

# 博士論文

レート歪み理論に基づいた画像のサブバンド符号化における  
空間-周波数領域分割の最適化に関する研究

Optimum Segmentation of Space-Frequency Domain  
in Subband Image Coding based on the Rate-Distortion Sense

岩手県立大学大学院  
ソフトウェア情報学研究科

宮崎春彦



2017年度博士後期課程(ソフトウェア情報学)論文

レート歪み理論に基づいた画像のサブバンド符号化における  
空間-周波数領域分割の最適化に関する研究

Optimum Segmentation of Space-Frequency Domain  
in Subband Image Coding based on the Rate-Distortion Sense

岩手県立大学大学院  
ソフトウェア情報学研究科

宮崎 春彦

2018年3月

主指導教員 亀田 昌志

副指導教員 伊藤 慶明

プリマ オキ ディッキ アルディアンシャー

小嶋 和徳



# 目次

<b>第1章</b>	<b>序論</b>	<b>1</b>
1.1	本研究の背景	1
1.1.1	画像符号化の歴史と背景	1
1.1.2	画像のサブバンド符号化	2
1.1.3	Wavelet Packet	4
1.1.4	画像のサブバンド符号化における最適帯域分割	5
1.2	本研究の目的	9
1.3	本論文の構成	11
<b>第2章</b>	<b>レート歪み理論に基づいた画像のサブバンド符号化における最適帯域分割</b>	<b>13</b>
2.1	序言	13
2.2	レート歪み理論に基づいた帯域分割と量子化の最適化	14
2.2.1	最適化問題の定式化	14
2.2.2	帯域分割と量子化の導出アルゴリズム	17
2.3	シミュレーション実験とその結果	21
2.4	結言	36
<b>第3章</b>	<b>レート歪み理論に基づいた画像のサブバンド符号化における空間-周波数領域の最適分割</b>	<b>37</b>
3.1	序言	37
3.2	レート歪み理論に基づいた周波数帯域分割・量子化・空間領域分割の最適化	39
3.2.1	周波数帯域分割の最適化	41
3.2.2	量子化の最適化	42
3.2.3	空間領域分割の最適化	43
3.2.4	アルゴリズム	45
3.3	シミュレーション実験とその結果	47
3.3.1	符号化性能	47

3.3.2	処理コスト	73
3.4	結言	75
<b>第4章</b>	<b>視知覚特性を考慮した画像のサブバンド</b>	
	符号化における空間-周波数領域の最適分割	<b>77</b>
4.1	序言	77
4.2	視覚の空間周波数特性と画像の顕著性を 考慮した重み付け関数の導出	79
4.2.1	視覚の二次元空間周波数特性による重み付け	80
4.2.2	Saliency Map による重み付け	83
4.3	視覚の空間周波数特性と画像の顕著性を 考慮した空間-周波数領域の最適分割	85
4.3.1	注目領域 (ROI) と非注目領域 (non-ROI) の分割	86
4.3.2	注目領域 (ROI) と非注目領域 (non-ROI) を考慮した 量子化の最適化	87
4.4	シミュレーション実験とその結果	90
4.4.1	周波数帯域分割・量子化・空間領域分割の結果	93
4.4.2	主観評価実験による画質評価	100
4.5	結言	108
<b>第5章</b>	<b>結論</b>	<b>109</b>
5.1	本研究で得られた成果	109
5.2	今後に残された課題	112
	謝辞	<b>113</b>
	参考文献	<b>115</b>
	本研究に関する研究業績	<b>121</b>

# 第 1 章

## 序論

### 1.1 本研究の背景

インターネットやマルチメディアコンテンツの急速な発展及び普及に伴い、膨大な量の静止画像や動画画像がデジタルデータとして流通している。デジタル化された画像情報そのものは、膨大な情報量を持つため、これらを効率よく処理、蓄積及び伝送するためには、画像情報の圧縮、すなわち、画像符号化の技術が必要不可欠となる。

#### 1.1.1 画像符号化の歴史と背景

画像符号化は、まず、波形符号化の技術を基盤にその研究が始められた。波形符号化とは、画像信号を不規則な二次元波形信号と見なした上で統計的な確率モデルを構築し、そのモデルに基づいて画像信号に含まれる統計的冗長性を取り除くことで情報圧縮を行う方式である。予測符号化 (Differential PCM) [1]～[3] は、波形符号化の中では最も歴史の古いものであり、画像の空間的あるいは時間的に近接する画素値の相関を利用し、符号化対象の画素値とその予測値との差分を符号化することで冗長性の除去を行う符号化方式である。ベクトル量子化 (Vector Quantization) [4]～[6] もまた、予測符号化と同様、画像の空間領域において冗長性の除去を行う符号化方式であるが、画素単位ではなく、 $N$  個の画素の情報を  $N$  次元ベクトルとして構成することで、より効率的に冗長性の除去を達成することができる。一方、直交変換符号化 (Orthogonal Transform Coding) [7]～[9] では、画像の一部あるいは全体を空間領域上の多次元ベクトルとみなし、座標軸の線形変換を行うことで特定の出力係数に信号電力を集中させ、この信号電力の偏りを利用することで冗長削減を実現している。このとき、変換後の領域が周波数領域となるものは、広義としてサブバンド符号化 (Subband Coding) [10]～[13] と呼ばれており、一般的な自然画像のように低域側に信号電力が集中する自己相関性の高い信号に対して、特に有効であることが知られている。

これら波形符号化は、現在においても画像符号化技術の中核を担っており、中でも特に、現在の画像符号化の主流となっているのは、画像符号化の国際標準方式である静止画像用の

JPEG [14], [15] や動画用の MPEG [16], [17] において、その技術の根幹を支えている直交変換符号化の一方式である DCT (Discrete Cosine Transform) [18], [19] に他ならない。国際標準に DCT が採用された理由は、圧縮効率に直接影響する低周波成分への電力集中の観点で最適な変換である Karhunen-Loève 変換 [20] とほぼ同等の性能を有することと、高速演算アルゴリズムが存在する [21] ことによるものである。DCT をベースとした符号化方式では、通常、画像を小領域のブロックに分割し、DCT によって各ブロック信号を一旦空間領域から周波数領域に変換した後、それぞれを量子化することによって情報量を削減する。このとき、JPEG や MPEG において行われる画像のブロック分割では、画像全体が  $(8 \times 8)$  画素の固定幅のブロックで均等に分割される。

この従来のブロック分割方式から一進して、2003 年に策定された国際標準の動画符号化方式 H.264/AVC [22], [23] では、対象となる画像領域の内容に応じてブロックサイズを  $(4 \times 4)$  または  $(8 \times 8)$  の 2 種類から選択できるようになっており、平坦領域等の画像内容が単純な部分に対しては大きいサイズのブロック、エッジやテクスチャ等の画像内容が複雑な部分に対しては小さいサイズのブロックが割り当てられるように、入力画像の内容を考慮して適応的にブロック分割を行うことによって、従来の MPEG-2 よりも約 2 倍となる高い圧縮効率を実現している。さらに、近年において国際標準化された動画符号化方式 H.265/HEVC [24], [25] では、ブロックサイズの候補として、更に  $(4 \times 4)$ ,  $(8 \times 8)$ ,  $(16 \times 16)$ ,  $(32 \times 32)$  の 4 種類が用意されており、入力画像の内容に応じて、より柔軟なブロック分割が可能になったことで、H.264/AVC よりも更に 2 倍の圧縮効率を得られることが報告されている。この可変ブロック型 DCT という考え方は、今から約 30 年前に既に議論されていたものであるが [26], [27]、今になってこの考え方が国際標準方式に採用された背景として、従来よりもハードウェアの性能が著しく向上したことが大きく関わっている。さらに、ムーアの法則 [28], [29] に従えば、ハードウェアの性能は今後も飛躍的に向上し続けると予想される。このことから、今後更にハードウェアの性能が向上し、処理コストの観点から従来においては実現が困難であった処理が将来的に可能となっていくことを考えたとき、画像符号化において、上述した可変ブロック型 DCT にあるような「入力画像の内容に対する適応的な処理」という考え方は、今後更に高い圧縮効率を実現するためには必ず考慮しなくてはならない必要不可欠なものになると考えている。

## 1.1.2 画像のサブバンド符号化

一方、JPEG の後継規格として、JPEG よりも優れた圧縮効率を実現できる国際標準方式に JPEG2000 [30]~ [32] があり、その基幹技術には、DCT に代わって高能率符号化の一手法であるサブバンド符号化 [10]~ [13] が採用されている。サブバンド符号化は、図 1.1 に示すように、画像全体に対してフィルタバンクと呼ばれる帯域通過フィルタを適用することで、画像を空間領域上ではなく、周波数領域上でいくつかの帯域 (サブバンド) に分割した後、分

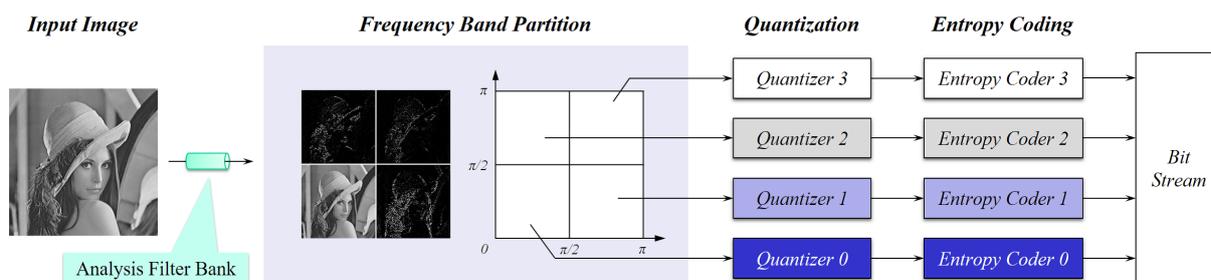


図 1.1 サブバンド符号化の符号化構成

割されたサブバンドごとに量子化及びエントロピー符号化を適用することで情報圧縮を実現する。DCTに基づいた符号化方式では、画像を空間領域上でブロック分割してから処理を行うため、再生画像にブロック歪みと呼ばれる格子状のアーティファクトが発生するという大きな問題があるが [33]、サブバンド符号化では、処理の特性上、原理的にそれが生じない。このブロック歪みが発生しない利点に加え、特に低符号化レートで DCT よりも高い圧縮性能を有することが、JPEG2000 にサブバンド符号化が採用された理由である [32]。しかしながら、JPEG2000 は高性能な画像符号化方式であるものの、JPEG と比べて大きな処理コストが要求される。この主な要因は、変換処理に際して、サブバンド符号化によるメモリの使用量が DCT よりも大きいというシステム構成上の根本的な問題によるものである。この処理コストの問題に起因して、JPEG2000 は、2001 年に策定されてから今日に至るまであまり普及には至っておらず、同様の理由から、その基幹技術であるサブバンド符号化をベースとした国際標準の新しい画像符号化方式は、それ以降提案されていない。しかしながら、1.1.1 で述べたように、今後期待される更なるハードウェアの性能向上に伴って、現在よりも大幅な高速処理が実現可能となった場合、画像符号化においては圧縮性能の高さがより重要視されることが予想されるため、将来的には、DCT よりも純粋に性能が高いサブバンド符号化に再び注目が集まり、サブバンド符号化に基づいた新しい画像符号化方式が開発され始められると考えている。

サブバンド符号化における情報圧縮、すなわち、冗長性の除去の基本原理は、特定のサブバンドに信号電力を集中させるための帯域分割と、分割後の各サブバンド信号に対する量子化及びエントロピー符号化にある。このうち、後者については、Coding Gain [34], [35] を最大とするように各サブバンドの信号電力に応じて最適なビット配分を行う手法 [36] や、レート歪み理論 [37], [38] の観点で最適なビット配分を行う手法 [39], [40] が提案されており、これらによって各サブバンドに対する適切な符号化ビットの配分量を決定することができる。一方、前者の帯域分割については、図 1.1 に例を示すように、画像の二次元周波数帯域を均等な帯域幅を有する 4 つのサブバンドに分割する方式が、サブバンド符号化の基本方式として用いられている。JPEG2000 では、このサブバンド符号化の基本方式の拡張とみなされる離散ウェーブレット変換 [41]～ [43] が採用されており、この方式では、一般的な自然画像

が低域側に信号電力が集中することを利用して、最低域のサブバンドを再帰的に分割する多重解像度分解により効率的な情報圧縮を実現している。すなわち、JPEG2000のように離散ウェーブレット変換に基づいた画像符号化方式では、自然画像が持つ統計的性質に整合した帯域分割を行うことによって高い圧縮性能を達成している。しかしながら、エッジやテクスチャ等の高周波成分が多く含まれるような自然画像の統計的性質をあまり満たさない入力画像が与えられた場合、離散ウェーブレット変換では高域側のサブバンドに対する分割が行われないため、その冗長性を取り除くことができず、圧縮性能を十分に改善することができない。したがって、サブバンド符号化において高い圧縮性能を実現するためには、入力画像が持つ固有の性質に整合して周波数帯域を適応的に分割することが重要であり、それはすなわち、離散ウェーブレット変換に基づいた現行の JPEG2000 についても、その圧縮性能を根本から改善できるという余地が未だ十分に残されていることを示唆している。H.265/HEVC 及び H.264/AVC が、ハードウェアの性能が従来よりも飛躍的に向上したことを踏まえて、DCT のブロック分割方式を固定型から可変型に切り替えたことで高い圧縮効率を実現したように、次世代画像符号化は、離散ウェーブレット変換のような帯域を固定の分割パターンにより分割する従来のサブバンド符号化方式に代わって、入力画像が持つ固有の性質に応じて帯域を分割する「適応型サブバンド符号化」により実現されると予想しており、そしてそれは、いずれ将来的な画像符号化の発展に大いに寄与する存在になるであろうと考えている。

### 1.1.3 Wavelet Packet

入力画像が持つ固有の性質に整合して適応的な帯域分割を実現するサブバンド符号化方式として、代表的なものに Wavelet Packet [44], [45] がある。Wavelet Packet は、多重解像度分解を一般化した方式であり、入力画像が持つ固有の性質を考慮して、低域のサブバンドのみならず、高域のサブバンドを含む全てのサブバンドを対象とした適応分割が行われる。Wavelet Packet は、まず、一旦入力画像の二次元周波数帯域を、四分木に基づいて指定された分解レベルの深さまで再帰的に分割した後、分割前のサブバンドのコスト関数と分割後に得られた4つのサブバンドのコスト関数の和をボトムアップ的に比較し、数多くある分割パターンの中からコストを最小とする分割パターンを探し求めることで実現される。分割パターンが決定された後は、各サブバンド信号に対して算術符号化 [46] 等のエントロピー符号化器が設計され、それぞれがビットストリームに出力される。Wavelet Packet の草創期において、R.R. Coifman らは、閾値処理された後の非ゼロ係数の個数や、レートまたは歪みの一方をコスト関数として最適な分割パターンを求める手法を提案している [44]。また、K. Ramchandran らは、画像符号化への実用を目的として、レート歪み理論の観点で最適な分割パターンと各サブバンドに適用する最適な量子化器を同時に導出する手法を提案している [45]。今日までに提案されてきた Wavelet Packet に基づく各種の画像符号化方式は、文献 [44] と [45] の手法をベースに発展してきたものとみなされる（例えば [47]～[54]）。

しかしながら，Wavelet Packet では，(1) 最適な分割パターンの導出と，(2) 各サブバンドに対して独立に適用されるエントロピー符号化器の設計に大きな処理コストが必要になるという欠点がある．前者の(1)における処理コストの問題を解決するための手法として，二次元 non-separable フィルタを使用して，繰り返し行われる四分木に基づいたサブバンドの再帰分割に必要な畳み込み演算の処理コストを削減するもの [55] や，各サブバンドへの再帰分割と同時に，最適な分割パターンの探索をトップダウン的に行うことで，分割パターンの自由度を制限する代わりに，畳み込み演算及び探索の処理コストを抑制するもの [56]～[58] 等が提案されている．一方，後者の(2)について，エントロピー符号化器の設計に要する処理コストは，決定された分割パターンのサブバンド数と直結しているため，二次元周波数帯域を離散ウェーブレット変換よりもはるかに多い数のサブバンドに分割することで圧縮性能の改善を実現する Wavelet Packet では，根本的にこの問題を解決することはできない．このとき，エントロピー符号化には，情報源全体の確率分布から事前に各シンボルに対する符号語を定義して割り当てる静的なものと，各シンボルに符号語を割り当てながら確率分布を逐次更新する動的なものがある．前者は，小さい処理コストでの実現が可能という利点がある一方で，サイド情報として確率分布の情報を復号化側に伝送する必要があるため，サブバンドの数に比例してサイド情報量が増大する．一方，後者は，復号化側の確率分布を使用するためにサイド情報を伝送する必要がないものの，確率分布の逐次更新を行わなくてはならないため，大きな処理コストを必要とする [12]．したがって，各サブバンドに対してエントロピー符号化を独立に適用するというサブバンド符号化の基本構成の下で，処理コストとサイド情報の両者を同時に低減可能な画像符号化を実現するためには，少数のサブバンドに対して静的なエントロピー符号化を適用するより他に方法はないと言える．

### 1.1.4 画像のサブバンド符号化における最適帯域分割

ハードウェア性能の著しい向上により，今後も更なる処理の高速化が予想されるが，適応型サブバンド符号化による次世代画像符号化の早期実現にあたっては，適応的な帯域分割をできる限り少ない処理コストにより実現することが望ましい．しかしながら，Wavelet Packet ではその分割の特性上，処理コストを低く抑えることは極めて困難である．そこで，1.1.3 の(2)に挙げた，サブバンド数の増加に伴ってエントロピー符号化器の設計に大きな処理コストが必要になるという Wavelet Packet 固有の問題に対して，あらかじめ少数のサブバンド数を指定した，すなわち，エントロピー符号化器設計の処理コストを低く抑えるとした条件下で，入力画像に対して二次元周波数帯域の分割パターンを適応的に決定する最適帯域分割が提案されている [59]～[61]．最適帯域分割は，サブバンド数  $M$  が指定されたとき，符号化後の再構成画像に含まれる量子化雑音電力を最小とすることを目的として，入力画像  $x(m, n)$  の二次元周波数帯域  $\Omega = \{(\omega_h, \omega_v) \mid 0 \leq \omega_h, \omega_v \leq \pi\}$  を互いに素な  $M$  個のサブバンド  $\Omega_k (k = 0, 1, \dots, M-1)$  に直和分割する方式である．このとき，入力画像を直接スカラー量子化

した場合（PCM 符号化）に生じる量子化雑音電力と，帯域分割処理後の各サブバンド信号をスカラ量子化した場合に生じる総量子化雑音電力の比によって定義される量子化雑音改善量

$$G_V = 10 \log_{10} \left( \frac{\sum_{k=0}^{M-1} \sigma_k^2}{\prod_{k=0}^{M-1} (\sigma_k^2 / \lambda_k)^{\lambda_k}} \right) [dB] \quad (1.1)$$

を評価値として導入することで，最適帯域分割は，サブバンド数が  $M$  である条件で  $G_V$  を最大とする帯域分割方式としてみなされる．ここで，式 (1.1) における  $\sigma_k^2$  と  $\lambda_k$  は，それぞれ式 (1.2) で表されるサブバンド  $\Omega_k$  の信号電力及び信号レートである．

$$\left. \begin{aligned} \sigma_k^2 &= \frac{1}{\pi^2} \iint_{\Omega_k} P_x(\omega_h, \omega_v) d\omega_h d\omega_v \\ \lambda_k &= \frac{1}{\pi^2} \iint_{\Omega_k} d\omega_h d\omega_v \end{aligned} \right\} (k = 0, 1, \dots, M-1) \quad (1.2)$$

ここで， $P_x(\omega_h, \omega_v)$  は入力信号の電力スペクトルである．なお，量子化雑音改善量  $G_V$  は Coding Gain [34], [35] と同意のものである．このとき，任意の帯域分割特性を持つフィルタバンクの設計は極めて困難であることから，文献 [60] では，QMF [62] 等の既存の帯域分割フィルタを用いて，二次元周波数帯域  $\Omega$  を帯域ブロック  $\Delta\Omega_i (i = 0, 1, \dots, N-1; N \geq M)$  と呼ばれる均等な帯域幅を持つ小領域の集合に一旦分割した後， $G_V$  の値が最大となるように， $N$  個の帯域ブロック  $\Delta\Omega_i$  を  $M$  個のサブバンド  $\Omega_k$  に分類することで，最適帯域分割を近似的に実現している．図 1.2 に，サブバンド数  $M = 4$ ，帯域ブロック数  $N = 64$  の条件を指定した場合の最適帯域分割の符号化構成を示す．図 1.3 は，SIDBA 標準画像 [63] の中から選択した画像 “Lenna” と画像 “Barbara”（いずれも  $256 \times 256$  画素，濃淡 8[bit/pel]）であり，図 1.4 は，これらを入力画像とした場合に， $M = 5$ ， $N = 64$  の条件で求められた最適帯域分割による二次元周波数帯域  $\Omega$  の分割パターンを示している．ここで，図 1.4 における点線の格子で区切られた小領域が帯域ブロックに対応しており，同じ色の帯域ブロック群で構成された領域をサブバンドとして表している．図 1.4 の結果から，最適帯域分割の適用により，各画像に対

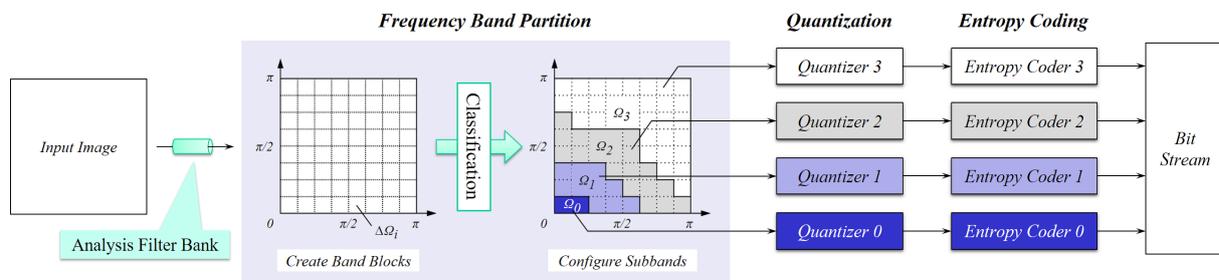


図 1.2 画像のサブバンド符号化における最適帯域分割の符号化構成 ( $M=4$ ,  $N=64$ )

して  $M$  個のサブバンドから成る固有な帯域分割パターンが得られていることがわかる。最適帯域分割を実画像に適用し、入力画像に固有な分割パターンを決定することにより、符号化画像に含まれる量子化雑音電力が低減され、少ないサブバンド数でも符号化性能を改善できることが明らかにされている。

しかしながら、従来の最適帯域分割において求められる分割パターンは、 $G_V$  を最大とするように、量子化前の各サブバンドの信号電力  $\sigma_k^2$  に基づいて決定されるため、符号化レートの変更に対しては不変である。そのため、特に低符号化レートが指定された場合には、量子化レベル数の小さい量子化器が各サブバンドに適用されることとなり、各サブバンドの信号電力が量子化前のものと比べて大きく変化してしまうので、真に最適な分割パターンが得られていないと推測される。また、分割パターンが決定された後は、サブバンドごとにスカラ量子化が行われるが、各量子化器は符号化レートの変更に対して変化するものの、サブバンドごとに自由に選択できるわけではなく、全てのサブバンドに対して同一の量子化器が一律に適用される仕様になっている。サブバンド符号化の性能を改善するためには、量子化による画質劣化への影響を考慮して、各サブバンド信号を適切な量子化器で量子化することが望ましいが、それは従来の最適帯域分割では実現されていない。以上のことから、従来の最適帯域分割によって決定された分割パターン及び各サブバンドに適用される量子化器は、符号化レートが変更された場合においても同様に最適なものであるとは言い難いため、符号化性能の改善が十分に達成されていなかったと考えられる。そのため、符号化レートが指定されることを考慮した上で、帯域分割と量子化の二つの最適化問題について改めて検討する必要がある。

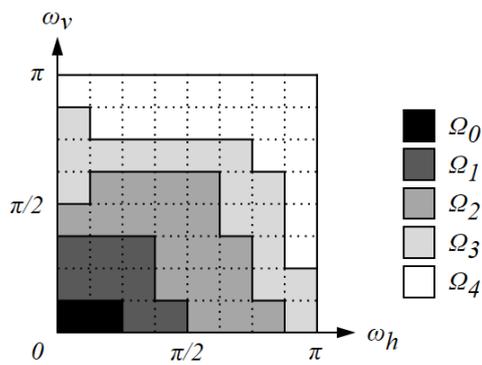


(a) 画像“Lenna”

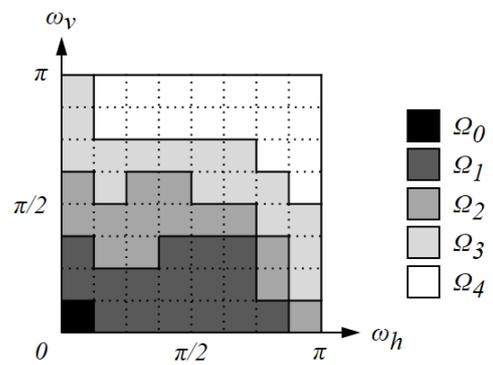


(b) 画像“Barbara”

図 1.3 テスト画像 ((256 × 256) 画素, 濃淡 8[bit/pel])



(a) 画像“Lenna”



(b) 画像“Barbara”

図 1.4 最適帯域分割により求められた二次元周波数帯域上の分割パターン ( $M = 5$ ,  $N = 64$ )

## 1.2 本研究の目的

現在の画像符号化の主流は、静止画像符号化方式 JPEG や、動画画像符号化方式 MPEG, H.264/AVC, H.265/HEVC において、その技術の基礎となっている DCT である。一方で、JPEG の後継規格である JPEG2000 では、サブバンド符号化がその根幹を支えている。サブバンド符号化は、DCT をベースとした画像符号化方式において特有に現れるブロック歪み発生の問題がないことと、特に低符号化レートでの圧縮性能が DCT よりも遥かに上回ることを理由に JPEG2000 に採用されたが、DCT と比べて変換処理への負荷が大きいために、あまり普及には至らなかった。しかし、ここ数年におけるハードウェア性能の著しい向上から、今後更に処理の高速化が実現されていくことを考慮すれば、純粋な圧縮性能の高さ故に、DCT に代わる方式としてサブバンド符号化が次世代画像符号化の中心的役割を担う存在になると考えている。さらに、サブバンド符号化は、上述した利点に加えて、DCT を適用した際には失われてしまう画像の空間領域情報を、変換後の領域において保持できるという大きな利点がある。これを活用すれば、DCT をベースとした場合よりも、更に高能率な画像符号化への発展が期待できる。そこで、本研究では、このサブバンド符号化に着目する。

サブバンド符号化は、画像信号に対してフィルタバンクを適用することにより、入力画像の二次元周波数帯域をいくつかのサブバンドに分割した後、各サブバンドに対して独立に量子化及びエントロピー符号化を適用することで情報圧縮を実現する。このとき、高い符号化性能を得るためには、入力画像が持つ固有の特性を考慮して画像の二次元周波数帯域を適応的に分割することが有効である。サブバンド符号化における適応分割の代表的な手法である Wavelet Packet [44], [45] は、入力画像の特性に応じて各サブバンドを必要なだけ再帰的に分割することで、高い符号化性能を実現できるものの、この再帰分割によって画像の二次元周波数帯域が多数のサブバンドに分割されるため、後の各サブバンド信号に対して独立に行われるエントロピー符号化器の設計に大きな処理コストが必要になるという欠点がある。

この問題に対して、エントロピー符号化器設計の処理コストを低く抑えるために、あらかじめ少数のサブバンド数を指定した条件の下で、入力画像の特性を考慮して二次元周波数帯域を適応的に分割する最適帯域分割が提案されている [59]～ [61]。しかしながら、従来の最適帯域分割は、非帯域分割方式からの量子化雑音電力の低減量として定義される量子化雑音改善量なる評価値を最大とすることを目的として、量子化が行われる前の各サブバンド信号の電力に基づき、指定されたサブバンド数による帯域分割パターンを求めるものであったため、符号化レートが変更されたときにも分割パターンは固定のままであった。また、各サブバンドに適用される量子化器は、符号化レートの変更に対して変化するものの、すべてのサブバンドで同じものが適用されるという仕様になっていた。すなわち、従来の最適帯域分割は、再生画像における画質劣化量に相当する量子化雑音電力のみを考慮して導出されたものであることから、符号化レートをも考慮した帯域分割及び量子化は行われておらず、十分な

符号化性能が達成されていなかった。

そこで、本研究では、従来の最適帯域分割における符号化性能を改善するために、従来の評価値である量子化雑音改善量に代わって、符号化レートと画質劣化量の相互関係を表すレート歪み理論 [37], [38] の観点で、帯域分割及び量子化を最適化する「レート歪み理論に基づいた画像のサブバンド符号化における最適帯域分割」を明らかにすることを目的とする。本論文は、筆者の行ったレート歪み理論に基づいた画像のサブバンド符号化における最適帯域分割の導出において得られた成果をまとめたものである。

まず初めに、レート歪み理論の観点から、二次元周波数帯域の分割パターン及び各サブバンドに適用する量子化器の組合せを最適化するための理論式を明らかにし、さらに、任意の符号化レートが指定されたときに、それら二つの解を同時に導出するアルゴリズムを開発する。そして、レート歪み理論に基づいた本最適帯域分割を実画像に適用し、その符号化性能を従来の最適帯域分割のものと比較することで、本最適帯域分割の有効性を明らかにする。次に、分割された各サブバンドにおいて画像の空間領域情報を保持できるというサブバンド符号化の利点から、周波数領域だけではなく空間領域にも着目し、各サブバンド信号における空間領域の冗長性が十分に除去しきれていないことを指摘した上で、帯域分割と量子化のみならず、不要な信号を切り捨てて必要な信号のみを保存することで冗長削減を実現する空間領域分割の三つを、レート歪み理論の観点で同時に最適化する「空間-周波数領域の最適分割」を提案する。そして、空間-周波数領域の最適分割を実画像に適用したときの符号化性能を、レート歪み理論に基づいた最適帯域分割や Wavelet Packet, 更には画像符号化の国際標準方式 JPEG 及び JPEG2000 のものと比較することで、空間-周波数領域の最適分割の有効性を明らかにする。最後に、視知覚特性を考慮した画質の観点で良好な再生画像を得ることを目的として、人間の視知覚特性の一種である視覚の空間周波数特性と画像の顕著性の二つを考慮した空間-周波数領域の最適分割を提案する。本空間-周波数領域の最適分割により得られた再生画像に対して、その画質を評価するための主観評価実験を行い、通常空間-周波数領域の最適分割及び JPEG2000 のものと比較することで、視知覚特性を考慮した空間-周波数領域の最適分割の有効性を明らかにする。

本研究は、その圧縮性能の高さから次世代画像符号化の根幹を支えるものとして有望視されているサブバンド符号化において、圧縮性能の改善と符号化器の設計コストの低減を同時に達成するために提案されている「最適帯域分割」の符号化性能の改善を目的としたものである。本研究によって得られた成果を、これまでに開発されてきたサブバンド符号化をベースとする全ての画像符号化方式に応用することで、符号化器の設計に必要な処理コストを低く抑えたまま、更なる符号化効率の改善が行われることが期待される。そしてそれは、適応型サブバンド符号化に基づいた次世代画像符号化の実現に際して、大いに寄与するものであると信じている。

## 1.3 本論文の構成

### 第1章：序論

画像符号化における背景と従来研究における問題を示し、本研究の目的と意義を述べる。

### 第2章：レート歪み理論に基づいた画像のサブバンド符号化における最適帯域分割

本研究の基礎となるレート歪み理論に基づいた最適帯域分割の理論的な導出と、その実現方法を明らかにする。まず、レート歪み理論における最適化問題が、ラグランジュの未定乗数法を適用することにより解が求められることを踏まえて、帯域分割と量子化の二つの最適化問題をそれぞれ定式化することで、レート歪み理論の観点で、二次元周波数帯域の最適分割パターンと各サブバンドに適用される最適な量子化器の組み合わせを導出するための理論式を明らかにする。次に、符号化レートが指定されたときに、最適な帯域分割パターン及び各サブバンドに対する最適な量子化器の組み合わせを求めるアルゴリズムを開発する。最後に、レート歪み理論に基づいた最適帯域分割を実画像に適用し、その符号化性能を、従来の最適帯域分割、JPEG2000における帯域分割方式として採用されている離散ウェーブレット変換、帯域分割の理論的な最適解を導出できる Wavelet Packet のものと比較することで、レート歪み理論に基づいた最適帯域分割の有効性を定量的に明らかにする。

### 第3章：レート歪み理論に基づいた画像のサブバンド符号化における空間-周波数領域の最適分割

第2章で提案したものも含めた従来の最適帯域分割では、サブバンド内にある全ての信号が符号化されるため、空間領域上の冗長性を十分に除去できていないと考えられる。そこで、第2章で議論した二次元周波数帯域の分割パターンと各サブバンドに適用される量子化器の組み合わせに加えて、帯域ブロック内の不要な信号を切り捨てて必要な信号のみを符号化することで冗長削減を実現する二次元空間領域分割の三つを、レート歪み理論の観点で最適化する空間-周波数領域の最適分割を提案する。このとき、空間-周波数領域の最適分割の実現にあたってサブバンド数が増えることはないため、エントロピー符号化器の設計に要する処理コストの増加は生じない。まず、三つの同時最適化は困難であることを指摘した上で、ある一つの最適化問題を解く際には残り二つの最適化問題は既に解かれていると仮定した上で、それぞれ三つの最適化問題を個別に定式化することにより、新しく空間領域分割が導入されたことを考慮した帯域分割及び量子化の最適解を導出するための理論式と、空間領域分

割の最適解を導出するためのアルゴリズムを明らかにする。次に、符号化レートが指定されたときに、最適な帯域分割パターン、各サブバンドに対する最適な量子化器の組み合わせ、そして、各帯域ブロックにおける二次元空間領域上の最適分割パターンを求めるアルゴリズムを開発する。最後に、実画像に対して空間-周波数領域の最適分割を適用したときの符号化性能を、第2章で提案したレート歪み理論に基づいた最適帯域分割、Wavelet Packet、提案手法と同様に空間領域上の冗長削減を実現する既存手法である Space-Frequency Quantization (SFQ)、画像符号化の国際標準方式 JPEG 及び JPEG2000 のものと比較することで、空間-周波数領域の最適分割の有効性を定量的に明らかにする。

## 第4章：視知覚特性を考慮した画像のサブバンド符号化における空間-周波数領域の最適分割

前章までに提案したレート歪み理論に基づいた最適帯域分割及び空間-周波数領域の最適分割は、サブバンド信号における量子化前後間の平均二乗誤差 (MSE) によって定義された歪みの量を最小化することを目的として導出されたものであることから、視知覚特性を考慮した画質という観点では、真に最適なものが得られているとは言い難い。そこで、第3章で提案した空間-周波数領域の最適分割よりも主観的に良好な再生画像を得ることを目的として、視知覚特性上で重要ではない信号の品質を犠牲にする代わりに重要な信号の品質を優先的に保つことで、主観的な画質の改善を達成する視知覚特性を考慮した空間-周波数領域の最適分割を提案する。本空間-周波数領域の最適分割は、人間の視知覚特性の一種である視覚の空間周波数特性と画像の顕著性の二つを考慮して、各サブバンド信号の歪み値に対する重み付けを行った後、その重み付けされた歪みの量を最小化するように、帯域分割、量子化、及び空間領域分割をそれぞれ最適化することによって実現される。最後に、視知覚特性を考慮した空間-周波数領域の最適分割により得られた再生画像の画質を評価するために、画質の主観評価実験を行い、得られた評価結果を、従来の視知覚特性を考慮しない通常的空間-周波数領域の最適分割方式及び画像符号化の国際標準方式 JPEG2000 のものと比較することで、本空間-周波数領域の最適分割の有効性を明らかにする。

## 第5章：結論

本論文の総括として、第2章から第4章までの結果を要約すると共に、今後に残された研究課題について述べる。

## 第 2 章

# レート歪み理論に基づいた画像のサブバンド符号化における最適帯域分割

### 2.1 序言

高能率画像符号化の一手法であるサブバンド符号化において、高い符号化性能を得るためには、まず、入力画像が持つ固有の特性に応じて画像の二次元周波数帯域を分割した後、分割された各サブバンドに対して、そのサブバンド信号の性質に整合した異なる量子化及びエントロピー符号化を適用することが重要である。本符号化において、エントロピー符号化器設計の処理コストを低く抑えつつ、高い符号化性能を達成するために、あらかじめ少数のサブバンド数を指定した条件の下で、入力画像の特性を考慮して二次元周波数帯域を適応的に分割する最適帯域分割が提案されている [59]～ [61]。しかしながら、従来の最適帯域分割は、量子化が行われる前の各サブバンドの信号電力に基づき、量子化雑音電力最小の観点で最適な帯域分割パターンを決定しているため、符号化レートが変更されたときにも帯域分割パターンは固定のままであった。また、各サブバンドに適用される量子化器は、符号化レートの変更に対して変化するものの、全てのサブバンドで同じものが適用される仕様となっていた。すなわち、従来の最適帯域分割は、符号化レートの変更に対して適切な帯域分割及び量子化が行われておらず、十分な符号化性能が達成されていなかった。

本章では、従来の最適帯域分割から符号化性能を改善することを目的として、符号化レートと画質劣化量の相互関係を表すレート歪み理論 [37], [38] の観点から、帯域分割及び量子化を最適化する「レート歪み理論に基づいた最適帯域分割」を提案する。提案手法では、レート歪み理論の観点で両者の最適化問題を定式化した後、所望の符号化レートに対する二次元周波数帯域の最適分割パターン及び各サブバンドに適用される最適な量子化器の組み合わせを求めるアルゴリズムを開発する。最後に、レート歪み理論に基づいた本最適帯域分割を実画像に適用し、その符号化性能を、従来の最適帯域分割、離散ウェーブレット変換、Wavelet Packet [45] のものと比較することで、本最適帯域分割の有効性を定量的に明らかにする。

## 2.2 レート歪み理論に基づいた帯域分割と量子化の最適化

従来の最適帯域分割における問題を解決するために、任意の符号化レートが指定されたときに、レート歪み理論の観点で、二次元周波数帯域の分割パターンと各サブバンドに適用される量子化器の組み合わせを同時に最適化する手法を提案する。このとき、レート歪み理論における最適化問題は、ラグランジュの未定乗数法の適用により解が求められることが知られている [38]。そこで、まずはレート歪み理論における帯域分割と量子化の二つの最適化問題を解くための理論式を明らかにし、その後、所望の符号化レートに応じてそれら二つの解を同時に導出するためのアルゴリズムを提案する。

### 2.2.1 最適化問題の定式化

一般に、レート歪み理論における最適化問題とは、符号化レート  $R$  が指定されているという制約条件の下で、符号化による画質の劣化量を表す歪み  $D$  を最小化する問題を指す。このような制約条件付きの最適化問題は、ラグランジュの未定乗数法を用いて、

$$J_{opt}(\lambda) = \min[D + \lambda R] \quad (2.1)$$

として定義される制約条件が無いコスト関数の最小化問題に置き換えることができる [38]。ここで、式 (2.18) における  $\lambda$  はラグランジュ乗数と呼ばれ、任意の符号化レートに応じた解を求めるための定数である。サブバンド符号化においては、分割された各サブバンドが互いに直交性を満たしているという条件の下で、各サブバンド信号をそれぞれ量子化した後に、帯域合成フィルタによって画像を再構成した場合、その再生画像において発生する全体の歪み  $D$  は、式 (2.2) のように各サブバンド  $\Omega_k (k = 0, 1, \dots, M-1)$  を量子化することによって生じる歪み  $D_k$  の総和として近似的に表すことができる [38], [45]。

$$D \approx \sum_{k=0}^{M-1} D_k \quad (2.2)$$

また、量子化後の各サブバンド信号は、それぞれ独立にエントロピー符号化が適用されるため、全体の符号化レート  $R$  は、各サブバンド  $\Omega_k$  における符号化レートを  $R_k$  として、

$$R = \sum_{k=0}^{M-1} R_k \quad (2.3)$$

により計算される。このとき、最適帯域分割によって求められる  $M$  個のサブバンド  $\Omega_k (k = 0, 1, \dots, M-1)$  は、画像信号に QMF [62] 等の帯域分割フィルタを適用することによって得ら

れた帯域ブロック  $\Omega_i (i = 0, 1, \dots, N-1)$  の集合で構成されていることから、互いに直交性を満たしている。したがって、レート歪み理論上の最適帯域分割の問題は、式 (2.2), (2.3) より、

$$J_{opt}(\lambda) = \min \left[ \sum_{k=0}^{M-1} D_k + \lambda \sum_{k=0}^{M-1} R_k \right] \quad (2.4)$$

として与えられる。

以上を留意して、まずは便宜上、先に各サブバンドに適用される量子化器の組み合わせにおける最適化問題を定式化する。ここで、提案手法において式 (2.4) を扱うことの利点は、帯域分割された複数のサブバンドをそれぞれ量子化するとき、各サブバンドのコスト関数を個別に最小化することで、それが最終的に最適な量子化器の組み合わせ結果となるところにある。加えて、 $D_k$  もまた、サブバンド  $\Omega_k$  に含まれる帯域ブロック  $\Delta\Omega_i (i \in I_k^{(M)})$  の歪み  $D_i$  の総和と等しいことを踏まえれば、量子化器の組み合わせの最適化問題は、式 (2.4) より、

$$J_{opt}(\lambda) = \sum_{k=0}^{M-1} \min \left[ \sum_{i \in I_k^{(M)}} D_i + \lambda R_k \right] \quad (2.5)$$

として定式化される。ここで、 $I_k^{(M)}$  は、サブバンド  $\Omega_k$  に含まれる帯域ブロック番号  $i$  の集合を表している。式 (2.5) は、当該の問題がサブバンド  $\Omega_k$  毎にコスト関数を最小化すればよいことを表しており、式 (2.5) を導入することで、各サブバンドに適用する最適な量子化器の組み合わせを求めることができる。このとき、式 (2.5) を解くためには、サブバンド  $\Omega_k$  を量子化したときに算出されるレート  $R_k$  と歪み  $D_k$  の各値が必要となるが、それらの値は、サブバンド  $\Omega_k$  に含まれる帯域ブロック  $\Delta\Omega_i$  の組み合わせに依存するため、一方の帯域分割パターンが未知である場合、式 (2.5) は不良設定問題となる。したがって、式 (2.5) を解く際には、帯域分割パターンが事前に求められている必要がある。

次に、帯域分割パターンの最適化問題の定式化を行う。最適帯域分割において、レート  $R_k$  はサブバンド  $\Omega_k$  の平均情報量（エントロピー）に、定数である帯域ブロックの信号数  $S_{\Delta\Omega}$  と、 $I_k^{(M)}$  の個数  $N_k$  を乗じたものとして定義されていることに留意すれば、式 (2.4) は、

$$J_{opt}(\lambda) = \min \left[ \sum_{k=0}^{M-1} \sum_{i \in I_k^{(M)}} D_i + \lambda \left( -S_{\Delta\Omega} \sum_{k=0}^{M-1} N_k \sum_{x=x_{min}^k}^{x_{max}^k} \Omega_k(x) \log_2 \Omega_k(x) \right) \right] \quad (2.6)$$

のように変形できる。ここで、式 (2.6) 中の  $\Omega_k(x)$  は、サブバンド  $\Omega_k$  における信号の確率分布を表し、 $x_{min}^k$  と  $x_{max}^k$  は、それぞれ  $\Omega_k(x) \neq 0$  となる  $x$  の最小値と最大値である。このとき、 $\Omega_k(x)$  は、サブバンド  $\Omega_k$  に属する帯域ブロック  $\Delta\Omega_i$  の確率分布  $\Delta\Omega_i(x) (i \in I_k^{(M)})$  を統合した

混合確率分布と等価であることから,

$$\Omega_k(x) = \frac{1}{N_k} \sum_{i \in I_k^{(M)}} \Delta\Omega_i(x) \quad (2.7)$$

