

大学院国際情報通信研究科

Master 's Thesis  
修士論文

Title

論文題目

ベクトル表現を用いた  
空間スケーラブル符号化に関する研究  
A Study on Spatial Scalable Coding  
with Vector Representation

Student ID 学籍番号	4205A180-9
Name 氏名	山本 勇樹

Supervisor 指導教員	渡辺 裕	印
--------------------	------	---

2007年3月15日

## Summary of Master 's Thesis

# 2007年3月修了修士論文概要書

Name 氏名	山本 勇樹	ID number 学籍番号	4205A180-9
題目 Title (日本語の場合は英文題目も記入)	Supervisor 指導教員 渡辺 裕 印		
日本語 Japanese	ベクトル表現を用いた空間スケーラブル符号化に関する研究		
英語 English	A Study on Spatial Scalable Coding with Vector Representation		

### 概要 Summary

近年、電子コンテンツの普及とその閲覧環境・端末の多様化により、映像符号化のスケーラビリティに注目が集められている。映像符号化におけるスケーラビリティとは、符号化データの再生時に、任意の品質や解像度で映像を取得可能にする仕組みである。特に、マルチメディアコンテンツの効率的で柔軟な扱いや配信、流通を実現する上で、異なる表示解像度に対応可能な空間スケーラビリティ機能を有する映像符号化の要求が高まっている。

従来のスケーラブルビデオ符号化方式の大部分はラスタ表現のみを対象としており、伝送時のスケーラビリティは確保されるが、表示時までの一貫した空間スケーラビリティの確保が十分ではない。このため、ディスプレイの解像度に合わせた縮小・拡大処理を行うと、縮小処理では、線分の欠落や、線幅の相対関係の変化が起こり、拡大処理では、曲線や斜線にジャギーが発生する。

一方、ベクトル表現は、ラスタ表現でのピクセルの概念から離れ、形状の情報を点と直線・曲線により表現する形式であり、空間スケーラビリティに優れている。ラスタ表現からベクトル表現を得る手法は一般的にラスタ・ベクタ変換と呼ばれる。この変換手法は、従来、フォントの形状や地図情報を対象に研究が始まり、現在では、画像に含まれる線要素や輪郭情報を効率的に表現することができる。つまり、ベクトル表現をエッジ成分の符号化に用いることで、ディスプレイに合わせた解像度変換時に線の品質保持が可能となる。

本論文では、伝送から表示までの一貫した空間スケーラビリティの確保を目的とし、ベクトル表現の優れた空間スケーラビリティとラスタ表現の符号化特性に対する優位性を利用したエッジの分離・再構成手法を検討する。ベクトル化はエイリアシングを考慮して、エッジ成分に対して行う。輝度情報を曲線で近似する手法を提案し、通常一本の線に対して一つの輝度しか持つことができないベクトル表現を、複数の輝度情報を持つように拡張する。位置情報と関数近似した輝度値を付加情報として持つことで、輝度値の再計算が容易となり、任意の解像度変換が可能となる。また、輝度の近似曲線における係数を変化させ、表示する解像度に最適な曲線の幅と振幅を与えることで、拡大・縮小表示時のエッジの保持ならびに強調が可能となる。提案手法のスケーラビリティを利用して再構成画像を生成することで、従来のラスタ表現ベースの解像度変換手法と比較して符号化特性が優位であることを示す。

以下に本論文の構成を述べる。第1章は本論文の序論であり、研究の背景、目的について述べている。第2章は、電子コンテンツの現状と、多様化する閲覧端末について述べ、スケーラビリティ機能に着目して、ラスタ・ベクトル表現のそれぞれの符号化方式を説明する。第3章は、ベクトル表現を用いた空間スケーラブル符号化に必要な要素技術、ならびに従来手法とその問題点について述べる。第4章は、空間スケーラビリティの向上を目的として、特に、解像度変換を行う際に劣化が目立つエッジ成分をベクトル化する手法を提案する。第5章は実験を行い、評価および考察を行う。第6章は本論文の結論であり、総括及び今後の課題について論じている。

注意 1. 提出期限 / Deadline of Submission 2007年1月25日(木)17:00/ January 25, 2007 17:00

2. 原本1部、コピー1部の計2部提出 / Must submit one original and one copy

平成18年度 修士論文

ベクトル表現を用いた  
空間スケーラブル符号化に関する研究  
A Study on Spatial Scalable Coding  
with Vector Representation

指導教授 渡辺 裕 教授

2007年3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

4205A180-9

山 本 勇 樹

Yuki YAMAMOTO

## 目次

第1章	序論	1
1.1	研究の背景	1
1.2	本研究の目的	3
1.3	本論文の構成	4
第2章	電子コンテンツをとりまく環境と画像符号化	5
2.1	はじめに	5
2.2	電子コンテンツの現状	5
2.2.1	電子コンテンツの需要	5
2.2.2	漫画・アニメーション産業のデジタル化	5
2.2.3	映画産業のデジタル化	6
2.2.4	テレビ産業のデジタル化	7
2.3	閲覧端末の多様化	7
2.3.1	閲覧端末の要求条件	7
2.3.2	電子ディスプレイ方式	10
2.4	ラスター表現	13
2.4.1	符号化技術の動向	13
2.4.2	二値画像符号化方式	14
2.4.3	多値画像符号化方式	16
2.4.4	動画画像符号化方式	19
2.5	ベクトル表現	20
2.5.1	PostScript (EPS)	21
2.5.2	SVG	22
2.5.3	Flash	22
2.5.4	符号化効率の比較	23
2.6	おわりに	25
第3章	空間スケーラブル符号化実現に向けた要素技術と従来手法	26
3.1	はじめに	26
3.2	解像度変換	26
3.2.1	補間フィルタ	26
3.2.2	マルチレート信号処理	29

3.2.3	直交変換の利用	30
3.3	エッジ検出	30
3.3.1	一次微分オペレータ	31
3.3.2	二次微分 (Laplacian) オペレータ	31
3.3.3	Canny オペレータ	32
3.4	ラスタ・ベクタ変換	34
3.4.1	曲線の近似関数	34
3.4.2	多角形近似	37
3.5	画質評価手法	38
3.6	画像の分離・再構成に関する従来手法と問題点	39
3.6.1	アニメーション符号化	39
3.6.2	骨格・テキストチャ分離	40
3.6.3	SVG レンダリング	42
3.7	おわりに	44
<b>第 4 章</b>	<b>提案手法</b>	<b>45</b>
4.1	はじめに	45
4.2	エッジ領域の上書きモデル	45
4.3	輝度近似モデル	48
4.3.1	エッジの分離・再構成	48
4.3.2	エッジ成分のベクトル化	50
4.3.3	解像度変換	53
4.4	おわりに	54
<b>第 5 章</b>	<b>評価実験</b>	<b>55</b>
5.1	はじめに	55
5.2	エッジ領域上書きモデル	55
5.3	輝度近似モデル	58
5.3.1	エッジの分離・再構成	58
5.3.2	エッジ成分のベクトル化	60
5.3.3	解像度変換	61
5.3.4	自然画への適用	63
5.4	おわりに	65
<b>第 6 章</b>	<b>結論</b>	<b>67</b>
6.1	総括	67
6.2	今後の課題	68
	謝辞	<b>70</b>

参考文献	71
圖一覽	77
表一覽	79
研究業績	80

# 第1章

## 序論

### 1.1 研究の背景

近年、デジタル技術の急速な発達と、広帯域化したネットワークの普及により、画像や映像、音声などを中心としたマルチメディアコンテンツが、身近な存在となっている。また、液晶ディスプレイをはじめとする表示デバイスの高解像度化により、コンピュータや、大画面テレビ、携帯端末で電子コンテンツを閲覧する機会が増えている。マルチメディアコンテンツを視聴するためのデバイスは多様化しており、現状の表示デバイスの画面解像度は、HDTV (High Definition Television) の全画素を表示できる  $1920 \times 1080$  [画素] から、PDA (Personal Digital Assistant) の  $640 \times 480$  [画素] (VGA)、携帯電話のディスプレイに搭載される  $320 \times 240$  [画素] (QVGA) まで、数多くの種類が存在している。大画面テレビの普及もめざましく、HDTV への対応機種が増え、100 インチを超える機種も発売されている。携帯電話機についても、処理能力の飛躍的な向上、高速なデータ通信が可能な第3世代携帯電話方式の普及、大容量記憶媒体のインターフェースの搭載、ワンセグ放送への対応などがあげられる。このため、様々なリッチコンテンツを取得・閲覧する端末として注目されている。また、電子書籍専用端末も発売され、コンテンツの電子化はますます進むものと考えられている。

このように、マルチメディアコンテンツを視聴する環境が多様化したことで、現在、映像符号化のスケラビリティに注目が集められている。符号化におけるスケラビリティとは、符号化データの再生時に、任意の品質や解像度で画像を取得することを可能とする仕組みである。特に、異なる表示解像度に対応可能な画像符号化技術として、時空間スケラビリティ機能を有するコンテンツ符号化の要求が高まっている。例えば、デジタルシネマでは、符号化データの空間スケラビリティが効率的な利用、ならびに映像配信を実現する上で大きな役割を果たしている。また、マルチメディアコンテンツの柔軟な扱いや流通を促進する上で、コンテンツの2次利用に対する要求も増加すると予想される。しかし、ワンソース・マルチユースを実現し、現状の様々な表示デバイスに向けた映像を作成するためには、技術的な課題が数多く残されている。

デジタルコンテンツの効率的な伝送・蓄積を行うためには、画像の符号化技術は欠くことができない。画像の符号化方式には、大きく分けてラスタ表現とベクトル表現がある。

ラスター表現は、一般的にディスプレイの解像度に合わせた縮小・拡大処理が十分でなく、ディスプレイの性能を最大限に引き出すことができないという問題がある。従来のスケラブルビデオ符号化方式の大部分はラスター表現であり、伝送時のスケラビリティは確保されるが、表示時までの一貫した空間スケラビリティの確保が十分でない。このため、縮小処理を行うと、線分が欠落し、線幅の相対関係が失われる。また、拡大処理を行うと、曲線や斜線にジャギーが発生する。また、現在、最も普及が進んでいる DCT (Discrete Cosine Transform) ベースの JPEG は、色の変化が滑らかで低周波成分が主である人物や風景などの自然画の特徴に着目して設計されているため、自然画像とは特性が異なる版面の画像では十分な性能が得られない。JPEG が周波数変換と量子化により高周波成分を削減していること、版面の画像にはエッジと呼ばれる信号レベルが急激に変化する部分が多く存在することが主な原因である。このため、セルアニメーションや CG、テロップなどのエッジを多く含む人工的なコンテンツを符号化すると、高周波成分がカットされ、復号時にモスキートノイズが発生するという問題点が生じる<sup>(1)(2)</sup>。

一方、ベクトル表現は、空間スケラビリティに優れ、表示するサイズや解像度に依存しないため、拡大・縮小表示時の線の品質保持が可能となる。ベクトル表現とは、直線や曲線の組み合わせで描いた画像であり、点の座標とそれを結ぶ方程式のパラメータ、線の色や太さ、塗りつぶしの色などの集合として表現される。ベクトル表現は、ラスター表現でのピクセルという概念から離れ、形状の情報を保存するという考え方に基づいた形式のため、エッジの記述などに適している。また、ベクトル表現は、オブジェクト単位の構造化を行うことができるため、ベクトル表現の通過点を用いてフレーム間の相関をとることで、時間スケラビリティの確保の可能性もある。さらに、優先度をつけた伝送は、ラスター表現では、領域単位でしかできないのに対して、ベクトル表現では内部を構成する部品単位で可能である。このため、ベクトル表現は、閲覧端末が多様化した現在、画像の表現手法として有効であると考えられている。近年、メモリ量の増大により、ビットマップをコンピュータのメモリ上に全て展開することが可能となり、また、フォントデータのベクトル化も実用化されたため、ベクトル表現を処理する環境は普及してきている。携帯電話機では、QVGA 液晶パネルの搭載にあわせ、Macromedia Flash をはじめとする標準的なベクトルグラフィックスの再生環境が整いつつある。ベクトルグラフィックスをリアルタイムでラスターライズすることで、液晶パネルの解像度や階調の性能を最大限に引き出すことができる。

ラスター表現からベクトル表現を得る手法は一般的にラスター・ベクタ変換と呼ばれる。ラスター・ベクタ変換のソフトウェアとしては Potrace<sup>(3)</sup> や Autotrace<sup>(4)</sup> などがある。この変換手法は、従来、フォントの形状や地図情報を対象に研究が始まり、現在では、画像に含まれる線要素や輪郭形状の情報を効率的に表現することができる。

このように、電子コンテンツとその閲覧端末の多様化は進んでおり、コンテンツの柔軟な扱いや流通を促進する上で、空間スケラビリティの確保はますます必要とされている。特に、空間スケラビリティ機能をもつベクトル表現は、その豊富な機能性と再生環境の普及に伴い注目を集めている。ベクトル表現をエッジ成分の符号化に用いることで、ディスプレイに合わせた解像度変換の際に線の品質保持が実現できると考えられる。



## 1.2 本研究の目的

空間スケーラビリティの確保に関する研究はこれまでも数多く行われてきているが、異なる表示解像度に対応可能な画像符号化技術の発展は十分でない。従来のスケーラブルビデオ符号化は、ワンソース・マルチユースを考慮して、伝送時のスケーラビリティ確保が重視されている。しかし、表示する際には、それぞれの閲覧端末の画面に応じた画像解像度変換が必須である。例えば、ある原画像を SVC で符号化後、 $1/2$  の解像度で伝送することを考える。この画像を、表示端末にあわせて 1.2 倍に拡大する場合には、原画像を直接 0.6 倍に縮小する場合と比較して、品質を十分に維持することができないという問題がある。つまり、表示までの一貫した空間スケーラビリティの確保が十分でなく、拡大・縮小表示時に線の品質が低下する。

本論文では、伝送から表示までの一貫した空間スケーラビリティの確保による、空間スケーラビリティの向上を目的としている。符号化データのスケラビリティを利用することで、従来行われていた解像度変換法と比較して、解像度変換後の画像において符号化歪みが優位となることを目指す。特に、ベクトル表現の空間スケーラビリティに着目し、解像度変換によって特に劣化が目立つエッジの分離・再構成手法を検討する。エッジ成分を抽出し、ベクトル表現に変換する。前提とする技術としては、

- エッジ検出
- ラスタ・ベクタ変換
- 解像度変換

などがあげられる。

まず、エイリアシングを考慮したエッジ領域へのベクトル表現の適用を目指し、アニメーション画像に対して、エッジ領域の上書きモデルを提案する。その後、自然画への適用も考慮したエッジ輝度の近似によるエッジの分離・再構成手法を提案する。ベクトル表現は、通常一本の線に対して一つの輝度しか持つことができない。そのため、本研究では特に、ベクトル表現が保有する輝度情報を拡張し、複数の輝度情報を持ったベクトル化手法を提案する。つまり、ベクトル情報として、エッジの位置と輝度値をもたせることで、アンチエイリアシング処理が適用されたエッジ領域に対してベクトル化を行うことが可能となる。さらに、任意の解像度変換を可能とするため、輝度値を近似曲線で表すことで、輝度値の再計算を容易にする。

このベクトル化手法を基に、スケーラビリティの確保と符号化歪みを考慮したエッジの分離・再構成手法を提案する。対象画像をエッジ成分とその差分成分に分離し、各々に適した符号化を行うことで、解像度変換後の画質ならびに符号化効率の向上を目指す。

最後に実験を行い、提案手法の性能を評価する。提案手法のスケーラビリティを利用することにより、解像度変換後の画像が従来手法のスケーラビリティ機能を利用した場合と比較して有効であることを示す。

### 1.3 本論文の構成

以下に本章以降の構成を示す。

第2章 電子コンテンツの現状として、その需要と各産業におけるデジタル化について説明する。次に閲覧端末の多様化について述べ、人間の視覚に影響を及ぼす要素と電子コンテンツに求められる機能、ならびに電子ディスプレイの現状を示す。スケーラビリティ機能に着目して、ラスタ表現の静止画像・動画像符号化技術を説明し、最後に、ベクトル表現の形式、特性を明らかにする。

第3章 ベクトル表現を用いた空間スケーラブル符号化に必要な既存技術、ならびに従来手法とその問題点について述べる。まず、空間スケーラビリティを確保するための解像度変換、ベクトル化する対象となるエッジの抽出技術、さらにラスタ・ベクタ変換について説明する。次に、画像のベクトル化、ならびに空間スケーラビリティ確保に向けた従来研究について述べ、その問題点を検討する。

第4章 空間スケーラビリティの向上を目的として、特に、解像度変換を行う際に劣化が目立つエッジ領域をベクトル化する手法を提案し、その内容を述べる。エッジ領域のエイリアシングを考慮して、エッジ成分の分離・再構成を行い、空間スケーラビリティの確保、ならびに符号化効率の向上を目指す。まず、ベクトル表現のエッジ領域への適用を目指し、エッジ領域の上書きモデルを提案する。そして、ベクトル表現が保有する輝度情報を拡張して、エッジ領域のベクトル化手法を検討し、分離・再構成手法を提案する。

第5章 実験を行い、評価および考察を行う。まず、アニメーション画像をエッジ領域の上書きモデルに適用し、有効性を検討する。次に、輝度近似モデルにおいて、輝度値を付加情報として保有したベクトル表現を用いることにより、エイリアシングを考慮したベクトル化が可能であることを示す。提案手法のスケーラビリティを利用することにより、解像度変換後の画像が、従来手法のスケーラビリティ機能を利用した場合と比較して有効であることを示す。

第6章 本研究の総括と今後の課題について述べる。

## 第2章

# 電子コンテンツをとりまく環境と画像符号化

### 2.1 はじめに

本章では、電子コンテンツをとりまく環境として、電子コンテンツの需要と各産業におけるデジタル化について述べ、多様化する閲覧端末について説明する。そして、符号化データのスケーラビリティ機能の観点から、ラスター表現とベクトル表現のそれぞれにおける画像符号化方式について説明する。

### 2.2 電子コンテンツの現状

#### 2.2.1 電子コンテンツの需要

近年、デジタル技術の急速な発達と、広帯域化したネットワークの普及により、画像や映像、音声などを中心としたマルチメディアコンテンツが、身近な存在となっている。ブロードバンド環境の普及や、配信対象デバイスのネット家電や携帯端末への拡大が、電子コンテンツ市場の成長を牽引している。平成13年に日本政府から発表された e-Japan 計画<sup>(5)</sup>では、情報流通を促進するための基盤整備や、コンテンツ・クリエイターの育成などが打ち出され、日本発の映画やアニメーションを代表するデジタルコンテンツが、海外でも広く受け入れられるようになってきている。また、知的財産の保護と活用に関連する法律整備に関する検討も行われ、マルチメディアコンテンツの2次利用に対する要求が増大すると考えられている。ワンソース・マルチユースの実現に向けて、スケーラビリティ機能を有する符号化方式を用いて映像ソースをアーカイブし、様々な視聴環境やネットワーク環境に合わせて配信用コンテンツを生成するといったシステムの重要性が高まっている。このような背景から、ネットワークによる映像配信の本格化が見込まれ、コンテンツ産業全体を取り巻く環境のデジタル化は今後ますます進むと予想される<sup>(6)</sup>。

#### 2.2.2 漫画・アニメーション産業のデジタル化

通常、フルデジタルでアニメーションコンテンツを製作すると、30分のTV用作品では30~50Gバイト、劇場用作品では1本あたり300Gバイトの大容量ストレージが必要であるとされる。近年の大容量ハードディスクや、高速CPUの登場、高速ネットワーク環境

の確保，ストレージ環境統合の実現によって，これらのデジタル化は急速に進んでいる．現在では，企画から，制作，ビデオ編集までの全工程を一括して行えるフルデジタル・アニメーション・システムまで登場している．日本では，早くからフルデジタル・アニメーションの世界に着目している株式会社 GDH が，グローバルな作品やコンテンツへの投資，作品ライセンス管理，海外への販売等を行っており，世界的に評価が高い日本のアニメ文化を積極的に国内外で展開している<sup>(7)</sup>．

また，漫画などの人工的に作成された電子コンテンツをコンピュータや携帯端末にダウンロードして閲覧する機会もますます増加している．eBOOK Initiative Japan の“電子書籍の本屋さん「10DaysBook」”<sup>(8)</sup> や，CELSYS の無料 WEB 漫画サイト「コミックアジト」<sup>(9)</sup> など，電子書籍を販売するサービスも提供されている．このように，テレビやコンピュータ上のみならず，携帯端末上で漫画をはじめとする電子コンテンツを閲覧する環境は整いつつある．

### 2.2.3 映画産業のデジタル化

デジタルシネマは，撮影・編集・配信・上映・アーカイブという映画制作における各段階の処理を，従来のフィルムではなく，デジタルデータで行うシステムの総称である．デジタルシネマの普及により，編集作業の効率化やネットワークを利用した映画配信などが可能となる．また，配給用フィルムの焼き増しが不要になるため，経済的なメリットも期待されている．

デジタルシネマの実現に向けては，超高精細画像をターゲットとした空間スケーラビリティを有する符号化の研究や，配信フォーマットの標準化が進められ，現状では，JPEG2000 Part1 のみを採用した標準化がされている．これまで，MPEG-4 FGS (Fine Granularity Scalability) などを代表とする動画像符号化においては，変動するネットワーク帯域に対して最適な符号化レートでコンテンツの伝送を可能とする，品質に対するスケーラビリティに注目が集められていた．しかし，デジタルシネマの実現には，空間スケーラビリティが求められている．これは，デジタルシネマの再生環境として，空間解像度が約 4000 × 2000[画素] のプロジェクタを所有する劇場と，約 2000 × 1000[画素] のプロジェクタを所有する劇場が混在するためである．再生時の空間解像度が異なる 2 つの劇場に対して，効率のよい映画配信を実現するためには，空間スケーラビリティが非常に大きな役割を果たす．

また，米国映画産業の大手配給会社は，劇場，テレビ，蓄積メディア (DVD) をターゲットとして，コンテンツを 2 次利用するウィンドウ戦略を経営戦略としてもっている．このため，スケーラビリティ機能をもつデジタルシネマの標準化を積極的に推進している．今後デジタルシネマが実現することで，フィルムベースと比較して，よりスムーズかつ効率的にコンテンツの 2 次利用を行うことができると考えられる．この点においても符号化データのスケーラビリティは必要不可欠の技術であると言える．

### 2.2.4 テレビ産業のデジタル化

映画産業全体を取り巻く環境がマルチメディアコンテンツを複合的に扱う環境に変化したことで、テレビ産業との親和性がより重要になってきている。テレビ産業におけるデジタル化は、1987年から検討が開始されたATV (Advanced Television) が始まりとされている。ATVによって始められた検討は、1995年に次世代テレビ・システム委員会(ATSC: Advanced Television System Committee)によって国際標準化された。日本におけるテレビ放送のデジタル化は、1995年に行われた狭帯域CSの規格化に始まり、現在は地上波のデジタル化が進められている。アナログ放送からのスムーズな移行を実現するために、2003年から2011年までの8年間をサイマルキャスト期間としている。テレビ放送に関する制作環境のすべてがデジタル化するわけではないが、すでにMotion JPEG 2000を用いた放送コンテンツのアーカイブ化の検討も始められている。

## 2.3 閲覧端末の多様化

近年、液晶ディスプレイをはじめとする表示デバイスの高解像度化により、コンピュータや、大画面テレビ、携帯端末で電子コンテンツを閲覧する機会が増えている。携帯電話機は、処理能力の飛躍的な向上、高速なデータ通信が可能な第3世代携帯電話方式の普及、大容量記憶媒体のインターフェースの搭載、ワンセグ放送への対応などにより、様々なリッチコンテンツを取得・閲覧できる。320 × 240[画素] (QVGA) のディスプレイを搭載する携帯電話など数多くの種類が存在しており、閲覧端末として注目されている。また、SONYのLIBRIe<sup>(10)</sup> や、PanasonicのBook<sup>(11)</sup> などの電子書籍専用端末、640 × 480[画素] (VGA) を表示するPDA (Personal Digital Assistant) も発売されている。

一方、近年の大画面テレビの普及もめざましく、これまで主流であったアナログテレビ放送用のブラウン管に加え、LCD (Liquid Crystal Display) やPDP (Plasma Display Panel) が登場し、EL (Electro-luminescence) ディスプレイなどの研究開発が進められている。HDTVの全画素を表示できる1920 × 1080[画素]の解像度を持った製品が増え、100インチを超える機種も発売されている。

図2-1に、閲覧端末の大きさと表示画素数の関係、ならびにそれぞれの用途を示し、以下で、閲覧端末に求められる条件ならびに多様化するディスプレイについて説明する<sup>(12)</sup>。

### 2.3.1 閲覧端末の要求条件

閲覧端末に対する性能要求は非常に高いといわれているが、これを評価するのはエンドユーザーであり、明解に説明し得る理論はない。しかし、一部の性能に関しては、人間工学的知見を取り入れ、規格が制定されている。閲覧端末の使いやすさや見やすさ、疲れにくさを決める一つの要因は画質であり、これらはディスプレイの物理的特性と人間の視覚特性や利用特性、感性などの心理的特性で決まる。画質を決定する要素としては、画面サイズ、精細度、詳細度、色再現性、コントラスト比、動画性能、雑音などがあるが、ここでは、画面

