

第 3 章

測定データとデータ解析

本章では、MPEG2 と呼ばれる動画像圧縮方式によって圧縮された VBR トラフィックが第 2 章で説明した自己相似性をもつかどうかを調べる。

MPEG2 の標準化によってインターネット上等での動画像の転送が可能になり、将来的には VOD 等にも利用されることが予想される。また DVD や SKY Perfect TV! 等の CS 放送においてもこの圧縮技術が既に用いられている。本章ではまず、MPEG2 動画像圧縮方式のもとになった MPEG1 について説明し、その後 MPEG2 によってどのように画像が圧縮されるかについて説明する。更に、そのように圧縮されたデータがどのように ATM 網上を流れるかについて説明する。

本論文ではこの MPEG2 によって VBR 符号化された実測データを用いる。画像のソースとして、郵政省通信総合研究所所有の資料映像の Laser Disk (LD) 3 枚と NEC C&C メディア研究所で実測した 10 時間の放送画像を用いた。これらの画像データを R/S 解析、分散解析、及びスペクトル解析と呼ばれる手法でそれぞれ解析し、その結果についてまとめる。

第 3.1 節では、MPEG1 ビデオについて説明し、それを拡張した MPEG2 ビデオについて第 3.2 節で説明する。第 3.3 節では、MPEG2 によって圧縮符号化されたデータがどのようにして ATM 網を流れるセルに分解されるかを説明する。第 3.3 節までが MPEG についての説明である。第 3.4 節は、MPEG2 によって圧縮されたデータを解析する際に用いた手法 (R/S 解析、分散解析、及びスペクトル解析) についての説明である。そして、今回用いたデータについて説明したのが第 3.5 節である。第 3.5 節で説明したデータに対して行った解析結果を第 3.6 節に示す。

3.1 MPEG1 ビデオ

まずはじめに、主として CD-ROM のための動画像圧縮方式として考えられてきた MPEG1 ビデオについて説明する。MPEG1 ビデオは、ISDN 対応のテレビ電話/会議用の動画像圧縮の国際標準 H.261 とカラーファックスや電子スチルカメラ (デジタルカメラ) のためのカ

ラー静止画像圧縮の国際標準 Joint Photographic coding Experts Group (JPEG) を継承して標準化されたものである [27, 29].

MPEG1 ビデオ符号化も H.261 で用いられたハイブリッド符号化の技術を用いている. ハイブリッド符号化は動画像圧縮の基本原理として考えられており, 空間的 (離散コサイン変換; Discrete Cosine Transfer (DCT)), 時間的 (動き補償), 及び統計的 (可変長符号化) 情報圧縮を用い, 全体の符号発生量を量子化ステップで制御するというものである. National Television System Committee (NTSC) を採用している地域では, 毎秒約 30 枚のフレームを符号化, 及び再生することで完全な動画像を実現している. NTSC 方式で伝送される情報量は, 1 画面の画素数が $84480 (= 352 \times 240)$ の場合約 60 Mega bits per second (Mbps)¹ であるが, これを MPEG1 を用いると 1.5 Mbps にまで圧縮することができる.

MPEG1 ビデオでは更に 1 枚の画面 (フレーム) を 3 つのタイプのフレームに分類している. 以下において, その 3 つのフレームについて説明する.

- Intra coded Frame (I フレーム)
時間的に他画面とは独立して符号化される画面
- Predictive coded Frame (P フレーム)
時間的に過去に位置する I フレームまたは P フレームから差分によって表現する画面
- Bidirectional predictive coded Frame (B フレーム)
時間的に前後に位置する I フレームまたは P フレームを用いてその間に挿入する画面

MPEG1 ビデオでの符号化順序は, まずその画面情報だけで完結した画面の I フレームを生成し, その I フレームから順方向予測により P フレームを生成する. 次に, I または P フレームの間に挿入する B フレームを双方向予測により生成する. これを示したのが図 3.1 である. この 3 つのフレームを適切に組み合わせて Group of Pictures (GOP) を作成する. GOP によってランダムアクセスが可能になる. 通常この GOP に含まれるフレームの数 (I フレームが現れる周期) は 15 であり, I フレームまたは P フレームが現れる周期は 3 である [27, 28, 29, 30].

3.2 MPEG2 ビデオ

MPEG2 ビデオは DVD のような高画質のデジタル映像蓄積メディアに対する標準化として考えられてきた. MPEG2 ビデオは MPEG1 ビデオの基本特徴をすべて備えている. そして, 更に以下のような新しい機能が付加されている.

- インタレース画像への対応

¹実際のテレビでは NTSC の有効走査線数は 480 本, 有効標本点数は 720 であり, それを R(赤), G(緑), B(青)それぞれ 8 bits で量子化するとフレーム 1 枚あたり 8.3 M bits/frame になる. NTSC 方式では単位秒あたりのフレーム数は 29.97 であるから 1 秒間に 250 M bits の情報量が必要になる. ここでは, 有効走査線数と有効標本点数をそれぞれ半分にした場合を考えている.

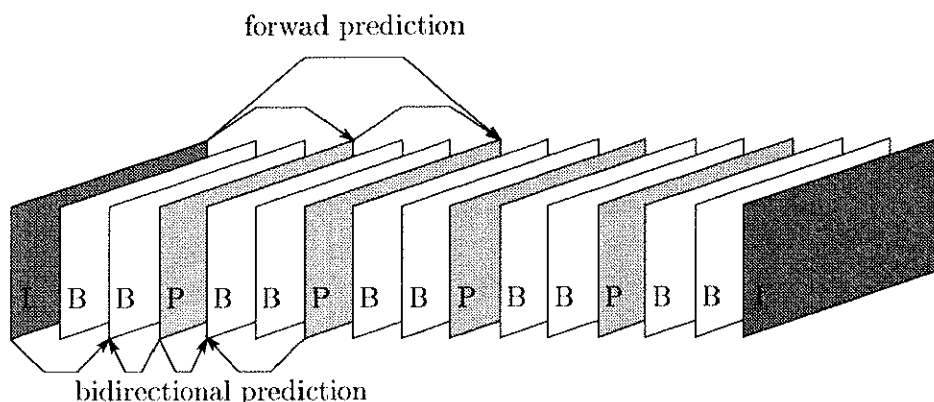


図 3.1: MPEG のフレーム列と GOP.

- スケーラビリティ
- High Definition Television (HDTV) と VBR への対応

これらについて簡単な解説を以下に示す [29].

3.2.1 インタレース画像への対応

インタレース走査とは、1枚のフレームを異なった時間(通常1/60秒ごと)でサンプリングして、1ラインごと(奇数番目の走査線と偶数番目の走査線)を交互に表示させたものである。こうすることによって、アナログテレビジョン信号の伝送帯域を半分にすることができる。インタレース画像に対応することによって、MPEG1ビデオで採用されていたフレーム予測の他にフィールド予測が可能になる。フィールドとは、インタレース走査で異なる時間にサンプリングされた画像である。MPEG1ビデオでも採用されていたフレーム予測では、2つの半画面(奇数ラインのものと偶数ラインのもの)から1つのフレームを作り、それを圧縮処理する。圧縮処理の手順は、

1. I フレームを DCT により生成 → 第 1 画面の DCT による圧縮情報.
2. I フレームをもとに P フレームを順方向予測により圧縮 → 第 1 画面からの圧縮情報.
3. I フレームと P フレームをもとに B1, B2 フレームを双方向予測により圧縮 → 第 1 画面と第 4 画面からの圧縮情報.

となっている。この一連の流れを示したものが図 3.2 である[27, 29].

それに対しフィールド予測は、各フィールドを走査線が奇数ラインだけのものと偶数ラインだけのものに分けて、一番初めのフィールドを I フィールドとして、次のフィールドを P フィールドにする。この 2 枚のフィールドを参照画面として次の P フィールドを作成する。更に、この 2 枚の P フィールドを参照画面としてその次の P フィールドを作成する。この

