

令和 3 年 6 月 3 日現在

機関番号：13401

研究種目：基盤研究(C) (一般)

研究期間：2018～2020

課題番号：18K11370

研究課題名(和文) 超高フレームレート動画像に対する時空間レート制御とその圧縮符号化への応用

研究課題名(英文) Spatio-temporal rate control techniques for very high-frame-rate videos and its application to video coding

研究代表者

吉田 俊之 (YOSHIDA, Toshiyuki)

福井大学・学術研究院工学系部門・教授

研究者番号：50240297

交付決定額(研究期間全体)：(直接経費) 2,100,000円

研究成果の概要(和文)：240フレーム/秒(240P)に及ぶ超高フレームレート動画像は、時間方向(動きの滑らかさ)の高画質化が期待される反面、従来の8倍に及ぶ情報量を限られた帯域に圧縮符号化し伝送/記録する際は、画質の最大化を図る必要がある。本研究では、代表者が過去に確立した「主観評価値の予測値の最大化に基づく時空間のレート制御手法」を2K240P動画像に拡張・適用し、その有効性を検証した。その結果、超高フレームレートは常に必要でなく、時空間のアクティビティに応じて適応的に制御することで大幅な画質改善が図れること、また提案法により当該制御が高精度に実現可能なことが確認された。

研究成果の学術的意義や社会的意義

動画像の高画質化は、これまで主に画素数の拡大方向に進んでおり、4Kや8K画像が身近となりつつある一方で、今後は時間方向の高フレームレート化も進むと予想される。高精細かつ240フレーム/秒に達する超高フレームレート動画像をネットワーク伝送あるいは保存する際は、高画質性を保った上で巨大なデータ量を削減する効率的なデータ圧縮・制御手法が不可欠である。本研究では、ビデオシーン単位の最適なフレームレート予測に基づく効率的な画質制御手法を検討・提案し、高画質動画像に対するデータ制御法への貢献を目的とする。

研究成果の概要(英文)：Very high frame-rate videos, running upto 240 frame/s, require a maximization of video quality for a transmission / storage within a prescribed band width. This study extends and applies the spatio-temporal rate control technique for videos based on a maximization of their predicted subjective quality metric, developed previously by the author, and validates the effectiveness of the approach. The results of this study have suggested that 240 frame/s is not always necessary in preserving the total video quality, which can drastically improved by applying the spatio-temporal adaptive rate control technique. They also have indicated that the proposed technique can precisely and efficiently realize the video quality maximization.

研究分野：画像/信号処理

キーワード：動画像符号化 高フレームレート 時空間画質 画質制御 レート制御 主観評価値 MOS

1. 研究開始当初の背景

本研究の背景を極めて平易に表現すると以下の通りとなる。動画像／ビデオのデータ量は非常に大きいので、その伝送や記録には圧縮符号化によってデータ量を目標値まで削減する処理が不可欠である。従来の圧縮符号化法は“動画像の毎秒当たりのコマ数”を一定（例えば毎秒 30 枚）に保って適用されるため、圧縮により“各コマの画質”は低下するが、コマ数一定のため“被写体の動きの滑らかさ”は不変に保たれる。一方、動画像は主として画素数の拡大（高精細化）と共に発展して来たが、近年ではコマ数の拡張（被写体の動きを滑らかにする効果）も顕著で、毎秒当たりのコマ数が 240 枚に及ぶ動画像も容易に利用可能となっている。こうした高コマ動画像は高速な被写体を滑らかに表示する目的には有効である一方、現状（毎秒 30 枚）の 8 倍に及ぶコマ数が常に必要とは限らない。

圧縮符号化の目標値が一定の場合、「各コマの画質」と「毎秒のコマ数」は補完（トレードオフ）関係にあり、一方を上げる（下げる）と他方は下がる（上がる）ため動画像の各シーン毎に両者がバランスする最適点が存在する。特に、240 枚に及ぶ高コマ動画像の場合は、“毎秒のコマ数の最適点”を予測し、この下で圧縮符号化することで従来手法に比べて画質の改善が図れる。

以下、本報告で用いる用語を整理する。動画像の各“コマ”を「フレーム」、毎秒当たりのコマ数を「フレームレート (FPS)」(t [frame/s] で表す) と呼ぶ。圧縮符号化の目標データ量を「ビットレート (BR)」(r [bit/s] で表す)、フレームレート t の動画フレーム n [枚] を r で符号化した場合の「単一フレーム当たりの平均ビット量 (BPF)」を $s = rn/t$ [bit/frame] と定義する。本研究は、240frame/s の「超高 FPS 動画像」を対象とした画質の最大化を目的とし、これを「時空間画質／レート制御」と呼ぶ。