に置き換えられる. さらに,  $-p \log q = -p \log p + p \log(p/q)$  という関係を考慮すれば, 式 (2.6) は,  $x_{min}^i$  と  $x_{max}^i$  を, それぞれ  $\Delta\Omega_i(x) \neq 0$  となる  $x$  の最小値と最大値として,

$$\begin{aligned} J_{opt}(\lambda) &= \min \left[ \sum_{k=0}^{M-1} \sum_{i \in I_k^{(M)}} D_i + \lambda \left( -S_{\Delta\Omega} \sum_{k=0}^{M-1} \sum_{i \in I_k^{(M)}} \sum_{x=x_{min}^i}^{x_{max}^i} \Delta\Omega_i(x) \log_2 \Omega_k(x) \right) \right] \\ &= \min \left[ \sum_{i=0}^{N-1} D_i + \lambda \left( - \sum_{i=0}^{N-1} S_{\Delta\Omega} \sum_{x=x_{min}^i}^{x_{max}^i} \Delta\Omega_i(x) \log_2 \Delta\Omega_i(x) \right) \right. \\ &\quad \left. + \lambda S_{\Delta\Omega} \sum_{k=0}^{M-1} \sum_{i \in I_k^{(M)}} \sum_{x=x_{min}^i}^{x_{max}^i} \Delta\Omega_i(x) \log_2 \frac{\Delta\Omega_i(x)}{\Omega_k(x)} \right] \quad (2.8) \end{aligned}$$

のように展開できる. ここで, 式 (2.8) 中の第 1 項は, 帯域ブロック  $\Delta\Omega_i$  を量子化したときに発生する歪み  $D_i$  の総和であり, 第 2 項は, 帯域ブロック  $\Delta\Omega_i$  のエントロピーと帯域ブロックの信号数  $S_{\Delta\Omega}$  の積の総和, すなわち, 帯域ブロック  $\Delta\Omega_i$  におけるレート  $R_i$  の総和である. 第 3 項は, 帯域ブロックの確率分布  $\Delta\Omega_i(x)$  とそれらを含むサブバンドの確率分布  $\Omega_k(x)$  への Kullback-Leibler Divergence [64] の総和に帯域ブロックの信号数  $S_{\Delta\Omega}$  を乗じたものである. ここで, 二つの確率分布  $P(x)$  と  $Q(x)$  が与えられたとき, それらの類似度を測るための距離として定義される Kullback-Leibler Divergence を

$$D_{KL}(P \parallel Q) = \sum_x P(x) \log \frac{P(x)}{Q(x)} \quad (2.9)$$

のように表記すれば, 式 (2.8) は,

$$J_{opt}(\lambda) = \min \left[ \sum_{i=0}^{N-1} D_i + \lambda \sum_{i=0}^{N-1} R_i + \lambda S_{\Delta\Omega} \sum_{k=0}^{M-1} \sum_{i \in I_k^{(M)}} D_{KL}(\Delta\Omega_i \parallel \Omega_k) \right] \quad (2.10)$$

と記述できる. このとき, 第 1 項と第 2 項は帯域分割パターンの変化に影響を受けず固定値となるが, 第 3 項は帯域分割パターンの変化に依存する. ここで, 第 3 項における Kullback-Leibler Divergence の計算では確率分布が参照されるが, その確率分布を構成するために帯域ブロック内の信号は全て事前に量子化されていなければならない. よって, 式 (2.10) を直接解くためには, 各帯域ブロック  $\Delta\Omega_i$  に対する量子化器の組み合わせを全て用意し, その中から各帯域ブロックの確率分布  $\Delta\Omega_i(x)$  を参照して第 3 項を最小化するものを探索しなくてはな

らないため、大きな処理コストを必要とする。しかしながら、事前に各帯域ブロック  $\Delta\Omega_i$  に対する量子化器が決定されているならば、レート  $R_i$  と歪み  $D_i$  は既知として扱われ、式 (2.10) における第 1 項と第 2 項は、いずれも定数となり無視できる。したがって、各帯域ブロックの量子化器が既に決定済みであることを条件として、式 (2.10) の解を求める問題は、第 3 項の最小化を目的とした Kullback-Leibler Divergence を距離尺度とした確率分布のクラスタリング問題に帰着される。

以上をまとめると、一つの最適化問題を解くためには、もう一方の最適化問題の解を必要とすることから、レート歪み理論における最適帯域分割の問題は、互いに適当な初期値を与え、指定された符号化レートに到達するまでラグランジュ乗数  $\lambda$  を更新させながら、

$$\omega_{opt}(\{\Delta\Omega_i(x)\}) = \min \left[ \sum_{k=0}^{M-1} \sum_{i \in I_k^{(M)}} D_{KL}(\Delta\Omega_i \parallel \Omega_k) \right] \quad (2.11)$$

$$\psi_{opt}(\{I_k^{(M)}\}, \lambda) = \sum_{k=0}^{M-1} \min \left[ \sum_{i \in I_k^{(M)}} D_i + \lambda R_k \right] \quad (2.12)$$

を反復することによって、それぞれの最適解を求めることが可能となる。次項では、式 (2.11) と式 (2.12) に基づいて、符号化レートが与えられたときに、最適な帯域分割パターンと各サブバンドに適用する最適な量子化器の組み合わせを決定するためのアルゴリズムを提案する。

## 2.2.2 帯域分割と量子化の導出アルゴリズム

ラグランジュの未定乗数法を用いてレート歪み理論における最適化問題を解く場合には、反復法による求根アルゴリズムを使用して、符号化レートに対応する最適なラグランジュ乗数  $\lambda$  を探索した後に、その  $\lambda$  値に基づいて式 (2.18) を解くことで、所望の符号化レートに応じた解の導出が行われる。これを踏まえて、提案手法では、求根アルゴリズム内に式 (2.11) と式 (2.12) を組み込み、指定された符号化レートに到達するまで反復処理による逐次的な最適化処理を行うことで、二次元周波数帯域の最適分割パターンと各サブバンドに適用される量子化器の組み合わせの両方の最適解を決定する。このとき、求根アルゴリズムには、文献 [65] において提案された二分法とニュートン法に基づく手法を採用する。図 2.1 に、指定の符号化レートに対応した二次元周波数帯域の最適分割パターン及び各サブバンドに適用される最適な量子化器の組み合わせを導出するアルゴリズムを示す。

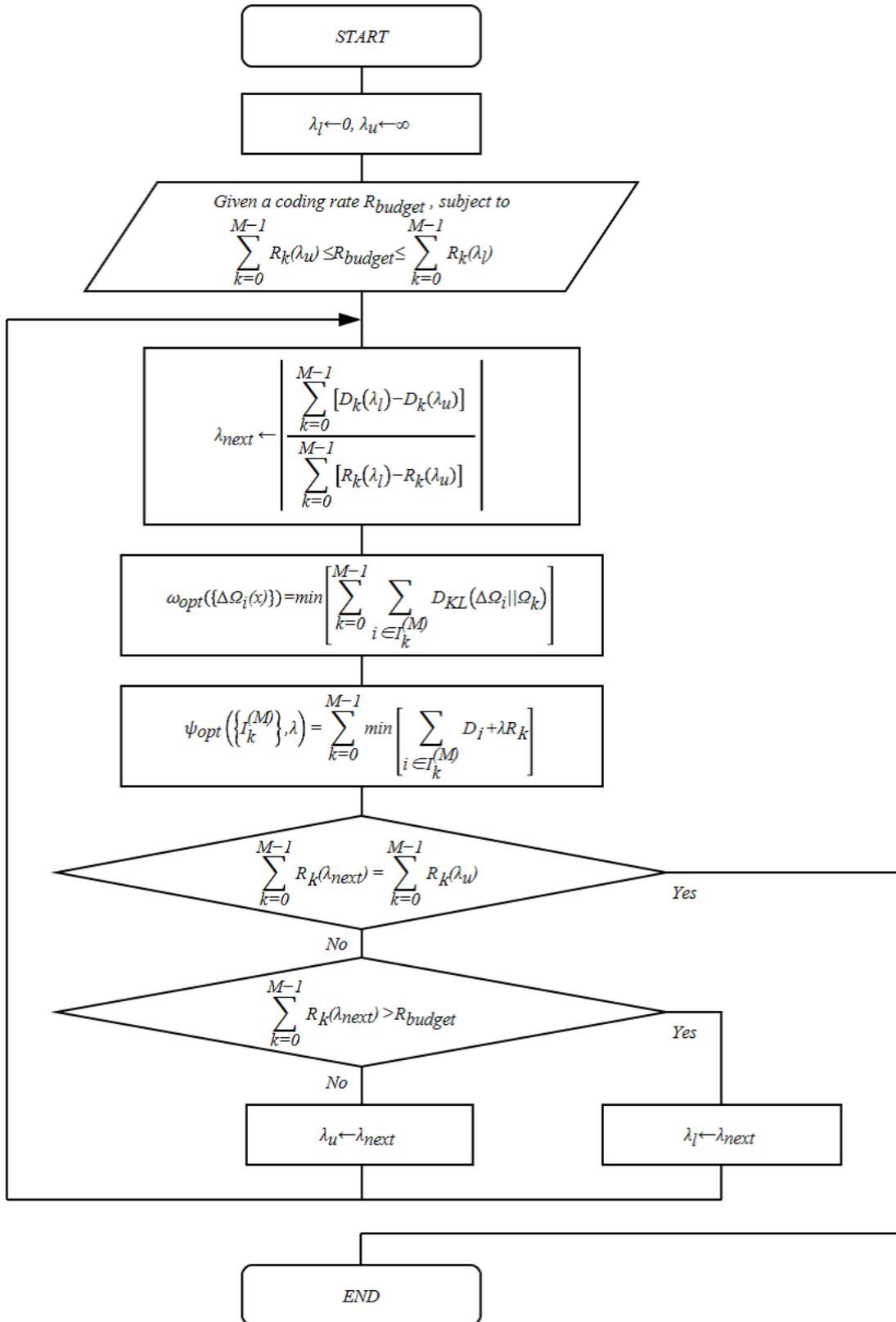


図 2.1 符号化レートに対応した二次元周波数帯域の最適分割パターン及び各サブバンドに適用される最適な量子化器の導出アルゴリズム

まず，入力画像の二次元周波数帯域  $\Omega$  に対して，QMF による一次元フィルタバンクを水平及び垂直方向に3段階適用することにより，互いに直交性を満たす  $N = 64$  個の帯域ブロック  $\Delta\Omega_i (i = 0, 1, \dots, 63)$  が得られているとする．次に，符号化レート  $R_{budget}$  が指定されると，帯域ブロックの確率分布  $\Delta\Omega_i(x)$  に対して，式(2.11)に示した最適化処理により各帯域ブロックに対する最適クラスタリングが行われ，その結果として  $M = 5$  の最適帯域分割が決定される．さらに，これらのサブバンドに対して，式(2.12)により，5個のサブバンドに対する最適な量子化器の組み合わせが決定される．ここで，各サブバンドに適用される量子化器には，量子化ステップ幅  $(1, 2, \dots, 16)$  を持つ16種類のミッドトレッド型の線形量子化器  $Q_1 \sim Q_{16}$  を用意した．量子化器の組み合わせが求められた後に算出されたレート  $R$  が  $R_{budget}$  に到達していなければ，前回求められたレート  $R(\lambda)$  と歪み  $D(\lambda)$  の値に基づいてラグランジュ乗数  $\lambda$  を更新した後，上述した二つの最適化処理が再帰的に行われる．このとき，

$$R(\lambda) = \sum_{k=0}^{M-1} R_k(\lambda) \quad (2.13)$$

$$D(\lambda) = \sum_{k=0}^{M-1} D_k(\lambda) \quad (2.14)$$

として計算され，各式中の  $R_k(\lambda)$  と  $D_k(\lambda)$  は，それぞれラグランジュの未定乗数法により求められる  $\lambda$  に対するサブバンド  $\Omega_k$  におけるレート  $R_k$  と歪み  $D_k$  となる．それぞれ，レート  $R_k$  は，サブバンド  $\Omega_k$  の平均情報量（エントロピー）にサブバンド  $\Omega_k$  の信号数と等価である  $N_k \cdot S_{\Delta\Omega}$  を乗じたもの，歪み  $D_k$  は，サブバンド  $\Omega_k$  に属する帯域ブロック  $\Delta\Omega_i (i \in I_k^{(M)})$  の量子化前後間の信号の平均二乗誤差（Mean Squared Error : MSE）の総和として定義した．

次に，それぞれの最適化処理についてその内容を詳述する．帯域分割パターンの最適化処理は，式(2.11)における Kullback-Leibler Divergence を距離尺度とした確率分布のクラスタリング問題を解くことと等価である．レート歪み理論に基づいた最適帯域分割では，解の収束が高速であり，劣悪な局所最適解に陥るのを防ぐために初期点の設定を工夫した k-means++ 法 [66] を使用して，その解を求めることとする．具体的には，まず，他とは確率分布の形状が大きく異なる直流成分の帯域ブロック  $\Delta\Omega_0$  を一つ目のクラスタ重心として選んだ後，残り 63 個の帯域ブロックの中から 4 個のクラスタ重心が選ばれるまで

$$\varphi(\Delta\Omega_i) = \max_{i \notin C^{(M)}} \frac{\gamma(\Delta\Omega_i)}{\sum_{j=0}^{N-1} \gamma(\Delta\Omega_j)} \quad (2.15)$$

を繰り返し解くことで，クラスタ重心の初期点を決定する．ここで，

$$\gamma(\Delta\Omega_i) = \min_{j \in C^{(M)}} D_{KL}(\Delta\Omega_i \parallel \Delta\Omega_j) \quad (2.16)$$

であり、 $C^{(M)}$  は、既にクラスタ重心として決定された帯域ブロック番号の集合を表している。その後、選ばれた5個のクラスタ重心を初期点として、通常のk-means法を実行することで、5分割の帯域分割パターンが決定される。このとき、解が収束するまで更新されるクラスタ重心は、各クラスタに所属する帯域ブロックの確率分布を統合した混合確率分布として扱われ、式(2.7)により求められる。量子化の最適化処理は、サブバンド $\Omega_k$ に16種類の線形量子化器 $Q_1 \sim Q_{16}$ をそれぞれ適用した後の信号からそれぞれの場合のレート $R_k$ と歪み $D_k$ を求め、その中でコスト関数 $J(\lambda) = D_k + \lambda R_k$ を最小にするときの量子化器の組み合わせを選択することで行われる。

前節で述べたように、帯域分割パターンの最適化処理では、一つ前の量子化の最適化処理で決定した量子化器の組み合わせを適用することで得られる各サブバンドの確率分布が必要とされる。このとき、初回の最適化処理では、各サブバンドに対する量子化器が求められていないため、量子化の初期状態として、16種類の量子化器のうち中央の量子化ステップ幅を持つ量子化器 $Q_8$ を全てのサブバンドに適用し、それによって得られた確率分布を確率分布の初期値として設定した。なお、いずれの初期状態を選択しても、同じ最適解に収束することを実験的に確認している。以上の処理によって、任意の符号化レートを指定するだけで、レート歪み理論の観点における最適な帯域分割パターンと、各サブバンドに適用する最適な量子化器の組み合わせを自動的に求めることができる。

## 2.3 シミュレーション実験とその結果

レート歪み理論に基づいた最適帯域分割を実画像に適用することにより、その符号化性能を明らかにした上で、従来の最適帯域分割との比較を行う。ここで、レート歪み理論に基づいた最適帯域分割及び従来の最適帯域分割は共に、帯域分割を求めるときの条件である帯域ブロック数  $N$  とサブバンド数  $M$  については、 $N = 64$ 、 $M = 5$  とした。このパラメータ値は、複数種類の実画像を用いた実験を行うことで、性能改善と処理コストの両方を考慮した適切な値として、最適帯域分割の既存研究の中で既に決定されているものである [61]。また、前述したように各サブバンドに適用する量子化器には、量子化ステップ幅  $(1, 2, \dots, 16)$  を持つ 16 種類のミッドレッド型の線形量子化器  $Q_1 \sim Q_{16}$  を用意した。さらに、レート歪み理論に基づいた最適帯域分割は、JPEG2000 で用いられている帯域分割手法である離散ウェーブレット変換、及びサブバンド数に制限を設けずに帯域分割の理論的最適解を導出できる Wavelet Packet [45] との性能比較も行う。なお、本論文では、これらの手法において、フィルタバンクによるサブバンドの再帰分解レベル  $d = 3$  に設定した。このとき、離散ウェーブレット変換は、サブバンド数を  $M = 10$  とした条件で、各サブバンドに適用される量子化器の組み合わせのみを、Wavelet Packet は、サブバンド数に制限を設けない条件下において、帯域分割パターンと各サブバンドに適用される量子化器の組み合わせを、レート歪み理論の観点で最適化できる手法としてみなされる。Wavelet Packet では全てのサブバンドに対して分解が行われるため、最大で  $M = 64$  のサブバンドに分割されることになる。また、各サブバンドに使用する量子化器には、二つの最適帯域分割法と同じく  $Q_1 \sim Q_{16}$  を用いることとした。

符号化性能の評価は、各手法を入力画像に適用したときの PSNR 対ビットレート特性により行われる。PSNR は、入力画像と符号化画像との間の再生誤差電力を  $\sigma_e^2$  として、

$$PSNR = 10 \log_{10} \frac{255^2}{\sigma_e^2} \quad (2.17)$$

により計算される。ビットレートについては、量子化された各サブバンド  $\Omega_k$  の信号が無記憶情報源から生じたものと仮定して発生確率  $\Omega_k(x)$  を求め、エントロピーの定義式に従って、式 (??) のように算出している。

$$bitrate = -\frac{1}{N} \sum_{k=0}^{M-1} N_k \sum_{x=x_{min}^k}^{x_{max}^k} \Omega_k(x) \log_2 \Omega_k(x) \quad (2.18)$$

なお、上述した 4 種類の帯域分割方式を用いて実際に符号化を行う場合、復号化側で再生するために、各サブバンドに適用される量子化器の組み合わせを伝送するためのオーバーヘッド情報が必要である。さらに、レート歪み理論に基づいた最適帯域分割、従来の最適帯域分割、及び Wavelet Packet においては、それに加えて決定された帯域分割パターンの情報が必

要となる。入力画像の画素数が  $(x \times y)$  画素，サブバンド数  $M$ ，各サブバンドに割り当てられる量子化器が  $q$  種類用意されていることを条件とした場合，前者の量子化器の組み合わせのオーバーヘッド情報は，4種類の帯域分割方式に共通して  $(\lceil \log_2 q \rceil [\text{bit}] \times M) / (x \times y) [\text{bit}/\text{pel}]$  となる。後者の帯域分割パターンに関しては，レート歪み理論に基づいた最適帯域分割及び従来の最適帯域分割の場合， $N$  個の帯域ブロックがどのサブバンドに属するかを示せばよいので， $(\lceil \log_2 M \rceil [\text{bit}] \times N) / (x \times y) [\text{bit}/\text{pel}]$  となる。一方，Wavelet Packet の場合は，サブバンド毎にフィルタバンクによる再帰分解レベルの深さが異なるため，帯域分割パターンのオーバーヘッド情報は，そのサブバンドが4つのサブバンドに分割されていることを“1”，逆に分割されていないことを“0”で示す二値情報によって表現され，全帯域が分解レベル  $d$  の深さまで分解されるときに最大となる。具体的に，本論文で使用している条件を当てはめた場合，全体のオーバーヘッド情報量は，レート歪み理論に基づいた最適帯域分割及び従来の最適帯域分割では， $x = y = 256$ ， $N = 64$ ， $M = 5$ ， $q = 16$  の条件下で  $0.0032 [\text{bit}/\text{pel}]$  となる。また，離散ウェーブレット変換では， $x = y = 256$ ， $d = 3$ （すなわち  $M = 10$ ）， $q = 16$  の条件下で  $0.0006 [\text{bit}/\text{pel}]$  となり，Wavelet Packet では，同条件下で  $0.0042 [\text{bit}/\text{pel}]$  が最大値となる。このことから，4種類の帯域分割方式のオーバーヘッド情報量は，画像自身の情報量と比較して無視できるほど小さな値となる。なお，以降より示す符号化特性にはこのオーバーヘッド情報量を含めた結果を表示する。

図 2.2(a)~(e) は，SIDBA 標準画像 [63] の中から選択した画像“Lenna”，画像“Barbara”，画像“Mandrill”，画像“Sailboat”，画像“Earth”（いずれも  $(256 \times 256)$  画素，濃淡  $8 [\text{bit}/\text{pel}]$ ）であり，図 2.3(a)~(e) は，図 2.2 に示す各画像に対して，レート歪み理論に基づいた最適帯域分割を適用することによって求められた PSNR 対ビットレート特性による符号化性能を，従来の最適帯域分割，離散ウェーブレット変換，Wavelet Packet によるものと比較した結果である。図 2.3 の結果から，レート歪み理論に基づいた最適帯域分割は，従来の最適帯域分割と比べて，入力画像の種類に依らず全ての符号化レートで符号化性能が改善されていることがわかる。量子化後のサブバンド信号の特性を考慮しながら，帯域分割と量子化の最適化を行ったことで，特に低符号化レートにおける符号化性能が従来の最適帯域分割よりも改善されており，レート歪み理論に基づいた最適帯域分割は，同程度の符号化レートを持つ従来の最適帯域分割よりも，PSNR 値において最大で  $1.0 [\text{dB}]$  程度の高い値を示している。

図 2.4(a)~(e) は，レート歪み理論に基づいた最適帯域分割において，(i) 低符号化レートと (ii) 高符号化レートをそれぞれ指定したときに，各入力画像に対して得られた二次元周波数帯域  $\Omega$  の分割パターンを示している。また，比較のために，従来の最適帯域分割による帯域分割パターンを，それぞれ図 2.5(a)~(e) に示す。ただし，これらの帯域分割パターンの形状は，符号化レートの変更に対しては不変となっている。さらに表 2.1(a)~(e) には，それぞれ図 2.4(a)~(e) に示した帯域分割パターンに対して，レート歪み理論に基づいた最適帯域分割により決定された各サブバンドに適用される量子化器の組み合わせ結果を示している。また，比較のために，図 2.5(a)~(e) に示した従来の最適帯域分割によって求められた帯域分

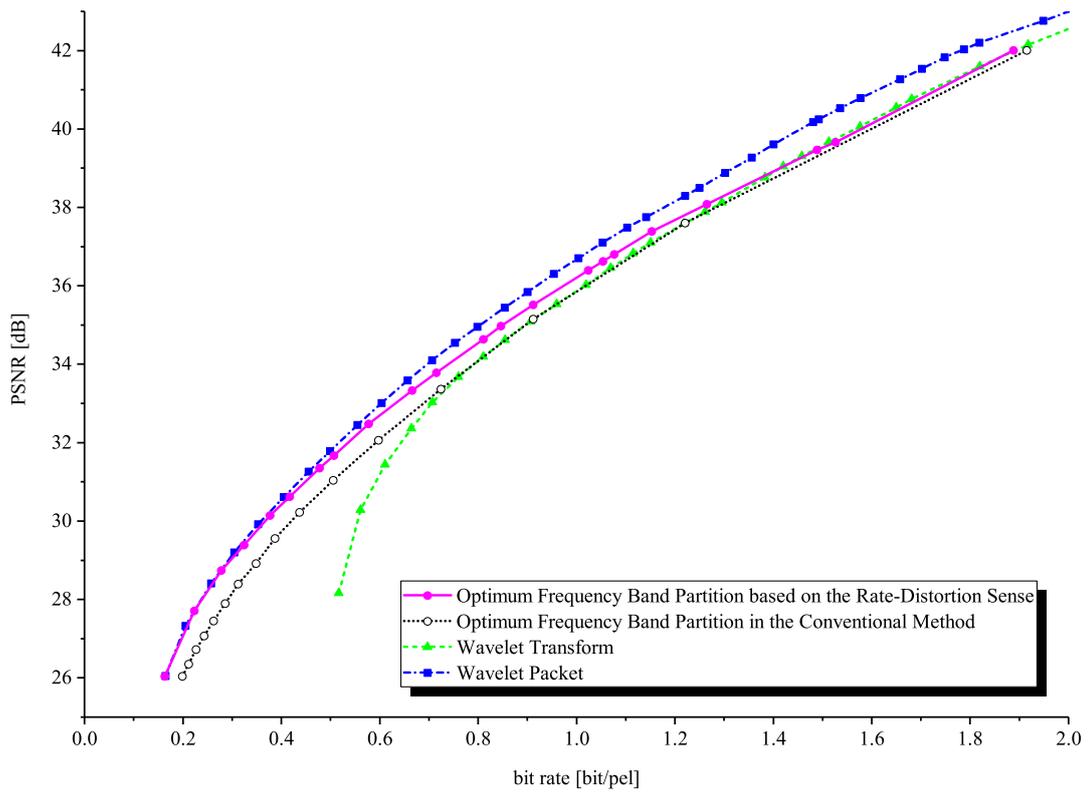
割パターンに対して、高符号化レートと低符号化レートそれぞれの場合に決定された各サブバンドに適用される量子化器の組み合わせ結果についても表示しているが、先述したように従来の最適帯域分割では、符号化レートの変更により変化するものの、全てのサブバンドにおいて同じ量子化器が適用されている。図 2.4(a)~(e) を比較したとき、レート歪み理論に基づいた最適帯域分割によって決定された帯域分割パターンは、それぞれ入力画像に対して固有な結果が求められており、さらに、各図における (i) と (ii) を比較すると、同じ入力画像であっても、低符号化レートと高符号化レートでは、それぞれ異なる形状の帯域分割パターンが求められていることがわかる。そして、これらは図 2.5 に示された従来の最適帯域分割による帯域分割パターンとも異なっていることから、レート歪み理論に基づいた最適帯域分割における確率分布の類似性に基づいた帯域ブロックのクラスタリングによって新しい最適帯域分割が求められていることがわかる。また、表 2.1(a)~(e) より、各サブバンドに適用される量子化器の組み合わせについても、符号化レートの変更に応じてサブバンド毎に異なる量子化器が選択されていることがわかる。これらの結果と図 2.3 に示す符号化性能の結果から、レート歪み理論に基づいた最適帯域分割では、符号化レートの変更に応じて帯域分割パターンと各サブバンドに対する量子化器を求めたことによって、従来の最適帯域分割から符号化性能を改善することができたと言える。

次に、図 2.3 において、レート歪み理論に基づいた最適帯域分割と、離散ウェーブレット変換及び Wavelet Packet の符号化性能を比較する。まず、離散ウェーブレット変換と比べると、レート歪み理論に基づいた最適帯域分割はいずれの入力画像においても、特に低符号化レートで符号化性能が大きく改善されていることがわかる。次に、レート歪み理論の観点で帯域分割と量子化の理論的な最適解が得られる Wavelet Packet と比較すると、低符号化レートの場合は全ての画像において同等の符号化性能が得られているものの、高符号化レートの場合は符号化性能が若干下回っていることが確認できる。



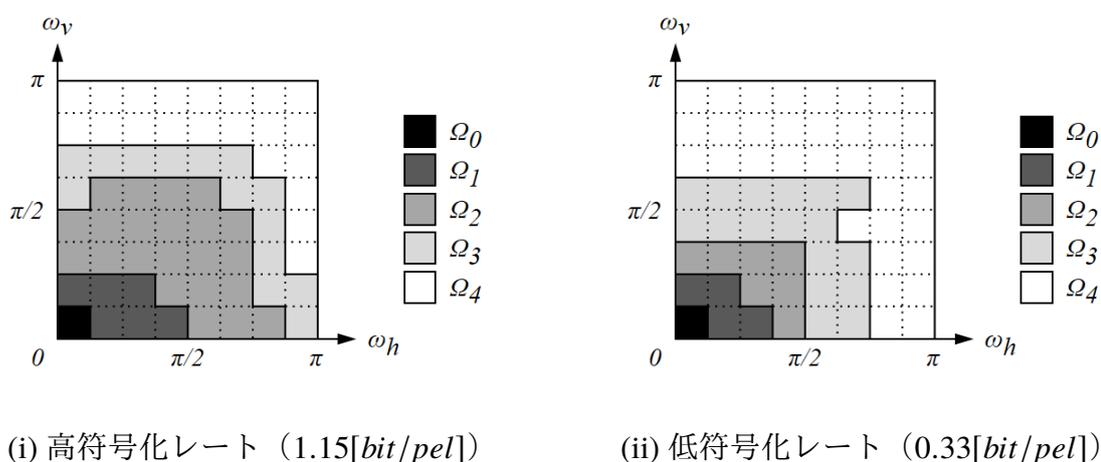
(a) 画像“Lenna” ((256 × 256) 画素, 濃淡 8[bit/pel])

図 2.2 テスト画像



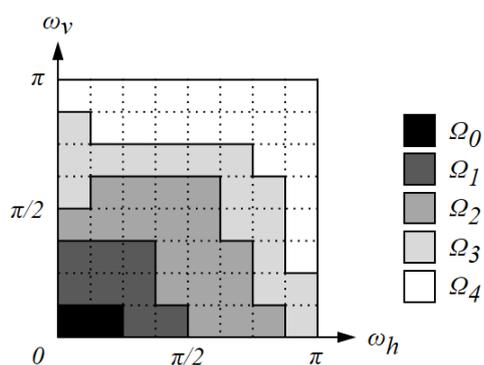
(a) 画像“Lenna”

図 2.3 PSNR 対ビットレート特性による符号化性能の比較



(a) 画像“Lenna”

図 2.4 レート歪み理論に基づいた最適帯域分割による帯域分割パターン ( $M = 5, N = 64$ )



(a) 画像“Lenna”

図 2.5 従来の最適帯域分割により求められた帯域分割パターン ( $M = 5, N = 64$ )

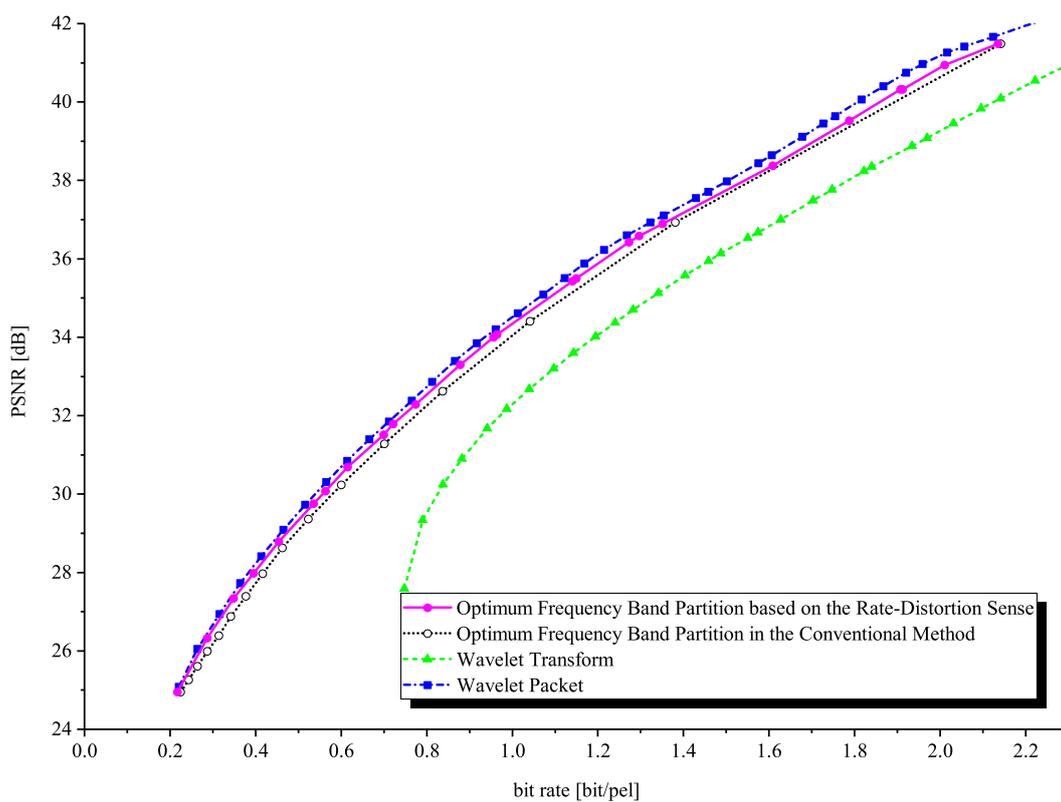
表 2.1 図 2.4(a) と図 2.5(a) に示す各サブバンド信号に適用される量子化器

subband	quantizer applied to each subband signal in frequency band partition pattern			
	Figure 2.4(a)		Figure 2.5(a)	
	(i) high bit rate (1.15[bit/pel])	(ii) low bit rate (0.33[bit/pel])	(i) high bit rate (1.22[bit/pel])	(ii) low bit rate (0.31[bit/pel])
$\Omega_0$	$Q_1$	$Q_6$	$Q_2$	$Q_{10}$
$\Omega_1$	$Q_2$	$Q_7$	$Q_2$	$Q_{10}$
$\Omega_2$	$Q_2$	$Q_8$	$Q_2$	$Q_{10}$
$\Omega_3$	$Q_2$	$Q_{10}$	$Q_2$	$Q_{10}$
$\Omega_4$	$Q_3$	$Q_{16}$	$Q_2$	$Q_{10}$



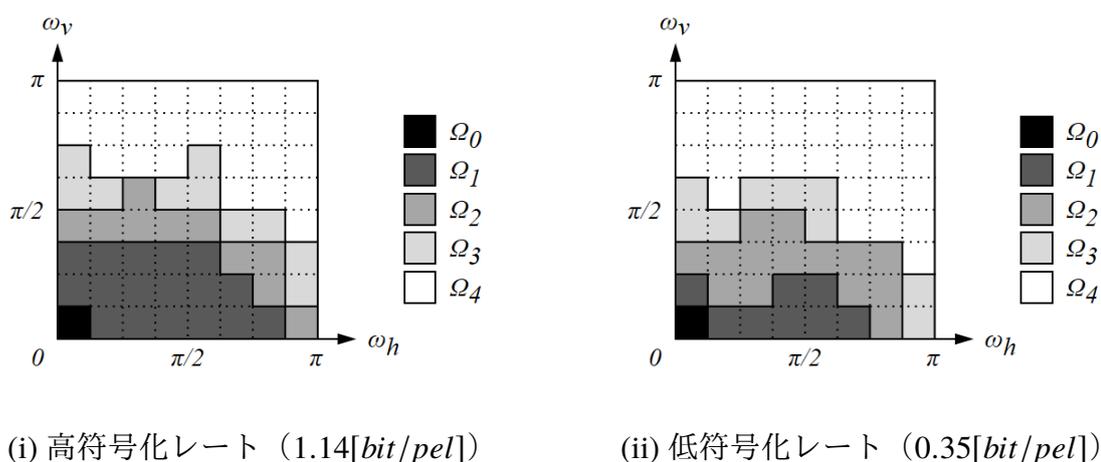
(b) 画像“Barbara” ((256 × 256) 画素, 濃淡 8[bit/pel])

図 2.2 テスト画像



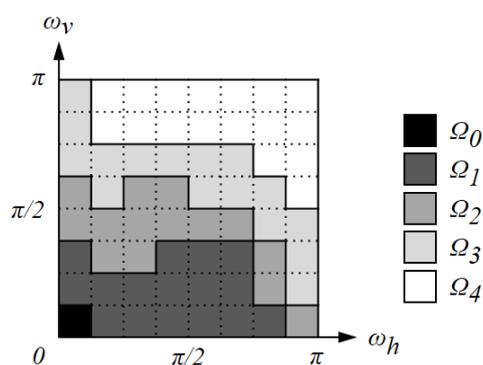
(b) 画像“Barbara”

図 2.3 PSNR 対ビットレート特性による符号化性能の比較



(b) 画像 “Barbara”

図 2.4 レート歪み理論に基づいた最適帯域分割による帯域分割パターン ( $M = 5, N = 64$ )

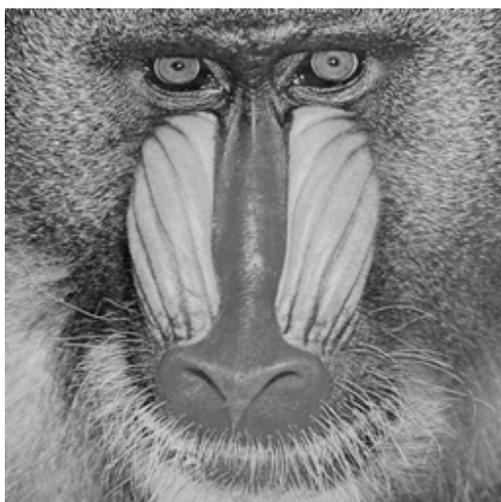


(b) 画像 “Barbara”

図 2.5 従来の最適帯域分割により求められた帯域分割パターン ( $M = 5, N = 64$ )

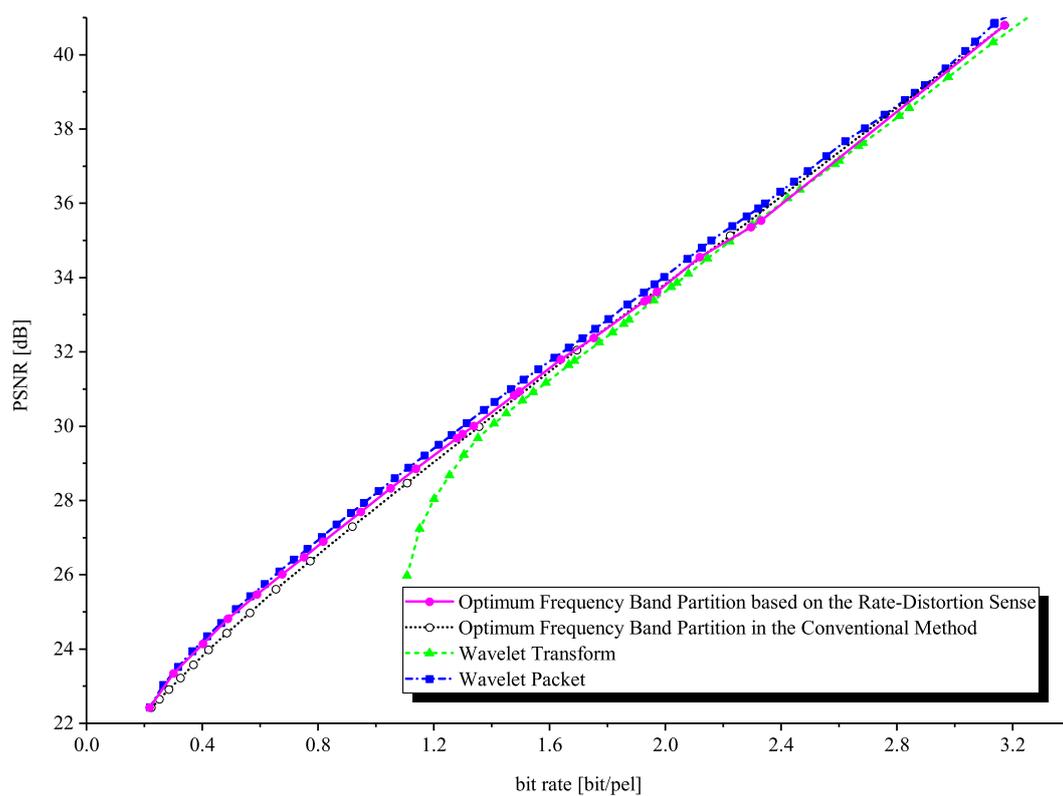
表 2.2 図 2.4(b) と図 2.5(b) に示す各サブバンド信号に適用される量子化器

subband	quantizer applied to each subband signal in frequency band partition pattern			
	Figure 2.4(b)		Figure 2.5(b)	
	(i) high bit rate (1.14[bit/pel])	(ii) low bit rate (0.35[bit/pel])	(i) high bit rate (1.04[bit/pel])	(ii) low bit rate (0.34[bit/pel])
$\Omega_0$	$Q_2$	$Q_8$	$Q_3$	$Q_{11}$
$\Omega_1$	$Q_2$	$Q_9$	$Q_3$	$Q_{11}$
$\Omega_2$	$Q_3$	$Q_{11}$	$Q_3$	$Q_{11}$
$\Omega_3$	$Q_3$	$Q_{16}$	$Q_3$	$Q_{11}$
$\Omega_4$	$Q_7$	$Q_{16}$	$Q_3$	$Q_{11}$



(c) 画像“Mandrill” ((256 × 256) 画素, 濃淡 8[bit/pel])

図 2.2 テスト画像



(c) 画像“Mandrill”

図 2.3 PSNR 対ビットレート特性による符号化性能の比較

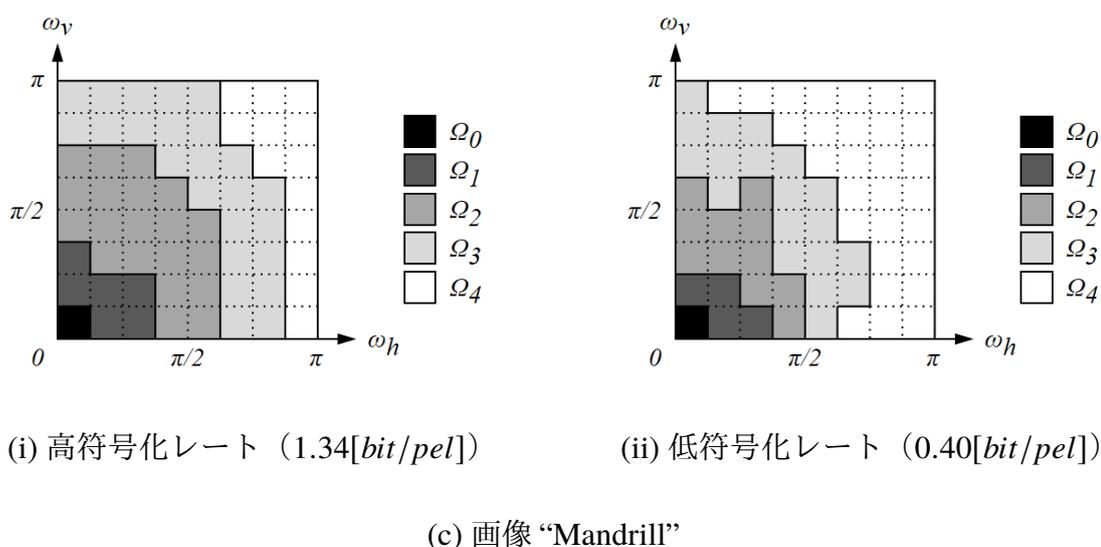


図 2.4 レート歪み理論に基づいた最適帯域分割による帯域分割パターン ( $M = 5, N = 64$ )

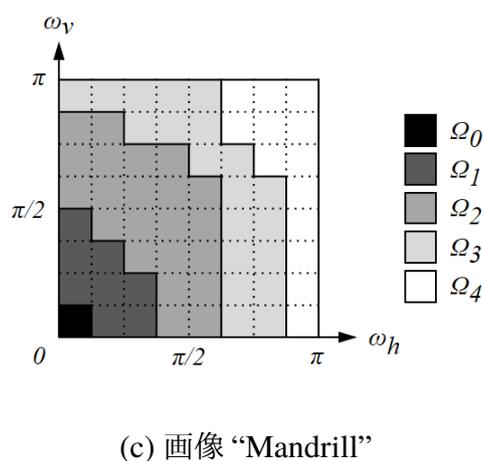


図 2.5 従来の最適帯域分割により求められた帯域分割パターン ( $M = 5, N = 64$ )

