テクスチャの再現性を考慮した H.264 符号化モード制御方式

H.264 Coding Mode Control Scheme for Effective Texture Reality Improvement

正会員 吉野 知伸[†], 正会員 内藤 整[†], 正会員 酒澤 茂之[†], 正会員 小池 淳[†]

Tomonobu Yoshino[†], Sei Naito[†], Shigeyuki Sakazawa[†] and Atsushi Koike[†]

Abstract A further improvement of subjective picture quality for H.264/AVC High Profile coding is required to maintain stable broadcasting quality even under low bit-rate transmission. The macroblock coding mode decision is one of the key factors in improving coding efficiency. In conventional methods, this mode is usually decided based on R-D optimization criteria, and the selected mode contributes to the high coding efficiency. However, it is not necessarily appropriate to avoid the subjective picture quality degradation so an improvement of the coding mode decision scheme is strongly required. Therefore, we propose an advanced coding mode decision scheme. The proposed scheme selects the coding mode based on new evaluating measure in order to pursue the texture reality. Experimental results show that our proposed scheme improves the subjective picture quality improvement, and the DSCQS(Double Stimulus Continuous Quality Scale) gain reaches a maximum 8.6 points.

キーワード: H.264/MPEG-4 AVC, 符号化モード選択, HDTV, 低ビットレート符号化, 主観画質

1. ま え が き

IP ネットワーク上で HDTV を配信するサービスの普及が期待されている。FTTH や DSL に代表されるコンシューマ回線を用いて同サービスを提供する場合、映像データに許容される帯域は 10Mbps 以下である。したがって、10Mbps 程度の低ビットレート環境において HDTV を安定的な品質で配信するためには、HDTV に対して高い符号化効率を実現する符号化方式が求められる。近年では、同符号化方式として、高精細映像に適した符号化ツールを備えた H.264/MPEG-4 AVC High Profile(以下、H.264) が最も有力である。

H.264 で高い符号化効率を得られる要因の一つに、マクロブロック (MB) 単位での符号化モード適応選択が挙げられる。H.264 では、入力信号の特徴に応じた符号化を可能とする様々な符号化モードが定義されており、同符号化モードの適応選択により、高精度の予測値を得ることができる。予測性能が符号化効率に与える影響は大きいので、H.264 エンコーダの性能は、符号化モード選択に関する符号化制

御の成否に強く依存する。

H.264 の代表的なリファレンスエンコーダである Joint Model(JM)¹⁾ では、MB 単位の符号化モード選択手法として、R-D 最適化法²⁾ が採用されている。同手法は決められた符号量の下で客観画質 (PSNR) を最大化する符号化モードを選択するため、同手法の符号化モード選択により高い符号化効率を得られる。同手法は、また、イントラ予測方向選択にも用いられており、同選択に特化した同手法の性能改善として、同等の客観画質を保持しつつ処理の高速化を実現する手法³⁾ も提案されている。しかしながら、文献 2)3) の手法は単一の MB のみにおける客観画質を評価するため、同手法の選択結果は必ずしも主観画質を最大化するものではなく、さらに近傍 MB も含めた局所領域における主観画質の低下を検知できない。したがって、同手法に基づく符号化モード選択に起因して主観画質の低下を招く恐れがある。その典型例としてテクスチャ再現性の低下があり、本稿では具体的な二つのケースに注目する。一つは、局所的に類似したイントラ予測方向のイントラ符号化モードが連続することにより、類似した絵柄の領域 (以下、類似領域) が拡大することに起因する、テクスチャ単調化である。もう一つは、符号量重視の符号化モード選択に起因する、オブジェクト境界成分の再現性低下である。特に後者については、類似する輝度値を持つほぼ平坦なオブジェクト同士の境界成分 (本稿では、同成分をエッジ成分と呼ぶ) において著しい。テクスチャ再現性低下の問題は、

2007 年 3 月, 情報処理学会 AVM 研究会にて発表

2007 年 6 月 8 日受付, 2008 年 3 月 12 日最終受付, 2008 年 5 月 22 日採録

[†] 株式会社 KDDI 研究所 映像通信グループ

(〒 356-8502 埼玉県ふじみ野市大原 2-1-15, TEL 049-278-7587)

[†] KDDI R&D Laboratories Inc. Visual Communication Laboratory (2-1-15 Ohara Fujimino-Shi Saitama 356-8502, Japan)

特に低ビットレート符号化において顕著である。

以上の背景から、近年では、H.264 符号化映像の主観画質を改善することを目的として、H.264 エンコーダにおける符号化モード選択手法を最適化する研究が行われている。文献4)では、B ピクチャにおけるダイレクトモードを適応的に用いることにより、主観画質を保ちつつ符号量を削減する手法が提案されている。しかしながら、同手法における符号化モード選択は R-D 最適化法の枠組みに基づいており、テクスチャ再現性に関しては考慮がなされていない。文献5)~7)では、主観画質の改善を目的とした符号化モード選択手法の枠組みが示されている。文献5)では、R-D 最適化法におけるコスト関数について、絵柄に応じて重み付けを行う手法が示されているが、テクスチャの再現性も含めて主観画質を改善するためのパラメータ設定に関して、具体的な方法が明示されていない。また文献6)および文献7)では、SSIM(Structure Similarity)⁸⁾に基づく符号化モード選択手法が示されている。しかしながら、同手法は単一の MB に対する評価のみしか行わないため、数個の MB で構成される局所領域でのテクスチャ再現性の低下を検知することができず、これに伴う主観画質の低下を抑制することは困難である。以上より、文献4)から文献7)で示した従来の符号化モード選択手法ではテクスチャ再現性の低下を十分に抑止することは困難である。