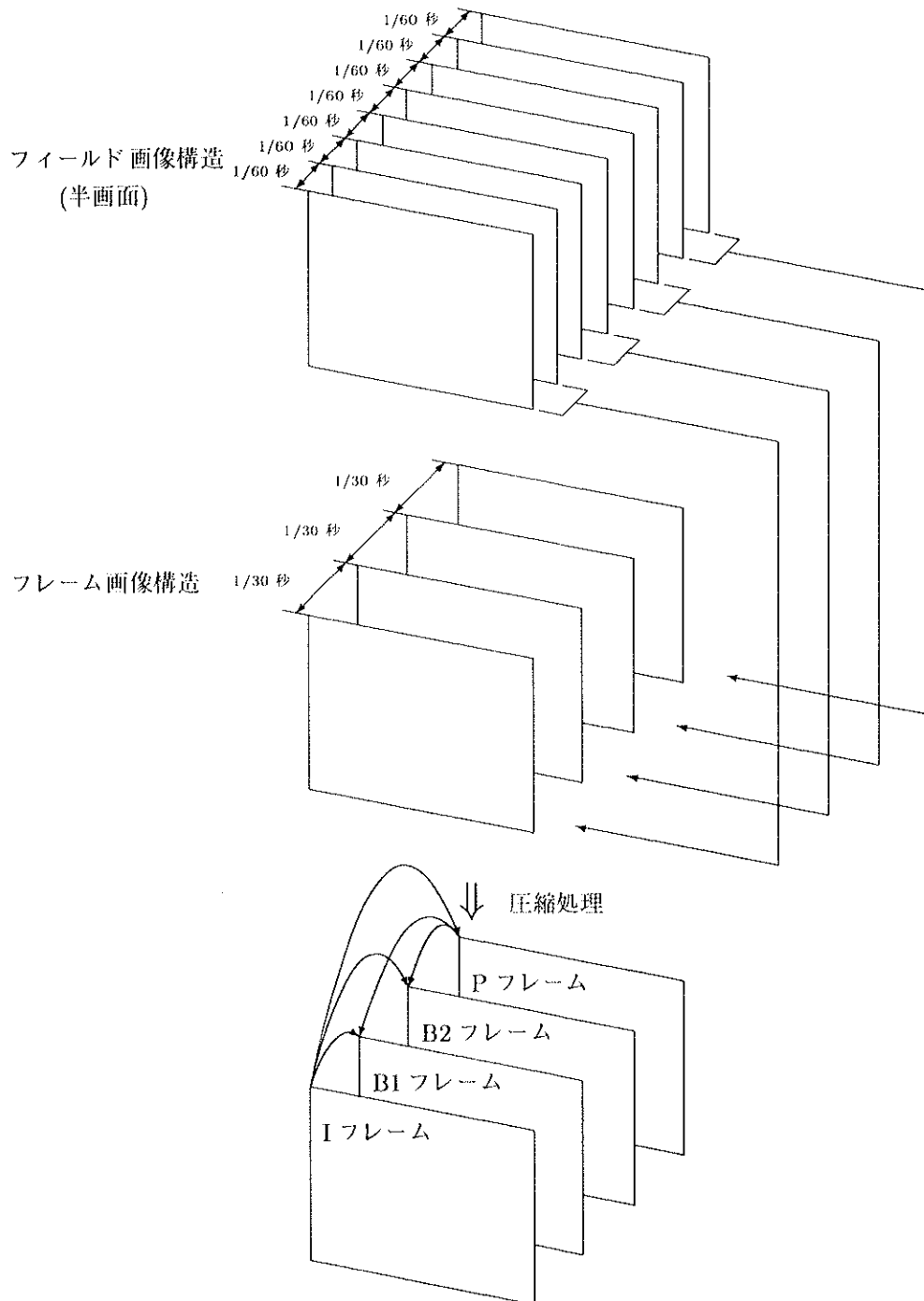


図 3.2: MPEG2 におけるフレーム構造による双方向予測の方法.

ようにして作られた、I、P フィールド 4 枚を参照して、それらの間に B フィールドを 4 枚作成する。圧縮処理の順序は以下のようにになっている。

1. 各画面を走査線が奇数ラインのもの (奇数ライン画面群) と偶数ラインのもの (偶数ライン画面群) に分ける (図 3.3).
2. I1 フィールドを DCT により生成 → 第 1 画面の DCT による圧縮情報画面.
3. I1 フィールドをもとに、P1 フィールドを順方向予測により生成 (図 3.3) → 第 2 画面の第 1 画面との差分情報.
4. I1 フィールド、P1 フィールドをもとに、P2 フィールドを順方向予測により生成 (図 3.4: 圧縮処理 1) → 第 7 画面の第 1、第 2 画面との各々の差分情報.
5. P1 フィールド、P2 フィールドをもとに、P3 フィールドを順方向予測により生成 (図 3.4: 圧縮処理 1) → 第 8 画面の第 2、第 7 画面との各々の差分情報.
6. I1、P1、P2、P3 から B1、B2 フィールドを双方向予測により生成 (図 3.4: 圧縮処理 2) → 第 3 画面の第 1、2、7、8 画面との各々の差分情報、第 4 画面の第 1、2、7、8 画面との各々の差分情報.
7. I1、P1、P2、P3 から B3、B4 フィールドを双方向予測により生成 (図 3.5: 圧縮処理 3) → 第 5 画面の第 1、2、7、8 画面との各々の差分情報、第 6 画面の第 1、2、7、8 画面との各々の差分情報.

この一連の流れを示したのが図 3.3、図 3.4、及び図 3.5 である [27, 29].

フィールド予測の場合、最初に符号化されるフィールドが I フィールドになり、同一フレームのもう一方のフィールドは I フィールドか P フィールドのいずれかを撰択できる。それ以外の場合、同一フレームでは 2 つのフィールドは同じタイプでなくてはならない。

また、フレーム予測とフィールド予測では、参照画面も変わってくる。フレーム予測の場合、圧縮方法は MPEG1 ビデオと同じなので P フレームを作成するとき最近の I もしくは P フレームを参照し、B フレームを作成するとき前後の I もしくは P フレームを参照する。フィールド予測の場合は、P フィールドは最近の I もしくは P フィールド 2 枚を参照する。但し、第 1 フィールドが I フィールドとして符号化され、第 2 フィールドが P フレームとして符号化される場合は例外である。この P フィールドは第 1 フィールドである I フィールドだけしか予測に使うことはできない。B フィールドは前後の I もしくは P フィールドを各 2 枚計 4 枚参照して作成される。これは MPEG1 ビデオからの拡張になっている。

3.2.2 スケーラビリティ機能

スケーラビリティとは、ビット列全体からある部分を取り出しても、部分的にデコードが可能となることである。MPEG2 ビデオで考えられているスケーラビリティには次の 3 つがある。

- 空間スケーラビリティ

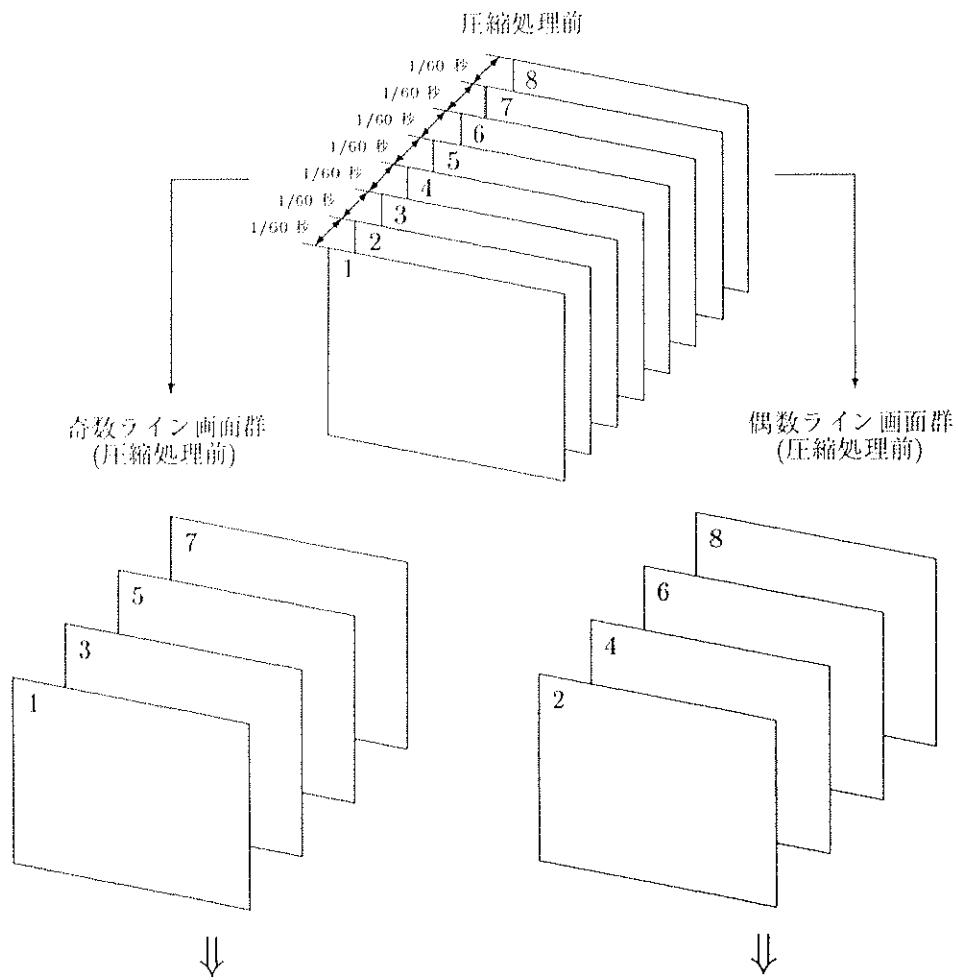


図 3.3: フィールド構造による双方向予測の方法 (I).