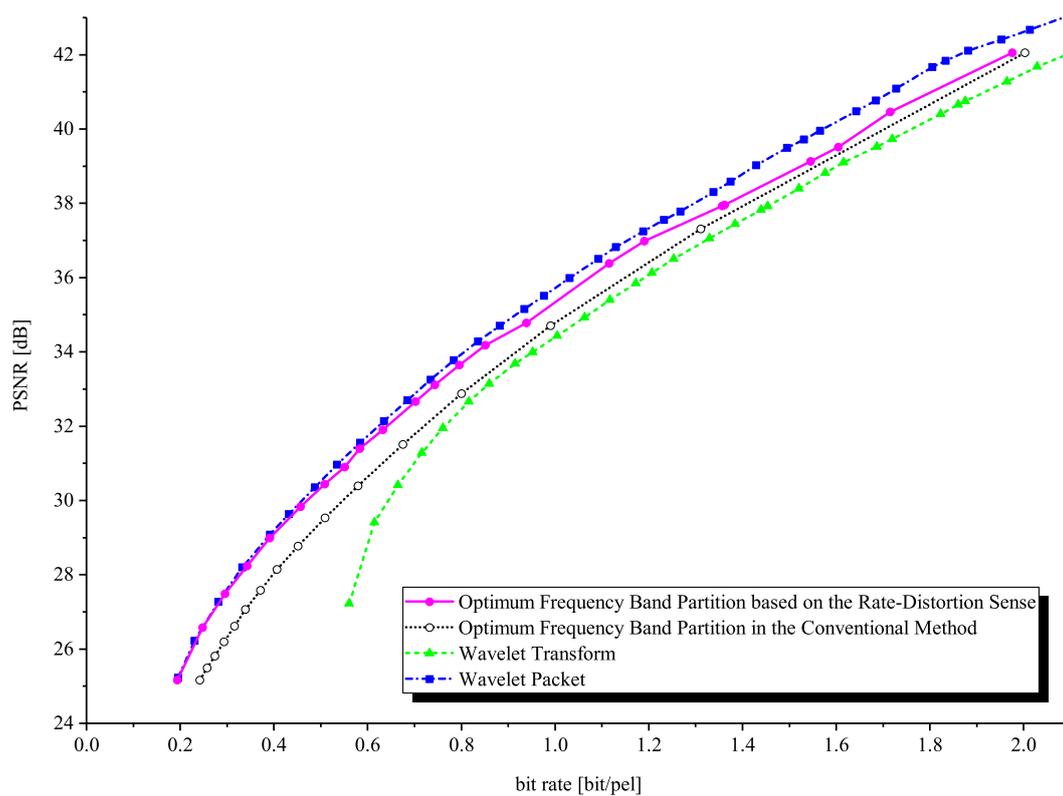
表 2.3 図 2.4(c) と図 2.5(c) に示す各サブバンド信号に適用される量子化器

subband	quantizer applied to each subband signal in frequency band partition pattern			
	Figure 2.4(c)		Figure 2.5(c)	
	(i) high bit rate (1.34[bit/pel])	(ii) low bit rate (0.40[bit/pel])	(i) high bit rate (1.36[bit/pel])	(ii) low bit rate (0.42[bit/pel])
$\Omega_0$	$Q_3$	$Q_7$	$Q_4$	$Q_{11}$
$\Omega_1$	$Q_3$	$Q_8$	$Q_4$	$Q_{11}$
$\Omega_2$	$Q_4$	$Q_{10}$	$Q_4$	$Q_{11}$
$\Omega_3$	$Q_4$	$Q_{13}$	$Q_4$	$Q_{11}$
$\Omega_4$	$Q_6$	$Q_{16}$	$Q_4$	$Q_{11}$



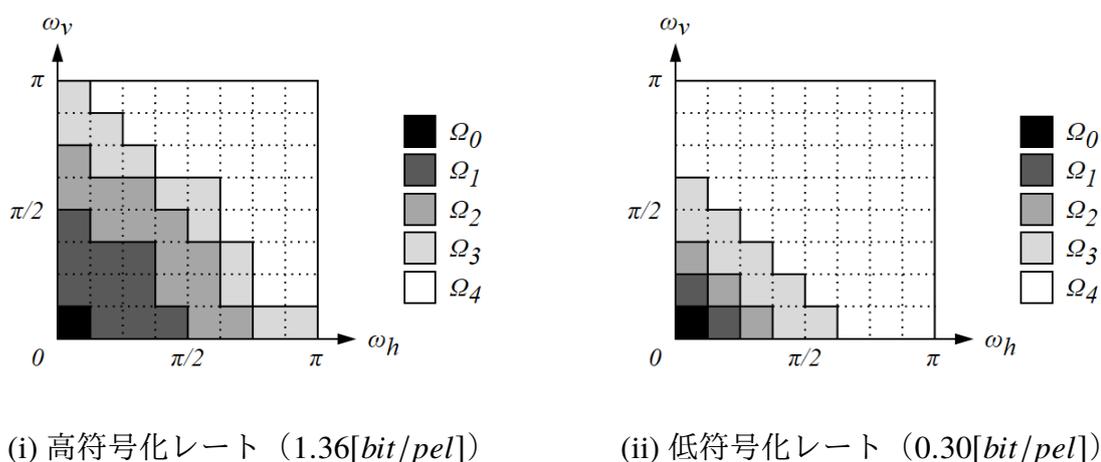
(d) 画像“Sailboat” ((256 × 256) 画素, 濃淡 8[bit/pel])

図 2.2 テスト画像



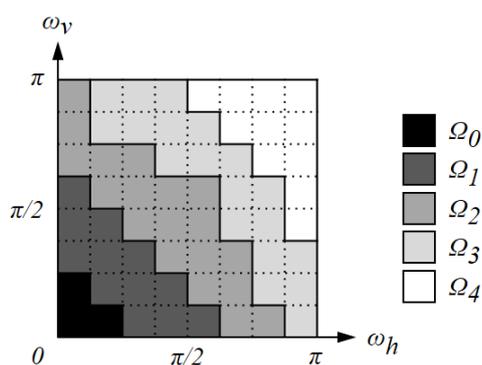
(d) 画像“Sailboat”

図 2.3 PSNR 対ビットレート特性による符号化性能の比較



(d) 画像“Sailboat”

図 2.4 レート歪み理論に基づいた最適帯域分割による帯域分割パターン ( $M = 5, N = 64$ )

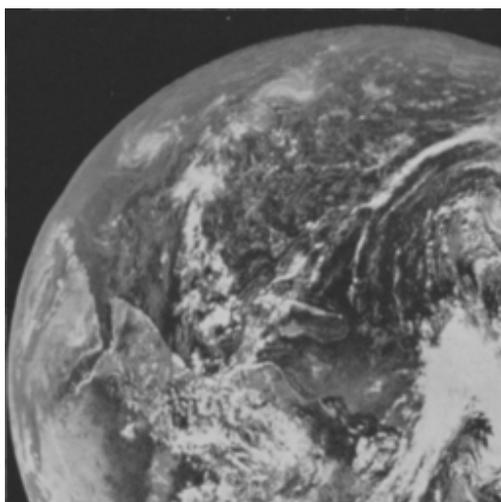


(d) 画像“Sailboat”

図 2.5 従来の最適帯域分割により求められた帯域分割パターン ( $M = 5, N = 64$ )

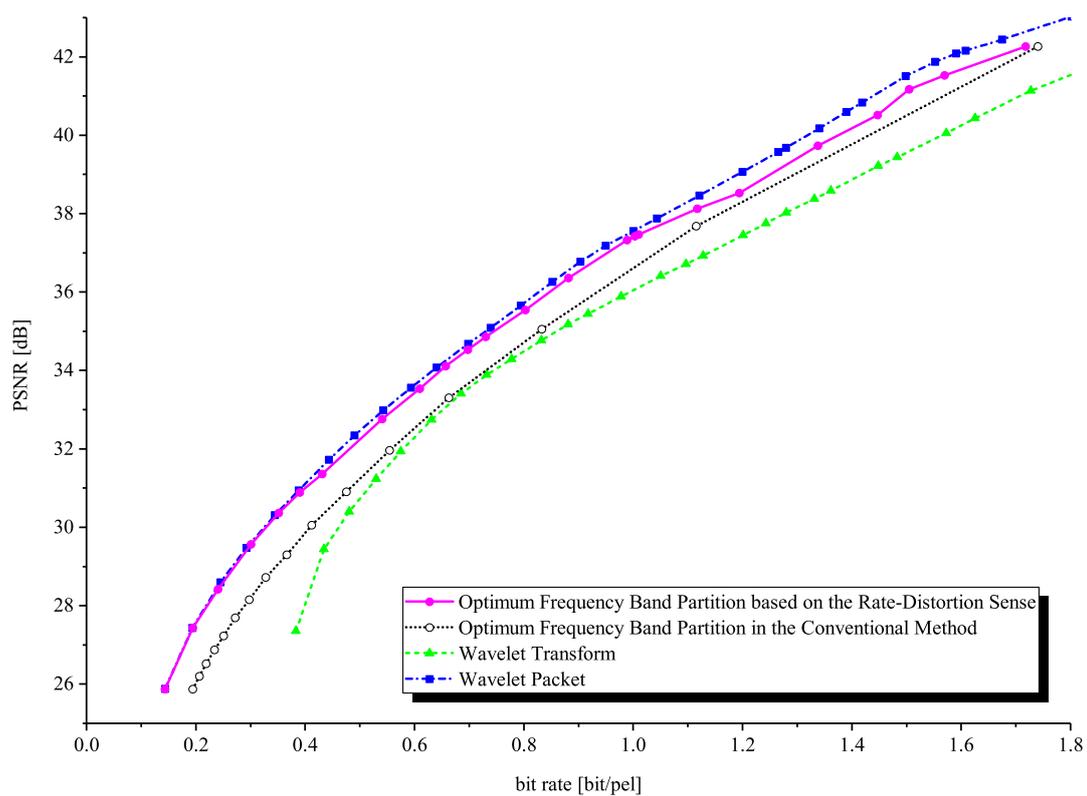
表 2.4 図 2.4(d) と図 2.5(d) に示す各サブバンド信号に適用される量子化器

subband	quantizer applied to each subband signal in frequency band partition pattern			
	Figure 2.4(d)		Figure 2.5(d)	
	(i) high bit rate (1.36[bit/pel])	(ii) low bit rate (0.30[bit/pel])	(i) high bit rate (1.31[bit/pel])	(ii) low bit rate (0.29[bit/pel])
$\Omega_0$	$Q_1$	$Q_7$	$Q_2$	$Q_{13}$
$\Omega_1$	$Q_1$	$Q_8$	$Q_2$	$Q_{13}$
$\Omega_2$	$Q_2$	$Q_9$	$Q_2$	$Q_{13}$
$\Omega_3$	$Q_2$	$Q_{12}$	$Q_2$	$Q_{13}$
$\Omega_4$	$Q_5$	$Q_{16}$	$Q_2$	$Q_{13}$



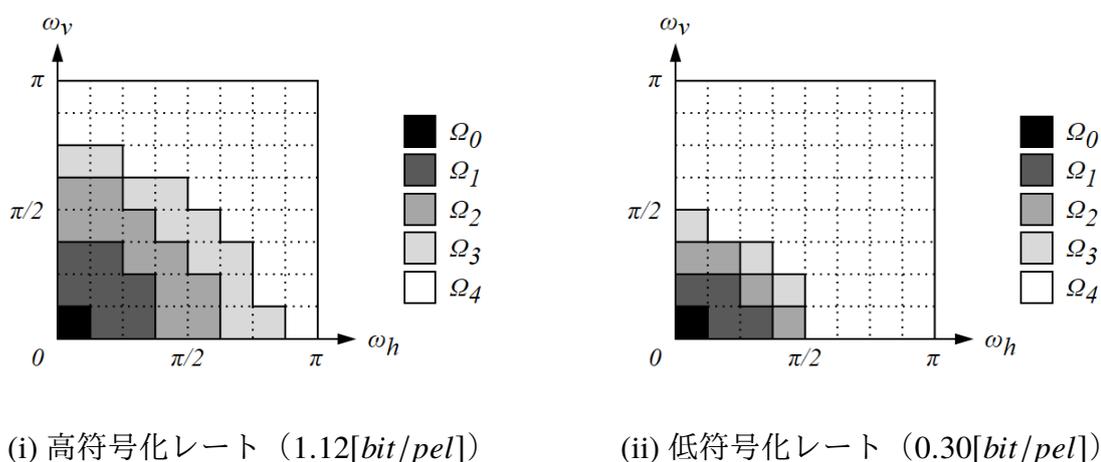
(e) 画像“Earth” ((256 × 256) 画素, 濃淡 8[bit/pel])

図 2.2 テスト画像



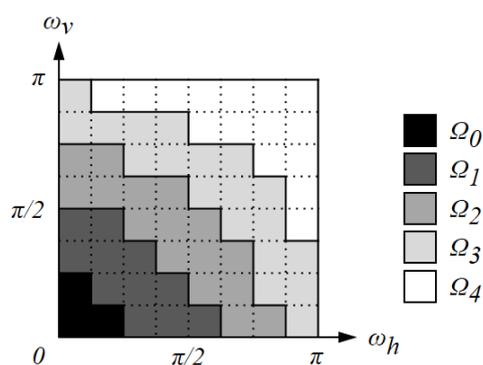
(e) 画像“Earth”

図 2.3 PSNR 対ビットレート特性による符号化性能の比較



(e) 画像“Earth”

図 2.4 レート歪み理論に基づいた最適帯域分割による帯域分割パターン ( $M = 5, N = 64$ )

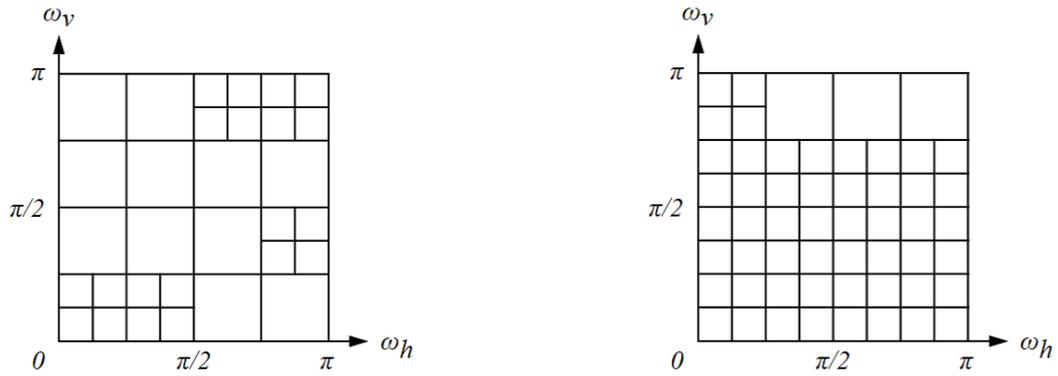


(e) 画像“Earth”

図 2.5 従来の最適帯域分割により求められた帯域分割パターン ( $M = 5, N = 64$ )

表 2.5 図 2.4(e) と図 2.5(e) に示す各サブバンド信号に適用される量子化器

subband	quantizer applied to each subband signal in frequency band partition pattern			
	Figure 2.4(e)		Figure 2.5(e)	
	(i) high bit rate (1.12[bit/pel])	(ii) low bit rate (0.30[bit/pel])	(i) high bit rate (1.12[bit/pel])	(ii) low bit rate (0.30[bit/pel])
$\Omega_0$	$Q_1$	$Q_6$	$Q_2$	$Q_{10}$
$\Omega_1$	$Q_1$	$Q_7$	$Q_2$	$Q_{10}$
$\Omega_2$	$Q_2$	$Q_7$	$Q_2$	$Q_{10}$
$\Omega_3$	$Q_2$	$Q_8$	$Q_2$	$Q_{10}$
$\Omega_4$	$Q_4$	$Q_{16}$	$Q_2$	$Q_{10}$



(a) 高符号化レート (1.14[bit/pel])

(b) 低符号化レート (0.31[bit/pel])

図 2.6 Wavelet Packet により求められた帯域分割パターン (画像 “Lenna”)

参考として、Wavelet Packet により求められた実際の帯域分割パターンを確認する。図 2.6 は、画像 “Lenna” に対して、高符号化レートと低符号化レートそれぞれの場合について、Wavelet Packet による二次元周波数帯域  $\Omega$  の分割パターンを示しており、実線の格子で区切られた各領域がサブバンドを表している。図 2.6 に示されているように、Wavelet Packet では、異なる大きさを持つ複数のサブバンドによって二次元周波数帯域  $\Omega$  が分割されている。これに対して、レート歪み理論に基づいた最適帯域分割では、図 2.4 に示すように、フィルタバンクの適用により二次元周波数帯域  $\Omega$  をあらかじめ指定された分解レベルの深さまで再帰的に分解し、それによって得られた一定のサイズを持つ帯域ブロックを基本要素として各サブバンドを構成していることから、Wavelet Packet のように各サブバンドの大きさを自由に選択できるわけではない。その結果として、高符号化レートの場合に、レート歪み理論に基づいた最適帯域分割の符号化性能が Wavelet Packet のものよりも若干低下したと考えられる。

一方、エントロピー符号化器設計の処理コストという観点から、レート歪み理論に基づいた最適帯域分割と Wavelet Packet の比較を行う。サブバンド符号化に基づく符号化システムでは、二次元周波数帯域をいくつかのサブバンドに分割及び量子化した後、量子化された各サブバンド信号に対して、符号量最小化の観点で適切なエントロピー符号化器が独立に設計される。このとき、設計される符号化器の数は、帯域分割パターンのサブバンド数と等しくなることから、サブバンド数が多くなればなるほど、それに伴ってエントロピー符号化器の設計に要する処理コストも増大する。したがって、エントロピー符号化器設計の処理コストを抑えた符号化システムを実現するためには、サブバンド数は少ない方が望ましい。この観点で Wavelet Packet を見た場合、Wavelet Packet は符号化性能の改善を優先してサブバンド数に制限をかけていないことから、決定される帯域分割パターンは多数のサブバンドによって構成される。図 2.7 は、図 2.2(a)~(e) に示した画像 “Lenna”, “Barbara”, “Mandrill”, “Sailboat”, “Earth” に対して Wavelet Packet を適用したとき、符号化レート毎に得られる帯

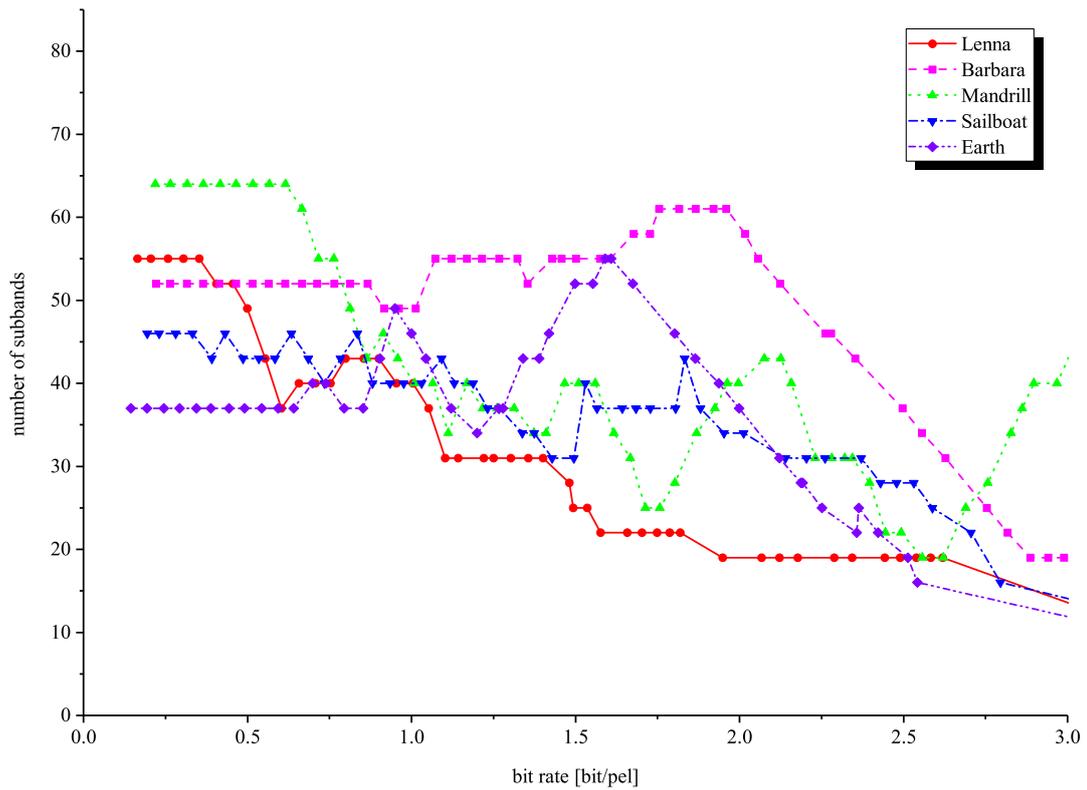


図 2.7 符号化レートと Wavelet Packet の帯域分割パターンにおけるサブバンド数の関係

帯域分割パターンのサブバンド数を示したものである。このとき、指定する符号化レートには、図 2.3 に示す PSNR 対ビットレート特性を算出するために設定した値と同じものを設定した。図 2.7 の結果から、Wavelet Packet における帯域分割パターンは、符号化レートの変更に対して変動するものの、20~64 個からなる多数のサブバンドによって構成されていることがわかる。それに対して、レート歪み理論に基づいた最適帯域分割及び従来の最適帯域分割では、全ての符号化レートにおいてサブバンド数  $M = 5$  と一定であるため、Wavelet Packet と比較して、エントロピー符号化器の設計に要する処理コストを 8~25% ( $5/64 \sim 5/20$ ) まで低く抑えることができる。

## 2.4 結言

本章では、従来において提案されている画像のサブバンド符号化における最適帯域分割の性能改善を目的として、未だ議論が行われていなかったレート歪み理論の観点から、所望の符号化レートに応じた二次元周波数帯域の最適な帯域分割パターンと、分割後の各サブバンドに適用する最適な量子化器の組み合わせを決定する「レート歪み理論に基づいた最適帯域分割」を提案し、実画像に適用することでその有効性を定量的に明らかにした。

まず、レート歪み理論における最適化問題が、ラグランジュの未定乗数法を適用することにより解が求められることを踏まえて、帯域分割パターンと量子化器の組み合わせにおける二つの最適化問題に対する定式化を行った。これにより、レート歪み理論上の帯域分割と量子化の最適化問題が、それぞれ、帯域ブロックをその確率分布の類似性に基づいてクラスタリングする問題と、サブバンド毎にラグランジュのコスト関数を最小とするように量子化器を選択する問題に帰着されることを明らかにした。次に、それを踏まえて、符号化レートが指定されたときに、これら二つの最適化問題を反復して解くことにより、符号化レートに応じた最適な帯域分割パターン及び各サブバンドに適用する最適な量子化器の組み合わせを自動的に求めるアルゴリズムを提案した。

実験においては、提案したレート歪み理論に基づいた最適帯域分割を、実画像に適用することでその符号化性能を求め、従来の最適帯域分割との比較を行った。その結果、レート歪み理論に基づいた最適帯域分割は、同程度の符号化レートを持つ従来の最適帯域分割と比べて、最大で約  $1.0[\text{dB}]$  の符号化性能の改善が得られることが明らかになった。また、JPEG2000で採用されている離散ウェーブレット変換との比較においては、特に低符号化レートでの符号化性能を大きく改善でき、さらに、帯域分割の理論的な最適解を導出できる Wavelet Packet との比較においては、サブバンド数の増加に伴うエントロピー符号化器設計の処理コストを  $8\sim 25\%$  まで低く抑えた上で、低符号化レートにおいてほぼ同等の符号化性能を達成できることが明らかになった。以上により、本章において提案したレート歪み理論に基づいた最適帯域分割は、従来の量子化雑音電力のみに基づいた評価値に代わって、レート歪み理論の観点から帯域分割と量子化の最適化を行うことで、従来の最適帯域分割から符号化性能を改善できる手法であることが明らかになった。

しかしながら、レート歪み理論に基づいた最適帯域分割は、サブバンド数が少数であるという制約の下で帯域分割及び量子化を求めているため、サブバンド数を制限せずに符号化性能の改善を優先する Wavelet Packet よりも、高い符号化性能を達成することは極めて困難である。そこで、次章では、サブバンド信号には空間領域情報が保持されているというサブバンド符号化における一つの大きな利点を活用し、帯域分割及び量子化による冗長性の削減だけでなく、空間領域上における冗長削減を実現することで符号化性能を改善する「空間-周波数領域の最適分割」について議論を行うこととする。

## 第 3 章

# レート歪み理論に基づいた画像のサブバンド符号化における空間-周波数領域の最適分割

### 3.1 序言

画像のサブバンド符号化において、エントロピー符号化器設計の処理コストに直結するサブバンド数を少数に抑えたまま、入力画像の特性に応じて二次元周波数帯域を分割する最適帯域分割が提案されている [59]～[61]。第 2 章では、従来の最適帯域分割の符号化性能を改善することを目的として、符号化レートと量子化による画質劣化量の相互関係を表すレート歪み理論の観点から、所望の符号化レートに応じて二次元周波数帯域の分割パターンと各サブバンド信号に適用する量子化器の組み合わせを最適化する「レート歪み理論に基づいた最適帯域分割」を提案した。本最適帯域分割を実画像に適用した結果、従来の最適帯域分割から、同符号化レートにおいて PSNR 値で約 1.0[*dB*] の符号化性能の改善が得られた。しかしながら、サブバンド数に制限がなく、サブバンド符号化における帯域分割の理論的最適解を導出できる Wavelet Packet [45] との性能比較においては、低符号化レートでは同等の符号化性能を得られるものの、高符号化レートでは符号化性能が劣化するという結果になった。このとき、最適帯域分割においては、あらかじめ指定するサブバンド数を増加させることにより更なる符号化性能の改善を得ることができるが、サブバンド数の増加はエントロピー符号化器の設計に要する処理コストの増加に直接結びつくものであり、さらに、サブバンド数の増加に伴うその符号化性能が比較的少ないサブバンド数において飽和する傾向にあるため [60]、符号化性能改善のための有効な手段であるとは言い難い。

ところで、最適帯域分割を含めたサブバンド符号化には、変換後の領域である各サブバンド信号において、周波数領域の情報だけではなく空間領域の情報を同時に保持できるという大きな利点がある。最適帯域分割を一般の自然画像に適用すると、各サブバンドに落ちる信号電力は空間領域上でエッジやテクスチャ等の部分に集中するため、サブバンド内にエネルギーが高い領域と低い領域が二次元空間領域上で局在する。第 2 章で提案したものも含め、

これまでに提案されてきた最適帯域分割法では、サブバンド内にある全ての信号が保存されるため、二次元空間領域の冗長性を除去しきれていないと考えられる。これを考慮して、各サブバンド信号を二次元空間領域上で保存すべき信号とそうでない信号に分類し、不要な信号を切り捨てて必要な信号のみを符号化すれば、更なる符号化性能の改善が期待できる。ここで、空間領域の冗長削減を考慮したサブバンド符号化に、Space-Frequency Quantization (SFQ) [67] がある。この方式もまた、帯域分割後の各サブバンド信号を保存すべきか否かで分類した後、必要な信号にのみ符号化を適用する。しかしながら、SFQ では、二次元周波数帯域の分割にウェーブレット変換が採用されており、入力画像や符号化レートの違いに対して帯域分割パターンは固定となっている。また、各サブバンド信号に適用される量子化器は、最低域のサブバンドを除いて全て同一のものが適用される仕様となっており。これらの同時最適化については議論されていない。

本章では、各サブバンド信号における空間領域の冗長性が十分に除去しきれていないことに着目して、第2章で議論した二次元周波数帯域の分割パターンと各サブバンド信号に適用する量子化器の組み合わせに加えて、各サブバンド内の不要な信号を切り捨てて、必要な信号のみを符号化することで冗長削減を実現する二次元空間領域分割の三つを、レート歪み理論の観点で最適化する「空間-周波数領域の最適分割」を提案する。空間-周波数領域の最適分割においてサブバンド数が増えることはないの、エントロピー符号化器設計のための処理コストの増加は生じない。まず、二次元周波数帯域分割、量子化器の組み合わせ、二次元空間領域分割の三つを同時に解くためのグローバルな最適化問題を定式化する。このとき、グローバルな最適化問題を解いて三つの解を同時に求めることは困難であるため、提案手法では、定式化されたグローバルな最適化問題から、それぞれ三つのローカルな最適化問題を個別に解くための理論式を導出する。その後、それら三つを指定された符号化レートに到達するまで反復して解くことによって、符号化レートに応じたそれぞれの最適解を導出するためのアルゴリズムを開発する。最後に、実画像に対して空間-周波数領域の最適分割を適用したときの符号化性能を PSNR 対ビットレート特性により評価し、その結果を第2章で提案したレート歪み理論に基づいた最適帯域分割、Wavelet Packet, SFQ, 更に画像符号化の国際標準方式である JPEG 及び JPEG2000 のものと比較する。さらに、各手法を実画像に適用したときの処理時間を計測し、処理コストの面からも性能評価を行うことで、空間-周波数領域の最適分割の有効性を定量的に明らかにする。

## 3.2 レート歪み理論に基づいた周波数帯域分割・量子化・空間領域分割の最適化

最適帯域分割によって分割される各サブバンド  $\Omega_k (k = 0, 1, \dots, M-1)$  は, QMF [62] 等の既存の帯域分割フィルタを用いて均等に分割された二次元周波数帯域  $\Omega$  の小領域である帯域ブロック  $\Delta\Omega_i (i = 0, 1, \dots, N-1; N \geq M)$  の集合によって構成される. 図 3.1 は, 第 2 章の図 2.2(a) に示した画像 “Lenna” ((256 × 256) 画素, 濃淡 8[bit/pel]) に対して,  $N = 64$  の条件で生成された二次元空間領域上における低域側と高域側の二つの帯域ブロック信号の係数値の例を示したグラフである. 図 3.1 を観察すると, 帯域ブロック内の信号は, 係数値の絶対値が大きい領域と小さい領域が二次元空間領域上で局在している. このことから, 従来の最適帯域分割のようにサブバンド内の全ての信号を符号化することは, 空間領域上における冗長削減の観点では望ましいことではない.

本節では, レート歪み理論の観点で, 二次元周波数帯域の分割パターン, 各サブバンド信号に適用する量子化器の組み合わせ, 空間領域上における冗長削減のためのサブバンド信号の二次元空間領域分割の三つを同時に最適化する空間-周波数領域の最適分割の理論的な導出と, その実現のために, 指定された符号化レートに応じてそれら三つの解を与えるためのアルゴリズムについて述べる. 図 3.2 に, サブバンド数  $M = 4$ , 帯域ブロック数  $N = 64$  の条件を指定した場合の空間-周波数領域の最適分割の符号化構成を示す. ここで, 空間領域分割によって削除された信号は, 復号化側でゼロ値に補間するものとする. このとき, 受信された符号情報から信号を再構成するためには, 各サブバンド信号が保存されているか否かを判別する必要がある. そのため, 各サブバンド信号に対する保存と削除の分類情報は, 付加情報として符号化し, 復号化側に伝送する.

一般に, 符号化におけるレート歪み理論の最適化問題は, 符号化レート  $R$  が指定されている制約条件の下で, 歪み  $D$  を最小化する問題とみなされ, ラグランジュの未定乗数法を適用

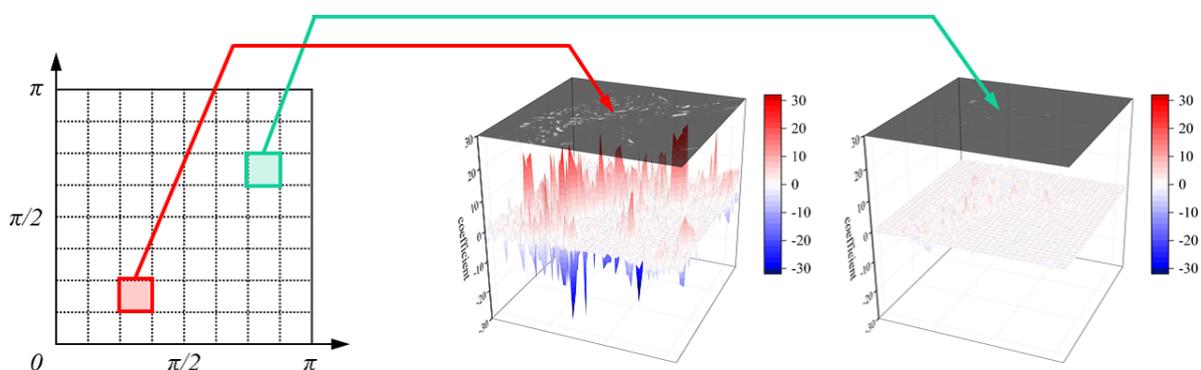


図 3.1 二次元空間領域上における帯域ブロック信号の係数値 (画像 “Lenna”)

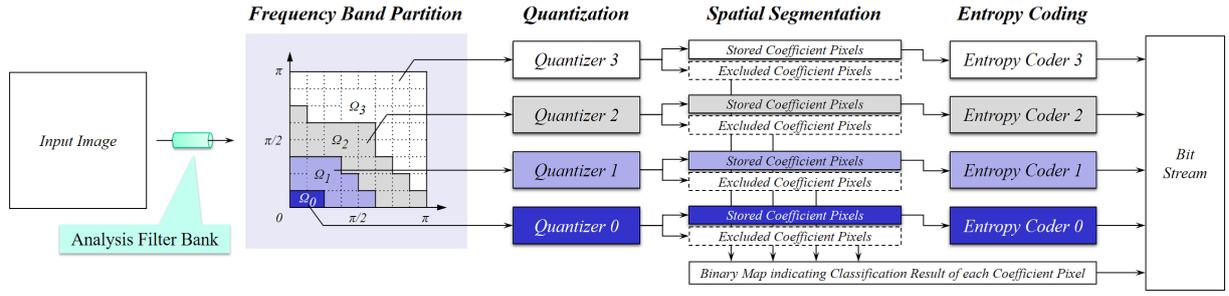


図 3.2 空間-周波数領域の最適分割の符号化構成 ( $M = 4, N = 64$ )

することで,

$$J_{opt}(\lambda) = \min[D + \lambda R] \quad (3.1)$$

として定義される制約条件のないコスト関数の最小化問題に置き換えることができる [38]. ここで,  $\lambda$  はラグランジュ乗数である. 最適帯域分割によって求められる  $M$  個のサブバンド  $\Omega_k$  は, 直交フィルタにより分割された帯域ブロック  $\Delta\Omega_i$  の集合によって構成され, 互いに直交性を満たすことから, 符号化後の再生画像において発生する全体の歪み  $D$  は各サブバンド  $\Omega_k$  における歪み  $D_k$  の総和により与えられる. さらに, サブバンド毎に独立にエントロピー符号化が適用されることを留意すれば, 全体のレート  $R$  についても, 各サブバンド  $\Omega_k$  におけるレート  $R_k$  の総和として求められる. ここで, 歪み  $D_k$  は量子化前後間の信号の平均二乗誤差 (Mean Squared Error : MSE), レート  $R_k$  はサブバンドの平均情報量 (エントロピー) に基づいて計算を行う. このとき, 空間領域分割の処理については, 信号を削除したときに発生する歪みは量子化前の信号そのものの二乗値として与えられ, レートはサブバンド  $\Omega_k$  内で保存される信号のみから算出されるエントロピーにその信号数  $N_k$  を乗じたものとして定義される. また, サブバンド  $\Omega_k$  で保存される信号群が, それに所属する全ての帯域ブロック  $\Delta\Omega_i (i \in I_k)$  毎に保存される信号群の集合と等価であることを踏まえれば, 空間-周波数領域の最適分割の問題は,

$$J_{opt}(\lambda) = \min \left[ \sum_{k=0}^{M-1} D_k + \lambda \sum_{k=0}^{M-1} R_k \right] \\ = \min \left[ \sum_{k=0}^{M-1} \left\{ \sum_{i \in I_k, j \in S_i} (w_{i,j} - \hat{w}_{i,j}^{q_k})^2 + \sum_{i \in I_k, j \notin S_i} w_{i,j}^2 \right\} \right. \\ \left. + \lambda \sum_{k=0}^{M-1} \left( -N_k \sum_{x=x_k^{min}}^{x_k^{max}} \tilde{\Omega}_k(x) \log_2 \tilde{\Omega}_k(x) \right) \right] \quad (3.2)$$

として定式化される. ここで,  $I_k$  は,  $\Omega_k$  に含まれる帯域ブロック番号  $i$  の集合,  $w_{i,j}$  と  $\hat{w}_{i,j}^{q_k}$  はそれぞれ, 帯域ブロック  $\Delta\Omega_i$  の  $j$  番目の信号における量子化前の係数値と, それをサブバンド  $\Omega_k$  に対する量子化器  $q_k$  で量子化したときの係数値であり,  $S_i$  は, 帯域ブロック  $\Delta\Omega_i$  内

で保存される信号番号  $j$  の集合を表す（逆に、 $S_i$  の補集合は帯域ブロック  $\Delta\Omega_i$  内で削除される信号番号  $j$  の集合に相当する）。また、 $\tilde{\Omega}_k(x)$  は、サブバンド  $\Omega_k$  で保存される信号群を  $q_k$  で量子化したときに求められる信号の確率分布であり、 $x_k^{min}$  と  $x_k^{max}$  は、それぞれ  $\tilde{\Omega}_k(x) \neq 0$  となる  $x$  の最小値と最大値である。式 (3.2) において、 $I_k$  は二次元周波数帯域  $\Omega$  上の分割パターン、 $q_k$  はサブバンド  $\Omega_k$  に対する量子化器、 $S_i$  は帯域ブロック  $\Delta\Omega_i$  における空間領域上の分割パターンの結果にそれぞれ対応しており、式 (3.2) の解を求めることでこれら三つを同時に最適化することが可能となる。

しかしながら、式 (3.2) 中の一つの最適化問題を解くためには、残り二つの最適化問題の解を必要とする相互の依存関係があることから、式 (3.2) を直接解いて三つの解を同時に最適化することは極めて困難である。具体的に、一つ目の二次元周波数帯域  $\Omega$  上の分割パターン  $I_k$  を決定するためには、式 (3.2) にあるように信号の確率分布を必要とするが、 $q_k$  と  $S_i$  が未知である場合には、確率分布を生成するための信号群とそれらの係数値が特定できないため、 $I_k$  を求めることはできない。二つ目の量子化器  $q_k$  を決定するためには、各サブバンド  $\Omega_k$  を量子化したときに算出される歪み  $D_k$  とレート  $R_k$  の各値が必要となるが、 $I_k$  と  $S_i$  が未知である場合、対象となる帯域ブロック及び信号群が特定できないため、 $q_k$  を求めることはできない。三つ目の空間領域上の分割パターン  $S_i$  を決定するためには、Space-Frequency Quantization (SFQ) と同様、各信号の歪みとレートの値に基づいて、信号を保存した場合のコスト関数を定義する必要があるが、 $I_k$  と  $q_k$  が未知である場合、歪みとレートの値が特定できないため、 $S_i$  を求めることはできない。そこで、一つの最適化問題を解く際には、残り二つの最適化問題は既に解かれていると仮定した上で、三つの最適化問題を個別に定式化し、それらを指定のレートに到達するまで反復的に解くことによってそれぞれの解を求める手法を提案する。このとき、反復の途中で仮決定された各最適化問題の解を、 $I_k^{(*)}, q_k^{(*)}, S_i^{(*)}$  のように  $(*)$  を付けて表すこととする。

### 3.2.1 周波数帯域分割の最適化

$q_k^{(*)}$  と  $S_i^{(*)}$  が既に決定済みであるという条件下で、二次元周波数帯域  $\Omega$  上の最適な分割パターン  $I_k$  を求める場合、画像全体の歪み  $D$  は分割パターンの違いに依存しないため、式 (3.2) における第 1 項は定数として扱われる。また、式 (3.2) における  $\tilde{\Omega}_k(x)$  は、サブバンド  $\Omega_k$  に属する帯域ブロック  $\Delta\Omega_i$  の中で、保存される信号群のみから求められた確率分布  $\Delta\tilde{\Omega}_i(x) (i \in I_k)$  を統合した混合確率分布と等価であることから、

$$\tilde{\Omega}_k(x) = \sum_{i \in I_k} \frac{\Delta N_i}{N_k} \Delta\tilde{\Omega}_i(x) \quad (3.3)$$

に置き換えられる。ここで、 $\Delta N_i$  は帯域ブロック  $\Delta\Omega_i$  内で保存されている信号数を表す。更に、 $q_k^{(*)}$  と  $S_i^{(*)}$  が既知であれば確率分布  $\Delta\tilde{\Omega}_i(x)$  と  $\Delta N_i$  が分割パターンの変化に対して不変に

なることから、 $-p \log q = -p \log p + p \log(p/q)$  という関係を考慮すれば、式 (3.2) は

$$\begin{aligned}
J_{opt}(\lambda) &= \min \left[ \lambda \sum_{k=0}^{M-1} \left\{ -N_k \sum_{x=x_i^{min}}^{x_i^{max}} \left( \sum_{i \in I_k} \frac{\Delta N_i}{N_k} \Delta \tilde{\Omega}_i(x) \right) \log_2 \tilde{\Omega}_k(x) \right\} \right] \\
&= \min \left[ \begin{aligned} &\lambda \sum_{i=0}^{N-1} \Delta N_i \left( - \sum_{x=x_i^{min}}^{x_i^{max}} \Delta \tilde{\Omega}_i(x) \log_2 \Delta \tilde{\Omega}_i(x) \right) \\ &+ \lambda \left( \sum_{k=0}^{M-1} \sum_{i \in I_k} \Delta N_i \sum_{x=x_i^{min}}^{x_i^{max}} \Delta \tilde{\Omega}_i(x) \log_2 \frac{\Delta \tilde{\Omega}_i(x)}{\tilde{\Omega}_k(x)} \right) \end{aligned} \right] \quad (3.4)
\end{aligned}$$

のように展開できる．ここで、式 (3.4) の第 1 項は、帯域ブロック  $\Delta \Omega_i$  内で保存される信号のみから算出されるエントロピーと  $\Delta N_i$  の積の総和、すなわち、帯域ブロック  $\Delta \Omega_i$  におけるレート<sup>1</sup>の総和であり、これも分割パターンの違いに依存しないので、定数となり無視できる．また、第 2 項は、帯域ブロックの確率分布  $\Delta \Omega_i(x)$  とそれらが所属するサブバンドの確率分布  $\Omega_k(x)$  への Kullback-Leibler Divergence [64] の総和に  $\Delta N_i$  を乗じたものである．このとき、二つの確率分布  $P(x)$  と  $Q(x)$  の間の距離として定義される Kullback-Leibler Divergence を

$$D_{KL}(P \parallel Q) = \sum_x P(x) \log \frac{P(x)}{Q(x)} \quad (3.5)$$

のように表記した上で式 (3.4) をまとめると、 $q_k^{(*)}$  と  $s_i^{(*)}$  が既に決定済みであるという条件下で、二次元周波数帯域  $\Omega$  の最適な分割パターン  $I_k$  を求めることは、 $\Delta N_i$  で重みを付けられた Kullback-Leibler Divergence を距離尺度として、確率分布  $\Delta \tilde{\Omega}_i(x)$  をクラスタリングする問題

$$I_k = \arg \min_{I_k} \left[ \sum_{k=0}^{M-1} \sum_{i \in I_k} \Delta N_i D_{KL}(\Delta \tilde{\Omega}_i \parallel \tilde{\Omega}_k) \right] \quad (3.6)$$

に帰着される．

### 3.2.2 量子化の最適化

$I_k^{(*)}$  と  $s_i^{(*)}$  が既に決定済みであるという条件下で、サブバンド  $\Omega_k$  に対する最適な量子化器  $q_k$  を求める場合、 $s_i^{(*)}$  によって保存される信号が全て特定されているため、式 (3.2) における第 1 項中の  $\sum_{i \in I_k, j \notin S_i} w_{i,j}^2$  は定数となり無視できる．また、 $I_k^{(*)}$  に従って構成されたサブバンド  $\Omega_k$  における歪み  $D_k$  とレート  $R_k$  は、サブバンド単位で独立に算出されるため、当問題は、サブバンド  $\Omega_k$  毎にラグランジュのコスト関数を最小とするような量子化器を個別に選

択する問題

$$q_k = \arg \min_{q_k} \left[ \sum_{i \in I_k^{(*)}, j \in S_i^{(*)}} (w_{i,j} - \hat{w}_{i,j}^{q_k})^2 + \lambda \left( -N_k \sum_{x=x_k^{min}}^{x_k^{max}} \tilde{\Omega}_k(x) \log_2 \tilde{\Omega}_k(x) \right) \right] \quad (3.7)$$

として与えられる.

### 3.2.3 空間領域分割の最適化

$I_k^{(*)}$  と  $q_k^{(*)}$  が既に決定済みであるという条件下では, サブバンド信号毎に歪みとレートの値が特定されているため, 信号を保存した場合のコスト関数を定義することが可能である. 帯域ブロック  $\Delta\Omega_i$  における空間領域上の最適な集合  $S_i$  は, 信号を保存した場合と削除した場合のコスト関数を比較し, 全体のコストが最小となるように全ての信号を保存あるいは削除の二つに分類することによって決定される. しかしながら, 復号化側において, 保存された信号を適切な位置に当てはめ, 削除された信号をゼロ値で補間するためには, 分類の結果を表す二値の付加情報が信号毎に必要となる. このとき, 全ての信号に対する付加情報を保存してしまうと, それだけで全体の情報量に  $1[\text{bit}/\text{pel}]$  が加わってしまう. SFQ ではこの問題を回避するため, ゼロツリー関係に基づいてサブバンド間の信号に木構造を定義した後, 親ノードの信号が削除された場合にはそれ以下にある子や孫ノードの信号を全て削除することと, 同じ親ノードを持つ複数の子ノード (ただし, 孫ノードは含まれない) 間で分類の情報を共有するという条件を付けることで, 符号化性能の低下を抑えつつ全体の付加情報量を大幅に削減している [67]. 空間-周波数領域の最適分割でもこの考え方を利用し, 各信号の分類結果を表す付加情報量を抑えながら, 空間領域上の最適な集合  $S_i$  を決定するためのアルゴリズムを提案する. 具体的な処理を以下に述べる.