以上の背景から、本稿では、HDTV に対する 10Mbps 程度の符号化ビットレートを対象として、主観画質の改善を目的とした符号化モード選択手法の提案を行う。提案手法は、処理 MB の近傍領域におけるテクスチャ再現性低下について、隣接する MB 間の輝度値の相違度および処理 MB における輝度値の分散を評価し、符号化モード選択制御に起因する類似領域の拡大によるテクスチャ単調化およびエッジ再現性の低下を検知する。テクスチャ再現性低下が見込まれる領域では、類似領域の拡大を抑制し、かつ高いエッジ再現性を得る符号化モード選択制御を行う。提案手法の性能評価として主観評価実験を行い、R-D 最適化法と比較して主観画質が改善することを確認する。

2. 従来手法の課題

2.1 R-D 最適化法

MB 単位の符号化において、符号化に起因する符号化歪み(具体的には原画像に対する差分2乗和)を D 、符号量を R 、割当て符号量を R_c とする。文献2)では、MB 単位の符号化における符号化モード選択において、割当て符号量の制約の下で客観画質を最大化することを目的とし、式(1)の最小化問題が示されている。すなわち、式(1)の解となる符号化モードが客観画質の観点から符号化効率が最大となる符号化モードであることを表している。また、式(1)にラグランジュ乗数を導入することで、式(2)のコスト関数が定義されている。式(2)は、コスト値 J が最小となる符号化モードを選択することで、符号化効率が最大となる

ことを表している。JM では、符号化モード選択手法として R-D 最適化法が適用されており、以下では JM を従来手法と呼ぶ。

$$\min\{D\}, \text{subject to } R < R_c \quad (1)$$

$$J = D + \lambda \times R \quad (2)$$

2.2 主観画質の低下

上述の R-D 最適化法により高い符号化効率が達成されるものの、同手法による MB 単位の符号化モード選択の結果、主観画質が著しく低下するケースがある。特に、動領域におけるテクスチャ再現性が低下することで、局所的に絵柄が不自然となり主観画質の著しい低下を招く。なお、本論文における動領域は、MB4 個から 8 個程度で構成され、画面間での動きを含む領域を指す。動領域に特有のテクスチャ再現性低下の代表例として、以下が挙げられる。

- 類似したイントラ予測方向を持つイントラ符号化モードの集中により、類似したテクスチャを持つ MB が連続する(類似領域が拡大する)ことで、局所領域のテクスチャが単調化する。

- 符号化モード選択において、符号量の抑制が過度に重視され、符号化歪みの抑制が軽視されることにより、エッジ成分の再現性が低下する。

(1) イントラ符号化モードの集中による類似領域拡大
H.264 におけるイントラ予測では、隣接する符号化済み MB の画素値を予測値として、予測方向に従って処理 MB 内の画素における予測値が設定される。ここで、高速なカメラワークなどにより絵柄がぼける映像においては、カメラワークの方向に沿って似た画素値を持つ画素が並ぶため、予測方向がカメラワークの方向と近いイントラ予測では、高精度な予測が期待できる。このため、低ビットレート符号化においても、符号化歪みは小さく保たれる。したがって、客観画質の観点でイントラ予測は非常に効果的である。しかしながら、予測性能が高いために残差信号の DCT 係数は小さく、粗い量子化により同係数が失われる可能性が高い。したがって、イントラ符号化を適用することにより、処理 MB における符号化映像の画素値は、周辺 MB における画素値と近い値をとる傾向がある。特に、処理 MB と周辺 MB で類似したイントラ予測方向が適用されている場合、両 MB 間のテクスチャ類似性が高まる。また、同絵柄では、カメラワーク方向に近いイントラ予測方向の予測精度が高いため、局所的に類似したイントラ予測方向を持つ MB が連続する可能性が高い。以上より、同映像では類似したイントラ予測方向を持つイントラ符号化モード MB が集中し易く、同 MB が集中することにより、類似領域が拡大しテクスチャが単調化する。

イントラ符号化モードでは予測方向単位で予測値を得るが、一方でイントラ符号化モードに関してはブロック単位で

予測値を得るため、インタ符号化モードは、イントラ符号化モードよりも近傍領域に対して処理 MB のテクスチャが類似する可能性は低い。したがって、インタ符号化モードが局所的に集中した場合でも、当該領域における動きの質感低下による主観画質の低下は避けられる。

テクスチャ単調化による主観画質低下は、インタ符号化フレームにおいて、テクスチャ単調化した領域が顕在化することにより発生する。また、同顕在化は、テクスチャ単調化した (MB 間の類似性が高い) 領域が、MB 間の類似性が低いインタ符号化モードの MB と混在することによって起因する。特に、動領域では、インタ符号化モード MB の類似性が低いために、同顕在化による主観画質低下が著しい。一方、イントラ符号化フレームでは、インタ符号化フレームにおいて、インタ符号化モードが適用されていた領域もイントラ符号化されるため、同領域の類似性も高くなる。結果、特定領域でのテクスチャ単調化は顕在化しづらくなり、主観画質に与える影響は小さくなる。

