

Title	画像のサブバンド符号化における帯域分割の最適化に関する研究
Author(s)	亀田, 昌志
Citation	
Issue Date	1998-03
Type	Thesis or Dissertation
Text version	author
URL	http://hdl.handle.net/10119/853
Rights	
Description	Supervisor:宮原 誠, 情報科学研究科, 博士

博士論文

画像のサブバンド符号化における帯域分割の
最適化に関する研究

指導教官 宮原 誠 教授

北陸先端科学技術大学院大学
情報科学研究科情報処理学専攻

亀田 昌志

1998年 1月16日

目次

1 序論	1
1.1 画像符号化の背景	1
1.2 本研究の目的	4
1.3 本論文の構成	6
2 1次元サブバンド符号化における最適帯域分割	10
2.1 序言	10
2.2 1次元サブバンド分割	11
2.3 最適帯域分割の導出	14
2.4 1次マルコフ信号に対する最適帯域分割	17
2.5 評価方法とその結果	20
2.5.1 帯域分割による改善量 G	20
2.5.2 改善量 G の理論限界	20
2.5.3 評価結果	22
2.6 結言	24
3 画像信号の2次元サブバンド符号化における最適帯域分割	25
3.1 序言	25
3.2 2次元サブバンド分割と分割出力レート	27
3.3 2次元最適帯域分割	29
3.3.1 量子化雑音電力	29
3.3.2 量子化ビット配分	30
3.3.3 最適帯域分割と最適ビット配分	31
3.3.4 最適帯域分割算出アルゴリズム	34
3.4 水平・垂直相関分離型画像モデルに対する最適帯域分割	36
3.5 水平・垂直相関分離型画像モデルにおける評価とその結果	39
3.5.1 帯域分割による改善量 G	39
3.5.2 DCT による改善量 G	41

3.5.3	改善量 G の理論限界	41
3.5.4	評価結果	42
3.6	水平・垂直相関非分離型画像モデルに対する最適帯域分割	46
3.6.1	実画像の自己相関特性	46
3.6.2	最適帯域分割の導出	53
3.7	水平・垂直相関非分離型画像モデルにおける評価とその結果	56
3.7.1	帯域分割による改善量 G	56
3.7.2	DCT による改善量 G	56
3.7.3	改善量 G の理論限界	56
3.7.4	評価結果	58
3.8	結言	60
	付録 A：式(3.17)，式(3.18)の導出について	62
	付録 B：式(3.38)の導出について	63
4	画像信号のサブバンド符号化における最適帯域分割の近似と実現	65
4.1	序言	65
4.2	帯域ブロックをベースとする最適帯域分割の近似	67
4.2.1	帯域ブロックをベースとする最適帯域分割法	67
4.2.2	$J_{opt}^{(M)}$ の算出法	70
4.2.3	G 特性と帯域分割パラメータの選定	72
4.3	帯域ブロックをベースとする最適帯域分割の実現	74
4.3.1	フィルタバンクの構成	74
4.3.2	非定常性に対する対策	76
4.3.2.1	直流成分の非定常性	76
4.3.2.2	相関係数の非定常性	78
4.3.3	符号化性能	79
4.4	適応分割	83
4.4.1	帯域ブロックをベースとする適応分割方式	83
4.4.2	帯域分割数の増加による対応	87
4.4.3	帯域ブロック数の増加による対応	90
4.5	結言	93
5	視知覚特性を考慮した画像信号の 2次元最適帯域分割	94
5.1	序言	94
5.2	視覚の空間周波数特性	96
5.3	視覚の空間周波数特性を考慮した最適帯域分割	97

5.3.1	帯域ブロックをベースとする視覚の空間周波数特性を 考慮した最適帯域分割	97
5.3.2	視覚の空間周波数特性を考慮した最適帯域分割 フィルタバンクの構成	99
5.4	帯域ブロックをベースとする視覚の空間周波数特性を考慮した 画像の最適帯域分割	101
5.4.1	帯域分割パラメータの選定	101
5.4.2	視覚の空間周波数特性を考慮した画像の 最適帯域分割パターン	108
5.5	評価方法とその結果	112
5.5.1	WMSEによるランダムノイズの定量的評価	112
5.5.2	主観評価実験による画質評価	116
5.6	結言	128
	付録C：式(5.12)，式(5.13)の導出について	129
6	結論	132
6.1	本研究で得られた成果	132
6.2	今後に残された課題	137
	謝辞	139
	参考文献	140
	本研究に関する研究業績	145

第 1 章

序論

1.1 画像符号化の背景

高度情報化社会（マルチメディア社会）へ向かっての発展は必至であるが，その中で，インターネットの爆発的な普及に伴う画像の伝送 / 交換サービスの充実，コンパクトディスク，デジタルビデオディスク等の画像記録 / 蓄積メディアの登場とその普及，画像データベース，ナビゲーションシステムの開発等，画像を用いたメディアは不可欠な存在である．また，画像信号のデジタルデータ化も，マルチメディア展開において不可欠な技術である．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である．近年では，画像符号化における国際標準化の作業も行われ[1],[2]，その方式は，上記の各種メディアにも応用されている．

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）[3]～[6]は，波形符号化の中では最も歴史の古いもので，画像の空間領域において波形レベルの直接的な冗長性の除去を行う，極めて直観的アイデアに基づく符号化方式である．ベクトル量子化（Vector Quantization）[7]～[11]もまた，空間領域における波形レベルの直接的な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 個の画素の情報をまとめる（ N 次元ベクトルの構成）ことで，より高能率な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）[12]～[15]やサブバンド符号化（Subband Coding）[16]～[20]では，画像信号を一旦空間領域とは異なる領域に変換した後，その領域において冗長性の除去を行う．特に，サブバンド符号化においては，変換後の領域が，周波数領域であることから，視覚の空間周波数特性を考慮することが容易であるという利点がある．近年，注目を集めた wavelet 変換[21]～[25]は，サブバンド符号化の特殊な場合であると見なすことができ，本変換は，周波数軸

上で帯域をオクターブ分割することで，多重分解能表現を実現することができる．これらの方式はいずれも「画像に存在する冗長性をいかに除去することができるか」という観点において，画像の持つ構造的な性質を統計的信号モデルという枠組みのもとで利用していると言える．

これに対して，上記の情報理論に基づく符号化方式から一世代進んで，画像内容と密接に関係する情報を利用する画像符号化技術も研究，開発され始められている．このような立場からの画像符号化方式の代表例としては，画像を均質な輝度成分を有する幾つかの領域に分割し，領域形状と共に各領域内の輝度成分を符号化する領域分割符号化[26]～[28]，画像の自己相似性を利用したfractal符号化[29],[30]，送信側と受信側とで画像の内容に関する共通のデータベースを持つという仮定のもとで，送信側では，データベースを辞書として用いることにより画像の内容に関するパラメータを抽出及び伝送し，受信側では，送られてきたパラメータを用いてデータベース内のデータを修正することで画像を再生するモデルベース符号化[31],[32]等がある．しかし，これらの方式は，いずれも画像自身の内容に依存する部分が大きく，領域分割符号化においては，領域の分割/統合の基準が曖昧であること，fractal符号化においては，自己相似性を認め難い画像の場合には，その符号化レートが極端に大きくなってしまふこと，モデルベース符号化においては，統一的なモデルを構成することが困難なため，未だその対象が顔画像に限定されないと実用的でないこと等，汎用性を含めた様々な点で今後の解決が望まれる課題が多く残されており，符号化方式として完成されているとは言い難い．ただ，一言付け加えるならば，これらの方式に見られる「画像個々の内容において何が重要であり，そしてそれをどのように保存するか」という考え方は，今後の画像符号化技術の発展において，特に重要なものになるであろうと考えている．

新世代画像符号化が実現困難であるが故に，波形符号化は，その歴史が長い技術であるにもかかわらず，現在でも画像符号化技術の中核を担っており，画像符号化の国際標準方式である静止画像用のJPEG[33],[34]，動画像用のMPEG[35],[36]において，その技術の基礎となっているのは，波形符号化の一方式であるDCT(Discrete Cosine Transform)[37]に他ならない．国際標準にDCTが採用された理由は，画像のように隣接標本値間の相関係数値が1.0に近い信号系列を入力信号に仮定した場合，最適な圧縮効率を有するKarhunen-Loève変換と比べて，DCTはほぼ同等の圧縮効率を実現できることに加えて，DCTには，FFT(Fast Fourier Transform)と同様に高速アルゴリズムが存在する[38]ことによるものである．しかし，DCTをベースとした符号化では，通常，画像を小領域のブロックに分割して処理を行うため，その再生画像にブロックひずみと呼ばれる格子状の画像ひずみが現れるという欠点があり，この符号化ひずみは，視知覚上最大の劣化要因となる[39],[40]．そこで，再生画像においてブロックひずみを発生しない利点故に注目をされているのが，DCTと遜色のない圧縮性能を有するサブバンド符号化である．サブバンド符号化は，画像全体を一括していくつかの周波数帯域に分割し，分割された帯域毎

に符号化を行うために，再生画像にブロックひずみを生じない．サブバンド符号化は，1986年にWoodsとO'Neil[16]によって初めて画像符号化に適用されて以来，今日までサブバンド符号化をベースとした多くの符号化方式が開発されている．これらの符号化方式は，サブバンド符号化のもう1つの利点である各帯域毎に独立の処理が適用できるということのみを応用したものがほとんどであり，帯域分割後の信号系列に，既存の画像符号化方式である予測符号化，DCT，ベクトル量子化等を適用する方式はハイブリッド符号化[41]～[43]と呼ばれている．

しかし，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，(1)適切な帯域分割と (2)帯域分割後の信号に対する適切なビット配分にある．このうち，後者のビット配分に対しては，分割された各帯域の信号電力に応じて最適なビット配分を行う方法が，1988年にWesterinkらによって明らかにされている[44]．残された重要な問題である前者の帯域分割に対しては，従来から今日に至るまで，帯域を周波数領域で均等に分割する方式が主たる帯域分割方式として用いられている．近年になって，1989年のMallat[45]，及び1992年のVetterliら[46]によって，wavelet変換をフィルタバンクによって実現する方法が明らかにされた後，wavelet変換は，周波数領域において帯域をオクターブ分割するので，そのことが，人間の視覚特性との整合の点においてよいと指摘され注目を集めたが，この議論はそこまでであり，その後この問題の解決はされていない．画像圧縮の目的を直接的に考慮した，入力信号の性質に応じた帯域分割の最適化については，その理論展開が困難であることから未だ議論が行われていない．そこで，本研究では，サブバンド符号化の本質とも言える帯域分割問題を解決することとした．画像信号の性質を直接的に考慮し，更に視覚の空間周波数特性までを考慮して最適帯域分割を明らかにすることを目的として，その解を与えた．これらの基礎理論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，更なる符号化効率の改善が行われることが期待される．

1.2 本研究の目的

現在の画像符号化の主流は、JPEG、MPEGにおいてその中心的役割を担っているDCTである。しかし、DCTに基づいた符号化方式では、視覚上最大の劣化要因となるブロックひずみ発生の問題が付きまとう。そこで、本研究は、ブロックひずみを生じないことから、DCTに代わる方式として有望視されているサブバンド符号化に注目する。

サブバンド符号化は、上記のブロックひずみ発生の問題がないことに加え、伝送誤りの影響が空間的、周波数的に制限されること、換言すれば、分割された各帯域毎に各々独立の処理が適用できることの利点から、画像の高効率符号化の一方式として、その検討が行われている。サブバンド符号化方式では、信号系列を帯域分割した後、各帯域に落ちる信号電力に対応した量子化ビット配分を行う。それ故、良い画質で高い圧縮効率を得るためには、適切な帯域分割及び適切な量子化ビット配分が重要となる。後者のビット配分については、分割後の各帯域の信号電力に応じて配分ビットを変えるビット配分の最適化がWesterinkらによって論じられている[44]。しかし、前者の帯域分割の最適化については、その理論展開が困難であることから未だ議論が行われていない。そこで、本研究では、画像符号化の原点である冗長性の除去の観点から、入力信号である画像信号の統計的性質に基づいて、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。

従来において、画像信号に対してサブバンド符号化を適用する際の帯域分割フィルタには、その扱い易さの理由のみから帯域を均等に分割する方式を用いることが一般的であった[16],[18]。均等分割以外の帯域分割方式としては、Mallat[45]、Vetterliら[46]によってそのフィルタバンクによる記述が明らかにされた後、wavelet変換に基づいた帯域のオクターブ分割が、視覚の空間周波数特性との整合がよいとして、画像符号化に適用されるようになった。また、近年では、画像符号化への適用を目的として、2次元の信号系列に直接適用することのできる2次元non-separableフィルタバンクを用いた帯域分割の研究も行われている[47],[48]。しかし、上述した入力信号の統計的性質に整合した最適帯域分割という観点からの検討は未だ行われていない。

本論文は、筆者の行っ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、理論展開の立場において、画像信号の統計的性質から導入された理論画像モデルに基づいて、原画像と符号化後の再生画像間のユークリッドノルムをベースに、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尺度を提案し、その最大値を与える最適帯域分割を明らかにする。次に、実画像への応用の立場において、導出された最適帯域分割の理論特性を実画像に適用するために、実画像に特有の非定常性と符号化システム実現の簡易性までを考慮した上で、最適帯域分割フィルタバンクの構成を明らかにする。また、その符号化特性をJPEGで用いられているDCTの場合と比較し、本最適帯域分割の有効

性を明らかにする．更に，量子化雑音電力にのみ注目して行われた議論を拡張し，視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，視覚の空間周波数特性を考慮した最適帯域分割を導出すると共に，それを実現するための新たな最適分割フィルタバンクの構成を明らかにする．本方式による再生画像に対して，上記画質の観点から評価を行い，DCT を適用した場合の再生画像と比較評価する．

本研究は，画像符号化においては過去の技術の感のある波形符号化の一方式であるサブバンド符号化を対象としているが，従来の議論において重要であるにもかかわらず着手されていなかった本質的な問題である「帯域分割の最適化」に理論的，实际的に解を与えることを目的としたものである．得られた成果を踏まえた上で，更なる実用面への応用を行うことで，高能率な画像符号化への更なる展開があるものと信じている．

1.3 本論文の構成

第 1 章：序論

画像符号化における現状を示し、本研究の目的と意義を述べる。

第 2 章：1 次元サブバンド符号化における最適帯域分割

最適帯域分割の導出において、その議論の理論的基礎となる1次元の入力信号系列に対する最適帯域分割法を明らかにする。ここでは、入力信号のグローバルな特性として定常性を仮定することで、定式化された入力信号の電力スペクトルに基づいて、最適帯域分割を導出するための理論式を明らかにする。最適帯域分割の評価を行うために、原信号と量子化（符号化）後の再構成信号との間のユークリッドノルムをベースに、非帯域分割方式からの量子化雑音電力の低減量を定量的に表し得る値として、量子化雑音改善量なる評価尺度を定義する。実際の信号例として、画像信号の統計的性質を考慮した隣接標本値間の自己相関係数値が0.9である1次マルコフ情報源系列を入力信号とした場合に、その最適帯域分割特性を求める。更に、最適帯域分割の性能を比較評価するために、入力信号である1次マルコフ情報源系列に対する量子化雑音改善量の理論限界値を明らかにした上で、既存の帯域分割方式との比較を行う。この結果、上記の1次元信号系列に対しては、wavelet変換によって実現される周波数軸上でのオクターブ分割による帯域分割が、量子化雑音改善量の観点において最適帯域分割とほぼ同等の性能を有していることが示された。このことは、1次元の画像信号系列に対しては、wavelet変換の適用が有効であることを意味している。

第 3 章：画像信号の 2 次元サブバンド符号化における最適帯域分割

現実の画像信号は、2次元の信号系列である。そこで、第2章で行われた議論を2次元へと拡張して、2次元サブバンド符号化における最適帯域分割を理論的に明らかにする。2次元周波数平面上において任意の形状の帯域分割が可能であることを仮定して、入力信号の統計的性質：自己相関関数に整合した2次元最適帯域分割、すなわち、量子化雑音改善量を最大とする帯域分割を得るための理論式の導出を行い、最適帯域分割算出アルゴリズムを明らかにする。まず、画像信号のモデルとして、2次元の自己相関関数が1次元系列の積として与えられる2次元分離型相関特性を有する水平・垂直相関分離型画像モデルを導入する。本画像モデルは、特に、空間領域において水平及び垂直方向の成分が多いような画像の場合にうまく適合するとされている。本画像モデルに基づいて、2次元周

波数平面上での最適帯域分割パターンを明らかにすると共に、量子化雑音改善量の観点から既存の2次元帯域分割方式との比較を行う。最終的に、より一般的な汎用画像モデルとして、自己相関関数が水平及び垂直方向の間で独立ではなく、等方性を有する空間的距離の関数で与えられる水平・垂直相関非分離型画像モデルを導入する。このとき、一般的な画像の代表として用いられている数枚の標準画像に対して、その自己相関特性を実験的に求めた結果、多くの画像の相関特性は、2次元空間上で水平及び垂直方向の距離に対しほぼ等方性の特性を有することが示された。実画像のモデルとして、ここに導入する水平・垂直相関非分離型画像モデルがより望ましいことを指摘している研究者はいるが、実際に計算を行い、このことを明らかにした。そこで、改めて本画像モデルを対象とした最適帯域分割の導出を行い、先に示した水平・垂直相関分離型画像モデルの場合と同様、その特性を既存の帯域分割方式と比較する。この結果、両画像モデルに対して、最適帯域分割時の量子化雑音改善量特性は、僅か帯域分割数=4なる条件で、帯域分割数=64に相当する(8×8)画素のブロック単位に適用されるDCTの性能を上回り、且つ、それは各画像モデルにおける量子化雑音改善量の理論限界値に漸近していることが明らかになった。この結果は、量子化雑音電力低減の観点から、画像信号のサブバンド符号化における理想的な帯域分割特性の解を理論的に明らかにしたという点において、大きな成果であると信じている。

第4章：画像信号のサブバンド符号化における最適帯域分割の近似と実現

第3章において、理論画像モデルに基づいて明らかにされた最適帯域分割の理論特性を実画像に適用することを目的として、最適帯域分割フィルタバンクの実現を行う。2次元周波数平面上において、帯域を任意の形状に分割する無損失で完全再構成可能な2次元non-separableフィルタの構成法は未だ確立されていないことから、理論解として得られた最適帯域分割特性を直接実現する2次元フィルタバンクの実用的な構成が極めて困難であるという問題がある。そこで、2次元最適帯域分割と比べて若干の量子化雑音改善量の低下を犠牲にすることとのトレードオフで、フィルタバンクの実現を可能にした帯域ブロック（周波数領域における基本ブロック）をベースにする2次元最適帯域分割法を提案する。このとき、符号化システム実現の簡易性を考慮することで、フィルタバンクの実用において効果的な量子化雑音の改善が得られる帯域分割パラメータの選定を行う。これらの各値に基づいて、画像モデルに水平・垂直相関非分離型画像モデルを用いることで導出された帯域ブロックをベースとする固定の最適帯域分割パターンを示すと共に、本最適帯域分割を実現するフィルタバンクの構成を明らかにする。更に、画像特有の性質である非定常性のうち、2次元空間領域の場所によって変動する直流成分と、画像内容の違いによって、或いは空間領域の場所によって変動する相関特性に着目し、これらに対

処するために，前者に対しては，最低域の帯域分割信号に対して小ブロック単位のDFT (Discrete Fourier Transform) を適用することで非定常直流成分を分離を行う方法を，後者に対しては，画像毎に固有な最適帯域分割を得るために，帯域ブロックをベースとした適応帯域分割法を提案する．汎用性の向上のために選ばれた画像数例について，固定及び適応分割方式による量子化雑音改善量特性の比較評価実験を行い，多くの画像については，固定の最適帯域分割により適応分割に代わり得る十分な改善が得られるが，水平・垂直相関非分離型画像モデルでのモデル化が特に困難な画像については，適応分割の適用が有効であることを示す．このとき，たとえ適応分割を行う場合においても，本処理に伴う演算コストと伝送すべき付加情報量は小さくなることを示す．また，帯域ブロックをベースとする最適帯域分割における符号化性能を調べるために， SNR とエントロピーの関係を求め，JPEGで用いられている (8×8) 画素のブロック単位で適用されるDCTとの比較を行った結果，本最適帯域分割は，非定常直流成分の分離を含めて最大でも僅か6分割(多くの画像については，5分割あれば十分である)という極めて少ない帯域分割数であるにもかかわらず， SNR 評価の観点で，64分割のDCTの符号化性能を優ることができるということが明らかになった．このことは，簡易的な符号化システムを構築するという目的において，画像信号に対して適切な帯域分割を行うことの重要性を示唆すると共に，本最適帯域分割方式が，実画像において有力な一手法であることを意味している．

第5章：視知覚特性を考慮した画像信号の 2次元最適帯域分割

第3章においては，実画像の統計的性質：自己相関特性に基づいて導入された画像モデルを対象に，最適帯域分割の理論的な導出について論じた．また，第4章においては，第3章で得られた最適帯域分割特性を実現するために，帯域ブロックをベースとする最適帯域分割法を提案し，そのフィルタバンクの構成を明らかにした上で，本フィルタバンクを実画像へ適用した場合の有効性を示した．しかし，前章までの議論において明らかにされた最適帯域分割及び帯域ブロックをベースとする最適帯域分割は，いずれも量子化雑音改善量なる SNR 値に基づいた評価値を最大とする条件で導出された帯域分割であることから，符号化後の再生画像に対して，視覚・知覚特性を考慮した画質という観点からは，真に最適な帯域分割が行われているとは言い難い．画像符号化の原点は，あくまでも視知覚的に良好な画質の再生画像を得ることにあり，その議論を避けて通ることはできない．そこで，前章までの議論を上記の画質という観点にまで拡張し，サブバンド符号化に人間の視知覚特性を導入するための第一段階として，視覚の空間周波数特性を考慮したサブバンド符号化方式を提案し，その議論において必要となる視覚の空間周波数特性を考慮した最適帯域分割の導出を行う．従来において，この視覚の空間周波数特性を考慮したいくつかの画像符号化方式が開発されているが，これらの各従来方式においては，視覚の空間周波数特性は量子化特性についてのみ考慮されており，本章で提案

するように帯域の分割パターンにまで踏み込んだ例は他にない。まず、本章において用いられる視覚の空間周波数特性について簡単に紹介した後、フィルタバンクの実現性を考慮して、帯域ブロックによる帯域分割をベースに、視覚の空間周波数特性によって重み付けられた量子化雑音改善量を最大にする最適帯域分割法を提案すると共に、本最適帯域分割を実現するDCTに基づいた新たなフィルタバンクの構成を明らかにする。次に、汎用性の向上を目的として、画像内容の異なる数枚の画像を対象に、本フィルタバンクを実用する上で効果的な評価改善量が見込まれる帯域分割パラメータの選定を行い、これらに基づき、各画像に対して視覚の空間周波数特性を考慮した2次元周波数平面上での最適分割パターンを明らかにする。最後に、視覚の空間周波数特性で重み付けられた量子化雑音電力($WMSE$)とエントロピーの関係を求め、本最適帯域分割の有効性をランダムノイズ低減の観点において定量的に評価する。視覚の空間周波数特性を考慮した本最適帯域分割による再生画像においては、その構成上、視覚的に最大の画質劣化要因となるブロックひずみを生じないが、ランダムノイズ($WMSE$)だけでは評価しきれない視覚的妨害を生じなくはない。そのような妨害として、例えば、誤差の間に強い自己相関性を有するテクスチャ状のパターンをもつ符号化ひずみは、ランダムノイズに比べて約10倍も知覚されやすく、視覚的に大きな妨害となる[39],[40]。符号化方式及び画像内容に依存することなく、これらの視覚的妨害の全てを含めたオーバーオール画質を評価するには主観評価法による以外になく、そこで、主観評価実験を行うことにより、本最適帯域分割の再生画像の画質を評価する。比較方式の1つに挙げるJPEGで用いられている(8×8)画素のブロック単位で適用されるDCTにおいては、再生画像に上記のブロックひずみを生じることから、その画質は大きく劣化する。この結果、視覚の空間周波数特性を考慮した本最適帯域分割による再生画像は、DCTのものと比べて、5段階評価に対応するMOS (Mean Opinion Score) 値において最大で1.0以上の改善があることが明らかになった。このことは、帯域ブロックをベースとする視覚の空間周波数特性を考慮した本最適帯域分割は、視覚・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力な一手法であることを意味している。

第6章：結論

本論文の総括として、第2章から第5章までの結果を要約すると共に、今後残された研究課題について述べる。

第 2 章

1 次元サブバンド符号化における 最適帯域分割

2. 1 序言

波形符号化の有力な一手法としてサブバンド符号化がある。本符号化法は、信号系列を周波数軸上で複数の帯域に分割し、各帯域に落ちる信号電力に対応した量子化ビット数を配分することにより能率の高い符号化を実現している。従来、この帯域分割方式としては均等分割（linear 分割）が主であった[16],[18]が、音声、画像等の現実の信号は非定常性を有するため、適応的な帯域分割、量子化ビット配分が有効となる。しかし、前者の適応帯域分割は実現が困難なため、入力信号系列のグローバルな特性に適合した帯域分割を行うことが重要と考えられる。

一方、人間の視聴覚特性に整合していると言われる対数スケール上で帯域を均等分割する wavelet 変換を用いたサブバンド分割法が近年注目を集めている[22],[23]。

以上の観点において、本研究では、画像圧縮の原点である冗長性の除去の観点から、画像のサブバンド符号化における最適帯域分割の導出を試みる。本章では、その議論の理論的基礎として、1次元の信号系列を対象とした場合に、入力信号のグローバルな性質として定常性を仮定し、与えられた入力信号の電力スペクトルを定式化した上で、最適な帯域分割を得るための理論式を明らかにする。特に、画像信号のグローバルな特性として近似できる隣接標本値間の相関係数値が 0.9 の 1 次マルコフ情報源系列を対象として、最適な帯域分割特性を具体的に求めると共に、上記の2種類の既存のサブバンド分割方式とのデータ圧縮性能を比較評価する。

2. 2 1 次元サブバンド分割

入力信号系列 $x(nT)$ の帯域 $[0, \pi]$ を M 分割するとき, linear 分割, log-linear 分割 (wavelet 変換) に対するフィルタの構成法は確立している [17]. しかし, 任意の帯域分割に対しては, 無損失で完全再構成可能なフィルタの構成法は未だ明らかにされていない. このため, 以下では帯域の分割に用いるフィルタは理想的なものであると仮定する. 現実には, 有限タップ長で近似的に構成することになるが, タップ長を十分大にとれば, 良い近似が得られるものと考えられる.

以上の仮定のもとで, 図 2.1 に示す理想的な 1 次元サブバンド分割を考える. 入力信号系列 $x(nT)$ は, 理想 BPF (Band Pass Filter) を用いて帯域分割した後に SSB 変調により基底帯域信号に変換し, それを再サンプリングすることにより帯域分割信号 $y_k(mT_k)$ に変換される.

一方, 受信側では, 各帯域信号 $y_k(mT_k)$ は, 内挿用理想 LPF (Low Pass Filter), 搬送帯域信号に変換する SSB 変調, 及び再サンプリング後の加算を経て, $x(nT)$ に変換される. その変換図を図 2.2 に, 変換式を以下に示す.

$$\begin{aligned} y_k(mT_k) &= \sum_{n=-M}^{n=-} A_k^*(mT_k, nT) x(nT) \\ x(nT) &= \sum_{k=1}^M \sum_{m=-} A_k(nT, mT_k) y_k(mT_k) \end{aligned} \quad (2.1)$$

ただし,

$$\begin{aligned} A_k(nT, mT_k) &= \frac{(-1)^m \sin(n - kT) - \sin(n - k-1T)}{(nT - mT_k)/T} \\ A_k^*(mT_k, nT) &= \frac{T_k}{T} A_k(nT, mT_k) \end{aligned} \quad (2.2)$$

$$T_k = \frac{2\pi}{k} \quad (2.3)$$

$$T_k = \frac{2\pi}{k} \quad (2.4)$$

とする. なお, 式(2.2)に示した $A_k(nT, mT_k)$ なる系列には, 次式に示す直交性がある.

$$\begin{aligned} \sum_{k=1}^M A_k(nT, mT_k) A_k^*(m'T_k, n'T) &= \delta_{mm'} \\ \sum_{n=-} A_k(nT, mT_k) A_{k'}^*(m'T_k, nT) &= \delta_{kk'} \delta_{mm'} \end{aligned} \quad (2.5)$$

故に, $x(nT)$ と $y_k(mT_k)$ ($k=1,2,\dots,M$) は直交変換対であるので, $x(nT)$ の各帯域に落ちる信号電力と $y_k(mT_k)$ の信号電力は等しくなる. サブバンド符号化においては, 各帯域の信号 $y_k(mT_k)$ は独立に, 且つ, 最適のビット配分で量子化が行われるものとする. このとき, 各帯域の信号に加わる量子化雑音電力の総和は, 同じくこの直交性のために合成フィルタ出力に現れる雑音電力と等しくなる.

均等分割及び対数スケール上での均等分割については Δ_k は次のようになる.

(a)帯域を均等に分割する方式 (linear 分割)

$$\Delta_k = \frac{1}{M} T \quad (k = 1, 2, \dots, M) \quad (2.6)$$

(b)帯域を対数スケール上で均等に分割する方式 (log-linear 分割)

$$\Delta_k = \begin{cases} \frac{1}{2^{M-1}} T & (k = 1) \\ \frac{1}{2^{M-k+1}} T & (k = 2, 3, \dots, M) \end{cases} \quad (2.7)$$

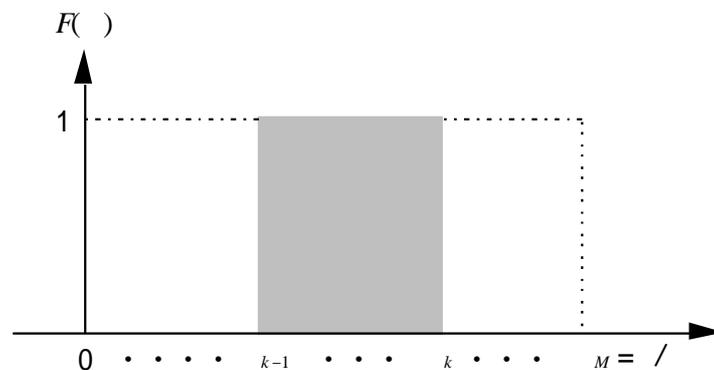


図 2.1 1次元サブバンド分割

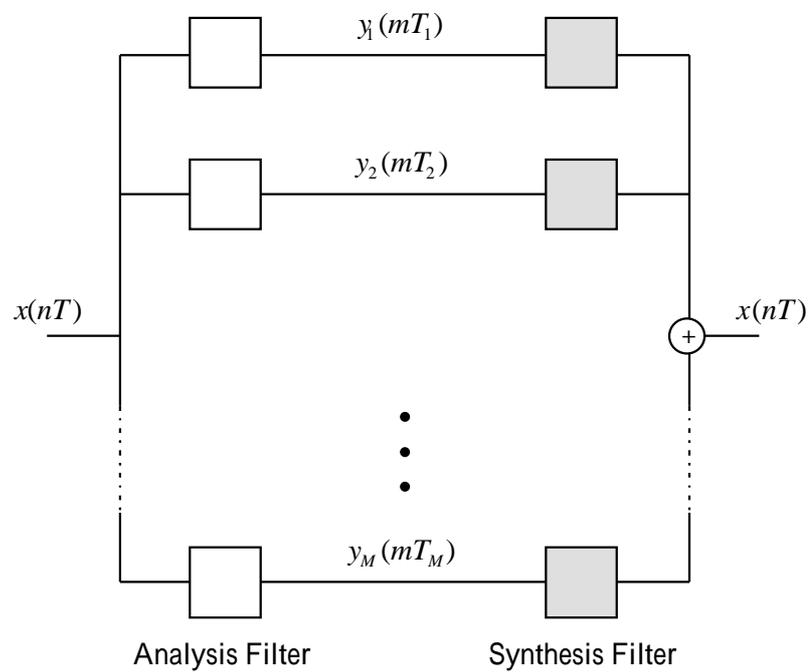


図 2.2 M 帯域の分割 / 合成方式の構成図

2.3 最適帯域分割の導出

各帯域の信号に対してスカラ量子化を適用したときの総量子化雑音電力 N が最も小さくなるような帯域分割 k 及び量子化ビット配分 R_k ($k=1,2,\dots,M$) を決定する。以下、式の簡単化のため $T=1$ に規格化して考える。

受信側における再構成信号に現れる全体の量子化雑音電力 N は、 R_k が十分大きいとしたときに、係数倍を除いて次式で近似的に表される[49]。

$$N = \prod_{k=1}^M \frac{2}{k} 2^{-2R_k} \int_{k-1}^k P_x(\omega) d\omega \quad (2.8)$$

ここで、

M : 帯域分割数

$\frac{2}{k}$: 各帯域の信号電力

$P_x(\omega)$: 入力信号の両側電力スペクトル

である。

信号全体に割り当てられるビットレート V を一定とすると

$$V = \sum_{k=1}^M R_k = \sum_{k=1}^M \frac{k - k-1}{k} R_k \quad (2.9)$$

ここで、

$$k : \text{各帯域の信号の速度比} \quad \prod_{k=1}^M k = 1$$

である。

式(2.9)の条件のもとで、 N を最小にする k 及び R_k の最適値 $k^{(opt)}$ 及び $R_k^{(opt)}$ を Lagrange の未定乗数法を用いて求める。このとき、 R_k は現実には非負整数であるが、理想条件で最適帯域分割を求めるため、実数も取り得るものとして定式化及び評価を行う。

その手順を以下に示す。

$$Q(k) = \frac{2}{k} \int_{k-1}^k P_x(\omega) d\omega \quad (2.10)$$

とおくことにより、式(2.8)は

$$N = \prod_{k=1}^M \{Q(k) - Q(k-1)\} 2^{-2R_k} \quad (2.11)$$

となる。

ここで,

$$F = N + \mu \sum_{k=1}^M \frac{R_k}{k} - V \quad (2.12)$$

として, Lagrange の未定乗数法を適用する.

$$\frac{\partial F}{\partial R_k} = Q'(k) \{e^{-R_k} - e^{-R_{k+1}}\} + \frac{\mu(R_k - R_{k+1})}{k} = 0 \quad (2.13)$$

$$\frac{\partial F}{\partial R_k} = \{Q(k) - Q(k-1)\} (-)e^{-R_k} + \frac{\mu(k - k-1)}{R_k} = 0 \quad (2.14)$$

ただし, $Q(k) = 2 \ln 2$ とする.

上式より,

$$R_k = \frac{1}{\mu} \ln \frac{B_k}{B_{k-1}} \quad (2.15)$$

ただし,

$$B_k = \frac{k - k-1}{Q(k) - Q(k-1)} \quad (2.16)$$

となる. この R_k を式(2.9)に代入して未定乗数 μ を求め, 更に, この μ を式(2.13) (2.15)に代入することにより次式を得る.

$$R_k = V - \frac{1}{M} \ln \frac{B_k}{\prod_{i=1}^M B_i} \quad (2.17)$$

$$Q'(k) = \ln \frac{B_{k+1}}{B_k} \frac{1}{B_{k+1} - B_k} \quad (2.18)$$

式(2.17)は, 帯域分割 $\{k\}$ に対する最適ビット数 $R_k^{(opt)}$ を表す. したがって, $R_k^{(opt)} > 0$ とするために, V は次式を満足せねばならない.

$$V > \frac{1}{M} \ln \frac{B_k}{\prod_{i=1}^M B_i} \quad (k = 1, 2, \dots, M) \quad (2.19)$$

また, 式(2.18)は, 式(2.16)と組み合わせて $\{k\}$ が最適帯域分割 $\{k^{(opt)}\}$ になるための $\{k\}$ の間の関係を表す. 条件 $R_0 = 0$, $R_M = \dots$ のもとで, この関係を満たす $\{k\}$ が最適帯域分割になり, これは以下に示すアルゴリズムにより求められる.

式(2.16) (2.18)及び $Q'(k) = P_x(k)/k$ より, R_1 に適当な初期値を与え,

$$B_1 = \frac{1}{\{Q(\omega_1) - Q(0)\}} \quad (2.20)$$

として,

$$\begin{aligned} f(\omega_k) &= \omega_k - \omega_{k-1} - B_k \{Q(\omega_k) - Q(\omega_{k-1})\} \\ g(B_{k+1}) &= \frac{P(\omega_k) \{B_{k+1} - B_k\}}{B_{k+1}} - \ln(B_{k+1}/B_k) \end{aligned} \quad (2.21)$$

なる2関数の根 ω_k 及び B_{k+1} ($k=2, \dots, M$) を交互に, 順次求める. しかし, 条件 $M=$ でなければ, 初期値である ω_1 を再度選び直し, 再演算を行う. ω_1 に対して M は単調増大するので, 演算を反復すれば常に解は収束し, $\omega_k^{(opt)}$ ($k=2, \dots, M-1$) が求められる.

入力信号系列の電力スペクトル $P_x(\omega)$ の微分 $P_x'(\omega)$ が, 区間 $(0, \omega_1)$ で常に負であるような信号に対しては, $f(\omega_k), g(B_{k+1})$ が共に凸関数であることが容易に証明できる. この場合には, 式(2.21)の根を求めるのにNewton法[50]を適用することができ, 次式の漸化式を反復して根が求められる.

$$\begin{aligned} B_{k+1}^{(n+1)} &= B_{k+1}^{(n)} \frac{\frac{B_k}{B_{k+1}^{(n)}} P(\omega_k) - \ln \frac{B_k}{B_{k+1}^{(n)}} - 1}{\frac{B_{k+1}^{(n)}}{B_{k+1}^{(n+1)}} P(\omega_k) - 1} \\ \omega_k^{(n+1)} &= \omega_k^{(n)} - \frac{B_k \{Q(\omega_k^{(n)}) - Q(\omega_{k-1}^{(n)})\} + \omega_k^{(n)} - \omega_{k-1}^{(n)}}{\frac{B_k}{B_{k+1}^{(n)}} P(\omega_k^{(n)}) - 1} \end{aligned} \quad (2.22)$$

2.4 1 次マルコフ信号に対する最適帯域分割

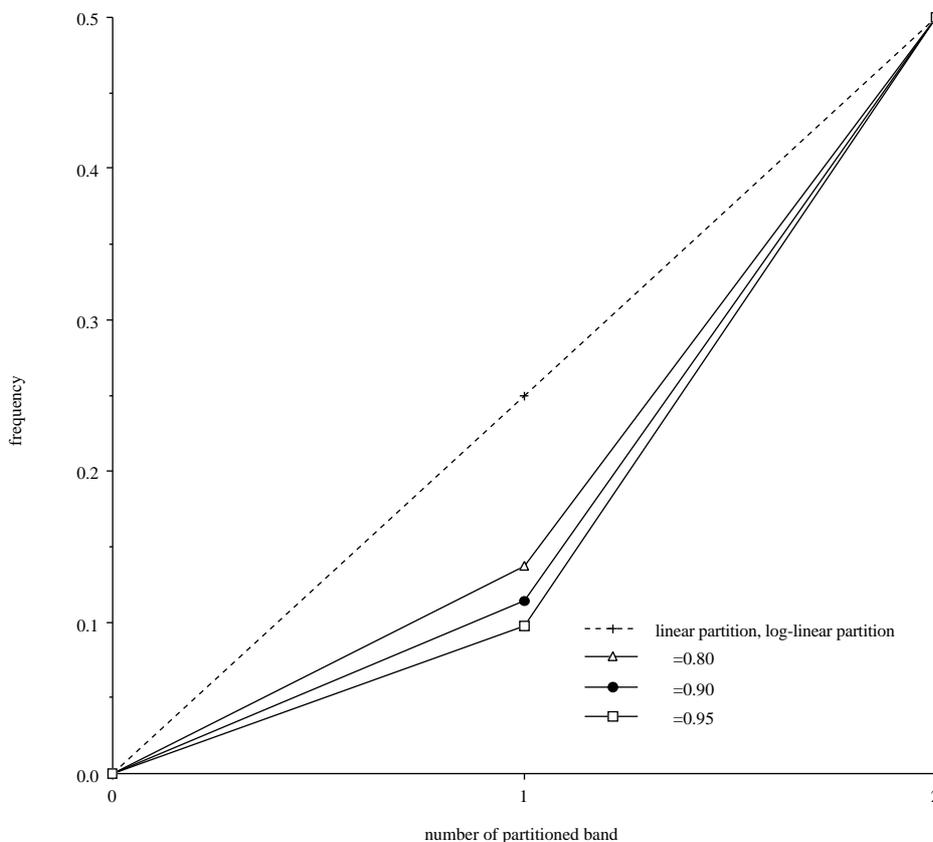
評価する信号を隣接標本値間の相関係数値が ρ である1次マルコフ情報源系列とする。この信号の規格化した電力スペクトル $P_x(f)$ は次式で与えられる[49]。

$$P_x(f) = \frac{1}{1 + \frac{4}{(1-\rho)^2} \sin^2 \frac{\pi f}{2}} \quad (2.23)$$

このとき、式(2.10)より、 $Q(f)$ は次式となる。

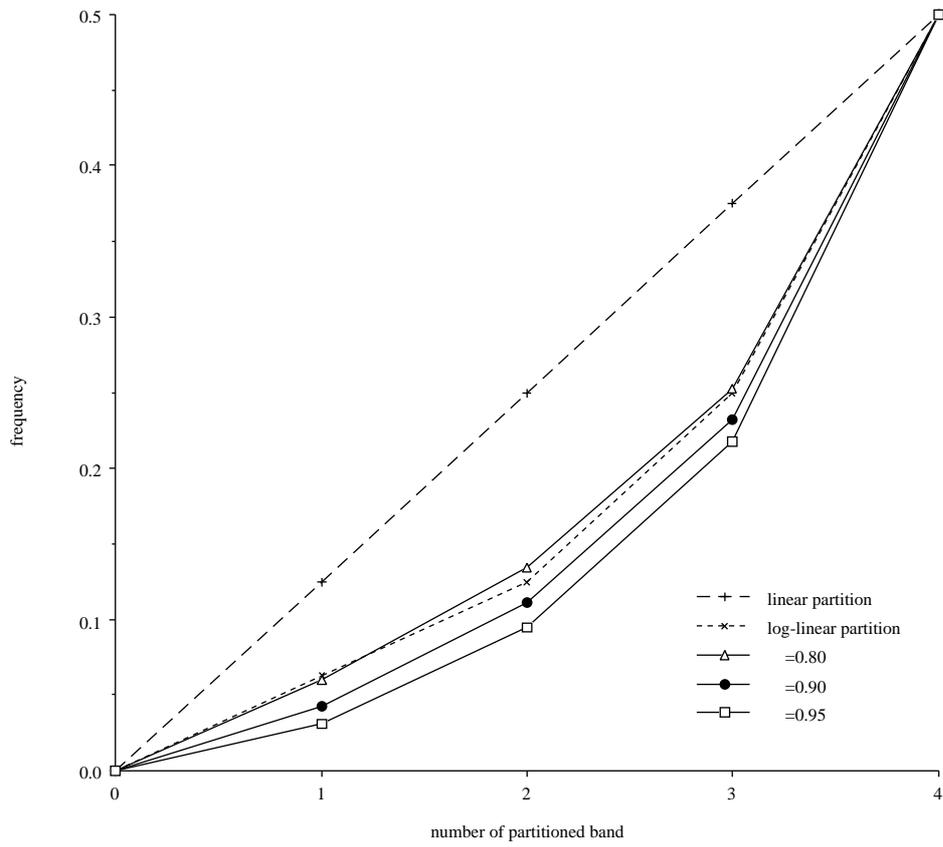
$$Q(f) = \frac{2}{1+\rho} \tan^{-1} \frac{1+\rho}{1-\rho} \tan \frac{\pi f}{2} \quad (2.24)$$

明らかに、 $P_x'(f) < 0$ ($0 < f < 1$) が成立するから、式(2.23) (2.24)を式(2.22)に適用して、分割数 $M=2, 4$ 及び 8 とした場合に求められた最適帯域分割における帯域(カットオフ周波数)の周波数配置を、既存の帯域分割方式である linear 分割, log-linear 分割と比較した結果を図2.3に示す。また、画像信号を仮定した上で、 $\rho=0.9$ [49],[51]なる上記信号系列について、 $M=2, 4$ 及び 8 それぞれの場合に、各 $f_k^{(opt)} (= f_k^{(opt)}/2)$ と $R_k^{(opt)}-V$ の値を求めた結果を表2.1に示す。

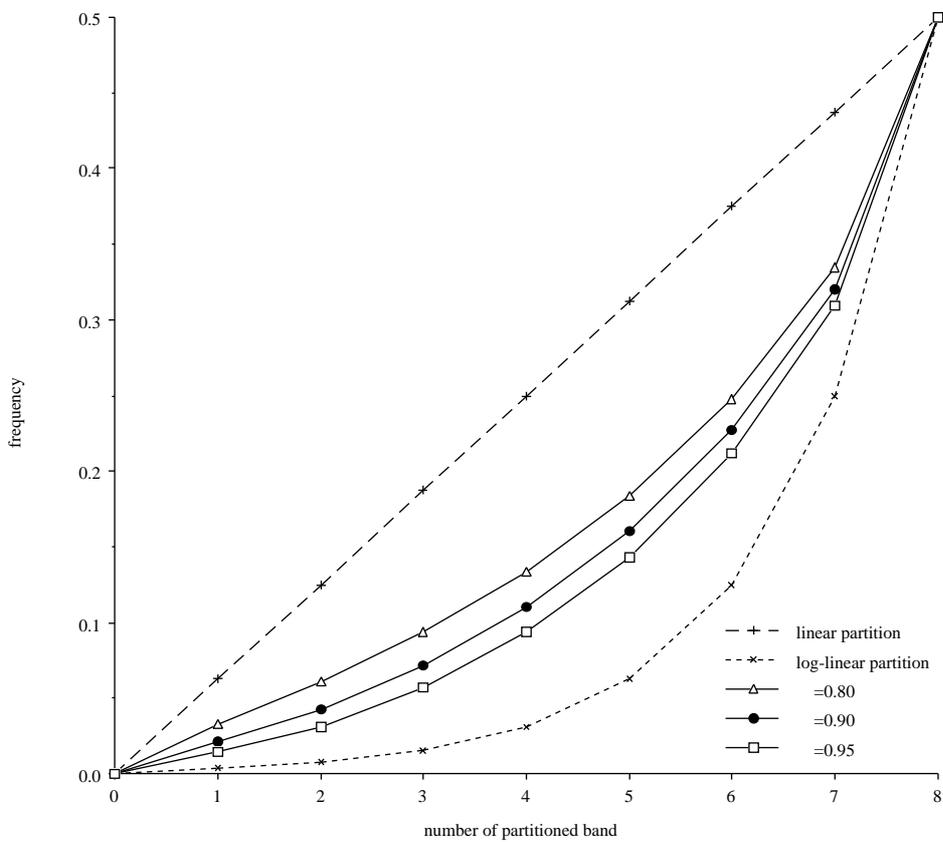


(a) 分割数 $M=2$

図2.3 分割された各帯域の周波数配置



(b) 分割数 $M=4$



(c) 分割数 $M=8$

図 2.3 分割された各帯域の周波数配置

表 2.1 最適帯域分割のカットオフ周波数及び最適ビット配分値 ($\alpha=0.9$ の場合)

(a) 分割数 $M=2$

k	f_k	$R_k - V$ (bit)
0	0.00000	—
1	0.11383	1.975809
2	0.50000	-0.582430

(b) 分割数 $M=4$

k	f_k	$R_k - V$ (bit)
0	0.00000	—
1	0.04237	2.733256
2	0.11080	1.199511
3	0.23190	0.069296
4	0.50000	-0.769426

(c) 分割数 $M=8$

k	f_k	$R_k - V$ (bit)
0	0.00000	—
1	0.02135	3.049384
2	0.04300	2.230149
3	0.07168	1.507634
4	0.10984	0.883615
5	0.16014	0.341858
6	0.22687	-0.129733
7	0.32001	-0.536112
8	0.50000	-0.856826

2.5 評価方法とその結果

式(2.17)の R_k を式(2.11)に代入すれば，量子化雑音電力 N は次式のように求められる．

$$N = 2^{-2V} \prod_{k=1}^M \frac{2^k}{k} \quad (2.25)$$

一方，帯域分割を行わず，入力された信号系列を直接スカラ量子化するときの量子化雑音電力 N_{PCM} は，次式で与えられる．

$$N_{PCM} = 2^{-2V} \prod_{k=1}^M \frac{2^k}{k} \quad (2.26)$$

この非帯域分割方式からの量子化雑音改善量 $G = N_{PCM}/N$ の大小により，各帯域分割方式に対する性能の評価を行う．この G は次式で定義される．

$$G = \frac{N_{PCM}}{N} = \frac{\prod_{k=1}^M \frac{2^k}{k}}{\prod_{k=1}^M \frac{2^k}{k}} \quad (2.27)$$

ここで，

$$\frac{2^k}{k} = \frac{2^{k-1}}{k-1} = \frac{2^k}{k} \quad (2.28)$$

$$\frac{2^k}{k} = Q\left(\frac{2^k}{k}\right) - Q\left(\frac{2^{k-1}}{k-1}\right)$$

