ユーザの視線動作に応じたインタラクティブな映像デコード手法

河野 翔太　　中村 朋生　　門本 淳一郎　　入江 英嗣　　坂井 修一

概要.　近年盛んに研究されている VR や AR といった映像技術を高品質かつより少ない計算資源で提供するために，映像の計算コストの削減を行う．映像の計算コストの削減方法の 1 つである Approximate Computing は計算コストを削減するが，出力映像の品質の低下をもたらす．Approximate Computing の適用対象をユーザの視界の範囲外に絞ることで，知覚品質の維持と計算コストの削減の両立が達成できる．本研究では，ユーザの状態と映像の情報を利用して，ユーザの知覚品質を維持して効率的に映像をデコードする手法を提案する．提案手法を実装し評価実験を行った結果，映像の動きが少ない場合にはユーザの知覚品質に影響を及ぼさずに計算コストを削減できることがあること，及び，動きベクトルを用いてデコード領域を調節すると，動きの激しい映像でもユーザの知覚する品質を低下させにくくできることが分かった．

1　はじめに

近年，Virtual Reality (VR)，Augmented Reality (AR) といった映像技術が盛んに研究されている．これらをより高品質に実現する場合，要求される計算資源は際限がない．ユーザの視界に合わせた最適な品質を決定することで，ユーザの知覚する品質を維持し，計算コストを削減できる可能性がある．

映像の計算コストの削減方法の 1 つに，Approximate Computing (AC) がある．AC は，計算誤差を許容し，計算コストの削減を得る計算パラダイムである [1, 5]．映像の全ての部分に AC を適用すると，計算コストは削減されるが，出力される映像の品質が低下する．AC の適用対象をユーザの視界の範囲外に絞ることで，知覚される品質を維持できる．

計算コストの削減については，実行時に動的に近似積極性を変更できるアーキテクチャフレームワーク [2] により達成できる．ユーザが現在最も必要としている箇所を特定すれば，この手法を利用し，その箇所に計算資源を集中させることができる．

本研究では，ユーザの知覚品質を維持しつつ，効率的に映像をデコードする手法を提案する．本研究の提案手法は，ユーザの視線の位置に加え，その時間変化と映像の情報を利用し，ユーザの要求する計算品質を過不足なく満たす．これにより，ユーザの知覚品質を維持した計算コストの削減を達成する．

2　提案手法

システムはユーザの視線位置を取得し，その時間変化と動画の情報からユーザの知覚しやすい領域を優先的にデコードする．ユーザの知覚しやすい領域はヒトの有効視野に基づいて推測され，視線を中心に図 1 に示す領域をデコードする．この領域の大きさは，以下の 3 点に応じて変動する．

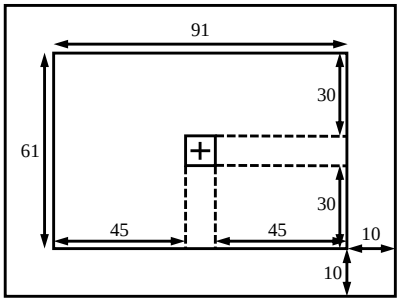

図 1. B フレームにおいてデコードされるマクロブロックの範囲と視線位置の関係.

1 つめは，ユーザ視線の位置変化から計算されるサッケードの情報である．瞬間的な視線移動であるサッケードの最中ではデコードする領域の縦横の長さをそれぞれ 2 分の 1 にする．フィクセーションの閾値は 20 °/s から 40 °/s の間に存在するため [4]，30 °/s 以上の速度の視線動作をサッケードとする．

2 つめは，フレームが含む動きベクトルの大きさの平均である．B フレームのデコードを削減した場合，動きの激しい場面でフレーム周辺部での輝度の変化が急激になる．これによる視聴体験の低下を防ぐため，動きベクトルの大きさを利用し，動きが激しい場面ではデコード領域を大きくする．

3 つめは，DCT 変換係数行列に含まれる高周波成分の値である．網膜周辺部は輝度を知覚しやすいため，図 1 のデコード範囲外周部では，輝度の急激な変化となるマクロブロックのみデコードする．マクロブロックが含む 1 つ以上のブロックの DCT 変換行列について，行と列の添え字の合計が 10 より大きい位置に非零の成分が存在する場合，そのマク

Copyright is held by the author(s). This paper is non-refereed and non-archival. Hence it may later appear in any journals, conferences, symposia, etc.

東京大学大学院情報理工学系研究科

表 1. 実験で使用した 4 通りの手法.

	手法			
	gaze_only	area_size	high_freq	full
視線周辺の選択的デコード	○	○	○	○
サッケード時のデコード領域の縮小		○		○
動きベクトルによるデコード領域の調節				
高周波マクロブロックの選択的デコード			○	○

ロブロックをデコードする.

3 評価

3.1 実験環境

提案手法をアイトラッカーと通信するデスクトップアプリケーションとして実装した. 実装にはC++言語, Tobii Pro SDK, そして Tobii Pro Fusion[3] のサンプリングレート 120 Hz のモデルを使用した. 実験時には, 解像度縦 1440 px, 横 2560 px のディスプレイを用い, 被験者の目とディスプレイ間の距離が 60 cm となるようにした.

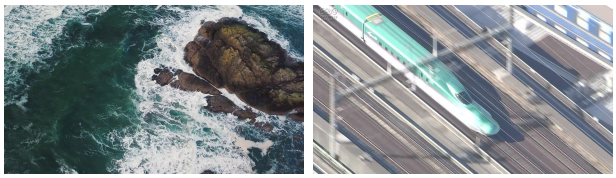

図 2. 実験に用いた 2 映像の第 1 フレーム. (左): 映像 (a), Ruvim Miksanskiy による. (右): 映像 (b), NHK Creative Video Library による.