イントラ符号化モードが集中することにより、主観画質が低下する問題を実画像上で確認するため、符号化映像における局所領域の類似性に関し、イントラ符号化およびインタ符号化双方の符号化結果について分析した。対象素材として、ITE HDTV 標準動画像シーケンス "Soccer action" を用いた。なお、同映像は、サッカー選手およびサッカーボールの動きに追従した高速なパニングを含んでいる。また、エンコーダとして JM を用いた。図 1 は同映像中のあるインタ符号化フレームを示しており、以下では同図中の枠で示す領域に着目する。同領域は、インタ符号化フレームにおいて、イントラ符号化モードの集中により類似領域の拡大が顕著となる箇所である。

ここで、類似領域の拡大を定量的に測るための指標を導入する。類似領域の拡大は、MB 単位の符号化モード選択において、類似したイントラ予測方向を持つイントラ符号化モードが隣接 MB 間で連続して選択され、同 MB 間で輝度値が類似することによって起因する。本稿では、類似領域の拡大を表す指標として、近傍 MB の輝度値に対する処理 MB の輝度値の相違度を用いる。相違度 s を式 (3) で定義する。

$$s = \min_{t \in T} \sum_{x,y} (p(C, x, y) - p(t, x, y))^2 \quad (3)$$

式 (3) において、 C は処理 MB を表し、 T は処理 MB の左、左上、上および右上に接する計四つの MB に相当する画素領域である。また、 t は T の内部に位置する任意の単位領域 (16 画素 $\times$ 16 ライン) を表し、MB 位置に限定されず画素単位でスライド可能である。図 2 に C および T の関係を示す。また、 $p(R, x, y)$ はある単位領域 R 内の輝度値であり、 x, y は同領域内における画素位置を指す。すなわち、 s は、処理 MB について隣接する領域に対する差分 2 乗和であり、同値が小さいほど類似性が高いことを表している。

図 3 に符号化モードとテクスチャ単調化の関係を見るための予備実験結果を示す。同図は、図 1 の枠で示す領域に対して、イントラおよびインタの各符号化モードのみで符号化した画像について、MB ごとの相違度の大きさを濃淡で塗り分けている。同図下段は、濃淡と相違度の大きさの対応を示している。また、比較のため原画像に対する結果も示す。符号化ビットレートは 8Mbps であり、符号化条件を表 1 に示す。

結果より、原画像では一定して相違度が高い傾向にあることに対し、イントラ符号化映像およびインタ符号化映像では相違度が低いことがわかる。また、特にイントラ符号化では、インタ符号化に対して、当該領域における MB 毎の相違度が低いことがわかる。実際、その相違度の低下は、当該領域における MB 毎の平均で 15% 程度であることを確認しており、類似性が高い領域が拡大しているといえる。

なお、符号化条件について動きベクトル探索範囲がテ

図 1 評価映像
Test sequence.

図 2 C および T
Location of C and T .

図 3 予備実験結果 (局所領域の相違度)
Pre-experimental results. (Difference of local region)

クスチャ単調化に与える影響を確認するため、同範囲を $\pm 256 \times \pm 256$ とし、同範囲以外の条件は表1として符号化実験を行った。結果、相違度について図3と同様の結果が得られ、テクスチャ単調化による主観画質低下が発生することを確認した。したがって、同範囲の設定は、表1における設定値で充分であるといえる。

また、テクスチャ単調化により主観画質が低下する問題の解決策の一つに、局所領域内における同一イントラ予測方向の連続を抑制するアプローチが考えられる。ここで、低ビットレート符号化におけるイントラ予測では、特に高い予測性能が得られる絵柄に対して、粗い量子化により残差信号のAC成分が失われ、その符号量は小さくなる。またこのとき、予測値と符号化結果の画素値がほぼ等価となる。本稿で対象とする絵柄では、カメラワークの方向に近い予測方向で高い予測性能が得られる。そのため、低ビットレート符号化におけるR-D最適化法に基づくイントラ予測方向の選択において、イントラ予測方向の選択肢は、カメラワークに近い二つもしくは三つ程度に限定される。その中で、同一イントラ予測方向の連続を制限しても、特定のイントラ予測方向が周期的に出現することとなり、テクスチャ単調化を回避することはできない。したがって、同一イントラ予測方向の連続を制限するだけでは、主観画質低下の問題を解決することはできない。さらに、同絵柄の領域に対して、カメラワークに適した複数のイントラ予測方向を任意に配置するアプローチも考えられる。しかしながら、同絵柄において、イントラ予測方向の選択肢は限定的であり、カメラワークの方向および速度によっては、特定のイントラ予測方向の周期的な出現を確実に排除することが困難であることが懸念される。以上より、イントラ符号化モードの枠内におけるイントラ予測方向の制御によるアプローチだけでは、テクスチャ単調化による主観画質低下の問題を確実に解決することは困難である。

また、R-D最適化法における λ を適応的に変更することにより、イントラ符号化モードおよびインタ符号化モードの選択を制御することで、類似領域の拡大を抑制する手法も考えられる。しかしながら、本論文で対象とする絵柄では、イントラ符号化モードは、インタ符号化モードに対して、符号化歪みおよび符号量とも符号化効率の観点から優れる傾向にあるため、コスト関数における λ の変更が意味

を持たない。したがって、イントラ符号化モードおよびインタ符号化モードの分布が期待通りとなるような λ を、符号化に先立って予測することによる解決は困難である。

以上より、類似領域の拡大抑制の観点から、従来のR-D最適化法の枠組みでは不十分である。したがって、上記のケースにおいて高い主観画質が得られる符号化モード選択手法の枠組みが求められる。