である．

2.5.1 帯域分割による改善量 G

まず，帯域分割による量子化雑音改善量 G の値を定義する． k として前節において導出された $k^{(opt)}$ を用いれば，最適帯域分割が行われ，その後の各帯域信号に対して最適ビット割当てが行われた場合の量子化雑音改善量 G_{opt} が求められる．また， k として式(2.6)，或いは式(2.7)を用いれば，各々 linear 分割，log-linear 分割が行われ，且つ，各方式による帯域分割後に最適なビット割当てをしたときの改善量 G_{lin} 及び G_{log} が求められる．

2.5.2 改善量 G の理論限界

不規則系列 $x(n)$ の自己相関関数を $r(n)$ ，電力スペクトルを $P_x(\omega)$ とするとき， $E[x(n)] = 0$ を

仮定すれば， $x(n)$ の電力²は，

$$r(0) = \frac{1}{2} \int_{-\pi}^{\pi} P_x(\omega) d\omega \quad (2.29)$$

と表される．

今，任意の1次元信号系列 $x(n)$ に対して，無限次数の線形予測フィルタを適用したときの予測誤差 $e(n)$ の電力²は次式で与えられる．

$$r_e = \frac{1}{2} \int_{-\pi}^{\pi} \left| 1 - \sum_{k=1}^{\infty} a_k e^{-j\omega k} \right|^2 P_x(\omega) d\omega \quad (2.30)$$

この予測誤差電力²を最小にする予測係数 a_1, a_2, \dots が最適予測係数であり，そのときの最小値を $r_{e,min}$ ²とすれば，量子化雑音改善量 G の理論限界値 G_{lim} は，次式で定義される．

$$G_{lim} = \frac{r(0)}{r_{e,min}} = \frac{r_x(0)}{r_{e,min}} \quad (2.31)$$

無限次数の最適予測を行うとすれば，予測誤差 $e(n)$ は完全に無相関な系列となり，その

電力スペクトル $\left| 1 - \sum_{k=1}^{\infty} a_k e^{-j\omega k} \right|^2 P_x(\omega)$ が白色化されることより，

$$P_x(\omega) = \frac{r_{e,min}}{\left| 1 - \sum_{k=1}^{\infty} a_k e^{-j\omega k} \right|^2} \quad (2.32)$$

なる関係が成立する．

ここで，系列 $x(n)$ の自己相関関数 $r(n)$ が与えられると，Wiener-Khinchineの関係から電力スペクトル $P_x(\omega)$ が求められ，それを式(2.32)のような形式で表現したときの係数 $r_{e,min}$ ²を求め，それを式(2.31)に代入することで G_{lim} が求められる．

例えば，2.4節において入力信号として用いた相関係数値 α の1次マルコフ情報源系列の場合，その電力スペクトルは，

$$P_x(\omega) = \frac{r(0)}{\left| 1 - \alpha e^{-j\omega} \right|^2} = \frac{1 - \alpha^2}{\left| 1 - \alpha e^{-j\omega} \right|^2} \quad (2.33)$$

と表すことができるから， $r_{e,min} = 1 - \alpha^2$ となることが分かり，また， $r(0) = 1.0$ であるから，この場合の理論限界値 G_{lim} は，

$$G_{lim} = \frac{1}{1 - \alpha^2} \quad (2.34)$$

として得ることができる．

2.5.3 評価結果

最適帯域分割に対する評価を行うために、 $M=2, 4$ 及び 8 の各場合について、に対する $G_{opt}[dB]$ の特性を図2.4に示す。同図には、各に対する量子化雑音改善量の理論限界値である G_{lim} の特性を併せて表示する。図2.4は、帯域分割処理による量子化雑音改善量の最適値、すなわち、帯域分割処理による量子化雑音改善量の理論限界を明らかにしたものである。その改善の効果は、入力信号系列の相関係数値が1.0に近くなるほど有効であることを示しており、特に画像信号を仮定した $\rho=0.9$ の信号系列に対する最適帯域分割による改善量は、 $M=4$ の場合において、約 $10[dB]$ である。また、 $M=4$ と $M=8$ の場合において、各 G_{opt} の大きさにほとんど差はなく、しかもこれらの値は、理論限界値である G_{lim} の値に漸近していることから、最適帯域分割を実現することを仮定した場合、システム実現の簡易化のためには、帯域分割数はできるだけ小さい方が望ましいことから、帯域分割数は4で十分であると結論づけられる。

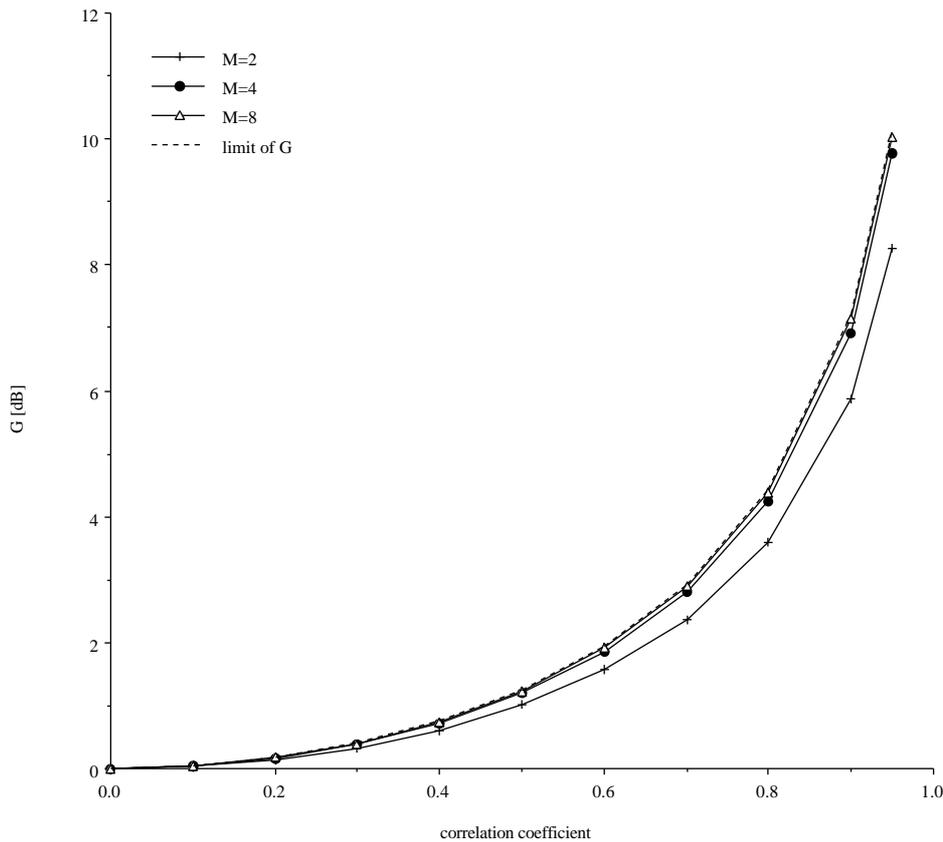


図2.4 G_{opt} 特性 (パラメータ: 相関係数値)

次に、既存の帯域分割方式であるlinear分割, log-linear分割における量子化雑音改善量と導出された最適帯域分割との比較を行うために、 $M=4, 8$ の場合の $G_{log}/G_{lin}, G_{opt}/G_{lin}$ [dB] の特性を図2.5に示す。図2.5において、帯域分割数の違いによって多少の差はあるものの、入力信号系列の相関係数値 ρ が0に近いところでは、linear分割による特性が最適帯域分割の特性を近似しており、逆に ρ が1.0に近いところでは、log-linear分割による特性が最適帯域分割のものにより漸近していることが分かる。この ρ が1.0に近い信号系列の場合において、 G_{opt} と G_{log} との差が僅少であることに加え、最適帯域分割の帯域分割特性を実現するフィルタバンクの構成が容易ではないことを考慮すれば、グローバルな特性として、 ρ が1.0に近い1次マルコフ情報源系列で近似される画像信号に対しては、log-linear分割を適用することによって、最適帯域分割に代わり得る十分な量子化雑音改善量を得ることができる。

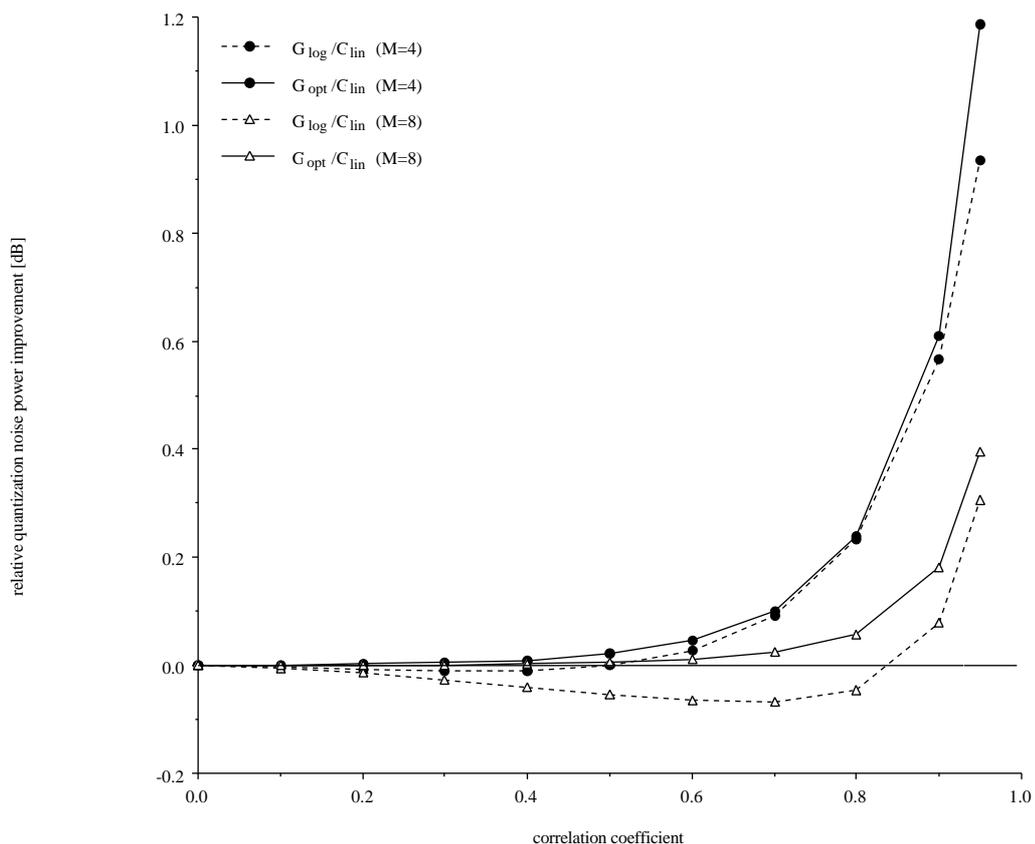


図 2.5 G_{log}/G_{lin} 及び G_{opt}/G_{lin} 特性 (パラメータ: 相関係数値 ρ)

2. 6 結言

画像信号のサブバンド符号化における最適帯域分割の導出を目的として、その議論の理論的基礎となる1次元の信号系列を対象とした最適帯域分割を与える理論式を明らかにした。特に、1次元信号系列の例として、画像信号をグローバルに近似する隣接標本値間の相関係数 $\rho = 0.9$ の1次マルコフ情報源系列を入力信号として、最適帯域分割の特性を具体的に算出した。

非帯域分割方式を基準として、帯域分割処理によって低減する量子化雑音電力を定量的に評価し得る値として提案された量子化雑音改善量なる評価尺度に基づいて、最適帯域分割の評価を行った結果、画像信号をグローバルに近似し得る相関係数 $\rho = 0.9$ の1次マルコフ情報源系列に対して、最適帯域分割は、帯域分割数が僅か4である場合に、上記の信号系列における量子化雑音改善量の理論限界値と匹敵する量子化雑音改善性能を有していることが明らかにされた。

また、フィルタの構成法が明らかである既存の帯域分割方式の性能評価を行うために、帯域を均等に分割するlinear分割と、帯域を対数スケール上で均等に分割するlog-linear分割、すなわち、wavelet変換に着目し、両帯域分割方式と最適帯域分割との性能比較を行った。この結果、1次マルコフ情報源系列に対しては、信号系列の相関係数 ρ が0に近いところでは、linear分割を用いることで、逆に ρ が1.0に近いところでは、log-linear分割を用いることで、最適帯域分割と上記の既存の各帯域分割方式との量子化雑音改善量の差は僅少となる。特に、 ρ が1.0に近い場合についての結果に基づいて、1次元の画像信号系列に対しては、フィルタバンクの構成法が明らかでない最適帯域分割を用いなくとも、フィルタバンクを実現が可能である既存のlog-linear分割を適用することで、有効な量子化雑音の低減を行うことができることが明らかにされた。

ここで得られた結論は、実際にサブバンド符号化を実現する上で、画像を帯域分割する際に、その帯域分割方式としてwavelet変換が多く適用されていることの1つの根拠ではあるが、本来、画像信号は空間領域における2次元の信号系列であるから、本章で行われた議論を2次元へと拡張し、2次元信号系列に対する最適帯域分割を直接的に導出した上で、その評価を行う必要がある。2次元サブバンド符号化における最適帯域分割の導出については、次章以降で述べられる。

第 3 章

画像信号の 2 次元サブバンド符号化における最適帯域分割

3.1 序言

画像信号の高能率符号化の有力な一手法であるサブバンド符号化は、DCTにより符号化された再生画像に現れる最大の画質劣化要因とされるブロックひずみの発生がなく、また、伝送誤りの影響が空間的、周波数的に制限される利点から、本符号化に基づいた数多くの画像符号化方式が検討されている。サブバンド符号化方式では、まず、入力された画像信号を帯域分割し、各帯域に落ちる信号電力に対応した量子化ビット数の配分を行う。良い画質で高い圧縮効率を得るためには、適切な帯域分割及び適切な量子化ビット配分が重要になる。各帯域の信号電力に応じて配分ビット数を変えるビット配分の最適化は、文献[44]で論じられているが、2次元の信号系列を対象とした帯域分割の最適化については、未だその議論は行われていない。現実の画像信号は2次元の信号系列であるから、入力信号系列である画像信号の統計的性質：2次元の自己相関関数、に整合した2次元帯域分割の最適化を行う必要がある。ところで、画像信号は非定常性があり、特に局所的特徴は重要であって無視することはできない[51]。本研究の目的は、非定常性：局所的特徴、を考慮し得る最適帯域分割の導出とその実現であり、本章では、それに到達するための理論的基礎として是非とも必要となる、統計的性質：2次元の自己相関関数、に整合した最適帯域分割の理論的な導出を行う。この結果は、これだけでも従来法より優れており、その効果も定量的に明らかにする。

従来、画像信号の帯域分割法として、1次元のQMF (Quadrature Mirror Filter) [52],[53]等を2次元空間領域において互いに直交する水平及び垂直軸方向に適用するSeparableフィルタバンクを用いることが一般的であった[16],[18],[54]。しかし、この処理においては、分割される2次元帯域の形状が長方形に限定されてしまい、任意の形状の帯域分割パターンが得られないという欠点が生じる。この欠点を補うことに併せて、人間の視覚特性が2次元空間領域の斜め成分に対して比較的敏感でない(具体的には、水平、垂直方向

の約1/2][55]ことから、近年、2次元フィルタを用いて帯域分割を行う non-separable フィルタバンクに関する研究が行われている[47],[48]。しかし、上述した入力信号の統計的性質に整合した最適帯域分割という観点からの検討は未だ行われていない。

本章では、第2章で行われた1次元信号系列に対する最適帯域分割の議論を2次元へと拡張することで、任意の帯域分割特性を有する2次元の非分離型フィルタの適用を前提に、2次元周波数平面の最適な帯域分割を理論的に明らかにする。まず、2次元周波数平面上において、任意の帯域分割フィルタバンクの構成に対する理論モデルを与え、分割フィルタの各帯域とその出力レートの関係について述べる。次に、分割後の各帯域信号がスカラ量子化されると仮定したとき、(1) 信号全体に割り当てられるビットレートが一定である、(2) 受信端での再構成信号の量子化雑音電力を最小にする、という条件のもとで2次元周波数平面の最適帯域分割及び分割後の各帯域信号への最適ビット配分を定める理論式を明らかにする。量子化雑音電力のみを考慮した理由は、サブバンド符号化では、ブロックひずみの発生がないために、符号化の誤差はランダムノイズのみを考慮すればよいと考えたからである。ここでは更に、最適帯域分割及び最適ビット配分の具体的な数値を算出するための数値計算アルゴリズムを明らかにする。そして、画像信号の統計的性質に基づいた理論的な画像モデルを導入して、本モデルを対象とした最適帯域分割と最適ビット配分の具体的な数値例を示す。画像モデルに関しては、まず初めに、2次元空間領域において、特に、水平及び垂直方向の成分が多い画像についてよく適合する画像モデルとして水平・垂直相関分離型画像モデルを導入する。本画像モデルは、自己相関関数が1次元の場合の積として与えられるために、最適帯域分割の具体的な導出を1次元の場合から比較的簡単に拡張することができる。最終的に、実際の画像信号に対してより一般的な画像モデルを与えるために、数枚の実画像に対して自己相関特性の実験的調査を行う。この結果、多くの画像の相関特性は、2次元空間領域で水平及び垂直方向の距離に対しほぼ等方性の特徴を有することから、一般の画像信号に対してより適切であると考えられる汎用的な画像モデルとして、自己相関関数が水平及び垂直方向の間で独立ではなく、等方性を有する空間的距離の関数で与えられる水平・垂直相関非分離型画像モデルを導入する。以降、これら2つの画像モデルを対象に、最適帯域分割と最適ビット配分の具体的な数値例を示すと共に、前章において定義された量子化雑音改善量を2次元へと拡張したものを評価尺度として、各画像モデルに対するこの値の理論限界値を明らかにした上で、導出された最適帯域分割と従来から用いられている separable フィルタバンク及び既存の実現可能な non-separable フィルタバンク、更にはJPEGで用いられている数画素毎のブロック単位で適用されるDCTとのデータ圧縮性能を比較評価し、最適帯域分割の有効性を定量的に明らかにする。

3.2 2次元サブバンド分割と分割出力レート

サンプリングレートを水平, 垂直方向共に1に規格化した2次元の入力信号系列 $x(m,n)$ を対象に, その帯域 $\Omega = \{(\omega_h, \omega_v) | 0 \leq \omega_h, \omega_v \leq \pi\}$ を式(3.1)のように互いに素な M 個の部分帯域 (サブバンド) に分割するフィルタの構成法について考える.

$$\left. \begin{aligned} \Omega &= \Omega_1 + \Omega_2 + \dots + \Omega_M \\ \Omega_i \cap \Omega_j &= \emptyset \quad (i \neq j) \end{aligned} \right\} \quad (3.1)$$

2次元周波数平面上において, 帯域 Ω を任意の形状の帯域 Ω_k ($k=1, 2, \dots, M$) に分割する無損失で完全再構成可能な2次元non-separableフィルタの構成法は, 未だ確立されていない. このため, ここでは原理的な構成モデルを考え, 本モデルが近似的に実現可能であることを示す.

まず, 図3.1に示す2次元周波数平面において, 水平及び垂直に微小帯域幅 $\Delta = \pi/N$ ($N \gg 1$) の間隔の格子で区切られた正方ますを考え, これを帯域ブロックと呼ぶことにする. 次に, それぞれの帯域ブロックを通過域とするBPFとそれらを通じた信号を, 水平及び垂直にそれぞれ $1/N$ サブサンプリングして出力する間引き回路で構成される N^2 個の2次元単位分割フィルタを用意する. ここで2次元帯域 Δ^2 を有する分割フィルタは, 1次元単位分割フィルタを水平及び垂直方向に適用する2次元separableフィルタで構成可能である. ここで, 1次元単位分割フィルタは, 既存のフィルタであるQMF[52], [53]やCQF (Conjugate Quadrature Filter) [17]を例えばtree状に多段に適用することで実現可能であり, 具体的な構成方法については, 上記の文献に示されている. このとき, タップ長を大きくとることで, 分割フィルタの遮断特性は急峻になり, 理想特性を限りなく近似することができる. また, フィルタの段数を大きくすれば, 帯域幅 Δ は任意に小さくすることができる.

k 番目の帯域 Ω_k を通過域とする分割フィルタは, 図3.1に示すように Ω_k を帯域ブロックの集まりで近似し, それらの帯域ブロックを通過域とする上記の単位分割フィルタ出力を寄せ集めることで構成できる. このとき, N の値を大きくしていけば, 帯域ブロックの集合による帯域 Ω_k の近似精度はいくらでも向上するため, 任意の帯域の2次元帯域分割特性が得られることになる.

以上の構成モデルにおいて, 水平及び垂直方向共にサンプリングレートを1に規格化した2次元の入力信号に対して, 各単位分割フィルタの2次元出力信号のレートは, $(\Delta/\pi)^2$ となる. したがって, 帯域 Ω_k の面積を S_k とすると, Ω_k の中には S_k/Δ^2 個の単位分割フィルタが含まれるから, k 番目の帯域である Ω_k を通過域とする2次元分割フィルタの出力レート λ_k は次式で表される.

$$\lambda_k = (\Delta/\pi)^2 \times (S_k/\Delta^2) = S_k/\pi^2 \quad (3.2)$$

ただし,

$$S_k = \iint_{\Omega_k} d\omega_h d\omega_v \quad (3.3)$$

$$\sum_{k=1}^M S_k = \pi^2 \quad (3.4)$$

2次元合成フィルタについても, N 倍のアップサンプリングと2次元帯域 Δ^2 を通過域にもつBPF からなる単位合成フィルタを用いて同様な原理的構成モデルを得ることができる.

以上の2次元分割/合成フィルタの構成において,分割/合成時にサンプリングレートは保存され,また,本フィルタは直交フィルタで構成されるため,分割/合成フィルタの入出力間で電力も保存される.以下では, N の値が十分大きいとして,理想2次元分割/合成フィルタを仮定した上で,最適帯域分割法について論じることにする.

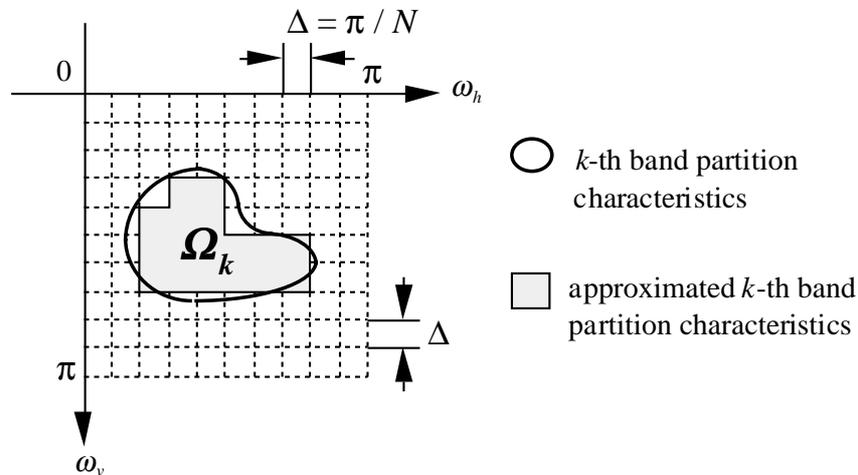


図 3.1 2次元 non-separable 分割フィルタの理論的構成モデル

3.3 2次元最適帯域分割

サブバンド符号化においては，ブロックひずみの発生がなく，符号化の誤差としてランダムノイズのみを考慮すればよいと仮定したから，量子化雑音電力に着目した最適帯域分割を考えればよい． M 個の帯域に分割した各信号にそれぞれスカラ量子化を適用したとき，全量子化雑音電力を最小にする条件から最適帯域分割及び各帯域への最適ビット配分を明らかにする．なお，前節で述べたように帯域分割フィルタは直交フィルタで構成するため，この量子化雑音電力は，伝送路上での付加雑音がないとした場合には，受信復号出力に現れる雑音電力と一致する．

3.3.1 量子化雑音電力

帯域 Ω_k ($k=1,2,\dots,M$)の信号の量子化ビット数 R_k が十分大きいときに，全体の量子化雑音電力 N は，係数倍を除いて次式で近似的に表される[49]．

$$N = \sum_{k=1}^M \sigma_k^2 2^{-2R_k} \quad (3.5)$$

ただし，

M ：分割数

σ_k^2 ： k 番目の帯域の信号電力

である．入力信号の電力スペクトルを $P_x(\omega_h, \omega_v)$ とすれば， σ_k^2 は次式で表すことができる．

$$\sigma_k^2 = \frac{1}{\pi^2} \iint_{\Omega_k} P_x(\omega_h, \omega_v) d\omega_h d\omega_v \quad (3.6)$$

また，信号全体に割り当てられるビットレート V は，次式で与えられる．

$$V = \sum_{k=1}^M \lambda_k R_k \quad (3.7)$$

ここで，式(3.2) (3.4)より次の関係が成立する．

$$\sum_{k=1}^M \lambda_k = 1 \quad (3.8)$$

以上より，最適帯域分割の問題は，入力信号の電力スペクトル $P_x(\omega_h, \omega_v)$ が与えられたときに，ビットレート V が一定の条件のもとで，式(3.5)の N を最小にする帯域分割 Ω_k 及びビット配分 R_k (共に， $k=1,2,\dots,M$)を求める問題として定式化され，これを以下の2段階に分けて解くこととする．

3.3.2 量子化ビット配分

まず，帯域分割 Ω_k ($k=1,2,\dots,M$) をある何らかの分割に固定するならば，式(3.2) (3.3) 及び式(3.6)より λ_k ， σ_k^2 の値が定まる．このとき， N を以下のように変形して $\lambda_k (> 0)$ で重み付けられた相加平均 \geq 相乗平均なる不等式[56]を適用すると，

$$N = \sum_{k=1}^M \lambda_k (\sigma_k^2 2^{-2R_k} / \lambda_k) \geq \prod_{k=1}^M (\sigma_k^2 2^{-2R_k} / \lambda_k)^{\lambda_k}$$

を得る．さらに上式の両辺に $\log_2[\cdot]/2$ の演算を適用し，式(3.7)の関係を用いて変形することで，次の不等式を得る．

$$\log_2 \sqrt{N} \geq \sum_{k=1}^M \lambda_k \log_2 [\sigma_k^2 / \lambda_k]^{1/2} - V \quad (3.9)$$

上式において等号は次式が満たされるときのみ成立する．

$$\sigma_k^2 2^{-2R_k} / \lambda_k = c \quad (k=1,2,\dots,M) \quad (3.10)$$

ただし， c は定数である．

したがって，式(3.8) (3.10)より各帯域へのビット配分を

$$R_k = V - \log_2 \left[\prod_{n=1}^M (\sigma_n^2 / \lambda_n)^{\lambda_n} / (\sigma_k^2 / \lambda_k) \right]^{1/2} \quad (k=1,2,\dots,M) \quad (3.11)$$

に選んだとき，上記の固定した帯域分割に対し，量子化雑音電力 N は最小になる．このとき，式(3.9)より

$$\log_2 \sqrt{N} = \sum_{k=1}^M \lambda_k \log_2 [\sigma_k^2 / \lambda_k]^{1/2} - V$$

が成立し，式(3.2) (3.6)を上式に適用して変形することで，次式を得る．

$$\begin{aligned} \log_2 \sqrt{N} &= \sum_{k=1}^M \frac{S_k}{\pi^2} \log_2 \left[\frac{1}{S_k} \iint_{\Omega_k} P_x(\omega_h, \omega_v) d\omega_h d\omega_v \right]^{1/2} - V \\ &= R_d + \frac{1}{\pi^2} \iint_{\Omega} \log_2 \sqrt{P_x(\omega_h, \omega_v)} d\omega_h d\omega_v - V \end{aligned} \quad (3.12)$$

ただし，式(3.12)において，

$$R_d = \sum_{k=1}^M \frac{S_k}{\pi^2} \log_2 \left[\frac{\frac{1}{S_k} \iint_{\Omega_k} P_x(\omega_h, \omega_v) d\omega_h d\omega_v}{\exp \left(\frac{1}{S_k} \iint_{\Omega_k} \ln P_x(\omega_h, \omega_v) d\omega_h d\omega_v \right)} \right]^{1/2} \quad (3.13)$$

である .

3.3.3 最適帯域分割と最適ビット配分

帯域 Ω_k ($k=1,2,\dots,M$) を変化させて , 量子化雑音電力 N を最小にする最適帯域分割を求める . 式(3.12)の右辺の第2項及び第3項は Ω_k と無関係な定数であるから , 第1項の R_d を最小にする Ω_k を求めればよい .

式(3.13)の右辺の $[\cdot]$ 内の分子及び分母は , それぞれ , $P_x(\omega_h, \omega_v)$ の帯域 Ω_k における相加平均及び相乗平均である . 周知のように相加平均/相乗平均 ≥ 1 であるから $\log_2[\cdot] \geq 0$ となり , 等号は $P_x(\omega_h, \omega_v)$ が Ω_k で定数となる時のみ成立する . したがって , 各帯域内での $P_x(\omega_h, \omega_v)$ の変動が少ないほど R_d は小さくなり , 全帯域 Ω でフラットであれば $R_d = 0$ になる . このことは , R_d を最小にする帯域分割は , 図3.2に示すように $P_x(\omega_h, \omega_v)$ が一定になる2次元周波数平面上の等電力スペクトル曲線

$$\left. \begin{aligned} P_x(\omega_h, \omega_v) &= C_k \quad (k=1,2,\dots,M-1) \\ (C_0 > C_1 > \dots > C_{M-1} > 0) \end{aligned} \right\} \quad (3.14)$$

を境界として持つような分割パターンであることを意味する . ここで , 対象とする2次元信号に画像信号を考慮することにより , 直流における $P_x(\omega_h, \omega_v)$ が最大になるものとして , その最大値 $C_0 = P_x(0,0) = 1.0$ に規格化している .

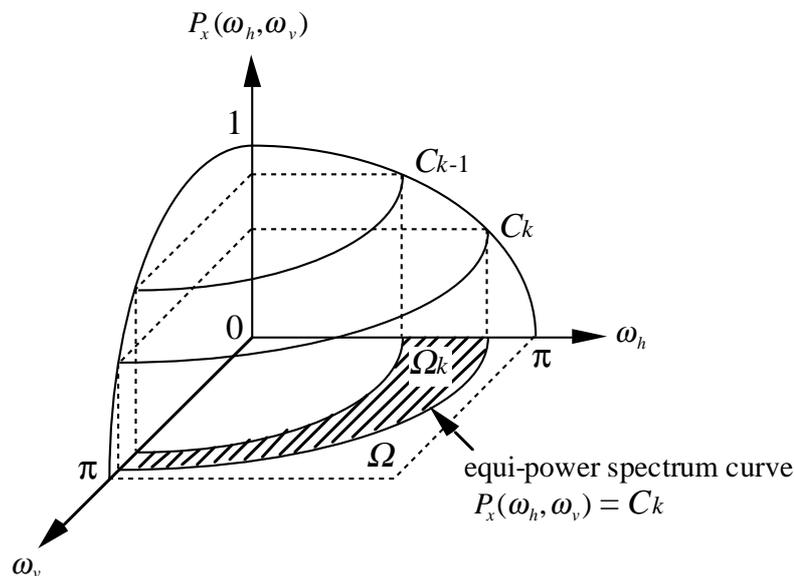


図 3.2 2次元周波数平面での帯域分割

以上より、量子化雑音電力 N を最小とする2次元周波数平面における最適帯域分割を決定する問題は、等電力スペクトル曲線を境界に持つような帯域分割パターンに限定した上で、 R_d を最小にする C_k ($k=1,2,\dots,M$)を求める問題に帰着される。このため以下、 C_k を帯域分割の境界値、 R_d を最小にする C_k を最適境界値と呼ぶことにする。

まず、

$$s(z) = \iint_{\{(\omega_h, \omega_v) | P_x(\omega_h, \omega_v) \geq z\} \cap \Omega} d\omega_h d\omega_v \quad (3.15)$$

なる関数を導入すると、式(3.14)を境界に持つ限定分割パターンでの帯域 Ω_k の面積 S_k は、次式で表される。

$$S_k = \iint_{\Omega_k} d\omega_h d\omega_v = \begin{cases} s(C_k) - s(C_{k-1}) & (k=1,2,\dots,M-1) \\ \pi^2 - s(C_{M-1}) & (k=M) \end{cases} \quad (3.16)$$

ただし、 $s(C_0)=0$ とする。

また、この $s(z)$ を用いて、以下の2式が導かれる。(付録A参照)

$$\iint_{\Omega_k} P_x(\omega_h, \omega_v) d\omega_h d\omega_v = \begin{cases} \int_{C_{k-1}}^{C_k} s'(z) z dz & (k=1,2,\dots,M-1) \\ \pi^2 \overline{P_x} - \int_{C_0}^{C_{M-1}} s'(z) z dz & (k=M) \end{cases} \quad (3.17)$$

$$\iint_{\Omega_k} \ln P_x(\omega_h, \omega_v) d\omega_h d\omega_v = \begin{cases} \int_{C_{k-1}}^{C_k} s'(z) \ln z dz & (k=1,2,\dots,M-1) \\ \pi^2 \overline{\ln P_x} - \int_{C_0}^{C_{M-1}} s'(z) \ln z dz & (k=M) \end{cases} \quad (3.18)$$

ただし、式(3.17) (3.18)において以下に定義される記号を用いた。

$$s'(z) = \frac{ds(z)}{dz}$$

$$\overline{P_x} = \frac{1}{\pi^2} \iint_{\Omega} P_x(\omega_h, \omega_v) d\omega_h d\omega_v$$

$$\overline{\ln P_x} = \frac{1}{\pi^2} \iint_{\Omega} \ln P_x(\omega_h, \omega_v) d\omega_h d\omega_v$$

式(3.16) ~ (3.18)を式(3.13)に代入して変形すれば、 R_d は C_k の関数として次式で表される。

$$\begin{aligned}
R_d = & \frac{1}{(2\ln 2)\pi^2} \sum_{k=1}^{M-1} \left\{ (s(C_k) - s(C_{k-1})) \ln \left[\frac{\int_{C_{k-1}}^{C_k} s'(z) z dz}{s(C_k) - s(C_{k-1})} \right] - \int_{C_{k-1}}^{C_k} s'(z) \ln z dz \right\} \\
& + \frac{1}{(2\ln 2)\pi^2} \left\{ (\pi^2 - s(C_{M-1})) \ln \left[\frac{\pi^2 \bar{P}_x - \int_{C_0}^{C_{M-1}} s'(z) z dz}{\pi^2 - s(C_{M-1})} \right] - \left(\pi^2 \ln \bar{P}_x - \int_{C_0}^{C_{M-1}} s'(z) \ln z dz \right) \right\} \quad (3.19)
\end{aligned}$$

次に，上式を $C_k (k=1,2,\dots,M)$ で偏微分してゼロとおくと， R_d の極値を与える C_k の漸化式

$$C_k = \frac{\ln[1/\bar{P}_x(C_{k+1})] - \ln[1/\bar{P}_x(C_k)]}{1/\bar{P}_x(C_{k+1}) - 1/\bar{P}_x(C_k)} \quad (k=1,2,\dots,M-1) \quad (3.20)$$

を得る．ここで，上式の $\bar{P}_x(C_k)$ は

$$\bar{P}_x(C_k) = \frac{1}{S_k} \iint_{\Omega_k} P_x(\omega_h, \omega_v) d\omega_h d\omega_v = \begin{cases} \frac{\int_{C_{k-1}}^{C_k} s'(z) z dz}{s(C_k) - s(C_{k-1})} & (k=1,2,\dots,M-1) \\ \frac{\pi^2 \ln \bar{P}_x - \int_{C_0}^{C_{M-1}} s'(z) \ln z dz}{\pi^2 - s(C_{M-1})} & (k=M) \end{cases} \quad (3.21)$$

であり，帯域 Ω_k における $P_x(\omega_h, \omega_v)$ の相加平均を表す．以上で帯域分割の最適境界値を定める理論式が得られた．

最後に，最適帯域分割パターンが決定された後の各帯域への最適ビット配分について述べる．帯域 Ω_k の信号電力 $\sigma_k^2 (k=1,2,\dots,M)$ は，式(3.6)に式(3.17)を代入することで

$$\sigma_k^2 = \begin{cases} \frac{1}{\pi^2} \int_{C_{k-1}}^{C_k} s'(z) z dz & (k=1,2,\dots,M-1) \\ \bar{P}_x - \frac{1}{\pi^2} \int_{C_0}^{C_{M-1}} s'(z) z dz & (k=M) \end{cases} \quad (3.22)$$

のように，式(3.20)により求められる C_k の関数の形で表すことができる．また，式(3.2)に示した帯域 Ω_k を通過域とする2次元分割フィルタの出力レート $\lambda_k (k=1,2,\dots,M)$ も，式(3.16)を代入することで，

$$\lambda_k = \begin{cases} \frac{s(C_k) - s(C_{k-1})}{\pi^2} & (k=1,2,\dots,M-1) \\ 1 - \frac{s(C_{M-1})}{\pi^2} & (k=M) \end{cases} \quad (3.23)$$

として C_k の関数の形で表すことができる．したがって，帯域 Ω_k への最適ビット配分 R_k ($k=1,2,\dots,M$) は，式(3.22) (3.23)を式(3.11)に代入して，

$$R_k = \begin{cases} V - \log_2 \left[\frac{R_M \{s(C_k) - s(C_{k-1})\}}{\int_{C_{k-1}}^{C_k} s'(z) z dz} \right]^{1/2} & (k = 1, 2, \dots, M-1) \\ V - \log_2 \left[\frac{R_M \left\{1 - \frac{s(C_{M-1})}{\pi^2}\right\}}{\bar{P}_x - \frac{1}{\pi^2} \int_{C_0}^{C_{M-1}} s'(z) z dz} \right]^{1/2} & (k = M) \end{cases} \quad (3.24)$$

として，求めることができる．ただし，上式において，

$$R_M = \prod_{k=1}^{M-1} \left(\frac{\int_{C_{k-1}}^{C_k} s'(z) z dz}{s(C_k) - s(C_{k-1})} \right)^{\frac{s(C_k) - s(C_{k-1})}{\pi^2}} \cdot \left(\frac{\bar{P}_x - \frac{1}{\pi^2} \int_{C_0}^{C_{M-1}} s'(z) z dz}{1 - \frac{s(C_{M-1})}{\pi^2}} \right)^{\left(1 - \frac{s(C_{M-1})}{\pi^2}\right)} \quad (3.25)$$

である．

3.3.4 最適帯域分割算出アルゴリズム

式(3.20)を満たす $\{C_k\}$ ($k=1,2,\dots,M$) は，2次元周波数平面において帯域の最適分割が行われたときの最適境界値 $\{C_k^{(opt)}\}$ を与え，これは以下のアルゴリズムによって算出できる．

まず初期条件として，電力スペクトル $P_x(\omega_h, \omega_v)$ の直流成分が最大値1.0に規格化されているとして， $C_0 = P_x(0,0) = 1.0$ とする．ここで，初期値として， C_1 に $P_x(\pi, \pi) = C_M < C_1 < 1.0$ を満たす適当な値を与え，式(3.21)より $\bar{P}_x(C_1)$ を求め，現段階においてこの値は固定しておく．次に， C_2 として $C_M < C_2 < C_1$ なる条件を満たす適当な値を与えれば，式(3.21)より $\bar{P}_x(C_2)$ が求められる．これらの値を用いて，式(3.20)が満足するか否かを調査し，もし満足していなければ，再度 C_2 を選び直す．このとき， $\bar{P}_x(C_2)$ の値は， C_2 の減少に対して単調に増加する傾向にあるので，解探索法の一つである挟み撃ち法[50]を用いて式(3.20)を満足する C_2 を決定することができる．この段階で $\bar{P}_x(C_2)$ の値を固定し， $C_M < C_3 < C_2$ なる条件を満足する値 C_3 を用意して，式(3.20)を満たす C_3 の探索を行う．以下同様の手順で C_{M-1} までの値を決定していく． C_{M-1} までの全ての値が固定値と見なされた段階で，最後の帯域である $\bar{P}_x(C_M)$ の値が，式(3.21)にあるように2次元周波数平面において，全領域 Ω から C_{M-1} まで

に考慮されている全ての帯域を差し引いた残りの部分についての演算により必然的に決定する。このとき式(3.20)の関係を満足しなければ、初期値である C_1 を再度選び直し、上記の解決定アルゴリズムを繰り返す。本アルゴリズムを簡略化したフローチャートの形で表したものを図 3.3 に示す。

以降の 3.4 節及び 3.6 節において導入される画像モデルのように、その電力スペクトルが単峰性を有し、領域 Ω において単調減少するような入力信号に対しては、 C_1 に対して各 C_k ($k=2,3,\dots,M-1$) は単調減少するので、演算を反復すれば常に解は収束し、最適境界値 $\{C_k^{(opt)}\}$ が求められる。

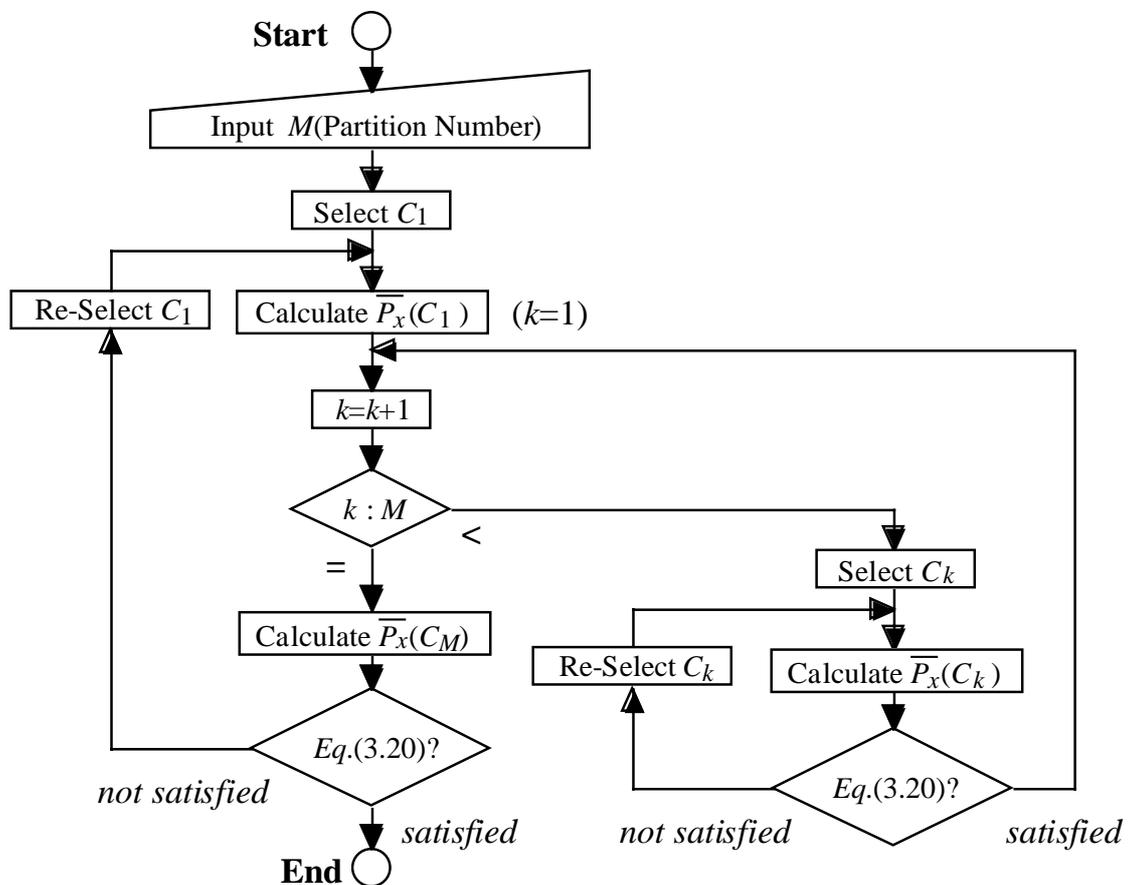


図 3.3 最適帯域分割の算出アルゴリズム

3.4 水平・垂直相関分離型画像モデルに対する 最適帯域分割

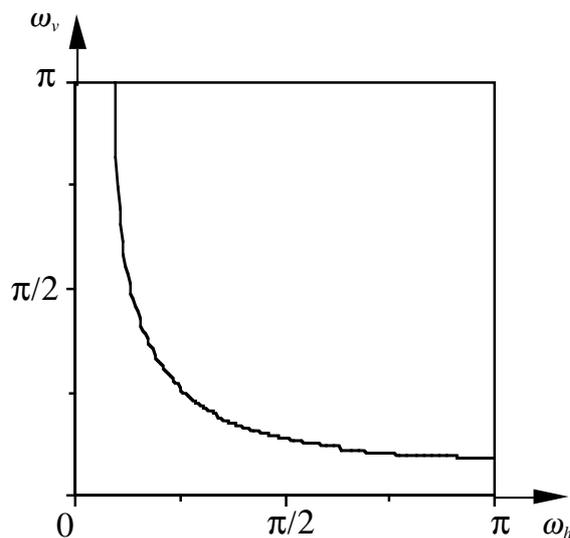
本研究において対象とする2次元信号系列は画像信号であり、本信号はその多くの区間で隣接する標本値間において高い相関を有する。そこで、まず、従来から用いられている最も単純な画像モデルとして、2次元自己相関関数の水平、垂直方向間が独立であるとした1次元系列の積の形で表される

$$r(m,n) = \sigma^2 \rho^{|m|+|n|} \quad (3.26)$$

なる水平・垂直相関分離型画像モデルを導入する[49]。ここで、 σ^2 は入力信号電力を、 ρ は入力信号の隣接標本値間の相関係数値を表し、ここでは、水平、垂直共に同じ値を有するものと仮定する[49],[51]。文献[49]では、画像のモデルとして、この信号の規格化された電力スペクトル $P_x(\omega_h, \omega_v)$ が次式で与えられている。

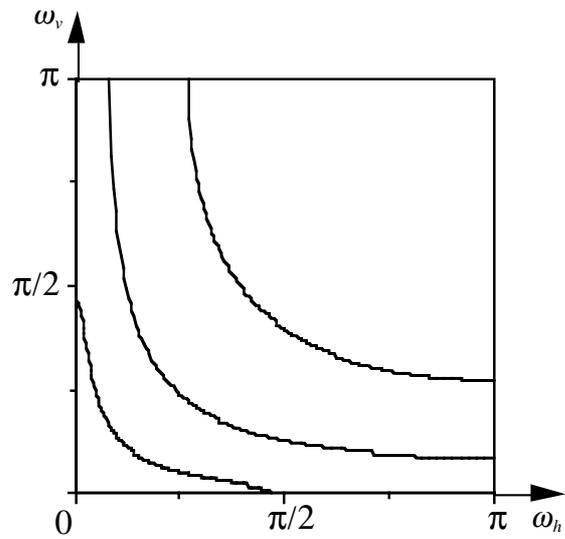
$$P_x(\omega_h, \omega_v) = \frac{1}{1 + \frac{4\rho}{(1-\rho)^2} \sin^2 \frac{\omega_h}{2}} \cdot \frac{1}{1 + \frac{4\rho}{(1-\rho)^2} \sin^2 \frac{\omega_v}{2}} \quad (3.27)$$

明らかに、この信号は領域 Ω において単峰性を有し、 $P_x(0,0)$ をピークとして単調減少するから、この信号に対して、前節に述べた最適帯域分割算出アルゴリズムを適用することができる。例えば、 $\rho=0.9$ としたとき[49],[51]、分割数 M をそれぞれ2、4、及び8とした場合の最適境界値 $\{C_k^{(opt)}\}$ と最適ビット配分 $\{R_k^{(opt)}-V\}$ ($k=0,1,\dots,M$) の値を表3.1に示す。また、各帯域分割数毎に、 $\{C_k^{(opt)}\}$ に対応する2次元周波数平面上の最適帯域分割パターンを図3.4に示す。

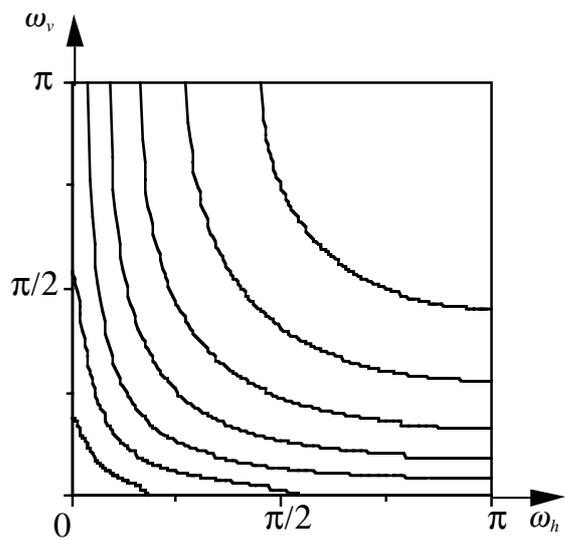


(a) 分割数 $M=2$

図3.4 2次元周波数平面上での最適帯域分割パターン
(水平・垂直相関分離型画像モデル, $\rho=0.9$)



(b) 分割数 $M=4$



(c) 分割数 $M=8$

図 3.4 2次元周波数平面上での最適帯域分割パターン
(水平・垂直相関分離型画像モデル, $\rho=0.9$)

表 3.1 最適帯域分割における電力スペクトルの境界値 $\{C_k^{(opt)}\}$ 及び最適ビット配分値 $\{R_k^{(opt)}\}$ (水平・垂直相関分離型画像モデル, $\rho=0.9$)

(a) 分割数 $M=2$

k	$C_k^{(opt)}$	$R_k^{(opt)} - V$ (bit)
0	1.000×10^0	—
1	3.173×10^{-4}	2.737750
2	7.673×10^{-6}	-0.967271

(b) 分割数 $M=4$

k	$C_k^{(opt)}$	$R_k^{(opt)} - V$ (bit)
0	1.000×10^0	—
1	6.089×10^{-3}	4.373449
2	3.841×10^{-4}	1.897247
3	4.510×10^{-5}	0.093980
4	7.673×10^{-6}	-1.325774

(c) 分割数 $M=8$

k	$C_k^{(opt)}$	$R_k^{(opt)} - V$ (bit)
0	1.000×10^0	—
1	3.260×10^{-2}	5.234628
2	4.883×10^{-3}	3.405831
3	1.182×10^{-3}	2.273370
4	3.444×10^{-4}	1.317706
5	1.151×10^{-4}	0.471213
6	4.307×10^{-5}	-0.280630
7	1.758×10^{-5}	-0.980977
8	7.673×10^{-6}	-1.571931

3.5 水平・垂直相関分離型理論画像モデルにおける評価とその結果

本最適帯域分割の性能を評価するために,1次元の場合と同様に,信号系列を直接スカラ量子化した場合(PCM符号化)に生じる量子化雑音電力 N_{PCM} と帯域分割処理後の各帯域信号にスカラ量子化を適用した場合に生じる全量子化雑音電力 N との比によって定義される量子化雑音改善量 G を評価尺度に用いる.この G は,帯域分割/合成時にサンプリングレート及び信号電力が保存される場合に,次式で与えられる.

$$G = \frac{N_{PCM}}{N} = \frac{\sum_{k=1}^M \sigma_k^2}{\prod_{k=1}^M \left(\frac{\sigma_k^2}{\lambda_k} \right)^{\lambda_k}} \quad (3.28)$$

ここで,上式の λ_k 及び σ_k^2 は,それぞれ式(3.2) (3.6)で示された帯域分割フィルタの出力レート及び各帯域の信号電力である.

