

低ビットレート動画通信における 最適フレームレートの画像依存性

稲積 泰宏* 磯谷 泰知* 吉田 俊之* 酒井 善則* 堀田 裕弘**
*東京工業大学大学院理工学研究科集積システム専攻 **富山大学工学部

152-8552 東京都目黒区大岡山 2-12-1

Tel:03-5734-3593

E-mail:{inazumi,tyoshida}@ss.titech.ac.jp

あらまし インターネットに代表される帯域非保証下での低ビットレート動画通信においてサービス品質を向上させるには、利用可能なビットレートや画像内容に応じて視覚的な品質が最も高くなるようにフレームレートやフレーム当たりのビット量をバランス良く制御することが必要である。そのため、指定されたビットレートによって最適フレームレートで画像を符号化すればよい。本稿では、主観評価実験により得られるフレームレート、フレーム当たりのビット量、視覚的な品質の関係から最適フレームレートを求め、その画像内容依存性について考察する。

Picture Dependency of the Optimal Frame Rate in a Low Bitrate Video Communication

Yasuhiro INAZUMI* Taichi ISOGAI* Toshiyuki YOSHIDA*
Yoshinori SAKAI* Yuukou HORITA**

*Dept. Communications and Integrated Systems,
Tokyo Institute of Technology

**Faculty of Engineering, Toyama University

O-okayama 2-12-1, Meguro-ku, Tokyo 152-8552, Japan

Tel:03-5734-3593

E-mail:{inazumi,tyoshida}@ss.titech.ac.jp

Abstract To improve QoS(Quality of Service) for a low bitrate video communication in a non-guaranteed network, it is necessary to control the frame rate and the amount of data per frame in such a way that perceptual video quality can be maximized according to a given bitrate and video contents. A target video they should be encoded with the optimal frame rate calculated from the given bitrate. This paper examines picture dependency of the optimal frame rate, which has been obtained by a subjective assessment test for several target videos considering the relation of the perceptual quality against the frame rate and the amount of data per frame.

1 はじめに

近年、計算機の処理能力・通信速度の向上を背景として動画のリアルタイム通信が注目されている。インターネットにおいては、通信速度が向上する一方でユーザ数の増加やユーザが利用する情報量の増大が顕著であり、現状では各ユーザの要求に従って帯域を確保することは難しく、将来的にユーザが利用可能な帯域が大きくなったとしても、要求帯域を常時確保することは困難であると考えられる。帯域非保証下での動画通信において、ユーザから見た QoS(Quality of Service) を向上するためには、伝送遅延を抑えるだけではなく、視覚的に高い品質を保つことが重要となる。一方で、従来から PSNR や MSE 等で表される誤差を抑えるようにレート制御を行う手法が用いられているが、このような誤差は品質を直接に表しているとは言い難いため、視覚的な品質に基づくレート制御技術が必要となる。これまで、文献 [1] 等において視覚的な品質に基づくレート制御が提案されているが、帯域非保証下ではフレームレート削減による視覚的な品質の影響を考慮する必要がある、そのような報告はほとんどされていない。

一般に、指定されたビットレートに対して、フレームレート（時間方向の情報量を規定）とフレーム当たりの平均ビット量（空間方向の情報量を規定）はトレードオフの関係となる。ビットレート一定の下で、視覚的な品質を最大にするようにフレームレートもしくはフレーム当たりの平均ビット量を求めることは互いに等価であるから、本稿ではフレームレートに着目する。指定されたビットレートに対する視覚的な品質が最も高くなるフレームレートを最適フレームレートと定義し、これが求めれば、フレーム当たりの平均ビット量は一意に決まる。そこで、フレームレートとフレーム当たりの平均ビット量を変化させた場合の品質の変化を品質の時空間特性と定義し、これを明らかにすることが必要である。

