

地上デジタル放送方式高度化に関わる
適用技術検討作業 最終報告

VVC 主観評価実験報告

2022年2月18日

デジタル放送システム開発部会／映像符号化方式作業班
一般社団法人 電波産業会

まえがき

総務省からの諮問第 2044 号「放送システムに関する技術的条件」(2019 年 6 月 18 日)を受け、情報通信審議会放送システム委員会に地上デジタル放送方式高度化作業班が設置され、技術的条件の検討が始まった。本活動の一環として、2020 年 6 月 22 日に、ARIB に対して映像符号化及び音声符号化方式の高度化に必要な技術的検討の依頼があった。

本依頼を受け、映像符号化方式作業班は、映像符号化方式の検討を開始している。2021 年 1 月には、VVC 規格を用いた際の所要ビットレートを求めるための主観評価実験の計画案を作成し、地上デジタル放送方式高度化作業班にて報告した。

本報告書は、上記計画案に従い実施した VVC 規格の主観評価実験の結果報告である。2021 年 10 月の中間報告では、所要ビットレートを報告した。本最終報告書では、特定のビットレートでの画質に関する考察を追加した。

内容

1. 目的.....	6
2. 実験計画.....	6
2.1. 実験会場、日時.....	6
2.2. 評価映像.....	6
2.3. 符号化条件.....	8
2.4. 評価実験方法.....	9
2.5. 所要ビットレート推定の基準.....	9
3. 符号化結果.....	10
3.1. エンコーダ A.....	10
3.2. エンコーダ B.....	11
4. 実験結果.....	13
4.1. 評価者のスクリーニング.....	13
4.2. ビットレートとMOS値との関係.....	15
4.2.1. 1080/60/P, SDR.....	15
4.2.2. 2160/60/P, SDR.....	16
4.2.3. 2160/60/P, HDR.....	16
5. 所要ビットレート.....	17
6. 考察.....	18
6.1. エンコーダ A の妥当性.....	18
6.2. 所要ビットレート推定(2160/60/P, SDR).....	19
6.3. VVC 規格の公称性能と所要ビットレートとの関係.....	19
6.4. 主観画質改善の可能性.....	20
6.4.1. 1080/60/P, SDR.....	20
6.4.2. 2160/60/P, SDR.....	20
7. 追加画質評価.....	21
7.1. 目的.....	21
7.2. 実験計画.....	21
7.2.1. 実験会場、日時.....	21
7.2.2. 評価映像.....	21
7.2.3. 符号化条件.....	21

7.2.4.	符号化映像.....	22
7.3.	評価方法.....	22
7.3.1.	客観画質.....	22
7.3.2.	主観画質.....	22
7.3.3.	評価基準及び提示方法.....	23
7.3.3.1.	評価 1: ビットレート毎の画質評価.....	23
7.3.3.2.	評価 2: 4K 放送品質映像との画質比較.....	23
7.4.	結果.....	24
7.4.1.	客観画質.....	24
7.4.2.	評価者のスクリーニング.....	28
7.4.3.	符号化難易度が高い映像のビットレートと MOS 値との関係.....	28
7.4.4.	画質改善手法の効果.....	29
7.4.5.	HEVC 30 Mbps 符号化映像との比較.....	32
7.4.6.	ヒアリング結果.....	33
7.5.	考察.....	35
7.5.1.	各ビットレートでの画質.....	35
7.5.2.	画質改善手法の効果.....	35
8.	付録.....	36
8.1.	評価映像の選定.....	36
8.1.1.	選定基準.....	36
8.1.2.	符号化難易度の導出方法.....	36
8.1.3.	符号化難易度分布と選定した評価映像.....	36
8.1.3.1.	1080/60/P, SDR.....	36
8.1.3.2.	2160/60/P, SDR.....	37
8.1.3.3.	2160/60/P, HDR.....	37
8.2.	前処理及び符号化解像度変更で適用したフィルタ.....	37
8.2.1.	3/4 縮小フィルタ.....	37
8.2.2.	1/2 縮小フィルタ.....	38
8.2.3.	4/3 拡大フィルタ.....	38
8.2.4.	2/1 拡大フィルタ.....	39

1. 目的

VVC 規格を適用した地上波デジタル放送の運用ガイドラインや技術基準の策定に寄与することを目的とし、UHDTV 及び HDTV の VVC 符号化映像の主観評価実験を実施して、所要ビットレートを明らかにする。所要ビットレートとは、放送されるほぼ全ての映像で一定水準以上の画質を確保可能なビットレートである。

2. 実験計画

2.1. 実験会場、日時

日本放送協会放送技術研究所(世田谷区砧)

2021年5月17日～28日

2.2. 評価映像

表 2-1 に示す、映像情報メディア学会の標準動画像を評価映像として用いた。

幅広い符号化難易度分布を持ち、かつ多様な絵柄を含む評価映像群を選定するため、絵柄の重複がなく様々な絵柄が含まれること、超高精細映像の評価映像として適切であること、符号化難易度の分布が従来の評価映像と類似することを条件とした。

表 2-1 評価映像

映像形式	準拠規格	シーン
1080/60/P, SDR	Rec. ITU-R BT.709	ハイビジョン・システム評価用標準動画像第二版 B シリーズの 8 映像(図 2-1)
2160/60/P, SDR	Rec. ITU-R BT.2020	超高精細・広色域標準動画像 A シリーズ、及び B シリーズの 8 映像(図 2-2)
2160/60/P, HDR	Rec. ITU-R BT.2100	超高精細・広色域 HDR 版標準動画像 C シリーズの 8 映像(図 2-3)
2160/60/P, SDR	Rec. ITU-R BT.2020	2160/60/P, HDR 評価映像を SDR 変換したもの 主観評価実験には用いていない

評価映像は、表 2-2 に示す映像形式にて実験で使用した。

表 2-2 映像形式

項目	値
映像信号形式	Y'CbCr 4:2:2
画素ビット数	10-bit
シーン長	10 秒 900 フレーム(15 秒)の内、第 180 フレームから第 779 フレームまでを使用

			
S201 Ginkgo trees	S202 Truck train	S204 Red leaves (pan up)	S209 Fountain (dolly)
			
S210 Studio concert	S214 Basketball	S218 Horse racing (dirt)	S265 Fountain (chromakey)

図 2-1 1080/60/P, SDR の評価映像

			
A03 Trains C	A05 Steel plant	A06 Festival	B06 Paddock
			
B07 Marathon (start)	B09 Marathon (panning)	B11 Water polo (scrolling)	B13 Drama (coffee)

図 2-2 2160/60/P, SDR の評価映像

			
C01 Fireworks (willow)	C05 Fireworks (barrage)	C06 Drama (standing up)	C08 Drama (sunset)
			
C11 Swim race (backstroke)	C12 Volleyball (fixed)	C15 Paddock (fixed)	C17 Horse race (homestretch)

図 2-3 2160/60/P, HDR の評価映像

2.3. 符号化条件

表 2-3 に示すエンコーダ A 及びエンコーダ B を用いて VVC 符号化を行った。エンコーダ A は、2025 年頃にハードウェアで実現可能なリアルタイムエンコーダの画質を実現するソフトウェアエンコーダである。エンコーダ B は、実装方法の違いによる符号化性能差の検証のために補助的に用いるものであり、一部のビットレートのみで使用した。実験では、エンコーダ A の符号化映像とエンコーダ B の符号化映像とを混ぜて評定者に提示した。この際、エンコーダ B の符号化映像は、符号化難易度が比較的高いもの(図 2-1 (1080/60/P, SDR) 及び図 2-2 (2160/60/P, SDR) それぞれの中の、下線付きの 4 評価映像)のみを使用した。

また、1080/60/P, SDR の評価映像 S265 (Fountain (chromakey)) に対してエンコーダ A の符号化制御のチューニングを行った映像が提供された。このチューニング効果を確認するビューイングを、本実験とは別に実施した。

表 2-3 VVC エンコーダ

エンコーダ	説明	符号化ビットレート
A	VVC エンコーダエミュータ 総務省の技術試験事務にて開発した、2025 年頃にハードウェアで実現可能なリアルタイムエンコーダの画質を実現するソフトウェアシミュレータ	表 2-4 に記載した全ビットレートで符号化
B	Fraunhofer HHI 研究所 VVC ソフトウェアエンコーダ VVenC (Version 0.2.1.0) VVC 標準化作業で開発された、最高性能を達成する参照ソフトウェア VTM (VVC Test Model) の最適化版リアルタイム動作よりも高性能化を主眼としており、将来の放送サービスにそのまま適用することは難しい	1080/60/P, SDR は 3Mbps 及び 7Mbps にて符号化 2160/60/P, SDR は 10Mbps 及び 20Mbps にて符号化

表 2-4 に示す条件にて VVC 符号化を行った。

ビットレートは、高度広帯域衛星デジタル放送の映像符号化方式検討時¹の値(HEVC を用い、1080/60/P, 2160/60/P でそれぞれ 10 Mbps – 15 Mbps, 30 Mbps – 40 Mbps で符号化)を参考に決めた。高ビットレートレンジでは VVC 方式の適用により 30% のビットレート削減が確実に見通せることから、高ビットレート側の値をそれぞれ 10 Mbps ($\approx 15 \times 0.7$), 30 Mbps ($\approx 40 \times 0.7$) とした。低ビットレートレンジの値は、これらの値の 70%, 50%, 30% とした。

表 2-4 符号化条件

項目	値	
プロファイル	Main 10 (10-bit, 4:2:0)	
ビットレート	1080/60/P 2160/60/P	3 Mbps, 5 Mbps, 7 Mbps, 10 Mbps 10 Mbps, 15 Mbps, 20 Mbps, 30 Mbps
符号化パラメータ	GOP 構造	階層 B 参照 (ARIB STD-B32 に記載の L3 構造)
	IRAP 間隔	32/60 sec
	GOP 長	8 フレーム
	CPB サイズ	1 秒分

¹ 実験報告書は https://www.soumu.go.jp/main_content/000262094.pdf である。

項目	値	
	色差信号 サンプリング位置	垂直・水平方向共、輝度信号位置と同じ (vui_chroma_sample_loc_type_frame = 2)
ツール設定	エンコーダ A	スクリーンコンテンツ向けツール(IBC)は無効化 他のツールは非公開
	エンコーダ B	Medium 設定、MCTF (プレフィルタ)は無効化

2.4. 評価実験方法

表 2-5 に示す方法にて実施した。

表 2-5 評価実験方法

項目	内容
評価方法	二重刺激劣化尺度(DSIS)法、5段階劣化尺度(表 2-6) 基準映像—評価映像のペアを一回提示
観視条件	Rec. ITU-R BT.500-14
評価者	専門家 38 人 注記: 装置故障により、1 人の 2160/60/P, SDR の実験データが欠損
ディスプレイ	PVM-X550 (55-inch LCD)
視距離	画面高の 1.5 倍(1.5H)(2160/60/P 時) 画面高の 3 倍(3H)(1080/60/P 時)

表 2-6 5段階劣化尺度

評点	評価語
5	劣化が分からない
4	劣化が分かるが気にならない
3	劣化が気になるが邪魔にならない
2	劣化が邪魔になる
1	劣化が非常に邪魔になる

一回の実験時間は、約 2 時間 30 分となった。内訳は、3 つの映像フォーマットそれぞれの観視・評価時間 (20 分×3)、説明時間 (30 分)、及び休憩時間 (30 分×2) である。なお、新型コロナウイルス対策のため、一日当たりの評価者数上限を 6 名 (午前 3 名、午後 3 名) とした。

2.5. 所要ビットレート推定の基準

放送品質を満足するサービスを提供するためには、画質の許容限界と見なされる平均評価値 (Mean Opinion Score, MOS) 3.5 以上がほぼ全ての映像で満足することが望まれ、また MOS 3 未満は放送品質として許容しがたいと考えられる。そこで、平均値と分散を元に、MOS 3.5 以上及び MOS 3 未満に該当するか否かを、有意水準 5% で検定する。

3. 符号化結果

各映像フォーマットの客観画質(15秒平均)を以下に示す。

3.1. エンコーダ A

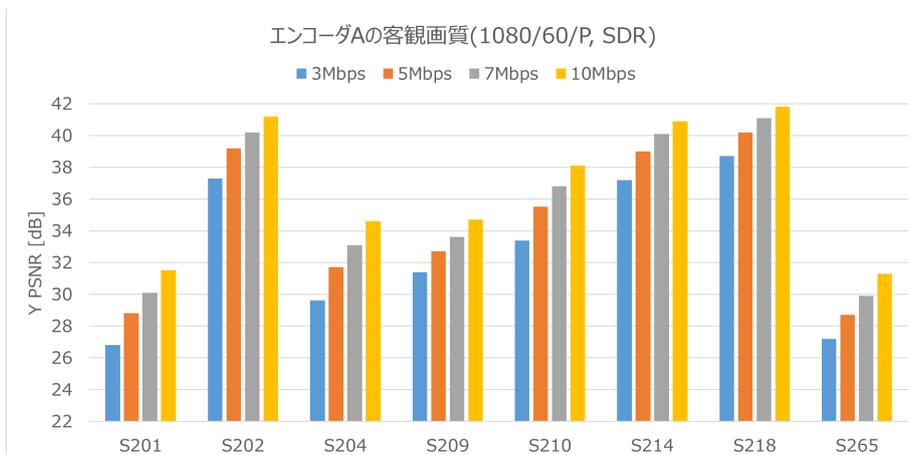


図 3-1 エンコーダ A の客観画質(1080/60/P, SDR)

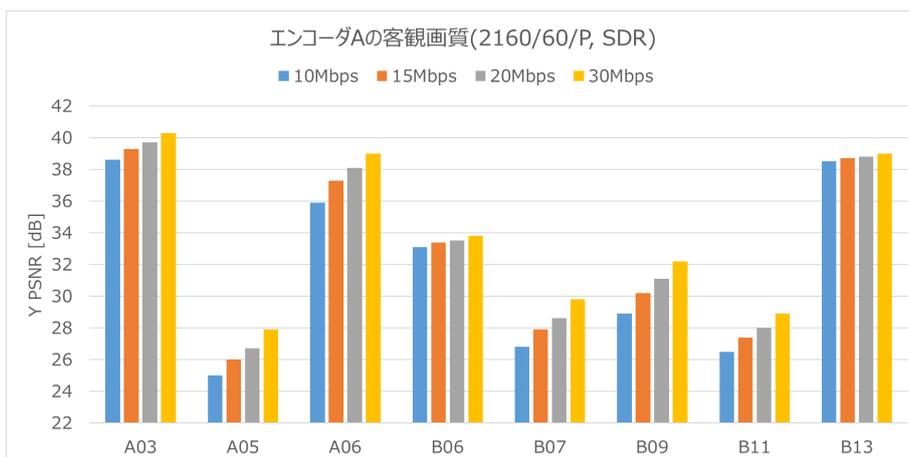


図 3-2 エンコーダ A の客観画質(2160/60/P, SDR)

