

多重記述符号化における冗長係数群の最適化手法

石川 孝明^{†1} 渡辺 裕^{†2}

本研究では、複素信号の位相成分に着目した多重記述符号化における冗長係数群の最適化手法を提案する。多重記述符号化は、入力信号を複数のビットストリームに分割し、独立に符号化する。各ビットストリームには、任意の量の冗長な情報が多重化されており、伝送時のエラー耐性を高めることができる。提案手法は、複素信号の振幅成分を復号器において補間し、位相成分を冗長な情報として多重化する。実験により、従来手法と比較して、冗長係数群を効率的に符号化可能なことを示すことを示す。

Optimization Method of Redundant Coefficients for Multiple Description Image Coding

TAKA AKI ISHIKAWA^{†1} and HIROSHI WATANABE^{†2}

We propose a new optimization method of redundant coefficients for multiple description image coding. Encoder of multiple description coding divide an input signal into various source of information, and then encode the information to multiple bitstreams individually. Through the encoding process, the encoder can multiplex arbitral redundant information in each bitstream. It increases robustness against packet losses during an image transmission. The proposed method interpolates the amplitude component of redundant coefficients, which is represented in complex values, and then it encodes and multiplexes the phase component. Experimental results show that the proposed method can encode the redundant coefficients of multiple description coding more efficiently compared to conventional method.

^{†1} 早稲田大学 国際情報通信研究センター

Global Information and Telecommunication Institute of Waseda University

^{†2} 早稲田大学 国際情報通信研究科

Graduate School of Global Information and Telecommunication Studies of Waseda University

1. ま え が き

画像や映像の伝送において情報の欠落や通信路の切断が生じると、受信側における映像再生品質が著しく低下する。通常は、通信プロトコルにより欠落した情報は補完される。しかし、リアルタイムアプリケーションでは欠落した情報を再送すると遅延が生じ、再生品質が向上するとは限らない。また、高い信頼性が求められるアプリケーションでは、通信路の切断による品質劣化は大きな問題となる。このような通信路の条件とその変化に対応するためには、情報源符号化された画像や映像に、適切な量の冗長な情報を付け加える必要がある。Forward Error Correction (FEC) とそのインターリーブを用いれば、任意の冗長な情報を付与することができる。しかし、パーストエラーなどの予測困難な情報の欠落には対応できない。これに対して、画像や映像を複数の経路を用いて伝送すれば、パーストエラーなどに対しても情報欠落のリスクを分散でき、映像再生品質の低下を抑制可能である。

複数経路を用いて情報伝送を効率良く行うための符号化手法に、多重記述符号化がある。多重記述符号化では、符号化対象となる画像や映像を複数の情報源に分割し、それぞれを独立に符号化する。このとき、各ビットストリームには、重複する情報が多重化されている。そのため、受信側で一部のビットストリームだけを受信したとしても、多重化された情報を元に欠落した情報を補完することで、FEC と同様に、再生品質の著しい低下を防ぐことができる。

多重記述符号化は、1980年代の音声通信におけるチャネル分離を目的とした研究として始った¹⁾。情報理論の研究としては、複数のチャネルに対する様々な歪み許容値を求める問題として位置づけることができ、El Gamal と Cover²⁾ がこの問題に言及して以来、様々な手法が提案され、画像や映像の符号化にも応用されている。

画像の多重記述符号化は、入力信号の分割を画素領域で行う場合と、冗長性の削減を目的として実行される、変換の後に係数領域で行う場合で特性が異なる。画素領域で分割を行う場合、分割後の画像間に高い相関が残されるため、ビットストリーム全体の冗長度は高くなる。そのため、パケットロス率が高い場合に有効である。具体的には入力画像をサブサンプリングする手法³⁾がある。一方、係数領域で行う場合は、係数間の冗長度が低く、パケットロス率が低い場合に有効である。具体的には、離散コサイン変換や離散ウェーブレット変換で得られる係数群を分割する手法⁴⁾⁵⁾がある。

近年になり、変換係数の表現について確率的な概念を導入した手法が提案されている。Sadri らは、オールパスフィルタを用いて信号の位相を変化させ、複数のビットストリーム

を生成する手法を提案している⁶⁾。しかし、無制限に位相を変化させると、符号化対象となる係数群において、係数間の相関が低くなり、符号化効率が低下する問題があった。これに対し、符号化効率の低下を抑制するように位相変化を限定する手法が提案されている⁷⁾。また、Wang らは、Compressive sensing を応用した多重記述符号化を提案している⁸⁾。