(2) エッジ成分の再現性低下

低ビットレート符号化でのR-D最適化法では、式(2)に示すコスト関数において符号量に大きな重み付けがされるため、符号量の少ない符号化モードが優先して選択される。これに起因する問題として、平坦領域の近傍にオブジェクトが存在する場合、平坦領域に接するオブジェクトのエッジ成分が欠落する恐れがある。以下では、低ビットレート符号化に特化して、同エッジ成分を含むMBにおける符号化モード選択の処理を解析する。

低ビットレート符号化では、R-D最適化法のコスト関数において、符号量に大きな重み付けが与えられる。したがって、符号化歪みの大小が符号化モード選択結果に与える影響は小さい。ここで、エッジ成分を含むMBにおける符号化モード選択を考える。符号化によりエッジ成分が欠落した場合、同ブロックの符号化結果では平坦なテクスチャとなり、符号化歪みの評価値は、符号化モード候補の中では最大の値となる。しかしながら、同領域における原信号の輝度値変化は小さいため、エッジ成分が欠落した場合でも、符号化歪みの評価値は、重み付けされた符号量の評価値に対して十分に小さい。結果、同領域の符号化におけるR-D最適化法に基づく符号化モード選択では、符号量最小化規範での選択と等価となり、エッジ成分の再現性が低下する。特に、エッジ成分に動きを伴う場合のインタ符号化モードでは、動きベクトル等のサイド情報に費やす符号量が多くなるため、重み付けされた符号量の評価値が増大し、符号量重視の傾向が顕著となる。

次に、符号量が少ない符号化モードが選択されたときの符号化結果について解析する。イントラ符号化に関しては、MBの分割ブロックサイズが大きいほど、サイド情報が少ないため符号化に要する符号量は少ない。インタ符号化に関しては、サイド情報(特に動きベクトル)の符号量が少ない符号化モードほど、符号化に要する符号量は少ない。したがって、エッジ成分を含むMBでは、分割ブロックサイズが大きいイントラ符号化モード、もしくはサイド情報が少ないインタ符号化モードが優先して選択される。同様の符号化モードは、他の符号化モードよりも、予測精度が低く符号化歪みが大きい。したがって、同領域における符号化モード選択では、符号化モード候補の中で符号化歪みが大きい符号化モードが優先されることとなる。

以上より、イントラ符号化モード、インタ符号化モードどちらについても、結果的に、エッジ成分が欠落する恐れ

表1 符号化条件
Coding conditions.

設定項目	設定値
プロファイル	HighProfile
レベル	4.0
符号化方式	CABAC
動き予測	UMHexagonSearch
GOP 構造	IBBPBBPBBPBBPBB
参照枚数	2
動きベクトル探索範囲	$\pm 64 \times \pm 64$
符号化構造	MBAFF
DCT ブロックサイズ	4×4 または 8×8

がある符号化モードが選択される。したがって、従来手法に基づく符号化モード選択では、いずれの符号化モードによる予測であってもエッジ成分の欠落を確実に避けることができず、エッジ成分の再現性低下により主観画質の低下を招く危険性が高い。

上記を実画像上で確認する。ITE HDTV 標準動画像シーケンス”Yachting”を対象として、JM 手法を用いて符号化を行った。なお当該映像は、ヨットの動きに追従した単調なパニングを含んでいる。図 4 に 1 フレーム分の画像を示しており、以下では同図中の枠で示す領域を対象とする。

図 5(a) は、図 4 の枠で示す領域であり、符号化によりエッジ成分欠落が発生する領域である。特に、同図中の点線枠においてエッジ成分の欠落が顕著であり、主観画質の低下を招いている。また、図 5(b) は、図 5(a) で示した領域に対する符号化結果について、MB ごとの符号化歪みの大きさを濃淡で塗り分けている。同図下段は、濃淡と符号化歪みの大きさの対応を示している。同図より、エッジ成分の欠落が顕著な領域の MB では、他の MB に対して、符号化歪みが大きい傾向にあることがわかる。なお、同領域に対して符号化歪み最小化規範の符号化モード選択制御を適用することにより、エッジ成分の欠落に起因する主観画質低下が抑制されることを確認した。

文献 9) では、医療映像を対象とした H.264 符号化映像の主観画質改善を目的として、映像特徴に応じて、R-D 最適化法におけるラグランジュ乗数 λ を制御する手法が提案されている。しかしながら、同手法では、MB 単位での判定であるため、MB を跨るエッジ成分を確実に保持することは困難である。また、同手法では、医療映像におけるほぼ静止した映像におけるエッジの厳密な再現性を追及している。一方、本研究では、自然画像における輝度レベルが互いに類似しているオブジェクト境界に注目し、同境界が時間的に移動する際に発生する主観画質の低下を避けることを目的としている。したがって、両手法では、注目するエッジ成分の性質がそれぞれ異なるため、文献 9) の手法における λ の設定値では、図 5(b) で示した問題に対して不十分である。実際、図 5(b) と同条件下で同手法を適用した場合でも、同問題が完全に解決しないことを確認している。

