

任意形状画像信号の符号化に関する基盤的研究とその点群映像符号化への適用

代表研究者 高 村 誠 之 法政大学情報科学部 教授

1 はじめに

近年、映像通信分野では、従来の平面映像から三次元空間そのものを伝送・共有する方向への発展が進んでいる。特にメタバースやデジタルツインでは、利用者が自由な位置・方向から映像を観察できる自由視点映像技術が重要となる。このような応用では、被写体の形状情報や表面属性を含む三次元データを効率的に取り扱う必要があり、点群やメッシュに代表される三次元メディアの圧縮技術が不可欠となっている。

MPEG では、点群データを対象とした Video-based Point Cloud Compression (V-PCC) および Geometry-based Point Cloud Compression (G-PCC) が標準化されている。特に V-PCC は、三次元点群を二次元画像へ投影し、既存の映像符号化技術を活用できることから、高い圧縮性能を実現している。一方で、点群を画像化する過程では複数のパッチが生成され、それらの間には符号化対象ではない領域 (DCR) が多数発生する。このような画像は一般的な矩形画像とは大きく異なる構造を有しており、従来の画像符号化手法をそのまま適用すると圧縮効率が低下するという課題が存在する。

また、近年標準化が進められている Video-based Dynamic Mesh Coding (V-DMC) においても、任意形状領域を含む画像が生成されるため、同様の課題が存在する。このことから、任意形状画像そのものを対象とした信号処理・符号化技術の重要性が高まっている。

本研究では、このような背景の下、「任意形状画像信号の符号化」を中心課題として研究を実施した。特に、任意形状画像において発生する不要領域をどのように扱うかという観点から、符号化志向パッチ補間技術、直交変換志向ブロック内パディング技術、最適パッチ補間画像生成技術、およびパッチ志向符号化最適化技術について研究を進めた。さらに研究を発展させる過程で、フレーム順序最適化、Occupancy 情報を利用したスパース信号分解型圧縮、さらには H. 266/VVC 符号化最適化へと研究を展開した。

これらの研究は、単に点群映像符号化の性能向上に留まらず、「信号が持つ構造情報を積極的に利用することで高効率圧縮を実現する」という共通理念に基づいている。本報告では、研究期間 (2024 年 4 月～2026 年 3 月) に実施した研究内容と得られた成果について報告する。

2 V-PCC における符号化志向パッチ補間技術

2-1 研究背景

V-PCC では、三次元点群を複数のパッチへ分解し、二次元の「パッキング画像」として配置することで既存映像符号化器を利用する。しかし、パッチ間には DCR が発生し、これらの領域も通常の画像符号化器によって処理されるため、符号化効率の低下を招く (図 1)。



図 1 V-PCC におけるパッキング画像の例。左: テクスチャ情報, 中央: 占有情報, 右: 幾何情報

従来方式として、Patch-Edge Extension (EdgeEx) や Smoothed Push-Pull (SPP) が提案されている。SPP は画像の滑らかさを考慮した補間方式であり、V-PCC 参照ソフトウェアでも採用されている。しかし、これらの方式は画素値の連続性のみに着目しており、映像符号化器内部の予測機構との整合性を十分考慮してい

2-2 提案方式

従来の DCR 補間方式は、補間後の画像が滑らかに見えることを主な目的として設計されており、符号化器内部で実行される予測処理との関係は必ずしも考慮されていなかった。そのため、人間の視覚的には自然な補間結果であっても、符号化器内部では大きな予測誤差が発生し、結果として余分な符号量を必要とする場合があった。

2023 年度までに画面内予測（イントラ予測）への適用を実現していたが、本研究ではさらに画面間予測（インター予測）へ拡張した。具体的には、動き補償予測によって参照フレームから生成される予測画像を利用し、DCR 領域の画素値を適応的に決定する方式を導入した。これにより、動き補償予測を伴う映像系列に対しても DCR 領域の符号量削減が可能となった。