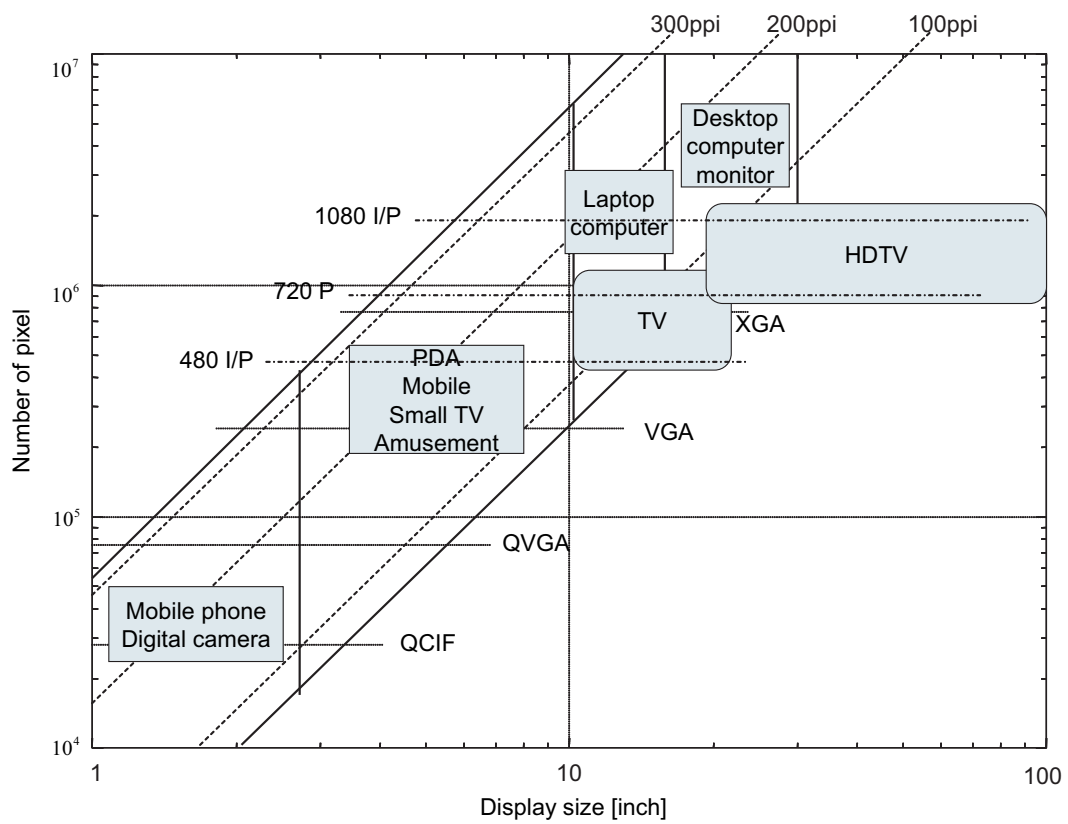


図 2-1: Applications of display (Reference [12])



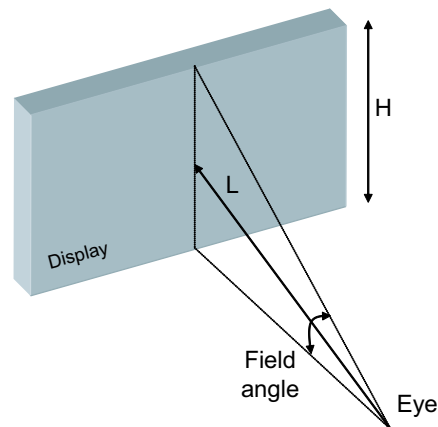


図 2-2: Field angle (Reference [12])

サイズと解像度，精細度について述べる．

### 画面サイズ

近年，映像における臨場感が重視されてきている．臨場感は，音・振動・匂い・味など人間の五感で感じられる外部刺激に対して，これらが現実のものに近い場合に特に強く人間に認知されるものである．映像の臨場感を高めるのに最も有効な方法は，画面の大型化である．画面サイズと臨場感の関係は，図 2-2 に示す画角を用いて表現される．人の目は，上下 120 度，左右 200 度程度の視野を持っている．画角は，人の目が捉えている全視野において，ディスプレイの表示画像が占める角度であり，ディスプレイの表示面の高さ  $H$  と視聴距離  $L$  で  $2\text{Arctan}(H/L)$  と表現される．画角が広くなるにつれ，人間の視覚特性上，明瞭に見えないものの物体の存在や動きが認知される領域が現れ，それに伴い臨場感が増す．テレビとして臨場感をもった映像を見る場合，画角は 30 度以上，映画やアミューズメント系コンテンツを楽しむためには，さらに大きな画角が必要とされている．このため，例えば，テレビの縦横比 9:16 のコンテンツを想定した場合，平均的な視聴距離の 2 m 前後では 50 インチ程度の画面サイズが必要となる．

### 解像度

解像度を考える際は，目の分解能が前提となる．目の空間分解能は通常角度で表され，視力 1.0 の人で 1 分である．視聴距離  $L$  で画面を測る際の空間分解能  $\Delta H$  を考えた場合， $\Delta H = 2L \tan(1^\circ/120)$  で表現される．必要な解像度は，ディスプレイの表示面の高さ  $H$  に対して， $H/\Delta H$  であり，視聴距離および画面サイズに依存する．従来の SDTV (Standard Definition TV) は，走査線数が 480 本，視野角が約 17 度程度であるが，50 インチのディスプレイを 2 m の視聴距離でみることを想定すると，臨場感を再現するには約 1100 本の走査線が必要となることがわかる．

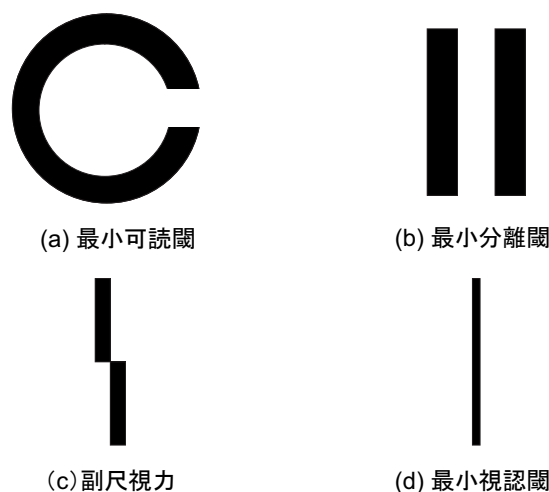


図 2-3: Eye sight (Reference [12])

### 精細度

人間の視力には、図 2-3 に示すように、最小可読閾、最小分離閾、副尺視力、最小視認閾がある。通常、視力と呼ばれるものは最小可読閾のことを表し、視力 1.0 の人は、視覚 1 度の隙間まで視認できる。これは、30cm の距離から観察した 250ppi (pixel per inch) 相当のディスプレイ 1 画素に相当する。しかし、副尺視力では視覚 2 秒、最小視認閾では、白地に黒の場合 0.5 秒、黒字に白の場合、一定以上の明るさであれば際限なく視認できる。つまり、副尺視力では 8000ppi、白地に黒の最小視認閾では 30000ppi のディスプレイ 1 画素の大きさまで認識可能であるといえる。

### 2.3.2 電子ディスプレイ方式

画像を表示する電子ディスプレイは、多様な方式が提案、実用化されており、構成材料や動作原理も多岐にわたっている。このため、利用される製品、用途のニーズに合わせて、各ディスプレイはそれぞれの特徴を活かしながら市場を拡大している。

### 液晶ディスプレイ

液晶は、薄型テレビをはじめ、携帯電話、ノート PC など幅広く利用されている。液晶には、多くの種類があり、現在主流で薄型テレビやパソコンモニタに利用されている液晶はネマティック液晶であり、駆動方式は TFT (Thin Film Transistor) を用いている。LCD は今後、さらなる大型化、高画質化が進む一方で表示デバイスの本来の枠を超えたさまざまな取り組みも始まっている。

例えば、人間の目がふたつあることを利用して、その形状や大きさをあらかず 3D ディスプレイが開発されている。また、多結晶シリコン技術を用いて、従来の液晶技術を LSI や電



子部品などの技術と融合することにより、新たな付加価値を提供するシステム液晶の開発も進んでいる。多結晶シリコン薄膜トランジスタは、ガラス基板上の低温で薄膜を形成でき、電子移動度も従来のアモルファス Si の百～数百倍と大きいことから、様々な回路や機能を表示部と一体形成できる能力を持っている。このため、すでに携帯電話やデジタルスチルカメラ、PDA などの多くのモバイル製品に搭載されている。

### プラズマディスプレイ

プラズマディスプレイは、1980 年代に、モバイル PC 用オレンジ単色パネルとして登場した。しかし、パッシブ型液晶パネルの登場により役目を終え、1990 年代のカラー PDP の登場により、表示用ディスプレイとして利用が始まった。テレビ用としては、NHK がハイビジョン用次世代大型ディスプレイとして開発を進め、長野オリンピックの 1998 年に始めて 42 型が登場した。PDP は、蛍光灯を並べたような構造の自発光型であり、その特性を活かして、高輝度、高画質、高い動画応答性能を持ち動画に強い。プラズマディスプレイは構造上、画面の小型が難しいという特徴があり、フル HD 化が現在可能な最小サイズは 40 型クラスである。

### 有機 EL

薄型・軽量かつ視野角も広く、高速応答、高コントラストなど多くの特徴をもつ有機 EL ディスプレイは、フレキシブル化することで、ファッション用途のウェアラブルディスプレイや広告媒体など、さまざまな用途がある。テレビ用途を目指した大型化が期待される中、温度変化などの過酷な状況にも強いデバイスとして、車載用ディスプレイとしても注目されている。液晶がバックライトの透過をオンオフする透過光型であるのに対して、有機 EL は素子そのものが光る自発光型である。記憶性はないが、自発光による高視認性、応答性の高さ、フレキシブルで超軽量という特徴から、大画面で書き換えの可能な電子ポスターとしての用途が期待されている。大画面の有機 EL パネルを実現するためには、印刷技術を利用できる高分子型の有機 EL 材料が必要である。しかし、実用に耐えうる寿命を備えた RGB の各色の材料は開発中であり、現状ではカラーフィルタを用いるディスプレイが試作されているのみである。それらのカラーフィルタは携帯電話機への搭載を想定して開発され、精細度が 200ppi 程度である。

### 電子ペーパー

文書データの電子化や蓄積については、大容量のツールがすでに整備されているのに比べ、それらを読む手段については、相対的に不備な現状がある。電子ペーパー技術は古くから開発が進められており、視認性という問題が解決して実用段階に入った現在、電子ブックをはじめ、値札や時計、掲示板など様々な用途で採用されはじめている。特に、携帯電話や RFID タグなどでの搭載は、主要なアプリケーションとして期待されている。電子ペーパー

の特徴の一つは、既存のディスプレイと比較して、消費電力が大幅に削減されていることである。加えて、フレキシブルディスプレイとしての活用により、デザイン性にとんだ製品設計が可能となる。2004年にはE Ink方式の電子ペーパー（解像度約170ppi、4階調グレースケール）やコレステリック液晶（解像度は180ppi、16階調グレースケール）を用いた電子的な読書用端末が実用化された。従来の携帯電話機やPCのディスプレイに比べても解像度が非常に高く、きれいな日本語表示が可能となっている。

電子ペーパーに期待される多彩な応用分野の中で、特に大きな市場が見込まれるのは、電子新聞と電子書籍である。情報の配信が通信網により迅速に、安価に、そして省資源的に行える状況下において、読みやすい電子表示媒体がないことは電子化のネックになっており、電子ペーパーはその解消の役割を担うものとして期待されている。

**高精細液晶** 高精細液晶は電子ペーパーの中で、最も完成され普及しており、インターネットなど既存のディスプレイコンテンツとの親和性も高い。特に、低温多結晶Si TFT技術を使った液晶パネルが精細度の点で最先端の表示デバイスである。2005年1月には、精細度が403ppiが、画面寸法2.0インチ、画素数はVGA（640×480画素）の液晶パネルが開発された。

**E Ink** マイクロカプセル型電気泳動方式の反射型・記憶型ディスプレイである。プラスチック基板の採用など現状では最も実用化が進んでいる電子ペーパーの1つであるが、カラー化や残像現象の解消などが課題とされている。カラー化技術としては、カラーフィルタを利用する方式、着色粒子を利用する方式、特殊な溶媒を利用してマイクロカプセルを不要にする方式の3方式が提案されている。このうちマイクロカプセルを不要にする方式は、最も安価に製造可能で柔軟性が高いため、紙のようなディスプレイに向けた有望な技術である。2004年4月にソニーが発売した読書用端末LIBRIeは、E Ink方式の電子ペーパーを1面使用しており、6インチ、SVGA（800×600画素）、4階調グレースケールで表示可能で、約170ppiの解像度を有する<sup>(10)</sup>。

**コレステリック液晶** 通常の液晶パネルに使うネマチック液晶分子にカイラル剤を加えた液晶を用い、電界によって選択的に特定光を反射させる反射型・記憶型ディスプレイである。2004年2月に松下電器産業が発売した読書用端末Bookは、コレステリック液晶を本と同じように見開きで2面配置しており、7.2インチ、XGA（1024×768画素）、16階調グレースケールで表示可能で、180ppiの解像度を有する<sup>(11)</sup>。

## LED

LEDディスプレイは通常のディスプレイと異なり、非常に大きな画面に迫力ある映像を写しだせるという特徴から、野球やサッカーなどの競技場や、ビル壁面やパチンコなどの宣伝、集客用途を中心に市場が形成されている。LEDを用いたディスプレイは、液晶のバックライトとして用いるもの、およびLEDの光そのものを用いるものがある。

### 無機 EL ディスプレイ

現在よく知られているフラットパネルディスプレイの中で最も歴史が長いのが無機 EL である。壁掛けテレビの有力候補として研究開発がなされたが、PDP や液晶、有機 EL の進展におされて開発が遅れた。しかし、高輝度青色発光という最大の懸念事項が解決されたことから、将来のディスプレイとして期待が再浮上している。無機 EL ディスプレイは、完全固体の発光型ディスプレイであり、視認性が高く、利用温度や圧力などの環境条件に対して堅牢なため、24 時間営業の店舗や、信頼性が要求される医療用、工業用、軍事用のディスプレイに用いられている。また、構造、駆動方法は単純で、高詳細化などの将来要求される仕様や価格に対しても優位であるといった特徴がある。

### FED/SED

FED/SED は、電子で蛍光体を発光させるというブラウン管と同じ原理で画像を表示し、高画質、低消費電力で、ブラウン管の真の後継フラットパネルディスプレイとして注目されている。FED (Field Emission Display) は、電解放出型ディスプレイ、SED (Surface-conduction Electron-emitter Display) は、表面伝導型電子放出素子ディスプレイと呼ばれるものである。技術的には、SED は FED の一種で、液晶テレビやプラズマテレビを上回る高画質、低消費電力が期待されている。

従来のブラウン管と比較した場合、従来のブラウン管では、発光面から十数～数十 cm 離れた位置に電子を放出する電子銃を一つ持ち偏向ヨークで画面全体を走査するのに対し、SED ではガラス基板上に画素数分の電子源を持つという違いがある。ブラウン管のように偏向が必要ないため、薄さを維持したまま大画面化が可能であり、画面の隅まで映像を映し出す事が可能となる。消費電力もブラウン管ディスプレイの半分程度で済むといった特徴がある。

## 2.4 ラスター表現

### 2.4.1 符号化技術の動向

デジタルコンテンツの効率的な伝送・蓄積を行うために、符号化技術は欠くことができない。静止画像圧縮技術としては、1999 年に二値画像を主な対象とした JBIG2、2000 年に多値画像を対象とした JPEG 2000 が、動画像圧縮技術としては、2003 年に MPEG-4 Part10 AVC (Advanced Video Coding) が、ISO/IEC JTC 1/SC 29 において国際標準化され、現在も日々進歩を遂げている。JBIG2 は、JBIG を高度化したもので、原画像に対して 1/2 の解像度変換を繰り返すことで階層的画像データを作成し、効率のよい伝送やプログレッシブ表示が可能である。JPEG 2000 は、可逆圧縮 (ロスレス) および非可逆圧縮 (ロッキー) の双方を実現する符号化方式で、現在デジタルカメラやインターネットで流通している JPEG よりも高い圧縮効率を実現すると共に、優れた品質スケーラビリティと空間スケーラビリティを実現している。また、MPEG-4 Part10 AVC は、ITU-T では H.264 として標準化されており、MPEG-2 と比較して約 2 倍の高い圧縮率を実現した高圧縮映像符号化方式

表 2-1: Characteristics of binary image coding

Coding system	Characteristic
MH	1 dimension Run Length coding
MR	2 dimension Run Length coding
MMR	2 dimension Run Length coding, Not insert EOL
JBIG	Markov-model, Arithmetic coding
JBIG2	Multi-page, Code representation

である。

符号化データのスケラビリティとは、符号化データの一部のみを復号することで、任意の品質のデータ再生や、低解像度の画像を取得することを可能とする仕組みである。マルチメディアコンテンツの柔軟な扱いや流通を促進する上で、今後の符号化方式にはこれまで通りの高い圧縮効率に加え、コンテンツの2次利用を容易に行うことを目的としたスケラビリティのような機能が要求されると予想される。事実、スケラビリティに関する研究はこれまで数多く報告されており、現在、ISO/IEC JTC 1/SC 29において、MPEG-4 Part10 AVC AMD1のSVC (Scalable Video Coding) として、具体的なアプリケーションを想定した標準化が進められている。以下に二値・多値静止画像、動画像それぞれにおける符号化方式とその機能について簡単に述べる<sup>(13, 14)</sup>。

#### 2.4.2 二値画像符号化方式

二値画像符号化方式としては、MH (Modified Huffman)、MR (Modified Read)、MMR (Modified MR) がFAXの符号化方式として利用されてきた。その後、FAXにおいては、画像通信の多様化と共に、2値画像だけでなく中間調画像も多く扱うようになり、その通信速度の高速化が望まれるようになった。通信速度を高めるためには、更に圧縮効率の高い符号化方式が望まれ、QM-Coderと呼ばれる圧縮効率の優れた符号化方式が検討され、JBIG (Joint Bi-level Image Experts Group) が標準化された<sup>(15)</sup>。JBIGは原画像に対して1/2の解像度変換作業を繰り返して階層的画像データを作成し、予測符号化するものであり、出力デバイスに応じて解像度を使い分けたり、逐次的な表示を実現するソフトコピーを可能とする。さらには、2000年5月にJBIGを高度化したJBIG2が標準化されている<sup>(16)</sup>。ここでは、それぞれの符号化方式の特徴を表2-1に示し、概要を説明する。

##### MH/MR/MMR

MHは、走査して得られる画像信号から白あるいは黒の連続する画素の集まり(ラン)を符号化対象とする一次元ランレングス符号化である。ラン長を64ごとのグループに分け、各グループを表すMake up符号 $m$ と、そのグループ内のどの長さのランかを表すTerminating

符号  $t$  を組み合わせて、ランレングス  $RL$  を  $RL = 64m + t$  と表す符号化方式である。

MR は二次元符号化であり、ライン間で白黒の境界が変化する様子を符号化する境界差分符号化を行う。符号化は、すでに符号化された一つ上の参照走査線と現在符号化しようとしている符号化走査線のデータを用いて行われる。一度符号誤りがあると、その影響が以後のラインに伝搬するため、これを防ぐために一定ラインごとに一次元符号化 (MH 符号化) を行う。

MMR は、全ラインを二次元符号化する。デジタル網に接続される G4 ファクシミリの符号化方式として規格化されたものである。誤り再送手順により、通信路誤りがないといった特徴がある。ラインの同期信号 EOL (End Of Line) を挿入しないことにより、圧縮率の向上を図っている。

## JBIG

ITU-T T.82, ISO/IEC 11544 に規定された符号化方式<sup>(15)</sup>で、以下の5つのブロックに分けられる。解像度を水平・垂直ともに  $1/2$  にする低解像度処理を行い、複数のレイヤを構築する。

1. 画像縮小 (解像度変換) (階層符号化のため)
2. 典型的予測 (TP) (近傍画素が全白のような符号化対象画素がかなり高い確率で予測できる場合に、ライン単位にそのような画素に予測はずれがあるかどうかを処理する)
3. 決定的予測 (DP) (解像度変換アルゴリズムにより、すでに符号化された縮小画像から当該画素の値が一義的に決まる。)
4. マルコフモデル符号化における参照画素モデル (モデルテンプレート)
5. 適応型算術符号化によるエントロピー符号化

ここで、シーケンシャルに符号化する場合は、第1と第3のブロックは利用されない。

送信はこの最低解像度のレイヤから開始し、順次高い解像度のレイヤに移り、原画像もしくは必要な段階の解像度までの画像を伝送する。二段目以降のレイヤの送信は、送信側で低解像度情報を共有する。差分情報のみで符号化するため、効率のよい伝送が行える。一方、受信側では、受信画像を逐次書き換えていくことで、プログレッシブ表示が実現される。オフィス文書や図面などをファクシミリ伝送するために開発された JBIG は、2値画像を対象としているため、既存の電子ペーパーや液晶パネルの階調を活用することができない。しかし、現在利用されている電子ファイリングシステムでは、JBIG のような階層的符号化はまだ行われておらず、プリント用と画面表示用の画像データを用意するか、必要に応じて原画像から生成するものがほとんどである。JBIG は原画像の  $1/2^n$  倍の解像度しか利用できないため、ユーザの要求する任意の解像度を得るためには、補間が必要となる。さらに、解像度変換 (縮小処理) により、特にコミック、印刷文書などでは網点領域でモアレとよばれるノイズが発生するという問題がある。



## JBIG2

ITU-T T. 88, ISO/IEC 14492 で規定された符号化方式<sup>(16)</sup>で、JBIGの高度化がなされている。JBIG2ではさらなる圧縮率向上を目指して、領域や辞書概念が導入され、文字以外を含む複合文書画像に対しても効率的な圧縮を行えるように、以下のような圧縮技術が採用された。

- 領域単位での、圧縮・伸張および段階的に再生画像を精緻化するプログレッシブ復号
- 文字などの小さなシンボルを辞書として圧縮するテキスト圧縮
- ハーフトーン画像をグレースケールに変換してから圧縮するハーフトーン圧縮

JBIG2はMulti-page概念を導入しており、オブジェクトとして認識することで、文字・ハーフトーンのコード表現が可能である。JBIG2は最大600dpiを想定して開発され、それ以上高い解像度の画像を想定していないため、より高解像度のデータを圧縮すると、圧縮率は低くなる。ロッキーとロスレスの二つの方式を備えている。

### 2.4.3 多値画像符号化方式

カラー静止画像の最初の国際標準符号化は、1986年にISO/IEC/JTC1/SC/WG1とITU-T SG 8 CCICの共同作業機関JPEG(Joint Photographic Experts Group)の設立後、1994年に離散コサイン変換(DCT)に基づくロッキー符号化を主要な用途とするJPEG Part 1(ISO/IEC 10918-1)として標準化された<sup>(17)</sup>。続いて、後続Partの標準化を行った後、高効率なロスレス/ニアロスレス(準可逆)符号化機能を持つJPEG-LSの標準化に取り組み、1999年にはその基本部分(ISO/IEC 14495-1)が出版された。このJPEG-LSの審議において、ロスレスに加えロッキー符号化も可能で、その圧縮性能や各種機能が従来のJPEGよりも優れる方式の存在が注目された。こうして、1997年に新たな静止画像符号化としてJPEG2000がISO/IEC JTC1/SC29/WG1において提案され、基本方式は2000年12月に標準化された<sup>(18)</sup>。ここでは、変換符号化方式として代表的なDCT変換を行うJPEGと、Wavelet変換を行うJPEG2000の特徴を表2-2にまとめる。また、アニメーション画像、自然画におけるJPEG、JPEG2000の符号化歪みを図2-4で比較した。図2-4(a)、(b)より、JPEG2000はJPEGと比較して符号化効率が良いことがわかる。また、アニメーション画像の歪み曲線は、自然画像の歪み曲線と比較して、急峻であることも確認できる。これは、アニメーション画像が、低ビットレートでは符号化ノイズによる劣化が起こりやすいことを示している。

次に、JPEG2000の機能について簡潔に述べる。

## JPEG2000

JPEG2000は2000年12月にISO/IEC JTC1/SC29/WG1において国際標準化された静止画像符号化<sup>(18)</sup>であり、低ビットレートからロスレスまでの広い範囲で高性能な段階的再

表 2-2: Characteristics of JPEG2000

	JPEG2000	JPEG
Scalability	Spatial/SNR	SNR
ROI		×
Binary Data		×
Lossless Coding		×

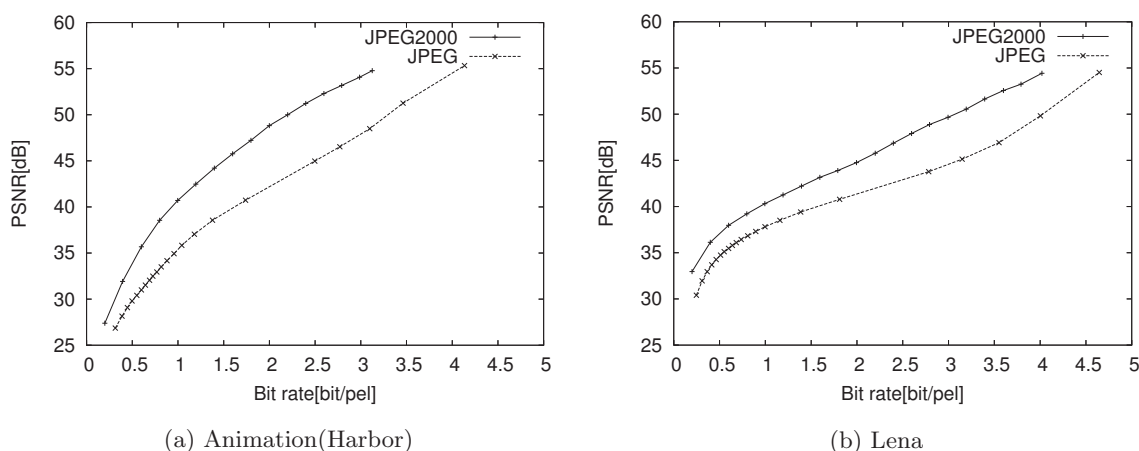


図 2-4: Rate-Distortion curve (JPEG, JPEG2000)