以上より、平坦領域における動オブジェクトのエッジ成分を含む MB についての符号化モード選択に関して、低ビットレート符号化でのエッジ成分保存の観点から、従来の R-D 最適化法の枠組みでは不十分であり、同領域において高い主観画質が得られる符号化モード選択手法が求められる。また、エッジ成分を含む MB の近傍 MB においてもエッジ成分の一部が含まれる場合がある。同 MB に含まれるエッジ成分はエッジとしての特徴は弱いものの、同エッジ成分の欠落は近傍のエッジ成分全体としての再現性を損ね、主観画質の低下を招く。したがって、複数 MB に跨るエッジ成分を保持する符号化モード選択手法が求められる。

なお、テクスチャ再現性が低下する問題に対する高画質

化要素として、画面内での QP 制御により当該領域における MB の符号量を優遇することで解決する手法などが考えられる。しかしながら、本研究では、H.264 に特有の符号化モード選択のみに着目し、他の高画質化要素に直接的な制約を与えないことで、これらの手法との併用が可能であることを前提とする。

3. 提 案 手 法

2 章では、R-D 最適化法により高い符号化効率が達成される一方で、同手法により主観画質が著しく低下するケースがあることを示した。本研究で提案する符号化モード選択手法は、テクスチャ再現性の低下を抑止する規範、および R-D 最適化規範を併せ持つものである。テクスチャ再現性の低下抑制に関しては、2 章で示した問題を解決するた

図 4 評価映像
Test sequence.

(a) 原画像
(a) Original image

(b) 符号化歪み
(b) Distortion

図 5 予備実験結果 (エッジ成分を含む領域での符号化歪み)
Pre-experimental results. (Coding distortion at the area including edge component)

めに、動領域に含まれ、かつ、以下のいずれかまたは両方に該当するケースについて、提案手法に基づく符号化モード選択を適用する。なお、提案手法で対象とするピクチャタイプは P ピクチャおよび B ピクチャであり、I ピクチャの符号化処理については、従来手法と等価な処理とする。

- 当該 MB の近傍でイントラ符号化が集中することにより類似領域が拡大する場合
- 平坦領域におけるオブジェクトに関するエッジ成分の再現性が低下する場合

最初に、処理フレームにおける局所的な動きの大きさに基づき、処理 MB が動領域に属するかを判定する。処理 MB および m 個の近傍 MB を検出対象領域とし、同領域に属する各 MB について、隣接するフレームの参照により動きベクトルを求める。ここで、検出対象領域において大きさが最大の動きベクトルを、同領域における mv_{local} とする。 $|mv_{local}|$ が閾値 T_{mv} より大きいとき、当該 MB は動領域に属すると判定する。

動領域に属する場合、以下の 3.1 および 3.2 の適応制御を行う。なお、動領域に属さない、もしくは 3.1 および 3.2 がともに適用されなかった場合は、従来の符号化モード選択と等価な制御となる。図 6 に提案手法における符号化モード選択の処理フローを示す。各パラメータの詳細については、以下の 3.1 および 3.2 で述べる。

3.1 イントラ符号化モードの制御

(a) 類似領域拡大の評価

原画像の当該 MB における相違度および当該 MB をイ

図 6 提案手法のフロー
Diagram of proposed scheme.

ントラ符号化した結果の相違度から、近傍領域における類似性の増大を評価し、類似性の高い領域が拡大する危険性を判定する。類似性増大については、式 (3) で定義される相違度について原画像に対する減少量が大きいほど類似性の高い領域が拡大する危険性があることを表している。したがって、類似性増大の評価尺度として、原画像とイントラ符号化画像における相違度に関する差分を用いる。

ここで、式 (3) より原画像の相違度 s_{org} およびイントラ符号化モードの相違度 s_{intra} を求める。なお、 s_{intra} は、当該 MB をイントラ符号化した場合の符号化映像における相違度を表しており、 T の輝度値として局所復号された輝度値を用い、 C の輝度値として当該 MB にイントラ符号化モードを適用した場合の局所復号された輝度値を用いるものとする。 s_{org} と s_{intra} の差分の絶対値を D_s とし、 D_s が閾値 T_s よりも大きいとき、イントラ符号化モードの局所的な集中によるテクスチャ単調化が起こると判断する。

(b) 符号化モード制御

3.1(a) において、イントラ符号化モードの局所的な集中によりテクスチャ単調化が発生すると判断された場合、当該 MB における符号化モードの候補からイントラ符号化モードを除外する。

3.2 符号化モード選択基準の変更

(a) エッジ成分再現性低下の評価

符号化モード選択基準を決定する MB(以下、当該 MB) にエッジ成分が含まれるとき、当該 MB ではエッジ成分再現性の低下を抑止する符号化モードを選択することが求められる。提案手法では、処理フレームの符号化に際して、MB 単位の符号化に先立つ前処理として、各 MB における符号化モード選択基準を決定する。図 7 に、提案手法における符号化モード選択基準の決定処理フローを示す。

当該 MB における符号化モード選択基準の変更は、当該 MB および近傍領域における入力映像の輝度値の分散から、平坦領域におけるオブジェクトのエッジ成分の再現性が低下する危険性を判定することにより行う。当該 MB に対して n 個の近傍 MB を判定領域として指定する。ここで、入力映像の輝度値について、判定領域における分散 e_n 、および当該 MB における分散 e_c を求める。分散 e_n が閾値 T_{en} 以下であり、かつ分散 e_c が閾値 T_{ec} 以上であるとき、エッジ成分の再現性低下が発生する危険性があると判断する。