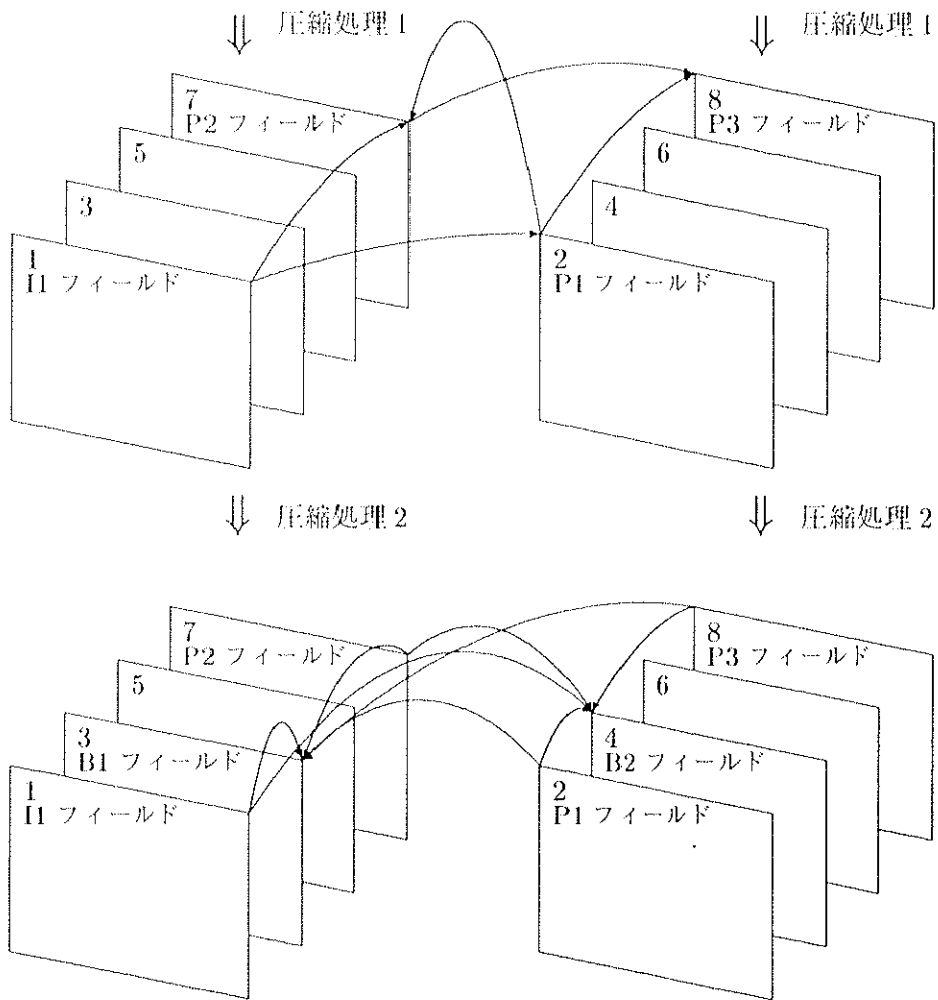


図 3.4: フィールド構造による双方向予測の方法 (II).

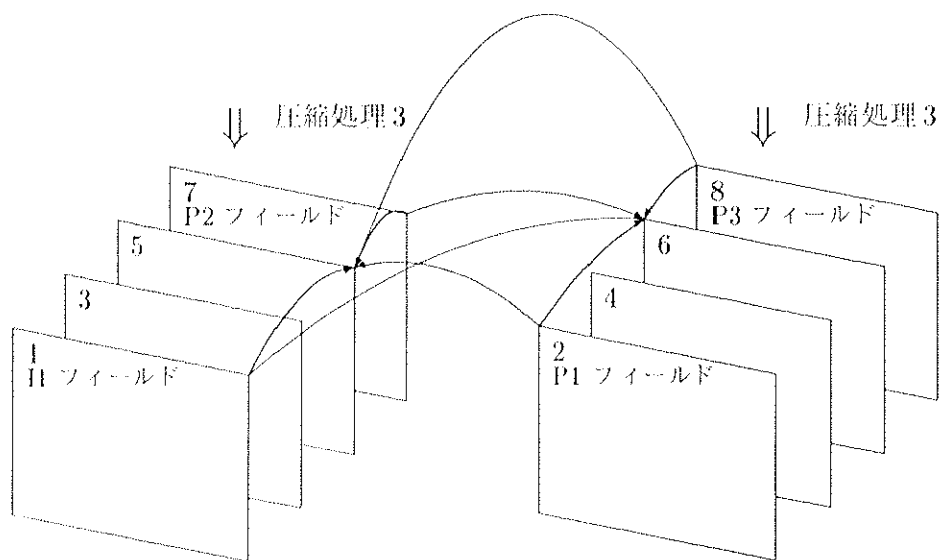


図 3.5: フィールド構造による双方向予測の方法 (III).

解像度の小さいレイヤを基本レイヤ、大きいレイヤを高位レイヤとする。

基本レイヤは通常 MPEG2 ビデオで符号化されていて、これをアップ・サンプルして高位レイヤと同じ画像を作り出す。² 高位レイヤは高位レイヤ中の画像のみから予測するのではなく、アップ・サンプルした画面からも予測することにより、より良く符号化することができる。空間スケーラビリティにより低解像度ディスプレイに対しては基本レイヤ情報だけから表示し、高解像度ディスプレイに対しては基本レイヤと高位レイヤの両方の情報を用いて高解像度表示することができる。

- Signal to Noise Ratio³ (SNR) スケーラビリティ

画質の低いレイヤを基本レイヤといい、高いレイヤを高位レイヤという。

基本レイヤは MPEG1 ビデオまたは MPEG2 ビデオで符号化されている。高位レイヤは基本レイヤで符号化されたデータを再現し、もとのデータから再現されたものだけを引き、この誤差分だけを基本レイヤの量子化ステップサイズより小さいステップサイズで量子化を行ない符号化する。高位レイヤではくっきりした画像になるのに対し、基本レイヤではぼんやりした画像になる。

- 時間スケーラビリティ

時間解像度の小さいレイヤを基本レイヤといい、大きいレイヤを高位レイヤという。

基本レイヤは通常 MPEG2 で符号化されている。これに対し、高位レイヤでは高位レイヤ中の I, P, B フレームを使うだけでなく、基本レイヤの画像を使って画像周期の短いフレームが予測可能になる。例えば、基本レイヤとして 1/30 秒周期とし、高位レイヤとして 1/60 秒周期とする。このとき、基本レイヤと高位レイヤを組み合わせると、高位レイヤ間のみの情報に基本レイヤからの予測情報を加えることができるようになる。これによって 1/60 秒周期の画像再生が可能になる。

これらのスケーラビリティは、ビット列全体を高位レイヤとし、その部分列を基本レイヤにしている [27, 29]。

3.3 MPEG2 と ATM

動画像の情報量は非常に多く、それを網へ流すには高速な回線が必要になる。その例として考えられるのが ATM 網である。ATM 網上では、すべてのデータは 53 バイトのセル⁴ と呼ばれる単位で伝送交換される。この ATM 網の上で VOD 等のサービスが提供されることが予想される。

²画素を補間して低解像度画面から高解像度画面を作り出すことをアップ・サンプルという。

³もとの高精細度画像に対して量子化を行なう際に、量子化値の粗さによって、もとの画素値とのズレ (雑音) が生じる。つまり、(原信号値/雑音値) のことを SNR という。この値が大きいほど高画質になる。

⁴ATM セルの 53 バイトのうち、48 バイトが実際のデータで、残りの 5 バイトは宛先の情報等を含むヘッダである。

そこで本節では、MPEG2によって符号化されたビット列がどのように ATM のセルに分割されるかを説明する。MPEG ビデオ等で符号化されたビット列を実際のアプリケーションに適用する場合には蓄積メディアやネットワークがもつ、固有の物理フォーマットやプロトコルに適合したデータ形式にする必要がある。このような作業を行なうのが MPEG システムの役割である [27]。

MPEG1 システムでは、任意の数のビデオとオーディオ等の、個別の符号化ビット列を多重化して、1組のプログラムとして1本化されたデータ列を構成する。主なメディアとしては CD-ROM が使用されていた。それに対して、MPEG2 システムは幅広いアプリケーションに対応している。そのため MPEG1 システムと同様に1つのプログラムを構成するプログラム・ストリーム (MPEG2 - Program Stream: MPEG2 - PS) と、複数のプログラムを構成することができるトランスポート・ストリーム (MPEG2 - Transport Stream: MPEG2 - TS) の2種類の方式が存在する。プログラム・ストリームは主に蓄積メディアに用いられ、トランスポート・ストリームは通信ネットワークや放送に利用される。トランスポート・ストリームにより、複数のプログラムを1本のデータ列にすることが可能になり、テレビ放送にも対応できるようになった。

MPEG2 システムによって ATM のセルは次のように生成される。符号化された各フレームデータをビデオ・エレメンタリー・ストリーム (Video Elementary Stream: ビデオ ES) という。このビデオ ES を更に適当なサイズに分割したものをパケタイズドエレメンタリー・ストリーム・パケット (Packetized Elementary Stream: PES パケット) という。MPEG システムでは、ビデオとオーディオの同期を取るため復号時刻情報 (Decoding Time Stamp: DTS) と提示時刻情報 (Presentation Time Stamp: PTS) を、各メディアのアクセス・ユニットごとに多重化し、メディア間の同期を実現している。アクセス・ユニットの提示単位はビデオではフレームであり、オーディオでは符号化フレームである。具体的には I フレームと P フレームには DTS と PTS の両方がつき、B フレームには PTS のみがつく。PES パケットを作るときビデオ PES やオーディオ PES の中の最初に現れるアクセス・ユニット (ビデオではフレーム、オーディオではシンクバイト) の先頭の時刻情報を PES パケットのヘッダに多重化する [27, 29]。