生が実現できる上，空間解像度・SNR (Signal to Noise Ratio) のスケーラビリティ，指定領域優先符号化 (ROI : Region Of Interest)，定レート・定サイズ性，伝送路誤り耐性の強化など，従来の JPEG 標準にはみられない多様な機能を有している．ここで，空間解像度・SNR のスケーラビリティは，サムネイルのような画像の概要から詳細な画像まで順次復号することができるため，インターネット上での画像の転送や，画像取得時の画像確認に適用できる．JPEG2000 エンコーダの基本ブロック構成を図 2-5 に示す．入力画像信号は，必要に応じ色空間座標の変換を経たのち，各色成分毎にウェーブレット変換 (DWT : Discrete Wavelet Transform) を行い，その係数を独立に符号化する．DWT は，周波数情報に加えて時間情報も解析できる変換手法として，DCT と比較して特に低ビットレートでの画質が良好であり，画質を徐々に高めるスケーラビリティ機能の実現が容易といった特徴がある．DWT には，ロスレス変換が可能で回路構成が比較的簡易な整数型フィルタと，ロスレス変換はできないが整数型より画質が良い実数型フィルタとがある．Part 1 では，その両方を規定し両者が使い分けられるようになっている．なお，整数型フィルタを使用してウェーブレット変換係数の下位  $N$  プレーンの符号化データを順に送信することで，ロッキーからロスレスへの段階的符号化も実現できる．量子化はスカラー量子化であるが，ロスレス符号化の場合には行わない．エントロピー符号化は新たに開発された EBCOT (Embedded Block

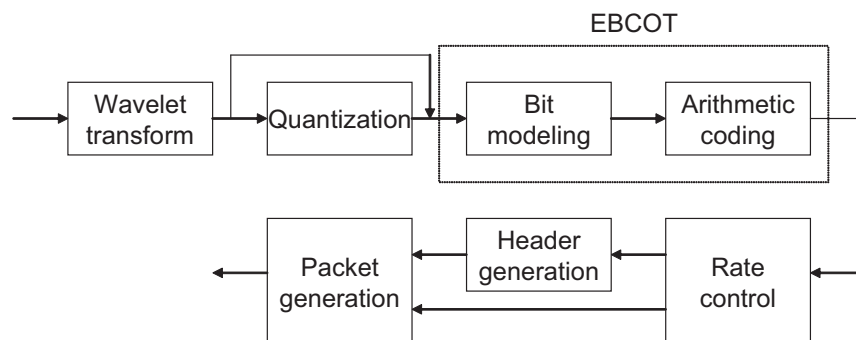


図 2-5: JPEG2000 Encoder

Coding with Optimized Truncation) 方式を用いる。

また、JPEG2000 は 2 値・多値混在データを統一的に符号化でき、ロスレス/ロッキーを同一アルゴリズムで実現できる。圧縮性能や豊富な機能を有するため、デジタルカメラ、インターネット、スキャナ、プリンタ、デジタルシネマや医療、電子商取引分野まで幅広い分野での普及が見込まれている。

具体的なアプリケーションとしては、デジタルシネマがあり、符号化データの空間スケーラビリティが大きな役割を果たしている。しかし、ワンソース・マルチユースを実現し、現状の様々な表示デバイスに向けた映像を作成するためには、技術的な課題が残されている。JPEG 2000 により符号化された映像ソースからは、デコード処理のみでは現行のテレビや様々な表示デバイスに対する映像ソースを取得することが困難であり、また、スケーラビリティにより取得可能な低品質画像における、文字の視覚的な劣化に関する問題も存在する。

これらの問題を解決するために、石川ら<sup>(19)</sup>は、JPEG 2000 の空間スケーラビリティにより取得する低域画像とデシメーションフィルタを組み合わせる映像解像度変換手法を提案し、有理数比の解像度変換を可能としている。デジタルシネマの空間解像度をソースとした場合や、現行の表示デバイスに焦点を絞り、デジタルシネマの空間解像度から、HDTV、NTSC、PAL などの空間解像度で表示するための解像度変換手法の検討を行っている。

次に、MRC (Mixed Raster Content) について説明する。

## MRC

MRC は、画像を前景、背景、マスクの 3 つのレイヤに分割して表現する符号化方式である<sup>(20)</sup>。レイヤに分かれた画像の合成方法を定義しており、各レイヤに対しての符号化は、最小符号量を得る手法を適用するのが一般的である。例えば、文字を持つ画像に対しては、前景を文字の色、背景を文字以外の自然画像、マスクをバイナリで表現が可能であり、前景と背景に対して JPEG 2000、マスクに対して JBIG/JBIG2 を適用すると最小符号量を得る<sup>(21)</sup>。

JPEG 2000 Part6 では、MRC の概念に従い、画像を複数のレイヤに分割し、それらを一



定の階調表現が可能なマスクを定義することで合成する事が出来る<sup>(22)</sup>。しかし、JBIGの最大レイヤリング数は3であるため、レイヤごとに異なる符号化方式を用いると、スケーラビリティが失われることがある。

#### 2.4.4 動画像符号化方式

ITU-T, ISO/IECなどで行われている動画像符号化の国際標準化活動は、開始から約20年たった現在も精力的に行われている。両標準化団体が共同で策定した最新の動画像符号化方式MPEG-4 AVC/H.264は、従来の符号化方式(MPEG-2など)の2倍の圧縮率を目指し規格化が進められ、2003年に国際標準として成立した。その後も、拡張方式として、高解像度画像の圧縮効率向上や主に業務用途の高品質化を実現するFRExt規格(Fidelity Range Extensions)の策定作業が進められ、2005年に規格化されている。

MPEG-4 AVC/H.264は、さまざまな映像アプリケーションに利用されるとともに、新たな映像アプリケーションの創出に寄与している。たとえば、低解像度かつ低ビットレートで配信される携帯端末向けのTV放送「ワンセグ」に代表されるサービスから、高解像度かつ高ビットレートが要求される次世代蓄積メディアサービスまで、幅広い分野で採用されている。

現在標準化団体での議論では高圧縮な符号化技術だけでなく、スケーラビリティを実現する符号化(SVC)<sup>(23)</sup>や3次元映像の符号化(MVC: Multiview Video Coding)などの高機能化に注目が集まっている。これらの符号化方式はMPEG-4 AVC/H.264の新たな拡張方式として規格化が予定され、SVCでは、時間、空間、SNRのスケーラビリティに関するさまざまな提案、議論が活発になされている。

#### SVC(Scalable Video Coding)

ISO/IECのワーキンググループであるMPEGは、2003年10月、スケーラブル映像符号化方式(SVC)の提案募集を行った。SVCは、従来、MPEG-21 Part13 SVCとして標準化が進められていたが、MPEG4 Part 10 (AVC) AMD 1に変更され、検討が進められている。SVCは、多様化するネットワーク環境や端末の能力などに応じた映像品質を、一つのストリームで提供できる符号化方式である。これにより、インターネットの回線容量に応じた映像配信や、端末の能力に応じた解像度変換、画像品質の選択が可能となる。SVCは、H.264/AVCをベースとした符号化技術でまとめつつあり、Wavelet変換をベースとしたスケーラブル符号化は今後の検討課題として位置づけられている。

このように、従来のSVCを始めとするスケーラブルビデオ符号化では、ワンソース・マルチユースを考慮して、伝送時のスケーラビリティ確保が重視されていることがわかる。しかし、表示する際には、それぞれの閲覧端末の画面に応じた画像解像度変換が必須である。例えば、ある原画像をSVCで符号化後、1/2の解像度で伝送することを考える。この画像を、表示端末にあわせて1.2倍に拡大する場合には、原画像を直接0.6倍に縮小する場合と

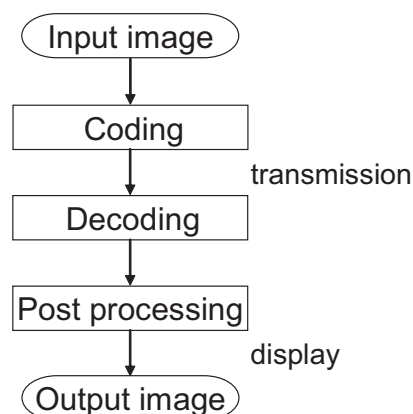


図 2-6: A flow of conventional method

比較して、品質を十分に維持することができないという問題がある。つまり、図 2-6 に示すように、表示までの一貫した空間スケーラビリティの確保が十分でなく、拡大・縮小表示時に線の品質が低下するといった問題が生じると考えられる。

## 2.5 ベクトル表現

一方、ベクトル表現は空間スケーラビリティに優れ、表示するサイズや解像度に依存しないため、拡大・縮小表示時の線の品質保持が可能となる。ベクトル表現は、直線や曲線の組み合わせで描いた画像であり、点の座標とそれを結ぶ方程式のパラメータ、線の色や太さ、塗りつぶしの色などの集合として表される。その最も基本的な構造は、パスであり、それぞれのパスが、一つもしくはいくつかの曲線や直線の線分を構成している。それぞれの線分はアンカーポイントと呼ばれる通過点を線分の両端に持つことで、画像を描くことができる。ラスター表現でのピクセルという概念から離れ、形状の情報を保存するという考え方に基づいた形式のため、エッジの記述などに適している。また、ベクトル表現は、オブジェクト単位の構造化を行うことができるため、ベクトル表現の通過点を用いてフレーム間の相関をとることで、時間スケーラビリティの確保の可能性もある。さらに、優先度をつけた伝送は、ラスター表現では、領域単位でしかできないのに対して、ベクトル表現では内部を構成する部品単位で可能である。このため、ベクトル表現は、閲覧端末が多様化した現在、画像の符号化に有効な手法のひとつであると考えられている。

ベクトル表現には、EPS(Encapsulated Post Script)<sup>(24)</sup>、SVG(Scalable Vector Graphics)<sup>(25)</sup>、Macromedia 社の Flash<sup>(26)</sup> などがある。ベクトル表現のそれぞれの符号化方式の特徴を表 2-3 に示し、以下に詳しく述べる。

表 2-3: Characteristics of vector coding

Coding system	Characteristic
EPS	3-dimension Bezier curve
SVG	2or3-dimension Bezier curve
Flash	B-spline curve

### 2.5.1 PostScript (EPS)

Postscript は Adobe Systems 社が開発したページ記述言語 (Page Description Language: PDL) であり, ページに表示されるグラフィックスやテキストを, コンピュータやプリンタが共通して理解できる標準的な形式で記述する<sup>(27)</sup>. ページを単位にするスタック型の言語であり, テキストやベクトル画像, ビットマップ画像を扱うための豊富な命令群を持つ. 文字や図形などの全ての要素のアウトライン (輪郭) を, 始点と終点, コントロールポイントの座標値から演算し, ベジエ曲線で描画する.

文字にはフォントや文字の大きさ, 字飾りなどを指定することができ, ヒント情報を用いて, サイズを問わず文字のデザインを維持し, 忠実に出力する. ポストスクリプト・ソフトウェアでは, アウトラインによる欧文フォントの技術は, Adobe Type1 フォントとして完全に公開されている. 図形は直線や円のほか, 自由曲線を表現することが可能になっている. カラーも RGB, CMYK をともに取り扱え, 現在コンピュータ上で実現可能な平面画像の構成をほぼすべて表現できる.

Level 1 と呼ばれている最初のバージョンは 1985 年に登場し, 1990 年にはカラー印刷や日本語などの 2 バイト言語に対応した Level 2 が, 1996 年にはインターネットへの対応や実装水準の段階化, PDF 形式への対応などを追加した Level 3 が発表されている. Level3 では, ページ単位のイメージをオブジェクト毎に処理できるようになり, イメージ処理の高速化, スムーズシェーディング品質の向上などが実現された.

パスは Newpath 命令により開始され, Closepath 命令によって終了する. 点の移動には Moveto 命令, 点から点への直線の描画には Lineto 命令を使う. 曲線を描く命令は中心指定と正接指定の円弧, 3 次ベジエ曲線の 3 種類がある.

### EPS

EPS は PostScript のページ記述にそった形式の一つである<sup>(24)</sup>. PostScript のデータの中から, ある図形を構成する部品を抜き出して, その部品全体の大きさなどの画像構成情報を加え, 他のアプリケーションで部品として利用できるようにカプセル化したものである. つまり, PostScript が解析できるデータであり, PostScript のプリンタドライバが作成するような PostScript ファイルと同じ扱いができる. したがって PostScript のプログラム記述に適合していれば, ビットマップデータや, フォント, ドローデータなどを混在させることが

できる。ただ、PostScript ファイル単体ではアプリケーションに貼り込むことができないため、表示用のプレビューデータ (PICT データ) や使用フォントなどのコメントを付けてカプセル化を行っている。

### 2.5.2 SVG

W3C (World Wide Web Consortium) で制定された XML 準拠の 2 次元グラフィックス記述言語であり、Web ページでを使用することを主に想定している<sup>(25)</sup>。PostScript のようにベクトル画像だけでなく、フォント情報やテキストデータも内部に保持でき、省データ性、高い操作性と拡張性を持つ。実際に SVG 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である“SVG Viewer”をインストールしておく必要がある。SVG は、ベクター形式のグラフィック、画像、テキストの 3 種類のグラフィックオブジェクトを扱う。オブジェクトをグループにまとめることや、スタイルの変更・変換、描画されたグラフィックへの合成ができる。XML のタグセットで表現されているので、データの加工性や編集性の向上、可読性の向上、汎用性や拡張性の向上、情報システムの構築コストの低減といった XML の長所を継承している。一方、SVG はプレーンテキスト形式であるため、データ格納効率が悪くデータサイズが肥大化することや、データ受信側の処理負荷が重いことなど、XML の短所も継承している。

SVG では、Moveto 命令によって現在の点が新しく設定され、Closepath 命令によってパスが終了し、現在の点から部分パスの始点へ直線が自動的に描かれる。種々の Lineto 命令によって直線を描かせることができる。曲線を描く命令は、3 次ベジエ曲線、2 次 B スプライン曲線、楕円弧曲線 (SVG Mobile では定義されない) の 3 種類がある。SVG 1.0 は 2001 年 9 月に、SVG 1.1 は 2003 年 1 月に W3C 勧告として公開された。携帯端末向けにカスタマイズされた SVG Mobile も勧告されている。

### 2.5.3 Flash

Flash は、マクロメディア社が開発しているインタラクティブなアニメーションを扱うための規格であり、アニメーションや音声などから作成されたベクトル表現の Web コンテンツである<sup>(26)</sup>。Web ブラウザで閲覧するためのプラグイン“Flash Player”は標準でインストールされているため、Flash の普及率は高い。内部では、B スプライン曲線を用いている。ベクター画像が規格の中心で、スクリプトで制御することにより、マウスの動きに合わせたアニメーションや、音を鳴らすなどのインタラクティブなサイトを作成するのに向いている。

Flash はタグ構造になっており、パスは Shape タグ内の Edge record タグに記述されている。Edge record には直線を描く Straight edge record と二次ベジエ曲線を描く Curved edge record の 2 種類がある。

表 2-4: Scanning condition

Scanner	Canon Canoscan LIDE30
Source images	Weekly magazine
Paper size	B5(4299 × 6090)
Scanning resolution	600dpi
Component/Bit depth	Gray scale/8bit
Binalize threshold	0.6 (0:white , 1:black)

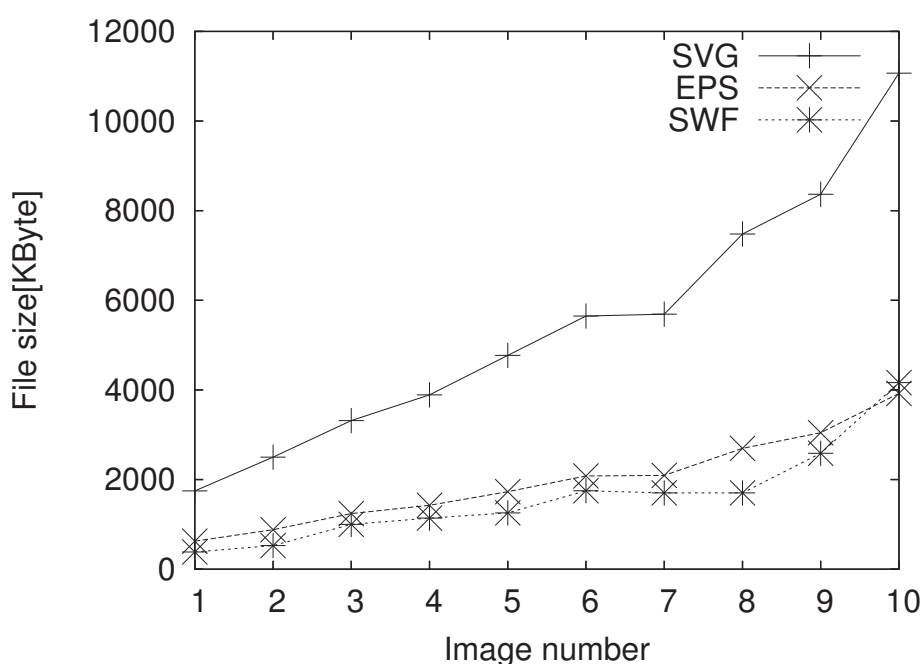


図 2-7: Coding efficiency of vector coding

#### 2.5.4 符号化効率の比較

以上述べた EPS, SVG, Flash (SWF) の3つのベクトル形式に10枚の二値画像をそれぞれ符号化した際のファイルサイズを図2-7に示す。実験は、週刊漫画雑誌を実際に10枚分スキャンし、それぞれの形式で符号化した。多値画像を2値画像に変換する際の閾値は、ファイルサイズと画質の観点から適切なものを実験により定めた。その上で、ラスター表現とベクトル表現の符号化効率について、ファイルサイズを比較した。実験条件を表2-4に示す。EPSへの変換はベクトル化ツールである Potrace<sup>(3)</sup>を用いた。また、Flashに関しては、AdobeのIllustrator<sup>(28)</sup>を用いて、Potraceで出力したEPSをFlash形式に変換した。この際、10段階中7の近似精度でEPS形式内のベジエ曲線をFlash形式内のBスプライン曲線で近似した。図2-7において、横軸に符号量が小さい順に画像を並べ、縦軸はファイルサイズとした。

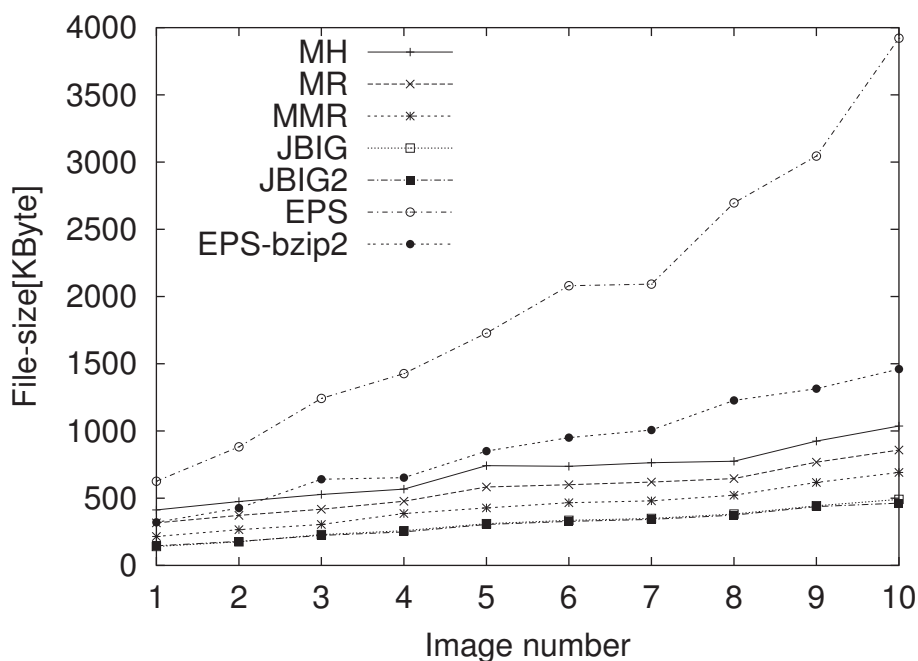


図 2-8: Coding efficiency for binary images

図 2-7より，EPS と Flash 形式はほぼ同じ符号化効率を得られたのに対し，SVG 形式は符号化効率が悪いことが明らかとなった．EPS と SVG は二値のフォーマットで記述されているのに対し，SVG はテキスト形式で記述されていることが，主な原因であると考えられる．

さらに，ベクトル表現として EPS の符号量を，ラスター表現の二値画像符号化方式である MH, MR, MMR, JBIG, JBIG2 の 5 つの符号化形式の符号量と比較した．JBIG がエントロピー符号化を行うのに対して，EPS はエントロピー符号化を行わないため，汎用の圧縮方式である bzip2<sup>(29)</sup> で圧縮した．結果を図 2-8 に示す．

スキャンした画像に 2 値化処理のみを施した状態で符号化すると，EPS に代表されるベクトル表現はラスター表現と比較して，符号化効率が悪いことがわかる．これは，スキャン時や紙の性質による輪郭線上へのノイズの混入などによる影響があり，これらの処理後の EPS の符号量は，JBIG と同等かそれ以下となることが示されている<sup>(30)</sup>．しかし，劣化をともなうという問題点がある．

以上述べた EPS に代表されるベクトル表現は，いずれも二値画像を対象としており，多値画像のベクトル表現としては，ポスタライゼーションを行うことで階調を量子化し，その輝度の境界線をベクトル化する SVG レンダリング<sup>(31)</sup> が挙げられる．これらの手法は，エイリアシング処理による輝度変化に対する考慮が十分でない．また，階調の量子化を伴うため，テクスチャなどの輝度情報が失われやすいという問題点がある．多値画像のベクトル表現の研究については，次章で詳しく説明する．

\*LZ 系のアルゴリズムと Burrows-Wheeler 変換 (ブロックソート) を併用し，その上でハフマン符号化を行う圧縮方法

## 2.6 おわりに

本章では、電子コンテンツの需要や閲覧端末の多様化などの観点から、空間スケーラビリティを持った符号化の重要性を説明した。そして、符号化データのスケラビリティ機能に着目し、ラスタ表現、ベクトル表現それぞれの符号化方式とその機能について述べた。さらに、符号化効率と機能の観点から検討を行い、ベクトル表現がエッジなどの輪郭の符号化に適切であることを述べた。



## 第3章

# 空間スケーラブル符号化実現に向けた要素技術 と従来手法

### 3.1 はじめに

本章では、ベクトル表現を用いた空間スケーラブル符号化に必要な要素技術，ならびに従来手法とその問題点について述べる．まず，空間スケーラビリティを確保するための解像度変換，ベクトル化する対象となるエッジの抽出技術，さらにラスタ・ベクタ変換について基本的な特性を説明する．また，画像の客観評価手法についてまとめる．次に，画像のベクトル化，ならびに空間スケーラビリティ確保に向けた画像の分離・再構成に関する従来研究について述べ，その問題点を検討する．

### 3.2 解像度変換

デジタル画像を扱う上で，画像もしくは映像の解像度変換は欠くことのできない技術であり，古くから研究が行われ様々な手法が存在している．デジタル信号処理における解像度変換は，レート変換によるエイリアシングやイメージングのハンドリングが重要な要素となる．特に縮小処理においては，エイリアシングの有無が出力画像の品質を大きく左右する．解像度変換の主な手法は，補間による手法，マルチレート信号処理に基づく手法，直交変換による手法の3つに大別される．それぞれの手法を以下にまとめる．

#### 3.2.1 補間フィルタ

画素で構成されるデジタル画像を縮小する場合，必ずしも入力画素に一対一対応する出力画素が存在するとは限らない．そのため，対応する画素が存在しない場合には，補間処理を行うことで出力画素を生成する．以下に示す3つの手法は，処理負荷が低いという理由から古くから用いられている補間手法である．拡大処理と縮小処理は，両者をまとめて，通常リサンプリングと呼ばれ，一般に広く利用されているリサンプリング手法として，ニアレストネイバー法，バイリニア法，バイキュービック法などがある．それぞれ以下に説明する(32)．



### ニアレストネイバー法

ニアレストネイバー法は輝度を求めるべき点の最近傍 (Nearest neighbor) の輝度値がそのまま輝度値になるという単純なアルゴリズムである。入力画像  $f(x, y)$  を  $A$  倍するとき、その出力画像  $g(x, y)$  は

$$g(x, y) = f([x/A + 1/2], [y/A + 1/2]) \quad (3.1)$$

と表せる。ここで、 $[x/A + 1/2]$  は  $x/A$  を四捨五入することを示す。

この手法は、処理が非常にコンピュータ向きであり高速に処理できる。出力される画像は比較的コントラストが高く入力画像の種類によっては良好に拡大できることもある。しかし、縮小処理においてはエイリアシングが頻繁に発生し画質は非常に悪い。

### バイリニア法

バイリニア法は一次補間の一種であり、入力画像  $f(x, y)$  を  $A$  倍するにあたって、 $(x/A, y/A)$  の四近傍を利用する。

$$\delta_x = x/A - [x/A] \quad (0 \leq \delta_x < 1) \quad (3.2)$$

$$\delta_y = y/A - [y/A] \quad (0 \leq \delta_y < 1) \quad (3.3)$$

$$f_{00} = f([x/A], [y/A]) \quad (3.4)$$

$$f_{10} = f([x/A + 1], [y/A]) \quad (3.5)$$

$$f_{01} = f([x/A], [y/A + 1]) \quad (3.6)$$

$$f_{11} = f([x/A + 1], [y/A + 1]) \quad (3.7)$$

と定義すると、

$$g(x, y) = \begin{pmatrix} f_{00} & f_{10} \\ f_{01} & f_{11} \end{pmatrix} \begin{pmatrix} 1 - \delta_x & \delta_x \\ \delta_x & 1 - \delta_x \end{pmatrix}^T \begin{pmatrix} 1 - \delta_y & \delta_y \\ \delta_y & 1 - \delta_y \end{pmatrix} \quad (3.8)$$

