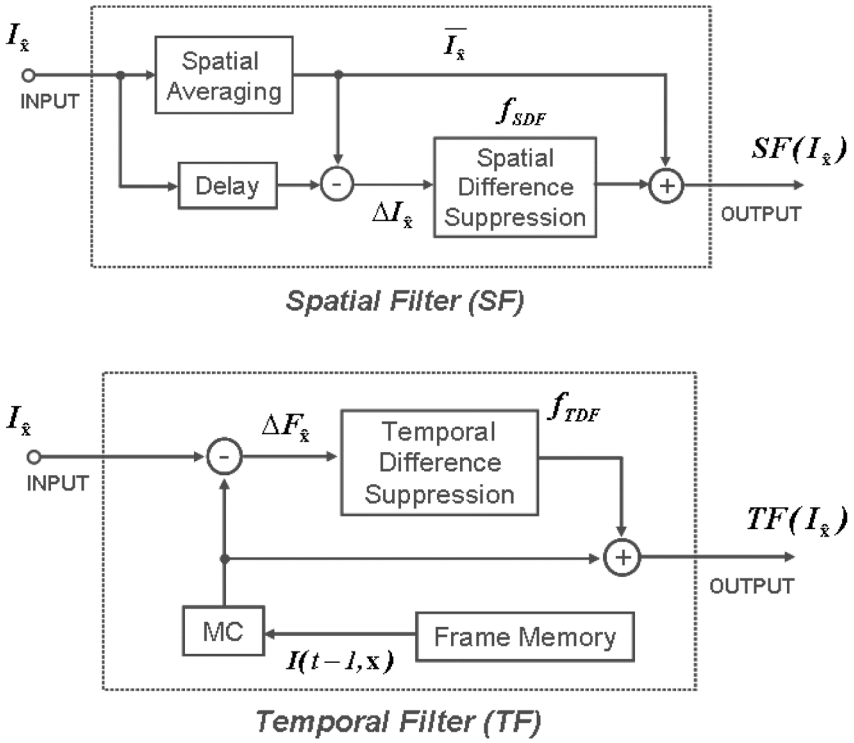


図 5.2: 従来の時空間平滑化フィルタの構成

空間平滑化フィルタの定式化　画像列 I のあるフレーム内において，画素 $\mathbf{x}$ の値を $I_{\hat{\mathbf{x}}}$ とすると，空間 4-近傍を用いた平均値フィルタで対象フレームを平滑化した画像 $\overline{I}_{\hat{\mathbf{x}}}$ は，

$$\overline{I}_{\hat{\mathbf{x}}} = \frac{1}{1+|\eta^{sp}_{\hat{\mathbf{x}}}|}\Big(I_{\hat{\mathbf{x}}} + \sum_{\hat{\xi}\in\eta^{sp}_{\hat{\mathbf{x}}}} I_{\hat{\xi}}\Big) \tag{5.6}$$

と書くことができる．本章では，特に空間 4-近傍を用いており，$|\eta^{sp}_{\hat{\mathbf{x}}}| = 4$ が成り立つことから，右辺の係数は $1/5$ となる．

このとき，空間平滑化フィルタは以下のように定式化される．

$$\begin{aligned} SF(I_{\hat{\mathbf{x}}}) &= \overline{I}_{\hat{\mathbf{x}}} + f_{SDS}\,(\Delta I_{\hat{\mathbf{x}}}), \\ &\qquad \text{where} \quad \Delta I_{\hat{\mathbf{x}}} = I_{\hat{\mathbf{x}}} - \overline{I}_{\hat{\mathbf{x}}}. \end{aligned} \tag{5.7}$$

$\Delta I_{\hat{\mathbf{x}}}$ は対象フレームの 2 次元座標位置 $\mathbf{x}$ における原画像と平滑化画像との差分値であり，これを引数に取る関数 $f_{SDS}\,(x)$ は，その差分値の微小な変動を抑制する役割を持つことから，空間差分抑制 (SDS: Spatial Difference Suppression) 関数と呼ばれる．

時間平滑化フィルタの定式化　画像列 I において，時刻 t のフレーム内画素を $\hat{\mathbf{x}} = (t, \mathbf{x})$，その画素値を $I_{\hat{\mathbf{x}}} = I(t, \mathbf{x})$ とする．このとき，時間平滑化フィルタは以下のように定式化さ

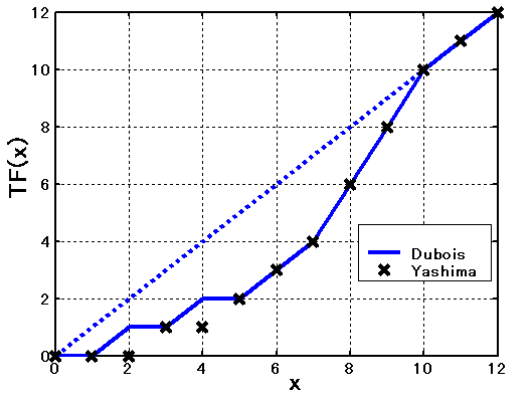


図 5.3: 従来法の TDS 関数

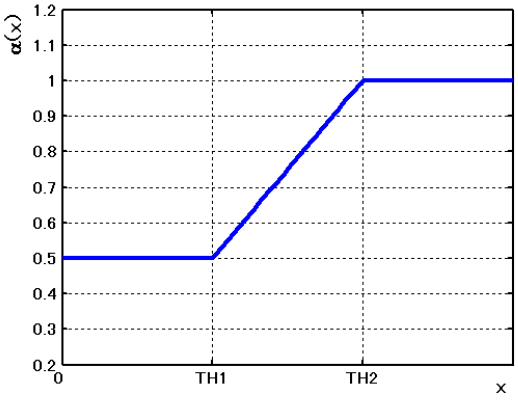


図 5.4: 係数 $\alpha(x)$ のプロット

れる．

$$TF(I_{\hat{\mathbf{x}}}) \quad = \quad I_{t-1}(\mathbf{x}) + f_{TDS}\,(\Delta F_{\hat{\mathbf{x}}}), \qquad (5.8)$$
$$\text{where} \quad \Delta F_{\hat{\mathbf{x}}} = I_t(\mathbf{x}) - I_{t-1}(\mathbf{x}).$$

$\Delta F_{\hat{\mathbf{x}}}$ は画素位置 $\mathbf{x}$ におけるフレーム間差分値であり，これを引数に取る関数 $f_{TDS}\,(x)$ はフレーム間における時間方向の細かな画素値変動を抑制する役割を持つことから，時間差分抑制 (TDS: Temporal Difference Suppression) 関数と呼ばれる．

従来法における差分抑制関数について　差分抑制関数は，フレーム間の画素値の微小変動をゼロの方向に抑制し，予測誤差信号の情報量を削減する役割を持つ．このため，八島らが [98] で提示したように，微小な変動値 x に対してのみ下方に膨らむ図 5.3 のようなグラフとして，差分抑制関数が定義されていた．なお，図は $0 \leq x \leq 10$ を微小な変動として抑制した TDS 関数の例である．

一方，Dubois ら [18] により導入された TDS 関数は次の式によって定義される．

$$f_{TDS}(x) = \alpha(x) \cdot x \qquad (5.9)$$

ここで，係数 α は x の関数であり，図 5.4 のグラフに基づいて定義される．$TH1 = 5, TH2 = 10$ として得られた Dubois の TDS 関数を floor 関数で量子化したプロットを図 5.3 に示す．

図 5.3 のプロットを比較すると，Dubois らの TDS 関数と八島らの TDS 関数は，定義の仕方こそ異なるものの，共にほぼ同じ特性を持つことがわかる．このため，以下の検討においては，数式としてより扱いやすい Dubois の定式化 (5.9) を採用する．

5.3.2　空間方向の非等方拡散法に関する理論的考察

ここでは，空間方向の差分方程式

$$I_{\hat{\mathbf{x}}}^{(n+1)} \;=\; I_{\hat{\mathbf{x}}}^{(n)} + \frac{1}{5}\sum_{\hat{\xi}\in\eta_{\hat{\mathbf{x}}}^{sp}} g^{sp}(\nabla_{\hat{\xi}} I_{\hat{\mathbf{x}}}^{(n)})\cdot\nabla_{\hat{\xi}} I_{\hat{\mathbf{x}}}^{(n)} \tag{5.10}$$

が，SDS 関数に基づく空間平滑化フィルタとして定式化できることを示す．簡単のため，$\Delta_\lambda \equiv \nabla_{\hat{\mathbf{x}}_\lambda} I_{\hat{\mathbf{x}}}^{(n)} = I_{\hat{\mathbf{x}}_\lambda}^{(n)} - I_{\hat{\mathbf{x}}}^{(n)}$, $\lambda\in\Lambda\equiv\{N,S,E,W\}$ と表記するとき，次の命題が成立する．

命題 5.1 空間 *4*-近傍を用いた非等方拡散法の反復ステップに関して，次の *i), ii)* が成り立つ．

i) 式 *(5.10)* は次の差分抑制関数に基づく空間平滑化フィルタである．

$$f_{SDS}\left(\Delta_N,\Delta_S,\Delta_E,\Delta_W\right) = -\frac{1}{5}\sum_{\lambda\in\Lambda}(1-g^{sp}(\Delta_\lambda))\cdot\Delta_\lambda \tag{5.11}$$

ii) 特にエッジ停止関数 g^{sp} が変数の総和 $\sum_{\lambda\in\Lambda}\Delta_\lambda$ の関数であるとき，差分抑制関数 f_{SDS} は $\Delta I_{\hat{\mathbf{x}}}$ の関数となり，次の式

$$f_{SDS}\left(x\right) = (1-g^{sp}(x))\cdot x \tag{5.12}$$

で与えられる．

[証明]

i) 簡単のため，$\Delta_\lambda \equiv \nabla_{\hat{\mathbf{x}}_\lambda} I_{\hat{\mathbf{x}}}^{(n)}$, $\lambda\in\Lambda\equiv\{N,S,E,W\}$ と表記する．式 (5.10) から，

$$\begin{aligned} I_{\hat{\mathbf{x}}}^{(n+1)} &= I_{\hat{\mathbf{x}}}^{(n)} + \frac{1}{5}\sum_{\lambda\in\Lambda} g^{sp}(\Delta_\lambda)\cdot\Delta_\lambda \\ &= \overline{I_{\hat{\mathbf{x}}}^{(n)}} + \Big(\Delta I_{\hat{\mathbf{x}}} + \frac{1}{5}\sum_{\lambda\in\Lambda} g^{sp}(\Delta_\lambda)\cdot\Delta_\lambda\Big) \end{aligned} \tag{5.13}$$

となる．ここで，

$$\Delta I_{\hat{\mathbf{x}}} = -\frac{1}{5}(\Delta_N+\Delta_S+\Delta_E+\Delta_W) \tag{5.14}$$

であることを考慮すれば，SDS 関数の定義式 (5.7) から，

$$\begin{aligned} f_{SDS}\left(\Delta_N,\Delta_S,\Delta_E,\Delta_W\right) &= -\frac{1}{5}\sum_{\lambda\in\Lambda}\Delta_\lambda + \frac{1}{5}\sum_{\lambda\in\Lambda} g^{sp}(\Delta_\lambda)\cdot\Delta_\lambda \\ &= -\frac{1}{5}\sum_{\lambda\in\Lambda}(1-g^{sp}(\Delta_\lambda))\cdot\Delta_\lambda \end{aligned} \tag{5.15}$$

が得られる．∎

ii) 式 (5.14) から，g^{sp} は $\Delta I_{\hat{\mathbf{x}}}$ の関数となり，その値は λ に関わらず等方的に定まることから，

$$\begin{aligned} f_{SDS}\,(\Delta I_{\hat{\mathbf{x}}}) &= (1 - g^{sp}(\Delta I_{\hat{\mathbf{x}}}))\Big(- \frac{1}{5}\sum_{\lambda\in\Lambda}\Delta_\lambda\Big) \\ &= (1 - g^{sp}(\Delta I_{\hat{\mathbf{x}}}))\cdot\Delta I_{\hat{\mathbf{x}}} \end{aligned} \tag{5.16}$$

がいえる．ゆえに，差分抑制関数 f_{SDS} もまた $\Delta I_{\hat{\mathbf{x}}}$ の関数となり，式（5.12）が成立する．∎

以上の命題から次のことがわかる．各近傍画素への差分値の総和に対して値が一意に定まるエッジ停止関数を導入することで，すべての近傍画素方向に対して等しい重みを与えたとき，式 (5.10) の反復ステップは SDS 関数に基づく従来の空間平滑化フィルタと同等になる．すなわち，空間方向の非等方拡散法は従来の空間平滑化フィルタに非等方性を導入して得られる自然な拡張手法と理解することができる．

次に，式 (5.10) の n 回反復後のフィルタ効果について検討する．以下では，エッジ停止関数 g^{sp} が変数の総和 $\sum_{\lambda\in\Lambda}\Delta_\lambda$ の関数であり，それゆえ命題 5.1 から，空間方向の拡散項が式 (5.12) の SDS 関数により定式化されている場合を考える．このとき，画素 $\hat{\mathbf{x}}$ の空間 4-近傍 $\eta^{sp}_{\hat{\mathbf{x}}}$ に閉じた画像[2]に対して，次の命題が成立する．

命題 5.2 画素 $\hat{\mathbf{x}}$ の空間 *4*-近傍 $\eta^{sp}_{\hat{\mathbf{x}}}$ に閉じた画像に対して，n 回の反復を伴う空間方向の非等方拡散フィルタは，次の差分抑制関数に基づく空間平滑化フィルタに等しい．

$$f^{(n)}_{SDS}\,(\Delta I_{\hat{\mathbf{x}}}) = (1 - g(\Delta I^{(n-1)}_{\hat{\mathbf{x}}}))\cdot\Delta I^{(n-1)}_{\hat{\mathbf{x}}} \tag{5.17}$$

ここで，$\Delta I^{(i)}_{\hat{\mathbf{x}}}\ (i = 0, 1, 2, ..., n-1)$ は，次の漸化式

$$\begin{aligned} \Delta I^{(n)}_{\hat{\mathbf{x}}} &= (1 - g(\Delta I^{(n-1)}_{\hat{\mathbf{x}}}))\cdot\Delta I^{(n-1)}_{\hat{\mathbf{x}}} \\ \Delta I^{(0)}_{\hat{\mathbf{x}}} &= \Delta I_{\hat{\mathbf{x}}} \qquad\quad (n = 1, 2, 3, ...) \end{aligned} \tag{5.18}$$

によって与えられる．

[証明]　$\Delta I^{(n)}_{\hat{\mathbf{x}}}$ と $\Delta I^{(n+1)}_{\hat{\mathbf{x}}}$ の関係を調べる．画素 $\hat{\mathbf{x}}$ における 1 回の拡散処理の結果 $I^{(n+1)}_{\hat{\mathbf{x}}}$ は式 (5.13) で与えられている．また，$\eta^{sp}_{\hat{\mathbf{x}}}$ の各画素における 1 回の拡散処理の結果は，

$$I^{(n+1)}_{\hat{\mathbf{x}}_N} = I^{(n)}_{\hat{\mathbf{x}}_N} - \frac{1}{5}g(\Delta_N)\cdot\Delta_N \tag{5.19}$$

$$I^{(n+1)}_{\hat{\mathbf{x}}_S} = I^{(n)}_{\hat{\mathbf{x}}_S} - \frac{1}{5}g(\Delta_S)\cdot\Delta_S \tag{5.20}$$

$$I^{(n+1)}_{\hat{\mathbf{x}}_E} = I^{(n)}_{\hat{\mathbf{x}}_E} - \frac{1}{5}g(\Delta_E)\cdot\Delta_E \tag{5.21}$$

$$I^{(n+1)}_{\hat{\mathbf{x}}_W} = I^{(n)}_{\hat{\mathbf{x}}_W} - \frac{1}{5}g(\Delta_W)\cdot\Delta_W \tag{5.22}$$

[2]近傍 $\eta^{sp}_{\hat{\mathbf{x}}}$ に閉じた画像とは，近傍外の画素に対する差分値をすべて 0 と評価できる画像と定義する．

で与えられる．ここで，$\hat{\mathbf{x}}$ の 4-近傍 $\eta^{sp}_{\hat{\mathbf{x}}}$ に閉じた画像を想定しているため，近傍外部に対する差分値はすべて 0 と評価している．このとき，以上の結果から，

$$\begin{aligned}\overline{I^{(n+1)}_{\hat{\mathbf{x}}}} &= \frac{1}{5}\Big[(I^{(n)}_{\hat{\mathbf{x}}}+\frac{1}{5}\sum_{\lambda\in\Lambda} g(\Delta_\lambda)\cdot\Delta_\lambda)+(I^{(n)}_{\hat{\mathbf{x}}_N}+I^{(n)}_{\hat{\mathbf{x}}_S}+I^{(n)}_{\hat{\mathbf{x}}_E}+I^{(n)}_{\hat{\mathbf{x}}_W})-\frac{1}{5}\sum_{\lambda\in\Lambda} g(\Delta_\lambda)\cdot\Delta_\lambda\Big]\\ &= \frac{1}{5}(I^{(n)}_{\hat{\mathbf{x}}}+I^{(n)}_{\hat{\mathbf{x}}_N}+I^{(n)}_{\hat{\mathbf{x}}_S}+I^{(n)}_{\hat{\mathbf{x}}_E}+I^{(n)}_{\hat{\mathbf{x}}_W})\\ &= \overline{I^{(n)}_{\hat{\mathbf{x}}}} \end{aligned} \tag{5.23}$$

と評価できることから，結局，$\Delta I^{(n)}_{\hat{\mathbf{x}}}$ と $\Delta I^{(n+1)}_{\hat{\mathbf{x}}}$ の関係は，

$$\begin{aligned}\Delta I^{(n+1)}_{\hat{\mathbf{x}}} &= I^{(n+1)}_{\hat{\mathbf{x}}}\\ &= (I^{(n)}_{\hat{\mathbf{x}}}-\overline{I^{(n)}_{\hat{\mathbf{x}}}})-g(\Delta I^{(n)}_{\hat{\mathbf{x}}})\cdot\Delta I^{(n)}_{\hat{\mathbf{x}}}\\ &= \Delta I^{(n)}_{\hat{\mathbf{x}}}-g(\Delta I^{(n)}_{\hat{\mathbf{x}}})\cdot\Delta I^{(n)}_{\hat{\mathbf{x}}}\\ &= (1-g(\Delta I^{(n)}_{\hat{\mathbf{x}}}))\cdot\Delta I^{(n)}_{\hat{\mathbf{x}}}\end{aligned} \tag{5.24}$$

となり，差分抑制関数 f_{SDS} として，式 (5.18) が得られる． ∎

5.3.3　時間方向の非等方拡散に関する理論的考察

次に，時間方向の差分方程式

$$I^{(n+1)}_{\hat{\mathbf{x}}} = I^{(n)}_{\hat{\mathbf{x}}}+\frac{1}{2}\sum_{\hat{\xi}\in\eta^{tp}_{\hat{\mathbf{x}}}} g^{tp}(\nabla_{\hat{\xi}}I^{(n)}_{\hat{\mathbf{x}}})\cdot\nabla_{\hat{\xi}}I^{(n)}_{\hat{\mathbf{x}}} \tag{5.25}$$

について，TDS 関数に基づく考察を行う．また，式 (5.25) を n 回反復して得られるフィルタの平滑化効果についてもあわせて考察する．

一般に，時間方向の非等方拡散フィルタによる n 回反復処理後の平滑化画像は前後 n フレームの画素の影響を受けるため，連続した 2 フレーム間の関係を表す TDS 関数では完全に記述することはできない．このため，以下の議論では，現在のフレーム I_t とその直前フレーム I_{t-1} の 2 フレームのみからなる時刻 t の瞬間の画像列を考え，必要に応じて $I_t = I_{t+1} = I_{t+2} = \cdots$，$I_{t-1} = I_{t-2} = I_{t-3} = \cdots$ のように画像列を時間軸前後に拡張することにより，時間方向の非等方拡散フィルタの特性を近似的に解析する．

まず，式 (5.25) に関する次の命題が成り立つ．

命題 5.3 連続する *2* フレーム I_{t-1}, I_t からなる時刻 t の瞬間の画像列に対して，時間 *2*-近傍を用いた非等方拡散法の反復ステップの式 *(5.25)* は，次の差分抑制関数

$$f_{TDS}(x) = \Big(1-\frac{1}{2}g^{tp}(x)\Big)\cdot x \tag{5.26}$$

に基づく時間平滑化フィルタに等しい．

[証明]　簡単のため，$\Delta_\lambda \equiv \nabla_{\hat{\mathbf{x}}_\lambda} I^{(n)}_{\hat{\mathbf{x}}}$, $\lambda \in \Lambda \equiv \{F, B\}$ と書く．式 (5.25) において，$\Delta_B = -\Delta F^{(n)}$，$\Delta_F = 0$ となることから，

$$\begin{aligned} I_t^{(n+1)} &= I_t^{(n)} + \frac{1}{2}\sum_{\lambda\in\Lambda} g^{tp}(\Delta_\lambda)\cdot\Delta_\lambda \\ &= I_t^{(n)} + \frac{1}{2} g^{tp}(-\Delta F^{(n)})\cdot(-\Delta F^{(n)}) \\ &= I_t^{(n)} - \frac{1}{2} g^{tp}(\Delta F^{(n)})\cdot\Delta F^{(n)} \end{aligned} \tag{5.27}$$

が得られる．さらに，式 (5.27) は，

$$\begin{aligned} I_t^{(n+1)} &= I_{t-1}^{(n)} + \Delta F^{(n)} - \frac{1}{2} g^{tp}(\Delta F^{(n)})\cdot\Delta F^{(n)} \\ &= I_{t-1}^{(n)} + \Big(1 - \frac{1}{2} g^{tp}(\Delta F^{(n)})\Big)\cdot\Delta F^{(n)} \end{aligned} \tag{5.28}$$

と書ける．したがって，TDS 関数の定義式 (5.8) から式 (5.26) が得られる．∎

また，式 (5.25) の n 回反復に関して，次の命題が成り立つ．

命題 5.4 連続する 2 フレーム I_{t-1}, I_t からなる時刻 t の瞬間の画像列に対して，n 回の反復を伴う時間方向の非等方拡散フィルタは，次の差分抑制関数に基づく時間平滑化フィルタに等しい．

$$f^{(n)}_{TDS}(\Delta F) = \Delta F - \frac{1}{2}\sum_{i=0}^{n-1} g^{tp}(\Delta F^{(i)})\cdot\Delta F^{(i)} \tag{5.29}$$

ここで，$F^{(i)}$ $(i = 0, 1, 2, ..., n-1)$ は，次の漸化式

$$\begin{aligned} \Delta F^{(i)} &= (1 - g^{tp}(\Delta F^{(i-1)}))\cdot\Delta F^{(i-1)} \\ \Delta F^{(0)} &= \Delta F \quad (i = 1, 2, .., n-1) \end{aligned} \tag{5.30}$$

によって与えられる．

[証明]　まず，式 (5.27) を繰り返し適用することによって，

$$\begin{aligned} I_t^{(n)} &= I_t^{(n-1)} - \frac{1}{2} g^{tp}(\Delta F^{(n-1)})\cdot\Delta F^{(n-1)} \\ &= I_t^{(n-2)} - \frac{1}{2} g^{tp}(\Delta F^{(n-2)})\cdot\Delta F^{(n-2)} \\ &\qquad - \frac{1}{2} g^{tp}(\Delta F^{(n-1)})\cdot\Delta F^{(n-1)} \\ &= ... \\ &= I_t^{(0)} - \frac{1}{2}\sum_{i=0}^{n-1} g^{tp}(\Delta F^{(i)})\cdot\Delta F^{(i)} \\ &= I_{t-1}^{(0)} + \Delta F^{(0)} - \frac{1}{2}\sum_{i=0}^{n-1} g^{tp}(\Delta F^{(i)})\cdot\Delta F^{(i)} \end{aligned}$$

となる．したがって，TDS 関数の定義式 (5.8) から式 (5.29) が得られる．

次に，ΔF から $\Delta F^{(i)}$ $(i = 0, 1, 2, ..., n-1)$ の導出を行う．式 (5.27) に相当する時刻 $t-1$ における時間方向の拡散から，

$$\begin{aligned} I_{t-1}^{(n+1)} &= I_{t-1}^{(n)} + \frac{1}{2}\sum_{\lambda\in\Lambda} g^{tp}(\Delta_\lambda)\cdot\Delta_\lambda \\ &= I_t^{(n)} + \frac{1}{2} g^{tp}(\Delta F^{(n)})\cdot\Delta F^{(n)} \end{aligned} \tag{5.31}$$

