

Video Coding for Machine によるサイバーフィジカル空間の効率的なデータ伝送

代表研究者	岩崎 裕江	東京農工大学 大学院工学研究院	教授
共同研究者	丹羽 直也	東京農工大学 大学院工学研究院	助教
共同研究者	新田 高庸	会津大学 コンピュータ理工学部	教授

1 はじめに

日本が目指すべき未来像の姿として提唱されている「Society5.0」(超スマート社会)では、すべての人とモノがつながり、様々な知識や情報が共有され、今までにない価値を産み出す。このような世界観では、IoT 機器が増加し、サイバー空間とフィジカル空間での通信トラフィックが増大すると予想され、近年、固定網・携帯網の両方でトラフィックの大幅な増大が確認されている。また、これまでは映像伝送は、遠隔にあるシーンを人が観るためのものが中心であったが、IoT 機器などの増加により、遠隔にある物を機械が観るためのものに変革し、IoT 機器の増加により、これからも、機械が観るための映像は増加の傾向になることには疑いの余地はない。そこで、本研究調査では、機械が観るための映像を効率よく伝送するための符号化方式について示す。

2 通信トラフィックの解析

2-1 通信トラフィック解析実験環境

アプリケーションが期待するモバイル環境などでの帯域を明確にするために、複数チャネルを同時に伝送する実験のトラフィックを解析する。本実験では、リアルな環境での伝送実験とするため、リアルなイベントにおけるトラフィックを測定した。実験環境を図1に示す。ここでは、コンテンツ保護などにも有効な閉空間を提供できるモバイル SINET を用いた実験環境を構築した。SINET (学術情報ネットワーク)とは国立情報学研究所(NII)が構築および運用しているネットワークである。セキュリティの頑強さと高速なネットワーク回線によって、大学等の研究機関に対して教育・研究に関する多くのコミュニティ形成を支援している。2022年から運用が開始した SINET6 は3G回線や4G回線に加えて5G回線が利用可能となった SINET であり、世界最高基準である400Gbps回線のネットワークを構成することで、約1000機関に及ぶ大学や研究機関等にハイレベルな学術情報基盤を提供している。また、SINET6は従来の SINET5 よりも回線の通信速度が向上しているため、膨大な情報量を扱う映像伝送に適しており、移動式実験施設による遠隔病理診断のための手術映像の伝送や、ドローンからの高精細映像伝送など幅広い用途で活用されている[1][2]。加えて、本研究で用いるモバイル SINET はネットワークを既存の民間キャリアから借りながら VPN を構築し、SINET と無線接続する携帯端末である。モバイル SINET は端末と SINET 及び大学等にある計算機との間のネットワーク環境を安全に構築できるという強みがあり、低遅延の5G回線が使用できる SINET6 が開発されてからはより一層実用化に向けた研究が進められている[3]。



図1 実験環境

2-2 実験結果

(1) ネットワークトラフィック

2つのエンコーダで記録した映像のビットレートと、設定したビットレートの最大値を示したグラフを図2と図3に示す。図中のTSrateは1チャンネルごとの伝送映像のビットレート値、TSMuxrateは設定したビットレートの最大値である。これらのグラフから設定したTSMuxrateに従って映像のビットレートが制御されていることが確認できる。また、伝送先で計測した4チャンネル合計のビットレート値は図4のようになった。

実験時には最終的に1チャンネル6Mbps、4チャンネル合計で24Mbpsの伝送までは映像の乱れがなく、適切に映像が伝送できていることが確認できた。しかし、1チャンネルにつき8Mbpsで伝送するようにTSMuxrateを設定したところ、映像に著しい乱れが見られた。このことは図2および図3においてTSMuxrateを8Mbpsまで引き上げた時刻に受信ビットレート値が4チャンネル合計で32Mbpsまで上昇しきっていないことから確認できる。

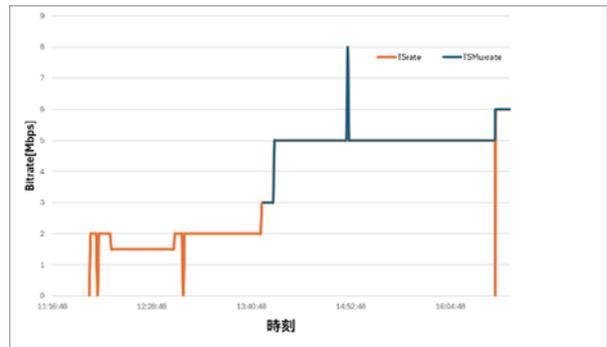


図2 映像ビットレート1

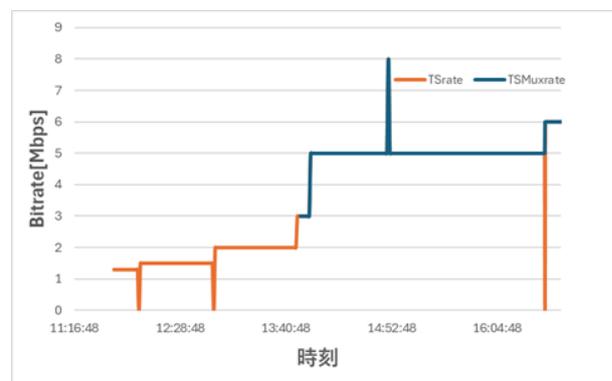


図3 映像ビットレート2

(2) パケットキャプチャ

本実験では16時40分から17時にかけて送信側と受信側の両方でパケットキャプチャを計測した。伝送の際の各パケットの遅延時間を表すヒストグラムを図5に示す。パケットキャプチャから伝送における平均パケット遅延は約36ミリ秒、パケットロス率は約0.44%という結果が得られた。

