

階層スケーラブル符号化における高速化手法の一考察 A study of speed-up methods in hierarchical scalable coding

小浦 陽[†] 丹羽 直也[†] 岩崎 裕江[†]
Yo Koura Naoya Niwa Hiroe Iwasaki

概要

ライブ配信等では、複数の解像度・ビットレートが同時に使われることが多い。階層符号化の中の一つである空間スケーラブル符号化は同じ映像を異なる解像度で同時に送信する場合において通常の符号化方式よりも高い符号化効率で圧縮できる。しかし、空間スケーラブル符号化は通常の符号化よりも符号化時間が大幅に増加してしまうという課題が存在する。この課題を解決するため、本研究ではブロックの境界線と形状を利用した空間スケーラブル符号化のブロック分割処理の高速化手法を提案する。提案手法は平均約 50%の高速化を達成し、符号化効率の低下を 3%未満に抑止することができた。

1. はじめに

動画配信サービスへの需要は、2020 年以降 急激に増大し、動画配信サービスの利用率が 30%以上増加している [1]。また、スマートフォンなどの携帯機器で動画を視聴するケースも増加しており、動画の受信デバイスの多様化が進んでいる [2]。大きさが異なる多様なデバイスに対する動画配信を実現するためには動画の送信元は同じ映像を複数の解像度で伝送する必要があり、これによって伝送に要する帯域が増加してしまうという問題がある。複数の解像度を効率的に符号化する空間スケーラブル符号化の実現が望まれている。

空間スケーラブル符号化は複数解像度の映像を効率的に符号化する技術であり、Joint Video Experts Team (JVET) によって 2020 年 7 月に規格化された次世代映像符号化方式である Versatile Video Coding (VVC) にも採用されている。近年は高解像度映像の普及や、複数解像度への配信への需要の高まりを背景に、その重要性和注目度が高まっている。一方、空間スケーラブル符号化は、符号化時間が増大するという課題が存在する。図 1 に通常の符号化を基準とした空間スケーラブル符号化による平均的な符号化時間の増加率を示す。空間スケーラブル符号化は、通常の符号化よりも符号化時間が増加することが分かる。

本研究では空間スケーラブル符号化において、解像度が異なる映像同士の類似性を用いて符号化時間の多くを占める探索処理を削減し、空間スケーラブル符号化を高速化する手法を提案する。

2. 空間スケーラブル符号化とその課題

2.1 空間スケーラブル符号化

空間スケーラブル符号化とは複数の解像度の映像を入力として、それらの映像を階層的に一つのビットストリーム

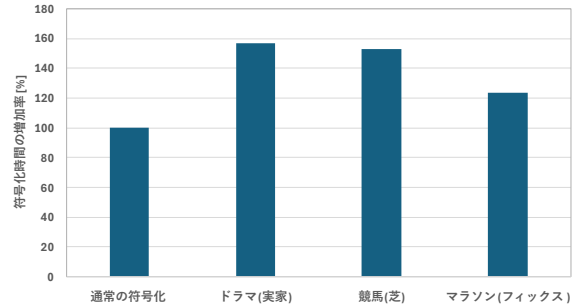


図 1. 符号化時間の増加率 [%]