The diagram illustrates the video encoding process flow. It starts with an input frame (a person) entering the 'DCR初期化' (DCR Initialization) block. The output of this block goes to the 'フレーム順序変更' (Frame Order Change) block, which is highlighted with a red border. From there, the data flows into the 'ビデオエンコーダ' (Video Encoder) block. A feedback loop exists from the 'ビデオエンコーダ' back to the 'フレーム順序変更' block, labeled '予測ブロック情報、イントラ・インター予測値' (Prediction block information, intra/inter prediction values). Additionally, an '占有情報' (Occupancy Information) input (a binary mask of the person) enters the 'DCR補間' (DCR Interpolation) block. The output of this block also feeds into the 'フレーム順序変更' block. A dashed blue line connects the 'フレーム順序変更' block to the 'ビデオエンコーダ' block, indicating a data path. The final output of the 'ビデオエンコーダ' block is the encoded video frame (a person).

図2 提案する符号化志向パッチ補間方式の概念図

2-3 フレーム順序最適化

従来の映像符号化では、撮影順あるいは生成順にフレームを符号化することが一般的である。しかし、点群映像ではフレーム間の見た目の類似度と生成順序が必ずしも一致しない。そのため、時間的に離れたフレーム同士が高い類似性を持つ場合でも、それらが予測に利用されないという問題があった。

また、提案方式は V-PCC の既存処理系を大きく変更することなく適用できるため、既存標準との親和性が高いという特徴を有する。さらに、符号化志向パッチ補間技術と組み合わせることで、空間方向および時間方向の双方から符号化効率改善を図ることが可能となる。

2-4 実験結果

提案方式を MPEG V-PCC 参照ソフトウェアに実装し、標準評価シーケンスを用いて評価を行った。評価では、従来方式である SPP 方式との比較を行い、符号化効率および生成画像の特性を検証した。

まず、符号化志向パッチ補間技術について評価した結果、DCR 領域における予測残差を大幅に削減できることを確認した。特にインター予測を利用する条件では、参照フレームから生成される予測信号を活用することで、従来方式よりも予測効率を向上できた。図 3 に示すように、提案方式では DCR 境界付近の画素値不連続性が抑制され、より符号化器に適した補間結果が得られた。



図 3 DCR 補間結果の例

次に、フレーム順序最適化技術について評価した。MS-SSIM に基づいて符号化順序を再構成することで、時間的に高い相関を有するフレーム同士を近接させることが可能となった。その結果、インター予測効率が向上し、予測残差の低減が確認された。また、提案方式は既存の V-PCC 処理フローに対して前処理として適用可能であり、符号化器構造を変更することなく利用できることを確認した。

さらに、両方式を組み合わせた場合の評価も行った。符号化志向パッチ補間は主として空間方向の構造情報を利用するのに対し、フレーム順序最適化は時間方向の相関を利用する。そのため両者は相補的な関係にあり、組み合わせることによってさらなる性能向上が期待できる。実験の結果、提案方式は従来の SPP 方式と比較して Y 成分 BD-rate において最大 6% 程度の改善を達成した (図 4)。また、フレーム順序最適化を併用することで追加的な符号化効率改善が得られた。