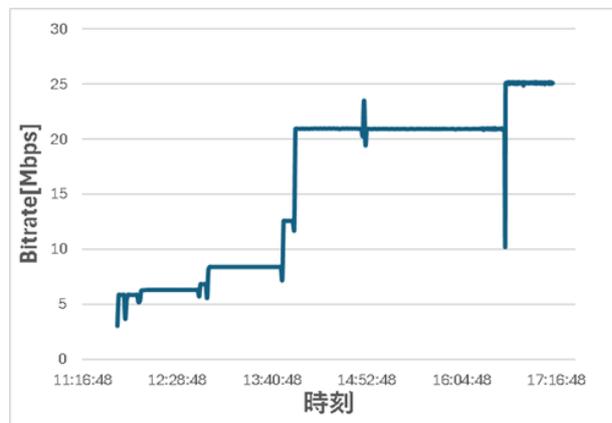


図4 4ch 合計ビットレート

2-3 通信トラフィック解析結果

(1) ネットワークトラフィック

映像の品質を確認しながらエンコーダのビットレート設定値を変更したところ、1つの映像ストリームで6Mbps、4つの映像ストリームの合計で24Mbpsとした場合は映像の乱れがなく、適切に映像が伝送できていることを確認した。1チャンネルにつき8Mbps、4つのストリームの合計で32Mbpsとした場合、伝送された映像に著しいノイズを確認。以上より、合計24Mbpsまでならば安定して実モバイル環境において伝送可能である。

(2) パケット遅延

情報通信技術委員会 (TTC) が発表している IP ベースサービスにおけるネットワーク性能目標では IP 通信サービスが満たすべき条件として平均的なパケット遅延は 100 ミリ秒以下である。本実験では、平均パケット遅延は約 36 ミリ秒であったため、本モバイル環境において、TTC の IP 通信サービスが満たす条件をクリア可能である。

(3) パケットロス率

情報通信技術委員会 (TTC) が発表している IP ベースサービスにおけるネットワーク性能目標では IP 通信サービスが満たすべき条件として、パケットロス率は 0.1%以下と規定されている。本実験では、パケットロス率は約 0.44%だった。この基準に照らし合わせると本実験で行ったモバイルSINETを用いた映像伝送は遅延の観点から見れば規定を満たしていない。本実験においては 5G 移動体通信を日比谷という都心で開催された非常に多くの人が集まるイベントにて用いている点から、同じ無線帯域を他の人が利用している可能性が高く、それら他の通信との衝突により基地局との通信に失敗している可能性が高いと考えられる。このため、サイバーフィジカル空間の効率的なデータ伝送には、解像度の低い少ないデータでパケットロスの確率の低いデータのみでもデコード可能な構成であるスケーラブル符号化が適している。

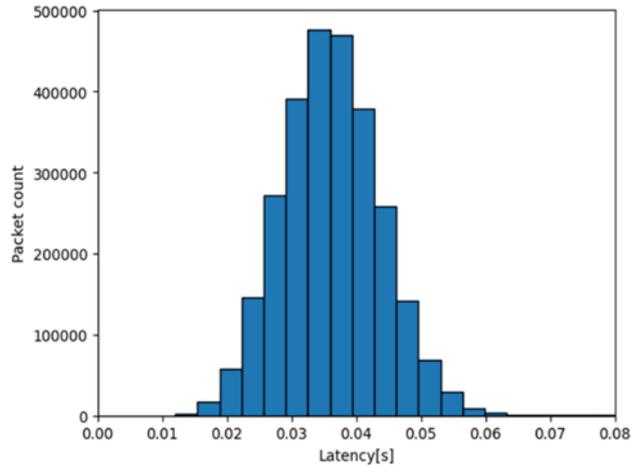


図 5 パケット遅延ヒストグラム

このため、サイバーフィジカル空間の効率的なデータ伝送には、解像度の低い少ないデータでパケットロスの確率の低いデータのみでもデコード可能な構成であるスケーラブル符号化が適している。

3 空間スケーラブル符号化における演算量の解析

2つの解像度に対して、それぞれに対して符号化を実施した場合の演算量と2つの解像度を1つのストリームにする空間スケーラブル符号化の演算時間の比較を図6に示す。2つの解像度をそれぞれ符号化した場合に対し、空間スケーラブルで符号化した場合の方が、符号化時間が長く、演算量が大きい。また、空間スケーラブル符号化の各処理時間の割合をグラフ1に示す。グラフ1に示す通り、各符号化モードを統一化する Unified Merge、AMVR(適応的探索精度探索)が大きな割合を示している。Unified Merge には、符号化モードの決定に関する処理が多く存在している。特に、ブロック分割については、スケーラブル符号化だけではなく通常の符号化でも大きな割合を占めている。このため、空間スケーラブル符号化の通常符号化が適用されているベースレイヤのブロック分割の高速化、さらに、エンハンスメントレイヤでのブロック分割の高速化の2つの高速化アルゴリズムについて提案する。