3.5.1 帯域分割による改善量 G

まず,式(3.20)によって求められた最適境界値 $\{C_k^{(opt)}\}$ を式(3.22) (3.23)に代入することにより求められた σ_k^2, λ_k の値を用いることで, G_{opt} を定義することができる. G_{opt} は,最適帯域分割を行い,その後の各帯域信号に最適ビット割当てを適用した時の量子化雑音改善量である.次に,最適帯域分割との比較を行うために,フィルタバンクの実現が可能である3種類の既存の帯域分割方式を例に挙げ,各々の方式の帯域分割パターンを図3.5~図3.7までに示す.ただし,帯域分割パターンの違いについてのみ評価を行うため,ここでは既存の帯域分割方式実現に用いられる分割フィルタは理想的であると仮定する.図3.5,図3.6の帯域分割は,1次元フィルタを2次元空間上で互いに直交する水平及び垂直方向に適用する separable フィルタバンクを用いて実現される.separable フィルタバンクの適用においては,図3.5のように帯域を均等に分割する linear 分割[16]と,図3.6のように帯域を対数スケール上で均等に分割する log-linear 分割[57],いわゆる wavelet 変換を比較の対象とする.しかし,separable フィルタバンクでは,帯域分割の形状が長方形に限られてしまう欠点があること,更には人間の視覚特性が空間上の斜め方向の成分に対して比較的敏感でないこと[55]を考慮して,2次元のサブバンド符号化に対して2次元 non-separable フィルタバンクの適用が検討され始めている[47], [48].上記の文献によれば,分割フィルタとして2次元 SSKF (Symmetric Short Kernel Filter) [19]を用いて, wavelet 変換を行う場合と同様にツリー状のフィルタバンクを構成することで 2次元周波数平

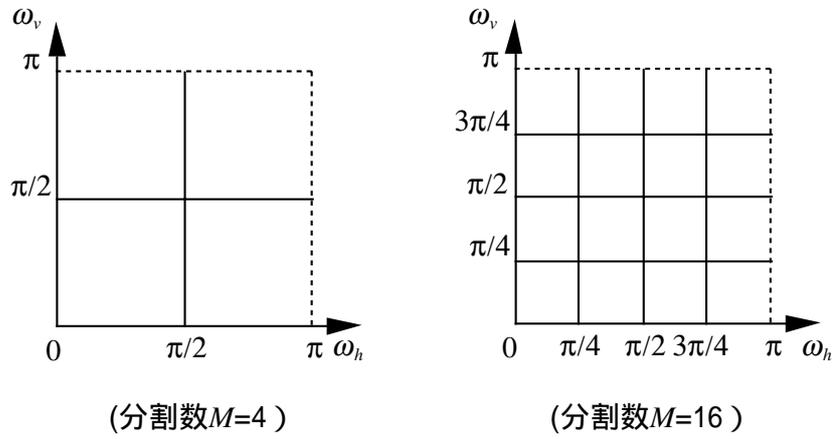


図 3.5 2次元 linear 分割

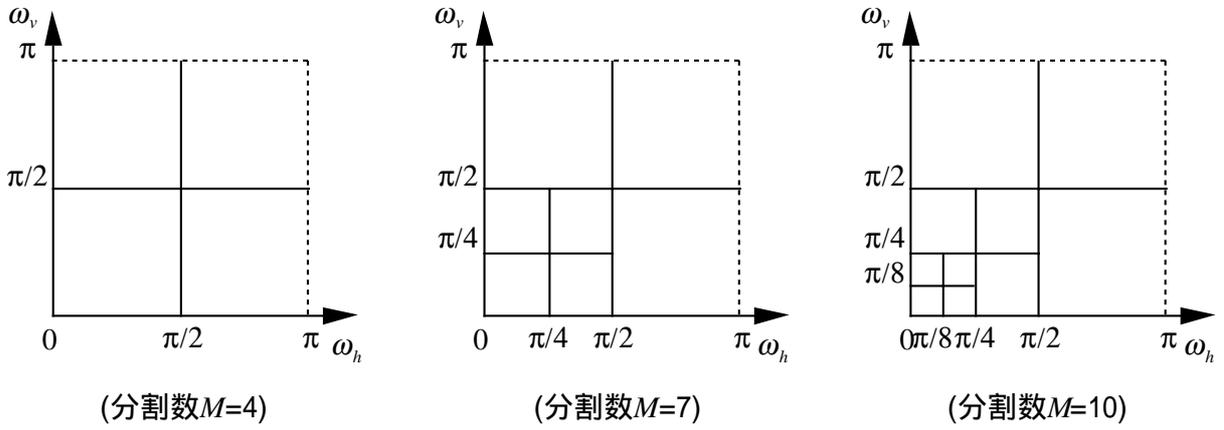


図 3.6 2次元 log-linear 分割

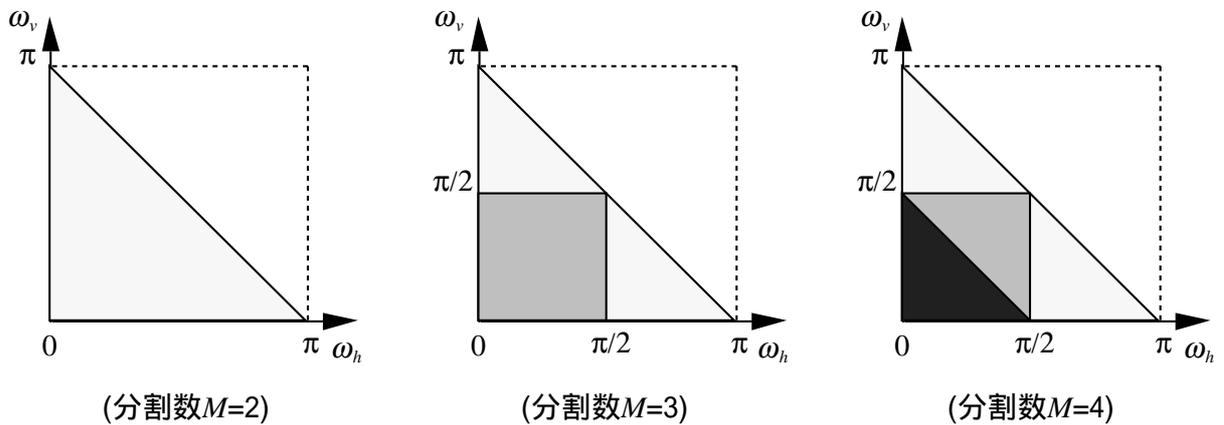


図 3.7 2次元 nonseparable フィルタによる帯域分割

面を両周波数軸に対して斜めに帯域を分割し，なおかつ低域のみを再帰的に分割する図 3.7 に示す帯域分割が可能である．

以上の linear 分割，log-linear 分割及び 2 次元 non-separable フィルタによる 3 種類の既存の帯域分割方式について，それぞれ帯域分割後に最適なビットが割り当てられるとした場合の量子化雑音改善量を G_{lin} ， G_{log} 及び G_{non} と表すものとする．

3.5.2 DCT による改善量 G

入力信号を $(M \times M)$ 画素毎にブロック化して DCT を適用する場合の量子化雑音改善量を $G_{DCT(M)}$ とすれば，それは， N_{PCM} と DCT を適用した場合の量子化雑音電力 N_{DCT} の比である

$$G_{DCT(M)} = \frac{N_{PCM}}{N_{DCT}} = \frac{\frac{1}{M^2} \sum_{u=0}^{M-1} \sum_{v=0}^{M-1} \sigma_{uv}^2}{\left(\prod_{u=0}^{M-1} \prod_{v=0}^{M-1} \sigma_{uv}^2 \right)^{1/M^2}} \quad (3.29)$$

として定義される．上式において，各変換係数電力 σ_{uv}^2 は，2 次元 DCT の基底が

$$a_{uv}(m, n) = \frac{2}{M} C(u)C(v) \cos \frac{(2m+1)u\pi}{2M} \cos \frac{(2n+1)v\pi}{2M} \quad (3.30)$$

ただし，

$$C(p) = \begin{cases} 1/\sqrt{2} & , p = 0 \\ 1 & , p \neq 0 \end{cases}$$

であるとき [37]，水平・垂直相関分離型画像モデルの適用において，入力信号の自己相関関数が式 (3.26) で与えられた場合には

$$\sigma_{uv}^2 = \sigma^2 \sum_{m=0}^{M-1} \sum_{n=0}^{M-1} \sum_{m'=0}^{M-1} \sum_{n'=0}^{M-1} a_{uv}(m, n) a_{uv}(m', n') \rho^{|m-m'|+|n-n'|} \quad (3.31)$$

によって求めることができる．

3.5.3 改善量 G の理論限界

2 次元不規則系列 $x(m, n)$ の自己相関関数を $r(m, n)$ ，電力スペクトルを $P_x(\omega_h, \omega_v)$ とするとき， $E[x(m, n)] = 0$ を仮定すれば， $x(m, n)$ の電力 σ^2 は，次式で表される．

$$\sigma^2 = r(0, 0) = \frac{1}{(2\pi)^2} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} P_x(\omega_h, \omega_v) d\omega_h d\omega_v \quad (3.32)$$

1次元の場合と同様にして，系列 $x(m,n)$ に対して無限次数の最適2次元線形予測フィルタを適用したときの予測誤差 $e(m,n)$ の電力を σ_e^2 とすれば，2次元の量子化雑音改善量 G の理論限界値 G_{lim} は，信号電力 σ^2 と予測誤差電力 σ_e^2 の比として

$$G_{lim} = \frac{\sigma^2}{\sigma_e^2} = \frac{r(0,0)}{\sigma_e^2} \quad (3.33)$$

によって定義される．また，最適2次元線形予測フィルタの最適予測係数を $a_{k,l}$ とすれば， $P_x(\omega_h, \omega_v)$ について

$$P_x(\omega_h, \omega_v) = \frac{\sigma_e^2}{\left| 1 - \sum_{\substack{k,l=0 \\ (k,l) \neq (0,0)}}^{\infty} a_{k,l} e^{-j(k\omega_h + l\omega_v)} \right|^2} \quad (3.34)$$

なる関係が成立する．

ここで，系列 $x(m,n)$ の自己相関関数 $r(m,n)$ が与えられると，Wiener-Khinchineの関係から電力スペクトル $P_x(\omega_h, \omega_v)$ が求められ，それを式(3.34)のような形式で表現したときの係数 $\sigma_{e,min}^2$ を求め，それを式(3.33)に代入することで G_{lim} が求められる．

例えば，3.4節で導入された相関係数値が ρ である水平・垂直相関分離型画像モデルの場合，その電力スペクトルは，

$$P_x(\omega_h, \omega_v) = \sum_{m,n=-\infty}^{\infty} \rho^{|m|+|n|} e^{-j(m\omega_h + n\omega_v)} = \frac{(1-\rho^2)^2}{\left| 1 - \rho e^{-j\omega_h} - \rho e^{-j\omega_v} + \rho^2 e^{-j(\omega_h + \omega_v)} \right|^2} \quad (3.35)$$

と表すことができるから， $\sigma_{e,min}^2 = (1-\rho^2)^2$ となることが分かり，また， $r(0,0)=1$ であるから，本画像モデルにおける理論限界値 G_{lim} は，

$$G_{lim} = \frac{1}{(1-\rho^2)^2} \quad (3.36)$$

として得ることができる．

3.5.4 評価結果

まず，最適帯域分割に対する評価を行う． $M=2, 4$ 及び8の場合に，入力信号系列の相関係数値 ρ に対する G_{opt} [dB]の特性を図3.8に示す．図3.8は，2次元信号系列に対する帯域分割処理による量子化雑音電力改善の理論限界を明らかにしたものである．また，同図には，本理論画像モデルにおける量子化雑音改善量の理論限界値である G_{lim} の特性を併せて表示している．この結果から，信号系列の相関係数値 ρ が1.0に近くなるほど，最適帯域分割による量子化雑音改善の効果が大きくなっており，特に画像信号をグローバル

に近似する $\rho = 0.9$ の信号系列に対する最適帯域分割による改善量は, $M=4$ の場合, 約 14 [dB] である. 更に, $M=4, 8$ における改善量の差がほとんどないこと, 及びこれらの場合の量子化雑音改善量特性と G_{lim} の特性との差は僅少であることに加え, 実際に最適帯域分割を実現するには帯域分割数はできるだけ少ない方が望ましいことを考慮して, 本最適帯域分割に必要なとされる帯域分割数は 4 で十分であると結論づけることができる.

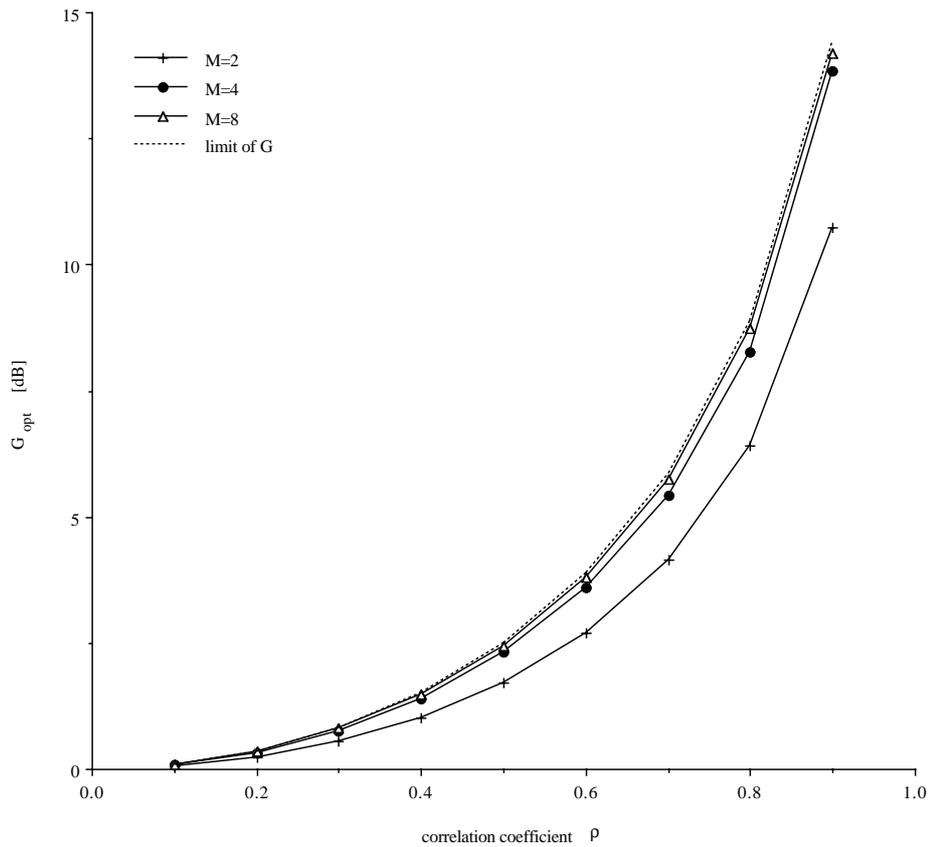


図 3.8 最適帯域分割の量子化雑音改善量 G_{opt} (パラメータ: 相関係数値 ρ)

次に, 従来から用いられている既存の帯域分割方式について, それらの量子化雑音改善量特性を最適帯域分割と比較することで, 既存の帯域分割方式による量子化雑音電力改善の性能を確認することができる. 図 3.5 ~ 図 3.7 までに示された linear 分割, log-linear 分割及び 2 次元 non-separable フィルタによる各帯域分割数 M に対する量子化雑音改善量特性を示した結果が図 3.9 である. このときの入力信号系列は, 式(3.27)の水平・垂直相関分離型画像モデルを用いて, その相関係数値 $\rho = 0.9$ としたものである[49],[51]. なお, 帯域分割数 $M=1$ の場合は, 当然のことながら各分割方式共に, 改善量は 0 [dB] である. 図 3.9 においては, 同信号系列における最適帯域分割による改善量 G_{opt} 及び 3.5.2 節に述べた $(M \times M)$ 画素単位に適用される DCT による改善量 $G_{DCT(M)}$, 更に 3.5.3 節に述べた本画像モデルにおける改善量の理論限界値 G_{lim} の値を併せて表示する.

この結果から， $M \leq 16$ の範囲内では，最適帯域分割に対するその他の既存の帯域分割方式の改善量は，separable フィルタを用いた場合には4 [dB]程度，non-separable フィルタを用いた場合には3 [dB]程度劣っていることが分かる．また，同じ M の範囲においては，最適帯域分割を含めた各帯域分割方式共に，分割数を4以上に増加していく場合の改善量の上昇はかなり小さくなり，特性が飽和傾向にあることが分かる．このことは，最適帯域分割については，その改善量の値が理論限界値 G_{lim} に近接していることから，帯域分割処理によっては理論的にこれ以上の改善が不可能であることを示唆しているが，既存の帯域分割方式については，理論限界値との差が未だ大きいため，各帯域毎に理想的な改善が行われていないことが推測される．特に，log-linear 分割の場合，低域のみを再帰的に分割していただくだけでは，その信号電力は小さいものの，帯域の面積が大きい高域部分での量子化雑音の影響を受け，全体として帯域分割による量子化雑音電力の抑圧がうまく行われていないと考えられる．その推測を根拠づけるものとして，帯域分割数 $M=16$ のときには，低域のみならず高域部分に対しても帯域分割を行うlinear 分割の方が，log-linear 分割よりも大きな改善の効果を得られている．non-separable フィルタによる帯域分割については， $M \leq 16$ の範囲においては，いずれの帯域分割数の場合においても separable フィルタの1つである log-linear 分割の場合よりも改善量が大きいことがわかる．ただし，non-separable フィルタについても分割数が4以上のところで改善量の上昇が見られないことの原因としては，non-separable フィルタは，図3.7に示されるように log-linear 分割の場合と同様に低域部分にのみ分割を繰り返すことで，帯域分割処理という意味で無視をした高域部分に含まれる量子化雑音電力の影響を受け，全体としてよい改善が行われなかったためであると考えられる．

JPEG で用いられている DCT との比較においては，最適帯域分割を適用することで，その帯域分割数が4である場合に，(8×8)画素のブロック単位での DCT，すなわち，64 分割の DCT の改善量特性を十分上回ることができる．既存の帯域分割方式との比較においても， $M \leq 16$ の範囲では，いずれの既存の帯域分割方式についても帯域分割数が同じであるという条件のもとでの DCT の改善量値を上回っていることが確認できる．このことは，DCT に比べてサブバンド分割の量子化雑音電力低減効果が優れていることを示唆しており，更に，その場合に適切な帯域分割，すなわち，最適帯域分割を行うことにより，最も効果的な量子化雑音電力の低減が実現できることになる．

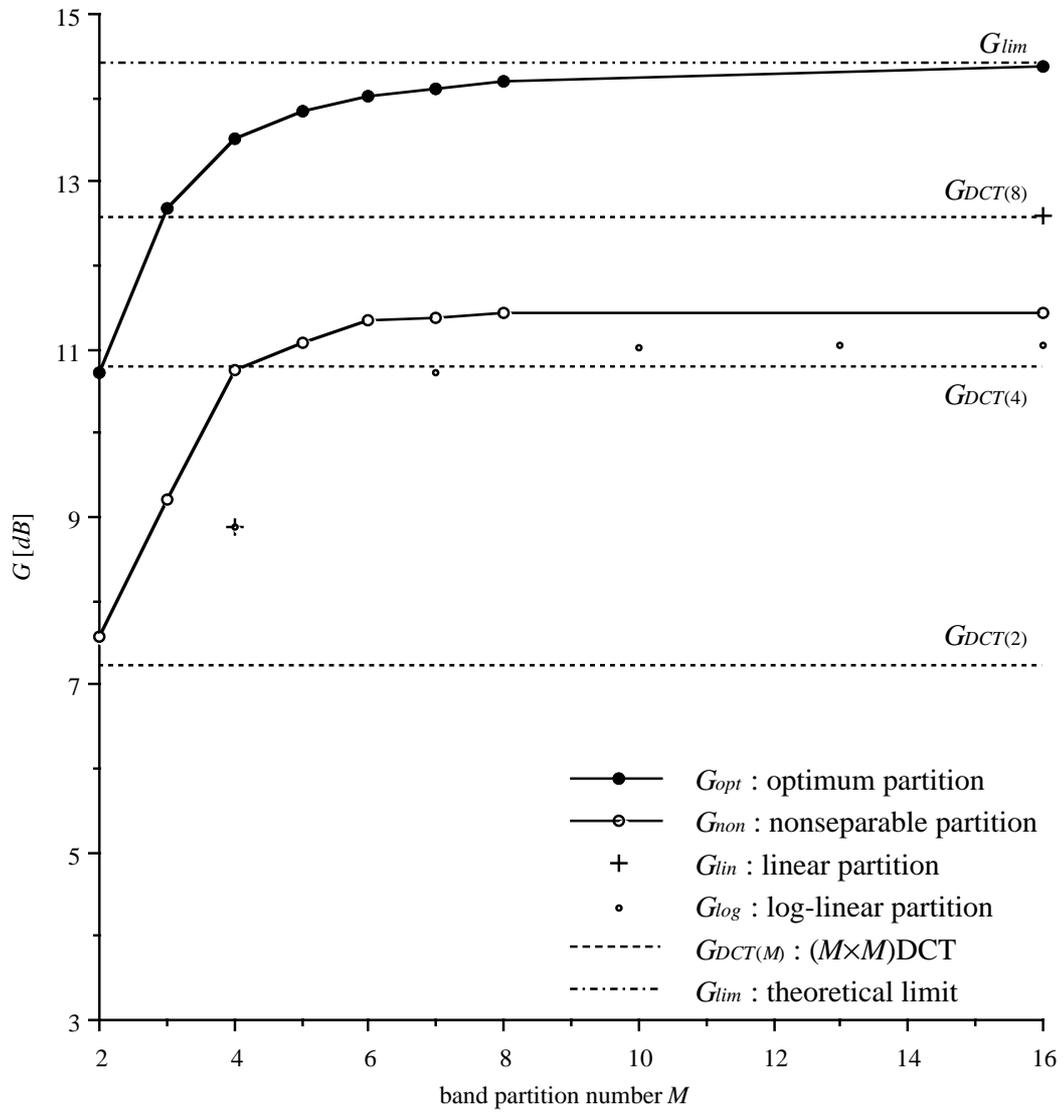


図 3.9 各帯域分割方式の量子化雑音改善量 G
 (水平・垂直分離型画像モデル, $\rho=0.9$)(パラメータ: 帯域分割数 M)

3.6 水平・垂直相関非分離型画像モデルに対する最適帯域分割

3.6.1 実画像の自己相関特性

3.4節において画像のモデルとして導入された水平・垂直相関分離型画像モデルは、建造物（特に、ビルディング）等のような水平及び垂直成分を多く含む特別な場合に、その自己相関特性をよりよく近似できると言われている[49],[51]。しかし、一般的な画像の場合には、それらの自己相関特性は、水平、垂直方向間で独立ではない場合が多く、且つ、それは等方性をもち、空間的距離の関数である

$$r(m,n) = \sigma^2 \rho^{\sqrt{m^2+n^2}} \quad (3.37)$$

なる等方性相関モデルがより汎用的であり、一般の画像に適合するとされている[49],[51]。ここで、 σ^2 は入力信号電力を、 ρ は入力信号の隣接標本値間の相関係数値を表している。このことを確認し、最終的により汎用性のある画像モデルを導入することを目的として、図3.10に示す12枚の画像に対して、2次元の自己相関関数特性を求める。図3.10に示された画像は、無数にある画像の代表として比較的多くの場面で用いられている標準画像の中から、人物(顔)の画像として、図3.10(a)~(f)に示した"Hada", "Hairband", "Wether", "Girl", "Lenna", "Barbara"の6枚を、風景(建造物)の画像として、図3.10(g)~(j)に示した"Chrch", "Tulip", "Cameraman", "Building"の4枚を、その他の種類のものには、静物の画像として、図3.10(k)に示した"Wine"と、特殊な画像として、図3.10(l)に示した月面の画像"Moon"の2枚を、汎用性の向上という観点から抜粋したものである^{注1}。なお、全ての画像のサイズは(256×256)画素、また、濃淡8ビット/画素である。図3.11は、図3.10に示された各画像を2次元の信号系列と見なし、2次元自己相関関数の等相関曲線を求めた結果である。図3.11において、横軸及び縦軸は、それぞれ水平及び垂直方向の画素数に換算した距離を、図中の数値は、各等相関曲線の相関係数値を表している。

以上の結果において、画像の自己相関関数は、水平、垂直方向間で独立ではなく、多くの場合については2次元空間領域で、距離に対してほぼ円形の分布特性を示すことから、画像に対する相関モデルを等方性相関モデルで近似することが望ましく、一般的な画像のモデルとして、式(3.37)に示した相関関数を有する水平・垂直相関非分離型画像モデルがより適切であることを確認することができる。次節以降、画像モデルとして、ここで導入した水平・垂直相関非分離型画像モデルを改めて適用し、その場合の最適帯域分割パターンを明らかにすると共に、量子化雑音改善量特性に対する評価を行う。

^{注1}：図3.10(a)"Hada", 図3.10(b)"Hairband", 図3.10(c)"Wether", 図3.10(g)"Church", 図3.10(h)"Tulip"は、ITEデジタル標準画像 ver.II[58]のものである。また、図3.10(d)"Girl", 図3.10(l)"Moon"は、SIDBA標準画像[59]のものである。



(a) 画像 "Hada"



(b) 画像 "Hairband"



(c) 画像 "Wether"



(d) 画像 "Girl"

図 3.10 テスト画像 (256×256)画素 , 濃淡 8 ビット / 画素)



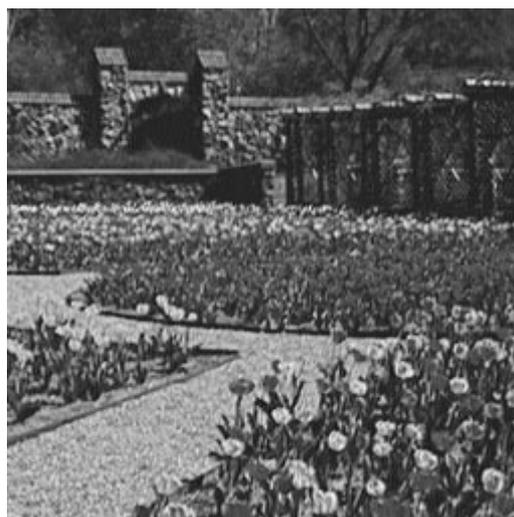
(e) 画像 "Lenna"



(f) 画像 "Barbara"



(g) 画像 "Church"



(h) 画像 "Tulip"

図 3.10 テスト画像 (256×256)画素, 濃淡 8 ビット / 画素)



(i) 画像 "Cameraman"



(j) 画像 "Building"

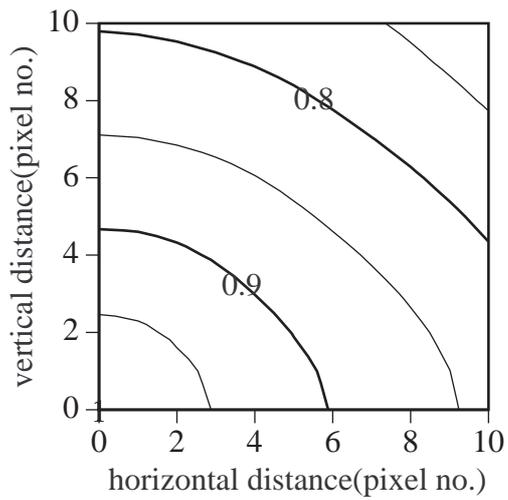


(k) 画像 "Wine"

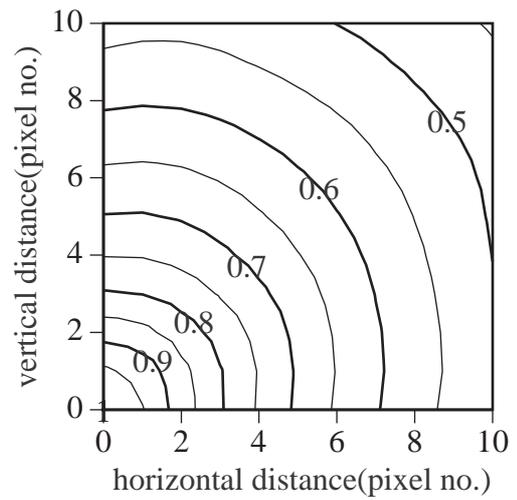


(l) 画像 "Moon"

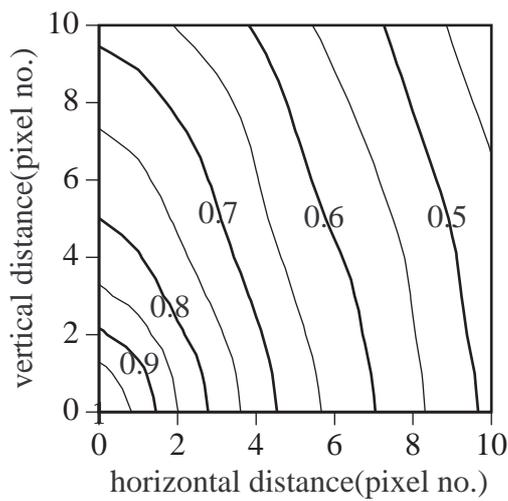
図 3.10 テスト画像 (256×256)画素, 濃淡 8 ビット / 画素)



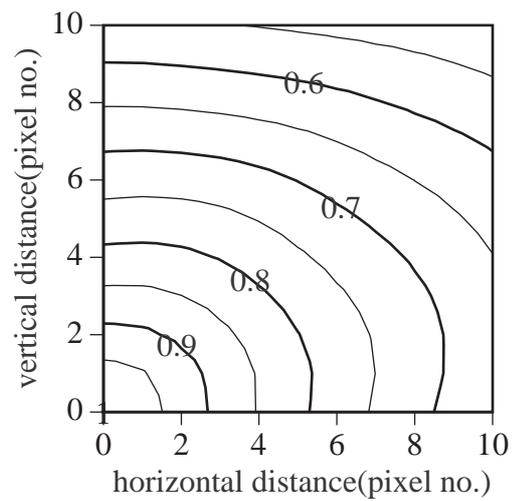
(a) 画像 "Hada"



(b) 画像 "Hairband"

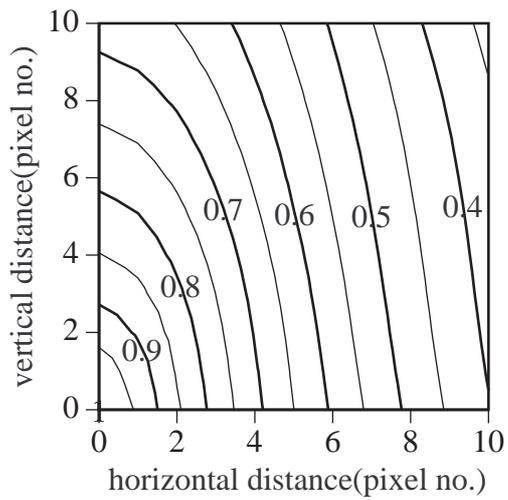


(c) 画像 "Wether"

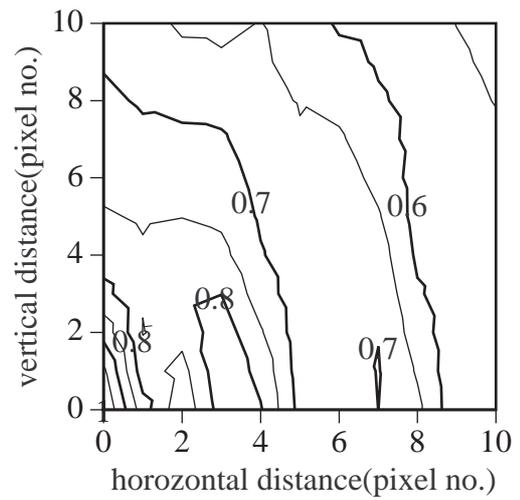


(d) 画像 "Girl"

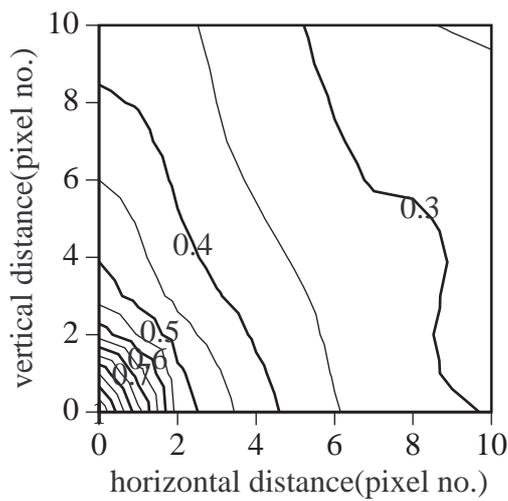
図 3.11 実画像における 2次元自己相関関数の等相関曲線



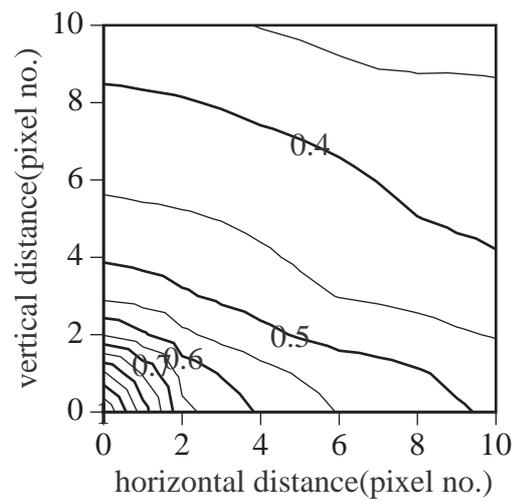
(e) 画像 "Lenna"



(f) 画像 "Barbara"

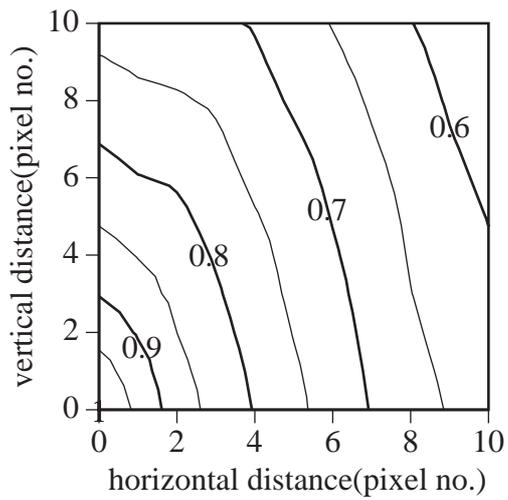


(g) 画像 "Church"

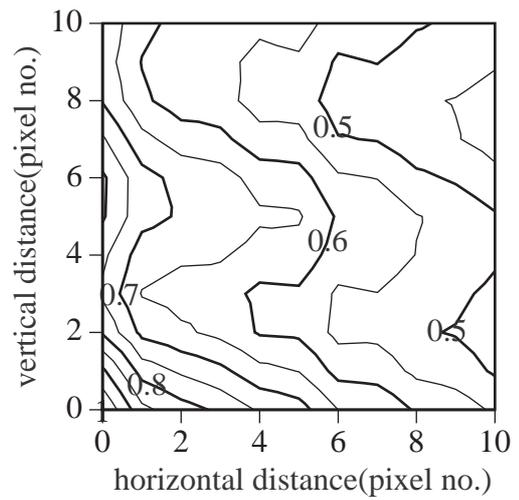


(h) 画像 "Tulip"

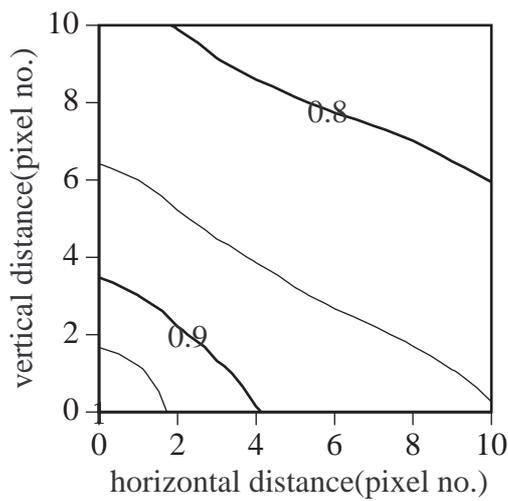
図 3.11 実画像における 2次元自己相関関数の等相関曲線



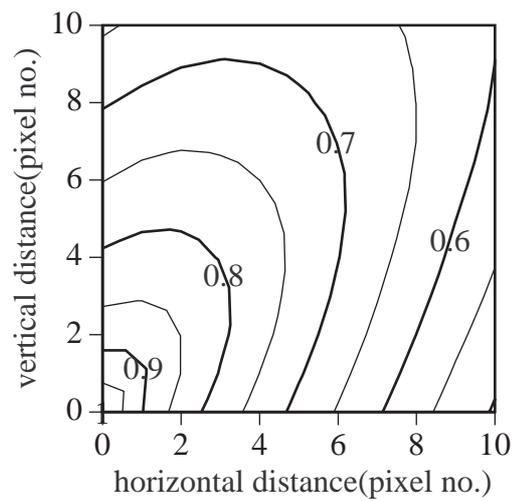
(i) 画像 "Cameraman"



(j) 画像 "Building"



(k) 画像 "Wine"



(l) 画像 "Moon"

図 3.11 実画像における 2次元自己相関関数の等相関曲線

3.6.2 最適帯域分割の導出

自己相関関数が式(3.37)で与えられる2次元信号系列の電力スペクトル $P_x(\omega_h, \omega_v)$ は, 最大値 $P_x(0,0)$ を1.0に正規化した場合に,

$$P_x(\omega_h, \omega_v) = \frac{\gamma^3}{[\gamma^2 + \omega_h^2 + \omega_v^2]^{3/2}} \quad (-\pi \leq \omega_h, \omega_v \leq \pi) \quad (3.38)$$

ただし,

$$\gamma = \ln(1/\rho) \quad (3.39)$$

と表すことができる。(付録B参照)

式(3.38)の $P_x(\omega_h, \omega_v)$ は, 領域 Ω において単峰性を有し, $P_x(0,0)$ をピークとして単調減少する。よって, 本画像モデルを対象として, 受信側での再構成信号に含まれる全量子化雑音電力を最小にする条件で最適帯域分割を導出するにあたり, 3.3.4節に述べた最適分割算出アルゴリズムを適用することができる。画像信号の統計的性質から, 一般的に式(3.37)の相関係数値 ρ は, 1.0に近い値をとることが知られている[51]。このことは, 図3.11の結果からも確認することができ, 各画像によってその値は異なるものの, 平均した場合に $\rho=0.9$ なる値を示している。そこで, 水平・垂直相関非分離型画像モデルの相関係数値を画像における平均的な値, $\rho=0.9$ に選んだ場合に, 式(3.20)及び式(3.38)を用いて, 最適境界値 $\{C_k^{(opt)}\}$ と最適ビット配分 $\{R_k^{(opt)}-V\}$ ($k=0,1,\dots,M$)の値を求めた結果を表3.2に示す。このとき, 帯域分割数 M は, それぞれ2, 4, 及び8とする。また, 各分割数毎に $\{C_k^{(opt)}\}$ に対応する2次元周波数平面上の最適帯域分割パターンを図3.12に示す。

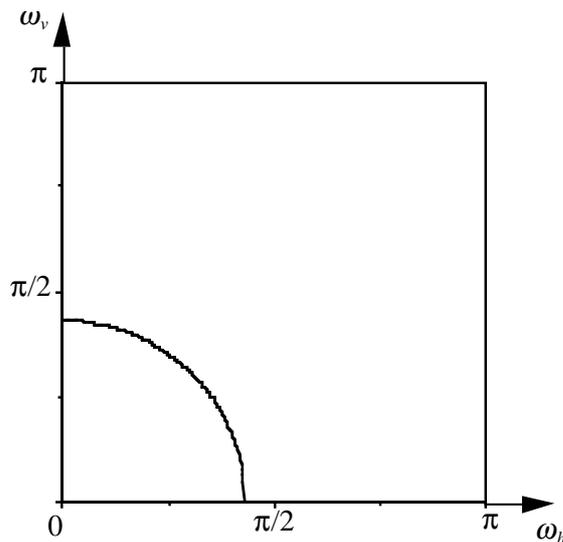
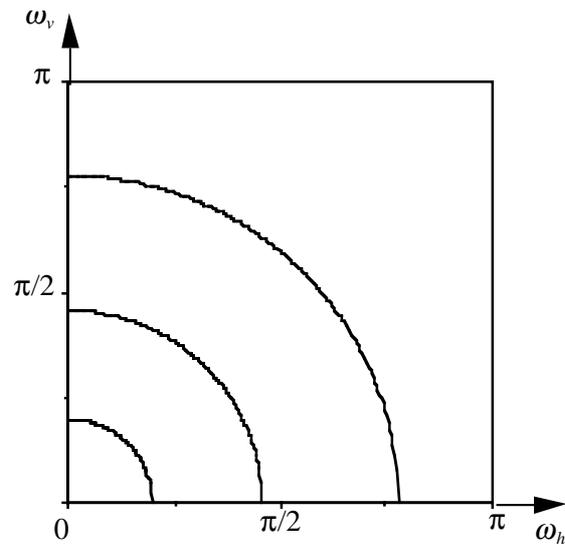
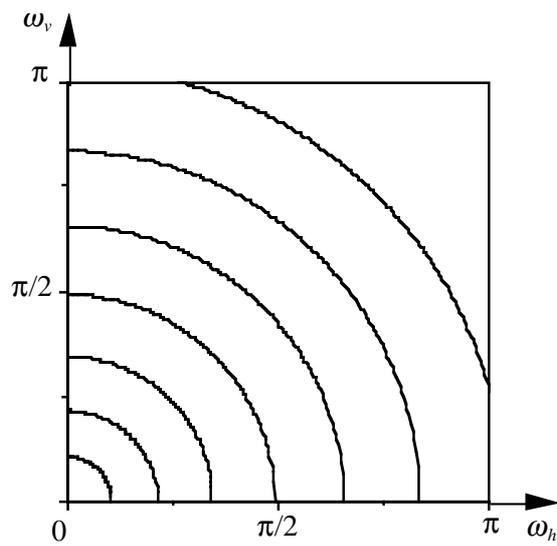


図 3.12 2次元周波数平面上での最適帯域分割パターン
(水平・垂直相関非分離型画像モデル, $\rho=0.9$)



(b) 分割数 $M = 4$



(c) 分割数 $M = 8$

図 3.12 2次元周波数平面上での最適帯域分割パターン
(水平・垂直相関非分離型画像モデル, $\rho=0.9$)

表 3.2 最適帯域分割における電力スペクトルの境界値 $\{C_k^{(opt)}\}$ 及び最適ビット配分値 $\{R_k^{(opt)}\}$ (水平・垂直相関非分離型画像モデル, $\rho=0.9$)

(a) 分割数 $M=2$

k	$C_k^{(opt)}$	$R_k^{(opt)} - V$ (bit)
0	1.000×10^0	—
1	4.699×10^{-4}	2.920566
2	1.333×10^{-5}	-0.496410

(b) 分割数 $M=4$

k	$C_k^{(opt)}$	$R_k^{(opt)} - V$ (bit)
0	1.000×10^0	—
1	4.692×10^{-3}	4.269850
2	3.951×10^{-4}	1.640076
3	7.980×10^{-5}	0.205017
4	1.333×10^{-5}	-0.786292

(c) 分割数 $M=8$

k	$C_k^{(opt)}$	$R_k^{(opt)} - V$ (bit)
0	1.000×10^0	—
1	2.732×10^{-2}	5.111477
2	3.663×10^{-3}	3.171352
3	9.120×10^{-4}	1.965675
4	3.151×10^{-4}	1.088821
5	1.334×10^{-4}	0.399232
6	6.478×10^{-5}	-0.169158
7	3.410×10^{-5}	-0.652627
8	1.333×10^{-5}	-1.098817

3.7 水平・垂直相関非分離型画像モデルにおける評価とその結果

水平・垂直相関非分離型画像モデルを入力信号系列に用いた場合について，最適帯域分割による量子化雑音低減効果を評価するために，3.5節で述べられた場合と同様に，式(3.28)に示される量子化雑音改善量 G を用いる．

3.7.1 帯域分割による改善量 G

最適帯域分割を行い，その後の各帯域信号に最適ビット割当てを適用した場合の量子化雑音改善量を G_{opt} で表すものとする．最適帯域分割との比較を行うために，それらのフィルタバンクが実現可能である既存の帯域分割方式として，図3.5に示した linear 分割，図 3.6 に示した log-linear 分割 (wavelet 変換)，そして図 3.7 に示した 2次元 non-separable フィルタによる3種類の帯域分割方式に対する評価を行う．各方式によって分割された帯域信号に最適ビット割当てが行われるとした場合の量子化雑音改善量をそれぞれ G_{lin} ， G_{log} ，そして G_{non} と表すものとする．

3.7.2 DCT による改善量 G

$(M \times M)$ 画素のブロック単位で DCT を適用した場合の量子化雑音改善量 $G_{DCT(M)}$ は，式(3.29)によって求めることができる．このとき，水平・垂直相関非分離型画像モデルとして，入力信号系列の自己相関関数が式(3.37)で与えられた場合には，式(3.29)の σ_{uv}^2 は，

$$\sigma_{uv}^2 = \sigma^2 \sum_{m=0}^{M-1} \sum_{n=0}^{M-1} \sum_{m'=0}^{M-1} \sum_{n'=0}^{M-1} a_{uv}(m,n) a_{uv}(m',n') \rho^{\sqrt{(m-m')^2 + (n-n')^2}} \quad (3.40)$$

によって求めることができる．ただし， $a_{uv}(m,n)$ は，式(3.30)に示された 2次元 DCT の基底である．

3.7.3 改善量 G の理論限界

量子化改善量の理論限界値 G_{lim} は，式(3.33)で定義される．しかし，ここで自己相関関数が式(3.37)で表される 2次元信号系列 $x(m,n)$ の生成モデルは不明である．勿論，無限次数の 2次元 AR モデルで表現することは可能であるが，その AR 係数 $a_{k,l}$ は簡単に求めることはできず，式(3.34)のような形式で表現することが困難であるため，理論限界値を求めるにあたって，3.5.3節に述べた方法を直接適用することができない．

そこで、この解決法として、今、信号系列の電力スペクトルが式(3.34)のような形式で表されるものとして、この式の両辺の対数をとって以下のように積分する。

$$\begin{aligned} \frac{1}{(2\pi)^2} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \ln[P_x(\omega_h, \omega_v)] d\omega_h d\omega_v &= \ln[\sigma_e^2] \\ &- \frac{1}{(2\pi)^2} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \ln \left[1 - \sum_{\substack{k,l=0 \\ (k,l) \neq (0,0)}}^{\infty} a_{k,l} e^{-j(k\omega_h + l\omega_v)} \right] d\omega_h d\omega_v \\ &- \frac{1}{(2\pi)^2} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \ln \left[1 - \sum_{\substack{k,l=0 \\ (k,l) \neq (0,0)}}^{\infty} a_{k,l} e^{+j(k\omega_h + l\omega_v)} \right] d\omega_h d\omega_v \end{aligned}$$

上式の右辺第2,第3項目の積分は、複素積分に変換することにより、 $x(m,n)$ が安定な(発散系列でない)系列であれば、全てゼロになることが証明できる。したがって、

$$\sigma_e^2 = \exp \left[\frac{1}{(2\pi)^2} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \ln[P_x(\omega_h, \omega_v)] d\omega_h d\omega_v \right] \quad (3.41)$$

を得る。式(3.41)を式(3.33)に代入することで、 G_{lim} は、

$$G_{lim} = \frac{r(0,0)}{\exp \left[\frac{1}{(2\pi)^2} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \ln[P_x(\omega_h, \omega_v)] d\omega_h d\omega_v \right]} \quad (3.42)$$