が得られる．したがって，式 (5.27) と式 (5.31) から，$I^{(n+1)}$ のフレーム間差分 $\Delta F^{(n+1)}$ を評価すると，

$$\begin{aligned} \Delta F^{(n+1)} &= I^{(n+1)}(t, \mathbf{x}) - I^{(n+1)}(t-1, \mathbf{x}) \\ &= \Delta F^{(n)} - g^{tp}(\Delta F^{(n)})\cdot\Delta F^{(n)} \\ &= (1 - g^{tp}(\Delta F^{(n)}))\cdot\Delta F^{(n)} \end{aligned} \tag{5.32}$$

となり，式 (5.30) が得られる． ∎

これらの命題により，時間方向の非等方拡散フィルタを差分抑制関数の枠組みで解析できるようになる．次章では，以上の結果に基づき，時間方向のエッジ停止関数の選定に関する理論的な検討を行う．

5.4　時間方向のエッジ停止関数に関する検討

時間方向のエッジ停止関数 g^{tp} は，フレーム間差分の絶対値が大きくなる画素の平滑化を抑制するため，時間的に不連続な画素間の誤った平滑化に伴う残像等の視覚的妨害を緩和する目的で使用することができる．このとき，フレーム間予測誤差の低減と主観画質の維持を両立させるためには，適切なエッジ停止関数を選定することが重要である．

以下では，平滑化効果の異なる 3 種類のエッジ停止関数を候補として用意し，これらの TDS 関数の挙動と平滑化画像の主観画質を調べることにより，時間フィルタとして最もふさわしい関数を選定する．

5.4.1　候補となるエッジ停止関数

時間方向のエッジ停止関数を選ぶには，フレーム間差分信号の低減とオブジェクト残像などの主観画質のトレードオフを考慮しなくてはならない．ここでは，平滑化抑制効果の異なるエッジ停止関数として，g^{PM}，g^{BS} 及び g^{TN} を時間方向のエッジ停止関数の候補とする．図 5.5 にこれら 3 種類のエッジ停止関数のプロットを示す．ここでは，$K = 10$ としたときの g^{BS} を基準とし，直線 $y = \epsilon$ を横切る際の x 座標がいずれも同じ値を取るように各プロットを調整した. なお，図は $\epsilon = 0.1$ とした場合の例である．

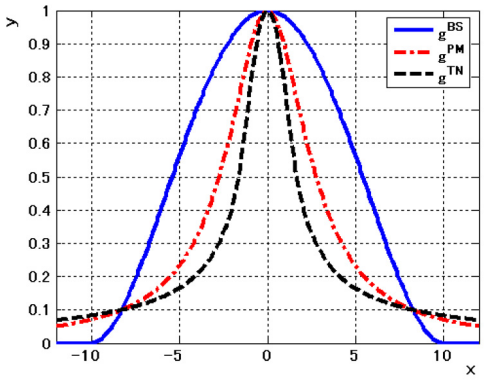


図 5.5: 候補となるエッジ停止関数

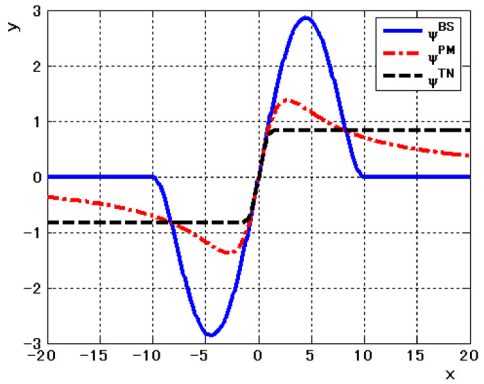


図 5.6: 各エッジ停止関数に対応する Ψ 関数

エッジ点における勾配値のような，はずれ値に対するエッジ停止関数の平滑化抑制効果は，influence function $\Psi(x) = g(x)x$（以下，Ψ 関数と呼ぶ）によって評価することができる [9]．図 5.6 に候補となる 3 種類のエッジ停止関数に対する Ψ 関数のプロットを示す．$|x| > 10$ に対するプロットを比較することにより，エッジをまたがる平滑化の抑制効果は，$g^{TN} < g^{PM} < g^{BS}$ の順で強くなっていることがわかる．以下では，理論的な考察とシミュレーション実験からこれらのエッジ停止関数の比較検討を行う．

5.4.2　TDS 関数の挙動に関する考察

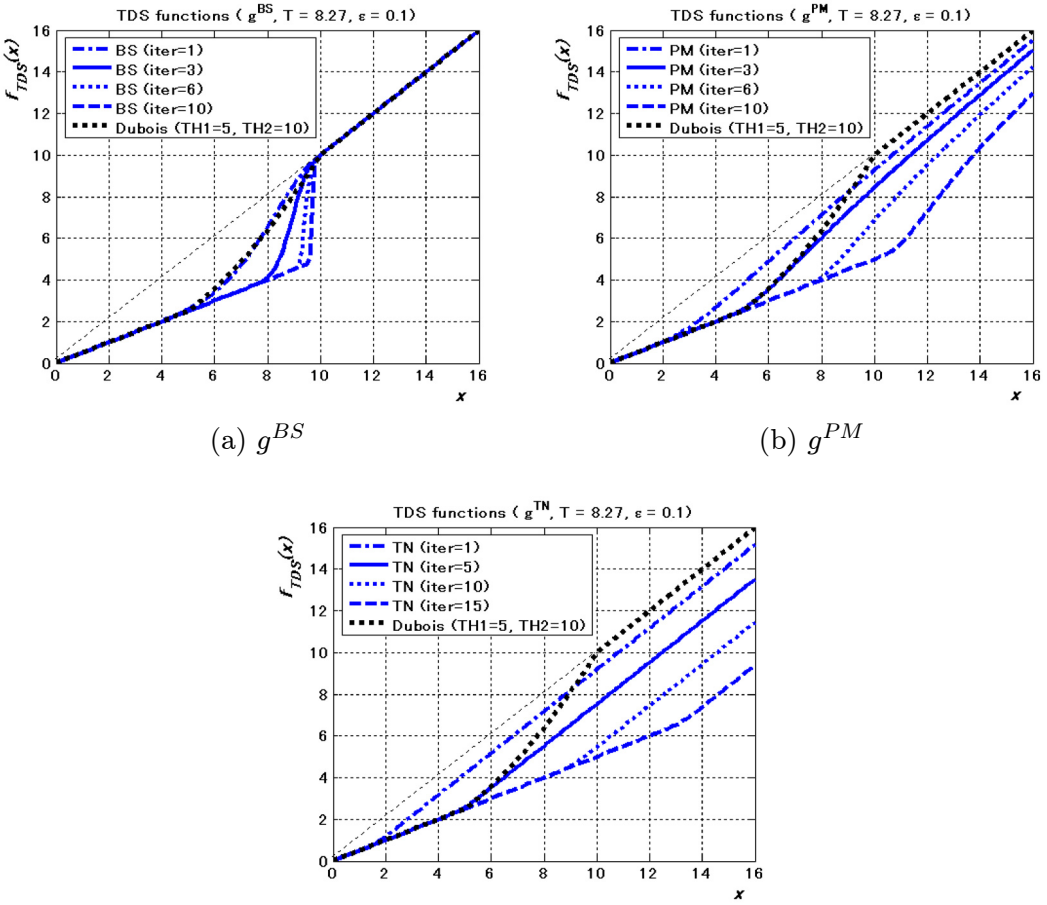
ここでは，上で提示した 3 種類のエッジ停止関数 g^{BS}，g^{PM}，g^{TN} から得られる時間方向の非等方拡散フィルタが，それぞれどのような TDS 関数に従うかを検討する．

先の考察で述べたように，式 (5.25) の n 回反復に基づく非等方拡散フィルタに対して，その TDS 関数は式 (5.29),(5.30) に従って順次計算することができる．計算機を用いてこの漸化式を解き，得られた TDS 関数をプロットした結果を図 5.7 に示す．平滑化抑制効果の高い順に，図 (a) が g^{BS}，(b) が g^{PM}, (c) が g^{TN} のプロットをそれぞれ示している．g^{BS}，g^{PM} に対しては繰り返し回数 $iter = 1, 3, 6, 10$ の結果を，g^{TN} に対しては繰り返し回数 $iter = 1, 5, 10, 15$ の結果をそれぞれ示した．また，Dubois の時間フィルタに対する TDS 関数 (5.9) において，関数 $\alpha(x)$ の閾値を $TH1 = 5$, $TH2 = 10$ として得られるプロットを比較のため併記してある．

図 5.7 を見ると，従来の時間平滑化フィルタに最も近い挙動を示すのは g^{BS} である．繰り返し回数 $iter = 1$ のプロットは，Dubois の時間フィルタに対する TDS 関数とほとんど重なっていることがわかる．さらに繰り返し回数を重ねると，プロットは急速に不連続関数

$$y = \frac{1}{2}x \ \ (|x| < 10), \text{ or } \ \ x \ \ (otherwise)$$

に収束することから，10 を超える差分値に対して抑制効果が働くことはない．このため，フレーム間予測誤差の低減に関する効果については不十分である可能性がある．



(a) g^{BS}　(b) g^{PM}

(c) g^{TN}

図 5.7: g^{tp} による時間拡散の TDS 関数

これに対し，g^{PM}，g^{TN} の場合は，10 を超える差分値に対してもある程度の差分抑制効果がある．g^{PM} に対する TDS 関数のプロットは，x が大きくなるにつれて漸近的に直線 $y = x$ に近づいてゆくため，繰り返し回数を増やすほど，より大きな差分値 x に対しても抑制効果が表れるようになる．一方，g^{TN} の場合，差分値 x がある一定の値を超えると，TDS 関数のプロットは直線 $y = x$ と完全に平行になってしまう．このため，常にすべての差分値に対して同程度の抑制効果が働いてしまうことから，平滑化に伴う残像などの弊害が懸念される．

このように，各 TDS 関数のプロットから判断する限り，g^{PM} の TDS 関数が，予測誤差の低減と主観画質の維持の両面において，最もよくバランスが取れていることがわかる．

5.4.3　主観画質への影響

時間平滑化に伴う主観画質への影響を確認するため，標準画像シーケンス Football に対して，式 (5.25) で与えられる時間方向のみの拡散法を適用した．ここでは，低ビットレート環境の動画像を想定しているため，フレームを 3 枚毎に採用した 10fps の画像列を対象に実験を行った．図 5.8 は平滑化処理後の第 19 フレームの結果を示したものである．図 5.8(a) に示す第 19 フレームとその前後のフレームの原画像から，画面内のプレイヤーが激しく動いている様子がわかる．

図 5.8(b) は，$g^{tp} \equiv 1$ として Gaussian と同等の時間平滑化を行った結果である．オブジェクトの動きに伴う残像が前後に現れていることが確認できる．これに対し図 5.8(c),(d),(e) は，エッジ停止関数としてそれぞれ g^{BS}, g^{PM}, g^{TN} を用いた，時間方向の拡散に基づく処理画像である．(c),(d) においては，エッジ停止関数の効果により，いずれも (b) に見られるような残像は解消されている．一方，(e) の場合にも残像はかなり解消されているが，よく観察すると，オブジェクトの前後にかすかな雑音が生じていることが確認できる．

各フィルタの平滑化効果を確認するため，16^{th}-19^{th} フレーム間差分信号に対して，原画像の RMSE (Root Mean Square Error) を基準とする，各平滑化画像の RMSE の比率を計算した．図 5.8 のキャプションにその値を示す．RMSE から比較する限り，g^{BS}, g^{PM}, g^{TN} の順に平滑化効果が高くなっていくことがわかる．また，このことは，画像の芝生部分の平滑化の様子から，主観的にも確認することができる．

5.5　まとめ

Perona-Malik による非等方拡散法を時空間 3 次元の画像列に拡張し，エッジ保存形の時空間平滑化フィルタを実現する方法について提案を行った．本フィルタは，特に低ビットレート動画像符号化の前処理における適用を想定したものである．

提案する時空間非等方拡散フィルタは，従来の空間方向の拡散項と新たに追加した時間方向の拡散項から構成される．本論文では，ある特定の条件の下で，これらの拡散項が差分抑制関数に基づいて定式化できることを数理的に明らかにした．この意味において，非等方拡散フィルタは従来法で用いられる非線形平滑化フィルタの自然な拡張とみなすことができる．

続いて，時間方向の拡散項に使用するエッジ停止関数の選定に関する検討を行った．ここでは，まず平滑化効果の異なる 3 種類の関数 g^{BS}, g^{PM}, g^{TN} を候補として提示し，これらの中から最も適当な関数を選択する方法を採った．差分抑制関数に基づく理論的な考察とシミュレーションに基づく主観的な画質評価の両面から検討した結果，時間方向のエッジ停止関数には g^{PM} が最も適切であることがわかった．

今後の課題としては，時間方向の拡散項の導入に伴う情報量削減効果を，様々な画像列に対するシミュレーションを通して実証することが挙げられる．また，Dubois ら [18] の手法に見られる motion trajectory に沿った時間平滑化のアイディアを，提案法である非等方拡散フィルタに導入する方法についても検討する必要がある．

(a) the 16^{th}, 19^{th}, and 22^{th} frames from Original Image Sequence

RMSE ratio: 0.30

(b) Gaussian(iter=1)

RMSE ratio: 0.84

(c) g^{BS}(iter=10)

RMSE ratio: 0.77

(d) g^{PM}(iter=10)

RMSE ratio: 0.64

(e) g^{TN}(iter=10)

図 5.8: 時間平滑化による主観画像品質の比較

第6章　ミッドレベル符号化における前処理技術 —VOP生成手法を中心に—

6.1　はじめに

MPEG-4はマルチメディアアプリケーションのためのオーディオ・ビジュアル符号化標準であり，モバイルネットワークやインターネットにおける低ビットレート映像通信・配信サービスから高品質なディジタルTV放送サービスに至るまで，その利用用途は広範囲にわたっている．MPEG-4ビジュアル符号化 (ISO/IEC 14496-2) の特筆すべき特徴の一つとして，その基本概念としてオブジェクトベース符号化が取り入れられたことが挙げられる．MPEG-4に準拠するエンコーダは，画像を構成する複数の前景オブジェクトと背景をそれぞれ個別に符号化したのち，各オブジェクトの圧縮ストリームを伝送路に多重して伝送する．一方，受信側のデコーダでは，送られてきた1本の多重化ストリームを各オブジェクトのストリームに分離し，それぞれを伸張したのち，一つの画面として合成して表示する．このとき，画面を構成するオブジェクトは特に矩形である必要はなく，テクスチャ情報に加えて形状情報を送信することにより，任意形状の領域を扱うことが可能である．

オブジェクトベース符号化を導入する第一のメリットとしては，ユーザによるオブジェクト単位のビデオ合成が可能になり，インタラクティブ性の高いコンテンツを提供できるようになることが挙げられる．さらに，もう一つの可能性として，符号化効率の向上への寄与というメリットも考えられる．たとえば，MPEG-4 Main Profileに提供されるスプライト符号化では，複数フレームから合成した背景画像を利用して，いったん背景を伝送した後は前景オブジェクトの情報のみを送ることにより，大幅な情報量削減を実現することが可能である．

しかしながら，一般の自然画像シーケンスにおいて，このようなオブジェクトベース符号化を適用することは容易ではない．MPEG-4の符号化単位であるVOP(Video Object Plane)を生成するためには，(1) 与えられた画像シーケンスからオブジェクトを検出し，(2) 検出したオブジェクトがビデオフレーム内に存在する間はこれを安定的に追跡し，(3) 各フレームからピクセル単位の正確なオブジェクト抽出を行う，という3つのステップが必要となるが，ここでいうオブジェクトの検出・追跡・抽出の問題はいずれ一つをとってもまだ技術が成熟しておらず，今なお活発に研究が行われている．

本章では，まず，オブジェクトベース符号化を実現する上で必要となるVOP自動生成技術の概要を述べるとともに，その研究事例として，エッジ情報に基づくVOP自動生成手法であるMeierの方法 [48] を取り上げ，そのアルゴリズムについて説明する．次に，本研究事例に基づく問題提起として，2値エッジ画像からの正確なオブジェクト輪郭抽出の問題を取

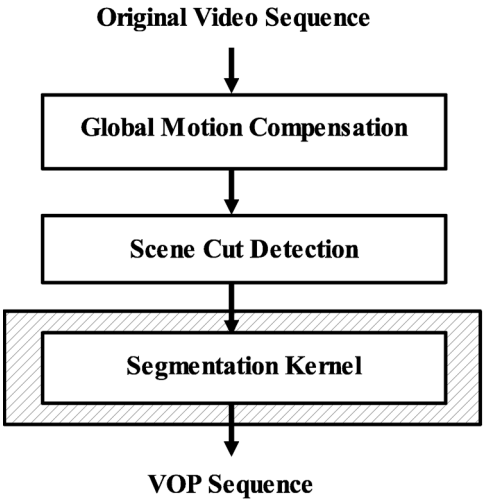


図 6.1: VOP 自動生成の共通フレームワーク [34]

り上げ，その従来研究について述べる．

6.2　VOP 自動生成技術の概要

6.2.1　共通フレームワーク

MPEG-4 の標準化文書 [34] には，VOP 自動生成手法の共通フレームワークに関する記述がある．これを図 6.1 に示す．このフレームワークによれば，VOP の自動生成にどのようなセグメンテーション・アルゴリズムを使用するにしても，その前処理としてグローバル動き補償とシーンカット点検出を行う必要がある．グローバル動き補償を行うことにより，画像シーケンスの背景は静止状態と考えることができるようになるため，カメラフレームに対して相対的に動いているオブジェクトを捉えやすくなる．また，シーンカット点検出が必要となるのは，カット点を境に画像の内容が変わることから，ビデオオブジェクトの定義をいったんリセットしなければならないためである．

本論文では，グローバル動き補償とシーンカット点検出についての検討は行わない．以下では，グローバル動き補償とシーンカット点検出が正しく行われた結果，背景が静止した連続なシーンが得られたものと仮定し，その後のセグメンテーションアルゴリズムについて検討を行うこととする．

6.2.2　従来研究

VOP 自動生成のためのセグメンテーション技術としては，MPEG-4 標準化の過程において 3 種類のアルゴリズムが提案されており，コア実験 N2 においてテストされた結果が参考資料として提示されている [34]．しかしながら，コア実験に提案された手法はいずれも完全に自動なアルゴリズムではなく，それぞれが固有のアルゴリズム上の弱点を抱えていたことが知られている．

一方，近年の VOP 自動生成アルゴリズムとしては，大きく分けて，領域情報に基づくアプローチとエッジ情報に基づくアプローチの 2 種類の方向性が存在する．領域情報に基づくアプローチとしては，Gu ら [26] による watershed アルゴリズムを用いた手法が提案され，最も有望な手法のひとつとして注目されている．また，エッジ情報に基づくアプローチとしては，Meier ら [48] による Hausdorff トラッキングを用いた手法が提案されている．

以下，本章では，VOP 自動生成手法の研究事例として，Meier らによるエッジベースの手法を取り上げ，そのアルゴリズムの説明を行う．また，次章以降の検討では，オブジェクト・トラッキングについては Meier の手法をそのまま利用しながら，主として 2 値エッジ画像からのオブジェクト輪郭抽出処理の部分を改善することが目的である．

6.3　エッジベースの VOP 自動生成（Meier の方法）

ここでは，Meier らによるエッジ情報に基づく VOP 自動生成アルゴリズム [48] の概要を述べ，本論文の目的となるオブジェクト抽出処理の位置づけについて説明する．

6.3.1　アルゴリズムの概要

Meier らによるエッジベースの VOP 自動生成手法のブロック図を図 6.2 に示す．また，本図で使用されている記号の定義を表 6.1 にまとめる．

表 6.1: 記号の定義 (図 6.2)

Notations	Definitions
$Im(i)$	i^{th} gray-scale image
$E(i)$	i^{th} binary edge image
$M(i)$	i^{th} object model
$MV(i)$	i^{th} motion vectors
$VOP(i)$	i^{th} VOP

オブジェクト抽出処理を行う前に必要となる前処理は，次の 3 つのステージから成る．すなわち，(1) 初期オブジェクトモデルの抽出 (Stage 1)，(2) エッジ画像の生成 (Stage 2)，(3) エッジベースのオブジェクト・トラッキング (Stage 3) である．これらの前処理のうち，最

も重要なパートは Stage 3 である．ここでは，オブジェクト抽出処理に先立ち，時間的に連続する 2 フレーム間において，ターゲットとなるオブジェクトのトラッキングを行う．このとき，1 フレーム先のオブジェクトを検出するために，2 種類の 2 値画像が使用される．一つはシーン全体のエッジを含む 2 値画像であり，これを “エッジ画像” と呼ぶ．もう一つは，トラッキングの対象となるオブジェクトのエッジのみを含む 2 値画像であり，これを “オブジェクトモデル” と呼ぶ．図 6.2 では，i 番目のフレームに対応するエッジ画像とオブジェクトモデルをそれぞれ $E(i)$, $M(i)$ と表記している．

Stage 3 の入力となる最初のオブジェクトモデル $M(1)$ は，Stage 1 において，2 枚の連続する濃淡画像 $Im(1)$, $Im(2)$ を用いて生成される．トラッキングに必要となるその後のオブジェクトモデルについては，Stage 3 ブロックの内部で生成され，自身の入力としてフィードバックされる．一方，各フレームのエッジ画像は Stage 2 において生成され，次の Stage においてオブジェクトモデルとともに使用される．本 VOP 自動生成手法の最終的な出力はビデオ・オブジェクト・プレーンであり，図 6.2 では $VOP(i)$ と表記されている．

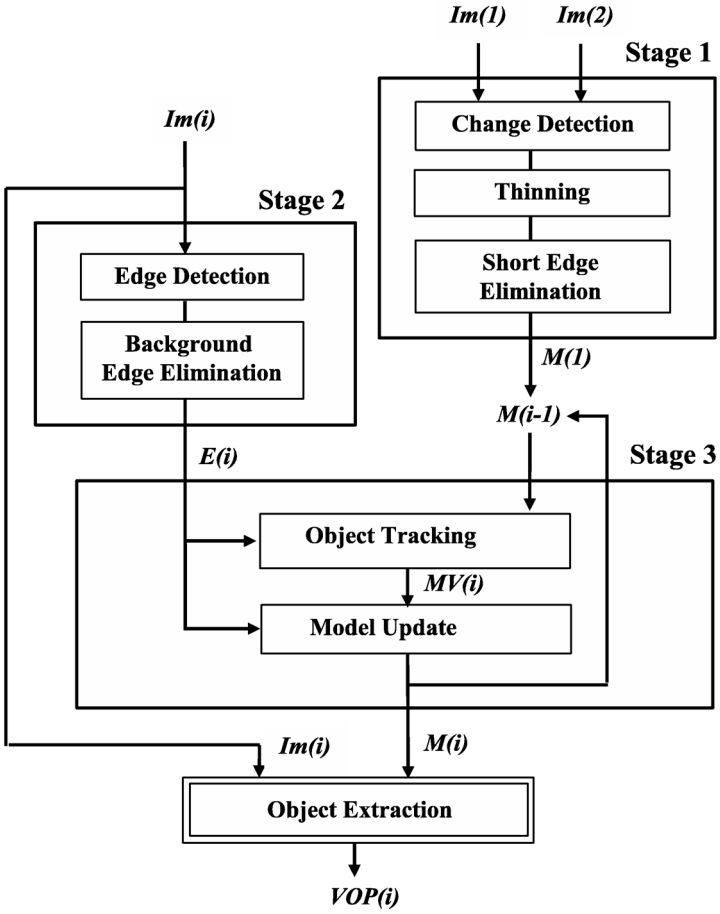
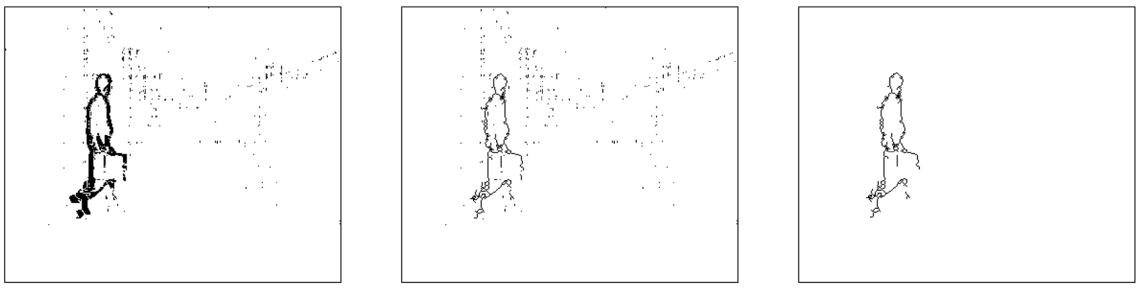


図 6.2: VOP 自動生成のブロック図



(a)　(b)　(c)

図 6.3: 初期オブジェクトモデルの抽出 (Stage 1): (a) 差分検出によるマスク (31^{st} - 32^{nd} frames), (b) マスクの細線化, (c) 短エッジの除去 (しきい値 8 pixels).