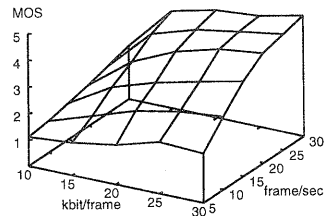
このような背景から、筆者らは視覚的な品質に基づくレート制御法を次の手順で実現することを目標としている。

- (1) 画像から動きやフレーム品質等の特徴量を抽出し、品質の時空間特性を求める。
- (2) (1) の結果を用いて、指定されたビットレートにおける最適フレームレートを求める。
- (3) 最適フレームレートによってレート制御を実現する手法を検討する。

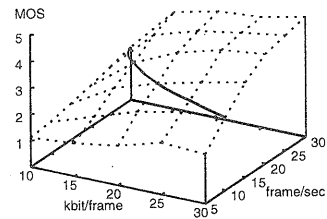
(1) の品質の時空間特性は、動きの大きさや画像の複雑さといった画像内容に依存するため、最適フレームレートも画像に大きく依存すると考えられる。そこで本稿では、最適フレームレートの画像依存性について検討する。

以下、2 章では最適フレームレートの導出方法を述べる。3 章では画像品質の時空間特性を得るための主観評価実験の方法について述べる。続く 4 章において 3 章の実験結果から最適フレームレートの画像内容依存性を考察すると共に、品質の時空間特性を導出する方法を検討する。5 章では、まとめと今後の課題を述べる。

(a) Spatio-temporal quality characteristic



(b) Quality on the available bit rate



(c) Peak quality on the available bit rate

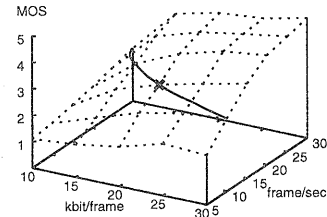


図 1: 最適フレームレートの導出法

2 最適フレームレートの導出法

2.1 品質の時空間特性の導出

最適フレームレートを導出する際に必要となる品質の時空間特性について説明する。品質の時空間特性は将来的には 1 の (1) のように画像の特徴量から求めるが、本稿では主観評価実験によって得られる MOS(Mean Opinion Score) を動画の品質とする。すなわち (1) を、

- (1)' 主観評価実験により品質の時空間特性を求める。

と置き換える。したがって、フレーム当たりの平均ビット量とフレームレートを変化させたときの MOS を図 1(a) の各点において求める。これらの MOS を線形補間することによって測定点以外の品質を求める。ここで、MOS の測定点をビットレート一定となる曲線上で求めないのは、品質の時空間特性を得るために必要な主観評価実験の回数を削減するためである。

以上のように得られる品質の時空間特性において、フレーム当たりのビット量が大きくなると静的な劣化が減少して品質が向上する。逆に、フレームレートが大きくなると動的な劣化の減少のため、品質が向上する [5]。

表 1: 5 段階品質尺度

評点	評価尺度
5	非常によい
4	よい
3	普通
2	悪い
1	非常に悪い

表 2: 主観評価実験の条件

視距離	40cm (7.3H 相当)
照明	蛍光灯
最大輝度	99cd/m ²
評定者数	11 名 (各自 2 回評価)
評価尺度	5 段階品質尺度
評価方法	単一刺激法

表 3: テスト画像のフレーム当たりの平均ビット量

画像	平均ビット量 [kbit/frame]
MC	25, 35, 45, 55, 65
TT	15, 20, 25, 30, 35
src13	20, 30, 40, 50, 60
src18	10, 15, 20, 25, 30
src19	20, 25, 30, 35, 40
src20	10, 15, 20, 25, 30
src21	10, 15, 20, 25, 30
src22	15, 20, 25, 30, 35

表 4: テスト画像の特徴

カメラワーク	画像	動きベクトルの平均値
静止	src20	0.0217
静止	TT	2.49
静止	src13	17.5
静止	src21	17.4
パンニング	MC	2.21
パンニング	src19	59.6
ズームアウト	src18	1.47
ズームアウト	src22	4.65

