

ROI 符号化を活用した ゲームキャラクタ領域の映像品質制御の研究

澤登光弘^{†1} 清原良三^{†2} 寺島美昭^{†1}
創価大学^{†1} 神奈川工科大学^{†2}

1. はじめに

低ビットレートなゲーム映像では、視聴者が関心のあるキャラクタ領域が画質劣化する。動画圧縮規格 H.264 では、関心領域(ROI)について考慮された ROI 符号化を利用でき、マクロブロックである 16x16px の領域毎に ROI の量子化パラメータ(QP)を低下させることで、ROI の画質を向上させる手法が Ferreira ら[1]により提案された。一方で、H.264 の ROI 符号化は、ROI の品質を向上させた分、非 ROI の画質が低下するため、トレードオフの関係にある。

だが、使用する際は視覚的に品質向上の効果が薄いにも関わらず、関心領域に多くの符号が割り当てられることが考えられるため、使用者が必要な品質分のみの QP の補正値を決定できる方法が必要である。本研究では、FPS(First Person Shooting)ゲーム映像を対象に、基準とする品質評価値に対応する QP の補正値を推定する手法を提案する。

2. ROI 符号化について

本研究で利用する ROI 符号化は、Ferreira らと同様の手法を利用した。機能の実装方法として、FFmpeg 内の `adroi` フィルタを用いて行う。QP の変更方法は、フィルタ内パラメータである QP 補正値 `qoffset` を -1 から 1 の範囲を実数値で選ぶ。qoffset は、ROI の QP について、圧縮規格の QP 範囲内でどれだけの割合を、非 ROI の QP から変化させるかどうかを決定するものである。値が小さいほど、指定した領域でより良い品質となる[2]。本研究では、映像圧縮規格として 8bit H.264 を使用する。8bit 下での QP の範囲は、0 から 51 であるから、ROI の QP である QP_{ROI} の計算は次のようになる。ただし、非 ROI の QP を QP_{bz} とする。

$$QP_{ROI} = QP_{bz} + qoffset * (51 - 0)$$

QP の範囲は、圧縮規格の扱うビット深度によって異なる。そのため、本研究では、QP の値ではなく QP の補正値(qoffset)をエンコードにおける要素として扱う。本研究の映像品質評価は SSIM(Structural Similarity)を用いる。SSIM 値が 1.0 に近いほど元映像から劣化が少ない指標である。

先行研究として、L.Yang ら[3]は SSIM と QP の関係を分析し、SSIM-QP モデルによる品質制御について提案した。L.Yang らの SSIM-QP 推定曲線では、QP が 30 未満の場合、主観的品質向上の効果が薄いことを示し、SSIM は 1 に収束するモデルであった。このモデルに使用したデータは 352x288px の映像符号化評価用標準動画画像である。そのため、ゲーム映像では激しい動きが起こるため、予測モデルが変わることが考えられる。また、基準 SSIM に対して必要な QP だけ割り当てることが課題とされている。本研究では、ROI における SSIM の必要量をビットレートが許す限り

SSIM=0.98 とし、非 ROI に対しては最大の SSIM を取る qoffset を考える。これは、小箱[4]による SSIM と主観評価の関係性を参考にした。この関係は、SSIM の値が 0.98 以上であれば、元画像と圧縮画像の区別がなくなり、0.90 以下であると明らかに劣化がわかるとされている。

3. 提案

本研究では、FPS ゲーム映像内の関心領域(ROI)の解像度とその個数、映像ビットレートを用いて、基準 SSIM に近い SSIM に対応する QP の補正値 qoffset を得た後、より細かく qoffset を推定する手法を提案する。これは後述する推定モデルから得た曲線の特徴を活かした。

