

動的メッシュ圧縮における変位量の量子化タイミングに関する検討 A Study on Quantization Timing of Displacements in Dynamic Mesh Compression

西村 仁志¹⁾ 加藤 晴久¹⁾ 河村 圭¹⁾
Hitoshi Nishimura Haruhisa Kato Kei Kawamura

1 はじめに

動的メッシュは時間的に頂点の接続関係が変化するテクスチャ付きの 3 次元メッシュで、VR のような没入型メディアでの利用が進んでいる。動的メッシュの幾何情報、つまり頂点の位置と接続関係はフレームごとに変化するため、データ量が膨大となる。従って、データの保管や伝送のためには、効率的な圧縮が必要不可欠となる。

2022 年 4 月に国際標準化団体 MPEG で動的メッシュ符号化の標準化が開始された (規格名: V-DMC)。参照ソフトウェアは VSMC [1] に基づいている。図 1 に示すように、入力メッシュを基本メッシュと変位量に分割して効率的な符号化を行っている。基本メッシュは、入力メッシュを間引いたものである。変位量は、基本メッシュを細分割し、その細分割した頂点と入力メッシュ表面との最近傍距離によって算出される。参照ソフトウェアでは、変位量を映像コーデックで符号化するか、直接算術符号化するかを選択できる。前者は、AVC [2]、HEVC [3] のような映像コーデックは既に普及が進んでいるため、既存のハードウェアを利用できることが利点である。一方、後者は我々が提案したものであり [4]、映像コーデックのような複雑な処理がなく計算量が少ないが、効率的に符号化できる点が利点である。直接算術符号化する手法では、符号化対象フレームの変換係数は参照フレームからインター予測し、残差のみを符号化する。しかし、インター予測の前に各フレームの変換係数を量子化するため、参照フレームと符号化対象フレームで 2 回分の量子化誤差が発生し、残差が大きくなることがある。

本論文では、量子化誤差の発生回数を抑制するため、先にインター予測を行った後に残差を量子化する手法を提案する。量子化された残差は、階層的に算術符号化することによって、少ない計算量で効率的に符号化する。提案手法によって、インターフレームを含むシーケンスにおいて符号化効率が改善することを示した点が本論文の貢献点である。また、インターフレームの割合が増えれば増えるほど、提案手法の有効性が向上することを確認した点も本論文の貢献点と言える。

2 関連研究

メッシュは幾何情報とテクスチャ情報で構成されている。本章では、提案手法のポイントであるメッシュの変位量と関連する幾何情報に焦点を当て、静的メッシュと動的メッシュに分けてメッシュ符号化の関連研究を紹介する。

2.1 静的メッシュ符号化

Draco [5] は静的メッシュ符号化の代表的な手法である。Draco では、頂点の接続関係は Edgebreaker [6] を用いて可逆で符号化し、頂点の位置は複数の平行四辺形に基づく予測 [7] を用いて可逆あるいは非可逆で符号化

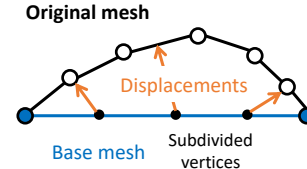


図 1: 基本メッシュと変位量に分割したメッシュ符号化。

する。Draco は多様体のメッシュのみしか符号化できないが、TFAN [8] は 1 つの頂点を共有する三角形のファン構造を利用することによって非多様体のメッシュを符号化することができる。しかし、これらの手法はフレーム内の情報しか利用できないため、時間的な相関が高い動的メッシュを効率的に符号化することができない。

2.2 動的メッシュ符号化

動的メッシュを効率的に符号化するためには時間的な相関を利用することが重要である。メッシュを 2 次元画像に変換し、映像コーデックによって符号化する手法が提案されている [9, 10]。しかし、これらの手法では、2 次元画像への投影誤差が発生するため、頂点の接続関係を効率的に符号化することができない。メッシュを 3 次元のままブロックに分割し、2 次元の映像コーデックと同様の予測を利用した手法も提案されている [11, 12]。しかし、これらの手法でも頂点の接続関係を効率的に符号化することができない。

