

8K 120Hz 映像符号化における所要ビットレートの検討

A Study of Required Bit-rate on 8K 120Hz Video Coding

杉藤 泰子† 神田 菊文† 境田 慎一†
Yasuko Sugito Kikufumi Kanda Shinichi Sakaida

1. はじめに

8K スーパーハイビジョン (以下, 8K) は, 臨場感の高い 8K 映像と立体音響を提供する次世代の放送方式である。日本では 2016 年に放送衛星による 8K 試験放送が開始され, 2018 年に 8K 実用放送が開始される予定である[1]。

8K の映像フォーマットは ITU-R 勧告 BT.2020[2]に規定されている。空間解像度は 7,680×4,320, 量子化ビット数は 10bit または 12bit, 走査方式はプログレッシブ, フレーム周波数は 60Hz ほか, 最大 120Hz となっている。フレーム周波数が高い 120Hz 映像は動きが速い被写体を鮮明かつなめらかに表示できることが特徴で, スポーツなどのコンテンツに適している。

8K 試験放送は当面の間, フレーム周波数 59.94Hz (以下, 60Hz とする) で実施される予定である[3]。筆者らはこれまでに 8K 放送規格に準拠したフレーム周波数 60Hz の映像符号化・復号装置を開発した[4]。今後 8K 120Hz 放送を目指して映像符号化・復号装置の開発を進める予定である。8K 120Hz 映像は非圧縮でのデータ量が約 144Gbps (8K/4:4:4/12bit/120Hz)と膨大であり, 8K 120Hz 放送を実現するためには, 映像を高効率かつリアルタイムに圧縮する符号化装置の開発が必要である。

本稿では, 8K 120Hz 映像符号化における所要ビットレートの検討について述べる。はじめに 8K 放送の映像符号化方式を紹介し, この方式に基づいて実施した 8K 120Hz 画像の符号化実験と実験結果を示す。さらに符号化実験で得られた画像の画質評価について述べ, これらの結果をもとに所要ビットレートについての考察を述べる。

2. 8K 放送の映像符号化方式

8K 放送の映像符号化方式は, 2014 年 7 月に ARIB (Association of Radio Industries and Businesses, 電波産業会) によって策定された ARIB STD-B32 3.0 版[5] (以下, ARIB 規格) に規定されている。以下, ARIB 規格の主な特徴を紹介する。

2.1. HEVC/H.265 方式の採用

ARIB 規格では, 8K 放送の映像符号化方式に HEVC (High Efficiency Video Coding)/H.265[6] (以下, HEVC) を採用している。HEVC は 2013 年に ISO/IEC と ITU-T によって共同で国際標準化された映像符号化方式である。特に 8K/4K 映像といった超高精細映像において, 従来の符号化方式に比べて符号化効率が良いことが知られている[7]。

2.2. 120/60Hz 時間階層符号化の採用

先述のように 2016 年に開始する 8K 試験放送ではフレー

† NHK 放送技術研究所, NHK STRL

ム周波数 60Hz を予定しており, 受信機には 8K 60Hz 映像用の復号装置が組み込まれる。ARIB 規格では, 8K 120Hz 放送が行われた際に, 8K 60Hz 放送用の受信機では 60Hz 映像を再生できる時間階層符号化を採用している。これは HEVC で規定のフレーム参照構造によって実現される。

図 1 に 120/60Hz 映像の時間階層符号化の例を示す。図において, 長方形はフレームを, 長方形内の数字はフレーム番号を示す。

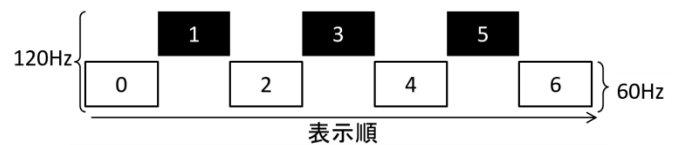


図 1 120/60Hz 映像の時間階層符号化

HEVC の時間階層符号化では, 120Hz 映像復号時にはすべての番号のフレームを再生し, 60Hz 映像復号時には偶数番号のフレームのみを再生する。本実験で使用したフレーム参照構造の詳細については 3.3.1 で述べる。

