

ベクタ化による人工画像の 高能率符号化

Efficient Coding of Artificial Still Images
by Vectorization

2013年2月

早稲田大学大学院 国際情報通信研究科
国際情報通信学専攻 オーディオビジュアル情報生成技術の研究 II

河 村 圭

目次

第1章	序論	1
1.1	研究の背景	1
1.1.1	画像表現方式の問題	2
1.1.2	ベクタ表現の問題	2
1.2	本論文の目的	2
1.3	本論文の構成	3
第2章	ベクタ変換技術と画像表現方式	6
2.1	まえがき	6
2.2	ベクタ表現と変換技術	6
2.2.1	ベクタ表現の種類	6
2.2.2	標準的なベクタ表現フォーマット	7
2.2.3	多角形近似を経るベクタ変換方式	7
2.2.4	ベジエ曲線の編集性と既存手法の問題点	8
2.3	網点による階調表現と網点の分離技術	10
2.3.1	網点による階調表現	10
2.3.2	網点の検出手法と分離手法, その問題点	10
2.3.3	画像のモデル化	11
2.4	全変動最小化の従来手法	12
2.4.1	画像分解問題	12
2.4.2	画像復元問題	13
2.4.3	従来の離散全変動の定義と問題点	14
2.4.4	グラデーションメッシュ	15
2.5	曲率と曲率スケール空間	16
2.5.1	曲率	16
2.5.2	曲率スケール空間	17
2.6	コミック画像からのメタデータ抽出方式	17
2.6.1	コマと視線遷移情報	17
2.6.2	コマ分割手法の問題点	18
2.6.3	文字切出し手法の問題点	19
2.7	むすび	20

第3章	網点分離による画像分解方式	21
3.1	まえがき	21
3.2	モアレの種類と発生原因	22
3.3	網点分離と階調近似の提案	23
3.3.1	網点中心の検出	24
3.3.2	網点領域の分離	25
3.3.3	線形勾配による近似	26
3.4	低解像度グレースケールへの拡張	26
3.4.1	2値・多値網点分離手法の統合	26
3.4.2	高解像度2値化	27
3.5	提案手法の評価	28
3.5.1	網点分離の特性	28
3.5.2	検出濃度の精度	29
3.5.3	モアレ低減の検証	30
3.5.4	2値・多値網点中心の検出	32
3.5.5	高解像度2値化の特性	32
3.6	むすび	33
第4章	線形勾配成分の抽出方式	35
4.1	まえがき	35
4.2	全変動最小化の高速計算手法	35
4.2.1	離散全変動の定義と凸性	36
4.2.2	劣勾配法の適用	37
4.3	線形勾配領域の抽出	38
4.4	グラデーションメッシュ	39
4.5	提案手法の評価	40
4.5.1	全変動最小化の高速計算手法の評価実験	40
4.5.2	線形勾配領域の抽出実験	45
4.5.3	グラデーションメッシュによる画像符号化	46
4.6	むすび	48
第5章	曲線の不変特徴量と編集性	51
5.1	まえがき	51
5.2	既存のスケール不変曲線特徴量	51
5.2.1	曲率スケール空間	53
5.2.2	Laplacian of Gaussian	54
5.2.3	濃度勾配に基づく曲率	54
5.3	スケール不変曲線特徴量	55
5.3.1	スケール正規化曲率	55

5.3.2	Difference of Curvature	57
5.3.3	スケール不変な曲率記述	58
5.4	曲線の頂点とベジエ曲線の編集性	59
5.4.1	ベクタ変換手法の概要	59
5.4.2	ベジエ曲線の編集性評価手法	59
5.5	提案手法の評価	60
5.5.1	スケール不変性の安定性	61
5.5.2	スケール不変性の頑健性	63
5.5.3	特徴点位置の位置特性	63
5.5.4	曲率と曲線の頂点	63
5.5.5	ベクタ変換の編集性	65
5.6	むすび	69
第 6 章	コミック画像からのメタデータ抽出方式	70
6.1	まえがき	70
6.2	文字領域の抽出	70
6.2.1	コミックにおけるセリフの特徴	70
6.2.2	文字切出し手法	71
6.2.3	文字サイズの自動決定手法	71
6.3	コマ分割手法	72
6.3.1	コミック画像におけるコマの特徴	72
6.3.2	高精度コマ分割手法	73
6.3.3	高速コマ分割手法	75
6.4	実験と考察	77
6.4.1	文字切出しの再現率	77
6.4.2	文字サイズの分布	78
6.4.3	コマ分割精度	78
6.4.4	コマ分割の処理時間	79
6.5	むすび	80
第 7 章	コミック画像符号化	82
7.1	まえがき	82
7.2	コミックの画像特性	82
7.2.1	構成要素	82
7.2.2	画像の入出力形式	84
7.2.3	符号化システムのブロック図	84
7.3	符号量制御	84
7.3.1	スキャン解像度と符号量制御	84
7.3.2	座標解像度の削減	85

7.3.3	直線・曲線の統合	85
7.4	エントロピー符号化	86
7.4.1	係数の分類	86
7.4.2	発生頻度と分布特性	87
7.5	実験と考察	88
7.5.1	座標解像度の削減と直線・曲線の統合	88
7.5.2	エントロピー符号化	89
7.5.3	ファイルサイズとモアレ低減	90
7.6	むすび	92
第 8 章	結論	93
8.1	総括	93
8.2	今後の課題	94
	謝辞	95
	参考文献	96
	図一覧	101
	表一覧	104
	研究業績	105

第1章

序論

1.1 研究の背景

電子ペーパーと呼ばれる高精細出力装置（ディスプレイ）が実用化され、書籍や文書などを閲覧する読書端末が急速に普及している。また、タブレット端末と呼ばれる大画面携帯端末が一般化しつつある。同時に、電子書籍や電子コミックなどのコンテンツ需要が増加している。

2004年に日本国内では電子ペーパーを用いた電子書籍専用の読書端末が相次いで発売された。その後、コンテンツ不足や購入の手間など複数の要因により、これら専用端末は普及することなく市場から姿を消していった。その一方で、第三代携帯電話に搭載された小型ではあるが高精細な液晶パネルが携帯電話と共に普及し、これを用いて電子書籍や電子コミックを閲覧する人が増加した。同時に、“ケータイ電子コミック”という新しいジャンルが勃興していった [1, 2]。2007年には米国でも電子ペーパーを用いた読書端末が発売され、日本とは対照的に大きく普及しつつある。2000年代後半には第三代携帯電話よりも大型の液晶パネルを搭載するスマートフォンが、2010年には大画面のタブレット端末が、全世界で急速に普及し電子書籍の読書端末としても注目されている。

上記の読書端末に搭載された E-ink [3] を代表とする電子ペーパーは、70～100dpi の PC 用モニタ（例えば CRT 表示器など）と比較して 200dpi 程度と解像度が高く、電源を切っても表示内容が保存されるなど消費電力が少ない。また、反射式であるため視野角が広く、強い日差しの下でも閲覧できるなど視認性も高い。しかしながら、階調表現は 4 階調や 16 階調グレースケールなど不十分である。一方、液晶パネルは 250～300ppi 程度と電子ペーパーよりも高い解像度が実用化されている。スマートフォン普及前には解像度や画面サイズが小さいために電子書籍の閲覧端末としての注目度は低かったが、スマートフォンやタブレット端末の普及を機会に画面が大型化・高精細化し、新聞紙面の配信なども開始された。これらの液晶パネルは電子ペーパーと比較して屋外での視認性は悪いが、階調表現やカラー表現に優れている。

このように解像度の異なる高精細出力装置が登場すると、電子書籍や電子コミックを閲覧するためにコンテンツの解像度変換が必須となる。文字はフォントが整備されているため、解像度に合わせた最適な閲覧ができる。一方、文字以外を含むコミックのようなコンテンツ

では、単純な解像度変換によってアーティファクトが発生して画像品質が大きく低下する。従って、電子コミックを効率よく蓄積・配信するためにはデータ圧縮（高能率符号化）と様々な解像度に対してワンソース・マルチユースできることが重要である。

端末が持つ出力装置の特性（特に解像度）に合わせた表示が可能な画像表現方式として、ベクタ表現がある。例えば、アウトラインフォントといわれるベクタ表現に基づくフォントを利用すると、拡大してもジャギーノイズが発生せず、さらにPCディスプレイのような低解像度な出力装置（例えば72ppi）とプリンタのような高解像度な出力装置（例えば600dpi）の両方に対応できる。従って、ベクタ表現のコンテンツ需要が今後増大していくと考えられる。

1.1.1 画像表現方式の問題

既存のコンテンツは大部分がラスタ表現で制作・蓄積・配信されている。さらに、ベクタ表現のコンテンツは、制作可能なクリエイタが少ない。

そこで、既存のラスタ表現コンテンツを自動的にベクタ表現に変換する技術が求められている。このような技術を活用すると、紙面をスキャナやデジタルカメラによりラスタ表現コンテンツとして取得（デジタル化）して、さらにベクタ表現に変換できる。

1.1.2 ベクタ表現の問題

ベクタ表現に向かない、もしくはラスタ表現の特性を利用した技術が使われている紙面やコンテンツも存在している。例えば、オフセット印刷において中間調を表現する網点が挙げられる。網点を単純にベクタ表現に変換すると、ファイルサイズが極めて大きくなるという問題が生じる。さらに、解像度変換を適用するとモアレとよばれるアーティファクトが生じ、画像品質が著しく低下する。

別の観点で、ベクタ化に適した画像特性が限定的であるという問題がある。すなわち、その仕組みに起因して600dpi~1200dpiという高解像度2値形式が求められる。しかしながら、スキャナやデジタルカメラによる取得においては、8bit階調や24bit階調の画像を取得でき、100dpi~200dpiと低解像度グレースケール画像の方が取得時間が大幅に短縮できる。このような低解像度グレースケール画像をベクタ化する技術はまだ確立されていない。加えて、JPEGやJPEG2000のようにラスタ表現のロッシェー符号化により蓄積されているコンテンツでもベクタ化の需要は大きい。

1.2 本論文の目的

高精細表示端末の普及拡大により、高精細コンテンツの需要増大が予想されている。本論文では、コミックに代表される紙面を含む既存の人工画像を対象に、ベクタ化を中核とした高能率符号化方式を考察する。さらに、その基盤となるベクタ変換方式及び入力画像分解方式について研究する。

処理レベル	コミック画像	アニメーション画像
ベクタ化のための画像モデル化	第3章 網点分離による 画像分解 [2.5, 2.6]	第4章 線形勾配成分の抽出 方式 [1.3]
低レベル解析	第5章 曲線の不変特徴量と 編集性 [1.2, 2.2]	
高レベル分析	第6章 コミック画像からの メタデータ抽出方式	
ベクタ化による 画像符号化	第7章 コミック画像符号化 [2.5, 2.6]	

図 1-1: 本論文の各章の位置づけ．括弧内は業績番号

先ほど述べたようにラスタ表現の特性を活用した網点を含む画像は、単純なベクタ表現に向かない。そこで、網点領域を分離し、網点画像とそれ以外の画像分解する手法を提案する。さらに、分離した網点画像については網点が存在していた領域と領域中の階調変化に復元してからベクタ変換する手法を提案する。

アニメーション画像やイラストのように、均一色・均一輝度領域や線形輝度勾配領域（線形グラデーション領域）を多く含む画像がある。このような画像に含まれる均一輝度領域や線形輝度勾配領域はベクタ表現に適しているため、これらの成分とテクスチャ成分に画像を分解する手法を提案する。

これら2種類の分解方式により得られるベクタ表現に適した成分に対して、実際にベクタ変換を行う。特に2次利用においては自動変換により得られたベクタ表現（ベジエ曲線）の調整、修正が要求されることが多い。そこで、再編集が容易なベクタ表現を得るための改善手法を提案する。さらに、ベクタ表現における特徴量に着目し、スケール不変特徴量を提案する。これを用いて、編集性の定量的に評価する手法を提案する。

対象コンテンツの一つであるコミック画像には、セリフなどの文字やコマレイアウトなど利用価値の高いメタデータが存在している。これを自動取得する手法を検討する。また、本論文で提案している個々の要素技術を統合して高能率符号化を具現化する仕組みを提案する。なお、符号量制御やエントロピー符号化についても述べる。

1.3 本論文の構成

以下に本章以降の構成を、それぞれの位置づけを図 1-1 に示す。

第 1 章「序論」は本章であり、本論文の背景と目的、研究の概要について述べた。

第2章「ベクタ変換技術と画像表現方式」では、本論文で前提とする技術について整理し、問題点を明らかにする。まず、平面曲線に限定してラスタ表現をベクタ表現に変換する基本的な方式と、すでに広く利用されているベクタ表現のファイルフォーマットを説明する。また、既存のベクタ変換により生成されるベジエ曲線は編集しにくいという課題を明らかにする。次に、ラスタ表現において濃淡を表現する網点手法と、これを画像中から分離する既存手法を説明し、高解像度2値画像からは分離できないという原理的な課題を明らかにする。さらに濃淡画像（グレースケール画像やカラー画像）からベクタ表現に適した均一輝度成分を取得する全変動最小化のフレームワークを説明し、離散化が不十分であることと反復処理が重いという問題点を明らかにする。再びベクタ表現のうち、平面曲線の特徴量を整理し、輪郭素片には適用できないことと特徴点の探索処理が必要であることを述べる。また、本論文の対象コンテンツの代表例であるコミック画像について、含まれているメタデータを整理し、その抽出方法をまとめる。

第3章「網点分離による画像分解方式」では、スキャナから取得したコミックのデジタルデータ（画像）から網点を分離する手法、及び分離した網点が存在していた領域の特定、領域内を連続階調で近似する手法を提案する。主に2値画像を対象とする画像分解手法を述べる。まず、網点を含む2値画像の縮小によってアーティファクトが生じる原因が、網点の周期的な配置にあることを確認する。次に、網点の正確な検出のために画像の x 軸、 y 軸にそって輝度変化の周期を検出し、これを統合して画像平面における周期を特定する。得られた周期を利用して、アーティファクトを発生させることなく網点を分離する。さらに、既存の階調画像における網点分離手法と統合して、2値画像とグレースケール画像に対応した網点分離手法を述べる。分離した網点にモルフォロジ処理を適用して網点領域を取得する。また、先の周期を利用して領域内の輝度勾配を線形輝度勾配で近似した場合のパラメータを取得する。最後に、提案手法により得られた線形輝度勾配パラメータを実験により評価する。

第4章「線形勾配成分の抽出方式」では、高能率符号化を実現するために二つの画像モデルを提案する。主にグレースケール画像を対象とする画像分解手法を述べる。一つは入力画像をベクタ表現可能な線形グラデーション領域とラスタ表現に適しているテクスチャ成分へ分離する画像モデルであり、もう一つは入力画像をグラデーションメッシュとテクスチャ成分に分離する画像モデルである。前者のモデルに応じた線形グラデーション領域を取得するために、画像に全変動最小化のフレームワークを適用し、高速計算手法を提案する。分離性能と処理時間について従来手法と比較評価を行う。さらに、提案したモデルによるベクタ表現可能な領域の割合について評価を行う。後者のモデルについては、目的関数を最小化するための2種類の手法について、近似性能と計算時間の評価を行う。

第5章「曲線の不変特徴量と編集性」では、ベクタ表現で利用される平面曲線に対して、スケール不変な特徴量を提案する。この特徴量を用いて、ベジエ曲線の編集性を定量的に評価する手法を示し、すでに我々が提案しているベクタ変換手法は編集性が高いことを示す。具体的には、離散パラメトリック曲線の曲率計算にはスケールの考慮が必須であるため、これを正規化するスケール正規化曲率を検討する。さらに、ベクタ表現の構成要素である通過点・制御点の配置と編集しやすさの関係について整理する。これまで、曲線の「頂点」とい

う主観的な特徴量を用いて、頂点とベクタ表現の通過点が一致しているほど編集しやすいことを示し、編集性を改善するベクタ変換手法を提案してきた。スケール正規化曲率を用いることで、曲線の頂点を定量的に検出可能となり、さらにベクタ変換により得られたベジエ曲線の定量的な評価実験を行う。

第6章「コミック画像からのメタデータ抽出方式」では、対象コンテンツをコミックに限定して、コミックから取得可能かつコンテンツ配信・閲覧の助けとなるメタデータとその取得手法について整理する。まず、コミック画像に文字やコマ配置など利用価値の高いメタデータが含まれることを指摘する。次に、これらのメタデータが自動的に取得できることを明らかにする。すなわち、手書き文字認識技術を応用して、文字の特徴を用いて切出す手法を提案する。また、画像分割に基づくコマ閲覧順序取得手法を提案する。

第7章「コミック画像符号化」では、コンテンツ配信に必要な符号化のフレームワークについてまとめる。コミック画像符号化では本論文で提案してきた個々の要素技術を統合するフレームワークを述べる。特に、網点分離処理やベクタ変換をはじめとする符号化技術と、コミック画像から取得可能なメタデータを統合する。さらに、高能率符号化の実用化に必須となる符号量制御の実現方法を提案する。また、エントロピー符号化について検討する。

第8章「結論」では、本論文で得られた成果を総括し、結論を述べる。

第2章

ベクタ変換技術と画像表現方式

2.1 まえがき

本章ではベクタ化による人工画像の高能率符号化に関する従来手法や関連手法を整理する。ベクタ化に関して、本論文で対象とするベクタ表現の種類とその特徴を明確にする。また、ベクタ表現において曲線やメッシュを表すベジエ曲線の特徴をまとめる。このベジエ曲線の取得を目標とする既存のベクタ変換手法について、その問題点を整理する。次に、人工画像のうちコミック画像に関して、中間調を実現する網点の特性や、既存の網点分離手法の問題点を整理する。また、人工画像のうちアニメーション画像に関して、既存の画像分解問題と画像復元問題を述べ、ベクタ変換を前提としたモデル化の観点から課題を整理する。さらに画像のモデル化と変換とを同時に実現するグラデーションメッシュ表現を紹介し、従来変換手法と課題を述べる。これらのモデル化を前提として、ベクタ変換後のベジエ曲線の曲率に関する特性を整理する。さらに、ベジエ曲線をオペレータが手動で編集するという観点から、ベジエ曲線の編集性についてまとめる。続いて、コミック画像から取得できるメタデータの自動抽出手法を説明し、問題点を明らかにする。

2.2 ベクタ表現と変換技術

2.2.1 ベクタ表現の種類

ベクタ表現とは、直線や曲線、塗りつぶしを用いて画像を描画する表現である。直線は始点と終点、線の太さを指定する。曲線には2次ベジエ曲線、3次ベジエ曲線などがよく利用される。これらの直線や曲線を用いて閉領域が構成される。この閉領域は特定の色やグラデーションで塗りつぶし可能となる。

グラデーションの種類としては線形グラデーションや放射状グラデーションがよく利用されている。より複雑なグラデーションとして、グラデーションメッシュ (Gradient Mesh) というベクタ表現手法がある。これは、メッシュの中をグラデーションで補間する手法である。この手法の登場により、これまでイラストのような人工画像に限られていたベクタ表現が、人物や風景のような自然画像を表現する手段としても現実的になった。

これらのベクタ表現は、特徴点だけを記述し、特徴点間や領域内は描画時に決定されるた

め、ディスプレイの解像度に合わせた描画が可能なる。同時に、適切なアンチエイリアシング処理によって、ジャギーの発生が低減可能となる。また、画像を拡大や縮小しても、再計算すれば最適な画像が取得できるため、階調についてもシャープな線や連続的な塗りつぶし効果が得られる。

2.2.2 標準的なベクタ表現フォーマット

ベクタ表現方式として PostScript とこれをベースとする Encapsulate PostScript (EPS) や Portable Document Format (PDF) が標準的に利用されている。特に PDF は国際標準化されている。

一方、PDF よりも早くに標準化されたベクタ表現として、Scalable Vector Graphic (SVG) が挙げられる [4]。SVG は XML ベースの 2D ベクタ表現記述言語である。2001 年 9 月に W3C 勧告として公開された。

また、現在最も広く普及しているベクタファイル形式として、Flash が挙げられる [5]。Flash は Macromedia 社 (現、Adobe 社) が開発した、音声やベクタ表現のアニメーションを組み合わせて Web コンテンツを作成するソフト、または作成されたコンテンツである。Flash ファイルを閲覧するためには、Web ブラウザに専用のプラグイン “Flash Player” をインストールしておく必要があるが、標準でインストールされているため普及率は高い。内部ファイル形式は公開されており、無償のオーサリングツールも開発されている [6]。

これらのファイル形式では、図形を塗りつぶす際に線形と放射状のグラデーションを利用できる。線形グラデーションでは、まず、矩形領域に対し左端から数点、アルファチャンネルを含む色を指定する。指定された色の間が線形補間される。次に、適用したい図形を含む矩形領域へのアフィン変換行列を指定する。実際に塗りつぶされる領域は図形の内部である。

なお、PostScript ベースでは、3 次ベジエ曲線を基本曲線としている。一方、Flash では 2 次ベジエ曲線を基本曲線としている。SVG はどちらも曲線も表現できる。本論文では 3 次ベジエ曲線を基本曲線としている。

2.2.3 多角形近似を経るベクタ変換方式

Selinger らは、グラフ理論に基づいて辺の数が最小となる多角形近似とベクタ変換を提案している [7]。入力された 2 値画像を輪郭パスに分解して、パスと辺の誤差が $1/2$ 画素以下とする制約条件を満たし、全ての点近傍を通過して辺の本数が最小となる多角形を得るという最短経路問題に帰着させる。さらに辺の数が同じ場合には、累積誤差が最小となる多角形を採用する。次に、多角形の頂点ごとにカーブ判定を行い、カーブの場合は辺の中点を固定点としてその頂点を 3 次ベジエ曲線で置換する。コーナーの場合には多角形のままとする。ベジエ曲線もパスとの誤差が $1/2$ 画素以下になるように置換する。最後に曲線最適化を行う。最適化とは、連続する複数の曲線を一つの曲線に統合する処理である。この処理により、近似精度を変えることなく通過点を減らせる。

曲線を統合する方針は以下の通りとなる。まず、連続する曲線だけを統合し、直線は含め

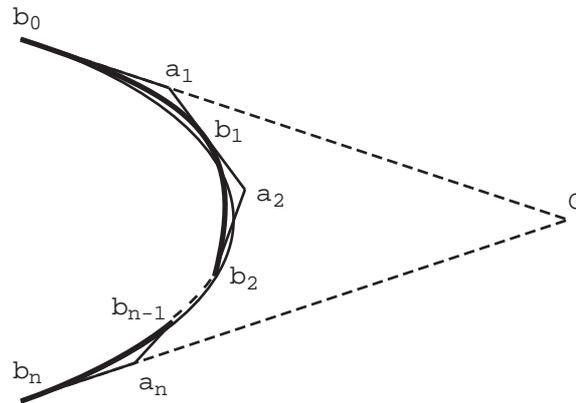


図 2-1: 従来手法における曲線最適化 (自発表 [8] より引用)

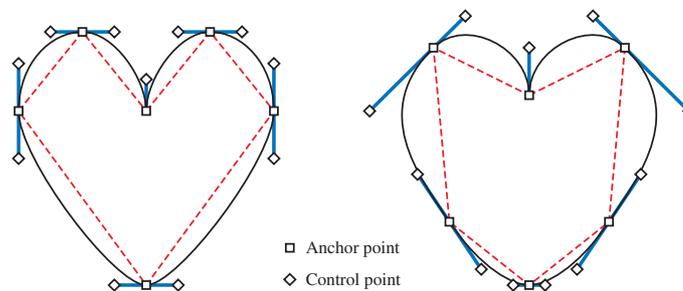


図 2-2: ベジエ曲線の描画方針．左は曲線の頂点に通過点，右は曲線の両端に通過点 (自発表 [10] より引用)

ない．次に，凸方向が一致する曲線だけを統合する．ただし，統合する曲線群の方向変化は180度未満とする．図 2-1 に具体例を示す． a_i は多角形の頂点を示し， b_i は辺の中点を示す． b_0 から b_n まで統合し，点 O を新しい頂点と見なした1本のベジエ曲線を得る．太線は統合前のベジエ曲線，細線は統合後のベジエ曲線を示す．

2.2.4 ベジエ曲線の編集性と既存手法の問題点

Medioni らは，B スプライン曲線の当てはめと，そこで得られる曲率を元に角 (Corner) を検出する手法を提案している [9]．これは曲率の計算に必用なデジタル曲線の微分を，B スプライン関数を用いて実現していることになる．この手法は曲率だけを基準にベクタ変換を実現しているため，通過点の個数が増大してしまう．また，パラメータの設定が困難なために変換精度も低いという問題がある．

Selinger らは，多角形近似と頂点の曲線近似によるベクタ変換を提案している [7]．この手法は，高い変換性能を有しているが，曲率の小さな位置に通過点を置くため，曲線操作が困難になるという問題がある．

ベジエ曲線を描画するには，通過点を曲線の頂点，または両端に置くという2種類の方針が知られている．前者は，描こうとするベジエ曲線の頂点に通過点を置き，制御点を頂点の

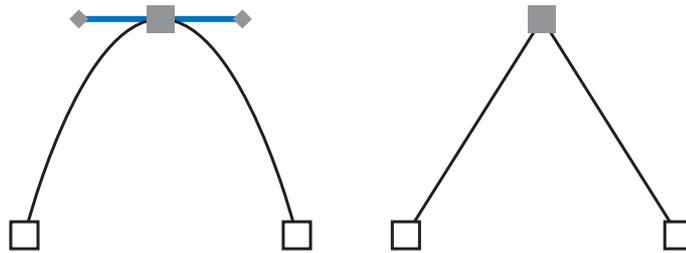


図 2-3: 滑らかな接続（左）と角としての接続（右）の例（自発表 [10] より引用）

左右に置いて曲線のふくらみを調節する．この方針はイラストなどを創作する場合に用いられることが多い．後者は，描こうとするベジエ曲線の両端に通過点を置き，それぞれの通過点の近傍に制御点を置いて曲線のふくらみを調節する．この方針は，線をベジエ曲線でなぞるときに用いられることが多い．図 2-2 にハート形を二つの方針によって描いた結果を示す．Medioni らの手法は前者であり，Selinger らの手法は後者である．前者の場合，曲線の膨らみの変更やカーブとコーナーの相互変換が容易という編集における有意性がある．

曲線のふくらみを制御するユーザインターフェースとしては，制御点を操作する方法と，曲線そのものを操作する方法がある．

前者を用いてベジエ曲線のふくらみを調節する際，もし通過点が曲線の頂点にあればユーザはベジエ曲線を滑らかに接続したままふくらみを調節できる．しかし，曲線の両端に通過点がある場合，滑らかな接続を破棄して制御点を移動する手法，滑らかな接続を維持するために隣り合う曲線にも影響を及ぼして制御点を移動する手法，曲線上の頂点部分に新たに通過点を追加する手法のいずれかを取らざるを得ない．いずれの手法も最初から頂点に通過点がある手法に比べて操作性が悪い．

後者のように曲線そのものを操作する場合には，滑らかな接続を維持するために二つの制御点が連動して移動するため，曲線操作の自由度が低下する．これは通過点の場所によらない．

さらに，頂点を滑らかに接続するか，角として接続するかを切り替えたい場合を考える．もし通過点が曲線の頂点にあれば，ユーザは単に通過点に隣接する制御点を削除すれば角になり，通過点に対して対象に制御点を置けば滑らかな接続になる．しかし，通過点が曲線の頂点ではなく両端にあれば，通過点を頂点付近に追加する必要がある．図 2-3 に頂点の滑らかな接続と角としての接続の切り替えを示す．

以上のように，ベジエ曲線で線をなぞり，その後編集を行わない場合は通過点の位置をどこに置いてもいい．しかし，ベジエ曲線により創作を行う場合には曲線の頂点に通過点を置く方が利便性が高い．

2.3 網点による階調表現と網点の分離技術

2.3.1 網点による階調表現

印刷，とくにオフセット印刷では白黒印刷しかできない．そのため，中間調を表現するために網点と呼ばれるテクスチャを利用する．点を周期的に配置し，さらに点の大きさを変えることで濃淡を表現する方式である．

最適な網構造を決定する研究や，網点画像から元の連続階調画像を復元する研究などがされてきた．また，効率的な伝送を実現するために，画像符号化方式として JBIG2 が標準化されている．

しかし，この符号化方式を復号する閲覧ソフトは普及していない．網点を含む画像であっても，低解像度グレースケール画像に変換して適当なローパスフィルタを適用した後，自然画像の符号化方式 JPEG 等を用いて蓄積，配信されることが多い．

2.3.2 網点の検出手法と分離手法，その問題点

ここでは，モアレ低減を実現するために，グレースケール画像において網点領域を分離する従来手法をまとめる．

例えば，ファクシミリやデジタル複写機における画像符号化の基礎として，網点周期を推定する相関法，パターンマッチング法などが提案されている [11–13]．また，連続階調用の画像処理を施す前処理として，網点画像から階調を復元するルックアップテーブル法，フィルタ法，射影法などが提案されている [14]．

一般に，網点写真をサンプリングすると，網点周期に相当する間隔で輝度が極大値，極小値を示す性質がある．さらに，2次元信号だけでなく1次元信号においても周期的に濃度が変化する．

上野らは，1次元信号における上記の網点の性質を利用して2値化する手法を提案している [15]．まず，画像の主走査方向の輝度について，輝度の増減を順次調べて上昇から下降または，下降から上昇に変化する画素（極点）を抽出する．次に，極点同士の間隔から網点周期を検出する．最後に，1周期ごとに輝度を平滑化して濃度を得る．ただし，主走査方向の1ラインだけで平滑化を行うと，濃度変化の小さな領域や粗い網点領域でモアレが発生する．そこで，副走査方向に4ラインを一括処理して，モアレを低減する．

大内らは，ブロック単位に網点領域の性質を利用して像域分離を行う手法を提案している [16]．まず， 3×3 画素ブロック内で最大，または最小の輝度を有し，かつ，ブロックの中央にある画素（ピーク画素）を検出する．次に，ピーク画素の有無により仮網点領域を検出・補正する．さらに，黒，白画素の連続性を検出してエッジ領域を検出する．最後に，仮網点領域とエッジ領域の両者の結果から，文字領域と網点領域を分離する．

これらの手法は，グレースケール画像に含まれる網点領域を分離する観点からは精度良く実現可能となる．しかし，網点写真は本来2値画像であり，さらに網点の輝度は一定となる．従って，スキャン解像度が高くなるにつれ極点やピーク画素を定義するのが困難になる．ま

た、濃度が高い領域や低い領域では網点周期を考慮しなくてもモアレが生じにくいいため、これらの領域の分離について十分考慮されていないなどの問題がある。

ところで、2値画像は白黒の境界線によって記述できるため、容易にベクタ表現に変換される。網点のようなラスタ表現特有の手法を含む画像を対象にして、境界線だけを用いて網点がベクタ変換すると、符号化効率が悪く、解像度変換時に著しく画質が劣化する。すなわち、網点領域を構成するそれぞれの点について、境界線をベクタ表現に変換すると非常に大きな符号量になる。また、解像度変換としてよく知られたバイキュービック法などを網点領域に適用すると、主観品質の劣化度の大きいモアレが生じる場合がある。

2.3.3 画像のモデル化

グレースケール画像をベクタ表現に変換するには、輝度の境界線情報が必要となる。2値画像の場合には白から黒に変化する画素の境界を輝度の境界線とみなせる。しかし、グレースケール画像の場合には、輝度の境界を決定する手法が画像処理の目的に応じて数多く提案されている。

グレースケール画像の2値化方式は文字認識の分野でも取り組まれている。光学文字認識では、300dpi～400dpi程度の解像度の2値画像が必要となる。藤本、鎌田等は200dpi未満のグレースケール画像を対象として、高精度な2値化手法を提案している [17,18]。本手法は、大津2値化を用いる文字線近傍抽出、線形補間による高解像度化、局所的2値化により構成されている。本方式は、文字認識率の向上には寄与するが、ベクタ表現に変換後の画像特徴や画質評価が不十分である。

同様に領域の明確化が求められる画像を対象とした先行研究として、SIC (Segmented Image Coding) [19] や MRC (Mixed Raster Content) [20] が挙げられる。

SICは衛星画像や自然画像に対して、輝度変化の少ない領域に分割し、その輪郭線とテクスチャに分解するモデルを採用している。輪郭線はChain Codeによりロスレス符号化する。また、MRCは文書画像と自然画像が混在する画像に対して、前景、背景、マスク画像に分離するモデルを採用している。マスク画像はJBIGなどの2値画像符号化方式で、マスク以外の画像は適度に縮小され、JPEGやJPEG2000などのグレースケール画像符号化方式が適用される。

これらのモデルでは、入力画像に含まれるアンチエイリアシング処理の扱いが十分に考慮されていない。そもそも低解像度画像においてこの処理は解像感の向上に必要である。これらの画像に対して整数画素精度の分割や分離を行うと、ジャギーやハロなどの副作用が生じる。さらに、1画素に満たない細線が欠落する。その結果、主観品質が著しく損なわれるという問題がある。また、輪郭線やマスク画像に歪みが許容されないため、解像度変換や符号量制御に関するスケラビリティが不十分である。

画像から強エッジと弱エッジを抽出し、エッジの中心線を符号化するスケッチコーディングがCarlssonによって提案されている [21]。本方式は低解像度グレースケール画像を対象としており、エッジもラスタ表現のままとなる。従って、本手法は解像度変換が適用できず、

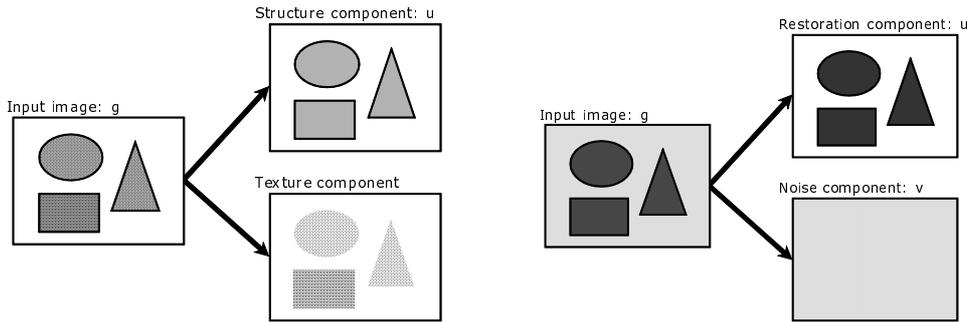


図 2-4: (a) Image Decomposition Problem and (b) Image Restoration. (自発表 [22] より引用)

本論文で想定するベクタ表現とはいえない。

2.4 全変動最小化の従来手法

2.4.1 画像分解問題

画像処理分野で全変動最小化 (Total Variation Minimization) を利用する問題として、骨格成分の取得が挙げられる。そこで、画像分解に基づく手法と画像復元に基づく手法の概要と、高速計算を実現するにあたっての問題点を述べる。

画像分解の目的は、図 2-4 (a) に示すように入力画像 g からエッジを保存しながらもノイズとテクスチャを除去した画像 u を得ることである。この画像 u を骨格成分と呼ぶ。Chambolle らは離散全変動 $J_{\text{tv}2}(u)$ を含む目的関数 (ROF モデル)

$$\min_{u \in X} \frac{\|u - g\|^2}{2\lambda} + J_{\text{tv}2}(u) \quad (2.1)$$

の双対問題を導出して、これを解く手法を提案した [23]。ここで、与えられる定数は入力画像 g と利用者によって選択された重み λ である。また、 X はユークリッド空間である。式 (2.1) のオイラー方程式を $w = (g - u)/\lambda$ で置き換えて

$$w - g/\lambda + \frac{1}{\lambda} \partial J_{\text{tv}2}^*(w) \ni 0 \quad (2.2)$$

と書き直すと、 w について

$$\min_{w \in X} \frac{\|w - g/\lambda\|^2}{2} + \frac{1}{\lambda} J_{\text{tv}2}^*(w) \quad (2.3)$$

という式 (2.1) に双対な最小化問題を得る。なお、 $J_{\text{tv}2}^*$ は $J_{\text{tv}2}$ の共役関数である。

ここで、 w はその定義より、入力画像から骨格成分を除去した振動成分となる。すなわち、双対問題では振動成分を求める問題が導出される。 $J_{\text{tv}2}^*$ はユークリッド空間の部分空間 K を規定しており、この空間は

$$\{\text{div } p : p \in Y, |p_{i,j}| \leq 1, \forall i, j = 1, \dots, N\} \quad (2.4)$$

で与えられる． Y は $X \times X$ のベクタを表している．従って，振動成分 w を求めるのではなく，その原関数となる p を求めることになる．

p は定義より画像の 2 倍の要素数になり，ラグランジュの未定乗数法と半陰的最急降下法を用いて求められる． $\tau > 0$ を選び， $p^0 = 0$ として，任意の $n \geq 0$ について p^{n+1} を計算する反復処理は

$$p_{i,j}^{n+1} = \frac{p_{i,j}^n + \tau(\nabla(\operatorname{div} p^n - g/\lambda))_{i,j}}{1 + \tau|(\nabla(\operatorname{div} p^n - g/\lambda))_{i,j}|} \quad (2.5)$$