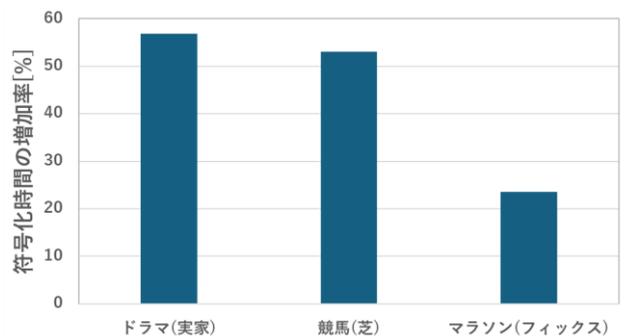


図 6 演算時間

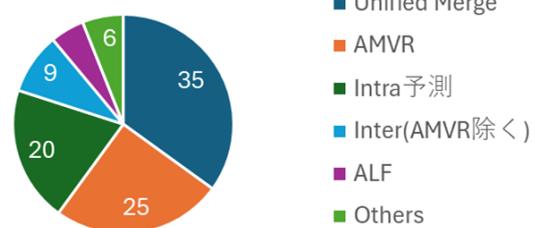
4 効率的軽量アルゴリズムの提案とその有効性

4-1 ベースレイヤにおける効率的ブロック分割アルゴリズム

(1) ベースレイヤにおける符号化概要

ベースレイヤにおける符号化は、2020年に標準化された国際標準符号化 VVC を用いている。VVCは、図7に示すように、ブロック分割、イントラ予測、インター予測、変換、量子化、逆変換、逆量子化、算術符号化の要素演算から構成される。入力映像は、時

グラフ1 空間スケーラブル符号化における各処理の割合(%)



間的・空間的に近接した画素間の相関が高くなる特徴がある。イントラ予測では、フレーム内の近接した画素から画素値を予測することにより空間的冗長性を排除し、インター予測では、前後のフレームから予測することにより時間的冗長性を排除している。

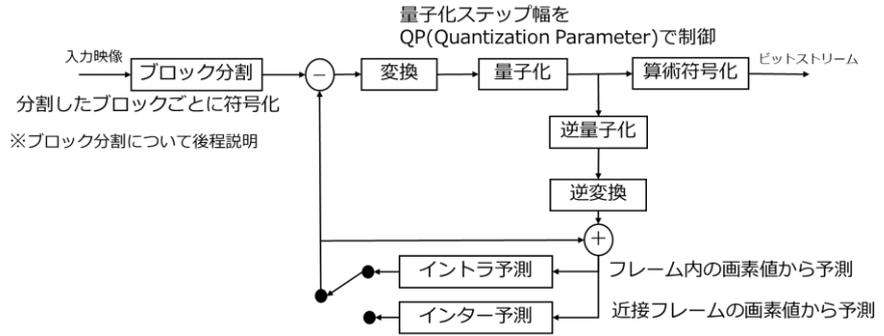


図 7 符号化の概要

VVC は、従来規格の HEVC から圧縮性能を向上させるために、新しい符号化ツールの採用や符号化ツールの改良がおこなわれている。特に、ブロック分割やイントラ予測では、大きな演算量を必要とする改良がおこなわれた。HEVC は正方ブロックのみのブロック分割であったのに対し、VVC はブロック分割モードが 2 分割や 3 分割など 5 種類に増えている。HEVC では 35 種類のモードでイントラ予測を行っていたが、VVC では、67 種類のイントラ予測モードから最適な予測モードを選択することができる。

(2) ブロック分割

VVC での符号化処理の基本単位は、Coding Tree Unit (CTU) であり、CTU のサイズは、8x8、16x16、32x32、64x64、128x128 画素から選択可能である。CTU 中のブロック分割は、図 8 に示すように、垂直 3 分木 (TV)、水平 3 分木 (TH)、垂直 2 分木 (BV)、水平 2 分木 (BH)、4 分木 (QT) の 5 種類のブロック分割であり、CTU を再帰的にブロック分割することが可能である。このことにより、VVC では表現できるブロック分割パターンが大幅に増加した。VVC のリファレンスソフトウェア (VVC Test Model, VTM) では、図 9 に示すように、128x128 から 4x4 まで、それぞれのブロックサイズにおける符号化効率を算出した後に、最適なものを選択している。128x128、64x64 のブロックは 4 分割し、32x32 以下のブロックでは、5 通りのブロック分割を試し、それぞれのブ

