

J-074

# 高フレームレート映像信号の符号量に関する理論モデルの構築 Theoretic encoding model of high-frame-rate video signal

坂東幸浩†,

高村誠之†,

上倉一人†,

八島由幸†

Yukihiro BANDO, Seishi TAKAMURA, Kazuto KAMIKURA and Yoshiyuki YASHIMA

## 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 [1]。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、充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 [2]。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。一方、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。

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。フレームレートの増加に伴いフレーム間予測誤差信号のフレームあたり情報量は減少することが予想される。しかし、その定量的な性質は従来把握されていなかった。従来の検討は、60 [フレーム/秒] 以下の映像に関するものが中心であり、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。

そこで、本報告では、高フレームレート映像信号の符号化を睨み、フレームレートとフレーム間予測誤差信号の情報量を結び付ける解析的なモデルの導出を行う。また、実映像を用いた符号化実験を通して、同モデルの妥当性を検証する。

## 2 フレームレートと情報量の関係

### 2.1 理論モデルの導出

フレームレートと情報量の関係を解析的に導出する。ここでは、簡単のため 1 次元信号を例にとり説明する。時刻  $t$  の位置  $x$  における信号  $f_t(x)$  に対して、区間  $B$  を単位として動き補償 (推定変位量  $\hat{d}$ ) を行った場合、そ

の区間内の動き補償後の予測誤差は次のように表現できる。なお、以下では、位置  $x$  における真の変位量を  $d(x)$  としている。

$$\begin{aligned}\sigma^2 &= \sum_{x \in B} |f_t(x) - f_{t-1}(x + \hat{d})|^2 \\ &= \sum_{x \in B} |f_{t-1}(x + d(x)) - f_{t-1}(x + \hat{d}) + n|^2 \\ &= \sum_{x \in B} \left| \left\{ \frac{d}{dx} f_{t-1}(x) \right\} (d(x) - \hat{d}) + \phi(x) + n \right|^2 \quad (1)\end{aligned}$$

ここで、 $\phi(x)$  はテイラー展開の 2 次以降の項である。また、 $n$  は雑音成分に起因する項である。雑音の要因としては、後述の高速度カメラにより撮影した映像の場合、シャッターの開口時間が短いことにより顕在化する撮像デバイス等の熱雑音があげられる。フレーム間での動きを等速運動とみなせば、変位量はフレーム間隔 (フレームレートの逆数) に比例することから、フレームレート  $F$  とは反比例の関係にある。このため、 $d(x) - \hat{d}$  は次式のように表せる。

$$d(x) - \hat{d} = (v(x) - \hat{v}) \cdot F^{-1}$$

ここで、 $v(x)$  は位置  $x$  の画素毎に定まる比例定数であり、 $\hat{v}$  は区間  $B$  に対して定まる比例定数である。両比例定数の物理的な意味は、いずれも速度となる。

上式を式 (1) に代入し、雑音成分の独立性、およびテイラー展開の一次近似を利用し、次式を得る。

$$\sigma^2 \simeq AF^{-2} + BF^{-1} + N \quad (2)$$

ここで、 $A, B, N$  は以下の通りである。

$$\begin{aligned}A &= \sum_{x \in B} \left\{ (v(x) - \hat{v}) \frac{d}{dx} f_{t-1}(x) \right\}^2 \\ B &= 2 \sum_{x \in B} \left\{ (v(x) - \hat{v}) \frac{d}{dx} f_{t-1}(x) \right\} \phi(x) \\ N &= \phi(x)^2 + n^2\end{aligned}$$

予測誤差がラプラス分布に従うと仮定し、同分布の differential entropy を用いて、予測誤差の情報量  $I(F)$  を見積もると次式の関係を得る。

$$\begin{aligned}I(F) &= \log(\sigma) \\ &= \frac{1}{2} \log(AF^{-2} + BF^{-1} + N) \quad (3)\end{aligned}$$