と与えられる．なお，離散全変動 $J_{\text{tv}2}$ ， div 演算子， ∇ 演算子，それぞれの定義と計算方法は第 2.4.3 節で述べる．

この手法の優位性は必ず収束値が存在することであり，その効率と安定性が保証されている．しかし，双対問題を用いる手法には三つの問題がある．第一に骨格成分 u が直接計算されるのではなく，振動成分 w （実際には原関数である p ）が計算されることである．第二に原関数 p の次元数は入力画像の次元数の 2 倍になることである．最後に半陰的な反復解法であるため，各項の役割が陽でないことである．

2.4.2 画像復元問題

画像復元の目的は， $g = Lu + v$ で与えられる画像劣化モデルが与えられた場合に，図 2-4 (b) に示すようにノイズ除去と原画像復元を同時に行うことである．ここで， g は入力画像， u は復元画像， v は加法性ノイズ成分， L は劣化フィルタを表す既知の線形演算子である．本論文では，劣化フィルタが全帯域フィルタであり，ノイズとテクスチャとがまとめてノイズであると想定する．このとき，画像復元による復元画像 u の取得は，画像分解により骨格成分の取得と同じになる．すなわち，復元画像 u が骨格成分となる．

Combettes らは，ノイズエネルギーが δ 以下であるという制約条件付き最小二乗問題と全変動を画像の分類（衛星写真や顔画像など）から決定される値にする凸最適化問題の二つを交互に解く手法を提案した [24]．まず， u 成分の統計的仮定として，以下の閉凸制約集合を構成する．

$$S = \{z \in X \mid \|Lz - g\|^2 \leq \delta\} \quad (2.6)$$

さらに全変動を最適な正則条件として採用する．画像復元は以下の制約条件付き最小二乗問題となる．

$$\text{Find } u^* \in S \text{ such that } J_{\text{tv}2}(u^*) = \inf J_{\text{tv}2}(S)$$

ここで， $\inf J_{\text{tv}2}(S)$ は画像の分類から決定されるスカラー値である．

任意の n について， u_{n+1} を計算する反復処理は

$$u_{n+1} = P_S \left(u_n + (\alpha_n - J_{\text{tv}2}(u_n)) \frac{\partial J_{\text{tv}2}(u_n)}{\|\partial J_{\text{tv}2}(u_n)\|^2} \right) \quad (2.7)$$

と与えられる． P_S は制約条件付き最小二乗問題に相当する射影関数であり，FFT と逆 FFT を含む手法で計算される． P_S の中の項は全変動を画像固有の値にする凸最適化問題であり，劣勾配法により計算される． α_n の更新式は文献 [24] による．なお，全変動の定義と計算方法は第 2.4.3 節で述べるが，画像分解問題におけるそれと同じである．

Combettes らの手法は骨格成分を直接求める点が，Chambolle らの手法と異なる．その結果，劣勾配法を含む反復処理において，骨格成分 u の変化が陽になっている．しかし，この手法には二つの問題がある．第一に，収束するまでの反復回数が Chambolle らの手法と比べて多いことである．第二に，1 回の反復処理ごとに FFT を含む射影関数を計算するため，計算コストが高いことである．

2.4.3 従来の離散全変動の定義と問題点

ここで，Chambolle らの手法と Combettes らの手法で共通となる離散全変動の定義と計算方法について述べる．画像は $N \times N$ の 2 次元行列であるとする． X はユークリッド空間 $\mathbb{R}^{N \times N}$ ， Y は $X \times X$ のベクタを表す．また， $y = (y_1, y_2) \in \mathbb{R}^2$ について， $|y| = \sqrt{y_1^2 + y_2^2}$ とする．

離散全変動を定義するために，線形の離散勾配演算子を導入する． $u \in X$ について勾配 $\nabla u \in Y$ は，

$$(\nabla u)_{i,j} = ((\nabla u)_{i,j}^1, (\nabla u)_{i,j}^2) \quad (2.8)$$

$$(\nabla u)_{i,j}^1 = \begin{cases} u_{i+1,j} - u_{i,j} & \text{if } i < N, \\ 0 & \text{if } i = N, \end{cases} \quad (2.9)$$

$$(\nabla u)_{i,j}^2 = \begin{cases} u_{i,j+1} - u_{i,j} & \text{if } j < N, \\ 0 & \text{if } j = N \end{cases} \quad (2.10)$$

となる．ただし， $i, j = 1, \dots, N$ とする．

離散発散 $\text{div} : Y \rightarrow X$ を連続系のアナロジーを用いて $\text{div} = -\nabla^*$ として定義すると， $p \in Y$ と $u \in X$ について $\langle -\text{div } p, u \rangle_X = \langle p, \nabla u \rangle_Y$ となる．このことは div が $p = (p^1, p^2) \in Y$ について，

$$\begin{aligned} (\text{div } p)_{i,j} = & \begin{cases} p_{i,j}^1 - p_{i-1,j}^1 & \text{if } 1 < i < N, \\ p_{i,j}^1 & \text{if } i = 1, \\ -p_{i-1,j}^1 & \text{if } i = N, \end{cases} \\ & + \begin{cases} p_{i,j}^2 - p_{i,j-1}^2 & \text{if } 1 < j < N, \\ p_{i,j}^2 & \text{if } j = 1, \\ -p_{i,j-1}^2 & \text{if } j = N \end{cases} \end{aligned} \quad (2.11)$$

となり，容易に確かめられる．

従って， u の離散全変動は

$$J_{\text{tv}2}(u) = \sum_{1 \leq i, j \leq N} |(\nabla u)_{i,j}| \quad (2.12)$$

と定義される．以下では， $J_{\text{tv}2}$ を L^2 全変動と呼ぶことにする．

この全変動の離散化には，劣微分係数が輝度値に依存するという問題が生じる．本来， L^1 ノルムとして全変動が定義されていたにもかかわらず，2次元へ拡張される際に L^2 ノルムが導入されたことに起因する．さらに，劣微分係数の計算には平方根計算が必須となり，計算コストがやや高くなる．

2.4.4 グラデーションメッシュ

画像をメッシュに分割し，それぞれの内部をグラデーションのベクタ表現に変換する手法が提案されている．グラデーションメッシュは不連続境界も精度良く表現できるため，画像全体を領域分割することなく符号化可能となる．商用ベクタ画像描画ソフトウェアにおけるグラデーションメッシュの仕様は明らかにされていないが，本論文では Sun らの仕様を土台とする [25]．また，彼らはインタラクティブな操作を伴うがグラデーションメッシュに変換する手法として，非線形最小二乗問題に帰結させて解く手法を提案している．

各メッシュは Ferguson Patch と呼ばれるパラメトリック平面によって構成されている．頂点の座標と，両辺の頂点における接線ベクタ，RGB 三原色がパラメータとして定義される．これを以下ではメッシュ点と呼ぶ．メッシュとメッシュ点の位置関係を図 2-5 に示す．また，媒介変数 u, v を用いた座標 m と色 f は

$$m(u, v) = VCQ_m C^t U^t, \quad f(u, v) = VCQ_c C^t U^t \quad (2.13)$$

で表される．ここで，

$$U = \begin{bmatrix} 1 & u & u^2 & u^3 \end{bmatrix}, V = \begin{bmatrix} 1 & v & v^2 & v^3 \end{bmatrix},$$

$$C = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -3 & 3 & -2 & -1 \\ 2 & -2 & 1 & 1 \end{bmatrix}, Q_m = \begin{bmatrix} m^0 & m^1 & m_u^0 & m_u^1 \\ m^2 & m^3 & m_u^2 & m_u^3 \\ m_v^0 & m_v^1 & 0 & 0 \\ m_v^2 & m_v^3 & 0 & 0 \end{bmatrix}$$

とする．また， Q_c は Q_m の要素を色にした行列である．

グラデーションメッシュの生成は，Ferguson Patch をラスタ画像にフィットさせる問題と見なせる．そこで，未知数ベクタを M として評価関数 E を

$$E(M) = \sum_{p=1}^P \sum_{0 \leq u, v < 1} \|I_p(m(u, v)) - f_p(u, v)\|^2 \quad (2.14)$$

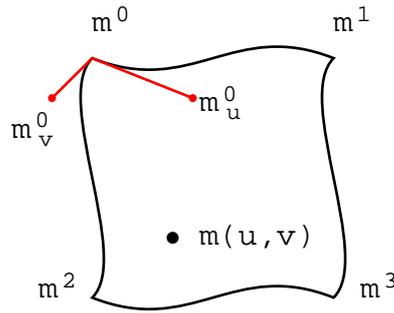


図 2-5: Ferguson Patch with Parameters (自発表 [26] より引用)

とおく．ここで， I は入力画像，添え字 p はメッシュ番号である．なお， Q_c の要素 c は画像から前景・背景を考慮したサンプリングにより決定するため，未知数ではない．

この形式の最小化問題は非線形最小二乗 (NLLS) 問題と呼ばれ，Levenberg-Marquardt (LM) 法によって数値計算可能となる [27]．LM 法では，評価関数の各項を各未知数で偏微分したヤコビ行列が必要である．評価関数 $E(M)$ には画像が含まれているため，その偏微分には画像と微分ガウス関数の畳込みを利用する．他の項では式 (2-13) から解析的に偏微分が計算可能となる．本方式は，大規模行列となるため非常に計算コストが高いという問題がある．例えば，実験で用いた対象画像の場合 (17×17 メッシュ， 6×6 サブメッシュ)，ヤコビ行列はおよそ $3,000 \times 1,000,000$ になる．

2.5 曲率と曲率スケール空間

2.5.1 曲率

曲率とは曲線の局所的な曲がり具合を定量的に表す指標である．ある曲線の微小区間を円で近似したとき，その円の半径を曲率半径，曲率半径の逆数を曲率と定義する．従って，円弧の長さを Δs ，中心角を $\Delta \alpha$ とすれば曲率半径 R は

$$R = \lim_{\Delta \rightarrow 0} \frac{\Delta s}{\Delta \alpha} = \frac{ds}{d\alpha} \quad (2.15)$$

と表せる．また，曲率 κ は

$$\kappa = \frac{1}{R} \quad (2.16)$$

となる．なお直線の曲率は 0 となる．さらに，右回りと左回りで正負が逆転する．

曲線が連続系の媒介変数表現で与えられるとき，曲率 κ は

$$x = x(t), \quad y = y(t) \quad (2.17)$$

$$\kappa = \frac{x'y'' - x''y'}{(x'^2 + y'^2)^{\frac{3}{2}}} \quad (2.18)$$

となる．なお，微分は媒介変数 t について行う．曲線が離散値として与えられるとき (デジタル曲線)，微分の計算方法に自由度がある．

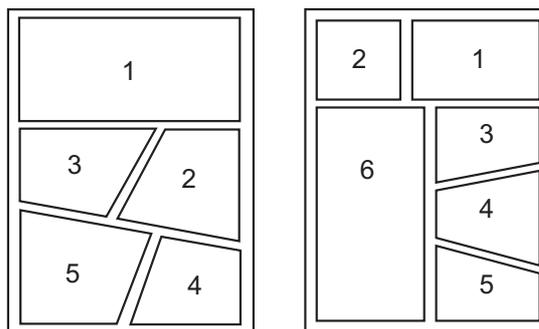


図 2-6: コミックのコマ配置例 [28] (著者の了解を得て抜粋)

2.5.2 曲率スケール空間

Asada と Brady はスケール空間の概念を，平面の曲線に沿った曲率の変化に拡張した．円弧の長さ s に対する接線 φ の向きを表す関数 $\varphi(s)$ として，曲線は表現される．そして， $\varphi(s)$ はガウシアン関数と畳み込まれる．畳み込まれた関数の一次と二次微分の局所的な最大値と最小値の位置が見出され， (s, σ) 平面の二つのスケール空間画像が結果となる．Mokhtaran と mackworth は，様々な詳細度における曲線の曲率ゼロ交差を見つけるために，曲率スケール空間 (CSS, Curvature Scape Scale) 手法を開発した．パス長の変数 u を考えるとき，曲線は二つの関数 $x(u), y(u)$ として表現される．そして， $x(u)$ と $y(u)$ はそれぞれ畳み込まれる．曲率ゼロ交差を用いた曲線の符号化によって，多重解像度形状表現が定式化された．

画像分類や形状分類など，大量のコンテンツを利用するためには，特徴量を抽出する必要がある．静止画像を含めた動画像の特徴量を記述するルールは MPEG-7 としてまとめられている．

形状特徴は MPEG-7 の一部であり，形状マッチングを行うための特徴量である．スケール空間フィルタリング (Scale Space Filtering) を用いて，直線の 2 次微分がゼロになる点が消失する位置とスケールを符号化している．

2.6 コミック画像からのメタデータ抽出方式

2.6.1 コマと視線遷移情報

コミックにおけるコマの重要な役割は，ページ内部における視線遷移情報を与えることである．コマの形状と読み順の関係について，隣接するコマ同士の相対的な枠線角度から順序づけを行う手法が提案されている [29]．

山田らの手法では，多角形のコマを考慮したコマ間の順序づけを行うことが可能となるが，今回処理対象としたコマの場合には，以下に示す，単純なルールにより順序を示すことができる．まず，それぞれのページでは，右上のコマから順に，右から左かつ上から下の順で読み進める．過程において，同一段に複数コマが存在する場合にはその段の右から左の順に読みすすめる．同一段に存在するすべてのコマが終わるとひとつ下の段へ移動する．左下

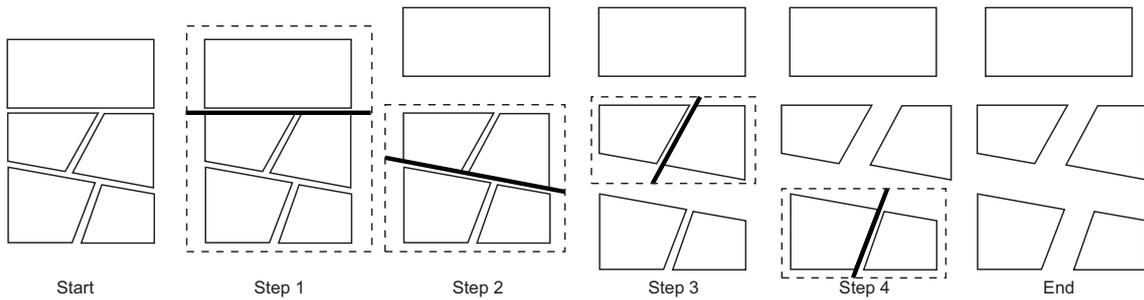


図 2-7: コマ分割のステップ [28] (著者の了解を得て抜粋)

のコマに到達するとそのページは終了となる。図 2-6 にコミック画像の一般的なコマ割りの一例を示す。図中の数字はそれぞれのページにおけるコマの読み順を示している。このコマの読み順が視線遷移情報である。

2.6.2 コマ分割手法の問題点

これまでに、文書画像を対象とした画像内のレイアウト解析手法が提案されている。文書画像のレイアウト解析処理では、文字や図などの画像上における空間的な配置や構造を領域解析によって認識する。投影法は、主に 2 値画像を対象とし、画像の垂直及び水平方向に黒画素の投影を行い、投票値の周辺分布から、ブロック境界を識別する [30]。図形融合では膨張、縮小処理や、Run Length Smearing Algorithm (RLSA) を用いて各図形の融合を行い、縦方向と横方向それぞれの結果から総合的にブロックを抽出する [31]。連結成分解析では、画像から抽出した各連結成分間のユークリッド距離に応じて結合処理を行うことで各オブジェクト単位にブロック化する [32]。しかしながら、コミック画像では、各オブジェクト間が明確に分離されていない場合があることや、コマ形状が矩形に限定されていないため、これらの手法では正しく認識が行えない。

一方、対象画像から、任意の形状を検出する一般化 Hough 変換が提案されている [33, 34]。任意の検出対象図形テンプレートを用意し、パラメータ空間上の評価により検出を行う。しかしながらコミック画像のコマ形状は矩形等に限定されず、高い自由度を持つため、テンプレートを用意できないという問題点がある。

コミック画像は文書画像と異なる特徴をもつため、レイアウト解析を行うためには特別な手法が必要となる。本論文では、画像の 2 辺を結び、始点座標及び角度によって指定される幅 1 画素の直線を検出線と呼ぶ。

田中らのコマ分割手法 [35] では、以下に示す手順でコマ分割処理を実行する。まず、分割に用いる重み付けされた濃度勾配値を決定する。画像全体に対して Sobel フィルタを適用し、得られた値に対して、ガウス関数を用いて画像中心に重みづけすることで、画像上の濃度勾配値を決定する。次に検出線の始点座標を画像の端部にそって 1 ピクセルずつ移動してゆく。そして各点それぞれにおいて検出線角度を変化させ画像の全探索を行う。ここで、角

度の変化量は1度刻みである。そして、得られた各検出線上の濃度勾配合計値によって分割を行う分割線を決定し、画像に対して2分割処理を行う。

全ての検出線上の濃度勾配合計値が小さい状態、つまり、分割対象画像内に分割線が存在しない状態となるまで、以上の処理を再帰的に適用して、各コマへの分割処理を行う。

2分割処理の再帰的適用による各コマへの分割ステップ及びその際に用いられた分割線の例を図2-7に示す。図中の点線で囲まれた部分はそのステップにおいて分割処理が行われた領域を示す。また、太線はそれぞれのステップにおいて分割に用いられた分割線である。

田中らの手法では検出線の始点座標及び角度を順次変更しながら画像内の探索を行うが、コマの枠線角度が検出線と若干ずれている場合にうまく検出が行えない。角度のずれを最小にするために、検出線のパラメータ変更きざみを小さく設定すると、結果として計算コストが非常に大きくなってしまふ。また、コマの枠線が歪んでいる場合にも検出が不可能という問題がある。そして、分割線の決定に濃度勾配値だけを用いるため、分割線の検出精度が低いという問題もある。

2.6.3 文字切出し手法の問題点

これまで多くの文字切出し手法、文字認識手法が提案されてきた。本論文では文字切出しだけに着目し、文字認識は対象外とする。ところで、文字認識結果や単語照合結果を文字切出しにフィードバックする方式も検討されている。しかし、形状的な特徴に基づく切出しがより正確になれば、フィードバック情報の確度が高まり、性能がより一層向上する。このため、文字の形状的特徴だけにに基づく文字切出しの高精度化は、重要な課題である。

文書画像を対象とする文字切出し手法においては、文字領域の抽出に先立ち、行を抽出する手法がある。具体的には、文字と思わしきもの（外接矩形）を文字行に垂直な直線に射影し、その分布密度を用いて行を抽出する。

後藤ら [36] は、罫線やノイズの混入を考慮しない単純な外接矩形が誤抽出の原因となることを示した。そこで、文書画像中の区分直線状の要素を文字行と仮定した抽出により、これを解決している。文書構造に関する知識を必要とせず、画像のゆがみにも耐性がある。本手法は必要なしきい値が多く、特に文字サイズに対するロバスト性が十分でない。また、罫線以外の要因に対する考慮が十分でなく、線画が混在すると誤検出が増加する。

線画の混在する画像として地図画像を対象とする文字切出し手法においては、文字枠図形をテンプレートとして用いる手法がある。テンプレートと原画像の適合度（黒画素密度）を抽出基準として文字を切出す。志久らは [37] は、地図画像から同一ポイント数の定型文字を切出すことを目的に、テンプレート数を1種類に削減する手法を提案している。背景との接触にも強いという特徴がある。本手法は、複数のポイント数が混在している画像から得られる結果の統合についての検討が不十分である。

2.7 むすび

本章ではベクタ表現の種類と，ラスタ表現における線画を，多角形近似と曲線置換によりベクタ表現に変換する技術を説明した．線画以外の画像をベクタ表現に変換するための手法についてまとめ，主観品質が著しく低下するという問題点を示した．次に，網点の特徴とその分離手法についてまとめ，ベクタ表現での階調再構成が考慮されていないことを明らかにした．また，階調画像から骨格成分と呼ばれる輝度が均一な成分を取得する全変動最小化フレームワークを説明し，計算コストが高いことを定性的に示した．加えて，グラデーションメッシュによるベクタ表現取得方法について説明した．さらに，曲率や曲率スケール空間など，平面曲線における特徴量の扱いを整理した．続いて，コミック画像から取得できるメタデータとしてコマレイアウトと文字切出しを説明し，その自動抽出手法の問題点を示した．最後に，電子コミックにおける配信や閲覧の既存システムについて言及した．

第3章

網点分離による画像分解方式

3.1 まえがき

本章では、人工画像の形式に関して高解像度2値画像、もしくは本来2値画像である印刷物をスキャンして電子化した低解像度グレースケール画像を対象として、2値画像で濃淡を表現する網点手法のモデル化を提案する。網点はコミックにおいて中間調を表現するために幅広く使われており、かつスクリーントーンとして様々な種類が市販されるなど、一般的な要素である。

対象画像が2値画像であれば網点かどうかによらず、容易にベクタ表現にできる。しかしながら、多数の小さな点をベジエ曲線で表現するのは符号化効率が悪い。さらに、点のモデル化はベクタ表現であっても、解像度変換によりモアレが発生する。従って、濃淡表現（グレースケール）の復元するモデル化が有効であると考えられる。

上記を踏まえて、網点手法が解像度変換にともなうアーティファクト発生の原因であることを示す。さらにアーティファクトを発生させないようにするため、高解像度2値画像において輝度変化の周期を取得する方法を提案する。これを用いて網点領域を分離し、元の濃淡（階調）を復元して線形勾配で近似する手法を提案する。なお、提案手法は特定の網点条件を仮定しておらず、汎用的に適用できる。

これらの提案手法は網点を構成する点が十分識別できるほどの高解像度2値画像を前提としている。しかし、汎用的なスキャナやデジタルカメラはグレースケールのような連続階調を取得できる。連続階調の場合、解像度が低くても十分な画像品質を達成できる。また、解像度の低下で画素数が少なくなり、計算機での取り扱いが容易になる。さらに、すでに蓄積された多くのコンテンツは、低解像度グレースケールであることが多い。この理由としては、画像符号化方式であるJPEGの普及が挙げられる。ただし、低解像度画像を2値化してベクタ表現に変換すると精度が低下するという課題がある。

そこで、低解像度グレースケール画像を対象に、網点を分離する手法と精度の高い線画を生成する手法を提案する。前者について、2値画像向けの提案手法とグレースケール画像向けの従来手法を、網点中心を取得するという観点で統合する。後者について、ベクタ変換に入力可能な2値画像を生成するために、解像度変換（高解像度化）と2値化手法のパラメータを最適化する。最後に、提案手法の有効性を示すために、以下の評価実験を行う。さまざま

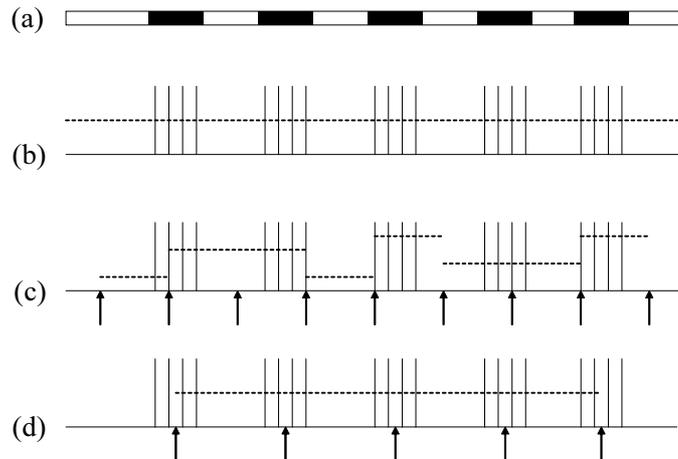


図 3-1: モアレ発生の模式図 (自発表 [38] より引用)

まな条件の網点テストチャートを作成し、網点の分離特性、網点の検出濃度の精度を定性的に評価する。さらに、実コンテンツを用いてモアレ低減の検証実験を行う。グレースケール画像についても、テストチャートを作成し、2値・多値統合網点中心の検出性能を評価する。さらに、線画だけを含む低解像度グレースケール画像（多値画像）を用いて、高解像度2値化手法の評価実験を行う。

3.2 モアレの種類と発生原因

オフセット印刷はインクを紙に載せるか否かによって紙面に画像を転写する。すなわち、電子的には2値画像に相当する。従って、オフセット印刷では本質的に階調を表現できない。オフセット印刷は極めて解像度が高く1200dpi程度であるため、黒インク領域をちりばめることで、視覚的には中間調を知覚させることができる。

2値画像で中間調を疑似表現する手法としてディザ法（誤差拡散法）が知られており、そのひとつに網点がある。網点の考え方は、無限小の点を周期的に配置し、さらに点の大きさを濃度によって変化させる。その結果、単位面積当たりの平均輝度が濃度となる。なお、網点は様々な線数で、均等濃度やグラデーションのスクリーントーンとして市販されている。なお、線数とは印刷用語で、1インチ当たりの点の個数である。一般に、モノクロ印刷では80線程度、カラー印刷では300線程度が利用されている。

ここで、画像を $1/n$ に縮小する場合を考える。図3-1(a)は、網点を含む画像の1列を抜粋したものである。図3-1(b)は、図3-1(a)の画素を濃度に置き換えて表している。また、点線は仮想的な濃度を示している。図3-1(b)を矢印で示す等間隔で網点面積率、すなわち濃度を計算したのが図3-1(c)である。濃度にうねりが生じているのが分かる。

縮小率 $1/n$ と網点周期の整合性がとれれば、モアレは生じない。例えば図3-1(d)に示すように等間隔で濃度を計算すれば、正しく求められる。網点周期は用いる画像によって異なる上に、同じ画像内でも網点領域毎に異なる。さらに網点周期が整数でない場合の対応は困

難である．網点はオフセット印刷において中間調を生成する極めて有用な技術であるが，解像度変換においてはモアレ発生の原因となる．そこで，これを分離する必要がある．

本論文で扱うモアレについて，用語と発生タイミングの整理をする．まず，画像をスキャナ等で入力する際に，網点と入力解像度の不整合がモアレの発生原因となる．これを入力時モアレと呼ぶこととする．次に，網点を含む2値画像を解像度変換する際に，網点と出力解像度の不整合がモアレの発生原因となる．このモアレを解像度変換モアレと呼ぶこととする．最後に，階調近似処理において，網点周期と整合性のない固定サイズで濃度を検出するときにも，モアレが発生する．これを検出時モアレと呼ぶこととする．本論文では入力時モアレの存在しない画像を対象としている．すなわち，入力時にモアレを発生させないことが前提条件である．

3.3 網点分離と階調近似の提案

本論文では，網点を分離してグレースケールで近似を行うと共に，網点領域の境界線と残差である線画をベクタ表現に変換して，モアレ・ジャギーの低減と網点領域の符号化効率改善を両立する画像モデルを提案する．

網点を分離せずに，点の位置と半径，もしくは点の境界線をベクタ表現に変換し，その後で解像度変換を行っても，モアレが発生する可能性がある．従って，2値画像から必ず網点を除去しなければならない．画像の再現性を補うために，網点を階調近似し，合成する手法を検討する．

以下では画像の制約や網点，線画のベクタ変換について述べる．

まず，これらの画像の解像度として300dpi～600dpiの高解像度を想定する．一般に，文書の2値画像は300dpi～400dpiの解像度が必要と言われているからである．さらに，コミックや書籍で利用されるルビのような小さい文字を鮮明に入力するためには，600dpiの解像度が必要である．

次に，本論文で対象とする網点として均等濃度を含めた線形グラデーションに限定する．それ以外の放射状グラデーションなどは利用頻度が少ないため，本論文では対象外とする．

最後に，以下の節では，画素の明るさを輝度と呼び2値画像とグレースケール画像の両方に対して用いる．また，ある領域の明るさを濃度と呼び，2値画像（網点）だけに対して用いる．さらに，2値画像において連続した黒画素を連結画素と呼ぶ．

提案手法の概念を図3-2に示す．入力された2値画像から網点を分離し，残差を線画とする．次に，分離した網点を階調近似して，ベクタ表現に適しているグラデーションに変換する．また，網点の存在する領域を定義し，その領域の境界をベクタ表現に変換する．

ここで，網点を分離する処理フローを具体的に述べる．まず，画像の水平，垂直方向それぞれについて，1次元信号における第3.3.1節で定義する網点中心を検出する．次に，この結果を統合して，2次元信号（平面）としての網点中心を検出する．さらに，1周期ごとに第3.3.2節で定義する反転率を計算し，網点を分離する．以下では，網点中心の検出，網点領域の分離，グラデーション近似を順に述べる．

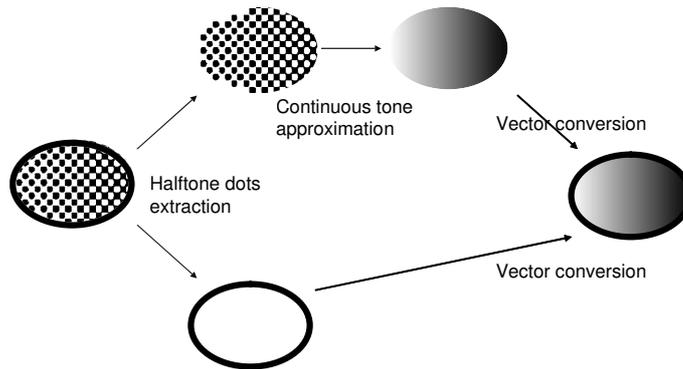


図 3-2: 網点分離による画像分解方式の概念図 (自発表 [39] より引用)

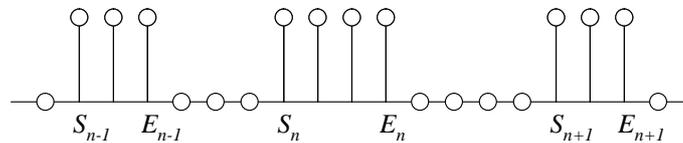


図 3-3: 網点領域における画素位置と輝度の関係 (自発表 [39] より引用)

3.3.1 網点中心の検出

本論文では入力として高解像度 2 値画像，もしくは本来 2 値画像である印刷物をスキャンして電子化した画像を想定する。

スキャンして得られる画像は階調画像であることが多いが，輝度ヒストグラムを取得すると，紙面の色とインクの色に相当する 2 カ所に頻度の偏った画像である．ところが，大内らの手法 [16] は輝度の変化を用いて極点やピーク画素を定義しているため，色頻度の偏った画像では網点中心を検出できない．そこで，網点領域を構成する点の幾何的な中心座標を網点中心と定義する．黒，または白の連結画素の中心座標を求め，隣接する中心座標間の距離（中心距離）が等しければ，網点中心であると判断する．これは，網点が周期的に配置されているという性質を利用している。

具体例として，網点領域の波形（1 次元）を図 3-3 に示す．また，中心距離が等しいという条件は

$$\left| \left(\frac{S_n + E_n}{2} - \frac{S_{n-1} + E_{n-1}}{2} \right) - \left(\frac{S_{n+1} + E_{n+1}}{2} - \frac{S_n + E_n}{2} \right) \right| < \frac{th}{2} \quad (3.1)$$

となる．ただし， S は連結画素の開始座標， E は連結画素の終了座標， n は連結画素の番号， th はしきい値である．さらに，式 (3.1) を満たす連結画素 n の中心座標 $(S_n + E_n)/2$ が，1 次元における網点中心の候補となる。

理想的には隣接した中心距離は一致するが，実際には一致しないことが想像される．その要因として，座標の値は整数化され座標ごとにたかだか $1/2$ 画素の誤差が生じることが挙げられる．さらに，この誤差を許容するために，しきい値 th はたかだか 4 となる．一方で，

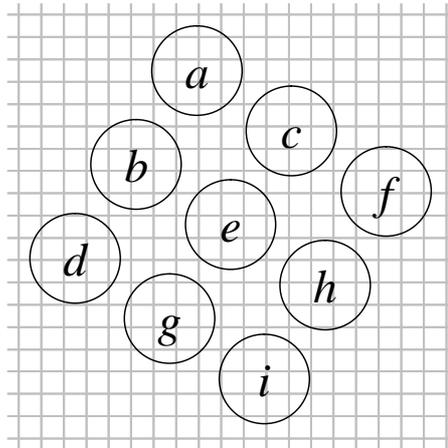


図 3-4: 仮想的な黒画素の大きさと画素位置の関係 (自発表 [39] より引用)

しきい値 th が大きいと誤検出が増加する．そこで，本論文では様々な画像に対する予備実験によりしきい値 $th = 3$ を採用した．

網点領域の網点と画素の模式図を図 3-4 に示す．それぞれの円は網点領域を構成する点，正方形は画素である．この図の水平，垂直方向の 1 次元信号について，先に述べた網点中心の候補を取得する．さらに，両方向共に網点中心の候補となった座標を取得する．この座標が網点中心である．

なお，提案手法において網点の配置は水平，垂直など特定の方向には限定していない．濃度が十分高ければ水平，垂直方向の 1 次元にスキャンするだけで，かならず網点周期が得られる．一方，濃度が小さい場合には 1 次元にスキャンするだけでは近傍の点との周期を取得できない可能性が高い．その場合では数個離れた網点との周期性が取得できる．さらに，特定の向きや濃度で周期が全く取れない可能性がある．この場合，解像度変換を適用してもモアレが発生しないため，そもそも分離する必要がない．このように，提案手法はモアレが発生しうる網点を必ず分離できる．網点を様々な方向に配置した場合，濃度が正しく得られているのか実験により評価を行う．

3.3.2 網点領域の分離

分離の処理単位として，図 3-5 に示す網掛け領域を用いる．この網掛け領域は網点 1 周期分に相当し，ある網点中心の周囲に四つの網点中心が存在し，かつそれらが正方形を構成している時の半分の領域である．この領域内における反転率がしきい値以上であるときに，その領域を網点として分離する．網点領域では網点中心が規則的に配置されているために，網点領域を隙間無く分離できる．

最後に，それぞれの手法で分離した網点を合成して網点画像を得る．また，網点画像に膨張処理を施して網点領域を，入力画像から網点画像を除いて線画を得る．

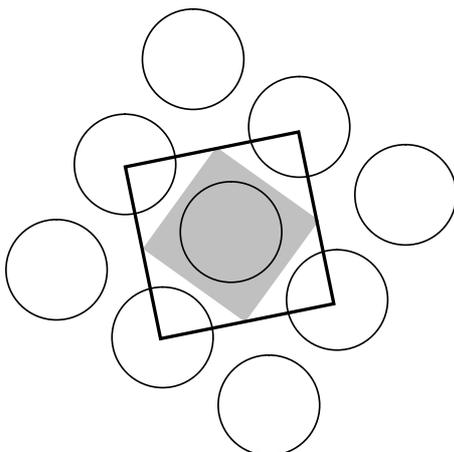


図 3-5: 網点抽出の基本処理単位 (自発表 [39] より引用)

3.3.3 線形勾配による近似

均等濃度と線形グラデーションは、画素の座標と輝度について平面の方程式で表現できる。そこで、網点領域を平面の方程式で近似し、グラデーションパラメータを得る。以下に具体的な処理手順を述べる。

まず、網点分離と同じ1周期分の領域を用いて、網点中心の座標ごとに濃度を求める。次に、それらの座標と濃度について平面の方程式 $z = ax + by + c$ を用いて最小二乗近似を行う。さらに、グラデーションパラメータ a, b, c を出力する。なお、予備実験により網点中心の座標における濃度 (0~1) の標準偏差は 0.0184 であった。従って、濃度を推定する領域として1周期を利用する手法は十分な精度を有している。

3.4 低解像度グレースケールへの拡張

画像入力 (スキャン) 時間を短縮するために、解像度は低い方が良い。また、JPEG 符号化方式の普及により、文書画像が低解像度グレースケール画像で保存される機会が増加している。本節では、低解像度グレースケール画像へのベクタ変換適用を目的とし、ベクタ表現への変換を考慮した小数画素精度 (サブピクセル精度) の2値化手法を提案する。また、ベクタ表現による画質評価を行う。

文字を高品質にベクタ表現に変換するには、300dpi から 600dpi 程度の解像度が必要である。対象画像の一般的な解像度は 120dpi 程度であるため、小数画素化率 n は 2~5 を目標とする。また、得られたベクタ表現について画質評価を行う。

3.4.1 2値・多値網点分離手法の統合

ここまで高解像度2値画像に対する網点分離手法を述べてきた。一方、スキャン速度の観点から印刷物を低解像度グレースケール画像として取得する場合がある。この場合、大内ら

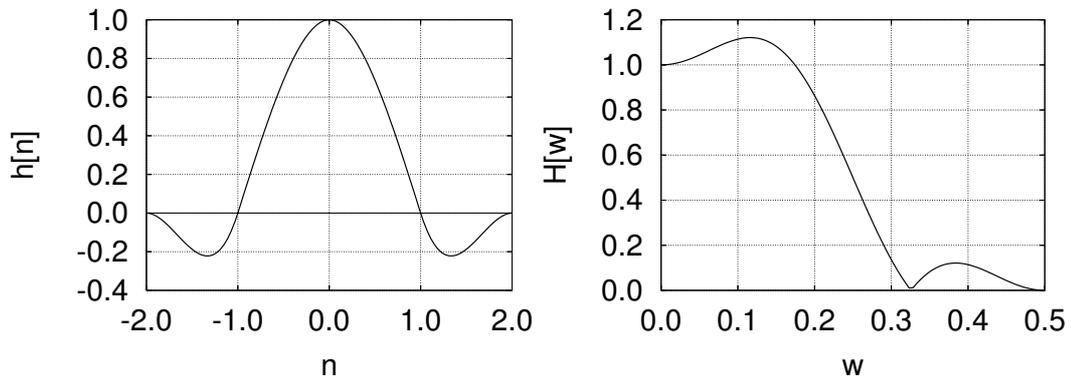


図 3-6: キュービック・コンボリューションの特性 (自発表 [41] より引用)

の手法 [16] が有効である .