空間領域分割の最適化は, 保存される信号群と削除される信号群それぞれから求められたコストの和を最小とするような  $S_i$  を求めることが目的である. しかし, 各信号のレートは  $S_i$  の決定後に求められた確率分布  $\tilde{\Omega}_k(x)$  に従って算出されるため, 分類を行う前のコスト関数を使用して決定された  $S_i$  は真に最適ではないと考えられる. そのため, 本アルゴリズムでは, 空間領域分割と  $\tilde{\Omega}_k(x)$  を逐次的に更新することでコストを最小とする  $S_i$  を決定する. 今, 上述した二つの反復処理を  $m$  回繰り返した時点での帯域ブロック  $\Delta\Omega_i$  における  $j$  番目の信号の分類状態を二値情報  $n_{i,j}^{(m)}$  として表す. ここで,  $n_{i,j}^{(m)} = 1$  のときは保存,  $n_{i,j}^{(m)} = 0$  のときは削除を表すものとする.

i) 初期化

まず,  $n_{i,j}^{(0)} \leftarrow \{1\}$  とおく.

ii) コスト関数の算出

$m$  回目の時点で帯域ブロック  $\Delta\Omega_i$  内に保存される  $j$  番目の信号に対してのみ, 量子化

器  $q_k$  で量子化したときのラグランジュのコスト関数

$$J_{i,j}^{(m)} = \left[ w_{i,j} - \hat{w}_{i,j}^{q_k} \right]^2 + \lambda \left[ -\log_2 \left\{ \tilde{\Omega}_k^{(m)}(\hat{w}_{i,j}^{q_k}) \right\} \right] \quad \text{subject to } \left[ n_{i,j}^{(m)} = 1 \right] \quad (3.8)$$

を求める．ここで  $\tilde{\Omega}_k^{(m)}(\hat{w}_{i,j}^{q_k})$  は，サブバンド  $\Omega_k$  内で係数値  $w_{i,j}^{q_k}$  を持つ信号が発生する確率である．

iii) 信号の保存と削除

$m$  回目の時点で帯域ブロック  $\Delta\Omega_i$  内に保存される  $j$  番目の信号を親ノードとした場合，それと対応付けられた全ての子や孫ノードに対して，信号を保存した場合と削除した場合のコスト関数を

if

$$\sum_{t \in C_i} w_{t,j}^2 \leq \sum_{t \in C_i} J_{t,j}^{(m)} \quad \text{subject to } \left[ n_{i,j}^{(m)} = 1 \right]$$

then

$$n_{t,j}^{(m)} \leftarrow 0 \quad (t \in C_i) \quad (3.9)$$

により比較し，削除した方がコストが小さくなる場合は，対象となる全ての子孫ノードの信号を削除する．ここで，式(3.9)中の  $C_i$  は，ある帯域ブロック  $\Delta\Omega_i$  を親としたとき，その子や孫に相当する帯域ブロック番号  $i$  の集合を表しており，図3.3はその具体例として  $N = 16$  個の帯域ブロック間に定義されたゼロツリー関係に基づく木構造を示している．図3.3より，1個の親ノードに対する子ノードの数は，最低域の帯域ブロックでは3個，それ以外の帯域ブロックでは4個になる．空間-周波数領域の最適分割において，全ての帯域ブロックは同サイズであるため，帯域ブロック間のゼロツリー関係は，空間領域上で同じ位置にある  $j$  番目の信号同士で定義される．したがって，木構造は信号単位ではなく， $C_i$  のように帯域ブロック単位で構成すればよい．式(3.9)の判

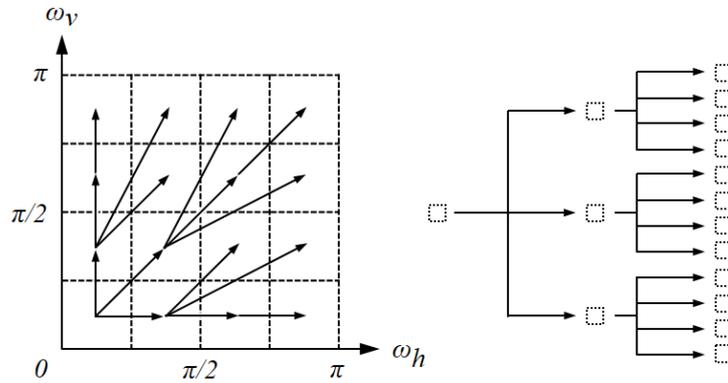


図 3.3 帯域ブロック間におけるゼロツリー関係に基づく木構造 ( $N = 16$ )

定処理は、初めは  $m$  回目の時点で保存されている全ての信号が対象であるが、処理の過程で親ノードが既に削除された子孫ノードに対しては式 (3.9) を実行しない。また、親ノードを唯一持たない直流成分を含む帯域ブロック信号は、他と比べてエネルギーが大きいために、削除するとコストが大きく増加してしまうことから、あらかじめ判定処理の対象外としている。

#### iv) 収束判定

処理 iii) によって求められた  $n_{i,j}^{(m)}$  が前回の結果  $n_{i,j}^{(m-1)}$  と比較して異なった場合、 $n_{i,j}^{(m+1)} \leftarrow n_{i,j}^{(m)}$  のように更新し、処理 ii) に戻る。  $n_{i,j}^{(m)} \neq n_{i,j}^{(m-1)}$  の場合、 $n_{i,j}^{(m)}$  のときに求められるコストは、 $n_{i,j}^{(m-1)}$  のものよりも必ず小さくなるのが文献 [67] において保証されているため、処理 ii) と iii) を逐次的に実行すればコストは単調に減少する。  $n_{i,j}^{(m)} = n_{i,j}^{(m-1)}$  となったときには、帯域ブロック  $\Delta\Omega_i$  における空間領域上の集合  $S_i$  と、各信号の分類状態を示す二値情報  $n_{i,j}$  を最適解として決定し、処理を終了する。

### 3.2.4 アルゴリズム

符号化レートが指定されたとき、それに応じて三つの最適解  $I_k, q_k, S_i$  を実際に決定するためのアルゴリズムを提案する。ラグランジュの未定乗数法により定式化されたレート歪み最適化問題においては、指定された符号化レートに到達するまでラグランジュ乗数  $\lambda$  を更新しながら式 (3.1) を反復的に解くことによって、所望の符号化レートに応じた最適解が求められる。空間-周波数領域の最適分割では、反復処理の中に式 (3.6) と式 (3.7)、及び前節 i)~iv) の処理を組み込み、それらを逐次的に解くことによって  $I_k, q_k, S_i$  を決定する。このとき、反復処理を行う求根アルゴリズムには、第 2 章で提案したレート歪み理論に基づいた最適帯域分割と同様、二分法とニュートン法に基づく手法 [65] を採用する。

まず、入力画像に対して、QMF による一次元フィルタバンクを水平及び垂直方向に適用することにより、互いに直交性を満たす  $N$  個の帯域ブロック  $\Delta\Omega_i (i = 0, 1, \dots, N-1)$  を得る。反復処理の前に、全ての帯域ブロック信号に対して、 $q_k$  の候補として用意された量子化器の異なる  $P$  種類の量子化器  $Q_p (p = 1, 2, \dots, P)$  を適用し、各帯域ブロック  $\Delta\Omega_i$  の  $j$  番目の信号を  $Q_p$  で量子化したときの係数値  $\hat{w}_{i,j}^p (p = 1, 2, \dots, P)$  を求めておく。そして、 $I_k^{(*)}, q_k^{(*)}, S_i^{(*)}$  にそれぞれ適当な初期値を与えた後、所望の符号化レートに到達するまで、以下 i)~iv) の処理を繰り返し行う。ここで、初期値として、 $q_k^{(*)}$  は中央の量子化器を持つ量子化器  $Q_{\lfloor P/2 \rfloor}$ 、 $S_i^{(*)}$  には全ての信号が保存されている状態として全ての信号番号  $j$  を与え、 $I_k^{(*)}$  は上述した  $q_k^{(*)}, S_i^{(*)}$  を使用して式 (3.6) を解いた結果を用いる。

- i)  $I_k^{(*)}, q_k^{(*)}$  を使用して、3.2.3 で述べた i)~iv) のアルゴリズムを適用し、各帯域ブロック  $\Delta\Omega_i$  における空間領域上の分割パターン  $S_i$  と、各信号の分類状態を示す二値情報  $n_{i,j}$  を求める。その後、 $S_i^{(*)} \leftarrow S_i$  と置く。

- ii)  $q_k^{(*)}, S_i^{(*)}$  を使用して,  $N$  個の帯域ブロックを  $M$  個のサブバンドに分割する式 (3.6) のクラスタリング問題を解くことで, 二次元周波数帯域上の分割パターン  $I_k$  を求める. 解の導出には, レート歪み理論に基づいた最適帯域分割と同様, k-means++法 [66] を適用する. その後,  $I_k^{(*)} \leftarrow I_k$  と置く.
- iii)  $I_k^{(*)}, S_i^{(*)}$  を使用して, 式 (3.7) を満たすように各サブバンド  $\Omega_k$  に対する量子化器  $q_k$  を求める. その後,  $q_k^{(*)} \leftarrow q_k$  と置く.
- iv) 処理 i)~iii) で決定された  $I_k, q_k, S_i$  の結果に従って, 全体の歪み  $D$  とレート  $R$  を計算する. このとき,  $D$  は式 (3.2) の第 1 項,  $R$  は式 (3.2) のラグランジュ乗数  $\lambda$  を除いた第 2 項と,  $n_{i,j}$  に当たる付加情報量の和から算出される. そして, レート  $R$  が所望の符号化レートに到達していれば, その時点での  $I_k, q_k, S_i$  をそれぞれの最適解として処理を終了し, そうでなければ, 求根アルゴリズムに従ってラグランジュ乗数  $\lambda$  を歪み  $D$  とレート  $R$  の値に基づいて更新し, 処理 i) に戻る.

最後に,  $n_{i,j}$  の符号化方法と, 復号化側での信号の復元方法について説明する. ゼロツリー関係により, 親ノードのエネルギーが高ければその子ノードの状態は  $n_{i,j} = 1$  となり, 逆に親ノードのエネルギーが低ければその子ノードの状態は  $n_{i,j} = 0$  となる傾向がある. 空間-周波数領域の最適分割では, この特徴を利用して子ノードの状態を予測することで,  $n_{i,j}$  を効率的に符号化している. まず親ノードが所属する帯域ブロックの各信号から分散値を求め, それを降順にソートした後, その並び方に対応付けて子ノードの  $n_{i,j}$  を並び替える. ここで, 分散は対象となる親ノードを中心に周囲  $3 \times 3$  画素の係数値から算出する. 並び替えられた  $n_{i,j}$  は, おおよそ前半に 1, 後半に 0 が集中するような列になるため, ランレングス符号化を用いて符号化される. 得られた  $n_{i,j}$  のバイナリデータは, 親ノードの帯域ブロック信号をエントロピー符号化することで得られるバイナリデータの直後に付加してビットストリームに送信する. 復号化側では, 受信した親ノードにあたる帯域ブロック信号のバイナリデータから各ノードの信号を再構成し, 符号化側と同様に各信号の分散値を求めて降順にソートする. そして,  $n_{i,j}$  の情報を表すバイナリデータを復号化し, それを並び替えられた各ノードに対応付けることで, 子ノードの  $n_{i,j}$  を復元する. その後, 子ノードが所属する帯域ブロック信号が受信されたら, 事前に求められている子ノードの  $n_{i,j}$  の結果に従って,  $n_{i,j} = 1$  のノードには受信した係数値を当てはめ,  $n_{i,j} = 0$  のノードはゼロ値で補間することで信号を再構成する. このとき, 復号化側で子ノードを持つ帯域ブロック信号を受信するより先に, 親ノードを持つ帯域ブロックの係数値を事前に全て再構成しておく必要がある. そのため, 二次元周波数帯域上の分割パターン  $I_k$  の決定では, 親ノードにあたる低域側の帯域ブロックが優先的に符号化・復号化されるようにサブバンドを構成している.

### 3.3 シミュレーション実験とその結果

実画像に空間-周波数領域の最適分割を適用し、その符号化性能を、第2章で提案したレート歪み理論に基づいた最適帯域分割及び従来のサブバンド符号化に基づく画像符号化方式と比較する。ここで、空間-周波数領域の最適分割及びレート歪み理論に基づいた最適帯域分割において、二次元周波数帯域上の分割パターンを求めるときの条件である帯域ブロック数  $N$  とサブバンド数  $M$  は、第2章と同様、それぞれ  $N = 64$ ,  $M = 5$  とした。また、各サブバンドに適用する量子化器には、量子化ステップ幅  $(1, 2, \dots, 16)$  を持つ  $P = 16$  種類のミッドトレッド型の線形量子化器  $Q_1 \sim Q_{16}$  を用意した。更に空間-周波数領域の最適分割は、Wavelet Packet [45] 及び SFQ [67] とその性能を比較する。このとき、いずれの手法とも各サブバンドに適用される量子化器には、空間-周波数領域の最適分割と同じ  $Q_1 \sim Q_{16}$  を用意した。なお、これらの手法においては共通して、フィルタバンクによるサブバンドの再帰分解レベルを3に設定している。この場合、SFQのサブバンド数は  $M = 10$ 、Wavelet Packet では最大で  $M = 64$  のサブバンドに分割される。

#### 3.3.1 符号化性能

符号化性能は、各手法を実画像に適用したときの PSNR 対ビットレート特性により評価する。ここで、PSNR は、入力画像と符号化後の再生画像との間の再生誤差電力を  $\sigma_e^2$  として、

$$PSNR = 10 \log_{10} \frac{255^2}{\sigma_e^2} \quad (3.10)$$

により計算される。ビットレートは、量子化や空間領域分割が行われた後の各サブバンド  $\Omega_k$  内の信号が、それぞれ無記憶情報源から生じたものと仮定して算出されたエントロピー  $H_k$  に、サブバンド  $\Omega_k$  に含まれる信号数  $N_k$  を乗じた総ビット数を入力画像の全画素数  $K$  で割った

$$bitrate = \frac{1}{K} \sum_{k=0}^{M-1} N_k H_k \quad (3.11)$$

、すなわち、1画素あたりの平均ビット数  $[bit/pel]$  として定義した。なお、上述した四つの手法によって実際に符号化を行うためには、復号化側で各サブバンドに対する量子化器の組み合わせを示す情報が必要となり、空間-周波数領域の最適分割とレート歪み理論に基づいた最適帯域分割及び Wavelet Packet については、二次元周波数帯域上の分割パターンの情報が必要となる。更に、空間-周波数領域の最適分割と SFQ では、各信号の分類状態を示す情報  $n_{i,j}$  が必要となる。

入力画像には、図3.4(a)~(f)に示す SIDBA 標準画像 [63] における画像“Lenna”，“Barbara”（いずれも  $(512 \times 512)$  画素，濃淡  $8[bit/pel]$ ），Kodak 標準画像 [68] における画像“kodim03”，“kodim23”（いずれも  $(768 \times 512)$  画素，濃淡  $8[bit/pel]$ ），AIC JPEG ad-hoc group が提供す

る高解像度の標準画像 [69] における画像 “bike”, “p01” (いずれも  $1280 \times 1600$  画素, 濃淡  $8[\text{bit}/\text{pel}]$ ) の 6 枚を選択した. 図 3.5(a)~(f) は, 図 3.4(a)~(f) の各画像に対して空間-周波数領域の最適分割を適用したときの符号化性能を, レート歪み理論に基づいた最適帯域分割, Wavelet Packet, SFQ によるものと比較した結果である. なお, 各手法による符号化性能の結果は, 必要とされる付加情報量を全て含めたものとして算出している. 図 3.5(a)~(f) の結果から, 空間-周波数領域の最適分割は他の手法と比べて, 入力画像の種類によらず全ての符号化レートで符号化性能が改善されていることがわかる. 空間-周波数領域の最適分割では空間領域分割についても最適化を行ったことで, 同じサブバンド数を指定した条件下でも, 同程度の符号化レートをもつレート歪み理論に基づいた最適帯域分割と比べて PSNR 値が最大で  $2.0[\text{dB}]$  程度の高い値を示している. 更に空間-周波数領域の最適分割は, サブバンド数に制限を設けない Wavelet Packet よりも高い符号化性能が得られており, サブバンド数を増加することなくレート歪み理論に基づいた最適帯域分割の問題を解決できることが明らかになった. また, SFQ との比較においても, 空間-周波数領域の最適分割は特に低符号化レートで符号化性能が大きく改善されていることが確認できる.

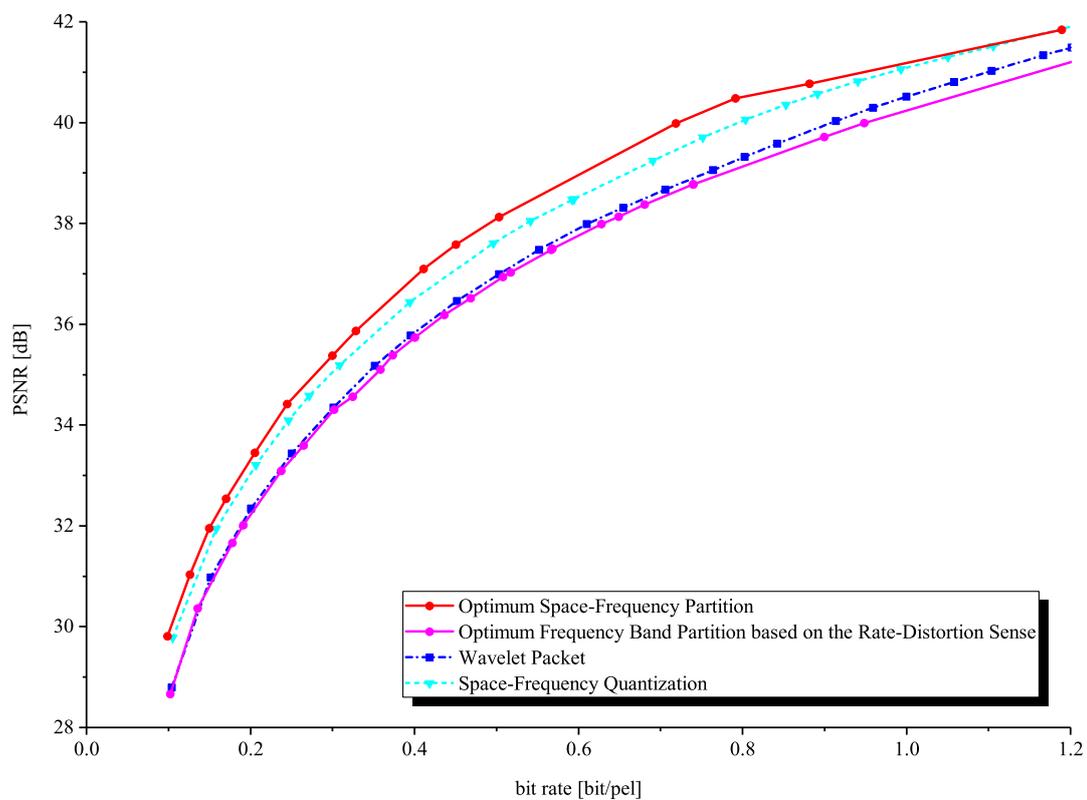
図 3.6(a)~(f) は, 図 3.4(a)~(f) の各画像に対して, (i) 高符号化レートと (ii) 低符号化レートとしてそれぞれ異なる符号化レートを指定したとき, 空間-周波数領域の最適分割によって求められた二次元周波数帯域上の分割パターンを示している. また, 図 3.7(a)~(f) は, 同じ二種類の符号化レートを指定したとき, 第 2 章で提案したレート歪み理論に基づいた最適帯域分割によって求められた二次元周波数帯域上の分割パターンを示している. 表 3.1~表 3.6 には, 図 3.6(a)~(f) 及び図 3.7(a)~(f) に示した各サブバンド信号に適用される量子化器の組み合わせ結果を示している. ここで, 表 3.1~表 3.6 における - の記号は, 該当するサブバンド内の信号が空間領域分割によって全て削除されたために, 量子化器が割り当てられていないことを表している. 図 3.8(a)~(f) は, 図 3.6(a)~(f) に示した各帯域ブロック信号における二次元空間領域上の分割パターンを示している. ここで, 図 3.6(a)~(f) に示した空間領域上の分割パターンでは, それぞれ黒画素が保存される信号, 図中の右側に示す各サブバンドに対応付けて色付けされた画素が該当するサブバンド内で削除された信号を表している. これらの結果において, 符号化レートの違いによって異なる内容が求められていることから, 三つの要素それぞれが指定の符号化レートに応じて最適化されていることがわかる. また, 空間-周波数領域の最適分割によって求められた二次元周波数帯域上の分割パターンの結果と, 各サブバンドに対する量子化器の組み合わせ結果は, 図 3.7(a)~(f) 及び表 3.1~表 3.6 に示したレート歪み理論に基づいた最適帯域分割によるものと異なっていることが確認できる. これは, サブバンド内の全ての信号ではなく, 空間領域分割によって保存された信号群のみに対して周波数領域分割と量子化の最適化処理を行ったことによるものである. 次に, 新たに加わった空間領域上の分割パターンの結果を考察する. 同じ符号化レートを指定したときに得られた 64 個の帯域ブロックにおける空間領域分割パターンの結果を比較すると, 低域側よりも高域側にある帯域ブロックの方が多くの信号が削除されている. これは, 低域側に存

在する親ノードにあたる信号が削除されたときには、高域側にある子ノードの信号も削除されるという条件によるものと、高域側の帯域ブロックでは、量子化によりゼロ値に丸め込まれる小さい係数値を持つ信号、すなわち保存するよりも削除した方がコストをより低減できる信号が多く含まれているためである。次に、(i) 高符号化レートと (ii) 低符号化レートの二つの結果から、二次元周波数帯域上で同じ位置にある帯域ブロックの領域分割パターンの結果を比較すると、高符号化レートよりも低符号化レートを指定したときに多くの信号が削除されている。これは、低符号化レートを指定した場合、各サブバンドに対して量子化ステップ幅が大きい量子化器が割り当てられ、ほとんどの信号の係数値がゼロ値に丸め込まれるため、削除した方がコストを低減できる信号が増加するためである。



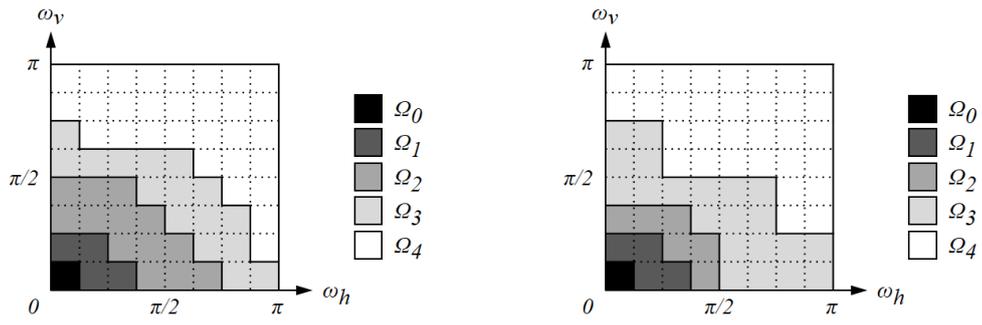
(a) 画像“Lenna” ((512 × 512) 画素, 濃淡 8[bit/pel])

図 3.4 テスト画像



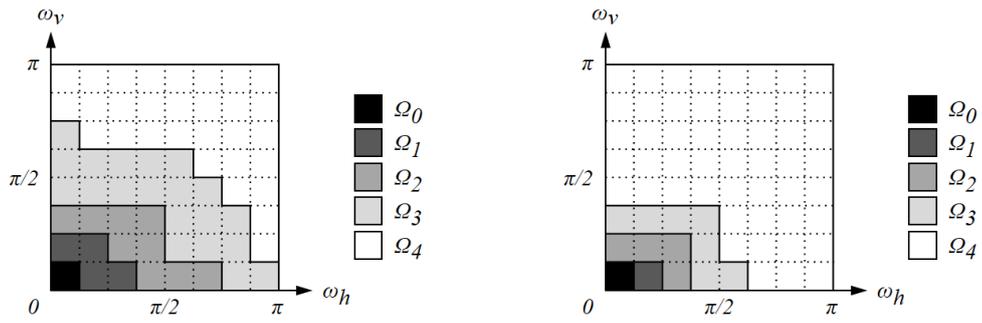
(a) 画像“Lenna”

図 3.5 PSNR 対ビットレート特性による符号化性能の比較



(i) 高符号化レート (0.72[bit/pel])      (ii) 低符号化レート (0.13[bit/pel])  
 (a) 画像“Lenna”

図 3.6 空間-周波数領域の最適分割による帯域分割パターン ( $M = 5, N = 64$ )

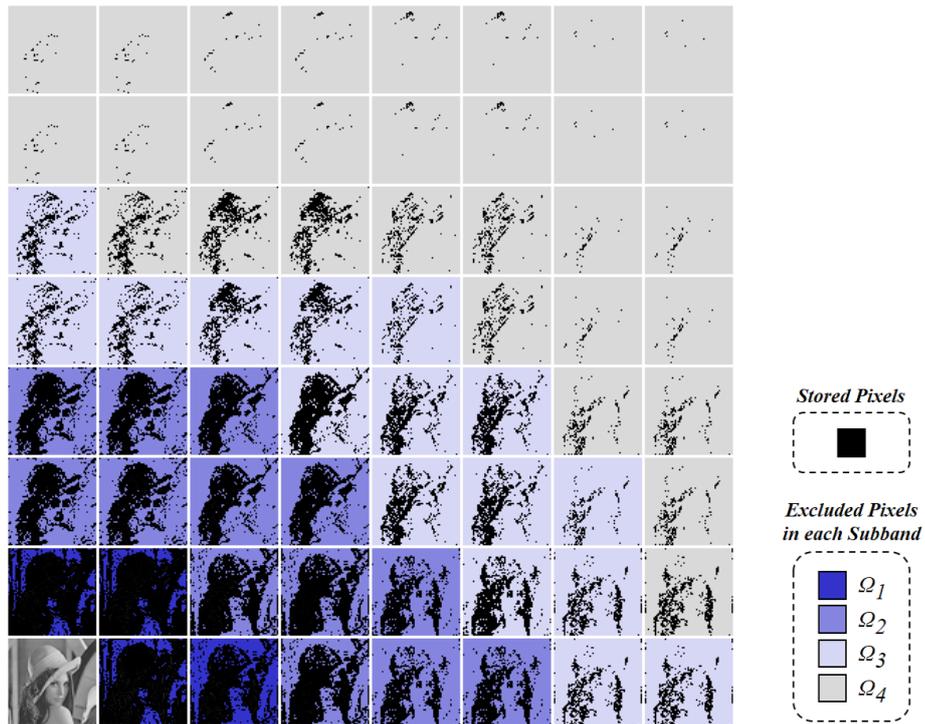


(i) 高符号化レート (0.74[bit/pel])      (ii) 低符号化レート (0.11[bit/pel])  
 (a) 画像“Lenna”

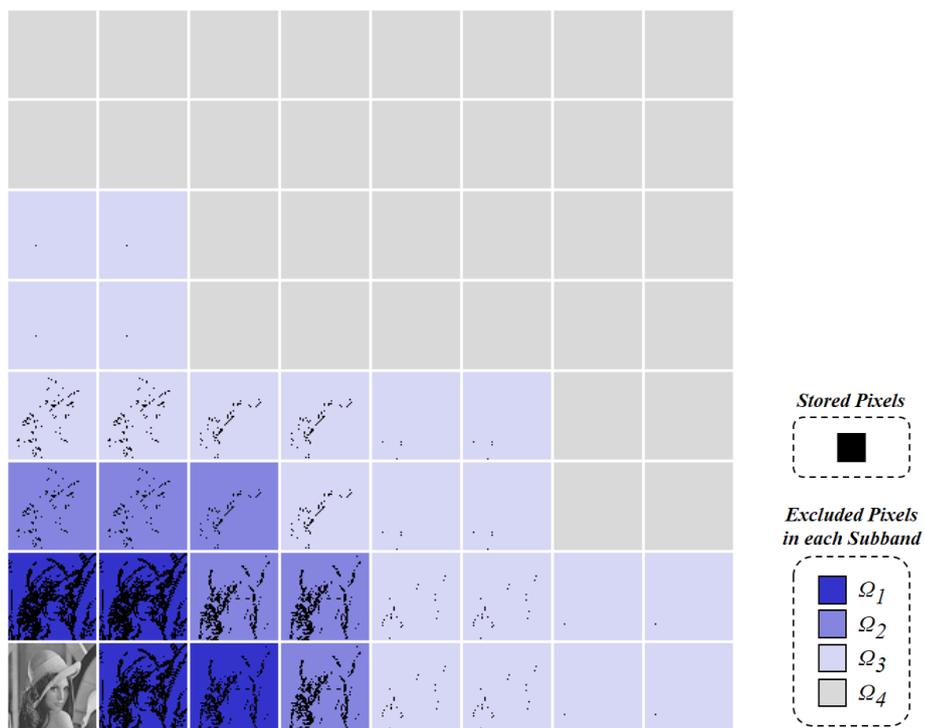
図 3.7 レート歪み理論に基づいた最適帯域分割による帯域分割パターン ( $M = 5, N = 64$ )

表 3.1 図 3.6(a) と図 3.7(a) に示す各サブバンド信号に適用される量子化器

subband	quantizer applied to each subband signal in frequency band partition pattern			
	Figure 3.6(a)		Figure 3.7(a)	
	(i) high bit rate (0.72[bit/pel])	(ii) low bit rate (0.13[bit/pel])	(i) high bit rate (0.74[bit/pel])	(ii) low bit rate (0.11[bit/pel])
$\Omega_0$	$Q_1$	$Q_8$	$Q_1$	$Q_{12}$
$\Omega_1$	$Q_1$	$Q_9$	$Q_1$	$Q_{14}$
$\Omega_2$	$Q_1$	$Q_9$	$Q_2$	$Q_{16}$
$\Omega_3$	$Q_1$	$Q_{10}$	$Q_2$	$Q_{16}$
$\Omega_4$	$Q_2$	—	$Q_3$	$Q_{16}$



(i) 高符号化レート (0.72[bit/pel])



(ii) 低符号化レート (0.13[bit/pel])

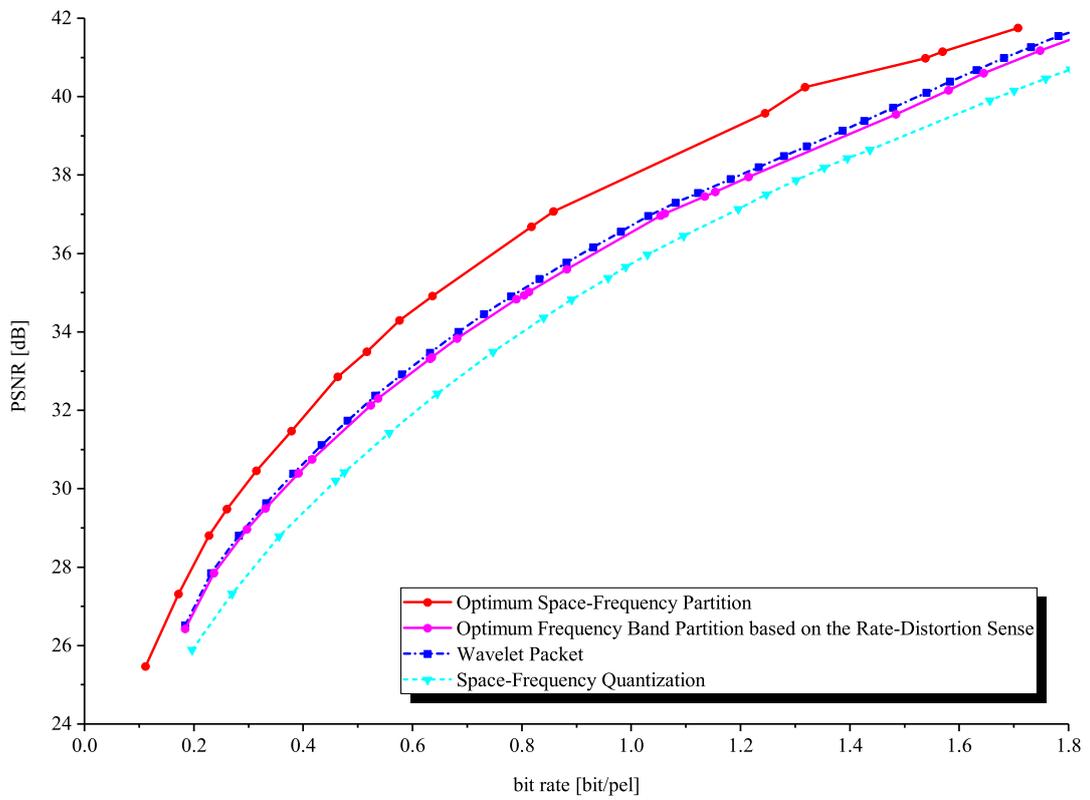
(a) 画像“Lenna”

図 3.8 空間-周波数領域の最適分割による空間領域分割パターン ( $M = 5$ ,  $N = 64$ )



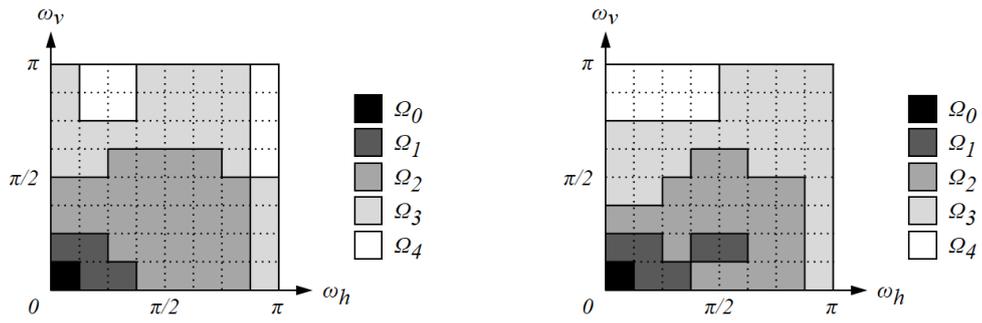
(b) 画像“Barbara” ((512 × 512) 画素, 濃淡 8[bit/pel])

図 3.4 テスト画像



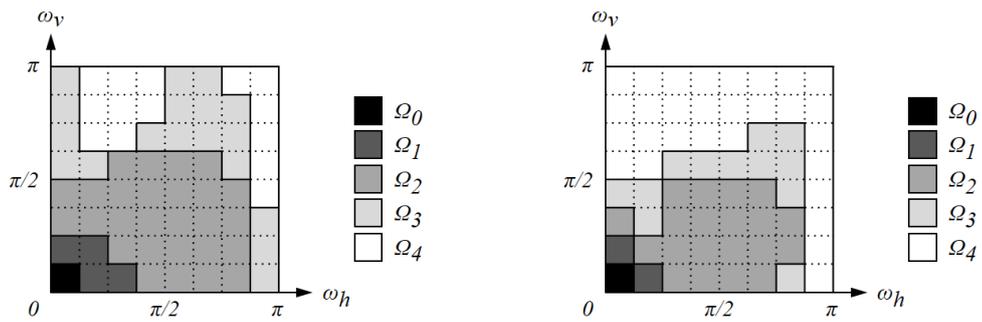
(b) 画像“Barbara”

図 3.5 PSNR 対ビットレート特性による符号化性能の比較



(i) 高符号化レート (0.82[bit/pel])      (ii) 低符号化レート (0.23[bit/pel])  
(b) 画像 “Barbara”

図 3.6 空間-周波数領域の最適分割による帯域分割パターン ( $M = 5, N = 64$ )

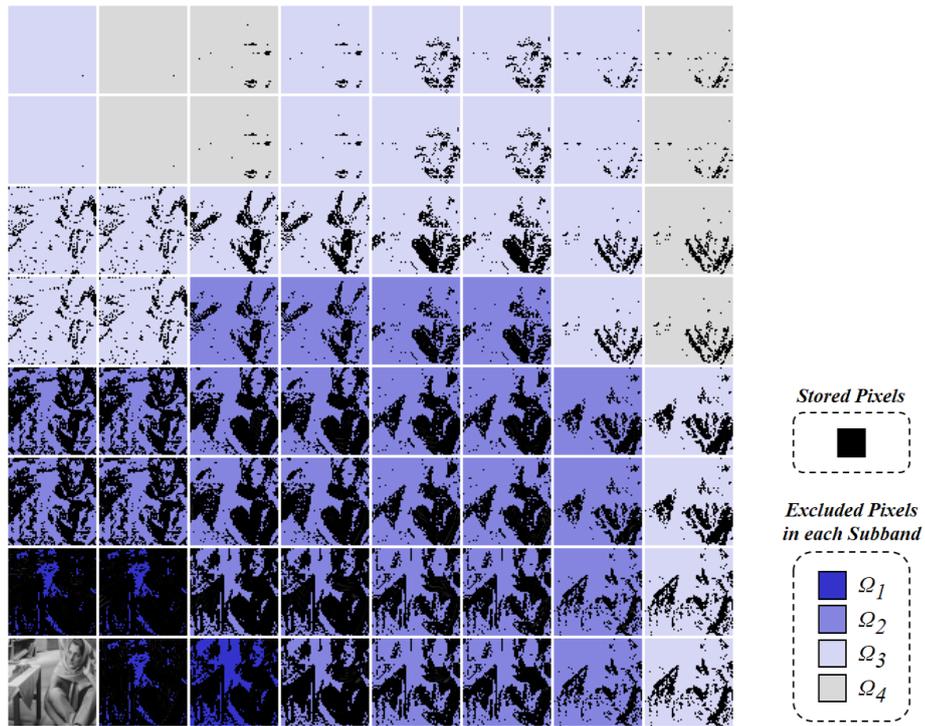


(i) 高符号化レート (0.81[bit/pel])      (ii) 低符号化レート (0.24[bit/pel])  
(b) 画像 “Barbara”

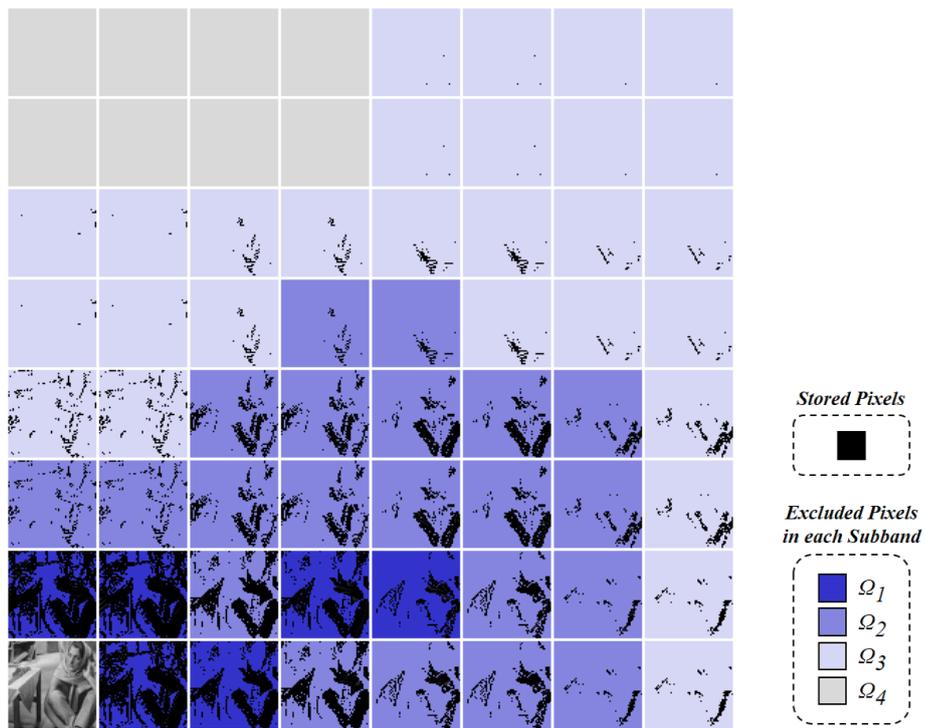
図 3.7 レート歪み理論に基づいた最適帯域分割による帯域分割パターン ( $M = 5, N = 64$ )

表 3.2 図 3.6(b) と図 3.7(b) に示す各サブバンド信号に適用される量子化器

subband	quantizer applied to each subband signal in frequency band partition pattern			
	Figure 3.6(b)		Figure 3.7(b)	
	(i) high bit rate (0.82[bit/pel])	(ii) low bit rate (0.23[bit/pel])	(i) high bit rate (0.81[bit/pel])	(ii) low bit rate (0.24[bit/pel])
$\Omega_0$	$Q_2$	$Q_8$	$Q_2$	$Q_8$
$\Omega_1$	$Q_2$	$Q_9$	$Q_3$	$Q_{11}$
$\Omega_2$	$Q_2$	$Q_9$	$Q_3$	$Q_{13}$
$\Omega_3$	$Q_3$	$Q_{10}$	$Q_4$	$Q_{15}$
$\Omega_4$	$Q_4$	—	$Q_6$	$Q_{16}$



(i) 高符号化レート (0.82[bit/pel])



(ii) 低符号化レート (0.23[bit/pel])

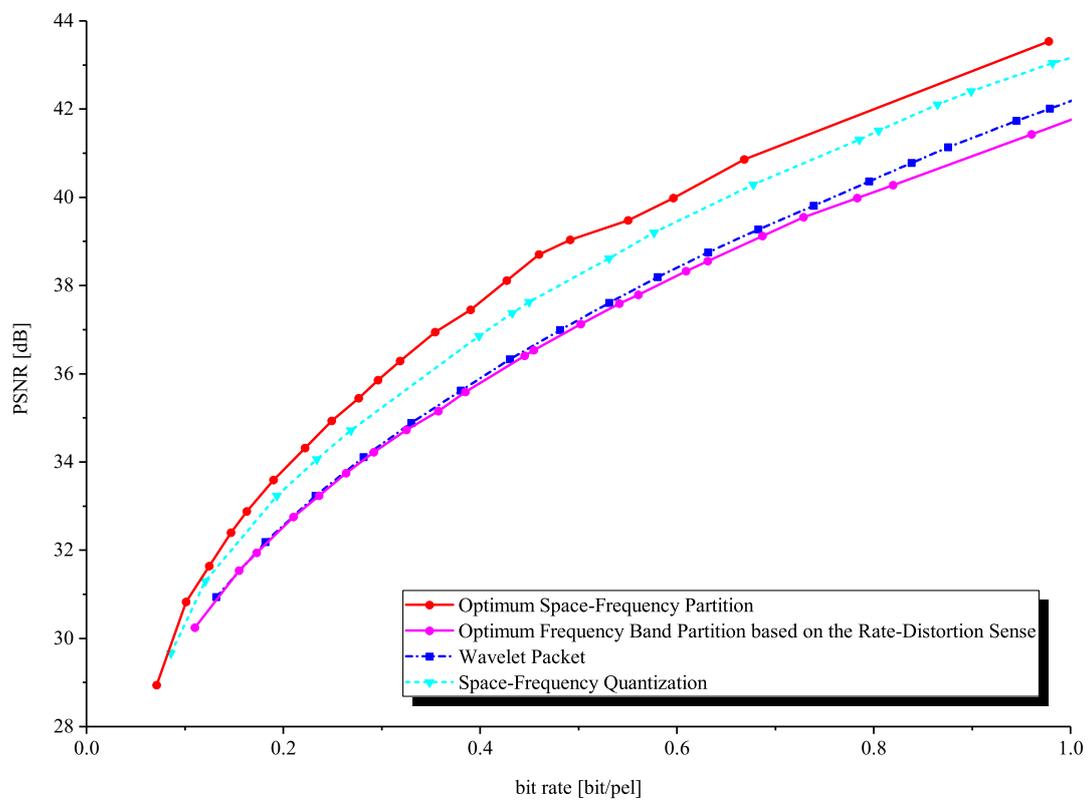
(b) 画像 “Barbara”

図 3.8 空間-周波数領域の最適分割による空間領域分割パターン ( $M = 5, N = 64$ )