変換係数を冗長に生成し、実際に符号化する係数を制限することで、冗長度を制御することも可能である。我々は以前、Goyal らのラティスによる係数分割を応用し、任意の冗長度を有する複数ビットストリーム生成手法を提案している⁹⁾。同手法は、パケットロスが生じない環境に対して有効な、冗長度が0に対応するビットストリームが入力信号の直交分解で導かれ、また、冗長度が1(2倍の情報量)に対応するビットストリームが、入力信号のすべてを重ねた状態であることに着目した。そして、0~1の間に対応する冗長度のビットストリームを生成するために、直交分解で得られる信号と入力信号の重みづけ線形和により、任意の冗長度を持つ信号を定義した。最終的に、各ストリームを独立なベクトルおよび共通するベクトルの合成により表現可能なことを示した。しかし、係数最適化を入力信号に対してのみに適用しており、共通成分となるベクトルに対しては、明示的な最適化処理を行っていない。そこで本検討では、共通ベクトルに対応する冗長係数群の最適化手法を提案する。

本稿は、以下のように構成される。2.では、複素ウェーブレット変換を用いた従来手法について述べる。3.では、提案手法について述べる。また、提案手法で用いるラティスと複素係数の線形量子化についても述べる。4.では提案方式の評価実験を行い、有効性を示す。5.で本検討をまとめる。

2. 従来手法

本章では、複素ウェーブレット変換を用いた従来手法について述べる。Li らは、複素ウェーブレット変換で得られる冗長な係数群の Noise Shaping による最適化過程を、多重記述符号化における係数の最適化に応用した¹⁰⁾。Noise Shaping については次節以降において述べる。典型的な二分割構成の多重記述符号化の符号化システムを図1に示す。図1では、互いに冗長な情報を共有していない複数ビットストリームを生成するための係数最適化部を central noise shaping と表している。一方、ビットストリームの一部が欠落し、復号器に伝送されない場合において、任意の冗長な情報が多重化されたビットストリームを生成するための符号化器に該当するブロックを side noise shaping と表している。

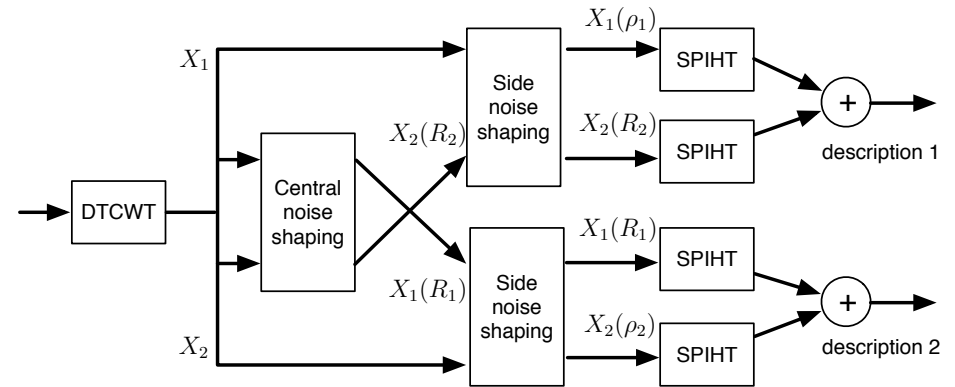


図1 NoiseShaping を応用した多重記述符号化器
Fig. 1 Multiple Descriptions Encoder with Noise Shaping.

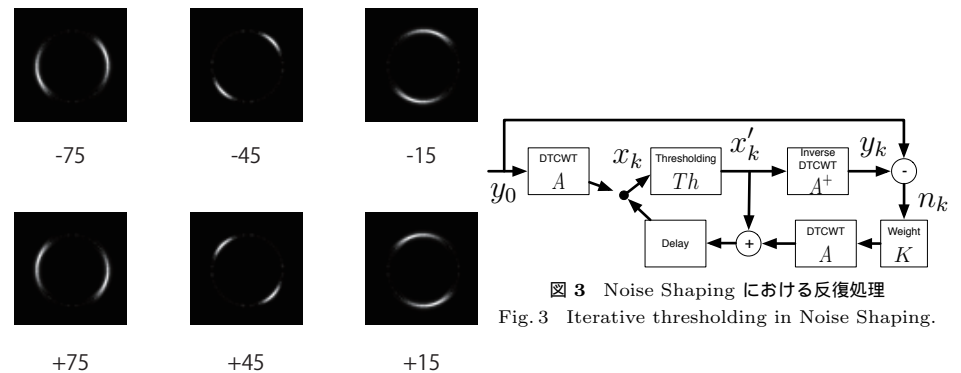


図3 Noise Shaping における反復処理
Fig. 3 Iterative thresholding in Noise Shaping.