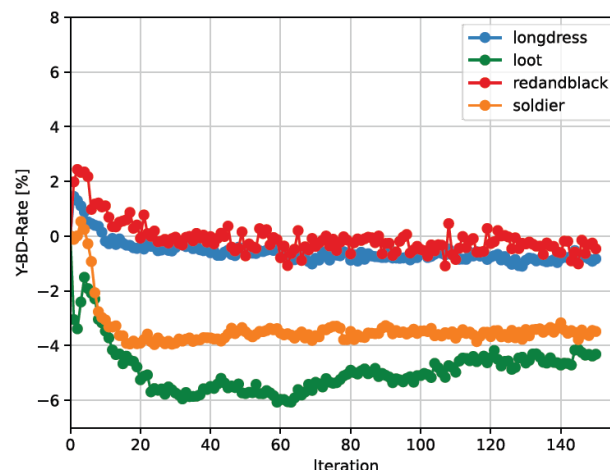


図 4 提案方式と従来方式の BD-rate 比較

これらの結果は、任意形状画像に対する符号化器内部処理と予測機構を統合的に最適化することの有効性を示している。特に、従来は画素値の滑らかさのみを考慮していた DCR 補間に対し、符号化器内部の予測特性や時間方向の相関構造を考慮することが重要であることを明らかにした。

関連成果として、PCSJ/IMPS2024 において「V-PCC 符号化における効率的なパッチ間補間方法に関する検討」を発表した。また、その発展成果を画像電子学会誌へ投稿し採録された[11]。

3 任意形状画像に対するスパース信号分解型圧縮

3-1 研究背景

V-PCC では、3 次元点群を複数のパッチに分割し、それらを 2 次元画像上に配置することで映像符号化技術を適用している。しかし、生成される画像は矩形画像とは異なり、占有領域 (Occupancy Region) と未使用領域 (DCR) が混在する任意形状画像となる。

従来の画像符号化では、DCT などの直交変換を矩形ブロック全体に適用するため、DCR を含むブロックでは不要な変換係数が発生し、圧縮効率が低下する。この問題を解決するため、本研究では、任意形状領域のみに着目した新たな符号化方式について検討した。

3-2 直交変換志向ブロック内パディング技術

当初は、DCR を適切な値で埋めるパディング技術について検討した。V-PCC において生成される任意形状画像は、一般的な矩形画像とは異なり、有効領域と DCR が混在する構造を有している。しかし、既存の画像符号化器では矩形ブロック単位で直交変換を適用するため、DCR を含むブロックに対しても通常の変換処理が実行される。

このとき、DCR 画素値の与え方によって変換係数の分布が大きく変化する。例えば、単純な一定値パディングでは境界部に急峻な不連続が発生し、高周波成分が増加する。一方、周辺画素からの補間を用いる方式では連続性は向上するものの、必ずしも変換係数のエネルギー集中度が最適化されるとは限らない。

そこで本研究では、直交変換後の係数分布に着目し、変換係数のエネルギーが少数の係数へ集中するようなパディング方式について検討した。具体的には、DCR 部分を未知領域とみなし、有効領域との整合性を保ちながら低ランク近似が成立するような画素値を推定する手法を検討した。その結果、従来の一定値パディングや単純補間方式と比較して、変換係数の低ランク性を利用したパディングが有効であることが示唆された。

しかし、任意形状画像では Occupancy 形状がブロックごとに大きく異なり、単一の変換基底だけでは十分な表現能力が得られない場合が存在することも明らかとなった。特に複雑な形状を有する領域では、最適な基底が画像ごとに変化するため、より柔軟な信号表現方式が必要であることが分かった。このことが、次節で述べるスパース信号分解型圧縮方式へ研究を発展させる契機となった。

3-3 スパース信号分解を用いた任意形状画像圧縮

前節の検討を通じて、任意形状画像では固定的な変換基底による表現に限界があることが明らかとなった。そこで本研究では、Occupancy Map に基づいて有効画素のみを対象としたスパース信号分解型圧縮方式を提案した。