PES パケットを 184 バイトずつ分割して 4 バイトのヘッダを付け、更にオーディオ PES を付加させたものがトランスポート・ストリーム・パケット (Transport Stream Packet: TS パケット) である。最後の余った部分にスタッフィングバイトを入れることによって TS パケットが生成される (図 3.6, [29])。スタッフィングバイトとは TS パケットが 184 バイトになるように挿入されるダミーバイトのことである。スタッフィングバイトは各フレームから作られる一番最初の TS パケットのヘッダのあとにあるアダプテーション・フィールドによって制御されている。TS パケットを 2 つ 1 組にして 8 バイトのトレイラを付ける。これを 48 バイトずつに分割し、それぞれに 5 バイトを付けることによって ATM セルが生成される (図 3.7) [27, 29, 92]。

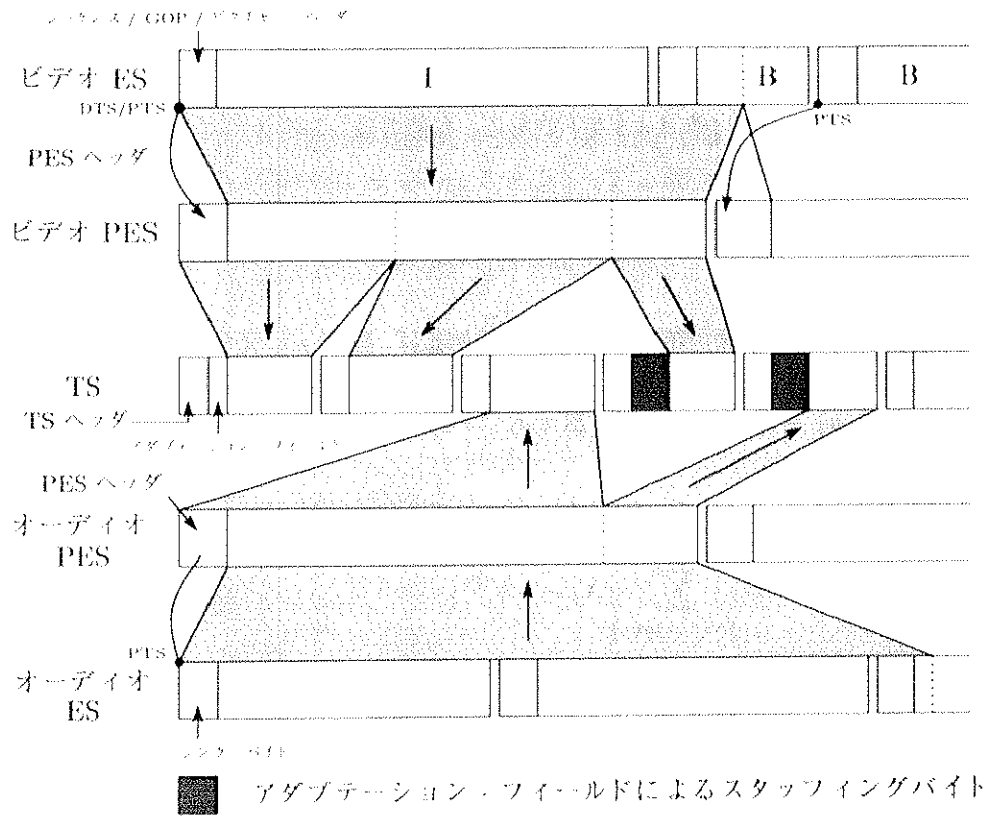


図 3.6: TS パケットの生成過程.

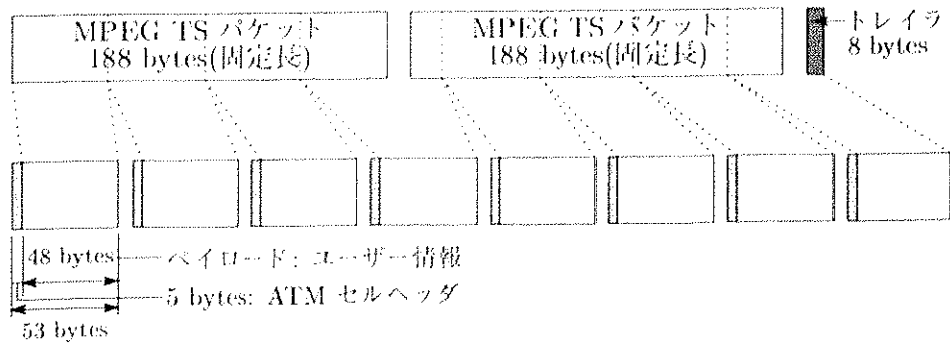


図 3.7: ATM セルの生成過程.

3.4 解析方法

第3.3節までは、MPEG 動画像圧縮方式について説明した。第3.4節からは実際の時系列データを解析する方法を説明する。本論文で用いた解析方法は、Hurst によって提案された R/S 解析 (第3.4.1節)、集積過程の分散を用いる分散解析 (第3.4.2節)、及びスペクトル密度関数を用いたスペクトル解析 (第3.4.3節) である。本節では、この3つ手法について説明する。

3.4.1 R/S 解析

確率変数 $\{X_j; j = 1, 2, \dots\}$ は時刻 j における量 (例えば、時刻 j に発生するセル数等) を表し、確率変数 $\{Y_n; n = 1, 2, \dots\}$ は時刻 1 から時刻 n までの合計、つまり、 $Y_n = \sum_{j=1}^n X_j$ とする。更に、 $Z(i, m, n)$ を以下のように定義する。

$$\begin{aligned} Z(i, m, n) &= Y_{(i-1)m+n} - Y_{(i-1)m} - \frac{i}{m} (Y_{im} - Y_{(i-1)m}) \\ &= \sum_{j=(i-1)m+1}^{(i-1)m+n} X_j - nX_i^{(m)} \end{aligned}$$

$Z(i, m, n)$ は時系列データを m 個ずつのグループに分けたとき (図 2.1 参照)、 i 番目のグループの中で n ($1 \leq n \leq m$) 番目までの確率変数の和がそのグループ内の平均に対してどの程度ずれているかを表す。 $Z(i, m, n)$ の最大値と最小値を用いて、adjusted range $R(i, m)$ という量を定義する。

$$R(i, m) = \max_{1 \leq n \leq m} Z(i, m, n) - \min_{1 \leq n \leq m} Z(i, m, n)$$

例えば、 $\{X_j; j = 1, 2, \dots\}$ を時刻 j におけるダムに入ってくる水量とし、一方で毎時 $\{X_i^{(m)}; i = 1, 2, \dots\}$ だけ水を放水すると仮定する。このとき、 $\max Z(i, m, n)$ はダムに貯まる最多貯水量で、 $\min Z(i, m, n)$ は最少貯水量である。濁水と洪水を防ぐには、最低 $R(i, m)$ だけダムの容量が必要になる [40, 101]。

標本標準偏差

$$S(i, m) = \sqrt{\frac{1}{m} \sum_{n=(i-1)m+1}^{im} (X_n - X_i^{(m)})^2}$$

に対する $R(i, m)$ の比を R/S 統計量という。つまり

$$\frac{R}{S} = \frac{R(i, m)}{S(i, m)}$$

が R/S 統計量である。

この統計量の対数を取ると

$$\log E \left[\frac{R}{S} \right] \approx a + \log m^H$$

と近似できる. ここで, a は定数とする. この H は Hurst パラメタと呼ばれ, 自己相似過程で議論してきた自己相似パラメタの H と同じ量を表す. R/S 統計量を両対数グラフにプロットするときその傾きが Hurst パラメタ H となることが分かる. Hurst は, ナイル川のデータ解析から経験的に $m \rightarrow \infty$ のとき, $1/2$ より大きい H に対して R/S 統計量が m^H に比例することを発見した [40, 101]. 理論的には, $\{X_n; n = 1, 2, \dots\}$ が i.i.d. ならば $m \rightarrow \infty$ のとき $(R/S)/\sqrt{m}$ はある分布に法則収束し,

$$\begin{aligned} E \left[\frac{R}{S} \right] &\approx \left(\frac{\pi}{2} \right)^{1/2} m^{1/2} \\ \text{var} \left[\frac{R}{S} \right] &\approx \left(\frac{\pi^2}{6} - \frac{\pi}{2} \right) m \end{aligned}$$

となることが示される [21, 23, 101].