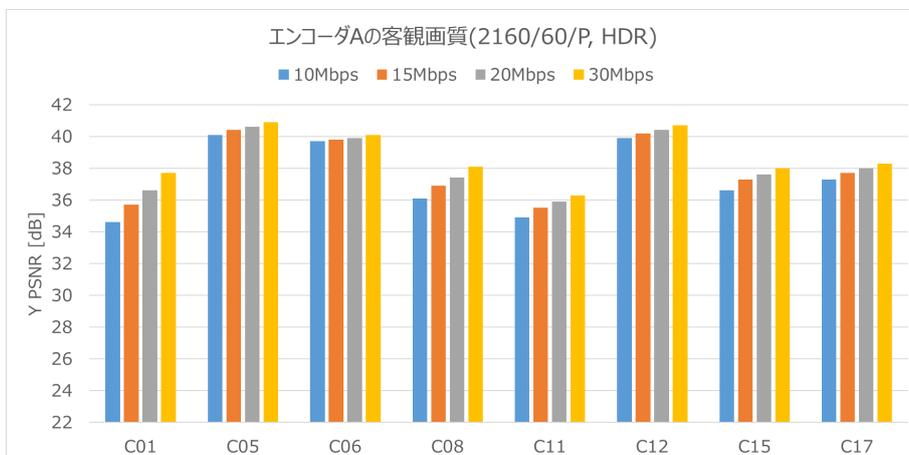


図 3-3 エンコーダ A の客観画質(2160/60/P, HDR)

3.2. エンコーダ B

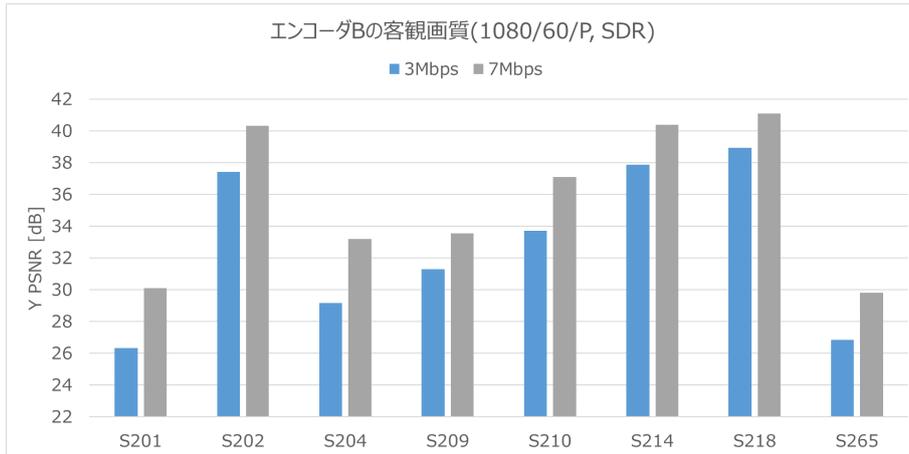


図 3-4 エンコーダ B の客観画質(1080/60/P, SDR)

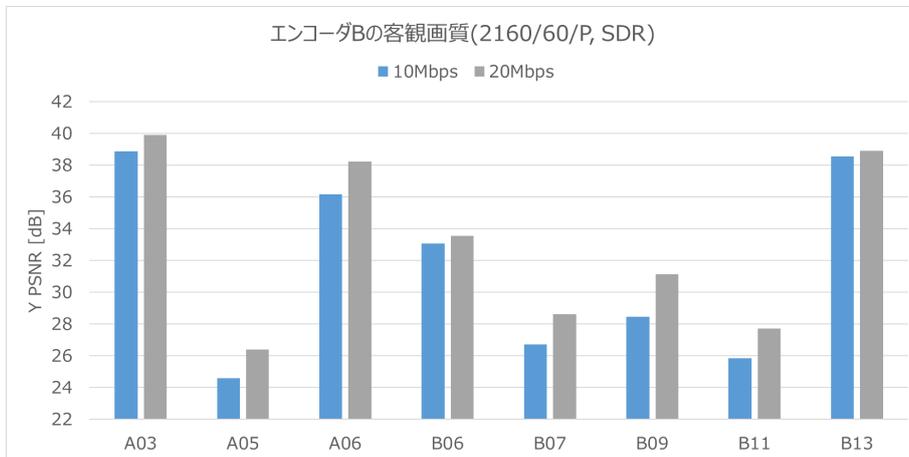


図 3-5 エンコーダ B の客観画質(2160/60/P, SDR)

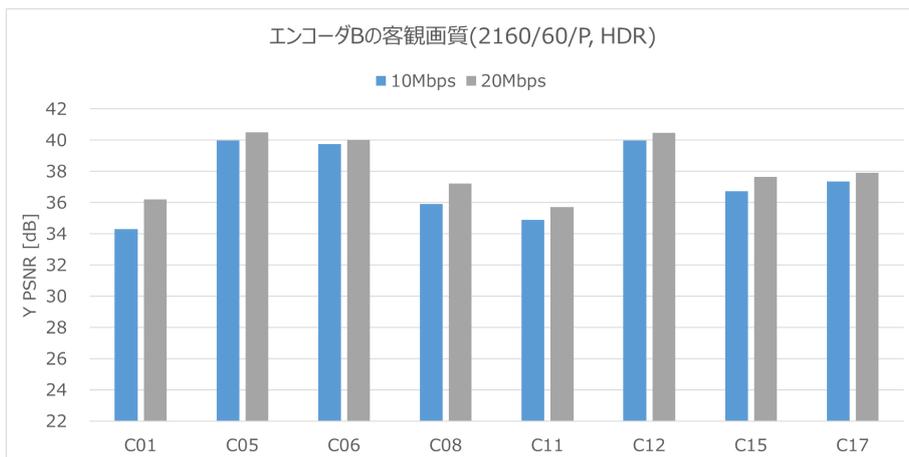


図 3-6 エンコーダ B の客観画質(2160/60/P, HDR)

2160/60/P, HDR 映像を SDR 変換した映像の客観画質を図 3-7、2160/60/P, HDR 映像と SDR 変換した映像との比較を図 3-8 に示す。SDR 変換により概ね 3 dB 低下する。

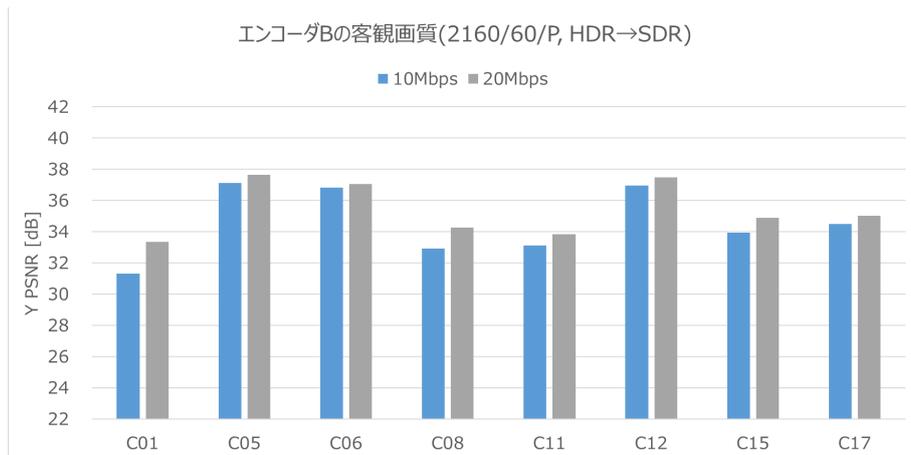


図 3-7 エンコーダ B の客観画質(2160/60/P, HDR→SDR)

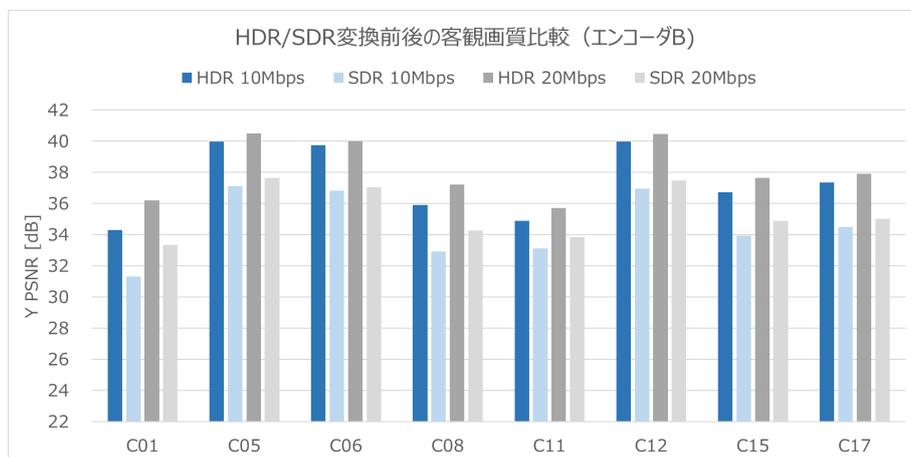


図 3-8 エンコーダ B における 2160/60/P, HDR 映像を SDR 変換した場合の客観画質の違い

4. 実験結果

4.1. 評価者のスクリーニング

3つの映像フォーマットそれぞれで、評価者のスクリーニングを行った。外れ値によって除外される評価者はいなかった。

更に、Pearson 相関に基づくスクリーニングを行った。図 4-1、図 4-2、図 4-3 に各映像フォーマットにおける、各評価者の評価結果と評価者平均値との Pearson 相関を示す。2160/60/P, HDR のみ、Pearson 相関が 0.4 未満の評価者が 7 名おり、これら評価者の評価結果を除外した。スクリーニング後の評価者数は 1080/60/P, SDR、2160/60/P, SDR、2160/60/P, HDR でそれぞれ、38 名、37 名、31 名となった。

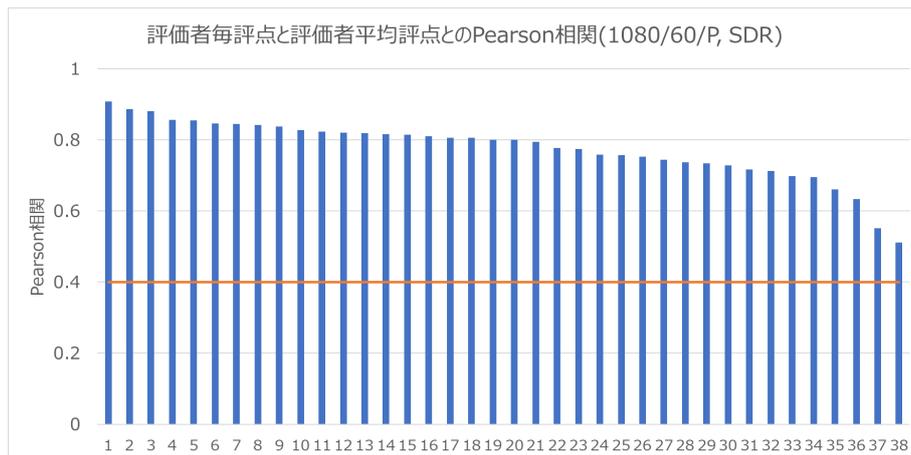


図 4-1 評価者毎評点と評価者平均評点との Pearson 分布(1080/60/P, SDR)

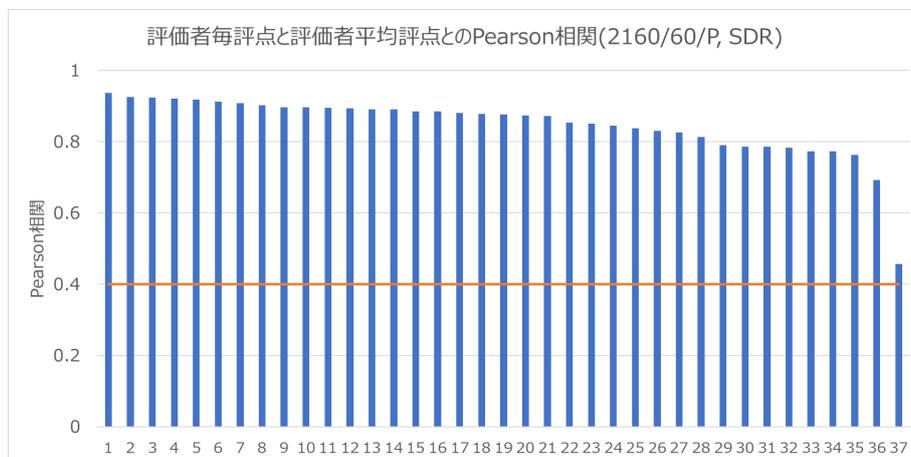


図 4-2 評価者毎評点と評価者平均評点との Pearson 分布(2160/60/P, SDR)

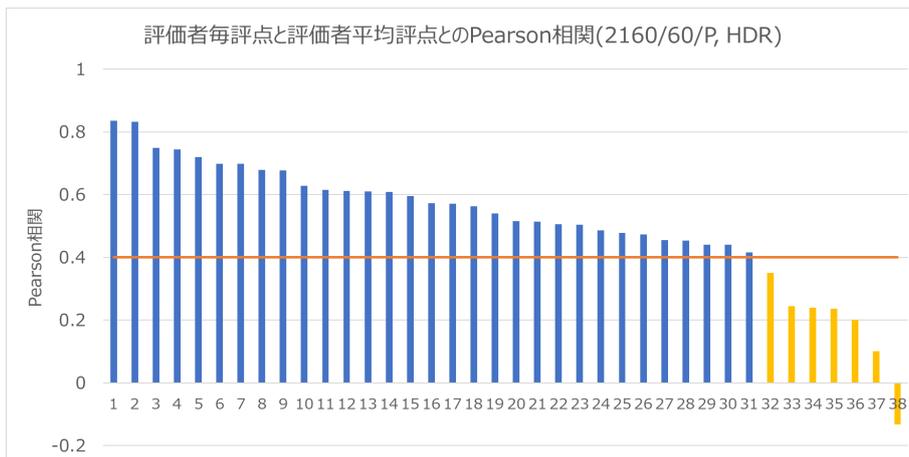


図 4-3 評価者毎評点と評価者平均評点との Pearson 分布(2160/60/P, HDR)

4.2. ビットレートと MOS 値との関係

エンコーダ A の符号化映像について、各映像フォーマットでのビットレートと MOS 値との関係、及び画質基準に対する評価映像の割合を示す。表の「MOS 値 3.5 以上」及び「MOS 値 3.0 未満」の算出には、エラー上限値を用いている。また、参考値である「MOS 値平均」の算出には、平均値を用いている。

4.2.1. 1080/60/P, SDR

最も符号化難易度が高い評価映像 S265 の MOS 値は、ビットレートが 7 Mbps 以上の時に 3.0 以上となる。また、5 Mbps 以上の時に、S265 以外の全評価映像の MOS 値が 3.5 以上となる。