図2 複素ウェーブレット変換のサブバンド画像
Fig. 2 Subband images of CWT.

2.1 複素ウェーブレット変換

2.1.1 変換係数の特徴

Dual Tree Complex Wavelet Transform (DTCWT) は、6種類の方向分離特性やシフト不変性などの特徴を有する変換手法である。複数の離散ウェーブレット変換により得られる変換係数を組み合わせることで、複素信号を構成することが可能である。実数の離散ウェーブレット変換と比較し、画像を表現する基底関数の種類が豊富であり、入力画像が曲線などエッジ成分を多く含む場合に効率よく表現することができる。図2に、離散ウェーブレット変換により得られるサブバンドの振幅成分画像を示す。入力画像は、一定の幅を持つ円が描かれた画像である。6種類のサブバンドに、それぞれ異なる方向のエッジ成分が検出されていることがわかる。

このような特性を持つDTCWTは、実数のウェーブレット変換と比較して4倍の変換係数を有する冗長な変換である。これは、複素ウェーブレット変換における複素信号が、複素空間の基底関数ではなく、実部と虚部を実数空間の基底関数の集合で展開しているためである。通常の画像符号化では、4倍の冗長度からの冗長度削減は、それ以外の目的を持たないが、多重記述符号化では、最適化の過程で冗長度が変化するため、最適化処理自体を冗長度制御に利用できる。

2.2 Noise Shaping

入力画像に対するDTCWT係数は信号に変化が無いような特別な場合を除いて零係数を含まない。そこでKingsburyらは、Noise Shapingと呼ばれる係数の最適化手法を提案している¹¹⁾。これは、複素ウェーブレット変換で得られる信号を少数の疎数列に加工する非線形な処理である。

Noise Shapingは、複素ウェーブレット変換で得られる係数群を初期元とし、非零係数を生成する空間と再構成時の品質を保存する空間を交互に射影することで非零係数の数を限定し、射影により生じるノイズを最小化している。このような最適化処理により、4倍に増加した変換係数を所定の数の非零係数に置き換えることが可能である。同方式を応用した画像符号化方式も提案されており、JPEG 2000などの符号化方式よりも高い符号化効率を実現している¹²⁾。

以下に、Noise Shapingのアルゴリズムを示す。また、Noise Shapingの回路図を図3に示す。ただし、複素ウェーブレットの順変換をAで表し、逆変換をA⁺で表す。

STEP1) 入力信号yを変換し、係数列を得る。この係数列x = Ayは非疎系列である。初めて得られる係数群をx₀とし、入力画像をy₀とする。

STEP2) x_kに閾値処理を行う。Th_kを下回る係数を0に置き換えることでx_kは、疎な系列x'_kとなる。

STEP3) x'_kを逆変換し、画素領域に戻し再構成画像y_kを得る。再構成画像は、閾値処理の影響により多くのノイズを含んでいる。

STEP4) 入力信号y₀とy_kとの差分をとり、ノイズ成分n_kを抽出する。

STEP5) 閾値処理で減衰したノイズ成分n_kのエネルギーを補うためにK倍する。その後、順変換し係数領域でのノイズ成分を求める。ノイズ成分を加算することで、新たな係数列x_{k+1}を得る。これは図3の遅延に該当する部分であり、次の反復処理に続く。すなわち、STEP2に戻り、閾値処理を施す。

STEP2で用いる閾値は、反復回数が増加する度に徐々に小さくする。Wangらは、初期閾値を256とし、反復ごとに閾値を1.0下げ、閾値が1になるまで処理を行えば、十分に比較可能な最適化した係数群を得られるとしている¹²⁾。

閾値を徐々に下げることで、ノイズ成分は徐々に小さくなり、同時に疎系列を画素領域に変換した画像は、徐々に入力信号に近づく。したがって反復により、疎系列はL0ノルムを増加させつつノイズの少ない系列になる。より詳しい説明は、文献¹¹⁾などに述べられている。