3.4.2 分散解析

第 2.2 節より, 時系列データが自己相似性をもつとき, 確率変数列 $\{X^{(m)}; m = 1, 2, \dots\}$ の標本分散はデータ数 m の逆数より遅く収束することが分かる.

$$\text{var} \left[X^{(m)} \right] \approx c_{\text{var}} m^{-\beta} \quad 0 < \beta < 1$$

このことは, $\text{cov}(X_{j_1}, X_{j_2})$ を計算することにより分かる [36].

$$\sum_{j_1=1}^m \sum_{j_2 \neq j_1}^m \text{cov}(X_{j_1}, X_{j_2}) = 2 \sum_{k=1}^m (m-k) \gamma(k) = 2\sigma^2 \sum_{k=1}^m (m-k) \rho(k)$$

となるので,

$$\begin{aligned} \text{var} \left[X^{(m)} \right] &= \frac{1}{m^2} \text{var} \left[\sum_{j=(i-1)m+1}^{im} X_j \right] \\ &= \frac{1}{m} \sigma^2 + \frac{1}{m^2} \sum_{j_1=1}^m \sum_{j_2 \neq j_1}^m \text{cov}(X_{j_1}, X_{j_2}) \\ &= \frac{\sigma^2}{m} \left[1 + \frac{2}{m} \sum_{k=1}^m (m-k) \rho(k) \right] \end{aligned} \quad (3.1)$$

となる. ここで, $\gamma(k)$ は自己共分散関数, $\rho(k)$ は自己相関係数である. $\gamma(k)$ と $\rho(k)$ の間には $\rho(k) = \sigma^{-2} \gamma(k)$ の関係がある. このとき, 2つの確率変数が独立ならば共分散 $\text{cov}(X_{j_1}, X_{j_2}) =$

0 となり自己相関係数が 0 になって、確率変数列 $\{X^{(m)}; m = 1, 2, \dots\}$ の分散は標本数 m の逆数に比例する.

$$\text{var} [X^{(m)}] = \sigma^2 m^{-1}$$

また、自己相似性をもつことが予想される場合は自己相関係数が $\rho(k) = c_\rho k^{-\beta}$ となるので、これを式 (3.1) に代入すると

$$\text{var} [X^{(m)}] = \frac{\sigma^2}{m} \left[1 + \frac{2}{m} \sum_{k=1}^m (m-k) c_\rho k^{-\beta} \right] \quad (3.2)$$

となり、また

$$\begin{aligned} \sum_{k=1}^m k^{-\beta} &= 1 + \sum_{k=2}^m k^{-\beta} \\ &< 1 + \int_1^m x^{-\beta} dx = 1 + \frac{1}{1-\beta} (m^{1-\beta} - 1) \end{aligned}$$

となることから、

$$\text{var} [X^{(m)}] \approx c_\rho m^{-\beta}$$

が得られる. β と 自己相似パラメタ H の間には $\beta = 2 - 2H$ の関係が知られている. つまり、ここで、 $\log_{10} \text{var} [X^{(m)}]$ に対する最小二乗法によって得られた直線の傾きから時系列データが自己相似性をもつかどうかを確認することができる.

式 (3.2) から集積過程の分散を用いて短記憶性と長記憶の性質を整理する. 短記憶性をもつ過程は以下の性質を満たす.

- $m \rightarrow \infty$ のとき $\text{var} [X^{(m)}] \approx \sigma^2 m^{-1}$

また、長記憶をもつ過程は

- $m \rightarrow \infty$ のとき $\text{var} [X^{(m)}] \approx \sigma^2 m^{-\beta}$ $0 < \beta < 1$

となる性質をもっている.

3.4.3 スペクトル解析

第 2.3 節に述べたように、スペクトル密度関数は

$$S(f) = \lim_{T \rightarrow \infty} E \left[\frac{1}{T} |X^T(f)|^2 \right]$$

で与えられるが、一般に有限個のデータしか得られないために、この定義をそのまま使うことはできない。有限個のデータからスペクトル密度関数を得るために次の2つの方法を組み合わせる。

1つは、有限個のデータから Fast Fourier Transform (FFT: 高速 Fourier 変換) を使ってピリオドグラム

$$S^T(f) = \frac{1}{T} |X^T(f)|^2$$

を得ることができるが、これ自身は $T \rightarrow \infty$ で $S(f)$ に平均収束せず、従って、かなり長いデータをもってきても、ゆらぎが存在する。しかしながら、任意の周波数区間で積分したものは、 $T \rightarrow \infty$ で収束し、

$$\lim_{T \rightarrow \infty} \int_{f_1}^{f_2} S^T(f) df = \int_{f_1}^{f_2} S(f) df$$

が成り立つ [64]。この性質を使って、ピリオドグラムをある周波数区間で平均したものでスペクトル密度関数を推定することができる。

もう1つは、無限個のデータの Fourier 変換 $X(f)$ が存在すると仮定したとき、有限区間の Fourier 変換 $X^T(f)$ は $X(f)$ に歪みの加わった Fourier 変換となっていることを利用する。この歪みをできるだけ軽減する方法として、データウィンドウを時系列データに掛けてから Fourier 変換をする方法が提案されている [12, 37]。これら2つの方法を組み合わせることにより、次の手順でスペクトル密度関数 $S(f)$ を計算する。

データ数の決定

通常の FFT のプログラムでは、データ数 N が 2^p (p : 整数) になっているものを選ぶ。今回はデータ数が $2^p \leq N < 2^{p+1}$ の関係を満たす 2^p までデータ数を削った。

データウィンドウ

本論文ではデータウィンドウとして、以下のようなものを用いた。

$$W(k) = \begin{cases} 0 & (k \leq 0) \\ \frac{1 + \cos\left(\frac{10}{N}k - 1\right)\pi}{2} & (0 < k \leq \lfloor N/10 \rfloor) \\ 1 & (\lfloor N/10 \rfloor + 1 < k < \lfloor 9N/10 \rfloor) \\ \frac{1 + \cos\left(\frac{10}{N}k - 9\right)\pi}{2} & (\lfloor 9N/10 \rfloor + 1 \leq k \leq N) \\ 0 & (k > N) \end{cases}$$

このデータウィンドウは Bingham ら [8] によって提案された方法である。図 3.8 は、この

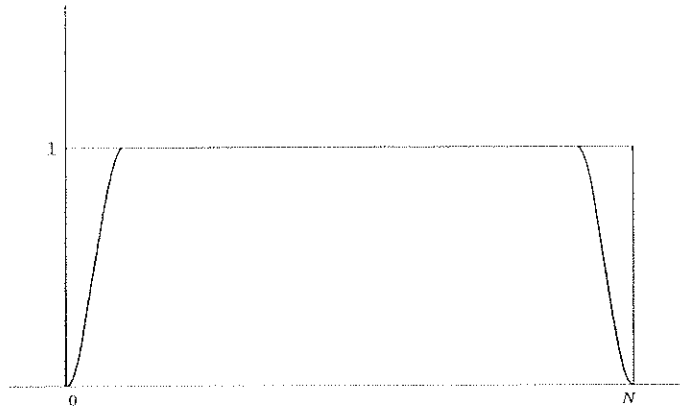


図 3.8: データウィンドウ.

データウィンドウを図示したものである. これを掛けることによってもとのデータの変動の強さは, Parseval の公式により

$$\begin{aligned}
 \frac{1}{\mu} &= \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{k=1}^N W^2(k) \\
 &= \frac{1}{T} \int_0^T W_c^2(t) dt \\
 &= \frac{1}{T} \int_0^{T/10} \left\{ \frac{1 + \cos(\frac{10x\pi}{T})}{2} \right\}^2 dx + \frac{4}{5} + \frac{1}{T} \int_{9T/10}^T \left\{ 1 + \cos(\frac{10x\pi}{T}) \right\}^2 dx \\
 &= \frac{1}{T} \int_0^{2T/10} \left\{ 1 + \cos(\frac{10x\pi}{T}) \right\}^2 dx + \frac{4}{5} \\
 &= \frac{1}{4T} \int_0^{2T/10} \left\{ 1 + \cos^2 \frac{10\pi x}{T} + 2 \cos \frac{10\pi x}{T} \right\} dx + \frac{4}{5} = 0.875
 \end{aligned}$$

倍に押えられるので, スペクトル密度関数の推定値に $\mu = 1/0.875$ を掛けておかななくてはならない. ここで, $W_c(t)$ は上の $W(k)$ の k を $t \in \mathbb{R}$, N を $T \in \mathbb{R}$ と置き換えたものである.

FFT と平滑化

FFT により得られた, $X^T(f)$ の実数部と虚数部よりピリオドグラム $S(f)$ を得る. 上述のデータウィンドウを用いているので, μ を掛ける. 最後に, ピリオドグラムを周波数分割し, その各々の区間で平均を取る. グラフには, 横軸に $\log_{10} f$ を取り, 縦軸に $\log_{10} S(f)$ を取る. 従って, $\log_{10} f$ が負の方向へ大きくなったときの $\log_{10} S(f)$ の傾を見ることにより, 短記憶性, 長記憶性, 及び非定常性が分かる.

理論モデルの解析例

図 3.9 に短記憶過程である 10 次の autoregressive 過程 (AR(10)) と典型的な自己相似過程の増分過程である自己相似パラメタ H が 0.8 の FGN のスペクトル解析を示す. 左側は AR(10), 右側は FGN をスペクトル解析したものである. AR(10) は $f \rightarrow 0$ のとき $S(f)$ が x 軸と水平になっていることが分かる. これは短記憶性を表す. 逆に FGN の場合は $f \rightarrow 0$ のとき $1/2$ と 1 の間に収束することが見てとれるため, 長記憶性をもつことが分かる.