大内らはブロック単位に網点領域の性質を利用して, 網点中心を検出する手法を提案している . まず, 3×3 画素ブロック内で最大, または最小の輝度を有し, かつ, ブロックの中央にある画素を検出する . 次に, 同じ画素ブロック内で, 中央の画素値と周囲の画素値がしきい値以上の画素を検出する . 両方を満たす画素をピーク画素と定義し, 網点中心と見なす .

これらを入力画像に適用すると, あらゆる解像度・階調について網点中心が抽出可能とある . さらに, それぞれの手法で検出された網点中心は, 論理和により統合する . 網点中心が得られた後は, 本論文の提案手法で分離可能となる .

3.4.2 高解像度 2 値化

提案手法の構成を以下に示す . グレースケール画像を高解像度化し, Niblack の 2 値化 [40] を適用する . ただし, 2 値化により強調されるモスキートノイズを低減するために, 中間調だけを 2 値化する . 生成された小数画素精度の 2 値画像をベクタ表現に変換する .

小数画素精度の画素値の取得には, 補間が用いられる . フィルタ処理によるリングング歪みは後段の 2 値化によって無視できるため, 周波数応答のうちカットオフ周波数が理想的であることが望ましい . 本論文では式 (3.2) に示すキュービック・コンボリューション補間法を用いる .

$$h(t) = \begin{cases} (a+2)|t|^3 - (a+3)|t|^2 + 1 & 0 \leq |t| < 1, \\ a|t|^3 - 5a|t|^2 + 8a|t| - 4a & 1 \leq |t| < 2, \\ 0 & 2 \leq |t|. \end{cases} \quad (3.2)$$

a は定数であり, 本論文では -1.5 を利用する . フィルタの時間, 周波数応答 (振幅) を図 3-6 に示す .

なお, JPEG 符号化やローパスフィルタなどにより高周波成分が十分に除去されている場合には, 線形補間でも同等の高解像度化性能が得られる [18] .

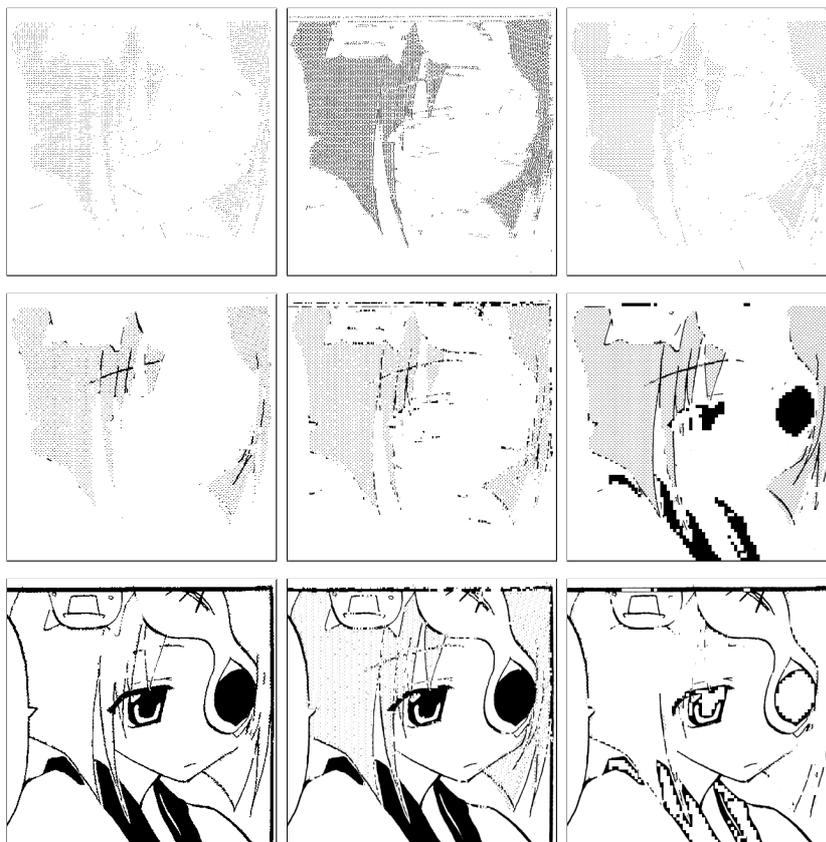


図 3-7: 網点分離結果画像の一部．上から網点中心，網点画像，線画，左から提案手法，従来手法 1，従来手法 2（自発表 [38] より引用）

2 値化に用いる Niblack 法は，窓内画素値の平均と標準偏差に基づいてしきい値を決定する手法である [40]．しきい値は式 (3-3) のように算出される．

$$T = E + k\sigma. \quad (3-3)$$

ただし， E は平均， σ は標準偏差， k は定数である．なお，平均や標準偏差を求める小領域の一辺は，小数画素化率 n に対して $2n + 1$ 画素とした [17]．これは，入力画像における 3 画素程度に相当する．

3.5 提案手法の評価

3.5.1 網点分離の特性

網点分離特性について提案手法，上野らの手法（従来手法 1）[15]，大内らの手法（従来手法 2）[16] の比較を行った．それぞれの手法について，検出された網点中心（極点，ピーク画素），分離された網点画像，残差である線画を図 3-7 に示す．

まず，抽出された網点中心について述べる．提案手法は画像の水平，垂直の両方向に着目して検出しているため，従来手法 1 と比較して過剰検出が削減されている．また，網点領域

の検出精度は従来手法 2 と同程度を有し、かつ線画領域の誤検出が減少している。従って、2 値画像に対して提案手法は、グレースケール画像を対象とした従来手法と同程度の検出精度を達成している。

次に、分離された網点画像と線画について述べる。提案手法は従来手法と同様に網点中の線画を誤分離しているが、網点領域に隣接した線画の誤分離は著しく低減されている。また、塗りつぶし領域を線画として分離できる。この網点中の線画については、周囲との濃度差を考慮するなどの、さらなる検討が必要である。従来手法 1 と比較して、線画における網点の分離漏れが著しく低減されている。また、従来手法 2 と比較して、分離の処理単位として網点 1 周期を利用しているため、網点と線画の境界領域における過剰分離が低減され、線画の品質が向上している。

3.5.2 検出濃度の精度

様々な濃度やグラデーションの網点を実際のコミックから収集することは困難であるため、計算機内で生成した網点画像を利用して評価を行う。スキャン解像度を 300dpi であると仮定し、300 画素四方のグレースケール画像を発生させ、網点法（網点角度は 45 度）により 2 値画像を生成した。また、他の網点角度や濃度勾配は、この画像を回転して生成した。

まず、60 線と 85 線、濃度 10% ~ 50% の均等濃度網点における、入力画像の網点角度と検出濃度の関係を図 3-8 に示す。図中の T は入力濃度、L は入力線数である。濃度 40% ~ 50%、網点角度 45 度の時に、濃度が 10% 程度低く検出されている。これは、2 値化したテスト画像の濃度が低下していたことに原因がある。他のケースでは線数や網点角度に依らず、3% 以内の誤差で濃度が検出されている。テスト画像の濃度は 2 値化（網点化）により 2% 程度の誤差を含んでおり、提案手法による濃度検出の精度は十分である。

次に、60 線と 85 線、網点角度 45 度と 90 度の均等濃度網点における、入力画像の既知濃度と検出濃度の関係を図 3-9 に示す。図中の A は入力角度である。先と同様に、濃度 40% ~ 50%、網点角度 45 度の時は 10% 程度濃度が低く検出されているが、他のケースでは線数や網点角度によらず理想値にそった濃度が検出されている。例えば、スクリーントーンの濃度は 10% 刻みで市販されているため、提案手法は十分実用的な検出精度が得られている。

さらに、60 線と 85 線、濃度 30%、60%、90% のグラデーション網点における、濃度勾配の既知方向と検出方向の関係を図 3-10 に示す。検出方向は理想直線上にのっており、誤差は最大でも 2 度であった。また、60 線と 85 線、濃度勾配 45 度と 90 度のグラデーション網点における、入力濃度と検出濃度の関係を図 3-11 に示す。濃度 50% 程度のグラデーションでは全体的に検出濃度が低下しているが、これは網点が結合し始める境界濃度だからである。二つの結果から、線数、グラデーション方向、勾配によらず、グラデーションの濃度を検出できることが確認できる。

上記の実験により、提案手法が分離された網点領域の濃度を、線数や網点角度などに対して頑健、かつ精度良く推定できることを明らかにした。

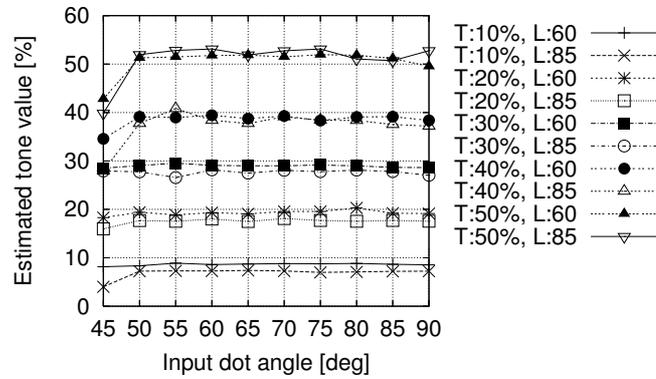


図 3-8: 濃度の異なる均等濃度網点における既知濃度と検出濃度の関係 (自発表 [42] より引用)

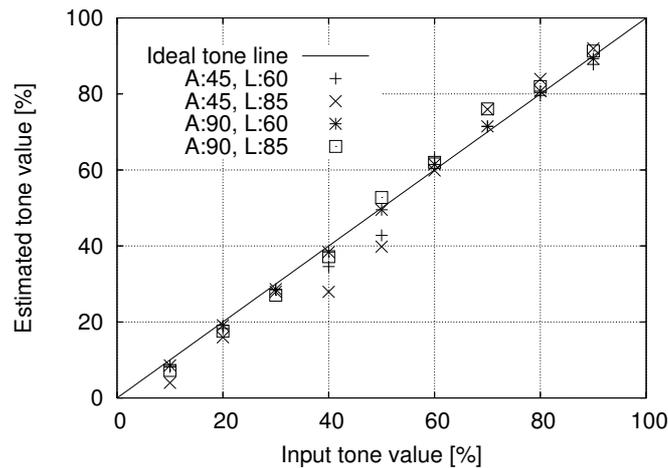


図 3-9: 網点角度の異なる均等濃度網点における既知濃度と検出濃度の関係 (自発表 [43] より引用)

3.5.3 モアレ低減の検証

解像度変換におけるモアレ低減について述べる．テスト画像の一部 (35.6mm×35.6mm) の原寸大と，67% に解像度変換する際，提案手法を適用した画像とバイキュービック法 [44] を適用した画像をそれぞれ図 3-12 に示す．なお，提案手法において中間調はグレースケールとして再現したが，再度網点法を適用するなどして質感の再現もできる．これらの図より，バイキュービック法では顕著なモアレが発生しているが，提案手法ではモアレが発生しないことを確認できる．また，拡大時にはジャギーが発生しないことも確認している．

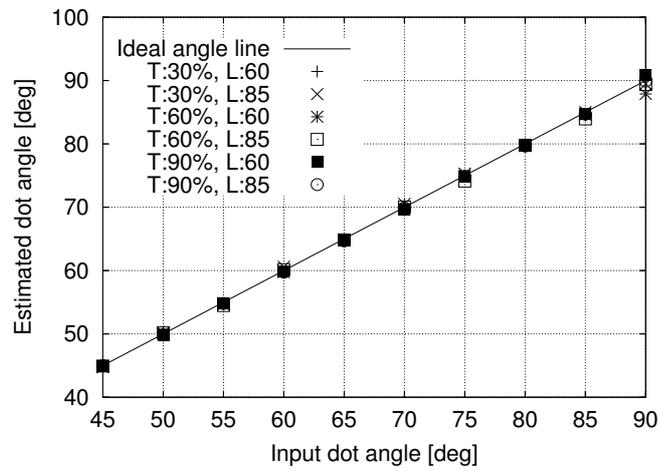


図 3-10: 線形濃度勾配網点における既知方向と検出方向の関係 (自発表 [43] より引用)

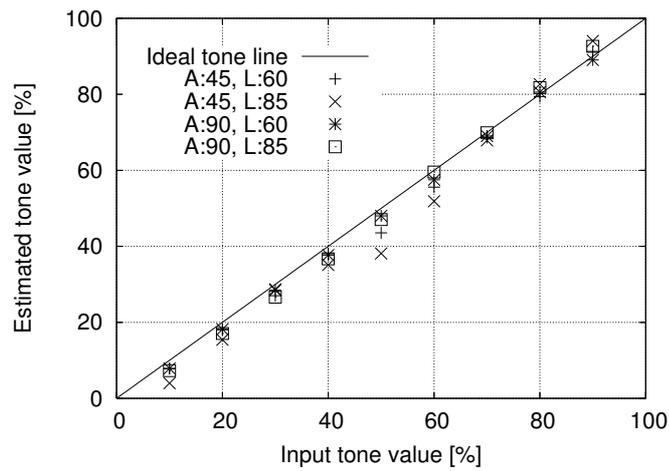


図 3-11: 線形濃度勾配網点における既知濃度と検出濃度の関係 (自発表 [43] より引用)

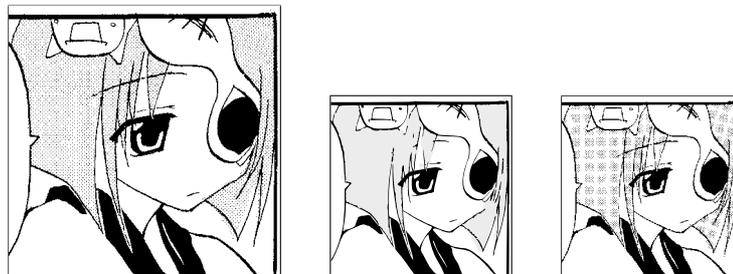


図 3-12: モアレ低減の検証. 左 原寸大, 中央 67%提案手法, 右 67%バイキュービック法 (自発表 [39] より引用)

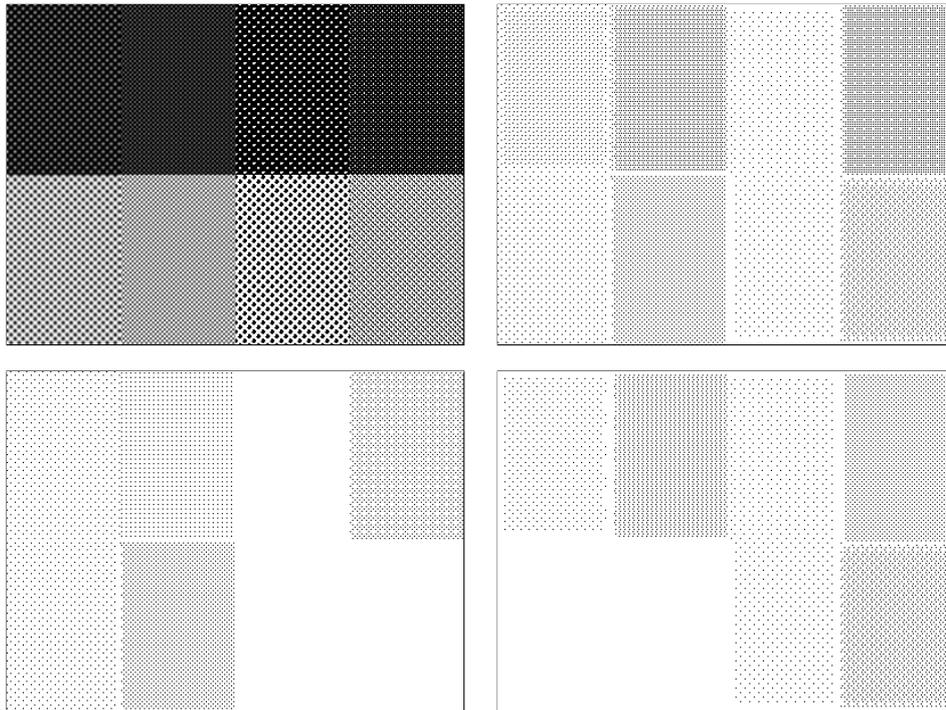


図 3-13: 2 値・多値統合網点中心の検出結果．左上から 2 値とグレースケールの網点画像，2 値とグレースケールの網点中心の検出結果，グレースケール画像用の従来手法，2 値画像用の提案手法（自発表 [45] より引用）

3.5.4 2 値・多値網点中心の検出

2 値画像とグレースケール画像のいずれの入力であっても網点中心を検出する手法を提案した．提案手法の有効性を検証するために，テスト画像を作成して性能評価を行った．テスト画像は濃度 75%（上段）と 25%（下段）に対して，グレースケール（左側）と 2 値（右側）及び 2 種類の線数による網点法を計算機内で適用した画像である．テスト画像と網点中心の検出結果を図 3-13 に示す．

この結果より，提案手法によりあらゆる解像度や階調の網点に対して，網点中心が取得できることが主観的に確認できる．また，2 値画像向けとグレースケール画像向けのそれぞれの手法が対象とする網点については必ず網点中心を取得している．さらに 2 値画像向けの手法でグレースケール画像の網点中心や，その逆について，一部条件について検出している．

3.5.5 高解像度 2 値化の特性

対象画像の一例としてコミックを想定する．B5 サイズ，300dpi（1486 × 2207 画素）のグレースケール画像を基本画像と呼ぶ．これを 150dpi に解像度変換し，JPEG で 1.0bit/pel に符号化した画像が実験の入力画像となる．Niblack 法において線のつながりを保持するために，しきい値が背景側になるように定数 $k = 0.1$ を採用する．実験条件より小数画素化率

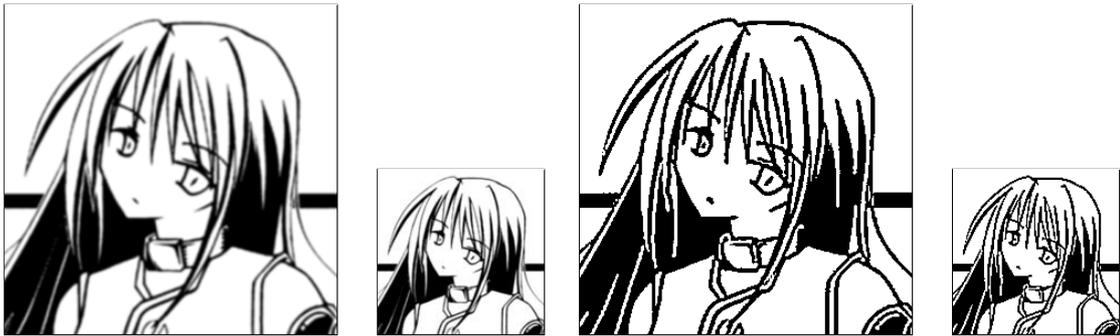


図 3-14: 高解像度 2 値化の実験結果例．左上から原画像，入力画像，高解像度 2 値化の結果，整数精度 2 値化の結果（自発表 [46] より引用）



図 3-15: ベクタ表現による比較．左から高解像度 2 値化，整数精度 2 値化，原画像（自発表 [46] より引用）

は $n = 2$ である．モスキートノイズの影響低減のため，輝度値が 128 ± 96 の画素について Niblack 法を適用する．整数画素精度の 2 値化を従来手法として比較を行う．

図 3-14 に基準画像と入力画像の一部，提案手法と従来手法により得られた 2 値画像の一部を示す．提案手法は文字領域や複雑な線画領域で，精度良く 2 値化できることが分かる．従来手法では，線画の欠落や曲線の単調化などにより，品質が低下している．

ベクタ表現による比較を行うために，提案手法，従来手法，基準画像をそれぞれベクタ表現に変換した結果を図 3-15 に示す．提案手法は基準画像と同様な線幅に変換されているが，従来手法では線幅が広がり，単調化している．また，複雑な線画領域における再現性が従来法よりも向上している．

3.6 むすび

本章では，人工画像の形式に関して高解像度 2 値画像と低解像度グレースケール画像を対象として，網点手法のモデル化を提案した．

まず，網点の間隔（周期）と解像度変換の比率とが整数倍でない場合にモアレが生じることを明らかにした．この性質を用いて，高解像度 2 値画像において輝度変化の周期を取得す

る方法を提案した。得られた網点周期を用いて網点領域を分離し、元の濃淡（階調）を復元して線形勾配で近似するモデル化を提案した。テストチャートを用いた実験により、提案手法は汎用的な網点に適用可能であり、近似精度と頑健性が高いことも示した。

さらに、低解像度グレースケール画像に対しては、上記提案手法と既存の網点検出手法とを、網点中心の検出という観点で統合する手法を提案した。また、線画の高精度なベクタ表現を得るために、線画は高解像度化とパラメータを最適化した2値化が有効であることを示した。

第4章

線形勾配成分の抽出方式

4.1 まえがき

本章では，人工画像の形式に関してグレースケール画像やカラー画像へ対象を拡大する．また，これらに適したベクタ表現である線形グラデーションやグラデーションメッシュに着目する．前者は入力画像を，ベクタ表現可能な線形勾配（グラデーション）領域・成分と，残差であるスカラ表現に適したテクスチャ成分に分離するモデルである．後者は，画像全体をグラデーションメッシュ成分とテクスチャ成分に分離するモデルである．

線形勾配領域・成分の抽出には，均一輝度成分を抽出する全変動最小化（Total Variation Minimization）のフレームワークを適用する．全変動はL1 ノルムのまま陽に離散化し，劣勾配法により数値計算できることを示す．劣微分値が $+1, 0, -1$ のいずれかになるため，従来手法と比較して除算と平方根計算を除去できる．また，提案手法を水平・垂直の微分画像に適用し，線形勾配領域・成分を抽出する．なお，劣勾配法や劣微分は文献 [47] に詳しい．

評価実験により提案手法は反復回数は増加するが，反復処理あたりの計算時間が削減されることを示す．また，従来手法と同程度の分離精度を維持して計算量の大幅な削減を明らかにする．提案手法はL1 ノルムのまま数値計算するため，従来手法よりも広い均一輝度成分の領域を得られる．その結果，従来手法よりも線形勾配領域が広いことを確認する．

さらに線形グラデーション成分をも内包するグラデーションメッシュ（Gradient Mesh）により画像全体を近似する手法について，画像符号化の基礎ツールとしての有効性を画像特性と計算量の観点から検討する．具体的には，画像中のオブジェクトだけではなく，画像全体を符号化するために，非線形最小二乗法への定式化を提案する．

4.2 全変動最小化の高速計算手法

全変動（Total Variation）は，逆問題を解くための正則化基準として，最初に Rudion, Osher, Fatemi によって画像処理に導入された（ROF モデル）[48]．グレースケール画像における全変動は，隣接画素における輝度値の差分絶対値の合計に相当する．たとえば，ある領域における輝度が単調に増加していれば，その領域の全変動は不連続点の有無にかかわらず一定値である．一方，テクスチャやノイズのように常に振動している領域では全変動が大きくなる．輝度変化が滑らかであるという仮定を導入せずに，全変動は画像を正則化する．

すなわち、オブジェクト境界のようなエッジを含む画像を取り扱うことができる。そのため、全変動を画像処理に適用する研究が盛んに行われている [23, 24, 48–52]。

全変動最小化により取得できる画像の中に骨格成分と呼ばれる画像がある。全変動を用いた画像分解問題において、骨格成分は画像内の幾何形状を単純に記述可能な、均等色オブジェクトをモデル化した画像と言われる。なお、入力画像から骨格成分を除去すると、テクスチャとノイズを含む振動成分が得られる。一方、全変動を用いた画像復元問題において骨格成分は、劣化フィルタが線形かつ既知であるが不良設定の逆問題を解いて得られる原画像に相当する。このとき、原画像の先見情報を全変動による正則化として画像復元問題が解ける。これらの骨格成分を取得する問題とは別に、領域分割問題への適用が検討されている。領域分割の入力として骨格成分を用いる手法や、領域の輪郭線を全変動を用いて正則化する手法は、テクスチャやノイズに頑健な領域分割手法となる [49]。

有界変動関数として定義される全変動及びこれを含む最小化問題に対して、これまで様々な離散化と数値計算手法が提案されてきた。Chan らや Vese らは、オイラー方程式を用いて最小化問題を直接解く手法を提案した [50, 51]。この手法には本質的にゼロ除算が含まれており、実装上の工夫により回避している。Carter らは、求める骨格成分とは双対の関係にある振動成分を表す双対変数の導入で、ゼロ除算を回避した [52]。Chambolle らも同様に、双対変数を含む双対問題を導出し、収束を保証した半陰的再急降下法を提案した [23]。Combettes らは、画像の先見情報を全変動による制約条件と見なし、制約条件付き最小二乗法と劣勾配法を用いて骨格成分を直接求める手法を提案した [24]。これらの中で、Chambolle らと Combettes らの手法は高い安定性と性能を有している。

しかし、上記の手法は計算コストが高いという問題がある。例えば、双対問題を用いる手法では、計算対象である振動成分の原関数の次元数が、入力画像のその2倍になる。制約条件付き最小二乗法を用いる手法では、1回の反復処理につきFFT (Fast Fourier Transform) と逆FFTの計算が必要になる。これまで利用されている離散全変動の定義は、画像へ拡張する際に L^2 ノルムを用いた輝度勾配を利用している。そのため、劣微分係数が輝度値に依存し、さらに平方根計算が必要になる。また、半陰的再急降下法では、反復処理における各項の役割が陽でない。

本論文では、全変動最小化を用いて画像から骨格成分の取得を想定し、全変動最小化を高速に計算する手法を提案する。まず、全変動を2次元画像に適用して離散化する際に、輝度勾配として平方根計算が必要ない L^1 ノルムを用いる。次に、劣勾配を用いた反復法で骨格成分を直接計算する。双対変数を計算しないこと、FFTを含む射影関数や平方根計算が不必要であること、さらに浮動小数点の計算回数の削減により計算量を削減できる。

4.2.1 離散全変動の定義と凸性

これまで述べてきた問題点を解決する全変動最小化の高速計算手法を提案する。まず、全変動を L^1 ノルムのまま2次元に拡張する方法を述べる。そして、この離散全変動が凸関数であることを明らかにする。さらに、提案した離散全変動の劣微分を導出する。次に、反復

処理における計算コストが低い Chambolle らの画像分解 (ROF モデル) と, 反復処理において各項の役割が陽である Combettes らの劣勾配法とを用いて, 提案する離散全変動最小化の数値計算を実現する.

全変動は隣接画素同士の差分絶対値の合計であるという概念を, 2次元画像への適用で離散化を実現する. 輝度勾配を L^2 ノルムではなく L^1 ノルムにより定義する. 従って, u の離散全変動として

$$J_{\text{tv1}}(u) = \sum_{1 \leq i, j \leq N-1} |u_{i+1, j} - u_{i, j}| + |u_{i, j+1} - u_{i, j}| \quad (4.1)$$

$$+ \sum_{1 \leq i \leq N-1} |u_{i+1, N} - u_{i, N}| + \sum_{1 \leq j \leq N-1} |u_{N, j+1} - u_{N, j}|$$

を提案する. ここで, 任意のベクタ $\mathbf{u}_1, \mathbf{u}_2$ と任意の実数 $0 < \alpha < 1$ に対して,

$$J_{\text{tv1}}(\alpha \mathbf{u}_1 + (1 - \alpha) \mathbf{u}_2) \leq \alpha J_{\text{tv1}}(\mathbf{u}_1) + (1 - \alpha) J_{\text{tv1}}(\mathbf{u}_2) \quad (4.2)$$

を満たしている. 従って, J_{tv1} は凸関数である [47]. 以下では, J_{tv1} を L^1 全変動と呼ぶことにする.

提案する離散全変動の劣微分は以下ようになる.

$$\begin{aligned} \partial J_{\text{tv1}}(\mathbf{u}_{i, j}) &= \text{sgd}(u_{i+1, j} - u_{i, j}) + \text{sgd}(u_{i, j+1} - u_{i, j}) \\ &+ \text{sgd}(u_{i-1, j} - u_{i, j}) + \text{sgd}(u_{i, j-1} - u_{i, j}) \end{aligned} \quad (4.3)$$

$$\text{sgd}(x) = \begin{cases} +1 & (x < 0) \\ 0 & (x = 0) \\ -1 & (x > 0) \end{cases} \quad (4.4)$$

ただし, 存在しない画素を含む項の sgd は 0 とする.

なお, 全変動を 2次元に拡張して離散化する際に 4近傍を用いる手法は比較的一般的であるが, 輝度勾配には L^2 ノルムを利用している [50, 53]. その結果, 第 2.4.3 節でも述べたように劣微分係数が輝度に依存するなどの問題が生じる. 提案手法は劣微分係数が 4近傍における輝度値の大小比較だけによって決定されるという特徴がある.

4.2.2 劣勾配法の適用

改めて, 全変動最小化の定式化を行う. 先に述べたように, 本論文では ROF モデルを利用する. 従って, 目的関数は

$$\min_{u \in X} \frac{\|u - g\|^2}{2\lambda} + J_{\text{tv1}}(u) \quad (4.5)$$

となる. 式 (2.1) と比較すると, 利用している全変動が異なる. この目的関数を計算するために, Combettes らと同様に劣勾配法を利用する. なお, 第一項の忠実化項は勾配法となる.

表 4-1: Characteristics Comparison between Discrete Methods of Total Variation

Methods / Characteristics	Iteration	Discritization	Computational cost
Proposed method	explicit subgradient	L^1 norm	low
Method 1 by Chambolle	semi-implicit gradient	L^2 norm	partial low
Method 2 by Combettes	explicit subgradient	L^2 norm	high

オイラー方程式を計算すると、勾配法による反復処理は

$$\mathbf{u}^{n+1} = \mathbf{u}^n - w^n (\mathbf{u}^n - \mathbf{g} + \lambda \partial J_{\text{tv1}}(\mathbf{u}^n)) \quad (4.6)$$

$$= (1 - w^n) \mathbf{u}^n + w^n \mathbf{g} - w^n \lambda \partial J_{\text{tv1}}(\mathbf{u}^n) \quad (4.7)$$

となる．ここに、 w^n は更新における重みで、反復処理が収束するように指数関数的に小さくする．

式 (4.7) の各項の役割は、以下のように解釈できる．第一項及び第二項は骨格成分 \mathbf{u} が入力画像 \mathbf{g} から離れすぎないようにする忠実化項である．第三項は劣微分係数を減算するベクタであるから、微分可能な点において一律に平滑化して骨格成分にする平滑化項である．

定義より明らかなように、提案した離散全変動の劣微分は、大小比較だけで決定されるために計算コストが低い．また、FFT などを含む射影関数がないため、1 回の反復処理における計算コストも低くなる．なお、定量的な評価は実験により行う．

表 4-1 に提案手法と従来手法の特徴まとめる．提案手法は反復処理における各項の役割を陽にしたままで数値計算を可能としている．また、輝度勾配にも L^1 ノルムを利用して計算を実現している．さらに、浮動小数点の計算回数低減により、計算量も削減される．これらの結果、エッジ強調など他のフィルタを加えた反復法の構成も可能になり、適用範囲が広がることが期待される．

これまでに劣勾配に局所微分を用いる手法も提案されている [50] が、提案手法は以下の優位性がある．第一に、除算がないため、0 除算を避けるための微少な定数等を加える必要がない．第二に、輝度勾配に L^2 ノルムを利用していないため、平方根の計算が全く必要なく、1 回ごとの計算コストも低い．収束速度が速い理由は、重み w を指数関数的に小さくしていることにあると考えられる．

4.3 線形勾配領域の抽出

本節では、自然画像から等輝度領域と線形グラデーション領域を抽出し、ベクタ変換による画像符号化を目的とする．提案手法の流れを図 4-1 に示す．勾配画像 (Gradient Images) の生成は方向 θ 毎に行う．入力画像から x 方向と y 方向の微分を算出し、さらに方向 θ に対する勾配値を算出する．入力画像と勾配画像に対してそれぞれ粗階調画像 (Posterized Image) を生成する．

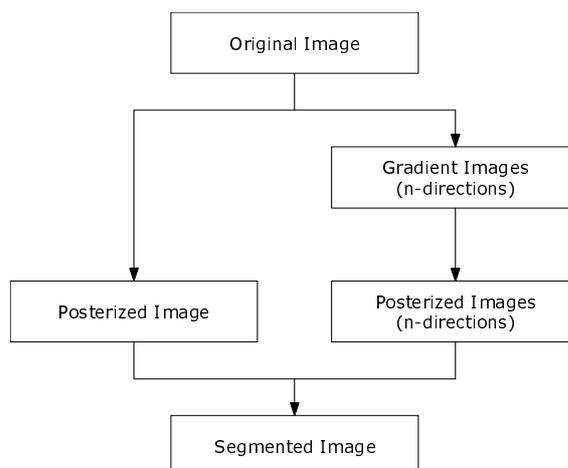


図 4-1: Posterization Procedure (自発表 [54] より引用)

ここで、ある領域の輝度が線形変化している場合、ある方向に対して勾配値が一定になると考えられる。そこで、入力画像の輝度値と方向毎の勾配値をそれぞれ粗階調化し、等輝度領域と線形グラデーション領域を抽出する。

得られた粗階調画像を合わせてクラスタリングし、等輝度領域及び線形グラデーション領域が抽出可能となる。線形グラデーション領域については、領域ごとにグラデーションの方向と大きさを最小二乗法により再計算する。また、除去されたテクスチャを含む残差はラスタ符号化を適用する。なお、本手法はカラー画像に容易に拡張できる。

4.4 グラデーションメッシュ

近年、ベクタ画像描画ソフトウェアではグラデーションメッシュと呼ばれる写実的な (photo-realistic) ベクタアートを作成する強力なツールが提供されている。オブジェクトを四角形メッシュに分割し、頂点に色を与えると同時にメッシュ内を滑らかに塗りつぶす手法である。

本論文では、画像符号化の基礎ツールとして、グラデーションメッシュの有効性を確認する。すなわち、人手をかけて作成するのではなく、与えられたラスタ画像をグラデーションメッシュに変換する手法を検討する。具体的にはグラデーションメッシュ生成における最適化手法として Fletcher-Reeves の共役勾配法 (FR 法) と Levenberg-Marquardt 法 (LM 法) を比較する [26]。

サンプリングによる色の誤推定を避けるために、数値計算により決定する手法を提案する。さらに色の接線ベクタも省略して、色行列を

$$Q_c = \begin{bmatrix} c^0 & c^1 & 0 & 0 \\ c^2 & c^3 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad (4.8)$$

とおく。色行列のヤコビ行列も解析的に計算できる。

画像符号化のツールとしては、オブジェクトの切出しをせずに適用できる方が扱いが容易になる。そこで画像全体をグラデーションメッシュに変換する。NLLS 問題の初期解として、画像を等分割してメッシュ点を配置する。

ヤコビ行列を用いた計算を回避するために共役勾配法の一つである FR 法を適用し、LM 法と比較を行う。FR 法では評価関数の微分行列が必要であるが、要素数はヤコビ行列と比較して大幅に削減できる。例えば、対象画像の場合およそ 3000 個である。しかしながら、大局的な最適解を得られるかは初期値に大きく依存している。そこで、実験により初期値の依存性を明らかにする。また、計算コストの比較を行う。

4.5 提案手法の評価

4.5.1 全変動最小化の高速計算手法の評価実験

性能評価項目と実験条件

提案手法の妥当性を示すために、全変動最小化を用いて静止画像から骨格成分を取得する実験を行う。ここで、本論文では第 2.4 節で述べた Chambolle らの画像分解手法（以下、従来手法 1）と Combettes らの画像復元手法（以下、従来手法 2）とにより得られる骨格成分を正解と見なして、アプリケーションを特定せずに骨格成分だけを評価する。そこで評価項目として、 L^1 全変動、 L^2 全変動、入力画像忠実度の 3 指標における一致度と骨格成分近似度を採用する。 L^1 全変動と L^2 全変動は得られた骨格成分の全変動を測る。入力画像忠実度は入力画像と得られた骨格成分との PSNR を測る。骨格成分近似度は、提案手法と従来手法のそれぞれから得られた骨格成分同士に対して、画像の構造についても評価できるといわれる Structural SIMilarity (SSIM) 指標を用いる [55]。なお、全変動と入力画像忠実度は提案手法と従来手法 1, 2 における目的関数や制約条件に該当する。最後に、提案手法より削減された計算量を、実際にかかった計算時間を比較して考察する。

対象画像として自然画像 2 枚、人工画像（アニメーションに含まれる静止画像）2 枚の計 4 枚を用いた。画像はすべてグレースケール (8bit/pel) である。各画像の名称、解像度、 L^1 全変動 (L^1 -TV)、 L^2 全変動 (L^2 -TV) を表 4-2 に示す。また、以下の表に示す全変動はすべて 1 画素あたりの値である。不連続点を含む画像を正規化可能な全変動により、平坦領域ではノイズに起因する振動成分が除去されること、急峻なエッジが骨格成分として保存されることを検討する。

自然画像 2 枚は Lenna と Barbara を用いた。これらの画像は自然画像の符号化効率評価として一般的に利用されている。背景と人物の境界における急峻なエッジや、細かいテクスチャ、肌のようなグラデーション領域といった様々な特性を持ち合わせている。Lenna は肌のグラデーションと髪飾りのテクスチャ、Barbara は服やテーブルクロスの方角のテクスチャに特徴がある。また、Anime 1 はキャラクタの全体像が描かれており、多数の小さな平坦領域から構成されている。Anime 2 はキャラクタの顔が背景と共に描かれており、少数の大きな平坦領域と複雑な背景から構成されている。Anime 1 と 2 の 240 行目における輝度

表 4-2: Characteristics of Target Images.