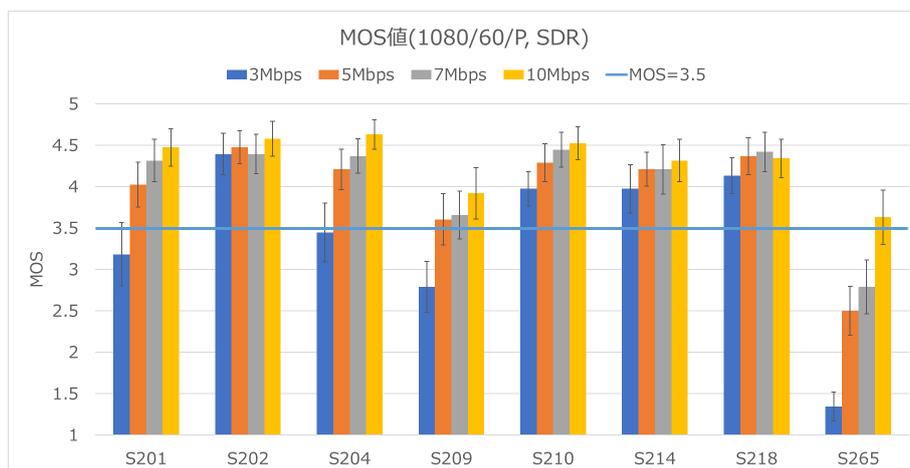


図 4-4 1080/60/P, SDR の評価映像におけるビットレートと MOS 値との関係(平均値と95%信頼区間)

表 4-1 1080/60/P, SDR の評価映像におけるビットレートと画質基準に対する評価映像の割合

	3 Mbps	5 Mbps	7 Mbps	10 Mbps
MOS 値 3.5 以上	6 / 8	7 / 8	7 / 8	8 / 8
MOS 値 3.0 未満	1 / 8	1 / 8	0 / 8	0 / 8
MOS 値平均(参考)	3.4	4.0	4.1	4.3

4.2.2. 2160/60/P, SDR

最も符号化難易度が高い評価映像 A05 の MOS 値は、ビットレートが 30 Mbps 以上の時に 3.0 以上となる。また、30 Mbps 以上の時に、A05 以外の全評価映像の MOS 値が 3.5 以上となる。

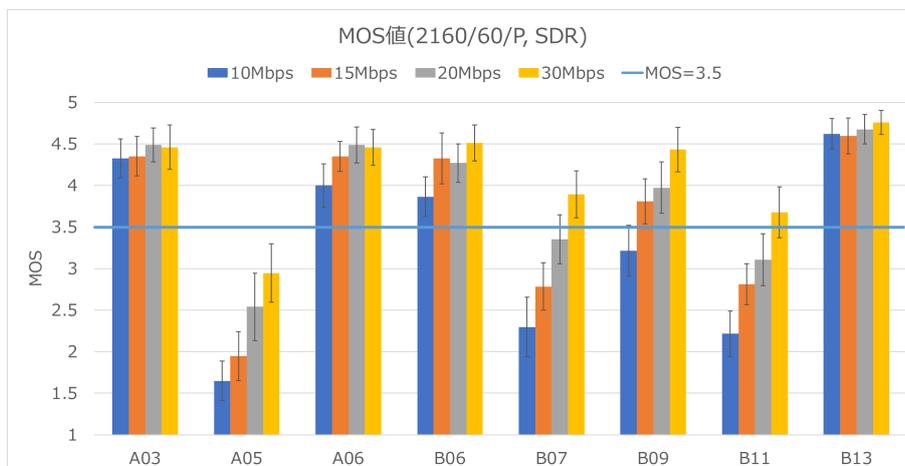


図 4-5 2160/60/P, SDR の評価映像におけるビットレートと MOS 値との関係(平均値と 95%信頼区間)

表 4-2 2160/60/P, SDR の評価映像におけるビットレートと画質基準に対する評価映像の割合

	10 Mbps	15 Mbps	20 Mbps	30 Mbps
MOS 値 3.5 以上	5 / 8	5 / 8	6 / 8	7 / 8
MOS 値 3.0 未満	3 / 8	1 / 8	1 / 8	0 / 8
MOS 値平均(参考)	3.3	3.6	3.9	4.1

4.2.3. 2160/60/P, HDR

ビットレートが 10 Mbps の時に、全ての評価映像の MOS 値が 3.0 以上となる。

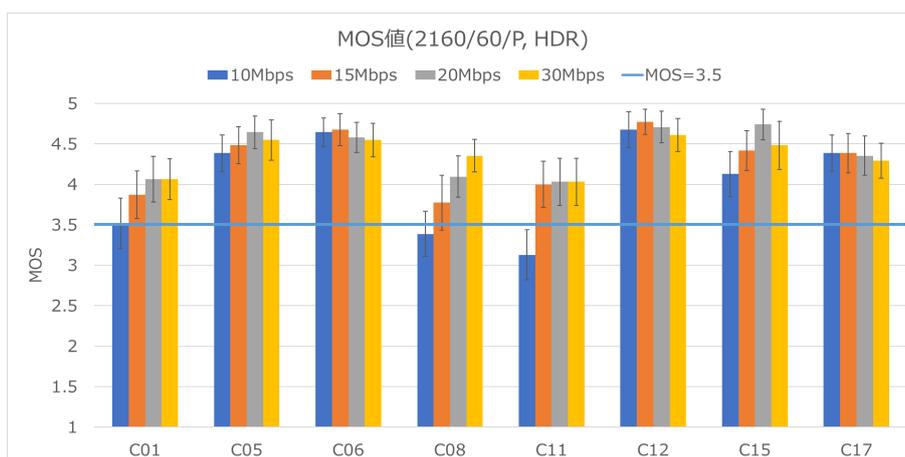


図 4-6 2160/60/P, HDR の評価映像におけるビットレートと MOS 値との関係(平均値と 95%信頼区間)

表 4-3 2160/60/P, HDR の評価映像におけるビットレートと画質基準に対する評価映像の割合

	10 Mbps	15 Mbps	20 Mbps	30 Mbps
MOS 値 3.5 以上	7 / 8	8 / 8	8 / 8	8 / 8
MOS 値 3.0 未満	0 / 8	0 / 8	0 / 8	0 / 8

MOS 値平均(参考)	4.0	4.3	4.4	4.4
-------------	-----	-----	-----	-----

5. 所要ビットレート

2.5 節の判断基準から、VVC 方式適用時の所要ビットレートは表 5-1 のように導出される。なお、2160/60/P 映像の所要ビットレートは、SDR 映像と HDR 映像で分けず、SDR 映像で得られた値とした。

今回用いた HDR 映像の符号化難易度は SDR 映像に比べて低く、図 8-2、図 8-3 に示すように SDR 映像における符号化難易度の低い 4 つのシーケンス(A03, A06, B06, B13)に相当する。これらのシーケンスは SDR 映像の主観評価実験においていずれも 10Mbps で MOS 値 3.5 以上の要求品質を満たしており(図 4-5)、HDR 映像の主観評価結果(図 4-6)と整合している。また、実際の運用では 2160/60/P, SDR 映像相当の符号化難易度を持つ 2160/60/P, HDR 映像が使用されると想定される。これらのことから、HDR 映像には SDR 映像と同等のビットレートが要求されると考えられる。

表 5-1 VVC 方式適用時の所要ビットレート

	1080/60/P	2160/60/P
所要ビットレート	7 Mbps	30 Mbps

6. 考察

6.1. エンコーダ A の妥当性

エンコーダ A の符号化映像と、エンコーダ B の符号化映像との MOS 値の違いを図 6-1 及び図 6-2 に示す。映像は、本実験の評価映像の中で符号化難易度が比較的高いものである。

これらの図の通り、エンコーダ A とエンコーダ B との MOS 値差は、シーンによって変わるものの、平均して同等と見なせる。このことから、本実験結果は、エンコーダ A に特化した結果ではないと言える。

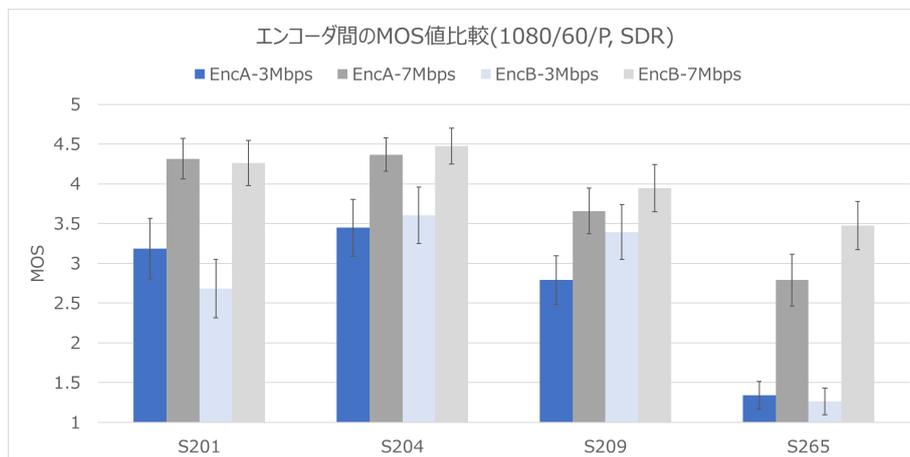


図 6-1 エンコーダ A とエンコーダ B の MOS 値比較 (1080/60/P, SDR 映像)

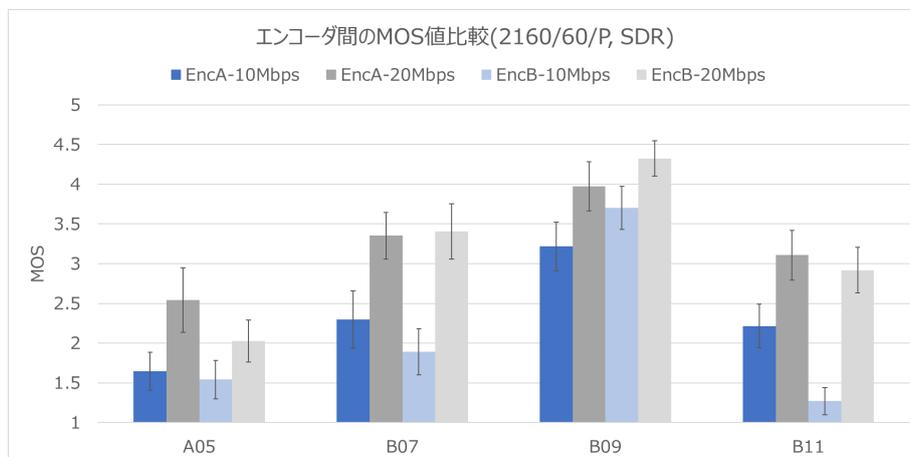


図 6-2 エンコーダ A とエンコーダ B の MOS 値比較 (2160/60/P, SDR 映像)

6.2. 所要ビットレート推定(2160/60/P, SDR)

5章では、2160/60/P, SDR の所要ビットレートを 30 Mbps とした。本実験における次に小さいビットレートは 20 Mbps であり、2.5 節の判断基準を満たすビットレートは 20 Mbps と 30 Mbps の間にあると考えられる。そこで、20 Mbps と 30 Mbps の間を直線近似し、判断基準を満足するビットレートを推定した。

図 4-5 及び表 4-2 から、評価映像 A05 の MOS 値が 3.0 以上、かつ評価映像 B11 の MOS 値が 3.5 以上となるビットレートが、推定するビットレートとなる。図 6-3 及び図 6-4 は、それぞれのビットレートと MOS 値(実線は平均値、点線はエラーバー上限及び下限)との関係である。エラーバー上限が上記の値となるビットレート 22 Mbps が、実際の所要ビットレートと推定される。

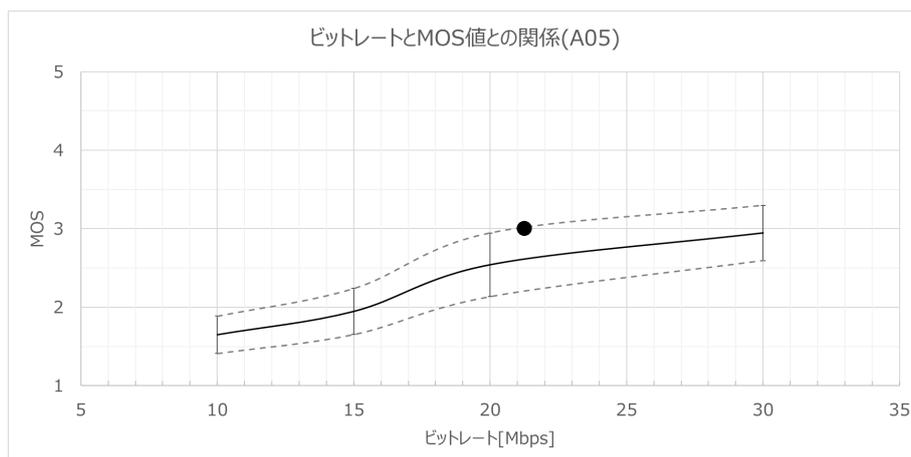


図 6-3 ビットレートと MOS 値との関係 (評価映像 A05)

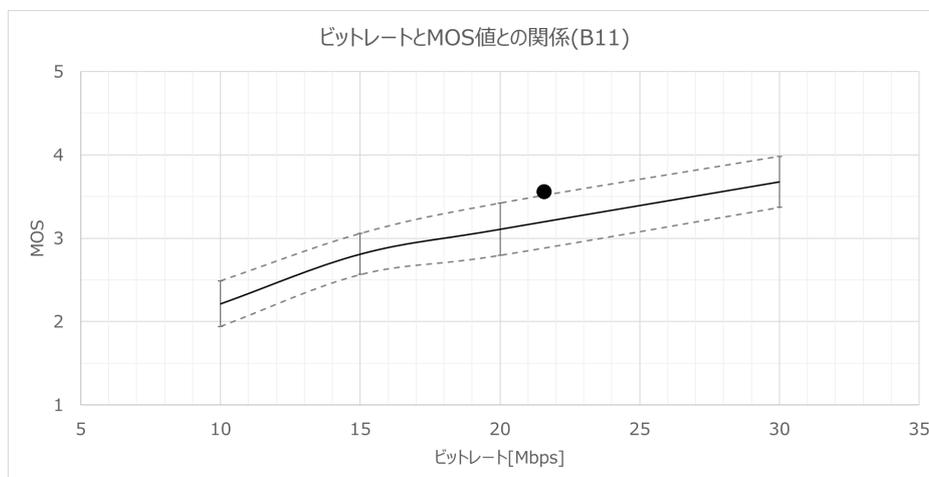


図 6-4 ビットレートと MOS 値との関係 (評価映像 B11)

6.3. VVC 規格の公称性能と所要ビットレートとの関係

VVC 規格の HEVC 規格からの圧縮効率向上は、客観画質評価に基づく分析によって 40%弱、主観画質評価に基づく分析によって 50%弱(共に SDR 映像)と、VVC 標準化を担う機関から報告されている。これらは、評価に用いられた少数の評価映像の平均的な数値である。一方、本報告における所要ビットレートは、符号化難易度の高い映像を含む評価画像のほぼ全てで一定水準以上の画質を確保可能なビットレートである。この違いに留意する必要がある。