3. 符号化実験

ARIB 規格に基づいて 8K 120Hz 放送を想定した符号化実験を行った。

3.1. 本実験の検討項目

本実験では, 以下の 3 点を検討した。

1. 8K 120Hz 時間階層符号化において, 所要ビットレートは 8K 60Hz 映像と同等で問題ないか。

8K 120Hz 放送では, 8K 60Hz で実施される試験放送と同じ伝送路を使用することが想定される。8K 120Hz 映像を 8K 60Hz 映像と同等のビットレートで時間階層符号化した際に, 画質が良好であるか検討することとした。

2. 8K 120/60Hz 時間階層符号化において, 符号量をどのように配分すれば適切であるか。

8K 120/60Hz 時間階層符号化では, 60Hz 映像部分 (図 1 の偶数番号フレーム) と, 60Hz 映像と 120Hz 映像との差分 (図 1 の奇数番号フレーム) に対する適切な符号量割当を検討する必要がある。8K 120Hz 映像のビットレートを上記 1. と同じ値に設定し, 8K 60Hz 映像に対するビットレートを変化させて, 120Hz 映像の符号量は 60Hz 映像の何%増であるか検討することとした。

3. 8K 120/60Hz 時間階層符号化において, 8K 60Hz 映像の画質は問題ないか。

時間階層符号化では 8K 120Hz および 60Hz 復号映像の双方が再生されるため, 60Hz 映像の品質も求められる。上記 2. の条件で符号化した際に, 60Hz 映像の画質が良好であるか検討することとした。

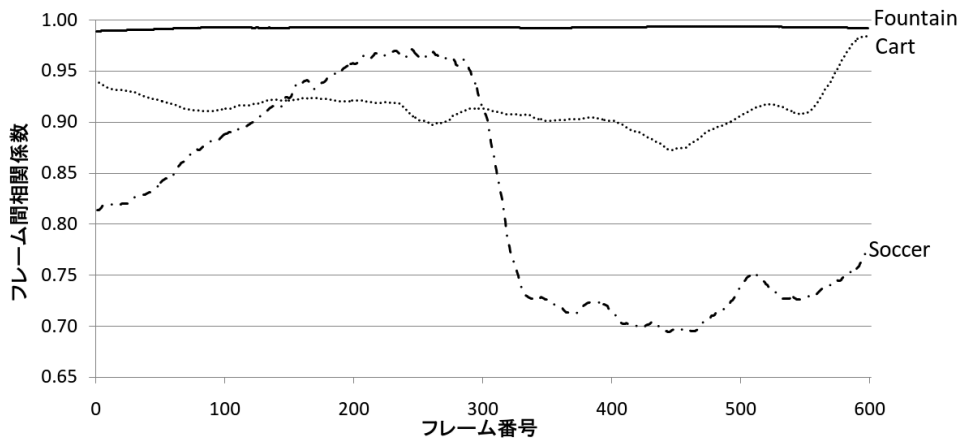


図 2 実験画像のフレーム間相関係数

3.2. 実験画像

実験には、性質が異なる 8K 120Hz 画像 3 種類、600 フレーム(5 秒分)を使用した。表 1 にこれらの画像の説明を示す。

表 1 実験画像の説明

画像名	説明
Fountain	固定撮影による噴水ショーの様子。水しぶきの動き、建物のオクルージョン、背景に平坦な空の部分を含む。
Cart	複数台のカートがサーキットを走る様子。カートの動きに合わせた左右方向のパンを含む。
Soccer	サッカーの試合の様子。後半でロングパスのボールをフォローパンする動きの大きい部分を含む。

画像の性質を調査するために、連続する 2 フレームの輝度値の相関係数 r を式(1)を用いて計算した。

$$r = \frac{\sum_{x,y} (f_1(x,y) - \bar{f}_1)(f_2(x,y) - \bar{f}_2)}{\sqrt{\sum_{x,y} (f_1(x,y) - \bar{f}_1)^2} \sqrt{\sum_{x,y} (f_2(x,y) - \bar{f}_2)^2}} \quad (1)$$