近年の手法 [13, 1, 4] では、各フレームでは Draco でメッシュを符号化し、別の手法によって時間的な相関を利用している。Graziosi ら [13] は、visual volumetric video-based coding (V3C) 規格 [14] に基づいて時間的な相関を利用しながら頂点の位置を符号化する手法を提案している。VSMC [1] は、メッシュの間引きと細分割を導入し、高い符号化効率を示している。VSMC では、各細分割頂点に対する変位量を算出し、ウェーブレット変換を行う。得られた変換係数は量子化し、2 次元の画像に変換した後、映像コーデックで時間的な相関を利用して符号化する。我々は、VSMC に対して、映像コーデックを用いずに変位量を直接算術符号化 [15] する手法を提案している [4]。そこでは、符号化対象フレームの変換係数は参照フレームからインター予測し、残差のみを符号化する。しかし、インター予測の前に各フレームの変換係数を量子化するため、参照フレームと符号化対象フレームで 2 回分の量子化誤差が発生し、残差が大きくなる。

3 提案手法

図 2 は、提案する変位量符号化器の構成を示す。まず、入力変位量に対してウェーブレット変換を適用し、ウェーブレット変換係数を算出する。フレーム種別がイントラの場合は、符号化対象フレームの係数をそのまま次の処理に入力する。フレーム種別がインターの場合は、参照フレームの対応する係数から符号化対象フレ

1) 株式会社 KDDI 総合研究所

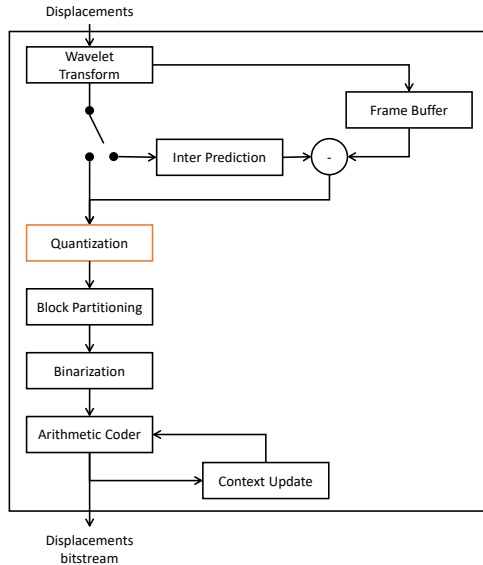


図 2: 提案する変位量符号化器の構成。量子化をインター予測の後に進めることが提案手法の特徴である。

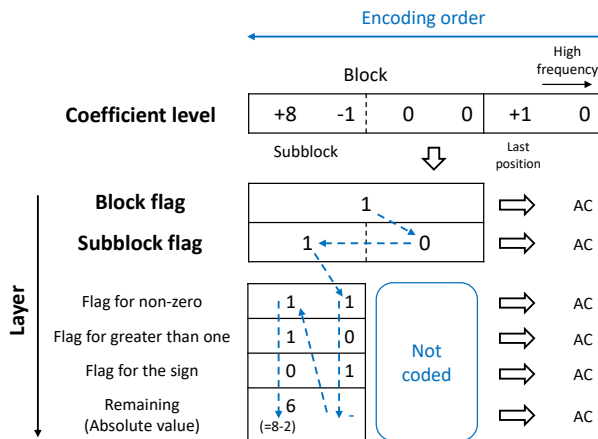


図 3: 提案する変位量符号化の例。"AC" は算術符号化器を示す。レベル値はサブブロックのフラグが 1 の場合にのみ符号化する。

ムにおける係数を予測し、残差を次の処理に入力する。次に、残差（あるいは係数）を量子化し、レベル値を算出する。そして、後述のブロック及びサブブロックに基づいてレベル値を二値化する。最後に、算術符号化器によってビットストリームを生成する。