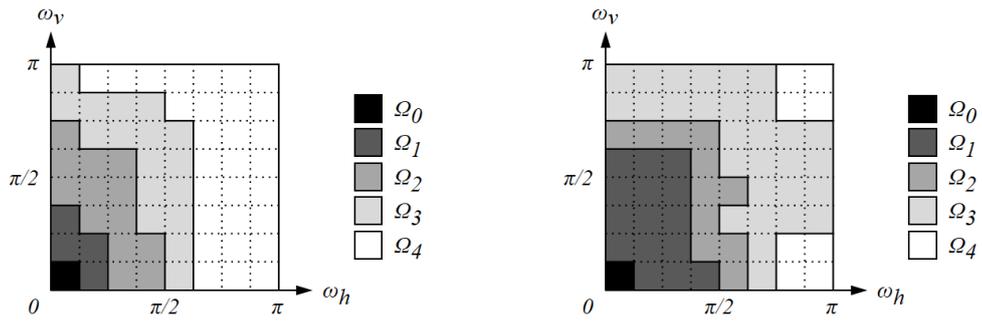
(c) 画像 “kodim03” ((768 × 512) 画素, 濃淡 8[bit/pel])

図 3.4 テスト画像



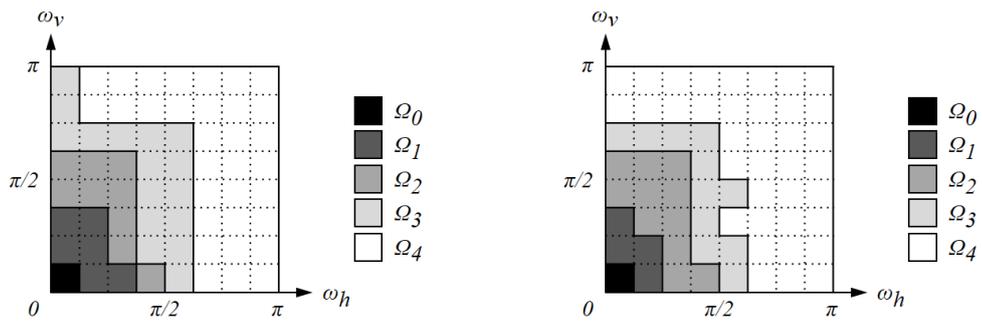
(c) 画像 “kodim03”

図 3.5 PSNR 対ビットレート特性による符号化性能の比較



(i) 高符号化レート (0.67[bit/pel])      (ii) 低符号化レート (0.22[bit/pel])  
(c) 画像“kodim03”

図 3.6 空間-周波数領域の最適分割による帯域分割パターン ( $M = 5, N = 64$ )

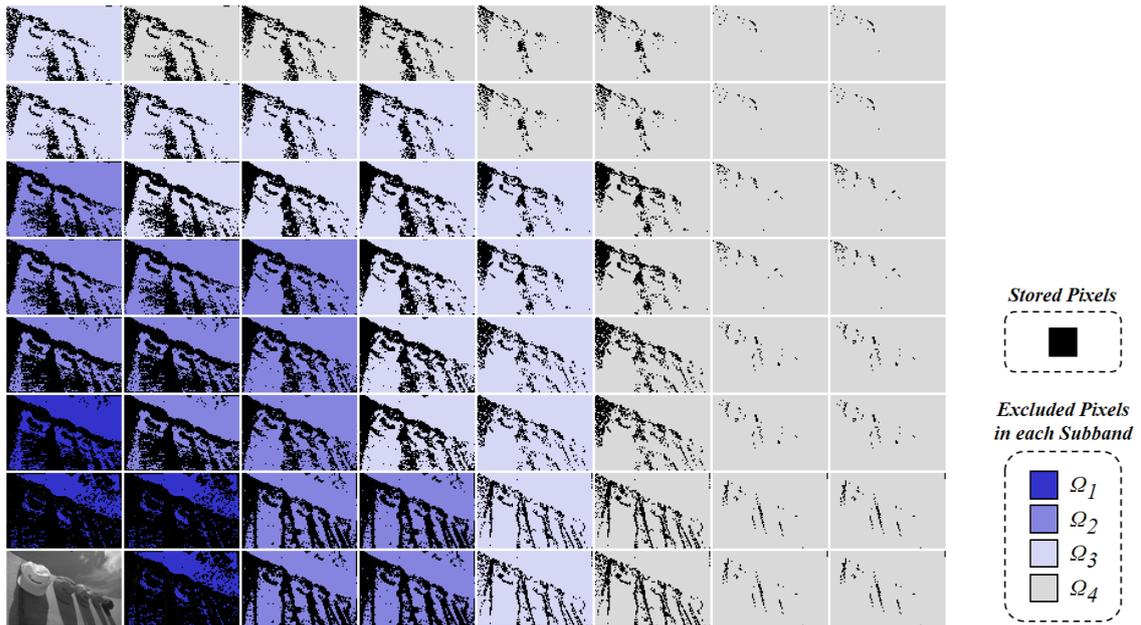


(i) 高符号化レート (0.68[bit/pel])      (ii) 低符号化レート (0.21[bit/pel])  
(c) 画像“kodim03”

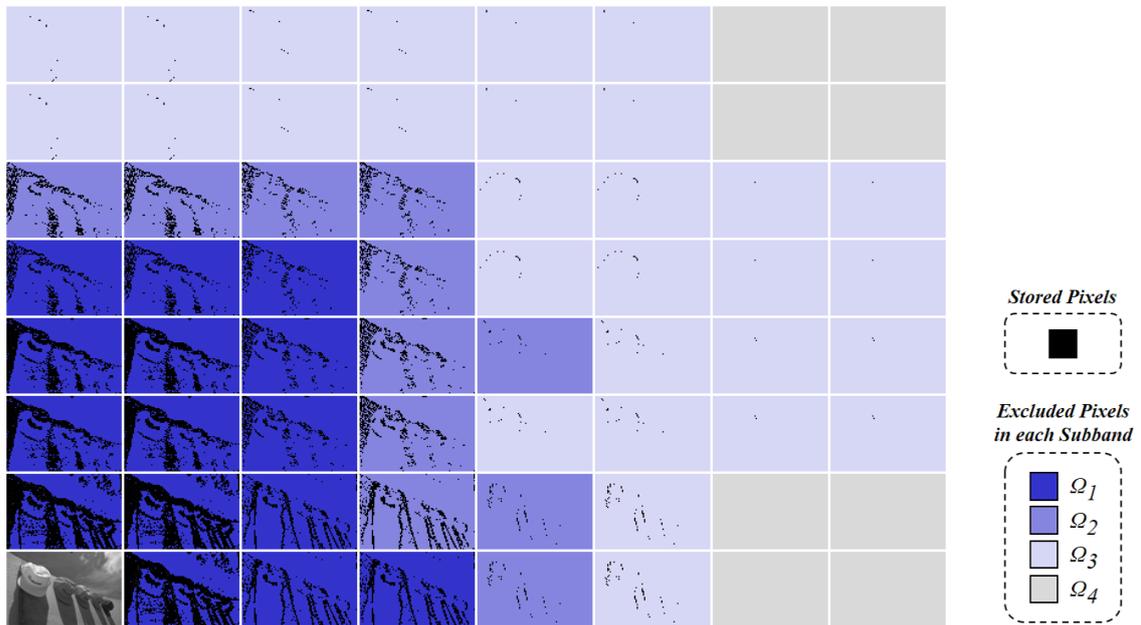
図 3.7 レート歪み理論に基づいた最適帯域分割による帯域分割パターン ( $M = 5, N = 64$ )

表 3.3 図 3.6(c) と図 3.7(c) に示す各サブバンド信号に適用される量子化器

subband	quantizer applied to each subband signal in frequency band partition pattern			
	Figure 3.6(c)		Figure 3.7(c)	
	(i) high bit rate (0.67[bit/pel])	(ii) low bit rate (0.22[bit/pel])	(i) high bit rate (0.68[bit/pel])	(ii) low bit rate (0.21[bit/pel])
$\Omega_0$	$Q_2$	$Q_4$	$Q_2$	$Q_5$
$\Omega_1$	$Q_2$	$Q_4$	$Q_2$	$Q_6$
$\Omega_2$	$Q_1$	$Q_5$	$Q_2$	$Q_7$
$\Omega_3$	$Q_1$	$Q_6$	$Q_2$	$Q_{11}$
$\Omega_4$	$Q_2$	—	$Q_3$	$Q_{16}$



(i) 高符号化レート (0.67[bit/pel])



(ii) 低符号化レート (0.22[bit/pel])

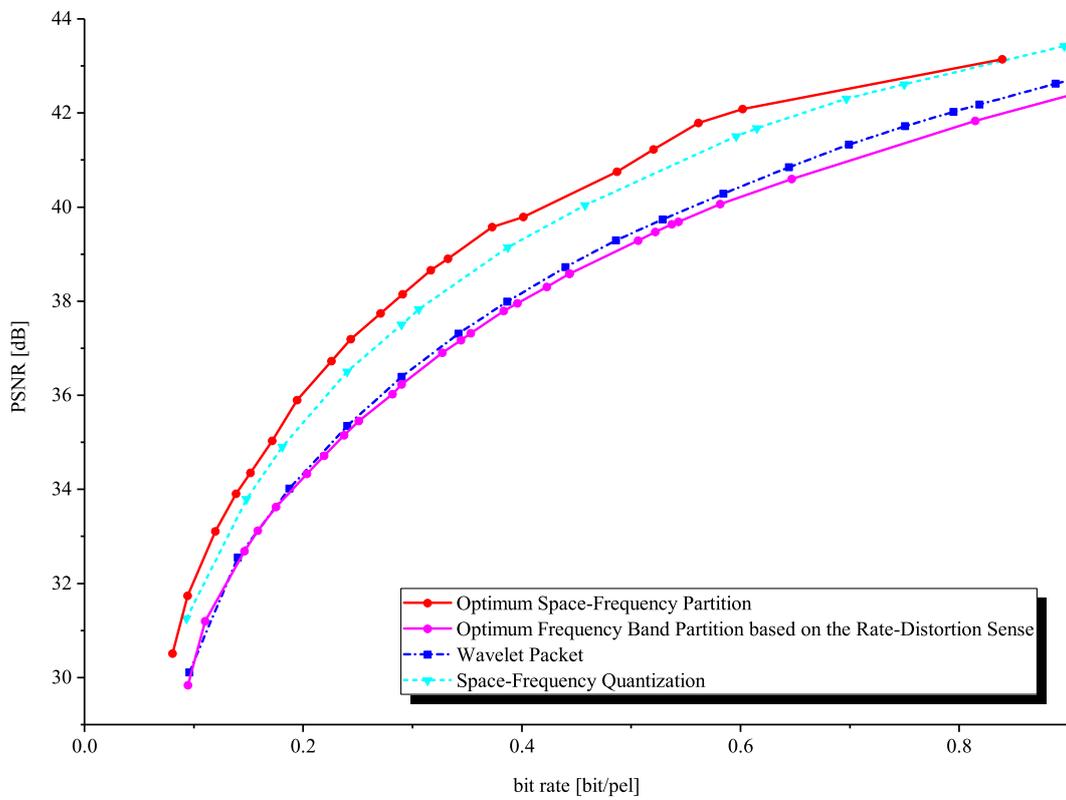
(c) 画像“kodim03”

図 3.8 空間-周波数領域の最適分割による空間領域分割パターン ( $M = 5, N = 64$ )



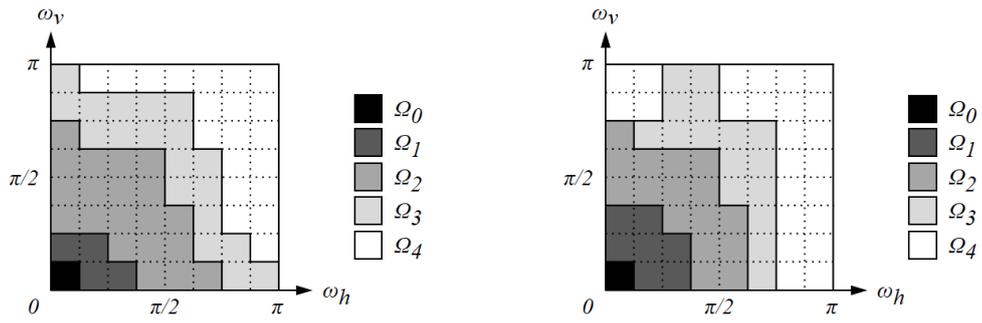
(d) 画像 “kodim23” ((768 × 512) 画素, 濃淡 8[bit/pel])

図 3.4 テスト画像



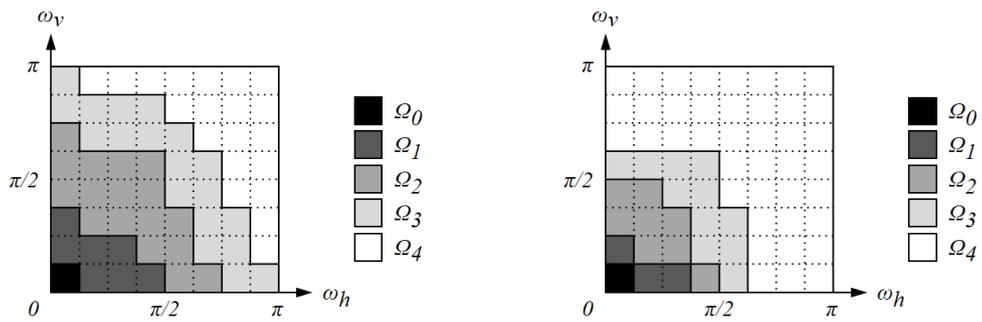
(d) 画像 “kodim23”

図 3.5 PSNR 対ビットレート特性による符号化性能の比較



(i) 高符号化レート (0.52[bit/pel])      (ii) 低符号化レート (0.12[bit/pel])  
(d) 画像“kodim23”

図 3.6 空間-周波数領域の最適分割による帯域分割パターン ( $M = 5, N = 64$ )

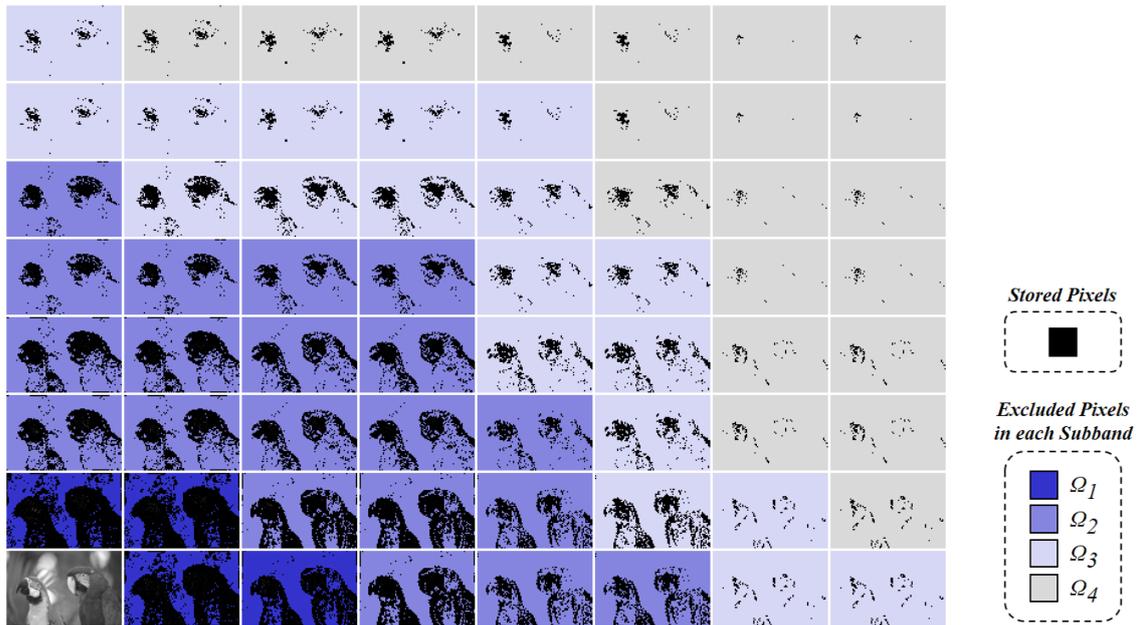


(i) 高符号化レート (0.54[bit/pel])      (ii) 低符号化レート (0.11[bit/pel])  
(d) 画像“kodim23”

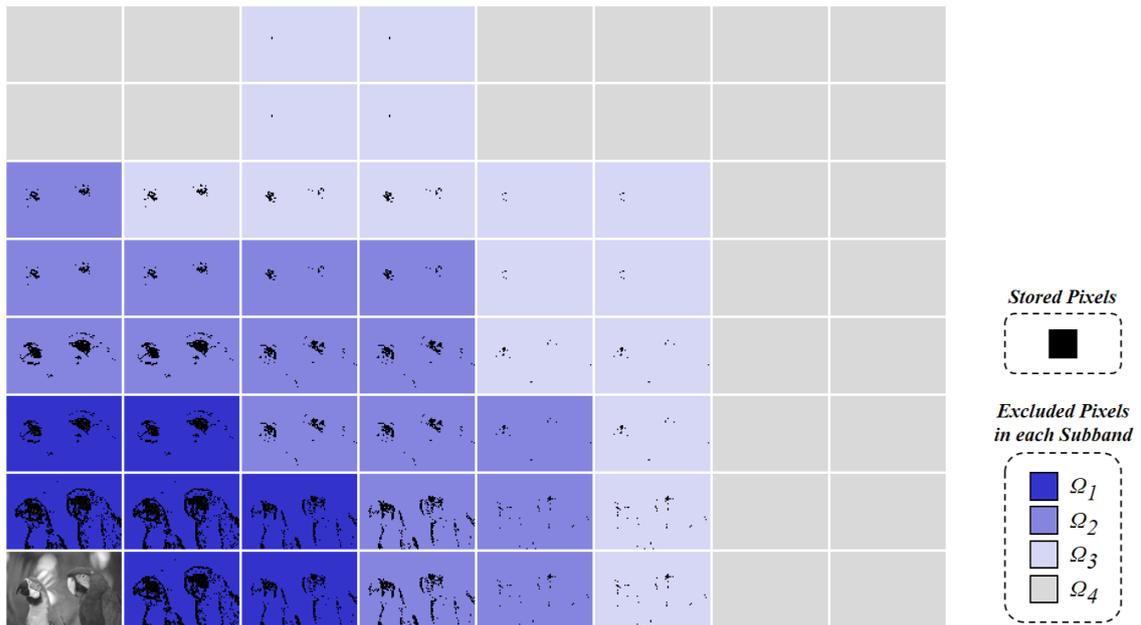
図 3.7 レート歪み理論に基づいた最適帯域分割による帯域分割パターン ( $M = 5, N = 64$ )

表 3.4 図 3.6(d) と図 3.7(d) に示す各サブバンド信号に適用される量子化器

subband	quantizer applied to each subband signal in frequency band partition pattern			
	Figure 3.6(d)		Figure 3.7(d)	
	(i) high bit rate (0.52[bit/pel])	(ii) low bit rate (0.12[bit/pel])	(i) high bit rate (0.54[bit/pel])	(ii) low bit rate (0.11[bit/pel])
$\Omega_0$	$Q_1$	$Q_6$	$Q_1$	$Q_{10}$
$\Omega_1$	$Q_1$	$Q_7$	$Q_2$	$Q_{13}$
$\Omega_2$	$Q_1$	$Q_8$	$Q_2$	$Q_{15}$
$\Omega_3$	$Q_2$	$Q_8$	$Q_3$	$Q_{16}$
$\Omega_4$	$Q_2$	—	$Q_4$	$Q_{16}$



(i) 高符号化レート (0.52[bit/pel])



(ii) 低符号化レート (0.12[bit/pel])

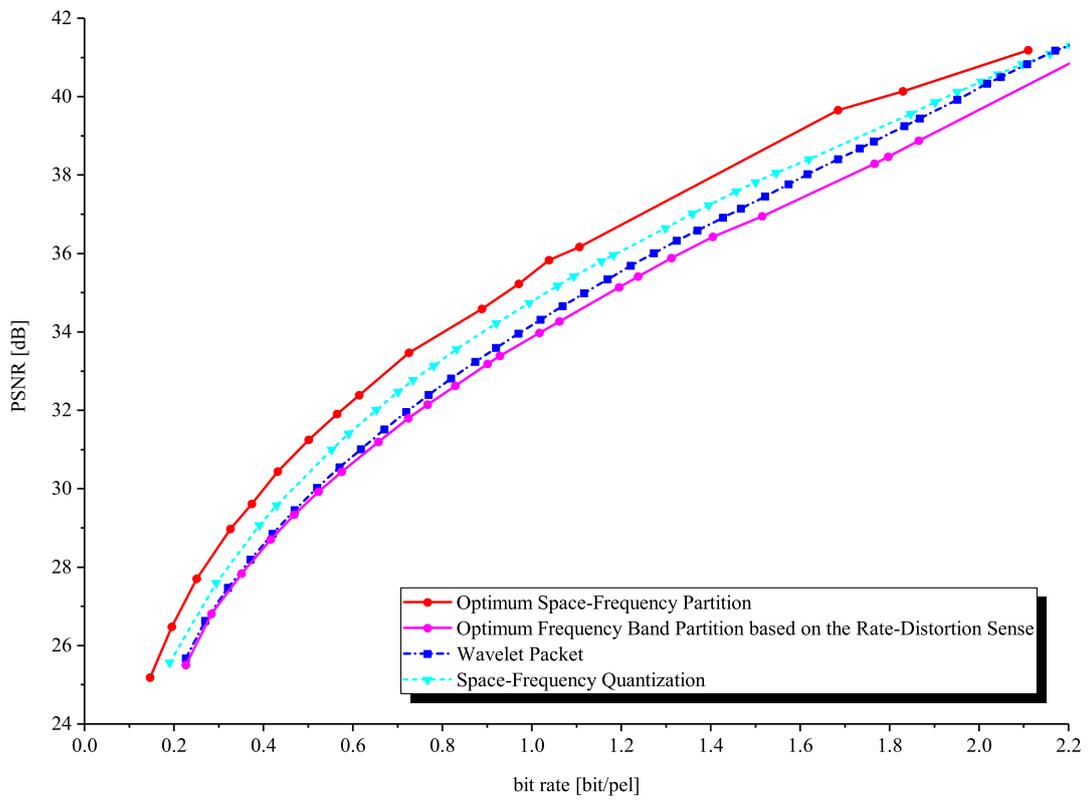
(d) 画像“kodim23”

図 3.8 空間-周波数領域の最適分割による空間領域分割パターン ( $M = 5, N = 64$ )



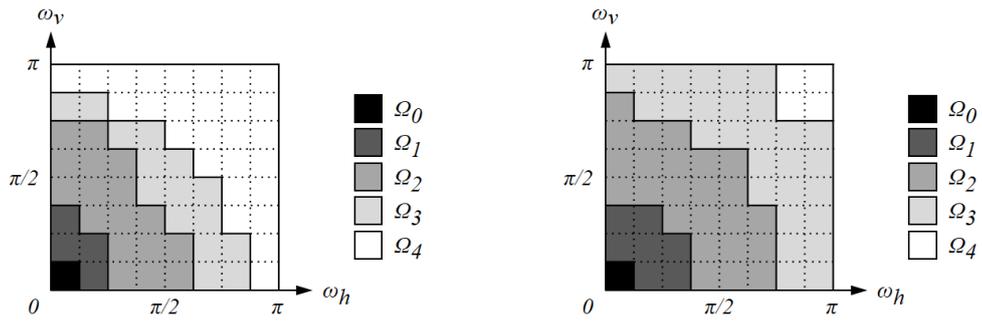
(e) 画像 “bike” ((1280 × 1600) 画素, 濃淡 8[bit/pe])

図 3.4 テスト画像



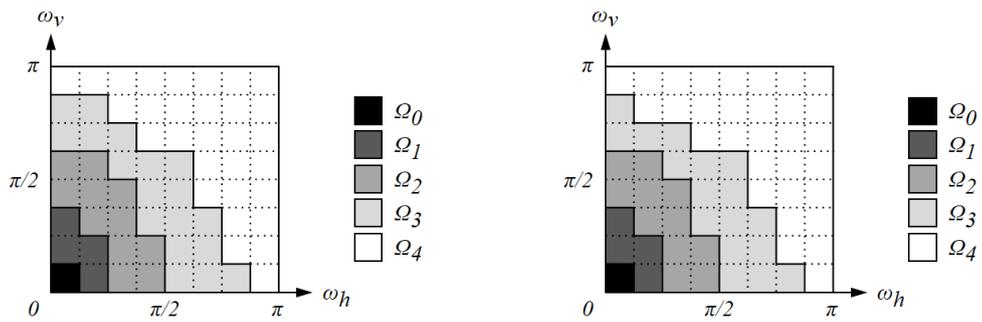
(e) 画像 “bike”

図 3.5 PSNR 対ビットレート特性による符号化性能の比較



(i) 高符号化レート (0.89[bit/pel])      (ii) 低符号化レート (0.32[bit/pel])  
(e) 画像“bike”

図 3.6 空間-周波数領域の最適分割による帯域分割パターン ( $M = 5, N = 64$ )

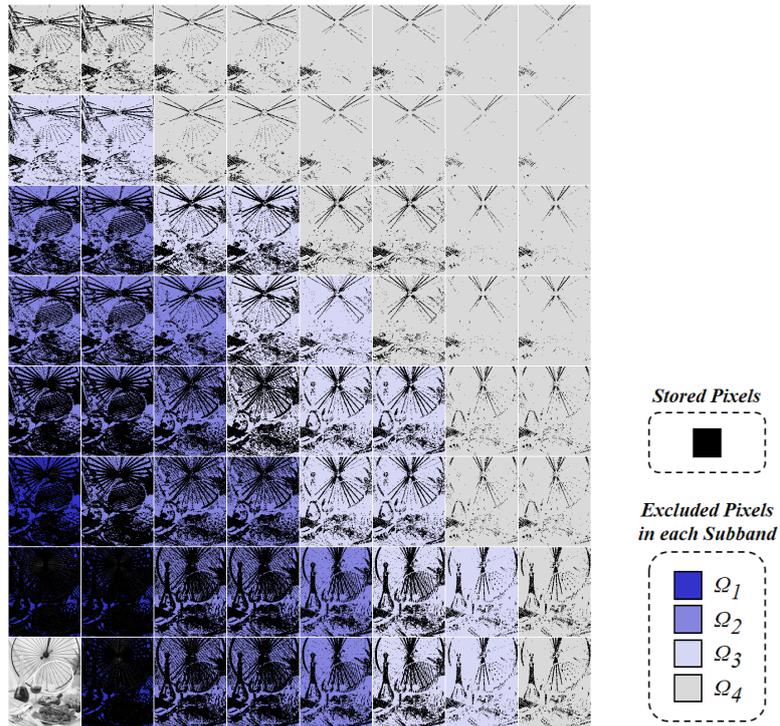


(i) 高符号化レート (0.89[bit/pel])      (ii) 低符号化レート (0.32[bit/pel])  
(e) 画像“bike”

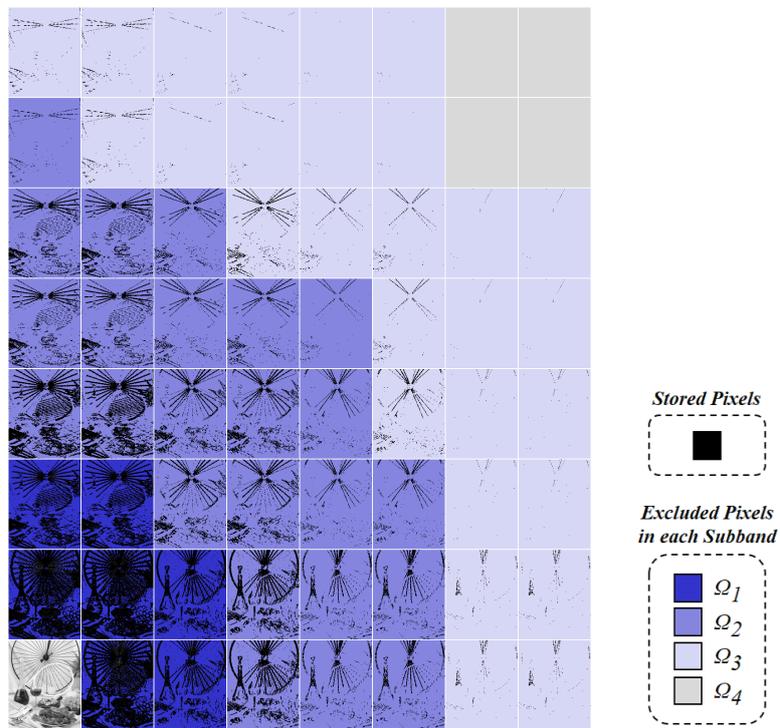
図 3.7 レート歪み理論に基づいた最適帯域分割による帯域分割パターン ( $M = 5, N = 64$ )

表 3.5 図 3.6(e) と図 3.7(e) に示す各サブバンド信号に適用される量子化器

subband	quantizer applied to each subband signal in frequency band partition pattern			
	Figure 3.6(e)		Figure 3.7(e)	
	(i) high bit rate (0.89[bit/pel])	(ii) low bit rate (0.32[bit/pel])	(i) high bit rate (0.89[bit/pel])	(ii) low bit rate (0.32[bit/pel])
$\Omega_0$	$Q_3$	$Q_7$	$Q_3$	$Q_7$
$\Omega_1$	$Q_3$	$Q_7$	$Q_3$	$Q_{10}$
$\Omega_2$	$Q_2$	$Q_7$	$Q_3$	$Q_{11}$
$\Omega_3$	$Q_3$	$Q_9$	$Q_4$	$Q_{14}$
$\Omega_4$	$Q_3$	—	$Q_5$	$Q_{16}$



(i) 高符号化レート (0.89[bit/pel])



(ii) 低符号化レート (0.32[bit/pel])

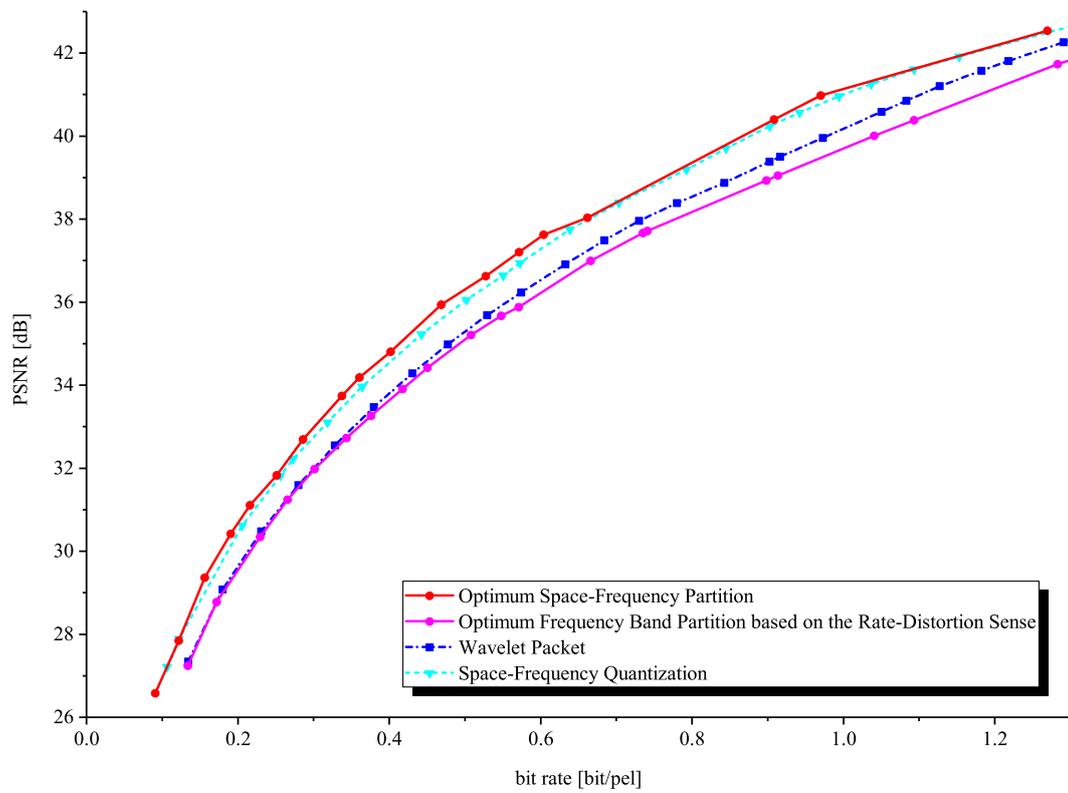
(e) 画像“bike”

図 3.8 空間-周波数領域の最適分割による空間領域分割パターン ( $M = 5, N = 64$ )



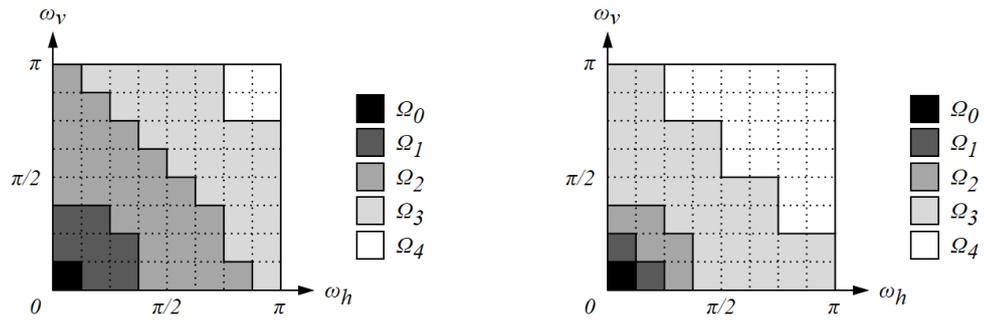
(f) 画像“p01” ((1280 × 1600) 画素, 濃淡 8[bit/pel])

図 3.4 テスト画像



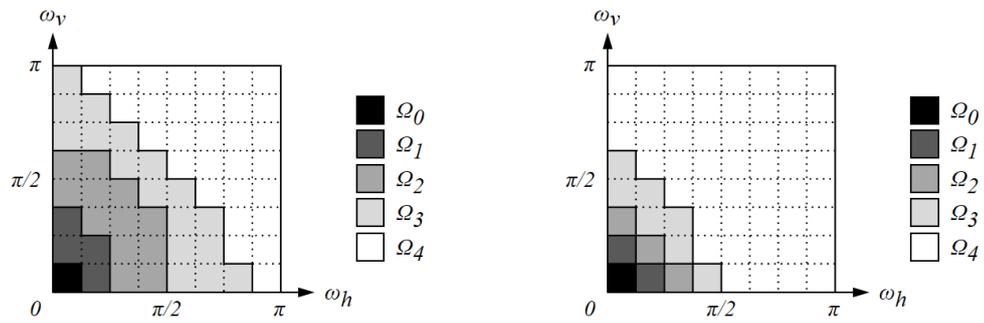
(f) 画像“p01”

図 3.5 PSNR 対ビットレート特性による符号化性能の比較



(i) 高符号化レート (0.91[bit/pel])      (ii) 低符号化レート (0.19[bit/pel])  
(f) 画像“p01”

図 3.6 空間-周波数領域の最適分割による帯域分割パターン ( $M = 5, N = 64$ )

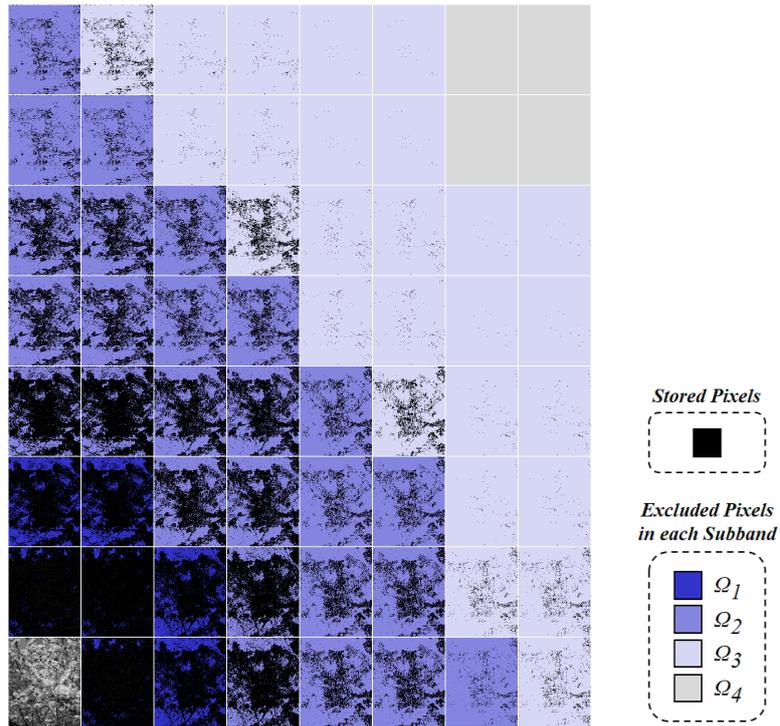


(i) 高符号化レート (0.90[bit/pel])      (ii) 低符号化レート (0.19[bit/pel])  
(f) 画像“p01”

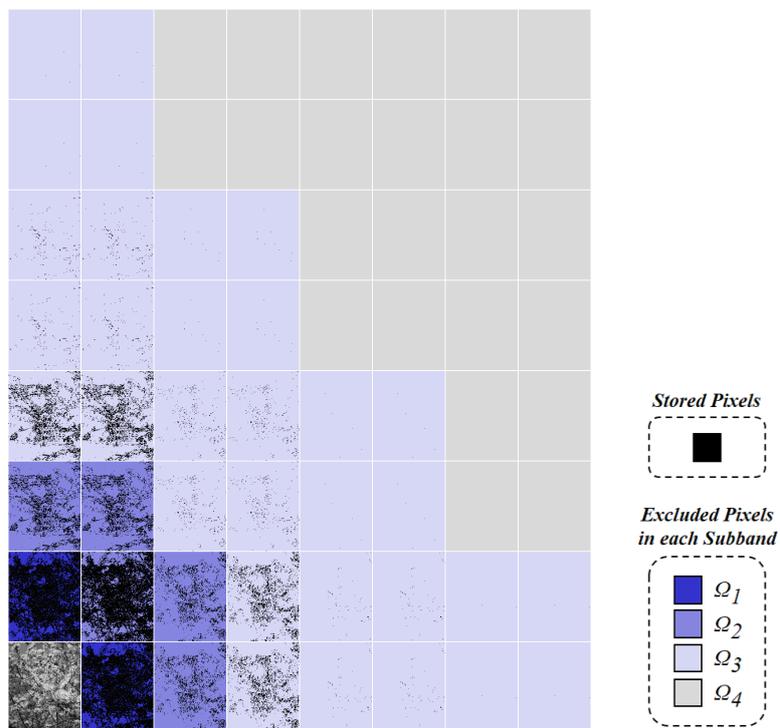
図 3.7 レート歪み理論に基づいた最適帯域分割による帯域分割パターン ( $M = 5, N = 64$ )

表 3.6 図 3.6(f) と図 3.7(f) に示す各サブバンド信号に適用される量子化器

subband	quantizer applied to each subband signal in frequency band partition pattern			
	Figure 3.6(f)		Figure 3.7(f)	
	(i) high bit rate (0.91[bit/pel])	(ii) low bit rate (0.19[bit/pel])	(i) high bit rate (0.90[bit/pel])	(ii) low bit rate (0.19[bit/pel])
$\Omega_0$	$Q_2$	$Q_8$	$Q_1$	$Q_8$
$\Omega_1$	$Q_1$	$Q_7$	$Q_1$	$Q_9$
$\Omega_2$	$Q_1$	$Q_8$	$Q_2$	$Q_{11}$
$\Omega_3$	$Q_2$	$Q_9$	$Q_2$	$Q_{16}$
$\Omega_4$	$Q_-$	-	$Q_3$	$Q_{16}$



(i) 高符号化レート (0.91[bit/pel])



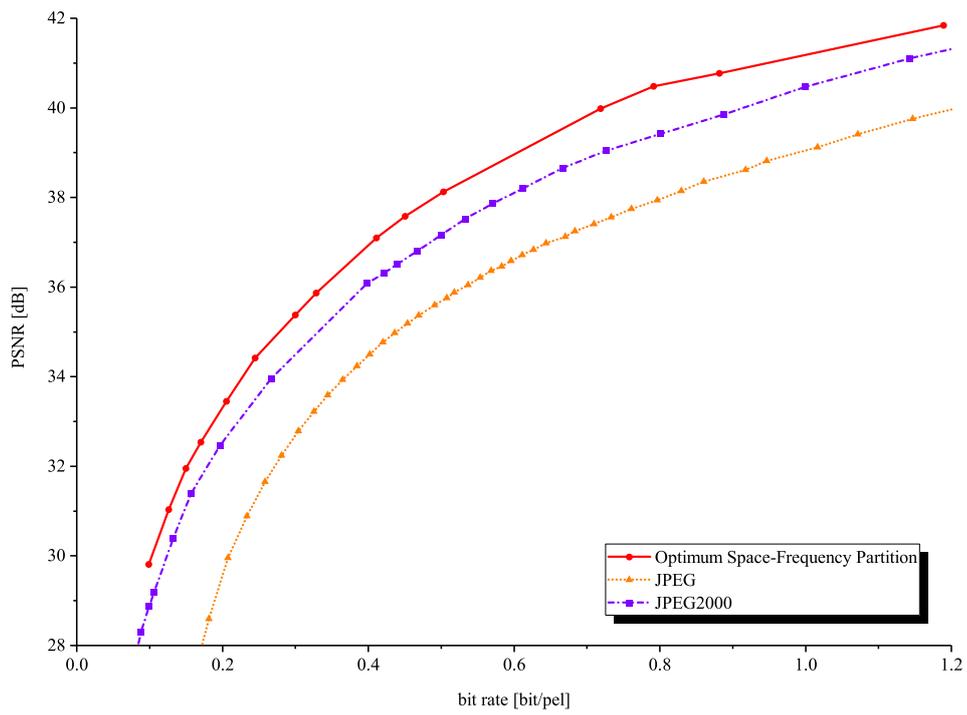
(ii) 低符号化レート (0.19[bit/pel])

(f) 画像“p01”

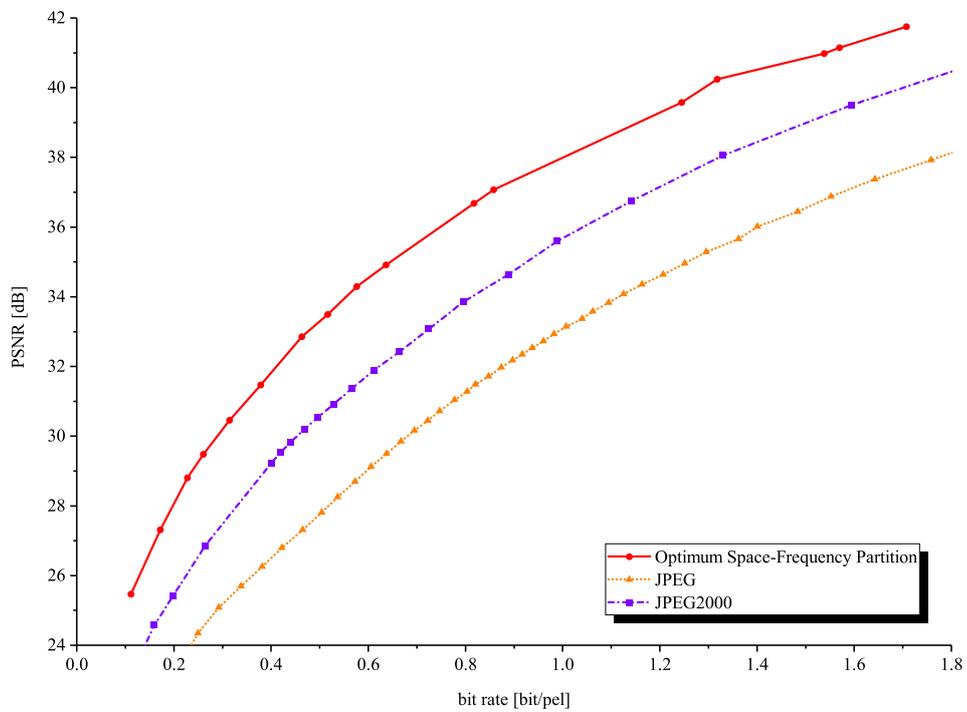
図 3.8 空間-周波数領域の最適分割による空間領域分割パターン ( $M = 5, N = 64$ )