上記判定をラスタスキャン順にすべての MB に施した後、エッジ成分の再現性低下が発生する危険性があると判断された MB について、同 MB の判定領域に属する近傍 n 個の MB についてエッジ成分の再現性が低下する危険性を判定する。具体的には、各近傍 MB における輝度値の分散 e_b が閾値 T_{eb} を下回るとき、エッジ成分の再現性が低下する危険性があると判定する。

(b) 符号化モード制御

3.2(a) においてエッジ成分の再現性が低下する危険性が

あると判断された MB では、エッジ成分再現性の低下を抑止するために、符号化に起因する符号化歪みを抑制する符号化モード選択制御を行う。ここで、符号化歪みに類似した指標として、予測誤差が挙げられる。本論文で対象とする低ビットレート符号化の条件下では、符号化歪みおよび予測誤差について、何れを最小化する規範で符号化モードを選択しても、選択結果はほぼ等価であることを予備実験にて確認している。また、符号化歪みは、予測誤差に対して DCT 処理および量子化処理を施すことにより得られるものであるため、算出するために多くの処理コストが求められる。

以下では、エッジ成分再現性の低下を抑止する符号化モード選択として、予測誤差を最小化する規範の符号化モードを選択する。すなわち、3.2(a)において、エッジ成分の再現性が低下する危険性があると判断された MB では、予測誤差を最小化する規範の符号化モードを選択する。

なお、動きベクトル決定の際に、R-D 最適化法に基づくコスト評価が行われるが、同決定は符号化モード選択結果に

影響を与えるものではないため、従来と同様の処理とした。

提案手法は、3.1 および 3.2 の処理により局所的な主観画質低下を評価することが可能であり、同評価に基づく符号化モード選択を行うことで、類似領域の拡大抑制およびエッジ成分の欠落防止が可能である。すなわち、本手法は R-D 最適化法に内在する低ビットレート符号化における主観画質低下に関する短所を補うものであり、本手法により主観画質の低下を抑制しつつ、高い符号化効率を実現する符号化モード選択が可能である。

4. 評価実験

提案手法について主観評価実験を行った。エンコードとして JM12.1 をベースとし、3 章に基づく手法を実装した。評価用映像として、HDTV 標準動画像シーケンス "Woman with Bird Cage", "Yachting", "Rustling Leaves", "Soccer action", "Bronze with credits" を用いた。各素材とも 300 フレーム分を選択して使用した。なお、各素材とも、フレームサイズは 1920 画素 × 1080 ラインである。各素材の映像特徴を表 2 に示す。符号化レートは 8Mbps, 12Mbps であり、符号化条件は表 1 に従う。なお、表 3 に、提案手法におけるパラメータを示す。同パラメータ値は、複数素材を対象として、実験的に得られた値である。また、符号化レート制御手法として、JM12.1 に実装されている手法を用いた。また、3.1(a) の評価領域に関しては、 $m = 4$ とし、処理 MB の左、左上、上、右上にそれぞれ接する MB で構成される領域とした。3.2(a) の評価領域に関しては、 $n = 8$ とし、処理 MB の周囲に接する MB で構成される領域とした。

提案手法による符号化映像について主観評価実験を行った結果を表 4 に示す。比較のため従来方式として JM12.1 による結果も同表に併せて示す。また、各結果についての 95%

図 7 提案手法のフロー (3.2(a))
Diagram of proposed scheme (3.2(a)).

表 2 素材の特徴
Feature of test sequences.

	素材名	特徴
(A)	Woman with Bird Cage	鳥かごを持った人に追随したカメラワーク (カメラパン)。単調な動き。
(B)	Yachting	ヨットの動きに追随したカメラワーク (カメラパン)。単調な動き。
(C)	Rustling Leaves	固定カメラ映像。葉が細かく動く。
(D)	Soccer action	サッカー選手の動きに追随したカメラワーク (カメラパン)。高速な動きが含まれる。観客の描画に関して詳細なテクスチャが含まれる。
(E)	Bronze with credits	緩やかなカメラワーク (カメラパンおよびズームアウト)。複数行に渡る字幕が流れる。

表 3 提案手法におけるパラメータ
Parameters of the proposed method.

パラメータ	設定値
Tmv	4
Ts	25600
Ten	40
Tec	64
Teb	40

信頼区間を表5に示す。また、表6は主観評価実験における測定条件を示しており、ITU-R BT.500¹⁰⁾に準拠した。原画像に対する符号化映像の劣化尺度として、DSCQS(Double Stimulus Continuous Quality Scale)を用いた。同評価は、原画像および対象映像の主観評価値をそれぞれ0から100で表現し、原画像の評価値に対する対象映像の評価値の差分を取ることで求められている。すなわち、値が小さいほど原画像からの劣化が小さいことを表している。結果より、素材(C)、素材(D)および素材(E)において、従来手法に対して有意な主観画質改善が得られている。特に、素材(D)において、従来手法に対してDSCQSが最大8.6ポイント程度改善している。以上より、本研究で想定する10Mbps前後の低レート符号化において、従来手法に対して同等以上の主観画質が得られることがわかる。

なお、ビットレートとDSCQSの推移について、素材(A)の提案手法、素材(B)の従来手法および提案手法において、8Mbpsの評価に対して12Mbpsの評価が低いが、信頼区間を考慮すると有意な差ではないことが分かる。

提案手法による主観画質の改善について効果分析を行う。素材(B)および素材(C)では、従来手法においてエッジ再現性の低下が発生する。図8は、素材(C)の1フレーム分