式(1)において、 f_1 、 f_2 は連続するフレームであり、 $f_i(x,y)$ 、 $i = 1, 2$ はフレーム f_i の座標 (x,y) における輝度値を、 \bar{f}_i はフレーム f_i の輝度値の相加平均をあらわす。種々の画像について実験した結果、相関係数 r は自然画像では 0.40~1.00 程度の値をとり、動きが大きい画像ほど小さい値をとることが確認できた。例えば、固定撮影の画像では 0.99 以上の値を、動きが非常に大きいシーンでは 0.70 程度の値を、シーンチェンジでは 0.40 程度の値をとる。

図 2 に本実験で使用した 3 種類の画像について、連続する 2 フレームの輝度値の相関係数を求めた結果を示す。縦軸は相関係数を、横軸はフレーム番号を示す。グラフより、画像 Fountain は固定撮影であるために動きが極めて小さく 1 に近い値をとり続けること、画像 Soccer は後半部分の動きが大きいため相関係数が 0.70~0.75 程度と小さくなること、画像 Cart は相関係数が 0.90 程度と他の 2 つの画像の中程度の動き量を持つことが確認できる。

3.3. 符号化条件

HEVC の標準化で使用されているソフトウェア HM (HEVC test Model) 16.6 版 (2015 年 6 月発行) を用いて 8K 120Hz 画像の符号化実験を行った。ARIB 規格と同様に Main 10 プロファイル (4:2:0/10 bit) に設定した。以下、本実験の符号化条件の詳細を説明する。

3.3.1. フレーム参照構造

本実験では ARIB 規格に記載された 120Hz 画像用の時間階層符号化フレーム参照構造[5] (図 3) を使用した。

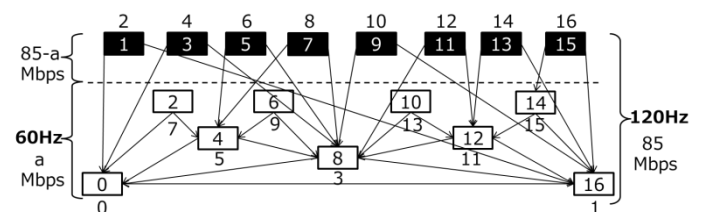


図 3 実験で使用したフレーム参照構造

図 3 において、長方形はフレームを、長方形内の番号はフレーム番号を示す。フレームの上または下の番号は符号化・復号順を示し、フレーム番号 0, 16, 1, 8, 3, 4, 5, 2, 7, 6... の順に並び替えて処理が行われる。復号処理の後には 0 から順にモニタに表示されるように並び替えられる。

HEVC では従来の映像符号化方式と同様に、近傍のフレームの符号化画像を参照して予測をする動き補償予測が使用されている。フレームから出ている矢印の先は、当該フレームを符号化する際の参照先を示す。例えば、フレーム番号 16 は 0 の符号化画像を参照して予測され、さらにフレーム番号 1 は、0 と 16 を参照して予測される。HEVC では、このように参照フレームから予測したフレームをさらに参照フレームとする階層構造をとることができる。一般に参照フレームの画質が高いほど予測精度が向上するため、参照される頻度の高い階層構造の下位ほど圧縮率を低く、下位のフレームを複数参照可能な上位ほど圧縮率を高く設定することでバランスのとれた符号化が可能となり、全体の符号化効率を改善することができる[8]。特に高フレームレート映像はフレーム間の相関が高いため、階層構造を用いることで効率よく圧縮することができる。

図 3 では、偶数番号のフレーム同士を参照する構造とすることで、偶数番号のフレームのみを再生して 60Hz 映像とする時間階層符号化を実現している。本実験ではこの 16 フレーム単位のフレーム参照構造を使用し、動き補償予測を使用しないイントラ画像(図中のフレーム番号 0)を、64 フレームごとに一回挿入した。これは現行デジタル放送が約 0.5 秒に一回であることに合わせた設定である。

