

映像シーンに適応したスケーラビリティ選択手法

D-11-19

Adaptive Selection of Scalability Method for Video Scene

高橋 良知 *1

石川 孝明 *2

渡辺 裕 *2

富永 英義 *1*2

Yoshitomo TAKAHASHI *1

Takaaki ISHIKAWA *2

Hiroshi WATANABE *2

Hideyoshi TOMINAGA *1*2

*1 早稲田大学 理工学部 電子・情報通信学科

*2 早稲田大学 国際情報通信研究センター

*1 Dept. of Elec. Info. and Comm. Eng., WASEDA Univ.

*2 Global Info. and Tele. Institute, WASEDA Univ.

1 はじめに

近年、ネットワーク帯域、および受信端末の解像度の多様化により、ワンソース・マルチユースの映像配信を実現するための技術として、映像のスケーラビリティ機能が注目されている。現在 MPEG では H.264/MPEG-4 AVC をベースとした SVC (Scalable Video Coding) の標準化が進められている [1]。SVC では、空間・時間・SNR スケーラビリティを提供することができる。しかし、一般的にスケーラブル符号化は非スケーラブル符号化に比べ符号化効率が低下することが知られている [2]。

そこで、本稿では、時間・SNR スケーラビリティに着目し、SVC におけるスケーラビリティの粒度と符号化効率の関係に基づき、SNR スケーラビリティを提供する FGS (Fine Granular Scalability) 機能を切り替えることによる主観品質の向上のための手法を検討する。

2 スケーラビリティと符号化効率

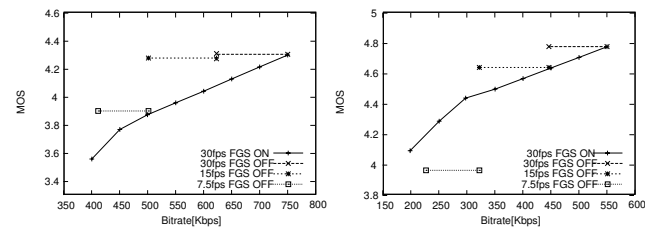
SVC は、時間スケーラビリティを実現するために MCTF (Motion Compensated Temporal Filtering) を用いる。MCTF は、時間方向のサブバンド分割を繰り返し実行することで、階層構造を構成し、時間スケーラビリティを実現している。MCTF では、GOP サイズに応じて、取得可能なフレームレートが変化する。一般に、 2^N ($N = 1, 2, 3, \dots$) の GOP サイズでは、 N 階層の時間解像度が得られる。動画の時間複雑度や動き補償予測の精度に応じて、フルデコードの際の平均 PSNR の最大値を与える GOP サイズが異なることが分かっている [3]。

また、SNR スケーラビリティの実現には FGS が用いられる。SVC は、Cyclic Block Coding と呼ばれる符号化を行うが、これは基本レイヤと拡張レイヤへ分割するために、MPEG-4 の Bitplane Coding と同様に、非 FGS 符号化と比較し、約 1~1.5[dB] 程の符号化効率の低下が見られる。

3 検討事項

FGS を用いたスケーラビリティ制御では、粒度の高い SNR スケーラビリティ制御が可能であるが、非 FGS に対し PSNR の低下が発生する。一方で、MCTF を用いたスケーラビリティ制御では、粒度の低い時間スケーラビリティ制御となるが、高い PSNR を維持することが可能である。そこで、動画の時間方向と SNR 方向の主観に対する重要性に基づき、これらのスケーラビリティを切り替えることで、主観品質の向上を検討する。

文献 [4] では、時間方向の主観品質を MOS_t 、空間方向の主観品質を MOS_s とし、それぞれの主観品質の推定式が提案され、これらの推定式に基づき、時間・SNR スケーラビリティの組み合わせ手法が検討されている。 MOS_t の推定の特徴量としては動きベクトルの最大値が用いられ、 MOS_s の推定の際の特徴量としては Sobel フィルタによるエッジの統計量が用いられている。本稿においても、同様の特徴量を用いているが、文献 [4] では、MPEG-4 をベースとした動き補償推定を基本とし



(a) FOOTBALL

(b) HARBOUR

図1 スケーラビリティの実現方法と MOS の関係

ているのに対し、本稿は H.264 を基本とする SVC 上で検討を行っているため、可変ブロックサイズや動き補償の参照フレーム間距離の変化などが生じる。そこで、今回は、可変ブロックについては、関連の強い動きベクトルを有するブロックのみを利用し、参照フレーム間の距離については、距離に応じた値で正規化を行う。

4 結果

図1は、標準動画 FOOTBALL, HARBOUR について、スケーラビリティを時間・SNR のいずれか一方により実現した場合の、推定 MOS を示している。なお、GOP サイズは、フルデコードの際の PSNR が最も高くなるサイズとしている。FOOTBALL などの時間複雑度の大きいシーケンスでは、フレームレートの低下が主観品質の低下に与える影響が小さく、かつ、低フレームレートにすることにより1フレームあたりに費やすビット量が増加する。このときの PSNR の向上が大きいため、時間スケーラビリティによる制御では、7.5[fps] においても FGS を用いた 30[fps] に比べて高い MOS が得られたと考えられる。

5 まとめと今後

スケーラビリティの粒度と符号化効率の関係に着目し、動画の特性に応じてスケーラビリティを時間・SNR のいずれか一方により実現することによる主観品質の向上を検討した。今後は、実際に主観評価実験を行い検討手法の有効性を確認するとともに、SNR・時間スケーラビリティの切り替え方法について検討を行う。

参考文献

- [1] J. Reichel et al., "Scalable Video Coding - Working Draft 4," Joint Video Team(JVT) of ISO/IEC MPEG & ITU-T VCEG 16th Meeting, Doc. JVT-Q201, Oct. 2005.
- [2] 木本, "MPEG におけるスケーラブルビデオ符号化の標準動向," 情処学, AVM-48, P-55.60, 2004.
- [3] 高橋, 石川, 後藤, 渡辺, "映像シーンに適応したスケーラビリティ選択手法に関する検討," PCSJ2005.
- [4] 宮田, 吉田, "要求された主観品質を実現する動画像符号化のレート制御法," PCSJ2005.