

高フレームレート映像信号に対する符号化特性解析

坂東幸浩[†], 高村誠之[†], 八島由幸[†]

[†] 日本電信電話株式会社 NTT サイバースペース研究所

〒 239-0847 神奈川県横須賀市光の丘 1-1

E-mail : {bandou.yukihiro, takamura.seishi,
yashima.yoshiyuki}@lab.ntt.co.jp

あらまし 近年、臨場感あふれる大画面のスポーツ映像やデジタルシネマに代表される超高画質映像への期待が高まっている。しかし、被写体の自然な動きを表現するために不可欠な時間解像度の向上、即ち、映像の高フレームレート化については、十分な検討がなされていない。現行映像のフレームレートはフリッカーの検知限から設定されており、自然な動き表現には、さらに高いフレームレートが必要となる。しかし、高フレームレート化により高画質化を追求しようとした場合、データ量の増大に対応するために効率的な符号化方式が必要となる。そこで、本報告では、高フレームレート映像の符号化を睨み、フレームレートと情報量の関係について定量的な評価を行い、差分信号のエントロピがフレームレートの対数に比例するという新たな知見を得た。

Encoding characteristic analysis of high frame-rate video signal

Yukihiro BANDO[†], Seishi TAKAMURA[†], and Yoshiyuki YASHIMA[†]

[†]NTT Cyber Space Laboratories, NTT Corporation

1-1 Hikarino-oka, Yokosuka, Kanagawa 239-0847, JAPAN

E-mail : {bandou.yukihiro, takamura.seishi,
yashima.yoshiyuki}@lab.ntt.co.jp

Abstract:

Realistic representations using extremely high quality images are becoming increasingly popular. For example, digital cinemas can now display moving pictures composed of high-resolution digital images. Although these applications focus on increasing the spatial resolution of the images, they ignore the temporal resolution. In order to faithfully represent the movement of an object in a video sequence, we need high-frame-rate images whose statistical properties have to be elaborated. In this paper, we report the statistical properties of such images and describe relationship between frame-rate and bit-rate.

1 はじめに

近年、臨場感あふれる大画面のスポーツ映像やデジタルシネマ [1] に代表される超高画質映像への期待が高まっている。これを受けて、映像の高精細化に関する研究が精力的に行われている [2]。超高画質映像の実現には次の四要素が必要である。すなわち、空間解像度、画素値深度、色再現性、時間解像度である。これを受けて、前者三要素についてはデジタルシネマ等の応用およびナチュラルビジョンプロジェクト [3] において検討が進められている。

しかし、被写体の自然な動きを表現するために不可欠な時間解像度の向上、即ち、映像の高フレームレート化については、十分な検討がなされていない。Spillmann らによるによれば、網膜の出力細胞である神経節細胞が出力するパルス数の上限は毎秒 300 ~ 400 個程度であるとする生理学的な知見が示されている [4]。このため、人の視覚系は 1/150 ~ 1/200 秒程度までの発光の違いを知覚できると推察される。これは、知覚可能なフレームレートの検知限が、150 ~ 200 [フレーム/秒]であることを意味する。現行映像のフレームレートである 30,60 [フレーム/秒] は、フリッカーの検知限から定められたものであり、自然な動きを表現するに十分な値ではない。

映像の超高画質化はデータ量の爆発的な増加を招くため、効率的な符号化法が不可欠となる。高フレームレート映像の符号化を検討する際、同映像信号の統計的性質を的確に把握することが必要となる。とりわけ、フレームレートと情報量の関係は重要である。従来のフレームレートと情報量についての検討は、60 [フレーム/秒] 以下の低フレームレート映像に関するものが中心であり [5, 6, 7]、符号化の面からみた高フレームレート映像信号の統計的性質については未知な部分が多い。例えば、高フレームレート映像は HDTV 等の現行映像と比較して高い時間相関を有することは予想されるが、その定量的な性質は把握されていない。

本報告では、高フレームレート映像信号の符号化を睨み、フレームレートと情報量の関係について定量的な評価を行う。フレームレートの増加に伴い、映像シーケンス全体でみた場合の情報量は増大する。一方、フレーム間の相関は大きくなるため、フレームあたりの情報量で見れば、減少することが予想される。フレームレートと情報量との関係を把握することは、将来の高フレームレート映像システムの構築に、