3.3.2. ビットレート設定

8K 120Hz 放送では 8K 60Hz 放送と同じ伝送路を使用することが想定される。2015 年 5 月に実施した試験放送と同条件の衛星による伝送実験では、8K 60Hz 映像のビットレートを 85Mbps に設定した[4]。本実験では 8K 120Hz 映像のビットレートを 85Mbps に設定し、HM に機能を追加して 60Hz 映像部分のビットレートを変更できるようにして、符号量配分を検討した。以下では符号量配分を 60Hz 映像を 100%とした増加量で表すこととする。例えば図 3 のように 60Hz 映像を a Mbps に設定した場合、符号量配分は $85/a \times 100$ (%)となる。表 2 に本実験で使用した符号量配分とビットレート設定の関係を示す。

表 2 符号量配分とビットレート設定の関係

符号量配分	60Hz 映像 (Mbps)	60Hz と 120Hz 映像の差分(Mbps)
103%	82.52	2.48
105%	80.95	4.05
108%	78.70	6.30
110%	77.27	7.73
113%	75.22	9.78
115%	73.91	11.09

本実験ではさらに 8K 60Hz 映像を 85Mbps に設定して符号化した。この条件は符号量配分 100%に相当し、8K 120Hz 映像を 85Mbps で時間階層符号化する際の画質の上限となる。

4. 実験結果

図 4~図 9 に 120Hz および 60Hz 映像の実験結果を示す。60Hz 映像の符号化結果は、120Hz 映像の符号化結果の偶数番号フレームのみを集計したものである。図 4~図 5 に画像 Fountain, 図 6~図 7 に画像 Cart, 図 8~図 9 に画像 Soccer のグラフを示す。グラフの縦軸は PSNR (Peak Signal to Noise Ratio), 横軸はビットレートで、左上にいくほど符号化効率が良いことを意味する。グラフの各点は、ラベルで示す符号量配分の条件におけるビットレートと、フレームごとの輝度値の PSNR を平均した値を示す。