Name	Size [pixel]	L^1 -TV	L^2 -TV
Lenna	512x512	11.32	8.93
Barbara	512x512	22.36	17.86
Anime 1	704x480	10.38	8.46
Anime 2	704x480	9.01	7.24

変化を図 4.5.1, 図 4-2 に示す.

実験に用いた計算機環境は, Intel Core 2 Duo (E8500) 3.16GHz, メモリ 8GB, 64bit Linux, Intel C++ Compiler 11.0 である. FFT 及び逆 FFT の計算には, FFTW3 ライブラリを利用した [56].

従来手法 1 に対する性能評価結果

従来手法 1 (Chambolle らの画像分解手法) により得られる骨格成分を正解とする場合の性能評価を行う. 従来手法 1 と提案手法におけるパラメータは $\lambda = 8$ で, 反復回数は 10 回である. 以下, この条件を提案手法 1 と呼ぶ.

提案手法 1 の性能評価結果を表 4-3 に示す. 左列から入力画像の名称と手法, L^1 全変動 (L^1 -TV), L^2 全変動 (L^2 -TV), 入力画像忠実度 (PSNR), 骨格成分近似度 (SSIM), 画像の入出力を除く計算時間 (Time) である.

この表より, L^1 全変動, L^2 全変動, 入力画像忠実度の 3 指標は誤差 3% 以下であり, かつ骨格成分近似度は 0.994 以上であることがわかる. 従って, 提案手法 1 から得られた骨格成分は従来手法 1 から得られたそれに十分近似していることが明らかとなった.

さらに, 提案手法 1 による計算時間は従来手法 1 と比べて, 53% ~ 58% と約半分に低減されている. 我々の実装において 1 画素ごとの浮動小数点計算のうち, 乗算が 7 回から 2 回へ, 除算や平方根計算が 2 回から 0 回に削減されたことが要因である.

Lenna について, 入力画像, 提案手法 1 と従来手法 1 により得られた骨格成分を図 4-4 に示す. Anime 2 について提案手法 1 と従来手法 1 により得られた骨格成分の輝度変化の一部を図 4.5.1 に示す.

これらの図より, 提案手法 1 は従来手法 1 と同様に, 自然画像と人工画像の両方においてテクスチャや平坦領域に含まれるノイズを平滑化する一方で, エッジを比較的保存していることが確認できる. ただし, 人工画像に含まれていた急峻なエッジについて, 少なからず滑らかになっている. 反復処理のどの項にもエッジ強調をする要因が含まれていないことが原因ある.

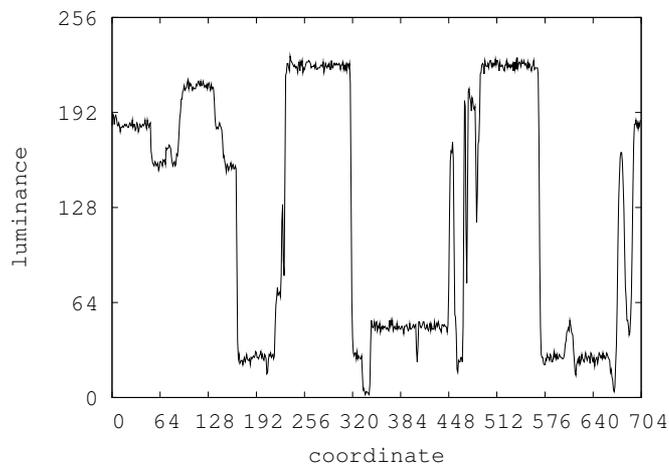


図 4-2: Anime 1 の 240 行目の画素位置と輝度値の関係 (自発表 [22] より引用)

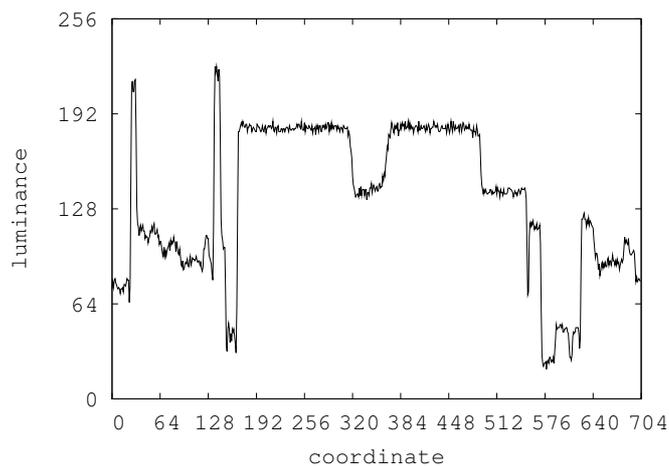


図 4-3: Anime 2 の 240 行目の画素位置と輝度値の関係 (自発表 [22] より引用)



図 4-4: 骨格画像の抽出結果 左から入力画像 Lenna, 提案手法 1 による結果, 従来手法 1 による結果 (自発表 [22] より引用)

表 4-3: 従来手法 1 に対する提案手法 1 の性能評価結果

Image / methods	L^1 - TV	L^2 - TV	PSNR [dB]	SSIM	Time [msec]
Lenna					
Prop. 1	5.79	4.82	35.9	0.996	59.8
Conv. 1	5.79	4.70	35.9	-	103.3
Barbara					
Prop. 1	11.19	9.42	30.7	0.994	59.5
Conv. 1	11.59	9.43	30.9	-	102.2
Anime 1					
Prop. 1	6.07	5.23	38.2	0.996	68.0
Conv. 1	6.03	5.11	38.2	-	128.4
Anime 2					
Prop. 1	4.92	4.16	38.9	0.997	73.6
Conv. 1	4.86	4.04	38.9	-	128.4

従来手法 2 に対する性能評価結果

従来手法 2 (Combettes らの画像復元手法) により得られる骨格成分を正解とする場合の性能評価を行う。従来手法 2 のパラメータは文献 [24] による。一方, 提案手法におけるパラメータ λ は, 入力画像忠実度 (PSNR) がほぼ一致するように決定した。以下, この条件を提案手法 2 と呼ぶ。

ところで, 本論文では ROF モデルにおける全変動最小化の計算手法を提案しているが, 画像復元における式 (2.7) の J_{tv2} は J_{tv1} に置換できる。ただし, FFT を含む射影関数は計算コストが高く, 高速計算に適さない。以下, この手法を提案手法 3 と呼び, 結果を参考値として示す。

提案手法 2, 3 の性能評価結果を表 4-4 に示す。左列から入力画像の名称と手法, L^1 全変動 (L^1 -TV), L^2 全変動 (L^2 -TV), 入力画像忠実度 (PSNR), 骨格成分近似度 (SSIM), 反復回数 (Iteration), 画像の入出力を除く計算時間 (Total time) である。さらにその右側に劣勾配法が占める計算時間と反復処理 1 回あたりの計算時間, FFT などを含む射影関数の計算時間とその反復処理 1 回あたりの計算時間を示した。

提案手法 2 は従来手法 2 と計算方針が大きく異なるが, Barbara を除いて骨格成分近似度は 0.982 以上であるため, 得られた骨格成分同士はかなり近似していると考えられる。Barbara の骨格成分近似度が低いのは, 画像分解に用いるパラメータ λ の大きさに比べてテクスチャの振幅が大きいため, テクスチャを十分除去できないことが原因である。さらに提案手法 2 の L^1 全変動と L^2 全変動は, 従来手法 2 のそれらと比べて低下している。従って, 提案手

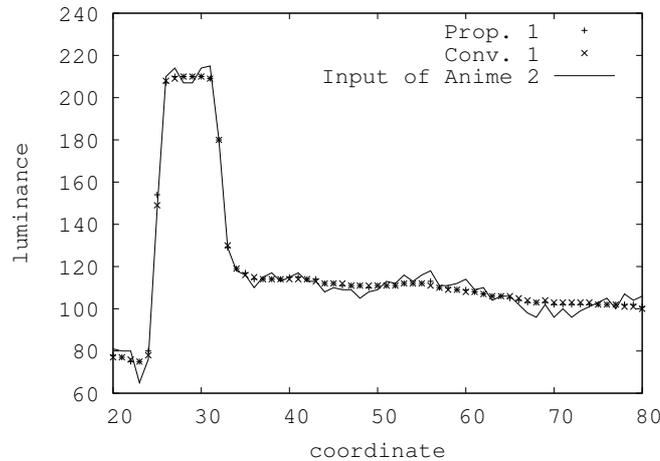


図 4-5: Anime 2 の 240 行目の画素位置と輝度値の関係．提案手法 1 と従来手法 1 によって得られた骨格成分と入力画像の比較（自発表 [22] より引用）

法 2 は振幅の小さなノイズやテクスチャに関する平滑化性能が高いと考えられる．

提案手法 2 による計算時間は従来手法 2 と比べて，5%～7%程度と大幅に削減されている．FFT を含む射影が必要なく，反復回数が削減されたためである．

参考値として結果を示した提案手法 3 について， L^1 全変動， L^2 全変動，入力画像忠実度の 3 指標は誤差 5%以下であり，かつ骨格成分近似度は 0.988 以上であることがわかる．従って，提案手法 3 から得られた骨格成分は従来手法 2 から得られたそれに十分近似していることが明らかとなった．

さらに，提案手法 3 の劣勾配法における 1 回あたりの計算時間は従来手法 2 と比べて，65%～70%に低減されている．我々の実装において 1 画素ごとの浮動小数点計算のうち，乗算が 5 回から 1 回へ，平方根計算が 1 回から 0 回に削減されたことが要因である．ただし，提案手法 3 の反復回数は従来手法 2 と比べて増減するため，全体計算時間は必ずしも低減しない．また，FFT などを含む射影関数にかかる計算時間が大部分を占めるため，計算量低減の影響もほとんど見られない．

Lenna について，提案手法 2，提案手法 3，従来手法 2 により得られた骨格成分を図 4-6 に示す．Anime 1 について，提案手法 2，提案手法 3，従来手法 2 により得られた骨格成分の 240 行目の輝度変化の一部分を図 4-7 に示す．

これらの図より，提案手法 2，提案手法 3 は従来手法 2 と同様に，自然画像と人工画像の両画像でテクスチャや平坦領域に含まれるノイズを平滑化する一方で，エッジの保存が確認できる．また，提案手法 1 と比較して提案手法 2 や提案手法 3 の方が，強く平滑化していることが分かる．

表 4-4: 従来手法 2 に対する提案手法 2, 3 の性能評価結果

Image / methods	L^1 - TV	L^2 - TV	PSNR [dB]	SSIM	Iteration [times]	Total time [msec]	Subgradient computa- tion [msec] ([msec/ite.])	Projection with FFT computation [msec] ([msec/ite.])
Lenna								
Prop. 2	4.69	3.97	33.0	0.985	30	178.3	-	-
Prop. 3	5.35	4.47	33.1	0.989	53	2772.2	270.7(5.11)	2194.3(41.40)
Conv. 2	5.37	4.35	33.1	-	50	2693.6	368.7(7.37)	2035.8(40.72)
Barbara								
Prop. 2	4.95	4.25	25.3	0.966	30	180.7	-	-
Prop. 3	5.65	4.78	25.1	0.988	50	2514.9	259.3(5.19)	1968.1(39.36)
Conv. 2	5.85	4.76	25.2	-	54	2810.1	404.0(7.48)	2104.7(38.98)
Anime 1								
Prop. 2	5.11	4.43	33.6	0.985	30	179.7	-	-
Prop. 3	5.83	5.07	33.7	0.991	49	3374.0	250.2(5.11)	2738.7(55.89)
Conv. 2	5.71	4.85	33.7	-	51	3680.5	391.4(7.67)	2888.6(56.64)
Anime 2								
Prop. 2	4.16	3.54	34.7	0.982	30	188.4	-	-
Prop. 3	4.77	4.06	34.6	0.991	54	3725.2	274.9(5.09)	3044.1(56.37)
Conv. 2	4.77	3.98	34.8	-	51	3650.3	396.1(7.77)	2858.4(56.05)



図 4-6: 骨格画像の比較 左から提案手法 2, 提案手法 3, 従来手法 2 (自発表 [22] より引用)

4.5.2 線形勾配領域の抽出実験

入力画像として自然画像 Lenna を用いた場合の, テクスチャ分離結果と粗階調画像を図 4-8 に示す. なお, 粗階調画像は値が完全に一致する連結領域を抽出し, その領域の面積がしきい値以上 (本実験では 256 画素とした) となる領域だけを示した. また, $\pi/4$ 毎に 4 方向の勾配画像を生成し, 勾配値を 4 倍にして 128 を加えてある. さらに, $q_{gstep} = 1/4$ とした.

ベクタ表現に適した領域が画像に占める割合を図 4-5 に示す. 対象画像は自然画像 Lenna と Barbara であり, 等輝度領域だけ (Homo. only) と線形グラデーション領域を加えた場合 (Homo. and Grad.) とを比較する.

これらの結果より, テクスチャ除去により等輝度領域が多く抽出できていることが確認できる. さらに, 線形グラデーション領域を追加したことにより, ベクタ表現に適した領域

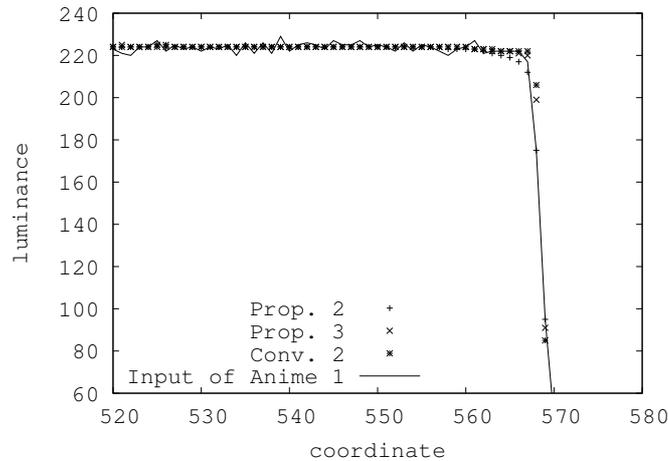


図 4-7: Anime 1 の 240 行における画素位置と輝度の一部分．提案手法 2, 3, 従来手法 2 により得られた骨格成分と入力画像の比較（自発表 [22] より引用）

表 4-5: Cover Ratio of Segmented Regions.

Image	Homo. only	Homo. and Grad.
Lenna	74.0%	90.5%
Barbara	74.8%	93.9%

（カバー率）が増加していることが確認できる．

4.5.3 グラデーションメッシュによる画像符号化

既知のグラデーションメッシュとして図 4-9 を，自然画像として図 4-12 を使用した．図中の青線は Ferguson patch の境界線，赤線と黄線は接線ベクタである．入力解像度はそれぞれ 48×48 画素， 128×128 画素，初期メッシュ点はそれぞれ 5×5 点， 17×17 点である．どちらも LM 法における反復回数は 30 回である．

既知のグラデーションメッシュにおける比較を行う．結果を図 4-10 に示す．そもそもこの図は，メッシュの幅を狭くするとエッジを表現できることを示している．実験結果からも図の中央上部と中央下部ではエッジが再現されている．また，入力画像と同様に滑らかな色変化も実現されている．接線ベクタは入力画像と比較して回転したり，極端に短くなっている場合が見受けられる．

自然画像の結果，メッシュ点を重畳した結果，媒介変数を $1/8$ 刻みで変化させたサブメッシュを重畳した結果をそれぞれ図 4-11 に示す．メッシュ点が十分に無いため，髪や帽子などのこまかいテクスチャは完全に平滑化されている．同様の理由で目や唇など細かいパーツも平滑化されている．一方主観的に比較する限り，エッジの再現性は非常に高い．サブメッシュがエッジ付近で大量に描画されていることから確認できる．なお，パラメータをテ

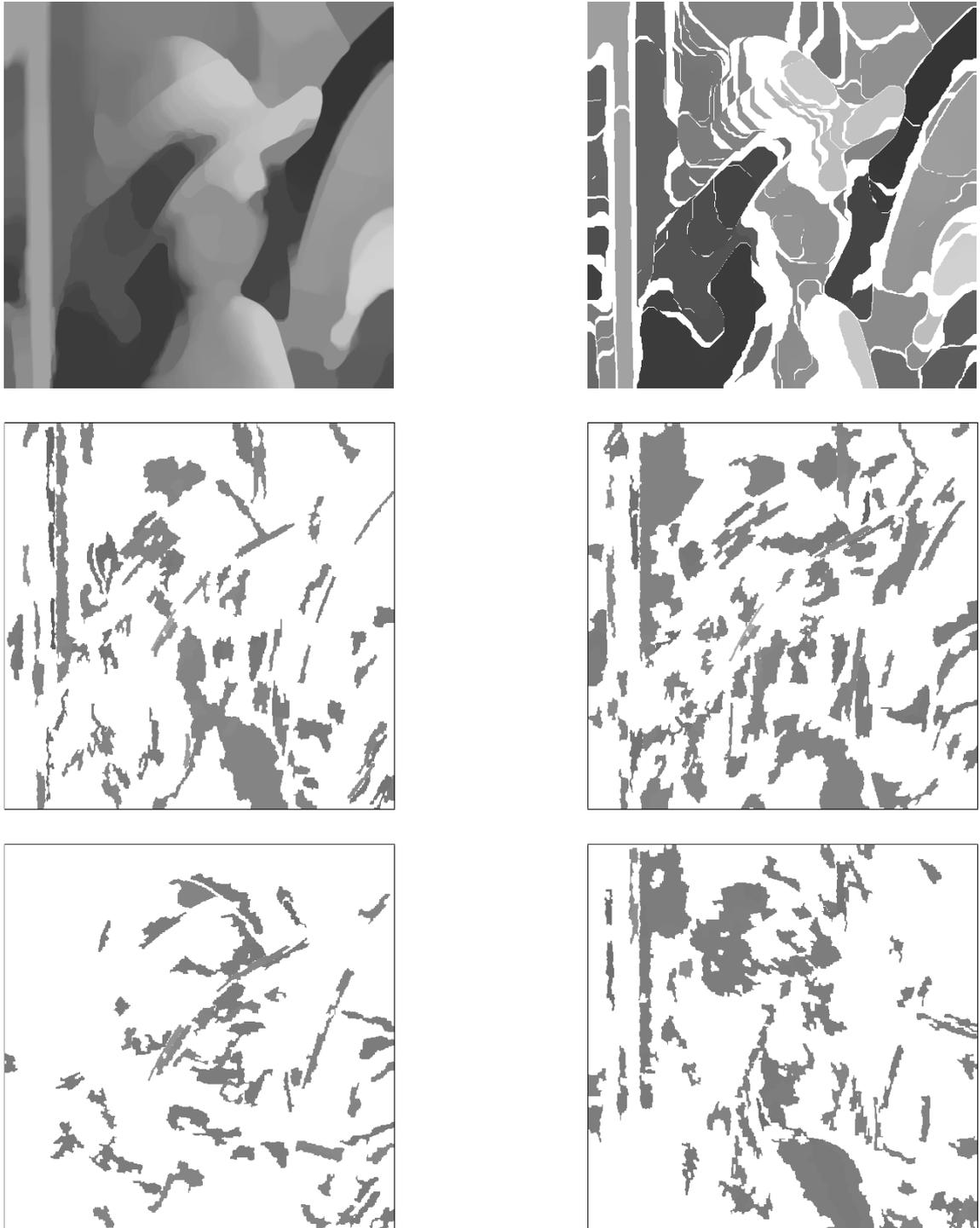


図 4-8: 平滑化画像と領域分割の結果 左上から全変動最小化による平滑化画像, 均一輝度領域, 角度 $0, \pi/4, \pi/2, 3\pi/4$ 方向のグラデーション領域 (自発表 [54] より引用)

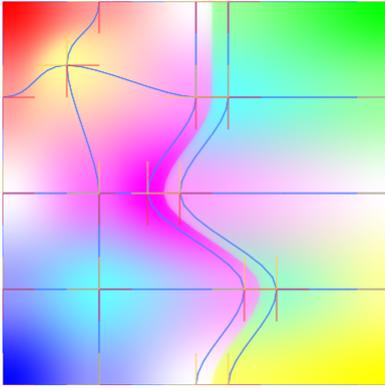


図 4-9: グラデーションメッシュのテスト画像 [25].

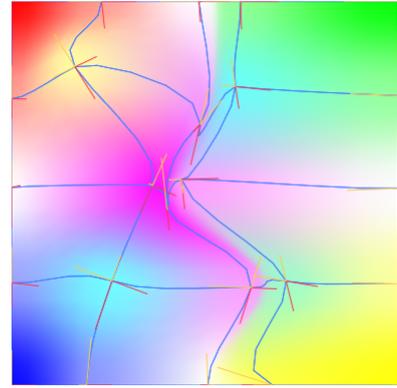


図 4-10: 生成されたメッシュデータ (自発表 [26] より引用)

キスト形式で出力し、汎用テキスト圧縮ツール (ZZIP) により圧縮した場合約 5.7KB、図 4-12 を JPEG 方式で符号化した場合約 4.7KB であった。

対象画像として図 4-12 を使用した。入力解像度は 128×128 画素、初期メッシュ点は 17×17 点であり、均等に配置した。図中の青線は Ferguson patch の境界線、赤線と黄線は接線ベクタである。

FR 法、LM 法による結果を図 4-12 にそれぞれ示す。FR 法は、メッシュ点が初期点からほとんど移動しておらず、近傍にエッジがある場合だけ接線ベクタが最適化されている。これは、サブメッシュを表示した図 4-12 から分かる。反復処理においても、収束せずに局所解に陥っていた。一方 LM 法は、多くのメッシュ点が移動し、かつ接線ベクタも最適化されている。

FR 法における反復回数は 26 回、LM 法における反復回数は 14 回である。1 回の反復処理について平均して、FR 法では 23 秒、LM 法では 3400 秒かかった。

4.6 むすび

本章では、人工画像の形式に関してグレースケール画像やカラー画像を対象として、二つのモデルを検討した。一つは入力画像を、ベクタ表現可能な線形勾配 (グラデーション) 領域・成分と、残差であるスカラー表現に適したテクスチャ成分に分離するモデルである。もう一つは、画像全体をグラデーションメッシュ成分とテクスチャ成分に分離するモデルである。

均一輝度成分を抽出する全変動最小化 (Total Variation Minimization) のフレームワークにおいて、高速計算手法を提案した。提案手法の特徴は、第一に 2 次元画像の離散全変動として平方根計算が必要ない L^1 ノルムを適用したことである。第二に双対問題や射影計算を経ることなく劣勾配法を用いて骨格成分を計算したことである。自然画像と人工画像を用いた骨格成分取得の実験により、本手法は従来手法と同等の性能を維持したまま高速計算を実現していることを示した。計算時間は画像分解手法と比較して 53% ~ 58% と約半分に、画像復元手法と比較して 5% ~ 7% へと大幅に削減された。また、反復処理を構成する項は、骨

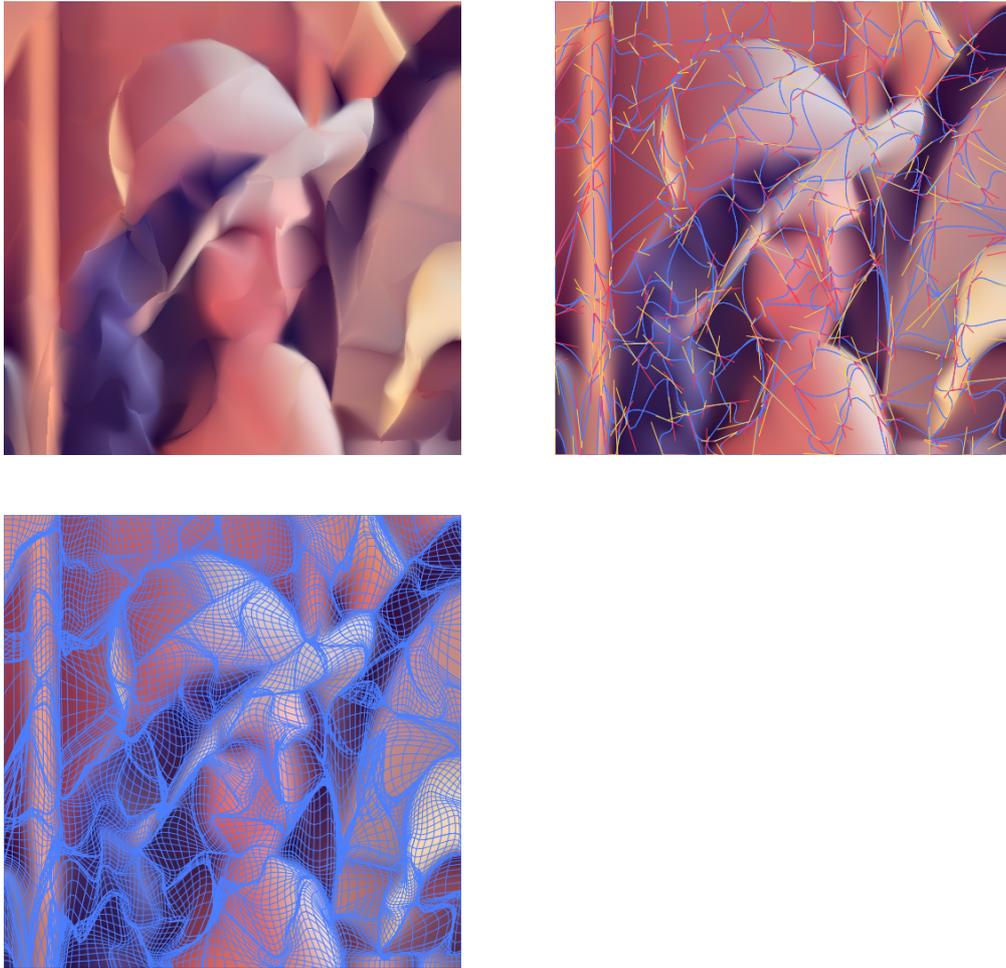


図 4-11: グラデーションメッシュによる近似結果 左から再現画像, メッシュを重畳, サブメッシュを重畳 (自発表 [26] より引用)

格成分を入力画像に近づける忠実化項と, 全変動の観点から骨格成分を平滑化する平滑化項であることを示した.

次に, ベクタ変換の対象となる均等輝度領域と線形グラデーション領域を抽出する手法を提案した. 実験により, 均一輝度領域に線形グラデーション領域を加えることでベクタ表現可能な領域 (カバー率) は 90% 以上となることを明らかにした.

さらに線形グラデーション成分をも内包するグラデーションメッシュ (Gradient Mesh) により画像全体を近似する手法について, 画像符号化の基礎ツールとしての有効性を画像特性と計算量の観点から検討した. 評価実験により, テクスチャや細かいパーツは平滑化される一方で, 滑らかな色変化 (グラデーション) やエッジの制限性が高いことを確認した.



図 4-12: 最適化手法の比較．左上から入力画像，FR 法の結果，LM 法の結果，LM 法の結果にサブメッシュを重畳（自発表 [26] より引用）

第5章

曲線の不変特徴量と編集性

5.1 まえがき

本章では、人工画像のベクタ化対象の中核となる平面曲線（線画）に着目する。まず、平面曲線において曲率に関する特徴量を整理する。これらは特徴点の探索が必要であること、スケール不変でないこと、開曲線には適用できない、または精度が著しく低下することを述べる。

濃淡画像においてこれらの問題点を解決した特徴点を抽出する SIFT アルゴリズムを紹介し、これを曲率に援用する手法を提案する。すなわち、スケール正規化曲率を導出し、さらに異なるスケール間で曲率差を算出し、その極値を最適スケールと定義する。これが平面曲線におけるスケール不変特徴量となる。

得られた特徴量は、適当なスケール範囲において曲率が大きな部分、すなわち曲線の頂点の定義に用いる。ここで、曲線のベクタ形式であるベジエ曲線は、曲線の頂点に通過点が存在していると編集しやすいことを述べる。また、多角形近似を経るベクタ変換手法において、これまで提案している多角形の辺ごとに曲線に置き換える手法は、編集しやすいベジエ曲線を生成することを明らかにする。なお、編集性については第 2.2.4 節に詳述している。

最後に、提案するスケール不変特徴量の安定性や頑健性を MPEG-7 のコア実験で規定された形状データベースを用いて評価する。また、提案しているベクタ変換により得られるベジエ曲線の編集しやすさを評価し、ベクタ変換の特徴を明らかにする。

5.2 既存のスケール不変曲線特徴量

スケール不変特徴量は画像検索や形状分類の分野で幅広く利用されている。平面曲線の曲率は基本的な特徴量であり、一般的には座標系によらず不変である。一方、多重解像度解析を適用すると、曲率に基づく特徴点の位置が変化する。そのため、特徴点を特定するためにスケール（解像度）の認識が重要である。

これまで曲率に基づく多くの形状記述子がパターン認識やコンピュータビジョンの分野で開発されてきた。曲率スケール空間（Curvature Scale Space; CSS）は曲線素片には適用できず、さらに特徴点の探索が必要という課題がある。濃度勾配に基づく曲率の導出は、勾配計算時にスケールを考慮するにもかかわらず、曲率値がスケールではなく曲線長で正規化する

という課題がある．スケール不変特徴変換 (Scale-Invariant Feature Transform; SIFT) アルゴリズムは画像から特徴点を特定するが [57]，それは微分ガウス (Difference of Gaussian; DoG) を用いて類似の問題を解決している．ただし，SIFT アルゴリズムを平面曲線へ特徴点検出のために単純に適用できない．

本章では，曲線素片に対する自動スケール特定手法を述べる．提案方式は適切なスケールと特徴点の位置の探索なしで，微分曲率 (Difference of Curvature; DoC) に基いて特定する．微分計算のために，スケールを正規化した曲率を導入する．DoC アルゴリズムの優位性は適切なスケールを曲線素片から局所特徴量として抽出可能な点にある．その結果，適用範囲の拡大が期待される．提案方式の妥当性は実験により確認される．曲率スケール空間法や勾配に基づく曲率計算などの方法の中で，提案方式は最も安定かつ頑健な特徴点を与える．

スケール不変特徴は画像検索や形状分類で幅広く利用されている．画像に対する SIFT アルゴリズムは内容に基づく画像検索や画像分類において非常に有効な技術のひとつである．SIFT アルゴリズムは画像の縮尺と回転に対して不変である．一方，形状検索や分類には輪郭や色，領域内の模様などの特徴量を用いる．これらの特徴量と記述子は MPEG-7 で標準化され，Multimedia Content Description Interface と呼ばれる [58–60]．本章では 2 次元形状の輪郭から得られる平面曲線を検討対象とする．

輪郭形状に対して多くの形状記述子がパターン認識やコンピュータビジョンの分野で開発されてきた．その初期において，Witkin らは 1 次元平面曲線に対するスケール空間フィルタリングを提案した [61]．平面曲線を表す関数 $f(x)$ は，偏差 σ^2 が小さな値から大きな値へと変化させたいくつものガウス関数と畳み込まれる．畳み込み後のそれぞれの関数の 2 次微分のゼロ交差が抽出され， $x - \sigma$ 平面中に記録される．得られた平面は関数のスケール空間画像である．この手法は平面曲線の構造特定に効果的であるが，探索と記述，特徴点の比較は負荷の高い問題である．

2 次元平面曲線について Mokhtarian らは曲率スケール空間 (CSS: Curvature Scale Space) を導入した [62]．輪郭形状の媒介変数表現はガウス関数と畳み込まれる．得られた曲線から曲率のゼロ交差が抽出され， $t - \sigma$ 平面に記録される．ここで， t は媒介変数である．CSS 表現は本質的に形状の回転や均一な拡大，移動に対して不変である．この表現は MPEG-7 の開発段階においてさらなる拡張や最適化がなされた [59]．MPEG-7 の輪郭形状記述において，ローパスフィルタの適用回数は輪郭線長に対して非線形正規化される．スケール不変を保証するために正規化された値によってスケール値は定義される．その結果，本手法は輪郭素片に対して適用できない．CSS の方針は継続的に改良されている．Zhong らは平面曲線ではなくその曲率とガウス核の畳み込みによって定義される，DCSS (Direct Curvature Scale Space) と呼ぶ方式を提案している [63]．また，Awrangjeb らは円弧上ではなくアフィン上の媒介変数を用いる改良 CSS 角検出器を提案している [64]．これらの方法は形状全体の円弧長に対しても，部分形状の変形に対しても独立していない．

Daliri らは勾配に基づく曲率 (GbC: Gradient-based Curvature) 計算を提案している [65]．輪郭線の長さについて接線ベクタの変化度合いという平面曲線に対する曲率の定義によれば， x 方向と y 方向に対する接線ベクタは計算できる．すなわち，平面曲線の接線ベクタは

勾配ベクタと直交しており、勾配ベクタは濃度画像と微分ガウス関数との畳込みによって得られる。勾配ベクタは輪郭に沿って最適なスケールで求められるが、曲率は輪郭線長によって正規化される。すなわち、曲率に対する適切なスケールは検出できない。

既存の形状記述子は上述したように輪郭素片に適用できない。さらに CSS 表現において、特徴点の追跡は必須である。SIFT アルゴリズムは類似の課題を、スケール正規化された LoG (Laplacian of Gaussian) の近似であるガウス差分 (DoG: Difference of Gaussian) の利用により解決している。

本章では、輪郭素片に対する自動スケール検出方式を提案する。提案方式は最適なスケールとその特徴点の位置を DoC (Difference of Curvature) に基いて探索なしで検出する。すなわち、DoG の Gaussian の代わりに、スケール正規化された曲率を DoC として用いる。DoC アルゴリズムの優位性は輪郭素片からでも局所特徴として最適なスケールが求められることにある。以下では、最適なスケールと位置をスケール位置と呼ぶ。

閉曲線だけではなく輪郭素片に対するスケール決定は適用領域を拡大する。例えば、全体形状ではなく部分形状を検索のクエリとして受理できる。特徴点の位置関係もまた受理される。ここで位置関係とは単一形状からだけでなく複数形状から得られる特徴点群と、スケールで正規化されたそれらの距離を意味する。適用範囲は輪郭素片に限定されない。複数かつ部分形状が単一形状に合成される時、いくつかの特徴点は合成の影響を受けることなく抽出できる。ここで、特徴点の記述方法については本論文の検討対象外とする。

5.2.1 曲率スケール空間

曲率スケール空間 (CSS: Curvature Scale Space) 表現は輪郭形状の構造を多重スケール解析によって検出する [58, 60]。本表現は形状検索や分類で非常によく使われている。

輪郭形状の CSS 表現を作るために、等間隔の N 点が輪郭から選択される。選択された N 点の x 座標と y 座標の集合は X, Y で表す。次に X と Y にローパスフィルタの繰り返し適用で輪郭は徐々に滑らかになる。平滑化の結果、輪郭が凸形状になるまでに輪郭線の凹部分は徐々に平らになってなくなる。

いわゆる CSS 画像は輪郭発展処理によって構成される。水平軸と垂直軸は輪郭点とフィルタの適用回数に対応する。CSS 画像中のそれぞれの水平線は、 k 回フィルタを適用して平滑化された輪郭線に対応する。それぞれの平滑化輪郭線について、曲率関数のゼロ交差が計算される。曲率ゼロ交差点は輪郭線の凹部と凸部を分割する点である。それぞれのゼロ交差点は水平線上に記録される。

CSS 画像は峰 (peak) によって特徴づけられる。CSS 画像中の突出した峰 (x_{css}, y_{css}) の座標値が抽出される。座標値は y_{css} の値によって降順に並び替えられる。 y_{css} の値は変換されて、量子化される。変換は非線形で、以下のように定義されている。

$$\text{peak} = 3.8 \left(\frac{y_{css}}{N^2} \right)^{0.6} \quad (5.1)$$

ここで、peak は [0,1.7] に切り詰められる。

CSS 表現は最適スケールを求めるにはいくつか問題がある．peak 値は特徴点のスケールを表し，それは画像の大きさに対して不変である．なぜなら，我々の実験において点の個数は固定だからである．その結果，CSS アルゴリズムは輪郭素片に適用できない．適切なスケールもまた検出できない．さらに，特徴点は凹部分または凸部分が消えることを示しているため，曲率の値そのものは利用されていない．

5.2.2 Laplacian of Gaussian

スケール不変特徴変換 (SIFT: Scale-invariant feature transform) アルゴリズムは特徴点を画像から抽出し，特徴を記述する．検出された特徴点は一般的に濃度勾配に基づく角に位置し，さらに画像の回転とスケールに対して頑健である．