近年の映像符号化方式は、映像に存在する冗長性を徹底的に取り除くことで圧縮効率を伸ばしている。冗長性が比較的多く存在し符号化難易度が低い映像では公称性能に近い圧縮効率向上となる一方、冗長性が少なく符号化難易度が高い映像では小さくなる。極端な例では、冗長性の無いノイズ映像は圧縮効率

の向上を期待できない。

今回の実験結果で得られた 1080/60/P 映像及び 2160/60/P 映像の VVC による所要ビットレートと高度広帯域衛星デジタル放送の映像符号化方式検討時の HEVC による所要ビットレートを表 6-1 に示す。1080/60/P では 30%以上の圧縮効率向上が認められる一方で、2160/60/P では圧縮効率の向上が小さい。

表 6-1 VVC と HEVC の所要ビットレート

	VVC	HEVC
1080/60/P	7 Mbps	10 – 15 Mbps
2160/60/P	30 Mbps	30 – 40 Mbps

符号化難易度が高い映像を除いた他の 2160/60/P 映像における、MOS 値 3.5 を超えるビットレートの比較を表 6-2 に示す。HEVC の場合の 15Mbps に対して、VVC では 10Mbps であり、30%以上の圧縮効率向上が認められる。

表 6-2 符号化難易度が高い映像を除いた他の 2160/60/P 映像における、MOS 値が 3.5 を超えるビットレート

VVC	HEVC	HEVC からの 圧縮効率向上率
10 Mbps	15 Mbps	約 33%

6.4. 主観画質改善の可能性

6.4.1. 1080/60/P, SDR

2.3 節のように、符号化制御チューニングを行ったエンコーダ A の符号化映像(評価映像 S265 のみ)が提供されたため、チューニング前(本実験対象)とチューニング後(本実験対象外)との主観的な画質比較を、実験参加者へのデモセッション(実験後、表 2-5 の条件に合致しない環境)にて実施した。

符号化制御チューニングを行うことによって、主観画質が概ね一つ上のビットレート(例えば、チューニング後の 5 Mbps 映像の主観画質は、チューニング前の 7 Mbps 映像のものと同様)となることが確認された。他の評価映像では未確認ではあるが、このような符号化制御チューニングは、HEVC 方式の実用化時にも同様な効果を生んだものであるため、他のシーンへの悪影響はないものと推定される。

4.2.1 節のように、所要ビットレートは評価映像 S265 の MOS 値によって 7 Mbps という結論となっているが、上記の結果を考慮すると、実用的には 5 Mbps でも基準を満足する可能性が高いものと思われる。

6.4.2. 2160/60/P, SDR

2160/60/P SDR の評価映像の中で、特に A05 の符号化難易度が高い。2160/60/P SDR の評価映像セットの中には、他にも A07 という非常に符号化難易度が高い映像がある(図 8-2 参照)。事前の主観画質確認では、本実験よりも更に高いビットレートでの符号化でないと十分な主観画質が得られないことが判明している。

VVC 規格は、従来規格にはない新たな符号化ツールとして、例えば RPR (Reference Picture Resampling)² が導入されている。RPR を用いて、符号化難易度が高いシーンでは空間解像度を落として(例えば 2160/P を 1080/P に)符号化することで、主観画質の向上が期待できる。RPR は本実験で使用したエンコーダに実装されていないため、使用していない。

² 動き補償時に、処理対象ピクチャと参照ピクチャとの画面サイズ比を考慮したスケーリングとフィルタリングを行う技術である。これにより、各ピクチャで異なる画面サイズを取ることが可能になる。

7. 追加画質評価

7.1. 目的

特定のビットレートでの 2160/60/P 映像の符号化画質の考察、及び画質改善手法(空間解像度を落として符号化歪を低減する二つの手法)の効果を確認する目的で、追加の画質評価を実施した。

情報通信審議会放送システム委員会地上デジタル放送方式高度化作業班で検討中の伝送方式による映像ビットレートを参考に、7 Mbps, 10 Mbps, 15 Mbps, 20 Mbps, 25 Mbps の画質を検討した。

7.2. 実験計画

7.2.1. 実験会場、日時

日本放送協会放送技術研究所(世田谷区砧)

2021年12月17日、20日

7.2.2. 評価映像

図 2-2 の 2160/60/P, SDR 映像(8 シーケンス)に、図 7-1 の映像 A07(6.4.2 節で言及した符号化難易度が非常に高い映像)を追加した。映像形式は表 2-2 の通りである。



A07
River

図 7-1 追加評価映像

7.2.3. 符号化条件

表 2-3 のエンコーダ B の 1.2.0 版を用いて、表 7-1 の条件で VVC 符号化を行った。ビットレート及びツール設定が表 2-4 とは異なる(表中下線部)。Slower 設定は、実時間動作可能なエンコーダの 2025 年以降の更なる性能改善を想定した設定であり、BD-Rate が Medium 設定よりも約 6%改善する一方、処理時間は約 14 倍増加する。

9 つの評価映像の内、符号化難易度が相対的に低い 5 つの映像(A03, A06, B06, B09, B13)は、図 4-5 に示す通り、10 Mbps で MOS 値が 3.5 を超えているため、7 Mbps のみの評価とした。

表 7-1 符号化条件

項目	値 (注:表 2-4 と異なる設定を下線で示す)	
プロファイル	Main 10 (10-bit, 4:2:0)	
ビットレート	7 Mbps, 10 Mbps, 15 Mbps, 20 Mbps, 25 Mbps	
符号化パラメータ	GOP 構造	階層 B 参照 (ARIB STD-B32 に記載の L3 構造)
	IRAP 間隔	32/60 sec
	GOP 長	8 フレーム
	CPB サイズ	1 秒分

項目	値（注：表 2-4 と異なる設定を下線で示す）	
	色差信号 サンプリング位置	垂直・水平方向共、輝度信号位置と同じ (vui_chroma_sample_loc_type_frame = 2)
ツール設定	<u>Medium 設定、及び Slower 設定</u> <u>MCTF (プレフィルタ) を有効化 (設定値 2)</u> <u>PerceptQPA (適応 QP 処理による主観画質向上オプション) を有効化 (設定値 1)</u>	

7.2.4. 符号化映像

符号化映像は表 7-2 の通りである。4K 解像度映像(3,840×2,160)をそのまま符号化した映像(以下、4K 符号化映像)に加え、符号化前に空間方向のローパスフィルタを適用した符号化映像(以下、前処理符号化映像)、及び RPR を想定し符号化画素数変更を行った符号化映像(6.4.2 節で言及、以下、RPR 符号化映像)の三種類の符号化映像を使用した。前処理符号化映像及び RPR 符号化映像は共に符号化する映像の空間解像度を低下させたものであるが、RPR 符号化映像は前処理符号化映像とは異なり、符号化画素数が少なく、符号化処理量も減少する。なお、シーン内で符号化解像度及び前処理フィルタ係数を固定し、映像に応じた適応的な制御は行っていない。

表 7-2 符号化映像

符号化映像	解像度縮小率(W, H)	符号化画素数	前処理
4K	(1/1, 1/1)	3,840×2,160	なし
前処理(2K 相当)	(1/1, 1/1)	3,840×2,160	フィルタ A
前処理(3K 相当)	(1/1, 1/1)	3,840×2,160	フィルタ B
RPR(2K)	(1/2, 1/2)	1,920×1,080	なし
RPR(3K)	(3/4, 3/4)	2,880×1,624	なし

表 7-2 のフィルタ A 及びフィルタ B は、RPR(2K)符号化及び RPR(3K)符号化にて符号化画素数を削減する際に使用するフィルタ(Lanczos-3)と同一の係数とした。これにより、元映像に対する 4K 再構成映像のボケ具合は、2K 同士及び 3K 同士で等しい。フィルタ係数と周波数特性は 8.2.1 節及び 8.2.2 節に記載の通りである。

7.3. 評価方法

7.3.1. 客観画質

復号映像の輝度成分 PSNR を元解像度・前処理適用前の原映像をリファレンスとして測定する。併せて、符号化ログからエンコーダ入力映像(前処理適用後もしくは画素数変更後)をリファレンスとした輝度成分 PSNR を測定する。

7.3.2. 主観画質

主観評価方法は表 7-3 の通りとする。時間等の制約を考慮し、比較的簡単な形式とした。符号化画素数を変更した場合には、復号映像を元の画素数に復元した上で、モニタに表示する。画素数の復元に用いたフィルタ(Lanczos-3)は 8.2.3 節及び 8.2.4 節に記載の通りである。

表 7-3 主観評価方法

項目	内容
提示方法	7.3.3 節の通り
観視条件	実験室環境
評価者	専門家 16 名
ディスプレイ	PVM-X550 (55-inch LCD)
視距離	画面高の 1.5 倍(1.5H)

7.3.3. 評価基準及び提示方法

以下の二つの評価を行った。

7.3.3.1. 評価 1: ビットレート毎の画質評価

符号化難易度が高い 4 つの映像(A05, A07, B07, B11)それぞれについて、表 7-4 のビットレート及びツール設定の順に提示・評価を行った。Slower 設定は、画質向上の効果がより明確になると想定される 7 Mbps 及び 10 Mbps のみとした。

表 7-4 絶対評価におけるビットレート・ツール設定

提示順番	ビットレート	ツール設定
1	7 Mbps	Medium
2	7 Mbps	Slower
3	10 Mbps	Medium
4	10 Mbps	Slower
5	15 Mbps	Medium
6	20 Mbps	Medium
7	25 Mbps	Medium

各ビットレート・ツール設定の組では、原画像の後に以下の順番で符号化映像を連続的に提示し、評定者は最後の符号化映像提示後に評価シートに評価値を記入する。なお、提示中の映像のビットレート、ツール設定及び符号化手法を評定者に事前に説明した上で評価を行い、また評定者が繰り返し映像を見ることを許容する。

表 7-5 各ビットレート・ツール設定での映像提示順番と評価内容

提示順番	符号化映像	評価値
1	4K 符号化映像	5 段階劣化尺度 (表 2-6、原画像比較)
2	前処理(2K 相当)符号化映像	4K 符号化映像との画質比較 ○:良い △:同等 ×:悪い
3	前処理(3K 相当)符号化映像	
4	RPR(2K)符号化映像	
5	RPR(3K)符号化映像	

7.3.3.2. 評価 2: 4K 放送品質映像との画質比較

リアルタイム HEVC ハードウェアエンコーダの 30 Mbps(4K 放送品質相当)での符号化映像(以降、HEVC

30Mbps)との画質比較を行う。

符号化難易度が高い4つの映像(A05, A07, B07, B11)については、表 7-1 の全てのビットレートにおいて評価する。なお、評価時間の制約から、事前確認の結果に基づき、ツール設定及び符号化映像はそれぞれ、Medium 設定、前処理(2K 相当)符号化映像(A05, A07 の場合)もしくは前処理(3K 相当)符号化映像(B07, B11 の場合)としている。

符号化難易度が相対的に低い5つの映像(A03, A06, B06, B09, B13)については、7 Mbps のみの評価とし、ツール設定及び符号化映像は Medium 及び 4K 符号化映像としている。

評価は、HEVC 30 Mbps と比較して、良い(○)、同等(△)、悪い(×)とする。

VVC 符号化映像と HEVC 30 Mbps 符号化映像は、隣接して並べた二つの 4K モニタに、シーン毎に同期する形で個別に提示する。1.5H の視距離を維持するため、評定者はそれぞれのモニタの前に適宜移動し、映像を繰り返し見ることが許容した。

7.4. 結果

7.4.1. 客観画質

図 7-2 から図 7-5 に、符号化難易度が高い映像(A05, A07, B07, B11)の客観画質(15 秒平均)を示す。各図には各ビットレート・ツール設定・符号化手法での輝度 PSNR(4K 原画像比較)に加え、VVC 符号化前の前処理適用映像及び RPR 適用映像の輝度 PSNR(4K 原画像比較)を示している。

前処理や RPR の効果は、映像、空間解像度、ビットレートによって異なるが、空間解像度を低下させるほど 4K 符号化映像よりも PSNR が低下する傾向があり、最大 1.5 dB 低下している(画像 B11)。同一空間解像度では、前処理よりも RPR の方が PSNR は低下し、最大 0.5 dB 低下している(画像 B11)。また、Medium 設定よりも Slower 設定の方が PSNR は増加し、最大 0.3dB 増加している(画像 B11)。

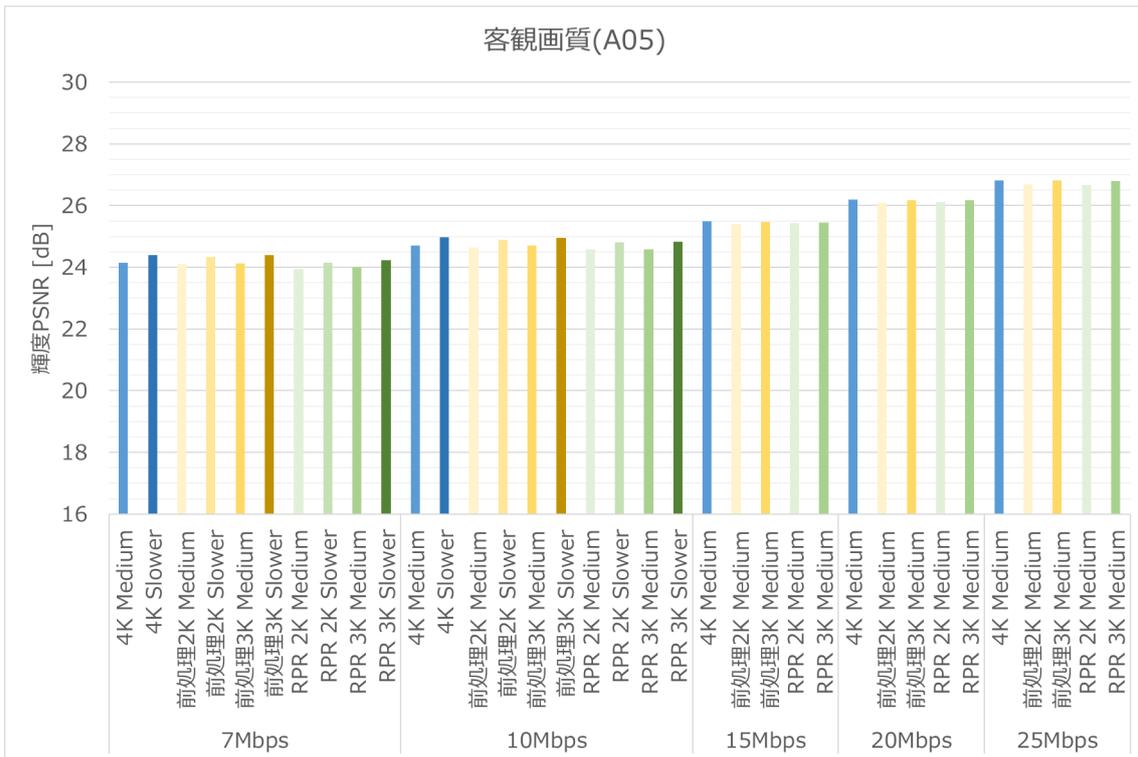


圖 7-2 客觀畫質(A05)