として求めることができる。ここで、式(3.38)を上式の $P_x(\omega_h, \omega_v)$ に代入して、

$$\begin{aligned} G_{lim} &= \frac{1}{2\pi\gamma} \exp \left[\frac{3}{2\pi^2} \int_0^{\pi} \int_0^{\pi} \ln[\omega_h^2 + \omega_v^2 + \gamma^2] d\omega_h d\omega_v \right] \\ &= \frac{1}{2\pi\gamma^2} \exp \left[\frac{3}{2\pi^2} \int_0^{\pi} \left[\pi \ln(x^2 + \pi^2 + \gamma^4) - 2\pi + 2\sqrt{x^2 + \gamma^4} \tan^{-1} \left(\frac{\pi}{\sqrt{x^2 + \gamma^4}} \right) \right] dx \right] \end{aligned} \quad (3.43)$$

となり、水平・垂直相関非分離型画像モデルにおける理論限界値 G_{lim} は、上式の数値積分によって求めることができる。ただし、式(3.43)中の γ は、式(3.39)に示されたものである。

3.7.4 評価結果

水平・垂直相関非分離型画像モデルを対象として、帯域分割数 M に対する最適帯域分割の改善量 G_{opt} の特性を既存の帯域分割方式における改善量特性である G_{lin} 、 G_{log} 、 G_{non} 及び $(M \times M)$ 画素の DCT における改善量特性 $G_{DCT(M)}$ と比較した結果を図 3.13 に示す。また、図 3.13 には、本画像モデルに対しての改善量の理論限界値である G_{lim} の値も併せて表示する。このとき、本画像モデルにおける隣接標本値間の相関係数値は、画像信号の統計的性質を考慮することにより、3.6.2 節と同様に $\rho=0.9$ を選択した。

この結果から、以下のことが分かる。最適帯域分割の改善量は、帯域分割数 M の値が大きくなるにつれて理論限界特性に近づき、 $M=16$ でほぼ理論限界値をとる。また、 $M=16$ の範囲内において、最適帯域分割は、既存の帯域分割方式に比べて 1 ~ 2[dB] 程度の改善があることがわかる。最適帯域分割の場合、及び既存の帯域分割方式の中の non-separable フィルタを用いた帯域分割方式の場合には、 $M \geq 4$ のところで量子化雑音改善量特性に飽和傾向が見られるため、符号化システム実現の簡易性までを考慮する場合においては、十分な改善量を得るための帯域分割数として、 $M=4$ が必要とされる。

また、最適帯域分割は、 (8×8) 画素の DCT に対しては 1 ~ 3[dB] 程度の改善が確認できる。DCT による改善量が、最適帯域分割、更には理論限界値 G_{lim} よりもかなり小さい理由としては、2次元の DCT は、式(3.30)にあるように1次元の DCT の組合せである separable モデルで記述されるため、このことに起因して、2次元の DCT は、入力信号として用いた等方性の相関特性を有する水平・垂直相関非分離型画像モデルに対してうまく適合できなかったと考えられる。このことは、既存の帯域分割方式の改善量特性の中で、non-separable フィルタを用いた場合の特性 G_{non} が最適帯域分割に最も近接した特性を示すことから考察することができる。

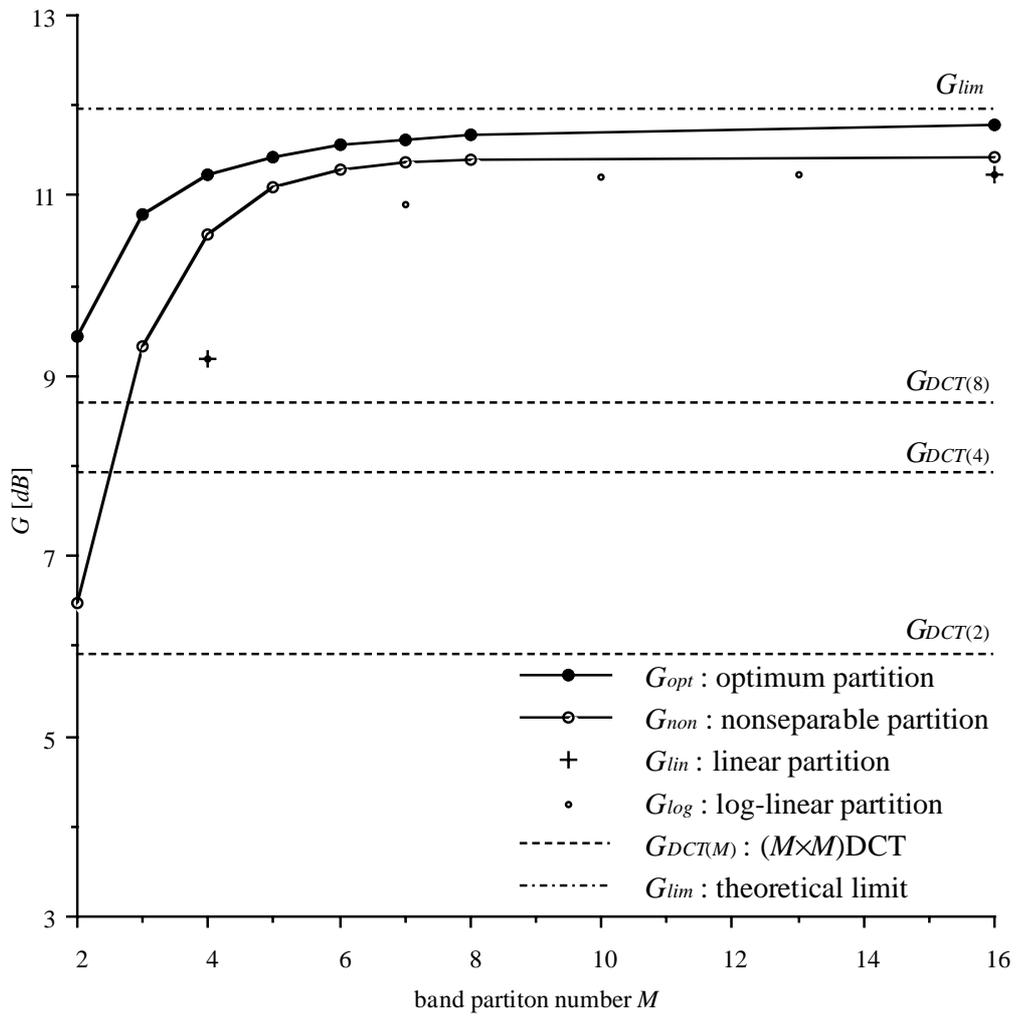


図 3.13 各帯域分割方式の量子化雑音改善量 G
 (水平・垂直非分離型画像モデル, $\rho=0.9$)(パラメータ: 帯域分割数 M)

3.8 結言

第2章で行われた1次元の信号系列に対する最適帯域分割の議論を2次元へと拡張することで、2次元信号系列に対する最適帯域分割の導出を理論的に行った。まず、2次元の帯域分割処理について、任意の帯域分割が可能であるフィルタバンク構成の理論モデルを提案し、分割フィルタの各帯域とその出力レートとの関係を示した。次に、分割後の各帯域信号がスカラー量子化されると仮定したとき、(1) 信号全体に割り当てられるビットレートが一定である、(2) 受信側での再構成信号に現れる量子化雑音電力を最小にする、という条件のもとで、2次元周波数平面における最適な帯域分割及び各帯域信号への最適ビット配分を定める理論式を導出すると共に、最適帯域分割を算出するための数値計算アルゴリズムを明らかにした。次に、2次元の入力信号系列として、水平・垂直相関分離型画像モデル、更には実際の画像信号の統計的性質を調査実験した結果に基づいて導入されたより一般的な画像モデルとして、水平・垂直相関非分離型画像モデルの2種類の画像モデルに対して、本アルゴリズムを適用することで、それぞれの場合の最適帯域分割に対応した帯域分割パターンを具体的に示した。

最適帯域分割を評価するために、2次元へと拡張された量子化雑音改善量を評価尺度に用いて、2次元の帯域分割を行う場合に従来から用いられている既存の方式として separable フィルタで実現される linear 分割及び log-linear 分割、そして2次元 non-separable フィルタで実現される帯域分割、また JPEG で用いられている空間領域の $(M \times M)$ 画素のブロック単位で適用される DCT との比較を行った。更に、最適帯域分割に対するより確かな性能評価を行うために、上記の2種類の画像モデルに対して、それぞれの場合の量子化雑音改善量の理論限界値を明らかにした。

この結果、水平、垂直成分が多い特殊な画像信号によく適合する水平・垂直相関分離型画像モデルに対しては、帯域分割数 $M \geq 4$ のところで、最適帯域分割の特性は、量子化雑音改善量の理論限界値をほぼ達成できることが明らかとなった。既存の帯域分割方式については、サブバンド符号化が画像符号化に適用される際に、一般的に用いられている separable フィルタによる帯域分割処理では、 $M=8$ の範囲では改善の効果が non-separable フィルタに比べて劣っていることが示され、特に、近年様々な分野で応用されている wavelet 変換に基づいた帯域分割方式については、低域成分のみを再分割の対象とし、高域成分に対しての処理を行わないことが、逆に量子化雑音電力改善の観点からは不利であることが示された。DCT との比較においては、サブバンド分割方式は、同じ帯域分割数の条件での DCT の特性を上回っているものの、JPEG で用いられている (8×8) 画素の DCT を用いた場合の改善量特性を上回るためには、既存の帯域分割方式を用いるならば、帯域分割数 $M \geq 16$ の linear 分割が必要であることが示された。これに対して、最適帯域分割を適用する場合には、同特性を上回るために必要な帯域分割数は、僅か $M=4$ でよいことが明らかになった。

次に，一般的な画像信号の大部分のものに対してより適切な近似を与える水平・垂直相関非分離型画像モデルに対しては，最適帯域分割が行われる場合，量子化雑音改善量の理論限界値をほぼ達成し得るに必要な帯域分割数は，水平・垂直相関分離型画像モデルの場合と同様に4程度であることが示された．また，既存の帯域分割方式の中では，non-separable フィルタを用いる帯域分割方式が，他の separable フィルタを用いる方式に比べてより大きな改善量を得ることができると示された．DCT に関しては，2次元 DCT の基底が，1次元 DCT の組合せである分 separable モデルで記述されることに起因して，本画像モデルを入力信号に仮定した場合には，最適帯域分割に比べて，その改善量特性が大きく劣っていることが示された．

本章において得られた結果は，量子化雑音電力低減の観点において，画像信号に対する2次元周波数平面上での理想的な帯域分割パターンを明らかにすると共に，帯域分割処理による量子化雑音電力改善の理論的限界を示したものである．更なる展開としては，理論解として導出された最適帯域分割特性を画像符号化システムへと適用するために，最適帯域分割を実現し得るフィルタバンクを構築する必要がある．また，その際，画像信号は非定常性を有しており，その局所的特徴を無視してはならない[51]．以上を考慮した上での最適帯域分割フィルタバンクの実現と，それを実際の画像信号への適用した場合の効果については，次章以降で述べられる．

付録 A：式(3.17)，式(3.18)の導出について

図 A.1 において，斜線部分 $\{(\omega_h, \omega_v) \mid z \sim P_x(\omega_h, \omega_v) \sim z + \Delta z\} \cap \Omega$ の面積は，式(3.15)に示した $s(z)$ を用いて，

$$\Delta s = s(z) - s(z + \Delta z) \cong -s'(z)\Delta z \quad (\Delta z \cong 0) \quad (\text{A.1})$$

と表される．したがって，この斜線部分を底面とし，高さ z の薄い柱の体積 ΔV は

$$\Delta V = \Delta s \cdot z \cong -s'(z)z\Delta z \quad (\text{A.2})$$

となる．

式(3.17)の左辺は，底面を Ω_k とする高さ $P_x(\omega_h, \omega_v)$ の柱状の体積を表すから，上式の薄い柱の体積を $z=C_k \sim C_{k-1}$ の範囲で総和をとり， $\Delta z \rightarrow 0$ とすれば，右辺の上側の積分式が導かれる．

なお， $k=M$ のときには，底面を Ω とする高さ $P_x(\omega_h, \omega_v)$ の立体の体積から， $k=1$ から $M-1$ 番目までの帯域に属する立体の体積を差し引くという考え方にに基づき，式(3.17)の右辺の下側の積分式を得る．

式(3.18)については， $P_x(\omega_h, \omega_v)$ を $\ln[P_x(\omega_h, \omega_v)]$ に置き換えた後，上述の議論を行うことで，同様に導出することができる．

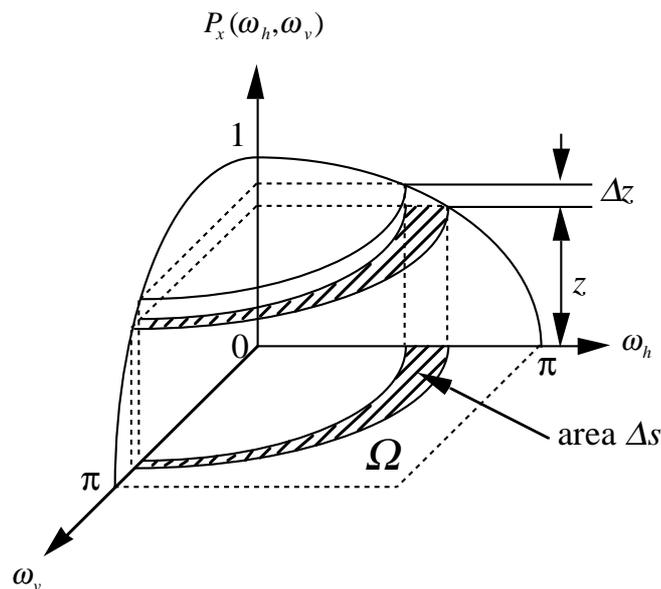


図 A.1 式(3.17)を導出するための説明図

(証明終)

付録 B：式(3.38)の導出について

離散系に対する Wiener-Khinchine の公式より，

$$P_x(\omega_h, \omega_v) = \sum_{m,n=-\infty}^{\infty} r(m,n) e^{-j(\omega_h m + \omega_v n)} \quad (\text{B.1})$$

を得る．式(B.1)は以下のように変形できる．

$$P_x(\omega_h, \omega_v) = \sum_{m,n=-\infty}^{\infty} P_0(\omega_h - 2\pi m, \omega_v - 2\pi n) \quad (\text{B.2})$$

ただし，

$$P_0(\omega_h, \omega_v) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} r(x,y) e^{-j(\omega_h x + \omega_v y)} dx dy$$

上式に式(3.37)を代入して，

$$P_0(\omega_h, \omega_v) = \sigma^2 \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-\gamma \sqrt{x^2 + y^2}} e^{-j(\omega_h x + \omega_v y)} dx dy$$

となる．ただし， γ は式(3.39)に示されたものである．ここで， $x=r\cos\theta$ ， $y=r\sin\theta$ とおけば，

$$\begin{aligned} &= \sigma^2 \int_{-\pi}^{\pi} \int_0^{\infty} e^{-[\gamma + \omega_h \cos\theta + \omega_v \sin\theta]r} r dr d\theta \\ &= \sigma^2 \int_{-\pi}^{\pi} \frac{d\theta}{[\gamma + j(\omega_h \cos\theta + \omega_v \sin\theta)]^2} \end{aligned}$$

を得る．さらに， $\omega_h = \lambda \cos\varphi$ ， $\omega_v = \lambda \sin\varphi$ とおけば，

$$\begin{aligned} &= \sigma^2 \int_{-\pi}^{\pi} \frac{d\theta}{[\gamma + j\lambda \cos(\theta - \varphi)]^2} \\ &= \frac{2\pi\sigma^2\gamma}{[\gamma^2 + \lambda^2]^{3/2}} = \frac{2\pi\sigma^2\gamma}{[\gamma^2 + \omega_h^2 + \omega_v^2]^{3/2}} \end{aligned} \quad (\text{B.3})$$

を得る[56]．

ここで， $\rho < 1.0$ かつ $\rho \approx 1.0$ より， $\gamma > 0$ で $\gamma \approx 0$ となるから， $P_0(\omega_h, \omega_v)$ の電力は $(\omega_h, \omega_v) = (0, 0)$ 付近に集中し，近似的に

$$P_0(\omega_h, \omega_v) \approx 0 \quad (|\omega_h| > \pi \cup |\omega_v| > \pi)$$

と見なせる．したがって，式(B.2) (B.3)より

$$P_x(\omega_h, \omega_v) \approx \frac{2\pi\sigma^2}{[\gamma^2 + \omega_h^2 + \omega_v^2]^{3/2}} \quad (-\pi \leq \omega_h, \omega_v \leq \pi) \quad (\text{B.4})$$

となり, $P_x(0,0) = 1.0$ となるように正規化を行えば,

$$\frac{2\pi\sigma^2}{\gamma^2} = 1$$

となるから, その結果, 式(3.38)を得ることができる.

(証明終)

第 4 章

画像信号のサブバンド符号化における 最適帯域分割の近似と実現

4. 1 序言

サブバンド符号化において、良い画質で高い圧縮効率を得るためには、まず、画像に整合した適切な帯域分割を行い、次に、分割された各帯域毎にその帯域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になる[18],[41],[60],[61]。したがって、処理コスト低減の観点において簡易化されたサブバンド符号化システムを実現するためには、できるだけ少ない帯域分割数が要求されるが、一般に、帯域分割数が少なくなると、符号化性能の評価尺度になる量子化雑音改善量は低下するという問題が生じる。

以上の観点から、前章までは、まず、最適帯域分割を理論的に導出することを目的に議論を進めてきた。第2章では、入力信号系列の統計的性質である自己相関関数に整合した、1次元信号系列における最適帯域分割を明らかにした。第3章では、2次元信号系列を対象として、(1)水平・垂直方向に多くの成分を有する特別な画像の場合に整合する水平・垂直方向の間の相関が独立であるとした水平・垂直相関分離型画像モデルに基づいた最適帯域分割、及び(2)多くの自然画像がそうであるように、水平・垂直方向の間の相関が独立ではなく、等方性を有する空間距離の関数で与えられたより汎用性のある一般的な画像モデルとして導入された水平・垂直相関非分離型画像モデルに基づいた最適帯域分割を明らかにした。

画像信号を対象とするサブバンド符号化の帯域分割としては、上記(2)の大部分の自然画像によく整合する水平・垂直相関非分離型画像モデルに基づく2次元周波数平面上での最適帯域分割が、少ない帯域分割数で最大の量子化雑音改善量を得る点で最適であるが、その帯域分割特性を実現する2次元フィルタバンクの実用的な構成が極めて困難であるという問題がある。

本章では、これまでに導出された最適帯域分割を理論的検討の範疇から実用面へ応用

することを目的として、実画像への適用を念頭においた高能率なサブバンド符号化システムの実現を行う。そのために、2次元最適帯域分割と比べて若干の量子化雑音改善量の低下を犠牲にして、帯域分割パターンに制限を加えることによりフィルタバンクの実現性を優先させた、周波数領域での基本ブロックとして定義される帯域ブロックをベースとした2次元最適帯域分割法を提案する。また、本最適帯域分割を実現するためのフィルタバンクの構成法を示し、実画像に適用してその有効性を明らかにする。

まず、第3章において導入された水平・垂直相関非分離型画像モデルを対象とした場合に、フィルタバンクの実現可能条件を考慮した帯域ブロックをベースとする最適帯域分割法を示すと共に、その最適帯域分割算出法を明らかにする。また、フィルタバンクを実用する上で、効果的な量子化雑音の改善が得られる帯域分割パラメータの選定を行い、これらの値に基づいて、本画像モデルを対象とした帯域ブロックをベースとする固定の最適帯域分割を明らかにする。次に、帯域分割フィルタとしてQMF[52],[53]を用いて、本最適帯域分割を実現するフィルタバンクの構成法を示し、それを実画像に適用して、量子化雑音改善効果を確認する。上述した画像モデルに基づき、最適化を行った帯域ブロックをベースとする固定な最適帯域分割に対して、更に、画像モデルでは考慮することのできない実画像特有の性質として、2次元空間領域の場所により変動する非定常直流成分、及び画像内容によって或いは空間領域の場所によって変動する相関特性に対処するために、前者に対しては、最低域の帯域分割信号に対してDFTを適用して非定常直流成分の分離を行う方法を、後者に対しては、帯域ブロックをベースとして、画像毎に固有な最適帯域分割を得る適応帯域分割法を提案する。そして、汎用性を高めることを考慮して選ばれた実画像数例について、本章で提案される固定及び適応最適帯域分割による量子化雑音改善量特性の評価実験を行い、多くの画像については、最適帯域分割を画像モデルにより求められた固定の最適帯域分割で代表することにより十分の改善が得られるが、本画像モデルでの近似が特に難しい画像については、適応帯域分割が有効であることを明らかにする。その際、たとえ適応帯域分割を行う場合においても、そのために要する処理コストの増分と伝送すべきサイド情報は小さくなることを示す。最後に、帯域ブロックをベースとする最適帯域分割による符号化性能を調べるために、SNRとエントロピーの関係を求め、JPEGで用いられている(8×8)画素のブロック単位で適用されるDCTと比較した場合の有効性を定量的に明らかにする。

4.2 帯域ブロックをベースとする最適帯域分割の近似

4.2.1 帯域ブロックをベースとする最適帯域分割法

2次元入力信号 $x(m,n)$ にサブバンド符号化を適用するときの最適帯域分割法として、入力信号の電力スペクトル $P_x(h', v')$ が与えられたとき (1)2次元帯域分割フィルタバンクは、任意の分割パターンを理想遮断特性で実現できる、(2)信号全体に割り当てられるビットレートが一定である、(3)受信側での再構成信号に含まれる量子化雑音電力を最小にする、という条件のもとで、2次元周波数帯域 $=\{(h', v') | 0 \leq h', v' \leq 1\}$ を、

$$= \sum_{k=0}^{M-1} \left(\int_{i_j}^{i_{j+1}} \int_{j_k}^{j_{k+1}} P_x(h', v') dh' dv' \right) \quad (4.1)$$

なる互いに素な M 個の帯域 k ($k=0, 1, \dots, M-1$) に分割する方法を第3章において詳述した。本最適帯域分割は、入力信号を直接スカラ量子化した場合 (PCM符号化) に生じる量子化雑音電力 N_{PCM} と帯域分割処理後にスカラ量子化した場合に生じる全量子化雑音電力 N の比として、次式で定義される量子化雑音改善量 G を最大にする分割と等価である^{注2}。

$$G = \frac{N_{PCM}}{N} = \frac{\sum_{k=0}^{M-1} \frac{2}{k}}{\sum_{k=0}^{M-1} \frac{2}{k}} \quad (4.2)$$

ただし、 $\frac{2}{k}$ は、それぞれ次式で表される帯域 k における分割フィルタの出力レートと帯域 k の出力電力である。

$$\begin{aligned} \frac{1}{k} &= \frac{1}{2} \int_{d_{h,k}}^{d_{h,k+1}} \int_{d_{v,k}}^{d_{v,k+1}} P_x(h', v') dh' dv' \\ \frac{2}{k} &= \frac{1}{2} \int_{d_{h,k}}^{d_{h,k+1}} \int_{d_{v,k}}^{d_{v,k+1}} P_x(h', v') dh' dv' \end{aligned} \quad (4.3)$$

今、直流成分を除いた大部分の自然画像信号に良く整合する自己相関関数として、信号電力を 1.0 に正規化した入力信号系列の隣接標本値間の相関係数値が $r(m,n)$ である

$$r(m,n) = \frac{1}{\sqrt{m^2+n^2}} \quad (4.4)$$

なる水平・垂直相関非分離型画像モデルを導入すれば、その電力スペクトルは、3.6.2節で述べたように、最大値 $P_x(0,0)$ を 1.0 に正規化した場合に、次式で与えられる。

^{注2}：前章の式(3.13)の R_d を最小とする k は、量子化雑音改善量 G を最大にする k と同一になることが容易に導かれる。

$$P_x(h, v) = \frac{3}{\left[2 + \frac{2}{h} + \frac{2}{v}\right]^{3/2}} \quad (-h, -v) \quad (4.5)$$

ただし，上式において

$$= \ln(1/ \quad) \quad (4.6)$$

である．この水平・垂直相関非分離型画像モデルにおいて，画像の統計的性質に基づいて選定された $\rho=0.9$ なる値を用いて，帯域分割数 $M=4$ とした場合に，その最適帯域分割を前章に示された理論式に基づいて求めれば，2次元周波数帯域の最適帯域分割パターンは図4.1のようになり，このような帯域分割特性を直接実現するフィルタバンクの構成は極めて困難であることが分かる．そこで，フィルタの実現可能条件を考慮して，帯域分割パターンに制限を加えたときの最適帯域分割を考える．

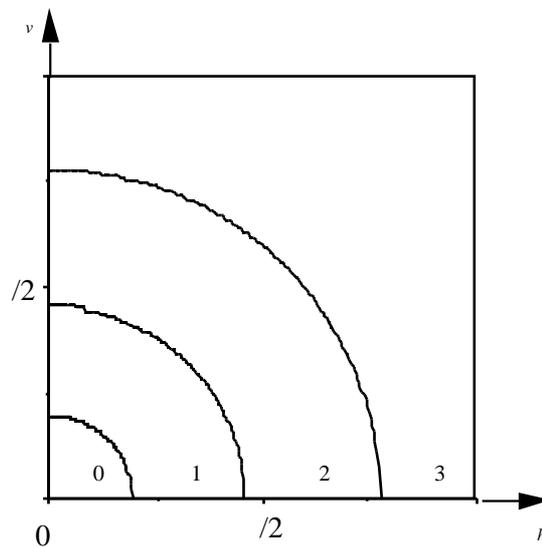


図 4.1 最適帯域分割パターン（水平・垂直相関非分離型画像モデル， $\rho=0.9$ ， $M=4$ ）

ここで，図4.2に示すように，2次元周波数帯域を水平及び垂直方向に均等に N 分割して， N^2 個の帯域ブロック

$$i,j = (h, v) \mid \frac{i}{N} \leq h < \frac{(i+1)}{N}, \frac{j}{N} \leq v < \frac{(j+1)}{N} \quad (4.7)$$

$$(i, j) \in I \quad (I = \{(i, j) \mid i, j = 0, 1, \dots, N-1\})$$

に分割することを考える．この帯域ブロックの集合 $\{(i, j) \mid (i, j) \in I\}$ のある直和分割により，次式に示すように I を M 個の帯域 I_k ($k=0, 1, \dots, M-1$) に分割する．ただし， M は $1 \leq M \leq N^2$ を満たす整数である．

$$= \sum_{k=0}^{M-1} k, \quad k = \left\{ i, j \mid (i, j) \in J_k^{(M)} \right\} \quad (4.8)$$

ここで、 $J_k^{(M)}$ は k に含まれる帯域ブロックの番号 (i, j) の集合を表し、それは2次元周波数帯域における分割の仕方を定める。このため、 $J^{(M)} = \{ J_0^{(M)}, J_1^{(M)}, \dots, J_{M-1}^{(M)} \}$ を、帯域ブロックによる (M) 分割と呼ぶ。相異なるこのような分割 $J^{(M)}$ の総数は、第2種Stirling数と呼ばれ、 $\frac{1}{M!} \sum_{k=0}^{M-1} (-1)^{M-k} \binom{M}{k} k^{N^2}$ となる[62]。また、帯域分割は、直和分割を仮定し、

$$\begin{aligned} J_k^{(M)} \cap J_l^{(M)} &= \emptyset \quad (k \neq l) \\ \bigcup_{k=0}^{M-1} J_k^{(M)} &= I \quad (J_k^{(M)} \text{ は } I \text{ の分割}) \end{aligned} \quad (4.9)$$

とする。以上のように、帯域分割を帯域ブロックによる (M) の分割に制限したときに、量子化雑音改善量を最大とするような分割 $J_{opt}^{(M)}$ を帯域ブロックをベースとする最適帯域分割と呼び、次節以降その分割法を検討する。

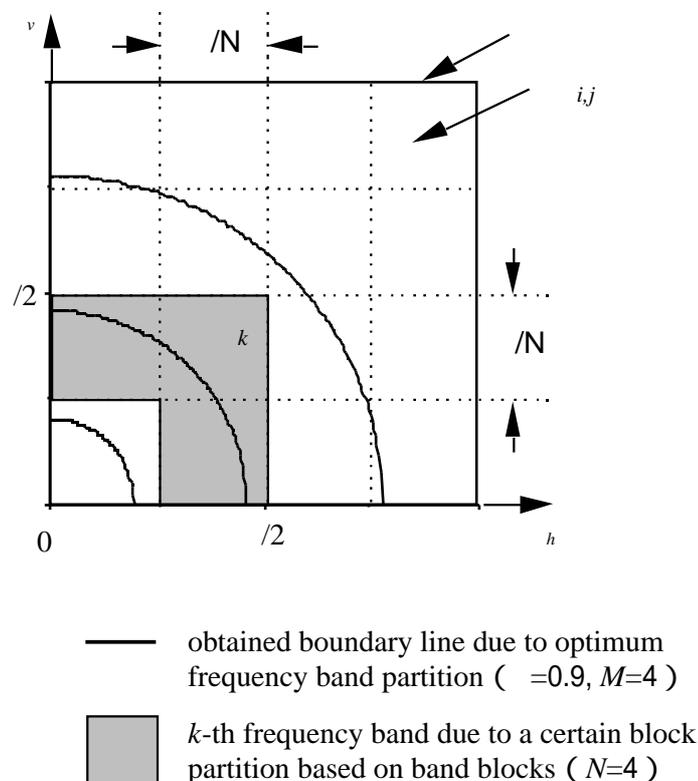


図 4.2 2次元周波数帯域の分割；最適帯域分割と帯域ブロックによる近似

4.2.2 $J_{opt}^{(M)}$ の算出法

ある M 分割 $J^{(M)}$ における帯域 k ($k=0,1, \dots, M-1$) の出力レート r_k と信号電力 P_k^2 は, 式(4.3)より次式となる.

$$r_k = |J_k^{(M)}| / N^2 \quad (k=0,1, \dots, M-1) \quad (4.10)$$

$$P_k^2 = \sum_{(i,j) \in J_k^{(M)}} P_{x(i,j)}^2 d_{i,j}^2$$

ただし,

$|J_k^{(M)}|$: 集合 $J_k^{(M)}$ の元の数 (帯域 k に含まれる帯域ブロックの数)

$$P_{x(i,j)}^2 = \frac{1}{2} P_x^2(d_{i,j}^2)$$

したがって, 式(4.10)を式(4.2)に適用することで, 特定の帯域分割における量子化雑音改善量 $G(J^{(M)})$ は求められるが, 逆に, $G(J^{(M)})$ を最大にするような最適分割 $J_{opt}^{(M)}$ を直接求めることは困難である. また, 相異なる分割 $J^{(M)}$ の総数は, N 及び M の値が少しでも大きくなると膨大となるため, 全探索を行うことは非現実的である. そこで, G の性質から分割の探索範囲を大幅に制限し, その中から最適分割解を求める.

$G(J^{(M)})$ は, 式(4.2) (4.10)より次式のように書き表される.

$$G(J^{(M)}) = \frac{G(J^{(N^2)})}{\prod_{k=0}^{M-1} [G(J_k^{(M)})]^{|J_k^{(M)}| / N^2}} \quad (4.11)$$

ここで, 上式において

$$G(J^{(N^2)}) = \frac{1}{N^2 \sum_{(i,j) \in I} P_{x(i,j)}^2} \quad (4.12)$$

$$G(J_k^{(M)}) = \frac{1}{\sum_{(i,j) \in J_k^{(M)}} P_{x(i,j)}^2} \quad (4.13)$$

である. 明らかに, $M=N^2$ のときの分割の仕方は一通りしかないため, 式(4.12)は確定した定数となる. また, 式(4.13)は, 帯域 k を $|J_k^{(M)}|$ 個の帯域ブロックに分割したときの G の値を表している.

以上の各 G の $[dB]$ 表示 ($10\log G$) を, $G_{(dB)}$ のように書き表すとした場合に, 式(4.11)より次式が得られる.

$$G(J^{(M)})_{(dB)} = G(J^{(N^2)})_{(dB)} - \sum_{k=0}^{M-1} \frac{|J_k^{(M)}|}{N^2} G(J_k^{(M)})_{(dB)} \quad (4.14)$$

上式の右辺第1項は定数であるから, $G(J^{(M)})_{(dB)}$ を最大とするためには, 第2項を最小とするような分割を行えばよい. ここで, $A.M.\{\}$ 及び $G.M.\{\}$ は, それぞれ, 相加平均及び相乗平均を表すものとして, 式(4.13)は,

$$G(J_k^{(M)}) = \frac{A.M.\left\{ \begin{matrix} 2 \\ i,j \end{matrix} \mid (i,j) \quad J_k^{(M)} \right\}}{G.M.\left\{ \begin{matrix} 2 \\ i,j \end{matrix} \mid (i,j) \quad J_k^{(M)} \right\}} \quad (4.15)$$

となることに注意すれば, $G(J_k^{(M)})_{(dB)}$ に関する以下の性質が得られる.

[性質]

- 1°) 帯域 k に1個の帯域ブロックしか含まなければ (すなわち, $|J_k^{(M)}|=1$), $G(J_k^{(M)})_{(dB)}=0$ である.
- 2°) 常に, $G(J_k^{(M)})_{(dB)} \geq 0$ が成立し, 等号は帯域 k に含まれる帯域ブロックの電力が全て等しいときに成立する. また, 各帯域ブロックの電力が互いに出来るだけ近い値であるほど, $G(J_k^{(M)})_{(dB)}$ の値は小さい.
- 3°) $G(J_{opt}^{(M)})_{(dB)} \geq G(J_{opt}^{(M-1)})_{(dB)}$ ($M=2,3,\dots,N^2$) なる関係が成立する.

以上の性質より, 各帯域ブロックの番号 $(i,j) \in I$ をその電力 P_{ij}^2 の値が大きい順に並べ, $\{(i,j)_0, (i,j)_1, \dots, (i,j)_{N^2-1}\}$ となったとすると (第 m 番目の番号を $(i,j)_m$ と表示), 各 $J_k^{(M)}$ ($k=0,1,\dots,M-1$) が隣接する番号のみを含むような

$$J^{(M)} = \underbrace{\{(i,j)_0, (i,j)_1, \dots, (i,j)_{N^2-1}\}}_{(J_0^{(M)})}, \underbrace{\{(i,j)_m, \dots, (i,j)_{N^2-1}\}}_{(J_1^{(M)})}, \dots, \underbrace{\{(i,j)_{N^2-1}\}}_{(J_{M-1}^{(M)})} \quad (4.16)$$

なる分割の中に $J_{opt}^{(M)}$ が存在することが分かる. 帯域ブロックをベースとする本帯域分割算出法における $J^{(M)}$ の相異なる総数は, ${}_{N^2-1}C_{M-1}$ であり, 当初の分割に比して大幅に削減される. 例えば, 次節で述べる $N^2=16, M=4$ の場合においては, その総数は約1.7億から455に低減する. 以下, $G(J^{(M)})$ を最大にする $J_{opt}^{(M)}$ は, このような分割の全探索により求めるものとする.

4.2.3 G 特性と帯域分割パラメータの選定

帯域ブロックをベースとする最適帯域分割において，帯域分割数 M 及び帯域ブロックの個数 N^2 は，共に大きいほど量子化雑音改善量 G は改善されるが，一方，処理コストの点からは両者は共に小さい方が望ましい．このため，4.2.1節に述べた相関係数値 $\rho=0.9$ の水平・垂直相関非分離型画像モデルを対象に， G をベースにして基本特性を検討し，それに基づき両パラメータの実用上有効な値を選定する．

図4.3は，帯域ブロックの個数 $N^2=16, 64$ とした場合に，帯域分割数 M をパラメータとして，帯域ブロックをベースとする最適帯域分割による $G(J_{opt}^{(M)})$ 特性を示したものである．同図には，本モデルに対して第3章で理論的に導出された最適帯域分割の G 特性，及び前章3.7.3節で述べた本画像モデルに対する G の理論限界値 G_{lim} の値を併せて表示している．この結果から，帯域分割数 M が大きくなることで最適帯域分割における G 及び帯域ブロックをベースとする最適帯域分割における $G(J_{opt}^{(M)})$ の値は増大し， G_{lim} に漸近するものの， $M=4$ の範囲では飽和傾向にあることが分かる．例えば，最適帯域分割において， $M=4, 8$ での G の差は約 $0.4[dB]$ である．この結果に加えて，符号化システム実現の簡易化のためには，特に帯域分割数は小さい方が望ましいことから， $M=4$ と選定できる．

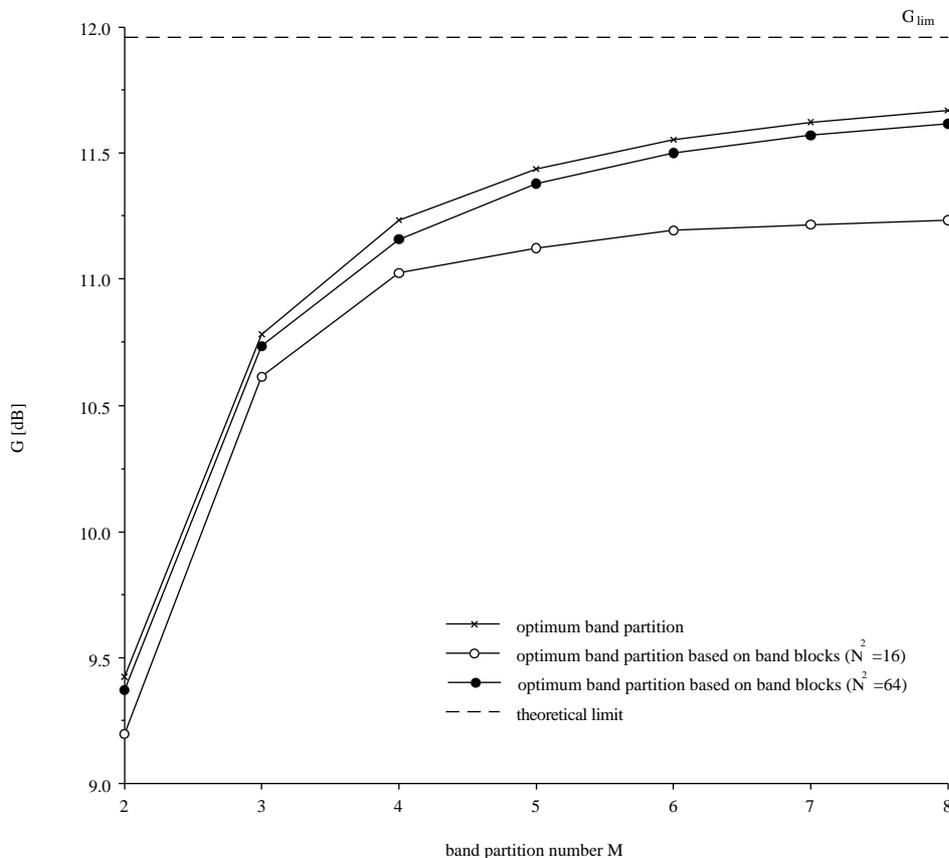


図 4.3 最適帯域分割と帯域ブロックをベースとする最適ブロック分割における量子化雑音改善量 G (パラメータ：帯域分割数 M)

次に、所要とする帯域ブロックの個数を検討する。帯域分割数を $M=4$ に固定した場合に、最適帯域分割と帯域ブロックの個数 N^2 を各々16及び64とした場合の帯域ブロックをベースとする最適帯域分割との G の差は、それぞれ、約 $0.2[dB]$ 及び約 $0.1[dB]$ であり、その差は僅か $0.1[dB]$ であることから、やはり符号化システム実現の簡易化を考慮して、帯域ブロックの個数 $N^2=16$ で十分であると結論づけることができる。

選定した帯域分割パラメータ $M=4$ 、 $N^2=16$ を適用したときの水平・垂直相関非分離型画像モデルに対して求められた帯域ブロックをベースとする最適帯域分割による2次元周波数平面上の帯域分割パターンを図4.4に示す。各帯域ブロックの中で、最も低周波の帯域ブロックである $(0,0)$ の電力 $P_{0,0}^2$ は、他と比較して突出して大きいため、本最適帯域分割の第0番目の帯域 B_0 には、この帯域ブロックのみが含まれることになる。このため、4.2.2節で述べた相異なる分割の探索数455は更に減少し、実際の探索に要する総数は、僅か91でよいことになる。フィルタバンクの実現性を考慮した上で、帯域ブロックをベースとして求められた最適帯域分割法に基づいて導出された本帯域分割パターンは、同じ $M=4$ の場合の理想的な最適分割パターンである図4.1に対して、 G の低下は $0.2[dB]$ と僅少である一方、フィルタバンクの実現性は極めて高い。次節以降では、本分割パターンをベースに、フィルタバンクの構成法及びその実画像への適用の検討を行う。

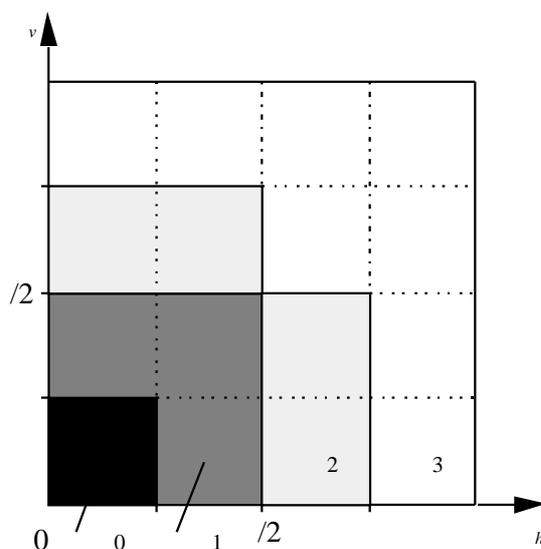


図 4.4 帯域ブロックをベースとする最適帯域分割パターン
(水平・垂直相関非分離型画像モデル, $\rho=0.9$, $M=4$, $N^2=16$)

4.3 帯域ブロックをベースとする最適帯域分割の実現

4.3.1 フィルタバンクの構成

帯域ブロックをベースとする最適帯域分割によって得られた各帯域 k ($k=0,1,\dots,M-1$) は、図4.4に示すように正方形の帯域ブロック $(i,j) \in I$ の適当な集合からなるので、この分割を実現するフィルタバンクは、図4.5に示すように簡易に構成することができる。同図で、切り替えスイッチ SW は固定 (fixed) 側に倒れているとする。

まず、画像入力信号 $x(m,n)$ ($NK \times NK$ 画素とする) を、1次元均等 N 分割フィルタバンクを用いて、水平及び垂直方向に帯域分割することにより、各帯域ブロック $(i,j) \in I$ を通過域とするフィルタ出力 $x_{i,j}(m,n)$ ($K \times K$ 画素) を作る。そして、帯域 k の出力信号 $x_k(m,n)$ は、帯域ブロックをベースとする最適帯域分割である $J_{opt}^{(M)}$ の k 番目の要素 $J_{k,opt}^{(M)}$ に含まれる帯域ブロック番号 (i,j) に対応したこのフィルタ出力の集合として

$$x_k(m,n) = \{ x_{i,j}(m,n) \mid (i,j) \in J_{k,opt}^{(M)} \} \quad (k=0,1,\dots,M-1) \quad (4.17)$$

のように得られる。なお、図4.5で $x_0(m,n)$ を更に2つの帯域に分割する点線以降の部分については、次節に述べられる。

4.2.3節で $N=4$ に選定したから、図4.5では上記の1次元均等 N 分割フィルタバンクは、タップ長 16 の QMF [52],[53] を 2 段の tree 状にして構成することができる。ここで、フィルタとして QMF を用いたのは以下の理由による。

- ・画像に対してサブバンド符号化を行う場合に、その直交性及び線形位相性から、一般に QMF が用いられている場合が多い [16],[18],[20],[41],[52],[63]。
- ・実画像には、次節で述べるように2次元空間領域での直流分の変動に代表される非定常性が含まれているが、QMFによる帯域分割は空間情報を保存するため、この非定常性に対応できるという利点がある。

また、タップ長の選定については、

- ・タップ長を大きくすることで、フィルタの遮断特性は急峻となり、4.2節で仮定した理想分割を近似することができるが、タップ長が 16 と 32 の場合でその量子化雑音改善量の特性にほとんど差がない [64]。
- ・一方、長いタップ長は畳込み演算の計算量が多くなること、再生画像の輪郭部周辺でのリングングが目立つ等の実用的な問題がある [65]。

を考慮して、タップ長として 16 を選択することにした。

以上で構成される帯域ブロックをベースとする最適帯域分割フィルタバンクを用い、帯域分割数 $M=4$ のときの最適帯域分割 $J_{opt}^{(4)}$ として図 4.4 に示した固定の分割パターンを適用し、実際の画像に対して量子化雑音改善量を求めた結果を表 4.1 の 2 列目に示す。

なお、テスト画像には、汎用性の向上を目的に選定された第 3 章において図 3.10(a) ~ (l) に示す 12 枚の標準画像（いずれも (256×256) 画素，濃淡 8 ビット / 画素^{注3}）を使用した。同表の 4 列目は、帯域ブロックをベースとする本最適帯域分割との比較のために、JPEG で用いられている (8×8) 画素毎の DCT を同画像に適用した場合の量子化雑音改善量の値を示している。この結果において、本実験に用いた半数以上の画像については、帯域ブロックをベースとする最適帯域分割は、その帯域分割数 M が僅か 4 であるにもかかわらず、 $(8 \times 8) = 64$ 分割の DCT と遜色のない量子化雑音改善効果があることを示しており、固定の分割パターンをもつ本最適帯域分割の有効性を確認することができる。

表 4.1 帯域ブロックをベースとする最適帯域分割（固定）における各画像の量子化雑音改善量 G [dB]

<i>image</i>	<i>fixed optimum band partition based on band blocks</i>	<i>with dc component separation</i>	<i>DCT (8x8 pixels)</i>
Hada	17.648	18.103	18.157
Hairband	14.224	14.449	14.289
Wether	13.468	13.742	14.251
Girl	14.471	14.767	14.708
Lenna	14.178	14.434	16.965
Barbara	9.402	9.768	13.114
Church	6.716	6.861	7.248
Tulip	7.342	7.515	7.883
Cameraman	12.287	12.619	13.670
Building	7.318	7.675	11.281
Wine	15.972	16.406	17.075
Moon	9.281	9.582	9.623

注3：以下 [bit/pe] と記すこととする。

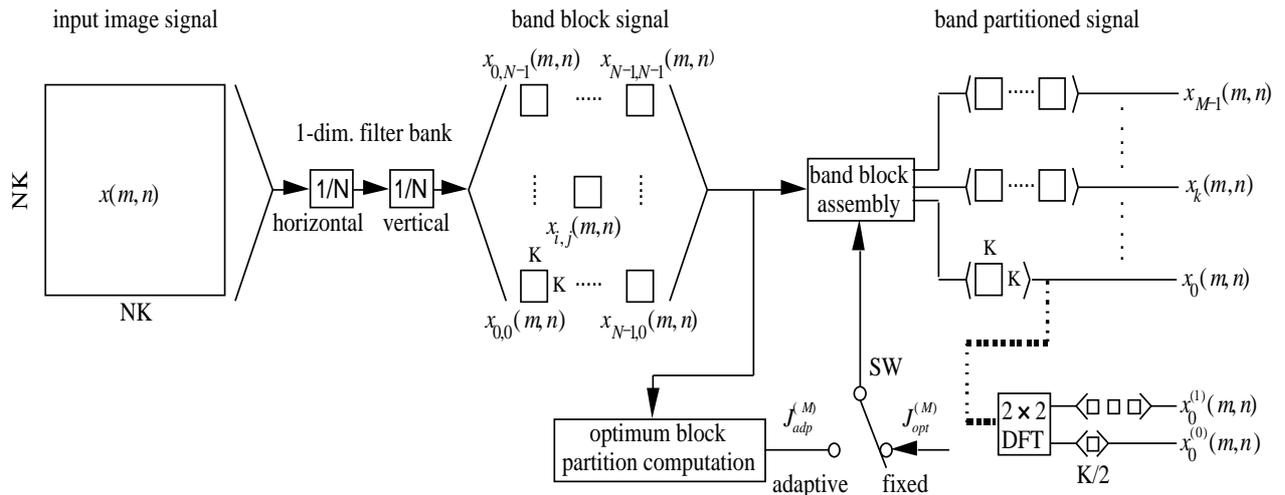


図 4.5 帯域ブロックをベースとする最適帯域分割を実現するフィルタバンクの構成

4.3.2 非定常性に対する対策

明らかに，現実の画像信号は非定常性を有する[51]．ここでは，2次元空間領域の場所によって変動する非定常直流成分と，画像内容によって，或いは空間領域の場所によって変動する相関係数に注目し，それらを考慮した上で，画像における帯域ブロックをベースとする最適帯域分割について考える．

4.3.2.1 直流成分の非定常性

式(4.5)に示された画像モデルには，直流成分は含まれていない．よって，4.3.1節において画像を対象に求められた帯域ブロックをベースとする最適帯域分割においては，直流成分は画像全体の平均値として1つの値で扱い，この直流成分が除かれた画像信号に対して帯域分割処理を行っている．この処理は，一般的に，画像のサブバンド符号化においては，同様に扱われていることが多い．しかし，上述したように，画像の直流成分は，空間的な場所によって変動しているために，それをたった1つの平均値で代表させることは，符号化において不利なことであると考えられる．そこで，以下に示す方法によって，画像に含まれる非定常直流成分の抽出及び分離を行う．