適切なスケールを抽出するために，sLoG (scale-normalized Laplacian of Gaussian) 演算子が用いられる．sLoG 演算子は計算負荷が高いため，DoG (Difference of Gaussian) によって以下のように近似される．

$$\text{sLoG} \approx \frac{G(x, y, k\sigma) - G(x, y, \sigma)}{k - 1}, \quad (5-2)$$

ここで， k は隣合うスケールの比で， G はスケールである標準偏差 σ を持つ 2 次元ガウスフィルタである．DoG の値が $\sigma = \sigma_{max}$ において局所最大値となる時，最適スケールは σ_{max} と定義される．画像のダウンサンプリングによって，DoG の計算負荷は著しく削減される．正確なスケールは特徴点の正規化にも使われる．すなわち，スケール不変性は最適なスケールの検出によって実現される．

特徴点抽出のために SIFT アルゴリズムの平面曲線への直接適用は難しい．平面曲線のラスタライズによって得られたビットマップ画像に対して SIFT 方針の間接的な適用は次節に述べるようにできる．

5.2.3 濃度勾配に基づく曲率

Daliri らは曲率計算のために輪郭線に沿った接線ベクタの利用を提案している．接線ベクタは勾配ベクタと直交している． L_x と L_y を画像と微分ガウス関数の畳込み結果と定義する．局所スケールは画像に対する以下の正規化した導出計算によって決定される．

$$G_\lambda = t^{\lambda/2} \sqrt{L_x^2 + L_y^2}, \quad (5-3)$$

$$\lambda = \frac{1}{2}. \quad (5-4)$$

ここで， $G_{1/2}(t)$ の値は $t = t_{max}$ で最大になるとする．最適スケールは $2\pi\sqrt{t_{max}}$ によって定義される． $G_\lambda(t)$ が単調増加するとき， t の最大値は最適スケールと見なせる．このスケールは輪郭に沿った画像のすべての点について計算できる．勾配 $G = (G_x, G_y)$ は最適スケールにおける 2 次元ガウス関数を x と y 方向に用いて計算される．接線ベクタは勾配ベクタに対して直交するため，最適スケールにおける接線ベクタは以下のように定義される．

$$T = (T_x, T_y) = (G_y, -G_x) \quad (5-5)$$

曲率 κ は円弧長 s に関する接線ベクトルの変化量と結び付けられる． x と y 方向に対する接線ベクトルが導出されなければならない．これは単純に接線ベクトルのそれぞれの成分に 1 次元の微分ガウス関数を畳込むことで実現される．

異なる形状は異なる輪郭線長であるから，ガウス関数のために適応的な σ_1 が用いられる．ガウス関数の標準偏差 σ_1 は輪郭線長に関係している．

$$\sigma_1 = \sigma_0 \frac{l}{l_0} \quad (5.6)$$

ここで， l は輪郭線長である．Daliri らによると， $l_0 = 200$ と $\sigma_0 = 3$ を用いた [65]．

接線ベクトルは二つの偏微分成分から成るため，曲率値は以下のように計算される．

$$\|\kappa\| = \sqrt{\left(\frac{\partial T_x}{\partial s}\right)^2 + \left(\frac{\partial T_y}{\partial s}\right)^2} \quad (5.7)$$

勾配に基づく曲率は輪郭素片に適用できない．接線ベクトルはスケールに関して不変であるように見えるが，曲率は輪郭線長に依存している．そのため適切なスケールは得られない．我々の実験において，得られた曲率値が便宜的にスケールであると考え，スケール不変の評価に用いた．さらに多重スケール解析が実現できないため，疎な輪郭形状の構造に対する考慮もまた不十分である．

5.3 スケール不変曲線特徴量

5.3.1 スケール正規化曲率

スケール不変特徴量を実現するために輪郭素片にも適用できる自動スケール検出方式を提案する．パラメトリック平面曲線 $(x(t), y(t))$ の曲率は以下のように定義される．

$$\kappa = \frac{x'y'' - x''y'}{(x'^2 + y'^2)^{3/2}}, \quad (5.8)$$

ここで， x' と x'' は t に関する x の 1 次微分と 2 次微分である．微分はパラメトリック曲線と微分ガウス関数の畳込みによって定義される．ガウス関数は以下のように定義する．

$$G(t, \sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{t^2}{2\sigma^2}\right), \quad (5.9)$$

ここで t と σ は媒介変数と標準偏差である．

この関数には標準偏差が含まれている．本論文では標準偏差をスケールと呼ぶことにする．離散パラメトリック曲線の微分もまたスケールを用いることになる．従って，曲率もまたスケールに依存する．ここで，一次微分を以下のように定義する．

$$x' = x'[t, \sigma] = G'[t, \sigma] * x[t], \quad (5.10)$$

$$y' = y'[t, \sigma] = G'[t, \sigma] * y[t] \quad (5.11)$$

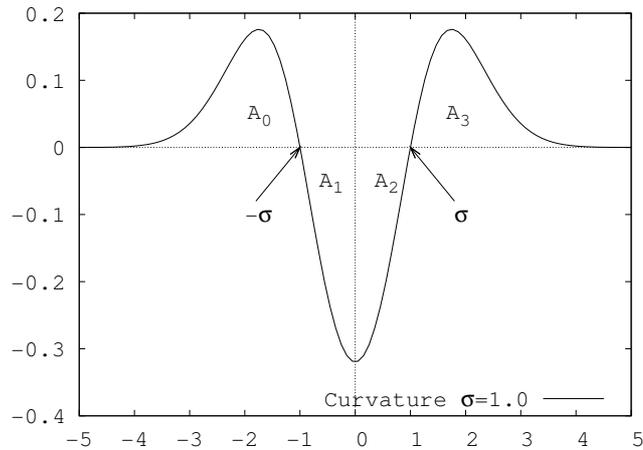


図 5-1: 曲率演算子の応答 (自発表 [66] より引用)

ここで “*” は t に関する畳込み演算子である。

様々なスケールにおける曲率は正規化されるべきである。ここで、以下の曲率演算子を導入する。

$$\frac{G'_x G''_y - G''_x G'_y}{(G'^2_x + G'^2_y)^{3/2}} \quad (5.12)$$

ここで、 G' と G'' はガウス関数の1次微分と2次微分である。この演算子はパラメトリック曲線に作用できる。スケール正規化係数を定義するためにこの演算子を簡略化する。すなわち、 $y[t] = \text{定数}$ とする。この場合、曲率演算子は簡略化され以下で与えられる。

$$\widehat{C}_1 = \frac{G''}{(1 + G'^2)^{3/2}}. \quad (5.13)$$

この演算子の応答を図 5-1 に示す。この図において4つの領域 A_0 と A_1, A_2, A_3 は同一の面積であるとする。この面積をスケール正規化規範として用いる。すなわち、この領域の面積をスケールによらず一定にする係数を選択する。例えば A_2 の面積は以下のように計算される。

$$A_2 = \int_0^\sigma \frac{-G''(t, \sigma)}{(1 + G'(t, \sigma)^2)^{3/2}} dt = \frac{1}{\sqrt{1 + 2\pi e \sigma^4}} \quad (5.14)$$

スケール正規化係数は $\sqrt{1 + 2\pi e \sigma^4}$ と定義される。以下の節では、式 (5.14) に基づくスケール正規化曲率を用い、その係数は曲率値同士の差を計算するために用いる。

ここで、スケール正規化係数には輪郭線長が含まれていないことが重要である。検出された曲率は形状全体の大きさとは独立している。従ってスケール正規化曲率は輪郭素片から局所特徴として取得できる。

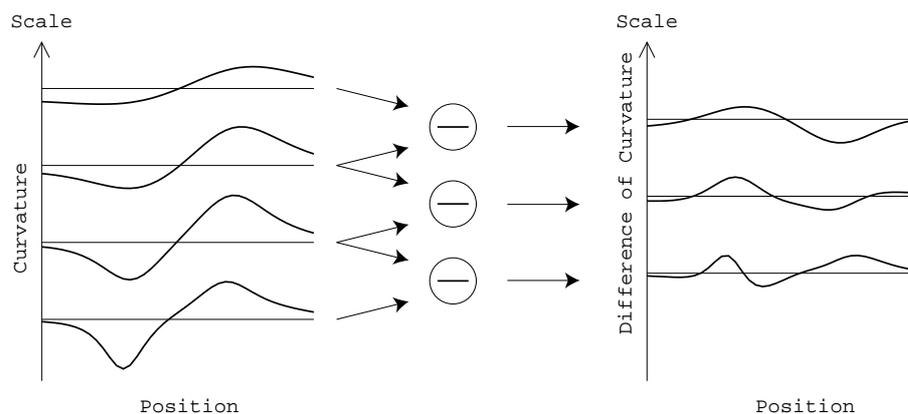


図 5-2: 曲率差 DoC の構築手順 (自発表 [66] より引用)

5.3.2 Difference of Curvature

提案するスケール正規化曲率を用いることで, DoC (Difference of Curvature) によるスケール検出手法を提案する. DoC は SIFT アルゴリズムにおける DoG のアナロジーである. DoC の構築手順を図 5-2 に示す. 初期平面曲線はスケール変数 σ を変えた微分ガウス関数と畳まれる. 図 5-2 の左側に示すように曲率スケール空間の集合を得るためにそれらの曲率が計算される. 図 5-2 の右側に示すように曲率差分 (DoC) を得るために隣合うスケールのスケール正規化曲率を引き算する. スケールが幾何級数的に増加するように, 隣合うスケールの比は一定とする.

DoC の中で極大値または極小値をとる点を特徴点として検出する. 曲率を比較する点を図 5-3 に示す. 横軸は媒介変数 t に基づく位置を表し, 縦軸は DoC の値を表す. それぞれの系列は異なるスケールを表す. x で印付けられた点は丸で印付けられた 8 近傍点と比較される.

極大値や極小値である特徴点のスケールと位置は線形補間を用いて精度を向上させる. 最終的に高精度な特徴点は DoC 空間での探索なしで検出できる. 提案するスケール検出の例を図 5-4 に示す. 一方を他方の 2 倍の大きさにした二つの輪郭線から極大スケールを検出した. 予想された通り σ_2 の値は σ_1 のちょうど 2 倍になっている. ここで, 隣接するスケールの比は予備実験から $2^{1/3}$ とした.

実用上, 再サンプリングやダウンサンプリング技法はノイズの多い輪郭線や巨大な輪郭線に対していくつかの優位性がある [57, 62]. MPEG-7 の形状記述子で用いられている再サンプリングは, 提案方式の中で使われていない. 再サンプリングやダウンサンプリングに起因する位置ずれの補正は大きな計算負荷を必要とする. なぜなら, DoC の計算は異なるスケールの同じ位置で実行しなければならないからである.

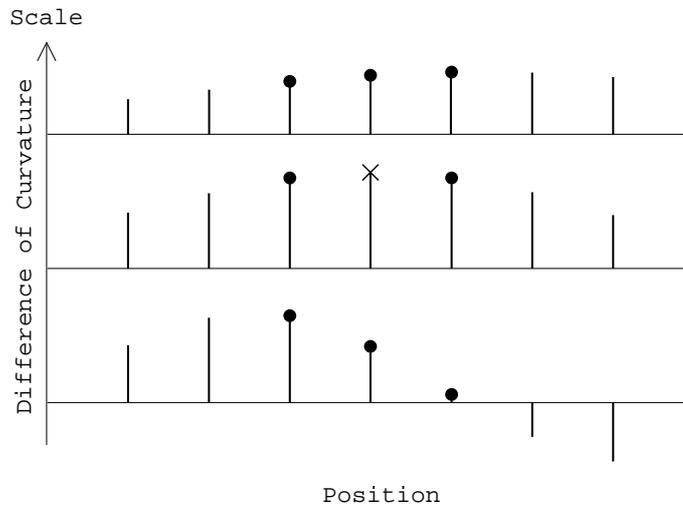


図 5-3: 曲率差 DoC の比較点 (自発表 [66] より引用)

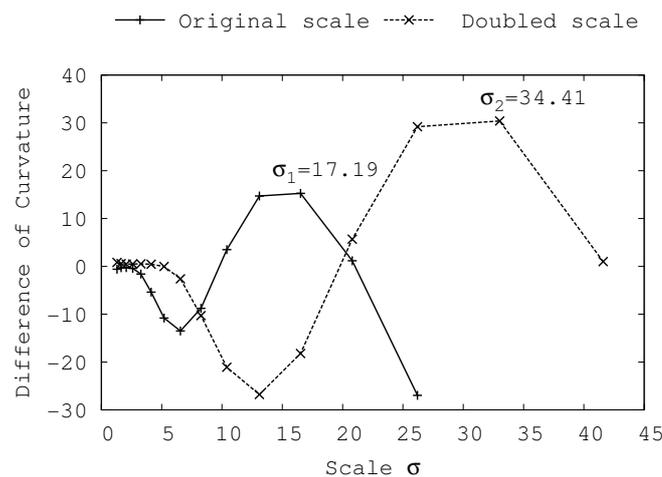


図 5-4: 一方を 2 倍に拡大した図形から検出されたスケールの例 (自発表 [66] より引用)

5.3.3 スケール不変な曲率記述

本論文では、特徴点のための記述方法には言及しない。記述方法は形状分類や形状検索、コンピュータビジョンのようなアプリケーションに強く依存するためである。

そこでいくつかの形状記述方針を述べる。ある特徴点から始めて、輪郭に沿って曲率が得られ、記述可能となる。ここで、特徴点の適切なスケールによる正規化が可能となる。この方針は SIFT 記述子のそれと同様である。他の方針は二つの特徴点の関係を記述である。この場合、二つの適切なスケールのうち一つを用いて、曲率値と二つの特徴点の距離を正規化する。特徴点は単一の形状から抽出されるが、これに限定はされない。すなわち、複数の形状から任意の特徴点の組合せが許される。

5.4 曲線の頂点とベジエ曲線の編集性

5.4.1 ベクタ変換手法の概要

これまでに蓄積されている多くの画像コンテンツはラスタ形式であるため、これを自動的にベクタ表現へと変換するベクタ変換技術が必須である。変換後のベクタ表現は多くの場合、人手による微調整が必要となるので、再編集が容易であることが重要である。

これまでのベクタ変換では、与えられたデジタル曲線の特徴点を抽出してベクタ変換を実現する方針や、誤差の最大値が $1/2$ 画素以下であることを保証する高精度なベクタ変換を実現する方針がある。安本らはデジタル曲線から支配点 (Dominant Point) と曲率の大きな点を検出して、B スプライン曲線を得る手法を提案している [67]。これらの特徴点は編集性に大きな影響を与えるため十分考慮されなければならない。しかし、変換精度や効率が不十分であるという課題がある。一方、Selinger らはグラフ理論を用いて高精度な多角形近似を取得し、これを 3 次ベジエ曲線を用いて滑らかなベクタ表現に変換する手法を提案している [7]。しかし、ベジエ曲線の編集性を十分考慮した変換ではないという課題がある。

そこで、我々は Selinger らの多角形近似に基づいて、再編集が容易なベクタ変換手法を提案している [10, 45]。多角形をベジエ曲線に置き換えることで滑らかな近似曲線が得られる。得られたベクタ表現の操作性は、第 2.2.4 節で述べたように、曲線と通過点の位置関係によって決まる。

従来手法 [7] により得られる最適多角形の辺を、図 5-5 のように 1 本のベジエ曲線で置き換え、頂点付近で滑らかに接続する手法について検討する。まず、頂点がコーナである場合を考える。通過点は頂点に置き、制御点も通過点に重ねる。次に、頂点がカーブである場合を考える。通過点は頂点の 2 等分線上で、頂点からの誤差が $1/2$ 画素以下の所に置く。制御点は、頂点を含む三角形の底辺に平行で、接続点を通る直線上の適当なところに置く。この条件により、二つのベジエ曲線は滑らかに接続する。最後に、ベジエ曲線の定義式と、ベジエ曲線が媒介変数 $t = 1/2$ で辺に接する、または通過する条件式を示す。ただし、 $y_2, y_3 = 0$ とし座標系を単純化している。

$$x(t) = x_1(1-t)^3 + 3x_2(1-t)^2t + 3x_3(1-t)t^2 + x_4t^3 \quad (5.15)$$

$$y(t) = y_1(1-t)^3 + 3y_2(1-t)^2t + 3y_3(1-t)t^2 + y_4t^3 \quad (5.16)$$

$$y(1/2) = 0 \quad (5.17)$$

$$\frac{dy(t)}{dx(t)} = \frac{dy(t)}{dt} \frac{dt}{dx(t)} = 0 \quad (t = \frac{1}{2}) \quad (5.18)$$

それぞれの辺に対して両端の条件の設定で、曲線の制御点の位置は一意に決定される。そして、操作性の高いベジエ曲線で置き換えられる。

5.4.2 ベジエ曲線の編集性評価手法

ベジエ曲線の編集には、曲線の膨らみを変更する場合と、カーブとコーナを相互変換する場合が想定される。いずれの場合も、曲線の頂点付近に通過点が存在していると、編集しや

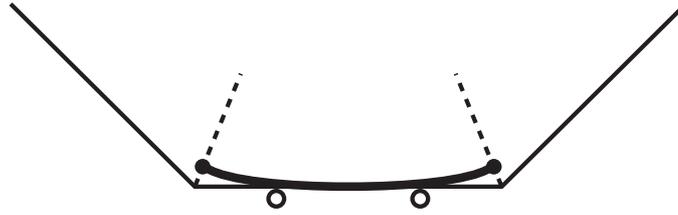


図 5-5: 辺の置き換え手法



図 5-6: Example of Test Shapes; Key, Hammer, and Apple. [59]

すいことが知られており、直感的な曲線操作を実現できる。

ここで、自由曲線のうち凸曲線区間に対して、凸の頂点が曲線の頂点であるといえる。しかし、凸曲線区間の取り方には自由度があり、さらに空間スケールによっても凸曲線区間が異なるため、曲線頂点は一意に決定できない。そこで、空間スケールを固定したときに曲率が極値（極大、または極小）になる点を曲線の頂点であると定義する。ただし、曲率が極値になっていても絶対値が小さいときは、曲線が穏やかに曲がっていることを表している。この場合、空間スケールが少し大きくなるだけで頂点とは定義されないと考えられる。そこで、曲率がしきい値 th より大きい場合だけ「頂点」とする。

この曲線の頂点付近にベジエ曲線の通過点が存在する場合に、曲線の頂点と通過点が一致しているとする。そして、全ての曲線の頂点に占める、対応する通過点が存在する頂点の割合を、編集性を考慮したベクタ変換の評価値であると定義する。なお、ベジエ曲線の通過点は、曲線の頂点以外にも存在しうるが、これは評価には影響を与えない。従って、変換性能自体については、あらかじめ評価しておくこととする。

5.5 提案手法の評価

本章では、画像に対するガウス差分 (DoG) と平面曲線に対する曲率差分 (DoC) の類似性を用いた。提案手法の有効性を二つの実験により確認する。第一実験は、ある形状の様々な大きさに対するスケール検出の安定性である。第二実験は、輪郭素片に対する頑健性である。

形状群は MPEG-7 コア実験で用いられた CE-Shape-1 [59] である。本章では 70 個の形状を用いた。いくつかの形状を図 5-6 に示す。

二つの従来手法と比較する。一つは MPEG-7 輪郭形状記述子で、曲率スケール空間に基

表 5-1: 正規化後のスケールの標準偏差

DoC	CSS	GbC
0.0396	0.0573	0.0651

づいた手法である．もう一つは輪郭画像に対する濃度勾配を用いた曲率である [65, 68]．以下，これらの従来手法を CSS と GbC，提案手法を DoC と呼ぶ．

5.5.1 スケール不変性の安定性

最初の実験は形状の様々な大きさに対するスケール検出の安定性を示す．本実験において，与えられた様々な大きさの形状同士から同一の場所に位置する検出された特徴点のスケール値は，与えられたスケールに対して線型であるか一定値であることが期待される．本実験の手順を以下に示す．輪郭線のテストセットを得るために，形状は与えられたスケールで縮小，または拡大される．最適スケールの集合は既知のスケールか，同一位置のスケール値の平均によって正規化される．従って，正規化されたスケール値は 1.00 に近づくはずである．最後に，安定性として正規化スケール値の集合の標準偏差を求める．

縮小と拡大はベクタグラフィックス上で実現される．このベクタグラフィックスは既存のベクタ化ツールによって生成される [7]．このベクタ化ツールは 2 値画像から得られる輪郭線を滑らかなベジエ曲線に変換する．ベジエ曲線の縮尺を変えて，再び 2 値画像へラスタライズする．既知のスケールは 0.1 から 10.0 へ変化させる．提案手法 (DoC) によって検出されたスケール値の特性は以下の通りである．スケール値は与えられたスケールに比例する一方で，同一形状の異なる場所におけるそれぞれのスケール値は異なる．従って，検出されたスケール値集合は既知のスケール値集合によって正規化し，さらに異なるスケール間の同一位置における検出スケール値の平均によって正規化する．CSS によって検出されたスケール値は既知のスケール値に依存するので，異なる既知のスケール値間の平均によって正規化する．GbC によって検出されたスケール値も DoC と同一の性質を持つ．同一箇所検出されたいくつかの特徴点について，正規化した結果を図 5-7 に示す．X 座標と Y 座標は既知のスケールと検出された正規化後のスケールを表す．それぞれの系はスケールの異なる複数形状における同一位置の特徴点集合を表す．系において点が存在していない箇所は検出されなかったことを表す．さらに，70 形状の正規化スケールの標準偏差を表 5-1 に示す．これらの結果から三つの考察が得られる．第一に提案手法は最適スケールを異なる大きさの形状から輪郭線の長さを使うことなく検出している．第二に提案手法は正しいスケールの特徴点だけを検出しており，誤検出がない．すなわち，他の手法ではいくつかの特徴点について誤ったスケールを検出しており，特に既知のスケールが小さい時に 1.00 からかなり離れた値になる．第三に提案手法は 30～40% 程度標準偏差が小さい．このように，提案手法は最も安定度の高い手法である．

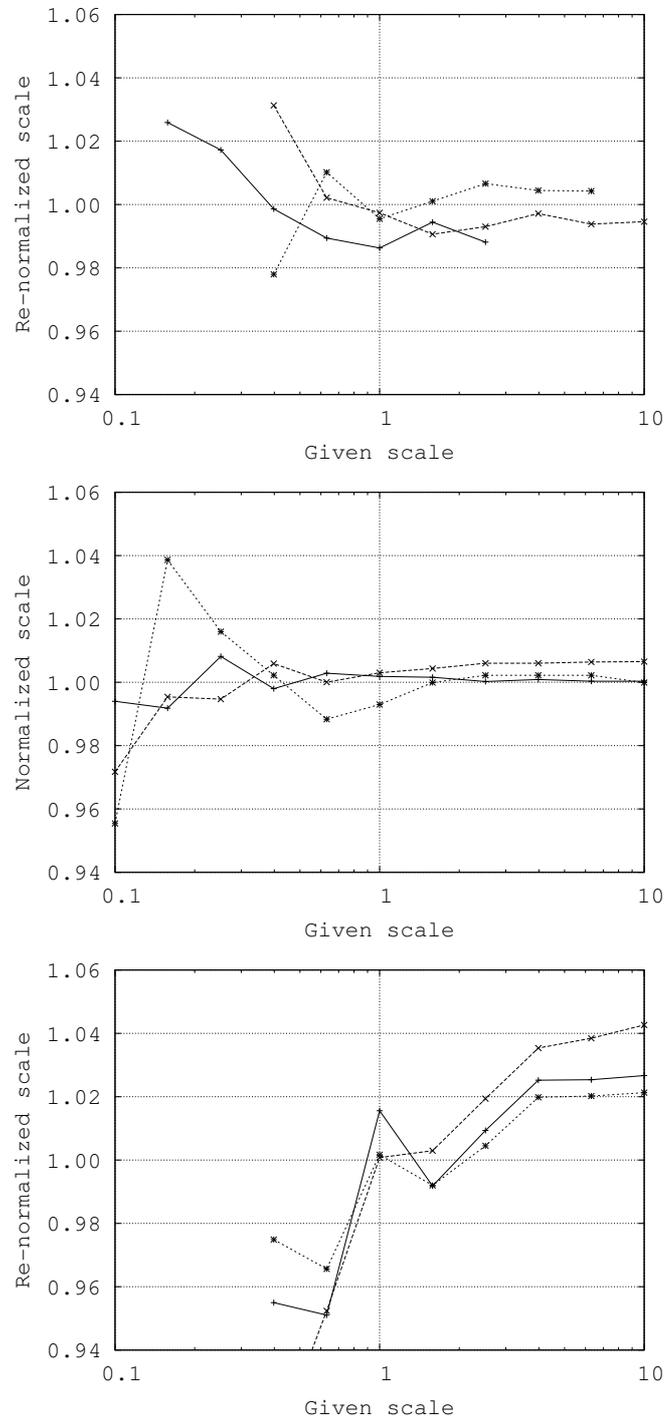


図 5-7: 再正規化後のスケール値 上から提案手法 (DoC), CSS, GbC (自発表 [66] より引用)

5.5.2 スケール不変性の頑健性

二番目の実験は輪郭素片に対する頑健性である。従来手法は輪郭素片に適用するのは困難であるが、比較のために、輪郭線全体の長さが短くなるように形状の一部を削った。以下では、形状全体の輪郭線の長さに対して短くした長さを部分長と呼ぶ。さらに、全ての手法は同じ輪郭素片に適用する。いくつかの特徴について、それぞれの手法によって得られた生データにおける部分長と検出スケールの関係を図 5-8 に示す。

提案手法はいずれの特徴についても一定のスケール値を検出している。一方で、従来手法により検出されたスケール値は変動している。GbC 手法の結果では、検出されたスケールは既知の部分長に比例する傾向がある。検出されたスケール値が同一輪郭素片のあるスケール値で正規化されたとしても、安定性は提案手法よりも低下する。

この結果より、提案手法は輪郭素片に対し、その素片に特徴点が残されてれば頑健であることが示された。

5.5.3 特徴点位置の位置特性

二つの形状について検出された特徴点の例を図 5-9 に示す。この図において、三手法によって検出されたスケール点は点線で表現され、太線で描かれた原輪郭線に重畳されている。それぞれの円の直径は検出されたスケールに比例しており、円の中心は特徴点の場所を表している。

これらの実験は二つの傾向を示している。第一に CSS 手法と比較して提案手法は異なるスケールを有する特徴点を検出する。第二に GbC 手法と比較して提案手法による特徴点の位置は曲率が視覚的に大きい所とは対応していない。これらの検出されたスケールは最終的に形状のスケール不変特徴の記述に用いられる。曲率に基づく特徴点の検出手法は滑らかな部分曲線において不利である。すなわち、曲率の変化が一定であると特徴点は定義できない。このような例を図 5-10 に示す。三手法による特徴点はなめらかな曲線部分には存在しない。この事実は滑らかな曲線には他の技術を用いて特徴づける必要があることを示唆している。

提案した DoC と従来の CSS とのそれぞれから得られた最適スケール位置の比較を行う。これらをオリジナルの図形に重畳している結果を図 5-11 にそれぞれ示す。これらの図形は Part B of CE-Shape-1 [59] の一部である。円の半径は、DoC において検出されたスケールを、CSS において適用したフィルタ回数にそれぞれ比例している。

5.5.4 曲率と曲線の頂点

デジタル曲線の曲率と、提案手法による曲線の頂点の位置を示す。ここで、パラメータ σ は、1 未満にすると微分が成立せず、大きすぎると大局的な頂点しか抽出できない。本論文では予備実験により $\sigma = 1$ 、 $th = 1/\pi$ を採用する。

図 5-12 に入力画像の多角形近似と、輪郭を構成する点番号（代表値）を示す。ここで、小黒点は入力画像の格子点を表している。従って、図中の四角は画素ではない。さらに、図

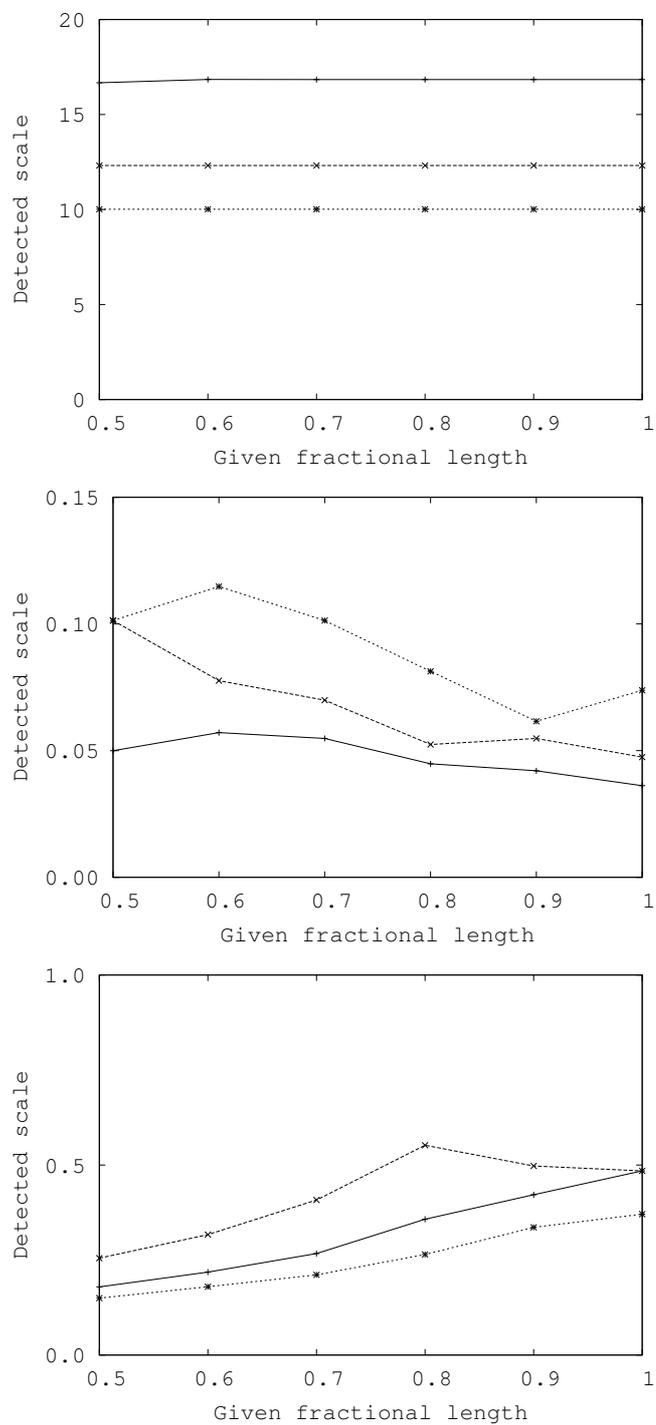


図 5-8: 輪郭素片から検出されたスケール 上から提案手法 (DoC), CSS, GbC (自発表 [66] より引用)

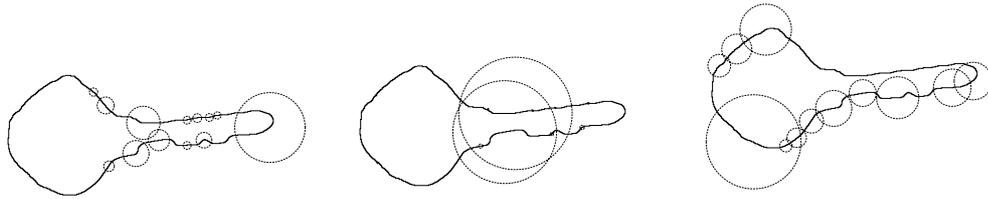


図 5-9: 特徴点が検出できる例．検出手法は左から順に DoC, CSS, GbC (自発表 [66] より引用)

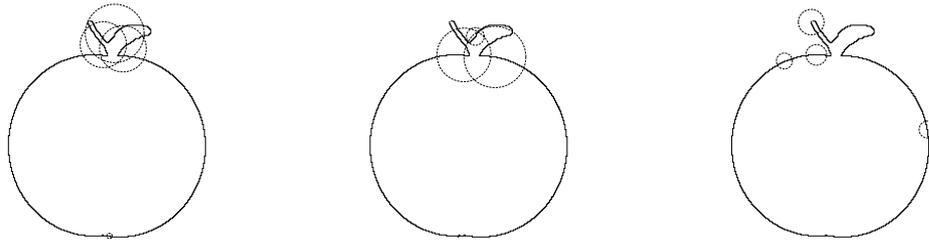


図 5-10: 特徴点が検出されない例．検出手法は左から順に DoC, CSS, GbC (自発表 [66] より引用)

5-13に入力画像の点番号と曲率の関係，及び提案手法により定義した頂点の位置を示す．

これらの図より，提案手法により得られた頂点は，入力画像を高精度に近似するスケールにおける曲線の頂点と一致が主観的に見て取れる．さらに，曲線の頂点付近には概ね多角形の頂点が位置していることも確認できる．これについては，しきい値を変化させて，かつ従来のベクタ変換手法とも比較して，定量的に検討する．

ただし，曲率が大きくても頂点が連続して検出されている部分が散見される．これは，曲線の向きが変化しているために生じると考えられる．これらの頂点については集約するなどの改善が必要である．

5.5.5 ベクタ変換の編集性

提案したデジタル曲線の頂点と，ベクタ変換によって得られるベジエ曲線の通過点について，比較を行う．ベクタ変換の従来手法として Selinger らの手法を用いる．提案した曲線の頂点 (Summits)，提案手法による通過点 (Prop.)，従来手法による通過点 (Conv.) を図 5-14に示す．なお，スケールとしきい値は前の実験と同じである．この図より，提案手法は多くの場合曲線の頂点に近い位置に通過点が存在していることが確認できる．一方，従来手法ではほとんどの通過点は曲線の頂点の間に存在していることも確認できる．

次に，しきい値 th を様々な値に変更したときの，検出された頂点の個数，提案手法と従来手法において頂点に対応する通過点の個数の関係を図 5-15に示す．この図より，しきい値が小さいと頂点が過剰に検出されることが確認でき，しきい値が大きくと大局的な頂点だけが検出されることが推定される．また，利用可能なしきい値には余裕があり， $\frac{3}{4\pi} \sim \frac{5}{4\pi}$ 程度である．

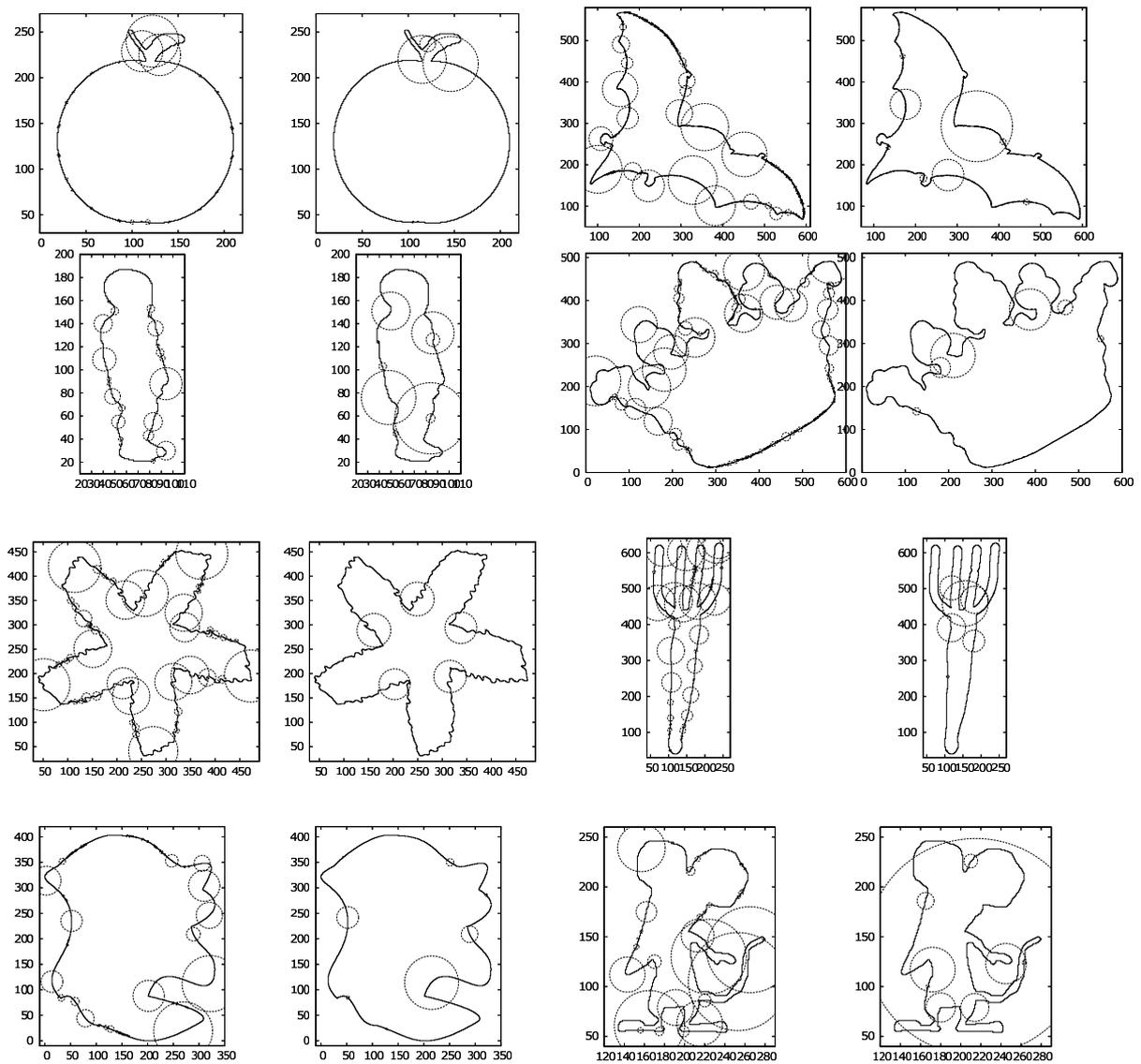


図 5-11: 特徴点の検出結果。左が DoC による提案手法で、右が CSS による従来手法。コンテンツは左上から apple, bat, children, crown, device0, fork, Misk, rat