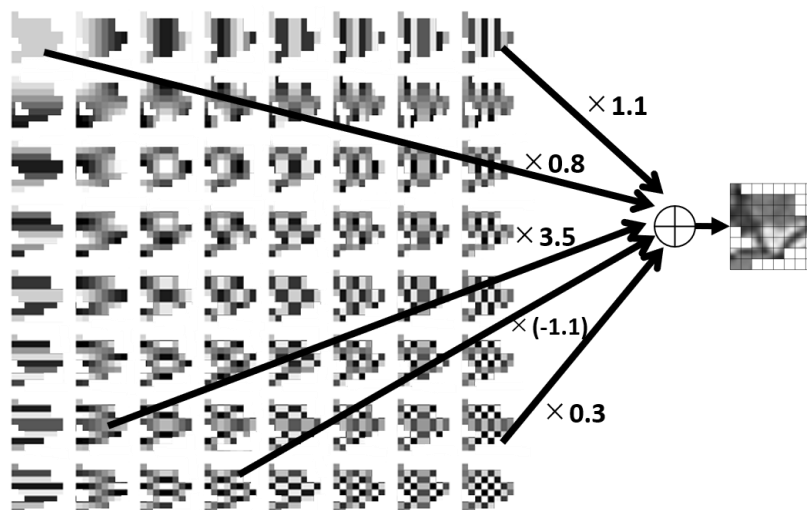


図 5 Occupancy 情報を利用したスパース信号分解の例

提案方式では、まず画像をブロック単位に分割し、Occupancy 情報を利用して有効画素のみを抽出する。

次に、あらかじめ用意した過完備辞書に対してスパース分解を行い、画像信号を少数の基底関数の線形結合として表現する。これにより、矩形ブロック全体を処理する従来方式とは異なり、実際に情報を有する領域のみを対象とした効率的な信号表現が可能となる（図 5）。

本方式の特徴は、Occupancy 形状に応じて適応的に基底を選択できる点にある。従来の DCT や DST では固定的な基底が用いられるが、スパース分解では信号構造に応じて最適な基底を選択できるため、複雑な形状を有する任意形状画像に対しても高い表現能力を有する。また、少数の係数のみを伝送することで高い圧縮性能が期待できる。

本方式は、

- DCR を含む矩形全体を処理する必要がない
- Occupancy 形状に適応した基底選択が可能
- 有効係数のみを伝送することで符号量削減が期待できる
- V-PCC や V-DMC など様々な任意形状画像へ適用可能である

といった利点を有する。

さらに、本研究では提案方式の基本原理解について知的財産化を進め、「圧縮方法及びコンピュータプログラム」として特許出願を行った[12]。提案技術は V-PCC だけでなく、Video-based Dynamic Mesh Coding (V-DMC)、AR/VR コンテンツ、自由視点映像、医療画像における関心領域 (ROI) 符号化など、任意形状領域を含む幅広い情報処理分野への応用が期待される。

また、本研究で検討したスパース信号分解の考え方は、近年注目されている深層生成モデルや学習型圧縮技術との親和性も高く、将来的には AI を利用した任意形状画像符号化への発展も期待できる。

4 H.266/VVC 符号化最適化への展開

4-1 研究背景

任意形状画像符号化において得られた知見は、「符号化器内部の自由度を利用して符号化効率を向上させる」というより一般的な問題へ発展した。

映像符号化国際標準 H.266/VVC では、復号器側の処理は規格によって厳密に規定されている一方、符号化器内部には一定の自由度が残されている。本研究では、この自由度を利用した符号化最適化について検討した。

4-2 変換係数摂動による最適化

VVC では、整数近似された変換行列が用いられるため、正変換と逆変換は厳密には一致しない。通常、VVC 参照ソフトウェア VTM では、逆変換行列の転置行列を正変換として利用しているが、整数化の影響により完全な直交変換とはなっていない。

本研究では、この性質に着目し、符号化器側で用いる変換係数を微小に調整することでレート歪最適化性能の改善を試みた。具体的には、変換行列要素に微小な摂動を与え、その結果生じる量子化係数の変化が符号量および復号画像品質に与える影響を解析した。復号器側の処理は一切変更せず、規格準拠性を維持したまま符号化効率改善を図ることが本手法の特徴である（図 6）。