有益な指針を与える。

2 高フレームレート化に伴う問題点

高フレームレート化により高画質化を追求しようとした場合、次の 2 つの問題が生じる。

一つ目は、データ量の増大である。単位時間あたりのデータ量は、フレームレートに比例して増加する。例えば、今回使用した高速カメラ（後述）を用いて、現行のフレームレート (30 [フレーム/秒]) の映像を撮影すれば、データ量は 120 [メガバイト/秒] である。これに対し、フレームレート 1000 [フレーム/秒] の映像を撮影すれば、データ量は約 4 [ギガバイト/秒] となる。このため、映像データの通信・蓄積を行う場合、データ量を圧縮できる効果的な符号化方式は必要不可欠と言える。本課題に対する検討の詳細は、3 章にて述べる。

二つ目は処理時間の増大である。処理すべきデータ量の増大は、処理速度の高速化も要求する。MPEG [8],[9] などの動画像符号化において、最も計算量を必要とする処理はフレーム間予測における動き推定である。シーケンス全体における動き推定の処理量は、フレームレートに比例して増大する。このため、フレームレートの増加は、符号化処理の増加に大きな影響を与える。この点に鑑みて、処理速度の高速化実現にあたり、まず取り組みべき課題は動き推定の高速化であるといえる。本課題に対する検討は、稿を改めて述べたいと思う。

3 フレームレートの符号化に及ぼす影響

3.1 高フレームレート映像ソース

高フレームレートの映像ソースは、表 1 に示す NAC 社製ハイスピードカメラ (MEMRECOM fx RX-3) を用いて、デジタルデータとして作成した。撮影時のフレームレートは 1000 [フレーム/秒] した。また、シャッター時間はフレーム間隔に等しいものとした。さらに、デジタルデータとして出力する際、画素値にはガンマ補正・ニー補正といった非線形変換は施していない。

表 1: ハイスピードカメラ (NAC 社製 MEMRECOM fx RX-3) のスペック

センサー素子	カラー CMOS
素子画素数	1280 × 1024 [画素]
画素値深度	10 [ビット/画素]
フレームレート	最大 1000 [フレーム/秒]
シャッター速度	1/1000 ~ 1/200000 [秒]
感度	ASA250 相当
レンズマウント	F マウント、C マウント

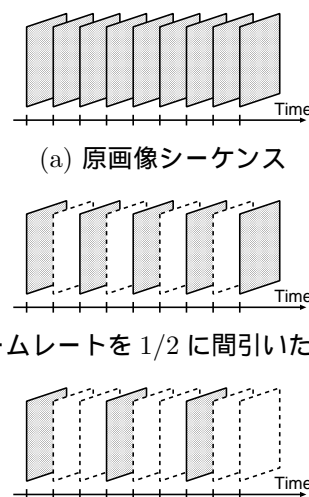
フレームレートと情報量の対応関係を求めるためには、画角が等しくフレームレートのみ異なる映像が必要となる。ここでは、図 1 に示すように、フレーム間引きにより、フレームレートの異なる映像を得る。同図 (b) は基準となる同図 (a) のシーケンスから、1 枚おきにフレームを間引き、フレームレートを半分にしたシーケンスを示しており、同図 (c) はフレームレートを 1/3 にしたシーケンスを示している。同図では、間引いたフレームを破線の矩形で示している。

フレーム間引きの基準となる映像は、表 1 の高速度カメラで撮影した RGB カラー映像であり、フレームレートは 1000[フレーム/秒]、総フレーム数は 480[フレーム]、解像度は 640 × 480[画素] である。また、映像素材はテニスと野球のスイングを撮影したスポーツシーンである。各シーケンスは、時間方向の相関の強さが異なるものを選んでいく。

3.2 H.264 を用いた符号実験

フレームレートの変化が符号化効率に及ぼす影響を調べるために、各シーケンスにおいてフレームレートを 1000, 125, 31[フレーム/秒] として、H.264[10] コーデック JM73 による符号化実験を行った。実験条件は表 2 の通りである。結果を図 2 に示す。ここで、横軸はフレーム当りの符号量とし、縦軸は SNR とした。なお、縦軸の SNR は原画像を $f_t(x, y)$ 、復号画像を $g_t(x, y)$ ($0 \leq x \leq X - 1, 0 \leq y \leq Y - 1$) として、次式により求めた。

$$\text{SNR} = 10 \log_{10} \left(\frac{255^2}{\frac{1}{XY} \sum_{y,x} \{f_t(x, y) - g_t(x, y)\}^2} \right)$$



(c) フレームレートを 1/3 に間引いたシーケンス

図 1: フレームレートの変換

表 2: H.264 の符号化条件

参照フレーム数	1 [フレーム]
動き探索範囲	±16 [画素]
シーケンスタイプ	先頭のみ I、それ以外は P
RD 最適化	ON
エントロピ符号化	UAVLC
Hadamard 変換	使用