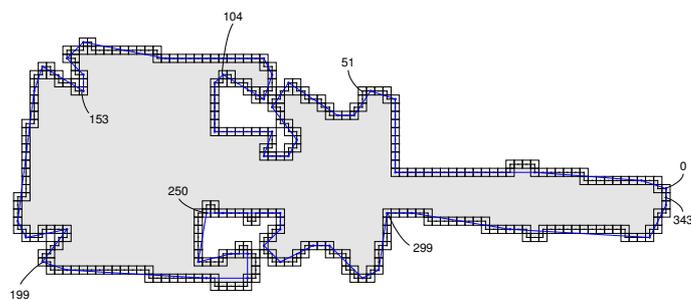


図 5-12: 多角形近似と点番号の関係 (自発表 [69] より引用)

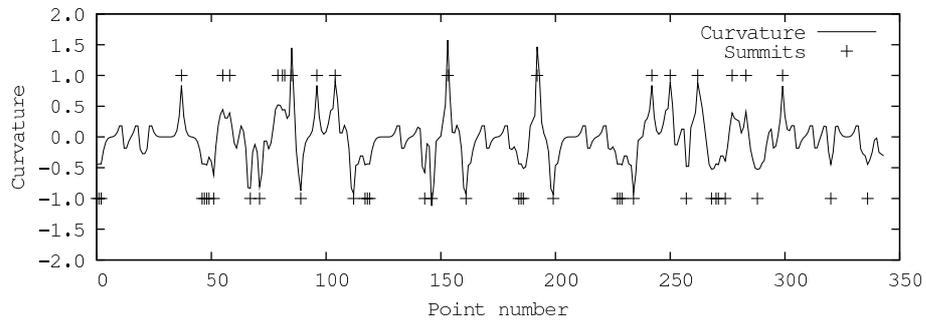


図 5-13: ドライバの曲率 (自発表 [69] より引用)

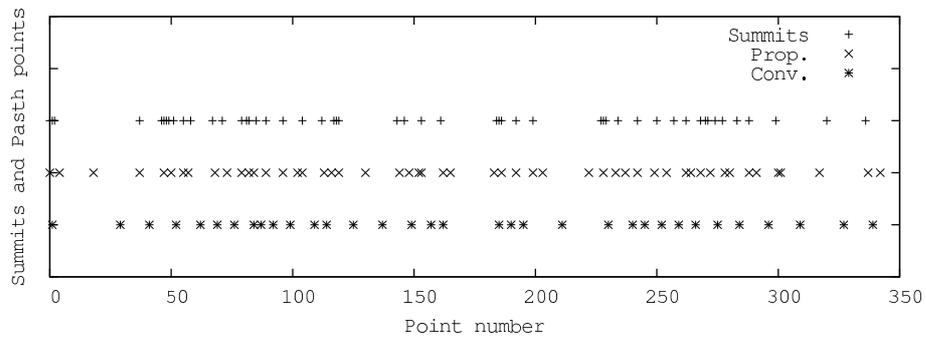


図 5-14: 頂点と通過点の位置関係 (自発表 [69] より引用)

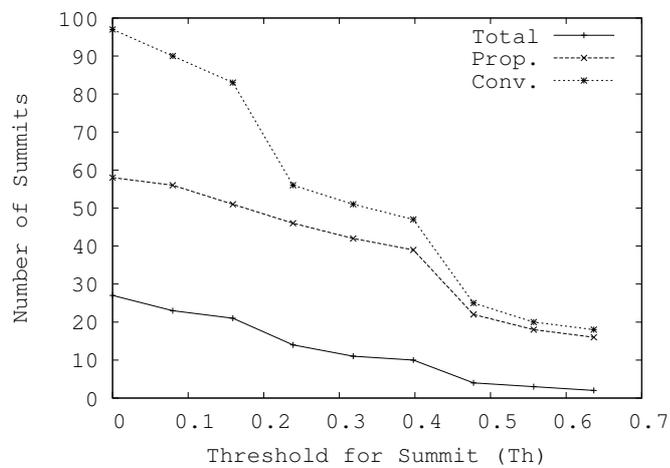


図 5-15: しきい値と頂点数の関係 (自発表 [69] より引用)

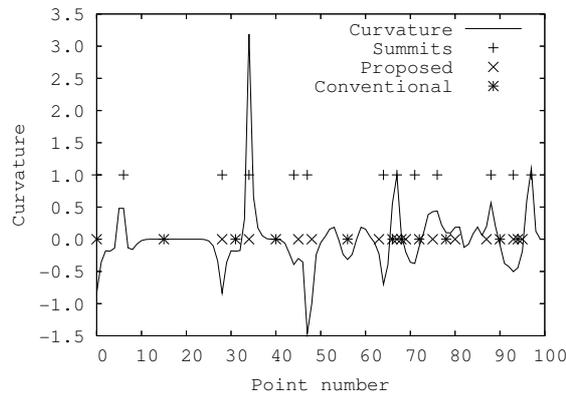


図 5-16: 曲率と通過点の位置関係 (自発表 [69] より引用)

表 5-2: Summits Ratio for each Bezier Curve

Proposed Method	Conventional Method
70.4%	32.3%

適切なしきい値を利用した場合，提案手法において曲線の頂点に通過点が存在する割合は 8 割程度である．なお，従来手法でも 2 割程度は曲線の頂点と通過点が一致している．これは，頂点がカーブではなくコーナと判定された場合と，先に述べたように連続して頂点を検出された場合であると考えられる．提案手法は従来手法に比べて割合が向上していることから，編集が容易な部分が大幅に増えているといえる．

約 260 個の頂点を有する 2 値のイラストを入力としてベクタ変換を適用し，曲率と通過点の位置関係を調査した．なお，ベクタ変換の従来手法として Selinger らの手法を用いる．パラメータ σ は，1 未満にすると微分が成立せず，大きすぎると大局的な頂点しか抽出できない．なお，スケールとしきい値は前の実験と同じである．

図 5-16 にデジタル曲線を構成する点の番号と，その点の曲率 (Curvature)，提案した曲線の頂点 (Summits)，提案手法による通過点 (Proposed)，従来手法による通過点 (Conventional) を示す．この図より，提案手法は多くの場合曲線の頂点に近い位置に通過点が存在していることが確認できる．一方，従来手法ではほとんどの通過点は曲線の頂点の間に存在していることも確認できる．

また，画像全体について，曲線の頂点付近に通過点が存在する頂点の割合を表 5-2 に示す．従来手法でも 30% 程度は曲線の頂点と通過点が一致している．これは，頂点がカーブではなくコーナと判定されたためであると考えられる．提案手法は従来手法に比べて割合が向上していることから，直感的な編集が可能な部分が大幅に増えているといえる．

5.6 むすび

本章では，人工画像のベクタ化対象の中核となる平面曲線（線画）に着目し，スケール不変特徴量とベジエ曲線の編集性の評価指標を提案した．また，これまで提案しているベクタ変換が編集性の高いベジエ曲線の生成を明らかにした．

スケール不変特徴量は SIFT アルゴリズムのうち Difference of Gaussian (DoG) を Difference of Curvature (DoC) に置換するフレームワークを提案した．これを実現するために，スケール正規化曲線を提案し，最適スケール曲率を定義した．

MPEG-7 のコア実験で規定された形状データベースを用いて，スケール不変特徴量の安定性と頑健性を評価した．また，閉曲線だけではなく，曲線の一部からでもスケール不変特徴量を抽出できることを示した．

得られた特徴点から曲線の頂点を定義し，ベジエ曲線の編集性を定性的に述べた．また，得られた特徴量は，適当なスケール範囲において曲率が大きな部分，すなわち曲線の頂点の定義に用いた．

最後に，これまで提案している多角形の辺ごとに曲線に置き換えるベクタ変換手法は，生成されるベジエ曲線の編集性が高いことを明らかにした．

第6章

コミック画像からのメタデータ抽出方式

6.1 まえがき

本章では、高能率符号化により得られたビットストリームの活用を促進するため、人工画像の代表例であるコミック画像からメタデータを抽出する手法を提案する。具体的には文字とコマ配置を対象とする。

コミック画像からの文字の切出しは、線画やテクスチャが混在しており、さらに文書と異なり文字数も少ないという困難性がある。そこで、対象は活字であるが、手書き文字の認識手法を組合せ、文字と文字列の特徴だけを用いて切出す手法を提案する。

さらに、コマ配置の取得には、コミック画像を再帰的に2分割してコマに分割する手法を説明する。分割線の探索時に線ではなく帯を用いて、さらにヒューリスティックな検査により高精度化する手法を提案する。さらに処理時間を削減するために、対象画像の低解像度化による計算量削減と、探索領域をヒューリスティックに限定する間引きとを組合せ、高速化する手法を提案する。

最後に、提案手法の有効性を示すために、文字切出しの再現率、コマ分割精度、コマ分割処理時間について評価実験を行う。

6.2 文字領域の抽出

6.2.1 コミックにおけるセリフの特徴

コミックにおける文字の役割は、キャラクタのセリフや物語に必要なト書きである。また、効果音や擬音を文字により表現している。本論文では、画像インデキシングや多言語化に必要と考えられるセリフやト書きなど、活字となっている文字だけを対象とする。

セリフやト書きは数文字以上からなる行を構成し、場合によっては複数行から構成されている。ただし、感嘆詞のような場合は、1, 2文字で構成されている。この場合、文字を見やすくするために周辺には空白領域が取られることが多い。

フォントサイズは、同じ画像中でも頻繁に変化する。ただし、同じ行の中でサイズが変わることはない。上記の他に、日本語文字の一般的な特徴である、文字はほぼ正方形であるという特徴 [70] を用いる。

6.2.2 文字切出し手法

線画や文字が混在する画像では、文字のピッチや大きさなどの事前情報から文字領域を抽出するのが困難である。そこで、以前より研究されている手書き文字列から文字を切出す方式に着目する。これらの手法では、外接矩形の面積や縦横比などの形状的特徴を文字の判定に用いる。さらに、文字認識結果や単語照合結果を文字切出しにフィードバックする方式も検討されている。しかし、形状的な特徴に基づく切出しがより正確になれば、フィードバック情報の確度が高まり、性能がより一層向上する。このため、文字の形状的特徴だけにに基づく文字切出しの高精度化は、重要な課題である。

本論文では、あらかじめ文字サイズを指定することなく、文字切出しを実現する手法を提案する。画像内に複数の文字サイズが混在している場合でも、文字列内では同一サイズになるという傾向を利用する。なお文字サイズとは、文字のポイント数ではなく、文字の外接矩形における長辺の画素数と定義する。

様々な文字サイズに対応するため、小さい文字から大きい文字まで順番に文字サイズ S を仮定して処理を進める。以下に S における抽出手順を述べる。

外接矩形の取得 連結黒画素から外接矩形を得る。矩形の長辺が S より長い場合は、その矩形を破棄する。

文字列中心の取得 注目する矩形の周辺を探索して、分布する矩形から文字列密度を計算する。得られた密度に対し、しきい値処理によって行が存在するかを判定し、存在する場合には文字列中心の座標を求める。なお、仮定した文字列に垂直方向へ $S \times 2/3$ 、水平方向に $S \times 4/3$ を探索範囲とする。予備実験から、しきい値 $th = 0.2 \times S^2$ とした。

分離した文字の統合 得られた文字列中心に対して、文字列と垂直方向に分離している矩形を統合し、より大きな矩形を得る。

文字列の取得 得られた文字列中心により、文字列と水平方向に分離している矩形を統合して、文字列を取得する。分離した文字を統合する条件は、矩形の中心同士の距離が文字列に垂直方向に $S \times 2/3$ 、水平方向に $S \times 4/3$ 以下で、かつ重なりが 0.5 以上とした [71]。

文字列群の統合 得られた文字列を、文字列方向と垂直に拡張して文字列群を取得する。統合した矩形の周囲に余白が存在する場合、孤立文字として抽出する。

本手法は、画像の局所的な特徴を利用しているため、画像のゆがみや文字の位置に対してロバスト性が高い。

6.2.3 文字サイズの自動決定手法

画像内においては様々な文字サイズが存在している。そこで、文字を結合する前後にそれぞれ文字サイズを付与する手法を提案する。

先に結合後における文字サイズ付与を述べる。探索範囲に含まれる矩形のうち、射影後の文字列中心が十分近く、射影後の文字中心が十分離れ、かつ矩形長辺の長さが近い場合に、同一文字列を構成していると判断する。探索範囲の大きさはおよそ3文字分であるから、行頭である場合を考慮して、文字数が2、または3文字の時に正しく文字列が検出されていると判断する。これらの矩形に対して矩形の長辺を文字サイズとして与える。

結合前においては、探索範囲を文字列の垂直方向へ $S \times 0.2$ 、水平方向に $S \times 3.2$ として、注目する矩形の周辺を探索する。これは文字列だけを検出するためである。範囲内に存在する矩形について、矩形の長辺の中央値を文字サイズとして与える。これにより、文字サイズが決定されると、そのサイズが文字列に伝搬される。

文字領域を階層的クラスタリングの一つである単連結法によりまとめる。単連結法では、文字領域を統合することに距離の再計算を行う必要がない、結合後の距離は0とする。なお、クラスタリング終了はしきい値以下の距離がなくなった時とする。

6.3 コマ分割手法

電子コミックの閲覧にあたり、携帯電話をはじめとする小型の端末では、画面サイズが非常に小さく、紙面全体を表示した際に、テキストの可読性が損なわれる問題がある。実際のコミック配信では、コマと呼ばれる枠の並びによって定義されている紙面内の読み順に沿って画面のスクロールを行うことで、文字の可読性を保ちつつ小型画面端末におけるコミック画像の閲覧を実現している。

本論文ではスクロール表示を行うための読み順を、ページ上における視線遷移情報と呼ぶ。現在、この視線遷移情報は人手で抽出され、配信データに重畳されている。コミック雑誌をはじめとする多量のページをもつコンテンツの配信を実現するため、視線遷移情報抽出における人的負担を低減する技術が求められている。

6.3.1 コミック画像におけるコマの特徴

コミック画像のレイアウトにおいて最も特徴的な点として、コマからの要素のはみ出しがある。要素のはみ出しにより、本来描かれるべきコマの枠線が隠れてしまい、枠線の一部が描かれていない場合がある。さらに、集中線などに代表される、コマの枠線と大きさ及び太さが似ている線分が画像内に多数存在する。これらがコミック画像のレイアウト解析を困難にさせる要因となっている。

多くのコミック画像には共通して以下のような特徴がみられた。

- (A) コマはある程度の大きさをもつ
- (B) コマは矩形に限定されない線分もしくは画像の縁からなる4辺に囲まれた領域である
- (C) コマとコマの間には空白部が存在する
- (D) コマ内部には絵やトーン等による様々な模様が存在する

- (E) コマの周囲に大きな余白が存在する場合がある
- (F) 電子化されたコミックではテキスト部分の可読性を保つため高解像度でスキャンされている
- (G) コマの枠線方向が水平もしくは垂直である割合が高い

6.3.2 高精度コマ分割手法

田中らのコマ分割処理 [35] では、検出線上の濃度勾配値だけを基準として分割線検出を行うため、分割線の検出精度が低いという問題があった。そこで検出線に幅をもたせることで、分割線の候補をよりロバストに検出する、高精度コマ分割手法を提案する。さらに、得られた分割線候補に対して、平均濃度勾配値の大きな順に、濃度勾配方向検査及びコマ内外検査からなる分割線適合検査を行うことで、最終的に決定される分割線の精度を向上させる。以下、分割線候補の検出を Detection of Separation-Line Candidates (DSL_C)、また分割線適合検査を Separation-Line Adjustment Test (SLAT) と記す。

高精度コマ分割手法の処理手順

処理フローを図 6-1 に示す。Separation Procedure (SP) は画像を引数とする分割処理のためのプロシージャである。SP を再帰的に実行してゆくことで各コマへの分割を行い、各コマの位置情報を抽出する。

Compute the Weighted Gradient Value (CWGV) は画像の濃度勾配値を求めるプロシージャである。濃度勾配値は従来手法 [35] と同様に求める。まず、画像に対して Sobel フィルタにより各画素における濃度勾配値を求め、さらにガウス関数を用いて画像中心に値の重みづけを行うことで取得する。

CWGV により得られた画像に対して、DSL_C をすべての方向に対して行う。そして得られた分割線候補に対して SLAT を行い、適合する分割線が発見された場合、その分割線を境に画像を 2 分割 (Separation) する。適合する分割線を Separation Line (SL) と表す。

分割後の画像は前後関係を与え、順序が先になる領域を Separated Image 1、後になる領域を Separated Image 2 として、SP を再帰的に実行する。適合する分割線が発見されなかった場合には SP は与えられた画像の位置情報を出力する。

以上の手順によって最初に読まれるコマから順番にその位置情報を取得することができる。

分割線候補の検出 (DSL_C)

DSL_C では以下で定義する検出帯を用いてコマの枠線位置を調査し、分割線の候補を決定する。検出線を構成する画素を検出線画素と呼ぶ、各検出線画素について、それぞれ検出線の角度が画像の横軸に対して ± 45 度以内であれば画像の縦軸方向に、それ以外の場合には

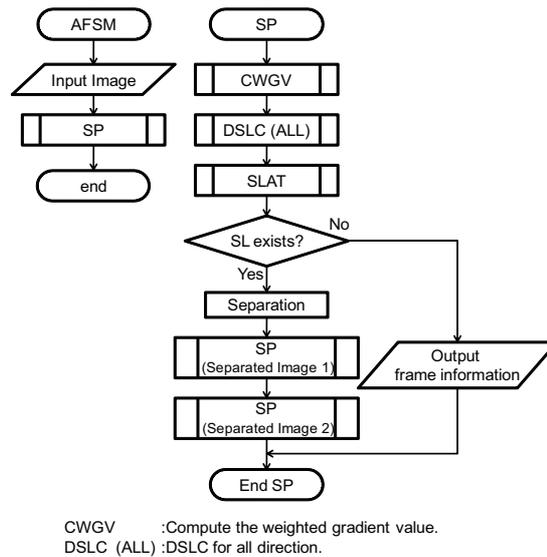


図 6-1: 高精度コマ分割手法の処理フロー [28] (著者の了解を得て抜粋)

画像の横軸方向に隣接している画素を含め、これを検査グループとする。一つの検出線に対応する検査グループを合わせたものを検出帯と定義する。検出帯の幅を ω と定義する。ここでは $\omega = L/250$ とした。ただし、 L は入力画像の長辺の画素数 [pixel]、250 は予備実験から決定した。

検出帯の始点座標及び角度を変化させ、画像の全探索を行う。ここで始点座標は 1 画素ごと、角度は 1 度ごとに変化させる。検出帯の各検査グループにおいて、濃度勾配の絶対値が最も大きい画素を選び、この画素をその検査グループの代表画素とする。各代表画素の濃度勾配の検出帯に対する垂直成分の値を計算し、その平均値を検出帯の平均濃度勾配値とする。平均濃度勾配値の大きい検出帯に対応する検出線を分割線候補とする。

分割線適合検査 (SLAT)

濃度勾配方向検査 濃度勾配方向検査では、分割線候補をいくつかの領域にわけ、それぞれの領域内の濃度勾配方向分布の一致割合を調べる。第 6.3.1 節 (D) に示した特徴より、コマの内部の絵やトーン等による様々な模様が存在する場所では、濃度勾配方向の分布が一定しない。これに対して、コマ枠線に沿った点では、濃度勾配方向はその分布が一定である。

具体的な処理を述べる。まず、対象検査帯を n 個の領域に等分し、それぞれの領域内で帯に対する各代表点の濃度勾配の角度を調べる。その角度が $(90 \pm \sigma)$ 度の範囲に収まる代表点の割合を調べる。この割合が $1 - \rho$ 以下の場合にはその領域を不適合領域とする。不適合領域数が m 以下である場合には、濃度勾配方向検査を合格とする。ただし $n > m + 2$ とする。

第 6.3.1 節 (E) に示した特徴として、コマの周囲に大きな余白が存在する場合がある。余白部分では濃度勾配値は 0 となり、代表点の濃度勾配方向が検出できない。これに対応するため、検査帯の最初と最後の領域では、代表点の除外処理を行う。最初の領域では検査帯の

始点，最後の領域では検査帯の終点に位置する代表点から順に濃度勾配値を検査してゆく．そして，濃度勾配値が0である代表点が続く限り，それらの代表点を評価から除外する．もしも，全ての代表点が除外された場合にはその領域自体の評価を行わない．そして，さらに余白が続くものとみなし，隣接する領域についても代表点の除外処理を行う．

コマ内外検査 コマ内外検査では，分割線上おけるコマ枠線の存在を調べる．第 6.3.1 節 (C) に示した特徴として，コミックの各コマ間には空白部分が存在する．このためコマの外枠に沿って分割線が検出されている場合には，その分割線の片側は必ず空白となる．これに対して，コマ内部に存在する直線などに沿って分割線候補を検出した場合には，この分割線候補はコマの枠線部分を通過することとなる．

具体的な処理を以下に述べる．まず，検査線の両隣に検査線と平行に仮想的な直線を定義する．そして仮想直線上の画素を検査線の始点側から終点側に向かって順に走査してゆき，それぞれの仮想直線上における最初と最後の黒画素位置を調べる．二つの仮想直線上における最初と最後の黒画素の出現位置がそれぞれ一致する場合にはコマ枠線上を通過しているとみなす．分割線がコマ枠線上を通過していない場合だけコマ内外についての条件を満たすとする．

6.3.3 高速コマ分割手法

高精度コマ分割手法は，精度が高い反面計算量が高い．そこで計算量の削減手法を検討する．

高精度コマ分割手法は，2分割処理を繰り返し適用するため，分割を進めるに従い処理対象画像が小さくなる．そのため検査帯の角度を1度ずつ変化させても，実際の検査対象画素の変化割合は小さく，検出処理が冗長となる．本特徴と，第 6.3.1 節 (A)，(F) 及び (G) で述べたコミック画像の特徴を組合せる．

高速コマ分割手法の処理手順

処理フローを図 6-2 に示す．Fast separation procedure (FSP) は画像を引数とする分割処理のためのプロシージャである．入力された画像に解像度変換を適用して低解像度にする．そして，この画像を FSP に渡す．

FSP では，分割線候補検出順序の最適化及び冗長な分割線検索の削減により，分割線の検出及び画像の分割処理を実行する．

画像解像度変換

第 6.3.1 節 (F) に示した特徴として，コミック画像は台詞部分における文字の可読性を保つため解像度が高いことが多い．一方，コマの枠線は文字と比較して低い解像度でも認識することができる．そのため，画像の解像度を下げることによって，実質的な検査量を削減できる．

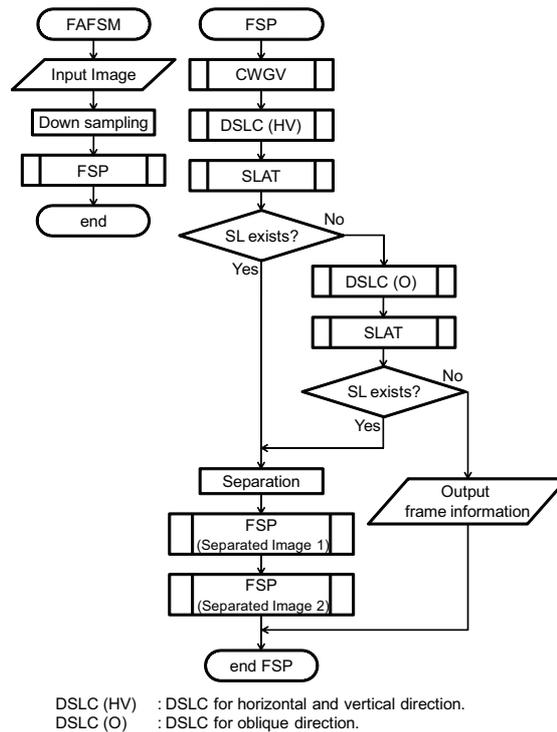


図 6-2: 高速高精度コマ分割手法の処理フロー [28] (著者の了解を得て抜粋)

具体的には入力画像に対して、 $k \times k$ サイズの平均値フィルタを適用し $1/k \times 1/k$ にダウンサンプリングして低解像度化する。

分割線候補検出順序の最適化

第 6.3.1 節 (G) に示した特徴として、コマの枠線方向は水平もしくは垂直方向である割合が高い。そこで、DSLС の順序を変更して早期打ち切りを導入し、処理量を削減する。分割線候補検出順序の最適化では、DSLС を画像の垂直方向及び水平方向だけに対して行う。ここで得られた分割線候補に対して SLAT を行い、候補中に適合する分割線候補があればその時点で画像の 2 分割処理を行う。

得られた分割線候補すべてが適合しない場合に、水平、垂直以外の分割線候補の検索を行う。新たに得られた分割線候補に対して、SLAT を行い、候補中に適合する分割線があれば画像の 2 分割処理を行う。水平、垂直以外の角度についても、適合する分割線が得られない場合にはこれ以上分割できないと判断し、与えられた画像の位置情報を入力する。

冗長な分割線検索の削減

第 6.3.1 節 (A) に示した特徴として、コマはある程度の大きさを持つため、画像の四隅近傍に複数のコマが存在することはない。つまり、画像の四隅近傍には分割線が存在しない。そこで画像の四隅近傍を始点とする分割線候補の検出処理の削減を行う。さらに、高精度コ

表 6-1: Recall Rate of Character Segmentation from Comic Images.

Method	Recall rate [%]
Multi size with detection	86.2
Multi size without detection	19.1
Single size	87.0

マ分割手法の特徴から，画像サイズに応じた検査帯の変化角度変更を行うことで，処理量の削減を図る．

まず，分割線候補の検出処理において画像の4隅から l [pixel] の範囲を始点座標から削減する． l の大きさは $l = L/E$ とする． E の値の変更で検出処理の削減度合いを制御可能となる．この値 E を小さくした場合には分割速度は向上する g ，ある大きさを超えると分割精度が低下する．予備実験では，多くの作品に対して， $E=10$ から 15 を超えると急速に分割精度の低下が確認された．そこで，本論文では $E = 10$ とした．

次に，画像サイズに応じて検出線角度の変化量を決定する．検出線角度の変化量を ξ 度とすると， ξ は以下の式で求める．

$$\xi = A\left(\frac{1}{\omega + 1}\right) + B \quad (6.1)$$

A 及び B は実験により最適な値である $A = 20, B = -2$ とした．

6.4 実験と考察

本節では，提案手法のうち文字列の中心を取得した結果，高精度コマ分割手法と高速コマ分割手法の分割結果を示し，考察を述べる．

6.4.1 文字切出しの再現率

対象画像の代表例であるコミック (B6 サイズ) を 300dpi でスキャンし入力画像とする．対象画像は 16 枚で，含まれる文字数は 1299 文字である．文字サイズを 1 種類だけに限定する場合，文字サイズの自動決定をする場合としない場合の 3 種類について比較を行う．実験結果を表 6-1 に示す．評価は一文字ごとに行い，ページ再現率 (1 ページに含まれるすべての文字のうち，実際に検出された文字数) の平均値を示した．

まず，文字サイズの自動決定手法を導入しない場合，文字サイズ S を変えると，再現率が著しく低下する．予備実験より， S が実際の文字サイズの 2 倍以上になった場合に顕著であった．複数の文字をまとめて一文字と判定してしまうことが原因である．

次に，文字サイズの自動決定手法を導入した場合，文字サイズ S を変えても再現率の低下は 1% 以下となる．依然として複数の文字が結合されてしまう場合がある一方で，タイトルページ含まれるような大きな文字サイズにおいても検出が可能となった．また，入力解像度の異なる場合，すなわち異なる文字サイズの画像においても，同様の再現率が得られた．

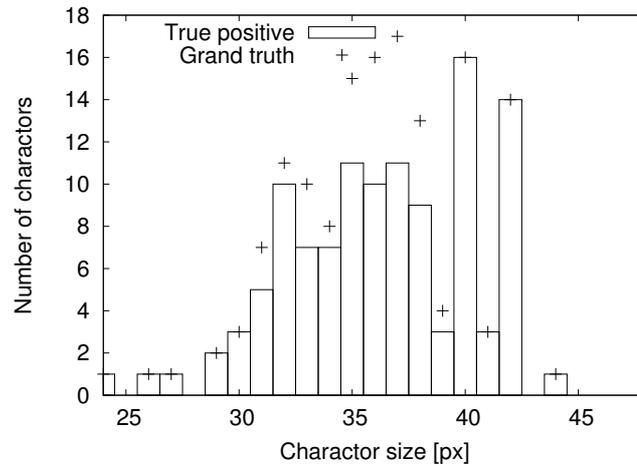


図 6-3: 文字サイズと再現数 (自発表 [72] より引用)

ところで、一般的な文字切出し手法と比べて、提案手法の再現率は高くない [36]。一般的な文書画像と比べて、コミックのセリフには三点リーダー、感嘆符、疑問符などの記号が多く含まれていることに起因する。これらの記号を考慮した文字切出しは今後の課題である。

6.4.2 文字サイズの分布

16枚のうち1枚について、文字サイズと含まれる文字数 (Grand Truth)、及び正しく検出された文字数 (True Positive) を図 6-3 に示す。なお、文字サイズと文字種の関係としては、32画素付近がセリフにおけるひらがな、36画素付近がセリフにおける漢字、40画素付近がト書きに相当している。この図より、提案している文字切出し手法は様々な文字サイズに対応していることが確認できる。

6.4.3 コマ分割精度

高精度コマ分割手法及び高速コマ分割手法の性能を調べるため、田中らの手法 [35] との比較実験を行った。分割精度の評価には以下の式を用いた。

$$P = \frac{\text{認識された正しい分割線数}}{\text{認識された全ての分割線数}} \quad (6-2)$$

$$R = \frac{\text{認識された正しい分割線数}}{\text{認識されるべき分割線数}} \quad (6-3)$$

$$F = 2 \times (P \times R) / (P + R) \quad (6-4)$$

P , R 及び F はそれぞれ検出された分割線の適合率、再現率及び F 値を示す。また、各々のページにおいて内包する全てのコマが正しい順番で得られた場合に評価値 1、そうでない場合は評価値 0 として、入力シーケンスにおける評価値の平均を平均成功率 S と定義する。さ

らに、一つの入力シーケンスにおける、1 ページを分割する為に要した時間の平均値を処理時間とする。

対象画像はコミック画像 4 シーケンス合計 110 ページとした。入力画像のサイズとページ数を表 6-2 に示す。実験に用いた計算機環境は Intel Pentium D 3.2GHz, メモリ 2GB である。ここで、コマ分割のパラメータは予備実験より $\rho = 15, \delta = 40, n = 7, m = 2$ とした。高速化の際の画像解像度変換パラメータである k は分割処理時間と精度に大きな影響を与えるため、値を 1 から 3 まで変化させた結果についてそれぞれ実験を行った。

実験結果を以下の表に示す。入力シーケンスごとの分割精度を表 6-3 から表 6-6 に、処理時間を表 6-7 に、処理時間の比較を表 6-8 にそれぞれ示す。

従来手法と比較して高精度コマ分割手法の分割精度は、すべてのシーケンスにおいて高い F 値及び S 値が得られた。高い F 値をもつことはページ内部のコマの多くが正しく分割されていることを示す。この要因は、分割線決定時に SLAT を行ったことによる分割線適合率の上昇である。

ただし、ComicC では、すべての手法において再現率が低い結果となった。ComicC では、コミックにコマからの要素のはみ出しが多数存在し、かつそれぞれのコマに対する大きさが大きいという特徴がある。SLAT の濃度勾配方向検査において、複数個所で濃度勾配方向の一致がとれないため、分割線候補から除外されたことが原因である。

6.4.4 コマ分割の処理時間

高精度コマ分割手法では、従来手法と比較しても 3 倍程度の処理時間が必要となる。実際の処理時間で見ても、1 枚当たり 2~4 分程度の処理時間を必要としており、分割精度が高くても処理時間の面で実用上現実的でない。

高精度コマ分割手法と比較して高速コマ分割手法の処理時間は $k = 1$ では 5~30% 程度、 $k = 2$ では 1~2%、さらに $k = 3$ では 1% 未満である。これは従来手法と比較しても 3~7% 程度の処理時間である。特に、 $k = 3$ の場合にはすべてのシーケンスで平均処理時間が 1 秒以下であった。

$k = 1$ 及び $k = 2$ の場合では高精度コマ分割手法による結果と同等の分割精度が維持されている。 $k = 3$ の場合に分割精度は低下するが、F 値の低下量はたかだか 7.1% であることが確認できた。この理由として、低解像度化による分割精度への影響が挙げられる。より高い低解像度化率で分割を行うと、分割精度はさらに低下する。

ComicC について他の結果よりも高速化率が低い結果となった。この理由は、ComicC に水平もしくは垂直以外の分割線により分割されるコマが多く存在しており、分割線検査順序の最適化による高速化の効果が少ないためである。

以上から、高速コマ分割手法では、高精度コマ分割手法とほぼ同程度の分割精度を保ちつつ、処理時間を高精度コマ分割手法の 1~2% に抑えることができる。また、若干の処理精度の低下が許容できる場合には、多くのシーケンスで 1 秒以下の非常に高速な分割処理を行うことができる。

表 6-2: コミック画像のサイズと枚数

	Image Size[pixel×pixel]	Numvers of Pages
Comic A	740×1200	20
Comic B	760×1200	30
Comic C	1024×1536	30
Comic D	840×1200	30

表 6-3: Comic A の分割精度

		<i>P</i> (%)	<i>R</i> (%)	<i>F</i> (%)	<i>S</i> (%)
高精度コマ分割手法		98.7	100.0	99.3	95.0
高速コマ分割手法	(k=1)	98.7	98.7	98.7	90.0
	(k=2)	98.7	98.7	98.7	90.0
	(k=3)	100.0	96.4	98.0	85.0
田中らの手法 [35]		69.4	76.6	72.8	60.0

6.5 むすび

本章では，人工画像の代表例であるコミック画像から，文字列とコマ配置の二つのメタデータを抽出する手法を提案した．

コミック画像からの文字の切出し手法は，文字は整然と並んでいる一方で，文字列は粗に分布しているという知見を利用した．また，グラフ理論により文字から文字列を構築するアプローチをとった．さらに，コマ配置の取得手法は，分割帯の探索とヒューリスティックな検査による高精度化手法と，低解像度化などによる高速化手法を提案した．

最後に，提案手法の有効性を示すために，文字切出しの再現率を評価した．また，コマ配置の取得手法は従来手法と比較して大幅に分割精度が向上すること，分割精度の低下なしに1～2%程度の処理時間で分割処理が行えることを確認した．

表 6-4: Comic B の分割精度

		<i>P</i> (%)	<i>R</i> (%)	<i>F</i> (%)	<i>S</i> (%)
高精度コマ分割手法		95.0	92.7	93.9	70.0
高速コマ分割手法	(k=1)	95.5	87.1	91.1	63.3
	(k=2)	95.4	83.9	89.2	60.0
	(k=3)	95.2	79.8	86.8	53.3
田中らの手法 [35]		74.6	81.3	77.8	46.7

表 6-5: Comic C の分割精度

		<i>P</i> (%)	<i>R</i> (%)	<i>F</i> (%)	<i>S</i> (%)
高精度コマ分割手法		98.6	79.8	88.2	73.3
高速コマ分割手法	(k=1)	92.3	80.9	86.2	73.3
	(k=2)	96.2	85.4	90.5	83.3
	(k=3)	94.6	78.7	85.9	76.7
田中らの手法 [35]		44.0	37.1	40.2	23.3

表 6-6: Comic D の分割精度

		<i>P</i> (%)	<i>R</i> (%)	<i>F</i> (%)	<i>S</i> (%)
高精度コマ分割手法		96.1	95.7	95.9	93.3
高速コマ分割手法	(k=1)	93.0	93.5	93.2	86.7
	(k=2)	95.7	95.7	95.7	93.3
	(k=3)	95.5	91.4	93.4	80.0
田中らの手法 [35]		88.4	95.1	91.7	50.0

表 6-7: コマ分割手法の処理時間

		ComicA (秒)	ComicB (秒)	ComicC (秒)	ComicD (秒)
高精度コマ分割手法		143.43	143.18	224.62	181.63
高速コマ分割手法	(k=1)	13.04	15.05	74.24	10.49
	(k=2)	1.76	2.00	5.59	2.02
	(k=3)	0.84	0.85	1.84	0.92
田中らの手法 [35]		58.95	56.88	82.31	75.97

