

超高精細画像の空間スケーラブル符号化時の特性評価

高木 幸一[†] 内藤 整[†] 松本 修一[†]

本稿では、HDTV より高い解像度を持つ超高精細画像に対し、空間スケーラブル符号化を適用した場合のレイヤ間予測の効果を評価したので報告する。空間スケーラビリティは低解像度画像の符号化結果をより高解像度の画像の符号化のために継承することで複数の解像度の画像を高効率に符号化する技術である。同技術を超高精細画像の画面内（イントラ）符号化に適用した場合、HDTV 解像度以下の画像に適用した場合と比較し、レイヤ間予測の効果が大きいことを確認した。さらに、マクロブロックのサイズを解像度にあわせて拡張することにより、高い効果が得られることを確認した。

An Evaluation of Spatial Scalable Coding for Ultra-high Resolution Video

KOICHI TAKAGI[†] SEI NAITO[†]
SHUICHI MATSUMOTO[†]

This paper presents the performance of inter-layer prediction in spatial scalable video coding which is applied to the ultra-high resolution video signal. The spatial scalable video coding (SSVC) can reduce the information bits of higher resolution video by predicting the image from lower resolution video. In case SSVC is applied to the ultra-high resolution video, inter-layer prediction is effective compared with the video whose resolution is less than HDTV. Moreover, better coding performance is obtained at lower bit-rate when MB size is extended.

1. はじめに

近年、IPTV やデジタル放送に代表される映像配信・放送サービスでは、HDTV クラスの解像度を持つコンテンツが主流となりつつある。さらに、将来的には、映像コ

ンテンツの高解像度化が進み、高臨場感が得られるサービスの実現が期待されており、その一つとしてスーパーハイビジョンに代表される走査線 4000 本/60fps クラスの超高精細映像(以下、8K 映像)が注目されている 1)。

このような超高精細映像を一般のコンテンツ配信網 (CDN) 等を通じて配信する際に、その配信環境として、超高精細映像そのものに対応したインフラ (伝送路、端末など)だけでなく、既存のデジタル放送回向けのインフラにも対応することが望ましい。

このように、伝送方式や表示環境が多様化する中で、単一の映像ソースを複数種類の環境に同時に配信したいという要求が存在し、それにこたえる形でスケーラブル符号化標準 SVC (Scalable Video Coding) が確立された 2)。SVC では空間/時間/SNR スケーラビリティなどをアプリケーションに応じて組み合わせることで、各環境に映像をパラレルに配信する場合と比較し、効率的に映像を配信することが可能となる。ところが、現行の SVC は、標準化プロセスにおいて、最大でも HDTV 解像度を対象として検討されたものであるため、上述の 8K 画像に対してこれを適用した時にどのような特性になるのかは確認されていない。

そこで、本稿では、現行のデジタル放送相当の走査線 1000 本クラスの映像 (以下 2K 解像度映像)をベースレイヤとし、高度 BS 放送方式に採用される予定の走査線 2000 本クラスの映像 (以下 4K 解像度映像)を中間レイヤとした上で、8K 解像度映像をこれらのレイヤにそれぞれスケーラブル符号化した場合の既存の SVC の特性を評価することを目的とする。特に本稿では、イントラピクチャにおいて、解像度が低い場合と比較し解像度が高い場合の空間スケーラブル符号化におけるレイヤ間予測の効果について検証する。

2. 空間スケーラブル符号化概要

SVC の符号化モデル JSVM における空間スケーラブル符号化のブロック図を図 1 に示す。図 1 は 3 レイヤの場合の例であるが、レイヤの数がそれ以外の場合も同様に符号化することができる。1 節の例で述べた、8K 解像度映像、およびそれを Down Convert (D/C)した 4K 解像度映像、2K 解像度映像の 3 レイヤをそれぞれ Layer 2、Layer 1、Layer 0 と呼ぶこととする。ここで、Layer 0 はすべての基本になる階層であるため Base Layer、それ以外の階層は Enhancement Layer と呼ばれる。

ところで、SVC において、下位レイヤがイントラ符号化されている場合、上位レイヤにおいて該当する場所 (マクロブロック (MB)) の符号化は、通常の 1) イントラ符号化、2) インター符号化以外に、3) 下位レイヤの複号画像を使って上位レイヤの画素値を予測し、その予測誤差を符号化する方法をとりうる。特にイントラピクチャでは 1 か 3 が選択されることになる。

[†](株)KDDI 研究所
KDDI R&D Labs. Inc.