実験では, 表 1 に示す 4 通りの手法でデコードした同一の映像を被験者に見せた. 実験に用いた映像の第 1 フレームを図 2 に示す. 映像 (a) は動きが少なく, 一貫して一定の画角で撮影されている. 映像 (b) は動きが激しいうえ, 画角の変更やシーンチェンジもあり, フレーム間での色の変化が激しい.

3.1.1 実験結果

図 3 及び図 4 に, 映像 (a) 及び映像 (b) に対する被験者 4 名の評価を示す.

映像 (a) では, 手法 gaze_only 及び high_freq について, 被験者によっては近似なしと同一の評価が与えられた. 動きの少ない映像 (a) では, 近似を行ったとしてもユーザの知覚品質に影響を及ぼさずに計算コストを削減できることがあるといえる.

一方映像 (b) では, 手法 gaze_only 及び high_freq の評価は低下した. これは, 映像 (b) の動きの激しさにより近似を行った際にユーザの知覚品質に影響を与えたといえる. しかし, 手法 area_size 及び full に関しては, 映像 (a) の評価を維持した. このことから, デコード領域の大きさを調節すると, 動きの

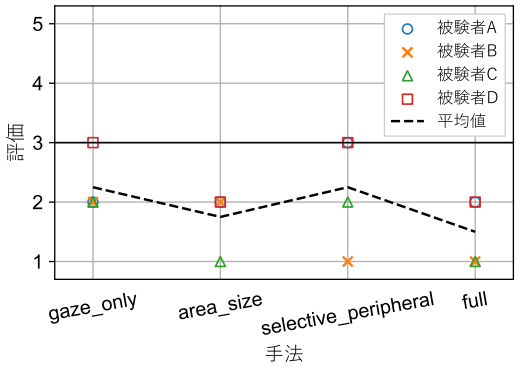

図 3. 映像 (a) に対する被験者の評価.

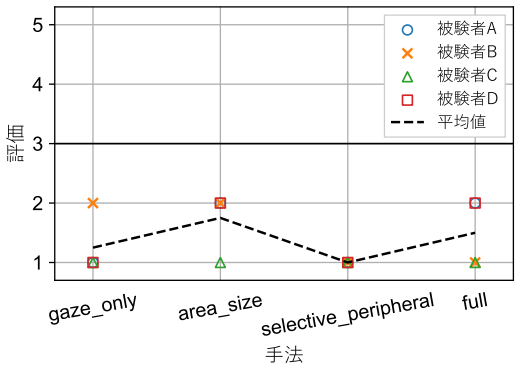

図 4. 映像 (b) に対する被験者の評価.

激しい映像でもユーザの知覚する品質を低下させにくくなることが分かる.

4 おわりに

本稿では, ユーザの視線に応じてインタラクティブに近似の積極性を動的制御する近似デコード手法を提案した. また, 提案手法を実装し, 評価実験を行った. この評価を通じて, 以下の 2 つが明らかになった. 1 つめは, 映像の動きが少ない場合, 近似を行ったとしてもユーザの知覚品質に影響を及ぼさずに計算コストを削減できることがある点である. 2 つめは, 動きベクトルを用いてデコード領域の大きさの調節を行うと, 動きの激しい映像でもユーザの知覚する品質を低下させにくくなる点である.

謝辞

本研究の一部は，JST さきがけ JPMJPR20M1，共同研究（富士通株式会社）「データ転送性能ボトルネックを解消する新アーキテクチャ研究」，JSPS 科研費 JP21J11687 による．

参考文献

[1] S. Mittal. A Survey of Techniques for Approximate Computing. *ACM Comput. Surv.*, 48(4):1–33, mar 2016.

[2] T. Nakamura, K. Tomida, S. Kouno, H. Irie, and S. Sakai. Stochastic Iterative Approximation: Software/hardware techniques for adjusting aggressiveness of approximation. In *2021 IEEE 39th International Conference on Computer Design (ICCD)*, pp. 74–82, 2021.

[3] Tobii Pro AB. Tobii Pro Fusion. `https://www.tobiipro.com/ja/product-listing/tobii-pro-fusion/`, 2022.

[4] Tobii® Technology. White Paper: Determining the Tobii I-VT Fixation Filter's Default Values. Technical report, August 2012.

[5] Q. Xu, T. Mytkowicz, and N. S. Kim. Approximate Computing: A Survey. *IEEE Design & Test*, 33(1):8–22, 2016.

未来ビジョン

VR や AR といった映像技術や，それを実現するデバイスは，まだまだ我々の手には届きにくいものである．しかし，それらがもしも非常に小さい計算コストで実現できれば，それらは低廉な計算機でも実現できることになり，結果としてより多くの人々が手に取ることができる．本研究では，映像の近似化を行っている．現時点では，本研究による性能の改善効果と，ユーザの注目度の推定は，非常にリアルな VR 映像を，認知の上で劣化させることなく低廉なノート PC で実現するような程ではない．しかし，将来的には，ハードウェアなどとの連携も含めて性能を強化すると同時に，よりユーザと近似システムのインタラクティブ性を強化することで，非常に小さな計算機で高品質なメタバース体験を実現できるようになると考える．そして，自分はその第一歩として研究に取り組んでいる．