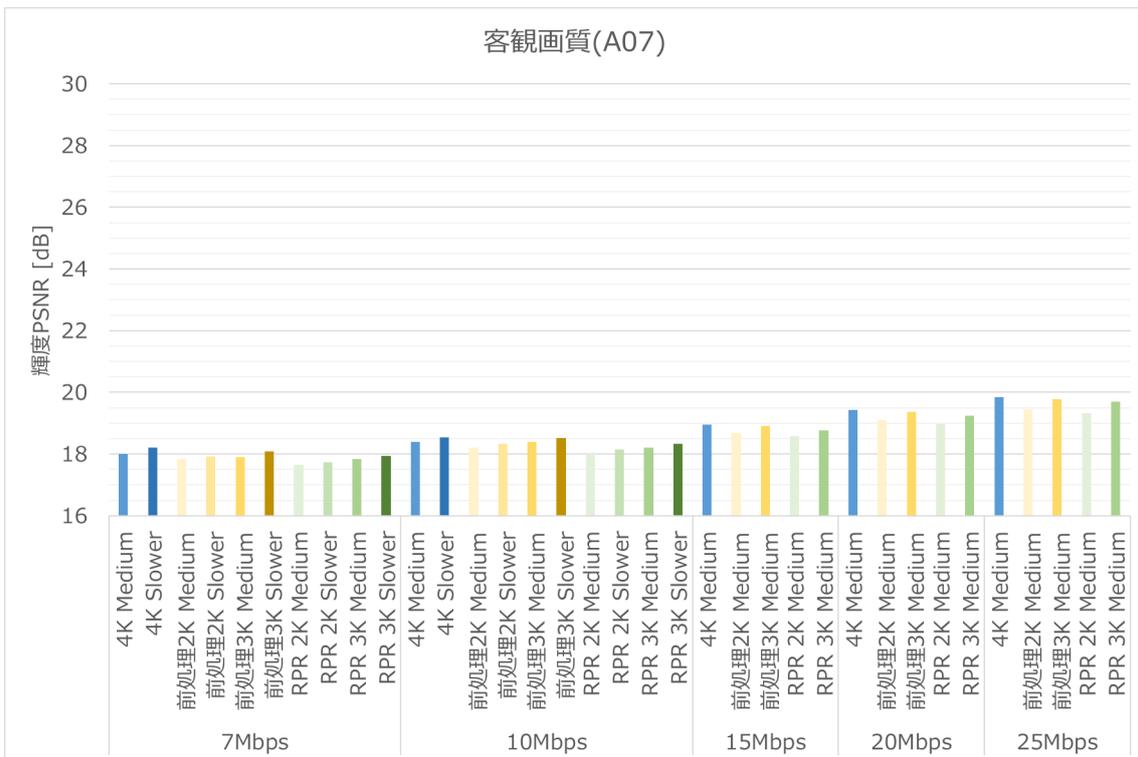


圖 7-3 客觀畫質(A07)

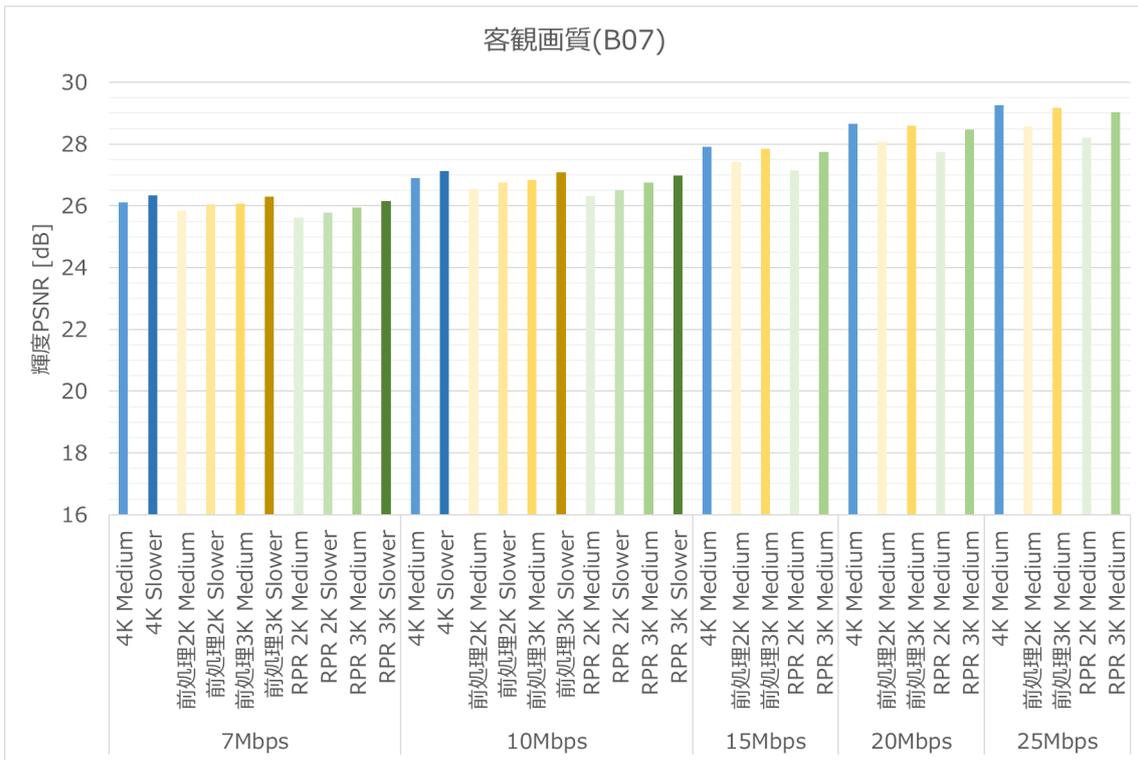


图 7-4 客觀畫質(B07)

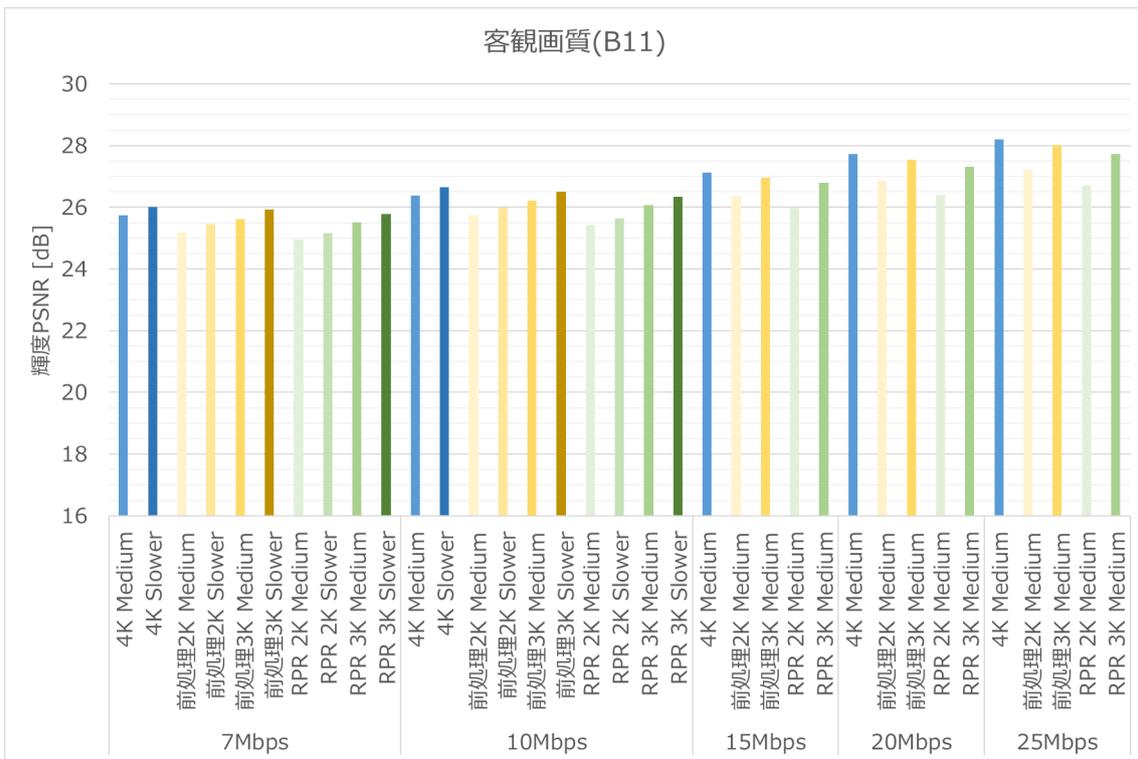


图 7-5 客觀畫質(B11)

図 7-6 に、符号化難易度が低い映像(A03, A06, B06, B09, B13)の 7 Mbps における客観画質を示す。B09を除き輝度 PSNR が 30 dB を超えている。また、3K 解像度相当の前処理フィルタが最も高い PSNR になっている。

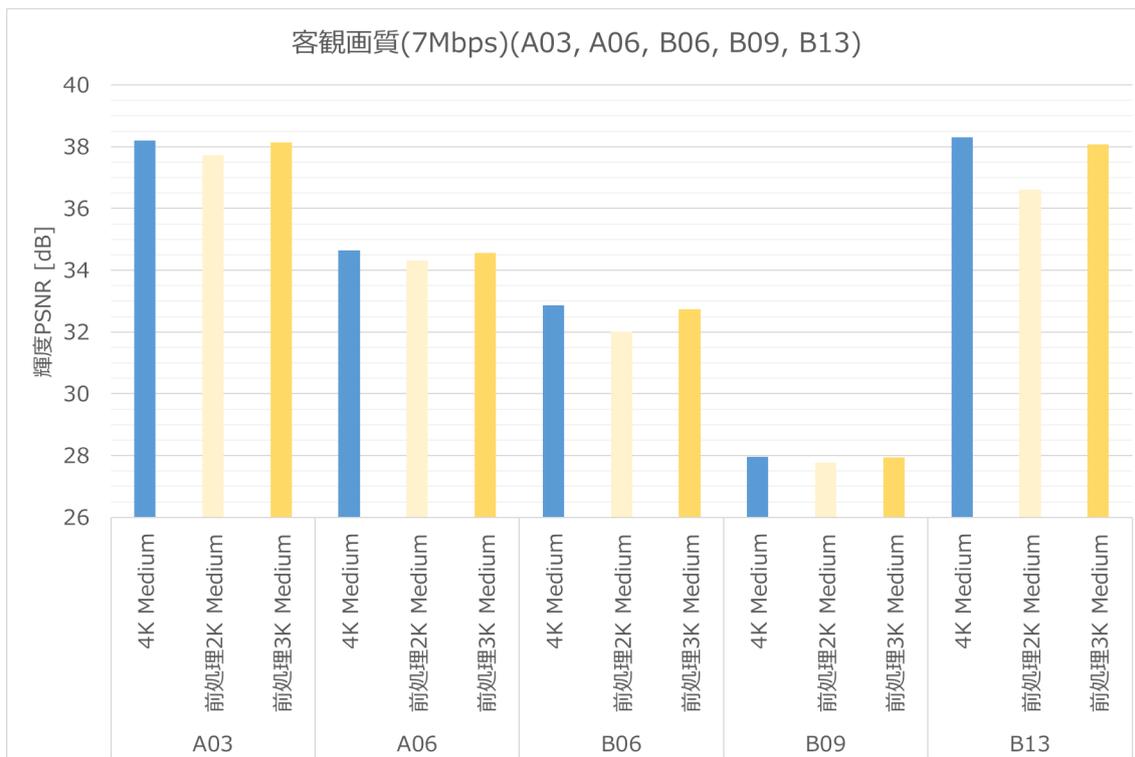


図 7-6 符号化難易度が低い映像の客観画質

図 7-7 に、圧縮前の前処理映像及び RPR 映像の輝度 PSNR を示す。RPR の方が PSNR は低下するが、これは画素数復元時のローパスフィルタの影響と考えられる。なお、評定者へのヒアリングにより主観画質(ボケ具合)は同等であることを確認した。

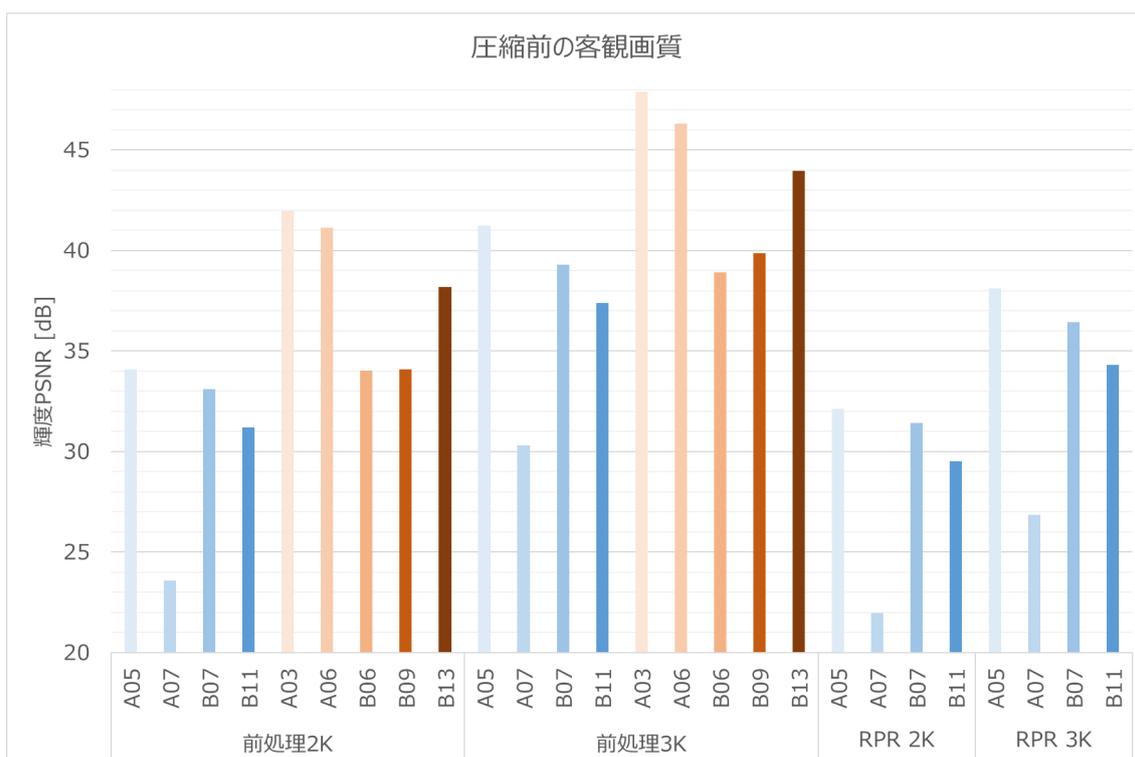


図 7-7 圧縮前の客観画質

図 7-8 に、符号化難易度が高い 4 つの映像(A05, A07, B07, B11)について、符号化入力映像(前処理映像及び 2K/3K 解像度映像)を基準とした場合の 7 Mbps における輝度 PSNR を示す。前処理及び RPR 共に、符号化解像度を 2K 相当まで落とすことで輝度 PSNR が 1 dB–2 dB 向上している。これは、同一レートでの条件下で量子化パラメータが小さくなって符号化歪が軽減したことを意味する。