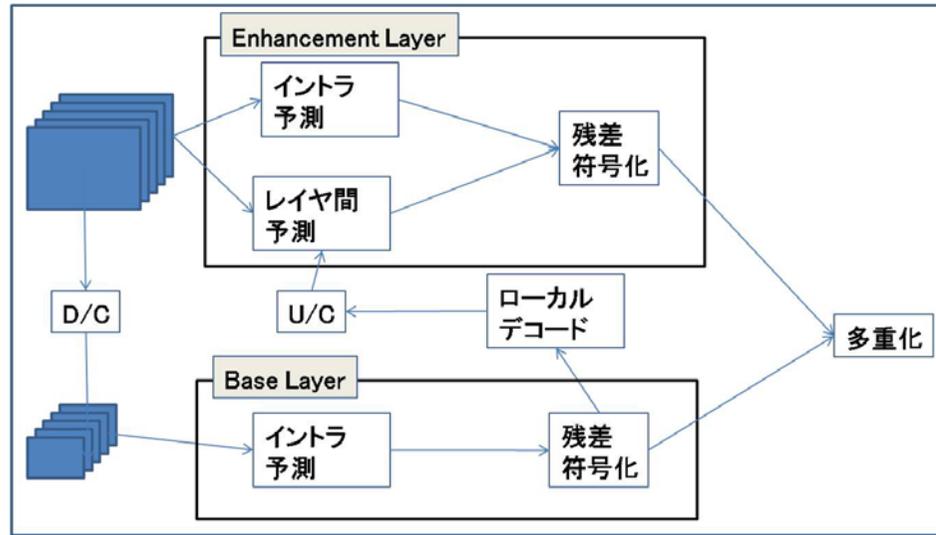


図1：空間スケラブルイントラ符号化ブロック図

表1 実験条件

項目	項目	内容
符号化	フレーム数, 構造	33, all intra
	エントロピー符号化	CABAC
	RD-opt	ON
素材	色空間	YUV4:2:0
	ビット深度	8bit
	素材	表2

表2 評価用素材の特徴

素材	特徴
A	静止カメラに桜の花と空が映っている映像. 全体的にほぼ静止.
B	紅葉した林を映した映像. カメラパンのカメラワークを伴う.
C	静止カメラに大量の向日葵が映る映像. 向日葵は風によりランダムに動く.
D	大量の蝶々が映る映像.

3. 解像度とレイヤ間予測選択率の関係

一般に、高精細になればなるほど、下位レイヤと上位レイヤの相関が高くなることが想定される。よって、今回対象とする超高精細画像に SVC を適用する場合、レイヤ間予測の選択率がより高くなることが予想される。よって、本節では、Base Layer がイントラ符号化されている際、レイヤ間予測が適用される割合について確認する。

具体的には、8K 解像度映像を水平・垂直それぞれ 1/4 に D/C した 2K 解像度映像を Base Layer としてイントラ符号化し、それを利用して 4K 解像度映像、8K 解像度映像をイントラ符号化する際にレイヤ内でイントラ予測を用いてイントラ符号化すべきか（“継承なし”）、2K 映像の複号画像を U/C したものをを用いるべきか（“継承あり”）を JSVM の RDO（Rate-Distortion Optimization）モードを用いて求めた。レイヤ間予測選択率を複数の QP、複数の素材に対して求めた結果を図2に示す。なお、全 Layer で同じ QP を使用している。また、それ以外の実験条件は表1の通りである。比較として、2K 解像度映像を最上位レイヤとして、同じ実験を行った結果をあわせて示す。