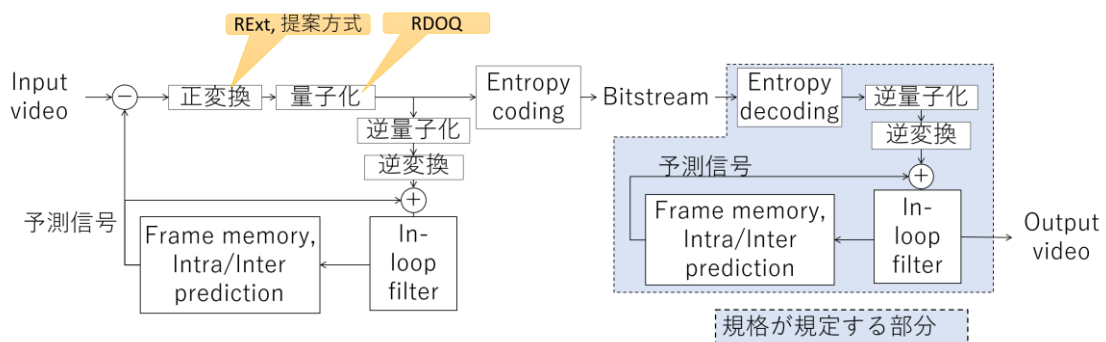


図 6 変換係数摂動符号化最適化の概念図

実験では、VTM に実装した提案手法を用いて標準評価シーケンスを符号化した。その結果、既存のレート歪最適量子化（RDOQ）を適用した場合でも、さらに Y-BD-rate で 0.20～0.35%程度の改善を確認した[13]。改善量自体は小さいものの、成熟した映像符号化規格において標準仕様を変更することなく性能向上を達成できることを示した点に意義がある。

4-3 被参照度マップに基づく局所 RD 最適化

さらに、画素間の予測依存関係に着目し、被参照度マップを利用した局所 RD 最適化について研究した。従来の映像符号化では、符号化単位ごとに単一のラグランジュ乗数を用いてレート歪最適化が行われる。しかし実際には、画像内の各画素が将来の予測に与える影響は様ではなく、重要度には大きな差が存在する。

提案方式では、イントラ予測過程において各画素が他画素からどの程度参照されるかを投票方式によって推定し、被参照度マップを生成する。被参照度が高い画素は後続画素の予測に大きな影響を与えるため重要度が高いと考えられる。一方、被参照度の低い画素は予測への寄与が小さい（図 7）。

そこで、本研究では被参照度に応じてラグランジュ乗数を局所的に変化させる方式を提案した。重要度の高い領域にはより多くの符号量を割り当て、重要度の低い領域では符号量を削減することで、画像全体として効率的なレート歪最適化を実現した。

実験では、VVC イントラ符号化において最大 0.38%の BD-rate 改善を確認した（表 1）[14]。また、被参照度マップはビットストリームから推定可能であり、追加のサイド情報を必要としないため、規格との互換性を維持したまま適用できる特徴を有する。

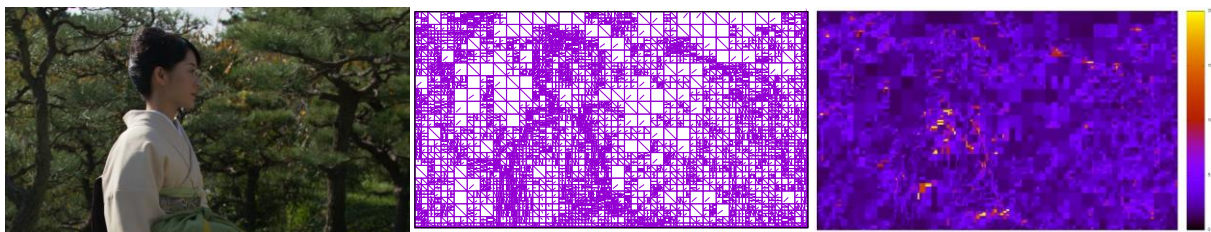


図 7 左：Kimono 原画、中：VVC 符号化後の Coding Unit 状態、右：導出された被参照度マップ

表 1 提案方式による BD-rate 改善結果

Sequence	Encoding pass	Y-BD-Rate ↓
Kimono	11	−0.17%
Parkscene	10	−0.01%
MarketPlace	23	−0.14%
EBUKidssoccer	30	−0.01%
Cactus	15	−0.05%
Vidyo	2	−0.23%
Jets	17	−0.38%