の画像である。素材(B)については図4の枠で示す領域、素材(C)については図8の枠で示す領域において、特に同問題が著しい。従来手法では同領域において、エッジ成分近傍の平坦領域におけるMBの符号化歪みが、近傍の平坦領域における符号化歪みより大きい。図9(a)および図10(a)は同領域における、原画像と符号化映像(JM手法)との差分2乗和をMBごとに集計した結果を濃淡で塗り分けており、上記傾向を読み取ることができる。すなわち、平坦領域におけるエッジ成分近傍において、符号化歪みが増大することでエッジ再現性が低下する危険性があることがわかる。

一方、提案手法によると、図4および図8の枠で示す領域において、3.2(b)の制御を適用し、予測誤差を最小化する規範の符号化モード選択制御を行うことで、符号化歪みを抑制できる。図9(b)および図10(b)は同領域における、原画像と提案手法による符号化映像との差分2乗和をMBごとに集計した結果を濃淡で塗り分けている。なお、提案手法では、同領域内のMBのうち52.7%程度のMBにおいて3.2(b)の制御が適用されている。以上より、提案手法により輪郭を含むMBにおいて予測誤差を最小化規範とする符号化モード選択を行うことで、同領域における符号化歪みが抑制され、特に同図中の枠で示す領域においてエッジ成分を安定的に保持していることが読み取れる。

素材(D)では、従来手法において動きの質感低下が発生し、特に図1の枠で示す領域付近において著しい。従来手法では、同領域における相違度が原画像よりも著しく小さく、類似領域が拡大する。図11(a)は、同領域における式(3)で求められるMB単位で相違度を濃淡で塗り分けており、上記傾向を読み取ることができる。すなわち、同領域において類似性が高まることで、動きの質感が低下する危険性があることがわかる。

一方、提案手法によると、図1の枠で示す領域付近では、3.1(b)の制御が適用されることによりインタフレームにおけるイントラ符号化モードの適用が制限される。2章より、イントラ符号化モードはイントラ符号化モードに対して類似性が低いことから、同制御により同領域における類似領域の拡大が抑制される。図11(b)は同領域におけるMB単位

表4 DSCQSの結果
Results of DSCQS.

素材	手法	DSCQS(%)	
		8Mbps	12Mbps
(A)	従来 (JM12.1)	4.37	-1.11
	提案手法	2.63	2.74
(B)	従来 (JM12.1)	3.64	4.50
	提案手法	2.79	4.29
(C)	従来 (JM12.1)	19.82	19.64
	提案手法	18.25	15.79
(D)	従来 (JM12.1)	23.89	22.32
	提案手法	19.42	13.74
(E)	従来 (JM12.1)	15.32	9.73
	提案手法	14.79	6.21

表5 信頼区間
Confidence interval.

素材	手法	Confidence interval	
		8Mbps	12Mbps
(A)	従来 (JM12.1)	±3.56	±3.82
	提案手法	±4.36	±3.16
(B)	従来 (JM12.1)	±4.12	±3.18
	提案手法	±3.16	±2.24
(C)	従来 (JM12.1)	±5.68	±4.39
	提案手法	±4.52	±3.79
(D)	従来 (JM12.1)	±4.57	±3.79
	提案手法	±3.60	±3.37
(E)	従来 (JM12.1)	±4.04	±3.32
	提案手法	±3.27	±2.06

表6 測定条件
Evaluating conditions.

視距離	3H
ディスプレイ	32 インチ CRT
被験者	非専門家 19 名
測定手法	DSCQS

図8 "Rustling Leaves" 原画像
"Rustling Leaves" original image.

の相違度を濃淡で塗り分けている．なお，提案手法では，同領域内の MB のうち 72.4% 程度の MB において 3.1(b) の制御が適用されており，同領域内における MB 単位の相違度の平均値が 13.3% 程度改善している．以上より，提案手法により同領域においてイントラ符号化モードを制限して類似領域の拡大を抑制することで，動きの質感を安定的に保持していることが読取れる．

提案手法における 3.2(b) の制御では，エッジ成分の有無を検知するための手法として，輝度値の分散を用いている．本手法では，エッジ成分の厳密な検知までは必要ないため，分散による簡易な検知でも有意であり，素材 (B) および素材 (C) に対して，主観画質改善効果があることからその有意性を確認できる．しかしながら，素材 (A) の 12Mbps において，信頼区間を考慮すると必ずしも有意ではないものの，従来手法に対して若干の主観画質低下がある．原因として 3.2(b) の制御が過剰に適用されたために符号化効率が低下したことが挙げられる．特に，カメラワークを含むシーンにおける背景に同制御が過剰に適用されており，同

領域に対して過度の符号量が割当てられる結果となった．同領域は 2 章で示したエッジ再現性が求められる領域ではなく，画質の改善には至らない．このため，3.2(b) の制御の

(a) 従来手法
(a) Conventional
(b) 提案手法
(b) Proposed scheme

(c) 原画像
(c) Original image

図 10 MB 毎の符号化歪み (素材 (C))
Coding distortion (Sequence (C)).

(a) 従来手法
(a) Conventional

(b) 提案手法
(b) Proposed scheme

(c) 原画像
(c) Original image

図 9 MB 毎の符号化歪み (素材 (B))
Coding distortion (Sequence (B)).

(a) 従来手法
(a) Conventional
(b) 提案手法
(b) Proposed scheme

(c) 原画像の相違度
(c) Difference of original image

図 11 MB 毎の相違度 (素材 (D))
Difference (Sequence (D)).