同図より、JSVM で用いられている RDO の観点で、一部を除き、解像度が高いほどレイヤ間予測が多く発生し、8K 解像度の場合、約9割はレイヤ間予測が使われていることがわかる。

また、QP の大きさとレイヤ間予測選択率の関係を見てみると、8K 解像度と 2K 解像度の間で QP が大きいほどレイヤ間予測選択率の差が小さくなることがわかる。これは、QP が大きくなるとレイヤ間予測に用いられる画像の精度が低くなるため、レイヤ間予測の当たりにくくなるためであると考えられる。

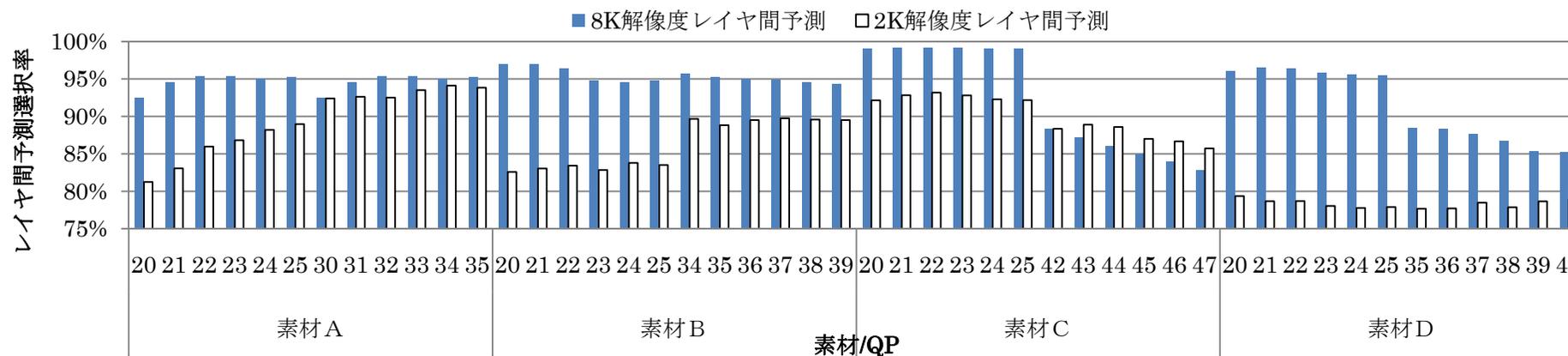


図2 レイヤ間予測モード選択率

4. MB サイズ拡張の適用

4.1 MB サイズ拡張の可能性

本稿で対象としている高解像度映像は、前節でも述べた通り空間解像度が高いことから画素間相関が非常に高い。特に、今回対象とする 8K 解像度映像では、2K 解像度以下の映像と比較すると、隣接する MB (16 画素×16 ライン (以下「16×16」と書く)) 間の相関も非常に高くなることが予想される。さらに、文献 2) に述べられているとおり、特に低ビットレートになればなるほど、制御用情報に要する符号量の割合が高くなるため、同一の性質をもつ領域はできる限り一括して扱えることが望ましい。よって、MB サイズとして既存の SVC で用いられている 16 画素×16 ラインより大きなものを適用することにより、制御用情報の大幅な削減が期待できる。

4.2 MB サイズ拡張方式の概要

MB サイズの拡張方式として、本稿では 16×16 の倍のサイズである 32 画素×32 ライン (以下「32×32」と書く) を適用し、その効果を検証することとする。簡単なた

め、MB サイズはレイヤごとに固定とする。

32×32 の MB における処理は、16×16 の MB における処理が 2 倍にスケールアップされたと考えればよい。例えば、以下の通りである。

- ・ 直交変換ブロックサイズについては、文献 2) 同様、低周波へのエネルギー集中性を考慮し、本稿では 8 画素×8 ライン を採用する。
- ・ 輝度のイントラ予測は 8 画素×8 ライン単位もしくは 32 画素×32 ライン単位で行う。

また、MB サイズをレイヤごとに固定とするとその組み合わせが何通りか考えられる。

Base Layer で 32×32 が、Enhancement Layer で 32×32 が用いられる場合は、基本的に既存の SVC で用いられている MB サイズを 2 倍にスケールアップしただけであるため既存の SVC 同様にレイヤ間予測を行うことができる。