以上の下で、 r 、 s 、 t の間には関係

$$r = s \cdot t \quad (1)$$

が成り立つため、時空間画質の最大化には、与えられた s と t で圧縮された符号化動画像に対する空間方向の画質（単一フレームの画質）および時間方向の画質（動きの滑らかさ）を総合的に評価する画質の評価尺度が必要になる。

研究代表者らは過去に、時空間画質の総合評価尺度として、主観評価値のひとつである **Mean Opinion Score (MOS)** に着目し、単一シーンの時間画質と空間画質を統一的な評価指標およびスケールの下で評価した **MOS** の予測値 **EMOS (Estimated MOS)** を用いた画質制御／レート制御手法を検討してきた（引用文献 [1-6]）。この手法では、図 1 のように対象動画像（単一シーン）を (s, t) の下で符号化した際の EMOS（5 段階スケール）を $EMOS(s, t)$ として予測し、 $EMOS(s, t)$ を最大化する (s, t) の下で符号化を行なうことに基礎をおく。ここで、引用文献 [1] では $EMOS(s, t)$ を

$$EMOS(s, t) \propto EMOS_s(s) \cdot EMOS_t(t) \quad (2)$$

のように、与えられた s に対するフレーム内空間画質を表す $EMOS_s(s)$ 、および与えられた t に対する時間画質（動きの滑らかさ）を表す $EMOS_t(t)$ に分離して推定可能なことを確認し、引用文献 [2-4] 等では $EMOS_s(s)$ と $EMOS_t(t)$ を対象動画像の特徴量から推定する手法を与えている。対象各シーンを単位として $EMOS(s, t)$ が推定できれば、図 1 に示すように、与えられた r によって定まる双曲線上で $EMOS(s, t)$ が最大となる $t = t_{opt}$ （最適 FPS）および対応する $s_{opt} = r/t_{opt}$ の組合せの下で符号化すれば、限られた BR 制約の下で画質を最大化する画質／レート制御および符号化が実現できる。

本研究では、画素数 1920x1080 (2K)、フレームレート 240frame/s (240P) である 2K240P 動画像に対して本手法を適用し、次章に示す目的の下に研究を進めることにより、最終的に 2K240P 動画像に対する時空間画質／レート制御手法の確立とその有効性の検証を目指した。

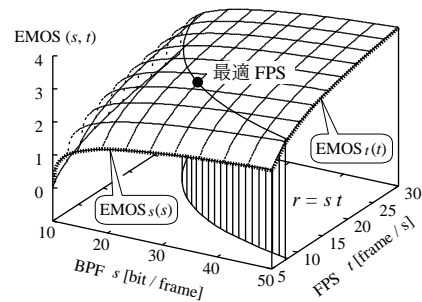


図 1 : $EMOS(s, t)$ の例と最適 FPS の決定.

2. 研究の目的

以上の背景に基づき、本研究では以下の3項目を研究目的とした。

- (1) 式(2)に基づき、2K240P 動画像に対する EMOS 予測手法を確立する
- (2) (1) の手法に基づく複数の時空間画質/レート制御法を H.265/HEVC 符号化器上を実現する
- (3) (2) の手法の有効性を 2K240P 試験動画像を用いて検証する

なお、対象動画像の FPS を低下させる際は、原画像の 240frame/s に対して選択可能な FPS を「2のべき乗分の1」、すなわち 30, 60, 120, 240 [frame/s] に制限し、フレームを間引く前にハーフシャッタをシミュレートする時間平均を施している。

3. 研究の方法

研究の目的(1)–(3)に対応し、以下の方法に従って研究を進めた。

(1) 2K240P 動画像に対する EMOS 予測法の確立

研究代表者らが引用文献 [2–4] で提案している手法を 2K240P 動画像に拡張・適用して $EMOS_s(s)$ および $EMOS_t(t)$ の各推定式を求め、式(2)と併せて 2K240P 動画像に対する EMOS 予測法を確立する。実際には、予め準備した 2K240P 試験動画像を相異なる s, t の下に各々符号化/フレーム間引きした動画像を用い、各 MOS を実測した結果をパラメトリック近似することで両推定式を導く。

(2) 時空間画質/レート制御法の実現と実装

(1) で求めた推定式により 2K240P 動画像に対してシーン単位で $EMOS(s, t)$ が予測可能となり、これを用いて種々の時空間画質/レート制御法が実現できる。本研究では、