3. 提案方式

本章では、ラティスによる係数分割と信号の位相成分に着目した冗長係数の最適化手法について述べる。提案手法は、以下に示す処理フローに基づいて複数のビットストリームを生成する。入力信号を二分割する場合の符号化器を図4に、description 1に対するサイドデコーダを図5示す。図中のTは、実部虚部と振幅位相を変換する行列であり、Iは、振幅の線形補間回路である。補間処理は、上下左右の4点を用いた単純な線形補間を行う。

STEP1) まず始めに、入力信号を複素ウェーブレット変換を用いて変換し、複素数の係数情報を得る。係数の総数は、入力信号の信号総数に対して四倍に増加する。実数値のツリーから求まる係数は複素信号であり、実部と虚部の二変数からなる。さらに、サブバンド上において正方向のエッジを表す複素数列と負方向を表す複素数列の二種類が得られる。

STEP2) 振幅と位相係数をラティスにより分割し、異なる格子点上の係数群に分割する。二分割のためのラティスは、図6のような構造である。格子点のラベルとして0,1を用いる。

STEP3) 得られた複素数列の実部と虚部を式(1)(2)を用いて振幅と位相に変換する。ただし、負の値が生じない様に条件が付けている。複素信号をA_n + B_niで表す。nは系列

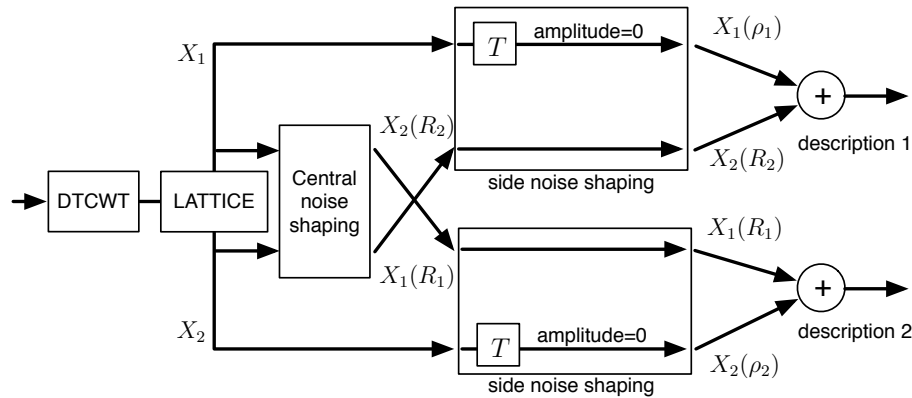


図 4 位相を冗長係数群として符号化する提案方式
Fig. 4 Multiple Descriptions Encoder using phase information.

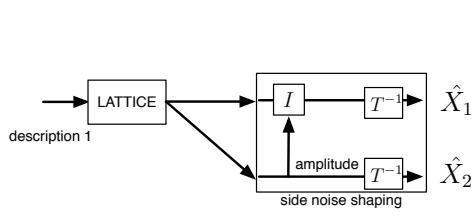


図 5 振幅を補間するサイドデコーダの提案方式
Fig. 5 Multiple Descriptions Side Decoder with amplitude interpolation.

0	1	0	1
1	0	1	0
0	1	0	1
1	0	1	0

Description 1

0	1	0	1
1	0	1	0
0	1	0	1
1	0	1	0

Description 2

図 6 ラティスを用いた係数の分割 (分割数が 2 の場合)
Fig. 6 Coefficients division by lattice (two division).

中のインデックスを表す.

$$amplitude_n = \sqrt{A_n^2 + B_n^2} \quad (1)$$

$$phase_n = \begin{cases} \tan^{-1}(\frac{A_n}{B_n}), & \tan^{-1}(\frac{A_n}{B_n}) \geq 0 \\ \tan^{-1}(\frac{A_n}{B_n}) + 2\pi, & \tan^{-1}(\frac{A_n}{B_n}) < 0 \end{cases} \quad (2)$$

STEP4) インデックス 0 とインデックス 1 の係数群について Noise Shaping を用いて最適化し, 少数の非零係数を導出する. これらを中央デコーダ (分割したすべてのストリームを受信する復号器) にて得られる符号化対象の係数群とする. なお, Noise Shaping において入力画像との差分を取る時には, 一時的にすべての係数を合成して複素ウェーブレット逆変換を施す.