提案手法は、複数の層を用いることで少ない計算量で効率的にレベル値を符号化できる階層的符号化 [4] に基づいている。図 3 は、提案する変位量符号化の例を示す。以下に、階層ごとのシンタックス要素を説明する。まず、レベル値全体に対して、高周波領域には多くの零が連続するため、非零レベル値の終端位置を導入する。なお、非零終端位置を特定するための実装が簡易なことから、レベル値は高周波から低周波の順に符号化する。次に、細分割の繰り返し回数を示す詳細レベル (LoD) に基づいてレベル値をブロックに分割する。各 LoD をブロックとして定義し、ブロック内に非零レベル値があるかどうかを示すブロックフラグを導入する。ブロックフラグが 0 の場合は、ブロック内のレベル値は符号化しない。さらに、各ブロックをより小さなサブブロックに

表 1: VSMC を基準とした際の全シーケンス平均 BD-Rate [%].

		D1	D2	Luma	Cb	Cr
AI	[4]	-0.2	-0.2	-0.8	-0.9	-0.9
	提案手法	-0.2	-0.2	-0.8	-0.9	-0.9
RA	[4]	-0.4	-0.4	-0.8	-0.8	-0.7
	提案手法	-1.5	-1.5	-1.4	-1.5	-1.4

分割し、同様にサブブロックフラグを導入する。サブブロックのサイズはあらかじめ定められた 2 以上の整数値とする。レベル値に 0 が多い場合は、サブブロックサイズを大きくした方が計算量を削減し符号化効率を向上させる効果が期待できる。そして、非零であるかどうか、2 以上であるかどうか、レベル値そのものの 3 種類を用いてレベル値を二値化する。正負符号はフラグとして二値化する。

レベル値の発生頻度の違いを考慮して、算術符号化に用いるコンテキストモデルは各シンタックス要素、各フレーム種別、そしてブロックごとに分割して定義する。特に、各ブロックで分割してコンテキストモデルを定義することによって、独立かつ並列に各ブロックを符号化することができる。各コンテキストモデルは、各レベル値を符号化するたびに更新し、算術符号化の効率を高める。

4 実験

提案手法の有効性を検証するため、V-DMC の Call for Proposal [16] と同様の実験条件で実験を行った。

4.1 実験条件

データセットは 1 名の人物を対象としたシーケンス 8 つで構成されており、各シーケンスは 150 あるいは 300 フレームが含まれている。ビットレート条件は 3 から 25 [Mbps] の範囲の 5 種類 (R1~R5) とし、符号化モードは All Intra (AI) モードと Random Access (RA) モードの 2 種類とした。AI モードのビットレート条件 R5 では、多くのフレームでレベル値が全て 0 になるため、変位量符号化は行わないように設定した。比較手法として、VSMC [1] と従来手法 [4] を用いた。VSMC では、変位量符号化に用いる映像コーデックとして HEVC を用いた。VSMC の実装はバージョン 3.0 を用いた。従来手法 [4] 及び提案手法において、サブブロックサイズは 100 に設定した。提案手法では、算術符号化の実装には Dirac [17] を用いた。

符号化効率は、VSMC と提案手法間の Bjøntegaard-delta (BD) rate [18] で評価した。BD-Rate は、2 つの手法に対して、同等の PSNR を実現するために必要な符号量を比較したものであり、小さいほど比較手法よりも性能が良い。PSNR は、基本メッシュと変位量を統合したメッシュ全体に関して、符号化前と復号後の間で算出した。PSNR の指標には、メッシュの位置を示す D1 と D2、メッシュの色を示す輝度、青色差、赤色差を用いた。D1 は点と点間の平均二乗誤差を示し、D2 は点と面間の平均二乗誤差を示す。なお、提案手法は、VSMC と比較して PSNR は低下しないが、ビット量は減少する。これは、両手法では変位量は可逆で符号化され、その他のアルゴリズムは同様であるためである。

表 2: 各シーケンスにおけるインターフレームの割合と D1 BD-Rate の関係。

	levi	thomas	mitch	soldier
インターフレームの割合	25%	65%	65%	80%
BD-Rate [%] [4]	-0.6	-0.7	-0.6	-0.7
提案手法	-0.7	-2.1	-1.9	-6.4