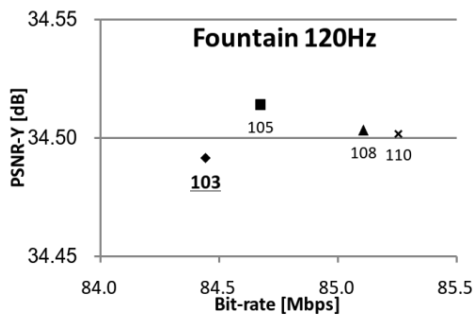


図 4 画像 Fountain 120Hz 映像符号化結果

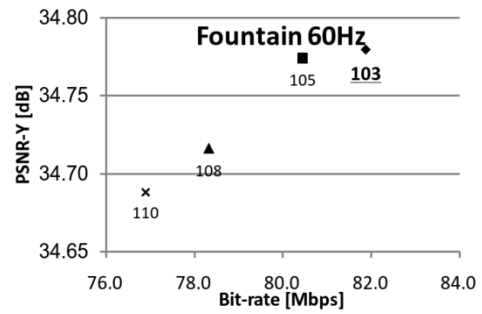


図 5 画像 Fountain 60Hz 映像符号化結果

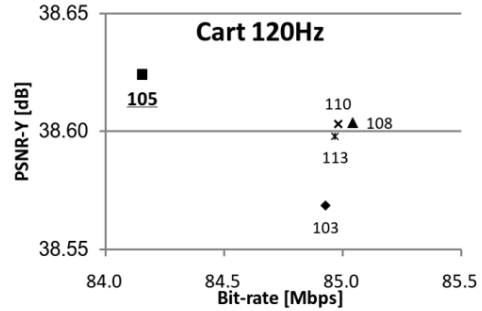


図 6 画像 Cart 120Hz 映像符号化結果

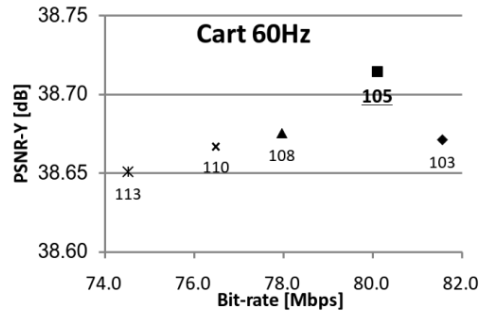


図 7 画像 Cart 60Hz 映像符号化結果

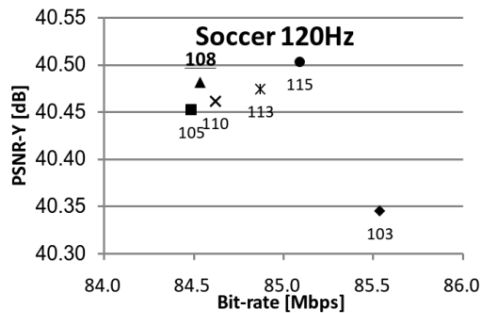


図 8 画像 Soccer 120Hz 映像符号化結果

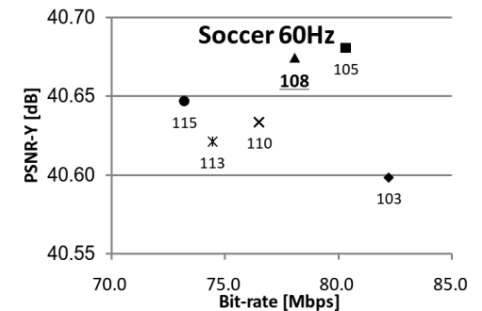


図 9 画像 Soccer 60Hz 映像符号化結果

実験結果から、画像に応じた最適な符号量配分があることが確認できた。図 4～図 9 のグラフで太字のラベルに下線を引いた点は、120Hz 映像の結果で符号化効率が最も良い最適な符号量配分を示す。画像 Fountain については、85Mbps での PSNR を比例で求めた結果から決定した。表 3 に画像ごとの最適な符号量配分とフレーム間相関係数の平均および標準偏差を示す。

表 3 最適な符号量配分とフレーム間相関係数

画像	最適な符号量配分	フレーム間相関係数 (平均, 標準偏差)	
		平均	標準偏差
Fountain	103%	0.9923	0.0009
Cart	105%	0.9127	0.0191
Soccer	108%	0.8210	0.0998

この結果より、フレーム相関が低く動きが大きい画像ほど符号量配分を高く、すなわち 60Hz 映像と 120Hz 映像の

差分に対して符号量を多くすると符号化効率が向上することが確認できた。

5. 画質評価

符号化実験で得られた画像の画質評価を行った。

5.1. 客観評価

符号化実験結果から各画像のフレームごとに輝度値の PSNR を求めた。図 10～図 12 に 120Hz 映像と 60Hz 映像を 85Mbps に設定して符号化した際のグラフを示す。グラフの縦軸は PSNR、横軸はフレーム番号である。120Hz 映像（実線）は最適な符号量配分の条件で符号化した。60Hz 映像（破線）はフレーム番号を 2 倍にした点にプロットしており、120Hz 時間階層符号化の 60Hz 映像部分と一致する。

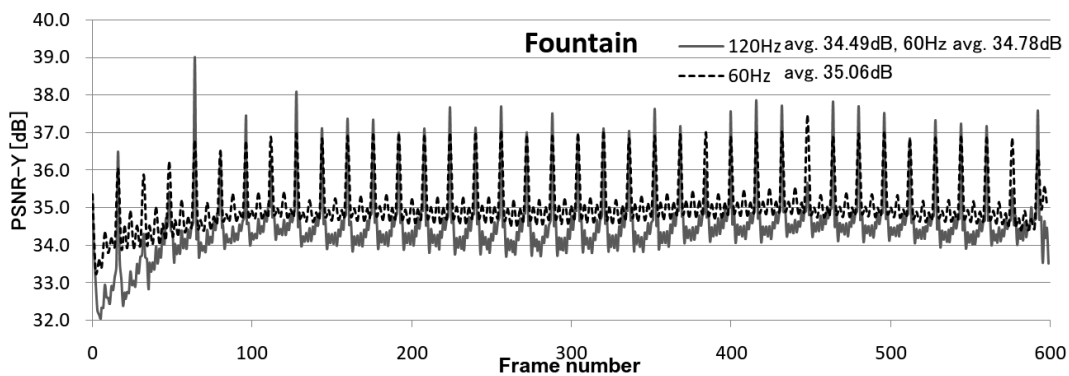


図 10 フレームごとの PSNR (画像 Fountain)