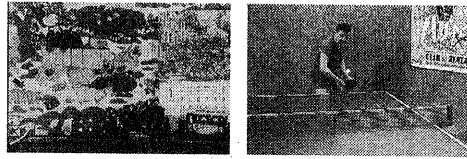
2.2 最適フレームレートの導出

1 の (2) に関連して、指定されたビットレートに対する最適フレームレートの導出について述べる。指定されたビットレートが c [kbps] となるときのフレーム当たりのビット量を s [kbit/frame]、フレームレートを t [frame/sec] とすると、フレームレートは、

$$t = \frac{c}{s} \quad (1)$$

となり、 c [kbps] における品質は図 1(b) の双曲線上で与えられる。したがって、最適フレームレートは以下の手順で求める。

- 2.1 で得られた品質の時空間特性上で指定されたビットレートを満たす双曲線を求める (図 1(b)).
2. の曲線上で品質が最大となる点におけるフレームレートが最適フレームレートとなる (図 1(c)).



MC

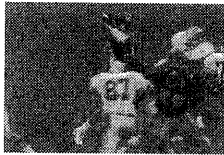
TT



src13



src18



src19



src20



src21



src22

図 2: テスト画像

3 主観評価実験

(1) の主観評価実験を行うために、本章では実験条件とテスト画像の作成について述べる。

3.1 実験条件

本実験はインターネット上での低ビットレート動画通信におけるレート制御への応用を考慮し、実際にユーザがパソコンで動画を見ることを想定する。そこで、標準観視条件 [3] を用いず、照明、視距離を通常のオフィス環境と同等に設定する。また、評価尺度は表 1 に示す 5 段階品質尺度を用いる。レート制御を行う際には、カメラワークや画像の動きに対して最適フレームレートが大きく依存すると考えられるため、各画像が 1 つのカメラワークで構成されるようにそれぞれ 1[sec] を取り出して実験を行う。提示回数は 1[sec] で構成される画像の判断は従来の DSCQS 法や EBU 法のような 2 回の評価では評価が困難であったため、経験的に十分に評価が可能となった 4 回とする。主観評価実験の条件を表 2 に示す。

符号化の条件として、フレームレートは、5,10,15,20,25,30[frame/sec] とした。また、画像によって必要とされるフレーム当たりの平均ビット量は大きく異なるため、予めフレームレート 15[frame/sec] にお

る MOS が 3 程度となるように予備実験を行うことにより、フレーム当たりの平均ビット量の範囲を画像ごとに定めた。これを表 3 に示す。

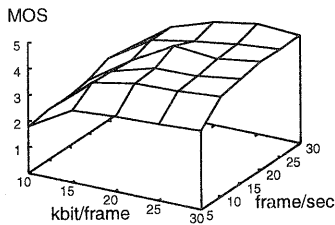


図 3: 品質の時空間特性 (src20)

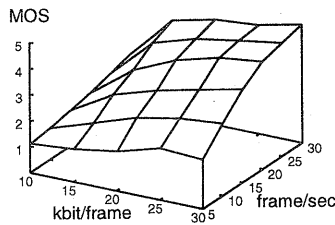


図 4: 品質の時空間特性 (src21)

3.2 テスト画像の作成

テスト画像の作成は以下の手順で行う。

1. フレーム番号の再割当によるフレームレートの削減
2. フレーム当たりの平均ビット量 s [kbit/frame], フレームレートを t [frame/sec] よりビットレート st [kbps] を求め、このレートで MPEG 符号化
3. 復号
4. フレームの再表示により 30[frame/sec] に復元