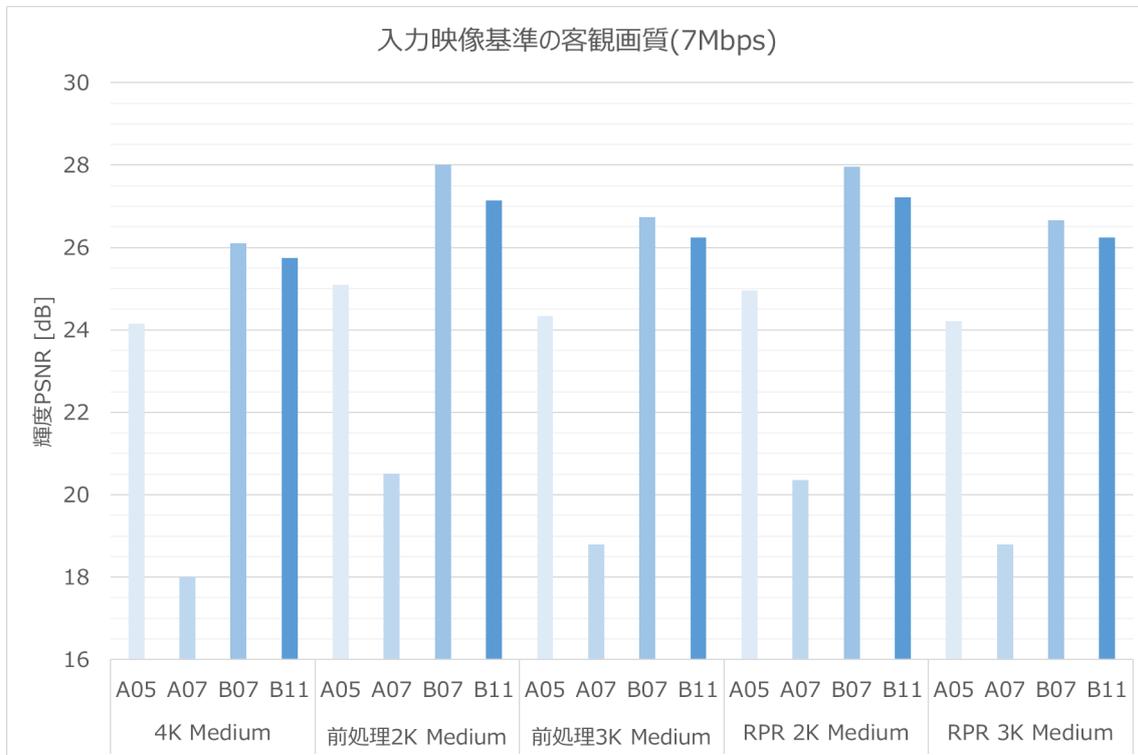


図 7-8 符号化入力映像基準の 7 Mbps 客観画質(A05, A07, B07, B11)

7.4.2. 評価者のスクリーニング

本実験は簡易的な主観画質評価のため、統計処理に基づいたスクリーニングは実施していない。

7.4.3. 符号化難易度が高い映像のビットレートと MOS 値との関係

図 7-9 に、符号化難易度が高い 4 映像の 4K 符号化映像の MOS 値を示す。全ての映像で MOS 値が 3.0 以上になるのは 20 Mbps から 25 Mbps の間である。10 Mbps 以下では全ての映像で MOS 値が 3.0 を下回る。Slower 設定は Medium 設定よりも若干の MOS 値改善があるが、一つ上のビットレート以上にまで改善することはなかった。今回の実験条件では Slower 設定による主観画質改善の効果は少なく、7.4.5 節に結果を示す HEVC 30 Mbps 符号化映像との比較において、評価を Medium 設定に限定したことは結論には影響しないと言える。

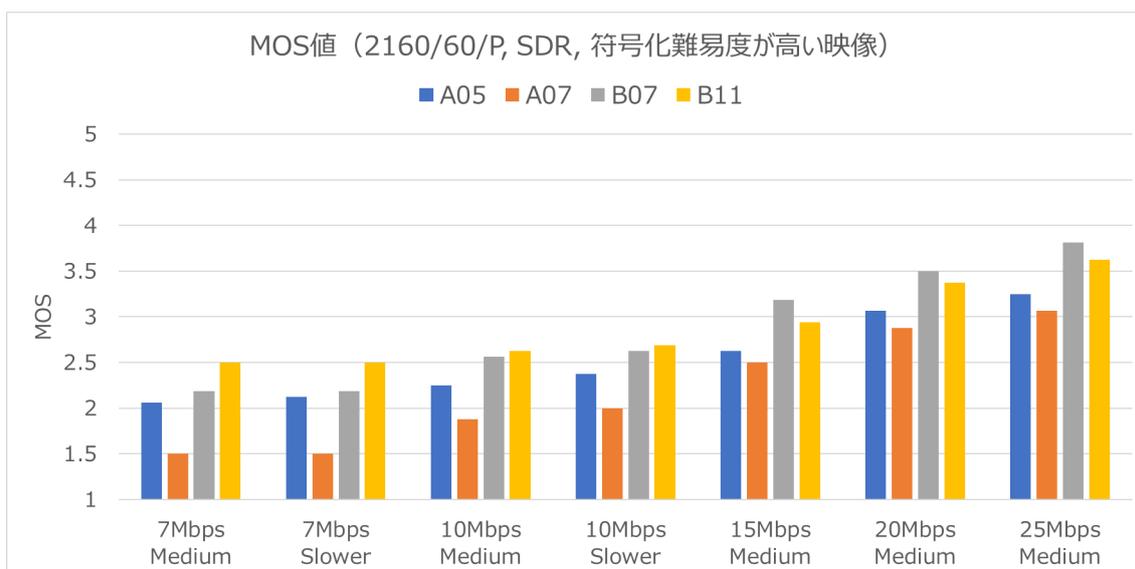


図 7-9 2160/60/P, SDR の符号化難易度が高い映像におけるビットレートと MOS 値との関係

7.4.4. 画質改善手法の効果

図 7-10から図 7-16に、評価1における、4K 符号化映像と比較した場合の、前処理符号化映像及びRPR 符号化映像の主観画質を良い(○)、同等(△)、悪い(×)の割合で示す。

以下の傾向が見られる。

- 符号化難易度の高さの順(A07, A05, B07, B11)で、4K 符号化映像よりも「良い」と回答した評定者が多くなる。ただし全評定者が「良い」と答えた条件は無い。
- 「良い」と回答した割合は、低ビットレート(15 Mbps 以下)では RPR(3K)符号化映像が最も高く、次いでRPR(2K)符号化映像、前処理(2K/3K 相当)符号化映像の順である。高ビットレート(20 Mbps 以上)では RPR(3K)符号化映像及び前処理(3K 相当)符号化映像、RPR(2K)符号化映像及び前処理(2K 相当)符号化映像の順である。
- 「悪い」と回答した割合は、RPR(2K)符号化映像及び前処理(2K 相当)符号化映像が高く、次いで RPR(3K)符号化映像及び前処理(3K 相当)符号化映像の順である。
- ビットレートが高くなるにつれて「悪い」の割合が増加する傾向がある。

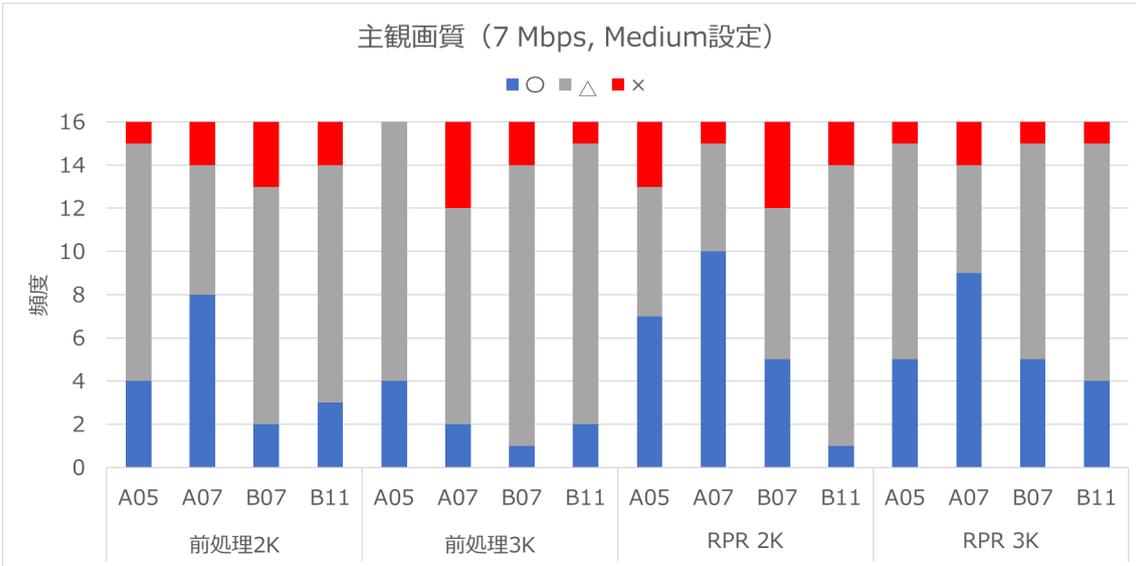


図 7-10 主観画質(7 Mbps, Medium 設定)

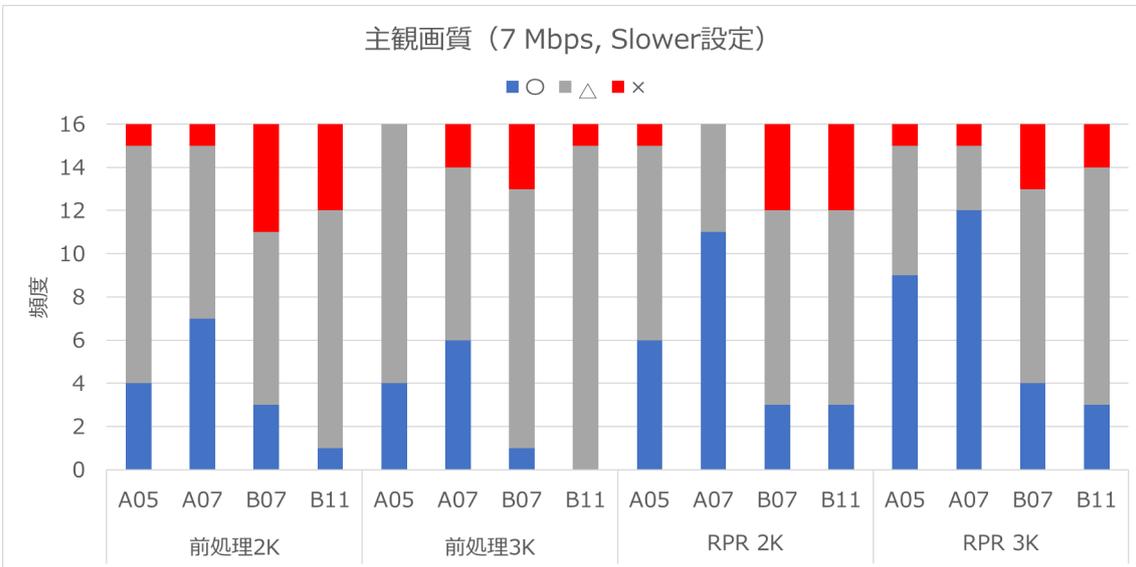


図 7-11 主観画質(7 Mbps, Slower 設定)

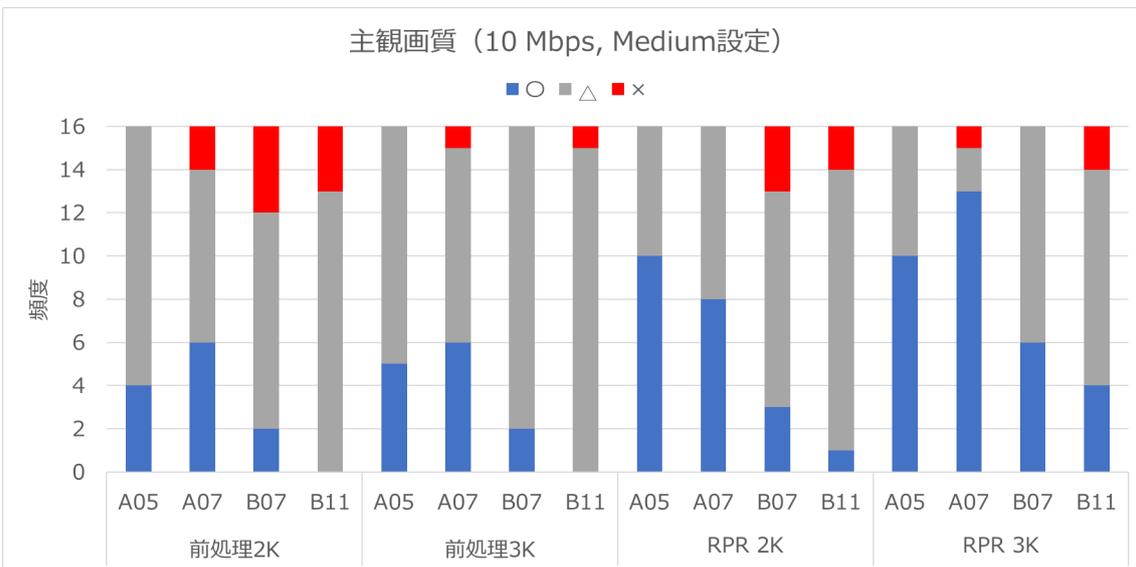


圖 7-12 主觀畫質 (10 Mbps, Medium 設定)