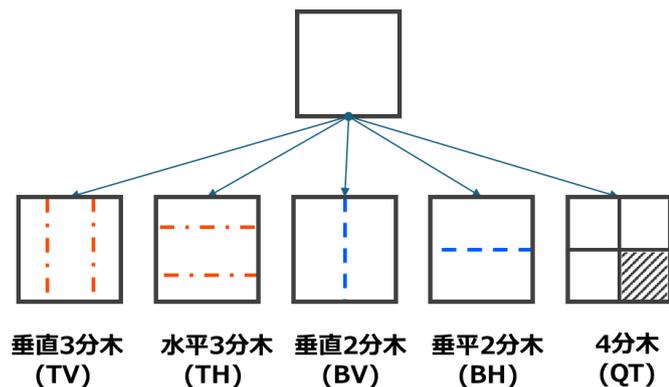


図 8 ブロック分割種別

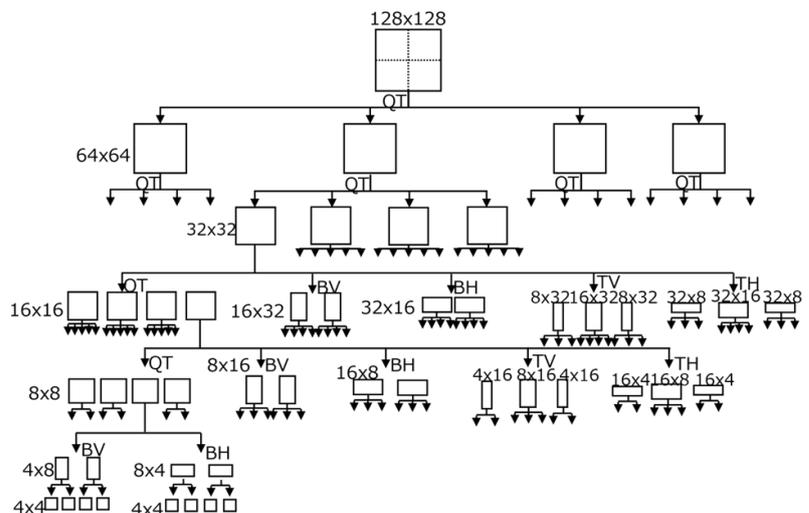


図 9 VTM におけるブロック分割決定手法

ブロック分割に対して、さらに、5通りのブロック分割を再帰的に試すことにより、最適なブロックを決定している。VTMにおけるブロック分割状況の例を図10に示す。青のブロックは64x64画素のブロックであり、水色、黄緑、黄色、オレンジ、赤の順にブロックサイズが小さくなっている。画素値の変化の少ない平坦な領域では、大きなブロック

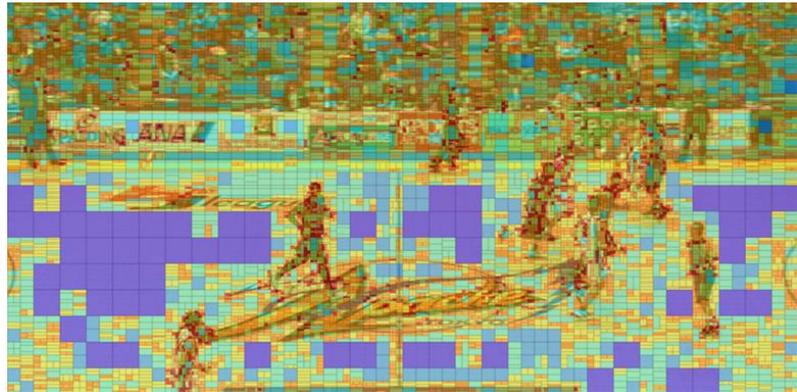


図10 ブロック分割状況

サイズが選択され、画素値の変化が激しい複雑な領域では、小さなブロックサイズが選択されていることがわかる。しかし、VTMでは、図9に示す通りの分割パターンにおけるビット量を、すべて計算した上で最終的な分割結果を決定する。このため、たとえ64x64が最終的に選択されていても、さらに小さいブロックで分割した場合のビット量を計算した上で、64x64が選択されている。このことから、可能なブロック分割パターンを全探索する処理がVTMのエンコード処理時間の増大の主要な原因となっている。このため、入力映像の当該ブロックの複雑さなどの特徴を利用し、分割パターンの全探索を行わないことにより、ブロック分割を高速化する手法が求められている。

(3) ベースレイヤにおける効率的ブロック分割アルゴリズム

VVCでは映像を事前に設定した固定サイズ、例えば128x128のCTUに分割し、そのCTUをブロック分割により高効率に符号化できる構成でCoding Unit(CU)に分割する。CUがどのようにブロック分割可能かは、ブロック分割の有無も含めてVVC規格で定められており、32x32より大きなCUはブロック分割しないかQT分割のみが選択可能だが、32x32以下のサイズのCUをブロック分割する場合は5種類のブロック分割モードが使用できる。そのため、32x32以下のブロック

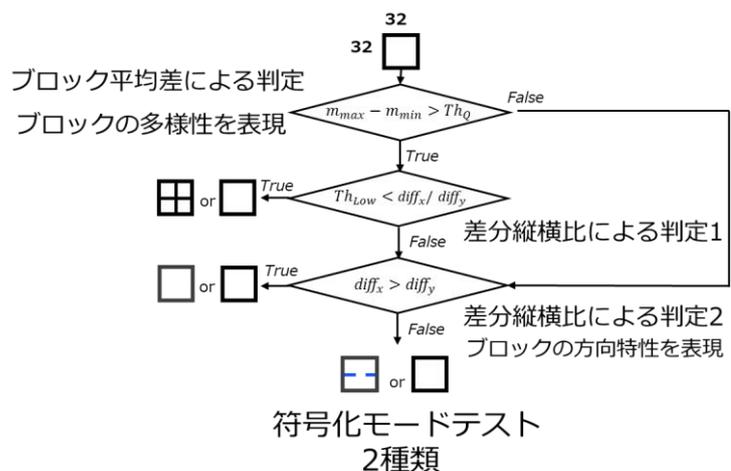


図11 ベースレイヤにおけるブロック分割アルゴリズム

分割の組み合わせは膨大であり、VTM等のVVC映像符号化実装において32x32以下のブロック分割探索に必要な演算量が占める割合が大きい。本研究では、画像特徴量を用いたブロック分割モード決定により高速化する。ブロック分割モードの探索中に複数の画像特徴量を算出し、探索するブロック分割モードを制限することで、探索木の枝刈りを行い、演算量を低減する。具体的には、効率的に演算量を削減するためにブロック分割モードが増える最大のブロックサイズである32x32ブロックの分割において、探索するブロック分割モードを絞り込む。この際、図11に示すように、より小さなブロック分割の探索を開始する前にブロック分割モードの絞り込みを行い、絞り込まれたブロック分割モードのみについて探索を行う。ブロック分割モードの絞り込みでは、VVCで規定されたブロック分割モードのうちQT、BT、BH分割の3種類の中から1つのモードを選択し、選択したブロック分割モー