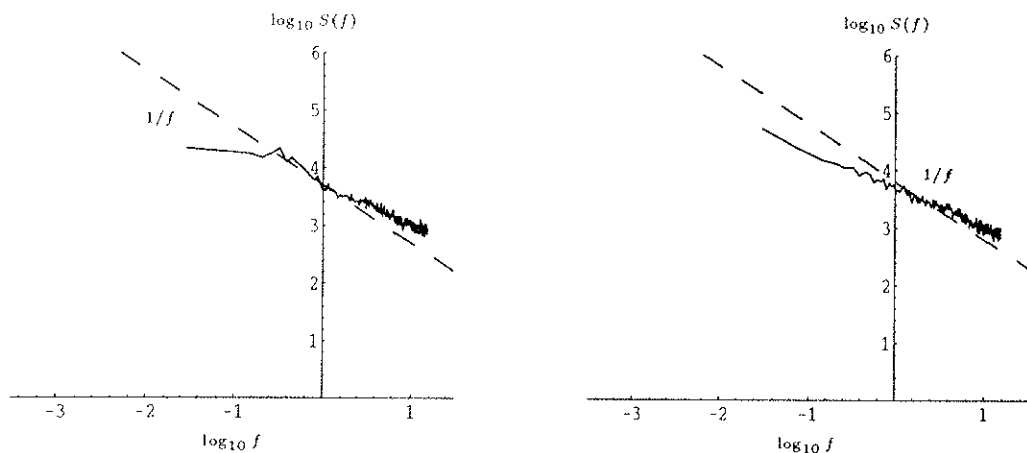


図 3.9: スペクトル解析: AR(10) (左) と FGN (右).

3.5 測定データ

本論文では, MPEG2 によって VBR 符号化された実測データを用いる. VBR で符号化する際, 通常は画質をある程度一定に保ちつつ, ビット数に制限を加える. こうすることによって, 圧縮したデータを通信網上へ伝送するとき, オーバーフローを起こさないようにする. しかし, 今回用いたデータはビット数の制限を一切与えていない, つまり, 画質を完全に一定に保ったデータである. これは, 画像データを伝送するときにある程度の画質を保ちながら, ビット数の上限を推定することを可能にする. 本論文では, 圧縮動画像データを ATM 網上に流すことを想定した. そのため, フレーム番号ごとのビット数という形で与えられたこのデータを ATM のセル数に変換し, その時系列データが自己相似性をもっているかどうかを解析した.

今回データとして用いたものは, 郵政省通信総合研究所所有の資料映像の LD 3 本 (便宜的にこれを LD1-LD3 と呼ぶ) と LD1 をソフト符号化させることによって MPEG2 のパラメータを変えて符号化したもの (LD1S と呼ぶ), 更に NEC C&C メディア研究所より提供された 10 時間の放送画像 (B4 と呼ぶ) とその一部, 及びそれから約 20 分間分を抜粋したもの (B4S と呼ぶ) である. 更に 3 本の LD を適当に組み合わせて, 次のようなデータ列を

表 3.1: 各画像データの詳細.

	時間	M	N	量子化パラメタ
LD1	約 20 分	3	15	6
LD1S	約 20 分	2	10	6
LD2	約 20 分	3	15	6
LD3	約 20 分	3	15	6
B4S	約 20 分	1	15	8
B4	約 10 時間	1	15	8
直列	約 1 時間	3	15	6
並列	約 20 分	3	15	6
混合	約 40 分	3	15	6

作った.

1. LD1, LD2, LD3 を繋げて 1 つの時系列データとする (直列データ).
2. LD1, LD2, LD3 を重ね合わせて (フレームごとの値を足し合わせて) 1 つの時系列データとする (並列データ).
3. LD1, LD2 を繋げ, その時系列データの上に LD 3 を重ね合わせて 1 つの時系列データとする (混合データ).

上のように考えたのは次のような場合を想定したからである. 直列データは単一ソースとしてある一定時間以上画像データを網に流した場合に, 並列データは同時刻に別のデータを多重化して網に流した場合に, 混合データはあるソースのデータが流れているときに別のソースが割り込んできた場合に, それぞれ対応する.

これらをまとめたものが表 3.1 である.

表中の M とは, MPEG で符号化される際, I フレームまたは P フレームのいずれかが出てくる間隔で N とは, I フレームが出てくる間隔である. ソフト符号化したデータ LD1S の画像のソースは LD1 と同じものであるが, 符号化の際のパラメタを変えた. 通常 MPEG で符号化される場合, M は 3 で N は 15 であるが, 今回ソフトエンコードしたデータはこの値を変え, $M = 2$, $N = 10$ とした. 今回のデータは量子化ステップ数を 6 もしくは 8 として VBR 符号化したものである. 量子化ステップ数とは, 符号化データをデジタル量で表現するときの精度である. この値が大きいときは, 解像度が粗くなるが情報量は少なく抑えることができ, 動きのよい映像が得られる. 一方, 小さくすると解像度が細くなる反面, 情報量が増えコマ送りの映像になってしまう. またスペクトル解析における平滑化の際の分割数は B4 のみが 64 で, それ以外のデータでは 8 としている. 図 3.9 に示した理論モデルにおけるは分割数は 64 である.

図 3.10 のグラフは今回用いた 3 種類の LD データを順番に流したときのセル数 (直列

データ) を 10 秒, 1 秒, 0.5 秒 (GOP 単位), 1/30 秒 (フレーム単位) ごとに数え上げたものである。また、右側の図は直列データと同じ平均をもつ Poisson 乱数を発生させ、それを同じ単位時間で描いたものである。また、それぞれの LD の自己相関係数を $k = 1$ から 300 まで計算したものを図 3.11 – 図 3.13 に示す。Poisson 乱数を発生させたものは、単位時間が大きくなるに従い発生するセル数が増えなくなるのに対し、VBR データはどの単位時間に対しても激しく変動していることが分かる。

本論文で考えるような自己相似性をもつトラヒックは、異なる単位時間 (10 秒, 1 秒, 0.1 秒等) に現れる "burstiness" によって特徴づけることができる。ここで、"burstiness" とは、長時間激しい変動のないデータの上に突発的に現れる先の尖ったようなデータ (状態) をいう [55, 56]。このような特性をもつトラヒックは、従来の電話のトラヒックやパケット交換網のトラヒックに対して考えられているモデル (Poisson, Batch Poisson, Markovian Arrival Process (MAP), Markov Modulated Poisson Process (MMPP) 等) とは異なるものである。

3.6 データ解析と考察

本論文では直列データ、並列データ、及び混合データに関してはフレーム単位並びに GOP 単位で R/S 解析、分散解析、スペクトル解析を行った。それ以外のデータについてはフレーム単位と GOP 単位ごとにスペクトル解析のみを行った。

3.6.1 R/S 解析と分散解析の結果

各 R/S 統計量と分散 $\text{var}[X^{(m)}]$ の値をプロットしたグラフが図 3.14 – 図 3.19 である。混合データにおいて LD3 の開始時点は乱数によってずらした。また比較のために、直列データと同じ平均の Poisson 乱数を発生させ、その乱数で同じ解析を行なったのが図 3.20 – 図 3.21 である。各グラフは、各ページの上側が R/S 解析、下側が分散解析である。そして、左側がフレーム単位 (1/30 秒間隔)、右側が GOP 単位 (1/2 秒間隔) の図である。これによって求められた Hurst パラメタは表 3.2 のとおりである。 R/S 解析と分散解析の結果、推定された自己相似パラメタ H を表 3.2 に示す。

R/S 解析の結果 (図 3.14, 図 3.16, 図 3.18) からフレーム単位においても GOP 単位においても推定された自己相似パラメタ H の値は 1/2 と 1 の間にある。このことから、MPEG2 によって画像圧縮された VBR トラヒックは長記憶性をもつことが確認された。フレーム単位ごとの自己相似パラメタ H の値がほぼ 1 になっているのは、MPEG2 の I フレームが周期的に現れることが原因と思われる。このことは GOP 単位で求められた自己相似パラメタ H の推定値と比較することによって確認できる。

分散解析 (図 3.15, 図 3.17, 図 3.19) においても、推定された Hurst パラメタの値は R/S 解析の推定値より小さく推定されているが、長記憶性を示している。分散解析の結果が R/S 解析の結果に比べ、過小推定されているのは、分割数 m の範囲の取り方に依存すると思われる。