一方、Base Layer で 16×16 が、Enhancement Layer で 32×32 が用いられる場合も、Base Layer の 1MB がちょうど Enhancement Layer の 1MB に相当するため、レイヤ間予測は MB ごとに 1 : 1 で行うことができる。

よって、どちらの場合も問題なくレイヤ間予測を構成することが可能である。

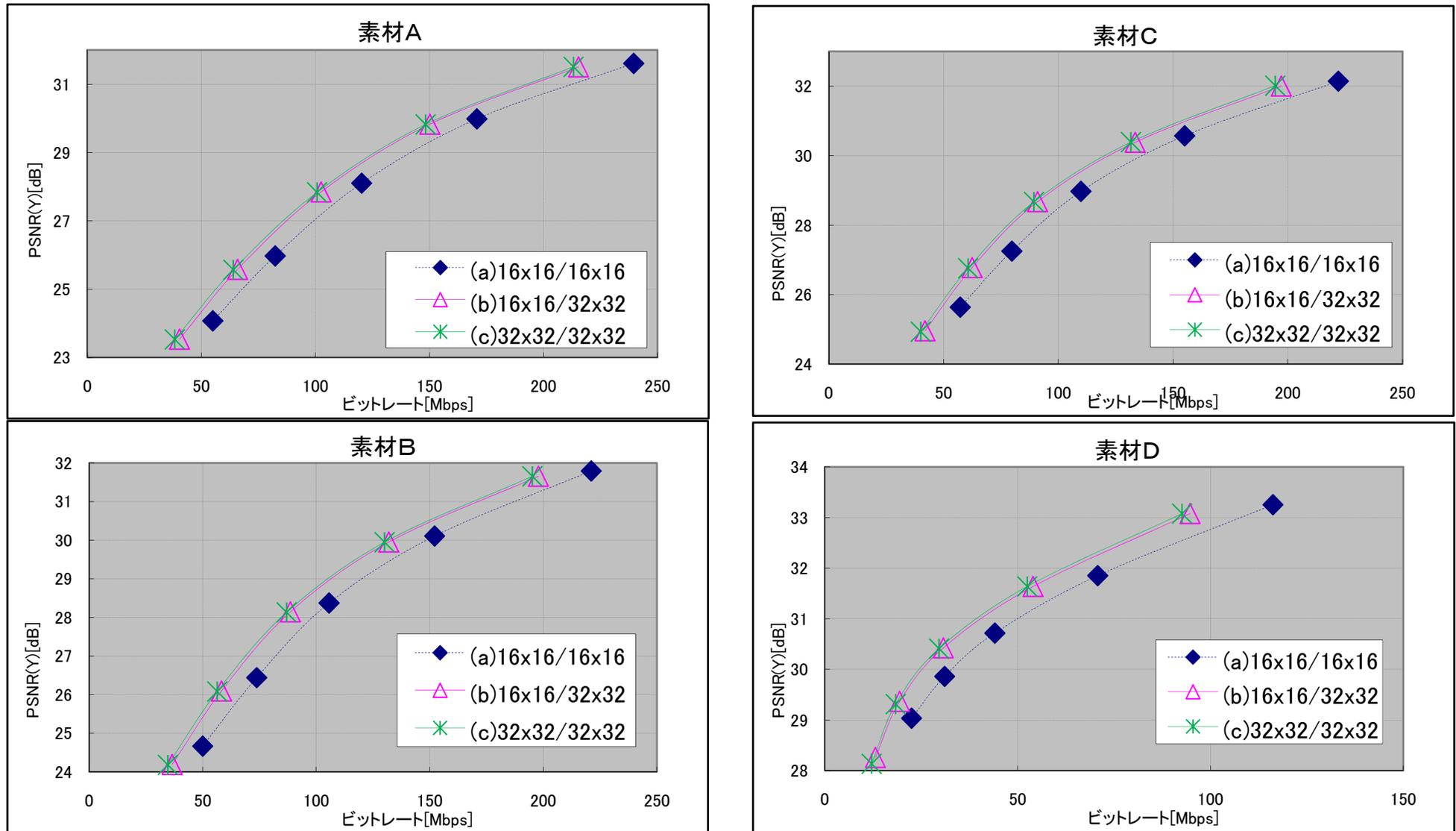


図3 拡張MB適用時のR-D特性

4.3 実験

符号化実験では、提案手法の局所的な効果を解析するため、8K 映像から 3840 画素×2160 ライン(4K 映像)の領域を切り出し、4K 映像を 1/2 に D/C した 2K 映像を Base Layer, 4K 映像を Enhancement Layer としてスケーラブル符号化を行った。MB サイズの組み合わせは (Base Layer, Enhancement Layer) = {(a)(16×16, 16×16), (b)(16×16, 32×32), (c)(32×32, 32×32)}の 3 通りとした (c)は既存の SVC に相当)。それ以外の条件は表 1 のとおりである。全素材にこれらを適用した結果を図 3 に示す。

同図より、MB サイズ拡張が適用された(b), (c)は既存方式(a)と比較しどの素材でも符号化効率が向上していることが確認できる。さらに、(b)と(c)を比較すると、若干ではあるが、(c)の方がより符号化効率が向上できていることが確認できる。

これらをビットレート削減率 BDBR, および PSNR の情報 BDPSNR4)で換算すると、表 2 のようになり、平均して効果が高いことが確認できる (QP は 36, 40, 44, 48 の 4 点を使用)。参考のために、単一レイヤ (上記 4K 映像) に MB サイズ拡張を適用した場合の効果をあわせて示す。この結果から、どの素材も単一レイヤでの効果も認められるが、SVC に適用することでさらに高い効果が得られていることがわかる。

表 2 SVC での MB サイズ拡張による BDBR, BDPSNR

	(b)(16×16, 32×32) v.s. (a)SVC		(c)(32×32, 32×32) v.s. (a)SVC	
	BDBR[%]	BDPSNR[dB]	BDBR[%]	BDPSNR[dB]
素材 A	▲12.3414	0.6172	▲14.3605	0.7114