ドとブロック分割しないモードを探索してより高効率に符号化できるモードを選択する。具体的には、5つのブロック分割モードとブロック分割しないモードとを合わせた6つの候補から、ブロック分割モード1つとブロック分割しないモードの計2つの候補に事前に絞り込む。なお、TH、TV 分割はほかのブロック分割モードに比べて、最適なブロック分割モードとして選択される割合が低いため、提案手法では除かれる。

以下に、画像特徴量について説明する。

【ブロック平均差】

ブロック平均差は、サブブロック間の輝度値の差を表し、QT 分割の是非を判定するために利用する。QT に分割してできる4つの16x16ブロックごとの輝度の平均値を求める。4つの平均値の中から、最大の平均値 m_{max} と最小の平均値 m_{min} を選ぶ。このときの、 $m_{max} - m_{min}$ をブロック平均差とする。 $m_{max} - m_{min}$ をブロック平均差として、ブロック平均差が大きいときに QT 分割が最適な分割モードだと判定する。

【差分縦横比】

差分縦横比は、水平方向の輝度の変化の総和 $diff_x$ と垂直方向の輝度の変化の総和 $diff_y$ の比率 $diff_x / diff_y$ で定義する。

【ブロック平均差と差分縦横比に基づく QT 分割の判定】

図7の1段目のブロック平均差を用いた条件と、2段目の差分縦横比を用いた条件の両方を満たすとき、最適なブロック分割が QT 分割であると判定する。

【ブロック平均差を用いた QT 分割の判定】

ブロック平均差を利用すると、サブブロックが異なった特徴をもつブロックであることを判定できる。ブロック平均差が大きい場合は、サブブロックの特徴が異なるため、ブロックを分割すると、サブブロックがそれぞれ最適なイントラ予測モードを選択できるようになり、圧縮性能が向上する。反対に、ブロック平均差が小さい場合は、サブブロックが同じ特徴を持つため、同一のイントラ予測モードによる予測の精度が高くなる。

上記のことから、ブロック平均差がしきい値 Th_Q 以下となるときには、4つのサブブロックに共通の特徴があるとして、QT 分割の可能性を排除する。

【差分縦横比に基づく QT 分割の判定】

差分縦横比の違いは高効率なブロック分割モードの違いと関係している。図12に差分縦横比とVTMが最適とする32x32ブロックの分割モードの関係を度数分布として示す。図12は提案手法の評価に使用した映像シーケンスの1フレームを映像の品質を制御する量子化パラメータ(QP)を22、27、32、37の4通りに変えてエンコードしたときの結果である。差分縦横比が大きい場合には、最適な分割モードとして、QT 分割が選択される確率が高くなっている。したがって、差分縦横比が大きい場合に QT 分割を選択する。ここでは閾値として $Th_{low}=0.5$ を設定し、差分縦横比がこの閾値を上回っているときに QT 分割を選

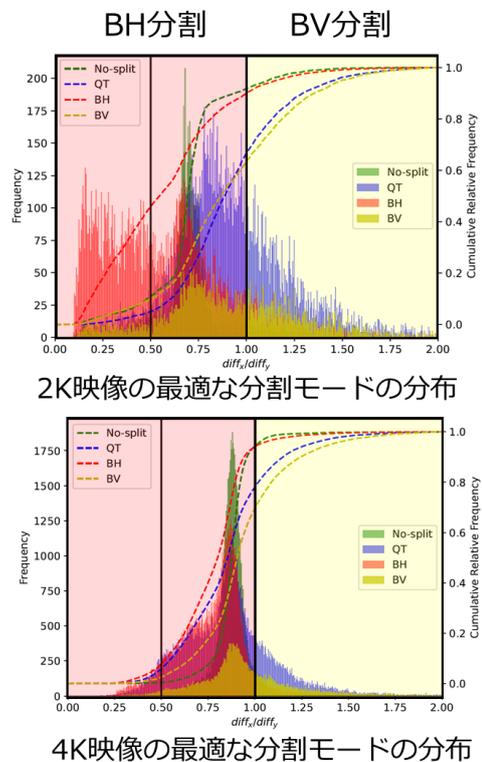


図12 差分縦横比分布

択する。図 12 で $\text{diff}_x / \text{diff}_y = 0.5$ より大きい場合に QT 分割が選択される。縦横差分比が 0.5 より小さい場合には、QT 分割が最適であると選択される確率が低いため、QT 分割を除外することによる符号化効率の損失は小さい。

【差分縦横比に基づく BH、BV 分割の判定】

QT 分割の是非を判定した後、ここでは BH 分割、BV 分割のどちらかを判定する。図 12 の BH 分割と BV 分割の分布を見ると、差分縦横比が小さいときには BH 分割が最適なモードとなる確率が高く、差分縦横比が大きくなるに伴い、BV 分割が最適なモードとなる確率が上がっていることがわかる。したがって、差分縦横比が小さいときには BH 分割が最適なモードと判定し、差分縦横比が大きいたまには BV 分割が最適なモードであると判定する。ここでは、 $\text{diff}_x > \text{diff}_y$ となるときには、BV 分割を選択し、 $\text{diff}_x \leq \text{diff}_y$ となるときには、BH 分割を選択する。図 12 で $\text{diff}_x / \text{diff}_y$ が 1 より大きい場合に BV 分割が選択され、小さい場合に BH 分割が選択される。