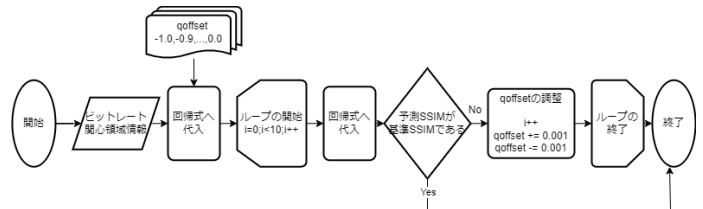


図2 推定フロー

図2より、始めにビットレートと ROI 情報(解像度と個数)を代入した回帰式に、qoffset を -1 から 0 の範囲かつ 0.1 刻みで離散的に代入する。その後、基準 SSIM に最も近い qoffset の予測 SSIM と基準 SSIM (=0.98)の大小関係によって、qoffset を調整する。この調整は、予測 SSIM が基準 SSIM を下回る場合には現在の qoffset から刻み値だけ低下させ、予測 SSIM が基準 SSIM を上回る場合には取得した qoffset から刻み値だけ上昇させる。この調整は最長で 10 回行い、刻み値は 0.001 とした。最終的に得られた qoffset を QP の補正値として使用する。なお、qoffset が -1 以下の場合には -1 として出力する。

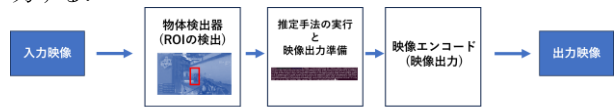


図3 提案システム

図3に推定手法を取り入れた配信システムを示す。入力映像は 3 秒のセグメント形式の映像を想定し、出力映像は ROI 符号化した映像を出力する。まず、物体検出器によって、入力映像から ROI とするキャラクタ領域を検出する。この領域情報から、推定手法によって、関心領域の QP を決定し、映像出力準備において、映像エンコードのためにラッパー処理を行い、映像エンコード部で映像を出力する。推定アルゴリズムを作成するために、SSIM を従属変数とした重回帰分析を行った。着目する点は、使用するゲーム、ROI 解像度及び個数の変化である。これは、使用ゲームによる品質向上の効果の変化や、映像上でのキャラクタ数・大きさの変化による qoffset の変動を考慮した。

A study of video quality control of game character domain using ROI encoding

^{†1} MITSUHIRO SAWANOBORI, Soka University

^{†2} RYOZO KIYOHARA, Kanagawa Institute of Technology

^{†3} YOSHIAKI TERASHIMA, Soka University

4. 回帰分析による推定モデル

従属変数は、SSIM とし、独立変数は、目標出力ビットレート、qoffset、ROI 解像度、ROI の個数とした次元数 2 の重回帰分析を行った。データセットは、ビットレート、ROI 解像度とその個数を変化させた 4334 個の ROI のデータを使用した。p<0.05 のとき、有意でない場合とし、該当する変数を取り除いた。図 4 にビットレートを 700kbps、ROI 解像度を 200×400px、ROI 数を 1 としたときの qoffset と SSIM のグラフを示す。図 4 より、最も品質が向上する qoffset は、必ずしも最も低い値から決定することはできないこと、SSIM が最大値になる qoffset の値から 0 までは必ず SSIM の値が下降することが分かった。

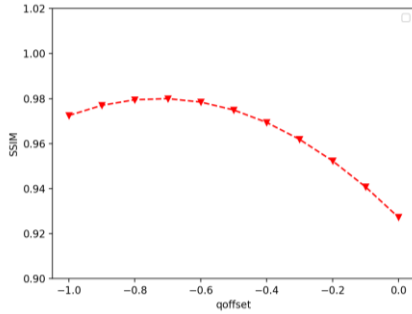


図 4 SSIM-qoffset 推定グラフ

5. 推定手法の性能評価

提案手法について、QP の補正値の予測結果を用いて 14~91 フレームの 32 個の映像を出力し、SSIM の実測値との誤差を評価するために実験を行った。なお、誤差は相対誤差を表し、正の値であれば予測 SSIM が実測 SSIM よりも大きい。図 5 に実験結果の分布図を示し、ROI 数が 2 以上の場合、同じ元映像である結果に直線を結んだ。また、表 1 に ROI 数毎の平均誤差を示す。また、ROI 数が 3 の場合は標準偏差が他の ROI 数と比較し 2.22~2.23 倍になった。