- ① 最適 FPS に基づく CBR 符号化
- ② EMOS を一定とする CMOS 符号化
- ③ 画質平坦化の適用

の3つの手法を検討し、各々を H.265/HEVC 符号化器に実装する。

(3) 2K240P 動画像に対する時空間画質/レート制御法の有効性の検証

(2) で実現した各手法を試験動画像に適用し、EMOS, BR, 選択される FPS の時間変化を相互に比較することで、提案手法の有効性を明らかにする。

4. 研究成果

(1) 2K240P 動画像に対する EMOS 予測法の確立

2K240P 動画像に対する $EMOS_t(t)$ と $EMOS_s(s)$ の推定式を求めた。なお、本研究で用いる MOS は、過去の研究 [1–5 等] と同様に 1~5 の5段階評価とする。

$EMOS_t(t)$ については、引用文献 [2] の手法を 240P に拡張して適用することにより、

$$EMOS_t(t) = \frac{4}{e^Q + 1} + 1 \quad (3)$$

$$Q = A_t d + B_t \varepsilon + C_t (t_{max}/t) + D_t \quad (4)$$

を導いた。ここで、 d, ε は各々、フレーム間差分及び空間 Sobel オペレータ出力に基づく特徴量で、 $t_{max} = 240$ とする。 $A_t \sim D_t$ は MOS 実測値との重回帰分析によって決定される定数で、各値を表 1 に示す (詳細については雑誌論文 ② 参照)。

表 1 : EMOS 推定式における定数値.

A_t	1.966×10^{-2}	B_t	-6.487×10^{-3}	C_t	3.745×10^{-1}	D_t	-2.123
A_s	2.435×10^{-5}	B_s	2.004×10^{-3}				
C_0	6.878×10^1	C_1	-4.222	C_2	-1.471	C_3	-7.510×10^{-1}

一方, $EMOS_s(s)$ については, 符号化器のパラメータ依存性を吸収するため PSNR(Peak Signal-to-Noise Ratio) [dB] を介在させた引用文献 [3,4] の手法を 2K 動画像に拡張した予測式

$$EMOS(x) = A_s T(x - T)^3 + B_s T(x - T) + C_s \quad (C_s = 5.000) \quad (5)$$

$$T = C_0 + C_1 \ln(\epsilon + 1) + C_2 \ln(E + 1) + C_3 \ln(L + 1) \quad (6)$$

を用いた. ここで, x [dB] は対象動画像を s で符号化した際の平均 PSNR, E 及び L は各々動き推定残差および平均輝度値に基づく特徴量である. $A_s \sim C_3$ は実測 MOS から導かれる定数で, 各値を表 1 に示す (詳細については引用文献 [4] および雑誌論文 ②参照). なお, s から PSNR x の予測は, 対象動画像を相異なる s で符号化して PSNR x を実測し, (s, x) 各点を

$$x = A_p (s - C_p)^{B_p} \quad (7)$$

で近似することによって行う. $A_p \sim C_p$ は当該近似により定まる定数である.

(2) 時空間画質/レート制御の実現と実装

(1) で求めた $EMOS_t(t)$, $EMOS_s(s)$, および式 (2) により, 2K240P 動画像の $EMOS(s, t)$ をシーン単位に予測できる. 次に, 対象原動画像から $EMOS$ 推定に必要な各特徴量を抽出して $EMOS(s, t)$ を求める処理を, 研究開始時点で最新の動画像符号化規格であった H.265/HEVC 符号化器に実装し, これに基づいて 3 章 (2) の ①~ ③ の時空間画質/レート制御法を以下の通り実現した.

① 最適 FPS に基づく CBR 符号化

図 1 に示す様に, 与えられたビットレート $r = s \cdot t$ の下で $EMOS(s, t)$ を最大化する $t = t_{opt}$ (最適 FPS) を, 30, 60, 120, および 240 [frame/s] の候補から選択し, t_{opt} および対応する $s_{opt} = r/t_{opt}$ の下にシーン単位に CBR (Constant Bit Rate) 符号化する.