表 6-8: 高精度コマ分割手法と高速コマ分割手法の処理時間比較

		ComicA (%)	ComicB (%)	ComicC (%)	ComicD (%)
高精度コマ分割手法		100.00	100.00	100.00	100.00
高速コマ分割手法	(k=1)	9.09	10.51	33.05	5.78
	(k=2)	1.23	1.40	2.49	1.11
	(k=3)	0.59	0.59	0.82	0.51

第7章

コミック画像符号化

7.1 まえがき

本章では、ベクタ化による人工画像の高能率符号化を実現する符号化フレームワークを提案する。符号化はモデル化とエントロピー符号化に分けて考えられる。モデル化として、コミックの特徴について再度まとめ、これを考慮して本論文で提案してきた個々の要素技術を統合する。エントロピー符号化として、特に発生符号量の大きな座標情報を対象として最適なロスレス符号化を模索する。また、座標情報の精度に基づく符号量制御を実現する手法を提案する。

符号化効率を追求せず、現在の環境において提案方式の一部を具現化できる事を示す検証実験を行う。具体的には得られたベクタ表現を、既存のファイルフォーマットに変換してPDAで表示する。

また、符号量制御の有効性を評価するために、レート歪み特性を取得する。同様に、入力画像を簡易化して既存の可逆画像符号化を組み合わせた方式を対象に、提案手法の符号化性能を評価する。

7.2 コミックの画像特性

7.2.1 構成要素

本論文で対象としているコミックは、信号特性の異なる様々な要素から構成されている。符号量が十分小さく機能性の高いという観点で最適なビットストリームを得るために、それぞれを分離して信号特性と利用方法を考慮した符号化を行う。

工学的な観点からコミックの構成要素を挙げると、1) セリフなどを表すための文字列、2) キャラクタや背景を構成する線画、3) 中間調を表現する網点、4) 上記以外のテクスチャとなる。

文字列

文字列は線画として扱うこともできる。従って簡易的なシステムでは特に分けて考える必要はない。しかし、文字の歪みは目立ちやすく、画像インデキシングやセリフの多言語化を

実現するためにも、文字列の分離の需要は大きい。

人は活字を見慣れているため、文字においてはわずかな歪みでも目立ちやすいという傾向がある。そのため、他の線画と同じ近似精度のベクタ表現は効率が悪い。文字の歪みが目立たない精度で線画をベクタ化すると、線画の近似精度も向上するが、それに見合うだけの主観品質が向上しない。逆に、線画を基準に近似精度を設定して文字をベクタ化すると、歪みが目立ち主観品質が著しく低下する。このような理由から、文字認識の有無にかかわらず、文字の分離は主観品質を向上させる上で有効である。文字認識だけではなくフォントも特定できれば、文字の輪郭をベクタ表現に変換する必要はなく、テキストとフォント情報に分離できる。

画像インデキシングのアプリケーション例として、セリフによるコマ検索が挙げられる。また、半自動的にコンテンツの多言語化を行う場合、あらかじめ文字列領域が抽出され自動翻訳による下訳が行われていると、作業者の負担が著しく軽減される。

線画

線画は解像度変換（拡大処理）によりジャギーが発生する。そのため、ベクタ化によって、ディスプレイの解像度で滑らかな線画が得られる。グレースケール表示可能な場合にはアンチエイリアシング処理により、さらに主観品質の高い線画が得られる。

線画はキャラクタや背景を構成しており、多くの場合それぞれの線幅は異なる。解像度が低い場合には線幅の差異は判別できないため、線幅を無視して中心線によって符号化できる。一方、解像度が十分に高い場合には線幅を考慮する必要があるため、中心線ではなく輪郭線（白と黒の境界線）による符号化が実用的である。このように、表示デバイスの解像度に対して、必要十分なベクタ変換の精度が存在する。

網点

網点とは、2値画像で中間調を擬似的に再現する印刷の手法である。広義にはテクスチャに包含されるが、解像度変換（縮小処理）によりモアレが発生するため、特に分けて考える必要がある。

単位面積あたりの黒領域の面積が濃度に相当する。また、点は格子状に配置されるため、周期的に濃度が変化する。点同士の距離は線数（1インチあたりの点の個数）で表される。縮小処理はダウンサンプリングと等価であり、網点の周期とサンプリング周波数が整数倍の関係にない場合、うねりという現象が生じる。これがモアレとして知覚され、主観品質を著しく低下させる。

テクスチャ

大部分のコミックは2値画像であるから、テクスチャもベクタ化できる。従って簡易的なシステムでは特に分けて考える必要はない。しかし、ベクタ表現は孤立点などのノイズや細

かい模様に対して、通過点が多数発生するため符号化効率が悪い。従って、テクスチャの品質が主観品質に与える影響を考慮して、ラスタ表現の符号化方式にゆだねることが有効である。

7.2.2 画像の入出力形式

入力画像には、高解像度2値画像と低解像度グレースケール画像を想定する。前者はスキャナを用いて得られた画像である。ここで、300dpi以上であると仮定する。後者はすでに蓄積されているデジタル画像である。1800×1200画素のグレースケール画像が多く、原稿がB6サイズの場合150dpiに相当する。

出力形式は既に広く普及しているベクタグラフィックスのプラットフォームであるSVGやFlash、PostScriptファイルとなる。また、テクスチャはラスタ表現のまま符号化され、ベクタグラフィックスに含めることもできる。ディスプレイが2値の場合にグラデーションを出力するにはディザ法の適用が考えられるが、提案システムは表示に関する具体的な手法を限定しない。

7.2.3 符号化システムのブロック図

コミックの構成をふまえ、信号特性と利用方法を考慮したコミック符号化システムを図7-1に示す。コミック符号化システムは、まず、網点の分離、線画の分離、文字領域の抽出から構成される。さらに、分離した網点はベクタ表現に最適な形式として、網点領域の境界線とグラデーションに分けて表現する。境界線と線画はそれぞれベクタ化によりベクタ表現に変換される。抽出された文字領域は既存のOCRにより文字として認識する。ただし、線画の分離手法は十分確立されておらず、本論文では網点以外をすべて線画として扱う。なお、ベクタ表現では、曲線の通過点と制御点の座標が符号量の大部分を占める。そこで、点の個数と座標精度に着目して符号量制御を実現する。

7.3 符号量制御

7.3.1 スキャン解像度と符号量制御

ベクタ表現では表示解像度によらず高品質な画像を得られるが、高解像度のラスタ表現からファイルサイズの小さい高品質なベクタ表現は得られない。ファイルサイズが増加する原因として、輪郭線上に表れるノイズにより曲線の通過点が増大すること、画素精度でベクタ変換が行われることにより、近似誤差の許容値が相対的に小さくなることが考えられる。低解像度のラスタ表現からはファイルサイズの小さい、実用的な品質のベクタ表現を得られる。すなわち、ラスタ表現の解像度によってベクタ表現の符号量と品質を制御できる。解像度変換は計算コストが高いため、ベクタ表現で閉じた符号量制御が求められる。

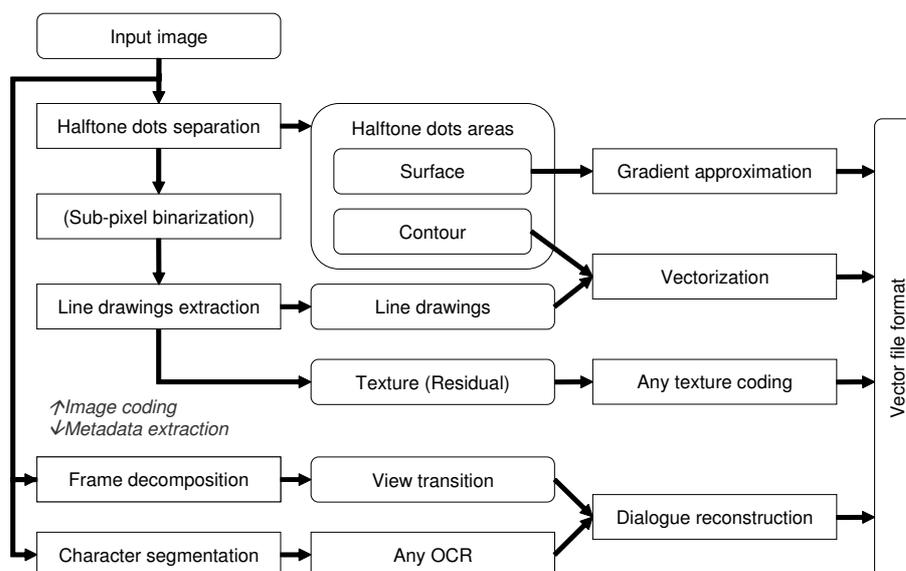


図 7-1: Block Diagram of Comic Image Coding System

7.3.2 座標解像度の削減

ベクタ表現は、画像を線、面、グラデーションなどの幾何的、数学的な要素を組み合わせで表現する方式である。そして、線や面の輪郭線は通過点と制御点を用いて表現される。通過点同士の間は数式により滑らかに補完される。また、通過点と制御点は適当な座標系で表現される。そこで、座標精度を落とすことで符号量を削減できる。

従来手法 [7] では、ベクタ変換において直線により輪郭線の最適近似を行う。そして、直線同士の交点の座標が小数点精度で求められる。次に、ファイルに出力する際にはスケールングを行い、座標を整数化して出力する。デフォルトでは 10 倍の座標精度を用いる。

しかし、入力解像度に対して出力解像度が低い場合には、10 倍の座標精度を確保しても有効に利用できない。実際、本論文の入力解像度として利用している 1200dpi のビットマップを表示できるディスプレイは現状では存在せず、150dpi や 200dpi 程度である。そこで、座標精度を等倍や 1/2 倍にして符号量を節約できる。

7.3.3 直線・曲線の統合

従来手法 [7] の曲線最適化では、統合できる線の組み合わせに様々な制約が課せられている。まず、連続する曲線だけが統合され、角を構成する直線は統合できない。次に、曲線の凸方向が異なる曲線同士、または変曲点を削減する統合はできない。最後に統合前後で大局的な曲線の凸方向が変わる統合はできない。

しかし、想定する入力画像では、輪郭線上に表れるノイズにより図 7-2(b) のように最適多角形の凸方向は頻繁に変化する。そこで、これらの制約を緩和する手法を検討する。また、ベクタ表現だけで符号量制御を行うために、曲線最適化処理を拡張する。最終的には統合さ

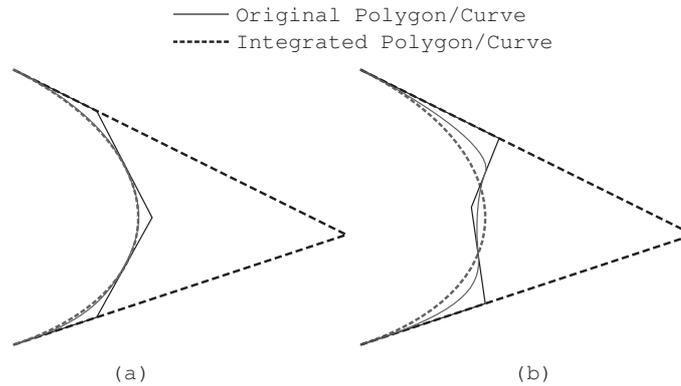


図 7-2: 凸方向の異なる頂点の統合 (自発表 [10] より引用)

れない曲線を含めて一貫性のあるベクタ表現を得るために、多角形をなめらかな曲線に置き換える手法を再適用する。

通過点削減に必要な緩和条件として、角を構成する直線も含めた統合と凸方向の異なる頂点の統合が考えられる。さらに、統合後に曲線が存在するためには、両端を構成する辺を延長し新たな頂点が作れることが必要条件となる (図 7-2)。これは、統合前後で大局的な曲線の凸方向を維持するのと同じである。

統合前の曲線ごとの頂点を代表点とする。次に代表点の接線に平行な直線が統合後曲線と接する点を求め、これらの距離を誤差として計算する。全ての点において誤差がしきい値以下であれば統合可能と判断し、誤差を累積する。

以上の誤差計算を任意の線の組み合わせに対して計算する。そして、統合後の本数が少なく、累積誤差の小さい最適な組み合わせを求める。誤差のしきい値を変化させることで、符号量が制御可能になる。

本手法は、ベクタ表現のまま通過点を削減し、符号量を削減する。そのため、ラスライズする必要がない。また、許容誤差に応じて統合される線が順次決定されるため、符号量制御が容易である。

7.4 エントロピー符号化

7.4.1 係数の分類

ベクタ表現は閉曲線の集合として表現される。それぞれの閉曲線は、絶対座標による開始点、直前の点との相対座標による通過点、各点間の曲線の種類と膨らみ具合を表す付加情報からなる。付加情報は本来 0~1 の値を持つが、適当な精度を維持するために近傍の相対座標値によりスケールされ、整数として表される [7]。以上の 4 種類について、係数の発生モデルを検討する。

本論文では汎用のテキスト符号化に用いられるモデルを利用し、BWT (Burrows-Wheeler Transform)、LZ77 (Lempel-Ziv 77)、PPM (Prediction by Partial Match) を比較する。

表 7-1: 各係数におけるモデルによる符号化効率の比較

	Symbol	Entropy	Codes	no model	BWT	LZ77	PPM
coorda_m	34,124	7.3	31,154	30,724 1.4%	30,066 3.5%	25,185 19.2%	33,764 -8.4%
coorda_s	34,124	7.2	30,431	30,255 0.6%	28,009 8.0%	24,511 19.5%	31,838 -4.6%
command	185,957	1.7	40,374	25,292 37.4%	23,429 42.0%	28,217 30.1%	23,296 42.3%
addinfo	257,554	3.9	122,491	118,932 2.9%	119,751 2.2%	124,596 -1.7%	134,804 -10.1%
coordr_m	320,728	6.6	262,490	263,382 -0.3%	259,259 1.2%	258,490 1.5%	287,654 -9.6%
coordr_s	320,728	6.6	261,210	257,111 1.6%	257,202 1.5%	252,970 3.2%	286,534 -9.7%
total	578,282	6.3	454,506	431,590 5.0%	428,391 5.7%	430,294 5.3%	476,472 -4.8%

いずれもエントロピー符号化には Rangelocoder を利用する。

7.4.2 発生頻度と分布特性

4種類の係数について、モデルなしを含めた四つのモデルの符号化効率を確認する予備実験を行う。

テストシーケンスは 1200dpi, B5 サイズの線画像を 2 値画像としてスキャンした画像である。3 枚分の画像に対してベクタ変換を適用して符号化を行い、発生符号量を合計した。表 7-1 に実験結果を示す。

Symbol は係数の個数, Entropy は平均符号長 (bit/symbol), Codes は平均符号量 (byte) を示す。表中の数値は、モデルごとの発生符号量 (byte) と平均符号量に対するゲイン (%) を示す。座標データに関して、 x 座標と y 座標をインタリーブの有無により比較を行った。coorda_m はインタリーブした開始点, coorda_s はインタリーブしない開始点, command は曲線の種類, addinfo は付加情報, coordr_m はインタリーブした通過点, coordr_s はインタリーブしない通過点, total は coorda_s, command, addinfo, coordr_s の合計を表す。

表 7-1 より、座標データはインタリーブしない方が効率が良いことがわかる。開始点は画像に依存するが、本実験では LZ77 方式が最適であった。曲線の種類を表す記号はただか 6 種類しかなく、同種の記号が連続するため圧縮されやすい。本実験では BWT モデルと PPM モデルが同程度で最適である。

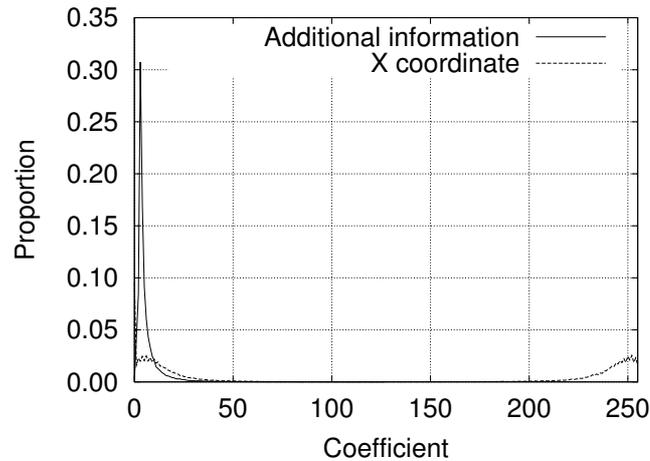


図 7-3: 曲線の付加情報と通過点 (x 座標) の係数分布 (自発表 [10] より引用)

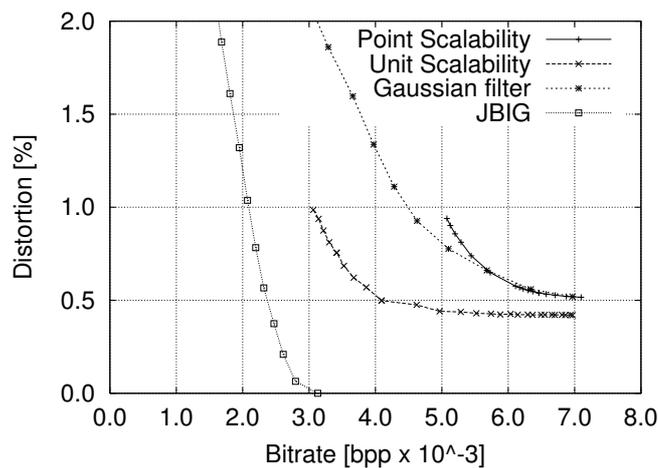


図 7-4: 各種法におけるレート歪み特性 1 (自発表 [10] より引用)

付加情報と通過点の x 座標の分布を図 7-3 に示す。付加情報は正の値だけを持ち、分布は非常に偏っている。平均符号量と比較してほとんどゲインが得られず、モデルなしが最適である。通過点は正負の値を持ち、分布は 0 近傍と 255 近傍に偏っている。ここで、255 は -1 を表す。いずれのモデルも平均符号量と比較してほとんどゲインを得られず、モデルなしで十分である。すべてを同じモデルで扱う場合には BWT モデルが最適である。

7.5 実験と考察

7.5.1 座標解像度の削減と直線・曲線の統合

テスト画像は以下のようにして得る。網点と線画から構成される印刷物を 2 値、1200dpi でスキャンする。本実験では線画に関する評価を行うため、網点をスキャン後に除去する。

提案する座標解像度の削減 (Unit Scalability) と直線・曲線の統合 (Point Scalability) によるレート歪み特性を図 7-4 に示す。提案手法と比較する従来手法として、分散値を変化させてガウスフィルタを適用した後、ベクタ表現に変換する手法 (Gaussian filter)、同様にガウスフィルタを適用した後、JBIG により可逆圧縮を行う手法 (JBIG) を用いる。

縦軸は、入力画像とベクタ変換後ラスタライズした画像の差分画素の全画素に対する割合を示す。横軸は、画素当たりのビットレートを示す。左下の系列ほど性能が良いことを示す。ベクタ表現には PostScript を利用し、さらにエントロピー符号化するために汎用圧縮方式の bzip2 を適用する。提案手法は差分の割合が 1% 程度で主観品質が大幅に悪化したため、打ち切っている。

座標解像度の削減による符号量制御は、 $4.0\text{bpp} \times 10^{-3}$ 以上ではビットレートだけが変化し、歪みはほとんど変化しない。さらに、 $4.0\text{bpp} \times 10^{-3}$ 以下では、ビットレートの減少と共に歪みも増加する。これは、ベクタ変換において座標解像度を 10 倍にスケールし、整数化して出力していることが影響している。座標解像度が等倍になるのが $4.0\text{bpp} \times 10^{-3}$ の時であり、座標解像度が等倍以下になると初めて歪みも増加する。

座標解像度を等倍より大きくスケールしても、歪みはほとんど減少できないことが明らかとなった。一方で、座標解像度を等倍以下にすると、ガウスフィルタより穏やかな歪みの増加が明らかとなった。また、提案手法は通過点同士の距離が十分離れている場合に有効である。さらに、ガウスフィルタや通過点の削減とは独立に制御できるため、いずれの方法とも組み合わせることができる。

直線・曲線の統合による符号量制御は、ビットレートが高い間はガウスフィルタによるレート歪み特性と同程度の性能である。しかし、ビットレートを低くすると性能が低下し、主観品質も悪化する。

ガウスフィルタは輪郭線上のノイズを削減するだけでなく、輪郭線そのものを削除できる。しかし、提案手法は輪郭線の削除を伴う符号量制御はできない。従って、大幅な符号量の削減には、輪郭線の平面的な位置関係から輪郭線そのものを削減する必要がある。同時に、細部をつぶさずに保存しながら、符号量を削減できる有効な手法である。

以上をまとめると、いずれの手法でも符号量制御を実現できることが明らかとなった。また、座標解像度の削減手法はガウスフィルタを適用するより性能が良く、直線・曲線の統合手法はガウスフィルタと同程度の性能が得られる。

7.5.2 エントロピー符号化

本節では、提案するモデル化を利用してエントロピー符号化を行った場合の、各種符号量制御の結果を示す、考察を述べる。

テスト画像は 1200dpi、B5 サイズの線画像を 2 値画像としてスキャンした画像である。座標解像度は 2~1/4 倍 (2, 1, 2/3, 1/2, 2/5, 1/3, 2/7, 1/4 倍) まで変化させた。テスト画像のうち 1 枚分のレート歪み特性を図 7-5 に示す。

縦軸は、入力画像とベクタ変換後ラスタライズした画像の差分画素の全画素に対する割合

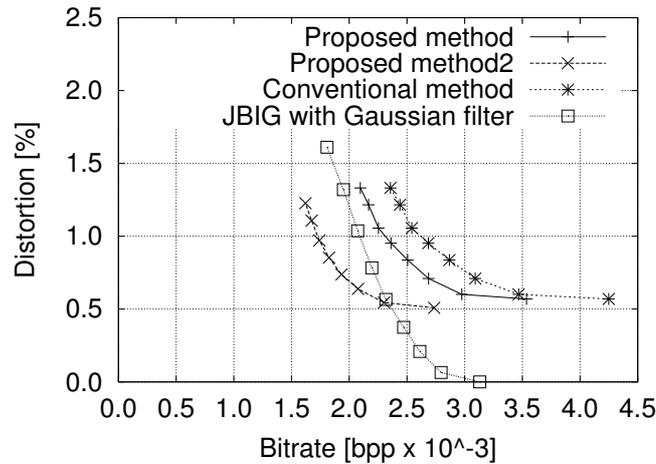


図 7-5: 各種法におけるレート歪み特性 2 (自発表 [10] より引用)

を示す。横軸は、画素当たりのビットレートを示す。左下の系列ほど性能が良いことを示す。Proposed method はモデル化だけを行った場合，Proposed method2 は分散 2 のガウスフィルタを適用した後，モデル化を行った場合，Conventional method はモデル化を行わずに汎用のテキスト圧縮プログラムを利用した場合，JBIG with Gaussian filter は分散の値を 0～4 まで 0.5 間隔で変化させたガウスフィルタを適用した後 JBIG によりロスレス符号化を行った場合である。

図 7-5 より，同品質で比較してモデル化の適用により $0.3 \sim 0.8 \times 10^{-3} \text{ bit/pel}$ ，ガウスフィルタの適用により $0.8 \sim 1.5 \times 10^{-3} \text{ bit/pel}$ の符号量をそれぞれ削減できることが確認できる。

座標解像度が等倍のベクタ表現は，ロスレスの JBIG よりビットレートが低くなり，符号化として実用的であると考えられる。

7.5.3 ファイルサイズとモアレ低減

提案したコミック符号化システムを計算機上に実装し，実際のコミックと計算機で生成した網点画像を用いて，評価実験を行った。コミックは B6 サイズの原稿を 300dpi でスキャンした。画像の概要を 7-6 に示す。

まず，ベクタ変換後のファイルサイズについて述べる。提案手法と“従来手法 1”によるファイルサイズを 7-2 に示す。また，それぞれの内訳として線画と網点領域のファイルサイズを示す。さらに，参考として JBIG により符号化した場合のファイルサイズも示す。なお，内訳はそれぞれを符号化した際のファイルサイズであるため，内訳の合計は全体と一致しない。

この結果より，提案手法はファイルサイズを半減し，ファイル全体に占める網点の割合も著しく減少させている。また，JBIG と比較してもファイルサイズが小さい。なお，No.1 の線画のファイルサイズが半減している要因は，一部のテクスチャが網点として分離されていることにある。ここで，2 値を多値に変換した場合の客観指標は十分確立していないため，

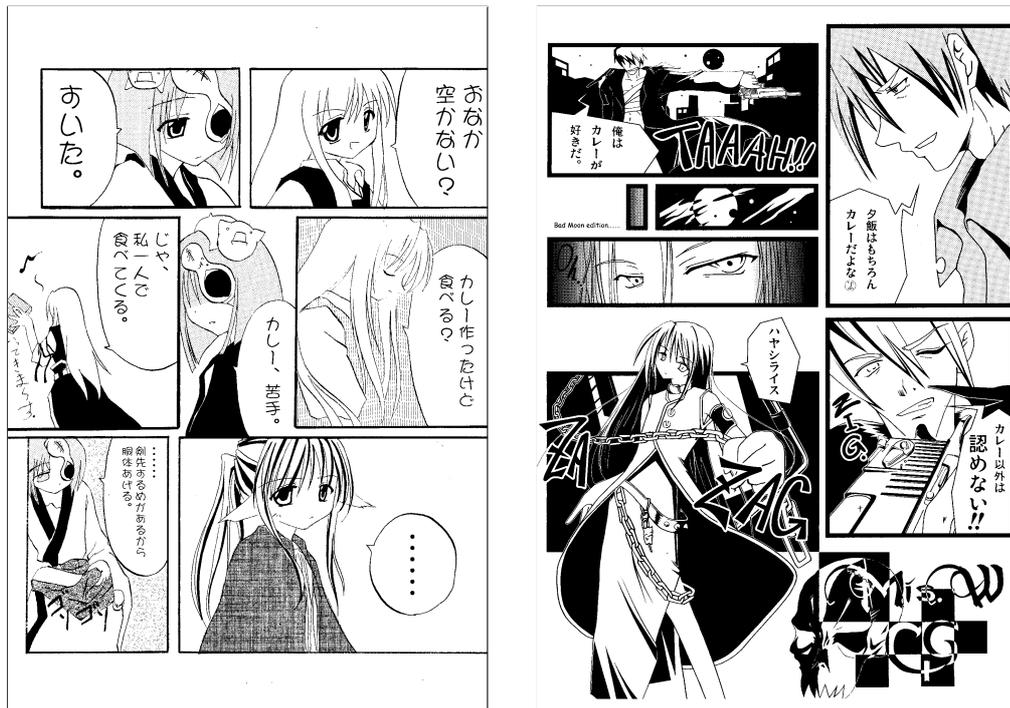


図 7-6: 入力するコミック画像のサムネイル (No. 1, No. 2) (自発表 [73] より引用)

表 7-2: File Size of Comic Image by each Coding Method [byte]

No.	1	2
Proposed method	40,691	36,732
Line Drawings	37,084	34,745
Halftone Areas	4,384	2,287
Conventional method 1	99,395	77,919
Line Drawings	77,278	39,211
Halftone Areas	21,046	40,631
JBIG	52,704	50,511

画像品質の客観評価が今後必要である。

また、PDA にベクタ表現のコミックを表示している様子を図 7-7 に示す。入力画像は B6 サイズの原稿を 300dpi でスキャンした 2 値画像である。出力は EPS を経由した Flash ファイルである。PDA のディスプレイは 320×480 画素 (約 140dpi) で、FlashPlayer を用いて表示している。

PDA のように携帯電話に比べて大きなディスプレイでは、ページ単位の閲覧ができることがわかる。逆に、これよりも小さなディスプレイではコマのようなページ未満の大きさで閲覧する必要がある。



図 7-7: PDA によるベクタ表現コミックの提示例 (自発表 [73] より引用)

7.6 むすび

本章では、ベクタ化による人工画像の高効率符号化を実現する符号化フレームワークを提案した。提案フレームワークは、コミック画像から、網点領域を分離してグラデーションで近似する一方で、線画像と網点領域をベジエ曲線に変換した。また、網点と線画以外の成分はテクスチャとして既存のラスタ符号化方式を適用した。これらを多重化してビットストリームとした。また、符号化に加えて文字列や視線遷移情報(コマ分割情報)をメタデータとして付加した。

高効率符号化を実現するために、座標の位置情報を対象として最適なロスレス符号化を検討した。また、座標情報を対象に符号量制御手法を提案した。ベクタ表現上で適用可能な符号量制御として、座標解像度の量子化と直線・曲線の統合手法を提案した。エントロピー符号化として、BWT アルゴリズムの適用が平均的に良い性能が得られることを明らかにした。さらに、従来のロスレス画像符号化方式とレート歪み特性を比較し、コミック符号化システムの有効性を示した。

第8章

結論

本章では本論文を総括し、今後の検討課題を述べる。

8.1 総括

本論文では、ラスタ表現の静止画像コンテンツをベクタ表現に変換し、符号化することを目的とした。網点分離及び線形勾配領域取得による画像分解方式と、ベクタ表現への自動変換方式を提案した。

第3章では、印刷物において階調表現を実現する網点技術が、解像度変換に伴うアーティファクト発生の原因であることを示した。原因となっている輝度変化の周期を取得し、得られた周期を利用して網点を分離するとともに、階調を復元して線形勾配で近似する手法を提案した。勾配変数が既知のテスト画像を用いた評価実験を行い、提案手法の近似性能が実用的であることを示した。

さらにコミックを代表とする需要の大きい印刷物を高品質なベクタ表現に変換可能とした。その結果、既存のコンテンツを読書に適した高精細出力装置で鑑賞できるようになった。特に、ワンソースで表示端末の性能（解像度）に適した表示ができるため、コンテンツ配信サーバにおいて効率的に蓄積できる。

第4章では、アニメーション画像やイラストのような人工的な画像をベクタ表現に変換するための画像分解方式を提案した。まず、画像をベクタ表現に適した線形勾配成分とベクタ表現に適さないテクスチャ成分に分解する画像モデルを想定した。そして、均一輝度成分を全変動最小化の枠組みで取得する手法において、高速計算手法を提案した。評価実験により提案手法は従来手法と同等の分解性能を維持し、処理時間を削減することを示した。次に、均一輝度成分を含む線形勾配成分を、先の高速計算手法によって取得する方式を提案した。評価実験により提案手法はベクタ表現可能な領域を増加させることを示した。さらに、線形勾配だけでなく、グラデーション全般をベクタ表現可能なグラデーションメッシュを用いた分解モデルを想定し、処理時間を検討した。

第5章では、平面曲線の曲率に関する不変特徴量を提案した。得られた最適スケールにおける曲率を用いて曲線の頂点を定義した。曲線の頂点とベジエ曲線の通過点が一致しているほど編集性が高いため、提案した曲線の頂点の定義を用いて、ベクタ変換によって得られる

ベジエ曲線の編集性を評価した。

第6章では、コミック画像に文字やコマ配置など利用価値の高いメタデータが含まれることを指摘した。文字切出し手法と画像分割に基づくコマ閲覧順序取得手法を提案し、これらのメタデータが自動的に取得できることを明らかにした。

特に、コマ閲覧順序は視線遷移情報と等価である。これを用いると、携帯電話端末のように小さなディスプレイでコマ単位のコミック閲覧を提供する場合、コマ分割にかかる作業者の負担を大幅に削減できると期待される。また、文字コードの取得によりセリフによるテキスト検索等が可能になり、2次利用が促進されうる。

第7章では、本論文で提案してきた個々の要素技術を統合し、コミック画像符号化のフレームワークを提案した。また、符号量制御手法を提案し、エントロピー符号化についてまとめた。

以上のように本論文では、取得が容易なラスタ表現の静止画像コンテンツを自動的にベクタ表現に変換し、高能率符号化を具現化する手法について論じた。すなわち、既存の蓄積コンテンツを将来の高精細出力装置で鑑賞に堪えうる品質に、低コストかつ自動で変換できることを示した。また、本論文で述べたベクタ変換はフォントが存在しない文字や図表においても有効である。

8.2 今後の課題

残された課題として以下のものが挙げられる。

本論文で提案したベクタ表現による静止画像符号化の動画像符号化への拡張が挙げられる。動画像の符号化においては、フレーム間の冗長度削減が必須である。MPEGに代表される既存の符号化方式では、オブジェクト符号化を含めて、画素領域における動き補償を用いた冗長度削減が基本である。今後の研究においては、時間方向予測とベクタグラフィックスの融合が必須となり、その際には各フレームに存在する人工画像特性の相関性の検討が必要である。

人工画像のモデル化としてグラデーションメッシュによる画像分解を検討した。本論文では数値計算手法に着眼して評価したが、実用化に際しては画像の特性を利用した計算の安定性向上や処理量削減が必要である。

コミック画像に含まれるメタデータとして文字とコマ配置を挙げた。コミック画像には複数のキャラクターが登場するなど、より意味レベルの高いメタデータが豊富に含まれている。また、「絵のタッチ」や「画風」と呼ばれるメタデータもある。このようなメタデータ抽出方式と、得られたメタデータを統合して作品の中身に基づく推薦システム等の実現が課題である。

謝辞

本研究の機会及び素晴らしい実験環境を与えて下さり、貴重な時間を割いて研究の方向性を御指導頂きました早稲田大学大学院 国際情報通信研究科 渡辺 裕 教授に心から感謝致します。

本研究を進めるにあたり、日頃から惜しみなく御指導して頂きました早稲田大学 富永 英義 名誉教授に深く感謝致します。

研究の方向性をはじめ、研究の細部に至るまで、数々の有意義な御意見、御助言を賜りました早稲田大学大学院 国際情報通信研究科 亀山 渉 教授に深く感謝致します。

本論文をまとめるにあたり貴重な御示唆を賜りました、早稲田大学大学院 国際情報通信研究科 津田 俊隆 教授ならびに早稲田大学大学院 基幹理工学研究科 甲藤 二郎 教授に心よりお礼申し上げます。

本研究のきっかけを与えて下さり、研究の進め方から文章の書き方まで丁寧に御指導下さった平成 15 年度 修士卒 の 中神 央二 氏にはこの場を借りて心から深く感謝致します。

貴重な御意見、様々な御提案を頂いた AVS ゼミの皆様にお礼申し上げます。特に、研究室でいろいろと御世話を下さった早稲田大学 国際情報通信研究センター 招聘研究員 石川 孝明 氏に心から感謝致します。

職場での研究活動において、常日頃より研究の方向性を導いて戴いた株式会社 KDDI 研究所 代表取締役所長 中島 康之 博士に心よりお礼申し上げます。また直属の上司として、超高精細映像符号化技術の細部にわたり有益かつ緻密な御意見、御助言を賜りました株式会社 KDDI 研究所 超臨場感通信グループリーダ 内藤 整 博士に深謝致します。

本研究における第 6 章は、石井 大祐 氏（現 早稲田大学大学院 国際情報通信研究科 助手）の平成 19 年度修士論文の研究を筆者と共同で行ったものであり、研究を進める上での有益な御討論、実験等にご助力を頂きました。ここに厚く御礼申し上げます。

最後に、研究室の皆様、株式会社 KDDI 研究所の方々、ならびに関係各位に心から感謝申し上げます。

平成 25 年 2 月

参考文献

- (1) 高木利弘, 電子書籍ビジネス調査報告書 2006, 株式会社インプレス R&D, 2006.
- (2) 高木利弘, 電子コミックビジネス調査報告書 2006, 株式会社インプレス R&D, 2006.
- (3) “E ink corporation,” <http://www.eink.com/>, 2009.
- (4) “Scalable Vector Graphics (SVG) XML Graphics for the Web,” <http://www.w3.org/Graphics/SVG/>, 2005.
- (5) “Macromedia flash file format (swf) specification,” <http://www.macromedia.com/software/flash/open/licensing/fileformat/>, 2005.
- (6) “Openswf.org,” <http://www.openswf.org/>, 2005.
- (7) P. Selinger, “Potrace: a polygon-based tracing algorithm,” <http://potrace.sourceforge.net/potrace.pdf>, 2003.
- (8) 河村圭, 渡辺裕, “ベクター変換における曲線最適化アルゴリズムの一検討,” FIT2004 第3回情報科学技術フォーラム, no.3 S-2, Sept. 2004.
- (9) G. Medioni, and Y. Yasumoto, “Corner detection and curve representation using cubic b-splines,” IEEE International Conference on Robotics and Automation, vol.3, pp.764-769, Apr. 1986.
- (10) 河村圭, “文字や網点を含む2値画像のベクター表現による画像符号化の検討,” 修士論文, 早稲田大学大学院 国際情報通信研究科, 2005.
- (11) 臼淵徹, “文字・写真混在画像のデータ圧縮符号化,” 電子情報通信学会技術研究報告 IE, vol.83, no.173, pp.19-24, Oct. 1983.
- (12) 村山登, 志村浩, “文字と写真の自動適応処理,” 画像電子学会誌, vol.15, no.4, pp.265-273, Oct. 1986.
- (13) 茨木久, 小林誠, 越智宏, “網点写真の識別処理方法,” 電子情報通信学会論文誌 B, vol.J70-B, no.2, pp.222-232, Feb. 1997.
- (14) P.W. Wong, and N. Memon, “Image processing for halftones,” IEEE Signal Processing Magazine, vol.20, no.4, pp.59-70, July 2003.
- (15) 上野博, 辻建三, “網点写真の二値化法,” 画像電子学会誌, vol.15, no.4, pp.377-388, Oct. 1986.
- (16) 大内敏, 今尾薫, 山田和作, “文字/絵柄(網点, 写真)混在画像の像域分離方式,” 電子情報通信学会論文誌 D-II, vol.J75-D-II, no.1, pp.39-47, Jan. 1992.