6.3.2　初期オブジェクトモデルの抽出 (Stage 1)

Stage 1 は，VOP 自動生成のための初期化に相当する処理を行う．ここでの目的は，時間的に連続する最初の 2 フレーム $Im(1)$，$Im(2)$ から，初期オブジェクトモデル $M(1)$ の生成を行うことである．図 6.3 に Stage 1 における画像処理の例を示す．ここで抽出したオブジェクトが，それ以降の画像シーケンスにおけるトラッキングの対象となる．

図 6.2 に示すように，Stage 1 は差分検出 (change detection)，細線化 (thinning)，短エッジの除去 (short-edge elimination) の，3 つの処理ブロックから構成される．以下に，各処理ブロックの内容について説明する．

まず，差分検出ブロックでは，最初の 2 フレームにおける動き領域を検出するため，$Im(1)$ と $Im(2)$ の間でフレーム間差分を取る．ただし，カメラ動きがある場合には，これに起因する差分を検出しないように，予めグローバル動き補償を行う必要がある．次に，画素値の変動が与えられたしきい値を超えた場合にその画素を動き領域として検出する．動き領域を検出しマスク画像を生成した例を図 6.3(a) に示した．

次に，細線化ブロックでは，差分検出によって得られたマスク画像を細線化し，オブジェクトの擬似輪郭線を生成する．これを実現するには，Holt の修正を加えた Zhang-Suen 細線化アルゴリズム [68] 等を使用すればよい．このアルゴリズムの主な利点として，マスク画像における連結性が細線化処理後のエッジ画像においてもそのまま維持される点が挙げられる．この性質により，マスク画像が連結している部分が細線化を通して分離してしまうことがなく，性質のよいオブジェクトモデルを生成することが可能となる．細線化によって得られたオブジェクトの擬似輪郭線の例を図 6.3(b) に示した．

図 6.3(b) を見ると，細線化後の画像には多くの短いエッジが混入していることがわかる．そこで，最後のブロックでは，上で得られた細線化画像に対して 8-近傍に基づく連結成分ラベリングを行い，連結成分を構成する画素の数が予め与えられたしきい値よりも少ない場合には，これをノイズとして除去するというフィルタリング処理を行っている．図 6.3(b) の擬似輪郭線に対して，しきい値を 8 画素に設定して，短エッジの除去を行った結果を図 6.3(c)

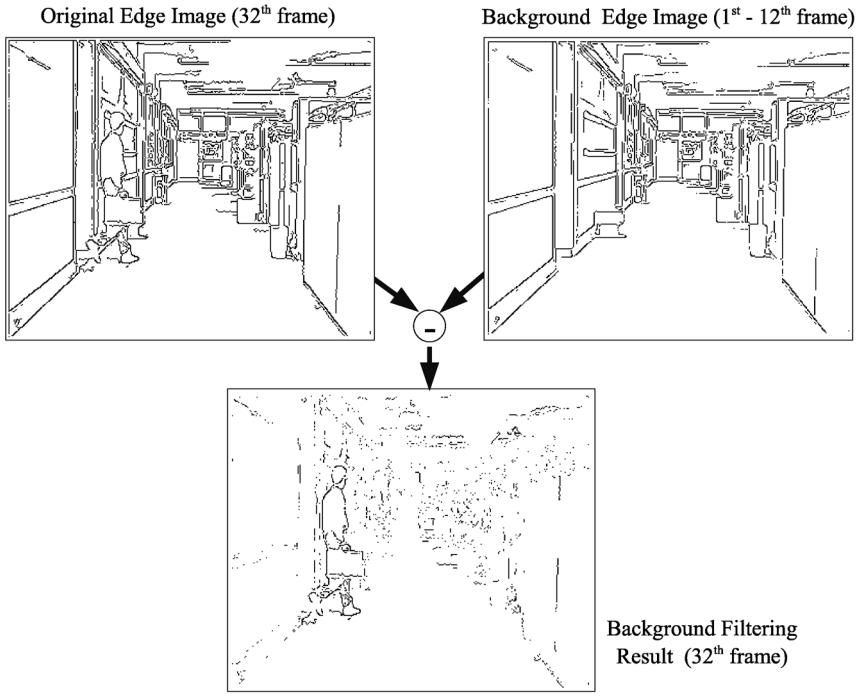


図 6.4: 背景エッジのフィルタリング (Stage 2)

に示した.

6.3.3　エッジ画像の生成 (Stage 2)

Stage 2 は，エッジ抽出 (edge detection)，背景エッジのフィルタリング (background filtering) の，2 つの処理ブロックから構成される．初めに，エッジ抽出ブロックでは，入力されたグレイスケール画像から，Canny オペレータ [10] を用いてエッジ抽出を行い，これを 2 値化してフレーム全体のエッジ画像を作成する．次に，背景エッジ・フィルタリングのブロックにおいて，背景領域のエッジを前もって取り除いておく．この処理は省略してもよいが，特に複雑な背景を持つ画像に対しては，これを行うことによって，次のステージで行うオブジェクトモデルのマッチング処理にかかる演算量を大幅に削減できるというメリットがある．図 6.4 に Stage 2 における背景エッジのフィルタリング処理の例を示す．ここでは，第 1-第 12 フレームから背景のエッジ画像を生成し，これとの差分を取ることによって，背景エッジのフィルタリングを行っている．なお，背景のエッジ画像は，全 12 フレームのうち 10 フレーム以上でエッジ点と判定された画素を選ぶことにより生成している．

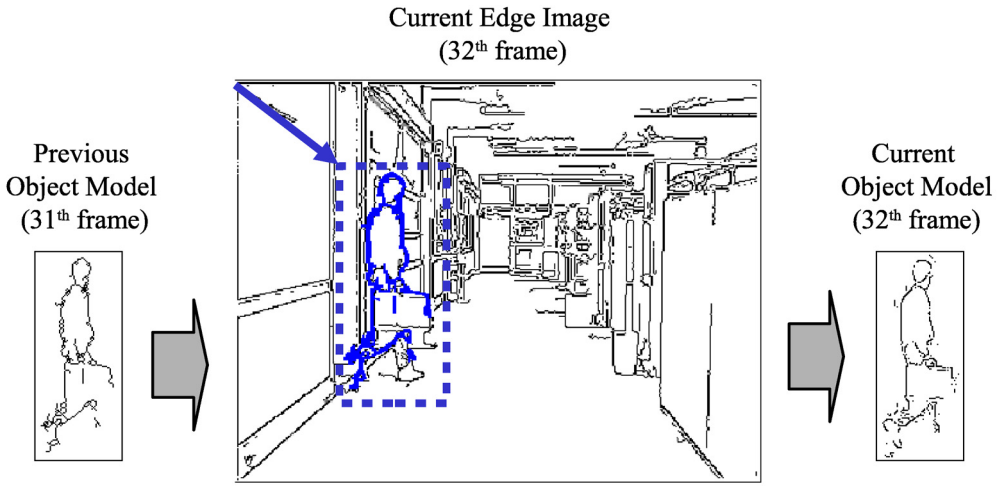


図 6.5: Hausdorff マッチングに基づくオブジェクトモデルの更新 (Stage 3)

6.3.4　エッジ情報に基づくオブジェクト・トラッキング (Stage 3)

Stage 3 は，オブジェクト探索 (object tracking) とモデル更新 (model update) の 2 つのブロックから構成される．

エッジ情報に基づく VOP 自動生成手法では，2 値エッジ情報のみを用いてオブジェクトのトラッキングを行う．これを実現するために，Meier の方法では，Hausdorff マッチングを用いたオブジェクトトラッキング手法 [30] が導入されている．Hausdorff マッチングでは，２つのエッジ画像を比較する際に，それらの類似度を判定する尺度として Hausdorff 距離 [29] を使用する．オブジェクト探索のブロックは，入力された i フレーム目のオブジェクトモデルが次の $(i+1)$ フレーム目のエッジ画像内のどの位置において最もよくマッチするかを，Hausdorff 距離に基づいて見つけ出し，その位置を動きベクトル $MV(i)$ として出力する．この手法の利点として，ターゲットとなるオブジェクトが移動しながら形を変えてもオブジェクトの追跡が可能であること，複雑な背景のシーンにおいてもロバストにオブジェクトを検出できることなどが挙げられる．

オブジェクトの位置が特定できたら，次に第 $(i+1)$ 番目のエッジ画像からオブジェクトに関わるエッジを切り出し，第 $(i+1)$ 番目のオブジェクトモデルを生成する必要がある．この処理をモデル更新（model update）と呼ぶ．モデル更新の手順は以下の通りである．まず，直前のモデル $M(i-1)$ を膨張させ，与えられたベクトル $MV(i)$ に従ってこれをシフトさせる．次に，現在のエッジ画像 $E(i)$ とシフトしたモデル画像 $M(i-1)$ がオーバラップしている部分のエッジを切り出し，現在のオブジェクトモデル $M(i)$ とする．Hausdorff マッチングに基づくモデル更新の例を図 6.5 に示す．

Stage 3 のモデル更新が終わると，オブジェクト抽出 (object extraction) のブロックにおいて，2 値エッジ画像として与えられる現在のオブジェクトモデル $M(i)$ を参照しながら，ビデオ・オブジェクト・プレーン $VOP(i)$ を生成する処理が行われる．ここでのオブジェク

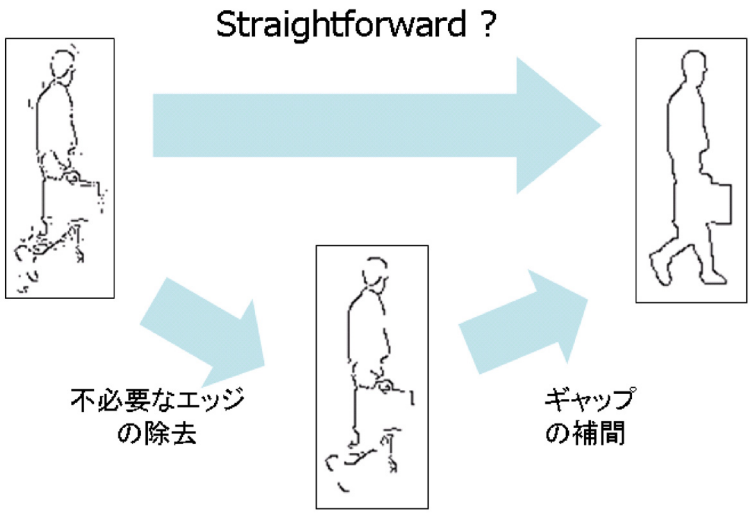


図 6.6: 2 値エッジ画像からのオブジェクト輪郭抽出の問題

ト抽出処理は，本論文の後半において取り扱う主要なテーマとなっているため，処理の詳細については次節で改めて述べることとする．

6.4　エッジ情報に基づくオブジェクト抽出

Meier の手法では，時間的に連続した 2 値エッジ画像に対する Hausdorff マッチングに基づき，オブジェクトトラッキングを実現する．このため，検出されるオブジェクトの画像もまた，図 6.5 に示すような 2 値エッジ画像として表現されており，これをオブジェクトモデルと呼んだ．図 6.2 のオブジェクト抽出ブロックにおいては，このオブジェクトモデルからオブジェクトの領域を推定し，VOP として出力する必要がある．

しかし，2 値オブジェクトモデルからオブジェクト輪郭を推定する問題は一見すると簡単なようだが，実はそれほど単純な処理ではない（図 6.6）．これは，主として以下に挙げる 2 つの理由による．第一に，Canny オペレータによって生成されるエッジは，通常，ところどころ途切れており，閉輪郭線をなしていない．主観的にはオブジェクトの輪郭があるはずと思われる場所であっても，実際にはエッジが抽出されているとは限らず，多くの場合，このようなギャップを補間しなければならない．第二に，オブジェクトモデルが，上とは反対に，過剰なエッジを含んでいる場合がある．特に複雑な背景を持つ画像においてトラッキングを行った場合は，背景に起因するエッジがオブジェクトモデルに取り込まれてしまう傾向がある．

本論文の後半部では，このような不完全な 2 値エッジ画像からオブジェクトの閉輪郭線を抽出する手法について検討している．以下では，2 値エッジ画像からのオブジェクト輪郭抽出に関する従来研究について述べる．また，Meier らが自身の VOP 自動生成手法のために

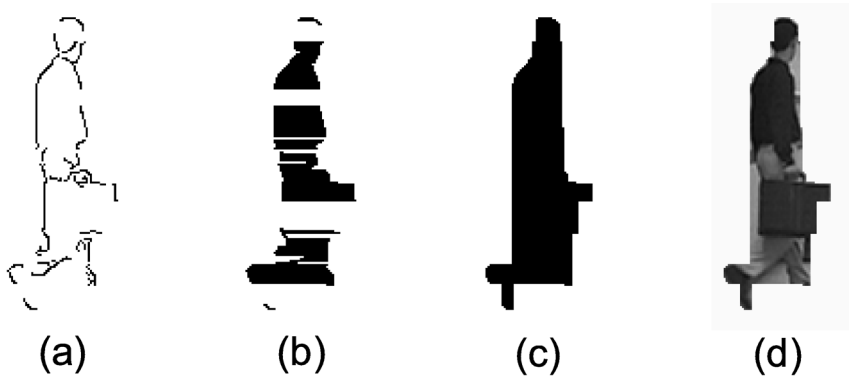


図 6.7: 単純な画素補填を用いたオブジェクト抽出例

考案した方法について説明し，その問題点を指摘する．

6.4.1　関連研究

2 値エッジ画像からの輪郭抽出 (boundary extraction) は，画像理解の分野で研究されてきた基本的な問題である．輪郭追跡法 (contour following)，グラフ探索法 (graph searching)，ハフ変換 (Hough transform) などは，古典的な技法としてよく知られている [36]．しかしながら，輪郭追跡法とグラフ探索法では，エッジ画像に大きなギャップが存在するケースを扱うことができない．ハフ変換はノイズやギャップに対してロバストであるが，オブジェクトの形や大きさに関する事前の知識が必要となる．

近年では，遺伝的アルゴリズムやシミュレーテッド・アニーリングを用いた輪郭抽出法 [1, 51] が提案されている．これらのアルゴリズムは，salt & pepper ノイズや小さなギャップを含むエッジ画像に対して適用可能であるが，エッジベースのオブジェクト・トラッキングにおいてしばしば見られる，背景やオブジェクト内部のテクスチャなどに起因する不要なエッジが多く含まれる画像に対する適用は，特に考慮されていない．

6.4.2　Meier らのアプローチ

ここでは，特にエッジベースの VOP 自動生成手法のために，Meier らが提案している方法について述べる．

Meier らが初めに提案した手法 [48] は，“simple filling-in” と呼ばれる，単純な画素補填に基づく手法である．本手法は，エッジ画像を単純に水平もしくは垂直方向にスキャンし，各スキャンラインにおいて，最初に出会ったエッジ点と最後に出会ったエッジ点の間の画素をすべて’1’ にセットすることにより，オブジェクトの内部領域を求める方法である．ここで，もしスキャンライン上にエッジ点が 1 点しか見つからない場合や 1 点も見つからない場合には，そのスキャンライン上ではどの画素も’1’ にはセットしない．スキャンラインを水平方

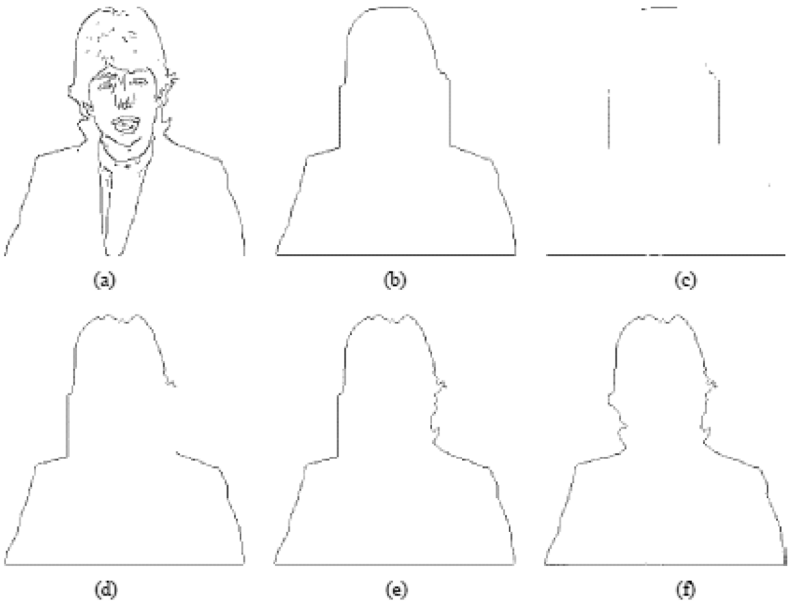


図 6.8: 境界線修正法による修正例 [49]

向に走らせる場合は水平画素補填 (horizontal filling-in)，垂直方向に走らせる場合は垂直画素補填 (vertical filling-in) と呼ぶ．図 6.7(a) の 2 値オブジェクトモデルに対して，水平方向の画素補填を行った例を同図の (b) に示す．単純画素補填 (simple filling-in) による方法は，まず水平方向に画素補填を行い，続いて垂直方向に画素補填を行うことにより実現される．単純画素補填によって得られたオブジェクトのマスク画像を図 6.7(c) に示す．また，このオブジェクトマスクを用いて生成した VOP を図 6.7(d) に示す．

単純画素補填は処理が非常に単純であり，処理速度も速い．この手法により得られるオブジェクトマスクは，与えられたエッジ画像において輪郭線が著しく失われているような場合を除けば，たいていの場合，オブジェクト輪郭を含んだやや大きめの領域となる．しかしながら，図 6.7(d) に見るように，特に 2 度目のスキャンが行われる結果として，オブジェクトの正確な形状が失われるという問題がある．さらに，背景などに起因する過剰なエッジが含まれている場合には，1 度目のスキャンにおいてさえ，正しい形状が失われることがある．このため，オブジェクト形状を正確に切り出し，より正確な形状の VOP を生成したい場合には，さらに高度な手法を採用する必要がある．

その後の論文において，Meier らは境界線修正法 (boundary correction method)[49, 55] と呼ばれる，より高度な手法を提案している．境界線修正法は，単純画素補填で得られたオブジェクト輪郭を最短経路アルゴリズムを用いて修正するという手法である．図 6.8 に境界線修正法によるオブジェクト輪郭の修正例を示す．図 6.8(a) が 2 値エッジ画像として与えられたオブジェクトモデルである．まず単純画素補填によって得られた図 6.8(b) の輪郭線において，オブジェクトモデルと一致しない図 6.8(c) のセグメントは本来のエッジとは無関係

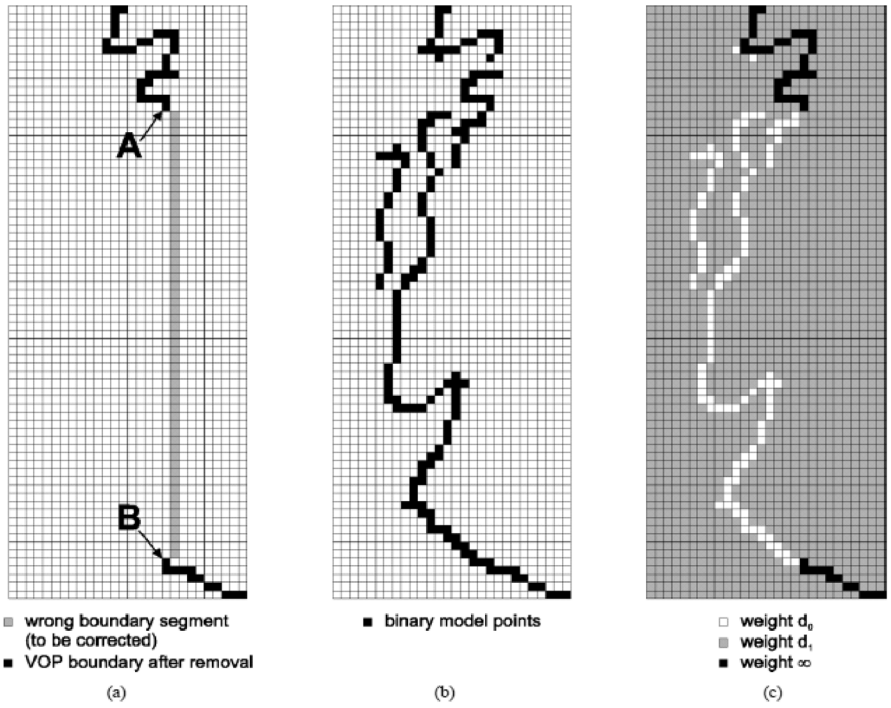


図 6.9: コストの設定 [49]

な誤った輪郭線であるとみなし，これを除去する．(c) の一部のセグメントを除去した結果が図 6.8(d) である．このとき，残った輪郭線の間に生じたギャップを，なるべくオブジェクトモデルのエッジに沿うように補間する処理を行う．その結果，得られた輪郭線が図 6.8(e) である．この処理を (c) のすべてのセグメントに対し繰り返すことにより，図 6.8(f) に示す最終的な輪郭を得る．

境界線修正法では，図 6.8(d) のようなギャップの補間をグラフの最短経路問題とみなし，Dijkstra 法 [17] を用いて解くことにより補間を実現している．この様子を図 6.9 に示す．図 6.9(a) の点 A と点 B を正しく結ぶ輪郭線を求める場合，まず画像の各画素をグラフの頂点とし，各画素の 8-近傍にあるすべての近隣画素に対してグラフの辺が存在するものと考える．各辺に対するコストは，移動先の画素が図 6.9(b) のエッジ点であれば d_0，そうでなければ d_1，既に確定した輪郭上の点であれば ∞ と設定する．すべてのコストを決定した結果が図 6.9(c) である．このとき，エッジ点に沿って進む経路がちょうど最短経路となるように，コストの設定を $d_0 \ll d_1$ として補間を行うというのが基本的なアイディアである．

図 6.8(a) のようにオブジェクトモデルが比較的良い状態にある場合に，境界線修正法はオブジェクト輪郭の形状を改善する上で非常に効果的である．

6.5　おわりに

結びとして，本章で取り上げた研究事例のまとめを述べ，それに基づく問題提起を行う．

本研究事例のまとめ

本章では，VOP 自動生成技術の従来研究の中から，特に Meier らによるエッジベースの VOP 自動生成手法を研究事例として取り上げ，そのアルゴリズムの概要について説明を行った．Meier の方法では，2 値エッジ画像を基に，Hausdorff トラッカーを用いてオブジェクト追跡を行うため，検出されるオブジェクトもまた 2 値エッジ画像として与えられる．このため，VOP 自動生成を実現するためには，このような 2 値画像から閉じたオブジェクト輪郭線を抽出する手法が必要となる．本章では，2 値エッジ画像からのオブジェクト輪郭抽出に関する従来研究として，Meier らがエッジベースの VOP 自動生成手法のために考案した単純画素補填とその境界線修正法について説明した．

本研究事例に基づく問題提起

上述したように，Meier の境界線修正法は，オブジェクトモデルが比較的良い状態にある場合にオブジェクト輪郭の形状を改善する上で非常に効果的である．しかしながら，オブジェクトモデルの状態が悪い場合についてはあまり考慮されておらず，原理的に様々な不都合が生じることが予想される．一例として，オブジェクトモデルにオブジェクト輪郭とは無関係な過剰なエッジ含まれている場合には，必ずしも正しい輪郭が選ばれる保証がない．また逆に，オブジェクトモデルのエッジが欠けている場合には，その部分の補間は単に直線で結ぶのが最短ということになり，精度のよい輪郭が得られない可能性がある．

次章では，以上の問題提起に基づき，Meier らとはまったく異なるアプローチとして，PDE に基づく曲線の時間発展処理を用いたオブジェクト輪郭抽出法の提案を行う．