② EMOS を一定とする CMOS 符号化

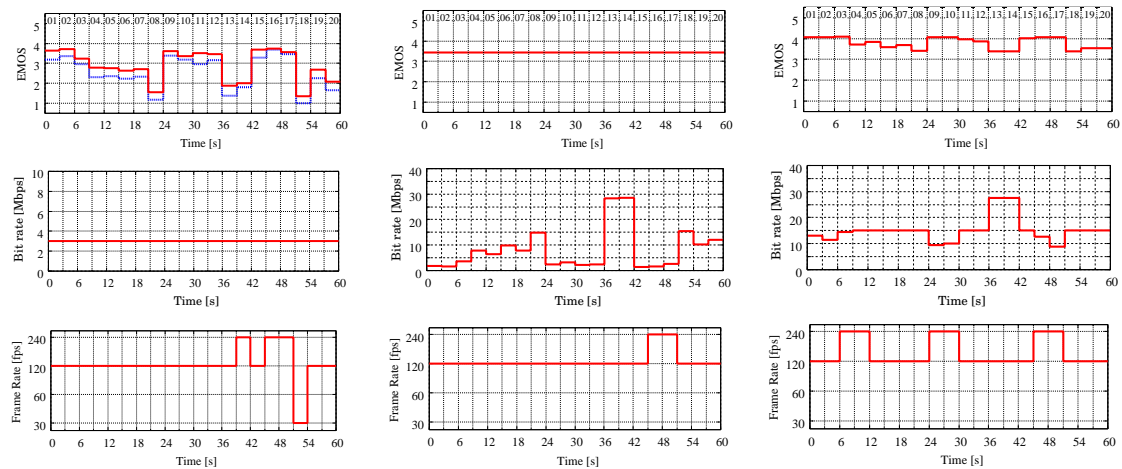
$EMOS(s, t)$ を $EMOS = E$ で切断すると, 当該 $EMOS$ を実現する (s, t) の組み合わせ候補が得られる. この中で式 (1) の r を最小にする (s, t) を選択し, シーン単位に VBR (Variable Bit Rate) 符号化することにより, 画質 (MOS) 一定 (Contant MOS) 符号化を実現する.

③ 画質平坦化の適用

② の CMOS 符号化では MOS 一定の制約が強くなり, 必要な BR がシーン間で大きく変動する. これを抑えるため, 引用文献 [5] で提案されている画質 (EMOS) の平坦化手法を実現した.

(3) 2K240P 動画像に対する時空間画質/レート制御法の有効性の検証

時空間の統計的性質が相異なる 20 種類の 2K240P 試験動画像 (各 720 フレーム) を接続した計 60 秒の動画像に対し, (2) の各制御法を適用した結果 ($EMOS$, ビットレート, 選択されるフレームレートの時間変動) を図 2 に示す.



(a) ①最適 FPS 下の CBR 符号化 (BR $r = 3$ [Mbit/s] 一定) (b) ②CMOS 符号化 (目標 $EMOS = 3.5$ 一定) (c) ③画質平坦化 (平均伝送 BR $r = 15$ [Mbit/s])

図 2 : ①~ ③ の各画質/レート制御/符号化手法の適用結果

図 2(a) の EMOS 変動における青い点線は、FPS を 240frame/s に固定し $r = 3$ [Mbit/s] で CBR 符号化した EMOS 変動 (60 秒間の時間平均 2.56) で、提案する最適 FPS で符号化 (赤線, 同時間平均 2.90) することにより、時空間画質 (EMOS) が改善されることが確認できる。一方で、シーン間で動画像の統計的性質 (時空間のアクティビティ) が変動するため対応する EMOS も大きく変動しており、シーン間で画質の変動が大きいことが判る。図 2(b) の CMOS 符号化では、EMOS が一定で完全に定画質符号化が実現されている反面、この場合は各シーン間で必要となる BR が大きく変動する。図 2(c) の画質平坦化は大規模バッファを利用して EMOS や BR 変動を抑圧する手法で、図 2(c) では $r = 15$ [Mbit/s] の定レート伝送を実現すると共に、(a) に比べて EMOS 変動が十分に抑圧されている。また、図 2(a)~(c) に共通して、シーン単位に適切な FPS が選択されており、以上より提案手法の妥当性と有効性が確認できる。

以上、本研究の成果は次のように総括される。本研究では、2K240P 動画像を対象として、

- (1) 2K240P 動画像に対する EMOS 予測手法の確立
- (2) EMOS 予測に基づく時空間画質/レート制御法の実現
- (3) (2) の各手法の有効性の検証

を行なった。その結果、

- 2K240P 動画像のシーンを単位として、時空間 MOS の予測値 $EMOS(s, t)$ を高精度に予測をする手法が確立された。
- $EMOS(s, t)$ を利用し、シーン単位で EMOS を最大化、一定化、あるいは平坦化する時空間画質/レート制御法が実現された。
- 当該画質/レート制御法を適用した結果、各手法によって、
 - － 対象シーンの統計的性質に応じて適切なフレームレートが選択されること
 - － フレームレート一定とする従来の符号化法に比べて画質 (EMOS) が改善されること
 - － 240P 等のフレームレート選択の自由度が高い場合は、提案法は非常に有効に機能すること
 - － 240frame/s は常に必要ではなく、例えば被写体が高速移動する場合等に適応的に選択することで、動画シーケンス全体の総合画質が改善されること