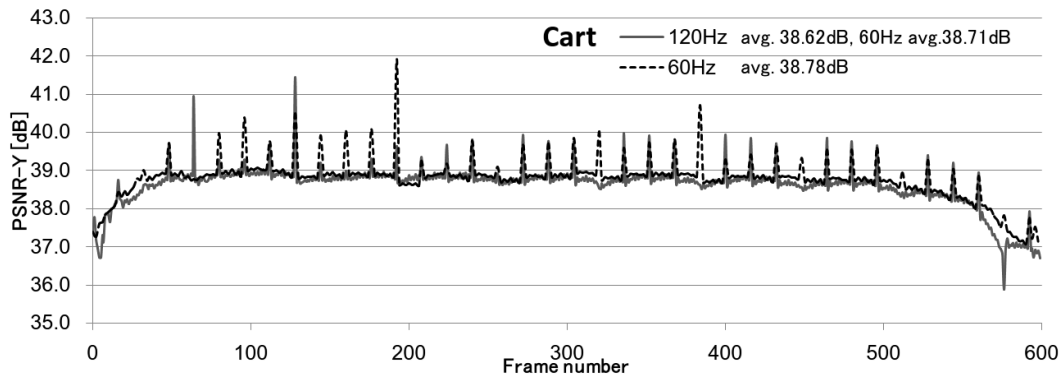


図 11 フレームごとの PSNR (画像 Cart)

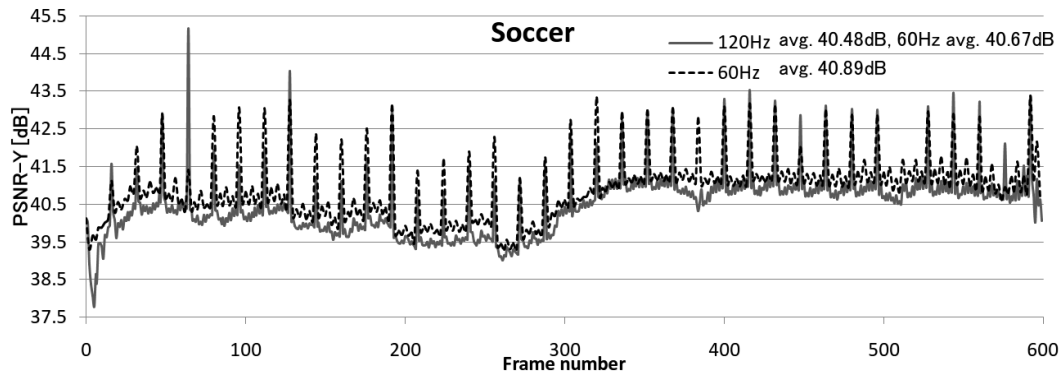


図 12 フレームごとの PSNR (画像 Soccer)

120Hz 映像を 85Mbps に設定して符号化した場合は、60Hz 映像を 85Mbps に設定して符号化した場合よりも PSNR の低下が見られる。グラフより、PSNR の低下が最も大きい画像 Fountain でもフレーム単位で最大約 1.0dB、平均約 0.6dB となっていることが確認できた。

5.2. 画質確認

符号化実験を行った条件のうち表 4 で示す条件について、8K 120Hz 映像と 8K 60Hz 映像の簡易な主観評価を行った。8K 60Hz 映像については、符号量配分 100% (8K 60Hz 映像を 85Mbps に設定して符号化した条件) も確認した。

表 4 画質確認を行った条件

画像	最適条件	比較条件	比較条件の選定理由
Fountain	103%	110%	符号量配分の差が大 符号化効率が最低
Cart	105%	113%	符号量配分の差が大
Soccer	108%	103%	符号化効率が最低