第7章　時間発展方程式に基づく輪郭線抽出 (I) – 輝度情報を用いる方法

7.1　はじめに

自然画像や画像シーケンスからのオブジェクト輪郭推定は，元来，コンピュータビジョン，画像理解等の分野において研究が進められてきたが，近年では動画像符号化においても重要な役割を担いつつある．

最近の映像符号化標準である MPEG-4 には任意形状オブジェクト符号化の機能が取り入れられ，エンコーダは映像中のオブジェクトと背景を個別に符号化して伝送することができる．MPEG-4 では，各オブジェクトに対応する符号化単位をビデオ・オブジェクト・プレーン (VOP) と呼び，デコーダはすべての VOP を復号した後，これらを一つのシーンに合成することになる．このような機能は画像をオブジェクトという視点から構造化しようという試みの一端であり，オブジェクト単位の映像合成やスプライト符号化などの効率的な映像符号化などを可能にするという利点を持っている．しかしながら，このような VOP を自動的に生成するためには，動物体の検出とその安定したトラッキング，及びオブジェクトの正確な抽出などを実現する高度な画像解析技術が必要であり，このことがオブジェクトベースの符号化を実現する上での障害となっている．

本論文では，そのような画像解析技術の一環として，特に 2 値エッジ画像からのオブジェクト輪郭抽出の問題を扱う．この研究は，Meier ら [48] の提案するエッジベースの VOP 自動生成手法における，オブジェクトモデルと呼ばれる 2 値エッジ画像からのオブジェクト輪郭抽出の問題をその動機付けとして行われたものである．第 6 章で既に述べたように，Meier の VOP 自動生成アルゴリズムにおいては，時間的に連続した 2 値エッジ画像の間で Hausdorff マッチング [29, 30] を行うことにより，オブジェクトのトラッキングを実現している．そのため，各フレームで検出されるオブジェクトもまた 2 値エッジ画像として表現される．一般に，エッジ検出フィルタの出力を閾値処理して得られる 2 値エッジ画像においては，弱いエッジが欠損して完全なオブジェクト輪郭を得ることができない．このため，正確なオブジェクト領域を抽出するには，不完全なエッジ情報を基に，オブジェクト領域を囲む完全な閉輪郭線を何とかして抽出しなければならない．本章では，このような問題を解決するため，曲線の時間発展方程式に基づき，2 値オブジェクトモデルからオブジェクトの輪郭抽出を行う方法を提案する．提案法では，Casseles ら [12] による，測地線に基づく動的輪郭法 (GAC) を利用してオブジェクト輪郭の抽出を行っている．

以下，本章の構成は次の通りである．まず 7.2 において提案手法の説明を行う．ここでは，

はじめに要素技術である GAC の概要をまとめた上で，提案法においてこれをどのように利用するかについての概要を述べる．アルゴリズムの詳細は擬似コードとして提示する．次に 7.3 において，シミュレーションによる検証結果について述べる．特に，Meier らによる従来手法との比較についても言及する．最後に 7.4 では，本章のまとめを述べる．

7.2　提案手法

7.2.1　GAC に関する設定

本章で提案するオブジェクト輪郭抽出法では，GAC に基づく曲線の時間発展処理を利用する．以下に，提案法で使用する GAC の設定について述べる．

まず，画像平面上の曲線 $\mathbf{c}(q)$ の長さを，

$$L(\mathbf{c}) = \int_a^b g(|\nabla I(\mathbf{c}(q))|)|\mathbf{c}'(q)|dq \tag{7.1}$$

と定義する．ここで，g はエッジ検出関数である．一般に，GAC では，適当に与えた画像平面上の初期輪郭線を，長さ $L(\mathbf{c})$ を極小とする測地線に向かって時間発展させることにより，オブジェクトの輪郭を検出する．

本章の提案手法では，GAC をレベルセット法を用いて実装する．まず，輪郭曲線 $\mathbf{c}(q)$ をゼロ交差とする曲面 $u(x, y)$ として，符号付き距離関数

$$u(\mathbf{x}) = \begin{cases} +d(\mathbf{x}, c) & \text{if } \mathbf{x} \text{ is outside } \mathbf{c} \\ -d(\mathbf{x}, c) & \text{if } \mathbf{x} \text{ is inside } \mathbf{c}, \end{cases} \tag{7.2}$$

を採用する．ここで，$d(\mathbf{x}, \mathbf{c})$ は，画像平面上の点 $\mathbf{x} = (x, y)$ から曲線 $\mathbf{c}(q)$ への距離とする．このとき，GAC に基づく曲線の運動は，曲面 u に関する次の時間発展方程式によって記述される．

$$\frac{\partial u}{\partial t} = g\,|\nabla u|\,\mathrm{div}\Big(\frac{\nabla u}{|\nabla u|}\Big) - \alpha\,g\,|\nabla u| + \langle \nabla g, \nabla u \rangle. \tag{7.3}$$

ここで，α は正の定数であり，曲線の時間発展を促進するためのパラメータである．α を大きくとると，曲線の時間発展のスピードは速くなるが，曲線の滑らかさに関する制約がゆるくなることが知られている．

エッジ検出関数 g としては，

$$g(x) = (1 + (x/K)^2)^{-1}. \tag{7.4}$$

を使用する．ここで，様々な入力画像からオブジェクトを精度よく抽出するためには，画像の勾配値の分布に基づいて，エッジ検出関数 g の定数 K を適切に決める必要がある．Black-Sapiro らは，ロバスト統計の知見に基づき，K の推定値を，対象とする画像の勾配ベクトル場 ∇I の MAD(Median Absolute Deviation) を用いて次の式で与えている [9, 75].

$$\begin{aligned} K &= 1.4826\ \mathrm{MAD}(\nabla I) \\ &= 1.4826\ \mathrm{median}_I(|\nabla I - \mathrm{median}_I(|\nabla I|)|). \end{aligned} \tag{7.5}$$

以下，本論文においても，式 (7.5) に基づいて定数 K を決定している．

7.2.2　アルゴリズムの概要

2 値オブジェクトモデルからオブジェクト輪郭を抽出する手順は以下の 4 ステップから成る：

(i) オブジェクトの内部に含まれる領域を生成し，その境界線を初期輪郭として設定する．この初期輪郭は複数の閉輪郭線から構成されていてもよい．

(ii) 初期輪郭として設定した複数の閉輪郭線を内部からオブジェクトの輪郭に向かって外方向に成長させる．ほとんどの輪郭はこの時点で融合される．

(iii) 一定速度により輪郭線を縮退して，融合されていない小領域の輪郭線を除去する．続いて，一定速度により輪郭線を膨張して，ステップ (ii) で融合しきれなかった領域の融合を促進し，真のオブジェクト境界よりもやや大きい単一輪郭線に成長させる．

(iv) 最後に，単一の閉輪郭線を外部からオブジェクトの輪郭に向かって内方向に成長させる．収束した時点での輪郭線をオブジェクト輪郭とする．

ここで述べた，オブジェクトの内部領域を外側に向かって成長させる時間発展アプローチは，特に，オブジェクトが強いエッジを多く含む複雑な背景をバックとしている場合に効果的な手法であると考えられる．

ステップ (i) において，初期輪郭を生成する手順を図 7.1 に示す．オブジェクトの内部領域に含まれる複数の輪郭線を得るために，まず図 7.1(a) のオブジェクトモデルに対して水平及び垂直方向に画素補填したマスク画像 (b),(c) を生成する．次に，これらの画像について論理的な AND を取ることにより，図 7.1(e) のような，オブジェクトの内部領域に含まれるマスク画像が得られる．このマスク画像は，単純画素補填によって得られるマスク画像 (d) に比べて，オブジェクトの形状をよく反映していることがわかる．そこで提案手法においては，マスク画像 (e) の境界線を初期輪郭線として採用する．一般に，マスク画像 (e) は複数の連結成分から構成されるため，その境界線もまた複数の閉輪郭線で構成されることとなる．マスク画像 (e) の境界として得られた，初期輪郭線を図 7.2(a) に示す．レベルセット法を用いた時間発展方程式では，図 7.2 の輪郭線 (a) の代わりに，これをゼロ交差とする超曲面を時間発展させる．符号付き距離関数に基づいて生成した超曲面を図 7.2(b) に示す．

ステップ (ii) では，前節において言及した，外側へ向かう前進 GAC(forward GAC) が使用される．前進法は，ステップ (i) で生成した初期輪郭を，オブジェクト領域の内側から膨張させながら単一オブジェクトの輪郭に統合，収束させる．ただし，オブジェクトには，その内部にエッジやテクスチャが存在するのが普通であり，そのような場合，輪郭線の成長が単一の領域に統合する前に止まってしまうことがしばしば起こる．また，オブジェクトモデルに含まれる不要な短エッジに起因する小領域が残っている場合がある．

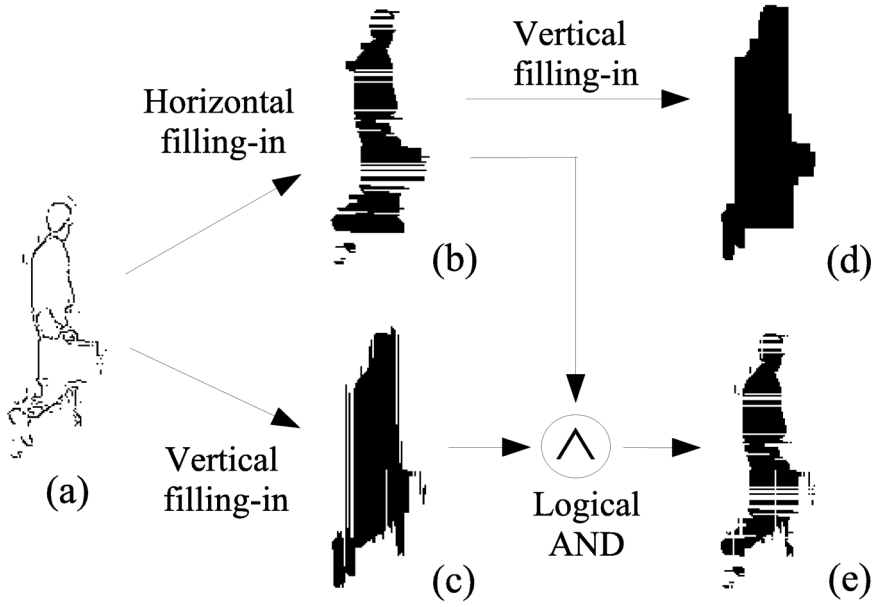


図 7.1: 単純な画素補填を用いたオブジェクト抽出例

ステップ (iii) では，これらの問題に対処するため，一定速度による輪郭線の縮退・膨張処理を導入する．これは，本質的には，モルフォロジーにおける縮退・膨張処理と同様の効果を持つ．

$$\frac{\partial u}{\partial t} = -c\,|\nabla u|, \quad \text{where } c = \pm 1. \tag{7.6}$$

ここで，縮退処理を実現するには $c = -1$，膨張処理を実現するには $c = 1$ とセットする．このとき，ステップ (iii) の処理としては，まず N 画素分の縮退処理を行い，不要な短エッジに起因する小領域の輪郭を除去する．次に $2N$ 画素分の膨張処理を行い，オブジェクト内部のエッジやテクスチャの影響で単一領域に統合できずに止まっている輪郭を成長させ，その統合を促す．$2N$ 画素分膨張させることにより，この時点で，実際のオブジェクト輪郭よりもやや大きな領域に成長していることが期待できる．

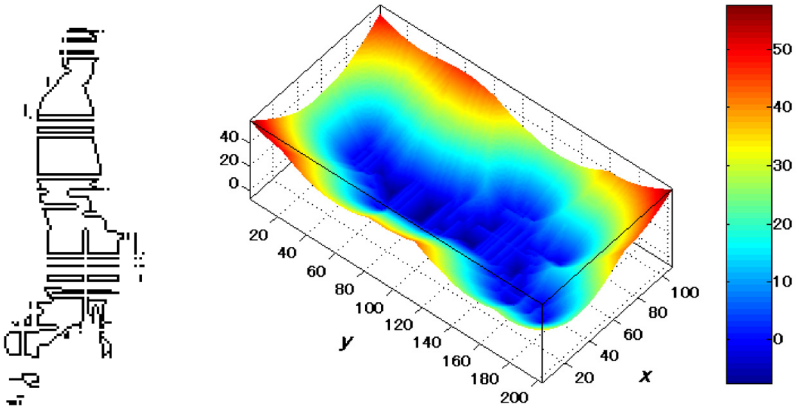
最後に，ステップ (iv) では，内側へ向かう後退 GAC(backward GAC) を使用し，輪郭線をオブジェクトに対して外側から収束させる．その目的は，輪郭線が正確なオブジェクトの形となるように微調整を行うことである．

7.2.3　擬似コード

提案するオブジェクト抽出アルゴリズムの擬似コードを以下に示す.

[オブジェクト抽出アルゴリズムの擬似コード]

M ← Object Model (binary edge image)



(a)　(b)

図 7.2: 初期輪郭線 (a) とこれに対応する符号付き距離関数 (b)

$I_m \leftarrow$ Object Image corresponding to M　(gray-scale image)

Setting constant parameters:
　$\alpha 1$: constant α in Eq.(7.3) for forward GAC
　$\alpha 2$: constant α in Eq.(7.3) for backward GAC
　N:　number of pixels to move in constant-speed evolution
　Δt: temporal interval in finite difference method

OBJECT_EXTRACT(M, I_m)
{
　M_{init}: Initial Object Model　(binary edge image)
　BM: Binary Mask
　SDF: Signed Distance Function　(2D array)
　VOP: Video Object Plane　(gray-scale Image)

　BM $\leftarrow$ HORIZONTAL_FILLING_IN(M) & VERTICAL_FILLING_IN(M)
　　▷ *Intersection area obtained by horizontal and vertical filling-in methods.*
　$M_{init} \leftarrow$ BOUNDARY(BM)
　　▷ *BM boundaries are used as initial contours.*
　SDF $\leftarrow$ SIGNED_DISTANCE_FUNCTION(M_{init})
　　▷ *Signed distance function is calculated with reference to M.*
　GEODESIC_ACTIVE_CONTOUR(SDF, I_m, 'f')
　　▷ *Evolve SDF in forward direction, detecting object contours from inside.*
　CONSTANT_SPEED_EVOLUTION(SDF, 'b', N1)

```
        ▷ Eliminate small contours with N1-pixel erosion.
    CONSTANT_SPEED_EVOLUTION(SDF, 'f', N2)
        ▷ Merge multiple contours with N2-pixel dilation.
    GEODESIC_ACTIVE_CONTOUR(SDF, I_m, 'b')
        ▷ Evolve SDF in backward direction, fitting object contours from outside.
    BM ← AREA(SDF ≤ 0)
        ▷ Binary mask of the obtained region.
    VOP ← IMAGE_MASK(I_m, BM)
        ▷ Extract object with reference to BM.
    return VOP
}

GEODESIC_ACTIVE_CONTOURS(U, I_m, flag)
{
  if flag = 'f'
    then  c ← +1, α ← α_1      ▷ forward scheme
    else  c ← −1, α ← α_2      ▷ backward scheme

  while STOPPING-CONDITION(flag) is false
    do for all (i, j)
      EVALUATE-G(I_m, i, j)
          ▷ Evaluate edge-detector g to obtain g_{i,j}, (g_x)_{i,j} and (g_y)_{i,j}.

      EVALUATE-CURVATURE(U, i, j)
          ▷ Evaluate U to obtain K_{i,j}.

      D_{i,j} ← g_{i,j} K_{i,j} sqrt((δ_x U_{i,j})^2 + (δ_y U_{i,j})^2)
          ▷ Evaluate the first term in Eq.(7.3). (paraboloc term)

      D_{i,j} ← D_{i,j} − α [ max(−c g_{i,j}, 0) ∇^+ U_{i,j} + min(−c g_{i,j}, 0) ∇^− U_{i,j} ]
          ▷ Evaluate the second term in Eq.(7.3). (hyperbolic term)

      D_{i,j} ← D_{i,j} + max((g_x)_{i,j}, 0) δ_x^− U_{i,j} + min((g_x)_{i,j}, 0) δ_x^+ U_{i,j}
                        + max((g_y)_{i,j}, 0) δ_y^− U_{i,j} + min((g_y)_{i,j}, 0) δ_y^+ U_{i,j}
          ▷ Evaluate the third terms in Eq.(7.3). (hyperbolic term)

      U_{i,j} ← U_{i,j} + D_{i,j} Δt
          ▷ SDF evolution for 1 step.
```

```
  return U
}

  CONSTANT_SPEED_EVOLUTION(U, flag, pixels)
{
  if flag = 'f'
    then c ← +1          ▷ forward scheme
    else c ← -1          ▷ backward scheme
  N ← pixels/Δt

  for k ← 0 to N-1
    do for all (i, j)
       D_{i,j} ← [max(c,0)∇⁺ + min(c,0)∇⁻] U_{i,j}
            ▷ Evaluate the right-hand side of Eq.(7.6).
       U_{i,j} ← U_{i,j} − D_{i,j} Δt
            ▷ SDF evolution for 1 step.
  return U
}
```

STOPPING_CONDITION は，GAC による曲線の時間発展をいつ止めるかを判定する関数である．基本的には，十分大きなステップ数の間に輪郭線で囲まれる領域の面積の変化がほとんどなかった場合に，GAC を停止している．今回の実装では，GAC の反復処理を 20 回繰り返すごとに，レベル曲面 u を符号付き距離関数によって再初期化しているため，その際に輪郭線で囲まれた領域の面積の変化をチェックし，停止するかどうかの判定をしている．なお，前進 GAC の場合は，領域の統合を促進する目的で α_2 を大きめに設定することが多い．このため，GAC の輪郭線がオブジェクト輪郭で停止せず，弱いエッジ部分から漏れ出してしまうことがある．このような場合にも，適当に処理を停止できるように，輪郭線で囲まれた領域の面積の最大値を決めておき，これを超えた場合にも，GAC の処理を停止するようにした．

EVALUATE_CURVATURE は，レベル曲面 u の等高線の曲率 $K_{i,j}$ を計算する関数である．曲率の計算には，次の公式を使用した [4]．

$$K = div\left(\frac{\nabla u}{|\nabla u|}\right) = \frac{u_{xx}u_y^2 + u_{yy}u_x^2 - 2u_x u_y u_{xy}}{(u_x^2 + u_y^2)^{2/3}}.$$

なお，上述した擬似コードで使用されている $\delta_* U_{i,j}$, $\delta_*^+ U_{i,j}$, $\delta_*^- U_{i,j}$ $(* = x, y)$ は，偏微分方程式の有限差分解法においてよく使用される，指定された軸方向の中心差分，前進差分，後退差分をそれぞれ表している．また，$\nabla^+ U_{i,j}$, $\nabla^- U_{i,j}$ は，双曲型方程式の差分解法にお

いてよく使用される，前進方向および後退方向の風上差分 (upwind scheme)

$$\nabla^+ U_{i,j} = \Big[max(\delta_x^- U_{i,j},0)^2 + min(\delta_x^+ U_{i,j},0)^2) + max(\delta_y^- U_{i,j},0)^2 + min(\delta_y^+ U_{i,j},0)^2)\Big]^{\frac{1}{2}}$$

$$\nabla^- U_{i,j} = \Big[max(\delta_x^+ U_{i,j},0)^2 + min(\delta_x^- U_{i,j},0)^2) + max(\delta_y^+ U_{i,j},0)^2 + min(\delta_y^- U_{i,j},0)^2)\Big]^{\frac{1}{2}}$$

をそれぞれ表している [4], [61].

7.3　シミュレーションによる検証

前章で述べた Hausdorff トラッキングによって生成されたエッジベースのオブジェクトモデルを用いて，オブジェクト抽出のシミュレーションを行った．ここでは，MPEG-4 の標準画像シーケンスである “Hall Monitor” を使用している．また，本節で行うすべての実験において，上述の擬似コードで設定が必要な固定パラメータ α_1, α_2, $N1$, $N2$, ΔT を

$$\alpha_1 = 1.5,\ \alpha_2 = 1.0,\ N = 2,\ \Delta T\ = 0.05.$$

と設定した．

7.3.1　オブジェクトの抽出過程

まず初めに，提案法によるオブジェクト抽出の方針がうまく機能しているかどうかを確認するため，輪郭曲線の時間発展の様子を観察した．図 7.2(a) は，“Hall Monitor” の第 32 フレームから抽出したオブジェクトモデルを基に生成された初期輪郭 M_{init} である．この時点で輪郭のほとんどがオブジェクト領域内部に含まれていることが確認できる．図 7.2(b) は，関数SIGNED_DISTANCE_FUNCTION によって M_{init} から得られた符号付き距離関数である．レベルセット法では，レベル曲面 (b) を GAC の方程式に基づいて時間発展させるため，輪郭線 (a) のように複雑な閉輪郭の時間変化を扱うことができる．

図 7.3 は，時間発展するレベル曲面のゼロレベルセットを画面上に白線で表示したものである．初期輪郭として与えられた複数の複雑な輪郭が次第に統合し，オブジェクト輪郭が抽出されていく様子がわかる．各画像の下に表示されたパラメータ *iter* は，そのオブジェクト輪郭を得るまでに要した，GAC の差分解法の反復回数を表す．

図 7.3 における輪郭曲線の時間発展の様子は次のように説明できる．

第 1 フェーズは $0 < \text{iter} \leq 300$ である．ここでは，前進 GAC により複数の輪郭線がオブジェクトの内側から外方向に膨張しながら次第に統合されていく様子が示されている．反復回数が 300 回の時点では，初期輪郭として与えられた複数個の輪郭のほとんどが単一輪郭に統合され，求めるオブジェクトに近い形状へと成長していることがわかる．しかしながら，この時点では，オブジェクト内部のエッジやテクスチャの影響で，いくつかの領域がまだ統合されずに残っていることが確認できる．さらに，オブジェクトモデルに混入していた背景

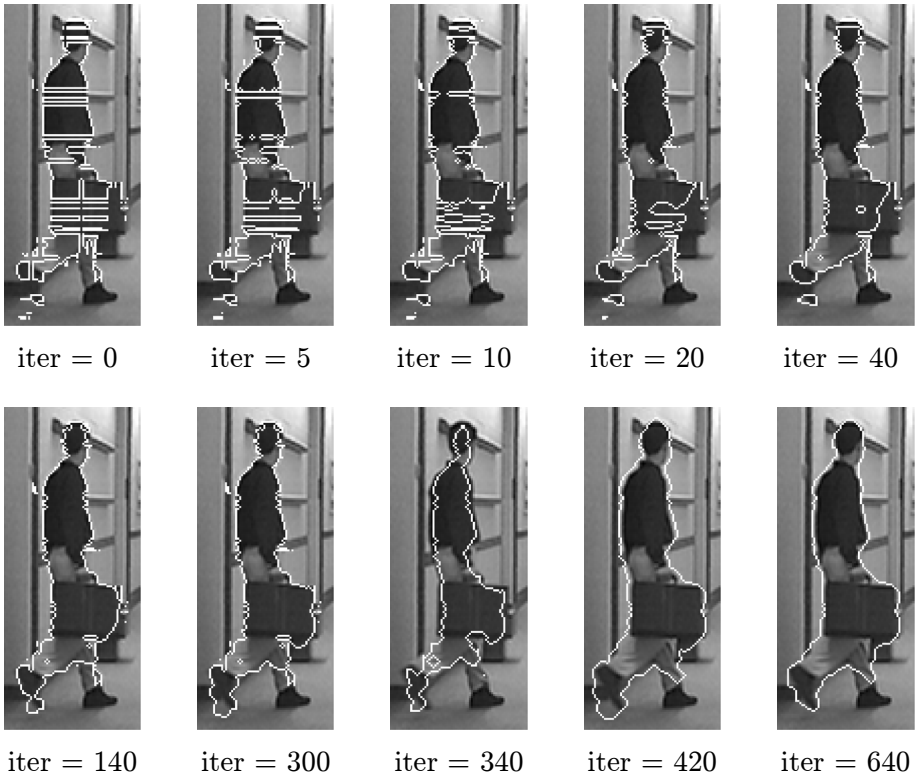


図 7.3: 提案法によるオブジェクト輪郭の抽出過程

エッジなどのノイズに起因する不要な小領域が，オブジェクト輪郭の周りに残っていることも確認できる．

第 2 フェーズは $300 < iter \leq 420$ である．ここでは，式 (7.6) に基づく一定速度の時間発展処理が行われている．$300 < iter \leq 340$ に対しては $c = -1$ と設定され，輪郭線は一定速度で N 画素分だけ縮退する．このとき，輪郭線の縮退に伴い，不要な小領域が除去されていることがわかる．続く $340 < iter \leq 420$ に対しては $c = 1$ と設定され，輪郭線は逆に一定速度で $2N$ 画素分だけ膨張する．これに刺激を受け，オブジェクト内部のエッジやテクスチャの影響で GAC では統合されなかったいくつかの領域が統合され，オブジェクトを取り囲む単一領域へと成長している．縮退の N 画素に対し，膨張を倍の $2N$ 画素とすることによって，第 2 フェーズの終了時には，オブジェクト領域よりもおよそ N 画素分だけ大きな輪郭線が得られることが期待される．図 7.3 では，第 2 フェーズの最後にあたる $iter = 420$ でこのことが確認できる．なお，この例では $N = 2$ に設定されている．