(4) 評価

提案手法による映像の符号化効率の変化と高速化率をソフトウェア実装により評価した。

VTM-20.2 のブロック分割手法を提案手法に置き換えたソフトウェアを用意し、変更を加える前の VTM と同じ映像を入力したときのビットレート、

PSNR や実行時間を計測した。入力する映像は 1920x1080(2K) と 3840x2160(4K) の 2 つの解像度の標準動画をを用い、解像度による提案手法の影響の違いを評価した。評価に用いた映像は、図 13、図 14 に示すように、2K の映像がハイビジョン・システム評価用標準動画像 第 2 版から 6 つ、4K の映像は超高精細 高色域標準動画像 B シリーズから 8 つ選択した。全ての映像のビット深度は 10bit であり、これらのクロマフォーマットを YUV420 形式に変換して使用した。ハイビジョン・システム評価用標準動画像の映像は、60 フレーム目から 10 フレームを、超高精細 高色域標準動画像 B シリーズの映像は、60 フレーム目から 5 フレームを使用した。すべてのフレームでイントラ符号化を用いて符号化し、QP は、22、27、32、37 の 4 点で評価する。また、ソフトウェアの評価環境を表 1 に示す。



図 13 HDTV 評価映像



図 14 4K 評価映像

表 1 評価環境

CPU	AMD Ryzen 5950X 16-Core Processor
クロック周波数	3.4GHz
メモリ	128GB
OS	Linux
VTM	Version 20.2

【評価指標】

評価指標として、VTM 及び提案手法のビットレート、PSNR、実行時間を用いる。ビットレートと PSNR から計算される BDBR と、実行時間から計算する Time-Saving の 2 つの指標により評価する。Peak Signal-to-Noise Ratio (PSNR) は画像がどの程度正確に再現されているかを示す品質指標

である。

【評価結果】

提案手法とVTM-20.2を比較したときのBDBRとTime-Savingを表2に示す。どちらの指標も基準をVTM-20.2として変化率を算出している。2K映像を用いた結果では、平均1.56%の符号化効率の59.45%のエンコード処理時間を短縮した。符号化効率の低下は最大で約3.1%であり、提案手法は幅広い映像に対してVVCが持つ高効率な符号化という特徴を損なっていない。また、すべての2K映像に対して55%以上の高速化ができており、多様な映像に対して高速化可能であると言える。4K映像を用いた結果では、平均1.25%の符号化効率の低下で、66.44%のエンコード処理時間を削減した。2K映像と比較すると符号化効率の低下が小さく、またエンコード処理時間の短縮率も大きいため、提案手法は解像度の高い映像に対する効果が大きいと言える。

また、2Kの時と同様に幅広い映像に対して有効である。表3に提案手法と既存手法の時間削減率とBDBRを示した。Fanらの手法、Fengらの手法は、BDBRが提案手法よりも高く符号化効率の低下が大きい。また、Yangらの手法は、2Kと4Kの映像でBDBRの差が大きい。検証に使用している映像は異なるものの、提案手法は、既存手法[4-6]と比べて、2K映像と4K映像の異なる解像度の映像に対して、符号化効率の低下を抑えながら実行時間を削減している。

表2 評価結果

(a) 2K 評価画像の結果

入力映像	BDBR (%)	Time-Saving (%)
No. 1 水球 (ゴール)	1.29	68.78
No. 2 水球 (開始)	1.60	69.81
No. 3 競馬 (ダート)	-0.39	65.63
No. 4 競馬 (芝)	1.67	70.42
No. 5 競馬 (ゴール)	-0.77	65.01
No. 6 パドック	1.61	64.26
No. 7 マラソン (スタート)	2.47	65.08
No. 8 マラソン (フィックス)	2.52	62.56
平均	1.25	66.44

(b) 4K 評価画像の結果

入力映像	BDBR (%)	Time-Saving (%)
No. 201 銀杏並木	1.22	65.22
No. 203 コスモス畑	0.86	60.65
No. 204 紅葉 (パンアップ)	0.92	59.31
No. 210 スタジオライブ (紙吹雪)	3.09	55.35
No. 214 バスケットボール	1.69	55.60
No. 218 競馬(ダート)	1.56	60.59
平均	1.56	59.45

表3 既存手法との比較

	Feng ら[4]		Yang ら[5]		Fan ら[6]		提案手法	
	2K	4K	2K	4K	2K	4K	2K	4K
BDBR (%)	2.36	1.03	1.87	N/A	2.45	2.96	1.56	1.25
Time-Saving (%)	68.0	59.3	53.7	N/A	74.0	57.2	59.5	66.4

4-2 エンハンスメントレイヤにおける効率的ブロック分割アルゴリズム

(1) アプローチ

空間スケーラブル符号化において、ベースレイヤとエンハンスメントレイヤでは解像度のみが異なり、映像のコンテンツに変化はない。このため、ベースレイヤとエンハンスメントレイヤにおける最適なブロック形状には一定の類似性があると考えられる。この類似性を活用し、ブロック分割形状を決定することによる高速化を行う。

(2) エンハンスメントレイヤにおける効率的ブロック分割アルゴリズム

図 15 に提案手法のアルゴリズムを示すフローチャートを示す。提案手法ではまず、現在の CU において幅または高さのいずれかが 32 ピクセル以上であるかを判定する。いずれかが 32 ピクセル以上である場合は、ブロック境界に基づく候補の絞り込みを行う。一方、32x32 ピクセル未満である場合、あるいはブロック協会に基づく絞り込み条件に該当しない場合には、ブロック形状に基づく候補の絞り込みを実施する。それでもいずれの条件にも該当しない場合は、通常の実行処理を実行する。 S_V は、領域内の縦長の CU の面積の総和、 S_H は、横長の CU の面積をとする。