素材 B	▲13.8273	0.6453	▲15.9629	0.7372
素材 C	▲12.8927	0.6176	▲14.8251	0.7059
素材 D	▲21.5898	0.5940	▲24.1318	0.6778

表 3 単一 Layer での MB サイズ拡張による BDBR, BDPSNR (参考)

	32×32 v.s. 16×16	
	BDBR[%]	BDPSNR[dB]
素材 A	▲7.0894	0.2956
素材 B	▲8.5280	0.3759
素材 C	▲6.5299	0.3532
素材 D	▲11.8710	0.5220

5. おわりに

本稿では、超高精細画像に対し、イントラ符号化時に空間スケーラブル符号化を適用した場合のレイヤ間予測の効果を評価した。超高精細画像に適用した場合、HDTV 解像度以下の画像に適用した場合と比較し、レイヤ間予測の効果が大きいことを確認した。さらに、マクロブロックのサイズを解像度にあわせて大きくすることにより、冗長となる制御情報を削減することができ、高い効果が得られることを確認した。

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「超高精細映像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。

参考文献

- 1) 菅原正幸, “走査線 4000 本級超高精細映像システム,” 映像学誌, 57,11, pp.1442-1444(2003).
- 2) H. Schwarz, D. Marpe, T. Wiegand, "Overview of the Scalable Video Coding Extension of the H.264/AVC Standard," IEEE Trans. CSVT, Vol. 17, Issue 9, pp. 1103-1120 (2007).
- 3) 吉野知伸, 内藤整, 酒澤茂之, “超高精細映像符号化におけるマクロブロックサイズ拡張手法に関する一検討,” 信学技報, vol. 108, no. 485, IE2008-256, pp. 19-22 (2009).
- 4) Gisle Bjontegaard, “Improvements of the BD-PSNR model,” ITU-T SG16/Q6, 34th VCEG Meeting, VCEG-A111, Berlin, Germanu. (2008).