ここでフレーム番号の再割当は、原画像におけるフレームレートを F_0 [frame/sec] (ここでは $F_0 = 30$)、フレーム番号を N_0 、また、削減後のフレームレートを F_1 [frame/sec]、フレーム番号を N_1 とすると、

$$N_1 = \lfloor \frac{N_0 F_1}{F_0} \rfloor \quad (2)$$

によって行う。

このとき、フレームレートの削減によってフレーム間の相関が低くなり、P,B ピクチャのマクロブロックがブロックマッチングに失敗してイントラ符号化される確率が高くなり、実験結果がエンコーダの性能に大きく依存するという問題が発生する。特に、イントラ符号化される確率が高い場合、フレームレートとフレーム品質がトレードオフの関係にならない場合が発生する。そこで、動きベクトルの探索範囲を 63[pixel] と大きくとることで B,P ピクチャにおいてイントラ符号化される確率を 10% 以下に保っている。

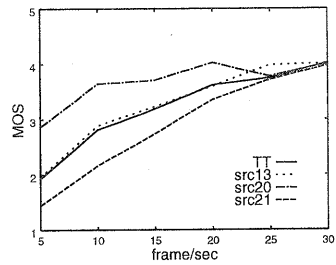


図 5: フレームレートと MOS の関係 (静止)

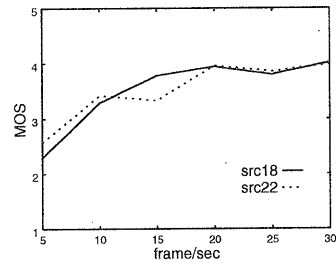


図 6: フレームレートと MOS の関係 (ズームアウト)

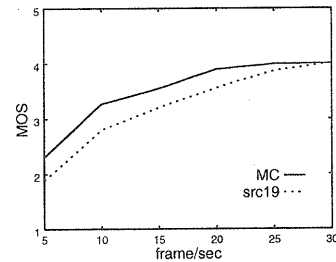


図 7: フレームレートと MOS の関係 (パンニング)

テスト画像は、MC(Mobile&Calendar), TT(Table Tennis), VQEG[2] の src13, src18, src19, src20, src21, src22 の 8 種類のそれぞれ 1[sec] を用いる。なお、画像サイズ 352 × 240 である。これらの画像を図 2 に、また画像の特徴として、カメラワークと動きベクトルの平均値を表 4 に示す。

4 実験結果

4.1 品質の時空間特性

主観評価実験によって得られた品質の時空間特性の例として、動きの小さい画像である src20 と、動きが激しい画像である src21 の結果をそれぞれ図 3, 図 4 に示す。両者の曲面の傾きは時空間方向にそれぞれ大きく異なる。したがって、同じカメラワークでも画像内容の複雑さや動きの大きさによって品質の時空間特性は大きく影響を

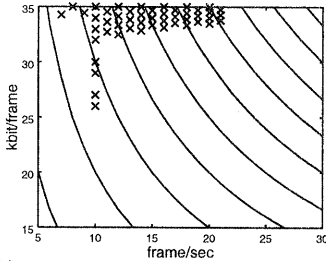


図 8: 最適フレームレートとフレームのビット量の関係 (TT)

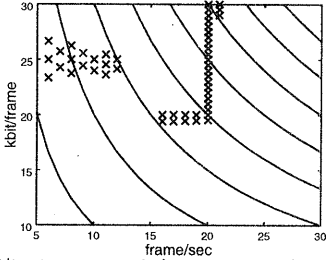


図 9: 最適フレームレートとフレームのビット量の関係 (src21)

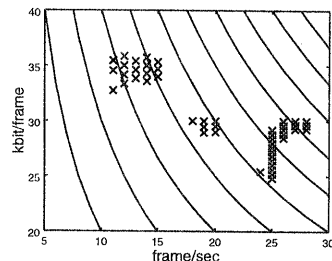


図 10: 最適フレームレートとフレームのビット量の関係 (src19)

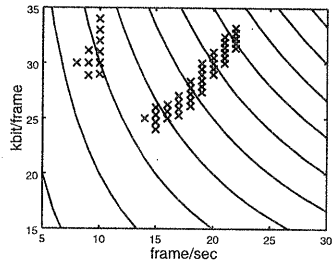


図 11: 最適フレームレートとフレームのビット量の関係 (src22)