- (17) 洋鎌田, 克仁藤本, “低解像度テキスト画像の高速かつ高精度な2値化方式,” 電子情報通信学会技術研究報告 PRMU, vol.98, no.490, pp.117-124, Dec. 1998.
- (18) 克仁藤本, 洋鎌田, “低解像度カラー文書画像から高品質な文字画像を抽出する二値化方式,” 電子情報通信学会技術研究報告 PRMU, vol.99, no.380, pp.7-14, Oct. 1999.
- (19) C. Christopoulos, W. Philips, A. Skodras, and J. Cornelis, “Segmented image coding: Techniques and experimental results,” Signal processing: Image communication, vol.11, no.1, pp.63-80, Nov. 1997.
- (20) R. de Queiroz, “Pre-processing for mrc layers of scanned images,” IEEE International Conference on Image Processing, pp.3093-3096, Oct. 2006.
- (21) S. Carlsson, “Sketch based coding of grey level images,” Signal Processing, vol.15, no.1, pp.57-83, July 1988.
- (22) 河村圭, 石井大祐, 渡辺裕, “全変動最小化の高速計算手法,” 電子情報通信学会論文誌 D, vol.J93-D, no.3, pp.326-335, March 2010.
- (23) A. Chambolle, “An algorithm for total variation minimization and applications,” Journal of Mathematical Imaging and Vision, vol.20, no.1-2, pp.89-97, Jan. 2004.
- (24) P. Combettes, and J.C. Pesquet, “Image restoration subject to a total variation constraint,” IEEE Transactions on Image Processing, vol.13, no.9, pp.1213-1222, Sept. 2004.
- (25) J. Sun, L. Liang, F. Wen, and H.Y. Shum, “Image vectorization using optimized gradient meshes,” ACM Transactions on Graphics, vol.26, no.3, p.11, July 2007.
- (26) 河村圭, 石井大祐, 渡辺裕, “グラデーションメッシュによる画像符号化の基礎検討,” 画像符号化シンポジウム PCSJ2008 (P-5.04), pp.71-72, Nov. 2008.
- (27) 金谷健一, これなら分かる最適化数学, 共立出版, 2005.
- (28) 石井大祐, “電子コミック画像のコマ分割処理,” 修士論文, 早稲田大学大学院 国際情報通信研究科, 2008.
- (29) M. Yamada, R. Budiarto, M. Endo, and S. Miyazaki, “Comic image decomposition for reading comics on cellular phones,” IEICE transactions on information and systems, vol.E87-D, no.6, pp.1370-1376, June 2004.
- (30) 辻善丈, “スプリット検出法による文書画像構造解析,” 電子情報通信学会論文誌 D-II, vol.J74-D-II, no.4, pp.491-499, Apr. 1991.
- (31) K. Wong, R. Casey, and F. Wahl, “Document analysis systems,” IBM Journal of Research and Development, vol.26, no.6, pp.647-656, Nov. 1982.

- (32) K. Kise, A. Sato, and M. Iwata, "Segmentation of page images using the area voronoi diagram," *Computer Vision and Image Understanding*, vol.70, no.3, pp.370-382, June 1998.
- (33) D. Ballard, "Generalizing the hough transform to detect arbitrary shapes," *Pattern Recognition*, vol.13, no.2, pp.111-122, 1981.
- (34) 木村彰男, 渡辺考志, "高速一般化ハフ変換-相似変換不変な任意図形検出法," *電子情報通信学会論文誌 D-II*, vol.J81-D-II, no.4, pp.726-734, Apr. 1988.
- (35) 田中孝昌, 東海林健二, 外山史, "マンガ画像のコマ割り構造解析," *電子情報通信学会総合大会講演論文集*, no.D-12-89, March 2006.
- (36) 後藤英昭, 阿曾弘具, "文字行の局所的な直線性を利用した頑健・高速な文字行抽出法," *電子情報通信学会論文誌 D-II*, vol.J78-D-II, no.3, pp.465-473, March 1995.
- (37) 志久修, 姉川正紀, 中村千秋, 中村彰, "地図からの文字の切り出し," *情報処理学会論文誌*, vol.34, no.2, pp.273-280, Feb. 1993.
- (38) 河村圭, "網点を含む2値画像のベクトル表現," *卒業論文*, 早稲田大学理工学部 電子・情報通信研究科, 2004.
- (39) 渡辺裕, "ベクター表現によるマンガ符号化の研究," *大川情報通信基金 2005年度研究助成 研究報告書*, 2006.
- (40) O. Trier, and A. Jain, "Goal-directed evaluation of binarization methods," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol.17, no.12, pp.1191-1201, Dec. 1995.
- (41) 河村圭, 石井大祐, 渡辺裕, "ベクター表現に適した小数画素精度の改良境界線検出に関する検討," *情報処理学会研究報告*, 2007-AVM-56, pp.61-64, March 2007.
- (42) 河村圭, 渡辺裕, 富永英義, "マンガ符号化における網点の多階調近似処理の評価," *情報処理学会全国大会講演論文集*, no.3 Z-1, pp.2.375-2.376, March 2004.
- (43) 河村圭, 渡辺裕, 富永英義, "マンガ符号化における濃度勾配検出法の改良," *情報処理学会研究報告*, 2004-AVM-44, pp.13-18, March 2004.
- (44) W.K. Pratt, *Digital Image Processing*, John Wiley & Sons, Inc., 1991.
- (45) 河村圭, 山本勇樹, 石井大祐, 渡辺裕, "ベクター表現による文字や網点を含む画像符号化システム," *情報処理学会全国大会講演論文集*, no.3 B-1, pp.2.13-2.14, March 2007.
- (46) 河村圭, 山本勇樹, 渡辺裕, "文字や網点を含む低解像度多値画像のベクター表現に関する検討," *FIT2006 第5回情報科学技術フォーラム*, no.J-032, Sept. 2006.

- (47) 田中謙輔, 凸解析と最適化理論, 牧野書店, 1994.
- (48) L.I. Rudin, S.J. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Physica D*, vol.60, pp.259-268, Nov. 1992.
- (49) T.F. Chan, S.E. Glu, and M. Nikolova, "Algorithms for finding global minimizers of image segmentation and denoising models," *SIAM Journal on Applied Mathematics*, vol.66, no.5, pp.1632-1648, 2006.
- (50) T.F. Chan, S. Osher, and J. Shen, "The digital TV filter and nonlinear denoising," *IEEE Transactions on Image Processing*, vol.10, no.2, pp.231-241, Feb. 2001.
- (51) L.A. Vese, and S.J. Osher, "Modeling textures with total variation minimization and oscillating patterns in image processing," *Journal of Scientific Computing*, vol.19, no.1, pp.553-572, Dec. 2003.
- (52) J. Carter, *Dual Methods for Total Variation-Based Image Restoration*, Ph.D thesis, UCLA, Jan. 2002.
- (53) 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 (画像・映像処理)," *電子情報通信学会論文誌 D-II*, vol.J88-D-II, no.6, pp.1024-1034, June 2005.
- (54) 河村圭, 石井大祐, 渡辺裕, "領域分割に基づくベクタ表現を用いた画像符号化方式の検討," *画像符号化シンポジウム PCSJ2009 (P-3.09)*, pp.41-42, Nov. 2009.
- (55) Z. Wang, A.C. Bovik, H.R. Sheikh, S. Member, E.P. Simoncelli, and S. Member, "Image quality assessment: From error visibility to structural similarity," *IEEE Transactions on Image Processing*, vol.13, no.4, Apr. 2004.
- (56) M. Frigo, and S.G. Johnson, "FFTW home page," <http://www.fftw.org/>, 2009.
- (57) D.G. Lowe, "Distinctive image features from scale-invariant keypoints," *International Journal of Computer Vision*, vol.60, no.2, pp.91-110, Jan. 2004.
- (58) M. Bober, "MPEG-7 visual shape descriptors," *IEEE Transactions on Circuits and Systems for Video Technology*, vol.11, no.6, pp.716-719, June 2001.
- (59) L. Latecki, R. Lakämper, and T. Eckhardt, "Shape descriptors for non-rigid shapes with a single closed contour," *IEEE Conference on Computer Vision and Pattern Recognition*, vol.1, pp.424-429, June 2000.
- (60) ISO/IEC JTC1/SC29/WG11, "ISO/IEC 15938-3:Information technology - Multimedia content description interface - Part 3: Visual," , 2002.

- (61) A.P. Witkin, "Scale space filtering," 8th International Joint Conference on Artificial Intelligence, pp.1019-1022, 1983.
- (62) F. Mokhtarian, and A. Mackworth, "A theory of multiscale, curvature-based shape representation for planar curves," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol.14, no.8, pp.789-805, Aug. 1992.
- (63) B. Zhong, and W. Liao, "Direct curvature scale space: Theory and corner detection," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol.29, no.3, pp.508-512, March 2007.
- (64) M. Awrangjeb, and G. Lu, "An improved curvature scale-space corner detector and a robust corner matching approach for transformed image identification," IEEE Transactions on Image Processing, vol.17, no.12, pp.2425-2441, Dec. 2008.
- (65) M.R. Daliri, and V. Torre, "Classification of silhouettes using contour fragments," Computer Vision and Image Understanding, vol.113, no.9, pp.1017-1025, Sept. 2009.
- (66) K. Kawamura, D. Ishii, and H. Watanabe, "Automatic scale detection for contour fragment based on difference of curvature," IEICE Transactions on Information and Systems, vol.E94-D, no.10, pp.1998-2005, Oct. 2011.
- (67) 安本芳雄, ジェラルドメディオニ, "B スプライン関数を用いたコーナー検出と曲線表現法," 電子情報通信学会論文誌 D, vol.J70-D, no.12, pp.2517-2524, Dec. 1987.
- (68) W. Vanzella, F. Pellegrino, and V. Torre, "Self-adaptive regularization," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol.26, no.6, pp.804-809, June 2004.
- (69) 河村圭, 石井大祐, 渡辺裕, "曲率尺度空間によるベクタ変換の評価手法に関する検討," 情報処理学会研究報告, 2008-AVM-60, pp.21-26, March 2008.
- (70) 馬場口登, 塚本正敏, 相原恒博, "手書き日本文字列からの文字切り出しの基礎的考察," 電子情報通信学会論文誌 D, vol.J68-D, no.12, pp.2123-2131, Dec. 1985.
- (71) 仲林清, 北村正, 河岡司, "あいまい用語検索を用いた高速枠なし手書き文字列読取り方式," 電子情報通信学会論文誌 D-II, vol.J74-D-II, no.11, pp.1528-1537, March 1991.
- (72) 河村圭, 石井大祐, 渡辺裕, "線画の混在する画像におけるテキスト領域抽出の改善手法," FIT2009 第8回情報科学技術フォーラム, no.H-027, Sept. 2009.
- (73) 河村圭, 石井大祐, 渡辺裕, "ベクター表現に適した小数画素精度の改良境界線検出に関する検討," 情報処理学会研究報告, 2006-AVM-54, pp.31-34, Sept. 2006.

図一覧

1-1	本論文の各章の位置づけ．括弧内は業績番号	3
2-1	従来手法における曲線最適化（自発表 [8] より引用）	8
2-2	ベジエ曲線の描画方針．左は曲線の頂点に通過点，右は曲線の両端に通過点 （自発表 [10] より引用）	8
2-3	滑らかな接続（左）と角としての接続（右）の例（自発表 [10] より引用）	9
2-4	(a) Image Decomposition Problem and (b) Image Restoration.（自発表 [22] より引用）	12
2-5	Ferguson Patch with Parameters（自発表 [26] より引用）	16
2-6	コミックのコマ配置例 [28]（著者の了解を得て抜粋）	17
2-7	コマ分割のステップ [28]（著者の了解を得て抜粋）	18
3-1	モアレ発生の模式図（自発表 [38] より引用）	22
3-2	網点分離による画像分解方式の概念図（自発表 [39] より引用）	24
3-3	網点領域における画素位置と輝度の関係（自発表 [39] より引用）	24
3-4	仮想的な黒画素の大きさと画素位置の関係（自発表 [39] より引用）	25
3-5	網点抽出の基本処理単位（自発表 [39] より引用）	26
3-6	キュービック・コンポリューションの特性（自発表 [41] より引用）	27
3-7	網点分離結果画像の一部．上から網点中心，網点画像，線画，左から提案手 法，従来手法 1，従来手法 2（自発表 [38] より引用）	28
3-8	濃度の異なる均等濃度網点における既知濃度と検出濃度の関係（自発表 [42] より引用）	30
3-9	網点角度の異なる均等濃度網点における既知濃度と検出濃度の関係（自発 表 [43] より引用）	30
3-10	線形濃度勾配網点における既知方向と検出方向の関係（自発表 [43] より引用）	31
3-11	線形濃度勾配網点における既知濃度と検出濃度の関係（自発表 [43] より引用）	31
3-12	モアレ低減の検証．左 原寸大，中央 67%提案手法，右 67%バイキュービッ ク法（自発表 [39] より引用）	31
3-13	2 値・多値統合網点中心の検出結果．左上から 2 値とグレースケールの網点 画像，2 値とグレースケールの網点中心の検出結果，グレースケール画像用 の従来手法，2 値画像用の提案手法（自発表 [45] より引用）	32

3-14	高解像度 2 値化の実験結果例．左上から原画像，入力画像，高解像度 2 値化の結果，整数精度 2 値化の結果（自発表 [46] より引用）	33
3-15	ベクタ表現による比較．左から高解像度 2 値化，整数精度 2 値化，原画像（自発表 [46] より引用）	33
4-1	Posterization Procedure（自発表 [54] より引用）	39
4-2	Anime 1 の 240 行目の画素位置と輝度値の関係（自発表 [22] より引用）	42
4-3	Anime 2 の 240 行目の画素位置と輝度値の関係（自発表 [22] より引用）	42
4-4	骨格画像の抽出結果 左から入力画像 Lenna，提案手法 1 による結果，従来手法 1 による結果（自発表 [22] より引用）	42
4-5	Anime 2 の 240 行目の画素位置と輝度値の関係．提案手法 1 と従来手法 1 によって得られた骨格成分と入力画像の比較（自発表 [22] より引用）	44
4-6	骨格画像の比較 左から提案手法 2，提案手法 3，従来手法 2（自発表 [22] より引用）	45
4-7	Anime 1 の 240 行における画素位置と輝度の一部分．提案手法 2，3，従来手法 2 により得られた骨格成分と入力画像の比較（自発表 [22] より引用）	46
4-8	平滑化画像と領域分割の結果 左上から全変動最小化による平滑化画像，均一輝度領域，角度 $0, \pi/4, \pi/2, 3\pi/4$ 方向のグラデーション領域（自発表 [54] より引用）	47
4-9	グラデーションメッシュのテスト画像 [25].	48
4-10	生成されたメッシュデータ（自発表 [26] より引用）	48
4-11	グラデーションメッシュによる近似結果 左から再現画像，メッシュを重畳，サブメッシュを重畳（自発表 [26] より引用）	49
4-12	最適化手法の比較．左上から入力画像，FR 法の結果，LM 法の結果，LM 法の結果にサブメッシュを重畳（自発表 [26] より引用）	50
5-1	曲率演算子の応答（自発表 [66] より引用）	56
5-2	曲率差 DoC の構築手順（自発表 [66] より引用）	57
5-3	曲率差 DoC の比較点（自発表 [66] より引用）	58
5-4	一方を 2 倍に拡大した図形から検出されたスケールの例（自発表 [66] より引用）	58
5-5	辺の置き換え手法	60
5-6	Example of Test Shapes; Key, Hammer, and Apple. [59]	60
5-7	再正規化後のスケール値 上から提案手法（DoC），CSS，GbC（自発表 [66] より引用）	62
5-8	輪郭素片から検出されたスケール 上から提案手法（DoC），CSS，GbC（自発表 [66] より引用）	64
5-9	特徴点が検出できる例．検出手法は左から順に DoC，CSS，GbC（自発表 [66] より引用）	65

5-10	特徴点を検出されない例．検出手法は左から順に DoC , CSS , GbC (自発表 [66] より引用)	65
5-11	特徴点の検出結果．左が DoC による提案手法で , 右が CSS による従来手法 . コンテンツは左上から apple, bat, children, crown, device0, fork, Misk, rat .	66
5-12	多角形近似と点番号の関係 (自発表 [69] より引用)	66
5-13	ドライバの曲率 (自発表 [69] より引用)	67
5-14	頂点と通過点の位置関係 (自発表 [69] より引用)	67
5-15	しきい値と頂点数の関係 (自発表 [69] より引用)	67
5-16	曲率と通過点の位置関係 (自発表 [69] より引用)	68
6-1	高精度コマ分割手法の処理フロー [28] (著者の了解を得て抜粋)	74
6-2	高速高精度コマ分割手法の処理フロー [28] (著者の了解を得て抜粋)	76
6-3	文字サイズと再現数 (自発表 [72] より引用)	78
7-1	Block Diagram of Comic Image Coding System	85
7-2	凸方向の異なる頂点の統合 (自発表 [10] より引用)	86
7-3	曲線の付加情報と通過点 (x 座標) の係数分布 (自発表 [10] より引用)	88
7-4	各種法におけるレート歪み特性 1 (自発表 [10] より引用)	88
7-5	各種法におけるレート歪み特性 2 (自発表 [10] より引用)	90
7-6	入力するコミック画像のサムネイル (No. 1, No. 2) (自発表 [73] より引用)	91
7-7	PDA によるベクタ表現コミックの提示例 (自発表 [73] より引用)	92

表一覧

4-1	Characteristics Comparison between Discrete Methods of Total Variation . . .	38
4-2	Characteristics of Target Images.	41
4-3	従来手法 1 に対する提案手法 1 の性能評価結果	43
4-4	従来手法 2 に対する提案手法 2, 3 の性能評価結果	45
4-5	Cover Ratio of Segmented Regions.	46
5-1	正規化後のスケールの標準偏差	61
5-2	Summits Ratio for each Bezier Curve	68
6-1	Recall Rate of Charactor Segmentation from Comic Images.	77
6-2	コミック画像のサイズと枚数	80
6-3	Comic A の分割精度	80
6-4	Comic B の分割精度	80
6-5	Comic C の分割精度	81
6-6	Comic D の分割精度	81
6-7	コマ分割手法の処理時間	81
6-8	高精度コマ分割手法と高速コマ分割手法の処理時間比較	81
7-1	各係数におけるモデルによる符号化効率の比較	87
7-2	File Size of Comic Image by each Coding Method [byte]	91

研究業績

査読付論文誌論文

- [1.1] 河村 圭, 加藤 晴久, 内藤 整, “CTB 構造におけるチャネル間予測符号化”, 電子情報通信学会論文誌, Vol.J95-D, No.5, pp.1286-1294, 2012 年 5 月 .
- [1.2] K. Kawamura, D. Ishii, H. Watanabe, “Automatic Scale Detection for Contour Fragment Based on Difference of Curvature,” IEICE Transactions on Information and Systems, Vol.E94-D, No.10, pp.1998-2005, Oct. 2011.
- [1.3] 河村 圭, 石井 大祐, 渡辺 裕, “全変動最小化の高速計算手法”, 電子情報通信学会論文誌, Vol.J93-D, No.3, pp.326-335, 2010 年 3 月 .
- [1.4] 石井 大祐, 河村 圭, 渡辺 裕, “(レター) コミックのコマ分割処理に関する一検討”, 電子情報通信学会論文誌, Vol.J90-D, No.7, pp.1667-1670, 2007 年 7 月 .

査読付国際会議

- [2.1] K. Kawamura, H. Kato, S. Naito, “Asymmetric Partitioning with Non-Power-of-Two Transform for Intra Coding,” 2012 Picture Coding Symposium (PCS2012), P4-9, May 2012.
- [2.2] K. Kawamura, D. Ishii, H. Watanabe, “Automatic Scale Detection Based on Difference of Curvature,” International Workshop on Advanced Image Technology (IWAIT2009), Jan. 2009.
- [2.3] Y. Yamamoto, K. Kawamura, H. Watanabe, “A Study on Spatial Scalable Coding Using Vector Representation,” IEEE International Conference on Multimedia and Expo (ICME2007), MA1-P2.4, July 2006.
- [2.4] Y. Yamamoto, K. Kawamura, H. Watanabe, “Bit Rate Reduction of Vector Representation of Binary Image,” IEEE International Conference on Image Processing (ICIP2006), WP-P3.7, Oct. 2006.
- [2.5] K. Kawamura, Y. Yamamoto, H. Watanabe, “Gradation Approximation For Vector Based Compression Of Comic Images,” IEEE International Conference on Image Processing (ICIP2005), WedPo6-6, Sept. 2005.

- [2.6] K. Kawamura, H. Watanabe, H. Tominaga, "Vector Representation of Binary Images Containing Halftone Dots," IEEE International Conference on Multimedia and Expo (ICME2004), TP2-1, June 2004.

国内研究会

- [3.1] 望月 慎太, 松田 一朗, 深井 寛修, 伊東 晋, 河村 圭, 内藤 整, "平均値座標に基づいた適応イントラ予測方式", 情処研報 2012-AVM-77, no.12, 2012年7月.
- [3.2] 河村 圭, 吉野 知伸, 内藤 整, "動き補償残差信号に対する方向性スキャンを用いた代替型1次元変換方式", 信学技報 IE2011-171, pp.49-54, 2012年3月.
- [3.3] 河村 圭, 加藤 晴久, 内藤 整, "CTB構造におけるチャンネル間予測符号化", 信学技報 IE2011-46, pp.223-228, 2011年7月.
- [3.4] 石井 大祐, 河村 圭, 帆足 啓一郎, 滝嶋 康弘, 渡辺 裕, "コミック画像におけるコマの角検出に関する一検討", 情処研報 2010-AVM-69, no.7, 2010年7月.
- [3.5] 石井 大祐, 河村 圭, 渡辺 裕, "コミック画像のコマ分割処理における制御パラメータに関する検討", 信学技報 PRMU2009-34, pp.187-192, 2009年5月.
- [3.6] 河村 圭, 石井 大祐, 渡辺 裕, "曲率尺度空間によるベクタ変換の評価手法に関する検討", 情処研報 2008-AVM-60, no.5, 2008年3月.
- [3.7] 河村 圭, 石井 大祐, 渡辺 裕, "疑似TVノルムの数値計算とその領域分割への応用に関する検討", 情処研報 2007-AVM-59, no.10, 2007年12月.
- [3.8] 河村 圭, 石井 大祐, 渡辺 裕, "人工的な多値画像のベクタ変換を考慮した領域分割に関する検討", 情処研報 2007-AVM-58, no.8, 2007年9月.
- [3.9] 河村 圭, 石井 大祐, 渡辺 裕, "ベクター表現に適した小数画素精度の改良境界線検出に関する検討", 情処研報 2007-AVM-56, no.11, 2007年3月.
- [3.10] 河村 圭, 山本 勇樹, 石井 大祐, 渡辺 裕, "ベクター表現による文字や網点を含む画像符号化システム", 情処研報 2006-AVM-54, no.6, 2006年9月.
- [3.11] 山本 勇樹, 河村 圭, 渡辺 裕, "ベクトル表現を用いた空間スケラブル符号化に関する一検討", 情処研報 2005-AVM-51, no.14, 2005年12月.
- [3.12] 河村 圭, 山本 勇樹, 渡辺 裕, "動画像圧縮におけるベクター表現の符号化に関する検討", 情処研報 2005-AVM-51, no.15, 2005年12月.
- [3.13] 山本 勇樹, 河村 圭, 渡辺 裕, 富永 英義, "ベクトル化した二値画像の画質評価に関する検討", 情処研報 2005-AVM-48, no.4, 2005年3月.

- [3.14] 河村 圭, 山本 勇樹, 渡辺 裕, “ベクター変換における階層符号化の改良”, 情処研報 2005-AVM-48, no.5, 2005 年 3 月 .
- [3.15] 河村 圭, 山本 勇樹, 渡辺 裕, “ベクター表現の階層符号化に関する検討”, 情処研報 2004-AVM-47, no.23, 2004 年 12 月 .
- [3.16] 河村 圭, 渡辺 裕, 富永 英義, “マンガ符号化における濃度勾配検出法の改良”, 情処研報 2004-AVM-44, no.3, 2004 年 3 月 .
- [3.17] 河村 圭, 渡辺 裕, 富永 英義, “マンガ符号化における網点のグラデーション処理の検討”, 情処研報 2003-AVM-43, no.8, 2003 年 12 月 .
- [3.18] 河村 圭, 渡辺 裕, 富永 英義, “マンガの超高圧縮符号化に関する検討”, 情処研報 2003-AVM-42, no.2, 2003 年 10 月 .

シンポジウム

- [4.1] 河村 圭, 吉野 知伸, 内藤 整, “Non-4:2:0 色形式対応のための HEVC 拡張の検討”, PCSJ2012 画像符号化シンポジウム資料 27th, P-1.03, 2012 年 10 月 .
- [4.2] 望月 慎太, 松田 一郎, 深井 寛修, 伊東 晋, 河村 圭, 内藤 整, “平均値座標に基づいた外挿 / 内挿適応イントラ予測方式”, PCSJ2012 画像符号化シンポジウム資料 27th, P-2.04, 2012 年 10 月 .
- [4.3] 河村 圭, 加藤 晴久, 内藤 整, “CTB 構造におけるチャンネル間予測の高速化手法”, PCSJ2011 画像符号化シンポジウム資料 26th, P-2.08, 2011 年 10 月 .
- [4.4] 河村 圭, 石井 大祐, 渡辺 裕, “領域分割に基づくベクタ表現を用いた画像符号化方式の検討”, PCSJ2009 画像符号化シンポジウム資料 24th, P-3.09, 2009 年 11 月 .
- [4.5] 石井 大祐, 河村 圭, 渡辺 裕, “アンチエイリアシングを考慮したベクタ表現による濃淡画像符号化に関する一検討”, PCSJ2008 画像符号化シンポジウム資料 23rd, P-2.04, 2008 年 11 月 .
- [4.6] 河村 圭, 石井 大祐, 渡辺 裕, “グラデーションメッシュによる画像符号化の基礎検討”, PCSJ2008 画像符号化シンポジウム資料 23rd, P-5.04, 2008 年 11 月 .
- [4.7] 石井 大祐, 河村 圭, 渡辺 裕, “動画像の時空間ベクタ表現に関する一検討”, PCSJ2007 画像符号化シンポジウム資料 22nd, P-1.03, 2007 年 11 月 .
- [4.8] 河村 圭, 石井 大祐, 渡辺 裕, “人工的な多値画像のベクター変換”, PCSJ2007 画像符号化シンポジウム資料 22nd, P-2.21, 2007 年 11 月 .
- [4.9] 河村 圭, 山本 勇樹, 石井 大祐, 渡辺 裕, “ベクター表現によるマンガ符号化システム”, PCSJ2006 画像符号化シンポジウム資料 21st, P-2.06, 2006 年 11 月 .

- [4.10] 石井 大祐, 河村 圭, 渡辺 裕, “コミックのコマ分割処理に関する一検討”, PCSJ2006 画像符号化シンポジウム資料 21st, P-2.07, 2006 年 11 月 .
- [4.11] 河村 圭, 山本 勇樹, 渡辺 裕, “イントラ符号化にベクター表現を用いた動画像圧縮に関する検討”, PCSJ2005 画像符号化シンポジウム資料 20th, P-2.15, 2005 年 11 月 .
- [4.12] 山本 勇樹, 河村 圭, 渡辺 裕, “ベクトル表現を用いた時空間スケーラブル符号化におけるエッジ再構成に関する検討”, PCSJ2005 画像符号化シンポジウム資料 20th, P-5.12, 2005 年 11 月 .
- [4.13] 山本 勇樹, 河村 圭, 渡辺 裕, 富永 英義, “二値画像のベクトル化における符号量削減に関する検討”, PCSJ2004 画像符号化シンポジウム資料 19th, P-2.04, 2004 年 11 月 .
- [4.14] 河村 圭, 山本 勇樹, 渡辺 裕, “ベクター表現における符号量制御に関する検討”, PCSJ2004 画像符号化シンポジウム資料 19th, P-5.04, 2004 年 11 月 .
- [4.15] 河村 圭, 渡辺 裕, 富永 英義, “網点を含んだ 2 値画像のベクトル表現に関する検討”, PCSJ2003 画像符号化シンポジウム資料 18th, P-2.13, 2003 年 11 月 .

国内大会

- [5.1] 河村 圭, 吉野 知伸, 内藤 整, “HEVC における非対称分割を用いたイントラ符号化方式”, 2012 映メ年次大会, 12-4, 2012 年 8 月 .
- [5.2] 浅川 和宏, 深井 寛修, 松田 一郎, 伊東 晋, 河村 圭, 内藤 整, “形状可変な連続関数群を用いた画像近似手法とその符号化への応用”, 2012 映メ年次大会, 3-6, 2012 年 8 月 .
- [5.3] 河村 圭, 加藤 晴久, 内藤 整, “CTB 構造におけるチャンネル間予測の早期打切り手法”, 2012 信学総大, D-11-84, 2012 年 3 月 .
- [5.4] 河村 圭, 加藤 晴久, 内藤 整, “CTB 構造におけるイントラ予測残差信号のチャンネル間予測符号化の評価”, FIT2011 第 10 回情報科学技術フォーラム, RI-001, 2011 年 9 月 .
- [5.5] 河村 圭, 加藤 晴久, 内藤 整, “CTB 構造におけるイントラ予測残差信号のチャンネル間予測符号化”, 2011 信学総大, DS-2-5, 2011 年 3 月 .
- [5.6] 河村 圭, 石井 大祐, 渡辺 裕, “コミック画像の文字位置特定に関する検討”, 2011 信学総大, D-12-52, 2010 年 3 月 .
- [5.7] 河村 圭, 石井 大祐, 渡辺 裕, “線画の混在する画像におけるテキスト領域抽出の改善手法”, FIT2009 第 8 回情報科学技術フォーラム, H-027, 2009 年 9 月 .

- [5.8] 河村 圭, 石井 大祐, 渡辺 裕, “グラデーションメッシュ最適化手法の検討”, 2009 信学総大, D-11-36, 2009 年 3 月 .
- [5.9] 谷 誠一, 河村 圭, 渡辺 裕, “角度と複比を用いた複数車線検出に関する一検討”, 2009 信学総大, D-11-68, 2009 年 3 月 .
- [5.10] 河村 圭, 石井 大祐, 渡辺 裕, “スケール不変な曲線特徴に関する検討”, FIT2008 第 7 回情報科学技術フォーラム, H-025, 2008 年 9 月 .
- [5.11] 谷 誠一, 河村 圭, 渡辺 裕, “複比を用いた複数車線検出に関する検討”, FIT2008 第 7 回情報科学技術フォーラム, H-040, 2008 年 9 月 .
- [5.12] 河村 圭, 石井 大祐, 渡辺 裕, “全変動ノルムの離散化手法に関する検討”, 2008 信学総大, D-11-82, 2008 年 3 月 .
- [5.13] 石井 大祐, 河村 圭, 渡辺 裕, “ベクタ表現を用いた画像のエッジ再現手法に関する一検討”, 第 70 回情処全大, 3Y-1, 2008 年 3 月 .
- [5.14] 河村 圭, 石井 大祐, 渡辺 裕, “直感的な曲線操作を実現するベクタ変換の検討”, 第 70 回情処全大, 4B-6, 2008 年 3 月 .
- [5.15] 河村 圭, 石井 大祐, 渡辺 裕, “ベクター表現に適した小数画素精度の改良境界線検出に関する検討”, FIT2007 第 6 回情報科学技術フォーラム, I-045, 2007 年 9 月 .
- [5.16] 山本 勇樹, 河村 圭, 渡辺 裕, “エッジ成分のベクトル化による空間スケラブル符号化”, 2007 信学総大, D-11-8, 2007 年 3 月 .
- [5.17] 河村 圭, 山本 勇樹, 石井 大祐, 渡辺 裕, “電子透かしを埋め込んだベクター画像の客観評価手法に関する検討”, 2007 信学総大, D-11-42, 2007 年 3 月 .
- [5.18] 高木 鉄平, 河村 圭, 渡辺 裕, 富永 英義, “複数画像によるダイナミックレンジ向上に関する検討”, 2007 信学総大, D-11-67, 2007 年 3 月 .
- [5.19] 石井 大祐, 河村 圭, 渡辺 裕, “コミック画像における高速コマ分割処理の改良に関する検討”, 2007 信学総大, D-12-158, 2007 年 3 月 .
- [5.20] 石井 大祐, 河村 圭, 渡辺 裕, “コミック画像におけるコマ分割処理の高速化に関する検討”, 第 69 回情処全大, 1P-2, 2007 年 3 月 .
- [5.21] 河村 圭, 山本 勇樹, 石井 大祐, 渡辺 裕, “ベクター表現による文字や網点を含む画像符号化システム”, 第 69 回情処全大, 3B-1, 2007 年 3 月 .
- [5.22] 高橋 良知, 河村 圭, 渡辺 裕, “複素ウェーブレット変換を用いた画像符号化に関する一検討”, FIT2006 第 5 回情報科学技術フォーラム, J-031, 2006 年 9 月 .

- [5.23] 河村 圭, 山本 勇樹, 渡辺 裕, “文字や網点を含む低解像度多値画像のベクター表現に関する検討”, FIT2006 第 5 回情報科学技術フォーラム, J-032, 2006 年 9 月 .
- [5.24] 石井 大祐, 河村 圭, 渡辺 裕, “分割線選択によるコミックのコマ分割処理に関する検討”, FIT2006 第 5 回情報科学技術フォーラム, J-033, 2006 年 9 月 .
- [5.25] 山本 勇樹, 河村 圭, 渡辺 裕, “ベクトル表現を用いた空間スケーラブル符号化”, 2006 信学総大, D-11-1, 2006 年 3 月 .
- [5.26] 早雨 美樹, 河村 圭, 永吉 功, 渡辺 裕, 富永 英義, “カラー画像の特性を考慮した超解像手法に関する検討”, 2006 信学総大, D-11-92, 2006 年 3 月 .
- [5.27] 河村 圭, 山本 勇樹, 渡辺 裕, “線画の混在する画像におけるテキスト領域の抽出に関する検討”, 2006 信学総大, D-11-99, 2006 年 3 月 .
- [5.28] 守屋 宏美, 河村 圭, 渡辺 裕, 富永 英義, “人物のふるまいに着目した動的特徴量からの人物照合に関する一検討”, 2006 信学総大, D-12-96, 2006 年 3 月 .
- [5.29] 本多 麻衣, 河村 圭, 渡辺 裕, 富永 英義, “逆転音声による音声・音楽の半開示スクランブルに関する検討”, 2006 信学総大, D-14-15, 2006 年 3 月 .
- [5.30] 河村 圭, 山本 勇樹, 渡辺 裕, “イントラ符号化にベクター表現を用いた動画像圧縮に関する検討”, 第 68 回情処全大, 2F-2, 2006 年 3 月 .
- [5.31] 山本 勇樹, 河村 圭, 渡辺 裕, “時空間スケーラビリティを考慮したセルアニメーションにおけるエッジ再構成に関する一検討”, FIT2005 第 4 回情報科学技術フォーラム, J-068, 2005 年 9 月 .
- [5.32] 河村 圭, 山本 勇樹, 渡辺 裕, “ベクター変換におけるパスの本数と点の個数に関する検討”, FIT2005 第 4 回情報科学技術フォーラム, J-075, 2005 年 9 月 .
- [5.33] 河村 圭, 山本 勇樹, 渡辺 裕, “ベクター表現における階層的変換手法の検討”, 2005 信学総大, D-11-6, 2005 年 3 月 .
- [5.34] 山本 勇樹, 河村 圭, 渡辺 裕, 富永 英義, “二値画像のベクトル表現における符号化効率改善”, 2005 信学総大, D-11-30, 2005 年 3 月 .
- [5.35] 山本 勇樹, 河村 圭, 渡辺 裕, 富永 英義, “通過点削減によるベクトル表現の符号化効率改善”, 第 67 回情処全大, 1L-5, 2005 年 3 月 .
- [5.36] 河村 圭, 山本 勇樹, 渡辺 裕, “ベクター表現の階層符号化の改良”, 第 67 回情処全大, 3W-4, 2005 年 3 月 .
- [5.37] 河村 圭, 渡辺 裕, “ベクター変換における曲線最適化アルゴリズムの一検討”, FIT2004 第 3 回情報科学技術フォーラム, J-025, 2004 年 9 月 .

- [5.38] 河村 圭, 渡辺 裕, “網点を含む 2 値画像のベクトル表現による高能率符号化に関する検討”, 画像電子学会第 32 回年次大会, 学生セッション 7, 2004 年 6 月 .
- [5.39] 河村 圭, 渡辺 裕, 富永 英義, “2 値画像のベクトル表現における網点の階調近似処理に関する検討”, 2004 信学総大, D-11-106, 2004 年 3 月 .
- [5.40] 河村 圭, 渡辺 裕, 富永 英義, “マンガ符号化における網点の多階調近似処理の評価”, 第 66 回情処全大, 3Z-1, 2004 年 3 月 .