図 13 は画質確認の様子の写真である。評価には 120/60Hz 映像ともに 85 インチ (約 1.8m×1.05m) の 8K LCD モニタを使用した。



図 13 画質確認の様子

画質確認は映像符号化専門家 8 名で行った。8K の視距離 0.75H (画面の高さの 0.75 倍) に基づいて画面から約 0.8m の位置で着座して観視したほか、画面の任意の部分を任意の距離および姿勢で観視して画質を確認した。

5.2.1. 8K 120Hz 映像

各画像について、グレイ画像 (2 秒間)、原画像、最適条件、比較条件の順に繰り返し提示した。

その結果、いずれの条件とも大幅な画質の劣化は見られないことを確認した。符号量配分の違いが視認できるのは画像 Fountain のみだった。画像 Fountain では 60Hz 映像に対するビットレートが少ない比較条件で、静止している背景の空の部分の劣化 (フリッカーやグラデーション部分のブロック歪) が目立った。これ以外の画像 Cart, Soccer では、静止画を切り出した際には違いがあるものの、動画では動きが速いために符号量配分の違いを視認することができなかった。

図 14 および図 15 は、画像 Cart を符号量配分 105% および 113% で符号化した奇数番号フレーム (60Hz 映像と 120Hz 映像との差分) の 100×137 画素部分を切り出したものである。符号量配分 105% では、60Hz 映像と 120Hz 映像との差分に対する符号量が少ないために予測残差信号を

十分に送ることができず、ドライバーのヘルメットの輪郭 (図中の破線部分) にブロック状の歪が見られる。

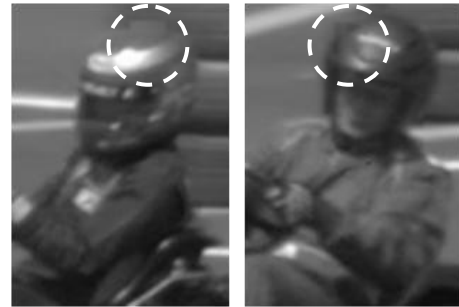


図 14 画像 Cart (符号量配分 105%)



図 15 画像 Cart (符号量配分 113%)

5.2.2. 8K 60Hz 映像

各画像について、グレイ画像 (2 秒間)、原画像 60Hz、符号量配分 100%、最適条件 60Hz、比較条件 60Hz の順に繰り返し提示した。実験に使用した 8K 60Hz 符号化映像はそれぞれの条件でビットレートが異なる。表 5 に実験に使用した符号化画像のビットレートを示す。

表 5 実験画像のビットレート

画像	符号量配分 100% (Mbps)	最適条件 60Hz (Mbps)	比較条件 60Hz (Mbps)
Fountain	85.07	81.87	76.89
Cart	85.01	80.11	74.51
Soccer	85.00	78.06	82.21

画質確認の結果、いずれの条件とも大幅な画質の劣化は見られないことを確認した。ビットレートによる違いが視認できたのは、8K 120Hz 映像と同様に画像 Fountain のみだった。この他の画像 Cart, Soccer については、ビットレートの違いを視認することができなかった。

6. 考察

符号化実験結果と画質評価結果から、3.1 で述べた本実験の 3 点の検討項目について考察する。

1. 8K 120Hz 時間階層符号化において、所要ビットレートは 8K 60Hz 映像と同等で問題ないか。

今回の実験では、8K 60Hz 映像と同等のビットレートで時間階層符号化した 8K 120Hz 映像は、画質評価の結果、大幅な品質低下は見られず、視覚的に認識できる大幅な劣化もなかった。

したがって、今回の実験画像および実験条件においては、8K 120Hz 時間階層符号化の所要ビットレートは 8K 60Hz 映像とほぼ同等で問題ないことが確認できた。

2. 8K 120/60Hz 時間階層符号化において、符号量をどのように配分すれば適切であるか。

今回の実験では、8K 60Hz 映像と同等のビットレートに設定した 8K 120Hz 時間階層符号化において、60Hz 映像に対する 120Hz 映像の符号量増加は最大 10% 程度であることが確認できた。