受けることが確認される。

次に、動きと品質の関係を各画像について比較するため、フレーム当たりの平均ビット量一定におけるフレームレートと MOS の関係を、カメラワークが静止 (TT,src13,src20,src21)、ズームアウト (src18,src22)、パンニング (MC,src19) の各場合について、それぞれ図 5, 6, 7 に示す。なお、ここでは MOS の画像依存性を考慮し、30[frame/sec]において MOS が 4 となるフレーム当たりの平均ビット量一定の直線上で特性を求める。

図 5 より src20 は動きが非常に小さいためフレームレートが 20[frame/sec] 以上としても MOS は改善されない。また、src21 と src13 の動きベクトルの平均値はほぼ同じであるが特性は異なり、その値が比較的小さい TT と src13 がほぼ同じ特性となった。これは、src13 では画面全体で複数の物体が動いているのに対し、src21 や TT は画面中央の 1 つの物体が動くことにより中央の動きに敏感となるため、フレームレートが低下するに従い MOS も大きく低下する。したがって、動きのばらつきの違いが品質に影響していると考えられる。ズームアウトの場合、図 6 と被験者の感想からフレームレートが 15[frame/sec] 程度あれば、フレームレートの違いがわかりにくいと言える。また、src18 と src22 は同様の結果となったことから動きベクトルの大きさの影響を受けにくいと考えられる。一方で、パンニングの場合はフレームレートが高いほど品質が向上し、ある値以上に品質が向上しなくなるフレームレートは動きの大きさに依存すると言える。

以上において、動きの平均値、動物体の動きのばらつき、およびカメラワークから特性が予測可能であると考えられる。また、これらは動きベクトルを用いた特徴量

によって表現可能である。したがって、フレームレートと MOS の関係は動きベクトルから得られる特徴量で表現可能であると考えられる。

4.2 最適フレームレートの画像依存性

最適フレームレートとフレーム当たりの平均ビット量の関係を図 8~11 に示す。ここではビットレートが一定となる曲線を 100[kbps] 間隔で描いた。また各 × 点は、最適フレームレートにおけるフレーム当たりのビット量を 10[kbps] 間隔で求めたものである。これらの図における各点の位置の変化から、品質に対するフレーム品質と動きの重要性が得られる。

これを図 8 を用いて説明する。まず、低ビットレート時においてビットレートが次第に大きくなると、フレームレートは 10[frame/sec] と一定であるのに対し、フレーム当たりの平均ビット量は増加する。その後、フレーム当たりの平均ビット量は 32[kbit/frame] で一定となり、この先はフレームレートが次第に増加する。すなわち、低ビットレート時はフレーム品質が重要であり、その後、フレーム当たりの平均ビット量が重要となると言える。これを各画像について同様に解析した結果、品質に対するフレーム品質と動きの重要性の変化は以下の 4 種類に分類可能なことが観測された。

- カメラが静止して動きが小さい TT と src20 では、図 8 のように低ビットレート時はフレーム品質が重要で、ビットレートが高くなることによりフレーム品質が得られると、フレームレートが重要となる。

- カメラが静止して動きが大きい src13 と src21 では、図 9 のように、低ビットレート時はフレーム品質がある程度必要で、ビットレートが高くなるに従って、フレームレートが重要となる。さらにビットレートが高くなるとフレームレートは一定(図 9 では 20[frame/sec])となり、フレーム品質が重要となる。
- パンニングの MC と src19 は、図 10 のように低ビットレート時はフレーム品質が重要であるが、ビットレートが高くなるに従って、フレームレートが重要となる。
- ズームアウトの src18 と src22 は、最初はフレーム品質が重要であるが、ビットレートが高くなるに従って、フレーム品質とフレームレートの両方が重要となる。