4.2 実験結果

提案手法の符号化効率を評価した。表 1 は、VSMC を基準とした際の全シーケンス平均 BD-Rate を示す。AI モードでは、提案手法の BD-Rate は従来手法 [4] と変わらなかった。これは、AI モードではインターフレームが存在せず、インター予測が行われなため、量子化タイミングの変更による影響がなかったからだと考えられる。一方、RA モードでは、従来手法 [4] と比較して、提案手法の BD-Rate は全ての指標において改善が見られた。また、表 2 は、各シーケンスにおけるインターフレームの割合と D1 BD-Rate の関係を示す。インターフレームの割合が増えれば増えるほど、提案手法の有効性が向上することが分かる。

次に、インターフレームにおいてビット数が減少する例 (thomas シーケンス) を詳細に調査した。図 4(a), 4(b) は、あるフレームにおけるレベル値の例を示す。図 4(a) は従来手法、図 4(b) は提案手法の例を示す。この例では、従来手法と比較して、提案手法では特に低周波領域において -1 のレベル値が 0 に変化していることが分かる。これによって、正負符号の符号化が不要となり、算術符号化の効率が向上したと考えられる。また、図 5(a), 5(b) は、2 フレーム分のレベル値を符号化する例を示す。図 5(a) は従来手法、図 5(b) は提案手法の例を示す。フレーム 619 において、従来手法では 1 が算術符号化されているのに対して、提案手法では 0 が算術符号化されている。これによって符号化効率が向上していると考えられる。ただし、上記は一例であるため、今後は他の例も含めて統計的に傾向を分析する予定である。

5 おわりに

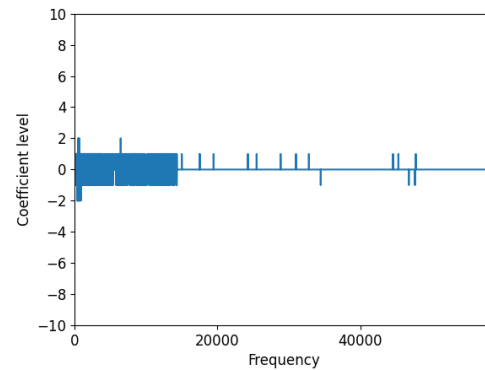
本論文では、動的メッシュ符号化における変位量の量子化誤差の発生回数を抑制するため、先にインター予測を行った後に残差を量子化する手法を提案した。実験では、RA モードにおいて、従来手法と比較して提案手法の BD-Rate が改善することを示した。また、インターフレームの割合が増えれば増えるほど、提案手法の有効性が向上することを確認した。今後は、標準規格採用に向け、提案手法による主観画質への影響を評価する。

謝辞

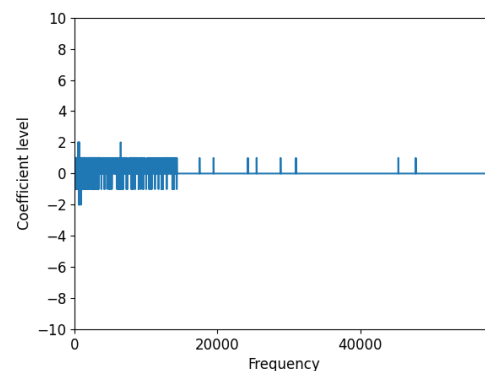
本研究開発は総務省 SCOPE (国際標準獲得型) JPJ000595 の委託を受けたものです。

参考文献

- [1] Khaled Mammou, Jungsun Kim, Alexis M Tourapis, Dimitri Podborski, and David Flynn, "Video and subdivision based mesh coding," in *10th European Workshop on Visual Information Processing (EUVIP)*, 2022, pp. 1–6.
- [2] Thomas Wiegand, Gary J Sullivan, Gisle Bjontegaard, and Ajay 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, 2003.
- [3] Gary J Sullivan, Jens-Rainer Ohm, Woo-Jin Han, and



(a) 従来手法 [4].



(b) 提案手法.