で表現される。この手法はニアレストネイバー法のようにエイリアシングが発生することはないが、輝度変化が直線的になりやすく、画像がぼやけるという欠点がある。

### バイキュービック法

バイキュービック法は、三次補間の一種であり、入力画像  $f(x, y)$  を  $A$  倍するにあたって、 $(x/A, y/A)$  の 16 近傍を利用する。この様子を図 3-1 に示す。

$Q$  は色を求める点であり、第一近傍点を  $P_1$  で、第二近傍点を  $P_2$  であらわしている。それぞれの近傍について、 $X$  方向、 $Y$  方向それぞれ独立に、距離  $d_x, d_y$  に対して、重み  $W_x, W_y$  を求め、各点に対する重み  $W = W_x W_y$  を得る。

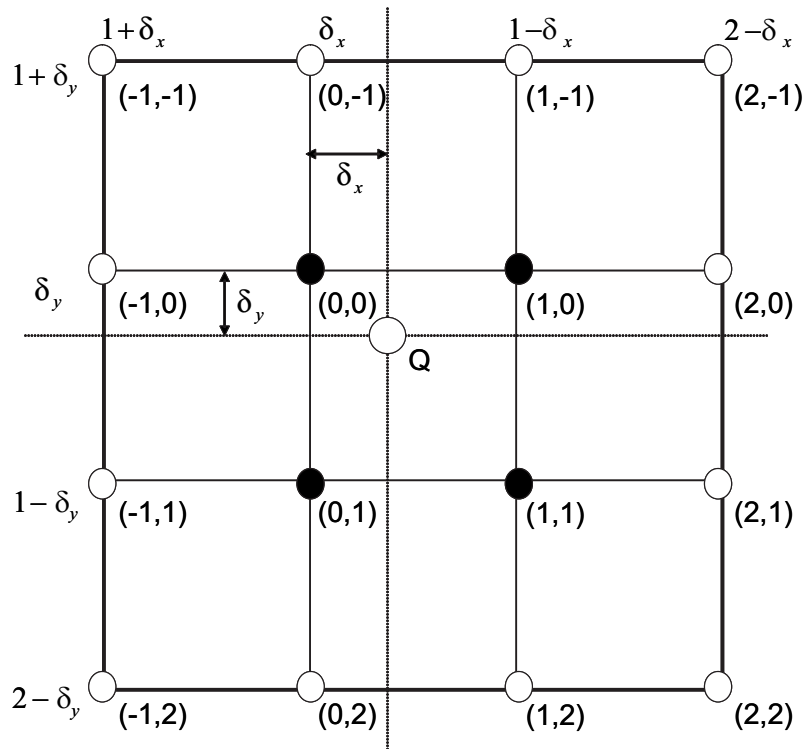


図 3-1: Bi-cubic method

$$W = \begin{cases} (d-1)(d^2-d-1) & \text{第一次近傍} \\ -(d-1)(d-2)^2 & \text{第二次近傍} \end{cases} \quad (3-9)$$

たとえば，点  $(-1,1)$ （第二次近傍）に対して， $W_x, W_y, W$  は，それぞれ，

$$W_x = \delta_x(\delta_x - 1)^2 \quad (3-10)$$

$$W_y = \delta_y(\delta_y - 1)^2 \quad (3-11)$$

$$W = \delta_x(\delta_x - 1)^2\delta_y(\delta_y - 1)^2 \quad (3-12)$$

とあらわせる．このとき， $W(x, y)$  を  $(x, y)$  の重みとすれば， $Q$  の輝度  $g(Q)$  は，

$$g(Q) = \sum_{x=-1}^2 \sum_{y=-1}^2 W(x, y)f(x, y) \quad (3-13)$$

と表せる．特徴としては，バイリニア法と比較してエイリアシングが発生しないことや，画像がぼやけないことなどがあげられる．しかし，計算量が多く，処理速度が遅い，画像に揺らぎが生じるなどの欠点がある．一般的には性能の良いリサンプリング手法の一つとされている．

これらの手法においては，バイキュービック法が，最も高画質であるが処理負荷も最も高い．逆にニアレストネイバー法は画質の面では劣るが，高速な処理が可能であるという特性

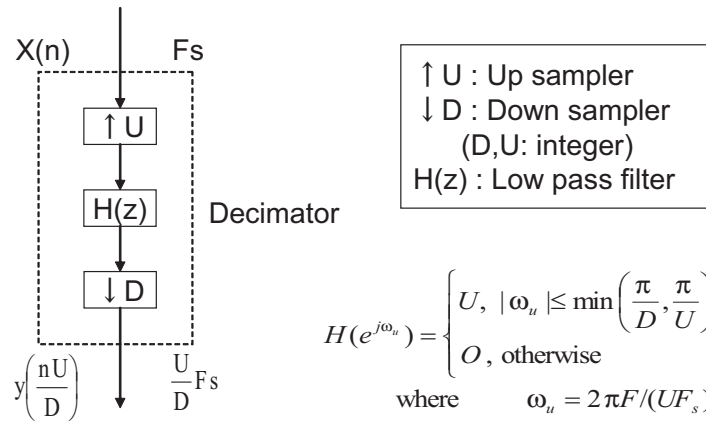


図 3-2: Decimator

がある．これらの手法は，明示的に縮小処理におけるエイリアシングを考慮していないため画質の面で不十分である場合がある．

### 3.2.2 マルチレート信号処理

エイリアシングを考慮した解像度変換は，マルチレート信号処理に基づくレート変換を行うことで可能となる (33, 34)．マルチレート信号処理における解像度変換は，図 3-2 のようなデシメータにより行われる．一般的に，デジタル画像の解像度変換は有理数比で行われるため，伝達関数が理想的な周波数特性を持つならば，有理数比  $U/D$  の理想的な解像度変換が行える．しかし，有理数比のダウンサンプリングを実現する単純なデシメータでは，カットオフ周波数  $\omega_c = \pi/D$  でフィルタリングを実行する必要があり，99/100 のように縮小率がほぼ 1 に近い場合，分母の値が非常に大きくなり演算コストが増加する．また，理論上，理想的な周波数特性を持つローパスフィルタを設計することは困難である．

有理数比のダウンサンプリングを実現する代表的なデシメーションフィルタとしては，平均画素法，Gaussian フィルタ，Lanczos フィルタなどがある．平均画素法は， $U/D$  の有理数比のダウンサンプリングを実行するために，入力信号の各値を  $U$  個だけ複製し，その後， $D$  個ごとの加算平均を取り出力信号の値を生成する．この処理過程はローパスフィルタに相当するため，デシメーションフィルタは，エイリアシングを考慮した補間フィルタとして位置づけることができる．また，Lanczos3 フィルタは，

$$Lanczos3(x) = \begin{cases} \frac{\sin(\pi x)}{\pi x} \frac{\sin(\pi x/3)}{\pi x/3}, & |x| < 3 \\ 0, & \text{otherwise} \end{cases} \quad (3-14)$$

と定義される．Lanczos フィルタのタップ数は，カットオフ周波数  $\omega_c = \pi U/D$  により決定されるため，低レートの場合にはより多くのタップ数が必要とされる．

### 3.2.3 直交変換の利用

信号処理における理想的な解像度変換は、信号の解析を周波数領域で行うことで可能となる。周波数領域においては、時間領域で無限タップのフィルタで表される理想ローパスフィルタが、式 3-15 のように表せる。

$$H(e^{j\omega}) = \begin{cases} 1, & |\omega| \leq \omega_c \\ 0, & \text{otherwise} \end{cases} \quad (3-15)$$

理想的な解像度変換を実現するためには、まず入力信号を周波数領域に変換し、理想的な帯域制限を行う。そして、解像度変換後に表現可能な周波数帯域のみを含む信号に置き換える。周波数領域への変換には、DFT を用いることでエイリアス成分を一切含まない出力画像を取得することができる。

しかし、この方法で有理数比のダウンサンプリングを実現するためには、FFT を使用する事が出来ず、演算量が膨大になる。また、カットオフ周波数付近の減衰が急激であるため、出力画像にギブス現象の影響が現れるなどの問題がある。特に画像の周囲において信号が不連続となる場合には、ギブス現象が目立つ。

この問題に対して、入力信号である画像を点对称に並べた画像に拡張する方法がある<sup>(35)</sup>。この手法では、入力信号を画像の周囲で折り返した信号と捉え、画像の周囲の不連続性を無くす。このため、ギブス現象の影響を受けることがなくなるが、最終的には DCT を用いる解像度変換に帰着することが明らかになっている。つまり、各周波数成分は実信号で表現されたため、位相成分が十分に考慮されず、複素信号で与えられる周波数解析とは異なるため、N 点周期信号に対する理想的な変換方法とならない。

## 3.3 エッジ検出

エッジは領域の境界を示す特徴であり、領域を規定する特徴値の不連続部を指す。または、信号において急激な変化がある場合に領域を分割する要素として表現される。各画素のエッジらしさを求める手法を、エッジ検出またはエッジ抽出と呼ぶ。エッジには、図 3-3 に示すように、ステップエッジ、ルーフエッジ、ラインエッジなど、いくつかのタイプが存在する。また、物理的にエッジを分類すると、シャドウエッジ、反射エッジ、方向エッジ、オクルージョンエッジ、そして specular エッジなどに分けられる<sup>(36)</sup>。

画像関数の不連続性を検出する手段であるエッジ検出は画像処理とシーン解析システムにおける基本的な操作の1つであり、特定の物体を取り出したり、複雑な画像の認識、理解のために使われる。コンピュータビジョンなどのアルゴリズムを用いる際には、エッジ検出の精度とこれらの操作を実装する効率性は重要な基準となる。エッジの検出・抽出手法は数多くの研究がなされている<sup>(37)</sup>。線形フィルタリングによる検出をはじめ、局所方位解析による手法<sup>(38)</sup>、分析的モデルあてはめによる手法<sup>(39)</sup>、局所エネルギーに基づいた手法<sup>(40)</sup>、エッジフローに基づく境界検出手法<sup>(41)</sup>などにいたるまで多岐にわたる。ここでは、最も一

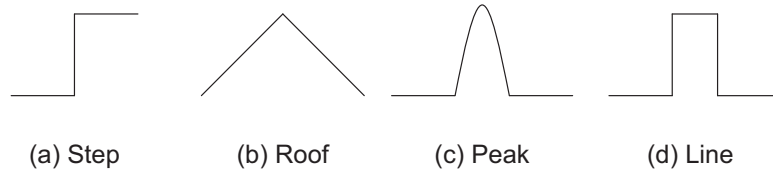


図 3-3: Edge type

般的で基礎となる微分オペレータについて述べる (42, 43) .

### 3.3.1 一次微分オペレータ

領域の境界では，画素の輝度値の変化が大きいため，画素値の変化に対して微分演算を行えば，エッジの検出を行うことができる．ここでの微分とは， $x$  成分および  $y$  成分により画像の濃淡変化を微分する空間微分のことであり，変化を検出する一般的演算子は微分オペレータと呼ばれる．

いま， $x$ - $y$  座標上での画像の濃度を  $f(x, y)$  とすると，空間一次微分 (gradient) は

$$\nabla f(x, y) = \frac{\partial f(x, y)}{\partial x} u_x + \frac{\partial f(x, y)}{\partial y} u_y \quad (3-16)$$

とベクトルであらわされる．ただし  $u_x, u_y$  はそれぞれ  $x, y$  方向の単位ベクトルである． $x$  方向の微分を  $f_x$ ， $y$  方向の微分を  $f_y$  で表すと，Gradient の強度，すなわち微分値と方向は，

$$|\nabla f(x, y)| = \sqrt{f_x^2(x, y) + f_y^2(x, y)} \quad (3-17)$$

$$\theta = \tan^{-1}\left(\frac{f_y}{f_x}\right) \quad (3-18)$$

で与えられる．

一般的にデジタル化された画像では，最小単位が 1 のため，微分演算の代わりに隣接画素間の濃淡レベルの差分演算を用いる． $f(x, y)$  における一次微分は，

$$f_x(x, y) = f(x+1, y) - f(x-1, y) \quad (3-19)$$

$$f_y(x, y) = f(x, y+1) - f(x, y-1) \quad (3-20)$$

となる．

しかし，通常差分演算を行うと画面に含まれる雑音成分に反応するため，雑音の低減とノイズの除去の両方の働きを持つ代表的な  $3 \times 3$  の一次微分オペレータとして，Prewitt operator や Sobel operator が存在する．それぞれのフィルタ係数を図 3-4，3-5 に示す．

### 3.3.2 二次微分 (Laplacian) オペレータ

また，変化率を特徴として，その不連続性を検出する空間二次微分 (Laplacian)

$$|\nabla^2 f(x, y)| = \sqrt{f_{x^2}^2(x, y) + f_{y^2}^2(x, y)} \quad (3-21)$$

-1	0	1	-1	-1	-1
-1	0	1	0	0	0
-1	0	1	1	1	1
(a) $f_x$			(b) $f_y$		

図 3-4: Prewitt operator

-1	0	1	-1	-2	-1
-2	0	2	0	0	0
-1	0	1	1	2	1
(a) $f_x$			(b) $f_y$		

図 3-5: Sobel operator

を用いる方法もある．フィルタ係数を図 3-6 に示す．一次微分オペレータと比較して，ノイズを含んだ画像や自然画像のエッジ検出に向いている．フィルタ数も 1 個で済むが，斜め方向のエッジの角度を簡単に算出することができないといった問題点がある．

### 3.3.3 Canny オペレータ

これらに対して，Canny operator<sup>(44)</sup> は，ノイズやテクスチャなどの微小な変化を検出せず，境界線のような主要なエッジだけを検出することを目的としている．ガウス関数による平滑化に一次微分オペレータを施すフィルタであり，勾配と濃度積によって抽出するため，エッジは細線化される．特徴は，

- エッジ検出率がよい．

0	1	0	1	1	1
1	-4	1	1	-8	1
0	1	0	1	1	1
(a) 4-neighbor			(b) 8-neighbor		

図 3-6: Laplacian operator

- エッジ位置を正確に検知．
- 一つのエッジに対して，線一本のつながったエッジを出力．

などがあげられる．以下に Canny operator の処理を示す<sup>(45)</sup>．

1. 画像平滑化と微分 ノイズ除去を含めたエッジ検出を行うため，ガウス関数の一次導関数により畳み込みを行う．一次元のガウス関数は，式 3-22 で表される．

$$G(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{x^2}{2\sigma^2}\right) \quad (3-22)$$

ここで  $\sigma^2$  は分散を表し，畳み込みの重み付けが決定される．また，検出するエッジの強度が調節できる．また，その一次導関数は，

$$G'(x) = \frac{\partial G(x)}{\partial x} = -\frac{x}{\sqrt{2\pi}\sigma^3} \exp\left(-\frac{x^2}{2\sigma^2}\right) \quad (3-23)$$

と表現される．

2. 勾配の強度と方向の決定 画像の勾配 (Gradient)  $\nabla f(x, y)$  の強度，すなわち微分値と方向は，x 方向の微分を  $f_x$ ，y 方向の微分を  $f_y$  で表すと，式 3-25, 3-25 同様に，

$$|\nabla f(x, y)| = \sqrt{f_x^2(x, y) + f_y^2(x, y)} \quad (3-24)$$

$$\theta = \tan^{-1}\left(\frac{f_y}{f_x}\right) \quad (3-25)$$

で与えられる．Canny operator では，エッジ検出率を高めるため，通常，エッジの方向に沿って二次元のオペレータを作る．そのため，画像の各点でオペレータを回転しながら，微分値を求める．対象画像  $f(x, y)$  と，回転角  $\theta$  のオペレータとの畳み込みは，オペレータ  $G'(x, y)$  の回転を画像の回転と等価とみなして，

$$f_{G'}(x, y) = \sum_i \sum_j f(x + \text{int}(i\cos\theta - j\sin\theta), y + \text{int}(i\sin\theta + j\cos\theta)) G'(x, y) \quad (3-26)$$

と計算できる．ここで， $\text{int}()$  は四捨五入を行う整数化関数である．角度  $\theta$  の分解能は，8 方向の探索で 22.5 度間隔となる．ガウス関数は，x 方向の関数と y 方向の関数の積に分解できるため，それぞれの一次導関数

$$\frac{\partial G(x)}{\partial x} = \frac{-x}{\sigma^2} \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{x^2}{2\sigma^2}\right) \quad (3-27)$$

$$\frac{\partial G(y)}{\partial y} = \frac{-y}{\sigma^2} \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{y^2}{2\sigma^2}\right) \quad (3-28)$$

を用いて，二次元の畳み込み積分を，x 方向の畳み込み積分とその結果に対する y 方向の畳み込み積分の二段階の一次畳み込み積分に分解することで，計算が容易となる．

式 3-26 より，その微分値が最大となるような方向を探索する．その最大方向の微分値が同じ方向の両側の隣のピクセルにおける微分値よりも大きい点をエッジ点とする．



3. 閾値処理 : エッジが存在する位置は強度  $|\nabla f(x, y)|$  が最高の点, すなわち強度の極大値, もしくは二次微分関数における 0 の点である. 通常 Canny operator では, 2 つの異なる閾値  $T_1, T_2$  ( $T_1 > T_2$ ) を設定し, 強いエッジと弱いエッジの検出を行い, エッジ画像  $f_c$  を得る.

$$f_c(x, y) = \begin{cases} 1 & \text{if } T_1 \leq |\nabla f(x, y)| \\ 1 & \text{if } T_2 \leq |\nabla f(x, y)| < T_1 \text{ and Connected with another edge} \\ 0 & \text{otherwise} \end{cases} \quad (3.29)$$

ここで, エッジ画素の輝度は  $f_c(x, y) = 1$  である. このように, 二つの閾値を使うことによって, 弱いエッジも, 強いエッジに結合している場合に出力できる.

このように, Canny operator では, 他の一次微分オペレータと異なり, ある輝度変化に対して, その変化の垂直な方向での変化のピークとなるピクセルだけを検出する. このため, ノイズの誤検出が少なく, 最も輝度に変化している階調値の尾根をつないだ一本線のエッジや, 幅 1 のピクセルのエッジなど, 真の弱いエッジを検出することができる.

K. Tombre らは, ベクトル化手法の様々な段階で考慮すべき定量的な評価方法に加え, いくつかの定性的要素を検討することで, Canny operator が, ベクトル化する際のエッジ抽出手法として, 最も良い手法の一つであることを示している<sup>(46)</sup>.

一方, カラー画像ではカラーベクトルを 1 つの画素に割り当てるため, カラー画像処理ではベクトル値画像関数を扱うことになる. カラー画像あるいは多重スペクトル画像におけるカラーエッジ検出には, 画像関数の第 2 偏微分に基づく Cumani オペレータや, ベクトル順序統計に基づくオペレータなどが存在するが, 平滑化画像関数の第一偏微分にもとづく Canny operator は, この中でも有効な手法であることが示されている<sup>(36)</sup>.

図 3-7 に, Sobel, Laplacian, Canny それぞれのオペレータによる処理画像を示す. 図 3-7 より, Canny operator を適用することにより, 一本線でつながったエッジを検出できていることが確認できた.

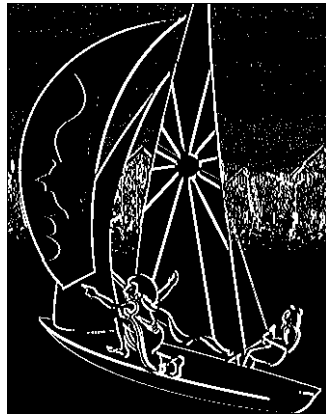
### 3.4 ラスタ・ベクタ変換

本節では, ラスタ・ベクタ変換時の機能として, まず近似関数曲線について述べ, 線画像を多角形により近似する手法について説明する. ベクトル化手法は, 従来フォントの形状や地図情報を対象として研究が進められたため, 画像に含まれる線要素及び輪郭形状の情報を効率的に符号化することができる.

#### 3.4.1 曲線の近似関数

曲線の近似には, 現在, 3 次関数のベジエ曲線と 2 次関数の B スプライン曲線が特に用いられている. 関数の次数が異なるため, ベジエ曲線を B スプライン曲線で完全に置き換え

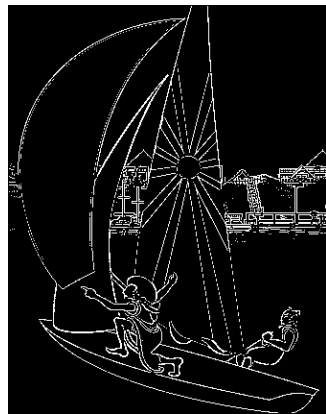




(a) Sobel operator (Animation)



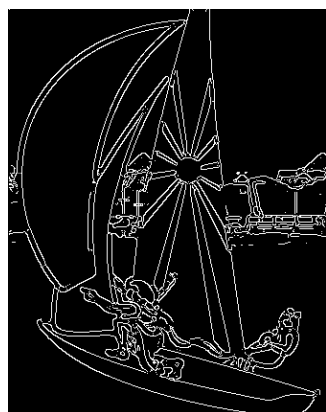
(b) Sobel operator (Lena)



(c) Laplacian operator (Animation)



(d) Laplacian operator (Lena)



(e) Canny operator (Animation)



(f) Canny operator (Lena)

図 3-7: Edge detection by differential operator

ることはできない．この2つの曲線には，共通する以下のような特徴がある．

- 凸閉包性: 曲線は制御多角形の凸包に含まれる．制御多角形の凸包とは，端点とすべての制御点を含む最小の凸領域のことである．したがって，各曲線は制御点の境界箱内に含まれる．
- 局所制御: 1つの制御点を移動すると周り数個の制御点が影響を受け，曲線が変更される．
- アフィン不変性: 各曲線に対してアフィン変換を適用した曲線は制御多角形をアフィン変換することによって得ることができる．

### B スプライン曲線

B スプライン曲線は，2つの端点（通過点）と1つの制御点によって，1つの区間曲線を二次式で定義しており，TrueType フォントの基本曲線となっている．2次関数を用いるので，計算コストは小さいが，通過点と制御点が多く必要となる．

2つの端点はそれぞれの点において滑らかに別の曲線に接続し，制御点は曲線上にないものとする．曲線を滑らかに接続するためには，2つの曲線の1次微分が端点において一致する必要がある．このため，端点における1次微分を，端点と隣接する制御点を結ぶ直線と一致させると，二次Bスプライン曲線は，端点を  $P_0(x_0, y_0)$ ,  $P_2(x_2, y_2)$ ，制御点を  $P_1(x_1, y_1)$  とすれば，媒介変数  $t$  を用いて，

$$x(t) = x_0(t-1)^2 + 2x_1t(t-1) + x_2t^2 \quad (3.30)$$

$$y(t) = y_0(t-1)^2 + 2y_1t(t-1) + y_2t^2 \quad (3.31)$$

と表現される．

### ベジエ曲線

一方，ベジエ曲線は，2つの通過点と2つの制御点によって1つの区間曲線を3次式で定義しており，PostScript の基本曲線となっている．3次関数を用いるので，計算コストがかかるが，少ない制御点で多くの表現ができる．

B スプライン曲線同様，2つの端点はそれぞれの点において滑らかに別の曲線に接続し，制御点は曲線上にないものとする．端点を  $P_0(x_0, y_0)$ ,  $P_3(x_3, y_3)$ ，制御点を  $P_1(x_1, y_1)$ ,  $P_2(x_2, y_2)$  とすれば，3次ベジエ曲線は媒介変数  $t$  を用いて，

$$x(t) = x_0(t-1)^3 + 3x_1t(t-1)^2 + 3x_2t^2(t-1) + x_3t^3 \quad (3.32)$$

$$y(t) = y_0(t-1)^3 + 3y_1t(t-1)^2 + 3y_2t^2(t-1) + y_3t^3 \quad (3.33)$$

と表現される．

### 3.4.2 多角形近似

線画の多角形近似は、パターン認識やコンピュータビジョンの領域で図形の形状記述手法として研究がされてきた。また、多角形近似を経ることなく、直接ベジエ曲線を用いてサブピクセル精度でベクトル変換を行う手法<sup>(47, 48)</sup>や、フォントが存在しないような古典的な文字を曲線近似によってベクトル変換する手法<sup>(49)</sup>なども提案されている。ここでは、古典的な手法である Ramer Algorithm, スケーラビリティを考慮した尺度空間フィルタリング, 近年議論が盛んなグラフ理論を用いた多角形近似手法について説明する。

#### Ramer Algorithm

曲線を直線で近似する古典的手法として、Ramer Algorithm がある<sup>(50)</sup>。このアルゴリズムでは、曲線上のある区間の両端  $P1, P2$  を結んだ直線から最も距離の離れた点を特徴点  $Pc$  として抽出し、端点と直線で結ぶ。曲線と直線に囲まれた領域の面積などを誤差として評価し、しきい値以下になるまでこれを繰り返す。

直線の長さを  $Length$ , 曲線関数を  $f_c(t)$ , 直線関数を  $f_s(t)$  とすれば、曲線と直線との誤差  $Error$  は、

$$Error = \int_{P1}^{P2} |f_c(t) - f_s(t)|^N dt^{1/N} \quad (3-34)$$

となる。誤差は曲線と直線のなす面積と見なすことができるので、 $N = 1$  と近似できる。しきい値を小さくすると、近似精度はあがるが、ノイズによる影響を受けやすくなり特徴点が過剰に抽出されるという特徴がある。