†日本電信電話株式会社 NTT サイバースペース研究所

## 2.2 理論モデルの漸近特性

式(3)のモデルは次式の関係満足する。

$$\lim_{F \rightarrow \infty} = \frac{1}{2} \log(N) \quad (4)$$

$$\frac{dI(F)}{dF} < 0 \quad (5)$$

式(4)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、フレーム間差信号の情報量として雑音成分の情報量が支配的になることを示している。式(5)は次式と等価である。

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

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信号の情報量が増加することを示している。このように、フレームレートの増減に伴う情報量の変動、およびその極限值に関して、式(3)のモデルの振る舞いは映像信号の物理的な性質と適合していることが確認できる。

## 2.3 実験

上述のモデルの妥当性を検証するために、実画像を用いた符号化実験を行った。高フレームレートの映像ソースは、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]とし、シャッター時間はフレーム間隔に等しいものとした。得られた画像はRGBカラー映像であり、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。この1000[フレーム/秒]の映像をソースとして、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た。

図1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。黒丸は実画像の符号化結果を示しており、実線は式(3)のモデルによる理論値である。ここで、モデルのパラメータ( $A, B, N$ )は最小二乗推定により求めた値を用いた。横軸はフレームレートを対数スケールで表わしている。縦軸は、映像の情報量として、動き補償誤差信号及び動ベクトルのエントロピの和を画素平均で表している。なお、動き補償時のブロックサイズは16×16[画素]とした。動き推定の精度は $\frac{1}{4}$ [画素]であり、小数画素位置の補間にはMPEG-4の補間フィルタを用いた。フレーム間予測は片方向予測とし、その際の参照フレームは直前のフレームとした。図1に示すように、導出したモデルの理論値はフレームレートの増加に伴うフレ

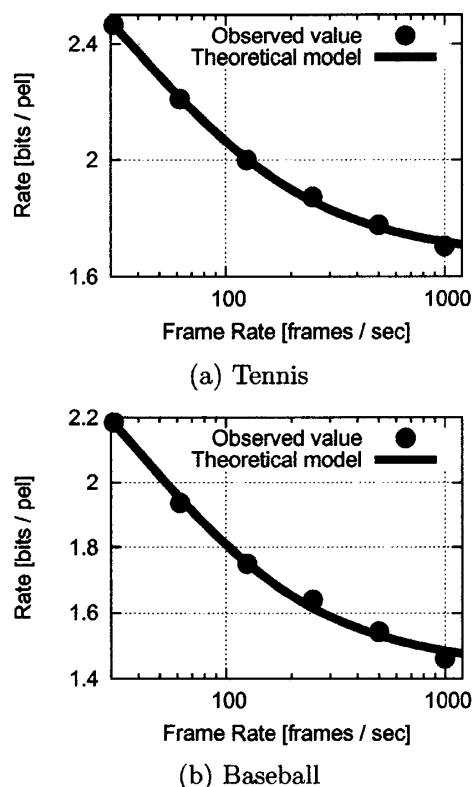


図1: フレームレートとフレーム当りの符号量

ム間予測誤差の情報量の減少を適切に表現できていることが確認できる。また、こうした理論値と実測値の関係は複数のコンテンツおよびシーンに対する符号化結果に対して確認できた。本実験結果は、モデルを導出する際の仮定、および、導出されたモデルの妥当性を裏付けるものと言える。

## 3 おわりに

本報告では、フレーム間予測誤差信号の情報量とフレームレートの関係を表すモデルを解析的に導出した。実画像を用いた符号化実験の結果、同モデルの理論値が実測値に対する良好な近似値を与えることを確認した。今後、本モデルを用いた高フレームレート映像システムの構築を検討する予定である。

## 参考文献

- [1] 中嶋正之他. 小特集 デジタルシネマ. 映メ誌, Vol. 59, No. 2, pp. 199-217, Feb. 2005.
- [2] L.Spillmann and J.S.Werner. *Visual perception the neurophysiological foundations*. Academic Press, 1990.