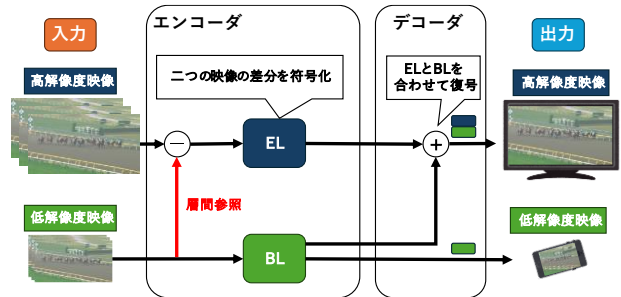


図 2. 空間スケーラブル符号化の構造

に圧縮する符号化方式である。図 2 は空間スケーラブル符号化の構造を表しており、送信側では高解像度映像を符号化する際に、Base Layer (BL) に含まれる低解像度映像との差分を Enhancement Layer (EL) として符号化し、低解像度デバイス向けには BL のみ、高解像度デバイス向けには BL と EL の両方を伝送する。この空間スケーラブル符号化の仕組みでは、EL の高解像度映像と低解像度映像の類似性を活かした高効率な符号化により、独立して低解像度映像と高解像度映像を符号化及び伝送する場合と比べて、より小さな帯域で複数解像度の映像を伝送できる。これらを同時に伝送することによって BL のみを復号すれば低解像度映像を、BL と EL を合わせて復号すれば高解像度映像を出力できる。この方式を用いることで解像度ごとに符号化を行った場合よりも小さな帯域で複数解像度の映像を符号化することが出来る。

2.2 空間スケーラブル符号化の課題

空間スケーラブル符号化には解像度ごとに独立して符号化を行った場合よりも符号化時間が増大してしまうという課題が存在する。BL の符号化は通常の符号化と差異はないが、EL では BL の情報を基にした予測を追加で行うため、処理時間が増加する。図 3 に空間スケーラブル符号化による処理の追加を示す。通常の符号化ではフレーム内の復号済みブロックの画素値や異なる時点のフレームの動き情報等からの予測を行う。空間スケーラブル符号化の EL では

[†] 東京農工大学 Tokyo University of Agriculture and Technology

それらに加えて、BLの同時点のフレームにおける動き情報等を用いる場合の符号化効率についても評価を行うため、通常の符号化よりも多くの探索を行うため計算量が增大する。

2.3 ブロック分割処理

ブロック分割処理とは、フレームを多様なサイズおよび形状の Control Unit (CU: 符号化の基本単位となる矩形ブロック) に分割する処理である。これにより、予測や変換、量子化などの一連の符号化処理を効率的に行うことが可能となる。

VVCで規定されている分割パターンを図4に示す。各CUにおける分割パターンは従来規格のHEVCでは分割をしない場合を除けばQuad Tree (QT) 分割のみだったが、VVCでは水平および垂直方向に二分分割、三分分割することが可能となった。分割方式の拡張により、VVCでは従来規格と比べて映像内容に応じたより柔軟なブロック分割が可能となり、圧縮性能の向上が実現された。一方で、分割候補の増加に伴って探索空間が拡大し、VVC全体の計算複雑度は従来規格よりも大幅に増加している。特にブロック分割処理は計算量増加の主要因となっており、全体の符号化時間のうち約90%以上を占めている[3]。

3. 先行研究

空間スケーラブル符号化の符号化時間の増大という課題を解決するため、様々な高速化手法が提案されている。

Liquan Shen と Guorui Feng らはCUの最適予測モードが空間、時間および層間で隣接しているCUの選択モードと類似している傾向があることから、それぞれの隣接CUの選択モードと動き情報によってELにおける探索する予測モードの絞り込みを行った[4]。これにより符号化効率の低下を1.5%以下に抑えつつ、60%以上の符号化時間の削減を達成した。

Wang Dayong らは階層間参照モードの早期決定やイントラ予測(フレーム内予測)におけるDirection Mode (DM)の探索の高速化等の手法を用いることで、ELのフレームにおいて符号化効率の低下を平均で約0.21%に抑えながら平均70%以上の高速化を実現した[5]。

4. 提案手法

4.1 提案手法のアプローチ

本研究ではELにおけるブロック分割候補の絞り込みによる高速化手法を提案する。

空間スケーラブル符号化において、BLとELでは解像度のみが異なり、映像の内容自体に変化はないため、最適となるブロックの分割方式には一定の類似性があると考えられる。このことからELのCUを符号化する際に、そのCUに対応するBLの領域にあるブロックの境界線と形状を活用することによって分割候補を効率的に絞り込む。