更に、空間-周波数領域の最適分割による符号化性能の結果を、画像符号化の国際標準方式である JPEG 及び JPEG2000 と比較する。このとき、JPEG2000 の符号化性能は、OpenJPEG [70] の JPEG2000 コーデックを使用して算出した。また、JPEG2000 における離散ウェーブレット変換は、空間-周波数領域の最適分割と同様に、サブバンドの再帰分解レベルを 3 に設定している。図 3.4(a)~(f) に示す各画像に対する比較結果を図 3.9(a)~(f) に示す。図 3.9(a)~(f) の結果から、空間-周波数領域の最適分割は、入力画像の種類によらず全ての符号化レートで、JPEG からは最大で約 5.0[*dB*]、JPEG2000 からは最大で約 2.5[*dB*] の符号化性能の改善が得られることが明らかになった。

最後に、各手法によって得られた符号化後の再生画像を観察し、主観評価の観点から空間-周波数領域の最適分割の有効性を明らかにする。図 3.10 は、図 3.4(a) に示す画像“Lenna”に対して、符号化レート 0.20[*bit/pel*] の条件で、(b) 空間-周波数領域の最適分割、(c) レート歪み理論に基づいた最適帯域分割、(d) Wavelet Packet、(e) SFQ、(f) JPEG、(g) JPEG2000 をそれぞれ適用することによって得られた再生画像の髪飾り部分を拡大したものである。このとき、図 3.10 には、各手法による PSNR 値を記載している。図 3.10 の結果から、空間-周波数領域の最適分割による再生画像は、レート歪み理論に基づいた最適帯域分割や Wavelet Packet のものと比べると、エッジ周辺に発生するアーティファクトが低減されていることがわかる。そして、SFQ による再生画像は、画像が全体的に不鮮明でテクスチャ部分が潰れてしまっており、JPEG による再生画像は、ブロック歪みが顕著に現れてしまっている。また、JPEG2000 による再生画像は、他の手法と比べて PSNR 値は比較的の高いものの、テクスチャ部分が潰れて不鮮明になってしまっており、更にエッジ部分でジャギーが発生してしまっている。これらに対して空間-周波数領域の最適分割では、特にテクスチャ部分が鮮明で、ブロック歪みやエッジ周辺のアーティファクトが発生していないため、他の既存手法と比べて主観的にも優れた再生画像が得られる手法であることが明らかになった。

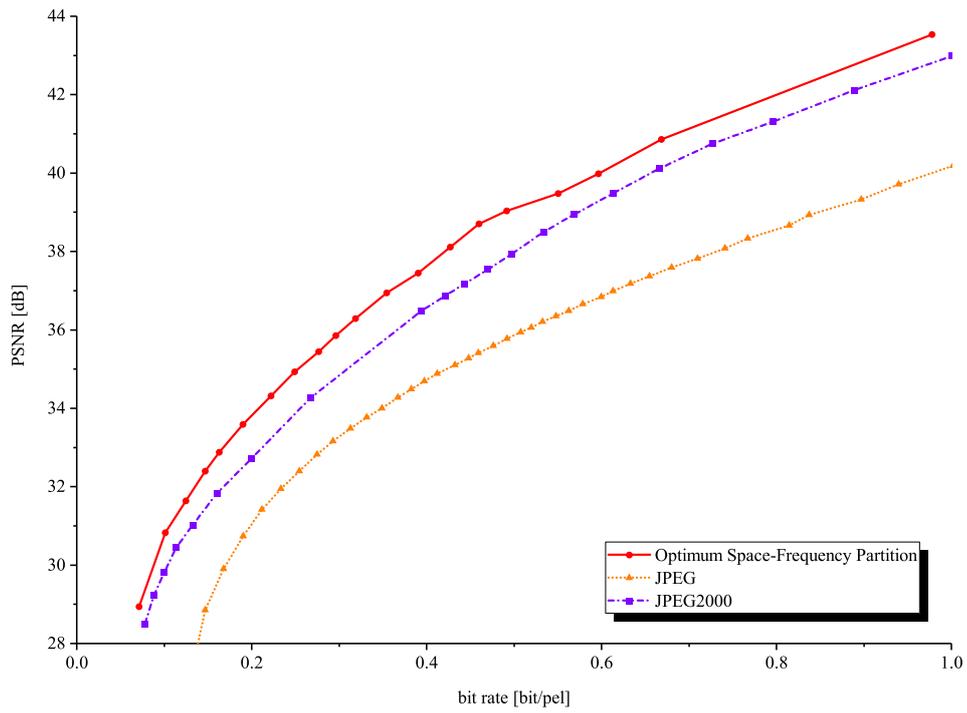


(a) 画像“Lenna”

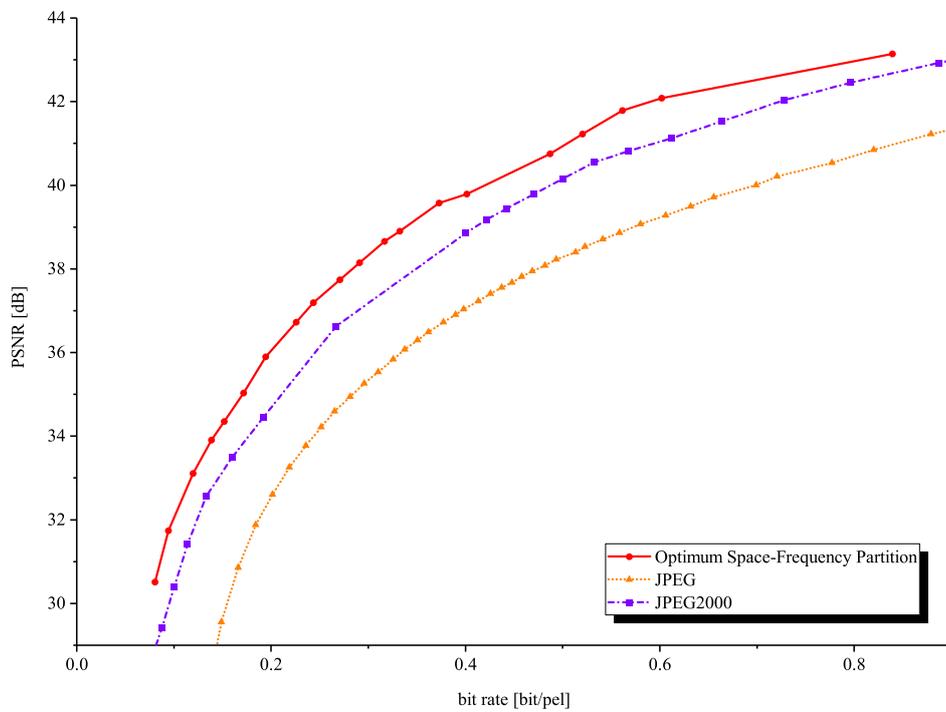


(b) 画像“Barbara”

図 3.9 空間-周波数領域の最適分割と JPEG 及び JPEG2000 との符号化性能の比較

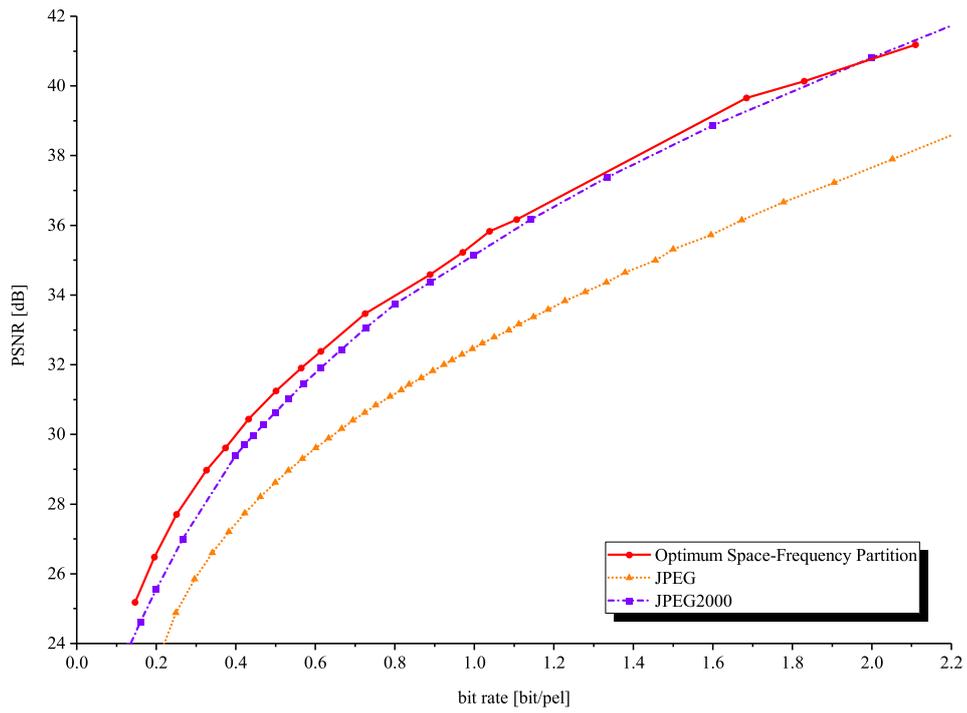


(c) 画像“kodim03”

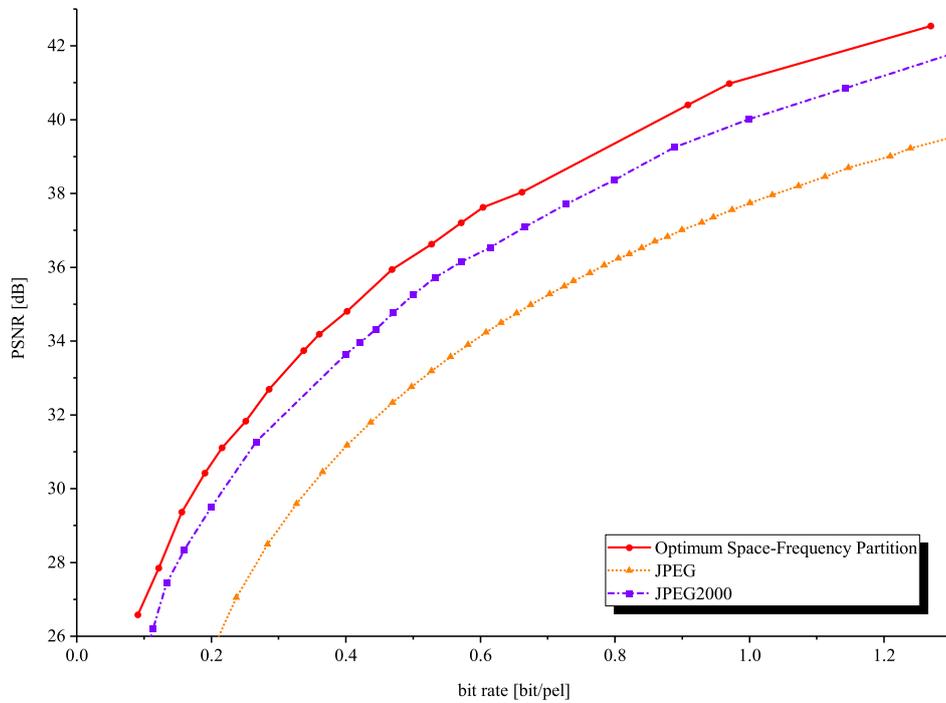


(d) 画像“kodim23”

図 3.9 空間-周波数領域の最適分割と JPEG 及び JPEG2000 との符号化性能の比較



(e) 画像“bike”



(f) 画像“p01”

図 3.9 空間-周波数領域の最適分割と JPEG 及び JPEG2000 との符号化性能の比較



(a) 原画像 (b) 空間-周波数領域の最適分割 (33.45[*dB*]) (c) レート歪み理論に基づいた最適帯域分割 (32.01[*dB*]) (d) Wavelet Packet (32.35[*dB*])



(e) SFQ (30.35[*dB*]) (f) JPEG (29.95[*dB*]) (g) JPEG2000 (32.46[*dB*])

図 3.10 符号化レート 0.20[*bit/pel*] を指定したときに各手法を適用することによって得られた符号化後の再生画像の比較 (画像 “Lenna”)

### 3.3.2 処理コスト

空間-周波数領域の最適分割及びレート歪み理論に基づいた最適帯域分割，Wavelet Packet [45]，SFQ [67]，JPEG2000 のサブバンド符号化に基づいた5種類の符号化手法それぞれを実画像に適用したときの処理時間を計測し，処理コストの面から性能評価を行う．処理コストは，指定された符号化レートに対する二次元周波数帯域上の最適分割パターンと各サブバンドに適用する最適な量子化器の組み合わせ，及び空間-周波数領域の最適分割とSFQに限っては空間領域上の最適分割パターンを求める処理を加えた符号化側の実行時間と，各最適化結果に基づいて復号化側で符号化画像を再構成するために要した実行時間の合計とした．ただし，JPEG2000については，JPEG2000コーデック [70] を使用して実際に符号化及び復号化したときの実行時間の合計とした．なお，動作環境としては，プロセッサ：Intel(R) Core(TM) i7-4790K CPU@4.00GHz，メモリ：16.0GBのPCを使用し，各手法はC言語により実装されている．

表 3.7 は，図 3.6(a)～(f) に示した各画像に各手法を適用し，異なる複数種類の符号化レートに対する結果を得るために要した処理コスト [秒] の平均値を示している．ここで，表 3.7 における括弧内は標準偏差である．表 3.7 の結果から，空間-周波数領域の最適分割では新たに空間領域分割の処理が加わるため，処理コストがレート歪み理論に基づいた最適帯域分割の約 1.25 倍程度に増加しているものの，Wavelet Packet と比べて約 0.75 倍程度の少ない処理

表 3.7 処理コスト [秒] の比較

method	“Lenna”	“Barbara”	“kodim03”
Optimum Space-Frequency Partition	1.020(±0.048)	1.059(±0.023)	1.578(±0.049)
Optimum Frequency Band Partition based on the Rate-Distortion Sense	0.848(±0.006)	0.855(±0.017)	1.288(±0.015)
Wavelet Packet	1.362(±0.022)	1.457(±0.036)	2.046(±0.041)
SFQ	5.595(±0.242)	6.828(±0.320)	9.700(±0.302)
JPEG2000	0.277(±0.276)	0.308(±0.281)	0.531(±0.469)

method	“kodim23”	“bike”	“p01”
Optimum Space-Frequency Partition	1.579(±0.046)	8.600(±0.312)	8.340(±0.533)
Optimum Frequency Band Partition based on the Rate-Distortion Sense	1.288(±0.010)	6.790(±0.180)	6.814(±0.138)
Wavelet Packet	2.095(±0.038)	10.687(±0.213)	10.720(±0.218)
SFQ	9.104(±0.331)	41.503(±2.062)	47.744(±2.327)
JPEG2000	0.517(±0.460)	1.548(±0.646)	1.575(±0.774)

コストで実現できていることがわかる。また、SFQは他の手法と比べて処理コストの大きさが顕著であるが、これはオリジナルのSFQでは、最低域のサブバンドを除いた各サブバンド信号に対して用意されている量子化器をそれぞれ適用した後、量子化された各結果に対して空間領域分割を行い、その中から空間領域上の最適分割パターンを探索するという処理コストの大きなアルゴリズムを用いているためである。表3.7と前節に示した図3.5(a)~(f)の結果より、空間-周波数領域の最適分割は、帯域ブロック数 $N$ やサブバンド数 $M$ を増加させずに、レート歪み理論に基づいた最適帯域分割から僅かに処理コストを増加させるだけで符号化性能を大きく改善できる手法であると言える。更に、Wavelet PacketとSFQとの比較においては、空間-周波数領域の最適分割は、少ない処理コスト及び $M=5$ という少ないサブバンド数で高い符号化性能を実現できていることから、符号化性能と処理コストの両者の観点で優れた手法であることが示された。しかしながら、JPEG2000と比較して、空間-周波数領域の最適分割は大きな処理コストを要することがわかる。これはJPEG2000では、帯域分割パターンが固定の離散ウェーブレット変換によるものであるため、高域側のサブバンドに対する分解が不要であることや、空間-周波数領域の最適分割や他の手法のように、レート歪み最適化問題を解くための繰り返し処理を行う必要がないためである。しかし、前節に示した図3.9(a)~(f)の結果から、符号化性能の観点では空間-周波数領域の最適分割がJPEG2000よりも優れている。実用化の観点では、例えば準最適解を得る代わりに処理コストを大幅に削減できるようにアルゴリズムを改良することで、図3.9(a)~(f)の結果から性能の劣化をできる限り抑えつつ、少ない処理コストで近似的に空間-周波数領域の最適分割を実現できるような符号化方式の提案が望まれるが、これについては今後の課題として検討したい。

## 3.4 結言

本章では、第2章で提案したレート歪み理論に基づいた最適帯域分割では、各帯域ブロック信号における二次元空間領域上の冗長性が除去されていないことに着目して、符号化レートが指定されたときに、二次元周波数帯域分割、各サブバンド信号に対する量子化器、そして、各帯域ブロック内の信号を保存と削除の二つに分類し、削除領域を切り捨てて保存領域のみを符号化することで空間領域上の冗長削減を実現する二次元空間領域分割の三つを、レート歪み理論の観点で最適化する「空間-周波数領域の最適分割」を提案した。このとき、第2章で提案したのもも含めた従来の最適帯域分割からサブバンド数が増えることはないのので、エントロピー符号化器設計に要する処理コストの増加は生じない。そして、空間-周波数領域の最適分割を実画像に適用することで、その有効性を定量的に明らかにした。

まず、レート歪み理論に基づいて、二次元周波数帯域分割、量子化器の組み合わせ、二次元空間領域分割の三つを同時に最適化するグローバルな最適化問題の定式化を行った。その後、それを直接的に解くことが現実的に困難であることを踏まえて、一つの最適化問題を解く際には、残り二つの解が既に与えられていると仮定した上で、グローバルな最適化問題に基づいて三つの各ローカルな最適化問題を個別に定式化した。そして、新しく空間領域分割が導入されたことを考慮した上で、帯域分割及び量子化の最適解を導出するための理論式と、分割されたサブバンド間におけるゼロツリー関係を考慮しながら空間領域分割に伴う付加情報量を抑えつつ、空間領域分割の解を導出するためのアルゴリズムを明らかにした。その後、これら三つを指定された符号化レートに到達するまで反復して解くことによって、所望の符号化レートに応じた二次元周波数帯域の最適分割パターン、各サブバンド信号に適用する量子化器の組み合わせ、各帯域ブロックにおける二次元空間領域分割の最適分割パターンの三つを求めるアルゴリズムを提案した。

最後に、空間-周波数領域の最適分割を実画像に適用し、PSNR 対ビットレート特性による符号化性能を求め、第2章で提案したレート歪み理論に基づいた最適帯域分割、Wavelet Packet, Space-Frequency Quantization (SFQ), 更に JPEG 及び JPEG2000 との比較を行った。その結果、提案手法は、サブバンド数を増加させずに、従来手法の僅か 1.25 倍程度の処理コストによって、同じ符号化レートをもつ従来手法と比べて PSNR 値において最大で 2.0[*dB*] 程度の改善が得られ、更に、全ての符号化レートで Wavelet Packet, SFQ, JPEG, JPEG2000 よりも高い符号化性能を得られることが明らかになった。以上の結果から、本章において提案した空間-周波数領域の最適分割は、サブバンド信号の領域分割による二次元空間領域の冗長削減を実現することで、エントロピー符号化器設計の処理コストに直結するサブバンド数を少数に抑えたまま、レート歪み理論に基づいた最適帯域分割の符号化性能を効果的に改善し、さらに、Wavelet Packet や SFQ 等の既存のサブバンド符号化方式よりも高い符号化性能を達成できる手法であることが明らかとなった。

本章までの議論で明らかにされたレート歪み理論に基づいた最適帯域分割及び空間-周波数領域の最適分割は、符号化レートが指定されている制約条件の下で、原画像と符号化後に得られる再生画像間における信号の平均二乗誤差（MSE）を最小化することを目的として導出されたものであることから、再生画像における視知覚特性を考慮した画質という点までは言及されていない。そこで、次章では、主観的な画質の改善を得ることを目的として、「人間の視知覚特性を考慮した空間-周波数領域の最適分割」について議論を行うこととする。

## 第 4 章

# 視知覚特性を考慮した画像のサブバンド 符号化における空間-周波数領域の最適分割

### 4.1 序言

画像のサブバンド符号化において、圧縮性能の改善と符号化器の設計コストの低減の二つを同時に達成するために、サブバンド数が少数であるという条件の下で、入力画像の内容を考慮して二次元周波数帯域を適応的に分割する最適帯域分割が提案されている [59]～ [61]. 本論文では、従来の最適帯域分割から符号化性能を改善させるために、第2章において、レート歪み理論の観点で二次元周波数帯域の分割パターンと各サブバンドに適用される量子化器の組み合わせを同時に最適化する新しい最適帯域分割法を提案した. また、第3章においては、各サブバンド信号における空間領域上の冗長性が十分に除去しきれていないことに着目し、二次元周波数帯域分割と量子化に加えて、各サブバンド内の不要な信号を切り捨てて、必要な信号のみを符号化することで冗長削減を実現する二次元空間領域分割の三つを、レート歪み理論の観点で同時に最適化する空間-周波数領域の最適分割を提案した. しかしながら、前章までに提案したレート歪み理論に基づいた最適帯域分割及び空間-周波数領域の最適分割は、サブバンド信号の量子化前後間の平均二乗誤差 (MSE) によって定義された歪みの量を最小化することを目的として導出されたものであることから、視知覚特性を考慮した画質という観点では、真に最適なものが得られているとは言い難い. 画像符号化における原点は、高い圧縮効率の下で、上記の画質の観点で良い評価を得ることにあり、画質に対する議論の展開が必要である.

本章では、第3章で提案した空間-周波数領域の最適分割よりも主観的に良好な再生画像を得ることを目的として、人間の視知覚特性を考慮した空間-周波数領域の最適分割を提案する. また、今回は入力画像としてカラー画像を対象とし、議論を行う. 本空間-周波数領域の最適分割における画質改善の原理は、視知覚特性上で画像中の重要ではない信号の品質を犠牲にする代わりに重要な信号の品質を優先的に保つことであり、そしてそれは、人間の視

知覚特性の一種である視覚の空間周波数特性 [71] 及び画像の顕著性の二つを考慮して、各サブバンド信号の歪み値に対する重み付けを行った後、その重み付けされた歪みの量を符号化レートが指定されているという制約条件の下で最小化することにより実現される。このとき、重み付け関数は、視覚の空間周波数特性と画像の顕著性を反映するために、画像中に含まれる各周波数成分に対する人間の視知覚感度をモデル化した視覚の二次元空間周波数特性と、入力画像に対する顕著性の分布とその高さを推定できる Saliency Map [72] に基づいて導出する。さらに、提案手法では、Saliency Map の結果に基づいて、各サブバンド信号を空間領域上で注目領域 (Region-of-Interest : ROI) と非注目領域 (non-Region-of-Interest : non-ROI) の二つに分割し、それぞれに対して異なる量子化器を選択できるように量子化の最適化処理を拡張する。これにより、視知覚感度が低い周波数成分を持つサブバンドや、空間領域上で顕著性が低い領域の品質を犠牲にする代わりに、人間の視知覚特性という観点で重要であると考えられる視知覚感度が高い周波数成分を持つサブバンドや、顕著性が高い領域の品質を優先的に保つことができるため、再生画像の主観的な画質の改善が期待できる。このとき、空間領域分割の最適化処理やエントロピー符号化は、ROI と non-ROI に分割せずに、従来と同様、サブバンド単位で行われるため、エントロピー符号化器設計のための処理コストの増加は生じない。最後に、複数種類の異なる符号化レートを指定した条件で、実画像に対して本空間-周波数領域の最適分割を適用し、得られた各再生画像の画質を評価するための主観評価実験を行う。そして、得られた評価結果を、従来の視知覚特性を考慮しない空間-周波数領域の最適分割方式及び画像符号化の国際標準方式 JPEG2000 のものと比較することで、本空間-周波数領域の最適分割の有効性を明らかにする。

## 4.2 視覚の空間周波数特性と画像の顕著性を考慮した重み付け関数の導出

第3章で提案した空間-周波数領域の最適分割は、周波数帯域分割、量子化、空間領域分割の三つの最適化問題を、所望の符号化レートに到達するまで反復して解くことにより、レート歪み理論の観点で、 $M$ 分割の二次元周波数帯域上の最適分割パターン  $I_k$ 、分割された  $M$ 個のサブバンド  $\Omega_k (k = 0, 1, \dots, M-1)$  に適用する最適な量子化器  $q_k$ 、 $N$ 個の帯域ブロック  $\Delta\Omega_i (i = 0, 1, \dots, N-1; N \geq M)$  における二次元空間領域上の最適分割パターン  $S_i$  の三つが決定される。ここで、 $I_k$  はサブバンド  $\Omega_k$  に属する帯域ブロック番号  $i$  の集合、 $S_i$  は帯域ブロック  $\Delta\Omega_i$  内で保存される信号番号  $j$  の集合にそれぞれ対応している。

一つ目の二次元周波数帯域上の最適分割パターン  $I_k$  は、帯域ブロック  $\Delta\Omega_i$  内で保存される信号  $j (j \in S_i)$  を量子化器  $q_k$  で量子化したときに得られる確率分布を  $\Delta\tilde{\Omega}_i(x)$ 、その信号数を  $\Delta N_i$  としたとき、 $\Delta N_i$  で重み付けられた Kullback-Leibler Divergence を距離尺度とする  $N$ 個の確率分布  $\Delta\tilde{\Omega}_i(x)$  のクラスタリング問題

$$I_k = \arg \min_{I_k} \left[ \sum_{k=0}^{M-1} \sum_{i \in I_k} \Delta N_i \sum_{x=x_i^{\min}}^{x_i^{\max}} \Delta\tilde{\Omega}_i(x) \log_2 \frac{\Delta\tilde{\Omega}_i(x)}{\tilde{\Omega}_k(x)} \right] \quad (4.1)$$

を解くことにより決定される。ここで、 $x_i^{\min}$  と  $x_i^{\max}$  は、それぞれ  $\tilde{\Omega}_i(x) \neq 0$  となる  $x$  の最小値と最大値である。また、 $\Delta\tilde{\Omega}_i(x)$  は、サブバンド  $\Omega_k$  内で保存される信号群から生成される信号の確率分布であり、サブバンド  $\Omega_k$  に属する帯域ブロックの確率分布  $\tilde{\Omega}_i(x) (i \in I_k)$  を統合した混合確率分布と等価である。

二つ目の各サブバンドに対する最適な量子化器  $q_k$  は、サブバンド  $\Omega_k$  毎にラグランジュのコスト関数を最小とする量子化器を個別に選択する問題

$$q_k = \arg \min_{q_k} \left[ \sum_{i \in I_k, j \in S_i} (w_{i,j} - \hat{w}_{i,j}^{q_k})^2 + \lambda \left( -N_k \sum_{x=x_k^{\min}}^{x_k^{\max}} \tilde{\Omega}_k(x) \log_2 \tilde{\Omega}_k(x) \right) \right] \quad (4.2)$$

を解くことにより決定される。ここで、 $w_{i,j}$  と  $\hat{w}_{i,j}^{q_k}$  はそれぞれ、帯域ブロック  $\Delta\Omega_i$  の  $j$  番目の信号における量子化前の係数値と、それを量子化器  $q_k$  で量子化したときの係数値である。また、 $\lambda$  はラグランジュ乗数、 $N_k$  はサブバンド  $\Omega_k$  内で保存される信号数であり、 $x_k^{\min}$  と  $x_k^{\max}$  は、それぞれ  $\tilde{\Omega}_k(x) \neq 0$  となる  $x$  の最小値と最大値である。

三つ目の二次元空間領域上の最適分割パターン  $S_i$  は、帯域ブロック  $\Delta\Omega_i$  における  $j$  番目の信号を保存した場合のコスト関数

$$J_{i,j} = (w_{i,j} - \hat{w}_{i,j}^{q_k})^2 + \lambda \left\{ -\log_2 \tilde{\Omega}_k(\hat{w}_{i,j}^{q_k}) \right\} \quad (4.3)$$

と，削除した場合のコスト関数  $w_{i,j}^2$  を比較し，保存するよりも削除した方がコストを小さくできる信号を，帯域ブロック内から除外することにより決定される．ここで，式(4.3)における  $\tilde{\Omega}_k(\hat{w}_{i,j}^{qk})$  は，サブバンド  $\Omega_k$  内で係数値  $w_{i,j}^{qk}$  を持つ信号が発生する確率である．

式(4.2)及び式(4.3)で示されているように，空間-周波数領域の最適分割では，帯域ブロック信号の量子化前後間における係数値の二乗誤差  $(w_{i,j} - \hat{w}_{i,j}^{qk})^2$ ，または，量子化前の係数の二乗値  $w_{i,j}^2$  によって定義された歪みの量を最小化することを目的としている．そのため，再生画像における視知覚特性を考慮した画質という観点では，求められる三つの解  $I_k, q_k, S_i$  が真に最適なものになっているとは言い難い．

本章で提案する視知覚特性を考慮した空間-周波数領域の最適分割は，人間の視知覚特性を考慮して各帯域ブロック信号の歪み値に対して重みを掛け，その重み付けされた歪みの量を最小化することによって，再生画像における主観的な画質を改善する．本節では，それを実現するための第一段階として，帯域ブロック信号の歪み値に対する重み付け関数の導出について述べる．このとき，重み付け関数は，視覚の二次元空間周波数特性と，画像に対する注意の向けられやすさを意味する顕著度の分布を算出する Saliency Map の二つに基づいて導出する．ここで，視覚の二次元空間周波数特性を導入する目的は，視知覚感度の低い周波数成分を有する帯域ブロックの品質を犠牲にする代わりに，視知覚感度の高い周波数成分を有する帯域ブロックの品質を優先的に保持するためであり，Saliency Map を導入する目的は，帯域ブロック内にある空間領域上で顕著度が低い信号の品質を犠牲にする代わりに，顕著度が高い信号の品質を優先的に保持するためである．

## 4.2.1 視覚の二次元空間周波数特性による重み付け

視覚の空間周波数特性は，測定条件等により異なるが，一般に帯域通過特性を持つことが知られており，従来において，いくつかの代表的なモデルが提案されている [71]．しかし，上記文献内に示されたこれらのモデルにおいては，その測定原理から直流成分に対する感度特性を求めることができない．そこで，提案手法では，文献 [73], [74] において引用されている網膜の神経回路をモデル化して求められた視覚の空間周波数特性を用いることとする．この空間周波数特性は，他のモデルと比べて低域の特性がそれほど落ちていないのが特徴である．以下，本特性に基づいた視覚の二次元空間周波数特性について簡単に説明する．

文献 [73], [74] において引用されている空間周波数特性  $S(f)$  は，式(4.4)で表される．

$$S(f) = 1.5 \exp(-\sigma^2 \omega^2 / 2) - \exp(-2\sigma^2 \omega^2) \quad (4.4)$$

ただし， $\sigma = 2$ ， $\omega = 2\pi f / 60$ ， $f = \sqrt{u^2 + v^2} / 2\pi$  であり， $u, v$  はそれぞれ水平，垂直方向の空間周波数である．ここで， $f$  は視覚1度あたりの縞の周期 (cycle per degree [cpd]) を意味している．また，文献 [75] によれば，視覚の二次元空間周波数特性は異方特性を有しており，具体的には45度傾いた縞模様に対する視知覚感度は，水平及び垂直方向に比べて最大1/2低

下するとされている。この異方特性は、 $S(f)$ のうちの高域特性にのみ影響を与えるものであり、この特性を

$$O(\theta) = \begin{cases} 1.0 & : f < f_p \\ 0.5(1 + \cos^4 2\theta) & : f \geq f_p \end{cases} \quad (4.5)$$

により近似する。ここで、 $\theta = \tan^{-1}(v/u)$ であり、 $f_p$ は、式(4.4)の $S(f)$ が最大レスポンスとなる空間周波数で、 $f_p = 3.86[\text{cpd}]$ である。よって、異方特性を有する視覚の二次元空間周波数特性は、式(4.4)と式(4.5)に示された関数の積

$$\psi(u, v) = S(f) \cdot O(\theta) \quad (4.6)$$

として定義することができる。このとき、ディスプレイの解像度や視距離などの測定条件に応じて角画素密度 (pixels per degree [ppd])

$$\text{ppd} = 1/\tan^{-1} \left( \frac{2.54 \cdot D}{\sqrt{H^2 + V^2} \cdot d} \right) \quad (4.7)$$

を求め、 $\pi$ のときに $f$ が最短周期(すなわち、2画素)になるように $u, v$ を正規化する。ここで、 $D$ はディスプレイの対角サイズ(inch)、 $H$ と $V$ はそれぞれ、ディスプレイの横と縦の画素数(pixels)、 $d$ は視距離(cm)である。例として、図4.1に、 $D=27(\text{inch})$ 、 $H=1920(\text{pixels})$ 、 $V=1080(\text{pixels})$ 、 $d=50(\text{cm})$ の条件で求められた視覚の二次元空間周波数特性を示す。図4.1より、視覚の空間周波数特性は、高域よりも低域の周波数成分に対して高い値を示しており、

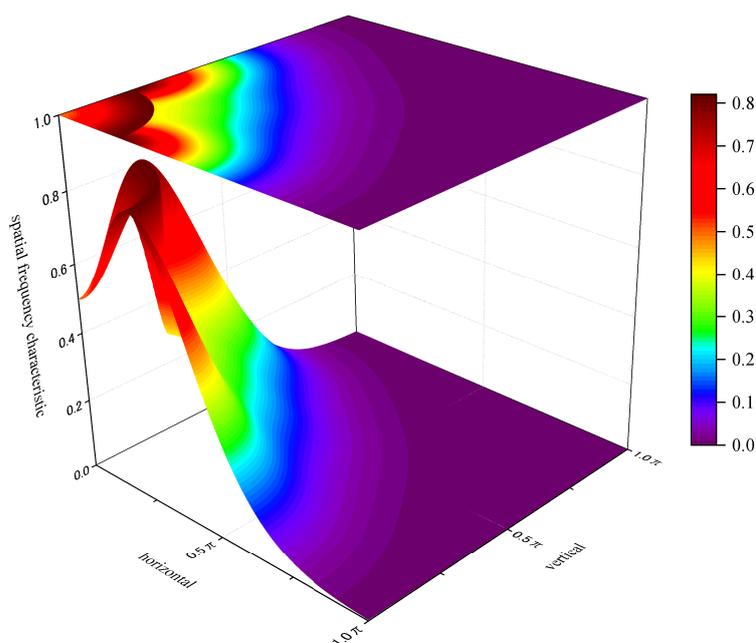


図 4.1 視覚の二次元空間周波数特性

さらに，異方特性が考慮されていることにより，水平及び垂直方向に比べて，斜め方向の周波数成分に対する視知覚感度が若干低下していることがわかる．

空間-周波数領域の最適分割においては，QMF等の帯域分割フィルタを用いて均等に分割された二次元周波数帯域の小領域である  $N$  個の帯域ブロック  $\Delta\Omega_i (i = 0, 1, \dots, N - 1)$  が，それぞれ異なる周波数成分を有する．したがって，提案手法では，式(4.6)によって得られる視覚の二次元空間周波数特性を反映するために，各帯域ブロックの歪み値に対して異なる重みが掛かるように重み付け関数を定義する．画素数  $(x \times y)$  から成る入力画像が与えられたとき，帯域ブロック  $\Delta\Omega_i$  に対する重みは，

$$\bar{\psi}_i = \frac{N}{xy} \sum_{(\pi m/x, \pi n/y) \in \Delta\Omega_i} \psi(\pi m/x, \pi n/y) \quad (4.8)$$

によって定義される．ただし， $(0 \leq m < x, 0 \leq n < y)$  である．式(4.8)を導入することによって，視知覚感度の低い周波数成分を有する帯域ブロックの品質を犠牲するに代わりに，視知覚感度の高い周波数成分を有する帯域ブロックの品質を優先的に保つことが可能となる．図4.2は，その具体例として， $N = 64$  個の各帯域ブロック信号の歪み値に対して重み付けされる重み係数値を示している．このとき，測定条件には，図4.1に示す視覚の二次元空間周波数特性を求めたときと同じものを指定した．

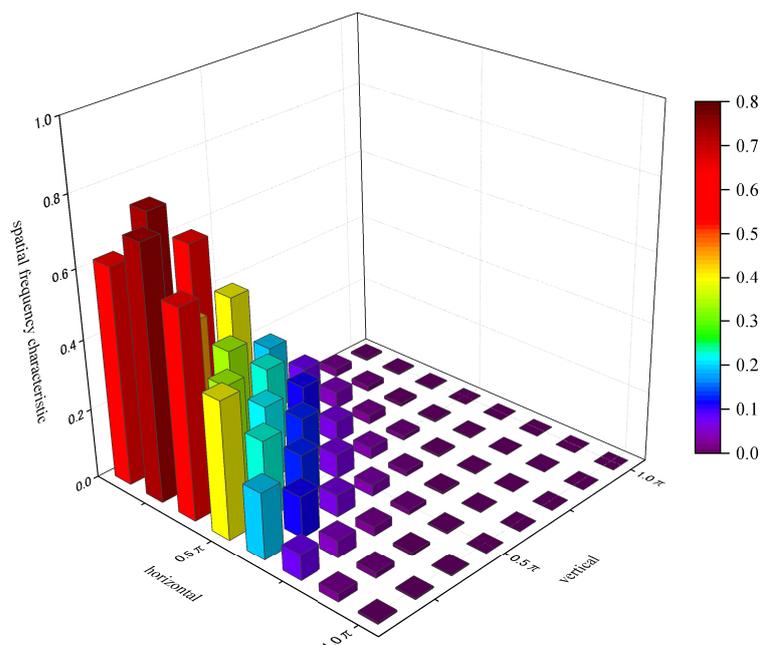


図4.2 視覚の二次元空間周波数特性による各帯域ブロックに対する重み付け係数値 ( $N = 64$ )

## 4.2.2 Saliency Map による重み付け

画像中には、人間の視覚的注意を引き付ける顕著領域が存在しており、従来において、被験者を必要とせずに画像特徴量からその顕著領域を推定する計算モデルが提案されている。Saliency Map は、この計算モデルに基づき、各画素に対して視覚的注意を引き付ける度合いを意味する顕著度が与えられた画像のことであり、画像認識や物体検出等のコンピュータビジョンの課題における前処理として利用することが期待されている。Saliency Map の先駆けとなった Itti らの手法 [72] では、入力画像から明度・色差・エッジ方向といった基本的な画像特徴量を複数スケールで抽出・統合することで、Saliency Map を算出している。文献 [72] の手法を含め、Saliency Map を算出する手法は、従来においていくつか提案されているが [76], [77], 画像中の前景や背景に高いコントラストを持つ小スケールの物体が含まれている場合、検出精度が著しく低下するという問題がある。そこで、提案手法では、この問題を根本的に解決し、従来よりも高い検出精度を持つ Yan らの手法 [78] を採用する。例として、図 4.3 に、(a)CIPR 標準画像 [68] における画像“girl” ((480×512) 画素, 24[bit/pel]) と、(b) その Saliency Map を示す。図 4.3(b) では、画素の輝度値の高さが顕著度の高さに対応しているため、画像“girl”においては顔領域が顕著性の高い領域として検出されている。

空間-周波数領域の最適分割において、帯域ブロック  $\Delta\Omega_i$  は空間領域情報を持つため、提案手法では、帯域ブロック内の各信号の歪み値に対して Saliency Map による重みが掛かるように重み付け関数を定義する。まず、画素数 ( $x \times y$ ) の入力画像に対する Saliency Map  $\phi(x \times y)$  を求めた後、平均画素法による縮小処理によって、 $\phi(x \times y)$  を帯域ブロックの信号サイズと



(a) 画像“girl”



(b) Saliency Map

図 4.3 画像“girl” ((480×512) 画素, 24[bit/pel]) とその Saliency Map

同じになるように縮小する．ここで，縮小された Saliency Map における  $j$  番目の信号を  $\bar{\phi}_j$  として表記することとし，これは帯域ブロックにおける  $j$  番目の信号と空間領域上で同じ位置に対応している．

最後に，帯域ブロック単位で定義された視覚の二次元空間周波数特性による重み  $\bar{\psi}_i (i = 0, 1, \dots, N - 1)$  と，帯域ブロック内の信号単位で定義された Saliency Map による重み  $\bar{\phi}_j$  を統合することによって，視覚の空間周波数特性と画像の顕著性を同時に考慮した重み付け関数を導出する．帯域ブロック  $\Delta\Omega_i$  における  $j$  番目の信号の歪み値に対する重み付け関数  $W_{i,j}$  は，式 (4.9) によって定義される．

$$W_{i,j} = [\|\bar{\psi}_i\| \cdot \|\bar{\phi}_j\|]^2 \quad (4.9)$$

ここで， $\|*\|$  は，係数値を 0~1.0 の範囲で正規化するための関数である． $W_{i,j}$  の重み係数値は，視知覚感度が高い周波数成分を有する帯域ブロック，且つ，その中の顕著度が高い信号が指定されたときに高い値を持つ．ただし，たとえ視知覚感度が高い帯域ブロックであっても，その中で顕著度の低い信号が指定された場合は，その重み係数値は低い値となる．また，空間領域上で顕著度が高い信号であっても，それが視知覚感度の低い帯域ブロック内のものであった場合，同様に低い値となる．図 4.4 は，その具体例として，帯域ブロック数  $N = 64$  の条件で，入力画像として画像 “girl” が与えられたとき，式 (4.9) によって算出された 64 個の帯域ブロック信号に対する重み付け関数  $W_{i,j}$  の係数値を可視化したものである．

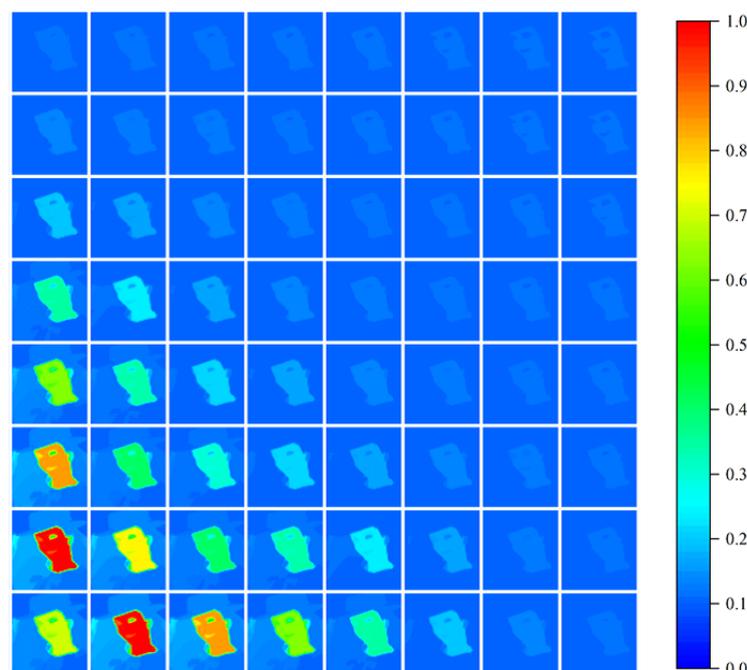


図 4.4 重み付け関数  $W_{i,j}$  の係数値 (画像 “girl”,  $N = 64$ )

### 4.3 視覚の空間周波数特性と画像の顕著性を考慮した空間-周波数領域の最適分割

視覚の空間周波数特性と画像の顕著性を考慮した空間-周波数領域の最適分割は、前節で導出した重み付け関数  $W_{i,j}$  によって重み付けされた歪みの量を、指定された符号化レートの制約条件の下で最小化することによって実現される。このとき、本章においては、入力画像としてカラー画像 (24[bit/pel]) を対象とする。入力画像が与えられたら、まず、式 (4.10) により、RGB 色成分を輝度成分と色差成分である YCrCb 色成分に変換する。

$$\begin{bmatrix} Y \\ Cr \\ Cb \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ 0.500 & -0.419 & 0.081 \\ -0.169 & -0.331 & 0.500 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (4.10)$$

その後、各三つの色成分をそれぞれ独立に符号化するという条件で、各色成分に対して空間-周波数領域の最適分割を適用し、色成分毎に、二次元周波数帯域上の分割パターン、各サブバンドに適用する量子化器、各帯域ブロックにおける二次元空間領域上の分割パターンを求める。このとき、式 (4.10) に示す YCrCb 色成分への変換は直交変換の一種であり、符号化後の再生画像において発生する RGB 色成分全体の歪みは、YCrCb の各色成分において発生する歪みの和で近似できるため、空間-周波数領域の最適分割における最適化処理は、各色成分に対して独立に行ってよい。