を確認した。

以上から、本研究の当初目標は達成されたと判断される。なお、提案手法を適用する際は、式 (7) に基づいて各対象シーンの PSNR を s で近似する必要があり、この処理は膨大な計算量を要するため提案手法の実時間適用は現状では困難となっている。逆に、式 (7) に対する高速な近似法が実現されれば、提案各手法は実時間適用が可能となる。本研究に関する残された課題として、式 (7) の処理の高速化、および提案各手法の実時間実現が挙げられる。

<引用文献>

- [1] 稲積泰宏 他, “ビットレート制限下における動画像通信のための最適フレームレート推定”, 電子情報通信学会論文誌 (B), vol.85-B, no.7, pp.1130-1142 (2002 年 7 月)
- [2] 吉田俊之 他, “4K60P 動画像に対する時間方向の主観画質推定”, 電子情報通信学会論文誌 (D), vol.J102 - D, no.10, pp.662-665 (2019 年 10 月)
- [3] 宮田高道 他, “SNR に基づく動画像の主観画質推定”, 電子情報通信学会論文誌 (A), vol.J88 - A, no.11, pp.1292-1296 (2005 年 11 月)
- [4] 吉田俊之, “4K 圧縮動画像に対する主観画質推定”, 電子情報通信学会論文誌 (D), vol.J100 - D, no.9, pp.835-838 (2017 年 9 月)
- [5] 吉田俊之, “予測 MOS の平坦化に基づく符号化画質改善法”, 電子情報通信学会論文誌 (D), vol.J93 - D, no.9, pp.1634-1637 (2010 年 9 月)

5. 主な発表論文等

〔雑誌論文〕 計3件（うち査読付論文 3件／うち国際共著 0件／うちオープンアクセス 0件）

1. 著者名 山口翔也, 吉田俊之	4. 巻 J102-D
2. 論文標題 4K60P動画像に対する時間方向の主観画質推定	5. 発行年 2019年
3. 雑誌名 電子情報通信学会和文論文誌D	6. 最初と最後の頁 662-665
掲載論文のDOI (デジタルオブジェクト識別子) 10.14923/transinfj.2019IEL0001	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 犬塚亮, 吉田俊之	4. 巻 J103-D
2. 論文標題 240P 動画像に対する時間方向の主観画質推定	5. 発行年 2020年
3. 雑誌名 電子情報通信学会和文論文誌D	6. 最初と最後の頁 747-750
掲載論文のDOI (デジタルオブジェクト識別子) 10.14923/transinfj.2020IEL0001	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 犬塚亮, 吉田俊之	4. 巻 未定
2. 論文標題 2K240P動画像におけるMOS推定に基づく時空間画質制御の有効性について	5. 発行年 2021年
3. 雑誌名 電子情報通信学会和文論文誌D	6. 最初と最後の頁 未定
掲載論文のDOI (デジタルオブジェクト識別子) なし	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

〔学会発表〕 計5件（うち招待講演 0件／うち国際学会 0件）

1. 発表者名 犬塚亮, 吉田俊之
2. 発表標題 2K240P動画像に対する時間方向のMOS予測手法
3. 学会等名 2019年度電気関係学会北陸支部連合大会
4. 発表年 2019年

1. 発表者名 犬塚亮, 吉田俊之
2. 発表標題 240P動画像に対する時間方向のMOS予測手法
3. 学会等名 2019年度画像符号化シンポジウム/映像メディア処理シンポジウム
4. 発表年 2019年

1. 発表者名 山口翔也, 吉田俊之
2. 発表標題 4K60P動画像を対象とした時間方向のMOS予測手法の精度向上について
3. 学会等名 画像符号化シンポジウム・映像メディア処理シンポジウム
4. 発表年 2018年

1. 発表者名 犬塚亮, 吉田俊之
2. 発表標題 EMOS最大化に基づく240P動画像の時空間レート制御について
3. 学会等名 2020年度画像符号化シンポジウム/映像メディア処理シンポジウム
4. 発表年 2020年

1. 発表者名 Ryo Inuzuka, Toshiyuki Yoshida
2. 発表標題 EMOS-based spatio-temporal rate control technique for 2K240P videos
3. 学会等名 International Workshop on Advanced Image Technology 2021
4. 発表年 2021年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

-

6. 研究組織

	氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
--	---------------------------	-----------------------	----

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8. 本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関
---------	---------