4.2 提案手法のアルゴリズム

本節では実装した高速化手法のアルゴリズムについて説明する。本研究の提案手法は二つの手法を組み合わせた形でブロック分割候補の探索範囲を絞り込んでいる。

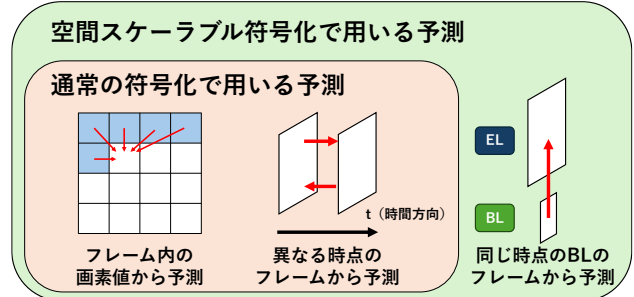


図3. 空間スケーラブル符号化による処理の追加

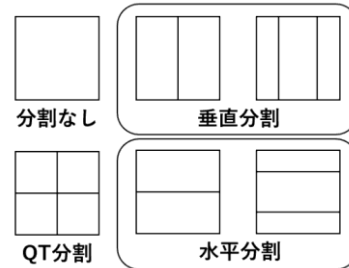


図4. VVCで規定されている分割パターン

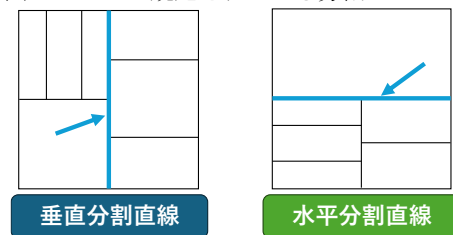


図5. 垂直分割直線及び水平分割直線の例

4.2.1 ブロック境界による絞り込み

提案手法では、まずBLのブロック分割の結果から、符号化対象領域を切断する直線を方向毎に検出し、ELの符号化時の探索範囲を絞り込む。本研究では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5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。提案手法では当該領域で取りえる最大の境界線を用いて、探索の初期段階から分割候補を絞り込むことで、効率的に探索空間を狭める。ELのCUの符号化時に対応するBLの領域を参照し、その領域内に垂直分割直線、もしくは水平分割直線のみである場合、その境界線と直交する分割候補の探索を打ち切ることで分割候補を適応的に絞り込む。なお、垂直分割直線および水平分割直線のいずれも検出できなかった場合、BLにおいてその領域はそれ以上の分割をしておらず、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。加えてサイズの小さいブロックや、水平と垂直のどちらの方向にも境界線のあるQT分割をしているブロック等においてはELとBLで最適となる分割方式が異なる場合が多いため、この手法は幅もしくは高さが32ピクセル以上かつ垂直分割直線か水平分割直線のいずれかが確認できた場合にのみ使用している。

4.2.2 BL のブロック形状による絞り込み

4.2.1 の条件に該当しなかった場合、BL の CU の形状を利用した分割候補の絞り込みを行う。EL の CU の分割時に対応する BL の領域を参照し、領域内のすべての CU の形状を調べることで領域内の縦長の CU の面積の総和である S_V と、横長の CU 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。

提案手法では現在の CU 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、 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。これによりどちらの方向にも境界線がある場合やサイズの小さいブロックなどの境界線だけでは分割候補の絞り込みが困難な場合においても分割候補を効率的に絞り込むことが可能となる。

4.2.3 提案手法の全体像

図 6 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。はじめに、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、4.2.1 で述べたブロック境界に基づく候補の絞り込みを行う。一方、幅・高さの両方が 32 ピクセル未満である場合、あるいは 4.2.1 の絞り込み条件に該当しない場合には、 S_V および S_H を用いた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、通常分割処理を実行する。

5. 評価

提案手法の高速化率とそれに伴う符号化効率の劣化を VVC の参照ソフトウェアである VVC Test Model (VTM) に搭載されている元の設定の空間スケーラブル符号化と、提案手法を導入した空間スケーラブル符号化を比較評価した。

5.1 評価環境

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 B シリーズ [6] 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の始めの 33 フレームを使用した。また、映像サイズは BL が 2K (1920x1080)、EL が 4K (3840x2160、元映像) とした。