空間-周波数領域の最適分割では、サブバンド毎に異なる量子化器を選択できるため、重み付け関数  $W_{i,j}$  を導入すれば、視覚の二次元空間周波数特性を考慮して視知覚感度の低い周波数成分を有するサブバンドの品質を犠牲にする代わりに、視知覚感度の高い周波数成分を有するサブバンドの品質を優先的に保つという処理を実現することができる。しかしながら、同一サブバンド内の信号に対しては全て同じ種類の量子化器が適用されるため、もう一方の

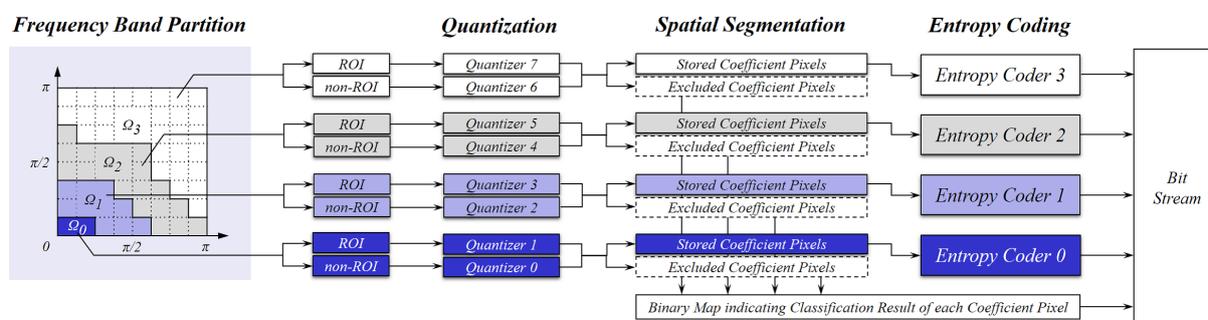


図 4.5 ROI と non-ROI を考慮した量子化の最適化を伴った空間-周波数領域の最適分割の符号化構成 ( $M = 4$ ,  $N = 64$ )

Saliency Map を考慮して空間領域上で顕著度が低い信号の品質を犠牲にする代わりに、顕著度が高い信号の品質を優先的に保つという処理に対応することができない。そこで、提案手法では、 $W_{i,j}$  を定義するために使用した  $\bar{\phi}_j$  の結果に基づいて、各帯域ブロック信号を空間領域上で注目領域 (Region-of-Interest : ROI) と非注目領域 (non-Region-of-Interest : non-ROI) の二つに分割し、各サブバンド内の ROI と non-ROI それぞれに対して、異なる量子化器を選択できるように量子化の最適化処理を拡張することで上記の問題を解決する。これにより、同一サブバンド内では異なる 2 種類の量子化器が割り当てられた信号が混在することとなるが、他二つの最適化処理である周波数帯域分割と空間領域分割、及びエントロピー符号化は、従来の空間-周波数領域の最適分割と同様にサブバンド単位で行われるため、エントロピー符号化器を設計するための処理コストの増加は生じない。図 4.5 には、サブバンド数  $M = 4$ 、帯域ブロック数  $N = 64$  の条件を指定した場合、ROI と non-ROI を考慮した量子化の最適化を伴った空間-周波数領域の最適分割の符号化構成を示す。

### 4.3.1 注目領域 (ROI) と非注目領域 (non-ROI) の分割

帯域ブロック信号における ROI と non-ROI の範囲は、全ての帯域ブロックにおいて共通に定義される。ROI と non-ROI の分割は、前節において、入力画像に対する Saliency Map を縮小することによって定義された  $\bar{\phi}_j$  の係数値を、以下の手順により二値化することで行われる。今、帯域ブロックにおける  $j$  番目の信号に対する ROI と non-ROI の分類状態を二値情報  $\alpha_j$  として表す。ここで、 $\alpha_j = 1$  のときは ROI、 $\alpha_j = 0$  のときは non-ROI を表すものとする。

#### i) 初期化

まず、 $\alpha_j \leftarrow \{0\}$  とおく。また、 $\bar{\phi}_j$  の係数値に対する閾値を  $T$ 、 $\bar{\phi}_j$  における係数の最大値を  $\beta_{max}$  として、 $T \leftarrow \lfloor \beta_{max} \rfloor$  とおく。ここで、 $\lfloor * \rfloor$  は床関数である。

#### ii) ROI の定義

$j$  番目の係数値  $\bar{\phi}_j$  について  $\lfloor \bar{\phi}_j \rfloor = T$  となる場合、 $\alpha_j \leftarrow 1$  として  $j$  番目の信号を ROI に分類する。

#### iii) 閾値の更新

non-ROI に分類されている信号 ( $\alpha_j = 0$ ) の数に対して、ROI に分類されている信号 ( $\alpha_j = 1$ ) の数が  $\rho$  以下である場合は、 $T \leftarrow T - 1$  として更新した後に処理 ii) に戻り、そうでないならば、処理 iv) を実行する。ここで、 $\rho$  は non-ROI に対する ROI の領域比であり、パラメータとして事前に与えられる。

#### iv) 膨張処理

分類された ROI ( $\alpha_j = 1$ ) に対して、 $(3 \times 3)$  の膨張処理を行う。そして、このときの  $\alpha_j$  を帯域ブロックに対する ROI と non-ROI の最終的な分類結果として決定する。



図 4.6 帯域ブロック信号に対する ROI と non-ROI の分割結果（画像 “girl”）

視知覚特性を考慮した空間-周波数領域の最適分割では，non-ROI の品質を犠牲にする分，ROI の品質を改善させるため，ROI に比べて non-ROI の範囲が広ければ広いほど，ROI の品質をより改善させることができる．これを踏まえて，提案手法では，上記の手順 iii) における ROI と non-ROI の領域比を決定するためのパラメータとして  $\rho = 1/5$  を設定した．これにより，処理 iii) が終了した時点で，各帯域ブロック信号はおおよそ 1:5 の領域比で ROI と non-ROI に分割される．また，処理 iv) で ROI の膨張処理を行う理由としては，ROI と non-ROI の境界はほとんどの場合，被写体と背景をまたぐような主要なエッジであることが多く，加えて再生画像におけるエッジ上での歪みは顕著に知覚されるため，エッジ領域を ROI で被覆し，高い量子化レベルの量子化器を割り当てることによってエッジ上での画質劣化を防ぐためである．例として，図 4.6 には，図 4.3(a) に示す画像 “girl” を入力画像として与えた場合に，上記のアルゴリズムによって決定された帯域ブロック信号に対する ROI と non-ROI の分割結果を示している．このとき，入力画像のサイズは  $(480 \times 512)$  画素であり，帯域ブロック数  $N = 64$  の条件を指定しているため，帯域ブロックの信号サイズは  $(60 \times 64)$  である．図 4.6 では，ROI を白画素，non-ROI を黒画素として表している．

### 4.3.2 注目領域（ROI）と非注目領域（non-ROI）を考慮した量子化の最適化

視知覚特性を考慮した空間-周波数領域の最適分割では，重み付け関数  $W_{i,j}$  における Saliency Map による効果を反映するために，サブバンド単位ではなく，各サブバンド内の ROI と non-ROI それぞれに対して異なる量子化器が適用される．このとき，量子化された各サブバンド信号の ROI と non-ROI に対して，後のエントロピー符号化を個別に適用してしまうと，全体の符号量は削減されるものの，エントロピー符号化器の数が単純に倍になってしまうため，符号化器の設計に要する処理コストが大きく増加してしまう．そこで，提案手法では，他二つの最適化処理である周波数帯域分割と空間領域分割，及びエントロピー符号化器の設計は，従来の空間-周波数領域の最適分割と同様，サブバンド単位で行われるように制約を設けた上で，サブバンド内の ROI と non-ROI に対してそれぞれ異なる量子化器が割り当てられるように量子化の最適化処理を拡張する．

空間領域分割やエントロピー符号化器設計のために使用される各サブバンドの確率分布には、量子化後の ROI と non-ROI それぞれの確率分布を統合した混合確率分布が参照される。そのため、量子化の最適化では、ROI と non-ROI とで量子化器の設計を独立に行うのではなく、サブバンド毎に  $W_{i,j}$  により重み付けされた歪みと混合確率分布のレートから成るコスト関数を最小とするように、ROI と non-ROI にそれぞれ対する量子化器の最適な組み合わせを求める必要がある。サブバンド  $\Omega_k$  の ROI と non-ROI に対する量子化器をそれぞれ  $q_{k,1}, q_{k,0}$  としたとき、その組み合わせの最適解は

$$(q_{k,1}, q_{k,0}) = \arg \min_{q_{k,1}, q_{k,0}} \left[ \sum_{i \in I_k, j \in S_i} W_{i,j} \left( w_{i,j} - \hat{w}_{i,j}^{q_{k, \alpha_j}} \right)^2 + \lambda \left( -N_k \sum_{x=x_k^{\min}}^{x_k^{\max}} p_k(x) \log_2 p_k(x) \right) \right] \quad (4.11)$$

により与えられる。ここで、 $\alpha_j$  は、4.3.1 で決定された帯域ブロックにおける  $j$  番目の信号が ROI と non-ROI のどちらに属しているかを判別する二値情報であり、 $\hat{w}_{i,j}^{q_{k, \alpha_j}}$  は、帯域ブロックにおける  $j$  番目の信号を  $q_{k,1}$  または  $q_{k,0}$  で量子化したときの係数値である。このとき、 $q_{k,1}$  と  $q_{k,0}$  のどちらを適用するかは、 $\alpha_j$  に従って決定される。また、

$$p_k(x) = \frac{1}{N_k} \sum_{i \in I_k} \Delta N_i \Delta p_i(x) \quad (4.12)$$

である。このとき、 $\Delta p_i(x)$  は、帯域ブロック  $\Delta \Omega_i$  における ROI と non-ROI の確率分布を統合した混合確率分布であり、

$$\Delta p_i(x) = \frac{\Delta N_{i,1} \cdot \Delta \Omega_{i,1}^{q_{k,1}}(x \cdot |q_{k,1}|) + \Delta N_{i,0} \cdot \Delta \Omega_{i,0}^{q_{k,0}}(x \cdot |q_{k,0}|)}{\Delta N_i} \quad (4.13)$$

により計算される。式(4.13)において、 $\Delta \Omega_{i,1}^{q_{k,1}}(-)$  と  $\Delta \Omega_{i,0}^{q_{k,0}}(-)$  は、帯域ブロック  $\Delta \Omega_i$  における ROI と non-ROI をそれぞれ量子化器  $q_{k,1}$  と  $q_{k,0}$  で量子化した後、ROI 内で保存される信号  $j(j \in S_i \cap \alpha_j = 1)$  から求められる確率分布と、non-ROI 内で保存される信号  $j(j \in S_i \cap \alpha_j = 0)$  から求められる確率分布であり、 $\Delta N_{i,1}$  と  $\Delta N_{i,0}$  は、帯域ブロック  $\Delta \Omega_i$  における ROI と non-ROI 内それぞれにおいて保存される信号数である。すなわち、 $\Delta N_{i,1}$  と  $\Delta N_{i,0}$  の和は、 $\Delta N_i$  と等価になる。また、 $|q_{k,1}|$  と  $|q_{k,0}|$  は、それぞれ量子化器  $q_{k,1}$  と  $q_{k,0}$  の量子化ステップ幅である。式(4.13)に示すように、混合確率分布  $\Delta p_i(x)$  は、ROI の確率分布  $\Delta \Omega_{i,1}^{q_{k,1}}(-)$  と non-ROI の確率分布  $\Delta \Omega_{i,0}^{q_{k,0}}(-)$  の各ヒストグラムビンをそれぞれ一旦ゼロ周りに集中させてから、二つの確率分布を統合することで構成される。このとき、仮に  $\Delta \Omega_{i,1}^{q_{k,1}}(-)$  と  $\Delta \Omega_{i,0}^{q_{k,0}}(-)$  を単純に加算して  $\Delta p_i(x)$  を構成してしまうと、 $q_{k,1}$  と  $q_{k,0}$  の量子化ステップ幅が異なる場合に、 $\Delta p_i(x)$  のヒストグラムビンが大量に増加し、エントロピーが著しく増加してしまう。 $\Delta p_i(x)$  を式(4.13)により構成する理由は、これを回避するためである。後に行われる各サブバンドに対するエントロピー符号化器の設計では、式(4.12)により与えられる混合確率分布  $p_k(x)$  が参照され

る。以上を踏まえて、復号化側では、再生される帯域ブロック  $\Delta\Omega_i$  における  $j$  番目の信号の係数値  $y_{i,j}$  を、逆量子化の工程において、

$$y_{i,j} = |q_{k,\alpha_j}| \cdot \hat{y}_{i,j} \quad (4.14)$$

として再構成する。ここで、 $\hat{y}_{i,j}$  は、エントロピー復号化の直後に復号化された帯域ブロック  $\Delta\Omega_i$  における  $j$  番目の信号の係数値である。式(4.14)による係数値の再構成について、以下より具体的な例を挙げて説明する。例えば、量子化器  $q_{k,1}$  と  $q_{k,0}$  の量子化ステップ幅としてそれぞれ  $|q_{k,1}| = 2$  と  $|q_{k,0}| = 5$  が決定され、復号化直後の係数値  $\hat{y}_{i,j} = 2$  が与えられた場合、 $j$  番目の信号が ROI に属するものである場合は  $y_{i,j} = 4$ 、non-ROI に属するものである場合は  $\hat{y}_{i,j} = 10$  がそれぞれ再構成される。

さらに、量子化の最適化だけではなく、他二つの最適化処理である周波数帯域分割と空間領域分割についても、 $W_{i,j}$  による歪み値への重み付けや、エントロピー符号化をサブバンド単位で行うために必要な混合確率分布  $p_k(x)$  及び  $\Delta p_i(x)$  が考慮されるように、処理を拡張する。二次元周波数帯域上の最適分割パターン  $I_k$  は、 $N$  個の帯域ブロック  $\Delta\Omega_i$  における ROI と non-ROI の混合確率分布  $\Delta p_i(x) (i = 0, 1, \dots, N-1)$  を、式(4.15)を満たすようにクラスタリングすることによって決定される。

$$I_k = \arg \min_{I_k} \left[ \sum_{k=0}^{M-1} \sum_{i \in I_k} \Delta N_i \sum_{x=x_i^{\min}}^{x_i^{\max}} \Delta p_i(x) \log_2 \frac{\Delta p_i(x)}{p_k(x)} \right], \quad (4.15)$$

一方、帯域ブロック  $\Delta\Omega_i$  における二次元空間領域上の最適分割パターン  $S_i$  は、 $j$  番目の信号を保存した場合のコスト関数を

$$J_{i,j} = W_{i,j} \left( w_{i,j} - \hat{w}_{i,j}^{q_{k,\alpha_j}} \right)^2 + \lambda \left\{ -\log_2 p_k \left( \hat{w}_{i,j}^{q_{k,\alpha_j}} \right) \right\} \quad (4.16)$$

、削除した場合のコスト関数を  $W_{i,j} \cdot w_{i,j}^2$  として定義し、それらの大小を比較することによって決定される。

視知覚特性を考慮した空間-周波数領域の最適分割では、YCrCb の色成分毎に、式(4.11)、式(4.15)、式(4.16)に示した三つの最適化問題を、第3章の??で提案したアルゴリズムを使用して反復的に解くことにより、 $W_{i,j}$  により重み付けられた歪みの量を最小化するという観点で、指定された符号化レートに応じた二次元周波数帯域上の最適分割パターン  $I_k$ 、各サブバンド  $\Omega_k$  の ROI と non-ROI に対する最適な量子化器  $q_{k,1}, q_{k,0}$ 、各帯域ブロック  $\Delta\Omega_i$  における二次元空間領域上の最適分割パターン  $S_i$  が決定される。また、復号化側で帯域ブロック信号の係数値を再構成するために必要となる、信号の ROI と non-ROI を判別するための二値情報  $\alpha_j$  は、ランレングス符号化を用いて符号化され、付加情報として復号化側に伝送される。

## 4.4 シミュレーション実験とその結果

実画像に対して、本章で提案する視知覚特性を考慮した空間-周波数領域の最適分割と、第3章で提案した視知覚特性を考慮しない通常的空間-周波数領域の最適分割を適用する。視知覚特性を考慮した空間-周波数領域の最適分割では、YCrCbの各色成分に対する二次元周波数帯域上の分割パターン、各サブバンドのROIとnon-ROIに適用される量子化器、各帯域ブロックにおける二次元空間領域上の分割パターンがそれぞれ決定される。一方、視知覚特性を考慮しない空間-周波数領域の最適分割では、歪み値に対して重み付け関数 $W_{i,j}$ による重み付けや、ROIとnon-ROIを考慮した量子化の最適化処理を行わず、YCrCbの各色成分毎に、帯域分割パターン、サブバンド単位で適用される量子化器、空間領域分割パターンが決定される。

入力画像には、視知覚特性を考慮した空間-周波数領域の最適分割において、Saliency Mapに基づいた $W_{i,j}$ による重み付けや、ROIとnon-ROIを考慮した量子化の最適化処理による効果を十分に発揮できるように、Saliency Mapによって画像中の顕著領域を特定することが比較的容易であると考えられる、図4.7に示すCIPR標準画像[68]における画像“girl”((480×512)画素, 24[bit/pel])と画像“model”((512×480)画素, 24[bit/pel])、Kodak標準画像[68]における画像“kodim04”((512×768)画素, 24[bit/pel])、AIC JPEG ad-hoc groupが提供する高解像度の標準画像[69]における画像“p22”をリサイズしたもの((640×800)画素, 24[bit/pel])の4枚のポートレート画像を選択した。参考として、図4.8には図4.7に示す各テスト画像に対して、Yanらの手法[78]を適用することで得られたSaliency Mapの結果を示す。図4.8の結果から、各Saliency Mapは、図4.7に示す各画像の顕著領域をおおよそ特定できていると言える。

二次元周波数帯域上の分割パターンを求めるときの条件である帯域ブロック数 $N$ とサブバンド数 $M$ は、視知覚特性を考慮した場合と考慮しない場合のいずれにおいても、それぞれ $N = 64$ ,  $M = 5$ とした。また、量子化器の候補には、量子化ステップ幅(1, 2, ..., 16)を持つ $P = 16$ 種類のミッドトレッド型の線形量子化器 $Q_1 \sim Q_{16}$ を用意した。視知覚特性を考慮した空間-周波数領域の最適分割において、重み付け関数 $W_{i,j}$ を導出するために用いられる視覚の二次元空間周波数特性は、4.2.1で例に挙げた図4.1に示す視覚の二次元空間周波数特性と同様、 $D=27(\text{inch})$ ,  $H=1920(\text{pixels})$ ,  $V=1080(\text{pixels})$ ,  $d=50(\text{cm})$ の測定条件で求めることとする。また、ROIとnon-ROIを考慮した量子化の最適化処理に際して必要となる、帯域ブロック信号におけるROIとnon-ROIの領域比を決定するためのパラメータ $\rho$ には、4.3.1で述べたように、 $\rho = 1/5$ を設定した。



(a) 画像“girl” ((480×512)画素)



(b) 画像“model” ((512×480)画素)



(c) 画像“kodim04” ((512×768)画素)

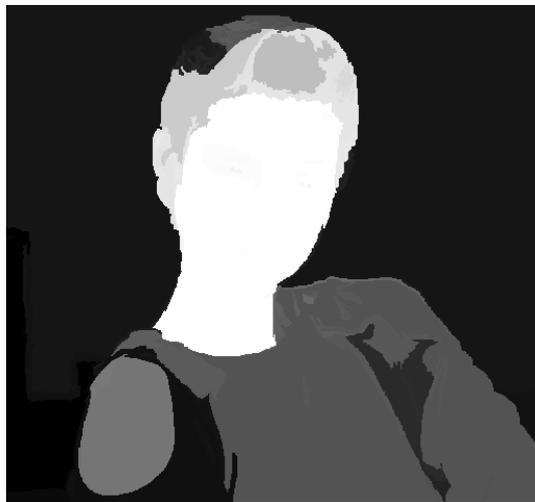


(d) 画像“p22” ((640×800)画素)

図 4.7 テスト画像 (24[bit/pel])



(a) 画像“girl”



(b) 画像“model”



(c) 画像“kodim04”



(d) 画像“p22”

图 4.8 Saliency Map

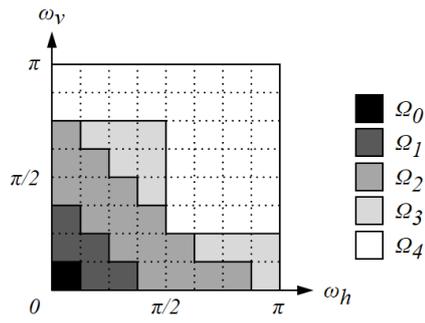
まず、視知覚特性を考慮した空間-周波数領域の最適分割の適用により、YCrCbの各色成分に対して求められた二次元周波数帯域上の分割パターン、各サブバンドのROIとnon-ROIに適用される量子化器、各帯域ブロックにおける二次元空間領域上の分割パターンを、視知覚特性を考慮しない空間-周波数領域の最適分割によるものと比較し、その違いについて考察する。次に、実際に得られた再生画像に対する主観評価実験を行い、視知覚特性を考慮した空間-周波数領域の最適分割による再生画像の画質を評価すると共に、視知覚特性を考慮しない空間-周波数領域の最適分割及び画像符号化の国際標準方式JPEG2000と比較する。

#### 4.4.1 周波数帯域分割・量子化・空間領域分割の結果

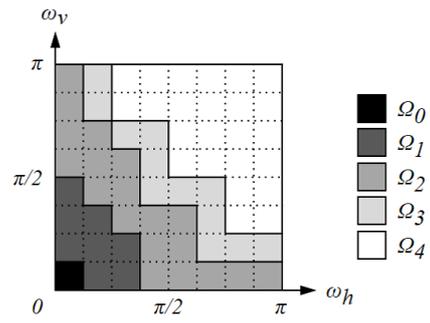
前節までにおいて、具体的な画像の例として取り上げた図4.7(a)に示す画像“girl”を入力画像としたとき、符号化レート0.10[bit/pel]の条件で求められた二次元周波数帯域上の分割パターン、各サブバンドのROIとnon-ROI（または、サブバンド単位）に適用される量子化器、各帯域ブロックにおける二次元空間領域上の分割パターンの結果について、比較及び考察を行う。図4.9は、YCrCbの各色成分に対して、(i)視知覚特性を考慮した空間-周波数領域の最適分割及び(ii)視知覚特性を考慮しない空間-周波数領域の最適分割によって求められた二次元周波数帯域上の分割パターンを示している。表4.1~4.3は、図4.9の(i)に示した各サブバンド信号におけるROIとnon-ROIに適用される量子化器の組み合わせと、図4.9の(ii)に示した各サブバンド信号全体に適用される量子化器の組み合わせを、色成分毎に示した結果である。ここで、表における-の記号は、該当する領域内の信号が空間領域分割によって全て削除されたために、量子化器が割り当てられていないことを表している。図4.10は、図4.9に示す各色成分毎に求められた各帯域ブロック信号における二次元空間領域上の分割パターンを示している。ここで、図4.10に示した空間領域上の分割パターンでは、それぞれ黒画素が保存される信号、図中の右側に示す各サブバンドに対応付けて色付けされた画素が該当するサブバンド内で削除された信号を表している。

これらの結果において、まず、図4.9に示す帯域分割パターンと図4.10に示す空間領域分割パターンの結果を比較すると、同じ符号化レートを指定したにも関わらず、視知覚特性を考慮した場合としない場合とで異なる内容が求められていることから、それぞれが視知覚特性を考慮した上で最適化されていることがわかる。まず、図4.9に示す帯域分割パターンの結果を観察する。Y成分に対する結果を観察すると、視知覚特性を考慮した場合に求められた帯域分割パターンは、視知覚特性を考慮しない場合よりも、二次元周波数帯域上で、低域及び中域側のサブバンドが占める面積が小さくなっており、その分、高域側におけるサブバンドの占める面積が大きくなっていることがわかる。これは、重み付け関数 $W_{i,j}$ における一つの効果である視覚の二次元空間周波数特性が考慮されたことによって、低域側のサブバンドが視知覚感度の高い周波数成分を有する帯域ブロック群、高域側のサブバンドが視知覚感度の低い周波数成分を有する帯域ブロック群により構成されたことによるものである。これ

とは逆に、Cr及びCb成分において視知覚特性を考慮した場合に求められた帯域分割パターンは、視知覚特性を考慮しない場合よりも、低中間域におけるサブバンドの面積が若干大きくなっている。これは、カラー画像における色成分のほとんどは、Y成分にエネルギーが集中するため、Cr及びCb成分ではあまりエネルギーを持たずに低域側に信号電力が集中しやすくなるものの、視覚の二次元空間周波数特性が考慮されたことで、低中間域における周波数成分が視知覚特性上で重要であると判断され、各サブバンドがその周波数成分を有する帯域ブロックを含めるように構成されたためである。次に、図4.10に示す空間領域分割パターンの結果を観察する。Y成分に対する結果を観察すると、視知覚特性を考慮した場合、低域側の帯域ブロックではより多くの信号が保存され、高域側にある帯域ブロックではより多くの信号が削除されている。一方、Cr及びCb成分においては、視知覚特性を考慮しない場合では削除の対象となっていた低中間域における帯域ブロックの信号が、視知覚特性を考慮したことで保存されていることがわかる。これらはいずれも、視覚の二次元空間周波数特性を考慮したことによって、視知覚感度が低い周波数成分を有する帯域ブロック信号の品質を犠牲にした分、視知覚感度が高い周波数成分を有する帯域ブロック信号の品質が優先的に保持されたためである。さらに、低域側の帯域ブロックについて詳細を観察すると、全ての色成分に共通して、視知覚特性を考慮した場合では、視知覚特性を考慮しない場合と比べて、画像“girl”の顔部分に相当する領域内でより多く信号が保存されている。これは、画像“girl”は顔領域に対する顕著性が高く、さらに、重み付け関数  $W_{i,j}$  における二つ目の効果である Saliency Map が考慮されたことによって、帯域ブロック内で顕著度が高い顔領域の空間領域情報を持った信号の品質が優先的に保たれたためである。次に、表4.1~4.3に示す量子化器の組み合わせ結果を比較すると、視知覚特性を考慮しない空間-周波数領域の最適分割ではサブバンド毎に1種類の量子化器が信号全体に割り当てられているのに対して、視知覚特性を考慮した空間-周波数領域の最適分割では、各サブバンドのROIとnon-ROIでそれぞれ異なる種類の量子化器が選択されていることがわかる。このとき、視知覚特性を考慮した空間-周波数領域の最適分割では、色成分に関わらずほとんどのサブバンドにおいて、non-ROIよりもROIに対して量子化ステップ幅が小さい量子化器が割り当てられている。これは、Saliency Map が考慮されたことによって、non-ROIよりもROI内にある信号の品質が優先的に保持されたためである。ただし、図4.9(a)に示すY成分に対する帯域分割パターンのように、低域側のサブバンドが帯域内で広がり、高域側のサブバンドの占める面積が比較的小さくなっている場合、表4.1の結果から、高域側のサブバンド対してはROIとnon-ROIに関わらず、どちらも小さい量子化ステップ幅の量子化器が割り当てられていることがわかる。これは、空間領域上の顕著度の高さに関わらずサブバンド自体が視知覚感度の低い周波数成分を持つものであった場合、重み付け関数  $W_{i,j}$  における視覚の二次元空間周波数特性及び Saliency Map の効果が同時に考慮されたことによって、そのサブバンド全体の品質を犠牲にする代わりに、視知覚感度の高い周波数成分を有し、且つ、顕著性が高い領域の品質が優先的に保持されたためである。以上の傾向は、図4.7に示す他のテスト画像においても同様に見られた。

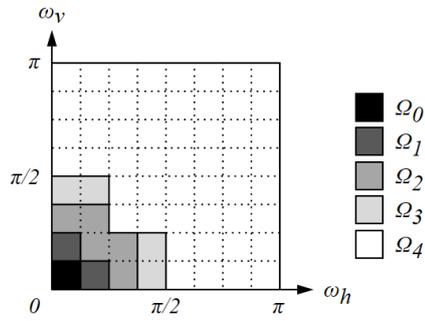


(i) 視知覚特性を考慮した場合

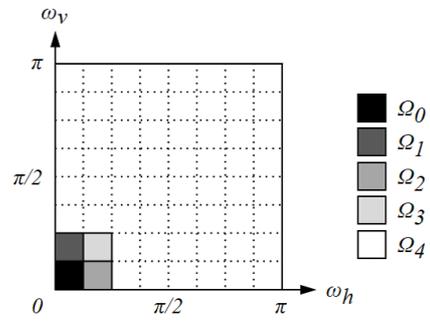


(ii) 視知覚特性を考慮しない場合

(a) Y 成分

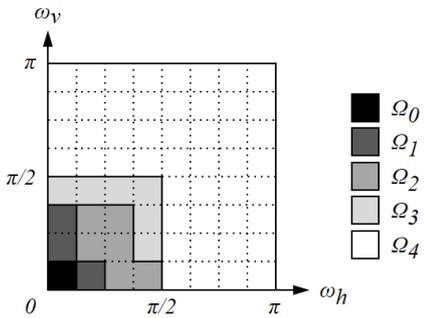


(i) 視知覚特性を考慮した場合

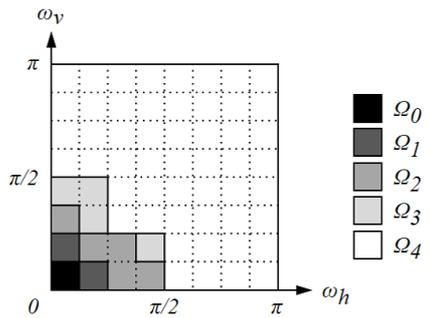


(ii) 視知覚特性を考慮しない場合

(b) Cr 成分



(i) 視知覚特性を考慮した場合



(ii) 視知覚特性を考慮しない場合

(c) Cb 成分

図 4.9 視知覚特性を考慮した場合と考慮しない場合の空間-周波数領域の最適分割による帯域分割パターン (画像 “girl”,  $M = 5$ ,  $N = 64$ )

表 4.1 図 4.9(a) に示す Y 成分の各サブバンド信号における ROI と non-ROI  
 (または, サブバンド信号全体) に適用される量子化器

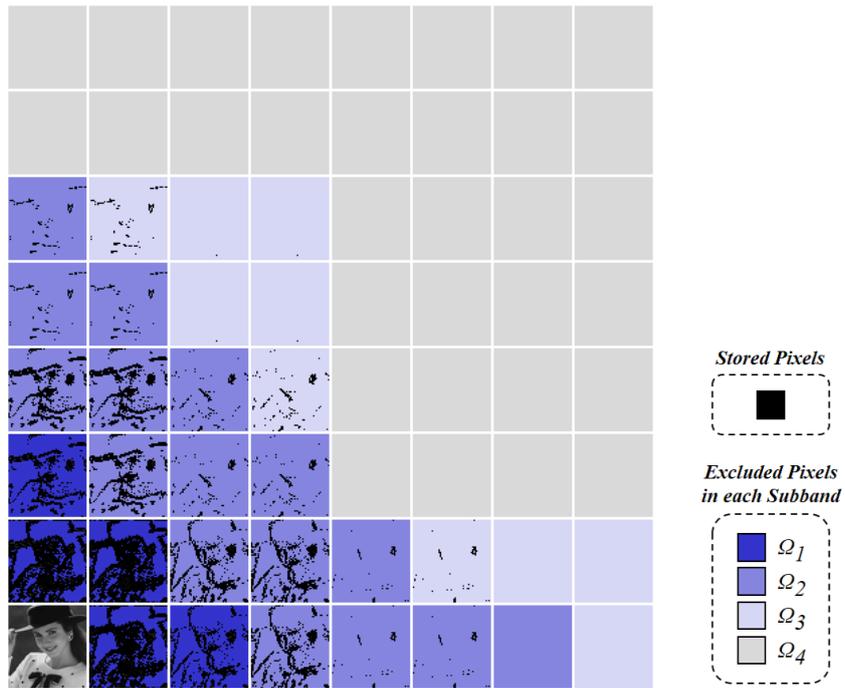
subband	quantizer applied to ROI and non-ROI on each subband (or the whole subband)		
	Figure 4.9(a)(i)		Figure 4.9(a)(ii)
	ROI	non-ROI	
$\Omega_0$	$Q_3$	$Q_5$	$Q_5$
$\Omega_1$	$Q_3$	$Q_6$	$Q_5$
$\Omega_2$	$Q_5$	$Q_8$	$Q_6$
$\Omega_3$	$Q_{16}$	$Q_{16}$	$Q_8$
$\Omega_4$	–	–	–

表 4.2 図 4.9(b) に示す Cr 成分の各サブバンド信号における ROI と non-ROI  
 (または, サブバンド信号全体) に適用される量子化器

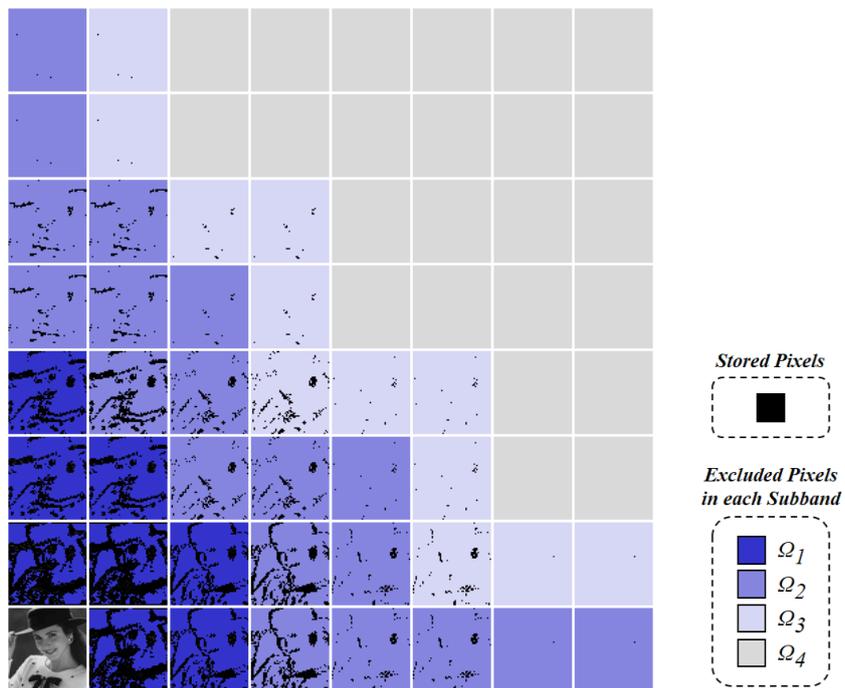
subband	quantizer applied to ROI and non-ROI on each subband (or the whole subband)		
	Figure 4.9(b)(i)		Figure 4.9(b)(ii)
	ROI	non-ROI	
$\Omega_0$	$Q_3$	$Q_5$	$Q_5$
$\Omega_1$	$Q_3$	$Q_7$	$Q_8$
$\Omega_2$	$Q_4$	$Q_9$	$Q_6$
$\Omega_3$	$Q_{16}$	–	$Q_8$
$\Omega_4$	–	–	–

表 4.3 図 4.9(c) に示す Cb 成分の各サブバンド信号における ROI と non-ROI  
 (または, サブバンド信号全体) に適用される量子化器

subband	quantizer applied to ROI and non-ROI on each subband (or the whole subband)		
	Figure 4.9(c)(i)		Figure 4.9(c)(ii)
	ROI	non-ROI	
$\Omega_0$	$Q_4$	$Q_5$	$Q_5$
$\Omega_1$	$Q_3$	$Q_7$	$Q_6$
$\Omega_2$	$Q_4$	$Q_8$	$Q_7$
$\Omega_3$	$Q_5$	$Q_7$	$Q_5$
$\Omega_4$	–	–	–



(i) 視知覚特性を考慮した場合



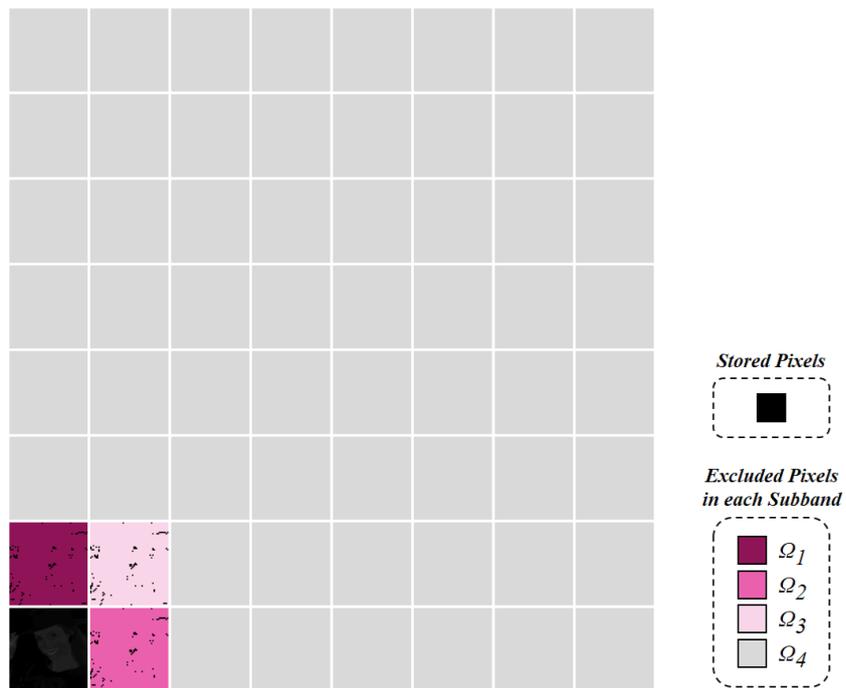
(ii) 視知覚特性を考慮しない場合

(a) Y 成分

図 4.10 視知覚特性を考慮した場合と考慮しない場合の空間-周波数領域の最適分割による空間領域分割パターン（画像“girl”， $M = 5$ ， $N = 64$ ）



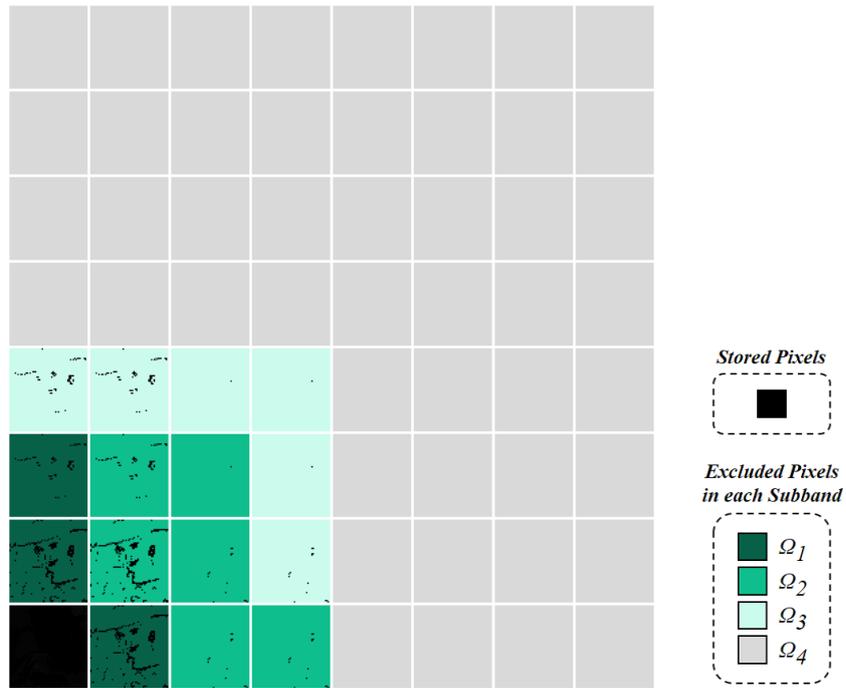
(i) 視知覚特性を考慮した場合



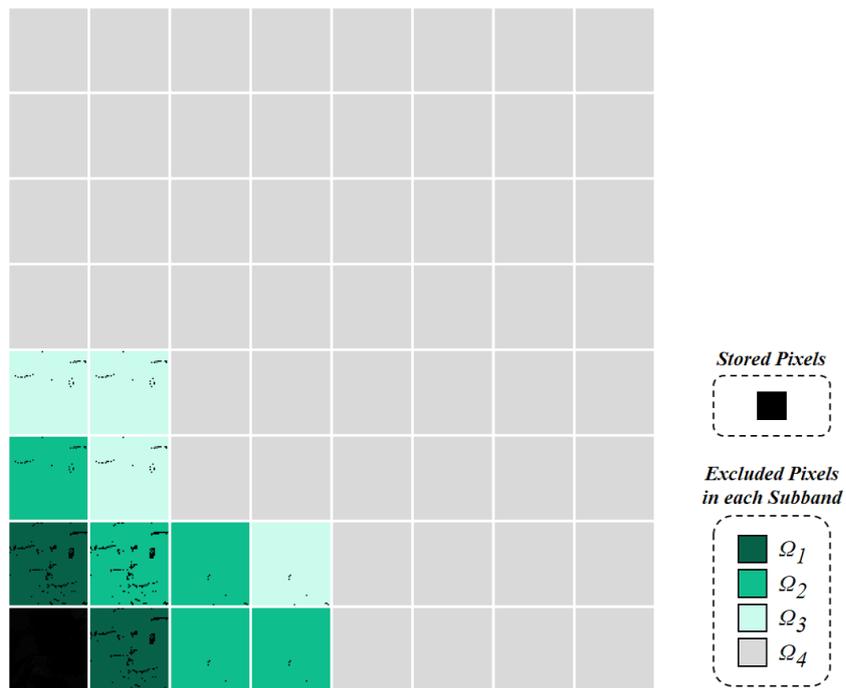
(ii) 視知覚特性を考慮しない場合

(b) Cr 成分

図 4.10 視知覚特性を考慮した場合と考慮しない場合の空間-周波数領域の最適分割による空間領域分割パターン (画像 “girl”,  $M = 5$ ,  $N = 64$ )



(i) 視知覚特性を考慮した場合



(ii) 視知覚特性を考慮しない場合

(c) Cb 成分

図 4.10 視知覚特性を考慮した場合と考慮しない場合の空間-周波数領域の最適分割による空間領域分割パターン (画像 “girl”,  $M = 5$ ,  $N = 64$ )