最後に，$420 < \text{iter} \leq 640$ にあたる第 3 フェーズでは，後退 GAC によりオブジェクトの外側から内方向に縮退しながら，輪郭線をオブジェクトへと近づける．ここでの目的は，第 2 フェーズで得られたややルーズなオブジェクト輪郭をより正確な形状へと整形することである．このとき，なるべく正確な形状を得るには，GAC の時間発展の速度を遅めに設定することが望ましい．このため，本実験においては，第 3 フェーズの後退 GAC に対しては

$\alpha_2 = 1.0$ と設定している．これに対し，第 1 フェーズでは，輪郭線の時間発展速度を速めるとともに複数領域の統合を促進させるため，$\alpha_1 = 1.5$ と設定している．

図 7.3 の $iter = 640$ の画像に描かれた白線が，提案手法により最終的に得られたオブジェクト輪郭である．左足の一部が検出されていないが，その他についてはおおよそ良好な結果であるといえるだろう．

7.3.2　従来手法との比較

我々の提案法と Meier らによる 2 つの従来法によって得られたオブジェクト輪郭を主観的に比較する実験を行った．ここで 2 つの従来法とは，一つは単純画素補填による方法 [48] であり，もう一つは Dijkstra の最短経路アルゴリズムに基づく境界線修正法 [49] である．本実験では，標準画像シーケンス “Hall Monitor” の第 31-80 フレームから生成した 2 値オブジェクトモデルを使用して，オブジェクト抽出を行った．

図 7.4 は，第 32, 50, 70 フレームのオブジェクトモデルから 3 種類の方法で抽出したオブジェクト輪郭を，グレイスケール画像上に白線で表示した結果を示している．図の上段，中断，下段には，それぞれ第 32 フレーム，第 50 フレーム，第 70 フレームの結果が提示されている．図の第 1 列にはオブジェクト抽出に使用したオブジェクトモデルを示している．また，第 2 列は単純画素補填によって得られたオブジェクト輪郭を示している．単純画素補填から得られるオブジェクト領域は，オブジェクトのほとんどすべての領域をカバーするが，得られたオブジェクト輪郭は所どころ角張っていて，決して自然な輪郭線とはいえない．

図 7.4 の第 3 列は，Meier らの境界線修正法によって得られたオブジェクト輪郭を示している．境界線修正法では，先の単純画素補填によって得られるオブジェクト輪郭を次のように修正する．まず，単純画素補填によって得られるオブジェクト輪郭のうち，オブジェクトモデルのエッジと一致するセグメントのみを残し，残りの部分を取り除く．するとオブジェクト輪郭にギャップが生じるため，これを Dijkstra の最短経路アルゴリズムを用いて正しい輪郭線となるようにリンクする．こうして得られた輪郭線（図左）には，背景に起因するノイズの影響から，しばしば不要な枝が生じる場合がある．そこで，これを除去して得られる閉輪郭線（図右）を本手法の最終結果とした．境界線修正法では，オブジェクト輪郭のギャップをつなぐ際に，エッジ点があれば必ずその点を通過するようにコスト関数を設定してあるため，成功した場合には，主観的に見て非常に良好な輪郭線を得ることができる．そのような例として，図 7.4(c) が挙げられる．しかしながら，オブジェクトモデルの状態が悪く，オブジェクト輪郭に相当するエッジがうまく検出されていない場合や，本来オブジェクトには属さない背景などのエッジが “ノイズ” として混入している場合には，近い位置にある誤ったエッジをつないでしまう結果として，図 7.4(g),(k) のように，不自然な凹凸を含む輪郭線が得られることになる．さらに，図 7.4(e) のように，本アルゴリズムで参照すべきオブジェクトモデルに大きなギャップが存在する場合には，最短距離として直線による補間が選択されてしまい，やはり主観的に不自然な輪郭となってしまう．以上の考察から，Meier の方法は，オブジェクトモデルのエッジがあまりよい状態にない場合（例えばノイズの混入やエッ

ジ画像のギャップなどが生じた場合）には，まだ不十分であると考えられる．

図 7.4 の第 4 列は，提案法によるオブジェクト輪郭抽出の結果である．時間発展処理をスタートする前の初期輪郭を図左に，収束後の採取結果を図右に示した．第 32 フレームのように与えられたオブジェクトモデルが比較的良い状態の場合には，最終結果として得られた輪郭線は，Meier の境界線修正法の方が提案法よりも若干正確であるように見える．しかしながら，第 50 フレームや第 70 フレームのように，オブジェクトモデルの状態が悪い場合には，提案法で得られた結果のほうが，より自然であることがわかる．特に第 50 フレームの例をみると，提案法ではオブジェクト頭部の抽出に成功しているが，Meier の方法ではこれに失敗している．このことから，オブジェクトモデルに混入するノイズやギャップに対しては，提案法の方がよりロバストであることがわかる．一方，第 70 フレームの例では，Meier の方法と提案法の両方において，オブジェクト頭部の抽出に失敗している．これは，単純画素補間（図 7.4(j)）の時点で，既に頭部がカバーされていないことが原因である．このことは，安定した VOP 自動生成を実現するためには，少なくとも単純画素補間の結果がオブジェクト領域全体をカバーできる程度の正確なオブジェクトモデルが得られるように，Hausdorff トラッカーの調整をする必要があることを示唆している．

7.4　おわりに

本章では，エッジベースの VOP 自動生成手法の要素技術の一つとして，与えられた 2 値オブジェクトモデルからオブジェクトの閉輪郭線を抽出する方法に関する提案を行った．

7.2 では，提案法のアルゴリズムについての説明を行った．エッジ情報を基にオブジェクトトラッキングを行う方法については Meier のフレームワークを踏襲し，得られたオブジェクトモデルを基にオブジェクト輪郭を抽出する問題を独自に解決するため，GAC に基づき輪郭線を時間発展させるアプローチについて提案を行った．提案法において時間発展を行う複数の初期輪郭線は，オブジェクトモデルを水平及び垂直方向に画素補填して得られた領域の論理積として定義されるマスク画像の輪郭線として与えられる．こうして与えられた初期輪郭線は，通常，オブジェクトの内部領域に含まれるため，提案法では，これを内側から外に向かって膨張させながら単一領域に統合してオブジェクト領域を得る inside-to-outside アプローチを採用した．

7.3 では，標準画像シーケンス “Hall Monitor” を用いてオブジェクト抽出実験を行い，提案法の有効性を確認した．また，提案法によって得られたオブジェクト輪郭と，Meier らが提案した単純画素補填及び境界線修正法によって得られたオブジェクト輪郭を主観的に比較したところ，オブジェクトモデルに混入するノイズやギャップに対して提案法の方がロバストであり，より自然なオブジェクト輪郭を与えることがわかった．

今後の課題として 2 つの項目が挙げられる．第一に，抽出したオブジェクト輪郭がまだ不十分で，改善の余地がある．例えば，図 7.4(d) では左足の一部が，(l) では頭部がそれぞれ欠けてしまっている．オブジェクトモデルの状態がある程度悪くても，正確なオブジェクト形状が得られるようにアルゴリズムを改善するには，画像の勾配だけでなく，動き，色テク

スチャなどの情報を統合して利用する必要がある．第二に，GAC の停止条件のロバスト性に関する検証がまだ不十分である．現在の提案法では，第 1 フェーズである前進 GAC と第 3 フェーズである後退 GAC の停止のタイミングによって，得られるオブジェクト輪郭が大幅に変わってしまうことがある．このため，より多くのシミュレーションを通して，その停止条件が適切であるかをより詳細に検討する必要がある．

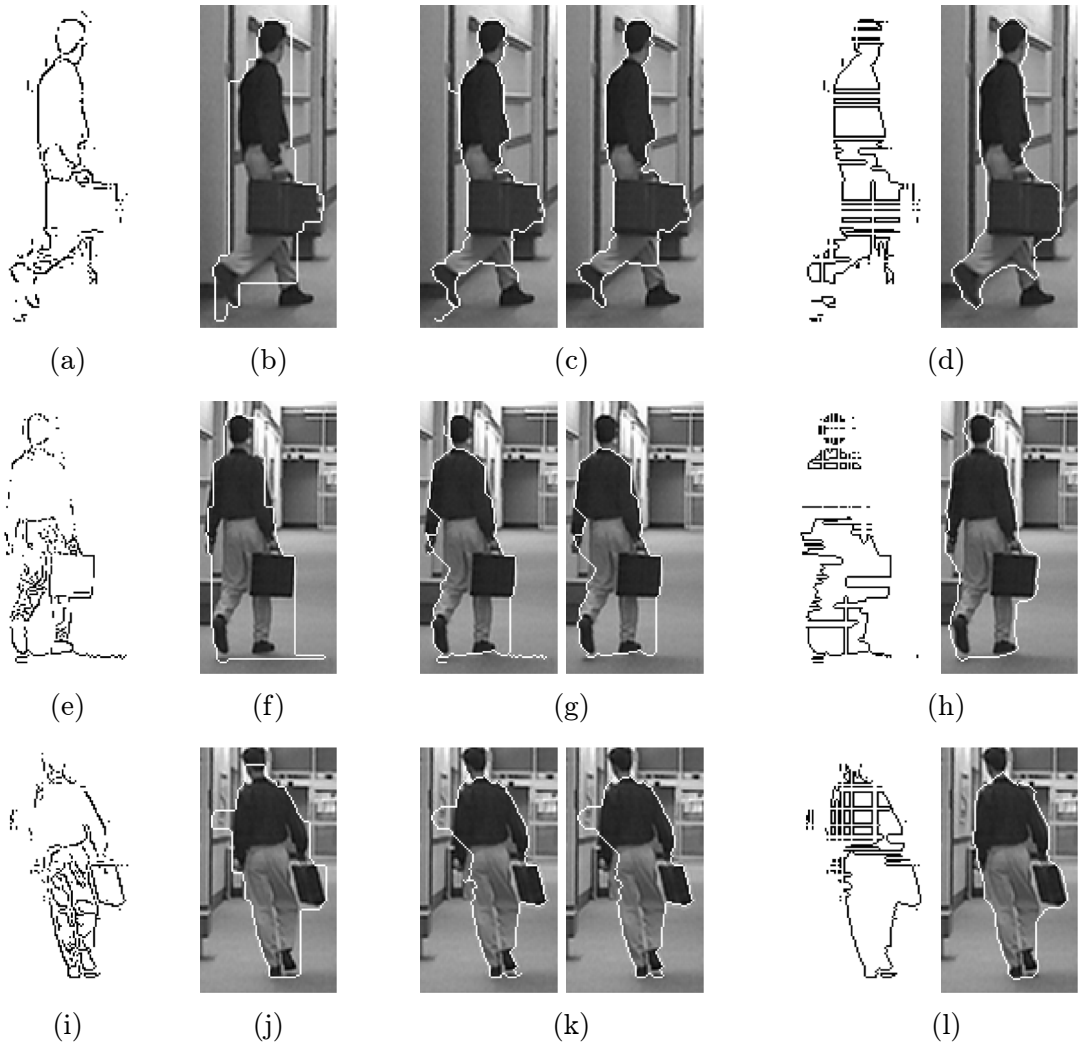


図 7.4: オブジェクト輪郭線の比較. (a),(e),(i) オブジェクトモデル from “Hall Monitor” (32^{nd}, 50^{th}, 70^{th} frames). (b),(f),(j) 単純画素補填による輪郭抽出結果. (c),(g),(k) 境界線修正法による輪郭抽出結果. (図左: オリジナル出力, 図右: 枝除去後の最終結果). (d),(h),(l) GAC に基づく提案法による輪郭抽出結果.

第8章　時間発展方程式に基づく輪郭線抽出 (II) – 2値エッジ情報を用いる方法

8.1　はじめに

2値エッジ画像からの輪郭抽出 (boundary extraction) は，コンピュータビジョンや画像理解の分野で研究されてきた基本的な問題である．これまで，輪郭追跡法 (contour following)，グラフ探索法 (graph searching)，ハフ変換 (Hough transform) などの古典的な技法 [36] から，遺伝的アルゴリズムやシミュレーテッド・アニーリングを用いた輪郭抽出法 [1, 51] に至るまで，様々な手法が提案されてきた．

近年では，動画像符号化や2次元CGアニメーションなどの従来とは異なる分野においても，このように高度な画像処理が求められるようになりつつある [48, 54]．MPEG-4における任意形状オブジェクト符号化を実現するには，入力された自然画像シーケンスを個々のオブジェクトと背景に分離し，ビデオ・オブジェクト・プレーン (VOP) を生成する必要がある．しかしながら，このようなVOPを自動的に生成するためには，動物体の検出とその安定したトラッキング，及びオブジェクトの正確な抽出などを実現する高度な画像解析技術が必要であり，このことがオブジェクトベース符号化を実現する上での障害となっている．

本章では，第7章に引き続き，Meierらによるエッジ情報に基づくVOP自動生成のフレームワーク [48] を想定した，2値オブジェクトモデルからのオブジェクト輪郭抽出の問題を扱う．第7章では，曲線の時間発展方程式を利用して，2値オブジェクトモデルからオブジェクトの輪郭抽出を行う方法について検討した [86]．ここで提案したアルゴリズムは，Casellesらによって提案されたGAC(Geodesic Active Contours) に基づくものであったため，結局のところ，グレイスケール画像の輝度情報を利用して，輪郭曲線の時間発展を停止していたことになる．これに対し，Meierの方法では，2値エッジ画像のみを用いてオブジェクト・トラッキングを行っており，本来はオブジェクト輪郭の抽出についても，グレースケール画像の情報を使わず，2値エッジ画像のみで実現できることが望ましい．

以上を踏まえ，本章の提案法としては，与えられた2値オブジェクトモデルから，その他の補助的な情報を一切使用することなく，オブジェクト輪郭を抽出する方法を検討する．前章と同様，曲線の時間発展に基づくアプローチを用いてこれを実現するためには，輝度勾配のかわりに，曲線の時間発展を止めるための場を2値エッジ画像から生成する必要がある．そこで，本章の提案法では，入力したオブジェクトモデルから各画素 (x, y) 上に輪郭線が存在する可能性の場を計算し，これに基づいて輪郭線を時間発展させる修正GACを導入する [88]．

曲線の存在可能性場という考え方は，もともと中村-齋藤ら [54] によって，2 次元セルアニメーションの塗りつぶし領域推定のために提案された．本章では，中村らの手法をベースに，これを若干修正する形で，曲線の存在可能性場の計算を行っている．本章で提案する手法は，この曲線の存在可能性場を勾配ベクトル場の代わりに用いることにより，GAC と同様の曲線の時間発展を実現するものである．なお，近年では，曲線の存在可能性場を利用した動的輪郭法に関する同様の趣旨の研究として，岡本-齋藤ら [56] による研究報告がある．

以下，本章の構成は以下の通りである．8.2 では，2 値のオブジェクトモデルから輪郭線の存在可能性を表す場を生成するアルゴリズムについて述べる．8.3 では，求めた存在可能性場を用いて GAC を修正する方法について述べ，8.4 においてそのシミュレーション結果について報告する．続いて 8.5 では，輪郭線の存在可能性場を計算する上で必要となる接ベクトルの導出に関する改善方法について述べ，8.6 でその効果をシミュレーションにより検証する．最後に 8.7 で本章のまとめを述べる．

8.2　輪郭線の存在可能性を表す場の生成

ここでは，2 値エッジ画像として与えられるオブジェクトモデルに対して，輪郭線の存在可能性場を導出する方法について述べる．

8.2.1　アルゴリズムの概要

輪郭線の存在可能性場を導出する手順は次の通りである．

まず前処理として，すべてのエッジ点における接ベクトルを算出しておく．ここでは，対象となるエッジ点を中心とする小さな窓を取り，そこに含まれるすべてのエッジ点に関する主成分分析を行い，その主方向を接ベクトルとして採用する．

以上の準備のもと，画像上のすべての点 P において輪郭線の存在確率 $pf(\mathbf{p})$ を導出する．各点 P における確率 $pf(\mathbf{p})$ は，次の 3-step に従って計算される．

Step 1:　探索窓のサイズを決定する．予め設定された窓サイズの初期値 h_{init} から始まり，探索窓内に少なくとも N_{min} 個のエッジ点が含まれるようになるまで窓サイズを徐々に増やしていく．最終的に決められた窓サイズを h とする．ここで窓サイズ h という場合は，現在確率を計算したい点 P を中心とする $(2h+1) \times (2h+1)$ の探索窓を指す．

Step 2:　確率を計算したい画素 P を中心とする窓を考え，画素 P と窓に含まれるエッジ点 Q，およびその接ベクトル $\mathbf{t}$ を用いて接触円による補間 [19] を行う．窓内のすべてのエッジ点 Q に対してこれを実施し，P における推定接ベクトル $\mathbf{t}^*$ の分布を調べる．接触円による補間の詳細については，8.2.2 で述べる．

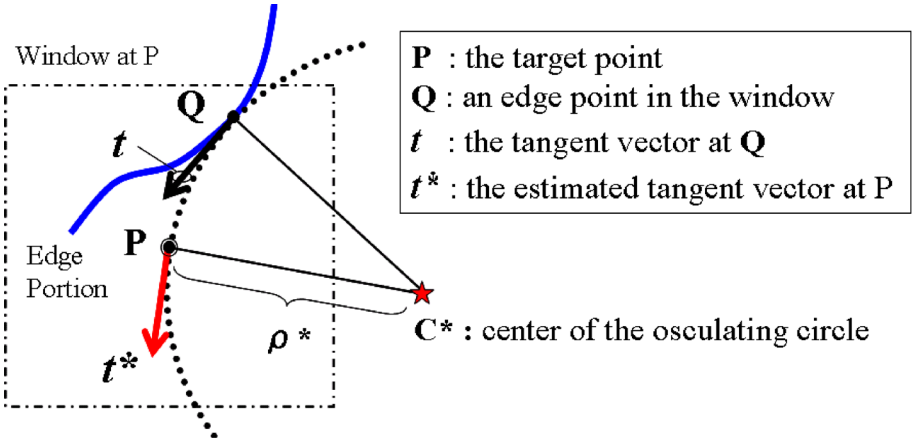


図 8.1: 接触円による補間

Step 3:　$\mathbf{t}^*$ の分布を基に P 上に輪郭線が存在する確率を計算する．P が輪郭線上にあれば，近傍の多くのエッジ点に支持され，傾きが本来の輪郭線の接線方向に集中するのに対し，輪郭点以外の点における傾きはかなりばらつく傾向がある．このため，傾きの集中度が高い画素ほど輪郭線の存在確率が高くなるよう存在可能性の場を構成すればよい．存在確率の具体的な計算方法については，8.2.3 で述べる．

8.2.2　接触円による補間

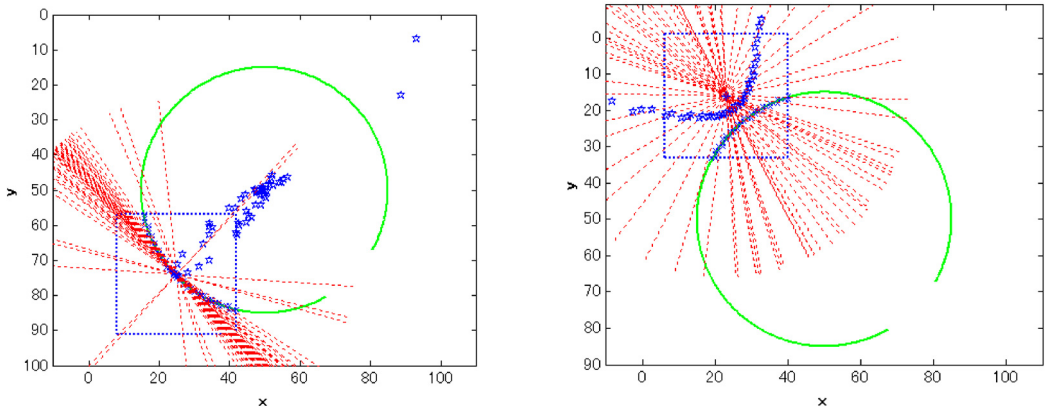
図 8.1 に示すように，確率を計算したい画素 P を中心とする窓を考え，画素 P と窓に含まれるエッジ点 Q，およびその接ベクトル $\mathbf{t}$ が与えられているものとする．このとき，2 点 P,Q を通り，点 Q において接ベクトル $\mathbf{t}$ に接する平面上の円が唯一つ存在する．この円を接触円 (osculating circle) と呼び，その中心座標 C^* 及び半径 ρ^* は次の式により計算される [19] :

$$\mathbf{c}^* = \mathbf{q} + \rho^* \mathbf{m}, \quad \rho^* = \frac{1}{2}\frac{\|\Delta\mathbf{q}\|^2}{\mathbf{m}\cdot\Delta\mathbf{q}}, \tag{8.1}$$

ここで，$\mathbf{m}$ は $\mathbf{t}$ に直交する単位ベクトルであり，$\Delta\mathbf{q}$ は差分ベクトル $\mathbf{p}-\mathbf{q}$ を表している．接触円の中心座標 C が計算されれば，ターゲットとなる画素 P における推定接ベクトル $\mathbf{t}^*$ は差分ベクトル $\mathbf{p}-\mathbf{c}^*$ に直交する方向として，容易に計算することができる．また，窓内のすべてのエッジ点 Q に対して同様の接触円補間を実施することにより，P における推定接ベクトル $\mathbf{t}^*$ の分布が得られる．

8.2.3　存在確率の算出

上述した接触円補間は点 P が直線 $l(\mathbf{q},\mathbf{t})$ 上に存在する場合には成立しない．ここで，$l(\mathbf{q},\mathbf{t})$ は，点 Q とその接ベクトル $\mathbf{t}$ によって定義される直線である．このため，確率計算を行う



(a) P on circle　　(b) P outside circle

図 8.2: 補間により求めた接線方向の分布

前に，点 P は直線補間すべきか，それとも接触円補間すべきかを考慮しておく必要がある．ここでは，この状況を次のような条件付き確率によってモデル化する．

$$\begin{aligned} &Prob\{P \text{ is on contour}\} \\ &= Prob\{\text{Line}\} \cdot Prob\{P \text{ is on line}|\text{Line}\} \\ &\quad + Prob\{\text{Curve}\} \cdot Prob\{P \text{ is on curve}|\text{Curve}\}. \end{aligned} \tag{8.2}$$

$Prob\{\text{Line}\}$ は，探索窓内にある何個のエッジ点が直線補間を支持しているかを支援する確率であり，次のように計算する．

$$Prob\{\text{Line}\} = \sum_{\substack{\mathbf{q} \ s.t. \\ P \in l(\mathbf{q},\mathbf{t})}} Score(\mathbf{q}) \Big/ \sum_{\mathbf{q}} Score(\mathbf{q}) \tag{8.3}$$