5.2 評価指標

本実験では符号化効率と EL における符号化時間の短縮率によって評価を行う。符号化効率の指標には同画質における平均的なビットレートの増加率を表す Bjontegaard Delta (BD) -rate を使用する。本実験における BD-rate は EL における Y-PSNR と、BL と EL のビットレートの合計値から算出する。また、EL における符号化時間の短縮率 TS_{EL} を

$$TS_{EL} = \frac{Time_{default} - Time_{proposed}}{Time_{default} - Time_{BL}} \dots (1)$$

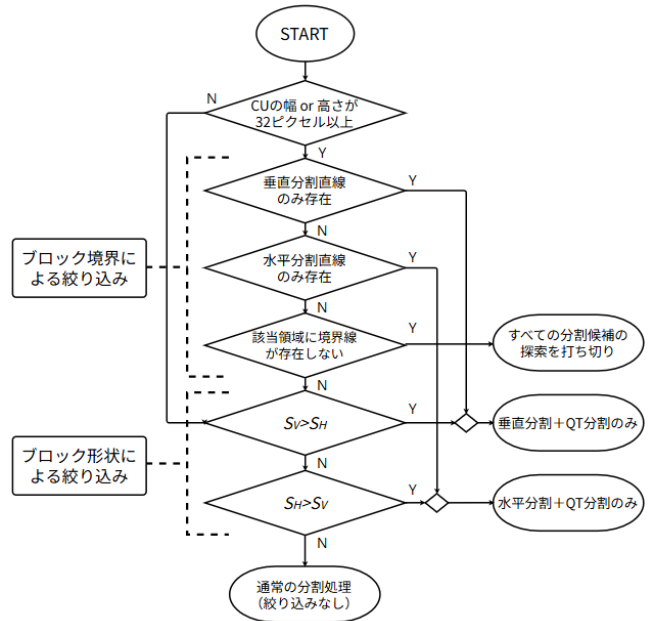


図 6. 提案手法のフローチャート

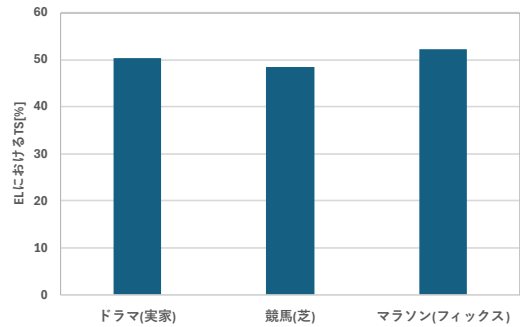


図 7. EL における時間短縮率の平均値 [%]

表 2. 動画ごとの BD-rate [%]

動画	競馬(芝)	マラソン(フィックス)
ドラマ(実家)	2.84	2.56

と定義する。ただし、 $Time_{proposed}$ は提案手法を使用した際の符号化時間、 $Time_{default}$ は元の設定の空間スケーラブル符号化を使用した際の符号化時間、 $Time_{BL}$ は BL の符号化時間を表す。この式によって BL の符号化時間に依らない EL のみの符号化時間の短縮率を導出する。

5.3 実験結果

それぞれの動画の EL における時間短縮率を図 7、BD-rate を表 2 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である。どの動画もおおよそ平均 50% 程度の時間短縮率が確認できる。

BD-rate は競馬 (芝) のみ 1% 未満、ドラマ (実家) およびマラソン (フィックス) では、3% 未満の低下が確認できた。それぞれの動画のビットレートと Y-PSNR の関係を表した RD 曲線を図 8、図 9、図 10 に示す。

6. 考察

EL における時間短縮率はどの動画も平均で 50% 前後であり、効果がみられた。

一方、BD-rate は、競馬（芝）と他の二つの動画で大きく異なる結果が得られた。これは競馬（芝）が他の二つの動画と比較して、平坦なテクスチャの領域が多かったことから BL と EL で分割方式が大きく変わる領域が少なかったため、符号化効率の低下が抑えられたと考えられる。加えて、他の 2 つの映像においても符号化効率の大幅な低下は抑制されていることが確認できた。