画像信号を図4.4に示した帯域ブロックをベースとする最適帯域分割 $J_{opt}^{(4)}$ に基づいて帯域を分割する場合，非定常直流成分が含まれるのは，最も周波数が低く，且つ，1つの帯域ブロックしか含まない $\omega_0 = \{ (h', v') | 0 \leq h', v' < N/4 \}$ なる帯域のみである．そこで，全体としての信号レートを保存したままで非定常直流成分を抽出するために，この帯域 ω_0 の信

号 $x_0(m,n)$ ($=x_{0,0}(m,n)$) を空間領域で (2×2) 画素のブロックに分割し、各ブロック毎に2次元2点DFTを次式のように適用する。

$$\begin{pmatrix} x_0^{(0,0)}(m,n) & x_0^{(0,1)}(m,n) \\ x_0^{(1,0)}(m,n) & x_0^{(1,1)}(m,n) \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 1 & 1 & 1 & -1 \\ 1 & -1 & 1 & -1 \end{pmatrix} \begin{pmatrix} x_0(2m,2n) & x_0(2m,2n+1) \\ x_0(2m+1,2n) & x_0(2m+1,2n+1) \end{pmatrix} \quad (4.18)$$

この変換によって得られた4つの信号成分 $x_0^{(p,q)}(m,n)$ ($p,q=0,1, K/2 \times K/2$ 画素, K : 偶数) を、以下のように、最も低周波の成分とそれ以外の成分の集合との2つに分ける。

$$\begin{aligned} x_0^{(0)}(m,n) &= x_0^{(0,0)}(m,n) \\ x_0^{(1)}(m,n) &= \{ x_0^{(p,q)}(m,n) \mid (p,q) = (0,1), (1,0), (1,1) \} \end{aligned} \quad (4.19)$$

以上の操作により、帯域 0 の信号 $x_0(m,n)$ は (2×2) 画素毎に抽出された非定常直流成分である $x_0^{(0)}(m,n)$ とそれ以外の周波数成分である $x_0^{(1)}(m,n)$ の2つの帯域に再分割されたことになる。図4.5において、 $x_0(m,n)$ から出ている点線以降の部分がこの帯域分割処理を表している。図4.6に、一例として、画像 "Hada" における帯域 0 の信号 $x_0(m,n)$ と、それから抽出された非定常直流成分である $x_0^{(0)}(m,n)$ の画像例を示す。この結果において、 $x_0^{(0)}(m,n)$ は、画像の空間領域でその値が大きく変動していることを確認することができる。

本処理を含めた帯域ブロックをベースとする最適帯域分割を適用することによって2次元周波数平面は合計5つの帯域に分割されることになる。なお、DFTを適用する (2×2) 画素というブロックサイズは、合成側のフィルタを通して画像を再生した場合には (8×8) 画素に相当する。これは、JPEGにおいて、画像のマルコフ性を考慮して決められたDCTを適用する際のブロックサイズに等しいものとなっている[49],[51],[63],[66]。

上述した非定常直流成分の分離を含めて、帯域分割数 $M=5$ とした帯域ブロックをベースとする最適帯域分割を適用した場合の量子化雑音改善量の値を表4.1の3列目に示している。画像によりその効果は異なっているが、非定常直流成分の分割処理を行うことで、量子化雑音改善量値において約 $0.3[dB]$ の更なる改善が得られていることが分かる。

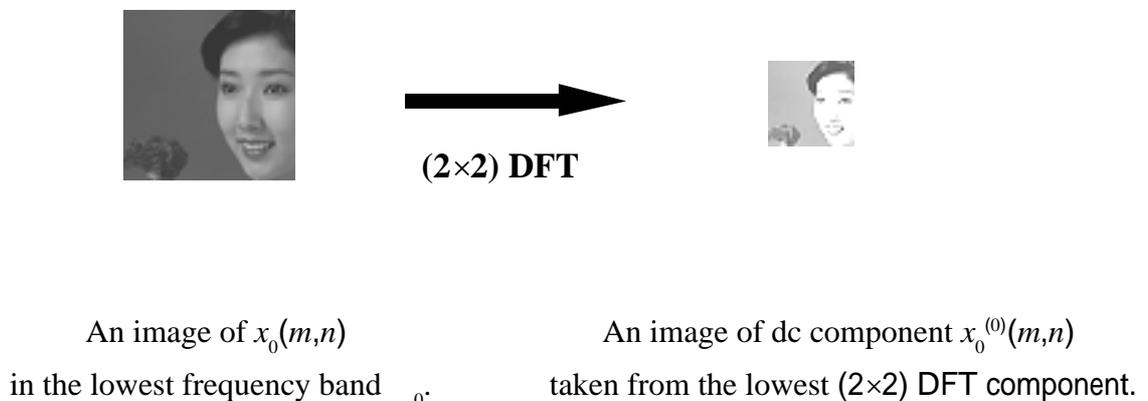


図 4.6 非定常直流成分の抽出

4.3.2.2 相関係数の非定常性

次なる非定常性として、相関係数の変動について考える。画像の相関係数は、画像内容によって、或いは同じ画像であっても空間的な場所によって変動している。図4.1及び図4.4に示した帯域分割パターンは、 $\rho=0.9$ なる特定の相関係数をもつ入力信号に対して得られた最適帯域分割、及び帯域ブロックをベースとする最適帯域分割であるから、本来ならば、画像内容によって、或いは空間的な場所によって、その相関係数に整合した最適帯域分割、及び帯域ブロックをベースとする最適帯域分割に適応的に切り換えるべきである。しかし、これによる処理コストの増加に見合う効果が得られるかどうかの問題であり、それを探るため、逆に、帯域分割パターンをある相関係数値をもつ入力信号に対して最適化を行った分割に固定したときに、それとは異なる相関係数をもつ信号が入力信号として与えられた場合の量子化雑音改善量の劣化特性という観点からの実験を行った。

図4.7は、水平・垂直相関非分離型画像モデルを入力信号として、その相関係数値 ρ を0.1～0.95まで変化させたとき、帯域分割パターンを $\rho=0.9$ 条件下での最適帯域分割（図4.1）及び帯域ブロックをベースとする最適帯域分割（図4.4）に固定した場合の量子化雑音改善量 $G_{I,fix}$ 及び $G_{B,fix}$ と、各 ρ の値に対して適応的に最適帯域分割及び最適ブロック分割を行った場合の量子化雑音改善量 $G_{I,adp}$ 及び $G_{B,adp}$ との差 $G_{I,fix} - G_{I,adp}$ （実線）、 $G_{B,fix} - G_{B,adp}$ （点線）を劣化特性として示したものである。

両特性曲線において、当然のことながら、入力信号の相関係数値 $\rho=0.9$ の場合の劣化は0[dB]となる。また、 ρ の値が0.9からずれるほど、この劣化度は大きくなる。 $G_{I,fix} - G_{I,adp}$ 特性が、 $\rho=0.3$ の範囲で劣化の度合いが逆に小さくなっているのは、 ρ の値が0に近い入力信号に対する量子化雑音改善量の値自身が小さいためである。また、 $G_{B,fix} - G_{B,adp}$ 特性が、 $\rho=0.4 \sim 0.95$ の範囲で劣化が0[dB]となっていることは、この範囲の ρ に対して求められた帯域ブロックをベースとする最適帯域分割の分割パターンが不変であることを示している。何れの場合においても、入力信号の ρ の変動による量子化雑音改善量の劣化は、0[dB]から-0.1[dB]の間にあり、これは無視できるほど小さい値であると考えられる。したがって、入力信号の相関係数の変動に対する特別な適応処理は不要であると判断できる。

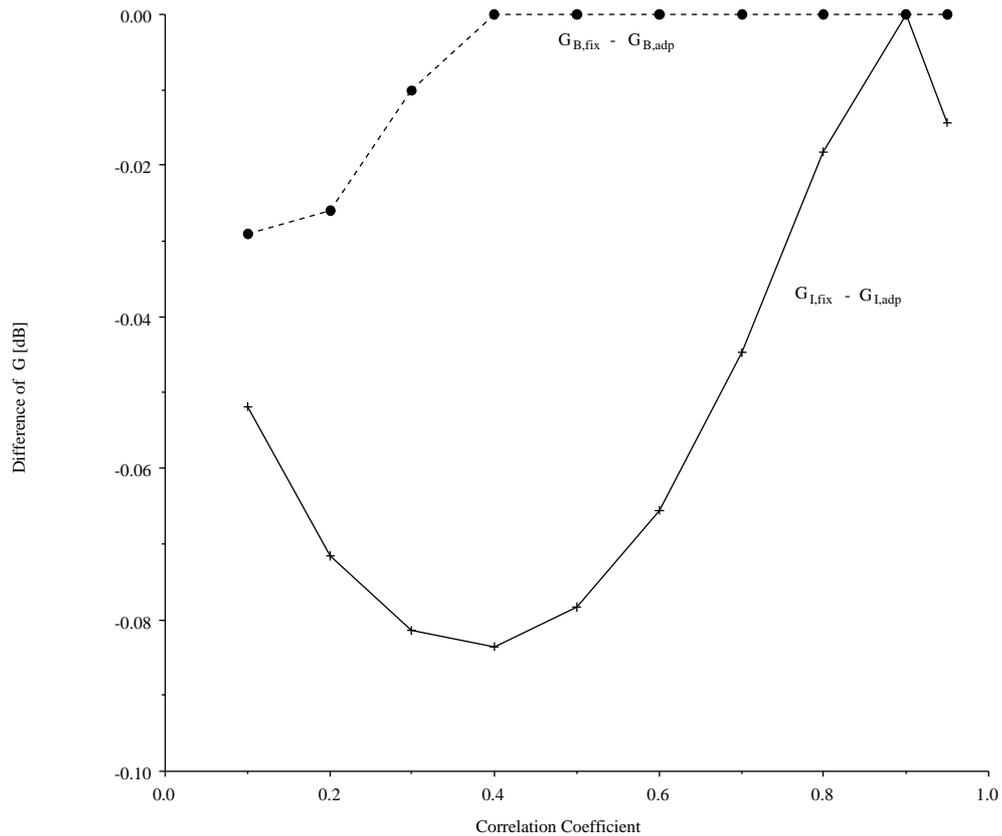


図 4.7 固定の最適帯域分割に対する量子化雑音改善量 G の変化 ($M=4$)
(パラメータ：相関係数値)

4.3.3 符号化性能

本章において提案した帯域ブロックをベースとする固定の最適帯域分割方式の符号化性能を評価するため、画像を用いて SNR 対エントロピー特性の評価実験を行った。

まず、図4.5の帯域ブロックをベースとする最適帯域分割フィルタバンクの構成で、分割パラメータは4.2.3節で述べた帯域分割数 $M=4$ 、帯域ブロック数 $N^2=16$ を用い、図4.4に示された固定の最適帯域分割を得た後、最も低い帯域信号 $x_0(m,n)$ には (2×2) DFT を適用して非定常直流成分を分離する構成とし、全体で5つの帯域分割信号を得た。

次に、非定常直流成分以外の各帯域分割信号は、2帯域ブロック以上の2次元信号系列の集合からなっているが、何れも適当なスキャン操作により1次元の帯域分割信号に変換した後、それぞれスカラ量子化を行う。ここで、量子化は、評価尺度が SNR であることを考慮して、再生誤差電力を最小とするように、ミッドトレッド型の線形量子化を用い、過負荷雑音が発生しないように各帯域分割信号の量子化レベル数を十分大きくして行った[65],[67]。なお、上記のスキャン法は多数存在するが、本 SNR 対エントロピー特性評価

においては，このスキャン方法には依存しない．

一方，非定常直流成分は，隣接ブロック間での相関が大きいことを考慮して，JPEGにおいて行われている方式と同様に，上と同じ線形量子化器によって量子化された後，直前のブロックとの差分をとった予測誤差信号を，非定常直流成分におけるエントロピーを求める対象の信号系列とした[2],[33],[34],[66]．

SNRは，原画像と再生画像との再生誤差電力を $\frac{1}{q}$ として，次式により求めた．

$$SNR = 10 \log_{10} \frac{255^2}{q} \quad (dB) \quad (4.20)$$

また，エントロピーについては，量子化された各帯域の信号が無記憶情報源から生じたものと仮定して算出を行った．

本節においては，固定の最適帯域分割により十分な量子化雑音改善量を得ることができ画像の代表として，図4.8に画像内容とその自己相関特性を示す画像"Hada"を例に挙げる．この自己相関特性は，式(4.4)に示された等方性相関モデルによって十分に近似可能であり，このことから，水平・垂直非分離型画像モデルに基づいて求められた固定の最適帯域分割の適用が有効であることが推測できる．なお，実験に使用した他のテスト画像の中で，前章の図3.10(b)の画像"Hairband"，同図(d)の画像"Girl"，同図(l)の画像"Moon"が，この分類に属する画像になる．

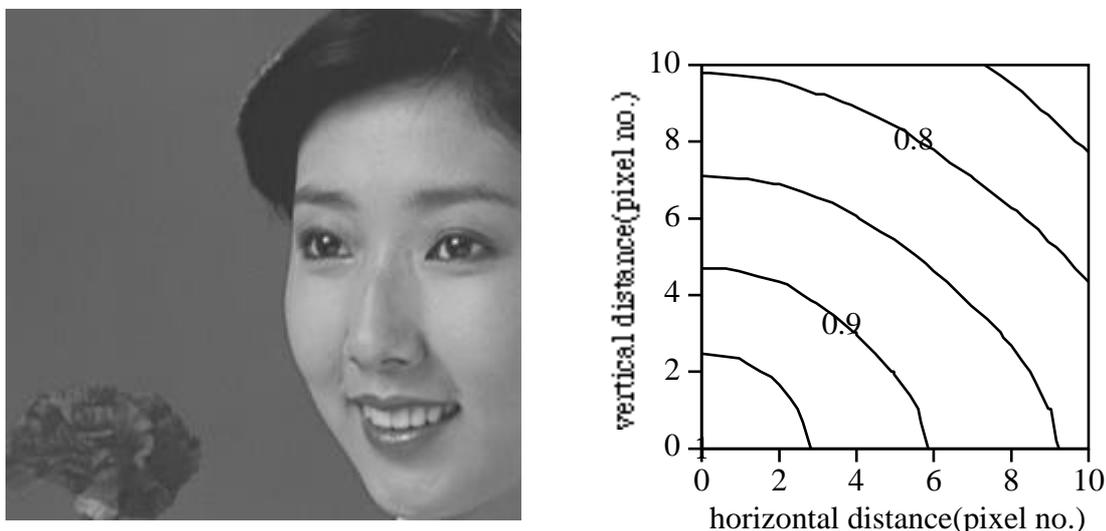


図 4.8 画像 "Hada" とその自己相関特性

画像"Hada"に固定の最適帯域分割を適用するにあたり，画像のサブバンド符号化において最も一般的に用いられている帯域を2次元周波数平面上で帯域を均等に分割するlinear

分割[16],[18]と比較した場合に，最適帯域分割は，4分割のlinear分割に比して，treeの段数が1段多くなるため，帯域分割に要する処理コストは2倍になるものの，量子化及び符号化に要する処理コストは変わらないという条件で，量子化雑音改善量 G において， $3.14[dB]$ の大きな改善を得ることができる．また，16分割のlinear分割と比較した場合には，帯域分割に要する処理コストは同じであるが，それに対して量子化及び符号化に要する処理コストが0.25倍に抑えられるにもかかわらず， G の低減は僅か $0.23[dB]$ にすぎない．このとき，量子化及び符号化器の設計法に依存するが，量子化及び符号化の処理は，帯域分割処理に比べてはるかに多くのコストを必要することを考慮されたい．

最後に，画像"Hada"に対して，帯域ブロックをベースとする固定の最適帯域分割を適用した場合のSNR対エントロピー特性を図4.9に示す．同図には，本最適帯域分割の評価を行うため， (8×8) 画素のDCTを適用した場合のSNRとエントロピーの関係を併せて表示している．なお，この特性はJPEGの方式に準拠して，画像に (8×8) 画素のDCTを適用して得られる64種の変換係数系列に，上記の量子化を適用した後，直流成分に対しては，直前のブロックとの差分を，交流成分に対しては，63個の帯域分割信号そのままの値を用いてSNRとエントロピーを求めることで得られたものである．更に，同図においては，後述する適応分割を同画像に適用した場合についてのSNR対エントロピー特性を表示する．画像"Hada"に対して，帯域分割数 $M=4$ ，帯域ブロック数 $N^2=16$ なる同条件の場合に，適応分割の適用によって得られる量子化雑音改善量 G の改善は，固定の最適帯域分割に比べて僅か $0.17[dB]$ である．図4.10に，本適応分割によって得られた画像"Hada"の帯域分割パターンを示す．図4.10は，図4.4の帯域分割パターンに酷似したものになっていることから，同画像に対しては，ほんの僅かな量子化雑音改善量の劣化で，固定の最適帯域分割によって近似することが十分可能であることを意味している．

図4.9において，帯域ブロックをベースとする最適帯域分割によるSNR特性は，固定及び適応両分割共に，同じビットレートの (8×8) 画素のDCTのSNRより $1[dB]$ 程度優れた値を示している．また，上述したように，固定分割と適応分割については，両分割方式による特性の差はほとんどなく，適応分割を行わなくとも，固定な最適帯域分割パターンによって近似できることが確認できる．

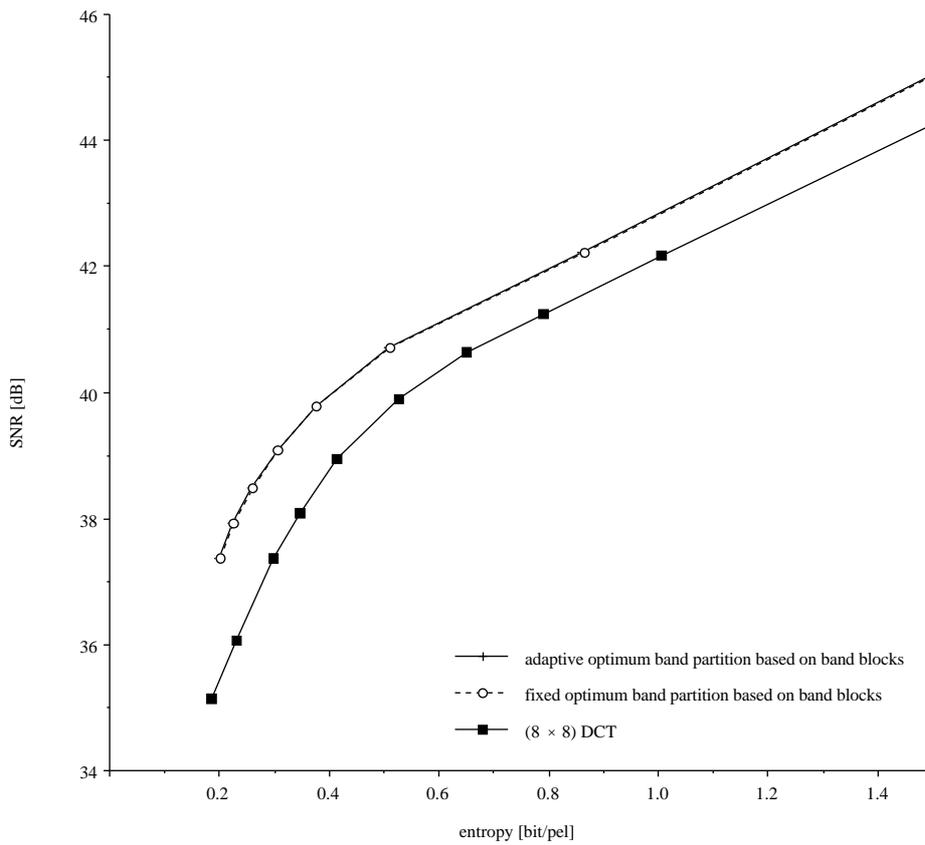


図 4.9 SNR 対エントロピー (画像 "Hada")

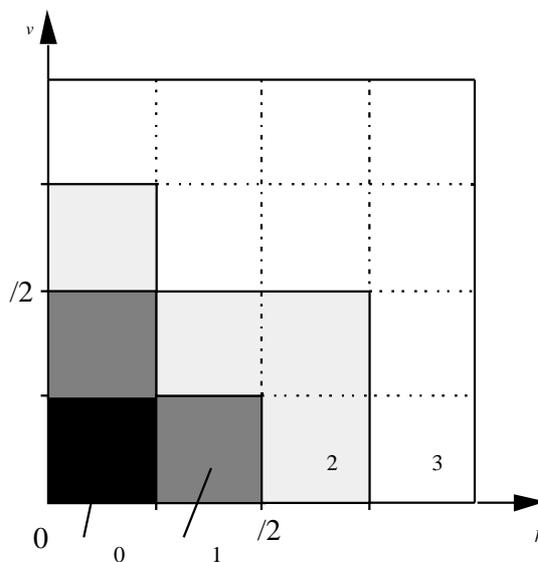


図 4.10 帯域ブロックをベースとする適応帯域分割パターン (画像 "Hada", $M=4$, $N^2=16$)

4.4 適応分割

本章においてこれまで扱ってきた帯域ブロックをベースとする最適帯域分割は，画像信号の自己相関関数が空間的に等方性の特性を有する水平・垂直非分離型画像モデルに基づいて導出されたものである．第3章に示したように，多くの画像に対して，この仮定は妥当なものであるが[49],[51]，画像によっては，自己相関特性が特別な形を有し，等方性の相関モデルが当てはまらないばかりか，そのモデル化が困難な場合がある．そのため，このような画像については，水平・垂直非分離型画像モデルに基づいて求められた帯域ブロックをベースとする固定の最適帯域分割を適用したとしても，実際には画像に対して真に最適な帯域分割は行われなかったと考えられ，表4.1に示したいくつかの画像においては，本最適帯域分割による量子化雑音改善量の値が，比較方式である(8×8)画素のDCTの場合の値に比して大きく下回っているものがある．

そこで，本来であれば，最適帯域分割は，画像毎に固有な帯域分割パターンであると考え，帯域ブロックをベースとする適応帯域分割方式を提案すると共に，本方式の適用による符号化性能について評価を行う．

4.4.1 帯域ブロックをベースとする適応分割方式

帯域ブロックをベースとする適応帯域分割方式は，図4.5において，切り替えスイッチSWを適応(adaptive)側に倒し，入力された画像信号の相関特性に整合した最適帯域分割 $J_{adp}^{(M)}$ に基づいて，2次元周波数帯域をブロック分割する方式である．ここで， $J_{adp}^{(M)}$ は次のように算出される．

図4.5において，各帯域ブロックの信号 $x_{i,j}(m,n)$ ($(i,j) \in I$) から，その信号電力を

$$P_{i,j} = \frac{1}{K^2} \sum_{m,n=0}^{K-1} x_{i,j}^2(m,n) \quad ((i,j) \in I) \quad (4.21)$$

により求める．この $P_{i,j}$ ($(i,j) \in I$) を用いて，4.2.2節で述べた算出法と同様にして，この場合の量子化雑音改善量 G を最大にする帯域ブロックをベースとする最適帯域分割 $J_{adp}^{(M)}$ を算出することにより，画像に固有な最適帯域分割を求めることができる．

まず，表4.1に示した12枚のテスト画像に対して，4.2.3節で得られた固定の最適帯域分割の場合と同一の帯域分割パラメータ条件(帯域分割数 $M=4$ ，帯域ブロック数 $N^2=16$)に基づき本適応帯域分割方式を適用し，その結果得られた量子化雑音改善量の値を表4.2に示す．なお，表4.2の3列目には，本適応帯域分割の後に，4.3.2.1節に述べたDFTを用いた非定常直流成分の分離法を適用し，合計5つの適応帯域分割処理を行った場合の量子化雑音改善量の値を示している．

表 4.2 帯域ブロックをベースとする最適帯域分割 (適応) における各画像の量子化雑音改善量 G [dB] ($M=4$, $N^2=16$)

<i>image</i>	<i>adaptive optimum band partition based on band blocks</i>	<i>with dc component separation</i>
Hada	17.819	18.274
Hairband	14.224	14.449
Wether	13.981	14.256
Girl	14.572	14.868
Lenna	17.346	17.601
Barbara	12.592	12.958
Church	7.251	7.396
Tulip	7.763	7.936
Cameraman	13.221	13.553
Building	9.926	10.283
Wine	16.433	16.868
Moon	9.425	9.725

以上の結果において、帯域ブロックをベースとする適応帯域分割を行うことにより、各画像に対する量子化雑音改善量の値は、表4.1に示した値から改善されている。非定常直流成分の分離を行うことで、ほとんどの画像においては、その値は (8×8) 画素のDCTによる量子化雑音改善量結果を若干上回るものとなっているが、未だDCTによる改善量に及ばない画像も存在する。そのような画像への対応は次節以降に述べるものとして、ここでは、帯域分割数 $M=4$ 、帯域ブロック数 $N^2=16$ なる条件をもつ本適応分割により効果的な量子化雑音改善量を得ることができる画像の代表として、図4.11に画像内容とその自己相関特性を示す画像 "Tulip" を例に挙げる。この自己相関特性は、相関特性の変化において等方性という性質を有しているものの、等方性相関モデルでの完全なる近似は幾分困難であることが分かる。このことから、本適応分割の適用が有効であることが推測できる。なお、実験に使用した他のテスト画像の中で、前章の図3.10(c)の画像 "Wether"、同図(g)の画像 "Church"、同図(k)の画像 "Wine" が、この分類に属する画像になる。

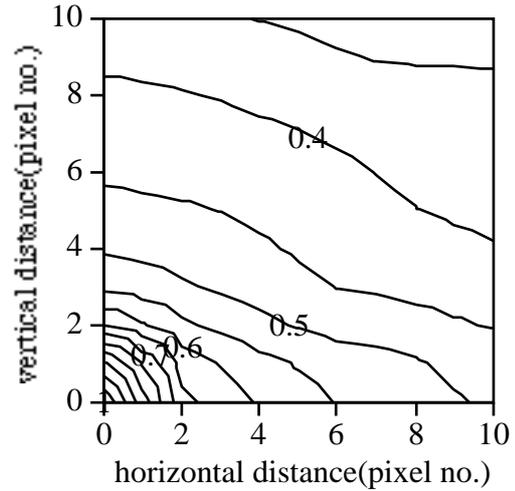
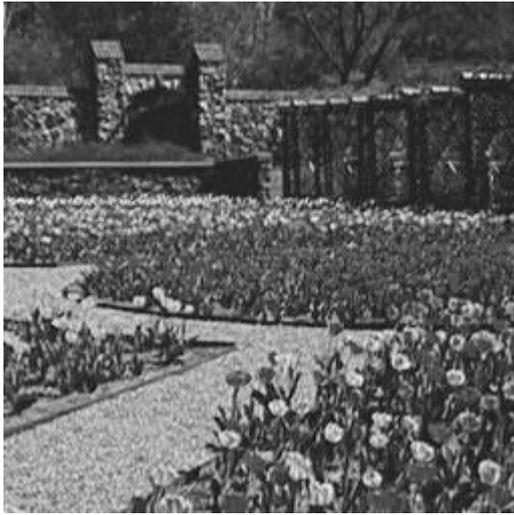


図 4.11 画像 "Tulip" とその自己相関特性

まず，画像"Tulip"に対して，本適応分割を適用することによって得られた帯域分割パターンを図4.12に示す．この結果は，図4.4に示した分割パターンとは少し異なっていることが確認され，固定の最適帯域分割の場合に比べて，量子化雑音改善量において $0.42 [dB]$ の改善を得ることができる．

本適応分割を実行するにあたっては，式(4.4)の $G(J^M)$ を最大にする最適解を探索する必要がある．しかし，そのために要する探索回数は僅か91回であり，これによる帯域分割処理に要するコストは，固定の最適帯域分割の場合に比して1.02倍にすぎない．また，画像"Tulip"の場合に，本適応分割は，4分割のlinear分割に比して，上述の最適解の探索を含めた帯域分割に要する処理コストは2.04倍になるものの，量子化及び符号化に要する処理コストは変わらないという条件で，量子化雑音改善量 G において， $1.58 [dB]$ の大きな改善を得ることができる．また，16分割のlinear分割と比較した場合には，帯域分割に要する処理コストは僅か1.02倍だけ増加するが，それに対して量子化及び符号化に要する処理コストは0.25倍に抑えられる．このときの G の低減は僅か $0.33 [dB]$ にすぎない．

更に，適応帯域分割を行う場合には，画像毎に求められた固有の帯域分割パターンを表す情報をサイド情報として受信側に伝送する必要がある．しかし， (256×256) 画素の画像に対して，帯域ブロックの数 $N^2=16$ ，帯域分割数 $M=4$ の条件で求められた最適帯域分割パターンを表すのに必要なビット数は，約 $0.0005 [bit/pel]$ となり，これは無視できるほど小さい値になる．

最後に，画像"Tulip"に本適応分割を適用した場合の SNR 対エントロピー特性を図4.13に示す．同図には，本分割による改善効果を確認するために，前節に示した固定の最適帯域分割を適用した場合，及び比較方式として (8×8) 画素のDCTを適用した場合の同特性を併せて表示している．

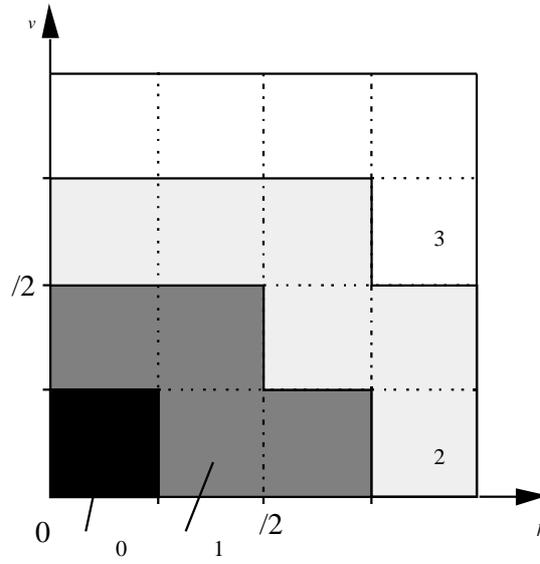


図 4.12 帯域ブロックをベースとする適応帯域分割パターン
(画像 "Tulip", $M=4$, $N^2=16$)

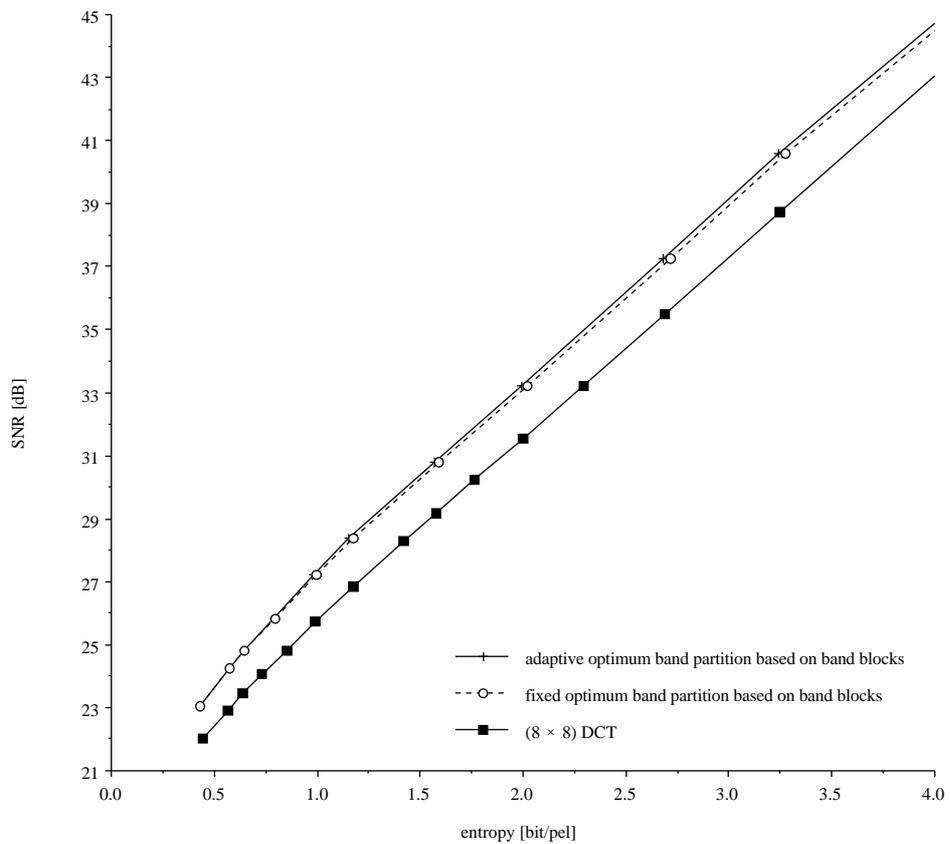


図 4.13 SNR 対エントロピー (画像 "Tulip")

図4.13において、帯域ブロックをベースとする最適帯域分割による SNR 特性は、固定及び適応両分割共に、同じビットレートの(8×8)画素の DCTの SNR より1 ~ 2[dB]程度優れた値を示している。このとき、固定分割と適応分割における差は、前節に示した画像 "Hada" の場合よりは若干大きくなっているものの、両分割方式による特性の差は未だ無視できるほど小さく、この場合においてもまた、適応分割を行わなくとも、固定な最適帯域分割パターンによってDCTに比べて十分な符号化性能を得ることが可能であり、本適応分割を固定分割によって近似することができる。このとき、更なる特性の改善を行うために、適応分割が有効であると考えられる。

4.4.2 帯域分割数の増加による対応

画像においては、その自己相関特性が更に特別な形を有することに起因して、前節に述べた固定の最適帯域分割と同一の帯域分割パラメータ条件を用いて、各画像に固有な最適帯域分割を得るために分割パターンのみを可変とした適応分割では対応できないものがある。上記の画像を対象とするとき、本節では、帯域分割パラメータの1つである帯域分割数 M の値を増加させることで対応可能な画像について考える。このような画像の代表として、図4.14に画像内容とその自己相関特性を示す画像 "Lenna" を例に挙げる。その自己相関特性においては、相関特性の変化が、原点に対して凸状に分布するという性質を残しているにとどまり、等方性相関モデルで近似を行うにはかなり困難な特殊な形の特性であることが分かる。なお、実験に使用した他のテスト画像の中で、前章の図3.10 (f)の画像 "Barbara"、同図(i)の画像 "Cameraman" が、この分類に属する画像になる。

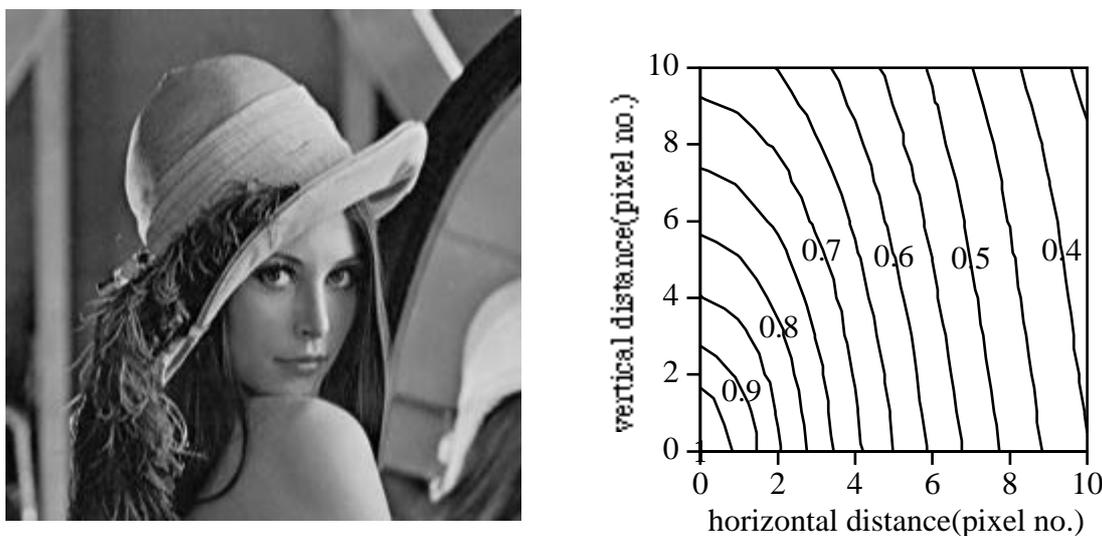


図 4.14 画像 "Lenna" とその自己相関特性

これらの画像に対して、帯域分割数 M を増加させることで、量子化雑音改善量 G の改善を試みるわけであるが、ここではまず、十分な改善量値を得るに必要な帯域分割数 M の選定を行う。上記のテスト画像に対して、帯域分割数 M を 4 から 5 に増加させた場合に、各画像に適応分割を適用することで、量子化雑音改善量において約 $0.4[dB]$ の比較的大きな改善が得られる。しかし、帯域分割数 M を更に 6 に増加させたとしても、量子化雑音改善量の改善は僅か $0.1[dB]$ にすぎない。この結果、帯域分割数 M の増加に伴う最適解の探索及び量子化、符号化に要する処理コストの増加を踏まえて、簡易的な符号化システムの実現を考慮することで、本適応分割において必要とされる帯域分割数として、 $M=5$ を選定する。

画像 "Lenna" に対して、本節において選定された帯域分割数 $M=5$ 、帯域ブロック数 $N^2=16$ の条件下で適応分割を適用することによって得られた帯域分割パターンを図 4.15 に示す。本適応分割後に 4.3.2.1 節に述べた DFT を用いた非定常直流成分の分離法を適用した場合に、画像 "Lenna" においては量子化雑音改善量 $G=17.974[dB]$ の値を得ることになり、表 4.1 に示された (8×8) 画素の DCT による改善量値を上回ることができる。また、固定の最適帯域分割と比較した場合、 $3.54[dB]$ もの大きな改善が行われていることになる。

本適応分割においては、4.5.1 節に述べた場合に比べて帯域分割数 M が 1 だけ増加しているために、式 (4.4) の $G(J^M)$ を最大にする最適解の探索回数が増加する。しかし、本適応分割に要する探索回数は 364 回であり、これによる帯域分割処理に要するコストは、固定の最適帯域分割の場合に比して 1.05 倍にすぎない。また、画像 "Lenna" の場合に、本適応分割は、4 分割の linear 分割に比して、上述の最適解の探索を含めた帯域分割に要する処理コストが 2.10 倍に、量子化及び符号化に要する処理コストが 1.25 倍に増加するという条件で、量子化雑音改善量 G において、 $4.64[dB]$ の非常に大きな改善を得ることができる。また、16 分割の linear 分割と比較した場合には、帯域分割に要する処理コストは僅か 1.05 倍だけ増加するが、それに対して量子化及び符号化に要する処理コストは 0.31 倍に抑えられる。このときの G の低減は僅か $0.29[dB]$ にすぎない。

更に、本適応帯域分割においても、画像毎に求められた固有の帯域分割パターンを表す情報をサイド情報として受信側に伝送する必要がある。しかし、 (256×256) 画素の画像に対して、各画像の最適帯域分割パターンを表すのに必要なビット数は、約 $0.0007[bit/pel]$ となり、これは無視できるほど小さい値になる。

最後に、画像 "Lenna" に本適応分割を適用した場合の SNR 対エントロピー特性を図 4.16 に示す。同図には、本適応分割による改善効果を確認するために、固定の最適帯域分割を適用した場合、及び比較方式として (8×8) 画素の DCT を適用した場合の同特性を併せて表示している。図 4.16 において、帯域ブロックをベースとする最適帯域分割による SNR 特性は、固定分割の場合において、同じビットレートの (8×8) 画素の DCT の SNR より $2[dB]$ 程度優れた値を示し、適応分割を行うことにより、更なる特性の改善が得られることが明らかになった。

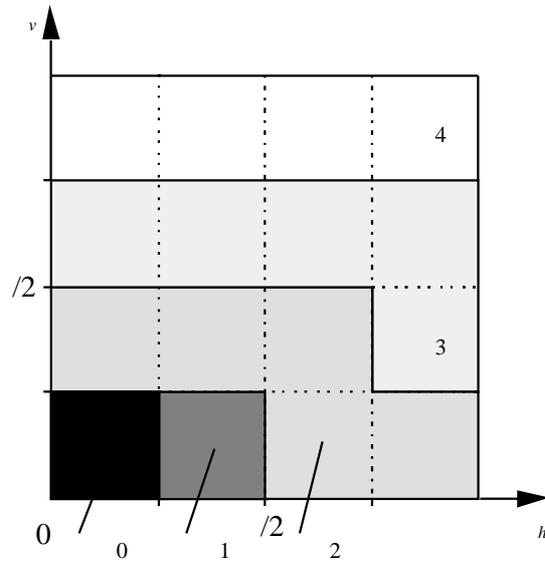


図 4.15 帯域ブロックをベースとする適応帯域分割パターン
(画像 "Lenna", $M=5$, $N^2=16$)

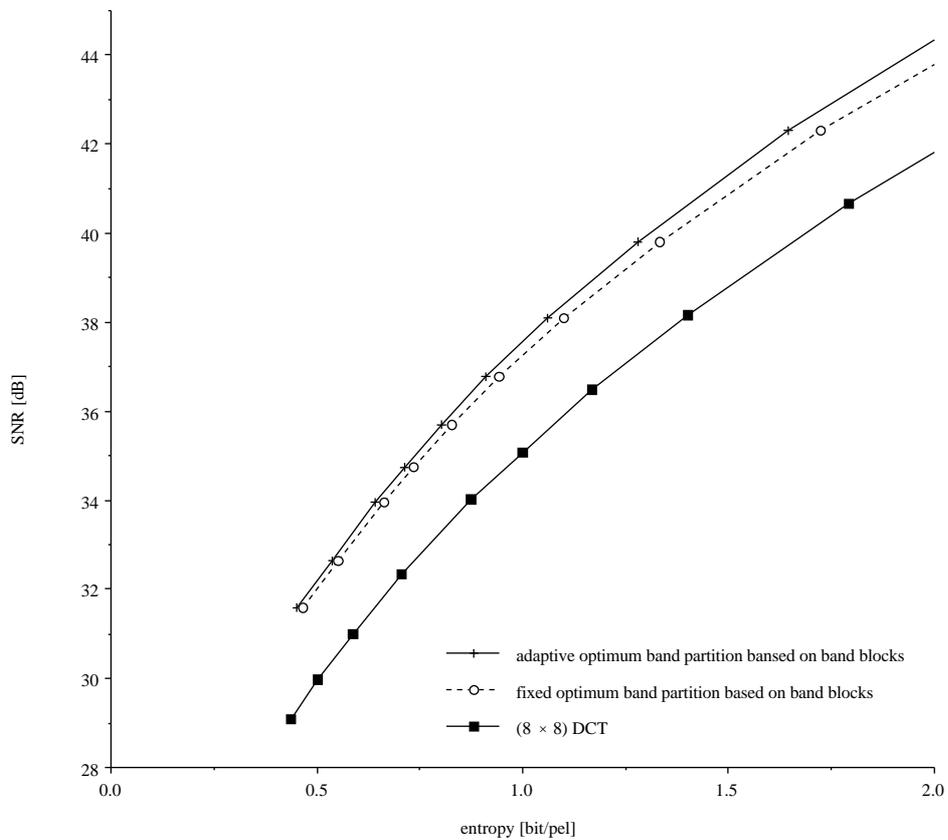


図 4.16 SNR 対エントロピー (画像 "Lenna")

4.4.3 帯域ブロック数の増加による対応

4.5.2節に述べた帯域分割数 M を増加した適応分割によっても十分な量子化雑音改善量を得られない画像については、もう1つの帯域分割パラメータである帯域ブロック数 N^2 を増加することによって対応する。図4.17に画像内容とその自己相関特性を示す画像"Building"において、その自己相関特性は、等方性相関モデルでの近似を議論するまでもない程の非常に複雑な形を有しており、一般的な画像の自己相関特性の性質に見られるような原点に対して凸状に分布する相関特性の変化さえ確認できない。

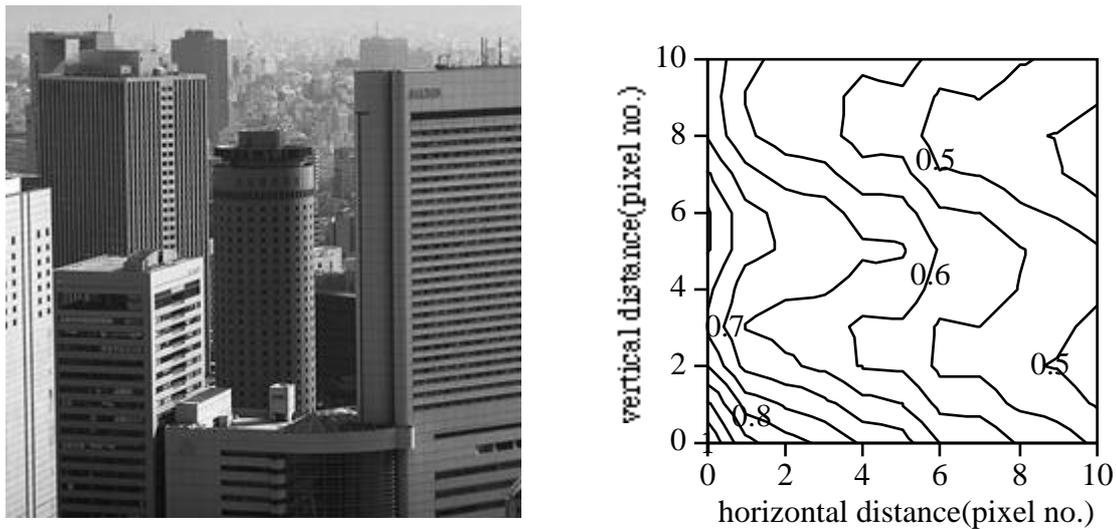


図 4.17 画像 "Building" とその自己相関特性

図4.17の画像に対しては、帯域ブロック数 N^2 を増加させることで、量子化雑音改善量 G の改善を得る。ここではまず、固定の最適帯域分割の条件を用いて、帯域分割数 $M=4$ とした場合に、図4.5に示したフィルタバンクの構成に基づいて、帯域ブロック数 N^2 を16から64に増加することを考える。上記の画像に対して、帯域分割数 $M=4$ 、帯域ブロック数 $N^2=64$ なる条件で適応分割を適用することで、4.5.1節で述べた帯域分割数 $M=4$ 、帯域ブロック数 $N^2=16$ の同画像に対する適応分割に比して、量子化雑音改善量において約 $1.0[dB]$ の大きな改善が得られる。次に、ここに述べた帯域ブロック数 $N^2=64$ なる条件のもとで、帯域分割数 M の増加による量子化雑音改善量の改善効果について考察を行う。このとき、帯域分割数 M を4から5に増加させた場合に、量子化雑音改善量の改善は僅か $0.3[dB]$ にすぎなかった。この結果より、帯域ブロック数 N^2 が64にまで増加していることを踏まえて、帯域分割数 M の増加に伴う最適解の探索及び量子化、符号化に要する処理コストの増加を考慮することで、本適応分割において必要とされる帯域ブロック数は $N^2=64$ を、帯域分割数は、 $M=4$ を選定する。

画像 "Building" に対して、本節において選定された帯域分割数 $M=4$ 、帯域ブロック数 $N^2=64$ の条件下で適応分割を適用することによって得られた帯域分割パターンを図4.18に示す。本適応分割後に4.3.2.1節に述べたDFTを用いた非定常直流成分の分離法を適用した場合に、量子化雑音改善量 $G=11.604[dB]$ の値を得ることになり、表4.1に示された(8×8)画素のDCTによる改善量値を上回ることができる。また、固定の最適帯域分割と比較した場合、本適応分割により $3.93[dB]$ もの大きな改善が行われていることになる。

本適応分割の実行においては、まず、64個の帯域ブロックを得るために、図4.5のフィルタバンクのtreeの段数を前節までに述べた場合と比べて1段増加し、3段のtree構成とする必要がある。また、帯域分割数 N^2 が64に増加しているために、式(4.4)の $G(J^M)$ を最大にする最適解の探索回数は更に増加し、3782回となる。これら全てを含めた帯域分割処理に要するコストは、固定の最適帯域分割の場合に比して2.0倍となる。また、画像 "Building" の場合に、本適応分割は、4分割のlinear分割に比して、帯域分割に要する処理コストが4.0倍になるものの、量子化及び符号化に要する処理コストは変わらないという条件で、量子化雑音改善量 G において、 $3.23[dB]$ の大きな改善を得ることができる。また、16分割のlinear分割と比較した場合には、帯域分割に要する処理コストは2.0倍だけ増加するが、それに対して量子化及び符号化に要する処理コストは0.25倍に抑えられるにもかかわらず、このときの G の値は、 $064[dB]$ も改善されている。

更に、本適応帯域分割においても、固有の帯域分割パターンを表す情報をサイド情報として受信側に伝送する必要がある。しかし、(256×256)画素の画像に対して、各画像の最適帯域分割パターンを表すのに必要なビット数は、僅か約 $0.0019[bit/pel]$ となる。

最後に、画像 "Building" に本適応分割を適用した場合の SNR 対エントロピー特性を図4.19に示す。同図には、本分割による改善効果を確認するために、固定の最適帯域分割を適用した場合、及び比較方式として(8×8)画素のDCTを適用した場合の同特性を併せて表示している。図4.19において、帯域ブロックをベースとする最適帯域分割による SNR 特性は、固定分割の場合に、同じビットレートの(8×8)画素のDCTの SNR 特性とほぼ同等性能を有しているが、本適応分割の特性は、これらの特性から更に $1[dB]$ 程度優れた値を示し、適応分割により効果的な特性の改善が得られることが明らかになった。