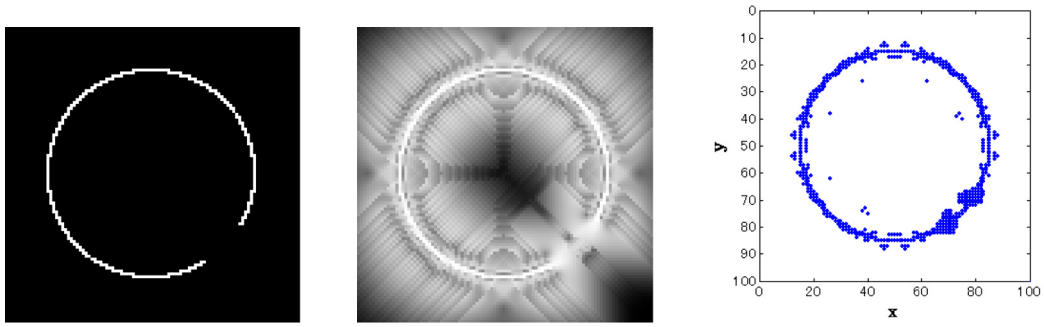
ここで，$\mathbf{q}$ は探索窓内のすべてのエッジ点を表し，関数 $Score(\mathbf{q})$ は点 Q のスコアを決める関数であり，Q が P から離れるにつれて小さく評価されるように設定されている．また，式 (8.2) に関して，さらに

$$Prob\{\text{Curve}\} = 1 - Prob\{\text{Line}\}, \tag{8.4}$$

$$Prob\{P \text{ is on line}|\text{Line}\} = 1. \tag{8.5}$$

として計算する．上の式 (8.5) は $P \in l(\mathbf{q}, \mathbf{t})$ を満たす任意のエッジ点 Q は，点 P において常に同じ接線方向 $\mathbf{t}$ を支持することが明らかであるという意味において，妥当な仮定であると考えられる．

最後に，上述した接触円補間によって得られた接ベクトル $\mathbf{t}^*$ の分布に基づいて，確率 $Prob\{P \text{ is on curve}|\text{Curve}\}$ を計算する．基本的には，より多くのエッジ点が点 P において



(a)　(b)　(c)

図 8.3: 欠けた円弧に対する「存在可能性を表す場」の導出 ($h_{init} = 3$, $N_{min} = 7$)

同じ方向の接ベクトルを支持しているほど，点 P が輪郭線上に存在する確率がより高くなる，という考え方に従って確率を決める．図 8.2 に，点 P おいて算出した接線の分布に関する 2 つの例を示す．ここでは，エッジ画像として，弧が一部欠けた円を使用している．(a) は P を円上にとった場合，(b) は P を円のやや外側に取った場合の事例を表している．図から明らかなように，(a) の場合は P における推定接ベクトルの方向が真の接ベクトルの周辺に集中しているのに対し，(b) の場合は推定接ベクトルはあらゆる方向に散らばっており特定の方向を見出すことはできない．

本章においては，存在確率を計算するにあたり，推定接ベクトルの分布そのものを使用せずに，むしろ接触円の中心座標 C^* の分布を用いることとする．その理由として，図 8.2(a) に見るように，たとえ P が正確に円弧の上に位置している場合にも，推定接ベクトルの方向は，離散化による影響により，かなりばらついてしまう傾向があることが挙げられる．接触円の推定中心 C^* もまたばらついているが，幸いなことに，それらは P と C を結ぶ直線の周辺に集中していることがわかる．そこで，推定中心 C^* に対して P を通る直線に関する主成分分析を行い，ここで得られた主方向への寄与率を，P における確率 $Prob\{P \text{ is on curve}|\text{Curve}\}$ として採用することとした．すなわち，いま対応する共分散行列の固有値を λ_1, λ_2 ($\lambda_1 \geq \lambda_2$) と書くことにすれば，

$$Prob\{P \text{ is on curve}|\text{Curve}\} = \frac{w_h \cdot \lambda_1}{\lambda_1 + \lambda_2}, \tag{8.6}$$

と書くことができる．ここで，$w_h \in [0, 1]$ は重み付け係数であり，窓サイズ h に関する単調減少関数として定義されるものとする．

図 8.3 は，弧の一部が欠けた円のエッジ画像に対して，上に述べた「輪郭線の存在可能性を表す場」の導出を行った例を示している．図 8.3(a) は入力されたエッジ画像，図 8.3(b) は，上の導出式に従って算出した存在可能性の場を表示した画像である．図 8.3(c) は，エッジ輪郭の存在可能性が 0.85 を超えた点をプロットしたものである．(b),(c) の画像から，欠けていた円弧がうまく補間されている様子がわかる．

8.3　存在可能性の場を用いた閉輪郭線の時間発展処理

本節では，曲線の時間発展方程式に基づきオブジェクト輪郭線を抽出する方法を検討する．以下，前節で検討した「輪郭線の存在可能性を表す場」を利用してオブジェクト輪郭が抽出できるよう，GAC を修正する．

8.3.1　測地線による動的輪郭モデルの修正

Caselles らによる GAC モデル [12] は，画像平面 $I(x,y)$ 上の曲線 $\mathbf{c}(q)=(x(q),y(q))$ の長さを

$$L(\mathbf{c})=\int_a^b g(|\nabla I(\mathbf{c}(q))|)|\mathbf{c}'(q)|dq \tag{8.7}$$

で定義したとき，これを最小とする測地線に向かって $\mathbf{c}(q)$ を収束させる幾何学的な動的輪郭法であった．ここで，関数 g はエッジ検出関数であり，通常，$\mathbf{R}^+$ 上で定義される単調減少関数が使用される．$L(\mathbf{c})$ は通常のユークリッド距離に重み係数 $g(|\nabla I(\mathbf{c}(q))|)|$ を掛けた一般化距離であり，このおかげで，オブジェクト輪郭に相当する閉曲線を $L(\mathbf{c})$ を最小とする曲線として特徴付けることができるようになる．

一方，閾値処理を行った 2 値のエッジ画像を対象とする場合，濃淡画像に基づく勾配 $|\nabla I|$ のような情報は既に失われ，利用することができない．そこで提案法では，前段で生成した輪郭線の存在可能性の場 $pf(\mathbf{p})$ を用いて，曲線の長さに関する最小化問題を以下のように修正することとした．

$$\min_{\mathbf{c}} L(\mathbf{c})=\min_{\mathbf{c}}\int_a^b\{1-pf(\mathbf{c}(q))\}|\mathbf{c}'(q)|dq. \tag{8.8}$$

ここで，従来法の重み付け係数 $g(|\nabla I(\mathbf{c}(q))|)|$ に代わって $1-pf(\mathbf{c}(q))$ が使用されている点に注意されたい．これにより初期輪郭線は，最終的に $pf(\mathbf{p})$ の存在確率がなるべく 1 に近い画素を通る最適な閉曲線に収束することが期待できる．式 (8.8) の解を数値的に求めるには，次の勾配降下フローに従って曲線を時間発展させればよい．

$$\frac{\partial\mathbf{c}}{\partial t}=(\kappa(1-pf)+\langle\nabla pf,N\rangle)N. \tag{8.9}$$

ここで，κ は曲線 $c(q)$ の曲率, N は曲線に対する単位法ベクトルである．また，ブラケット $\langle\cdot,\cdot\rangle$ は R^2 における内積を表す．なお，式 (8.9) の右辺は，式 (8.8) に関する Euler-Lagrange 方程式を計算することによって得られる．

8.3.2　レベルセット法による定式化

ここでは，修正 GAC の式 (8.9) のレベルセット法による定式化を考える．レベルセット法では，曲線 $\mathbf{c}(q)$ を画像平面上で定義される曲面 $u(x,y)$ のゼロ交差とみなし，曲面 u 全体を時間発展させることにより曲線 $\mathbf{c}(q)$ の変化の様子を表現する．曲面 u を以下の符号付き

距離関数として定義する．

$$u(\mathbf{x}) = \hat{d}(\mathbf{x}, \mathbf{c}) = \begin{cases} +d(\mathbf{x}, \mathbf{c}) & \text{if } \mathbf{x} \text{ is outside } \mathbf{c}, \\ -d(\mathbf{x}, \mathbf{c}) & \text{if } \mathbf{x} \text{ is inside } \mathbf{c}, \end{cases} \tag{8.10}$$

ここで，$d(\mathbf{x}, \mathbf{c})$ は曲線 c へのユークリッド距離で，$\mathbf{x} = (x, y) \in \Omega \subset R^2$ とする．

このとき，式 (8.9) に対応する曲面 u に関する時間発展方程式は次式で与えられる．

$$\frac{\partial u}{\partial t} = (1 - pf) div\Big(\frac{\nabla u}{|\nabla u|}\Big)|\nabla u| - \langle \nabla pf, \nabla u \rangle. \tag{8.11}$$

レベルセット法を導入することのメリットとして，得られる曲線が常に閉曲線であること，曲線の分離・統合などのトポロジーの変化に対応可能であることなどが挙げられる．

8.4　シミュレーションによる検証 (1)

上で述べた修正 GAC を用いて，2 値エッジ画像からオブジェクト輪郭を抽出するための実験を行った．本シミュレーションで使用したエッジ画像は，Hausdorff トラッカー [48] によって得られた，2 値エッジ画像に基づくオブジェクトモデルである．図 8.4(a) 示した画像は，MPEG-4 標準画像シーケンス “Hall Monitor” の第 32 フレームから切り出したオブジェクトモデルの例である．

今回，輪郭線の存在可能性を表す場の導出にあたり，次の関数を使用した．

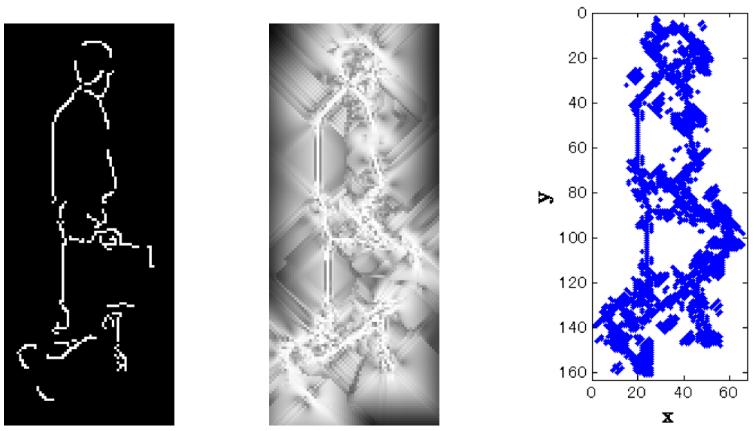
$$Score(\mathbf{q}) = w(\|\mathbf{q} - \mathbf{p}\|_\infty, 2h), \;\; w_h = w(h - h_{min}, 30),$$

ここで，$w(n, N)$ は Hanning 窓 [31] を次のように修正した関数として定義される．

$$w(n, N) = 0.5 - 0.5cos(\pi + \pi n/N).$$

このとき，結果として得られた存在可能性を表す場 $pf(x, y)$ を図 8.4(b) に示す．ただし，本図においては，$pf(x, y)$ が 1.0 を取るときは白，0.0 を取るときは黒として表示している．続く，輪郭線の時間発展処理においては，(b) を若干平滑化した画像を用いる．図 8.4(c) は，存在可能性 $pf(x, y)$ が 0.85 を超えた点をプロットしたものである．この図から，不完全なオブジェクト輪郭を持つエッジ画像 (a) に対して，輪郭線の存在可能性の場を考えると，ギャップが補間され，閉輪郭が得られることがわかる．

図 8.5 は，レベル曲面 u のゼロレベルの等高線が，初期輪郭である矩形の輪郭からオブジェクト輪郭まで時間発展していく様子を示した図であり，反復回数が 50, 100, 200, 400, 1000 のときの輪郭を提示している．図 8.5(f) に示す 1000 回反復後の最終結果を見る限り，オブジェクトの形状をよく捉えた輪郭線が得られていることがわかる．少なくとも，輝度勾配を用いる通常の GAC によって得られたオブジェクト輪郭と同程度には，良好な結果が得られている．しかしながら，得られたオブジェクト輪郭の正確性については，まだ若干不満が残る箇所も残されている．たとえば，図 8.5(f) の例では，輪郭線が顔の前方にやや膨らんでいたり，左足の先に当たる部分が抽出できなかったり点については，まだ改善の余地がある．



(a)　(b)　(c)

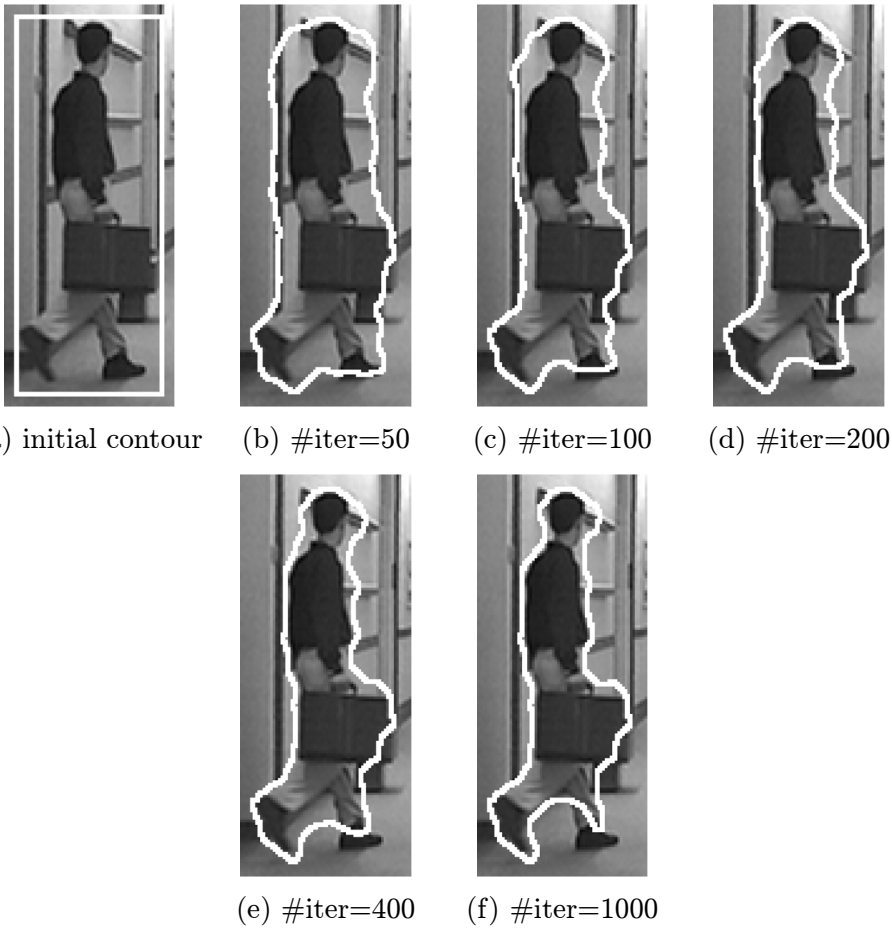
図 8.4: オブジェクトモデルに対する「存在可能性を表す場」の導出 ($h_{init} = 3$, $N_{min} = 5$)

8.5　接ベクトル導出手法の改善

8.4 まで述べてきたオブジェクト抽出法において正確なオブジェクト領域が抽出できるかどうかは，結局のところ，存在可能性の場の性質の良し悪しにかかっている．さらに言えば，提案法における存在可能性の場生成は接触円によるエッジ補間に基づくことから，前処理で求めた各エッジ点における接ベクトルの精度を高めることが，直接，存在可能性の場 pf の精度を高めることにつながることが容易に予想される．そこで本章では，従来法における接ベクトルの導出に関する問題点を指摘するとともに，その改善手法を提案する．

8.5.1　従来法の問題点

各エッジ点における接ベクトルの導出は主成分分析に基づいて計算される．従来法においては，各エッジ点を中心とする小さな窓を考え，これに含まれるすべてのエッジ点の位置に関する主成分分析を行い，その主方向を接ベクトルの方向として採用していた．この方法では，オブジェクト輪郭が滑らかな曲線として与えられる場合には良好な接線方向が得られるが，角形の頂点に代表される接ベクトルが不連続となる点や，線分が交差した点などにおいては，すべての方向を平均化した接線方向がひとつ得られるのみである．このような場合，真のエッジの接線方向は複数存在しており，これらは通常，得られた平均的な接線方向とは一致しない．真のエッジの方向と一致しない平均的なベクトルを用いて接触円補間を行うことは，角形のコーナー付近の存在可能性の場を拡散させ，オブジェクトの正確な形状の抽出を困難にする可能性がある．



(a) initial contour　(b) #iter=50　(c) #iter=100　(d) #iter=200

(e) #iter=400　(f) #iter=1000

図 8.5: 修正 GAC によるオブジェクト輪郭の抽出過程

8.5.2　改善手法の提案

従来法の問題点は，すべてのエッジ点に対して 1 つのベクトルしか割り当てなかったことに起因する．例えば，図 8.6 に示す角形エッジの頂点 P において主成分分析を行うと，主方向として図左に示すような斜め方向のベクトル $\mathbf{v_p}$ が得られる．しかし，接触円による補間を行う際には，真のエッジの方向を反映した図右に示す 2 方向のベクトルを使用することが望ましいことは直感的に明らかである．以上の考察を踏まえ，本論文では，接線方向が不連続となるエッジ点に対して真の接線方向を反映した複数の接ベクトルを割り当てられるようアルゴリズムの改善を行った．

改善した接ベクトル導出アルゴリズムは次のような再帰処理により実現される．

(1) P を中心とするウィンドウに含まれるすべてのエッジ点 Q の集合に対して主成分分析を行い，主方向 $\mathbf{v_p}$ を求める．

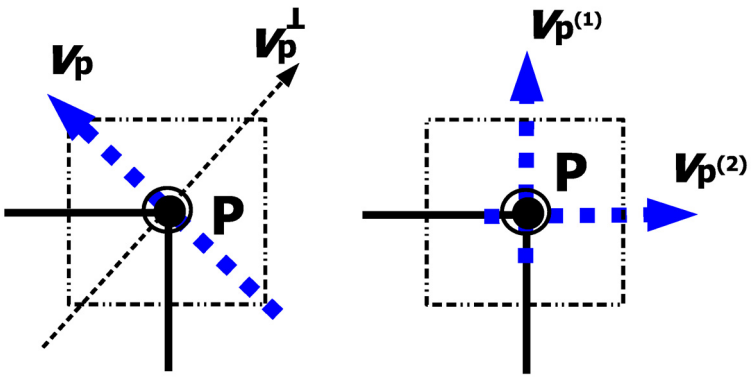


図 8.6: P における接ベクトルの導出

(2) 主方向に対する寄与率が 95%以上であれば，主方向 $\mathbf{v_p}$ を接ベクトルとして採用して処理を終了する．

(3) そうでない場合，再帰の段数が 3 以上であれば主方向 $\mathbf{v_p}$ を接ベクトルとして採用せずに処理を終了する．

(4) (2),(3) の条件がいずれも成立しない場合，主方向 $\mathbf{v_p}$ に直交する $\mathbf{v_p^{\perp}}$ 方向の直線を境としてエッジ点の集合を 2 つに分割する．

(5) 分割したそれぞれのエッジ点集合において，同様の接ベクトル導出処理を再帰的に行う．

以上のアルゴリズムの主なアイディアは，十分信頼できる寄与率が得られるまでエッジ点集合を分割していく点にある．なお，この際，主方向に直交する直線を用いてエッジ点集合を分割するのは，主方向 $\mathbf{v_p}$ に関してエッジ点の分散が最も大きくなっているためである．図 8.6 の例では，点 P を通り $\mathbf{v_p^{\perp}}$ 方向に沿った直線に基づくエッジ点の分割が，結果的に図右の 2 つの接ベクトル $\mathbf{v_p^{(1)}}$，$\mathbf{v_p^{(2)}}$ を導出する適切な選択になっていることが確認できる．

8.6　シミュレーションによる検証 (2)

エッジ画像から得られる存在可能性の場とこれに基づくオブジェクト抽出精度が，前章で提案した接ベクトル導出手法によってどの程度改善されるかを確認するためにシミュレーション実験を行った．

8.6.1　基本図形による実験

実験に使用した基本図形（円，扇形，角形）に関するテスト画像を図 8.7 に示す．いずれも輪郭の一部が欠損したエッジ画像である．また，扇形と角形のエッジには接ベクトルが不連続となる頂点が含まれている．これらのエッジ画像を入力として，存在可能性を表す場の

(a) 円　(b) 扇形　(c) 角形

図 8.7: 基本図形のエッジ画像

導出を行った結果を図 8.8 に，得られた場に基づいてオブジェクトを取り囲む閉輪郭線を時間発展させ，オブジェクト輪郭の推定を行った結果を図 8.9 に示す．

図 8.8 上段は従来手法に基づき，各エッジ点で 1 つの接ベクトルのみを用いた場合の存在可能性の場であり，図 8.8 下段は提案手法に基づき，接ベクトルが不連続となる点において複数のベクトルを割り当てた場合の存在可能性の場である．図左の円弧の場合は，基本的にすべてのエッジ点において接ベクトルは連続に変化することから，従来法でも提案法でも処理の内容はほぼ同等であり，得られた場にも顕著な差は見られない．これに対し，図中央の扇形および図右の方形のエッジに対する存在可能性の場においては，コーナー付近の確率分布の様子が顕著に異なることがわかる．特に従来法で求めた存在可能性の場には各頂点の主方向に対して直線状の強い確率が分布しているが，提案法ではこれがかなり解消されている．

このことはまた，図 8.9 の輪郭推定の結果からも確認できる．図 8.9 上段は従来手法を，図 8.9 下段は提案手法をそれぞれ用いて推定されたオブジェクト輪郭を示している．なお，オブジェクトを取り囲む初期輪郭の時間発展処理については，GAC の差分方程式に基づく繰り返し演算処理をいずれも 1200 回実施し，輪郭の時間発展が完全に止まった（解に収束した）ことを確認した上で最終結果として提示している．図 8.9 の扇形および方形の輪郭推定結果に着目すると，従来法ではコーナーを正確に抽出できていないのに対して，提案法ではこの点がよく改善されていることが主観的に確認できる．

8.6.2　標準画像による実験

次に MPEG-4 の標準画像 “Hall Monitor”の第 36 フレームから抜き出した人物のエッジ画像を入力として，本稿で提案した手法に基づき，存在可能性の場の生成・輪郭の抽出を行った結果を図 8.10 に示す．かなり良好な結果が得られたものの，頭部やカバンの上側など正確な輪郭が得られていない箇所もあることから，自然画像から得られる複雑なエッジ画像に対しては，より性質のよい存在可能性の場を導出するための更なる工夫が必要である．

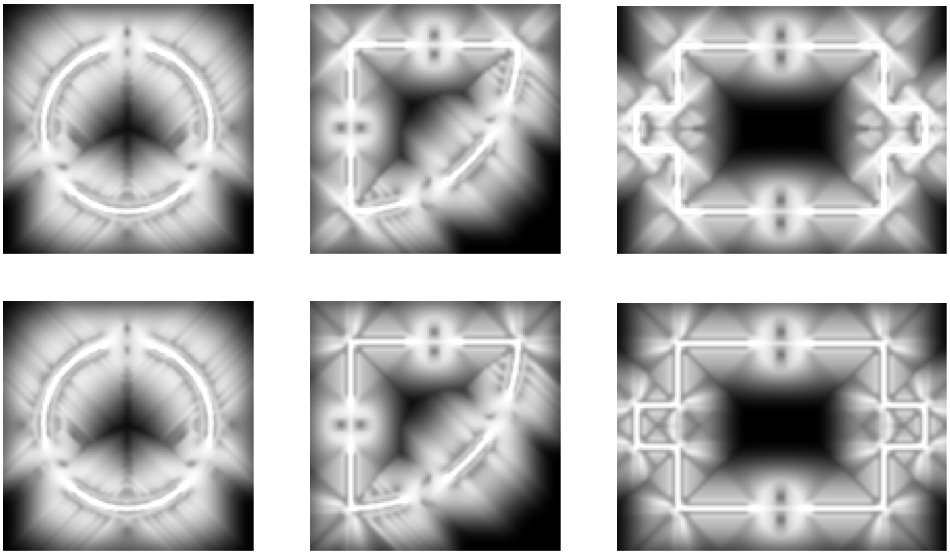


図 8.8: 存在可能性の場の導出結果（上段：従来法，下段：提案法）

8.7　おわりに

本章では，閾値処理を行った 2 値のエッジ画像からオブジェクト輪郭を抽出するための新しい手法の提案を行った．

提案手法は，与えられた 2 値のエッジ画像から輪郭線の存在可能性を表す場 $pf(x, y)$ を生成するフェーズとこの場に基づいて曲線の時間発展を行うフェーズの 2 つから構成される．新たに，輪郭線の存在可能性を表す場を導入したことにより，輝度勾配などの補助的な情報を一切用いることなく，与えられた 2 値のエッジ画像のみからオブジェクト輪郭を抽出できるという点が，本章における提案法の特長である．シミュレーション実験を行った結果，輝度勾配を用いる通常の GAC によって得られたオブジェクト輪郭と同程度には，良好な結果が得られることがわかった．