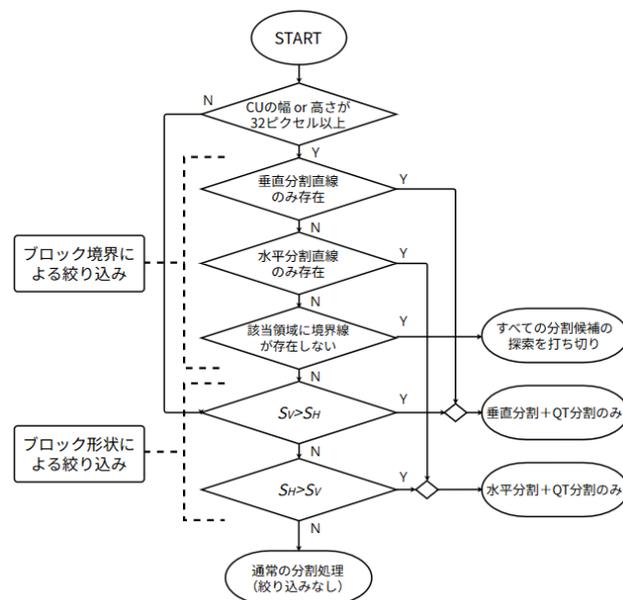


図 15 提案手法の流れ

【ブロック境界を用いた判定】

ベースレイヤのブロック分割の結果から、符号化対象領域を分割する直線を方向毎に検出し、エンハンスメントレイヤの符号化時の探索範囲を縮小する。領域を左右に分割する直線を垂直分割直線、上下に分割する直線を水平分割直線とする。図 16 に垂直分割直線及び水平分割直線の例を示す。VVC では矩形領域を繰り返し分割する仕組みを取っているため、長い境界線が探索初期の分割に相当する。該当領域における最大の境界線を用いて、分割候補を絞り込むことで、効率的に最適な候補を決定する。エンハンスメントレイヤのブロックと同じ位置に対応するベースレイヤの領域の垂直分割直線、もしくは水平分割直線のみが存在する場合、その境界線と直交する分割候補の探索を打ち切る。また、垂直分割直線および水平分割直線のいずれも存在しない場合、ベースレイヤにおいて当該領域は、平坦なテクスチャをしていると考えられるため、全ての分割候補への探索を打ち切ることで高速化を図る。

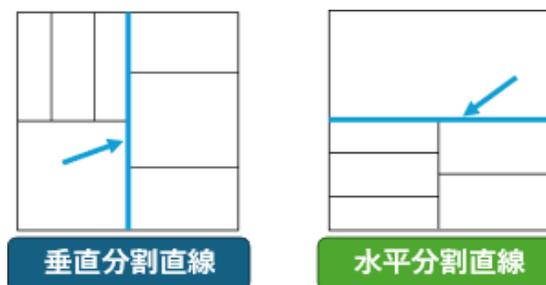


図 16 垂直/水平分割直線の例

【ブロック形状による判定】

ブロック境界を用いた判定の条件に該当しなかった場合、ベースレイヤのブロックの形状を利用した分割候補の絞り込みを行う。エンハンスメントレイヤのブロックを分割する時に対応するベースレイヤの領域を参照し、領域内のすべてのブロックの形状を調べることで領域内の縦長のブロックの面積の総和である S_V と、横長のブロックの面積である S_H を算出する。これにより、当該

領域内における支配的な分割方向を推定する。現在のブロックの幅か高さが 32 ピクセル未満かもしくは垂直分割直線および水平分割直線のいずれも確認できなかった場合場合、 S_V と S_H の大小によって水平方向もしくは垂直方向の分割候補の探索を打ち切る。

(3) 評価

本実験では、参照ソフトウェアである VTM-23.6 を AMD Ryzen 9 5950X プロセッサを搭載した PC 上で実行し、ランダムアクセス構成において量子化パラメータ (QP) を 37、32、27、22 に設定した 4 つの条件で符号化を行った。加えて、本実験では映像情報メディア学会が提供する超高精細・広色域標準動画像 B シリーズより競馬 (芝)、マラソン (フィックス)、ドラマ (実家) の特徴の異なる三つの動画像の始めの 33 フレームを切り抜いたものを対象として使用した。また、映像サイズはベースレイヤが 2K (1920x1080)、エンハンスメントレイヤが 4K (3840x2160、元映像) とした。

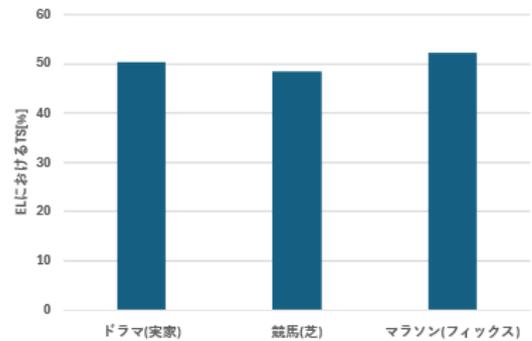


図 17 エンハンスメントレイヤにおける符号化演算短縮率

表 4 BDBR (%)

コンテンツ	競馬 (芝)	マラソン (フィックス)
ドラマ (実家)	2.84	2.56
競馬 (芝)	0.97	

【評価指標】

評価指標として、VTM 及び提案手法のビットレート、

PSNR、実行時間を用いる。ビットレートと PSNR から計算される BDBR と実行時間から計算する Time-Saving の 2 つの指標により評価する。Peak Signal-to-Noise Ratio (PSNR) は画像がどの程度正確に再現されているかを示す品質指標である。