STEP5) インデックス 0 の係数群を, Noise Shaping を用いて最適化する. このとき, インデックス 1 の係数群は, 所定のビットレートを実現する係数群であり, すでに中央デコーダ用に生成されている. サイドデコーダ (分割した一部のストリームのみを受信する復号器) に対するビットストリームは, 最適化の対象の係数群として位相成分だけを用いる. 具体的な処理は次節で述べる. 最適化処理では, インデックス 0 のみに閾値処理を施し, インデックス 1 の係数群は不変とする. 最終的に所定のビットレートを達成するインデックス 0 の信号が, 冗長係数群となる.

以上のようにして得られるインデックス 0 の信号と, インデックス 1 の信号を一つのビットストリームに合成し, 任意の冗長な情報量を有する信号として出力する.

3.1 最適化する冗長係数群の構成

提案手法では, 最適化する係数群を複素係数から得られる位相成分のみに限定する. これは, 複素数信号の実部と虚部がどちらも符号情報を有し係数値が正負に分布するのに対して, 振幅と位相はどちらも正値に限定可能なためである. 正の値で表される信号の振幅は, 欠落した係数を線形補間により再構成しやすいという特徴がある. 一方, 位相係数は, 任意の振幅に対して単円上の回転角を決定している. これは実信号における係数の符号情報に対応している. したがって, 線形補間が当てはまらない場合には補間した係数値が真値から大きく離れることになる. そのため, 必ずしも補間がうまくいくとは限らない. そこで, 提案手法では, 欠落した情報を補うために用いる冗長な情報として, 振幅成分よりも位相成分を優先して符号化する.

3.2 係数量子化と逆量子化

振幅と位相の量子化は, 複素数平面における母点を表すインデックスで表される¹³⁾. 母

点の密度が高ければ、係数の精度が高くなるが、インデックスの総数も増加するため情報量が増える。なお、振幅位相成分から実部虚部に戻す過程において振幅がゼロの場合、式(3)(4)にあるように、位相成分に値が存在しても実部虚部の係数は常に零係数になる。提案方式では位相成分だけを冗長係数として符号化し、振幅成分をゼロとしている。したがって、補間処理を正しく行うためには、係数の補間処理を逆量子化直後の係数群に対して行う必要がある。

$$A_n = \text{amplitude}_n \times \cos(\text{phase}_n) \quad (3)$$

$$B_n = \begin{cases} \text{amplitude}_n \times \sin(\text{phase}_n), & \text{phase}_n \leq \pi \\ \text{amplitude}_n \times \sin(\text{phase}_n - 2\pi), & \text{phase}_n > \pi \end{cases} \quad (4)$$

4. 評価実験

提案手法の有効性を確認するための評価実験を行う。実験では、ビットストリームの分割数を2とし、評価は次の2点について行う。一つは、中央デコーダにおける符号化効率である。これは係数の欠落が生じない場合である。もう一つは、サイドデコーダにおける符号化効率である。これは、一方のインデックスから得られる係数群を、冗長な係数群として多重化している場合である。

4.1 実験条件

実験には、 512×512 [pel] (グレースケール) の Lena と Barbara を用いた。従来手法は、複素ウェーブレット変換を、正の角度のエッジ成分を含むサブバンドから構成されるビットストリームと、正の角度のエッジ成分を含むビットストリームに分離する。復号器における合成では、係数補間を行わず、欠落係数は冗長に符号化された係数群をそのまま適用することで再構成画像を生成する。一方、提案手法は、係数分割にラティスを用い、欠落した係数は振幅成分を周囲係数から補間生成し、位相成分は冗長に符号化された係数群を直接用いることで再構成画像を得る。なお、複素ウェーブレット変換には、第1レベルを13-19タップの双直交フィルタ、第2レベル以降を18タップの直交フィルタを使用する。Noise Shaping の閾値処理は、すべての実験で共通して初期値を256、終了値を1とし、反復時に1づつ閾値を下げる。また、従来手法ではSPIHTを用いてビットストリーム生成しているが、本実験では係数を線形量子化し、情報量は量子化係数のエントロピーで表す。

4.2 中央デコーダにおける符号化効率

二つのビットストリームの両方を利用可能な場合の符号化効率を計測する。Lena の符号

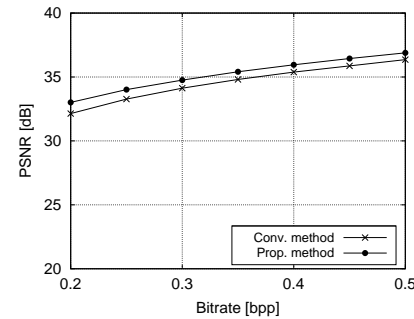


図7 中央デコーダにおける符号化効率 (Lena)
Fig.7 Coding performance at central decoder (Lena).