以上の点から提案手法は高速化と符号化効率のバランスにおいて一定の有効性があると考えられる。

7. おわりに

本研究では、BL のブロック境界線および形状情報を活用した、VVC における空間スケラブル符号化の高速化手法を提案した。実験の結果、提案手法を適用することによって平均約 50% の高速化を実現し、加えて符号化効率の低下も最大で 2.84% に抑えられることを確認した。今後は他の映像についても同様の検証を行うことで、提案手法が効果的に働く条件を分析し、予測モードやイントラ予測における DM の絞り込み手法等も取り入れることによってより優れた高速化手法の開発していく。

謝辞

本研究の一部は JSPS 科研費 23K24827, 24K20751 及び一般財団法人テレコム先端技術研究支援センター(SCAT)、公益財団法人放送文化基金、公益財団法人電気普及財団、2024 年度国立情報学研究所公募型共同研究 (24S0103, 24S108) の支援による。

参考文献

- [1] 総務省情報通信政策研究所. 令和 5 年度情報通信メディアの利用時間と情報行動に関する調査報告書, 2024.
- [2] ICT 総研, 2025 年 有料動画配信サービス利用動向に関する調査, <https://ictt.co.jp/report/20250423.html>, (最終閲覧日 2025/06/04)
- [3] A. Tissier, A. Mercat, T. Amestoy, W. Hamidouche, J. Vanne and D. Menard, "Complexity Reduction Opportunities in the Future VVC Intra Encoder," *2019 IEEE 21st International Workshop on Multimedia Signal Processing (MMSP)*, Kuala Lumpur, Malaysia, pp. 1-6, 2019, doi: 10.1109/MMSP.2019.8901754.
- [4] L. Shen and G. Feng, "Content-Based Adaptive SHVC Mode Decision Algorithm," in *IEEE Transactions on Multimedia*, vol. 21, no. 11, pp. 2714-2725, Nov. 2019, doi: 10.1109/TMM.2019.2909859.
- [5] D. Wang et al., "Hybrid Strategies for Efficient Intra Prediction in Spatial SHVC," in *IEEE Transactions on Broadcasting*, vol. 69, no. 2, pp. 455-468, June 2023, doi: 10.1109/TBC.2022.3222997.
- [6] 映像情報メディア学会, 電波産業会. 超高精細・広色域標準動画像—B シリーズ, 2017.

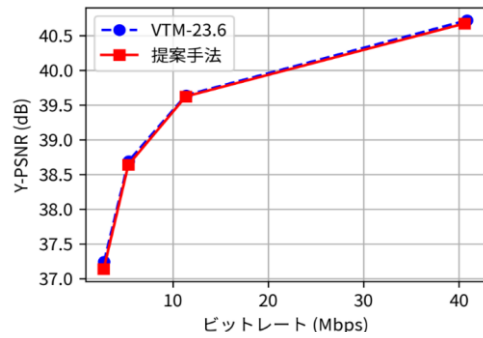


図 8. RD 曲線 (ドラマ (実家))

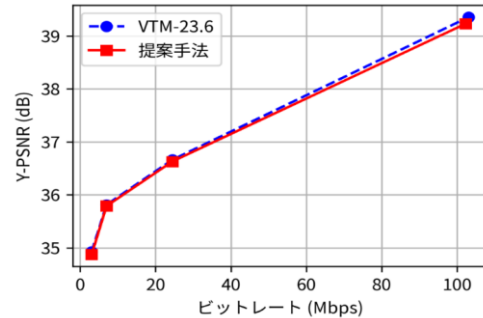


図 9. RD 曲線 (競馬 (芝))

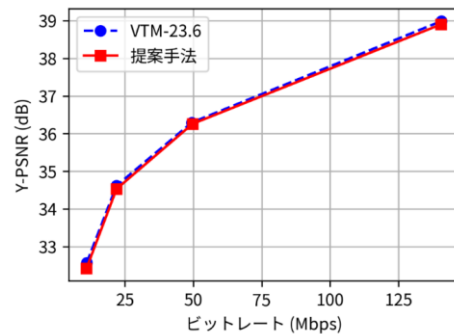


図 10. RD 曲線 (マラソン (フィックス))