本節を終えるにあたって、帯域ブロックをベースとする最適帯域分割に基づいた固定及び適応分割の適用による再生画像と、それらの比較方式である(8×8)画素のDCTを適用したときの再生画像に生じる符号化ひずみについて簡単な考察を行う。エントロピーレートが低い場合に、DCTによる再生画像には、画面全体においてブロックひずみが現れており、それに起因する画質劣化が著しい。一方、本最適帯域分割による再生画像には、同等のエントロピーレートにおいて、サブバンド符号化特有の符号化ひずみである輪郭部周辺のリングングが若干現れているものの、DCTの場合に比べて主観的に比較的良好な再生画像を得ることができる。

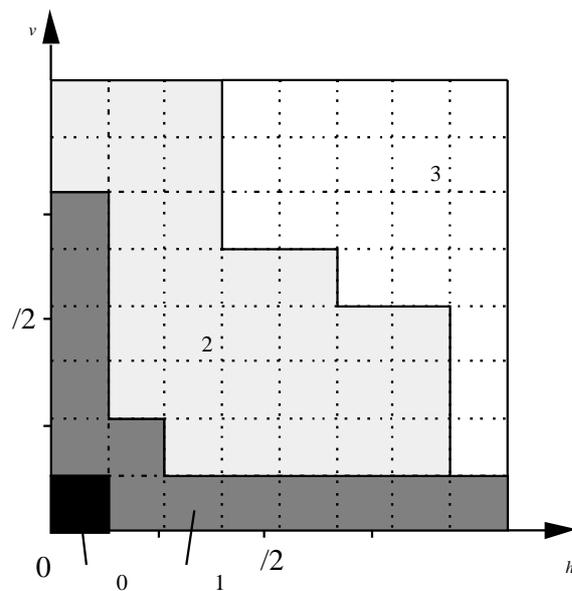


図 4.18 帯域ブロックをベースとする適応帯域分割パターン
(画像 "Building", $M=5$, $N^2=64$)

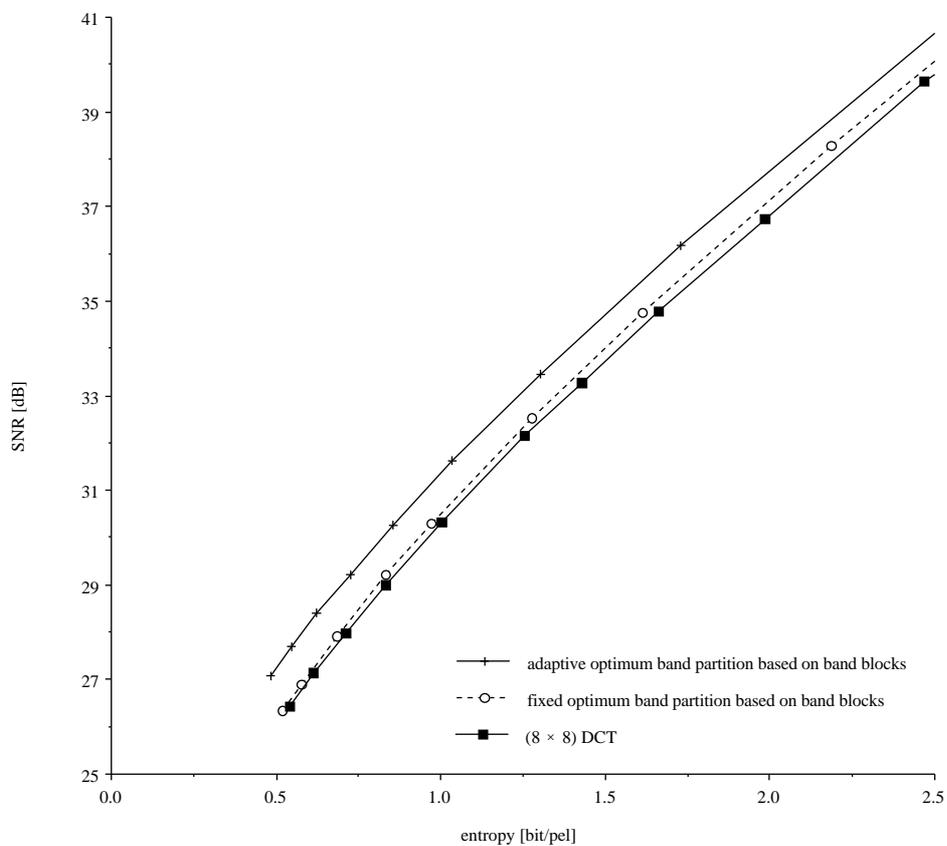


図 4.19 SNR 対エントロピー (画像 "Building")

4.5 結言

本章では、サブバンド画像符号化システムの簡易化・高性能化を目指し、第3章で理論的に導出された最適帯域分割特性を実現するために、少ない帯域分割数で、高い量子化雑音改善量を達成できる帯域ブロックをベースにした2次元最適周波数分割法を提案し、画像に適用してその有効性を明らかにした。まず、多くの一般的な画像に適合する等方性相関特性を有する水平・垂直相関非分離型画像モデルを対象に、量子化雑音改善量を最大にする帯域ブロックをベースとする最適帯域分割法を提案し、それに基づいてフィルタバンクを実用する上で、効果的な量子化雑音の改善が得られる帯域分割パラメータとして、帯域分割数=4、帯域ブロック数=16を選定した。次に、本最適帯域分割を実現するフィルタバンクの構成法を示し、本フィルタバンクを実際の画像に適用して、量子化雑音改善効果を確認した。以上は、画像モデルにおいて最適化を行った帯域ブロックをベースとする固定な最適帯域分割であるが、実際の画像には上記画像モデルに含まれない非定常性として、直流成分と相関特性の画像内容、或いは空間的な場所による変動がある。このため、前者に対しては、最低域の帯域分割信号に対して(2×2)画素のDFTを適用して非定常直流成分の分離を行う方法を、後者に対しては、本来は各画像に対して固有なものになる最適帯域分割を得るために、帯域ブロックをベースとする適応帯域分割法を提案した。汎用性を高めることを考慮して選ばれた複数の画像について、固定及び適応両帯域分割方式による量子化雑音改善量特性の比較評価実験を行い、その結果、大半の画像については、固定最適帯域分割で十分の改善が得られるが、本画像モデルでの近似が特に困難である画像については、適応分割が有効であることを明らかにした。また、たとえ適応分割を行う場合においても、その実現に要する処理コストの増加と伝送しなければならないサイド情報の増加は小さくなることを示した。最後に、帯域ブロックをベースとする最適帯域分割における符号化性能を調べるために、SNRとエントロピーの関係を求め、JPEGで用いられている(8×8)画素のDCTとの比較を行った。この結果、本最適帯域分割は、同じビットレートのDCTに比べて1~2[dB]程度のSNRの改善が得られることを確認した。このことは、画像に対して最適な帯域分割を実現することで、(8×8)画素のDCT、すなわち64分割のDCTと比して、はるかに小さな帯域分割数であるにもかかわらず、DCTによる符号化性能を優ることができていることを意味しており、画像に対して適切な帯域分割を行うことの重要性を示唆すると共に、本章において提案された帯域ブロックをベースとする最適帯域分割の有効性を明らかにすることができた。

本章までの議論で明らかにされた最適帯域分割及び帯域ブロックをベースとする最適帯域分割は、量子化雑音改善量なる評価尺度に基づいて導出されたものであるから、その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、上記画質の観点で良好な再生画像を得ることを目的として、視覚の空間周波数特性を考慮した最適帯域分割について議論を行うことにする。

第 5 章

視知覚特性を考慮した画像信号の 2次元最適帯域分割

5.1 序言

画像のサブバンド符号化において，入力信号（画像）に整合した帯域分割の最適化を実現するために，第3章においては，実画像の統計的性質である自己相関特性に基づいて画像モデルを導入することで，最適帯域分割を理論的に導出した．また，第4章においては，最適帯域分割を実現するフィルタバンクの構成を明らかにすると共に，実画像への適用における本最適帯域分割の有効性を示した．しかし，前章までに導出された最適帯域分割は，量子化雑音改善量なる SNR 値に基づいた評価値を最大とする帯域分割であることから，視知覚特性を考慮した画質という点では，真に最適な帯域分割が行われているとは言い難い．画像符号化における原点は，高い圧縮効率のもとで，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．

従来より人間の視知覚特性を考慮した画像符号化方式の検討が行われている[68]～[71]．本研究で対象としているサブバンド符号化においては，JPEG等の空間領域のブロック処理において生じる視知覚上最大の劣化要因となるブロックひずみの発生はなく，ランダム状の符号化雑音とフィルタのタップ長に依存して再生画像の輪郭部周辺に発生するリングングが画質劣化の主たる要因となる．

本章では，サブバンド符号化に，人間の視知覚特性を導入するための第一段階として，ランダム状の符号化雑音に着目した上で，視覚の空間周波数特性を考慮した符号化方式を提案し，その議論において必要となる視覚の空間周波数特性を考慮した最適帯域分割の導出を行う．上記の文献に示された従来の符号化方式においては，この視覚の空間周波数特性は量子化ビット配分についてのみ考慮されており，本研究のように帯域の分割パターンにまで踏み込んだ例は他にない．量子化のみならず，帯域の分割の仕方に対して視覚の空間周波数特性を考慮することにより，特に，視知覚特性を考慮したランダムノイズ削減の観点において再生画像の更なる画質改善が期待される．

まず、本章において用いられる視覚の空間周波数特性を簡単に紹介する。次に、フィルタバンクの実現可能条件を考慮して、第4章で詳述された手法に基づいて、帯域ブロックをベースに、視覚の空間周波数特性を考慮し得る最適帯域分割法を提案すると共に、本最適帯域分割を実現するフィルタバンクの構成を明らかにする。ここでは、汎用性の向上を目的として、その内容が異なる5枚の画像を用いて、視覚の空間周波数特性を考慮した最適帯域分割フィルタバンクを実用する上で望ましい帯域分割パラメータの選定を行い、これらに基づいた視覚の空間周波数特性を考慮した最適帯域分割法を適用することにより、各画像毎に固有な分割パターンとして、2次元周波数平面上での最適帯域分割を明らかにする。最後に、視覚の空間周波数特性で重み付けられた量子化雑音電力($WMSE$)とエントロピーの関係を求め、本最適帯域分割の適用による視覚の空間周波数特性を考慮したランダムノイズ低減における定量的評価を行う。しかし、本最適帯域分割による再生画像の画質を視知覚特性までを考慮して評価するにあたっては、上述した $WMSE$ だけでは不十分である。例えば、再生画像において、ランダムノイズとは異なるテクスチャ状のパターンノイズが現れている場合には、このようなノイズは、ランダムノイズに比べて約10倍も知覚されやすく、視知覚的に大きな妨害となる[39],[40]。そこで、符号化方式及び画像内容に依存することなくオーバーオール画質を評価することを目的として、主観評価実験を行い、本最適帯域分割による再生画像の画質を評価すると共に、視覚の空間周波数特性を考慮しない場合の最適帯域分割方式及びJPEGで用いられている(8×8)画素のブロック単位で適用されるDCTによる再生画像との比較を行う。

5. 2 視覚の空間周波数特性

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も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[72]。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分に対する感度特性を求めることができない。そこで、本章では、文献[71]、[73]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、他のモデルに比べて特性の低域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2次元空間周波数特性について簡単に述べることにする。

文献[71]、[73]において引用されている空間周波数特性 $S(f)$ は、式(5.1)で表される。

$$S(f) = 1.5 \exp\left(-\frac{f^2}{2}\right) - \exp\left(-2f^2\right) \quad (5.1)$$

ただし、 $f = \sqrt{u^2 + v^2}$ 、 $f = f/60$ 、 $f = (u^2 + v^2)^{1/2}$ (f : cycle/degree) 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、文献[55]によれば、視覚の2次元空間周波数特性には解像度方向性があり、具体的には45度傾いたパターンに対する視覚感度は、水平、垂直方向に比べて約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

$$O(\theta) = \begin{cases} 1.0 & : f < f_p \\ 0.5(1 + \cos^4 2\theta) & : f \geq f_p \end{cases} \quad (5.2)$$

で近似することができる。ここで、 $\theta = \tan^{-1}(v/u)$ であり、 f_p は、式(5.1)の $S(f)$ が最大レスポンスとなる空間周波数を表す。

よって、異方特性を有する視覚の2次元空間周波数特性 $W(u, v)$ は、式(5.1) (5.2)に示された関数の積として、

$$W(u, v) = S(f) O(\theta) \quad (5.3)$$

により定義することができる。次節以降では、視覚の空間周波数特性として、式(5.3)に示す $W(u, v)$ を用いるものとする。

5.3 視覚の空間周波数特性を考慮した最適帯域分割

M 分割のサブバンド分割によって得られる帯域 k ($k=0,1,\dots,M-1$) において, k 番目の帯域 k におけるスペクトルを, 前節に示した水平及び垂直方向の空間周波数 u, v を用いて, $x_k(u,v)$ ($k=0,1,\dots,M-1$) と表すものとする. この $x_k(u,v)$ に対して, 視覚特性を考慮した画質の議論を行うにあたって, ここでは, その第一段階としてランダム状の符号化雑音に着目する. そのために, 視覚の空間周波数特性を考慮した量子化を行う際に, 送信側において, 式(5.3)に示した視覚の空間周波数特性 $W(u,v)$ で重み付けした $W(u,v)x_k(u,v)$ なる信号を量子化した後, 受信側において, 復号値を $1/W(u,v)$ 倍して再生すればよい. 本処理によって, 信号成分は元の値 $x_k(u,v)$ が再生されるが, 雑音成分のみが $1/W(u,v)$ 倍され, 視覚の空間周波数特性を考慮した量子化が可能となる.

このとき, $x_k(u,v)$ なる信号系列について考える. 例えば, 第4章において詳述された量子化雑音改善量 G を最大にする条件を満足するように求められた帯域ブロックをベースとする最適帯域分割は, そこで得られた各帯域 k ($k=0,1,\dots,M-1$) に分散のほぼ等しい帯域ブロックが寄せ集められる. しかし, この $x_k(u,v)$ を量子化するために $W(u,v)$ で重み付けられた $W(u,v)x_k(u,v)$ の分散は, 同じ帯域内で大きく変動してしまい, これらを1つの符号化器で量子化, エントロピー符号化を行うことは極めて不利である. そこで, 同一帯域内での信号の分散が小さくなるように, 視覚の空間周波数特性を考慮した最適帯域分割を導入し直す必要がある.

5.3.1 帯域ブロックをベースとする視覚の空間周波数特性を考慮した最適帯域分割

視覚の空間周波数特性を考慮した画像の最適帯域分割は, 入力画像の電力スペクトル $P_x(u,v)$ が与えられた場合に, 信号全体に割り当てられるビットレート一定の条件のもとで, 再生画像において $W(u,v)$ で重み付けられた量子化雑音電力 ($WMSE$) を最小にするような互いに素な M 個の帯域 k ($k=0,1,\dots,M-1$) に分割することで定義される. これは, 次式で定義される重み付き量子化雑音改善量 $G_{(w)}$ を最大にする分割と等価である.

$$G_{(w)} = \frac{\sum_{k=0}^{M-1} \frac{2}{k(W)}^2}{\sum_{k=0}^{M-1} \frac{2}{k(W)}^k} \quad (5.4)$$

ただし, $\frac{2}{k(W)}$ 及び $\frac{2}{k(W)}^k$ は, 帯域 k の出力レート及び視覚の空間周波数特性によって重み付けられた帯域 k の重み付き信号電力であり, それぞれ次式で表される.

$$k = \frac{1}{2} \int_k dudv \quad (5.5)$$

$$G_{k(W)}^2 = \frac{1}{2} \int_k W^2(u,v) P_x(u,v) dudv \quad (5.6)$$

今、視覚の空間周波数特性を考慮した最適帯域分割を導出するにあたり、フィルタバンクの実現性を優先させるために、式(5.4)の $G_{(W)}$ に若干の犠牲を払うことで、帯域分割パターンに以下のような制限を加える。まず、2次元周波数帯域を水平及び垂直方向に N 分割することで、 N^2 個の帯域ブロック

$$i,j = (u,v) \mid \frac{i}{N} \leq u < \frac{(i+1)}{N}, \frac{j}{N} \leq v < \frac{(j+1)}{N} \quad (5.7)$$

$$(i,j) \in I \quad (I = \{(i,j) \mid i,j = 0,1, \dots, N-1\})$$

と呼ばれる周波数平面上で定義される小領域に分割する。次に、次式に示すように、この帯域ブロックの集合の直和分割によって、 J_k を M 個の帯域 J_k ($k=0,1,\dots,M-1$)に分割する。ただし、 M は $1 \leq M \leq N^2$ を満たす整数である。

$$J_k = \bigcup_{i,j} J_{k(W)}^{(M)}(i,j) \quad k=0,1,\dots,M-1 \quad (5.8)$$

上式において、 $J_{k(W)}^{(M)}$ は J_k に含まれる帯域ブロックの番号 (i,j) の集合を表し、これが J_k における帯域分割パターンを定めることになる。このとき、式(5.4)の $G_{(W)}$ を最大とするような分割 $J_{opt(W)}^{(M)}$ が、帯域ブロックをベースとする視覚の空間周波数特性を考慮した最適帯域分割となる。

N, M の値が与えられたとき、ある特定の帯域分割が定められた場合には、式(5.4)~(5.6)を用いることで本分割における視覚の空間周波数特性によって重み付けられた量子化雑音改善量 $G_{(W)}(J_{k(W)}^{(M)})$ を求めることができる。しかし、逆に $G_{(W)}(J_{k(W)}^{(M)})$ を最大にするような分割 $J_{opt(W)}^{(M)}$ を直接求めることは困難である。そこで、 $G_{(W)}$ の性質から分割の探索範囲を大幅に制限し、その中から最適解を求めることとする。最適解の導出における詳細は前章4.2.2節に譲るが、具体的には、各帯域ブロックの番号 $(i,j) \in I$ を、視覚の空間周波数特性 $W(u,v)$ によって重み付けられた各帯域ブロックにおける信号電力 $P_{i,j(W)}^2$ の値が大きい順に並べ、 $\{(i,j)_0, (i,j)_1, \dots, (i,j)_{N^2-1}\}$ となったとすると(第 m 番目の番号を $(i,j)_m$ と表示する)、各 $J_{k(W)}^{(M)}$ ($k=0,1,\dots,M-1$)が隣接する番号のみを含むような

$$J^{(M)} = \underbrace{\{(i,j)_0, (i,j)_1, \dots, (i,j)_{m-1}\}}_{(J_{0(W)}^{(M)})}, \underbrace{\{(i,j)_m, (i,j)_{m+1}, \dots, (i,j)_{m+n-1}\}}_{(J_{1(W)}^{(M)})}, \dots, \underbrace{\{(i,j)_{N^2-1}\}}_{(J_{M-1(W)}^{(M)})} \quad (5.9)$$

なる分割の中に $J_{opt(W)}^{(M)}$ が存在することになる。

5.3.2 視覚の空間周波数特性を考慮した最適帯域分割 フィルタバンクの構成

式(5.7)で定義された帯域ブロックを得るための具体的な手法を提案し、本手法に基づいて、視覚の空間周波数特性を考慮した最適帯域分割フィルタバンクを構成する。視覚の空間周波数特性 $W(u,v)$ は、2次元周波数帯域上で定義されている。それ故、入力画像の電力スペクトル $P_x(u,v)$ を求め、 $W^2(u,v)$ を乗じることで、視覚の空間周波数特性で重み付けられた画像の電力スペクトルを得ることができる。ここで、画像のような有限長の信号系列のスペクトルの算出にあたっては、系列の左右（上下）両端の外で突然に値をゼロにすると、その結果スペクトルが広がることになる。そこで、 $(K \times K)$ 画素の入力画像 $x(m,n)$ ($m,n=0,1,\dots,K-1$) に対して、画像の左右（上下）を滑らかに接続するため、次式に示すように、画像端でmirror symmetry条件を考慮して、更に水平及び垂直方向に周期 $2K$ の周期性を仮定した画像

$$\tilde{x}(m,n) = \begin{cases} \tilde{x}(m,n) & : 0 \leq m < K, 0 \leq n < K \\ \tilde{x}(2K-1-m,n) & : K \leq m < 2K, 0 \leq n < K \\ \tilde{x}(m,2K-1-n) & : 0 \leq m < K, K \leq n < 2K \\ \tilde{x}(2K-1-m,2K-1-n) & : K \leq m < 2K, K \leq n < 2K \end{cases} \quad (5.10)$$

ただし、

$$\begin{aligned} \tilde{x}(m,n) &= \tilde{x}(m+2K,n) \\ \tilde{x}(m,n) &= \tilde{x}(m,n+2K) \end{aligned} \quad (5.11)$$

を考え、この $\tilde{x}(m,n)$ の2次元片側電力スペクトルを求めると、次式となる。(付録C参照)

$$P_x(u,v) = \frac{(2K)^2}{K^2} \sum_{p,q=0}^{K-1} X^2(p,q) \cos\left(u - \frac{p}{K}\right) \cos\left(v - \frac{q}{K}\right) \quad (5.12)$$

ただし、

$$X(p,q) = \frac{2}{K} \sum_{m=0}^{K-1} \sum_{n=0}^{K-1} x(m,n) \cos\left(\frac{(2m+1)p}{2K}\right) \cos\left(\frac{(2n+1)q}{2K}\right) \quad (5.13)$$

$$k = \begin{cases} 1/\sqrt{2} & : k = 0 \pmod{2K} \\ 1 & : \text{otherwise} \end{cases} \quad (5.14)$$

上式より、画像の電力スペクトルは、空間周波数 $(\frac{p}{K}, \frac{q}{K})$ ($0 \leq p,q < K-1$) における線スペクトルとなり、そのスペクトルの電力は $X^2(p,q)$ となる。ここで、式(5.13)より、 $X(p,q)$ は、画像 $x(m,n)$ の2次元DCTに他ならない。(付録C参照)

以上より、帯域ブロックをベースとして、視覚の空間周波数特性を考慮した最適帯域

分割を実現し得るフィルタバンクの構成は図5.1のようになる。視覚の空間周波数特性を考慮した本最適帯域分割を得るためには、まず、入力された画像信号 $x(m,n)$ を2次元($K \times K$)点DCTを用いて周波数領域に変換した後、式(5.7)に基づいて2次元周波数帯域を帯域ブロック i_j に分割し、各帯域ブロックのスペクトル $x_{i,j}(p,q)$ ($i,j = 1, \dots, M$)、各ブロックのデータサイズ： $K/N \times K/N$)を得る。各帯域ブロックのスペクトルは、視覚の空間周波数特性によって $W(\frac{p}{K}, \frac{q}{K})x_{i,j}(p,q)$ のように重み付けられ、視覚の空間周波数特性によって重み付けられた信号電力 $\int \int |x_{i,j}(p,q)|^2 W(\frac{p}{K}, \frac{q}{K}) dp dq$ が求められる。このとき、5.3.1節で述べた本最適帯域分割の導出法に基づいて、帯域 k のスペクトル $x_k(p,q)$ は、本最適帯域分割 $J_{opt(W)}^{(M)}$ の k 番目の要素 $J_{k,opt(W)}^{(M)}$ に含まれる帯域ブロック番号 (i,j) に対応した信号の集合として

$$x_k(p,q) = \{x_{i,j}(p,q) \mid (i,j) \in J_{k,opt(W)}^{(M)}\} \quad (k = 0, 1, \dots, M-1) \quad (5.15)$$

のように得ることができる。その後スペクトル $x_k(p,q)$ は、視覚の空間周波数特性によって $W(\frac{p}{K}, \frac{q}{K})x_k(p,q)$ のように重み付けられ、各帯域毎に量子化及びエントロピー符号化される。受信側では、復号信号を $1/W(u,v)$ 倍した後、各帯域を合成し、2次元($K \times K$)点IDCTを適用することで画像を再構成することができる。

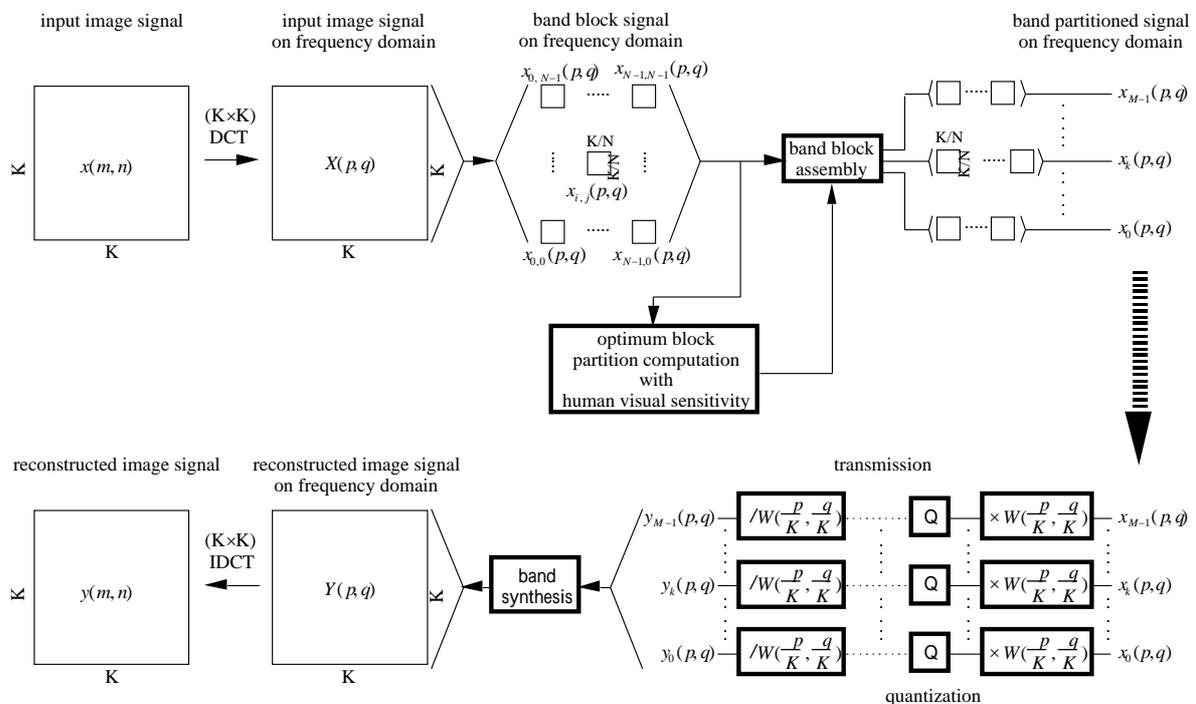


図 5.1 帯域ブロックをベースとする視覚の空間周波数特性を考慮した最適帯域分割を実現するフィルタバンクの構成

最後に，本方式における利点を挙げておく．

- ・変換時にmirror symmetry条件が考慮されているので，画像端における偽エッジの発生がない．
- ・周波数領域で処理を行うために，視覚の空間周波数特性による正確な重み付けを実現できる．
- ・変換後の係数を全て実数で扱うことができるために，後の量子化を簡易的に実現することができる．
- ・DCTは，画像全体に対して適用されるので，ブロックひずみを発生しない．

5. 4 帯域ブロックをベースとする視覚の空間周波数特性を考慮した画像の最適帯域分割

5.4.1 帯域分割パラメータの選定

帯域ブロックをベースに，視覚の空間周波数特性を考慮した最適帯域分割を実現するにあたって，帯域分割数 M 及び帯域ブロック数 N^2 は，共に大きいほど重み付き量子化雑音改善量 $G_{(w)}$ の値は向上する．一方，処理コストの点からは，両者は共に小さい方が望ましい．また，各帯域ブロックにおいて，視覚の空間周波数特性によって重み付けられた信号電力 $\sigma_{ij}^2(w)$ の値を計算する場合に，その推定精度を高めるためには，1つの帯域ブロックに含まれるデータの個数をある程度を大きくとる必要がある[74]，その観点からもあまり大きな N^2 の値は選択すべきではない．

本来ならば，あらゆる種類の画像に対して汎用的な帯域分割パラメータを決定することが網羅性の観点からは望ましいわけであるが，膨大な数の画像全てについて実験を行うことは困難である．そこで，本節では，画像内容が大きく異なる図5.2(a)～(e)に示す5枚の画像を用いた．各画像は，いずれも(256×256)画素，濃淡8[bit/pel]のもので，人物画像(平坦部の多い比較的単純な画像とそうでないもの)として図5.2(a)の画像"Hada"(ITEデジタル標準画像ver.II[58])と図5.2(b)の画像"Lenna"を，風景画像(平坦部の多い比較的単純な画像とそうでないもの)として図5.2(c)の画像"Cameraman"と図5.2(d)の画像"Tulip"(ITEデジタル標準画像ver.II[58])を，静物画像として図5.2(e)の画像"Wine"を採択した．これらの画像を対象として，汎用的な帯域分割パラメータを求めることができれば，他の画像についても十分カバーできると考え $G_{(w)}$ をベースに基本特性を検討し，それに基づき両パラメータの実用上有効な値を選定する．

図5.3(a)～(e)は，それぞれ図5.2(a)～(e)に示した画像について，帯域ブロック数 N^2 を各々16及び64とした場合に，帯域分割数 M をパラメータとして，帯域ブロックをベー

スとする視覚の空間周波数特性を考慮した最適帯域分割による改善量特性 $G_{(w)}(J_{opt(w)}^{(M)})$ を求めたものである。なお、各特性図には、各画像における $G_{(w)}$ の理論限界値を推定するために、帯域ブロックのみによって分割が行われる場合の $G_{(w)}$ の値である $G_{(w)}(N^2)$ 特性を併せて表示している。

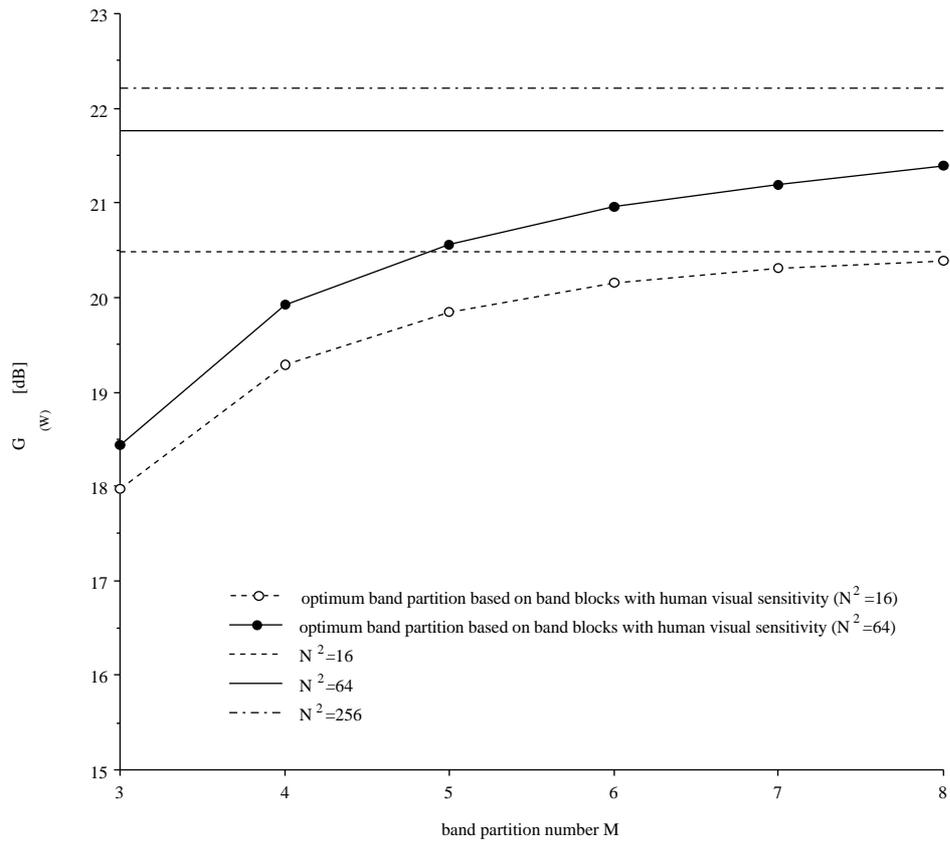
まず、所要とされる帯域ブロックの個数 N^2 について検討する。帯域ブロック数 N^2 に対する $G_{(w)}(N^2)$ は、 N^2 を16から64まで増加させた場合、その改善が最大である画像においては1.7[dB]、最小である画像においても0.8[dB]なる平均して1.3[dB]もの大きな改善がなされる。それに対して、更に N^2 を64から256にまで増加させても、それに伴う改善は全ての画像において僅か0.5[dB]前後にすぎない。なお、今回は、実画像を用いた検討のために、各画像に対する $G_{(w)}$ の理論限界値を正確に求めることはできないが、以上の結果から、 N^2 の値を今後更に増加させたとしても、 $G_{(w)}$ の値においてこれ以上の大きな改善が得られるとは考えにくい。よって、帯域ブロック数として、 $N^2=64$ を選定する。

次に、帯域分割数 M について検討を行う。上記で選定された $N^2=64$ の条件下で、各 M が与えられた場合に各画像において求められた $G_{(w)}(J_{opt(w)}^{(M)})$ 特性は、 M の増加に伴い、改善量の値は増大するものの、図5.3の各図共に、 M 6の範囲において飽和傾向が見られる。各図において、 M の値を64とした場合の $G_{(w)}(J_{opt(w)}^{(64)})$ なる値は、 $G_{(w)}(64)$ の値と一致することは明らかであるから、 $M>8$ の範囲においては、 $G_{(w)}(J_{opt(w)}^{(M)})$ 特性は $G_{(w)}(64)$ に対して、緩やかに漸近していくものと推測される。しかし、上述したように、符号化システム実現の簡易化のためには、帯域分割数 M はできるだけ小さい方が望ましいことから、ここでは、妥当と思われる所要帯域分割数として、 $M=8$ を選定する。全ての画像において、帯域ブロックをベースとする視覚の空間周波数特性を考慮した場合に、帯域分割数 M が8であるときの最適帯域分割による改善量 $G_{(w)}(J_{opt(w)}^{(8)})$ と64分割に相当する改善量 $G_{(w)}(64)$ との差は、平均して僅か0.5[dB]となっている。



(a) 画像 "Hada"

図 5.2 テスト画像



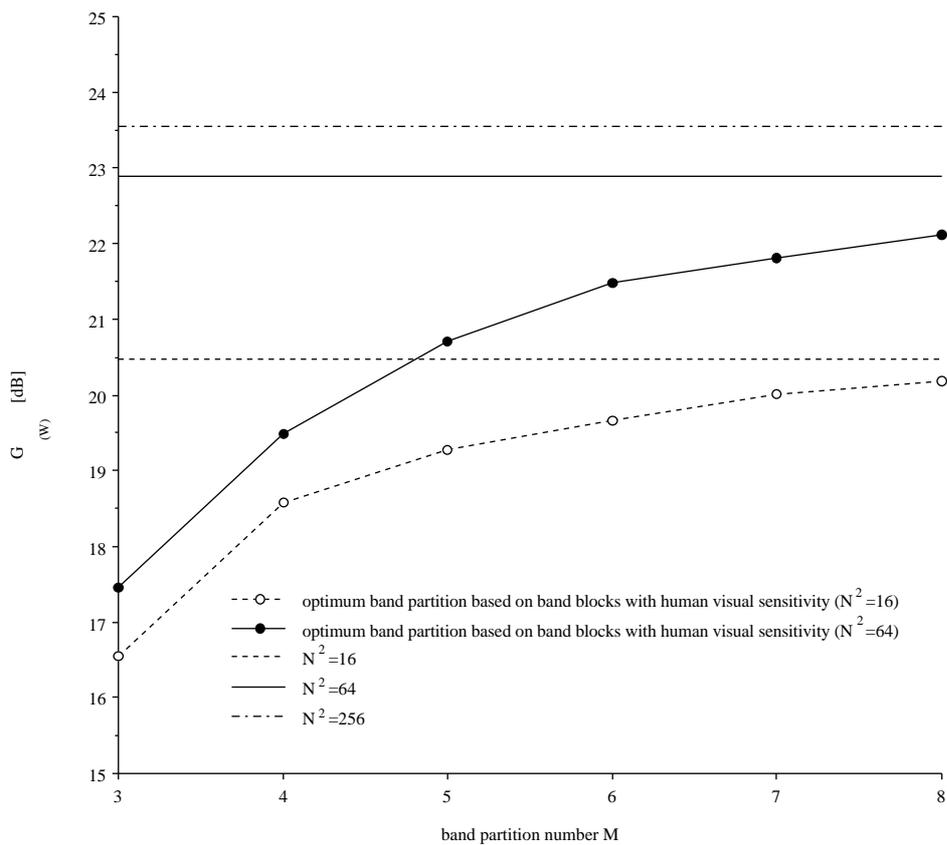
(a) 画像 "Hada"

図 5.3 帯域ブロックをベースとする視覚の空間周波数特性を考慮した最適帯域分割における $G_{(W)}$ (パラメータ: 帯域分割数 M)



(b) 画像 "Lenna"

図 5.2 テスト画像



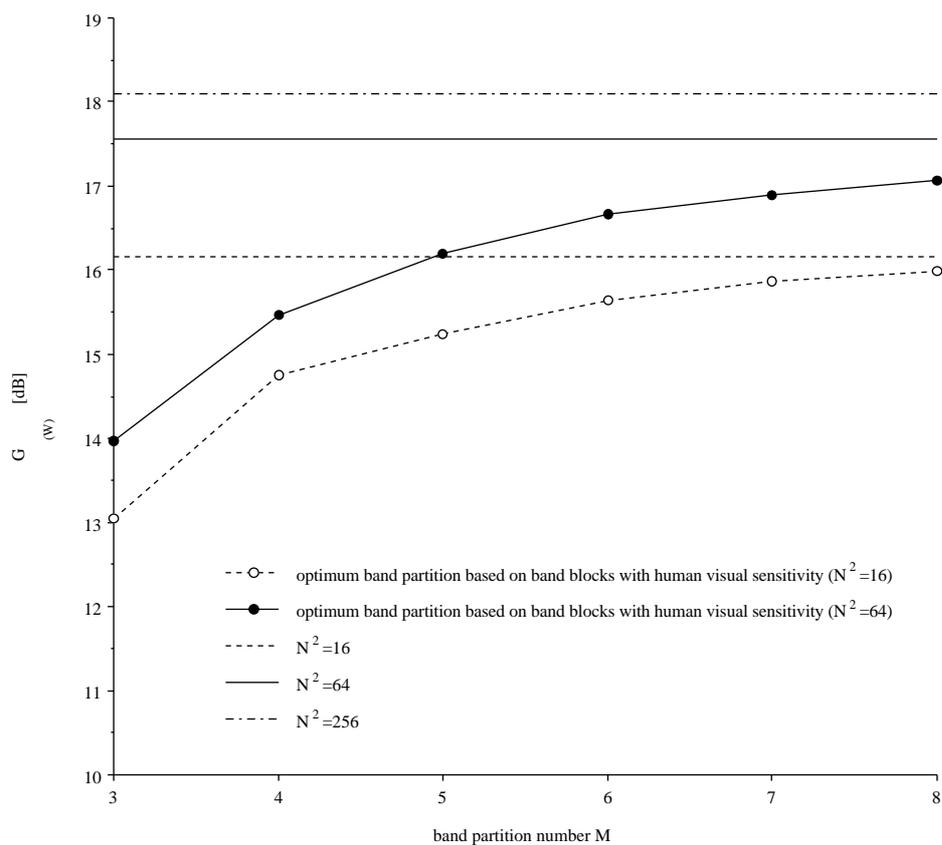
(b) 画像 "Lenna"

図 5.3 帯域ブロックをベースとする視覚の空間周波数特性を考慮した最適帯域分割における $G_{(w)}$ (パラメータ: 帯域分割数 M)



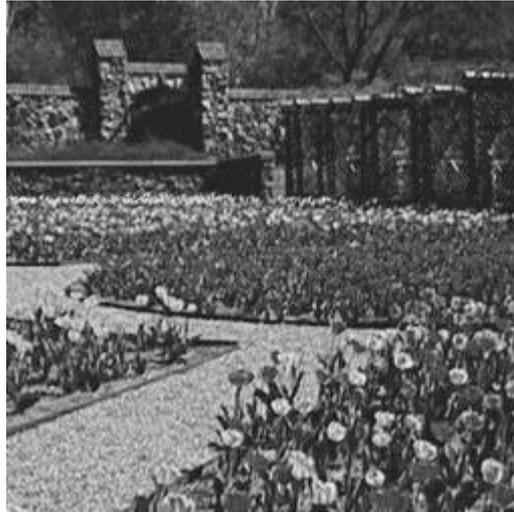
(c) 画像 "Cameraman"

図 5.2 テスト画像



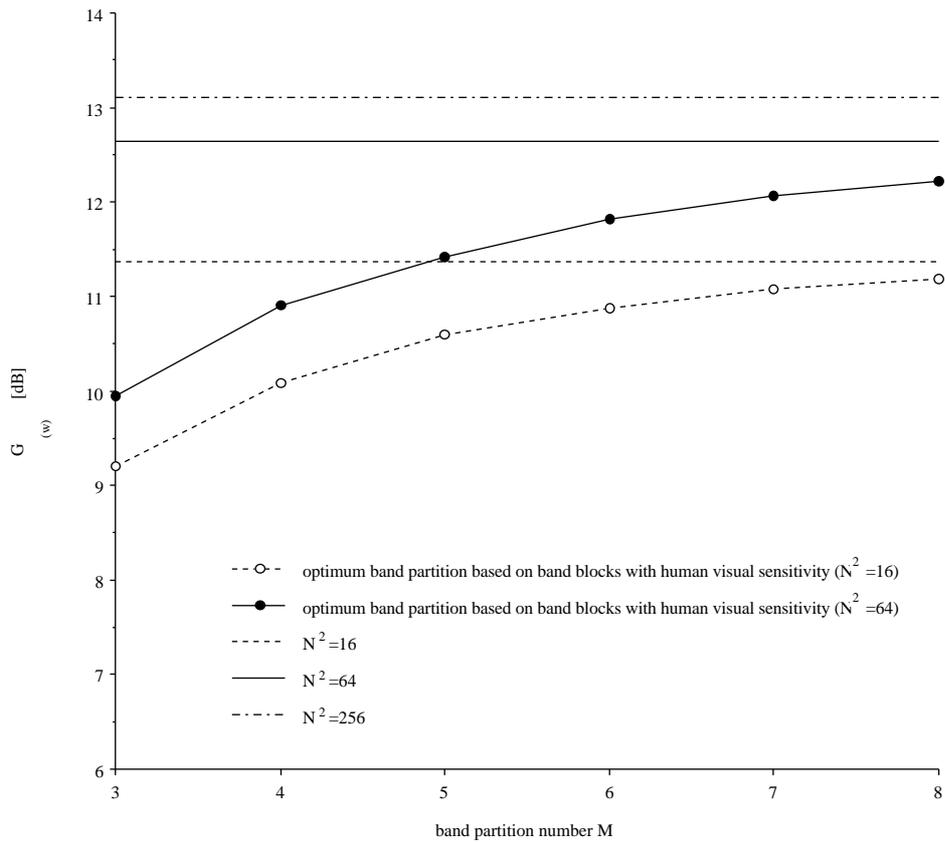
(c) 画像 "Cameraman"

図 5.3 帯域ブロックをベースとする視覚の空間周波数特性を考慮した最適帯域分割における $G_{(W)}$ (パラメータ: 帯域分割数 M)



(d) 画像 "Tulip"

図 5.2 テスト画像



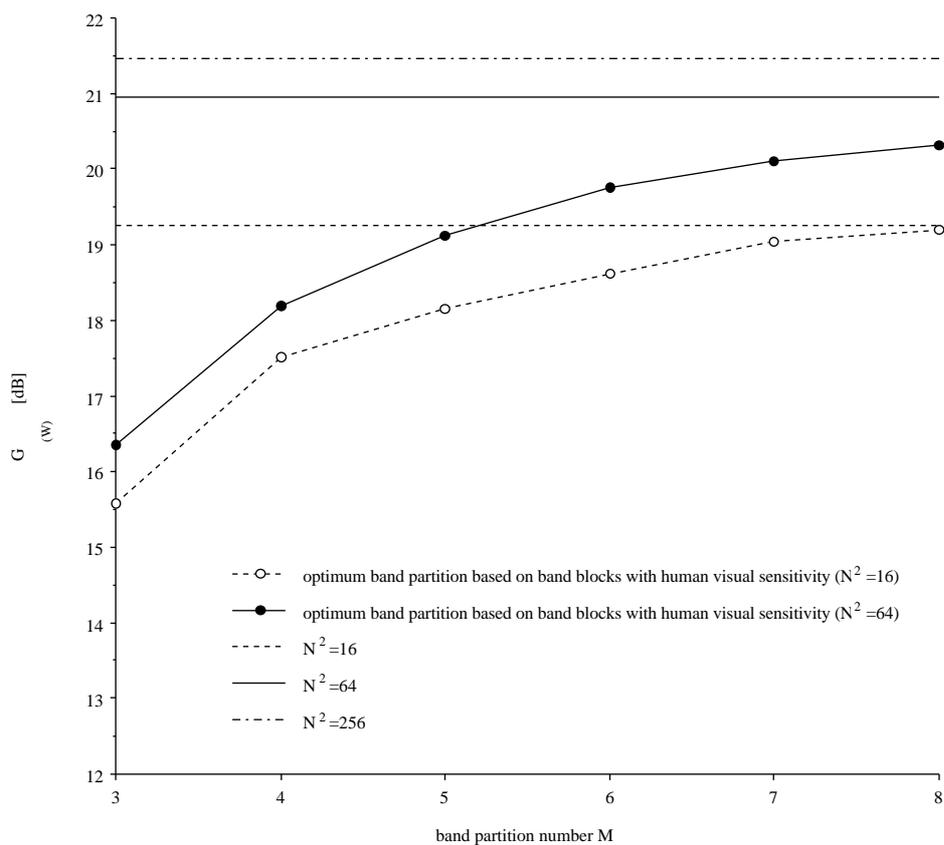
(d) 画像 "Tulip"

図 5.3 帯域ブロックをベースとする視覚の空間周波数特性を考慮した最適帯域分割における $G_{(w)}$ (パラメータ: 帯域分割数 M)



(e) 画像 "Wine"

図 5.2 テスト画像



(e) 画像 "Wine"

図 5.3 帯域ブロックをベースとする視覚の空間周波数特性を考慮した最適帯域分割における $G_{(W)}$ (パラメータ: 帯域分割数 M)

5.4.2 視覚の空間周波数特性を考慮した画像の最適帯域分割パターン

前節において選定された帯域分割パラメータ $M=8$, $N^2=64$ の条件のもとで, 図5.2に示した5枚の各画像に対して, 帯域ブロックをベースとする視覚の空間周波数特性を考慮した最適帯域分割法を適用することによって求められた帯域分割パターンをそれぞれ図5.4(a)~(e)に示す.

ここで, 比較のために, 同じ帯域分割パラメータの条件下で, 視覚の空間周波数特性を考慮しない場合, すなわち, 第4章の式(4.2)に示した量子化雑音改善量

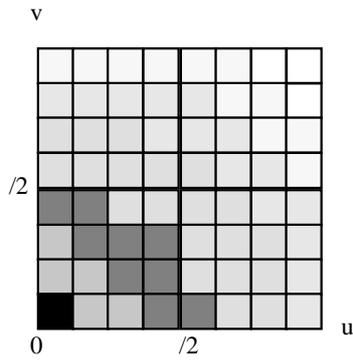
$$G = \frac{\sum_{k=0}^{M-1} \frac{1}{k^2}}{\sum_{k=0}^{M-1} \frac{1}{k}} \quad (5.16)$$

を最大にする条件において求められた帯域分割パターンを図5.5(a)~(e)に示す. ここで, k は, 式(5.5)で示された帯域 k の出力レート, k^2 は, 次式に示す帯域 k の信号電力

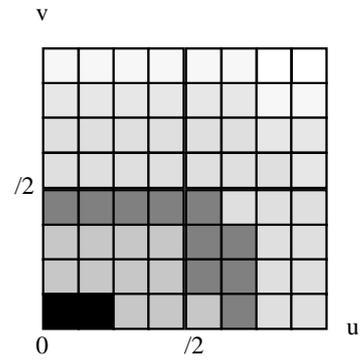
$$k^2 = \frac{1}{2} \int_k P_x(u,v) du dv \quad (5.17)$$

である. 以降, 図5.5に示した帯域分割パターンを得るための帯域分割方式を, "視覚の空間周波数特性を考慮しない最適帯域分割方式" と称することにする.

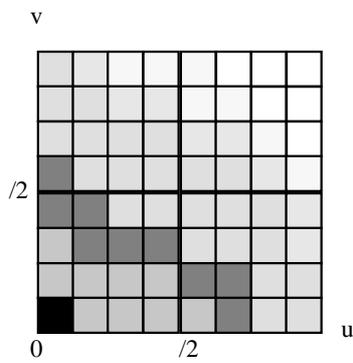
以上の結果において, まず, 図5.5に示される"視覚の空間周波数特性を考慮しない最適帯域分割方式"によって得られた各画像における帯域分割パターンが画像毎に大きく異なっていることが確認できる. しかし, 視覚の空間周波数特性を考慮することで, 図5.4に示した各画像の帯域分割パターンは, 本特性の影響を受け, 2次元周波数領域において, 低中間域の帯域の占める面積が大きくなっていると共に, 高域側では帯域の形状が, 水平, 垂直方向において対称な形に近づいている. この影響は, 画像信号の性質に依存しており, それは, 信号電力が低域に集中している, 例えば, 画像"Hada"において大きく, このことから, 図5.4(a)及び図5.5(a)に示したそれぞれの場合の帯域分割パターンは大きく異なっている. 一方, 図5.5(e)に示した画像"Wine"の結果のように, "視覚の空間周波数特性を考慮しない最適帯域分割方式"において, その帯域分割パターンが水平, 垂直方向について最初から対称に近い形になっている場合には, 視覚の空間周波数特性による重み付けの影響は小さく, 図5.4(e)に示した帯域分割パターンは, 図5.5(e)に示したものと大きくは変わらない. "視覚の空間周波数特性を考慮しない最適帯域分割方式"によって求められた図5.5に示した帯域分割パターンについて, 式(5.4)の $G_{(w)}$ を用いて評価を行い, その値を図5.4に示した帯域ブロックをベースとする視覚の空間周波数特性を考慮した最適帯域分割による $G_{(w)}(J_{opt(w)}^{(8)})$ と比較した結果を表5.1に示す.