これに対して、許容される誤差を座標系の精度から繰り返し定義するプログレッシブ境界線符号化<sup>(51)</sup>などがあり、許容される誤差を0とすれば、ロスレスな変換が実現できる。

#### 尺度空間フィルタリング

尺度空間フィルタリングとは、ガウスフィルタの帯域幅を変化させて、ひとつの波形を様々な尺度を持った波形集合に拡張する方式である。特徴点は、それぞれの波形が極大値、極小値を示す点、すなわち2次微分がゼロになる点と定義する。

A. Pinheiro らは空間尺度フィルタリングにより特徴点を抽出し、順次符号化することで、スケーラビリティを実現できる手法を提案している<sup>(52)</sup>。大きな尺度で得られる特徴点を基本として基本階層の多角形近似を得て、さらに、小さな尺度で得られる特徴点を適宜追加することで、より精度の高い多角形近似を得ることができる。

#### グラフ理論

グラフ理論における Shortest path algorithm を利用した境界線符号化手法<sup>(53)</sup>が提案されている。Shortest path algorithm とは、多角形の頂点となりうる全ての点(ノード)同士

を接続して得られる直線と元の曲線の誤差の総和を許容誤差以下にして、エッジの本数を最小にするパスを求めるアルゴリズムである。ここで、多角形の頂点は境界線を構成する点、及びその周辺画素を候補としている。許容誤差を指標として、最小の本数の辺からなる多角形を求めることができる。A. Kolesnikovらは、頂点をあらかじめ削減しておくことで高速で準最適な多角形を最小本数の辺で得る手法を提案した<sup>(54)</sup>。

P. Selingerは、ノードを輪郭パス上に限定し、辺と輪郭パスの誤差を1/2画素以下と固定し、1/2画素以上の誤差が生じるエッジは採用しない手法を提案している<sup>(3, 55)</sup>。またこのとき、辺の数が同数のパスが複数存在するときには、累積誤差が最も小さいパスを選択する。さらに、得られた多角形の近似精度を向上させるために、多角形の頂点となった輪郭線上の点間において直線による回帰分析を行う。これにより最適な直線を取得し、直線同士の交点を新たな頂点に変更するといった特徴がある。

### 3.5 画質評価手法

近年、デジタル放送やマルチメディア配信の進展に伴い、コンテンツの圧縮符号化や伝送処理による劣化の客観評価尺度が注目を集め、様々な評価手法が提案されている<sup>(56, 57, 58)</sup>。画質の客観評価手法としては、PSNRが最も一般的で広く使われているが、実際には、完全参照(FR)型、限定的参照(RR)型、無参照(NR)型およびマルチメディア評価などの方式が存在している。PSNRはFR型である。FR方式は適用範囲が限定的ながら、ITU-T勧告J.144として標準化されている。RR法は参照画像の特徴情報を利用するため、FR法より適用範囲は広いが、NR型と同様に現在具体的方式をビデオ品質専門家グループ(VQEG: Video Quality Experts Group)で検討中である。また、ANSI客観評価パラメータとしては、最大付加動き電力、最大消失動き電力、平均動き電力差分、雑音除去平均消失動き電力、繰返しフレーム率、最大付加エッジ電力、最大消失エッジ電力、平均エッジ電力差分、付加エッジ電力周波数、最大付加空間周波数、および最大消失空間周波数などの多種多様な評価手法が存在する。

特に、エッジは画像中の識別対象に対して幾何学的な情報を与える最も知覚的に重要な特徴の一つであり、解像度変換による線の欠落やジャギーの発生によって主観品質は著しく劣化する。このため、PSNRと比較してより主観品質に近い評価手法の提案もなされている。具体的には、FR型の画質評価手法として、BT方式、延正大方式、CPqD方式、NTIA方式などが提案されており、線の品質に重みをおいている点が特徴である。また、人間の視覚特性を活用するため、視覚の空間周波数特性とマスキング効果をふまえ、DCT係数の視覚感度を統一的に表現する式も示されている。

また、低ビットレート符号化においても、PSNRは知覚画質に対する良好な測度とは言えず、画像信号の知覚歪を計算する新しい手法が提案されている<sup>(59)</sup>。この手法では、評価基底として、エッジ損傷、ブロック性、リンギングという三つの主要な知覚的攪乱アーティファクトを検出している。このように、電子コンテンツの増加、電子ディスプレイの多様化に合わせ、主観品質により近づいた客観評価手法が確立されてきているといえる。

### 3.6 画像の分離・再構成に関する従来手法と問題点

画像の復元・再構成に関する従来研究としては、一枚の画像を任意形状のオブジェクトに分割し再構成する研究が広く行われ、MPEG-4などで導入されている。ベクトル表現をこのオブジェクト符号化における輪郭線（境界線）符号化に適用するための研究も、情報量が削減される点に着目し数多く行われてきたが、主な対象はフォントや地図情報、人物の輪郭などの比較的単純な図形であった。また、空間スケーラビリティも考慮されず、複雑で入り組んだ線画が対象にされることはほとんどなかった。

一方、階層的エッジ抽出を用いたエッジ指向のプログレッシブ画像符号化機構<sup>(60)</sup>は、エッジと平滑要素の2要素モデルに基づいており、エッジ抽出に近似位置近傍探索アルゴリズムを導入し、250:1という高い圧縮率での画像復元を可能としている。また、堀内らは、対象画像をエッジ画像と低周波成分に分割し、それぞれに対してスプライン曲線で近似することでスケーラビリティの確保と高い画質維持を実現している<sup>(61)</sup>。さらに、パルスエッジとステップエッジの二つのエッジをいくつかの解像度で独立に抽出し、背景とともに符号化するMPS (Multiscale Primal Sketch) も提案されている<sup>(62)</sup>。低ビットレートでは有効であるが、高品質の画像を得るためにはテクスチャ成分を伝送する必要があり、符号量に大きく影響する。

また、エッジの復元に関する研究として、空間領域でリングングとボケを表現できる項を有するエッジモデルを導入した研究がある<sup>(63)</sup>。この手法では、原画のモデルパラメータを推定するとともに、局所的にGaussフィルタリングを行ってエッジを復元する。反復計算の必要が無いので、他の後処理法に比べて計算量が少なく済むと同時に、高いPSNRを達成できる。さらに、エッジの補間により、高解像度のディスプレイ向けに、低解像度画像を改善する手法も存在する<sup>(64)</sup>。

本節では、アニメーション符号化、骨格・テクスチャ分離、SVGレンダリングについて詳しく述べる。

#### 3.6.1 アニメーション符号化

アニメーションに特化した符号化方式として、線画と均等色領域、背景に分割して符号化する研究がある<sup>(1)</sup>。アニメーション画像は線画や境界線などの高周波成分が多く、DCTベースのMPEGで符号化すると、モスキートノイズがエッジ周辺に頻出して画質劣化を招くなどの問題点がある。これは、MPEGが、色の変化が滑らかで低周波成分が主である人物や風景などの自然画の特徴に着目して設計されているため、周波数変換と量子化により高周波成分を削減していることが原因である。このため、自然画像とは特性が異なるエッジを多く含む人工的なコンテンツをMPEGで符号化すると、高周波成分がカットされ、復号時にモスキートノイズに加え、ブロック歪みや動物体像の劣化も生じる。

これらの問題点の解決を目的として、アニメーション画像の特性に着目し、画像に含まれる線要素および輪郭形状の情報を符号化する方式を検討している。この手法では、まず、抽出した線画均等色領域に対して、線画を基準に均等色領域を補間する。次に、線要素群を線



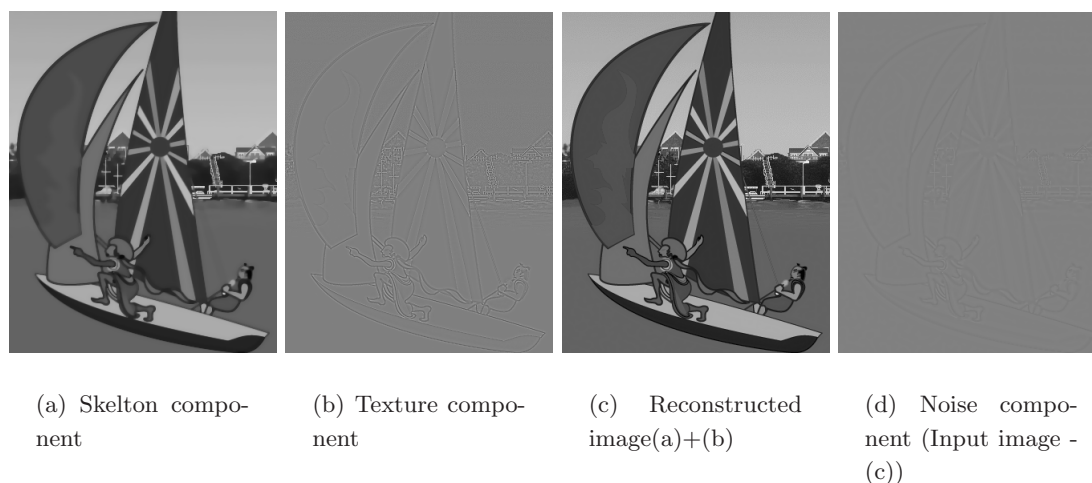


図 3-8: Skelton-texture separation for animation

長線種によって分類することで、それぞれに適した方式で線要素を符号化する。有意点のみによってアニメーション画像を符号化するために、平滑化を行いエッジを保存する。そして、平滑化によって失われた情報を補うため、抽出された均等色領域における元画像との差分情報を DCT によって符号化する。しかし、エッジ領域のエイリアシング処理による輝度変化に対する考慮が十分でなく、品質保持が十分になされないという問題点がある。

### 3.6.2 骨格・テクスチャ分離

画像の復元あるいは再構成を完全に実現するために、Vese と Osher は、雑音除去問題に関して、入力画像を骨格成分とテクスチャ成分とに分離する手法を考案した<sup>(65, 66)</sup>。この手法は、エネルギー最小化と規則化に基づくものであり、画像が骨格成分とテクスチャ成分の和で表現されると仮定している。しかし、この加法的モデルに基づく分離手法では、十分な骨格/テクスチャ分離が実現されないという問題点があった。斎藤らは、光学的画像生成過程を考慮して、観測画像が骨格画像とテクスチャ生成子との積で記述される方がより合理的であると仮定し、乗法的モデルに基づく骨格/テクスチャ分離法を検討し、分離能力が向上することを示している<sup>(67)</sup>。この研究では、分離された骨格画像とテクスチャ生成子とに対し、各々に適した補間法を適用することで、折り返し歪、リングング歪、画像ぼけ等のアーティファクトを発生させることなく、ユーザの好みに応じてエッジの鮮鋭さやテクスチャの鮮明さを制御しつつ、画像を補間拡大することが可能であることを明らかにしている。しかし、計算量、テクスチャ画像の符号化の面で数多くの課題があるといえる。図 3-8、3-9 に、分離画像を示す。

図 3-8、3-9 における処理画像は、それぞれ (a) 骨格成分、(b) テクスチャ成分、(c) 再構成画像、(d) ノイズ成分を示している。(b)、(d) はいずれも表示のために輝度 128 を処理画像に付加している。図 3-8 (c) における再構成画像の入力画像との PSNR は 44.30[dB]、図



図 3-9: Skelton-texture separation for graphics (Lena)



3-9 (c) における PSNR は 45.30[dB] であった。符号量は、テクスチャ成分に対する適切な圧縮方式が存在しないため、測定していない。図 3-8, 3-9 より、(b) のテクスチャ成分、(d) のノイズ成分の符号化を十分に考慮する必要があるといえる。

### 3.6.3 SVG レンダリング

また、ポスタリゼーションを行うことで階調を量子化し、その輝度の境界線をベクトル化する手法が挙げられる。離散サンプル点を用いて連続的輝度平面を近似するこの再構成法は、すべてのデジタル画像に適用でき、画像全体をベクトル化するため、空間スケーラビリティの面で有利であるなどの特徴がある。代表的な汎用のソフトウェアとしては、Adobe の Illustrator に搭載された Livetrace<sup>(31)</sup> があるが、一般画像から XML を用いる SVG を取得する研究も数多く行われている。SVG ファイルは、コンピュータや PDA、携帯電話などのディスプレイシステムでの利用が可能であるため、非常に有効であるといえる。

S. Battiato らは、一定の三角形分割によって生成したエッジを交換しながらデータ依存性三角形分割 (DDT:Data Dependent Triangulation)<sup>(68)</sup> を行う SVGenie を提案している<sup>(69)</sup>。通常的手法では双線形補間を用いるため、補間で発生するサンプリング格子点位置が強調され、画質の面で問題が発生するが、この手法では、三角形分割したデータのエッジ情報に基づいて輝度を決定するため、双線形補間で問題は生じない。また、ウェーブレットマルチレベル変換から収集したデータに基づいてウェーブレット利用三角形分割 (WBT: Wavelet Based Triangulation)<sup>(70)</sup> を行う SVGWave も提案されている<sup>(71)</sup>。S. Battiato らは、他にも、最新の流域分割法とアドホック発見的手法の組合せによる SWaterG を提案している<sup>(72)</sup>。このシステムは、流域分解法<sup>(73, 74)</sup> を用いて画像を一樣な領域と隣接している領域に分割し、特定のチェーンコードを構築することによって、SVG 表現を取得する。

また、ラスタ・ベクタ変換の無料の汎用ソフトウェアとしては、Potrace や Autorace<sup>(4)</sup>、KVec<sup>(75)</sup>、Ras2Vec<sup>(76)</sup>、Vector Eye<sup>(77)</sup> などがある。

これらの手法により、デジタル写真の高画質レンダリングを行うことが可能であるが、階調の量子化を伴うため、テクスチャなどの輝度情報が失われやすいという問題点がある。また、SVG の符号量が大きくなるといった問題点も指摘されている。

そこで、Vector Eye を用いて、実際に符号化ひずみを測定した。実験はアニメーション画像と自然画像の二枚の画像を用い、階調の量子化、曲線近似精度、輪郭精度のパラメータを変化させた。表現されるベクトルの本数と符号量の関係を表 3-1 に、R-D 曲線による評価を図 3-10 に、実際の処理画像を図 3-11 に示す。

実験より、表現されるベクトルの本数と符号量は、ほぼ比例することが確認できた。また、符号量は大きく、画質が十分でないこと、処理時間にも問題があることが明らかとなった。

#### Potrace

P. Selinger が提案した、グラフ理論を基に多角形近似を経て曲線で近似する手法<sup>(55)</sup> を基に作られている。GNU GPL(General Public License) の元で配布されており、ソースコー

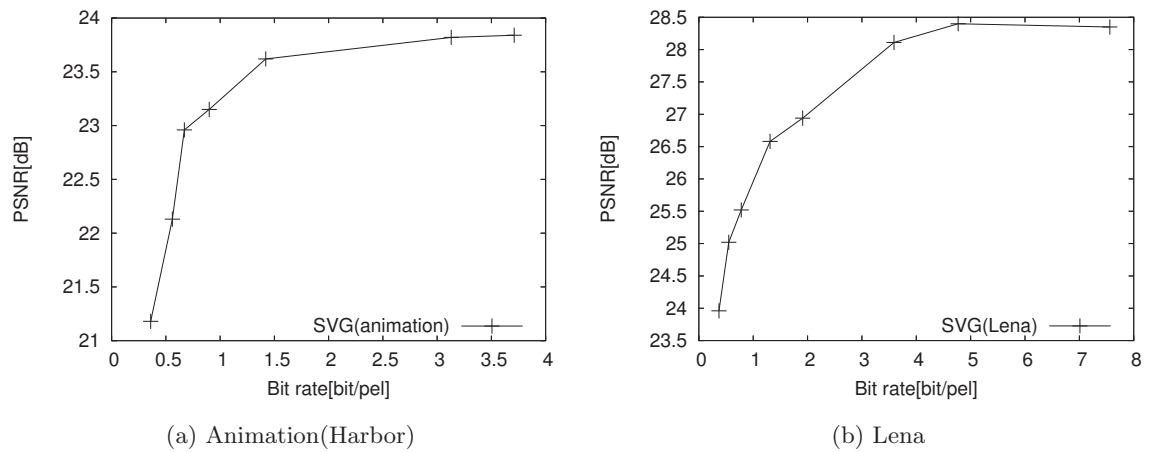


図 3-10: Rate-Distortion curve of SVG Rendering

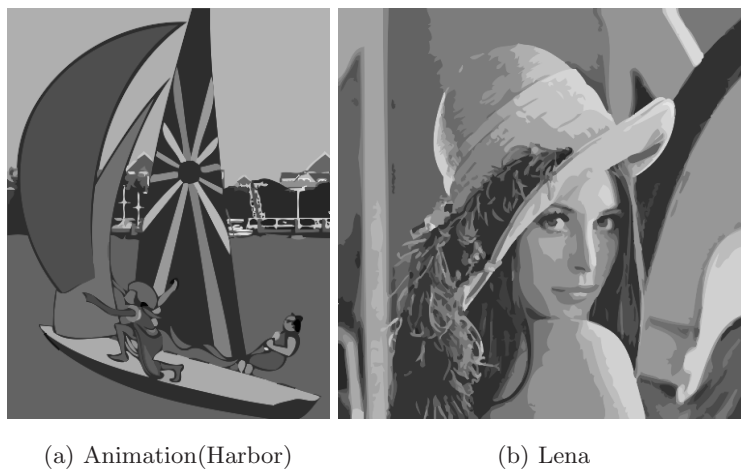


図 3-11: Modified images by SVG Rendering

表 3-1: Evaluation of SVG Rendering (animation)

Operation	Number of vectors	Bitrate[bit/pel]	PSNR[dB]
Detail	717	3.71	23.84
Normal	180	0.90	23.15
Simplified	137	0.67	22.96
More simplified	81	0.36	21.18

ドが自由に入手できる。デフォルト出力形式はEPS形式である。既存の画像入出力ライブラリとリンクすることで、数多くのフォーマットに対応しているが、Flash形式への対応は十分でない。内部では、ベジエ曲線が用いられている。入力パラメータの1つに座標解像度があり、近似の精度を変えることができる。デフォルトの座標解像度は10倍の精度があるが、これを変更することで、制御点を変えることができる。

### 3.7 おわりに

本章では、まず、本提案手法で必要となる既存技術として、解像度変換、エッジ検出、ラスタ・ベクタ変換について説明した。解像度変換としては、補間による手法、マルチレート信号処理に基づく手法、直交変換による3つの手法について、その特徴を述べた。エッジ検出手法は、一般的な微分をつかったオペレータについて述べ、Cannyオペレータがベクトル化を行う際のエッジ検出に有効であることを明らかにした。次に、ラスタ・ベクタ変換の基礎として、近似関数曲線について述べ、線画像を多角形により近似する手法について説明した。また、デジタル放送やマルチメディア配信の進展に伴い、コンテンツの圧縮符号化などの客観評価尺度も注目を集めており、PSNRと比較してより主観品質に近い評価手法が提案されていることを述べた。従来画像の分離・再構成手法として、アニメーション符号化、骨格・テクスチャ分離、SVGレンダリングの3手法について述べた。そして、符号化歪み特性が悪いなどの問題点をあげた。

## 第4章

### 提案手法

#### 4.1 はじめに

本章では、空間スケーラビリティの向上を目的として、特に、解像度変換を行う際に劣化が目立つエッジ領域をベクトル化する2つの手法を提案する。エッジ領域のエイリアシングを考慮して、抽出したエッジ領域をベクトル表現に変換することで、空間スケーラビリティの向上ならびに、符号化効率の向上を目指す。

- エッジ領域の上書きモデル
- 輝度近似モデル

それぞれの手法について、その詳細を以下に示す。

#### 4.2 エッジ領域の上書きモデル

本手法は、主にアニメーション画像への適用を想定している。アニメーション画像に代表される人工的なコンテンツは線画や境界線などの高周波成分が多く、自然画像とは特性が異なる。従来のDCTベースのMPEGは、人物や風景などの、色の変化が滑らかで主に低周波成分から成る自然画の特徴に着目して設計され、周波数変換と量子化により高周波成分を削減する。このため、アニメーション画像をDCTベースの符号化方式で圧縮すると、モスキートノイズがエッジ周辺に頻出したり、ブロック歪みや動物体像の劣化を招くなどの問題点がある。また、解像度変換を行うと、拡大処理によりジャギーが、縮小処理により線の欠落が生じる。これを解決するための手法として、アニメーション画像を、線画と均等色領域、背景に分割し、線画に対してベクトル表現を適用するアニメーション符号化<sup>(1)</sup>があることを前章で述べた。しかし、この手法は、エッジ領域のエイリアシング処理による輝度変化に対する考慮が十分でなく、品質保持が十分になされないという問題点があった。

本手法では、エイリアシングを考慮しつつ、エッジ領域をベクトル表現に変換することで、空間スケーラビリティの確保、ならびにエッジ周辺のモスキートノイズをおさえることを目指す。エッジ成分をベクトル表現で圧縮し、DCT変換とベクトル表現によるハイブリッドな符号化方式を提案する。処理フローを図4-1に示す。

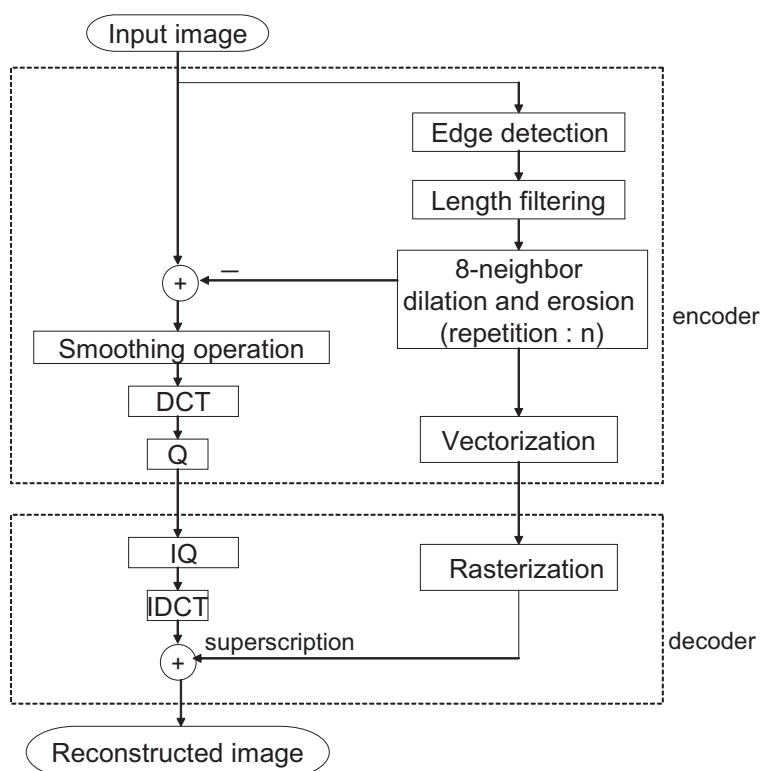


図 4-1: A flow of Superscription Model

本手法は、エンコーダ側では、抽出したエッジに対してベクトル化を行う一方で、原画像のエッジ領域を平滑化する。この平滑化画像に対して、DCT 変換ならびに量子化を行う。デコーダ側でエッジ成分を上書きすることで、エッジの保持、ならびにモスキートノイズの削減が可能となり、符号化効率の向上も期待できる。

まず入力画像  $f(x, y)$  からエッジ抽出を行う。効率のよいベクトル化を行うため、エッジ抽出は、フィルタによる処理と、周囲長の短い画素集合の除去という2つのフェーズから成る。エッジを、図 4-2 (a) のように、大きく線画 (ラインエッジ) と境界線 (ステップエッジ) とに分ける。ここで、次の処理でエッジ領域を平滑化することを考慮すると、隣接した境界線を包括的に線画として扱い、高周波成分を減らすことが重要となる。このため、抽出したエッジに対して、さらに8近傍に数回膨張・収縮を行うことで、隣接した境界線を線画として扱うことが可能となる。これをエッジ領域とし、ベクトル化を行う。エッジの輝度情報は、エッジをラベリングしてそれぞれ与える。デコーダ側には、ベクトル情報と、輝度情報を伝送する。

一方、入力画像において、エッジ領域の平滑化を行う。図 4-2 (b) に平滑化の様子を示す。これは、高周波成分除去による、DCT 変換における符号化効率向上ならびにモスキートノイズ低減のための操作である。平滑化は、図 4-3 のように、ガウス関数を利用して、式 4-1 で

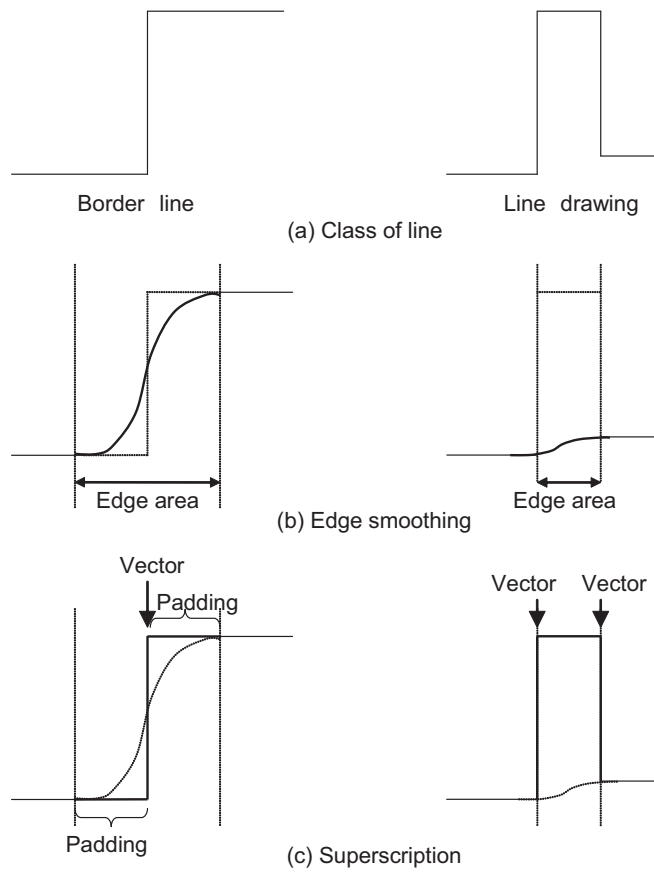


図 4-2: Superscription of edge

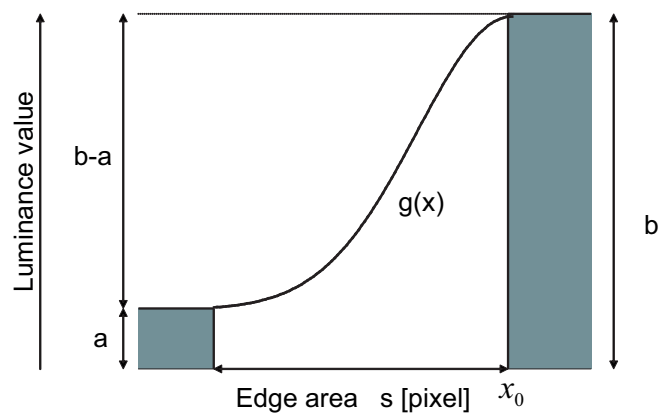


図 4-3: Smoothing method

x 方向, y 方向にそれぞれ行う.

$$g(x) = a + (b - a) \exp\left(-\frac{(x - x_0)^2}{2\sigma^2}\right) \quad (4.1)$$

$$\text{where } \begin{cases} \exp\left(-\frac{s^2}{2\sigma^2}\right) < 1.0 * 10^{-3} \\ \sigma = \frac{1}{\sqrt{14}}s \end{cases} \quad (4.2)$$

ここでガウス関数の標準偏差  $\sigma$  の値は, 式 4.2 の上式の条件を与え, 平滑化を行う距離  $s[\text{pixel}]$  を用いて算出を行う. この平滑化により, DCT 変換に適した低周波成分から構成される画像を得る. この画像に対して, DCT 変換ならびに量子化を行う.