さらに，提案法において，接ベクトル導出の改善手法を新たに提案し，その存在可能性の場生成とオブジェクト輪郭抽出への効果について検証を行った．

自然画像から切り出した複雑なエッジ画像から良好なオブジェクト輪郭抽出を行うには，さらに正確かつ性質のよい存在可能性の場の生成が必要である．このため，今後の課題としては，オブジェクト輪郭の時間的なコヒーレンス性などを利用した，さらに高度なオブジェクト輪郭抽出法の検討が必要と考えられる．

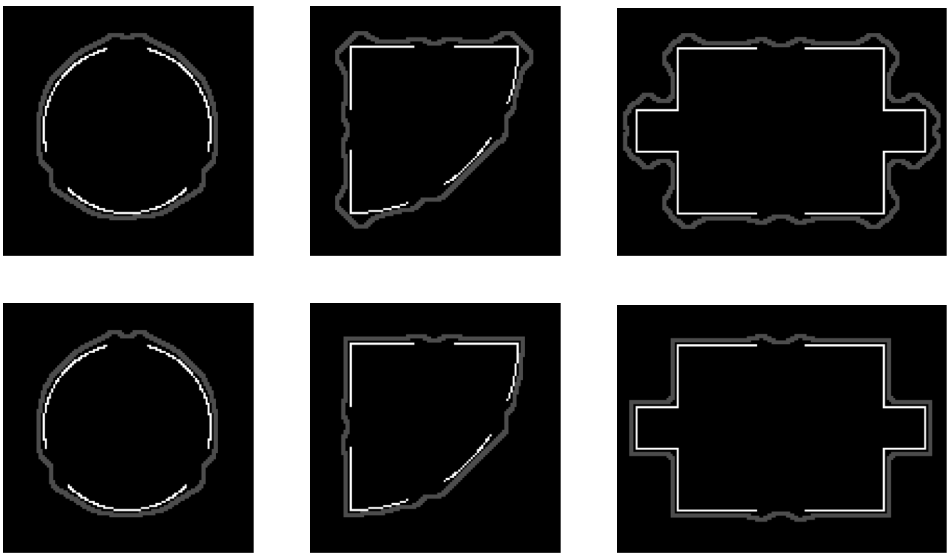


図 8.9: 輪郭の推定結果（上段：従来法，下段：提案法）

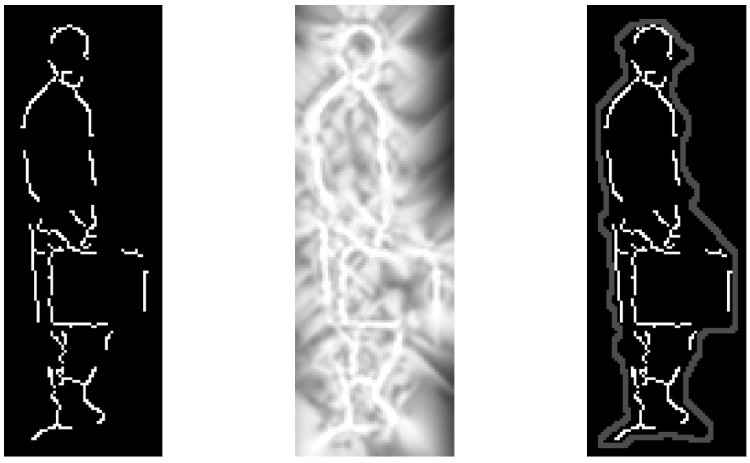


図 8.10: 人物の輪郭推定（左から，エッジ画像，存在可能性の場，輪郭推定結果）

第9章　結論

9.1　まとめ

本論文では，低ビットレート映像符号化の高度化を実現する上で重要となる前処理技術として，ローレベル符号化に対するエッジ保存平滑化処理とミッドレベル符号化に対するオブジェクト抽出処理を取り上げ，これらを2つの柱として検討を行った．双方の処理を統一的に扱うことを可能とする要素技術として，PDE (Partial Differential Equations) に基づく画像処理を積極的に導入している点が，本論文の特徴である．

第1章は序論であり，低ビットレート映像符号化において画像解析を伴う高度な画像処理の重要性を明確化し，本論文の背景及び目的について述べた．併せて本論文の全体構成を示した．

第2章では，PDEに基づく画像処理の従来研究として，スケールスペース理論，画像修復，セグメンテーション，オブジェクト・トラッキング，画像編集への適用事例について述べた．特に，エッジ保存平滑化フィルタの構築に重要な役割を果たす「非等方拡散法 (PMD)」とエッジ情報に基づくオブジェクト抽出処理において重要な役割を果たす「測地線に基づく動的輪郭法 (GAC)」については詳細に述べた．

第3章では，H.263をはじめとするローレベル符号化において，プレフィルタ処理が入力画像の情報量を減らし符号化歪みの発生を緩和する効果があることを具体的な事例を通して示した．本事例では，パラメータ b により平滑化の強弱を設定できる線形フィルタ群を用いて，これを適切に制御しながらエンコーダのプレフィルタとして使用した．提案手法は，入力画像の複雑度とVBVバッファ残量に応じて，b を $1 \leq b \leq 8$ の範囲で変動させることにより，Iピクチャにおける発生符号量を抑えて，フレームスキップの発生を抑制するものである．本章の結びとして，本事例を通して，画像全体にボケを伴う線形プレフィルタの問題点を指摘し，視覚的に重要なエッジはぼかさずに残し，他のテクスチャのみを選択的に平滑化する非線形フィルタの必要性を主張している．

続く第4章－第5章は，本論文の1つの柱であるエッジ保存平滑化処理をPDEに基づき実現するための検討を行っている．

第4章では，PMDを用いてエッジ保存空間平滑化フィルタを構築するための検討を行った．正則化PMDはもともとエッジ保存性を持っており，PMDの欠点であった振幅の大きい雑音 (salt & pepperノイズ) などの除去にも対応できるため，一見，平滑化プレフィルタとしての要件を満たしているように見える．しかし，テスト画像を用いた予備実験の結果から，過度のエッジ強調効果により，ステップエッジ部分にピークが残留してしまうことがわかった．そこで，正則化処理を行わなくても，エッジを保存しながら振幅の大きいノイズの

除去ができ，かつエッジ強調効果が起こらない非等方拡散を実現するために，エッジ停止関数が満たすべき数学的な条件を明らかにした．また，これらの条件を満たすエッジ停止関数として g^{TN} を具体的に導出するとともに，シミュレーション実験を通して，導出した関数のノイズ除去性能とエッジ保存性及び非強調性の確認を行った．

第 5 章では，動画像符号化におけるフレーム間予測誤差信号の低減を考慮し，第 4 章で提案した空間 2 次元の PMD を，時空間 3 次元に拡張するための検討を行った．Perona-Malik によって提示された非等方拡散の PDE は一般に n 次元で成立するため，時空間 3 次元への拡張は近傍系を時空間 6-近傍に取ることによって容易に実現される．本章では，まず，新規に導入した時間方向の拡散処理が，従来から動画像符号化で使用されてきた差分抑制関数による時間フィルタ (Temporal Filter) の自然な拡張とみなせることを数理的に明らかにした．さらにこの事実に基づき，時間方向のエッジ停止関数 g^{tp} の選定を行った．差分抑制関数に基づく理論的な考察とシミュレーション実験を通して，従来法に登場する 3 種類のエッジ停止関数 g^{BS}, g^{PM}, g^{TN} の中では，g^{PM} が時間方向のエッジ停止関数として最も適していると結論した．

第 6 章では，ミッドレベル符号化に位置づけられるオブジェクトベース符号化の実現に必要となる VOP (Video Object Plane) 自動生成の研究事例として，エッジ情報に基づく Meier の方法について述べた．Meier の方法では，2 値エッジ画像に対する Hausdorff マッチングを用いてオブジェクトトラッキングを行うことから，得られたオブジェクトモデルもまた 2 値エッジ画像として表現される．このため，オブジェクトを抽出するには，輪郭が不完全で過剰なエッジを含んだ 2 値エッジ情報を頼りに，オブジェクト領域を抽出する必要がある．本章の結びとして，従来手法を踏まえながら，このようなオブジェクト抽出の問題が簡単に見えて実はそうではないことを指摘し，続く章への問題提起としている．

続く第 7 章－第 8 章では，以上の問題提起を受け，本論文のもう 1 つの柱であるエッジ情報に基づくオブジェクト抽出処理を PDE に基づき実現するための検討を行っている．

第 7 章では，第 6 章で提示されたオブジェクト抽出の問題を，GAC を用いて解決する方法について検討を行った．提案法では，まずオブジェクトモデルとして与えられた 2 値エッジ画像から，オブジェクト内部に含まれる領域を生成する．ここで，この領域は複数個の非連結成分を含んでいてもよい．次に，これらの輪郭を取って初期輪郭線とし，これを，レベルセット法により，オブジェクトの内側から外側に向かって膨張・統合させて，オブジェクト輪郭を得る．このとき，輪郭線の時間発展を停止するために，グレイスケール画像を参照して輝度勾配を計算している．シミュレーション実験を通し，本提案法は，特にエッジの検出状態が悪い場合に，従来法よりも主観的に優れたオブジェクト輪郭を生成する傾向があることを確認した．

第 8 章では，第 7 章と同様の問題を，輝度情報を参照することなく，2 値情報のみを用いて解決する方法について提案を行った．第 7 章のように輝度信号の勾配を用いて輪郭線の運動を止める場合は，勾配強度の弱い部位から輪郭線がはみ出し，誤って背景のエッジを取り込んでしまう場合があった．ここでは，輝度勾配に代わって新たに「オブジェクト輪郭線の存在可能性を表す場」を導入し，これに基づいて輪郭線を時間発展させることにより，上述

のような事態を回避できることを確認した．本手法は場の定義を改善することにより，更に正確なオブジェクト輪郭を得られるポテンシャルを持っている．

以上，本論文は，PDE に基づく画像処理手法を要素技術として，低ビットレート映像符号化に有効な前処理技術の提案を行った．提案したエッジ保存平滑化処理及びオブジェクト抽出処理は，従来は困難であった画像解析を伴うオブジェクト単位の処理を動画像符号化に導入するための試みであり，今後の動画像符号化アルゴリズム研究における 1 つの重要な方向性として意味を持つとともに，低ビットレート環境の動画像通信サービスにおける主観画質と機能性の向上に貢献できると考えられる．

9.2　今後の課題 — 動画像符号化への応用

前半で検討を行った時空間平滑化フィルタに関しては，現在，エントロピーなどを使った机上でのシミュレーションでの確認のみに止めている．動画像符号化への応用を考える場合には，実際に MPEG-4 などのエンコーダを通して，フィルタによる主観画質の向上を確認する必要があるが，パラメータの組み合わせが膨大となると，主観評価ですべて行うのは困難である．どのような評価方法が妥当になるのか，検討を進めなければならない．

また，後半のオブジェクト輪郭の比較にしても，現時点では目視で確認し，主観的に良し悪しを論じているのみであり，これについてもいくつかの客観評価尺度を導入し，厳密な比較を進めていく必要がある．その際にも，評価の基準となる正しいセグメンテーションとは何かを検討しておかなければならない．

謝辞

本研究を進めるにあたりご指導いただきました，指導教官である東京工業大学大学院情報理工学研究科計算工学専攻の中嶋正之教授に深く感謝いたします．社会人として仕事と家庭を持ちながらの研究は想像以上に厳しく，研究室に顔を出すこともままならない私に対しても，学位取得に向けていろいろと気にかけていただきました．また，研究テーマについても，経験の浅い私の自主性を尊重していただき，自由に進めさせていただきました．

本論文の審査をお引き受けいただきました東京工業大学大学院情報理工学研究科計算工学専攻の亀井宏行教授，同大学院理工学研究科集積システム専攻の酒井善則教授，同大学院総合理工学研究科物理情報システム専攻の羽鳥好律教授，同大学院情報理工学研究科計算工学専攻の齋藤豪助教授に感謝いたします．予備審査会及び博士論文発表会においては，拙い研究内容にもかかわらず熱心に耳を傾けていただき，数々の有益なコメントをいただきました．なお，本論文の第８章で提案したアルゴリズムの基礎となる「輪郭線の存在可能性を表す場」に関するアイディアは齋藤先生の論文を参考にさせていただきました．国内大会と国際会議に論文を投稿する際には，場の精度を高める方法について多くのヒントをいただきました．

東京工業大学大学院情報理工学研究科計算工学専攻の高橋裕樹助手に感謝いたします．研究の進め方から研究室の行事に至ることまで，常日頃からいろいろと相談に乗っていただきました．また，論文投稿の際には拙い論文を添削して頂くとともに，有益な助言をいただきました．

今回の学位取得にあたり，私が東京工業大学大学院情報理工学研究科に在籍する３年の間，同時に社会人として助手を務めた神奈川工科大学情報学部情報工学科においては，多忙のため様々なご迷惑をおかけしました．ご協力いただきました先生方に感謝いたします．特に，常日頃から研究の場を提供していただきます神奈川工科大学情報学部情報工学科の徳増眞司教授に感謝いたします．

最後に，博士論文の執筆のため，家族と過ごす時間があまり取れない状況の中にも，惜しみない協力をしてくれた妻と長男に深く感謝いたします．

参考文献

[1] 阿部正英，大内隆志，川又政征，“遺伝的アルゴリズムを用いた 2 値画像からの任意形状の輪郭線抽出”, 信学論 (D-II), vol.J86-D-II, no.7, pp.1036-1048, July 2003.

[2] L. Alvarez, P.L. Lions, and J.M. Morel, “Image Selective Smoothing and Edge Detection by Nonlinear Diffusion II,” SIAM J. Numer. Anal., vol.29, no.3, pp.845-866, March 1992.

[3] L. Alvarez, F. Guichard, P.L. Lions, and J.M. Morel, “Axioms and Fundamental Equations of Image Processing”, Arch. Ration. Mechan. Anal., vol.123, no.3, pp.199-257, 1993.

[4] G. Aubert and P. Kornprobst, Mathematical Problems in Image Processing, Springer, New York, 2002.

[5] D. Barash, “A Fundamental Relationship between Bilateral Filtering, Adaptive Smoothing, and the Nonlinear Diffusion Equation,” IEEE Trans. Pattern Anal. Mach. Intell., vol.24, no.6, pp.844-847, June 2002.

[6] D. Barash and D. Comaniciu, “A Common Framework for Nonlinear Diffusion, Adaptive Smoothing, Bilateral Filtering and Mean Shift,” Image and Video Computing, vol.22, no.1, pp.73-81, Jan. 2004.

[7] M. Bertalmio, G. Sapiro, and G. Randall, “Morphing Active Contours”, IEEE Trans. Pattern Anal. Mach. Intell., vol.22, no.7, pp.733-737, July 2000.

[8] M. Bertalmio, G. Sapiro, V. Caselles, and C. Ballester, “Image Inpainting”, Proceedings of the SIGGRAPH, pp.417-424, 2000.

[9] M.J. Black, G. Sapiro, D.H. Marimont, and D. Heeger, “Robust Anisotropic Diffusion”, IEEE Trans. Image Process., vol.7, no.3, pp.421-432, March 1998.

[10] J. Canny, “A Computational Approach to Edge Detection,” IEEE Trans. Pattern Anal. Mach. Intell., vol.8, no.6, pp.679-698, Nov. 1986.

[11] V. Caselles, F. Catte, T. Coll, and F. Dibos, “A Geometric Model for Active Contours in Image Processing”, Numer. Math., vol.66, pp.1-33, 1993.

[12] V. Caselles, R. Kimmel, and G. Sapiro, “Geodesic Active Contours”, Int. J. Comput. Vision, vol. 22, no. 1, 1997.

[13] F. Catte, P.L. Lions, J.M. Morel, and T. Coll “Image Selective Smoothing and Edge Detection by Nonlinear Diffusion,” SIAM J. Numer. Anal., vol.29, no.1, pp.182-193, Feb. 1992.

[14] T.F. Chan, S. Osher, and J. Shen, “The Digital TV Filter and Nonlinear Denoising,” IEEE Trans. Image Process., vol.10, no.2, pp.231-241, Feb. 2001.

[15] T.F. Chan and L.A. Vese, “Active Contours Without Edges,” IEEE Trans. Image Process., vol.10, no.2, pp.266-277, Feb. 2001.

[16] D. Comaniciu and P. Meer, “Mean Shift: A Robust Approach Toward Feature Space Analysis,” IEEE Trans. Pattern Anal. Mach. Intell., vol.24, no.5, pp.603-619, May 2002.

[17] T.H. Cormen, C.E. Leiserson, and R.L. Rivest, Introduction to Algorithms, MIT Press, 1990.

[18] E. Dubois and S. Sabri, “Noise Reduction in Image Sequences Using Motion-Compensated Temporal Filtering,” IEEE Trans. Communications, vol.32, no.7, July 1988.

[19] G. Farin, Curves and Surfaces for CAGD (5th edition), Morgan Kaufmann, 2002.

[20] スタンリー・ファーロウ（伊理正夫，伊理由美 訳），偏微分方程式 科学者・技術者のための使い方と解き方，朝倉書店，1996.

[21] 藤原洋，安田浩（監修），ポイント図解式　ブロードバンド＋モバイル　標準 MPEG 教科書, ASCII, 2003.

[22] 藤原洋（監修），画像＆音声圧縮技術のすべて, CQ 出版社, 2000.

[23] S. Geman and D. Geman, “Stochatic Relaxation, Gibbs Distributions, and the Bayesian Restoration of Images”, IEEE Trans. Pattern Anal. Mach. Intell., vol.14, no.3, pp.721-741, March. 1993.

[24] 儀我美一，陳蘊剛，“動く曲面を追いかけて”, 日本評論社, 1996.

[25] R.C. Gonzalez, R.E. Woods, Digital Image Processing (second edition), Prentice Hall, 2002.

[26] C. Gu and M. Lee, “Semiautomatic Segmentation and Tracking of Semantic Video Objects”, IEEE Trans. Circuits and Systems for Video Technology, vol.8, no.5, pp.572-584, Sept. 1998.

[27] M.T. Heath, Scientific Computing - An Introductory Survey (2nd edition), MacGraw-Hill, 2002.

[28] B.B. ハバード, “ウェーブレット入門 —数学的道具の物語—” , 朝倉書店, 2003.

[29] D.P. Huttenlocher, G.A. Klanderman, and W.J. Rucklidge, “Comparing Images Using the Housdorff Distance,” IEEE Trans. Pattern Anal. Mach. Intell., vol.15, no.9, Sept. 1993.

[30] D.P. Huttenlocher, J.J. Noh, and W.J. Rucklidge, “Tracking Nonrigid Objects in Complex Scenes”, Proc. IEEE Int. Conf. Computer Vision (ICCV’93), Berlin, Germany, May 1993.

[31] 飯國洋二, “基礎から学ぶ信号処理”, 培風館, 2004.

[32] 井上誠喜, 八木伸行, 住吉英樹, “C 言語で学ぶ実践 CG・映像合成”, オーム社, 2005.

[33] ISO/IEC JTC1/SC29/WG11, “Test Model 5,” MPEG93/457, April 1993.

[34] ISO/IEC JTC1/SC29/WG11, “Description of Automatic Segmentation Techniques Developed and Tested for MPEG-4 Version1”, MPEG97/2702, Oct. 1997.

[35] ISO/IEC JTC1/SC29/WG11, “MPEG-4 Video Verification Model version 17.0,” MPEG/N3515, July 2000.

[36] A.K. Jain, Fundamentals of Digital Image Processing, Prentice Hall, 1989.

[37] P.V. Karunaratne, C.A. Segall, and A.K. Katsaggelos, “A Rate-distortion Optimal Video Pre-processing Algorithm,” Proc. IEEE Int. Conf. Image Process. (ICIP’01), vol.1, pp.481-484, Thessaloniki, Greece, Oct. 2001.

[38] M. Kass, A. Witkin, and D. Terzopoulos, “Snakes: Active Contour Models”, Proc. IEEE Int. Conf. on Conputer Vision (ICCV’87), pp.259-268, London, UK, 1987.

[39] I. Kim, Y. Jeong, and K.T. Park, “The Block-based Preprocessing System for the Coding Performance Improvement”, IEEE Trans. Comsumer Electronics, vol.44, no.3, pp.1048-1053, March 1998.

[40] R. Kimmel, Numerical Geometry of Images, Springer, New York, 2004.

[41] 木全英明, 清水淳, 八島由幸, 小林直樹, “低ビットレート映像符号化におけるエッジ情報を利用した時空間平滑化プレフィルタの一検討”, 1999 年画像符号化シンポジウム (PCSJ99), P-1.08, Nov. 1999.

[42] 木全英明，八島由幸，小林直樹，“低ビットレート映像符号化におけるエッジ保存形プレフィルタ制御法の一検討”, 2000 年画像符号化シンポジウム (PCSJ00), pp.57-58, Nov. 2000.

[43] H. Kimata, Y. Yashima, and N. Kobayashi, “Edge Preserving Pre-post Filtering for Low bitrate Video Coding,” Proc. IEEE Int. Conf. Image Process. (ICIP’01), vol.3, pp.554-557, Thessaloniki, Greece, Oct. 2001.

[44] J. Koenderink, “The Structure of Images,” Biol. Cybern., no.50, pp.363-370, 1984.

[45] R. Malladi, J.A. Sethian, and B.C. Vemuri, “Shape Modelling with Front Propagation: a Level Set Approach”, IEEE Trans. Pattern Anal. Mach. Intell., vol.17, no.2, pp.158-175, Feb. 1995.

[46] A-R. Mansouri, A. Olivier, and J. Konrad, “Topology-independent Region Tracking with Level Sets”, Proc. IEEE Int. Conf. Image Process. (ICIP’00), vol.3, pp.66-69, Vancouver, Canada, Sept. 2000.

[47] A-R. Mansouri, “Region Tracking via Level Set PDEs without Motion Computation”, IEEE Trans. Pattern Anal. Mach. Intell., vol.24, no.7, pp.947-961, July 2002.

[48] T. Meier and K.N. Ngan, “Automatic Segmentation of Moving Objects for Video Object Plane Generation”, IEEE Trans. Circuits and Systems for Video Technology, vol.8, no.5, pp.525-538, Sept. 1998.

[49] T. Meier and K.N. Ngan, “Video Segmentation for Content-Based Coding”, IEEE Trans. Circuits and Systems for Video Technology, vol.9, no.8, pp.1190-1203, Dec. 1999.

[50] F. Meyer and P. Paragos, “Nonlinear Scale-Space Representation with Morphological Levelings”, J. Visual Communication and Image Representation, vol.11, no.2, pp.245-265, 2000.

[51] 湊淳，小澤哲，“2 値画像の輪郭線抽出へのシミュレーテッドアニーリングの応用”，信学論 (D-II), vol.J88-D-II, no.1, pp.148-153, Jan. 2005.

[52] N. Mukawa, H. Kuroda and T. Matsuoka, “An Interframe Coding System for Video Teleconferencing Signal Transmission at a 1.5 Mbit/s Rate”, IEEE Trans. COM, vol.32, no.3, 1984.

[53] 長尾智晴，進化的画像処理，昭晃堂，2002.

[54] 中村亜希子，齋藤豪，中嶋正之，“セルアニメーションのための曖昧輪郭線に対する塗りつぶし領域の推定のための場の検討”，情報処理学会第 65 回全国大会，vol.2, pp.265-266, March 2003.

[55] K.N. Ngan, T. Meier, and D. Chai, Advance Video Coding: Principles and Techniques, Elsevier, 1999.

[56] 岡本穏，齋藤豪，高橋裕樹，中嶋正之，“線の延長可能性を表す場を用いた不連続な輪郭線の延長法”, 情報処理学会第 67 回全国大会，vol.4, pp.243-244, March 2005.

[57] 大久保榮（監修），角野眞也，菊池義浩，鈴木輝彦（共著），H.264/AVC 教科書，インプレス，2004.

[58] 小野定康，鈴木純司，わかりやすい JPEG2000 の技術，オーム社，2003.

[59] S. Osher and J.A. Sethian, “Fronts Propagating with Curvature Dependent Speed: Algorithms Based on Hamilton-Jacobi Formulation”, Journal of Computational Physics, vol.79, pp.12-49, 1988.

[60] S. Osher and L.I. Rudin “Feature-Oriented Image Enhancement Using Shock Filters”, SIAM J. Numer. Anal. vol.27, no.4, pp.919-940, Aug. 1990.

[61] S. Osher and R. Fedkiw, Level Set Methods and Dynamic Implicit Surfaces, Springer, New York, 2002.

[62] N. Paragios and R. Deriche, “A PDE-based Level-Set Approach for Detection and Tracking for Moving Objects”, Proc. IEEE Int. Conf. Computer Vision (ICCV’98), Bombay, India, Jan. 1998.