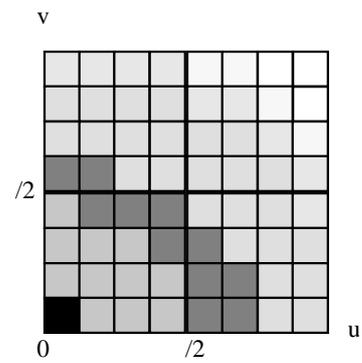
(a) 画像 "Hada"



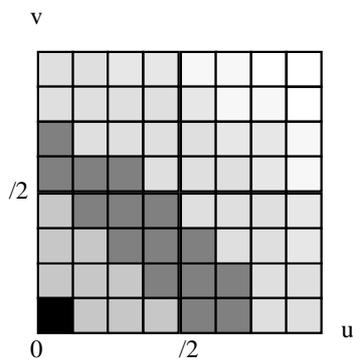
(b) 画像 "Lenna"



(c) 画像 "Cameraman"

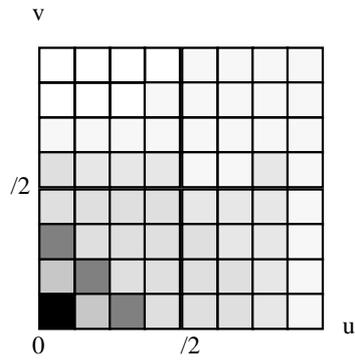


(d) 画像 "Tulip"

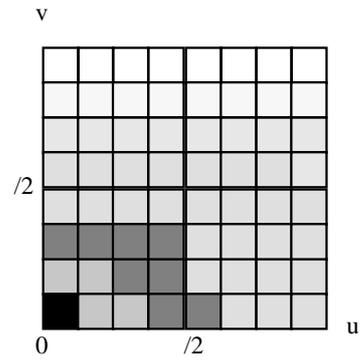


(e) 画像 "Wine"

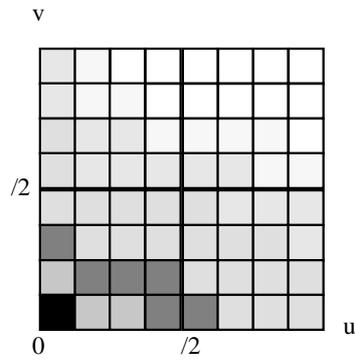
図 5.4 帯域ブロックをベースとする視覚の空間周波数特性を考慮した最適帯域による帯域分割パターン ($M=8, N^2=64$)



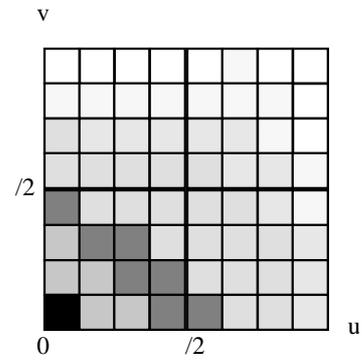
(a) 画像 "Hada"



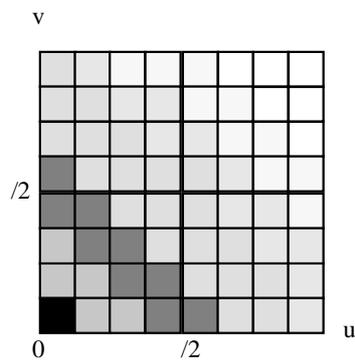
(b) 画像 "Lenna"



(c) 画像 "Cameraman"



(d) 画像 "Tulip"



(e) 画像 "Wine"

図 5.5 " 視覚の空間周波数特性を考慮しない最適帯域分割方式 " による
帯域分割パターン ($M=8$, $N^2=64$)

表5.1より,画像"Wine"のように改善の効果が少ない場合はあるものの,帯域分割パターンの決定に視覚の空間周波数特性を考慮することによって, $G_{(w)}$ の評価において約1[dB]の改善が得られることが明らかになった.

表 5.1 視覚の空間周波数特性を考慮する最適帯域分割と " 視覚の空間周波数特性を考慮しない最適帯域分割方式 " における $G_{(w)}$ の比較 (単位[dB])

<i>image</i>	<i>optimum band partition based on band blocks with human visual sensitivity</i>	<i>optimum band partition based on band blocks (maximized G)</i>
Hada	21.396	19.974
Lenna	22.112	21.113
Cameraman	17.066	16.217
Tulip	12.214	11.307
Wine	20.309	20.217

5. 5 評価方法とその結果

本章で提案する帯域ブロックをベースとする視覚の空間周波数特性を考慮した最適帯域分割法を適用した場合に，再生画像のランダムノイズの低減を定量的に評価するために，図5.2に示した5枚の画像に対して，量子化を行った後，視覚の空間周波数特性によって重み付けられた量子化雑音電力 ($WMSE$) 対エントロピー特性を求める．更に，実際に得られた再生画像のオーバーオール画質について，視知覚特性までを考慮した評価を行うには，上記のランダムノイズのみによる評価，すなわち， $WMSE$ 評価のみでは不十分であると考えられる．例えば，再生画像において，テクスチャ状のパターンをもつ符号化ひずみは，その視知覚的妨害がランダムノイズの約10倍あるとされている[39],[40]．そこで，符号化方式及び画像内容に依存することなく再生画像の画質を評価するために，主観評価実験を行い，視覚の空間周波数特性を考慮した本最適帯域分割による再生画像の画質を評価すると共に，図5.5に分割パターンを示した"視覚の空間周波数特性を考慮しない最適帯域分割方式"及びJPEGで用いられている(8×8)画素のDCTとの比較を行う．

5.5.1 $WMSE$ によるランダムノイズの定量的評価

図5.4に示すように，各画像に対して求められた帯域ブロックをベースとする視覚の空間周波数特性を考慮した最適帯域分割は，8つの帯域分割信号からなる．各帯域分割信号は，2次元信号系列の集合であるが，何れも適当なスキャン操作により1次元の帯域分割信号に変換し，それぞれスカラ量子化を適用する．ここで，量子化は，再生誤差電力を最小とするように，ミッドトレッド型の線形量子化を用い，過負荷雑音が発生しない範囲まで各帯域分割信号の量子化ステップ幅を一定とすることで行った[65],[67]．このとき，直流成分については他の帯域信号とは別途に，いかなる場合においても量子化ステップ幅=1で量子化されるものとして扱う．視覚の空間周波数特性を考慮した本最適帯域分割において， $(K \times K)$ 個の全帯域信号の中で，直流成分に相当する信号はたった1つであるから，本処理に伴って増加するビットレートの値は，無視できるほど小さいものになる．

今，図5.1における信号系列 $x_k(p,q)$ と復号された信号 $y_k(p,q)$ (共に， $k=0,1,\dots,M-1$) との間の量子化雑音誤差を $e_k(p,q)$ ($k=0,1,\dots,M-1$) で表すものとすれば，視覚の空間周波数特性によって重み付けられた量子化雑音電力 ($WMSE$) は，

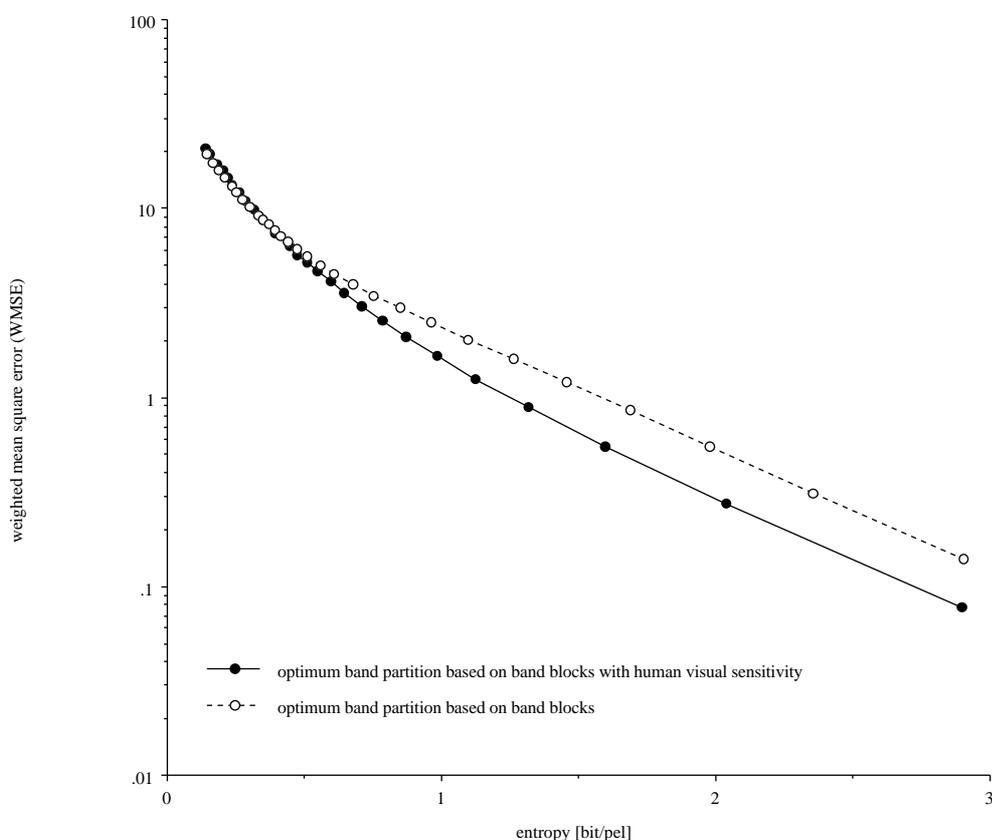
$$WMSE = \frac{\sum_{k=0}^{M-1} \sum_{(p,q)} W^2\left(\frac{p}{K}, \frac{q}{K}\right) e_k^2(p,q)}{K^2} \quad (5.18)$$

によって定義される．このとき，量子化器のステップサイズを変化させた場合に，各帯域における量子化後の信号系列が，無記憶情報源から生起されたものと仮定することで

算出されたエントロピーと $WMSE$ との関係を求める。なお、上記のスキャン法は多数存在するが、算出されるエントロピーの値は、このスキャン法には依存しない。

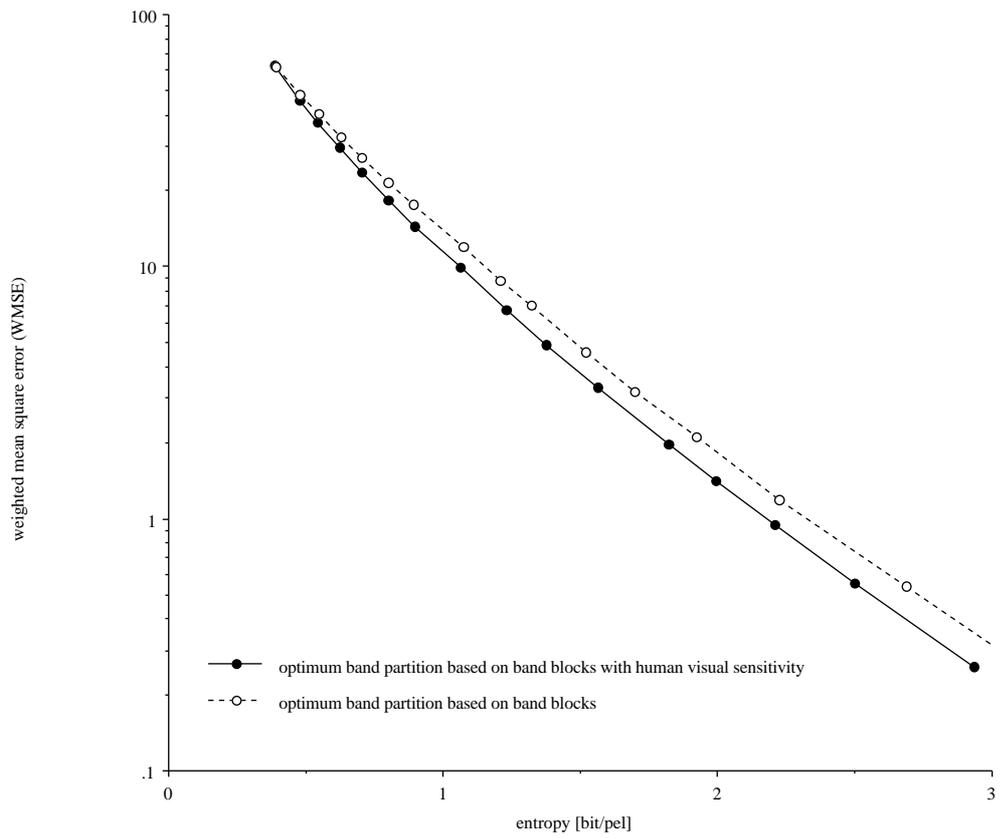
図5.2(a)~(e)に示した各画像に対して、 $WMSE$ とエントロピーの関係を求めた結果を図5.6(a)~(e)に示す。同図においては、図5.4と図5.5に示した2種類の帯域分割方式により得られた帯域分割パターンを適用した場合の特性を比較している。

図5.6において、まず、視覚の空間周波数特性を考慮した本最適帯域分割は、“視覚の空間周波数特性を考慮しない最適帯域分割方式”と比べて $WMSE$ の値が小さくなっていることを確認することができる。2種類の帯域分割方式によるこの差は、画像によって若干異なっており、これは、前節の表5.1に示した $G_{(w)}$ についての比較評価結果に対応するものとなっている。また、各画像における2種類の帯域分割方式における両特性は、エントロピーの値が小さくなったところで漸近する。これは、ビットレートが小さくなる、すなわち、量子化ステップ幅が大きくなることで、分散の小さな信号を多く含む高域の帯域信号は、そのほとんど全ての値が代表値ゼロに量子化されてしまい、視覚の空間周波数特性を考慮することの効果小さくなるためである。この傾向は、信号電力が低域に集中している、例えば、図5.6(a)にその特性を示した画像“Hada”においてより大きくなっている。以上より、視覚の空間周波数特性を考慮することは、特に、ビットレートが比較的大きい場合に、 $WMSE$ の低減、すなわち、ランダムノイズを低減するために有効であることを定量的に評価することができた。

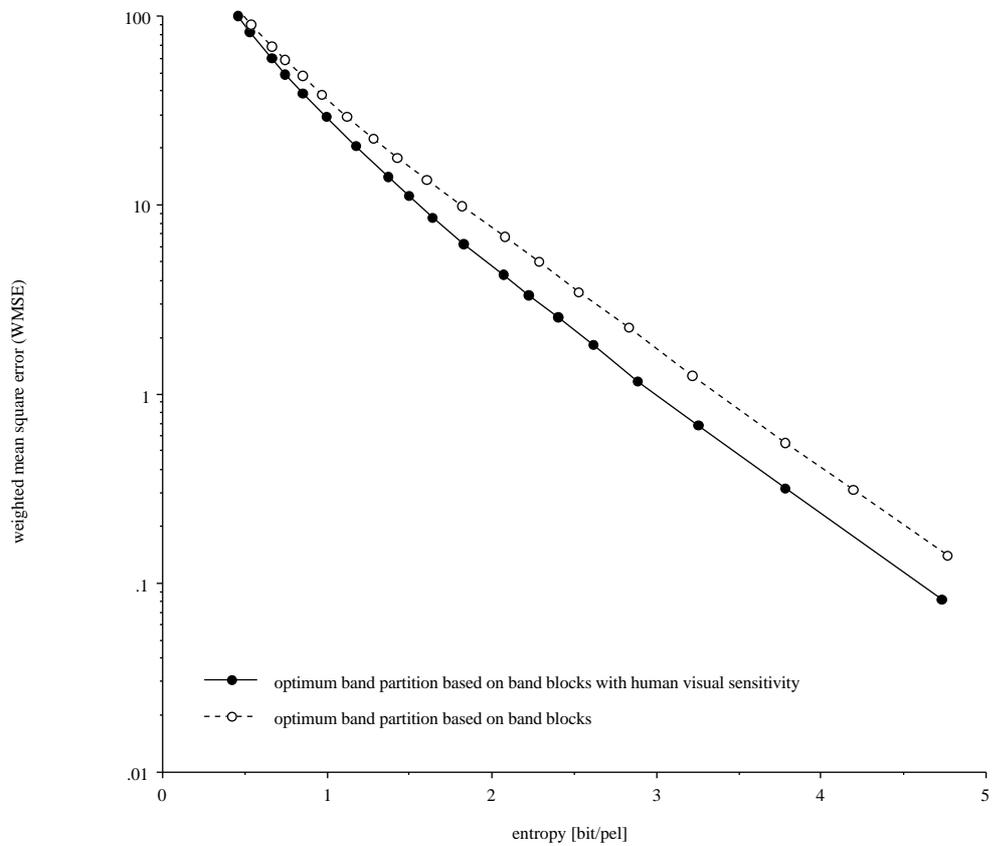


(a) 画像 "Hada"

図 5.6 $WMSE$ 対エントロピー

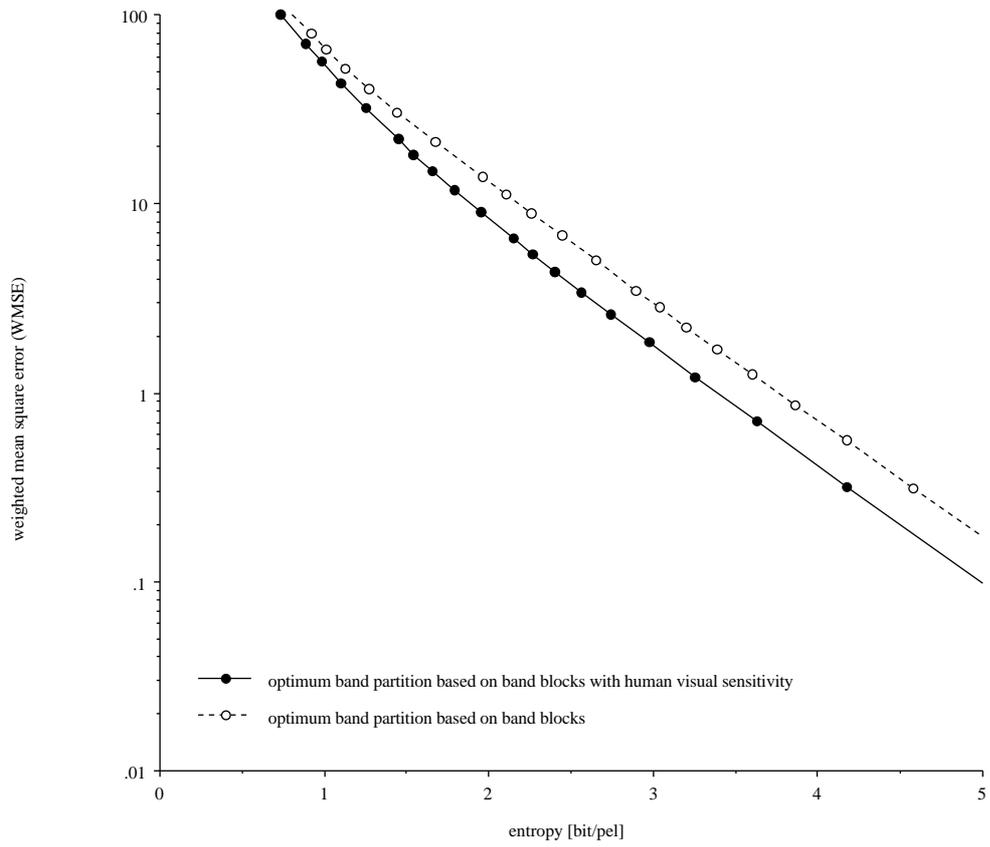


(b) 画像 "Lenna"

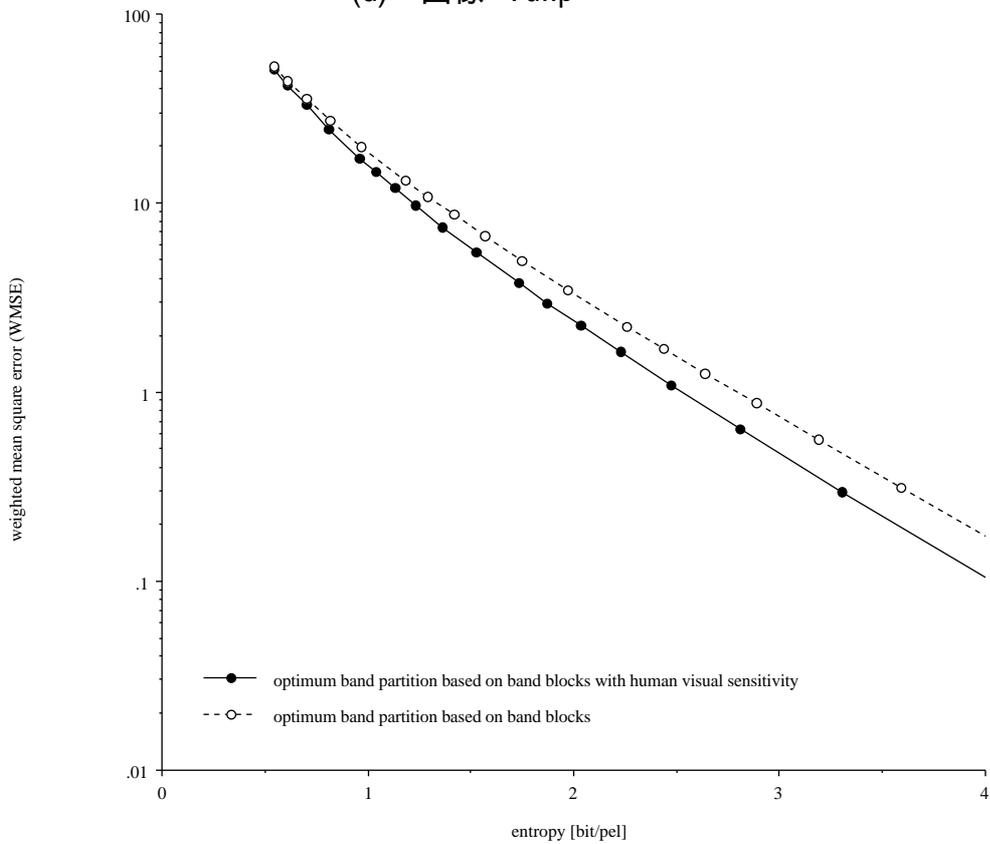


(c) 画像 "Cameraman"

図 5.6 WMSE 対エントロピー



(d) 画像 "Tulip"



(e) 画像 "Wine"

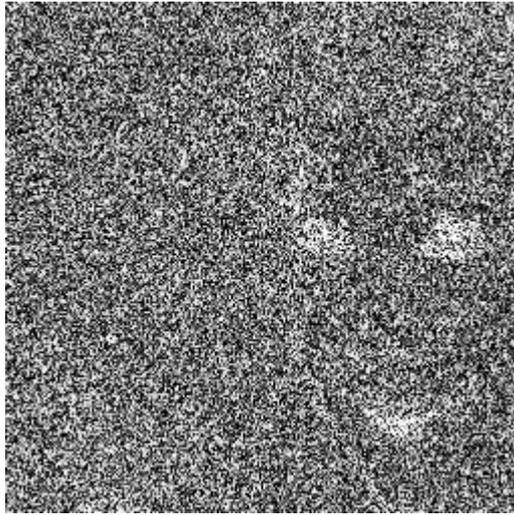
図 5.6 WMSE 対エントロピー

5.5.2 主観評価実験による画質評価

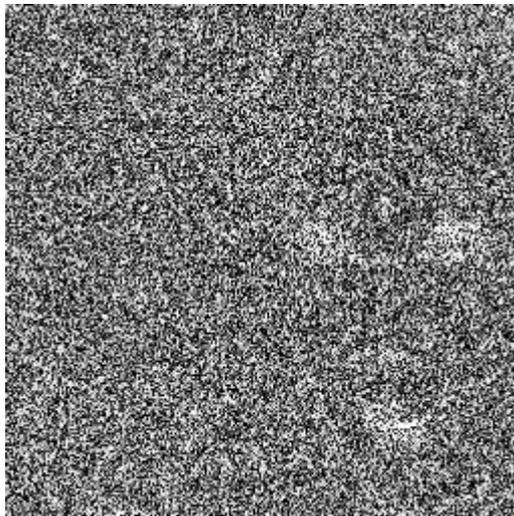
視覚の空間周波数特性を考慮した本最適帯域分割法の適用によって得られた再生画像に対して、視覚特性を考慮した画質の評価を行い、また、JPEGで用いられている(8×8)画素のDCTを適用した場合と比較する。本評価を行う理由は、前節において評価尺度として導入した *WMSE* では、再生画像の画質に対して、ランダムノイズの大小についてのみの評価しか行えず、正確な画質評価においては不十分なものと考えたからである。視覚の空間周波数特性を考慮した本最適帯域分割による再生画像には、その構成上、視覚的に最大の画質劣化要因となるブロックひずみは生じないものの、ランダムノイズ以外の符号化ひずみが含まれていないとは言えない。特に、そのひずみがテクスチャ状のパターンをもっているような場合には、誤差の間に強い自己相関性を有することから、ランダムノイズに比べて、約10倍の視覚的妨害があるとされている[39],[40]。そのため、これら全ての視覚的妨害を十分考慮した上で、本最適帯域分割による画質の評価を行う必要がある。

まず、視覚の空間周波数特性を考慮した本最適帯域分割による再生画像において、上述したテクスチャ状のノイズが含まれているか否かを調査する。図5.2(a)に示した画像 "Hada" を一例として、ある特定のビットレートにおいて、視覚の空間周波数特性を考慮した本最適帯域分割法を適用することにより得られた再生画像と原画像との差分画像を図5.7に示す。同図においては、比較のために、同じビットレートのもとで、"視覚の空間周波数特性を考慮しない最適帯域分割"を適用した場合の再生画像と原画像との差分画像、また、JPEGで用いられている(8×8)画素のDCTを適用した場合の再生画像と原画像との差分画像を示している。図5.7(a)に示した視覚の空間周波数特性を考慮した本最適帯域分割における差分画像より、その符号化に起因して生じる雑音のほとんどは、ランダムノイズであることが分かる。しかし、画像内の人物の顔の部分に強い自己相関性をもつパターン状のひずみが現れている。次に、2つの比較方式の符号化ひずみについても考察を行えば、図5.7(b)に示した"視覚の空間周波数特性を考慮しない最適帯域分割方式"は、図5.7(a)の場合と同様、その構成上の理由から、差分画像に現れる符号化雑音は、ランダムノイズが大半を占める。図5.7(c)に示した(8×8)画素のDCTによる再生画像は、図5.7(a),(b)のものに比して、ランダムノイズは小さくなっているものの、逆に、ブロックひずみに代表されるはるかに大きなテクスチャ状のパターンノイズが現れている。

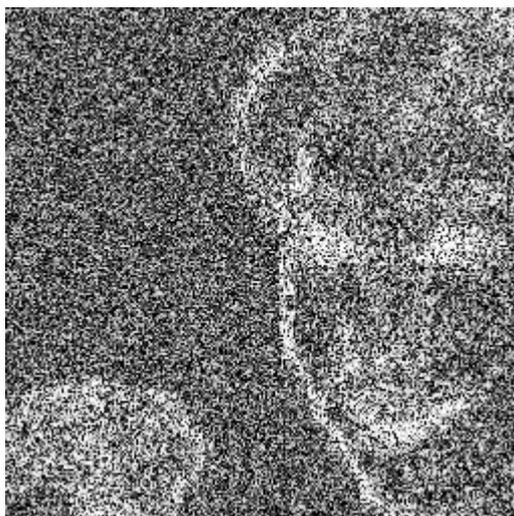
このことから、視覚の空間周波数特性を考慮した本最適帯域分割及び比較方式による再生画像に対して、視覚特性を考慮した画質を評価するためには、前項の *WMSE* 評価のみでは完全なものであるとは言えない。また、符号化方式や画像内容が異なる場合に、それらの全てを網羅して、上記のテクスチャ状のパターンノイズに代表される全ての視覚的妨害を含めてオーバーオール画質を評価し得る客観的画質評価尺度は、未だ明らかにされていない。そこで、上記の3種類の帯域分割方式による再生画像の画質を比較・評価するために、主観評価実験を行う。



(a) 帯域ブロックをベースとする視覚の空間周波数特性を考慮した最適帯域分割



(b) "視覚の空間周波数特性を考慮しない最適帯域分割方式"



(c) (8×8)DCT

図 5.7 各帯域分割方式における再生画像と原画像との差分画像
(画像 "Hada" , 0.6[bit/pel])

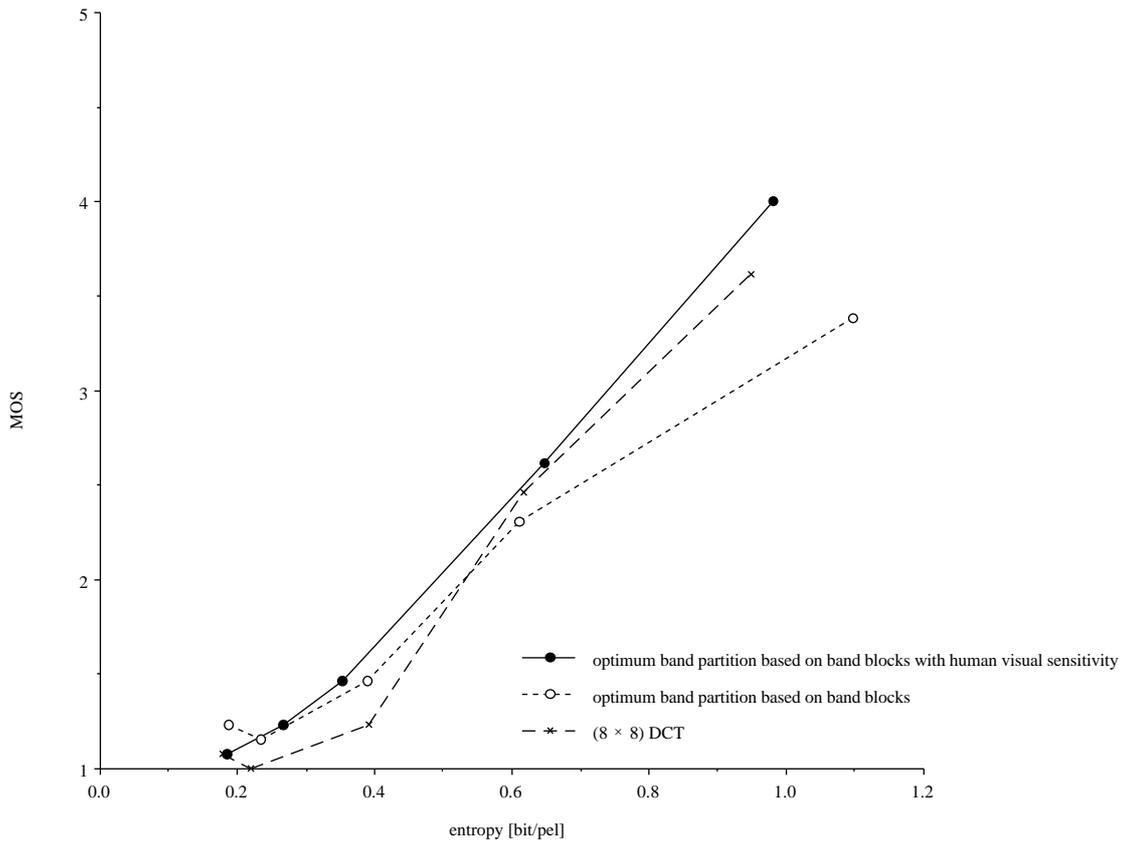
主観評価実験は、表5.2に示した条件に基づいて、既に評価訓練を終えた評定者に対して実験の目的を明示することで行った。表5.2の条件は、CCIR勧告に記載されたものと若干異なっているが、画面全体の印象というよりむしろ画像の重要な中心部分における誤差の影響を解析的に見極めるために誤差が目立つよう周囲を暗室とした。また、(256×256)画素で表示された画面の周囲は全く見えないから、CRTのサイズは意味をもたない。評定者には、2重刺激妨害尺度法に基づき、原画像と各帯域分割方式によって得られた再生画像を交互に提示し、表5.2に示した5段階評価によって再生画像に生じる妨害の程度を評価させた。評定者数を n 、 k 番目の評定者の評価値を a_k とすれば、 MOS (Mean Opinion Score) は、次式により計算される。

$$MOS = \frac{1}{n} \sum_{k=1}^n a_k \quad (5.19)$$

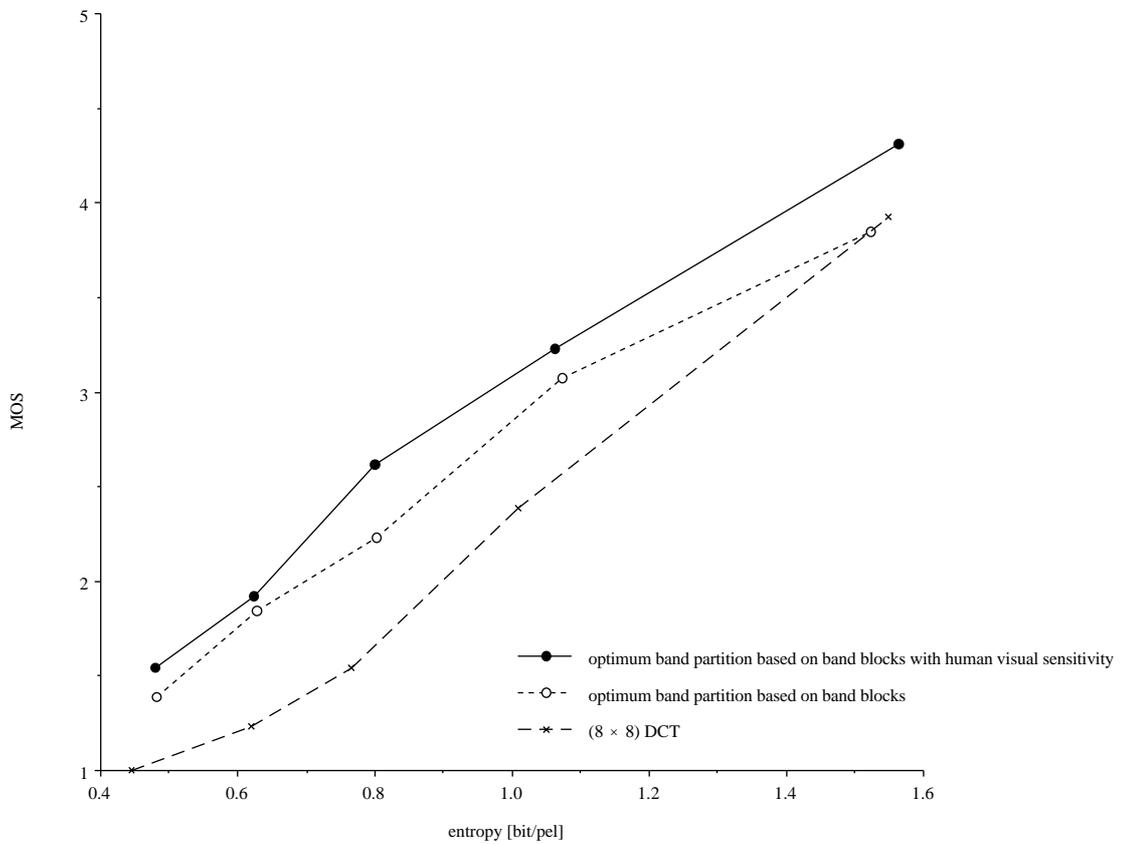
図5.8(a)～(e)は、図5.4(a)～(e)に帯域分割パターン示した視覚の空間周波数特性を考慮した本最適帯域分割法を適用した場合の図5.2(a)～(e)に示した各画像に対する MOS とエントロピーの関係を示したものである。なお、比較のために、同図においては、図5.5(a)～(e)にその帯域分割パターンを示した"視覚の空間周波数特性を考慮しない最適帯域分割方式"の同特性を、また、JPEGで用いられている(8×8)画素のDCTを適用し、且つ、JPEGで推奨されている視覚特性を考慮した量子化テーブル[2],[66]を用いて量子化を行った場合の同特性を併せて表示する。

表 5.2 主観評価実験の条件

評価法	2重刺激妨害尺度法
視距離	4H (H: 画像高)
照明	なし (暗室)
最高輝度	68.0 [cd/m ²]
最低輝度	0.01 [cd/m ²]
画像提示時間	10 [sec]
提示方法	連続表示
評定者数	13
評価尺度	評価語
	5: 妨害がわからない
	4: 妨害がわかるが、気にならない
	3: 妨害が気になるが、我慢できる
	2: 妨害が我慢できない
	1: 妨害が全く我慢できない

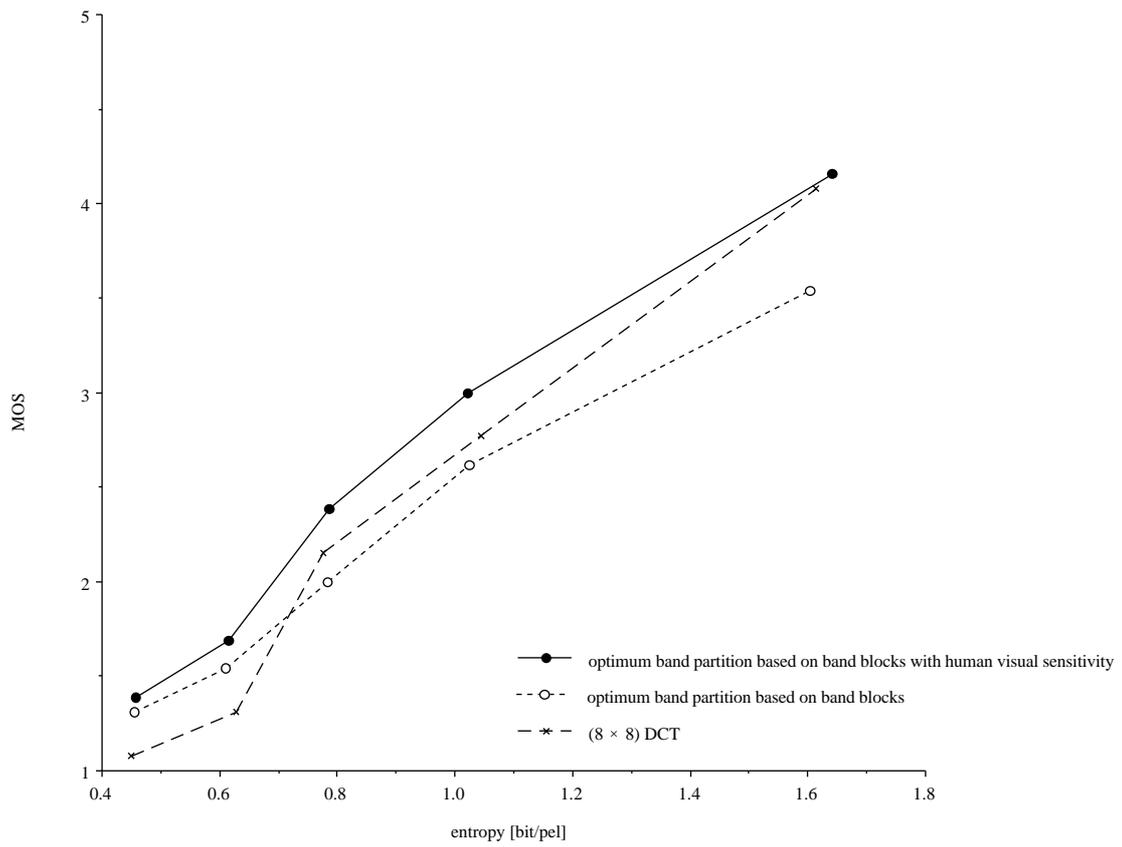


(a) 画像 "Hada"

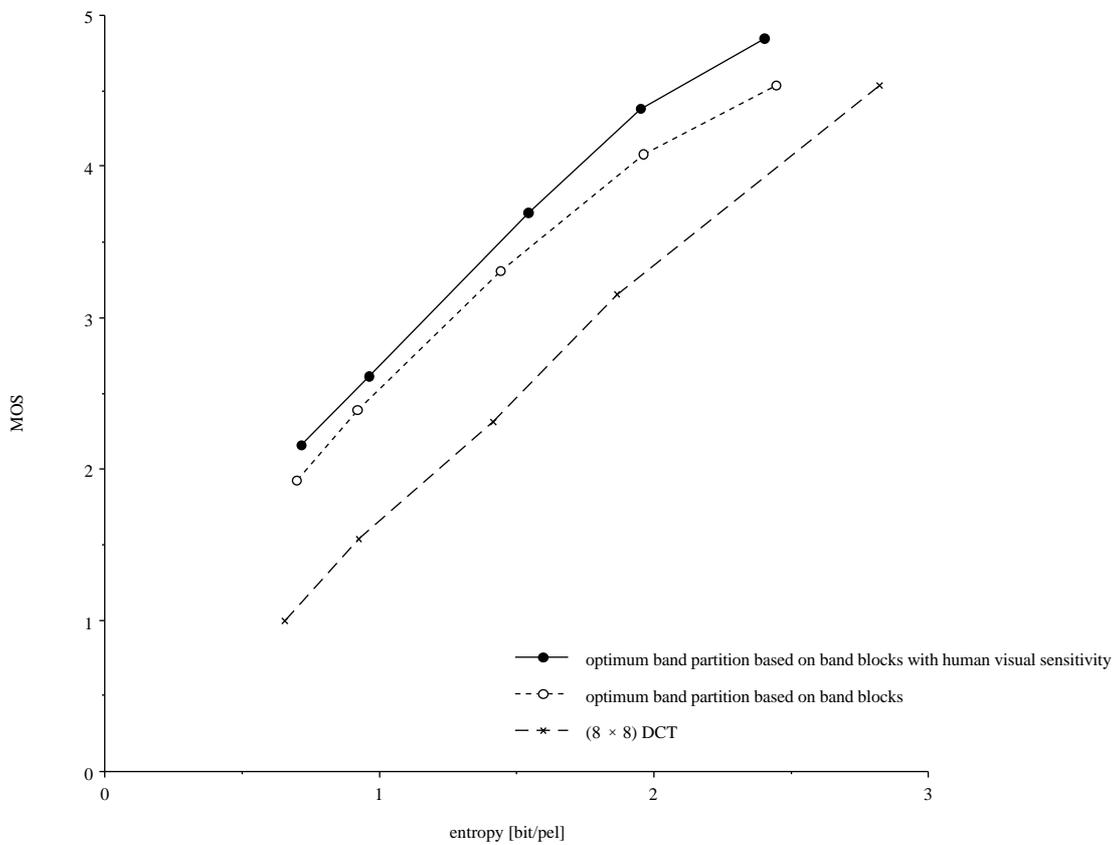


(b) 画像 "Lenna"

図 5.8 MOS 対エントロピー

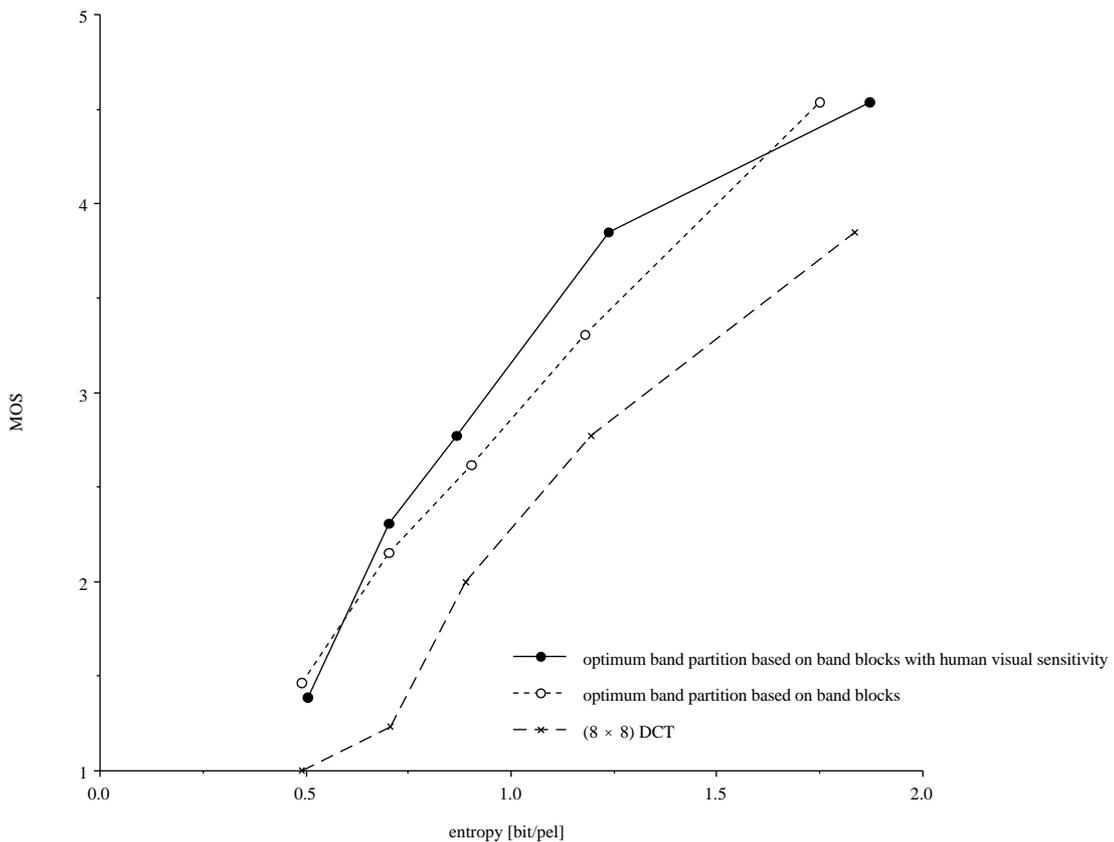


(c) 画像 "Cameraman"



(d) 画像 "Tulip"

図 5.8 MOS 対エントロピー



(e) 画像 "Wine"

図 5.8 MOS 対エントロピー

この結果、視覚の空間周波数特性を考慮した本最適帯域分割を適用することにより、図 5.8 に示した各テスト画像のほとんど全てのビットレートの範囲において、最も高い MOS 評価値が得られていることが分かる。"視覚の空間周波数特性を考慮しない最適帯域分割方式"の特性は、視覚の空間周波数特性を考慮した本最適帯域分割と比較した場合にやや劣っており、この差は、5段階評価に対応する MOS 値において 0.3 程度である。また、(8×8) 画素の DCT に対する主観評価値は、特に、比較的多くの detail を含む画像 "Lenna"、画像 "Tulip"、画像 "Wine" においては大きく劣っており、この場合は視覚の空間周波数特性を考慮した本最適帯域分割に比べて、5段階評価に対応する MOS 値において 1.0 以上の劣化が確認される。画像内容が単純である画像 "Hada" 及び画像 "Cameraman" においては、比較を行った 3 方式の特性は非常に近接し、特に、ビットレートが大きいところでは、DCT による MOS 値が "視覚の空間周波数特性を考慮しない最適帯域分割方式" のものを上回る場合がある。しかし、この場合においても提案する視覚の空間周波数特性を考慮した本最適帯域分割の特性を上回ることはない。以上のことから、視覚の空間周波数特性を考慮した本最適帯域分割を適用することにより、視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ることができることが明らかにされた。

最後に、図5.2に示した各画像について、ある特定のビットレート値を与えた場合に、視覚の空間周波数特性を考慮した本最適帯域分割法の適用によって得られた再生画像を図5.9～図5.13に示す。なお、各図においては、本手法との比較のために、同ビットレート条件下で"視覚の空間周波数特性を考慮しない最適帯域分割方式"の適用によって得られた再生画像及び(8×8)画素のDCTを適用した場合の再生画像を表示している。また、主観評価実験の結果得られた MOS 値を各再生画像の下部に記している。

この結果、(8×8)画素のDCTを用いた場合には、その再生画像にブロックひずみが現れることにより、局所的に視覚上画質劣化の著しい部分が観測される。これに対して、視覚の空間周波数特性を考慮した本最適帯域分割及び"視覚の空間周波数特性を考慮しない最適帯域分割方式"を適用した場合には、割り当てられたビットレートが小さいときに、高域に対するビット配分が小さくなることに起因して、両方式のフィルタバンクの構成上、画像全体に広がったモスキート雑音が発生し、それがランダムノイズのように観測される。この符号化ひずみは、画像内容が単純である画像 "Hada" 及び画像 "Cameraman" において、各画像の平坦部分に現れた場合に、比較的大きな画質劣化の要因となり、図5.8に示したこれらの画像に対する各帯域分割方式による特性が近接したと考えられる。逆に、多くの detail を含むその他の画像においては、この符号化ひずみの出現が画像の detail によってマスクされてしまい、画質の劣化が目立たない。これに対して、(8×8)画素のDCTに現れるブロックひずみは、画像の連続性を失わせ、再生画像が不自然なものになってしまう。この結果、図5.8に示した画像 "Lenna"、画像 "Tulip"、画像 "Wine" に対する各帯域分割方式による特性において、DCTによるものが大きく劣化したと考えられる。また、主観評価実験において得られた結果が示すように、本最適帯域分割は、視覚の空間周波数特性を考慮して帯域分割を最適化することにより、"視覚の空間周波数特性を考慮しない最適帯域分割方式"に比べて、上記のモスキート雑音、すなわち、ランダムノイズによる視覚的妨害がやや小さくなっているように観測される。