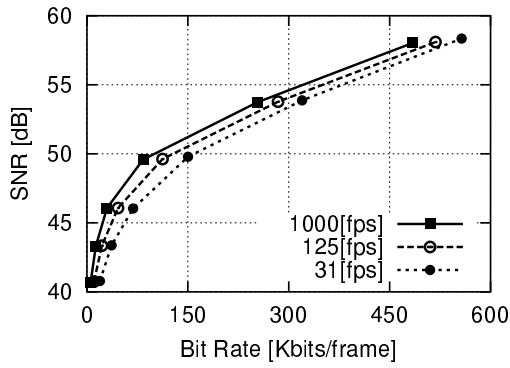
いずれのシーケンスにおいても、フレームレートの増加と共に、グラフが左上方に移動しており、少ない符号量で高い画質を実現できたことが確認できる。

3.3 フレームレートと情報量の関係

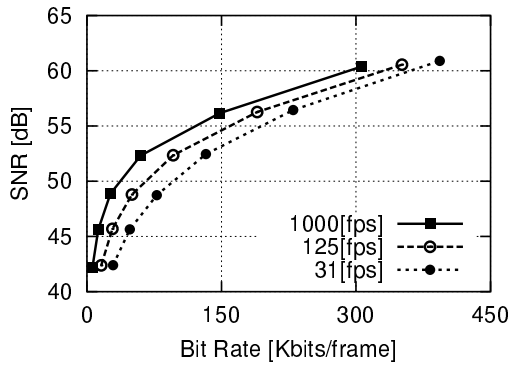
図 3 にフレームレートを変化させた場合の情報量の変化を示す。ここでは映像のフレーム当りの情報量として、次の 2 つを用いた。

- (i) フレーム間差分信号 (動き補償無し) のエントロピ
- (ii) 動き補償誤差信号及び動ベクトルのエントロピ

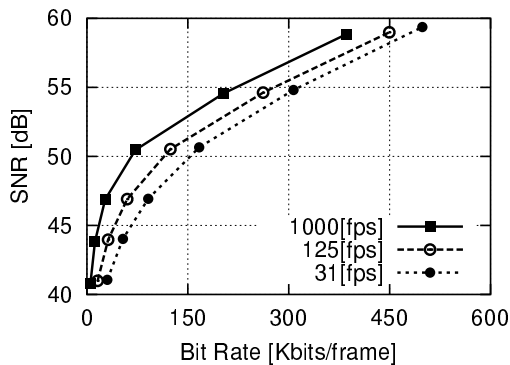
なお、動き補償のブロックサイズは 16×16[画素] とした。この結果、(i) と (ii) のいずれの情報量もフレームレートの増加とともに単調に減少することが確認できる。ただし、両者の減少率には違いがあり、高いフレームレートでは、(i)(ii) の情報量はほぼ等しくなっている事が確認できる。これは、高フレーム



(a) テニス [シーン 1]



(b) テニス [シーン 2]

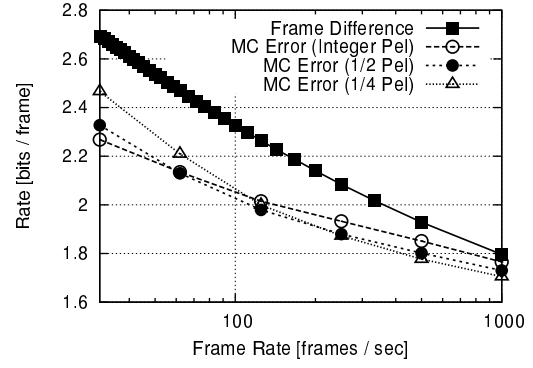


(c) 野球

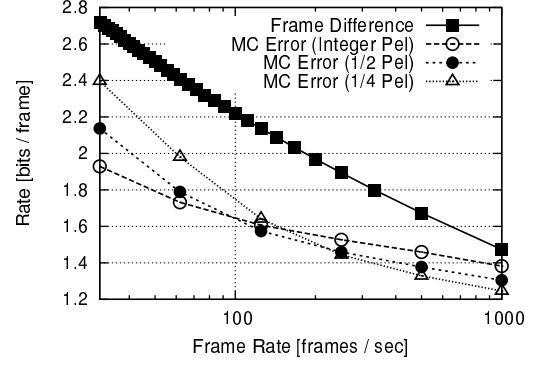
図 2: フレームレートの異なるシーケンスに対する H.264 による符号化結果