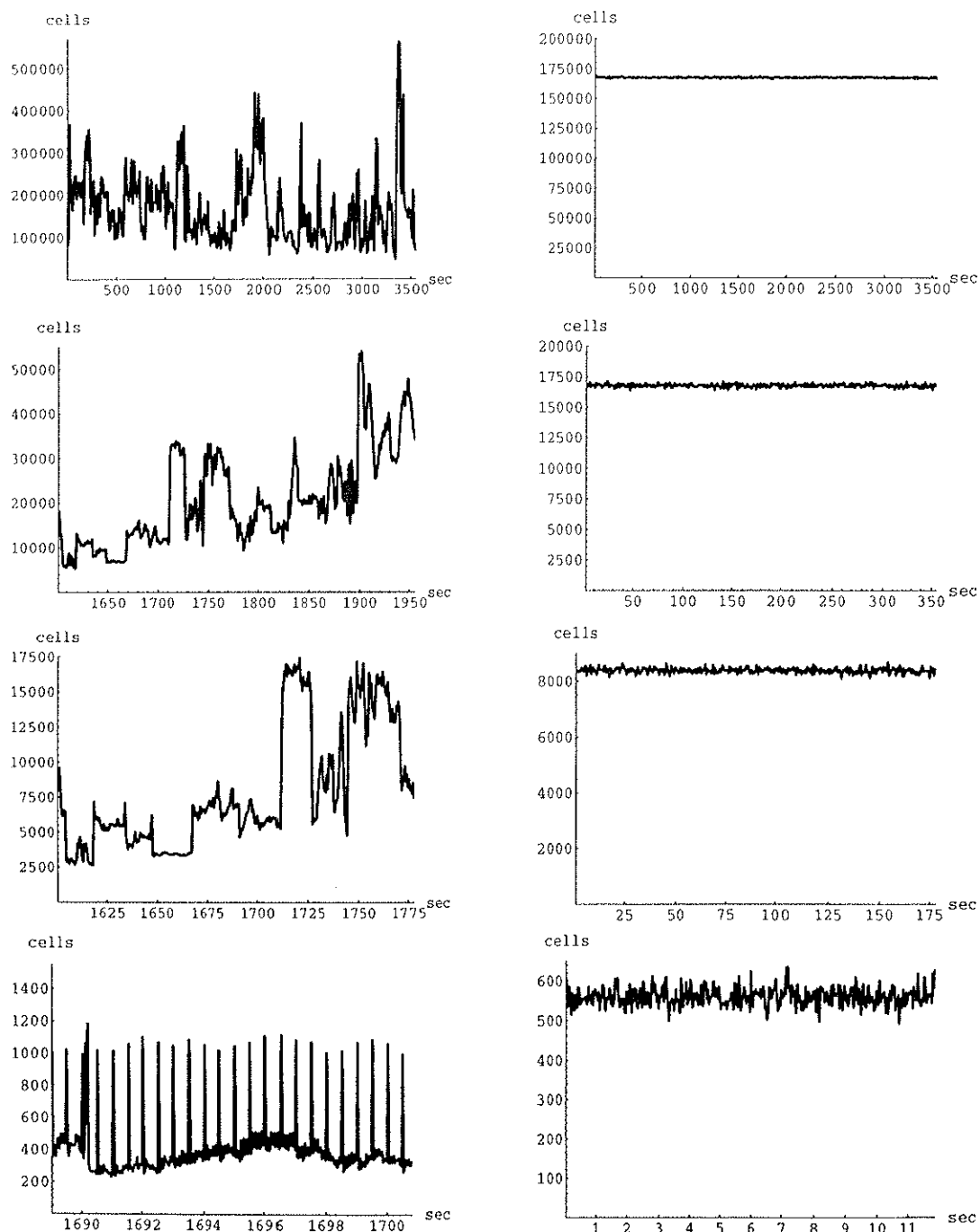


図 3.10: 単位時間 (上から 10 秒, 1 秒, 0.5 秒 (GOP), 1/30 秒 (フレーム)) あたりのセル数 (左) 画像データ, (右) 画像データと同じ平均をもつ Poisson 乱数.

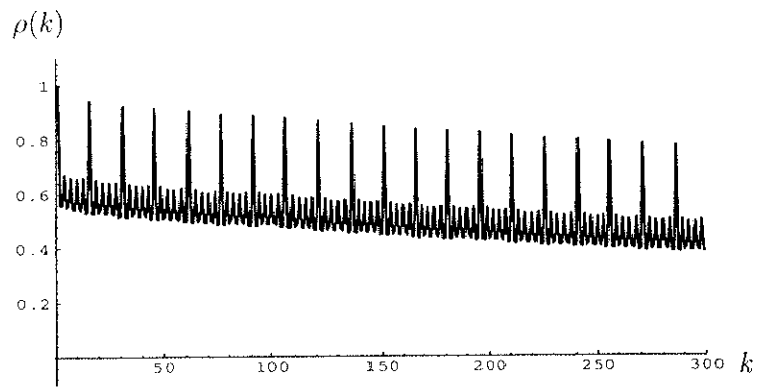
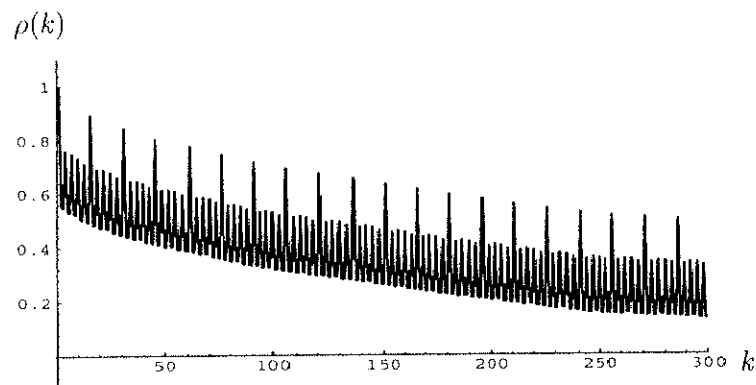
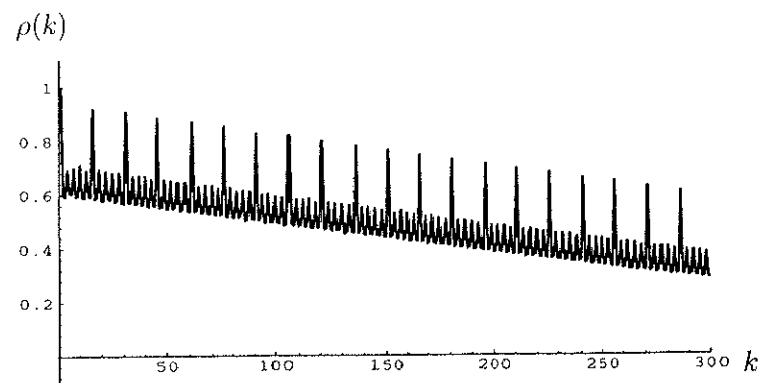
図 3.11: LD1 の自己相関係数 ($k = 1 \sim 300$).図 3.12: LD2 の自己相関係数 ($k = 1 \sim 300$).図 3.13: LD3 の自己相関係数 ($k = 1 \sim 300$).

表 3.2: R/S 解析と分散解析による自己相似パラメタ H の推定結果.

時系列データ	R/S 解析		分散解析	
	フレーム単位	GOP 単位	フレーム単位	GOP 単位
直列データ	0.970	0.868	0.836	0.783
並列データ	0.983	0.897	0.840	0.784
混合データ	0.995	0.887	0.806	0.790
Poisson 乱数	0.513	0.537	0.465	0.450

る. 分割数 m が小さいときは, 各集積過程 $X_i^{(m)}$ に含まれるデータが少ないため, 定常状態に達していないと考えられる. 一方, 分割数 m が大きいときは, 集積過程 $X_i^{(m)}$ 自体の数が少なくなってしまう. そのため, 集積過程の分散は集積過程の値そのものになり, 変動が大きくなる. このような, 振る舞いは画像データに限らず, Poisson 乱数 (図 3.21) や第 4.4 節のシミュレーションにおいても見てとれる. 最小二乗法を行う際の m の範囲の取り方により, R/S 解析より過小推定してしまっていると考えられる.

R/S 解析, 分散解析の結果から, MPEG2 VBR 符号化された画像データが長記憶性をもつことが示された. フレーム単位と GOP 単位を比較してフレーム単位の方が Hurst パラメタの値が高いのは, 極端に情報量の大きな I フレームが周期的に現れるからであると思われる.

また, 参考のために示した Poisson 乱数では Hurst パラメタの値は約 0.5 で, これは短記憶性を示していることになる.

3.6.2 スペクトル解析の結果

スペクトル解析の結果を図 3.22 – 図 3.31 に示す. 各図において左側がフレーム単位, 右側が GOP 単位である. フレームの図におけるスパイクは, 周期的に I フレームが現れることにより, 生じるものであるが, f の大きい領域だけで生じるので関知する必要はない. フレーム単位の図では, フレームの間隔が 1/30 秒であるため最大周波数が 15 Hz となり, $\log_{10} 15 = 1.176$ までプロットされるが, GOP 単位では GOP の間隔が 1/2 秒のため最大周波数は 1 Hz となり, $\log_{10} 1 = 0$ までしかプロットされない. 横軸は $\log_{10} f$ であるから, $f \rightarrow 0$ での振舞いは $\log_{10} f \rightarrow -\infty$ のときに対応する.