(a) 原画像



(b) 帯域ブロックをベースとする視覚の
空間周波数特性を考慮した最適帯域分割
(MOS=2.61)



(c) " 視覚の空間周波数特性を考慮しない
最適帯域分割方式 "
(MOS=2.31)



(d) (8×8) DCT
(MOS=2.46)

図 5.9 各帯域分割方式における再生画像 (画像 "Hada" , 0.6[bit/pel])



(a) 原画像



(b) 帯域ブロックをベースとする視覚の
空間周波数特性を考慮した最適帯域分割
(MOS=3.23)



(c) " 視覚の空間周波数特性を考慮しない
最適帯域分割方式 "
(MOS=3.08)



(d) (8×8) DCT
(MOS=2.38)

図 5.10 各帯域分割方式における再生画像 (画像 "Lenna", 1.0[bit/pel])



(a) 原画像



(b) 帯域ブロックをベースとする視覚の
空間周波数特性を考慮した最適帯域分割
(MOS=3.00)

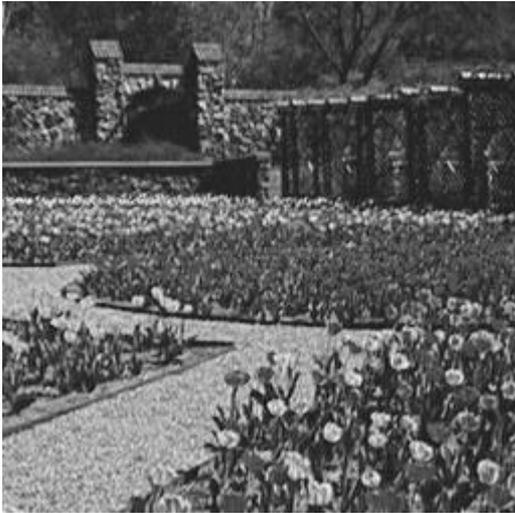


(c) " 視覚の空間周波数特性を考慮しない
最適帯域分割方式 "
(MOS=2.62)

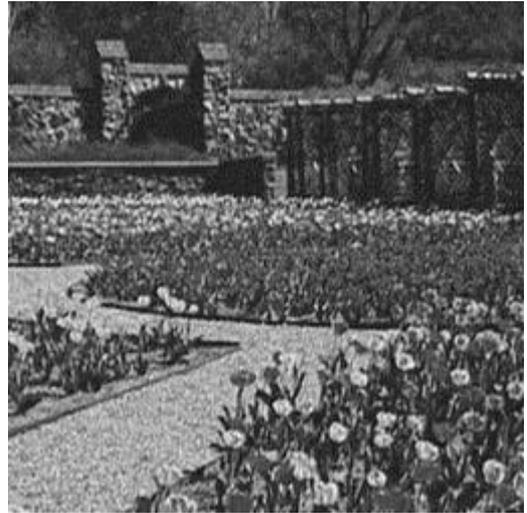


(d) (8×8) DCT
(MOS=2.77)

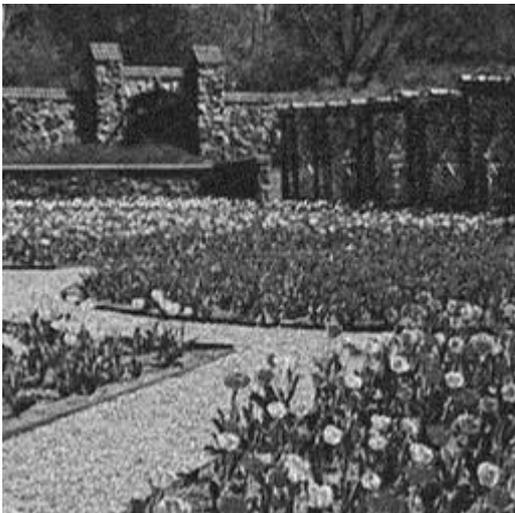
図 5.11 各帯域分割方式における再生画像 (画像 "Cameraman", 1.0[bit/pel])



(a) 原画像



(b) 帯域ブロックをベースとする視覚の
空間周波数特性を考慮した最適帯域分割
(MOS=3.69)



(c) "視覚の空間周波数特性を考慮しない
最適帯域分割方式"
(MOS=3.31)



(d) (8×8) DCT
(MOS=2.31)

図 5.12 各帯域分割方式における再生画像 (画像 "Tulip" , 1.5[bit/pel])



(a) 原画像



(b) 帯域ブロックをベースとする視覚の空間周波数特性を考慮した最適帯域分割

(MOS=3.84)



(c) "視覚の空間周波数特性を考慮しない最適帯域分割方式"

(MOS=3.31)



(d) (8×8) DCT

(MOS=2.77)

図 5.13 各帯域分割方式における再生画像 (画像 "Wine", 1.2[bit/pel])

5. 6 結言

本章では、サブバンド符号化において、視知覚特性を考慮した画質の観点から、良好な再生画像を得ることを目的として、帯域ブロックをベースとする視覚の空間周波数特性を考慮した2次元最適帯域分割法を提案し、汎用性向上の目的において選ばれた画像内容の異なる5枚のテスト画像に本手法を適用してその有効性を明らかにした。まず、視覚の空間周波数特性によって重み付けられた量子化雑音改善量を評価パラメータとして、本評価値を最大とする帯域ブロックをベースとする視覚の空間周波数特性を考慮した最適帯域分割法を提案し、視覚の空間周波数特性を考慮した本最適帯域分割を実現し得るフィルタバンクの構成を明らかにした。次に、上記5枚の画像を対象として、本フィルタバンクを実用する上で、効果的な評価改善量を得ることができる帯域分割パラメータとして、帯域分割数=8、帯域ブロック数=64を選定した。更に、選定された帯域分割パラメータを用いて、分割パターンは各画像毎に固有なものであるとして、画像それぞれについて帯域ブロックをベースとする視覚の空間周波数特性を考慮した最適帯域分割解を求め、2次元周波数平面上での最適帯域分割パターンを明らかにした。

視覚の空間周波数特性を考慮した本最適帯域分割法の適用によって得られた再生画像について、視知覚特性を考慮した画質の観点から評価を行った。まず、ランダムノイズの低減を定量的に評価することを目的として、視覚の空間周波数特性によって重み付けられた量子化雑音($WMSE$)とエントロピーの関係を求め、同じ帯域分割パラメータ条件下で求められた、視覚の空間周波数特性を考慮しない場合、すなわち、単なる量子化雑音改善量を最大にする条件を満たすよう導出された帯域分割方式との比較を行った。この結果、視覚の空間周波数特性を考慮した本最適帯域分割は、ビットレートが比較的大きい場合に、 $WMSE$ 削減、すなわち、ランダムノイズの低減において有効であることが明らかにされた。再生画像の画質を視知覚特性までを考慮して評価するにあたり、ランダムノイズによる考察のみでは不十分である。例えば、画像のひずみとしてテクスチャ状のパターンをもつ符号化ひずみが現れている場合には、ランダムノイズに比べて約10倍もの視知覚的妨害があるとされている。しかし、あらゆる符号化方式及び画像内容に依存することなく、正確な画質評価値を得ることができる汎用的な客観的画質評価尺度は未だ明らかにされていない。そこで、視覚の空間周波数特性を考慮した本最適帯域分割による再生画像に対して主観評価実験を行い、上述の画像ひずみを考慮した画質評価を行った。この結果、視覚の空間周波数特性を考慮した本最適帯域分割による再生画像は、5段階評価である MOS 値において、視覚の空間周波数特性を考慮しなかった場合に比べて約0.3の改善があり、また、JPEGで用いられている(8×8)画素のブロック単位でDCTを適用した場合と比べて、最大1.0以上もの改善があることが明らかになった。以上により、視覚の空間周波数特性を考慮した本最適帯域分割法は、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力な一手法であることが示された。

付録 C : 式(5.12) , 式(5.13)の導出について

式(5.10) ,式(5.11)に示された $\tilde{x}(m,n)$ の2次元フーリエ変換 $\tilde{X}(u,v)$ を求めると以下のようになる .

$$\begin{aligned}
 \tilde{X}(u,v) &= \sum_{m,n=-} \tilde{x}(m,n) e^{-j(um+vn)} \\
 &= \sum_{p,q=-} e^{-j(u2Kp+v2Kq)} \sum_{m,n=0}^{2K-1} \tilde{x}(m,n) e^{-j(um+vn)} \\
 &= \frac{1}{K^2} \sum_{p,q=-} \left(u - \frac{2p}{2K}\right) \left(v - \frac{2q}{2K}\right) \sum_{m,n=0}^{2K-1} \tilde{x}(m,n) e^{-j(um+vn)} \\
 &= \sum_{u,v=-} \frac{1}{K^2} \sum_{m,n=0}^{2K-1} \tilde{x}(m,n) e^{-j\frac{2}{2K}(pm+qn)} \left(u - \frac{2p}{2K}\right) \left(v - \frac{2q}{2K}\right) \\
 &= \sum_{p,q=-} \frac{1}{K^2} \tilde{X}(p,q) \left(u - \frac{2p}{2K}\right) \left(v - \frac{2q}{2K}\right)
 \end{aligned} \tag{C.1}$$

ただし , $\tilde{X}(p,q)$ は , 次式に示す $\tilde{x}(m,n)$ の $2K$ 点の2次元 DFT である .

$$\tilde{X}(p,q) = \sum_{m,n=0}^{2K-1} \tilde{x}(m,n) e^{-j\frac{2}{2K}(pm+qn)} \tag{C.2}$$

ここで , $\tilde{x}(m,n)$ と $\tilde{X}(p,q)$ は ,

$$\begin{aligned}
 \tilde{X}(p,q) &= \sum_{m,n=0}^{2K-1} \tilde{x}(m,n) e^{-j\frac{2}{2K}(pm+qn)} \\
 \tilde{x}(m,n) &= \frac{1}{(2K)^2} \sum_{p,q=0}^{2K-1} \tilde{X}(p,q) e^{j\frac{2}{2K}(pm+qn)}
 \end{aligned} \tag{C.3}$$

なる DFT の変換対で結ばれ , また , 次式に示す Parseval の等式が成立する .

$$\sum_{m,n=0}^{2K-1} \tilde{x}^2(m,n) = \frac{1}{(2K)^2} \sum_{p,q=0}^{2K-1} |\tilde{X}(p,q)|^2 \tag{C.4}$$

更に , 式(5.10)に示したmirror symmetry条件を考慮すると , 次の関係式が成立することが分かる .

$$\begin{aligned}
\tilde{X}(2K - p, q) &= \tilde{X}(-p, q) = e^{-j\frac{2p}{2K}} \tilde{X}(p, q) \\
\tilde{X}(p, 2K - q) &= \tilde{X}(p, -q) = e^{-j\frac{2q}{2K}} \tilde{X}(p, q) \\
\tilde{X}^*(p, q) &= \tilde{X}(-p, -q) = \tilde{X}(p, q) e^{-j(\frac{2p}{2K} + \frac{2q}{2K})}
\end{aligned} \tag{C.5}$$

また，上式より以下の関係式も得ることができる．

$$\begin{aligned}
|\tilde{X}(2K - p, q)|^2 &= |\tilde{X}(p, 2K - q)|^2 = |\tilde{X}(-p, q)|^2 = |\tilde{X}(p, -q)|^2 = |\tilde{X}(p, q)|^2 \\
|\tilde{X}(K, q)|^2 &= |\tilde{X}(p, K)|^2 = |\tilde{X}(K, K)|^2 \\
\tilde{X}(p, q) e^{-j\frac{(p+q)}{2K}} &= \text{real}
\end{aligned} \tag{C.6}$$

よって，式(C.5)，式(C.6)より，DFT係数 $\tilde{X}(p, q)$ ($p, q=0, 1, \dots, 2K-1$)の $4K^2$ 個の複素数のうち，独立なものは K^2 個の実数 $\tilde{X}(p, q) e^{-j\frac{(p+q)}{2K}}$ ($p, q=0, 1, \dots, 2K-1$)のみであり，これから全ての係数 $\tilde{X}(p, q)$ が求まることが分かる．

以上に基づいて，mirrior symmetry 条件を考慮した画像の電力スペクトルを導出する．

まず，式(C.1)より， $\tilde{X}(p, q)$ が空間周波数 $u = \frac{p}{K}$ ， $v = \frac{q}{K}$ におけるスペクトルを表すことに注意し，また，式(5.10)，式(C.6)を用いて，式(C.4)を以下のように変形する．

まず，

$$\tilde{x}^2(m, n) = 4 \sum_{m, n=0}^{2K-1} x^2(m, n) = \frac{1}{(2K)^2} \sum_{p, q=0}^{2K-1} |\tilde{X}(p, q)|^2 = \frac{4}{(2K)^2} \sum_{p, q=0}^{K-1} |\tilde{X}(p, q)|^2 \tag{C.7}$$

なる関係が成立する．ただし，上式において

$$k = \begin{cases} 1/\sqrt{2} & : k = 0(\text{mod } 2K) \\ 1 & : \text{otherwise} \end{cases} \tag{C.8}$$

である．上の関係式より，式(C.7)は更に以下のように変形できる．

$$\begin{aligned}
\frac{1}{K^2} \sum_{m,n=0}^{K-1} x^2(m,n) &= \frac{1}{4K^4} \sum_{p,q=0}^{K-1} \sum_{p,q=0}^{K-1} |\tilde{X}(p,q)|^2 \\
&= \frac{1}{(2K)^2} \sum_{p,q=0}^{K-1} \left| \frac{1}{2K} \tilde{X}(p,q) \right|^2 \left(u - \frac{p}{K}\right) \left(v - \frac{q}{K}\right) dudv \\
&= \frac{1}{(2K)^2} \sum_{p,q=0}^{K-1} \frac{1}{K^2} \left| \frac{\sum_{p,q=0}^{K-1} \tilde{X}(p,q) e^{-j\frac{(p+q)}{2K}}}{2K} \right|^2 \left(u - \frac{p}{K}\right) \left(v - \frac{q}{K}\right) dudv
\end{aligned} \tag{C.9}$$

式(C.9)の左辺は、画像 $x(m,n)$ の平均電力を表すから、当然、右辺の積分値もまたこの平均電力を表している。このため、式(C.9)において[]に示す被積分関数は、 $x(m,n)$ の両側電力スペクトルを表すと考えられる。それ故、この電力スペクトルは、以下のようになる。

$$P_x(u,v) = \frac{1}{K^2} \sum_{p,q=0}^{K-1} \sum_{p,q=0}^{K-1} X^2(p,q) \left(u - \frac{p}{K}\right) \left(v - \frac{q}{K}\right) \tag{C.10}$$

ただし、 $X(p,q)$ は次式で定義され、それは式(C.6)より実数値をとる。

$$X(p,q) = \frac{1}{2K} \sum_{p,q=0}^{K-1} \tilde{X}(p,q) e^{-j\frac{(p+q)}{2K}} \tag{C.11}$$

式(C.9)より、

$$\sum_{m,n=0}^{K-1} x^2(m,n) = \sum_{p,q=0}^{K-1} X^2(p,q) \tag{C.12}$$

なる関係が成立するから、次なる変換

$$x(m,n) \quad (m,n = 0,1, \dots, K-1) \quad X(p,q) \quad (p,q = 0,1, \dots, K-1) \tag{C.13}$$

は、実数系列間の正規直交変換になっていることが分かる。

なお、この直交変換は、式(5.10)、式(C.3)、式(C.11)より以下のように変形され、

$$\begin{aligned}
X(p,q) &= \frac{1}{K} \sum_{p,q=0}^{K-1} x(m,n) \cos \frac{(2m+1)p}{2K} + \cos \frac{(2n+1)q}{2K} \\
&= \frac{2}{K} \sum_{p,q=0}^{K-1} x(m,n) \cos \frac{(2m+1)p}{2K} \cos \frac{(2n+1)q}{2K}
\end{aligned} \tag{C.14}$$

結局、この $X(p,q)$ は、 $x(m,n)$ の2次元DCTに他ならない。

(証明終)

第 6 章

結論

6.1 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、帯域分割の最適化を行う理論とその手法について述べた。ここではまず、画像信号の統計的性質に基づいた画像モデルを導入した上で、最適帯域分割を与える理論式を導出した。また、その結果を実画像に適用するための手法として、帯域ブロックをベースとする最適帯域分割法を提案し、本最適帯域分割に基づいたフィルタバンクの構成を明らかにした。更に、実画像への適用においては、画質の観点から良好な再生画像を得るために、視覚の空間周波数特性を考慮した最適帯域分割法を提案した。

第2章においては、本研究の理論的基礎となる1次元の入力信号系列に対する最適帯域分割の導出を述べた。最適帯域分割は、入力信号に割り当てられたビットレートが一定であるとの条件のもとで、受信側での再構成信号に含まれる総量子化雑音電力を最小とする帯域分割として定義した。ここでは、入力信号のグローバルな特性として定常性を仮定し、最適帯域分割を導出するための理論式を明らかにした。特に、1次元信号系列の例として、画像信号をグローバルに近似する隣接標本値間の相関係数値 $\rho = 0.9$ の1次マルコフ情報源系列を入力信号として、具体的な最適帯域分割特性を示した。導出された最適帯域分割を評価するために、帯域分割を行わず、入力信号系列を直接スカラ量子化する場合の量子化雑音電力値からの改善量として定義される量子化雑音改善量なる評価尺度を提案した。上記の信号系列に対しては、最適帯域分割を適用した場合には、帯域分割数は僅か4であるにもかかわらず、量子化雑音改善量の理論限界値に匹敵する性能を有していることが明らかにされた。また、既存の帯域分割方式に対しての性能評価を行うために、帯域を均等に分割するlinear分割と、帯域を対数スケール上で均等に分割するlog-linear分割、すなわち、wavelet変換に着目し、量子化雑音改善量の観点から両帯域分割方式と最適帯域分割との比較を行った。この結果、1次マルコフ信号に対しては、相関係数が0に近いところでは、linear分割を用いることで、逆に、 ρ が1.0に近いところでは、

log-linear 分割を用いることで，最適帯域分割の量子化雑音改善量特性を近似し得ることが明らかにされた．以上において，画像信号を仮定した1次元の信号系列に対しては，そのフィルタバンクの実現が困難である最適帯域分割を行わなくとも，log-linear 分割を適用することによって十分な量子化雑音の低減効果を得られることが示された．

第3章においては，第2章で1次元信号系列を対象に行われた最適帯域分割導出の議論を2次元へと拡張し，2次元の最適帯域分割の導出について述べた．ここでは，まず，2次元の帯域分割処理について任意の帯域分割を仮定したフィルタバンク構成の理論モデルを提案し，分割フィルタにおける各帯域とその出力レートの関係を示した．次に，2次元の信号系列に対して，分割後の各帯域信号がスカラ量子化されると仮定したとき，(1) 信号全体に割り当てられるビットレートが一定である，(2) 受信側での再構成信号に含まれる総量子化雑音電力を最小にする，という条件のもとで，2次元周波数平面における最適帯域分割及び各帯域信号への最適ビット配分を定める理論式を導出すると共に，最適帯域分割を算出するための数値計算アルゴリズムを明らかにした．2次元の入力信号系列として，画像の統計的性質に基づいてモデル化された水平・垂直相関分離型及び水平・垂直相関非分離型なる2種類の画像モデルを導入して，各モデルに対する2次元の最適帯域分割特性を具体的に示した．最後に，2次元へと拡張された量子化雑音改善量を評価尺度として，最適帯域分割の評価を行った．まず，上記の2種類の画像モデルに対する量子化雑音改善量の理論限界値を明らかにした上で，最適帯域分割による量子化雑音改善量特性の観点から評価を行った結果，両画像モデル共に，帯域分割数=4の条件で得られた最適帯域分割は，各画像モデルにおける量子化雑音改善量の理論限界値をほぼ達成できることが明らかにされた．次に，2次元の帯域分割を行う際に従来から用いられている方式として，separableフィルタで実現されるlinear分割及びlog-linear分割，そして2次元non-separableフィルタで実現される帯域分割方式，更にはJPEGで用いられている空間領域の数画素毎のブロック単位で適用されるDCTと本最適帯域分割との比較を行った．入力信号として水平・垂直相関分離型画像モデルを用いた場合には，最適帯域分割は，既存の帯域分割方式に比べて，量子化雑音改善量の観点において3~4[dB]の改善があることが明らかにされた．既存の帯域分割方式については，画像符号化等において一般的に用いられているseparableフィルタによる帯域分割処理は，帯域分割数 $M=8$ の範囲では，改善の効果がnon-separableフィルタに比べて劣っていることが示され，特に，近年様々な分野で応用されているlog-linear分割（wavelet変換）に基づいた帯域分割方式については，低域成分のみを再分割の対象とし，高域成分に対しての分割処理を行わないことが，逆に量子化雑音電力改善の観点からは不利であることが分かった．DCTとの比較において，既存の分割方式も含めてサブバンド分割方式は，同じ帯域分割数の条件でDCTの特性を上回っているものの，JPEGで用いられている(8×8)画素のDCTによる量子化雑音改善量特性を上回るためには，既存の帯域分割方式を用いるならば，帯域分割数 $M=16$ のlinear分割が必要である．これに対して，本最適帯域分割が適用されるならば，同特性を上回

るために必要な帯域分割数は、僅か4でよい。次に、入力信号として、一般の画像に対してより汎用性のある水平・垂直相関非分離型画像モデルを用いた場合には、最適帯域分割は、既存の帯域分割方式に比べて、量子化雑音改善量の観点において1~2[*dB*]の改善があることが明らかにされた。また、既存の帯域分割方式の中では、2次元non-separableフィルタによる帯域分割が、他のseparableフィルタを用いた方式に比べて大きな改善量を有していることが分かった。DCTについては、2次元DCTの基底が、1次元DCTの組合せ、すなわち、separableモデルで構成されていることに起因して、水平・垂直相関非分離型画像モデルに対しては、その量子化雑音改善量特性がサブバンド分割方式と比べて大きく劣ってしまうことが示された。以上により、最適帯域分割が理論的に明らかにされたことで、画像信号を対象とした帯域分割処理による量子化雑音電力改善の理論的限界が示される共に、量子化雑音電力低減の観点において、2次元周波数平面上における理想的な帯域分割特性が明らかにされた。

第4章においては、第3章において理論解として得られた最適帯域分割特性を実現するために、サブバンド画像符号化システムの簡易化・高性能化を念頭において、少ない帯域分割数であるにもかかわらず、高い量子化雑音改善量を達成できる帯域ブロックをベースにした2次元最適帯域分割法について述べた。まず、大部分の自然画像によく適合する水平・垂直相関非分離型画像モデルを用いるという仮定のもとで、帯域ブロックをベースにする最適帯域分割法を提案した。帯域ブロックをベースにする本最適帯域分割は、理論的に量子化雑音電力の最大値を取り得るものの、そのフィルタバンクの実現が極めて困難である第3章で求められた最適帯域分割に対して、若干の量子化雑音改善量の犠牲を許すことで、そのフィルタバンクの実現を可能とした帯域ブロック（周波数領域における基本ブロック）に基づいた最適帯域分割として位置付けられる。ここでは、本手法に基づいてフィルタバンクを実用する上で、効果的な量子化雑音の改善が得られる帯域分割パラメータとして、帯域分割数=4、帯域ブロック数=16を選定した。このとき、上記の画像モデルにおいて、理想的な最適帯域分割と帯域ブロックをベースとする最適帯域分割との量子化雑音改善量の差は、僅か0.2[*dB*]となり、精度の高い近似が行われていることを明らかにした。次に、帯域ブロックをベースとする最適帯域分割を実現し得るQMFを用いたフィルタバンクの構成法を示し、それを実画像に適用した場合の量子化雑音改善効果を確認した。この結果、実験に用いた大半の画像において、帯域分割数=4の本最適帯域分割は、(8×8)画素のDCTと遜色ない量子化雑音改善性能を有していることが示された。以上は、水平・垂直相関非分離型画像モデルにおいて最適化された固定な最適帯域分割であるが、実際の画像には本モデルに含まれない非定常性として、直流成分と相関特性の画像内容による或いは空間的な場所による変動がある。このため、前者に対しては、最低域の帯域分割信号に対して(2×2)画素のDFTを適用して非定常直流成分の分離を行う方法を、後者に対しては、画像毎に固有な最適な分割を得るために、帯域ブロックをベースとする適応帯域分割法を提案した。画像例について、本固定及び適応分

割方式による量子化雑音改善量特性の比較評価実験を行い，多くの画像については，固定な最適帯域分割により十分な改善が得られるが，水平・垂直相関非分離型画像モデルでの近似が特に困難である画像については，適応帯域分割が有効であることを明らかにした．このとき，たとえ適応分割を行う場合においても，それに要する処理コストと付加情報量の増加は僅かであることを示した．最後に，帯域ブロックをベースとする最適ブロック分割における符号化性能を調べるために， SNR とエントロピーの関係を求め，JPEG で用いられている (8×8) 画素のDCTとの比較を行った．この結果，本最適帯域分割法は同じビットレートのDCTに比べて約 $1[dB]$ の SNR の改善が得られることが示された．以上により，最適帯域分割を実現し得るサブバンド符号化システムの簡易的な構成法を明らかにすると共に，本手法を画像への適用した場合の有効性が示された．

第5章においては，サブバンド符号化において，視知覚特性を考慮した画質の観点から良好な再生画像を得ることを目的として，視覚の空間周波数特性を考慮した2次元最適帯域分割法について述べた．まず，フィルタバンク実現を可能とするために，帯域ブロックをベースに帯域分割を検討し，視覚の空間周波数特性によって重み付けられた量子化雑音改善量を評価値に，これを最大にする分割として，視覚の空間周波数特性を考慮した最適帯域分割法を提案した．更に，本最適帯域分割を実現し得るフィルタバンクとして，DCTに基づいた新たなフィルタバンクの構成を明らかにした．次に，汎用性向上の目的において選ばれた画像内容の異なる5枚の実画像を対象として本フィルタバンクを実用する上で，効果的な評価改善量を得ることができる帯域分割パラメータとして，帯域分割数=8，帯域ブロック数=64を選定した．更に，選定された帯域分割パラメータのもとで，実画像それぞれに対して，視覚の空間周波数特性を考慮した最適帯域分割法を適用することで，各画像に固有なものとして，2次元周波数平面上での最適帯域分割パターンを明らかにした．最後に，視覚の空間周波数特性を考慮した本最適帯域分割による再生画像の画質を，まず，視知覚特性を考慮したランダムノイズの観点から定量的に評価するために，視覚の空間周波数特性によって重み付けられた量子化雑音電力 ($WMSE$) とエントロピーの関係を求め，同じ帯域分割パラメータ条件において求められた視覚の空間周波数特性を考慮しない場合，すなわち，単なる量子化雑音改善量を最大にする条件で導出された帯域分割方式との比較を行った．この結果，視覚の空間周波数特性を考慮した本最適帯域分割は，ビットレートが比較的大きい場合に，効果的な $WMSE$ の低減を実現できることが明らかにされた．しかし，本来の目的である視知覚特性を考慮した画質評価を行うにあたり，上記のランダムノイズ($WMSE$)だけで十分な評価であるとは言えず，例えば，画像ひずみとして，テクスチャ状のパターンノイズが現れている場合には，誤差の間に強い自己相関性を有するため，ランダムノイズに比べて約10倍も知覚されやすく，大きな視知覚的妨害となる．本最適帯域分割との比較方式に挙げたJPEGで用いられている (8×8) 画素のブロック単位で適用されるDCTによる再生画像においては，上記のパターンノイズの代表とも言えるブロックひずみが現れることは必然で

ある．本最適帯域分割による再生画像においては，そのフィルタバンクの構成上ブロックひずみは発生しないものの，その他のパターンノイズが現れている．そこで，符号化方式及び画像内容に依存することなく，上記のパターンノイズ以外の全ての視知覚的妨害を含めたオーバーオール画質を評価するために，主観評価実験を行った．この結果，視覚の空間周波数特性を考慮した本最適帯域分割による再生画像は，5段階評価におけるMOS値において，視覚の空間周波数特性を考慮しなかった場合に比べて約0.3の改善があり，また，JPEGで用いられている(8×8)画素のDCTのものとは比べて，最大で1.0以上の改善があることが示された．以上において，帯域ブロックをベースとする視覚の空間周波数特性を考慮した最適帯域分割法は，視知覚を考慮した画質の観点においても，有力な一手法であることが明らかにされた．

画像のサブバンド符号化において，その情報圧縮のための本質は，適切な帯域分割と各帯域信号への適切なビット配分にある．本研究は，その理論展開が困難であることから議論されないままであった帯域分割の最適化問題に取り組んだものであり，画像信号を対象に理論的に求められた最適帯域分割は，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，本研究におい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と信じている．

6.2 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、最適帯域分割の適用により、量子化雑音電力を最も低減可能であることを明らかにした本手法をベースとして更に高能率な符号化システムを構成するためには、いくつかの課題が残されている。1つ目は、最適帯域分割によって得られた各帯域信号に対するスキャン法を含めたエントロピー符号化器の開発である。JPEGで用いられているジグザグスキャンとEOB(End of Block)コードの組合せに相当する各帯域信号の性質に適合したスキャン処理を行うことができれば、本符号化における更なる圧縮効率が実現されるものと考えられる。2つ目は、第4章で述べた固定分割と適応分割の適用における符号化システムの実現である。この問題については、本論文で示したように、画像毎に判別を行い、固定分割と適応分割を切り換え、更に適応分割においては、望ましい帯域分割パラメータを選定する方法と、あらゆる画像をも網羅し得る最も厳しい帯域分割パラメータを用意しておき、全ての画像に対応できるようにする方法が考えられる。前者の方法の場合には、入力された画像に対して、自動判別を行なうための処理コスト分が付加されることになる。しかし、本論文で示された結果によれば、厳しい帯域分割パラメータを必要とする画像は、画像全体に対してごく一部であると考えられ、多数の画像を符号化することを想定したときに、自動判別の結果、その他の大半の画像は短時間で処理することが可能となり、1画像当たりの平均処理時間が短縮されることが期待できる。前者の方法を選択するとした場合に、現状においては、入力された画像信号に対して、固定、適応各々の最適帯域分割を実際に適用した場合の量子化雑音改善量の値を比較して判断する最も原始的な方法で対応している。固定分割は、水平・垂直相関非分離型画像モデルに基づいて導出されたものであるから、本論文で述べたように、入力された画像の自己相関特性に基づいたある程度の判別は可能であると考えられる。しかし、より高能率な符号化システムの実現を目指して、システムに画像が入力された時点で精度の高い自動判別を行うためには、自己相関特性のみならず、電力スペクトル、或いは画像の非定常性までを考慮した画像の特徴を表し得る判定パラメータの考察が必要であり、それに基づいた自動判別技術の確立が望まれる。
- ・本研究では、最適帯域分割を実現するためのフィルタバンクは、第4章においてはQMFを、また第5章においてはDCTをベースに構築されている。これらは、いずれもseparableフィルタバンクの適用を必要とする帯域ブロックに基づいて最適帯域分割を実現しているために、帯域分割パラメータの条件が厳しくなった場合には、それに伴う処理時間に問題が残されている。そこで、2次元周波数平面上で、最適帯域分割特性を直接実現し得るフィルタバンクの構築が可能ならば、この処理時間の問題は解決されるものと考え

ており,2次元non-separableフィルタの適用を前提とした新たなフィルタバンクの構成法についての検討が望まれる。

- ・ 本研究では,その統計的性質がある程度明確であることから,濃淡静止画像に限定して,最適帯域分割の検討を行ってきた。昨今,画像符号化の対象は,カラー画像,動画画像,そしてステレオ画像へと移っており,これらの各種画像信号に対しても最適帯域分割の開発が必要である。そのためには,上記の各画像信号の統計的性質は明らかにされているが,まず,これを再検討した上で,本研究で提案した手法に基づいて,最適帯域分割の導出と各画像における最適帯域分割の実現を行えばよいと考えている。また,カラー画像においては,輝度信号と色信号の,ステレオ画像においては,L-R間信号の間の相関の大きさは利用できると考えている。
- ・ 第5章では,再生画像におけるランダムノイズの低減を目的として,人間の視知覚特性の1つである視覚の空間周波数特性を考慮した最適帯域分割の導出を行ったが,視知覚特性については,今回考慮されていないものが他に多く存在する。特に,サブバンド符号化において,再生画像の輪郭部周辺に発生するリングングは,その誤差の間に強い自己相関性を有しているため,このような誤差は視知覚上大きな妨害となると考えられる。リングングの大小については,帯域分割/合成フィルタの特性に依存しているため,これらの各フィルタをうまく選択し,その条件において帯域分割の最適化を実現することで,リングングの発生をある程度抑えられるものと考えている。このとき,逆に,輪郭近傍の誤差は知覚されにくいという視覚のマスクングの効果があり,本評価においては,例えば,客観的画質評価尺度の1つであるPQS[73]の算出において行われているように,マスクング効果を表すvisibility関数を考慮した定量的な画質の評価のもとでの検討が必要である。

謝辞

本研究を遂行するにあたり，長きにわたって懇切なる御指導と御鞭撻を賜りました北陸先端科学技術大学院大学 教授 宮原 誠 博士に深甚の謝意を表します．

本研究を推進するにあたって，長岡技術科学大学修士課程在学中から今日に至るまで，終始ひとかたならぬ御指導と御教授を賜りました法政大学 教授 大竹 孝平 博士に深甚の謝意を表します．

本研究を推進するにあたり，日頃から熱心に御討論いただき，有益なる御助言を賜りました北陸先端科学技術大学院大学 助教授 小谷 一孔 博士に心より感謝いたします．

本論文をまとめるにあたり，草稿の段階から貴重な御助言と御指導を賜りました金沢大学 教授 橋本 秀雄 博士，並びに北陸先端科学技術大学院大学 助教授 赤木 正人 博士に心より感謝いたします．

副テーマの遂行という立場ではありましたが，研究を理論的に体系づけるということの美しさ，且つ，難しさについて，御教授を賜りました元北陸先端科学技術大学院大学 教授 飯島 泰蔵 博士に謹んで感謝の意を表します．

神戸大学に在職中，数多くの御教授を賜ったと共に，進学を決意するにあたって多大なる激励をいただきました神戸大学 教授 平野 浩太郎 博士をはじめとする諸先生方に感謝の意を表します．

また，日頃より多大なる議論と激励をいただきました神戸大学 助手 黒木 修隆 博士をはじめとする諸先生方，並びに北陸先端科学技術大学院大学 宮原・小谷研究室の諸氏に厚く御礼申し上げます．

末筆ながら，私の研究生活を温かく見守ってくれた父と母，妻とその両親，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] 安田 浩 編著, " マルチメディア符号化の国際標準," 丸善, (1991)
- [2] 遠藤 俊明, " カラー静止画像の国際標準符号化方式," インターフェース, (1991-12)
- [3] J.B.O'Neal, "Predictive quantizing system(differential pulse code modulation)for the transmission of television signals," Bell Syst. Tech. J., vol.45, pp.689-721, (1966)
- [4] 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)
- [5] 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. Accoust., Speech, Signal Processing, vol. ASSP-32, pp.1213-1229, (1984)
- [6] Z.Lin, Y.Attikiouzel, "Two-dimensional linear prediction model-based decorrelation method," IEEE Trans. Pattern Anal. Mach. Intel., vol. PAMI-11, pp.661-665, (1989)
- [7] Y.Linde, A.Buzo and R.M.Gray, "An algorithm for vector quantizer design," IEEE Trans. Commun., vol. COM-28, pp.84-95, (1980)
- [8] R.M.Gray, "Vector Quantization," IEEE ASSP Mag., vol.1, pp.4-28, (1984)
- [9] 田崎, 山田, " ベクトル量子化," 信学誌, vol.67, pp.532-536, (1984)
- [10] N.M.Nasrabadi, R.A.King, "Image coding using vector quantization: A review," IEEE Trans. Commun., vol. COM-36, pp.957-971, (1988)
- [11] C.F.Barnes, S.A.Rizvi and N.M.Nasrabadi, "Advances in residual vector quantization: A review," IEEE Trans. Image Processing, vol.5, pp.226-262, (1996)
- [12] P.A.Wintz, "Transform Picture Coding," Proc. IEEE, vol.60, pp.809-820, (1972)
- [13] N.Ahmed, T.Natarajan, and K.R.Rao, "On image processing and a discrete cosine transform," IEEE Trans. Computer., vol. C-23, pp.90-93, (1974)

- [14] R.J.Clarke, "Transform Coding of Images," Academic Press, (1985)
- [15] H.S.Malvar, D.H.Staelin, "The LOT: Transform coding without blocking effects," IEEE Trans. Accoust., Speech, Signal Processing, vol. ASSP-37, pp.553-559, (1989)
- [16] J.W.Woods, S.D.O'Neil, "Subband coding of images," IEEE Trans. Accoust., Speech, Signal Processing, vol. ASSP-34, pp.1278-1288, (1986)
- [17] M.J.T.Smith, T.P.Barnwell, III, "Exact reconstruction techniques for tree-structured subband coders," IEEE Trans. Accoust., Speech, Signal Processing, vol. ASSP-34, pp.434-441, (1986)
- [18] H.Gharavi, A.Tabatabai, "Sub-band coding of monochrome and color images," IEEE Trans. Circuits and Systems, vol. CAS-35, pp.207-214, (1988)
- [19] D.L.Gall, A.Tabatabai, "Subband coding of digital images using symmetric short kernel filters and arithmetic coding techniques," Proc. of Inter., Conf., Acoust., Speech, Signal Processing, ICASSP'88, pp.761-764, (1988)
- [20] J.W.Woods, "Subband Image Coding," Kluwer Academic Publishers, (1990)
- [21] I.Daubechies, "Orthogonal bases of compactly supported wavelets," Commun. Pure Appl. Math., vol.XLI, pp.909-996, (1988)
- [22] S.G.Mallat, "Multifrequency channel decompositions of images and wavelet models," IEEE Trans. Accoust., Speech, Signal Processing, vol. ASSP-37, pp.2091-2109, (1989)
- [23] O.Rioul, M.Vetterli, "Wavelets and signal processing," IEEE SP Magazine, pp.14-38, (1991)
- [24] 太田 睦, "画像符号化における直交ウェーブレット," 数理科学, no.354, pp.24-30, (1992)
- [25] チャールズ K. チュウイ 著 / 桜井 明, 新井 勉 共訳, "ウェーブレット入門," 東京電機大学出版局, (1993)
- [26] U.Franke, R.Master, "Region based image representation with variable reconstruction quality," SPIE Visual Communication and Image Processing'88, pp.178-186, (1988)
- [27] 堀田, 宮原, "均等色空間に基づくカラー画像の領域分割符号化," 信学論B, vol. J75-B-I, pp.422-430, (1992)
- [28] 松田, 伊東, 宇都宮, "画像の適応的可変ブロック形状KL変換符号化," 信学論B, vol. J76-B-I, pp.399-408, (1993)
- [29] 齊藤, 鄭, "新しい画像符号化技術 - フラクタル理論を中心として," 信学誌, vol.75,

- pp.1343-1355, (1992)
- [30] M.F.バーンスレイ, L.P.ハード 共著 / 蔡 東生, 江守 正多 共訳, "マルチメディアフ
ラクタル画像圧縮," トッパン, (1995)
- [31] 原島 博, "知的画像符号化と知的通信," テレビ誌, vol.42, pp.519-525, (1988)
- [32] 相澤, 原島, 齊藤, "構造モデルを用いた画像の分析合成符号化方式," 信学論 B, vol.
J72-B-I, pp.200-207, (1989)
- [33] (JPEG) ISO CD 10918, "Digital compression and coding of continuous-tone still images,"
ISO / IEC JTC1 / SC2 / WG10, (1991)
- [34] G.K.Wallace, "The JPEG still picture compression standard," Communication of the ACM,
vol.34, pp.30-44, (1991)
- [35] D.L.Gall, "MPEG: A video compression standard for multimedia applications," Communi-
cation of the ACM, vol.34, pp.46-58, (1991)
- [36] ISO / IEC CD 13818-2, "Generic coding of moving pictures and associated audio," ISO /
IEC JTC1 / SC29 / WG11, (1993)
- [37] 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)
- [38] W.H.Chen, C.H.Smith and S.C.Fralick, "A fast computational algorithm for discrete cosine
transform," IEEE Trans. Commun., vol. COM-25, pp.1004-1009, (1977)
- [39] M.Miyahara, "Quality assessments for visual service," IEEE Communications Magazine,
vol.26, no.10, pp.51-60, (1988)
- [40] M.Miyahara, K.Kotani, "Block distortion in orthogonal transform coding - analysis, minimi-
zation, and distortion measure," IEEE Trans. Commun., vol. COM-33, pp.90-96, (1985)
- [41] T.P.Barnwell, III, "Subband coder design incorporating recursive quadrature filters and opti-
mum ADPCM coders," IEEE Trans. Accoust., Speech, Signal Processing, vol. ASSP-30,
pp.751-765, (1982)
- [42] N.Sakurai, K.Irie and R.Kishimoto, "Subband DCT codec applied to HDTV transmission
system," IEICE Trans. Commun., vol. E76-B, pp.431-437, (1993)
- [43] P.H.Westerink, D.E.Boekee, J.Biemon and J.W.Woods, "Subband coding of images using
vector quantization," IEEE Trans. Commun., vol. COM-36, pp.713-719, (1988)

- [44] P.H.Westerink, J.Biemoed and D.E.Boeke, "An optimal bit allocation algorithm for sub-band coding," Proc. of Inter., Conf., Acoust., Speech, Signal Processing, ICASSP'88, pp.757-760, (1988)
- [45] S.G.Mallat, "A theory for multiresolution signal decomposition: the wavelet representation," IEEE Trans. Pattern Anal. Mach. Intel., vol. PAMI-11, pp.674-693, (1989)
- [46] M.Vetterli, C.Herley, "Wavelets and filter banks: theory and design," IEEE Trans. Signal Processing, vol. SP-40, pp.2207-2232, (1992)
- [47] J.Kovacevic, M.Vetterli, "Nonseparable multidimensional perfect reconstruction filter banks and wavelet bases for R^n ," IEEE Trans. Info. Theory, vol. IF-38, pp.533-555, (1992)
- [48] 福田, 吉岡, 川中, "サブバンド画像符号化用2次元SSKFの設計," 信学春季全大, D-246, (1993)
- [49] N.S.Jayant, P.Noll, "Digital Coding of Waveforms," Prentice-Hall, (1984)
- [50] 吉川 敏則, 川橋 正昭, 東原 紘道 共著, "工学における数値計算法," 日本理工出版会, (1974)
- [51] 宮原 誠 著, "系統的画像符号化," アイピーシー, (1990)
- [52] H.Gharavi, A.tabatabai, "Application of quadrature mirror filtering to the coding of monochrome and color images," Proc. of Inter., Conf., Acoust., Speech, Signal Processing, ICASSP'87, pp.2384-2387, (1987)
- [53] P.P.Vaidyanathan, "Quadrature mirror filter banks, M-band extensions and perfect-reconstruction techniques," IEEE ASSP Magazine, vol.4, pp.4-20, (1987)
- [54] K.Irie, R.Kishimoto, "A study on perfect reconstructive subband coding," IEEE Trans. Circuit and Systems for Video Tech., vol.1, pp.42-48, (1991)
- [55] 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)
- [56] 小林 幹雄, 福田 安蔵, 鈴木 七緒, 安岡 善則, 黒崎 千代子 共編, "数学公式集," 共立全書, (1959)
- [57] M.Antonini, M.Barlaud, P.Mathieu and I.Daubechies, "Image coding using wavelet transform," IEEE Trans. Image Processing, vol.1, pp.205-220, (1992)
- [58] 日下, 湯山, "テレビジョン・システム評価用デジタル標準画像," テレビ誌, vol.39,

- pp.1010-1013, (1985)
- [59] M.One, et al., "SIDBA standard image data base," MIPC Report79-1, (1979)
- [60] F.Kossentini, W.C.L.Chung and M.J.T.Smith, "A jointly optimized subband coder," IEEE Trans. Image Processing, vol.5, pp.1311-1323, (1996)
- [61] C.K.Cheong, K.Aizawa, T.Saito and M.Hatori, "Subband image coding with biorthogonal wavelets," IEICE Tran. Fundamentals, vol. E75-A, pp.871-881, (1992)
- [62] 高橋 磐郎, 藤重 悟, "離散数学," 岩波講座 情報科学 -17, 岩波書店, (1981)
- [63] 原島 博 監修, "画像情報圧縮," オーム社, (1991)
- [64] C.Calvagno, G.A.Mian and R.Rinaldo, "Computation of the coding gain for subband coders," IEEE Trans. Commun., vol. COM-44, pp.475-487, (1996)
- [65] 小松, 瀬崎, 安田, "画像のサブバンド符号化に用いる低次の完全再構成フィルタの設計," 信学論 A, vol. J77-A, pp.1100-1109, (1994)
- [66] 小野 定康, 鈴木 純司, "わかりやすい JPEG / MPEG2の実現法," オーム社, (1995)
- [67] 甲藤, 安田, "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," 信学技報, IE91-10, pp.33-40, (1991)
- [68] 浜田, 松本, "画像の局所的変化度による雑音マスキング効果を考慮した直交変換係数の最適量子化," 信学論 B-I, vol. J75-B-I, pp.791-801, (1992)
- [69] 松田, 伊東, 森, 宇都宮, "視覚特性を考慮した可変ブロック形状 KL 変換符号化法," TV 学会誌, vol.49, pp.931-934, (1997)
- [70] T.Fukunaga, Y.Yamada, S.Tazaki, "Subband coding of images using VQ with the weighted distortion measure," Proc. of Inter. Picture Coding Sympo., PCS '96, pp.605-610, (1996)
- [71] 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)
- [72] B.Chitprasert, K.R.Rao, "Human visual weighted progressive image transmission," IEEE Trans. Commun., vol. COM-38, pp.1040-1044, (1990)
- [73] 宮原, 小谷, 堀田, 藤本, "客観的画質評価尺度 (PQS) - local feature の考慮と汎用性 -, " 信学論 B-I, vol. J73-B-I, pp.208-218, (1990)
- [74] 有本 卓, "信号・画像のデジタル処理," 産業図書, (1980)

本研究に関する研究業績

1 . 発表論文

- [1] 亀田 昌志 , 大竹 孝平 , "Subband Coding における最適周波数分割 , " 信学論(A), vol. J74A, no.5, pp.808-811, (1991)
- [2] 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 , " 信学論(A), vol. J80A, no.1, pp.237-247, (1997)
- [3] 亀田 昌志 , 大竹 孝平 , 宮原 誠 , " 等方性相関モデルに基づいたサブバンド符号化における最適周波数分割 , " 信学論(A), vol. J80A, no.11, pp.2036-2040, (1997)
- [4] 亀田 昌志 , 大竹 孝平 , 宮原 誠 , " 画像信号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 , " 信学論(D-II) (採録決定)

2 . 国際会議

- [1] M.Kameda, K.Ohtake, "Optimum frequency band partition for subband image coding," Proc. of IEEE International Conference on System Engineering, pp.600-603, (1992)
- [2] M.Kameda, K.Ohtake, M.Miyahara, "Optimum partition of frequency band in 2-dimensional subband coding," Proc. of International Picture Coding Symposium, PCS '96, pp.497-502, (1996)

3 . 研究会・口頭発表

- [1] 亀田 昌志 , 大竹 孝平 , "波動変換符号化の基礎的検討,"信学信越支部大, 158, (1990)
- [2] 亀田 昌志 , 大竹 孝平 , "Subband Coding における最適周波数分割 , " 信学技報 , CS90-51, pp.37-41, (1990)

- [3] 亀田 昌志, 大竹 孝平, "サブバンド符号化における冗長電力最小帯域分割方式," 1991 信学春季全大, A-284, (1991)
- [4] 亀田 昌志, 大竹 孝平, "画像のサブバンド符号化の実験的考察," 信学信越支部大, 110, (1991)
- [5] 亀田 昌志, 宮原 誠, 大竹 孝平, "2次元サブバンド符号化における最適周波数分割," 平7北陸連大, F-43, (1995)
- [6] 亀田 昌志, 大竹 孝平, 宮原 誠, "等方性相関モデルを用いた2次元サブバンド符号化における最適周波数分割," 信学情報・システムソサイエティ大会, D-181, (1996)
- [7] 亀田 昌志, 大竹 孝平, 宮原 誠, "等方性相関モデルによる最適周波数分割の導出とフィルタバンクの実現," 画像符号化シンポジウム, PCSJ '96, P-3.15, (1996)
- [8] 亀田 昌志, 大竹 孝平, 宮原 誠, "実画像における最適周波数分割の導出," 信学情報・システムソサイエティ大会, D11-39, (1997)
- [9] 亀田 昌志, 大竹 孝平, 宮原 誠, "帯域ブロックに基づいた画像の最適周波数分割," 画像符号化シンポジウム, PCSJ '97, P-2.5, (1997)
- [10] 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視知覚特性を考慮した最適周波数分割," 信学技報, CQ97-65, pp.45-52, (1997)