レートの場合、単純差分による符号化が効果的であることを示唆している。

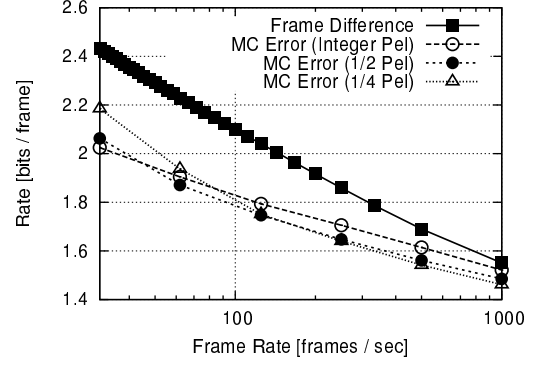
図 3 に示した単純差分信号のエントロピは、フレームレートの対数に比例することが確認できる。これは、従来の低フレームレートの観測では得られていない新たな知見である。そこで、フレームレート F [frames/sec] と差分信号のフレーム当りのエントロ



(a) テニス [シーン 1]



(b) テニス [シーン 2]



(c) 野球

図 3: フレームレートとフレーム当りの符号量 (動き探索のブロックサイズは 16×16 [画素] とした)

ピ $I(F)$ [bits/frame] の関係を次式のようにモデル化する。

$$I(F) = \alpha \log\left(1 + \frac{\beta}{F+1}\right) + n \quad (1)$$

ここで、 α 、 β は画像信号に応じて定まる定数であり、 n は誤差信号の情報量である。

以下では、式 (1) のモデルの物理的な妥当性に関して検証する。式 (1) のモデルは次式の間係を満足

する。

$$\lim_{F \rightarrow \infty} = n \quad (2)$$

$$\frac{dI(F)}{dF} < 0 \quad (3)$$

$$\lim_{F \rightarrow 0} = \alpha \log(1 + \beta) \quad (4)$$

式(2)は、フレーム間隔が極限まで小さくなる(フレームレートが極限まで大きくなる)と、フレーム間差分信号の情報量として誤差信号の情報量が支配的になることを示している。式(3)は

$$\frac{dI(F)}{dT} > 0$$

と等価である。ここで、 T は隣接フレーム間の時間間隔であり、フレームレート F の逆数で表わされる。つまり、フレーム間隔の増大(フレームレートの減少)と共にフレーム間差分信号の情報量も増加することを示している。ただし、この情報量は無限に増加することはない、式(4)が示すように、フレーム間隔が極限まで大きくなった場合(フレームレートが極限まで小さくなった場合)であっても、フレーム間差分信号の情報量は一定値で抑えられる。これは、フレーム内の情報量が有限であることと符合する。このように、式(1)のモデルは、増加率のみならず、極限值に関しても動画像の物理的な特性と適合していることが確認できる。

4 おわりに

本報告では、高フレームレート映像の高エネルギー符号化を睨み、フレームレートと情報量の関係について定量的な調査を行なった。この結果、差分信号の情報量がフレームレートの対数に比例するという関係を突き止めた。さらに、同情報量とフレームレートとの関係を表わすモデルを構築するとともに、その妥当性を検証し、構築したモデルが動画像の物理的な特性と適合していることを確認した。

参考文献

- [1] 中嶋正之他. 小特集 デジタルシネマの最新事情. 映メ誌, Vol. 57, No. 2, pp. 178-202, Feb. 2003.
- [2] 高畑文雄他. 小特集 大画面ディスプレイ. 映メ誌, Vol. 56, No. 8, pp. 1209-1232, Aug. 2002.

- [3] 羽石秀昭. マルチスペクトルに基づく映像技術. In *PCSJ2003*, pp. 1-6, 2003.
- [4] L.Spillmann and J.S.Werner. *Visual perception the neurophysiological foundations*. Academic Press, 1990.
- [5] 武川直樹, 川西久, 松岡毅. フレーム間符号化における駒落とし方式の検討. 信学論, Vol. J65-B, No. 4, pp. 455-462, Apr. 1982.
- [6] H. Song and C.-C.J.Kou. Rate control for low-bit-rate video via variable encoding frame rates. *IEEE trans. CSVT*, Vol. 11, No. 4, pp. 512-521, 2001.
- [7] 稲積泰宏, 吉田俊之, 酒井善則, 堀田裕弘. 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の推定. 信学論, Vol. J85-B, No. 7, pp. 1130-1142, 2002.
- [8] ISO/IEC 13818-2. *Information technology - Generic coding of moving pictures and associated audio information : Video*, 1996.
- [9] ISO/IEC JTC 1/SC 29/WG11 N2725. *Overview of the MPEG-4 Version 1 standard*, 1999.
- [10] Joint Video Team. *Draft ITU-T Recommendation and Final draft international standard of joint video specification*. ITU-T Rec.H.264 and ISO/IEC 14496-10 AVC, 2003.