4-4 学術的・産業的意味

これらの研究は、

- 標準規格を変更しない
- ビットストリーム互換性を維持する
- 符号化器のみの改良で性能向上を実現する

という点に特徴がある。また、「構造を考慮した符号化最適化」という本研究の基本理念を、任意形状画像から一般映像符号化へ拡張した成果と位置付けられる。

本研究で提案した手法は、いずれも復号器側の変更を必要とせず、既存規格との互換性を維持したまま適用できる。このため、標準規格改訂を伴わずに実用システムへ導入可能であり、産業界における実装障壁が低いという利点を有する。また、符号化器内部の自由度を積極的に利用するという考え方は、今後の映像符号化規格や AI 支援符号化技術にも応用可能である。

5 まとめと今後の展望

本研究では、任意形状画像信号の符号化を中心課題として、V-PCC における DCR 補間、直交変換志向パディング、スパース信号分解型圧縮、VVC 符号化最適化について研究を行った。

特に、V-PCC においては、符号化器内部の予測機構を考慮した DCR 補間技術およびフレーム順序最適化技術により、従来方式を上回る圧縮性能を実現した。また、Occupancy 情報を利用した任意形状領域圧縮方式を提案し、特許出願に至った。さらに、本研究を通じて得られた知見は、VVC 符号化最適化へも展開され、変換係数摂動や局所 RD 最適化による符号化性能向上を実現した。

本研究の学術的意義は、従来は矩形画像を前提として発展してきた画像符号化技術に対し、任意形状画像という新たな信号モデルを導入し、その構造を積極的に利用する符号化技術を体系的に検討した点にある。特に、Occupancy 情報やパッチ構造といった信号固有の特徴を利用することで、従来方式では利用されていなかった情報を圧縮効率向上へ活用できることを示した。

また、本研究で得られた成果は、点群映像符号化だけでなく、動的メッシュ符号化、自由視点映像、AR/VR コンテンツ、医療画像における関心領域符号化など、任意形状領域を含む幅広い情報処理問題への応用が期待される。特に近年注目されているメタバースやデジタルツインでは、三次元空間情報を効率的に伝送・蓄積する技術が不可欠であり、本研究成果はその基盤技術の一つとして貢献できる可能性がある。

今後は、本研究で検討した任意形状画像符号化技術をさらに発展させ、深層学習を利用した予測技術や生成モデルとの融合を進める予定である。また、点群映像や動的メッシュのみならず、マルチモーダルセンシングデータや高次元情報への適用についても検討したい。さらに、MPEG 等の国際標準化活動への展開やリアルタイム符号化システムへの実装を通じて、実用的なイマーシブメディア通信技術への応用を目指す。

以上のように、本研究では任意形状画像信号の構造情報を利用した高効率符号化技術の有効性を示すことができた。これらの成果は、将来の高臨場感映像通信や三次元メディア処理を支える基盤技術として発展することが期待される。