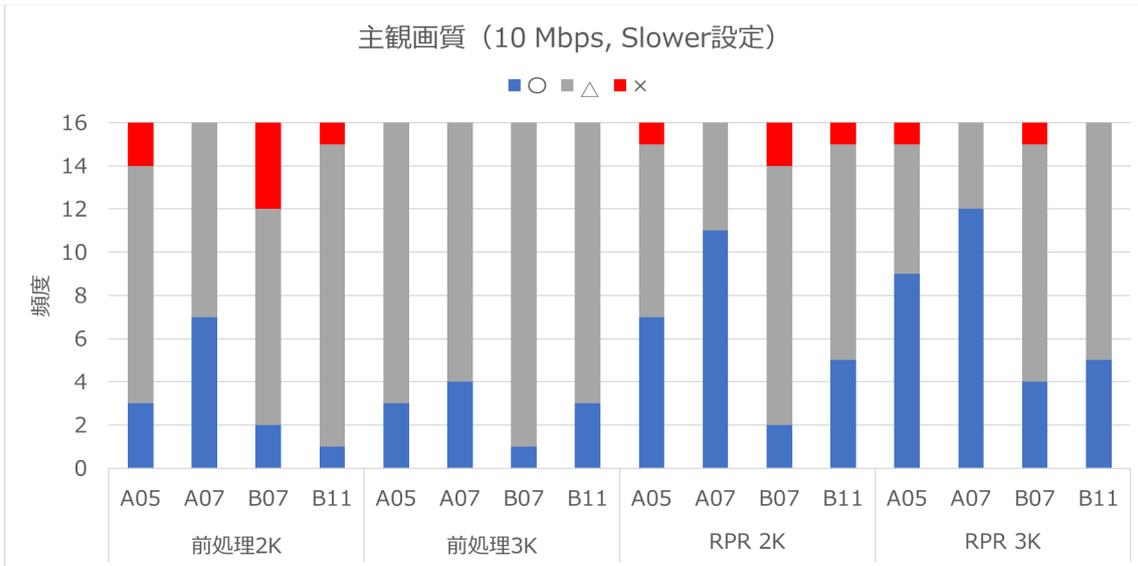


圖 7-13 主觀畫質 (10 Mbps, Slower 設定)

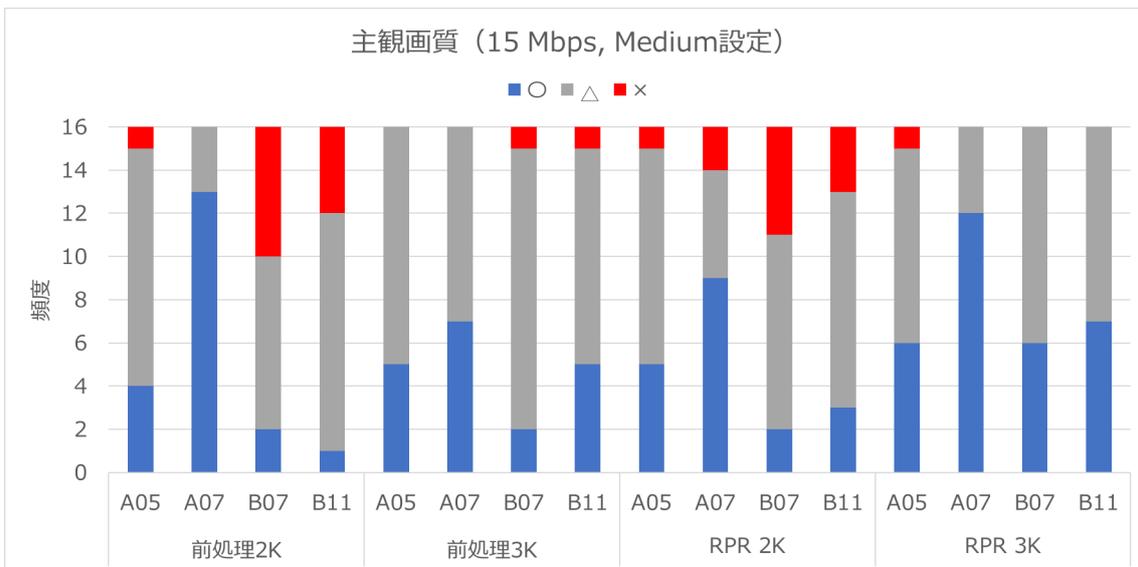


圖 7-14 主觀畫質 (15 Mbps, Medium 設定)

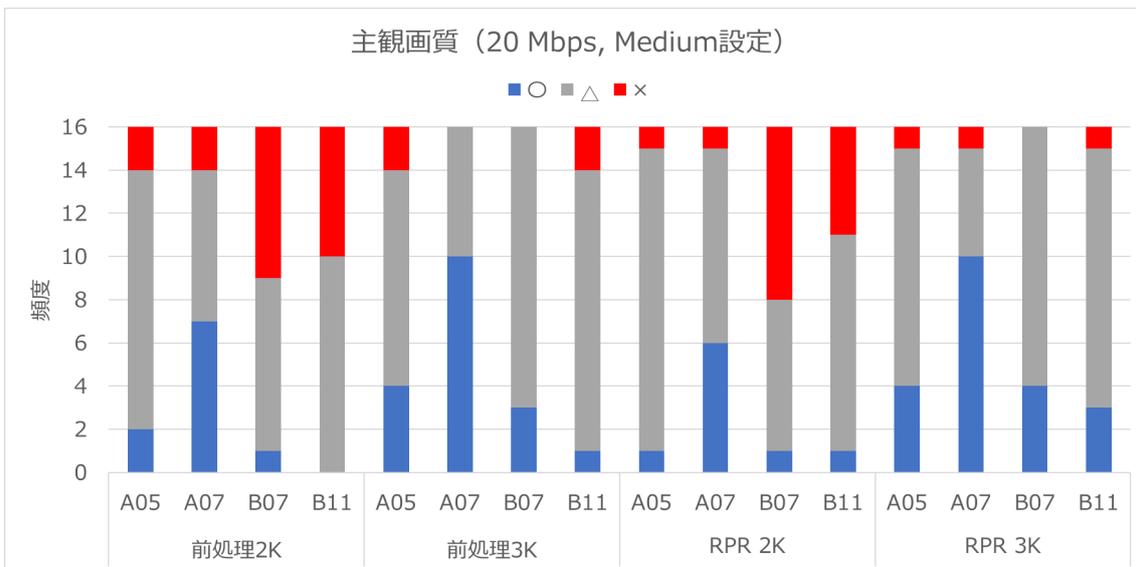


図 7-15 主観画質 (20 Mbps, Medium 設定)

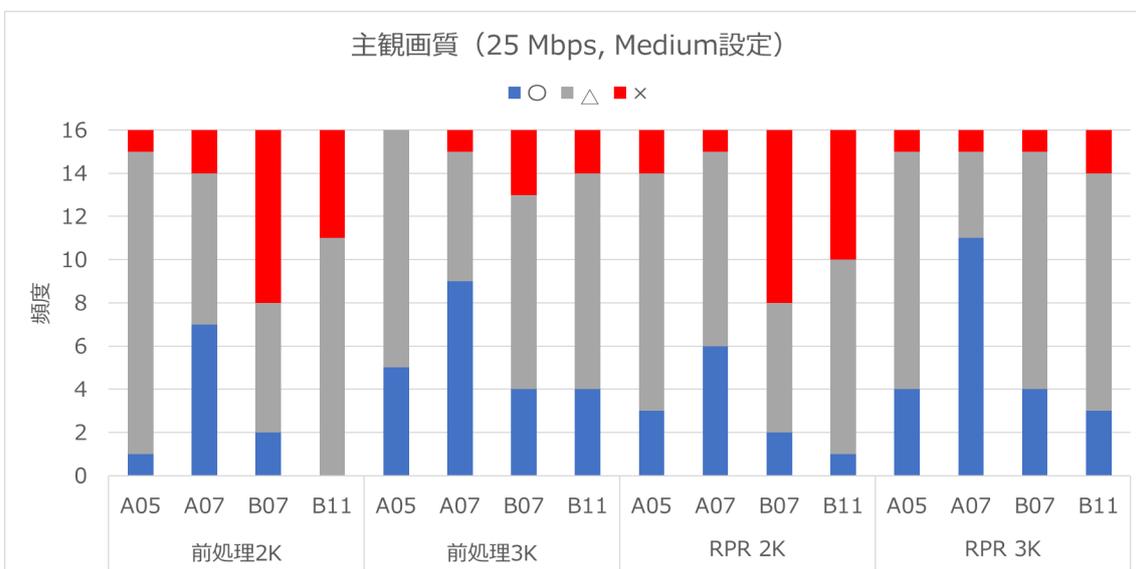


図 7-16 主観画質 (25 Mbps, Medium 設定)

7.4.5. HEVC 30 Mbps 符号化映像との比較

評価 2 での VVC 符号化映像は、事前確認を行い、7.3.3.2 節の通り符号化難易度が高い映像では前処理符号化映像、符号化難易度が低い映像では 4K 符号化映像に絞っている。図 7-17 に、映像毎に HEVC 30 Mbps 符号化映像と比較した画質を示す。符号化難易度が高い映像 (A05, A07, B07, B11) では、15 Mbps で概ね半数以上の評定者が、20 Mbps 及び 25 Mbps では大多数の評定者が、VVC 符号化映像は HEVC 30 Mbps 符号化映像と「同等」または「良い」と回答した。

図 7-18 に、4 つの映像をまとめて HEVC 30 Mbps 符号化映像と総合的な画質を比較した場合の、HEVC 30 Mbps 符号化映像と「同等」以上と評価した回答数を示す。評定者全員が HEVC 30 Mbps 符号化映像と同等以上と評価したビットレートは 20 Mbps であり、15 Mbps では 44% (7 名) が「同等」と回答した。

符号化難易度が低い映像 (A03, A06, B06, B09, B13) では、7 Mbps において全ての映像で半数以上の評定者が VVC 符号化映像は HEVC 30 Mbps 符号化と「同等」または「良い」と回答し、大多数の評定者が「同等」と回答した。

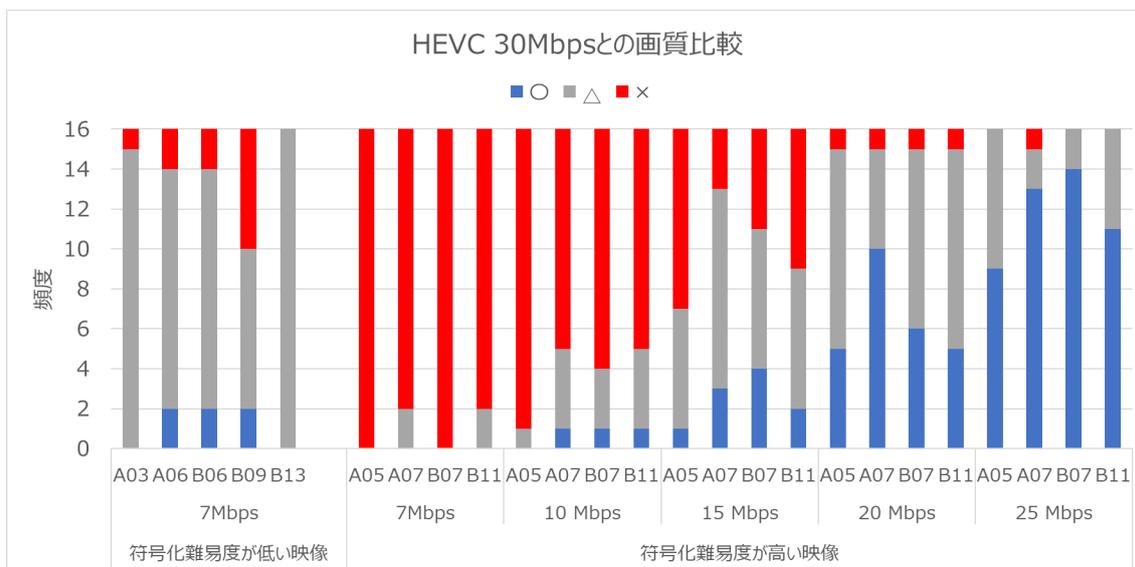


図 7-17 HEVC 30 Mbps 符号化映像との画質比較(映像別)

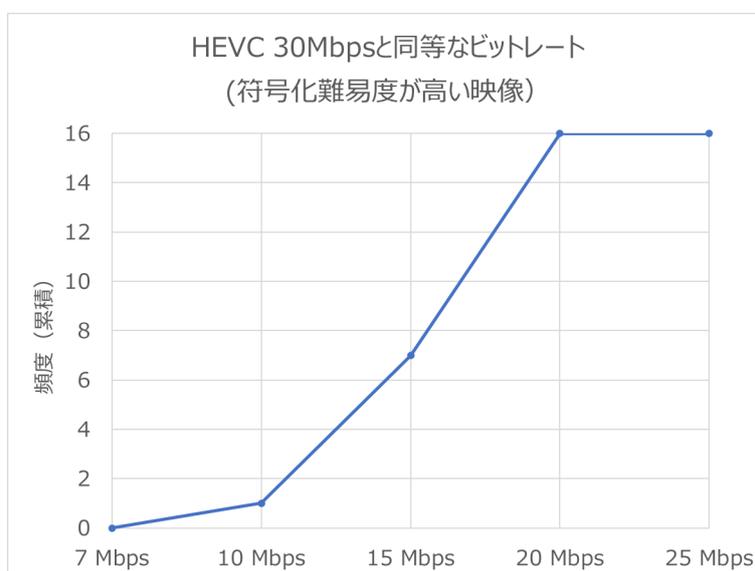


図 7-18 HEVC 30 Mbps 符号化映像と同等以上と評価した回答数 (符号化難易度が高い映像を前処理によって空間解像度を落とした場合)

7.4.6. ヒアリング結果

全評価映像のビューイング後に、評定者にヒアリングを行った。回答の概要は以下の通りである。

- 大半の評定者が、低ビットレート(15 Mbps 未満)では前処理や RPR による画質改善が見られたと回答した。ただし、一つ上のビットレート(例えば 7 Mbps なら 10 Mbps)相当以上まで画質が向上したと評価した評定者はいなかった。
- 一部の評定者からは Slower 設定による画質改善があることが指摘された。ただし、図 7-9 が示すように Medium 設定との主観画質差は僅かであった。
- 放送局の評定者からは、10 Mbps 以下では符号化難易度が高い映像の劣化が大きく、放送品質の観点で疑問視する声が複数あった。前処理や RPR は、画質向上の観点で適用が望ましいとの意見であった。

- 前処理や RPR は絵柄やビットレートによって主観画質への効果が変わり、低ビットレートでは改善を示すものの、高ビットレートでは前処理による劣化が気になることが指摘され、絵柄やビットレートに応じて適切に制御する必要があるとの意見があった。

7.5. 考察

7.5.1. 各ビットレートでの画質

上記の結果を基に、表 7-6 に各ビットレートでの画質をまとめる。

表 7-6 ビットレート毎の画質

ビットレート	画質	画質改善の可能性
7 Mbps, 10 Mbps	符号化難易度が低い映像の場合、4K 放送品質相当の画質となる。 符号化難易度が高い映像の場合、4K 放送品質未満の画質となる。	前処理や RPR で画質改善を確認した。 (映像とビットレートの組み合わせによって効果が異なる)
15 Mbps	符号化難易度が高い一部の映像の場合、4K 放送品質未満の画質となる。	
20 Mbps, 25 Mbps	4K 放送品質相当もしくは高い画質となる。	

注 1: 「4K 放送品質」とは、リアルタイム HEVC ハードウェアエンコーダ (2021 年現在) の 30 Mbps の画質である (7.3.3.2 節参照)。