図 4: あるフレームにおけるレベル値の例。シーケンス: thomas, フレーム: 619 (インターフレーム), モード: RA, ビットレート: r1. 横軸はウェーブレット係数の周波数, 縦軸は横軸に対応する係数のレベル値を示す。

Thomas Wiegand, "Overview of the high efficiency video coding (HEVC) standard," *IEEE Transactions on circuits and systems for video technology*, vol. 22, no. 12, pp. 1649–1668, 2012.

- [4] Hitoshi Nishimura, Kei Kawamura, and Koki Kishimoto, "Block-based context-adaptive arithmetic coding of displacements," ISO/IEC JTC1/SC29/WG7/m62115, 2023.
- [5] Google, "Draco," <https://github.com/google/draco/>.
- [6] Jarek Rossignac, "Edgebreaker: Connectivity compression for triangle meshes," *IEEE transactions on visualization and computer graphics*, vol. 5, no. 1, pp. 47–61, 1999.
- [7] Costa Touma and Craig Gotsman, "Triangle mesh compression," in *Proceedings-Graphics Interface*, 1998, pp. 26–34.
- [8] Khaled Mamou, Titus Zaharia, and Françoise Prêteux, "TFAN: A low complexity 3D mesh compression algorithm," *Computer Animation and Virtual Worlds*, vol. 20, no. 2-3, pp. 343–354, 2009.
- [9] Hitoshi Habe, Yosuke Katsura, and Takashi Matsuyama, "Skin-off: Representation and compression scheme for 3D video," in *Picture Coding Symposium*, 2004, pp. 301–306.
- [10] Junhui Hou, Lap-Pui Chau, Ying He, and Nadia Magnenat-Thalmann, "A novel compression framework for 3D time-varying meshes," in *2014 IEEE International Symposium on Circuits and Systems (ISCAS)*. IEEE, 2014, pp. 2161–2164.
- [11] Seung-Ryong Han, Toshihiko Yamasaki, and Kiyoharu Aizawa, "Time-varying mesh compression using an ex-

		Frame 618 (Intra)	Frame 619 (Inter)
Input Coefficient		-1.415	-0.504
Encoder	Quantization	-1	0
	Inter prediction Subtraction		1
Arithmetic coding		-1	1
Decoder	Inter prediction Addition		0
	Inverse quantization	-1.260	0

(a) 従来手法 [4].

		Frame 618 (Intra)	Frame 619 (Inter)
Input Coefficient		-1.415	-0.504
Encoder	Inter prediction Subtraction		0.755
	Quantization	-1	0
Arithmetic coding		-1	0
Decoder	Inverse quantization	-1.260	0
	Inter prediction Addition		-1.260

(b) 提案手法.

図 5: 2 フレーム分のレベル値を符号化する例. シーケンス: thomas, モード: RA, ビットレート: r1.

- tended block matching algorithm,” *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 17, no. 11, pp. 1506–1518, 2007.
- [12] Toshihiko Yamasaki and Kiyoharu Aizawa, “Patch-based compression for time-varying meshes,” in *17th IEEE International conference on image processing*, 2010, pp. 3433–3436.
- [13] Danillo Bracco Graziosi, “Video-based dynamic mesh coding,” in *28th IEEE International Conference on Image Processing (ICIP)*, 2021, pp. 3133–3137.
- [14] “Information technology – coded representation of immersive media – Part 5: Visual volumetric video-based coding (V3C) and video-based point cloud compression (V-PCC),” ISO/IEC 23090-5:2021, 2021.
- [15] Jorma Rissanen and Glen G Langdon, “Arithmetic coding,” *IBM Journal of research and development*, vol. 23, no. 2, pp. 149–162, 1979.
- [16] “CfP for dynamic mesh coding,” ISO/IEC JTC1/SC29/WG7/N00231, 2021.
- [17] Tim Borer and Thomas Davies, “Dirac video compression using open technology,” *BBC EBU technical review*, vol. 303, 2005.
- [18] Gisle Bjontegaard, “Calculation of average PSNR differences between RD-curves,” *VCEG-M33*, 2001.