対象外とすることが求められる。制御が過剰に適用されないよう、適用条件を最適化することが今後の課題である。

本研究では、符号量制御に関して、画面単位でのQP更新を前提とした。しかしながら、映像配信サービスなどにおける実用のエンコーダでは、高画質化を目的として、画面内(MB単位)でのQP更新制御も広く使われている。提案手法とMB単位でのQP更新制御を併用した場合について、JM12.1に実装されている手法に基づき、考察する。MB単位のQP更新では、ピクチャの目標符号量を定めた上で、各MBにおいて目標符号量を設定しQPを導出する。一例として、図12に、素材(B)について、あるフレームにおけるピクチャ内未符号化MBの目標符号量推移を示す。図中、横軸は符号化MB位置を指し、矢印は、提案手法が有効に働いた箇所を示している。同図より、提案手法が有効となるMB位置において、過度に符号量を消費する傾向は見られず、後続の符号化に与える影響は小さいといえる。これは、同MB位置において、従来手法と比べた発生符号量の差異が少ないことによるものである。しかしながら、同傾向は素材に依存するものであり、MB単位のQP更新を行う条件下では、ピクチャの目標符号量が不足しないようコスト関数における符号量の重みを適切に設定する仕組みが求められ、本件については今後の課題である。

5. む す び

本研究では、HDTVを対象としたH.264低ビットレート符号化において、主観画質を改善するモード選択手法を提案した。提案手法は、局所領域における主観画質低下要因を検知する新たな評価尺度を備え、これに基づいたモード選択を行うことで、テクスチャの再現性を追及したものである。主観評価実験より、提案手法の適用によって、従来のモード選択方式に対する主観画質の改善効果があることを確認した。

最後に、日頃ご指導いただく株式会社KDDI研究所代表取締役所長 秋葉重幸博士に感謝の意を表する。本研究は、独立行政法人情報通信研究機構による委託研究「ソフトウェア符号化技術に関する研究開発」として実施したものである。

図12 目標符号量
Target bit amount.

のである。

〔文 献〕

- 1) Karsten Suhring, H.264/AVC Software Coordination, <http://ip.hhi.de/index.htm>
- 2) Gary J. Sullivan, Thomas Wiegand, "Rate-Distortion Optimization for Video Compression", IEEE Signal Processing Magazine, pp.74-90 (Nov. 1998)
- 3) Chao-Hsuing Tseng, Hung-Ming Wang, and Jar-Ferr Yang, "Enhanced Intra-4 × 4 Mode Decision for H.264/AVC Coders", IEEE Trans. on CSVT2006, pp.1027-1032, 16, 8 (Aug. 2006)
- 4) Adriana Dumitras and Barry G. Haskell, "Enhancement of Direct Mode Selection in B Pictures for Bit Rate Reduction of Compressed Video Sequences", ICIP2003, III-825-8 (Sep. 2003)
- 5) Koohyar Minoo and Trung Q. Nguyen, "Perceptual Video Coding with H.264", Signal, Systems and Computers 2005, pp.741-745 (Oct. 2005)
- 6) Zhi-Yi Mai, Chun-Ling Yang, Kai-Zhi Kuang and Lai-Man Po, "A Novel Motion Estimation Method Based on Structural Similarity In H.264 Inter Prediction", ICASSP2006, II-913 - II-916 (May 2006)
- 7) Zhi-Yi Mai, Chun-Ling Yang, Sheng-Li Xie, "Improved Best Prediction Mode(s) Selection Methods Based on Structural Similarity In H.264 I-Frame Encoder", ICSMC2005, pp.2673-2678, 3 (Oct. 2005)
- 8) Z. Wang, A. C. Bovik, H. R. Sheikh, and E.P. Simoncelli, "Image quality assessment from error visibility to structural similarity", IEEE trans. Image Processing, 13, 4, pp. 600-612 (Apr. 2004)
- 9) H. Yu, Z. Lin, F. Pan, "Applications and Improvement of H.264 in Medical Video Compression", IEEE Trans. on Circuits and Systems, 52, 12 (2005)
- 10) ITU-R BT.500-11, "Methodology for the subjective assessment of the quality of television pictures" (2002)

吉野 とも のぶ
2001年、東京工業大学工学部情報工学科卒業。2003年、同大学大学院修士課程修了。同年4月KDDI(株)入社。2005年以来、(株)KDDI研究所にて、主に画像符号化方式に関する研究・開発に従事。正会員。

内藤 せい
1994年、早稲田大学理工学部電子通信工学科卒業。1996年、同大学大学院修士課程修了。同年、国際電信電話(株)に入社。現在、(株)KDDI研究所にて、主にHDTV、SDTVの動画像符号化および伝送方式の研究・開発に従事。正会員。

酒澤 茂之
1990年、神戸大学工学部電子工学科卒業。1992年、同大学院修士課程修了。同年、国際電信電話株式会社に入社、研究所勤務。現在、(株)KDDI研究所映像通信グループ主任研究員。動画像符号化、映像通信システム、コンテンツセキュリティ、自由視点映像の研究・開発に従事。博士(工学)。正会員。

小池 あつ し
1983年、新潟大学工学部情報工学科卒業。1985年、東北大学大学院修士課程修了。同年、国際電信電話(株)に入社。以来、テレビ会議・電話システム、動画像の高効率符号化方式の研究・開発に従事。現在、(株)KDDI研究所映像通信グループリーダー。博士(情報学)。正会員。