表 1 ROI 数と平均誤差

ROI 数	データ数	平均誤差	標準偏差
1	20	0.0167	0.0315
2	10	0.0264	0.0314
3	15	0.0547	0.0701

6. 配信システムの処理時間

試作した提案システムを用いて、提案手法の実現可能性を評価するために、処理時間を計測した。物体検出アルゴリズムに YOLOv7 を使用し配信システムを実装した。使用した映像の解像度は 1920×1080px、フレームレートは 60fps、映像時間は 3 秒である。目標出力ビットレートは 700kbps である。表 2 に、各処理部分について処理時間を示す。1 フレーム毎の物体検出の時間について、計測フレームは 3990 フレームであり、その他の計測項目のデータ数は一律で 22 である。表 2 より、1 フレーム毎の物体検出にかかる時間は 0.262 秒であり、180 フレームの動画の場合では 47.2 秒である。

表 2 システムの処理時間

計測項目	時間(s)
1フレーム毎のキャラクタ検出時間	0.262
1映像毎の推定手法と映像出力準備時間	0.0128
1映像毎の映像エンコード時間	1.75
1映像毎の総出力時間	49.2

7. 結論

本研究では、ゲーム映像における関心領域の画質劣化を抑制するために、ROI 符号化を利用した際の関心領域毎に QP の補正値を決定する推定手法を提案した。

推定手法で用いる予測モデルを作成するため、重回帰分析を行った。結果として QP の補正によって SSIM が向上したが、不適切な QP の補正は品質低下を招くため、QP 補正値の設定が重要であることが分かった。

推定手法の性能評価では、ROI 数が 1 つの正確度は他の場合と比べ高く、QP の補正がわずかに不足したが、多くの場合で十分な品質で映像を出力できた。一方で ROI が複数の場合、QP の補正不足が増加し、予測難度が増加した。特に ROI 数が 3 つの場合では他の場合と比較し、精度は半減した。また、映像の局所的な動きの激しさは考慮していないため、圧縮技術で扱われている動き補償やシーン検出のようなフレーム間での局所的な動きに関するファクターを導入する必要がある。

試作システムを用いた推定手法の実現可能性を測る実験では、ライブ配信を想定した場合に推定手法の処理時間は 0.0128 秒以下であり、遅延に関与しないため、推定手法自体に実現可能性はある。一方で、物体検出の処理時間が全体の出力時間の 96%の割合を占めているため、処理するフレームを限定することで処理時間を速める必要がある。

今後の課題は、推定手法を組み込んだ配信システムの実装上の問題を解決することである。現時点では、フレーム毎にフィルタを掛けた場合、途中のフレームが欠落する問題を確認した。

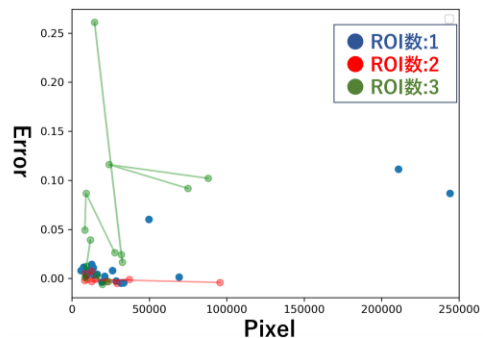


図 5 ピクセル数と推定手法に関する誤差

参考文献

- [1] L. Ferreira, L. Cruz and P.A. Assunc.ao, "H. 264/SVC ROI ~ encoding with spatial scalability," in Proc. of International Conference on Signal Processing and Multimedia Applications, 2008, p. 212-215.
- [2] FFmpeg - AVRegionOfInterest Struct Reference <https://ffmpeg.org/doxygen/trunk/structAVRegionOfInterest.html#ae7421571aad4d48b63e91274ba424ee0> (閲覧日:2023年12月15日)
- [3] L. Yang, L. Zhang, et al, "A ROI quality adjustable rate control scheme for low bitrate video coding", in Picture Coding Symposium, Chicago, U.S., 2009
- [4] 小箱 "電子化文書の画像圧縮ガイドライン" 標準化への道 標準化委員会報告 3, Vol.50, No.5, 2011-5月号