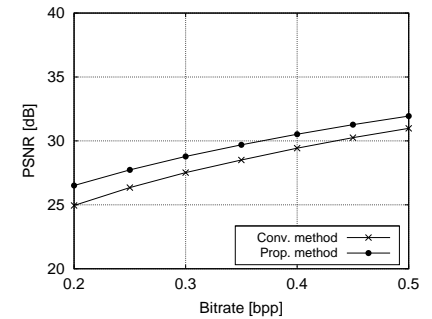


図8 中央デコーダにおける符号化効率 (Barbara)
Fig.8 Coding performance at central decoder (Barbara).

化効率を図7に、Barbara の符号化効率を図8に示す。図の横軸は、インデックス0およびインデックス1のそれぞれの情報量であり、該当ビットレートを実現するように線形量子化の量子化パラメータを調節している。具体的には、小数点以下4桁の精度でビットレートが等しくなるパラメータを探索した結果である。図より提案手法と従来手法の中央デコーダにおける符号化効率はほぼ同程度か、提案手法が若干上回る結果になった。

なお、0.1[bpp]以下のビットレートにおいては、従来手法と提案手法の効率が完全に一致する場合があったため、結果から除外している。これは、Noise Shaping の初期閾値が適切に与えられていないためと考えられる。初期閾値で導出されるL0ノルムの大きさが不十分なためであり、適切な値の導出は今後の課題である。

4.3 サイドデコーダにおける符号化効率

次に、二つのビットストリームのうち、いずれか一方のビットストリームだけが利用可能な場合の符号化効率を計測する。Lena の符号化効率を図9と11に、Barbara の符号化効率を図10と12に示す。図より、冗長係数群の情報量によらず、提案手法が従来手法の符号化効率を大きく上回ることがわかる。これは、提案方式が欠落する情報を適切に補間し、かつ、位相成分だけを符号化することで、補間に必要な情報を効率的に伝送できるためと考えられる。

5. むすび

本検討では、多重記述符号化における冗長な情報を表す係数群の最適化手法を提案した。

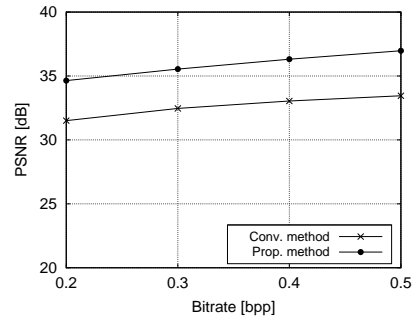


図 9 サイドデコーダに対する符号化効率
($X_1 = X_2 = 0.5$ [bpp], Lena)

Fig. 9 Coding performance at a side decoder
($X_1 = X_2 = 0.5$ [bpp], Lena).

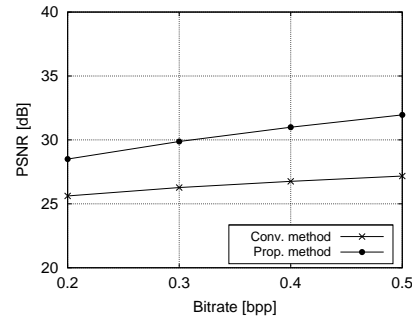


図 10 サイドデコーダに対する符号化効率
($X_1 = X_2 = 0.5$ [bpp], Barbara)

Fig. 10 Coding performance at a side decoder
($X_1 = X_2 = 0.5$ [bpp], Barbara).