【評価結果】

エンハンスメントレイヤにおける符号化処理時間の短縮率を図 17、BDBR を表 4 に示す。なお、時間短縮率は 4 種類の QP (37、32、27、22) における結果の平均値となっている。時間短縮率はコンテンツごとの差異はほとんどなく、どのコンテンツもおおよそ平均 50%前後の符号化実行時間を削減できている。BDBR は、評価したすべてのコンテンツにおいて 3%以下である。競馬 (芝) のみ 1%未満であった。それぞれのコンテンツのビットレートと Y-PSNR の関係を表した RD 曲線を図 18、図 19、図 20 に示す。どの RD 曲線も概ね既存の方式と重なっており、符号化効率の低下が微小であることが確認できる。

エンハンスメントレイヤにおける時間短縮率は

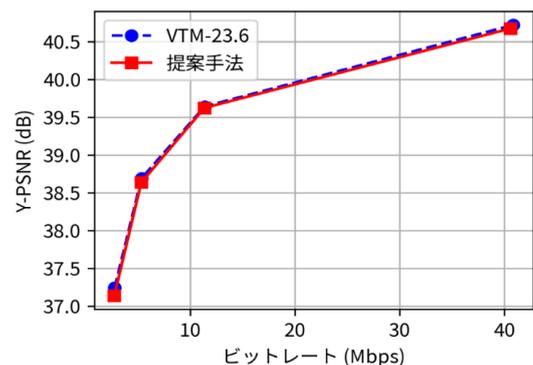


図 18 RD 曲線(ドラマ(実家))

本評価において用いた映像では、平均 50%前後であり、映像による違いはほとんどなかった。

このことから、本手法は、おおよそのコンテンツにおいて一定の高速化可能であると言える。BDBR に関しては、すべての映像で 3%未満であり、大きな画質の劣化ないため、本高速化アルゴリズムは有効である。

5 おわりに

本調査研究では、機械が観るための映像を効率よく伝送するための符号化方式に関する研究を実施した。

本研究は、リアルな通信トラフィックを解析し、帯域が制限される通信基盤と広帯域である通信基盤での共通に有効な符号化として期待できる空間スケーラブル符号化に対する高速化アルゴリズムを提案した。今後は、空間スケーラブル符号化の更なる高速化アルゴリズムなどを検討していく予定である。

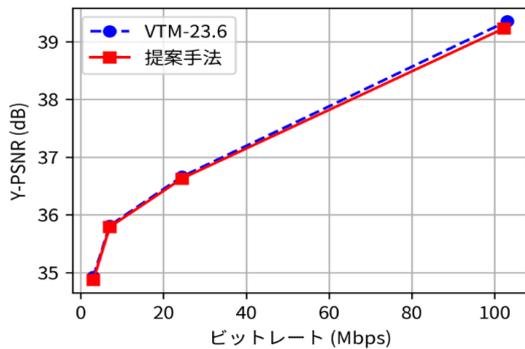


図 19 RD 曲線(競馬(芝))

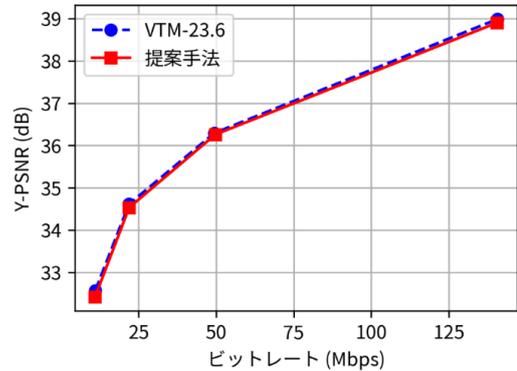


図 20 RD 曲線(マラソン)

【参考文献】

- [1] 国立情報学研究所『SINET6 の概要』
URL: https://www.nii.ac.jp/openforum/upload/01_sinnet6-gaiyou.pdf
(最終閲覧日 2025 年 6 月 28 日)
- [2] 学術情報ネットワーク SINET6 『学術情報ネットワークとは』
URL: <https://www.sinet.ad.jp/aboutsinet>(最終閲覧日 2025 年 6 月 28 日)
- [3] NII Today 第 95 号『5G モバイルと 400Gbps の融合』
URL: <https://www.nii.ac.jp/today/95/4.html>(最終閲覧日 2025 年 6 月 28 日)
- [4] Feng, A., Liu, K., Liu, D., Li, L. and Wu, F.: Partition map prediction for fast block partitioning in vvc intraframe coding, IEEE Transactions on Image Processing (2023).
- [5] Yang, H., Shen, L., Dong, X., Ding, Q., An, P. and Jiang, G.: Low-complexity CTU partition structure decision and fast intra mode decision for versatile video coding, IEEE Transactions on Circuits and Systems for Video Technology, Vol. 30, No. 6, pp. 1668–1682 (2019).
- [6] Fan, Y., Sun, H., Katto, J., JingMing, E ほか: A fast QTMT partition decision strategy for VVC intra prediction, IEEE Access, Vol. 8, pp. 107900-107911 (2020).

〈発表資料〉

題名	掲載誌・学会名等	発表年月
画像特徴量を用いたブロック分割によるVVC イントラエンコーダの高速化	研究報告システム・アーキテクチャ (ARC)	2024年8月
Adaptive Parallelization based on Frame-level and Tile-level Parallelisms for VVC Encoding	The Twelfth International Symposium on Computing and Networking	2024年11月
A Fast Block Partitioning Decision Method Using Luminance Textures for VVC Encoders	2025 IEEE International Conference on Consumer Electronics	2025年1月