デコーダ側では, 平滑化後に符号化した画像に対して, 逆量子化ならびに逆 DCT を施す. エッジ成分に対しては, ラスタライズし輝度情報を加える. 図 4-2 (c) のように, エッジ成分の上書きを行い, 残りのエッジ領域に対しては, 平滑化によるパディングを行うことで, 再構成画像を得る.

この手法により, エッジ領域の高周波成分を DCT 変換することがなくなるため, 周辺のもスキートノイズの発生をおさえることができ, 符号化効率の改善も可能となる. しかし, この手法は, エイリアシングを考慮したエッジ領域のベクトル化は成されるが, 解像度変換への適用は十分でない. また, 輝度をラベリングし, 平滑化によるパディングを行うため, 輝度値がほぼ一定のアニメーション画像への適用には十分であるが, 輝度値の変化しやすい自然画像への適用は適切でないと考えられる.

### 4.3 輝度近似モデル

前節の手法が, 上書きモデルであったのに対し, さらなる空間スケーラビリティ向上と符号化歪み特性の向上を目的とした, エッジ領域の加算モデルを提案する. また, 前手法での課題であった自然画への適用も同時に目指す. この手法は, ベクトル表現のスケーラビリティに対する優位性とラスタ表現の優れた符号化効率を考慮して, エッジ領域における高周波成分相当の画像を, 輝度近似をすることでベクトル化する. 入力画像との差分画像は, 従来の変換符号化を用いて圧縮を行う. デコーダ側で, ラスタライズ処理や解像度変換後に, これらを加算することで, 画像品質の向上を目指す.

#### 4.3.1 エッジの分離・再構成

まず, 本手法におけるエッジの分離・再構成の処理フローを図 4-4 に示す. 入力画像をエッジ画像と, 主に低周波成分からなる差分画像 (LFC: Low Frequency Component) に分離する. それぞれ最適な符号化を行った後に伝送し, デコーダ側で再構成を行う. ベクトル化を行うエッジ画像は, エイリアシングを考慮するために, 抽出したエッジ領域に, その画素に対応する高周波成分の輝度値を加えた画像とする.

入力画像を  $f(x, y)$  とすると, エッジ画像  $f_H(x, y)$  は,

$$f_H(x, y) = \text{Edge}\{f(x, y)\} \quad (4.3)$$



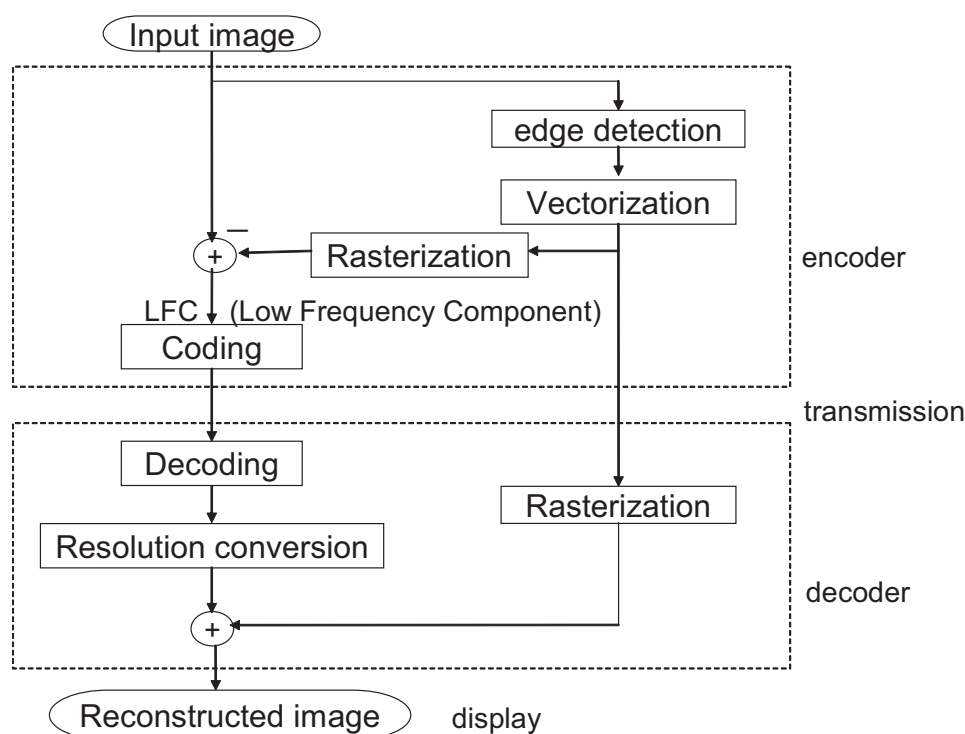


図 4-4: A flow of proposed method

と表せる．Edge{ $f$ } はラスタ画像  $f$  からエッジ画像を抽出する関数とする．また，ベクトル情報  $V_H(n)$  を，

$$V_H(n) = \text{Vectorize}\{f_H(x, y)\} \quad (4.4)$$

と表す．Vectorize{ $f$ } は，ラスタ画像  $f$  をベクタライズする関数とする．一方，Rasterize{ $v, s$ } をベクター画像  $v$  をスケール  $s$  でラスタライズする関数とすると，ベクトル情報  $V_H(n)$  と，4.3.2 で述べる輝度の近似曲線の係数  $h(n)$  を利用して，ラスタライズ時に  $s$  倍の拡大・縮小処理を施したエッジ画像  $R_{sH}(x, y)$  は，

$$R_{sH}(x, y) = \text{Rasterize}\{h(n)V_H(n), s\} \quad (4.5)$$

となる．

LFC 成分  $f_L(x, y)$  は，式 4-6 に示すように，輝度の近似曲線とベクトル情報により  $s = 1$  でラスタライズしたエッジ画像と，入力画像との差をとることで生成する．

$$f_L(x, y) = f(x, y) - R_{1H}(x, y) \quad (4.6)$$

LFC 成分は既存の手法により符号化  $X$  を行い， $Xf_L(x, y)$  を得る． $Xf_L(x, y)$  と，エッジ画像の情報として，ベクトル情報  $V_H(n)$  ならびに輝度の付加情報  $h(n)$  をそれぞれ伝送する．

次にデコータ側の処理として，まず拡大・縮小処理を行わない場合を述べる．受信側が持つ情報は，従来方式では信号処理で取得した信号系列であるのに対して，提案方式は信号系

列  $Xf_L(x, y)$  ならびにベクトル情報  $V_H(n), h(n)$  となる．符号化ノイズを  $n_c(x, y)$  とすると，LFC 成分  $Xf_L(x, y)$  をデコードした画像

$$X^{-1}Xf_L(x, y) = f_L(x, y) + n_c(x, y) \quad (4.7)$$

と，等倍の拡大・縮小処理によりラスタライズしたエッジ画像

$$R_{1H}(x, y) = \text{Rasterize}\{h(n)V_H(n), 1\} \quad (4.8)$$

との和をとることで再構成画像  $f_R(x, y)$

$$\begin{aligned} f_R(x, y) &= R_{1H}(x, y) + f_L(x, y) + n_c(x, y) \\ &= R_{1H}(x, y) + f(x, y) - R_{1H}(x, y) + n_c(x, y) \\ &= f(x, y) + n_c(x, y). \end{aligned} \quad (4.9)$$

を得る．

以上の処理により，エッジ成分を分離し，それぞれ符号化し，デコーダ側で再構成することが可能となる．

また，特に LFC 成分に対してロスレス符号化を適用すると，

$$n_c(x, y) = 0 \quad (4.10)$$

であるので，

$$\begin{aligned} f_R(x, y) &= R_{1H}(x, y) + f_L(x, y) + 0 \\ &= f(x, y). \end{aligned} \quad (4.11)$$

より，可逆な分離・再構成が可能となる．

次に，この分離・再構成を実現するため，エイリアシングを考慮した，エッジ情報の保持を目的としたエッジ成分のベクトル化手法を示す．

### 4.3.2 エッジ成分のベクトル化

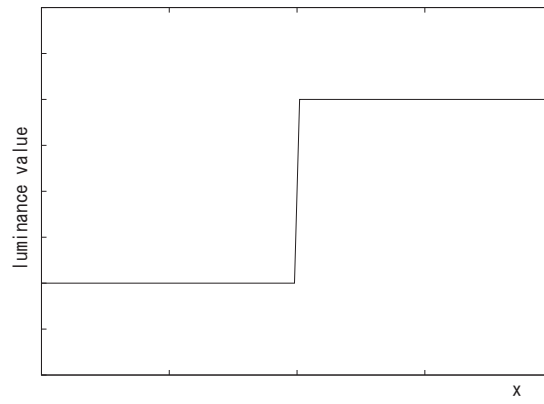
ここで，ベクトル化を行うエッジ画像は，抽出したエッジ領域に対応する高周波成分と定義する．つまり，エッジ画像は，主として高周波成分であり，輝度情報を持つ．しかし，ベクトル表現は，通常一本の線に対して一つの輝度しか持つことができない<sup>(24)</sup>．今回は，複数の輝度情報を持つように，ベクトル表現を拡張する．

一般的に境界線に対する高周波成分の輝度は一次元では図 4-5(c) のように表され，式 4-12 の曲線  $g(x)$  で近似できる．

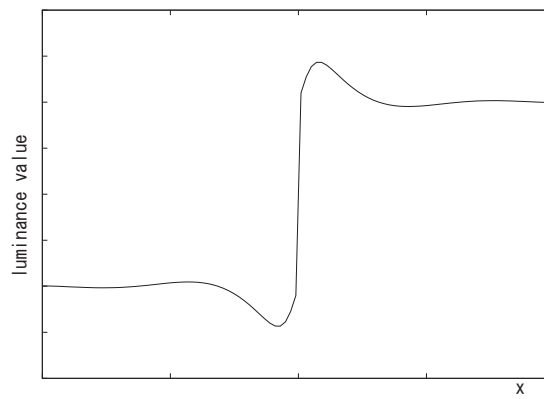
$$g(x) = \frac{a \sin(bx + c)}{x^2 + dx + e} \quad (4.12)$$

ここで， $a, b, c, d, e$  は変数であり，各値はエッジの強度によって決定される．ベクトル方向に沿った輝度値の取得を行い，近似曲線で輝度値を表すことで，ベクトル表現時の輝度値の付加が可能となる．

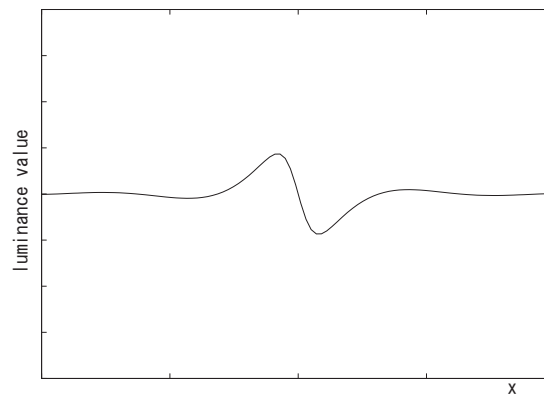
近似曲線の係数取得は，パラメータ  $m, w$  を用いて以下の手順で行う．



(a) Luminance value of input image around edge



(b) Luminance value of LFC



(c) Luminance value of Edge component  
 $(c)=(a)-(b)$

☒ 4-5: Luminance value of each component

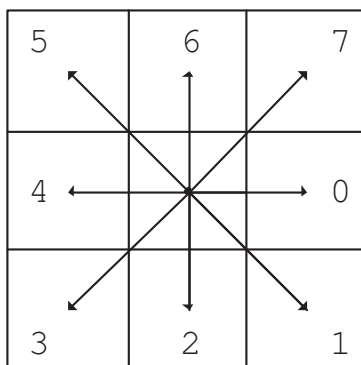


図 4-6: Freeman chaincode

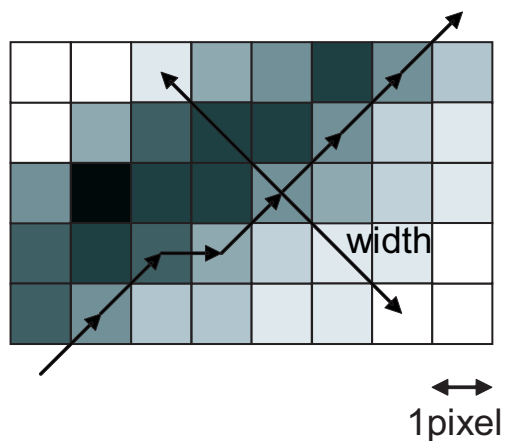


図 4-7: Acquisition of luminance values

1. ベクトル化する多値のエッジ画像を細線化する
2. Freeman の chain code<sup>(78)</sup>(図 4-6) を用いて, 線に沿って方向を探索する
3.  $m$  画素ごとの平均方向を算出する
4. 方向が同じ隣接集合画素をまとめる
5. 算出した方向に対して直角方向に, 輝度を幅  $w$  画素で取得し, 集合ごとに平均化する
6. 得られた輝度値を用いて, 最小二乗法により  $g(x, y)$  の係数を取得する.

この操作の様子を図 4-7 に示す. 次に,  $g(x, y)$  をベクトル座標  $n$  にあわせて  $h(n)$  に変換する. デコーダ側でラスタライズする際, この手法で取得した係数  $h(n)$  を利用して, 輝度値をベクトル情報  $V_H(n)$  に付加することで, 図 4-8 に示すように多値のエッジ画像の再現が可能となる.

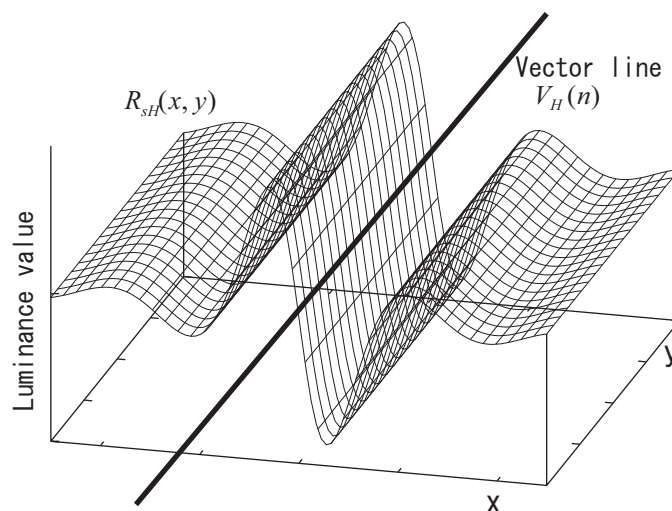


図 4-8: Rasterization

### 4.3.3 解像度変換

次に，近似曲線の係数を利用して，デコーダ側で任意の解像度変換を行う手法を示す． $s$  倍の拡大・縮小処理を行う場合，ラスタライズ後のエッジ画像  $R_{sH}(x, y)$  は，伝送後のベクトル情報  $V_H(n)$  を用いて，

$$R_{sH}(x, y) = \text{Rasterize}\{h(n)V_H(n), s\} \quad (4.13)$$

と表せることはすでに述べた．また，LFC 成分に対しては， $\text{Rconv}\{f_L(x, y) + n_c(x, y), s\}$  の処理を行う． $\text{Rconv}\{f, s\}$  は，ラスタ画像  $f$  をスケール  $s$  で解像度変換する関数とする．エッジ画像は，ベクトル画像をラスタライズする際に，目標とする解像度にあわせた輝度を，前節で求めた近似曲線を使って画素ごとに求めることで生成する．LFC 成分に対しては，デコード後に解像度変換を施す．

以上の処理により，再構成画像  $f_R(x, y)$

$$f_R(x, y) = R_{sH}(x, y) + \text{Rconv}\{f_L(x, y) + n_c(x, y), s\} \quad (4.14)$$

を得る．

また，ラスタライズ時に，近似曲線の係数を操作し，表示する解像度に最適な曲線の幅と振幅を与えることで，拡大・縮小表示時のエッジの保持ならびに強調を行うことも可能である．このとき，エッジ画像  $R_{sH}(x, y)$  は，

$$R_{sH}(x, y) = \text{Rasterize}\{h'(n)V_H(n), s\} \quad (4.15)$$

で与えられる．

#### 4.4 おわりに

本章では、空間スケーラビリティの向上を目的として、ベクトル表現を用いたエッジの分離・再構成手法について提案した。特に、解像度変換時に特に劣化の原因となるエッジ領域をベクトル化する手法について検討した。まず、エイリアシングを考慮したベクトル表現のエッジ領域への適用を目指し、エッジ領域の上書きモデルを提案した。次に、エッジ輝度の近似によるエッジの分離・再構成手法を提案した。付加情報として複数の輝度情報を持ったベクトル化手法を提案し、アンチエイリアシング処理が適用されたエッジ領域に対してベクトル化を行うことが可能となることを示した。また、解像度変換を行わない場合、完全再構成が可能な手法であることを明らかにし、解像度変換を行った際の特性について説明した。

## 第5章

### 評価実験

#### 5.1 はじめに

本章では，前章で提案した手法を実装し，評価を行う．まず，アニメーション画像に対してエッジ領域の上書きモデルを適用し，有効性を検討する．次に，輝度近似モデルにおいて，輝度値を付加情報として保有したベクトル表現を用いることにより，エイリアシングを考慮したベクトル化ならびにラスタライズが可能であることを示す．解像度変換を行い，ベクトル表現のスケーラビリティを利用して再構成画像を得ることで，従来の信号の補間手法と比較して有効であることを示す．

#### 5.2 エッジ領域上書きモデル

提案手法を実装し，実験を行った．実験条件を表 5-1 に示す．入力画像は，図 5-1(a) のように急峻なエッジを含むアニメーション画像を用いた．Canny operator は，第三章でも示したように，ベクトル化する際の最も良いエッジ抽出手法の一つであるため，まず Canny operator によってエッジを抽出した．その後，膨張・収縮処理は各一回ずつ施した．

ベクトル表現はベクトル化ツール Potrace を用いて EPS で符号化した．また，EPS はエントロピー符号化を行わないため，汎用のテキスト圧縮ツールである bzip2 で圧縮した．今回はイントラ画像を対象とするため，DCT 変換ならびに量子化は JPEG を利用した．

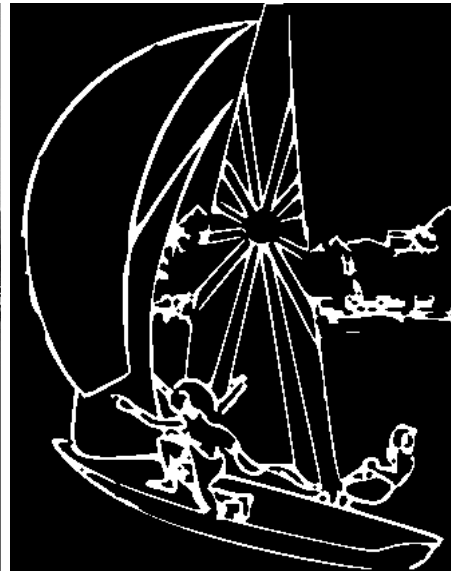
表 5-1: Experiment condition of Superscription Model

Image size	width:334 hight:426[pixel]
Luminance value	0 ~ 255
Edge detection	Canny operator
	Length filtering : 30[pixel]
	Repetition : n=1
Vector coding	EPS(by Potrace)
DCT coding(intra)	JPEG

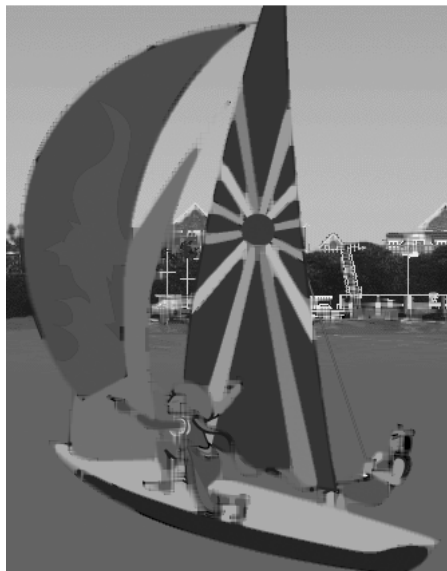




(a) Input image



(b) Edge image  
EPS:0.194[bit/pel]



(c) Smoothed image  
JPEG:0.501[bit/pel]



(d) Reconstructed image  
EPS+JPEG:0.695[bit/pel]  
PSNR:31.90[dB]

☒ 5-1: Modified images of Superscription Model

表 5-2: Bit rate of Superscription Model

Operation	Bit rate[bit/pel]			PSNR [dB]
	Vector	Smoothed image	Total	
Proposed method	0.194	0.501	0.695	31.90

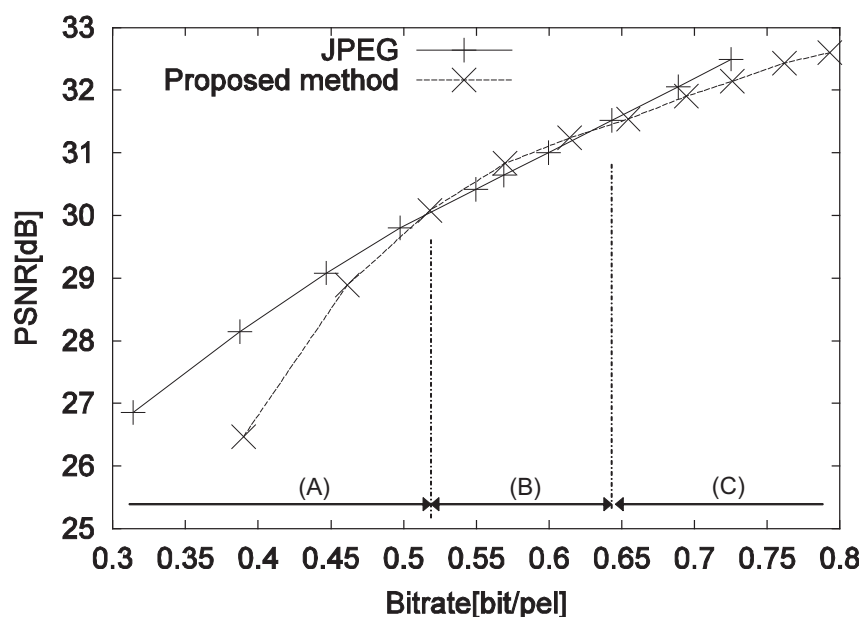


図 5-2: Evaluation of Superscription Model

図 5-1において，エッジ画像を (b) に，平滑化画像を (c) に，再構成画像を (d) に示す．図 5-1より，(b) のエッジ画像と (c) の平滑化画像とを再構成することにより，(d) の画像が得られることが分かる．図 5-1における，それぞれの符号量を表 5-2に示す．

また，本実験での R-D 曲線を図 5-2に示す．ここで，提案手法における符号量は，JPEG 画像と，ベクトル画像のファイルサイズを足し合わせたものとする．

本手法では，エッジ除去後の画像に平滑化処理を施すことにより，高周波成分が削減され，量子化後の符号量削減が可能となった．また，平滑化によるモスキートノイズの低減で，品質が向上することを確認した．一方，ベクトル表現の導入により，抜き出すエッジの量に比例して符号量は増加する．

図 5-2より，(B) の範囲において，提案手法は既存の JPEG 符号化と比べて最大で約 0.2[dB] のゲインを得られることがわかった．(C) の範囲では，モスキートノイズによる劣化はあまり起きず，平滑化による効果が十分に表れていない．一方，(A) の範囲では，ハイブリッドな符号化におけるベクトル表現の占める符号量が相対的に大きくなり，符号量を十分に削減できないと考えられる．本手法によって，時空間スケーラビリティの確保と線の保持が可能になり，エッジ付近に生じる多数のモスキートノイズを抑えられることがわかった．R-D 曲線を利用した判定によるフレキシブルな符号化方式の選択により，効果的な符号化が可能に