[63] N. Paragios and R. Deriche, “Unifying Boundary and Region-based Information for Geodesic Active Tracking”, Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR’99), Ft. Collins, CO, June 1999.

[64] N. Paragios and R. Deriche, “Geodesic Active Regions for Supervised Texture Segmentation”, Proc. IEEE Int. Conf. Computer Vision (ICCV’99), Kerkyra, Greece, Sept. 1999.

[65] N. Paragios and R. Deriche, “Geodesic Active Regions for Motion Estimation and Tracking”, Proc. IEEE Int. Conf. Computer Vision (ICCV’99), Kerkyra, Greece, Sept. 1999.

[66] N. Paragios and R. Deriche, “Geodesic Active Contours and Level Sets for the Detection and Tracking of Moving Objects”, IEEE Trans. Pattern Anal. Mach. Intell., vol.22, no.2, pp.266-280, March 2000.

[67] N. Paragios, O. Mellina-Gottardo, and V. Ramesh, “Gradient Vector Flow Fast Geometric Active Contours”, IEEE Trans. Pattern Anal. Mach. Intell., vol.26, no.3, pp.402-407, March 2004.

[68] J.R. Parker, Algorithms for Image Processing and Computer Vision, Jhon Wiley & Sons, New York, 1997.

[69] P. Perez, M. Gangnet, and A. Blake, “Poisson Image Editing”, ACM Transactions on Graphics (SIGGRAPH’03), vol.22, no.3, pp.313-318, July 2003.

[70] P. Perona and J. Malik, “Scale-Space and Edge Detection Using Anisotropic Diffusion,” IEEE Trans. Pattern Anal. Mach. Intell., vol.12, no.7, pp.629-639, July 1990.

[71] L.I. Rudin, S. Osher, and E. Fatemi, “Nonlinear Total Variation Based Noise Removal Algorithm,” Physica D, vol.60, pp.259-268, 1992.

[72] L.I. Rudin and S. Osher, “Total Variation Based Image Restoration with Free Local Constraints,” Proc. IEEE Int. Conf. Image Process. (ICIP’94), vol.1, pp.31-35, Austin, Texas, Nov. 1994.

[73] P. Saint-Marc, J-S. Chen, and G. Medioni, “Adaptive Smoothing: A General Tool for Early Vision,” IEEE Trans. Pattern Anal. Mach. Intell., vol.13, no.6, pp.514-529, June 1991.

[74] 齊藤隆弘，“小特集 映像高能率符号化技術，2. 映像符号化の分類・要素技術”，映像情報メディア学会誌, vol.54, no.5, pp.629-635, May 2000.

[75] G. Sapiro, Geometric Partial Differential Equations and Image Analysis, Cambridge University Press, 2001.

[76] G. Sapiro,“From Active Contours to Anisotropic Diffusion: Connections between the Basic PDE’s in Image Processing”, Proc. IEEE Int. Conf. Image Process. (ICIP’96), vol.1, pp.477-480, Lausanne, Switzerland, Sept. 1996.

[77] J.A. Sethian, “A Fast Introduction to Fast Marching Methods and Level Set Methods”, http://math.berkeley.edu/ sethian/Explanations/fast_marching_explain.html.

[78] J.A. Sethian, Level Set Methods and Fast Marching Methods: Evolving Interfaces in Computational Geometry, Fluid Mechanics, Computer Vision, and Material Sciences, Cambridge University Press, 1999.

[79] M. Sonka, V. Hlavac, and R. Boyle, Image Processing, Analysis, and Machine Vision(2nd edition), Thompson Publishing Company, 1998.

[80] G.J. Sullivan and T. Wiegand, “Rate-Distortion Optimization for Video Compression”, IEEE Signal Processing Magazine, pp.74-90, Nov. 1998.

[81] J. Sun, J. Jia, C-K. Tang, and H-Y. Shum, “Poisson Matting”, ACM Transactions on Graphics (SIGGRAPH’04), vol.23, no.3, pp.315-321, Aug. 2004.

[82] C. Tomasi and R. Manduchi, “Bilateral Filtering for Gray and Color Images,” Proc. IEEE Int. Conf. Computer Vision (ICCV’98), pp.839-846, Bombay, India, Jan. 1998.

[83] E. Trucco and A. Verri, Introductory Techniques for 3-D Computer Vision, Prentice Hall, 1998.

[84] H. Tsuji, T. Sakatani, Y. Yashima, and N. Kobayashi, “A Nonlinear Spatio-temporal Diffusion and its Application to Prefiltering in MPEG-4 Video Coding,” Proc. IEEE Int. Conf. Image Process. (ICIP’02), vol.1, pp.85-88, Rochester, NY, Sept. 2002.

[85] 辻裕之，徳増眞司，高橋裕樹，中嶋正之，“非等方拡散法に基づく時空間平滑化処理とその動画像符号化への応用”, 第 19 回 NICOGRAPH 論文コンテスト論文集，pp.175-180, Nov. 2003.

[86] H. Tsuji, S. Tokumasu, H. Takahashi, and M. Nakajima, “Object extraction method based on contour evolutions in edge-based object tracking,” *5th Int. Conf. Intel. Tech.*, Houston, Texas, Dec. 2004.

[87] 辻裕之，徳増眞司，高橋裕樹，中嶋正之，“非等方拡散法に基づく空間平滑化フィルタの構築に関する基礎検討”, 信学論 (D-II), vol.J88-D-II, no.6, pp.1024-1034, June 2005.

[88] H. Tsuji, S. Saito, H. Takahashi, and M. Nakajima, “Estimating Object Contours from Binary Edge Images”, IEEE International Conference on Image Processing (ICIP’05), Genova, Italy, Sept. 2005.

[89] 上田剛弘，金子正秀，齋藤隆弘，原島博，“平面の 3 次元配置に基づく動画像の構造利用符号化”，信学論 (D-II), vol.J80-D-II, no.2, pp.487-496, Feb. 1997.

[90] 内田顕貴，田中清，砂押貴之，河尻寛之, “伝送符号量削減を目的とする適応型プレフィルタ”, 画像電子学会誌，vol.32, no.5, pp.589-600, May 2003.

[91] 梅原雅顕，山田光太郎，曲線と曲面 – 微分幾何的アプローチ –，裳華房，2002.

[92] N. Vasconcelos and F. Dufaux, “Pre and Post-filtering for Low-bitrate Video Coding”, Proc. IEEE Int. Conf. Image Process. (ICIP’97), vol.1, pp.291-294, Santa Barbara, CA, Oct. 1997.

[93] J.W.A. Wang and E.H. Adelson, “Representing Moving Images with Layers”, IEEE Trans. Image Process., vol.3, no.5, pp.625-638, Sep. 1994.

[94] J. Weickert, “Coherence-enhancing Diffusion filtering”, Int. J. Comput. Vision, vol. 31, no. 2/3, April 1999.

[95] A.P. Witkin, ”Scale-space Filtering,” International Joint Conference in Artificial Intelligence, pp.1019-1021, 1983.

[96] C. Xu and J.L. Prince, “Snakes, Shapes, and Gradient Vector Flow”, IEEE Trans. Image Process., vol.7, no.3, pp.359-369, March. 1998.

[97] L. Yan, ”Noise Reduction for MPEG Type of CODEC,” IEEE Int. Conf. Acoustics, Speech, and Signal Process. (ICASSP’94), vol.V, pp.429-432, Adelaide, Australia, April 1994.

[98] 八島由幸，沢田克敏，”HDTV 信号のフレーム内／フレーム間適応外挿内挿予測符号化”，信学論 (B), vol.J70-B, no1, pp.96-104, 1987.

[99] 安田浩，渡辺裕，ディジタル画像圧縮の基礎，日経 BP，1996.

[100] Y. You, W. Xu, A. Tannenbaum, and M. Kaveh, “Behavioral Analysis of Anisotropic Diffusion in Image Processing”, IEEE Trans. Image Process., vol.5, no.11, pp.1539-1553, Nov. 1996.

関連研究業績

論文

1. **辻裕之**，徳増眞司，高橋裕樹，中嶋正之，“非等方拡散法に基づく空間平滑化フィルタの構築に関する基礎検討”, 信学論 (D-II), vol.J88-D-II, no.6, pp.1024-1034, June 2005.
2. *Hiroyuki Tsuji*, Shinji Tokumasu, Hiroki Takahashi and Masayuki Nakajima, “Extracting Objects Using Contour Evolutions in Edge-based Object Tracking”, Journal of Advanced Computational Intelligence and Intelligent Informatics (JACIII), vol.10, no.3, May 2006. (to appear)
3. **辻裕之**，徳増眞司，高橋裕樹，中嶋正之，“非等方拡散法に基づく時空間プレフィルタの検討”, 信学論 (D-II), 投稿中.

国際会議

1. *Hiroyuki Tsuji* and Shinji Tokumasu, “A Combined Spatio-temporal Anisotropic Diffusion for MPEG-type Video Compression”, the 6th International Conference on Engineering Design and Automation (EDA’02), pp.753-758, Aug. 2002.
2. *Hiroyuki Tsuji*, Toru Sakatani, Yoshiyuki Yashima and Naoki Kobayashi, “A Nonlinear Spatio-Temporal Diffusion and Its Application to Prefiltering in MPEG-4 Video Coding”, IEEE International Conference on Image Processing (ICIP’02), vol.1, pp.85-88, Sept. 2002.
3. *Hiroyuki Tsuji*, Shinji Tokumasu, Hiroki Takahashi and Masayuki Nakajima, “Spatio-temporal Prefiltering Scheme Based on Anisotropic Diffusion in Low Bitrate Video Coding”, International Workshop on Advanced Imaging Technology (IWAIT’04), pp. 141-145, Jan. 2004.
4. *Hiroyuki Tsuji*, Shinji Tokumasu, Hiroki Takahashi and Masayuki Nakajima, “Object Extraction Method Based on Contour Evolutions in Edge-based Object Tracking”, the 5th International Conference on Intelligent Technologies (Intech’04), Dec. 2004. *(Best Research Paper Award)*
5. *Hiroyuki Tsuji*, Suguru Saito, Hiroki Takahashi and Masayuki Nakajima, “Estimating Object Contours from Binary Edge Images”, IEEE International Conference on Image

Processing (ICIP'05), vol.3, pp.453-456, Sept. 2005.

技術報告

1. **辻裕之**，徳増眞司，高橋裕樹，中嶋正之，“非等方拡散法に基づく空間平滑化フィルタの構築に関する検討”，信学技報，IE2004-19，pp.23-28，May 2004.
2. **辻裕之**，徳増眞司，高橋裕樹，中嶋正之，“低ビットレート映像符号化への適用を目的とした非等方拡散法の時空間３次元への拡張に関する検討”, 信学技報，IE2004-20，pp.1-6, July 2004.

国内大会，シンポジウム

1. **辻裕之**，阪谷徹，八島由幸，小林直樹，“非等方拡散に基づく低ビットレート符号化用プレフィルタの検討”，第 16 回画像符号化シンポジウム (PCSJ2001), pp.109-110, Nov. 2001.
2. **辻裕之**，徳増眞司，“非等方拡散に基づく時空間フィルタの改善に関する一検討”， 第 17 回画像符号化シンポジウム (PCSJ2002)，pp.95-96，Nov. 2002.
3. **辻裕之**，徳増眞司，八島由幸，中嶋正之，“低ビットレート映像符号化の画質改善に関する検討 ーピクチャスキップ抑制のためのプレフィルタ制御法ー”，情報科学技術フォーラム（FIT2003）情報技術レターズ，vol.2，LJ-017，pp.255-257，Sept. 2003.
4. **辻裕之**，徳増眞司，高橋裕樹，中嶋正之，“非等方拡散法に基づく時空間平滑化処理とその動画像符号化への応用”，第 19 回 Nicograph 論文コンテスト　論文集，pp.175-180, Nov. 2003. （**最優秀論文賞**）
5. **辻裕之**，徳増眞司，高橋裕樹，中嶋正之，“不完全なエッジ情報に基づくオブジェクト領域抽出法の検討”, 第８回映像メディア処理シンポジウム（IMPS2003），pp.55-56，Nov. 2003.
6. **辻裕之**，徳増眞司，高橋裕樹，中嶋正之，“非等方拡散法に基づく平滑化フィルタ制御に関する基礎検討”, 電子情報通信学会 総合大会，D-11-35, March 2004.
7. **辻裕之**，徳増眞司，齋藤豪，高橋裕樹，中嶋正之，“２値エッジ画像に基づくオブジェクト輪郭推定法の検討”, 電子情報通信学会 総合大会，D-12-55, March 2005.
8. **辻裕之**，徳増眞司，齋藤豪，高橋裕樹，中嶋正之，“２値エッジ画像に基づくオブジェクト輪郭推定法 ―接ベクトル導出の改善に関する検討―”，情報科学技術フォーラム（FIT2005）一般講演論文集，vol.3，pp.57-59，Sept. 2005.

その他の研究業績

論文

1. Yasuyuki Murai, Koki Matsumura, Hisayuki Tatsumi, *Hiroyuki Tsuji* and Shinji Tokumasu, “Accelerated Genetic Programming for Intelligent Fuzzy Robots”, Journal of Advanced Computational Intelligence & Intelligent Informatics (JACIII), vol.8, no.6, pp.582-590, Nov. 2004.
2. Yasuyuki Murai, *Hiroyuki Tsuji*, Hisayuki Tatsumi and Shinji Tokumasu, “Fast Placement Algorithm for Rectilinear Jigsaw Puzzle”, Journal of Advanced Computational Intelligence & Intelligent Informatics (JACIII), vol.10, no.3, May 2006. (to appear)

国際会議

1. Junichi Nakajima, *Hiroyuki Tsuji*, Yoshiyuki Yashima and Naoki Kobayashi, “Motion vector Re-estimation for Fast Video Transcoding from MPEG-2 to MPEG-4”, the 2nd Workshop and Exhibition on MPEG-4, June 2001.
2. Yasuyuki Murai, *Hiroyuki Tsuji*, Hisayuki Tatsumi and Shinji Tokumasu, “Solution of Rectilinear Jigsaw Puzzle with Pairwise Preference Constraints between Pieces”, the 6th International Conference on Engineering Design and Automation (EDA2002), pp.601-609, Aug. 2002.
3. Hisayuki Tatsumi, Yasuyuki Murai, *Hiroyuki Tsuji*, Shinji Tokumasu and Masahiro Miyagawa, “Logic Circuit Fault Models Detectable by Neural Network Diagnosis”, the 4th International Symposium on Advanced Intelligent Systems (ISIS2003), pp.154-157, Sept. 2003.
4. Yasuyuki Murai, Koki Matsumura, Hisayuki Tatsumi, *Hiroyuki Tsuji* and Shinji Tokumasu, “Introduction and Improvement of Genetic Programming for Intelligent Fuzzy Robots”, the 4th International Symposium on Advanced Intelligent Systems (ISIS2003), pp.388-391, Sept. 2003.
5. Yoshitaka Nagata, *Hiroyuki Tsuji*, Yasuyuki Murai and Shinji Tokumasu, “3D Passage Navigation under Unknown Environments Based on Distance Field Space Model”, the 4th International Symposium on Advanced Intelligent Systems (ISIS2003), pp.500-

503, Sept. 2003.
6. Masatoshi Ohno, *Hiroyuki Tsuji*, Yasuyuki Murai and Shinji Tokumasu,"Fast Placement Algorithm for Rectilinear Jigsaw Puzzle", the 5th International Conference on Intelligent Technologies (Intech'04), Dec. 2004.

技術報告

1. **辻裕之**，古角康一，端山聡，小林直樹，"インタラクティブ映像・音声配信システムの開発"，画像電子学会 研究会予稿 00-01-02, pp.7-10, May 2000.
2. **辻裕之**，阿久津隆史，清水淳，八島由幸，小林直樹，"低レート用 MPEG-2 トランスコーダ"Trampeg"の開発と性能評価"，信学技報 IE2000-16, vol.100, no.179, pp.1-7, July 2000.
3. 阿久津隆史，**辻裕之**，清水淳，八島由幸，小林直樹，"低レート用 MPEG-2 トランスコーダ"Trampeg"における動きベクトル変換方式"，信学技報 IE2000-17, vol.100, no.179, pp.9-14, July 2000.
4. 中嶋淳一，**辻裕之**，八島由幸，小林直樹，"MPEG-2/MPEG-4 トランスコーダにおける適応的再探索の検討"，画像電子学会 研究会予稿 01-02-02, pp.7-10, July 2001.

国内大会，シンポジウム

1. **辻裕之**，阪谷徹，八島由幸，小林直樹，"DCT ドメインにおける MPEG ビデオトランスコードの検討"，第 15 回画像符号化シンポジウム (PCSJ2000)，pp.129-130, Nov. 2000.
2. 中嶋淳一，**辻裕之**，阪谷徹，八島由幸，小林直樹，"MPEG-2/MPEG-4 トランスコーダにおける動きベクトル統合方法の比較"，第 15 回画像符号化シンポジウム (PCSJ2000)，pp.41-42, Nov. 2000.
3. 永田善隆，村井保之，**辻裕之**，徳増眞司，"ピース間の配置制約を考慮した平面ポリオミノ箱詰め問題の解法", 情報処理学会 第 65 回全国大会 講演論文集，vol.1, pp.241-242, March 2003.
4. 村井保之，奥山大輔，**辻裕之**，徳増眞司，巽久行, "電子白板を利用した教育支援システム", 第 11 回全国大学情報教育方法研究発表会 予稿集, pp.142-143, July 2003.
5. 巽久行，村井保之，**辻裕之**，徳増眞司，宮川正弘, "ニューラルネットワーク診断で検出できる論理回路故障モデルの一考察", 情報科学技術フォーラム（FIT2003）一般講演論文集, vol.1, pp.307-308, Sept. 2003.
6. 村井保之，松村幸輝，**辻裕之**，徳増眞司, "進化論的手法に基づくファジィロボットの障害物回避モデル", 情報科学技術フォーラム（FIT2003）一般講演論文集, vol.2, pp.289-290, Sept. 2003.
7. 大野雅俊，**辻裕之**，村井保之，徳増眞司, "可変サンプリングによる静止画像符号化の主観画質向上", 情報科学技術フォーラム（FIT2003）一般講演論文集, vol.3, pp.273-274,

Sept. 2003.
8. 村井保之，巽久行，宮川正弘，**辻裕之**，徳増眞司, “電子白板を用いた弱視学生の教育支援”, 情報科学技術フォーラム（FIT2003）一般講演論文集, vol.4, pp.395-396, Sept. 2003.

付 録A　命題の証明

A.1　PMDの安定性に関する命題 (第2章)

命題 A.1 *Perona-Malik* による差分方程式 *(2.45)* が安定に収束するための必要条件は，安定性パラメータ λ が $0 \leq \lambda \leq 1$ を満たすことである．

[**証明**]　時間方向のサンプル間隔を Δt，水平・垂直方向のサンプル間隔を $\Delta x = \Delta y = h$ とするとき，Perona-Malik の原論文に習って，

$$c_{N_{i,j}} = g(|I^t_{i-1,j} - I^t_{i,j}|), \quad c_{S_{i,j}} = g(|I^t_{i+1,j} - I^t_{i,j}|),$$
$$c_{E_{i,j}} = g(|I^t_{i,j+1} - I^t_{i,j}|), \quad c_{W_{i,j}} = g(|I^t_{i,j-1} - I^t_{i,j}|)$$

とおけば，Perona-Malik の差分方程式は次のように書くことができる．

$$\frac{I^{t+\Delta t}_{i,j} - I^t_{i,j}}{\Delta t} = \frac{c_{N_{i,j}}(I^t_{i-1,j} - I^t_{i,j}) + c_{S_{i,j}}(I^t_{i+1,j} - I^t_{i,j}) + c_{E_{i,j}}(I^t_{i,j+1} - I^t_{i,j}) + c_{W_{i,j}}(I^t_{i,j-1} - I^t_{i,j})}{h^2} \tag{A.1}$$

ここで，$\Delta t = \lambda/|\eta_{\mathbf{s}}|$，$h = 1$ とおけば，式 (2.45) と等価な離散式が得られる．さて，このとき $I_i(t) = (I^t_{i,1}, I^t_{i,2}, ..., I^t_{i,n})^T$ （$i = 1, 2, ...m$）と書くことにすれば，式 (A.1) は次の方程式の時間微分を前進オイラー法で離散化した式に他ならないことがわかる．

$$\frac{\partial I_i}{\partial t} = \begin{cases} (B_i I_i + B_{i+1} I_{i+1})/h^2 & (i = 1) \\ (B_{i-1} I_{i-1} + B_i I_i + B_{i+1} I_{i+1})/h^2 & (i = 2, ..., m-1) \\ (B_{i-1} I_{i-1} + B_i I_i)/h^2 & (i = m) \end{cases} \tag{A.2}$$

ただし，ここでは

$$B_{i,j-1} = \begin{pmatrix} 0 & 0 & 0 & \dots & 0 \\ 0 & c_{N_{i,2}} & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & c_{N_{i,n-1}} & 0 \\ 0 & \dots & 0 & 0 & c_{N_{i,n}} \end{pmatrix},$$

$$B_{i,j+1} = \begin{pmatrix} c_{S_{i,1}} & 0 & 0 & \dots & 0 \\ 0 & c_{S_{i,2}} & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & \dots & c_{S_{i,n-1}} & 0 \\ 0 & 0 & \dots & 0 & 0 \end{pmatrix},$$

$$B_{i,j} = \begin{pmatrix} -D_{i,1} & c_{E_{i,1}} & 0 & \dots & 0 \\ c_{W_{i,2}} & -D_{i,2} & c_{E_{i,2}} & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & c_{W_{i,n-1}} & -D_{i,n-1} & c_{W_{i,n-1}} \\ 0 & \dots & 0 & c_{W_{i,n}} & -D_{i,n} \end{pmatrix},$$

$$D_{i,j} = c_{N_{i,j}} + c_{S_{i,j}} + c_{E_{i,j}} + c_{W_{i,j}}$$

とした．

さて，PMD の差分方程式 (2.45) が安定となるためには，上の連立微分方程式 (A.2) の時間微分に関する前進オイラー法が数値解法として安定であることを保証すればよい．Gershgorin の定理 [27] によれば，連立微分方程式 (A.2) によって定まるヤコビ行列

$$\mathbf{J} = \frac{1}{h^2} \begin{pmatrix} B_{1,1} & B_{1,2} & 0 & \dots & 0 \\ B_{2,1} & B_{2,2} & B_{2,3} & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & B_{m-1,m-2} & B_{m-1,m-1} & B_{m-1,m} \\ 0 & \dots & 0 & B_{m,m-1} & B_{m,m} \end{pmatrix}$$

の固有値は，$-D_{i,j}/h^2$ を中心とする半径 $R = (c_{N_{i,j}} + c_{S_{i,j}} + c_{E_{i,j}} + c_{W_{i,j}})/h^2 = D_{i,j}/h^2$ のディスク内部に存在することが保証される．エッジ停止関数 g の値域が $[0, 1]$ であることを考慮すれば，任意の i, j に対して $0 < D_{i,j} \leq 4$ と評価され，$\mathbf{J}$ のすべての固有値は実数軸上の $-4/h^2$ を中心とした半径 $4/h^2$ のディスクに含まれる．一般にヤコビ行列 $\mathbf{J}$ で与えられる常微分方程式に対する前進オイラー法が安定となるには，$\Delta t\mathbf{J}$ のすべての固有値が複素空間における実数軸上の -1 を中心とする半径 1 の円の内部に入っていなければならない [27]．ゆえに，Δt としては，

$$0 \leq \frac{4\Delta t}{h^2} \leq 1 \quad \Leftrightarrow \quad 0 \leq \Delta t \leq \frac{h^2}{4}$$

を満たすように取らなければならない．式 (2.45) においては，$\Delta t = \lambda/|\eta_{\mathbf{x}}|$，$h = 1$ であることから，

$$0 \leq \frac{\lambda}{|\eta_{\mathbf{x}}|} \leq 1/4 \;\; \Leftrightarrow \;\; 0 \leq \lambda \leq 1$$

を満たすように λ を設定する必要がある． ∎

以上の議論により，PMD の差分方程式 (2.45) の数値解が，時間微分の離散化に伴うノイズの増幅により不安定になるのを回避するためには，安定性パラメータ λ を $0 \leq \lambda \leq 1$ に設定しておく必要がある．このため本論文で扱う PMD では，特に断らない限り，$\lambda = 1$ に固定して議論を進めることとする．