## 4.4.2 主観評価実験による画質評価

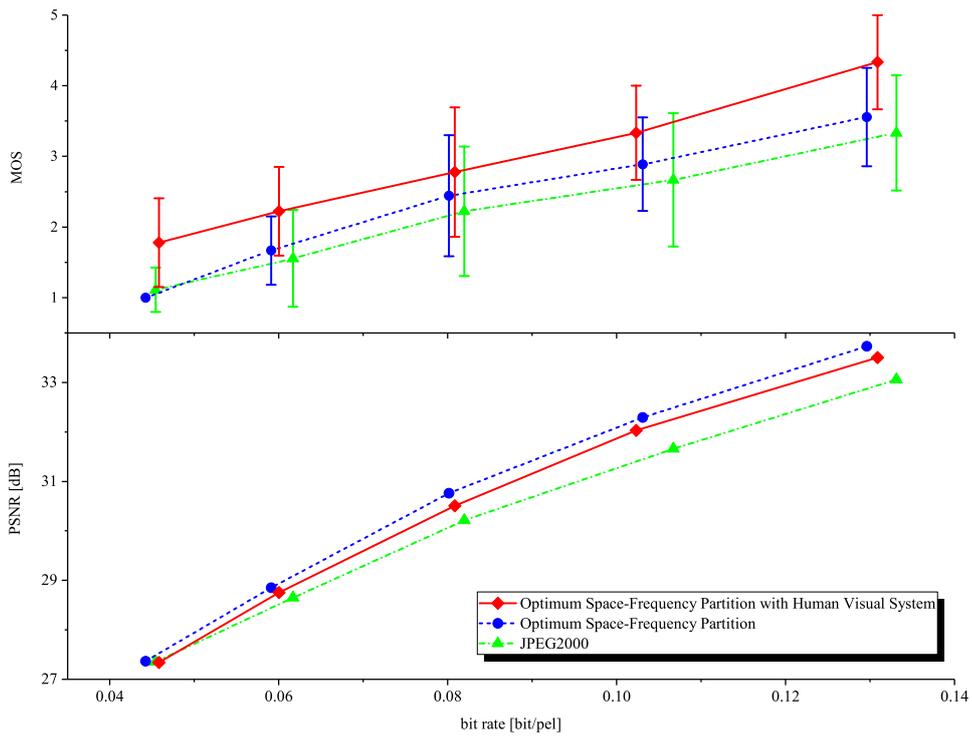
図4.7に示す各テスト画像を入力画像として与えたとき、視知覚特性を考慮した空間-周波数領域の最適分割の適用によって得られた再生画像の画質を主観評価実験により評価し、また、視知覚特性を考慮しない空間-周波数領域の最適分割及びJPEG2000を適用した場合と比較する。このとき、JPEG2000における離散ウェーブレット変換は、空間-周波数領域の最適分割と同様に、サブバンドの再帰分解レベルを3に設定した。主観評価実験は、表4.4に示した条件に基づいて、評定者に対して実験の目的を明示することで行った。評定者には、ITU-T勧告P.910で規定される絶対範疇尺度法（Absolute Category Rating Method：ACR法）に基づき、三つの各符号化手法によって得られた再生画像を提示し、それらの再生画像の画質を表4.4の下部に示した5段階評価により評価させた。このとき、提示する再生画像には、テスト画像1枚あたりに5種類の異なる符号化レートを指定した条件で三つの各符号化手法を適用し、得られた計15枚の再生画像を用いることとし、評価値は平均オピニオン評点（Mean Opinion Score：MOS）

$$MOS = \frac{1}{K} \sum_{n=0}^{K-1} a_n \quad (4.17)$$

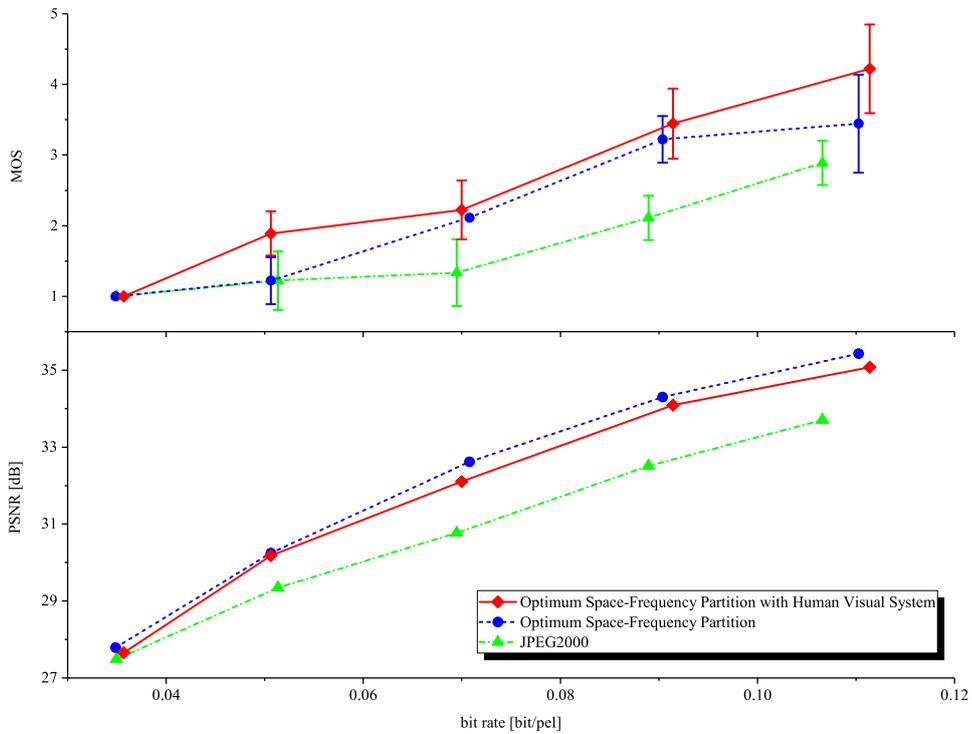
により算出した。ここで、 $K$ は評定者数、 $a_n$ は $n$ 番目の評定者の評点である。図4.11は、図4.7に示す各画像に対して、視知覚特性を考慮した空間-周波数領域の最適分割を適用した場合に求められたMOS対ビットレート特性を、視知覚特性を考慮しない空間-周波数領域の最適分割及びJPEG2000によるものと比較した結果である。なお、同図には、このときに求められた画質の客観評価尺度PSNR対ビットレート特性を併せて表示する。

表 4.4 主観評価実験の条件

評価法		絶対範疇尺度法（ACR法）
ディスプレイ解像度	$D$ (inch)	27 (inch)
	$H \times V$ (pixels)	1920 × 1080 (pixels)
視距離 $d$ (cm)		50 (cm)
画像提示時間 (sec)		10 (sec)
評定者数 $K$		10
評点	評価語	
5	非常に良い（Excellent）	
4	良い（Good）	
3	普通（Fair）	
2	悪い（Poor）	
1	非常に悪い（Bad）	

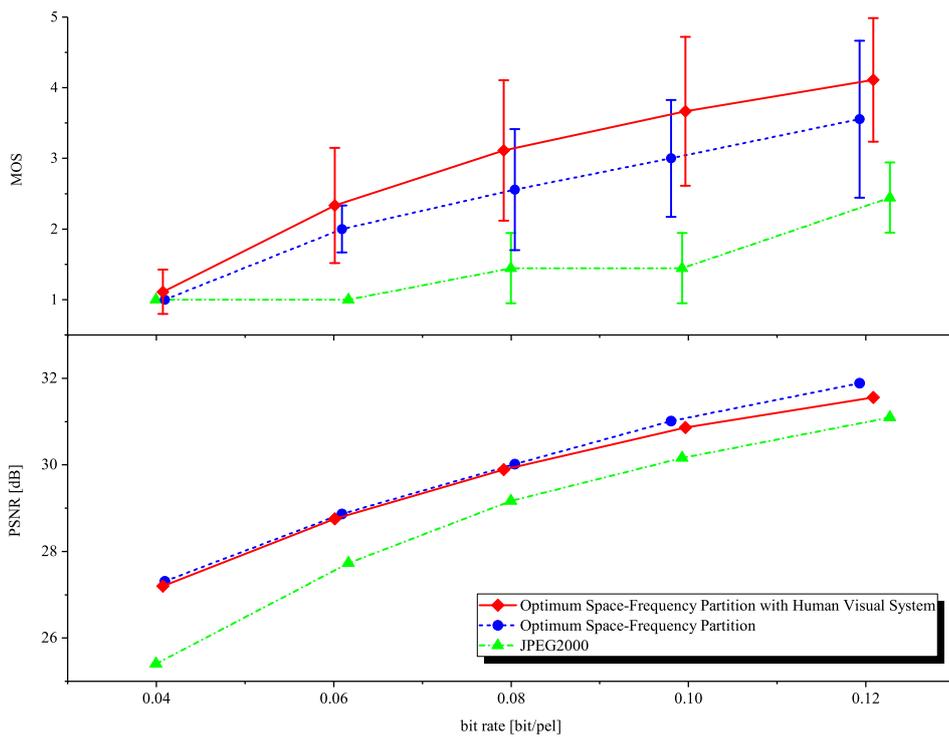


(a) 画像 "girl"

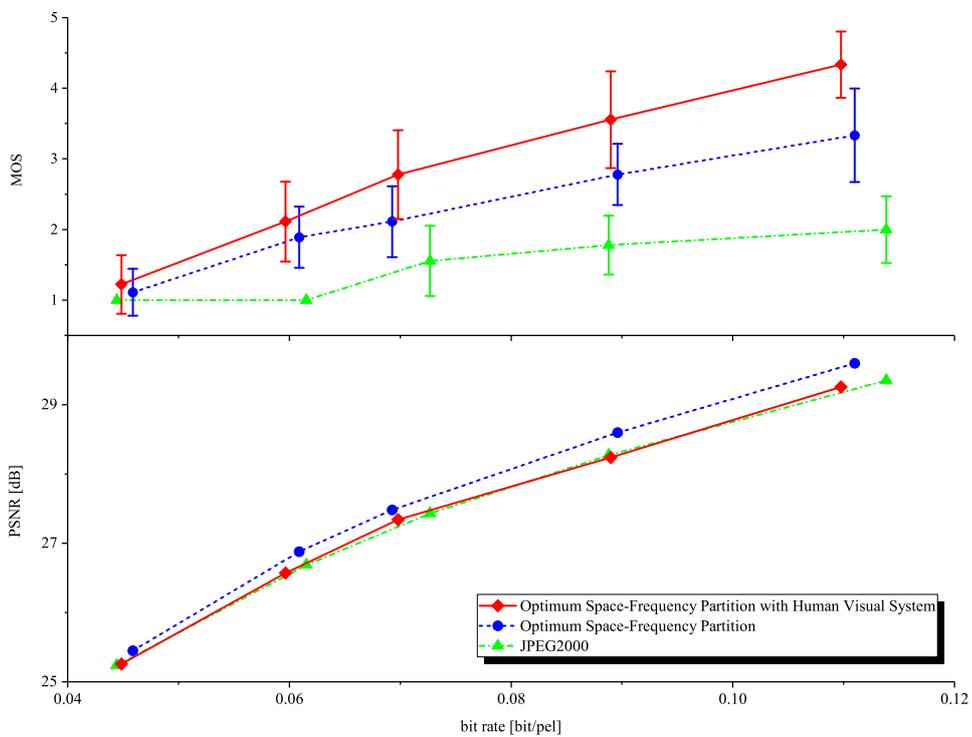


(b) 画像 "model"

図 4.11 MOS 対ビットレート特性及び PSNR 対ビットレート特性



(c) 画像“kodim04”



(d) 画像“p22”

図 4.11 MOS 対ビットレート特性及び PSNR 対ビットレート特性

図 4.11 の結果から、視知覚特性を考慮した空間-周波数領域の最適分割は、重み付け関数  $W_{i,j}$  によって重み付けされた歪みの量を最小とするように最適化を行っているため、PSNR 値は視知覚特性を考慮しない空間-周波数領域の最適分割と比べて若干低下しているものの、MOS 値は各テスト画像のほとんど全ての符号化レートにおいて、最も高い値が得られていることがわかる。このとき、その差は最大で 0.8 程度であった。また、JPEG2000 との比較においては、MOS 値と PSNR 値はいずれもほぼ全てのテスト画像において高い評価値を得ることができており、特に MOS 値に関しては、最大で 2.2 程度の高い値を示している。以上のことから、視知覚特性を考慮した空間-周波数領域の最適分割を適用することにより、視知覚特性を考慮した画質の観点においても、良好な再生画像を得られることが明らかにされた。

最後に、図 4.7 に示した各画像について、ある特定の符号化レートを与えた場合に、視知覚特性を考慮した空間-周波数領域の最適分割の適用によって得られた再生画像を図 4.12～図 4.15 に示す。なお、各図においては、本手法との比較のために、同符号化レートの条件下で、視知覚特性を考慮しない空間-周波数領域の最適分割の適用によって得られた再生画像及び JPGE2000 を適用した場合の再生画像を並べて表示している。また、主観評価実験の結果得られた MOS 値を記載している。まず、JPEG2000 により符号化された再生画像を観察すると、顔や背景等にある平坦領域での疑似輪郭や、斑点模様のような色のアーティファクトが発生しており、局所的に視知覚上画質劣化の著しい部分が観測される。これと比較して、視知覚特性を考慮した空間-周波数領域の最適分割及び視知覚特性を考慮しない空間-周波数領域の最適分割による再生画像を観察すると、上述したようなアーティファクトが発生しておらず、主観的に高い画質が得られていることがわかる。次に、視知覚特性を考慮した場合と考慮しない場合の空間-周波数領域の最適分割それぞれによって得られた再生画像を比較すると、視知覚特性を考慮した場合、4 枚のテスト画像全てに共通して、顔領域の画質が著しく改善されていることがわかる。これは、重み付け関数  $W_{i,j}$  を導入して歪み量を最小化したことで、顕著性が低い背景等の non-ROI に相当する領域の品質や、画像全体に含まれている視知覚感度が低い周波数成分を有する信号の品質が犠牲にされた分、画像中で特に顕著性が高い ROI に相当する顔領域の画質が優先的に保持されたことによる効果である。



(a) 原画像



(b) 視知覚特性を考慮した空間-周波数領域の最適分割 (MOS=3.33)



(c) 視知覚特性を考慮しない空間-周波数領域の最適分割 (MOS=2.89)



(d) JPEG2000 (MOS=2.67)

図 4.12 各手法によって得られた再生画像 (画像“girl”, 0.10[bit/pel])



(a) 原画像



(b) 視知覚特性を考慮した空間-周波数領域の最適分割 (MOS=3.44)



(c) 視知覚特性を考慮しない空間-周波数領域の最適分割 (MOS=3.22)



(d) JPEG2000 (MOS=2.11)

図 4.13 各手法によって得られた再生画像 (画像“model”, 0.09[bit/pel])



(a) 原画像



(b) 視知覚特性を考慮した空間-周波数領域の最適分割 (MOS=3.67)



(c) 視知覚特性を考慮しない空間-周波数領域の最適分割 (MOS=3.00)



(d) JPEG2000 (MOS=1.44)

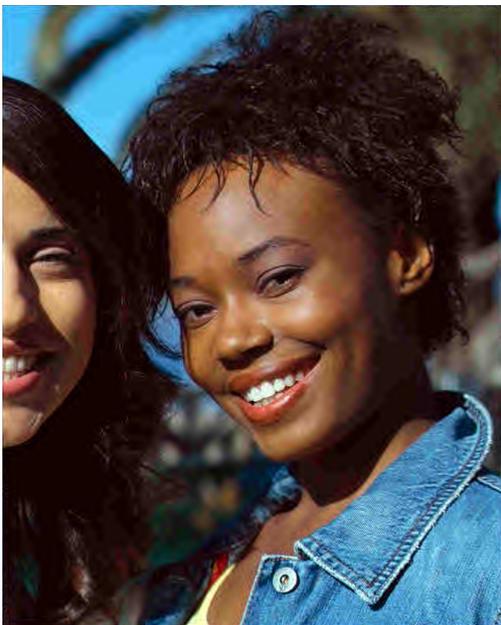
図 4.14 各手法によって得られた再生画像 (画像“kodim04”, 0.10[bit/pel])



(a) 原画像



(b) 視知覚特性を考慮した空間-周波数領域の最適分割 (MOS=3.56)



(c) 視知覚特性を考慮しない空間-周波数領域の最適分割 (MOS=2.78)



(d) JPEG2000 (MOS=1.78)

図 4.15 各手法によって得られた再生画像 (画像“p22”, 0.09[bit/pel])

## 4.5 結言

本章では、第3章で提案した空間-周波数領域の最適分割よりも主観的に良好な再生画像を得ることを目的として、人間の視知覚特性上で重要ではない信号の品質を犠牲にする代わりに重要な信号の品質を優先的に保つことで、主観的な画質の改善を達成する「視知覚特性を考慮した空間-周波数領域の最適分割」を提案した。そして、本空間-周波数領域の最適分割を実画像に適用し、得られた再生画像の画質を評価するための主観評価実験を行うことで、その有効性を明らかにした。

本空間-周波数領域の最適分割は、人間の視知覚特性を考慮して各帯域ブロック信号の歪み値に重み付けを行い、その重み付けされた歪みの量を最小化とするようにレート歪み最適化問題を解くことで、再生画像における主観的な画質を改善する。まず、人間の視知覚特性を反映するものとして視覚の空間周波数特性と画像の顕著性の二つに着目し、まずはそれらの特性を考慮するために、画像信号に対する視覚の二次元空間周波数特性と、入力画像の顕著性を推定する Saliency Map に基づいて、各帯域ブロック信号の歪み値に対する重み付け関数を定義した。次に、より画像の顕著性を考慮するために、Saliency Map の結果に基づいて各帯域ブロック信号を空間領域上で注目領域 (Region-of-Interest : ROI) と非注目領域 (non-Region-of-Interest : non-ROI) の二つに分割し、各サブバンド内の ROI と non-ROI それぞれに対して、異なる量子化器を選択できるように量子化の最適化処理を拡張した。最後に、各サブバンドの ROI と non-ROI にそれぞれ異なる量子化器が割り当てられることを考慮した上で、重み付けられた歪みの量を最小化するように、所望の符号化レートに応じた二次元周波数帯域の最適分割パターン、各サブバンドにおける ROI と non-ROI それぞれに適用する最適な量子化器の組み合わせ、各帯域ブロックにおける二次元空間領域上の最適分割パターンを求めるアルゴリズムを提案した。

最後に、視知覚特性を考慮した空間-周波数領域の最適分割を実画像に適用し、得られた再生画像の画質を評価するために主観評価実験を行った。その結果、5段階評価に対応する MOS 値において、視知覚特性を考慮しない通常的空間-周波数領域の最適分割と比較して最大で約 0.8 程度、JPEG2000 と比較して最大で約 2.2 程度の改善があることが示された。以上により、視知覚特性を考慮した空間-周波数領域の最適分割は、視知覚特性を考慮した画質の観点において、良好な再生画像を得るための有効な一手法であることが示された。

## 第 5 章

### 結論

#### 5.1 本研究で得られた成果

本論文では、画像のサブバンド符号化において、少ないサブバンド数で適応的な帯域分割を実現する最適帯域分割の符号化性能の改善を目的として、未だ議論が行われていなかったレート歪み理論の観点から、帯域分割及び量子化の最適化を行う理論とその手法について述べた。また、周波数領域だけではなく空間領域にも着目し、帯域分割と量子化のみならず、空間領域上の冗長性を除去するための空間領域分割の三つをレート歪み理論の観点で同時に最適化する空間-周波数領域の最適分割を提案した。さらに、主観的な画質の観点から良好な再生画像を得るために、視知覚特性を考慮した空間-周波数領域の最適分割を提案した。

第2章においては、本研究の基礎となるレート歪み理論に基づいた最適帯域分割の導出を述べた。まず、レート歪み理論における最適化問題が、ラグランジュの未定乗数法を適用することにより解が求められることを踏まえて、帯域分割と量子化の二つの最適化問題をそれぞれ定式化することにより、レート歪み理論の観点で、二次元周波数帯域の最適分割パターンと各サブバンドに適用される最適な量子化器の組み合わせを導出するための理論式を明らかにした。次に、符号化レートが指定されたときに、導出された二つの理論式を所望の符号化レートに到達するまで反復して解くことで、符号化レートに応じた最適な帯域分割パターン及び各サブバンドに対する最適な量子化器の組み合わせを求めるアルゴリズムを開発した。最後に、提案したレート歪み理論に基づいた最適帯域分割を実画像に適用した結果、従来の最適帯域分割と比較して、PSNR 値において最大で約 1.0[*dB*] 程度の符号化性能の改善が得られることが示された。また、離散ウェーブレット変換及び帯域分割の理論的な最適解を導出できる Wavelet Packet との性能比較においては、本最適帯域分割法は少ないサブバンド数であるにも関わらず、特に低符号化レートにおいて、離散ウェーブレット変換よりも符号化性能を大きく改善でき、Wavelet Packet とほぼ同等の符号化性能を得られることが示された。

第3章においては、第2章で提案したのもも含めた従来の最適帯域分割では、各帯域ブロック信号における空間領域上の冗長性を十分に除去しきれていないことに着目し、帯域分割と

量子化に加えて、帯域ブロック内の不要な信号を切り捨てて必要な信号のみを保存することで冗長削減を実現する空間領域分割の三つを、レート歪み理論の観点で最適化する空間-周波数領域の最適分割の導出について述べた。このとき、三つの同時最適化は困難であることから、まず、ある一つの最適化問題を解く際には残り二つの最適化問題は既に解かれていると仮定した上で、各最適化問題を個別に定式化することにより、新しく空間領域分割が導入されたことを考慮した帯域分割及び量子化の最適解を導出するための理論式と、空間領域分割の最適解を導出するためのアルゴリズムを明らかにした。その後、これら三つの最適化問題を、指定された符号化レートに到達するまで反復して解くことによって、所望の符号化レートに応じた二次元周波数帯域の最適分割パターン、各サブバンドに適用する最適な量子化器の組み合わせ、各帯域ブロックにおける二次元空間領域上の最適分割パターンを求めるアルゴリズムを開発した。最後に、提案した空間-周波数領域の最適分割を実画像に適用し、その符号化性能を求めた結果、第2章で述べたレート歪み理論に基づいた最適帯域分割から、僅か1.25倍程度の処理コストによって、PSNR値において最大で約2.0[*dB*]程度の符号化性能の改善が得られることが示された。さらに、Wavelet Packet、提案手法と同じく空間領域上の冗長削減を実現するSpace-Frequency Quantization (SFQ)、画像符号化の国際標準方式JPEG及びJPEG2000よりも高い符号化性能が得られることが示された。

第4章においては、第3章で提案した空間-周波数領域の最適分割よりも主観的に良好な再生画像を得ることを目的として、視知覚特性上で重要ではない信号の品質を犠牲にする代わりに重要な信号の品質を優先的に保つことで、主観的な画質の改善を達成する視知覚特性を考慮した空間-周波数領域の最適分割について述べた。このとき、視知覚特性として、視覚の空間周波数特性と画像の顕著性に着目し、まずはそれらを考慮するために、画像信号に対する視覚の二次元空間周波数特性と、入力画像の顕著性を推定するSaliency Mapに基づいて、各帯域ブロック信号の歪み値に対する重み付け関数を定義した。次に、特に画像の顕著性を反映するため、Saliency Mapの結果に基づいて各サブバンド信号を空間領域上で注目領域(ROI)と非注目領域(non-ROI)に分割した後、重み付けられた歪みの量を最小化するように、二次元周波数帯域の最適分割パターン、各サブバンドにおけるROIとnon-ROIそれぞれに適用する最適な量子化器の組み合わせ、各帯域ブロックにおける二次元空間領域上の最適分割パターンを求める手法を提案した。最後に、視知覚特性を考慮した空間-周波数領域の最適分割により得られた再生画像の画質を評価するために、画質の主観評価実験を行った結果、5段階評価に対応するMOS値において、視知覚特性を考慮しない通常の空間-周波数領域の最適分割と比較して最大で約0.8程度、JPEG2000と比較して最大で約2.2程度の改善があることが示された。

本研究は、従来研究にあたる「画像のサブバンド符号化における最適帯域分割」の符号化性能を改善することを目的として、符号化性能に直結するにも関わらず未だ議論が行われていなかったレート歪み理論の観点から、帯域分割及び量子化、さらには空間領域上の冗長性を除去するための空間領域分割の最適化問題に取り組んだものである。その中で、最適解の

理論的な導出，及びそれを実現するためのアルゴリズムを構築することにより，画像符号化へ適用した場合の有効性を明らかにしたことは，大きな成果であると考えている．以下に示すように，まだいくつかの研究課題は残されているが，本研究によって得られた重要な成果は，今後の画像符号化技術の発展に大いに寄与するものであると信じている．

## 5.2 今後に残された課題

今後の課題としては、以下に示すものが挙げられる。

- 本研究では、レート歪み理論に基づいた最適帯域分割及び空間-周波数領域の最適分割を実現する際に必要なパラメータであるサブバンド数  $M$  と帯域ブロック数  $N$  には、先行研究 [61] での検討結果に基づき、符号化性能の改善量と処理コストの両方を考慮して、入力画像や符号化レートに関わらず汎用的にそれぞれ  $N = 64$  と  $M = 5$  を設定している。しかしながら、厳密には画像の内容に依存して、それらの最適なパラメータ値は入力画像及び符号化レート毎に異なっていることから、より高能率な符号化システムを実現するためには、帯域分割、量子化、及び空間領域分割だけではなく、サブバンド数  $M$  と帯域ブロック数  $N$  についても、入力画像の特性に応じて最適化を行う必要があると考えられる。このとき、パラメータ値の候補をいくつか与え、それぞれの条件で求められた符号化性能を比較して最適なパラメータを決定するという原始的な方法で対応することは、処理コストの観点からは望ましいことではない。ここで、入力画像が持つ様々な特性の中で、最終的に算出される符号化性能及びその改善量に直接的に影響を与えるのは、信号を量子化した際に算出される符号量と歪みの量、すなわち、レートと歪みの関係である。これらを踏まえて、システムに画像が入力された時点で、そのレートと歪みの関係、あるいは、それと密接に関わる信号の確率密度関数に基づいて、入力画像毎に最適なパラメータ値を事前に推定できれば、処理コストの増加を最小限に抑えつつ、より高い符号化性能を達成できると考えられる。
- 本研究では、所望の符号化レートに対する最適な帯域分割、量子化、及び空間領域分割を決定するために、レート歪み最適化処理を導入している。しかしながら、レート歪み最適化は、各帯域ブロック信号に対して、量子化器の候補としてあらかじめ用意されている量子化器を全て適用し、それにより算出されるレートや歪みの値を事前に用意しておく必要があるため、符号化性能を大幅に改善できる反面、処理コストが膨大になるという問題がある。そこで、レートと歪みの値を、演算量が少ない他の何らかの特徴量により推定できれば、符号化性能の低下をできる限り抑えながらレート歪み最適化処理に要する処理コストを大幅に削減できると考えており、実用化までを視野に入れるならば、それを実現するためのアルゴリズムの提案が望まれる。
- 第4章において、入力画像の対象を、初めてグレースケール画像からカラー画像へと拡張したが、現状の符号化システムでは、YCrCb の各色成分に対して共通の処理を行っているため、Cr 及び Cb 成分は Y 成分よりも知覚に鈍感である、といったような色情報に対する人間の視知覚特性は考慮されていない。そのため、色情報に対する視知覚特性をも同時に考慮した空間-周波数領域の最適分割の検討が望まれる。

# 謝辞

本研究を遂行するにあたり、長きに渡って懇切なる御指導と御鞭撻を賜りました岩手県立大学 教授 亀田昌志 博士に深甚の謝意を表します。亀田先生におかれましては、研究の進め方から文章の書き方まで、丁寧にご指導していただき、心より感謝しております。亀田先生からご教授頂いた専門知識や研究に対する考え方は、今後の人生において大きな糧となるものです。

そして、本論文の審査にあたり、有意義な御指摘と広い視点からのご助言を賜りました岩手県立大学 准教授 プリマ オキ ディッキ アルディアンシャー 博士、講師 小嶋和徳 博士に深く感謝申し上げます。また、研究活動を進めるにあたり、貴重なご意見を下さいました岩手県立大学 講師 松田浩一 博士、講師 塚田義典 博士に心からお礼を申し上げます。

日頃より多大な議論を頂きました岩手県立大学 亀田研究室の皆様感謝いたします。特に鶴崎裕貴さんにおきましては、同じ博士後期課程の学生、また、友人として共に切磋琢磨し、同じ悩みを共有できたことは、私の学生生活において大きな原動力となるものでした。感謝申し上げます。

末筆ながら、私の研究生活を暖かく見守ってくれた父と母、祖父母、兄弟、並びに全ての友人達に心より感謝いたします。



## 参考文献

- [1] J.B. O'Neil, "Predictive quantizing systems (differential pulse code modulation) for the transmission of television signals," *Bell Syst. Tech. J.*, vol.45, pp.689-721, (1966).
- [2] P.R. Abbott, "A differential pulse-code-modulation codec for videotelephony using four bits per sample," *IEEE Trans. Commun.*, vol. COM-19, pp.906-912, (1971).
- [3] P.A. Maragos, R.W. Shafer and R.M. Mersereau, "Two-dimensional linear prediction and its application to adaptive predictive coding of images," *IEEE Trans. Acoustics, Speech, and Signal Processing*, vol. ASSP-32, pp.1213-1229, (1984).
- [4] Y. Linde, A. Buzo and R.M. Gray, "An algorithm for vector quantizer design," *IEEE Trans. Commun.*, vol. COM-28, pp.84-95, (1980).
- [5] R.M. Gray, "Vector quantization," *IEEE ASSP Mag.*, vol.1, pp.4-28, (1984).
- [6] N.M. Nasrabadi, R.A. King, "Image coding using vector quantization: A review." *IEEE Trans. Commun.*, vol. COM-36, pp.957-971,(1988).
- [7] P.A. Wintz, "Transform picture coding," *Proc. IEEE*, vol.60, pp.809-820, (1972).
- [8] R.J. Clarke, "Transform coding of images," Academic Press, (1985).
- [9] H.S. Malvar, D.H. Staelin, "The LOT: Transform coding without blocking effects," *IEEE Trans. Acoustics, Speech, and Signal Processing*, vol. ASSP-37, pp.553-559, (1989).
- [10] J.W. Woods and S.D. O'Neil, "Subband coding of images," *IEEE Trans. on Acoustics Speech and Signal Processing*, vol.34, no.5, pp.1278-1288, (1986).
- [11] K. Irie and R. Kishimoto, "A study on perfect reconstructive subband coding," *IEEE Trans. Circuits Syst. Video Technol.*, vol.1, no.1, pp.42-48, (1991).
- [12] T.A. Ramstad, S.O. Aase, and J.H. Husoy, "Subband compression of images: Principles and examples," New York: Elsevier, (1995).

- [13] 貴家仁志, よくわかるデジタル画像処理, pp.119-161, CQ 出版社, (1996).
- [14] G.K. Wallace, "The JPEG still picture compression standard," *Communication of the ACM*, vol.34, pp.30-44, (1991).
- [15] W.B. Pennebaker, J.L. Mitchell, "JPEG: Still image data compression standard," Springer Science & Business Media, (1992).
- [16] D.L. Gall, "MPEG: A video compression standard for multimedia applications," *Communication of the ACM*, vol.34, pp-30-44, (1991).
- [17] I.E. Richardson, "H. 264 and MPEG-4 video compression: video coding for next-generation multimedia," John Wiley & Sons, (2004).
- [18] N. Ahmed, T. Natarajan and K.R. Rao, "Discrete cosine transform," *IEEE trans. on Computers*, vol.100, no.1, pp90-93, (1974).
- [19] K.R. Rao, P. Yip, 共著 / 安田 浩, 藤原 洋 共訳, "画像符号化技術 - DCT とその国際標準 -,;" オーム社, (1992).
- [20] R. Dony, "Karhunen-loeve transform," *The transform and data compression handbook*, vol.1, pp.1-34, (2001).
- [21] W.H. Chen, C.H. Smith and S.C. Fralick, "A fast computational algorithm for the discrete cosine transform," *IEEE Trans. on communications*, vol.25, no.9, pp.1004-1009, (1977).
- [22] T. Wiegand, G.J. Sullivan, "Overview of the H. 264/AVC video coding standard," *IEEE Trans. on circuits and systems for video technology*, vol.13, no.7, pp.560-576, (2003).
- [23] H. Schwarz, D. Marpe and T. Wiegand, "Overview of the scalable video coding extension of the H. 264/AVC standard," *IEEE Trans. on circuits and systems for video technology*, vol.17, no.9, pp.1103-1120, (2007).
- [24] G.J. Sullivan, J. Ohm, et al. "Overview of the high efficiency video coding (HEVC) standard," *IEEE Trans. on circuits and systems for video technology*, vol.22, no.12, pp.1649-1668, (2012).
- [25] M.T. Pourazad, C. Doutre, et al. "HEVC: The new gold standard for video compression: How does HEVC compare with H. 264/AVC?," *IEEE consumer electronics magazine*, vol.1, no.3, pp.36-46, (2012).
- [26] D.J. Vaisey, A. Gersho, "Variable block-size image coding," *Acoustics, Speech, and Signal Processing, IEEE International Conference on ICASSP'87*, vol.12, pp.1051-1054, (1987).

- [27] D.J. Vaisey, A. Gersho, "Image compression with variable block size segmentation," IEEE Trans. on Signal Processing, vol.40, no.8, pp.2040-2060, (1992).
- [28] ゴードン・ムーア, "インテルとともに—ゴードン・ムーア 私の半導体人生," 玉置直司 取材・構成, 日本経済新聞社, (1995).
- [29] R.R. Schaller, "Moore's law: past, present and future," IEEE spectrum, vol.34, no.6, pp.52-59, (1997).
- [30] A. Skodras, C. Christopoulos, and T. Ebrahimi, "The JPEG 2000 still image compression standard," IEEE Signal Processing Magazine, vol.18, no.5, pp.36-58, (2001).
- [31] M. Rabbani and R. Joshi, "An overview of the JPEG2000 still image compression standard," Signal Process., Image Comm., vol.17, no.1, pp.3-48, (2002).
- [32] 小野 定康, 鈴木 純司, "わかりやすい JPEG2000 の技術," オーム社, (2003).
- [33] Y. Luo, R.K. Ward, "Removing the blocking artifacts of block-based DCT compressed images," IEEE trans. on Image Processing, vol.12, no.7, pp.838-842, (2003).
- [34] N.S. Jayant and P. Noll, "Digital coding of waveforms," Prentice Hall, (1984).
- [35] 甲藤二郎, 安田靖彦, "サブバンド符号化の特性評価とそのフィルタ係数の最適化について," 電子情報通信学会技術研究報告, IE91-10, pp.33-40, (1991).
- [36] P.H. Westerink, J. Biemond and D.E. Boekee, "An optimal bit allocation algorithm for sub-band coding," Acoustics, Speech, and Signal Processing, ICASSP'88, pp.757-760, (1988).
- [37] T. Berger, "Rate- Distortion Theory," Encyclopedia of Telecommunications, (1971).
- [38] G.J. Sullivan, T. Wiegand, "Rate-distortion optimization for video compression," IEEE signal processing magazine, vol.15, no.6, pp.74-90, (1998).
- [39] P.A. Chou, L. Tom and R.M. Gray, "Optimal pruning with applications to tree-structured source coding and modeling," IEEE trans. on information theory, vol.35, no.2, pp.299-315, (1989).
- [40] E.A. Riskin, "Optimal bit allocation via the generalized BFOS algorithm," IEEE Trans. on information Theory, vol.37, no.2, pp.400-402, (1991).
- [41] A.S. Lewis and G. Knowles, "Image compression using the 2-D wavelet transform," IEEE Trans. on Image Processing, vol.1, no.2, pp.244-250, (1992).

- [42] M. Antonini, M. Barlaud, P. Mathieu, and I. Daubechies, "Image coding using wavelet transform," *IEEE Trans. on Image Processing*, vol.1, no.2, pp.205-220, (1992).
- [43] G. Strang and T.Q. Nguyen, "Wavelets and filter banks," Wellesley-Cambridge Press, (1996).
- [44] R.R. Coifman and M.V. Wickerhauser, "Entropy-based algorithms for best basis selection," *IEEE Trans. on Information Theory*, vol.38, no.2, pp.713-718, (1992).
- [45] K. Ramchandran and M. Vetterli, "Best wavelet packet basis in a rate-distortion sense," *IEEE Trans. on Image Processing*, vol.2, no.2, pp.160-175, (1993).
- [46] I.H. Witten, R.M. Neal and J.G. Cleary, "Arithmetic coding for data compression," *Communications of the ACM*, vol.30, no.6, pp.520-540, (1987).
- [47] K. Ramchandran, Z. Xiong, K. Asai, and M. Vetterli, "Adaptive transforms for image coding using spatially varying wavelet packets," *IEEE Trans. on Image Processing*, vol.5, no.7, 1197-1204, (1996).
- [48] 大竹敢, 福田光, 川中彰, "ゼロツリーにより画像の自己相似性を考慮したウェーブレットパケット符号化方式," *テレビ誌*, vol.50, no.11, pp.1798-1804, (1996).
- [49] N.M. Rajpoot, R.G. Wilson, F.G. Meyer, and R.R. Coifman, "Adaptive wavelet packet basis selection for zerotree image coding," *IEEE Trans. on Image Processing*, vol.12, no.12, 1460-1472, (2003).
- [50] N. Sprljan, S. Grgic, and M. Grgic, "Modified SPIHT algorithm for wavelet packet image coding," *Real-Time Imaging*, vol.11, no.5, pp.378-388, (2005).
- [51] T. Stütz, B. Mühlbacher, and A. Uhl, "Efficient and rate-distortion optimal wavelet packet basis selection in JPEG2000," *IEEE Trans. on Multimedia*, vol.14, no.2, pp.264-277, (2012).
- [52] K.S. Kumar, A. Pravin, T.S.S. Phani, and G.S. Babu, "Performance analysis of wavelet packet based SPHIT algorithm," *IJSEAT*, vol.2, no.2, pp.101-106, (2014).
- [53] V. Gopinath, N. Kanvel, and T.K. Shanthi, "A novel image compression scheme with wavelet packets optimization using artificial bee colony algorithm," *IJETT*, vol.4 no.8, pp.3448-3454, (2014).
- [54] F. Wenchao and X. Wang, "An image compression method based on orthogonal wavelet packet transform," *Boletín Técnico*, vol.55, no.1, pp.18-25, (2017).
- [55] F.G. Meyer, A.Z. Averbuch, J.O. Strömberg, "Fast adaptive wavelet packet image compression," *IEEE Trans. on Image Processing*, vol.9, no.5, pp.792-800, (2000).

- [56] G.K. Kharate, V.H. Patil, “Color image compression based on wavelet packet best tree,” *IJCSI International Journal of Computer Science Issues*, vol.2, no.3, pp.31-35, (2010).
- [57] J. Sethi, S. Mishra, P.P. Dash, S.K. Mishra, S. Meher, “Image compression using wavelet packet tree,” *ACEEE International Journal on Signal & Image Processing*, vol.2, no.1, pp.41-43, (2011).
- [58] K.S. Kumar, A. Pravin, T.S.S. Phani, G.S. Babu, “Performance analysis of wavelet packet based SPHIT algorithm,” *IJSEAT*, vol.2, no.2, pp.101-106, (2014).
- [59] 亀田昌志, 大竹孝平, 宮原誠, “画像信号の2次元サブバンド符号化における最適周波数分割,” *電子情報通信学会論文誌 A*, vol.J80-A, no.1, pp.237-247, (1997).
- [60] 亀田昌志, 大竹孝平, 宮原誠, “画像のサブバンド符号化における帯域ブロックをベースにした2次元最適周波数分割,” *電子情報通信学会論文誌 D*, vol.J81-D-II, no.6, pp.1095-1107, (1998).
- [61] 亀田昌志, 大竹孝平, “画像のサブバンド符号化における空間-周波数領域の最適分割,” *電子情報通信学会論文誌 A*, vol.J90-A, no.2, pp.142-156, (2007).
- [62] P.P. Vaidyanathan, “Quadrature mirror filter banks, M-band extensions and perfect-reconstruction techniques,” *IEEE ASSP Magazine*, vol.4, no.3, pp.4-20, (1987).
- [63] M. One, “SIDBA Standard Image Database,” *MIPC Report 79-1*, (1979).
- [64] J.R. Hershey and P.A. Olsen, “Approximating the Kullback Leibler divergence between Gaussian mixture models,” In *ICASSP*, vol.4, pp.IV-317-IV-320, (2007).
- [65] Y. Shoham, A. Gersho, “Efficient bit allocation for an arbitrary set of quantizers (speech coding),” *IEEE Trans. on Acoustics, Speech, and Signal Processing*, vol.36, no.9, pp.1445-1453, (1988).
- [66] D. Arthur and S. Vassilvitskii, “k-means++: The advantages of careful seeding,” *Proceedings of the eighteenth annual ACM-SIAM symposium on Discrete algorithms*, Society for Industrial and Applied Mathematics, pp.1027-1035, (2007).
- [67] Z. Xiong, K. Ramchandran, and M.T. Orchard, “Space-frequency quantization for wavelet image coding,” *IEEE Trans. on Image Processing*, vol.6, pp.677-693, (1997).
- [68] “CIPR still images,” <http://www.cipr.rpi.edu/resource/stills/>, (March 2018 access).
- [69] AIC AhG, “JPEG XR subjective assessment: Core Experiments Description 4.1,” *Tech. Rep. WG1N5001*, ISO/IEC JTC1/SC29/WG1 (JPEG), (2009).

- [70] H. Drolon, F.O. Devaux, A. Descampe, Y. Verschuere, D. Janssens, and B. Macq, “Open-JPEG library: an open source JPEG 2000 codec,” Image and Signal Processing Group (ISPGGroup), <http://www.openjpeg.org>, (March 2018 access).
- [71] B. Chitprasert, K.R.Rao, “Human visual weighted progressive image transmission,” IEEE Trans. Commun., vol. COM-38, pp.1040-1044, (1990).
- [72] L. Itti, C. Koch, E. Niebur, “A model of saliency-based visual attention for rapid scene analysis,” IEEE Trans. on pattern analysis and machine intelligence, vol.20, no.11, pp.1254-1259, (1998).
- [73] M. Miyahara, H. Otani, H. Horita and T. Fujimoto, “Objective evaluation scale for image qualities (PQS); Local features and general-purpose applications,” IEICE Trans., vol.J73-B-I, pp.208-218, (1990).
- [74] 堀田, 宮原, “画像 OTC 量子化器の最適化 - 統計量, 視覚特性の厳密な考慮 -,” 信学技報, IE85-92, pp.13-18, (1985).
- [75] 樋渡, 渡部, 森, 長田, “視覚の空間正弦波レスポンス,” NHK 技術研究 1, pp.38-60, (1964).
- [76] R. Achanta, F.J. Estrada, et al. “Salient region detection and segmentation,” International conference on computer vision systems, pp.66-75, (2008).
- [77] M.M. Cheng, N.J. Mitra, et al. “Global contrast based salient region detection,” IEEE Trans. on Pattern Analysis and Machine Intelligence, vol.37, no.3, pp.569-582, (2015).
- [78] Q. Yan, L. Xu, J. Shi and J. Jia, “Hierarchical Saliency Detection,” IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp.1155-1162, (2013).

# 本研究に関する研究業績

## 1. 発表論文

- [1] 宮崎春彦, 亀田昌志, “レート-歪理論に基づいた画像のサブバンド符号化における最適帯域分割,” 電子情報通信学会論文誌 A, Vol. J100-A No.5, pp.205-217, (2017).

## 2. 国際会議

- [1] Haruhiko Miyazaki, Masashi Kameda, “The Optimum Frequency Band Partition based on Kullback-Leibler Divergence in Subband Image Coding,” International Conference on Intelligent Systems and Image Processing 2014 (ICISIP 2014), pp.336-340, (2014).
- [2] Haruhiko Miyazaki, Masashi Kameda, “The Optimum Space-Frequency Partition for Subband Image Coding in a Rate-Distortion Sense,” Proc. of Picture Coding Symposium 2016 (PCS 2016), P-3-6, (2016).
- [3] Haruhiko Miyazaki, Masashi Kameda, “Optimum Space-Frequency Partition in Subband Image Coding with Human Visual Sensitivity and Region-of-Interest,” Proc. of Human Vision and Electronic Imaging 2018 (HVEI 2018), HVEI-521, (2018).

## 3. 研究会

- [1] 宮崎春彦, 亀田昌志, “サブバンド符号化の分割パターンと量子化器選択の最適化における性能改善,” 電子情報通信学会, 画像工学研究会, IE2012-137, pp.333-338, (2013).
- [2] 宮崎春彦, 亀田昌志, “信号の電力と尖度をパラメータとする画像のサブバンド符号化における最適帯域分割,” 情報処理学会第 75 回全国大会, (2013).
- [3] 宮崎春彦, 亀田昌志, “一般化ガウス分布モデルを用いた最適帯域分割における符号化性能改善のための検証,” 電子情報通信学会, 画像工学研究会, IE2013-27, pp.7-12, (2013).
- [4] 宮崎春彦, 亀田昌志, “画像のサブバンド符号化における分布形状の類似性に基づく最適帯域分割,” 画像符号化シンポジウム 2013 (PCSJ 2013), P-2-02, (2013).

- [5] 宮崎春彦, 亀田昌志, “符号化レートに対応した信号分布の類似性に基づく画像の最適帯域分割,” 電子情報通信学会, 画像工学研究会, IE2013-123, pp.165-170, (2014).
- [6] 宮崎春彦, 亀田昌志, “確率分布の類似性に基づいた画像のサブバンド符号化における最適な帯域分割,” 映像情報メディア学会, AIT2014-110, pp.33-36, (2014).
- [7] 宮崎春彦, 亀田昌志, “確率分布の類似性に基づいた符号化レートを可変とする画像の最適帯域分割,” 画像符号化シンポジウム 2014 (PCSJ 2014), P-2-14, (2014).
- [8] 宮崎春彦, 亀田昌志, “画像のサブバンド符号化におけるカラー静止画像に対する最適帯域分割,” 画像符号化シンポジウム 2015 (PCSJ 2015), P-3-06, (2015).
- [9] 宮崎春彦, 亀田昌志, “確率分布のクラスタリングに基づく Wavelet Packet 画像符号化の演算量削減法,” 電子情報通信学会, 画像工学研究会, IE2015-113, pp.107-112, (2016).
- [10] 宮崎春彦, 亀田昌志, “画像のサブバンド符号化における視知覚特性を考慮した空間-周波数領域の最適分割,” 画像符号化シンポジウム 2016 (PCSJ 2016), P-3-06, (2016).