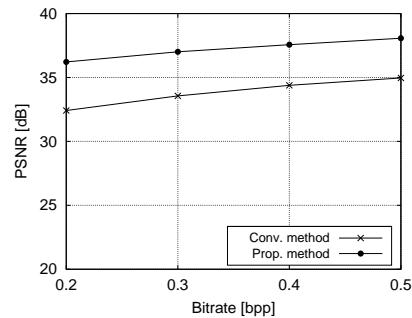


図 11 サイドデコーダに対する符号化効率
($X_1 = X_2 = 1.0$ [bpp], Lena)

Fig. 11 Coding performance at a side decoder
($X_1 = X_2 = 1.0$ [bpp], Lena).

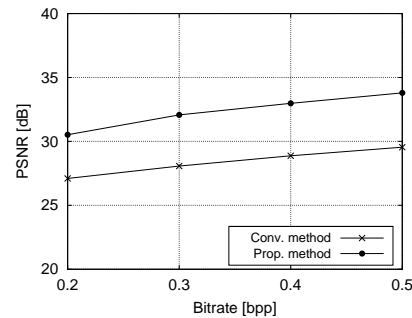


図 12 サイドデコーダに対する符号化効率
($X_1 = X_2 = 1.0$ [bpp], Barbara)

Fig. 12 Coding performance at a side decoder
($X_1 = X_2 = 1.0$ [bpp], Barbara).

参 考 文 献

- 1) Goyal, V.: Multiple Description Coding: Compression meets the network, *IEEE Signal Processing Mag.*, Vol.18, pp.74–93 (2001).
- 2) Gamal, A.E. and Cover, T.M.: Multiple User Information Theory, *Proceedings of the IEEE*, Vol.68, No.12, pp.1466–1483 (1980).
- 3) 松村宏基, 藤井俊章, 谷本正幸: サブサンプリングを用いたマルチストリーム動画画像伝送方式の検討, *AVM39*, pp.29–34 (2002).
- 4) Matty, K.R. and Kondi, L.P.: Balanced Multiple Description Video Coding Using Optimal Partitioning of the DCT Coefficients, *IEEE Trans. on Circuits and Systems for Video Technology*, Vol.15, No.7, pp.928–934 (2005).
- 5) Bajic, I.V. and Woods, J.W.: Domain-Based Multiple Description Coding of Image and Video, *IEEE Trans. on Image Processing*, Vol.12, No.10, pp.1211–1225 (2003).
- 6) Sadri, K. and Shirani, S.: Multiple description coding of images using phase scrambling, *Acoustics, Speech, and Signal Processing, 2004. Proceedings. (ICASSP '04). IEEE International Conference on*, Vol.3, pp.iii – 41–4 vol.3 (2004).
- 7) Uto, T. and Ohue, K.: Multiple Description Coding Based on Phase Scrambling with Adjustable Spread Range, *EURASIP2006 European Signal Processing Conference*, Vol.5 (2006).
- 8) Wang, L., Wu, X. and Shi, G.: A compressive sensing approach of multiple descriptions for network multimedia communication, *Multimedia Signal Processing, 2008 IEEE 10th Workshop on*, pp.445 – 449 (2008).
- 9) 石川孝明, 渡辺裕: フレーム展開による画像の多重記述符号化に関する検討, 情報処理学会 AVM 研究会研究報告, Vol.2009-AVM65, No.16, pp.1–6 (2009).
- 10) Li, L. and Cai, C.: Multiple description image coding using dual-tree discrete wavelet transform, *Intelligent Signal Processing and Communication Systems, 2009. ISPACS 2009. International Symposium on*, pp.655 – 658 (2009).
- 11) Reeves, T.H. and Kingsbury, N.G.: Overcomplete Image Coding Using Iterative Projection-Based Noise Shaping, *IEEE ICIP*, Vol.3, No.3, pp.597–600 (2002).
- 12) J. Yang, Y. Wang, W.X. and Dai, Q.: Image Coding Using Dual-Tree Discrete Wavelet Transform, *IEEE Trans. on Image Proc.*, Vol.17, No.9, pp.1555–1569 (2008).
- 13) Reeves, T.H. and Kingsbury, N.G.: R-D quantisation of complex coefficients in zerotree coding, *Proceedings 11th IEEE Workshop on Statistical Signal Processing 2001* (2001).

提案手法は、入力信号を複素ウェーブレット変換を用いて複素信号で表現し、ラティスにより係数群を分割した。また、冗長な情報を表す係数群には、複素信号の位相成分だけを符号化し、振幅成分は復号器において線形補間した。実験により、従来手法と比較して冗長な情報を表す係数群を効率良く符号化できることを示した。