【参考文献】

- [1] M. Preda, T. Zaharia, F. Preteux, "3D Media Compression for Virtual and Augmented Reality Systems," Wiley, 2018.
- [2] MPEG 3DG Ad-Hoc Group, "G-PCC and V-PCC Codec Description v2," ISO/IEC JTC1/SC29/WG7, m53622, Oct. 2019.
- [3] H. Schwarz, D. Marpe, T. Wiegand, "Overview of the Versatile Video Coding (VVC) Standard and its Applications," IEEE Transactions on Circuits and Systems for Video Technology, vol. 31, no. 10, pp. 3736–3764, Oct. 2021.
- [4] A. Fu, Z. Li, F. Li, et al., "Video-based Point Cloud Compression: A Survey," IEEE Transactions on Multimedia, vol. 25, pp. 3223–3243, 2023.
- [5] A. Dricot, G. Lafruit, "Smoothed Push-Pull Padding for Patch-based Point Cloud Compression," ISO/IEC JTC1/SC29/WG11 MPEG Input Document, m40092, Jan. 2017.
- [6] T. Wiegand, G. J. Sullivan, G. Bjøntegaard, A. Luthra, "Overview of the H.264/AVC Video Coding Standard," IEEE Transactions on Circuits and Systems for Video Technology, vol. 13, no. 7, pp. 560–576, Jul. 2003.
- [7] B. Bross, J. Chen, S. Liu, Y.-K. Wang, "Versatile Video Coding (Draft 10)," Joint Video Experts Team (JVET), JVET-S2001, Jul. 2020.
- [8] R. Rubinstein, A. Bruckstein, M. Elad, "Dictionaries for Sparse Representation Modeling," Proceedings of the IEEE, vol. 98, no. 6, pp. 1045–1057, Jun. 2010.
- [9] M. Aharon, M. Elad, A. Bruckstein, "K-SVD: An Algorithm for Designing Overcomplete Dictionaries for Sparse Representation," IEEE Transactions on Signal Processing, vol. 54, no. 11, pp. 4311–4322, Nov. 2006.
- [10] D. Taubman, M. Marcellin, "JPEG2000: Image Compression Fundamentals, Standards and Practice," Springer, 2002.

- [11] 藤原章大, 高村誠之, "V-PCC 符号化における効率的なパッチ間補間法," 画像電子学会誌, vol. 54, no. 3, pp. 304–312, July 2025.
- [12] 高村誠之, 特願 2025-147542, 「圧縮方法及びコンピュータプログラム」, 2025
- [13] 高村誠之, 手嶋悠人, "変換係数摂動による H.266/VVC 符号化最適化," 電子情報通信学会論文誌 D, vol. J109-D, no. 10, 掲載予定, Oct. 2026.
- [14] T. Mashino and S. Takamura, "Pixel-Wise Reference Voting-Based Optimization for VVC Intra Coding," Proceedings of the International Conference on Image Electronics and Visual Computing (IEVC 2026), Kitakyushu, Japan, Mar. 2026.

〈発 表 資 料〉

題 名	掲載誌・学会名等	発表年月
V-PCC 符号化における効率的なパッチ間補間方法に関する検討	PCSJ/IMPS2024	2024 年 11 月
Lossless image compression using multi-band prediction for visible and infrared spectra	Proc. SPIE, IWAIT2025	2025 年 1 月
Lossless Visible and Infrared Image Coding based on Multiple Inter-Band Predictors	Data Compression Conference (DCC 2025)	2025 年 3 月
V-PCC 符号化における効率的なパッチ間補間法	画像電子学会誌, Vol. 54, No. 3, pp. 304-312	2025 年 7 月
H. 266/VVC 符号化における変換係数調整に関する一検討	画像電子学会年次大会	2025 年 8 月
H. 266/VVC 符号化器における変換係数摂動に関する一考察	PCSJ/IMPS2025	2025 年 11 月
Pixel-Wise Reference Voting-Based Optimization for VVC Intra Coding	IEVC2026	2026 年 3 月
Quaternion-based Deformation of Cloth Point Clouds	IWAIT2026	2026 年 1 月
Quaternion-based Deformation of Cloth Point Clouds and Its Verification	Proc. SPIE, IWAIT2026	2026 年 1 月
変換係数摂動による H. 266/VVC 符号化最適化	電子情報通信学会論文誌 D, Vol. J109-D, No. 10	2026 年 10 月(採録決定)
圧縮方法及びコンピュータプログラム	特許出願 (特願 2025-147542)	2025 年