注 2: 「4K 放送品質」との比較では、符号化難易度が高い映像の VVC 符号化において、前処理を適用している。

注 3: VVC 符号化には VVenC (1.2.0 版) を使用 (7.2.3 節参照) しており、所要ビットレート導出 (1 章から 6 章) で用いた、VVC エンコーダエミュレータ (2.3 節参照) ではない。

7.5.2. 画質改善手法の効果

入力映像への前処理 (ローパスフィルタによる画面内高周波成分の抑制) 及び RPR (符号化画素数の削減) は共に、符号化難易度が高い映像を低ビットレートで符号化する場合に主観画質の改善効果がある。両手法共、映像とビットレートの組み合わせによって効果が異なるが、これらを適切に行うことで主観画質が改善されることが確認された。

両手法は、エンコーダの符号化制御の観点でも効果がある。図 7-8 に示すように、7 Mbps では 2K 相当まで解像度を落とすことで輝度 PSNR が 1 dB–2 dB 向上している。これは同一レートでの条件下で平均量子化パラメータが小さくなった (即ち符号化しやすくなった) ことを意味し、各ブロックの量子化パラメータ制御の自由度が増え、エンコーダの符号化制御破綻 (符号化バッファのオーバーフロー) を抑えやすくなったと言える。本実験で用いた映像よりも符号化難易度が高い映像が入力された際にも、絵柄に応じて適切に制御すれば、低ビットレートでの符号化破綻を抑制できることが期待される。

また、RPR は前処理と比べて画面内ブロック数が少なくなり、ブロックヘッダのオーバーヘッドが下がるため、低ビットレートでの符号化制御破綻の回避はより容易になると考えられる。

8. 付録

8.1. 評価映像の選定

8.1.1. 選定基準

高度広帯域衛星デジタル放送の映像符号化方式検討(以下、HEVC 実験)の時と同様に、多様な符号化難易度や絵柄を含む評価映像を選択した。符号化難易度は、輝度信号の PSNR とする。

8.1.2. 符号化難易度の導出方法

VVenC で符号化した際の符号化難易度とした。符号化条件を表 8-1 に示す。

表 8-1 符号化条件

項目	内容
エンコーダ	VVenC version 0.2.1.0
ツール設定	Medium 設定、MCTF は無効化
符号化パラメータ	Random access GOP 長 8 フレーム IRAP 間隔 32 フレーム
ビットレート	1080/60/P: 8 Mbps 2160/60/P: 20 Mbps
処理フレーム数	900 フレーム

8.1.3. 符号化難易度分布と選定した評価映像

映像情報メディア学会標準動画像における符号化難易度の分布と、選定した映像を映像フォーマット毎に示す。符号化難易度が高いシーンを中心に、中間のシーンも含めて選定している。図右側は、HEVC 実験時の 4320/60/P 評価映像を各解像度にダウンサンプルした後に、同一条件で符号化した場合の符号化難易度である(2160/60/P, HDR は除く)。

8.1.3.1. 1080/60/P, SDR

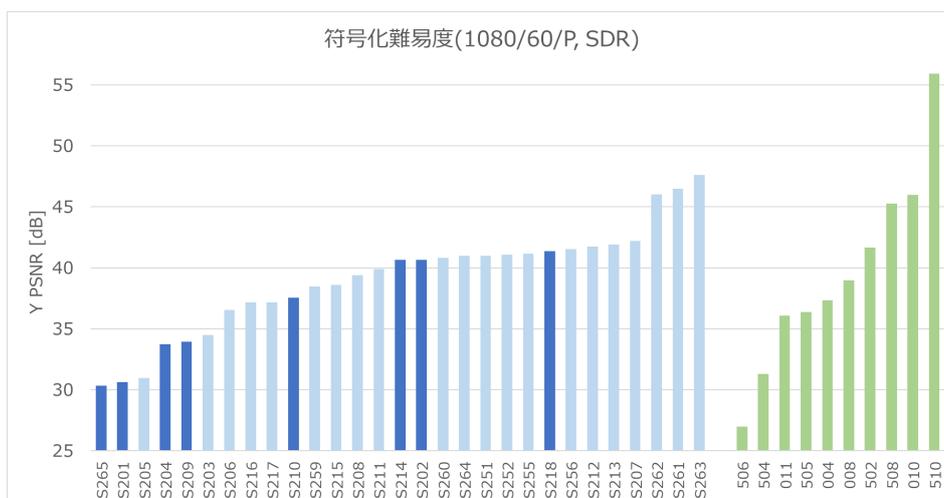


図 8-1 1080/60/P 映像の符号化難易度、選定した評価映像(濃青線で表示)

8.1.3.2. 2160/60/P, SDR

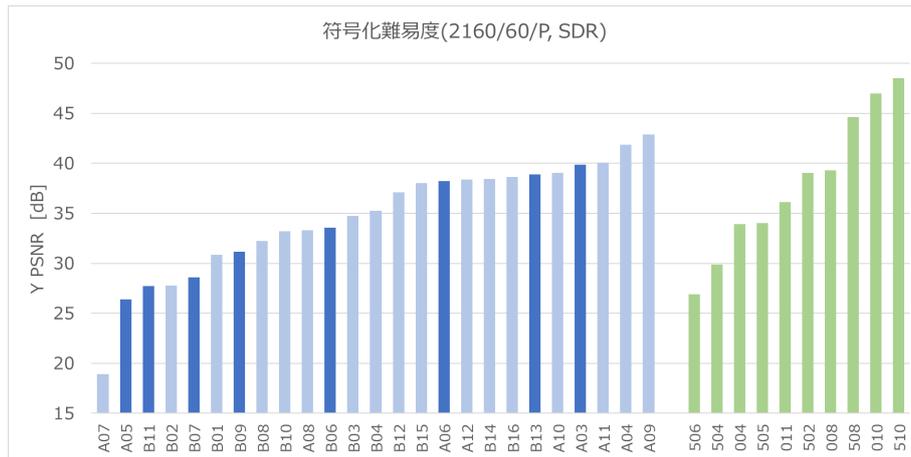


図 8-2 2160/60/P, SDR 映像の符号化難易度、選択した評価映像(濃青線で表示)

8.1.3.3. 2160/60/P, HDR

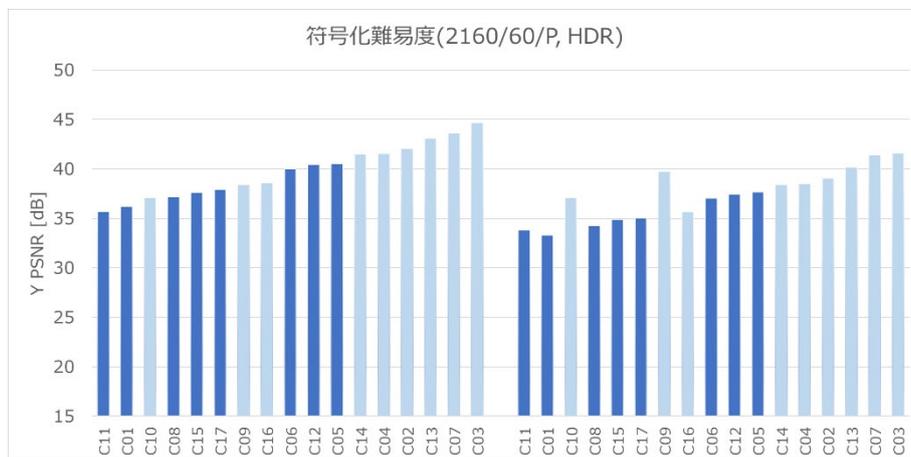


図 8-3 2160/60/P, HDR 映像の符号化難易度、選択した評価映像(濃青線で表示)
右側は SDR 変換した場合の符号化難易度

8.2. 前処理及び符号化解像度変更で適用したフィルタ

符号化解像度変更で用いる縮小・拡大フィルタとして、Lanczos-3 フィルタを用いた。一次元フィルタであり、画像に対して水平方向、垂直方向の順に適用する。

前処理フィルタは、フィルタ A(2K 相当)は表 8-3 の C1、フィルタ B(3K 相当)は表 8-2 の C1 とした。

8.2.1. 3/4 縮小フィルタ

縮小後の画素位置によって、適用するフィルタ C1, C2, C3 を切り替えて適用する。他のフィルタも同様である。

表 8-2 3/4 縮小フィルタ係数(分母 256)

	[0]	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]
C1	0	5	-25	51	194	51	-25	5	0	0
C2	0	4	-12	-1	170	119	-25	0	1	0
C3	0	1	0	-25	119	170	-1	-12	4	0

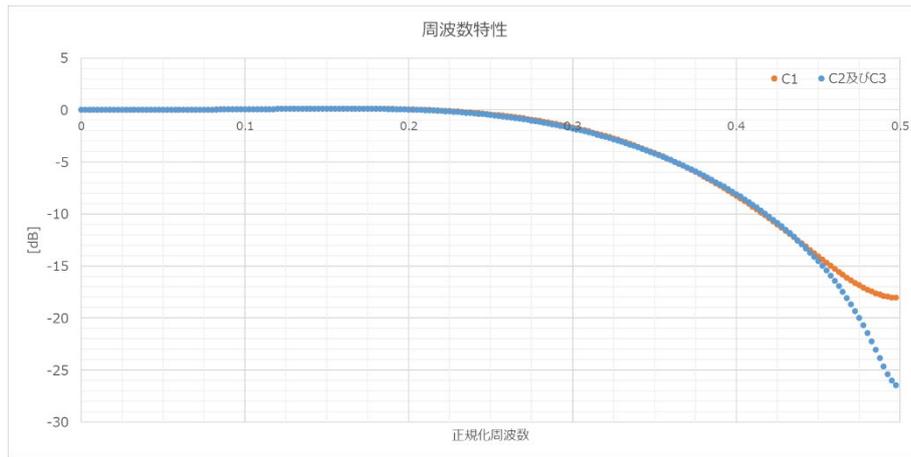


図 8-4 3/4 縮小フィルタの周波数特性

8.2.2. 1/2 縮小フィルタ

表 8-3 1/2 縮小フィルタ係数(分母 256)

	[0]	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]	[11]	[12]	[13]
C1	0	3	0	-17	0	78	128	78	0	-17	0	3	0	0

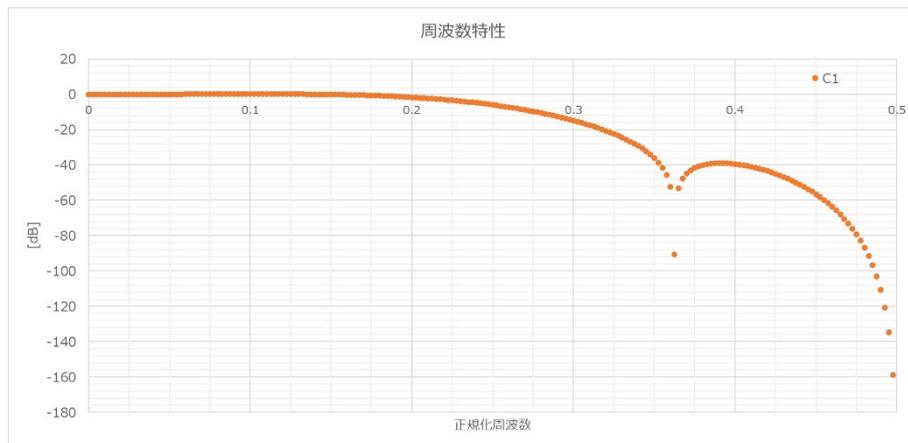


図 8-5 1/2 縮小フィルタの周波数特性

8.2.3. 4/3 拡大フィルタ

表 8-4 4/3 拡大フィルタ係数(分母 256)

	[0]	[1]	[2]	[3]	[4]	[5]
C1	0	0	256	0	0	0
C2	2	-18	74	225	-34	7
C3	6	-34	156	156	-34	6
C4	7	-34	225	74	-18	2

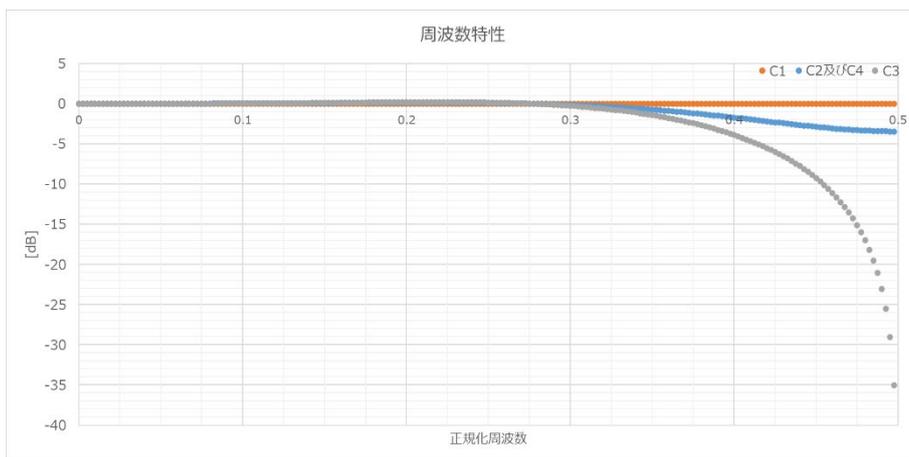


図 8-6 4/3 拡大フィルタの周波数特性

8.2.4. 2/1 拡大フィルタ

表 8-5 2/1 拡大フィルタ係数(分母 256)

	[0]	[1]	[2]	[3]	[4]	[5]
C0	0	0	256	0	0	0
C1	6	-34	156	156	-34	6

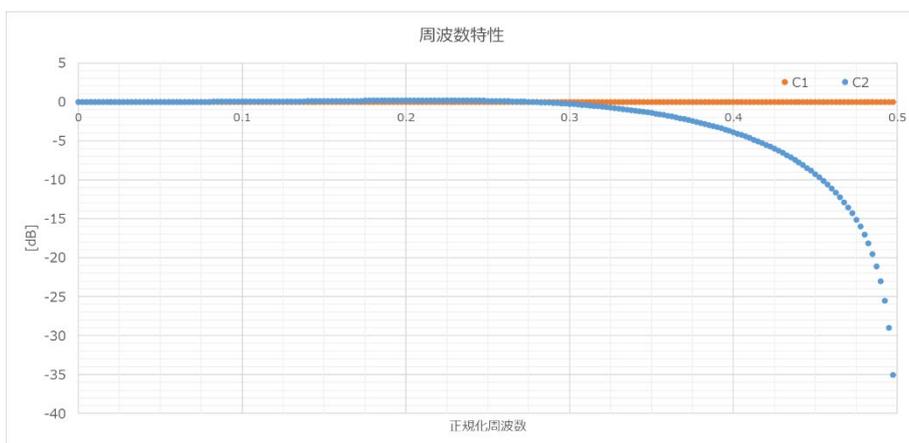


図 8-7 2/1 拡大フィルタの周波数特性