表 5-3: Experiment condition

Image size	width:334 hight:426[pixel]
Liminance value	0 ~ 255
Edge detection	Canny operator
	Length filtering : 30[pixel]
Vectorization	EPS(by Potrace <sup>(3)</sup> )
	Length :n= 3[pixel]
	Width :w =5[pixel]
Lossless Coding	DPCM

なると考えられる。しかし、この手法では、解像度変換への適用は十分でないと考えられる。また、輝度のラベリングと、平滑化によるパディングを行うため、輝度値の変化しやすい自然画像への適用は適切でないと考えられる。

### 5.3 輝度近似モデル

#### 5.3.1 エッジの分離・再構成

提案手法を実装し、実験を行った。実験条件を表 5-3 に示す。

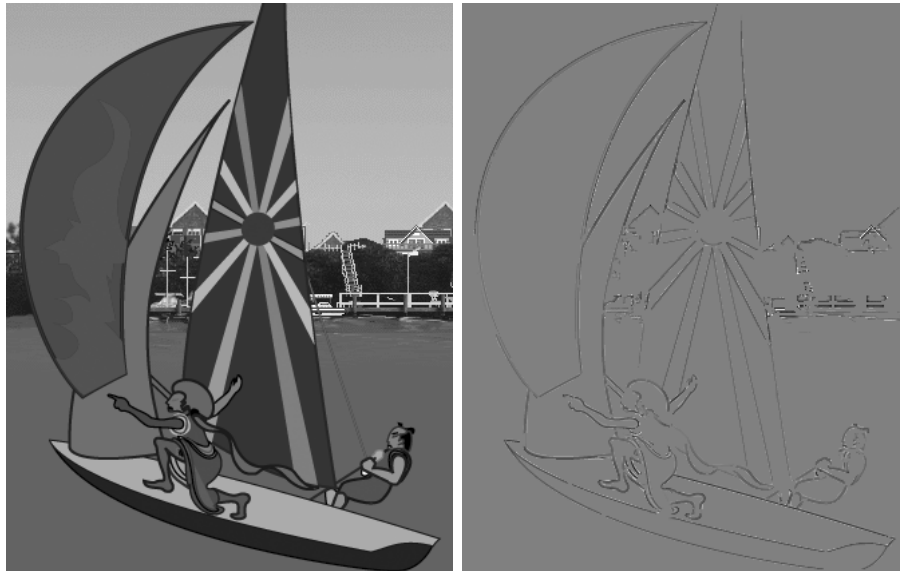
入力画像は、前節の実験同様、まずは図 5-3 (a) のように急峻なエッジを含むアニメーション画像を用いた。エッジ画像は、以下の手順によって求めた。

1. 高周波成分の画像を求める
2. 入力画像に対して、Canny operator を用いてエッジ領域を抽出する
3. 周囲長の短い画素集合を取り除く
4. (3) で求めた画素領域に対応する高周波成分をエッジ画像とする。

Canny operator における各パラメータは実験値を用いた。ベクトル表現は Potrace を用いて EPS で符号化し bzip2 で圧縮した。ベクトル表現に付加された輝度情報  $h(n)$  も同様に bzip2 で圧縮した。

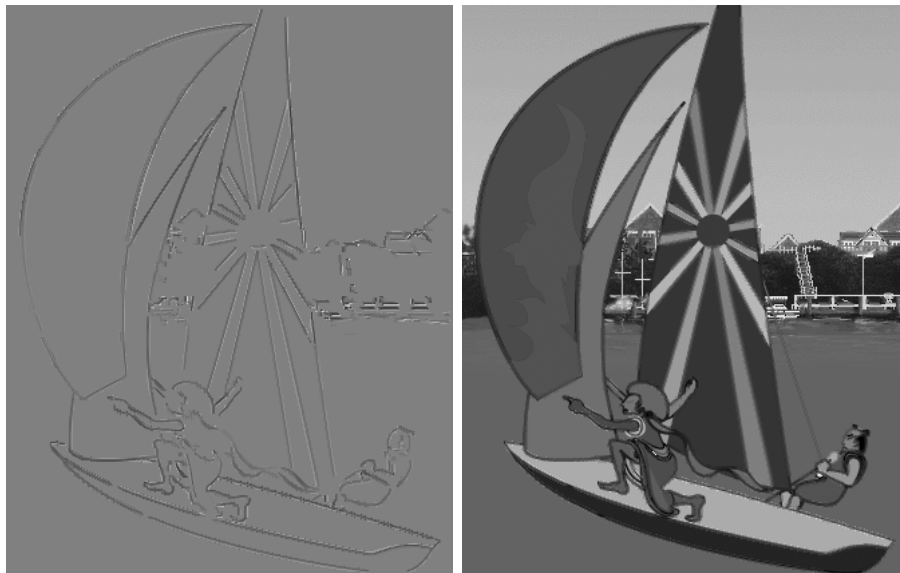
図 5-3 において、エッジ画像を (b) に、ラスタライズ画像を (c) に、LFC 成分の符号化前の画像を (d) に示す。ここで、図 5-3 (b), (c) のエッジ画像は、それぞれ輝度値 128 を付加したものを表示している。図 5-3 より、ベクトル画像をラスタライズする際に、近似曲線の係数を与えることで、多値画像のベクトル化が可能であることがわかる。これにより、縮小・拡大表示時の輝度の計算が可能となり、空間スケーラビリティが確保できる。

また、LFC 成分にロスレス符号化を適用した際には、再構成画像は入力画像と等しくなることも確認できた。



(a) Input image

(b) Edge component



(c) Edge component after rasterization (proposed method)

(d) Low Frequency Component (LFC): (a)-(c)

図 5-3: Modified images (animation)

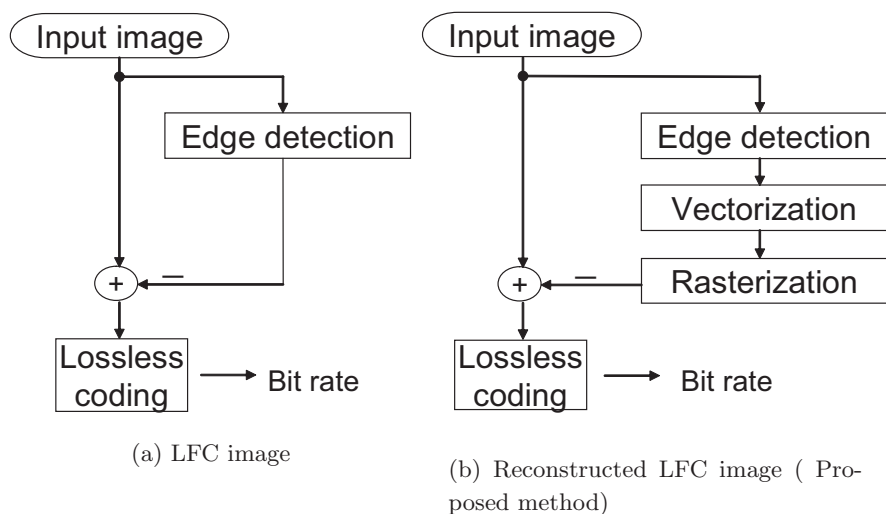


図 5-4: Evaluation method of vectorization

表 5-4: Evaluation of rasterized image

Image	Bitrate[bit/pel]
Input image	4.34
LFC image	3.94
Reconstructed LFC image(proposed method)	4.33

### 5.3.2 エッジ成分のベクトル化

多値で表されたエッジ画像のベクトル化における再現性を測るために、LFC成分のロスレス圧縮符号化時の符号量を比較した。測定方法を図5-4に示す。入力画像から、単純に抽出したエッジ画像を引いた差分画像と、エッジ画像をベクトル化しラスタライズしたものを引いた差分画像の符号量を比較することで、再現性を示す。

各符号量を表5-4に示す。表5-4より、提案手法によって生成されるLFC成分の符号量は、入力画像の符号量とほぼ同等であることがわかった。これは、ラスタライズして得たエッジ成分画像は輝度を一画素ずつ与えるため、参照画素が少しでもずれると、図4-5(c)のような波形によって、ハイパス成分を強めるような輝度が与えられることが原因にあると考えられる。近似曲線の係数の与え方についてのさらなる検討を行い、再現性を高めていく必要がある。

表 5-5: Bit rate of edge components

Resolution 334x428	Bit rate[bit/pel]		
	Vector coding: $V_H(n)$	Sub data: $h(n)$	Total
668x856(x2.0)	0.65	0.44	1.09
334x428(x1.0)	0.38	0.91	1.29
167x214(x0.5)	0.17	0.22	0.39

### 5.3.3 解像度変換

#### エッジ成分の符号量

解像度を変化させた際のエッジ成分の符号量を測定した。エッジ成分の符号量は、ベクトル情報  $V_H(n)$  とその輝度情報  $h(n)$  の和であるが、ベクトル化を行う解像度によって変化する。つまり、

- ベクトル情報におけるパス数の変化
- ベクトル付加情報である輝度情報量  $h(n)$  の変化

が考えられる。解像度とエッジ成分の符号量の関係を表 5-5 に示す。なお、輝度を求める際の長さ、ならびに幅は、それぞれの解像度に応じて変換を行った。また、少数点以下に対しては四捨五入した。

表 5-5 より、解像度と符号量との関係は、非線形であることがわかる。また、この情報を使って、任意の解像度で画像を表示する際は、一般的に高い解像度で抽出したエッジ情報を使った方がより再現性は高いといえる。

#### 解像度変換における画質評価

任意の解像度の画像を生成する際に、ベクトル情報を使った再構成手法が有効であることを示すために、従来の信号の補間手法と比較した。原画像を伝送時に 1/2 に縮小することを仮定し、デコーダ側で再び 2 倍に拡大した際の品質を測定した。(図 5-5)

従来手法は、原画像を符号化し、伝送時に 1/2 に縮小した。一方、提案手法は、原画像からエッジ画像を引いた LFC 成分を符号化し、伝送時に 1/2 に縮小した。デコーダ側では、従来手法、提案手法ともに、それぞれ二倍に拡大し、原画像との PSNR を測定した。縮小画像の生成には JPEG2000 で導入されている 9/7 タップの Daubechies Filter を用いた。Daubechies Filter の係数を表 5-6 に示す。

拡大処理には、

- ニアレストネイバー法
- バイキュービック法

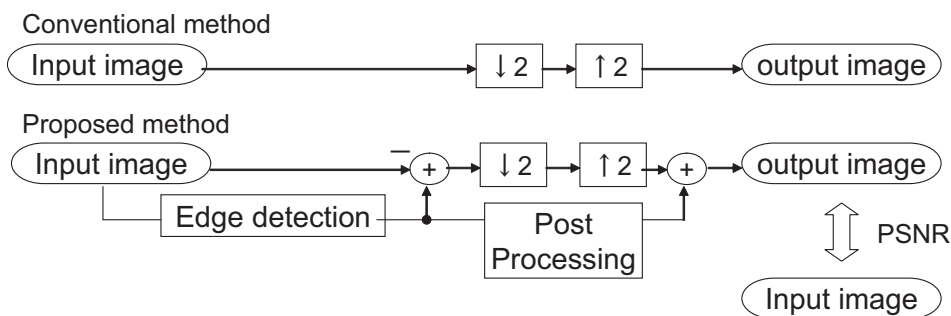


図 5-5: Resolution conversion

表 5-6: 9/7 Daubechies filter coefficients

k	Low Pass	High Pass
0	0.6029490182363579	1.115087052456994
+/- 1	0.2668641184428723	-0.5912717631142470
+/- 2	-0.07822326652898785	-0.05754352622849957
+/- 3	-0.01686411844287495	0.09127176311424948
+/- 4	0.026748741080976	-

- バイリニア法

を利用した。

実験は符号化ノイズの影響を考慮して二通り行った。

- ロスレス符号化
- ロッシー符号化

ロッシー符号化には JPEG2000 を使い、LFC 成分を 1.0[bit/pel] で圧縮した。従来の信号の補間手法における符号化レートは、提案手法の合計ビットレートと同一のものとした。それぞれの PSNR を測定した結果を表 5-7, 5-8 に示す。また、その際のビットレートを表 5-9 に示す。

表 5-7, 5-8 より、提案手法は従来手法と比較して、いずれも良い結果が得られた。特に、拡大処理にバイリニア法を利用すると、提案手法は従来手法と比べて約 2dB 向上することが確認できた。また、符号化ノイズ  $n_c(x, y)$  による再構成時の影響は特に見受けられないことがわかった。

ロスレス変換において、バイリニア法を用いた際のそれぞれの出力画像を図 5-6 に示す。バイキュービック法、ニアレストネイバー法においてもほぼ同等の出力画像を得た。

図 5-6 より、従来手法では、エッジの部分にジャギーが見られ、複雑なエッジに対しては線の欠落やぼけが生じているが、提案手法ではいずれも線の品質保持が可能となっていることがわかった。



表 5-7: PSNR of reconstructed images (Lossless coding)

Operation	PSNR[dB]	
	Conventional method	Proposed method
Nearest neighbor	23.47	24.66
Bi-cubic	24.40	25.63
Bi-linear	24.43	26.54

表 5-8: PSNR of reconstructed images (Lossy coding)

Operation	PSNR[dB]	
	Conventional method	Proposed method
Nearest neighbor	23.31	24.57
Bi-cubic	24.19	25.61
Bi-linear	24.23	26.26

#### 5.3.4 自然画への適用

さらに、 $512 \times 512$ pixel の自然画 Lena に対しても同様に実験を行った。アニメーション画像を用いた実験より、提案手法は、符号化ノイズの影響を特に受けないことは確認できているが、拡大手法としてバイリニア法を用いてさらに詳しく検討を行った。LFC 成分の符号量を変化させて測定を行うと同時に、従来手法を提案手法と同じビットレートで符号化することにより、R-D 特性を調べた。測定結果を図 5-7 に示す。また LFC 成分を  $1.0[\text{bit}/\text{pel}]$  で符号化した際の、各符号量ならびに PSNR を表 5-10 に示す。

図 5-7 より、提案手法による再構成画像は従来手法と比較して、最大で約  $1.5[\text{dB}]$  向上することが確認できた。また、特に、符号量  $0.6[\text{bit}/\text{pel}]$  における PSNR の値は、提案手法と従来手法でほぼ等しいという結果が得られた。 $0.61[\text{bit}/\text{pel}]$  におけるそれぞれの処理におけ

表 5-9: Bit rate of each component (animation)

Component		Bit rate[bit/pel]	
		Conventional	Proposed
Vector: $V_H(n)$		-	0.38
Sub data: $h(n)$		-	0.91
Lossless coding	LFC	-	4.33
	total	4.34	5.61
Lossy coding	LFC	-	1.00
	total	2.29	2.29

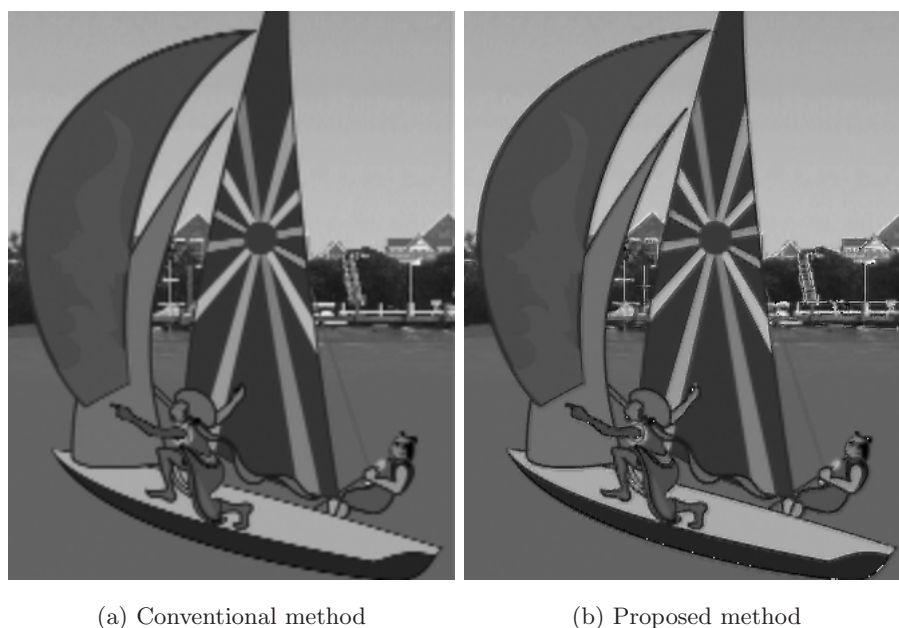


図 5-6: Reconstructed images

表 5-10: Bit rate of each component (Lena)

Operation (Method)	Bit rate[bit/pel]				PSNR [dB]
	Vector: $V_H(n)$	Sub data: $h(n)$	LFC	Total	
Conventional	-	-	-	1.31	30.05
Proposed	0.23	0.08	1.00	1.31	31.50

る出力画像を図 5-8 に示す。図 5-8 において、エッジ画像を (b) に、LFC 成分の符号化後の画像を (c) に、再構成画像を (d) に示す。ここで、図 5-8 (b) のエッジ画像は、それぞれ輝度値 128 を付加したものを表示している。また、0.61[bit/pel] での符号化後、従来の補間手法を適用した画像を図 5-8 (e) に示す。

低レートにおいては、図 5-8 より、従来の補間手法では全体的に劣化し、エッジ領域がぼけた画像となることがわかる。一方、提案手法では、LFC 成分のビットレートはさらに低いため著しく劣化するものの、エッジ画像を加算し再構成することで、特にエッジ領域の画質が向上することが確認できた。エッジ領域の画質のみを評価する場合、提案手法はさらに優れた性能を有していることがわかる。PSNR を用いて、画像全体の品質を評価した場合についても、0.6[bit/pel] 以上の符号化レートにおいて、提案手法は従来手法と比較して有効であることが確認できた。

本実験では、算出したエッジ輝度の近似曲線係数をそのまま利用し、解像度での輝度を与えたが、ラスタライズ時に、表示する解像度に最適な曲線の幅と振幅を与えることで、拡大・縮小表示時のエッジの保持ならびに強調が実現できると考えられる。

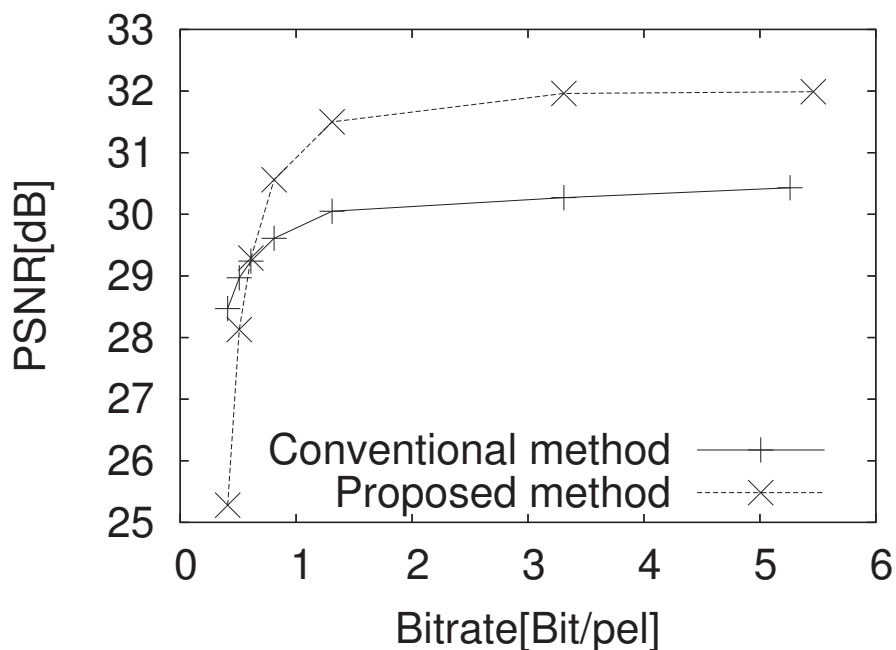


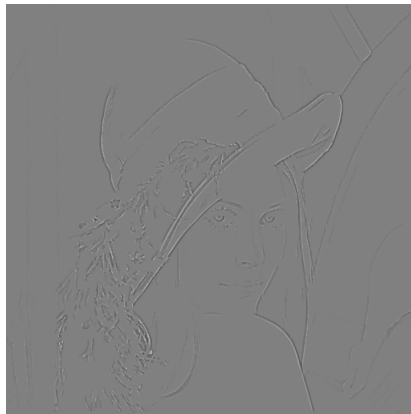
図 5-7: Rate-Distortion of proposed method (Lena)

#### 5.4 おわりに

本章では、空間スケーラビリティの向上を目的として、エッジ成分をベクトル化する手法を実装し、評価を行った。まずエッジ領域の上書きモデルをアニメーション画像に適用し、有効性を検討した。この手法により、時空間スケーラビリティの確保と線の保持が可能になり、エッジ付近に生じる多数のモスキートノイズを抑えられることを示した。また、R-D 曲線を利用した判定によるフレキシブルな符号化方式の選択により、効果的な符号化が可能であることを示した。しかし、この手法は、解像度変換への適用は十分でなく、輝度値の変化しやすい自然画像への適用は適切でないことを明らかにした。次に、輝度近似手法について評価ならびに考察を行った。輝度値を付加情報として保有することにより、エイリアシングを考慮したベクトル化が可能であることを示した。また、ベクトル表現の符号量に大きく左右されることを示した。解像度変換を行い、スケーラビリティを利用して再構成画像を得ることで、従来の信号の補間手法と比較して最大約 2[dB] の品質向上が実現することを明らかにした。また、提案手法は、符号化雑音に影響されず、自然画にも適用可能であることを示した。



(a) Input image



(b) Edge component  
 $V_H(n)$ : 0.23,  $h(n)$ : 0.08[bit/pel]



(c) Low Frequency Component  
 LFC: 0.30[bit/pel]



(d) Reconstructed image  
 EPS+LFC: 0.61[bit/pel]  
 PSNR:29.29[dB]



(e) Reconstructed image  
 (Conventional) 0.61[bit/pel]  
 PSNR:29.24[dB]

☒ 5-8: Modified images (Lena)

## 第6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。

#### 6.1 総括

本論文では、ベクトル表現の空間スケーラビリティに着目し、伝送から表示までの一貫したスケーラビリティの確保を目的として、符号化歪み特性の改善を目指した。ベクトル表現のスケーラビリティに対する優位性とラスタ表現の優れた符号化効率を考慮して、エッジ成分をベクトル化し再構成する手法について提案した。実験では、アニメーション画像に加え、従来ベクトル表現では不適切とされてきた自然画像に対しても提案手法を適用し、有効性を検証した。

本研究により、多様化した閲覧端末における様々な解像度での表示技術として必要な、伝送から表示までの一貫したスケーラビリティの確保に関する基礎検討がなされた。ベクトル表現のスケーラビリティ機能を用いたエッジ成分の符号化により、ディスプレイに合わせた解像度変換時に線の品質保持が達成された。また、従来の手法と比較して符号化歪み特性が改善することが確認できた。これによって、表示解像度によらずマルチメディアコンテンツを再生することが可能となり、柔軟な扱いや流通、二次利用によるコンテンツのワンソース・マルチユースがますます実現に向かうと考えられる。

第2章では、電子コンテンツの需要や閲覧端末の多様化などの背景より、空間スケーラビリティを持った符号化の必要性を説明した。また、アニメーション画像と自然画像の電子コンテンツにおける信号特性の違いと符号化への影響を明らかにした。そして、符号化データのスケーラビリティ機能に着目し、ラスタ表現の二値・多値静止画像、動画像それぞれにおける符号化方式とその機能について述べた。さらに、ベクトル表現の特徴を明らかにし、符号化効率と機能の観点から検討を行い、ベクトル表現がエッジなどの輪郭の符号化に適切であることを述べた。

第3章では、まず、本提案手法で必要となる要素技術として、解像度変換、エッジ検出、ラスタ・ベクタ変換について説明した。解像度変換では主な手法として、補間による手法、マルチレート信号処理に基づく手法、直交変換による3つの手法について、その特徴を述べた。エッジ検出手法は、一般的な微分オペレータについて述べ、Canny オペレータがベクト

ル化を行う際のエッジ検出に有効であることを明らかにした。ラスタ・ベクタ変換については、その基礎として、近似関数曲線について述べ、線画像を多角形により近似する手法について説明した。従来の画像の分離・再構成手法として、アニメーション符号化、骨格・テクスチャ分離、SVG レンダリングの3つについて述べ、問題点を明らかにした。

第4章では、空間スケーラビリティの向上を目的として、解像度変換時に特に劣化の原因となるエッジ領域をベクトル化する手法を提案した。まず、エイリアシングを考慮したベクトル表現のエッジ領域への適用を目指し、アニメーション画像に対して、エッジ領域の上書きモデルを提案した。次に、自然画への適用も考慮したエッジ輝度の近似によるエッジの分離・再構成手法を提案した。ベクトル表現は、通常一本の線に対して一つの輝度しか持つことができないため、付加情報として複数の輝度情報を持ったベクトル化手法を提案し、アンチエイリアシング処理が適用されたエッジ領域に対してベクトル化を行うことが可能となることを示した。また、解像度変換を行わない場合、完全再構成が可能な手法であることを明らかにし、解像度変換を行った際の特性について説明した。