以上より、ビットレートに対する最適フレームレートの変化は、カメラワークや動きの大きさを表すような動きベクトルの特徴量とフレーム品質に依存すると言える。

4.3 最適フレームレートの推定

これまでの実験結果と考察から、フレーム品質が十分に高い場合のフレームレートと品質の関係(動的な劣化特性と定義する)は動きベクトルのみで得られると考えられる。また、フレームレートが十分に高い場合のフレーム当たりの平均ビット量と品質の関係(静的な劣化特性と定義する)はフレーム品質を視覚特性に基づいて評価する文献 [4] 等の手法で求められる。そこで、動的な劣化と静的な劣化を表すそれぞれの特性を画像から求めることにより、品質の時空間特性を予測する方法を検討する。その一例として、src18 における品質の時空間特性を 30[kbit/frame] におけるフレームレートと MOS の関係(動的な劣化特性)、および 30[frame/sec] におけるフレーム当たりの平均ビット量と MOS の関係(静的な劣化特性)の両者を関数近似し、品質の時空間特性を予測する。ここで、動的な劣化と静的な劣化が互いに独立であると仮定し、両者の積で表す。すなわち、フレームレートを f 、フレーム当たりの平均ビット量を b として、

$$MOS = 4.26 \left(\frac{4.8(f - 3.7)^{0.1} - 2.4}{4.26} \right) \times \left(\frac{3.1(b - 9.5)^{0.13} - 0.335}{4.26} \right) \quad (3)$$

を用いて MOS を近似する。この関数によって予測した src18 における品質の時空間特性を図 12 に示す。また、比較のため、src18 の主観評価実験による品質の時空間特性を図 13 に示す。これらの図より、動的な劣化特性と静的な劣化特性より品質の時空間特性は式 (3) の関数をさらに工夫することで推定可能であると考えられる。

5 まとめ

本稿では、帯域保証のない動画通信において、指定されたビットレートにおける視覚的な品質が最も高くな

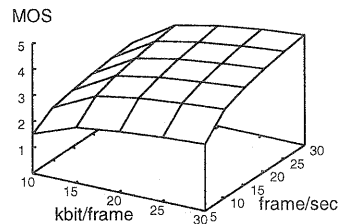


図 12: 関数による品質の時空間特性 (src18)

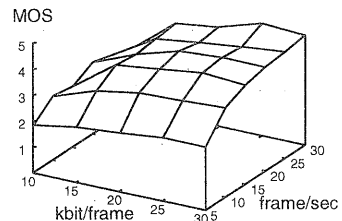


図 13: 主観評価実験による品質の時空間特性 (src18)

る最適フレームレートについて、その画像依存性を検討した。まず、主観評価実験によって得られる品質の時空間特性の画像依存性を示した。次に、フレームレートと品質の関係は、画像の動きベクトルの平均値、動物体の動きのばらつき、カメラワークの影響が大きいことを示した。また、ビットレートに対する最適フレームレートの変化は、動きベクトルの特徴量とフレーム品質に依存することを確認した。さらに、動的な劣化特性と静的な劣化特性より、品質の時空間特性の推定可能性を簡単に示した。

今後、動きベクトルとフレームの品質から品質の時空間特性を求め、最適フレームレートを推定する具体的な手法を検討する予定である。

謝辞 分光輝度計による CRT の輝度測定に御協力頂いた東京工業大学像情報工学研究施設大山・山口研究室の村上百合助手に感謝致します。

参考文献

- [1] Fu-Huei Lin, Russell M. Mersereau, "Rate-quality tradeoff MPEG video encoder" Signal Processing 14(1999) 297-309.
- [2] <http://www.crc.ca/vqeg/>
- [3] ITU-R Recommendation BT.500-7, ITU, Geneva, Switzerland, 1995.
- [4] 堀田, 稲積, 村井, "フレームごとの品質を用いた白黒動画の画質評価法" 信学論 Vol. J81-B-I No.11 pp.652-660, Nov.1998.
- [5] 中須, "画質評価 -画質とユーザ要求-" テレビ誌 Vol.49 No.4 pp.497-501, Oct.1993.