さらに、画像によって最適な符号量配分が異なること、動きが大きい画像ほど 60Hz 映像と 120Hz 映像の差分に対する符号量を大きくすることが適切であることが確認できた。

本実験において、画像に応じた最適な符号量配分が存在する理由について考察する。符号量配分が小さすぎる場合は、奇数番号フレーム (60Hz 映像と 120Hz 映像との差分) に対する符号量が足りないために予測残差信号を十分に送ることができず、例えば図 14 のように、動きが大きい画像で品質が低下する可能性がある。反対に符号量配分が大きすぎる場合は、60Hz 映像部分に対する符号量が足りないために画質が十分でないことが考えられる。この場合、60Hz 映像部分を参照して符号化する奇数番号フレームの符号量を多くしても、予測残差が大きいため割り当てられた符号量の予測残差信号では十分な画質が得られない可能性がある。また、動きが小さい画像では、奇数番号フレームに割り当てられた符号量で予測残差信号による画質向上が十分である一方、60Hz 映像部分の画質が十分ではなく、画像全体の品質が低下することが考えられる。フレームごとの PSNR を調査した結果から、画像 Fountain の比較条件では後者の理由によって劣化が見られたものと推察される。

したがって、今回の実験画像および実験条件において、時間階層符号化による 8K 120Hz 映像のビットレートは、8K 60Hz 映像に対するビットレートが十分であれば、8K 60Hz 映像の最大 10% 増程度であることが確認できた。

3. 8K 120/60Hz 時間階層符号化において、8K 60Hz 映像の画質は問題ないか。

今回の実験では、8K 120Hz 映像の時間階層符号化から得られた 8K 60Hz 映像は、画質評価の結果、8K 60Hz の原画像と比べて大幅な品質低下は見られなかった。

したがって、今回の実験画像および実験条件では、8K 120/60Hz 時間階層符号化における 8K 60Hz 映像の画質に問題がないことが確認できた。

7. まとめ

8K 120Hz 放送の実現に向けて、8K 120Hz 映像の時間階層符号化における所要ビットレートを検討した。その結果、今回の実験では 8K 60Hz 映像と同等のビットレートで大幅な画質低下が見られないことを確認した。

今後、その他の画像も使用して同様の実験を行い、所要ビットレートの検討を進めるとともに、120Hz 映像符号化に適したパラメータの検討を進めていく予定である。併せて 8K 120Hz 映像符号化・復号装置の開発を行う予定である。

参考文献

- [1] 総務省, "4K・8K ロードマップに関するフォローアップ会合 第二次中間報告", http://www.soumu.go.jp/main_content/000370906.pdf (2015).
- [2] Rec. ITU-R BT.2020-2 Parameter values for ultra-high definition television systems for production and international programme exchange (2015).
- [3] 次世代放送推進フォーラム, "BS/C S110 4K・8K 放送技術仕様の特徴", http://www.soumu.go.jp/main_content/000392424.pdf (2015).
- [4] Y. Sugito, K. Iguchi, A. Ichigaya, K. Chida, S. Sakaida, H. Sakate, Y. Matsuda, Y. Kawahata, N. Motoyama, "HEVC/H.265 codec system and transmission experiments aimed at 8K broadcasting", The Best of IET and IBC 2015-2016, vol.7, p.24-29 (2015).
- [5] ARIB STD-B32 3.0 版, "デジタル放送における映像符号化, 音声符号化及び多重化方式" (2014).
- [6] ISO/IEC 23008-2:2013 - Information technology -- High efficiency coding and media delivery in heterogeneous environments -- Part 2: High efficiency video coding | Rec. ITU-T H.265 (2013).
- [7] Y. Sugito, K. Iguchi, A. Ichigaya, K. Chida, S. Sakaida, Y. Shishikui, H. Sakate, T. Itsui, N. Motoyama, S. Sekiguchi, "Development of the super hi-vision HEVC/H.265 realtime encoder", SMPTE Conf. Proc., 2013, Vol. 2013, No. 10, pp. 1-16 (2013).
- [8] 大久保榮, "H.265/HEVC 教科書", インプレスジャパン (2013).