第5章では、まずアニメーション画像をエッジ領域の上書きモデルに適用し、有効性を検討した。この手法により、時空間スケーラビリティの確保と線の保持が可能になり、エッジ付近に生じる多数のモスキートノイズを抑えられることを示した。また、R-D 曲線を利用した判定によるフレキシブルな符号化方式の選択により、効果的な符号化が可能であることを示した。しかし、この手法は、解像度変換への適用は十分でなく、輝度値の変化しやすい自然画への適用は適切でないことを明らかにした。次に、輝度近似手法について評価ならびに考察を行った。輝度を付加情報として保有することにより、エイリアシングを考慮したベクトル化が可能であることを示した。提案手法は、ベクトル表現の符号量に大きく左右されるが、解像度変換時には、スケーラビリティを利用することにより、最大約 2[dB] の品質向上が実現することを明らかにした。また、提案手法は、符号化雑音に影響されず、自然画にも適用可能であることを示した。

第6章は結論であり、本論文の総括および今後の課題について述べている。

## 6.2 今後の課題

今後検討すべき課題を以下に示す。

本論文では、符号化歪み特性を考慮した、エッジ画像のベクトル表現による空間スケーラビリティの確保を目的として検討を行った。このため、ベクトル表現のスケーラビリティに対する優位性とラスタ表現の優れた符号化効率を考慮して、対象画像をエッジ成分と LFC 成分に分離し、符号化を行っている。しかし、これ以外にも、分離方法としてはいくつかの方法が考えられる。例えば、1. LFC 成分を、純粋な低周波成分とその差分のテクスチャ成分に分離する方法、2. 画像を低周波成分と高周波成分に分離後、高周波成分に対してベクトル化を施す手法、3. SVG レンダリングで生成された画像を、その差分のテクスチャ成分とともに再構成する手法、などが考えられる。しかし、テクスチャに対する十分な符号化技術は依然確立されておらず、圧縮が十分になされないといった問題点がある。符号化歪み



特性を重視するか、スケーラビリティを確保するかで、アプローチは変わると考えられ、今後の情報通信産業の動向をふまえる必要があると考えられる。

本実験では、エッジ画像の抽出には、符号化歪みが最適となるような実験値を用いている。エッジ検出は、特に自然画像の提案手法への適用において、符号量ならびに品質を左右する重要な要素のひとつと考えられるため、最適なエッジ検出手法を今後継続して検討していく必要がある。

また、輝度近似手法の問題点としては、エッジ成分はラスタライズ時に輝度を一画素ずつ与えるため、参照画素が少しでもずれると、ハイパス成分を強めるような輝度を与えられるということがある。ベクトル化の際に、近似曲線の係数の与え方と算出方法、サブピクセル単位での輝度値の取得についてのさらなる検討を行い、ロバスト性を高めていくことが求められる。

また今回、解像度変換と再構成のフェーズにおいては、その再現性の評価に PSNR を用いている。つまり、入力画像との比較が必要であり、再構成画像を入力画像と等倍にして画質を評価した。このため、 $1/2$  倍にダウンサンプリングしたものを伝送後、 $2$  倍にアップサンプリングした。しかし、実際の環境では、様々な解像度への変換が想定されるため、PSNR 以外の客観評価手法として、例えば、無参照 (NR) 型の客観評価手法を用いて、様々な有理数比での検証をする必要があると考えられる。

さらに、本実験では再構成時に、算出したエッジ輝度の近似曲線係数をそのまま利用し、任意の解像度での輝度を与えたが、提案手法では、線の保持や強調などの機能性の付加が可能である。拡大・縮小表示時に、エッジの保持ならびに強調を実現するためには、ラスタライズ時に、表示する解像度に最適な曲線の幅と振幅を与えることが重要であると考えられる。ここで与える近似曲線の係数を算出する方法が検討課題のひとつとしてあげられる。

最後に、本手法のアニメーションを含む動画像への適用に向けて、ベクトル表現の機能を活用した時間スケーラビリティの確保が大きなテーマになると考えられる。ベクトル表現は、オブジェクト単位の構造化を行うことができるため、ベクトル表現の通過点を用いてフレーム間の相関をとることで、これが実現できる可能性がある。時間方向に一本のベクトルで物体の動きを表すことが可能となれば、時空間スケーラビリティはより向上すると考えられる。



## 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はじめ、研究の細部まで御指導頂きました渡辺 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国際情報通信科 富永 英義 教授に深く感謝致します。

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河村 圭 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オーディオビジュアルゼミの皆様に御礼申し上げます。特に、研究の方向性について数々の有意義なご意見をいただいた石川 孝明 氏に心から感謝致します。また、英語論文執筆時に、細部にいたるまでチェックいただいた Abhay Ghatpande 氏に深く感謝いたします。オーディオビジュアルゼミで3年間研究を共にした同期の、数学者イタサンこと伊谷 裕介 君、画像符号化に関していろいろと教えてくれた研究肌の後藤 崇行 君、サポートセンター 加藤 幸一 君、そして、大学1年からの長い長いお付き合い 秋間 佑輔 君に心からお礼申し上げます。さらに、学部4年の頃から研究生活を充実したものにしてくれた富永研究室の同期のみなさん、先輩・後輩のみなさんに深く感謝します。

最後に、私をここまで育てて下さった家族に深く感謝します。

平成 19 年 3 月 15 日

## 参考文献

- (1) O. Nakagami, T. Miyazawa, H. Watanabe, and H. Tominaga, "A Study on Two-layer Coding for Animation Images," IEEE International Conference on Multimedia and Expo. 2002 (ICME2002), Vol. 1, pp. 26 - 29, Aug. 2002.
- (2) K. Kawamura, H. Watanabe, and H. Tominaga, "Vector Representation of Binary Images Containing Halftone Dots," IEEE International Conference on Multimedia and Expo. 2004 (ICME2004), Vol. 1, pp. 335 - 338, Jun. 2004.
- (3) "Potrace," <http://potrace.sourceforge.net/>
- (4) "Autotrace," <http://autotrace.sourceforge.net/>
- (5) "高度情報通信ネットワーク社会推進戦略本部 (IT 戦略本部)," <http://www.kantei.go.jp/jp/singi/it2/index.html>
- (6) 経済産業省, コンテンツ産業の現状と課題, Feb. 2005.
- (7) "GDH," <http://www.gdh.co.jp/>
- (8) "10DaysBook," <http://bb.10daysbook.com/shop/>
- (9) "「マンガのコンテンツ流通支援ソリューションの提供」を開始," <http://www.celsys.co.jp/pdf/company/press/2003/031022contents.pdf>
- (10) "LIBRIe," <http://www.sony.jp/products/Consumer/LIBRIE/>
- (11) "シグマブック," <http://www.sigmabook.jp/>
- (12) 内田龍男 (監), 図解 電子ディスプレイのすべて, 工業調査会, Oct. 2006.
- (13) 小野文孝, 渡辺裕, 国際標準画像符号化の基礎技術, コロナ社, Mar. 1998.
- (14) 亀山渉, 金子格, 渡辺裕, 情報圧縮技術, IDG ジャパン, Mar. 2003.
- (15) ISO/IEC 11544, International standard : "Coded Representation of Picture and Audio Information-Progressive Bi-Level Image Compression," Dec. 1992.
- (16) ISO/IEC 14492, International standard : "Information Technology - Coded Representation of Picture and Audio Information - Lossy/Lossless Coding of Bi-Level Images," 1999.
- (17) ISO/IEC 10918-1, International standard : "Information Technology - Digital Compression and Coding of Continuous Still Images," 1994.
- (18) ISO/IEC 15444-1, International standard : "Information Technology - JPEG 2000 Image Coding System - Part 1: Core Coding System," 2000.

- (19) 石川孝明, 渡辺裕, デジタルシネマを想定した解像度変換システム, 情報処理学会研究報告 Vol.2005, No.23 (AVM-48), pp.7 - 12, Mar. 2005.
- (20) ITU-T Recommendation T. 44, "Mixed Raster Content," Apr. 1999.
- (21) D. Mukherjee, C. Chrysafis, and A. Said, "Low Complexity Guaranteed Fit Compound Document Compression," IEEE Transactions on Image Processing, Vol. 1, pp.22-25, Sep. 2002.
- (22) ISO/IEC 15444-6, International standard, "Information Technology - JPEG 2000 Image Coding System - Part 6 : Compound Image File Format," Dec. 2002.
- (23) J. Reichel, H. Schwarz, and M. Wien, "Joint Scalable Video Model JSVM-6," ISO/IEC JTC1/SC29/WG11, N8015, Apr. 2006.
- (24) Adobe Developer Support, "Encapsulated PostScript File Format Specification," May. 1992.
- (25) "Scalable Vector Graphics (SVG) XML Graphics for the Web," <http://www.w3.org/Graphics/SVG/>, 2005.
- (26) "Macromedia Flash File Format (SWF) Specification," <http://www.macromedia.com/software/flash/open/licensing/fileformat/>
- (27) A.S.Incorporated, PostScript Language Reference (3rd Edition), Addison Wesley Professional, 1999.
- (28) "Adobe Illustrator," <http://www.adobe.co.jp/products/illustrator/main.html>
- (29) J.Seward, "bzip2, ver.1.0.2." <http://sources.redhat.com/bzip2/>
- (30) Y. Yamamoto, K. Kawamura, and H. Watanabe, "Bit Rate Reduction of Vector Representation of Binary Image," IEEE International Conference on Image Processing (ICIP2006), Oct. 2006.
- (31) "Adobe Illustrator 新機能," <http://www.adobe.co.jp/products/illustrator/newfeatures.html>
- (32) 高木幹雄, 下田陽久, 画像解析ハンドブック, 東京大学出版, 1991.
- (33) 貴家仁志, マルチレート信号処理, 昭晃堂, 1995.
- (34) R. E. Crochiere, and L. R. Rabiner, Multirate Digital Signal Processing, Prentice-Hall, 1983.
- (35) N. A. Dodgson, "Image Resampling," Technical Reports, UCAM-CL-TR-261, University of Cambridge, Aug. 1992.

- (36) A. Koschan, and M. Abidi, "Detection and Classification of Edges in Color Images," IEEE Signal Processing Magazine, Vol. 22, No. 1, pp. 64 - 73. Jan. 2005.
- (37) D. Ziou, and S. Tabbone, "Edge Detection Techniques - An Overview," International Journal of Pattern Recognition and Image Analysis, Vol. 8, No. 4, pp. 537 - 559, 1998.
- (38) J. B. Martens, "Local Orientation Analysis in Images by Means of the Hermite Transform," IEEE Transactions on Image Processing, Vol. 6, No. 8, pp. 1103 - 1116, Aug. 1997.
- (39) G. Chen, and Y. Yang, "Edge Detection by Regularized Cubic B-Spline Fitting," IEEE Transactions on Systems, Man and Cybernetics, Vol. 25, No. 4, pp. 636 - 643, Apr. 1995.
- (40) M. C. Morrone, and R. A. Owens, "Feature Detection from Local Energy," Pattern Recognition Letters, Vol. 6, No. 5, pp. 303 - 313, Dec. 1987.
- (41) W. Y. Ma, and B.S. Manjunath, "EdgeFlow : A Technique for Boundary Detection and Image Segmentation," IEEE Transactions on Image Processing, Vol. 9, pp. 1375 - 1388, Aug. 2000.
- (42) 南畝, 中村納, 画像工学 -画像のエレクトロニクス- コロナ社, Aug. 1989.
- (43) 土屋裕, 深田陽司, 画像処理, コロナ社, Jan. 1990.
- (44) J. Canny, "A Computational Approach to Edge Detection," IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 8, No. 6, pp. 679 - 698, Nov. 1986.
- (45) "Edges:The Canny Edge Detector," <http://homepages.inf.ed.ac.uk/rbf/CVonline/LOCALCOPIES/MARBLE/low/edges/canny.htm>
- (46) K. Tombre, C. Ah-Soon, Ph. Dosch, G. Masini, and S. Tabbone, "Stable and Robust Vectorization : How to Make the Right Choices," Graphics Recognition Recent Advances : Third International Workshop (GREC'99), pp. 3 - 18, Sep. 1999.
- (47) 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.
- (48) 森浩一, 和田耕一, 寅市和男, "関数化図形表現を用いた紙文書のデジタル化," 情報処理学会デジタル・ドキュメント研究報告, 99 - DD - 19, pp. 17 - 23, Jul. 1999.
- (49) H. M. Yang, J. J. Lu, and H. J. Lee, "A Bezier Curve-Based Approach to Shape Description for Chinese Calligraphy Characters," Sixth International Conference on Document Analysis and Recognition, pp. 276 - 280, Sep. 2001.

- (50) U. E. Ramer, "An Iterative Procedure for the Polygonal Approximation of Plane Curves," *Computer Graphics and Image Processing*, Vol. 1, pp. 224 - 256, 1972.
- (51) C. L. B. Jordan, and T. Ebrahimi, "Progressive Polygon Encoding of Shape Contours," *Sixth International Conference of Image Processing and Its Applications*, Vol. 1, pp. 14 - 17, Jul. 1997.
- (52) A. Pinheiro, and M. Ghanbari, "Scalable Coding of Shape Contours in Scale Space," *IEEE International Conference on Image Processing (ICIP2002)*, Vol. 1, pp. 165 - 168, 2002.
- (53) G. Schuster, and A. Katsaggelos, "An Optimal Segmentation Encoding Scheme in the Rate Distortion Sense," *IEEE International Symposium on Circuits and Systems (ISCAS1995)*, Vol. 2, pp. 640 - 643, May. 1995.
- (54) A. Pinheiro, and M. Ghanbari, "A Fast Near-optimal Algorithm for Approximation of Polygonal Curves," *16th International Conference on Pattern Recognition*, Vol. 4, pp. 335 - 338, 2002.
- (55) P. Selinger, "Potrace : A Polygon-Based Tracing Algorithm," <http://potrace.sourceforge.net/potrace.pdf>, 2003.
- (56) ITU-T Recommendation J. 144, "Objective Perceptual Video Quality Measurement Techniques for Digital Cable Television in the Presence of a Full Reference," Jan. 2001.
- (57) ITU-T/COM9-C60-E, "Final Report from the Video Quality Experts Group on the Validation of Objective Models of Video Quality Assessment," FR - TV2, Sep. 2003.
- (58) 川田亮一, "デジタル動画像の画質評価 2. デジタル動画像画質評価の国際標準化動向," *映像情報メディア学会誌*, Vol. 59, No. 8, pp. 1136 - 1139, Aug. 2005.
- (59) E. P. One, W. Lin, Z. Lu, S. Yao, and X. Yang, "Low Bit Rate Video Quality Assessment based on Perceptual Characteristics," *Proceedings of International Conference on Image Processing (ICIP2003)*, Vol. 3, pp. 189 - 192, Sep. 2003.
- (60) Y. Itoh, "An Edge-Oriented Progressive Image Coding," *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 6, No. 2, pp. 135 - 142, Apr. 1996.
- (61) R. Haruki, and T. Horiuchi, "Scalable Image Coding by Spline Approximation for a Gray-scale Image," *Proceedings of the 4th International Conference on Document Analysis and Recognition*, Vol. 1, pp. 407 - 411, 1997.

- (62) X. Xiaohui, and W. Xiaolin, "Directly Operable Image Representation of Multiscale Primal Sketch, " *IEEE Transactions on Multimedia*, Vol. 7, No. 5, pp. 805 - 816, Oct. 2005.
- (63) G. Fan, and W. K. Cham, "Model-based Edge Reconstruction for Low Bit-rate Wavelet-compressed Images, " *IEEE Transactions on Circuits System Video Technology*, Vol. 10, No. 1, pp. 120 - 132, Feb. 2000.
- (64) T. S. Rzeszewski, "Method for Improving Low-Resolution Video on Higher-Resolution Displays, " *IEEE Transactions on Consumer Electronics*, Vol. 45, No. 4, pp. 1030 - 1037, Nov. 1999.
- (65) L. Vese, and S. Osher, "Modeling Textures with Total Variation Minimization and Oscillating Patterns in Image Processing, " *Journal of scientific computing*, No. 19, pp. 553 - 572, 2003.
- (66) S. Osher, L. Vese, and A. Sole, "Image Decomposition, Image Restoration, and Texture Modeling using Total Variation Minimization and the  $H^{-1}$  Norm, " *Proceedings of IEEE International Conference on Image Processing (ICIP2003)*, Vol. 1, pp. 689 - 692, 2003.
- (67) 齊藤隆弘, 石井勇樹, 中川陽介, 小松隆, "骨格/テクスチャ分離に基づく画像拡大, " 2006 年 電子情報通信学会総合大会, D-11-90, Mar. 2006.
- (68) N. Dyn, D. Levin, and S. Rippa, "Data Dependent Triangulation for Piecewise Linear Interpolation, " *IMA Journal of Numerical Analysis*, Vol. 10, No. 1, pp. 137 - 154, 1990.
- (69) S. Battiato, G. Gallo, and G. Messina, "SVG Rendering of Real Images using Data Dependent Triangulation, " In *Proceedings of the 20th Spring Conference on Computer Graphics (ACM/SCCG2004)*, pp. 185 - 192, 2004.
- (70) S. Lee, "Wavelet-Based Multiresolution Surface Approximation from Height Fields, " Ph. D. Thesis, Virginia Polytechnic Institute and State University, Feb. 2002.
- (71) S. Battiato, G. Barbera, G. Di Blasi, G. Gallo, and G. Messina, "Advanced SVG Triangulation Polygonalization of Digital Images, " In *Proceedings of SPIE Electronic Imaging 2005 Internet Imaging VI - Vol. 5670.1*, pp. 1 - 11, Jan. 2005.
- (72) S. Battiato, A. Costanzo, G. Di Blasi, G. Gallo, and S. Nicotra, "SVG Rendering by Watershed Decomposition, " In *Proceedings of SPIE Electronic Imaging 2005 - Internet Imaging VI - Vol. 5670.3*, pp. 23 - 32, Jan. 2005.

- (73) R. C. Gonzales, and R. E. Woods, Digital Image Processing - Second Edition, Prentice Hall, 2002
- (74) L. Vincent, and O. Soille, "Watersheds in Digital Spaces: an Efficient Algorithm based on Immersion Simulations," IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 13, No. 6, pp. 583 - 598, Jun. 1991.
- (75) KVec, "Raster Vector Conversion," <http://www.kvec.de>, 2005.
- (76) Ras2Vec, "Raster to vector conversion program," <http://xmailserver.org/davide.html> 2005
- (77) VectorEye, "Raster to Vector Converter," <http://www.siame.com/index.html> 2005.
- (78) H. Freeman, "Computer Processing of Line-Drawing Images," ACM Computing Surveys, Vol. 6, No. 1, pp. 57 - 97, Mar. 1974.



## 図一覽

2-1	Applications of display (Reference [12])	8
2-2	Field angle (Reference [12])	9
2-3	Eye sight (Reference [12])	10
2-4	Rate-Distortion curve (JPEG, JPEG2000)	17
2-5	JPEG2000 Encoder	18
2-6	A flow of conventional method	20
2-7	Coding efficiency of vector coding	23
2-8	Coding efficiency for binary images	24
3-1	Bi-cubic method	28
3-2	Decimator	29
3-3	Edge type	31
3-4	Prewitt operator	32
3-5	Sobel operator	32
3-6	Laplacian operator	32
3-7	Edge detection by differential operator	35
3-8	Skelton-texture separation for animation	40
3-9	Skelton-texture separation for graphics (Lena)	41
3-10	Rate-Distortion curve of SVG Rendering	43
3-11	Modified images by SVG Rendering	43
4-1	A flow of Superscription Model	46
4-2	Superscription of edge	47
4-3	Smoothing method	47
4-4	A flow of proposed method	49
4-5	Luminance value of each component	51
4-6	Freeman chaincode	52
4-7	Acquisition of luminance values	52
4-8	Rasterization	53
5-1	Modified images of Superscription Model	56
5-2	Evaluation of Superscription Model	57
5-3	Modified images (animation)	59

5-4	Evaluation method of vectorization . . . . .	60
5-5	Resolution conversion . . . . .	62
5-6	Reconstructed images . . . . .	64
5-7	Rate-Distortion of proposed method (Lena) . . . . .	65
5-8	Modified images (Lena) . . . . .	66

## 表一覽

2-1	Characteristics of binary image coding . . . . .	14
2-2	Characteristics of JPEG2000 . . . . .	17
2-3	Characteristics of vector coding . . . . .	21
2-4	Scanning condition . . . . .	23
3-1	Evaluation of SVG Rendering (animation) . . . . .	44
5-1	Experiment condition of Superscription Model . . . . .	55
5-2	Bit rate of Superscription Model . . . . .	57
5-3	Experiment condition . . . . .	58
5-4	Evaluation of rasterized image . . . . .	60
5-5	Bit rate of edge components . . . . .	61
5-6	9/7 Daubechies filter coefficients . . . . .	62
5-7	PSNR of reconstructed images (Lossless coding) . . . . .	63
5-8	PSNR of reconstructed images (Lossy coding) . . . . .	63
5-9	Bit rate of each component (animation) . . . . .	63
5-10	Bit rate of each component (Lena) . . . . .	64

## 研究業績

	題 名	発表年月	発表掲載誌	連名者
海外				
(1)	Gradation Approximation for Vector Based Compression of Comic Images	2005年9月	IEEE International Conference on Image Processing (ICIP2005)	K.Kawamura <b>Y.Yamamoto</b> H.Watanabe
(2)	A Study on Spatial Scalable Coding Using Vector Representation	2006年7月	IEEE International Conference on Multimedia & Expo (ICME2006)	<b>Y.Yamamoto</b> K.Kawamura H.Watanabe
(3)	Bit Rate Reduction of Vector Representation of Binary Image	2006年10月	IEEE International Conference on Image Processing (ICIP2006)	<b>Y.Yamamoto</b> K.Kawamura H.Watanabe
国内				
(4)	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	2004年11月	PCSJ2004 画像符号化シンポジウム	山本 勇樹 河村 圭 渡辺 裕 富永英義
(5)	ベクター表現における符号量制御に関する検討	2004年11月	PCSJ2004 画像符号化シンポジウム	河村 圭 山本 勇樹 渡辺 裕
(6)	ベクター表現の階層符号化に関する検討	2004年12月	情報処理学会研究報告 第47回オーディオビジュアル複合処理研究会	河村 圭 山本 勇樹 渡辺 裕

	題 名	発表年月	発表掲載誌	連名者
(7)	通過点削減によるベクトル表現の符号化効率改善	2005年3月	情報処理学会 第67回総合大会	山本 勇樹 河村 圭 渡辺 裕 富永英義
(8)	ベクター変換における符号化効率の改善	2005年3月	情報処理学会 第67回総合大会	河村 圭 山本 勇樹 渡辺 裕
(9)	ベクトル化した二値画像の画質評価に関する検討	2005年3月	情報処理学会研究報告 第48回オーディオビジュアル複合処理研究会	山本 勇樹 河村 圭 渡辺 裕 富永英義
(10)	ベクター変換における階層符号化の改良	2005年3月	情報処理学会研究報告 第48回オーディオビジュアル複合処理研究会	河村 圭 山本 勇樹 渡辺 裕
(11)	二値画像のベクトル表現における符号化効率改善	2005年3月	電子情報通信学会 2004年総合大会	山本 勇樹 河村 圭 渡辺 裕 富永英義
(12)	ベクター表現における階層的変換手法の検討	2005年3月	電子情報通信学会 2004年総合大会	河村 圭 山本 勇樹 渡辺 裕
(13)	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	2005年9月	第四回情報科学技術フォーラム (FIT2005)	山本 勇樹 河村 圭 渡辺 裕
(14)	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	2005年9月	第四回情報科学技術フォーラム (FIT2005)	河村 圭 山本 勇樹 渡辺 裕
(15)	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	2005年11月	PCSJ2005 画像符号化シンポジウム	山本 勇樹 河村 圭 渡辺 裕

	題 名	発表年月	発表掲載誌	連名者
(16)	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	2005年11月	PCSJ2005 画像符号化シンポジウム	河村 圭 山本 勇樹 渡辺 裕
(17)	ベクトル表現を用いた空間スケーラブル符号化に関する一検討	2005年12月	情報処理学会研究報告 第51回オーディオビジュアル複合処理研究会	山本 勇樹 河村 圭 渡辺 裕
(18)	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	2005年12月	情報処理学会研究報告 第51回オーディオビジュアル複合処理研究会	河村 圭 山本 勇樹 渡辺 裕
(19)	イントラ符号化にベクター表現を用いた動画像符号化に関する検討	2006年3月	情報処理学会 第68回総合大会	河村 圭 山本 勇樹 渡辺 裕
(20)	ベクトル表現を用いた空間スケーラブル符号化	2006年3月	電子情報通信学会 2005年総合大会	山本 勇樹 河村 圭 渡辺 裕
(21)	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年3月	電子情報通信学会 2005年総合大会	河村 圭 山本 勇樹 渡辺 裕
(22)	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	2006年9月	第五回情報科学技術フォーラム (FIT2006)	河村 圭 山本 勇樹 渡辺 裕
(23)	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	2006年9月	情報処理学会研究報告 第54回オーディオビジュアル複合処理研究会	河村 圭 山本 勇樹 石井 大祐 渡辺 裕
(24)	ベクター表現によるマンガ符号化システム	2006年11月	PCSJ2006 画像符号化シンポジウム	河村 圭 山本 勇樹 石井 大祐 渡辺 裕

	題 名	発表年月	発表掲載誌	連名者
(25)	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	2007年3月 (発表予定)	情報処理学会 第69回総合大 会	河村 圭 山本 勇樹 石井大祐 渡辺 裕
(26)	エッジ成分のベクトル化 による空間スケーラブル 符号化	2007年3月 (発表予定)	電子情報通信学 会 2006年総 合大会	山本 勇樹 河村 圭 渡辺 裕
(27)	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年3月 (発表予定)	電子情報通信学 会 2006年総 合大会	河村 圭 山本 勇樹 石井大祐 渡辺 裕