3.6.3 非定常性に関する考察

スペクトル解析の結果からは, 測定時間の長いデータを扱うことができた. R/S 解析, 分散解析では扱っていない LD のそれぞれの画像データ (LD1 – LD3) についてまず, スペク

トル解析を行ってみたところ、データの時間が 20 分 (表 3.1) と短いためその特性を捉えることはできなかった。そこで R/S 解析、分散解析を行った。直列データ、並列データ、混合データについてスペクトル解析を行ってみた。その結果を見ても、直列データ以外は明らかに非定常性が現れているように思われる。直列データに関しては、 $1/f$ の直線に漸近的に近づいている。これは、Hurst パラメタが 1 に近づくことと同じで、 R/S 解析や分散解析の結果とも一致する。そして更に長い時間に渡るデータについても調べてみた。それが、10 時間の放送画像データ (B4) である。測定時間が非常に長いため、このデータでは明らかに $1/f$ に漸近的に近づいていることが見てとれる。この性質は、第 2.3 節で説明したように $1/f$ 特性と呼ばれるものである。 $1/f$ 特性が現れるとき、自己相似パラメタ H の推定値は 1 に近づく。 $f \rightarrow 0$ のときスペクトル密度関数が $1/f$ に近づくということは、もとのデータが長記憶と非定常の境界にあることを意味する。

また、Poisson 乱数の場合は短記憶性の特徴を示しているが、理論モデルの解析例でしめした、図 3.9 の AR(10) とは若干違う振る舞いを示している。AR(10) では、 f が大きいとき、 $1/f$ に沿った振る舞いをしているが、Poisson 乱数は f の値に関わらず、ほぼ同じ値の付近を振動している。これは Poisson 乱数が無記憶性をもっているためにこのように振る舞いをするとと思われる。

以上の解析結果より、MPEG2 によって VBR 符号化されたデータは、明らかに長記憶性もっているとはいうことは難しいが、少なくとも短記憶性の特性は示したいないことが分かった。このことから、MPEG2 等のデータを網に流すとき従来の Poisson タイプのモデルを用いて性能評価やシミュレーションを行うのはかなり問題があるように思われる。また、現在のところいくつかのモデルは提唱されている [15, 31, 39, 34, 46, 47, 59] が、Poisson タイプに代わるモデルはまだ研究段階である。今後高速化するであろう情報通信網において画像データ等の長記憶性 (自己相似性) をもつトラヒックのモデルを考えることは重要である。

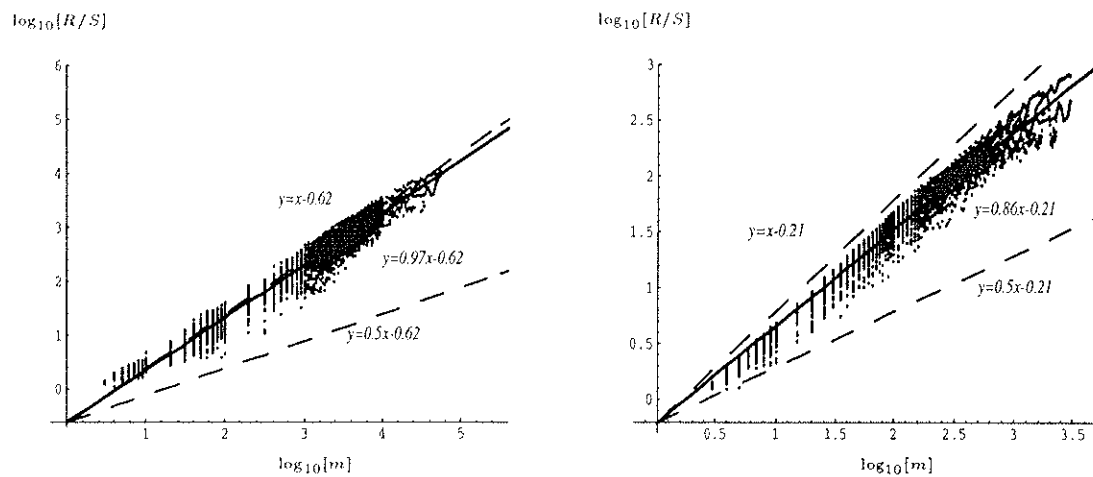


図 3.14: 直列データの R/S 解析. (左): フレーム単位, (右): GOP 単位.

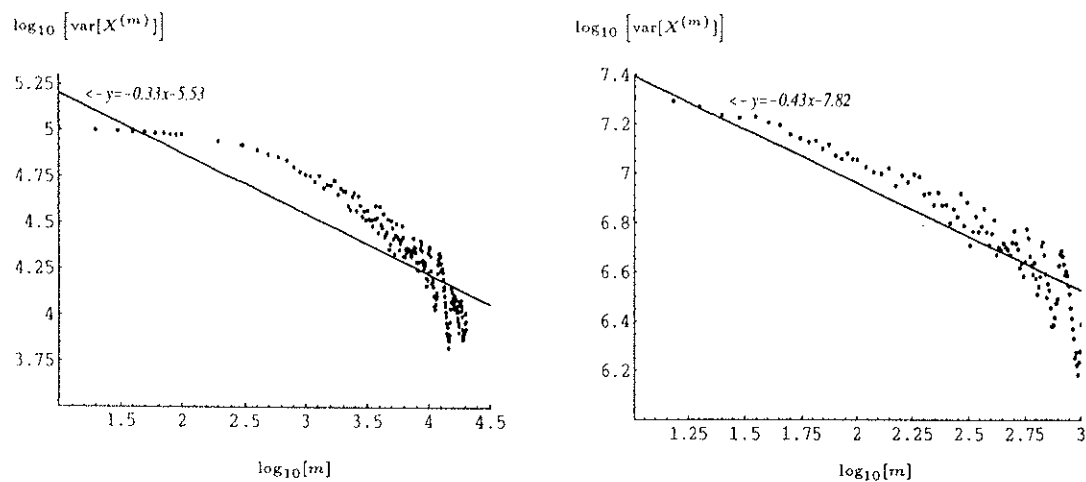


図 3.15: 直列データの分散解析. (左): フレーム単位, (右): GOP 単位.

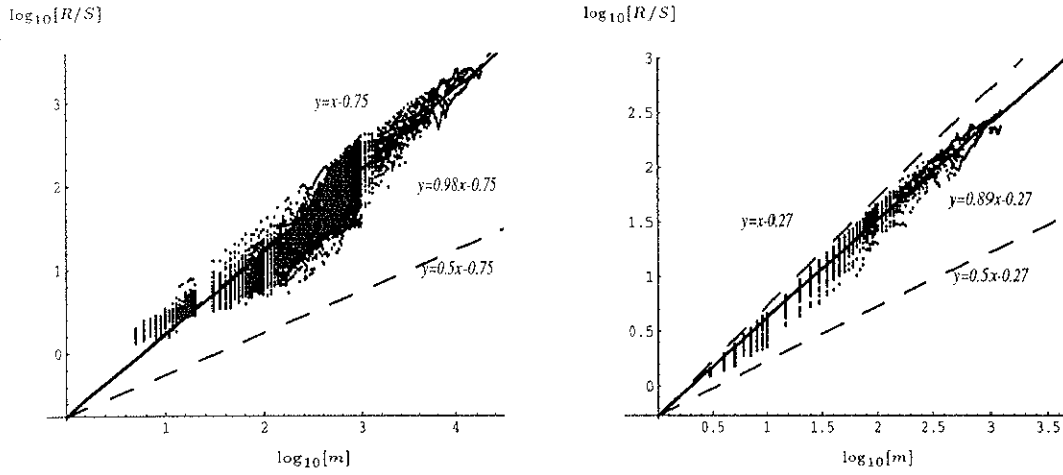


図 3.16: 並列データの R/S 解析. (左): フレーム単位, (右): GOP 単位.

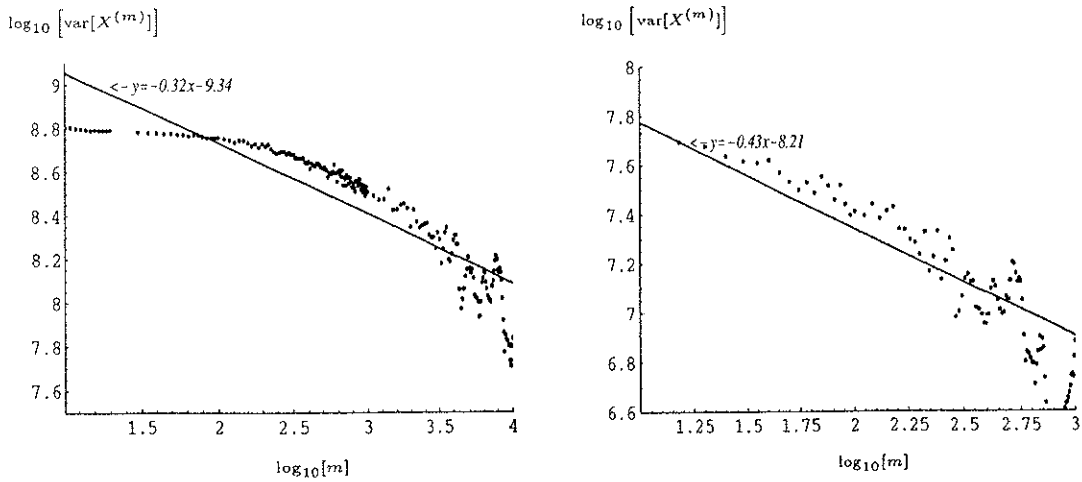


図 3.17: 並列データの分散解析. (左): フレーム単位, (右): GOP 単位.

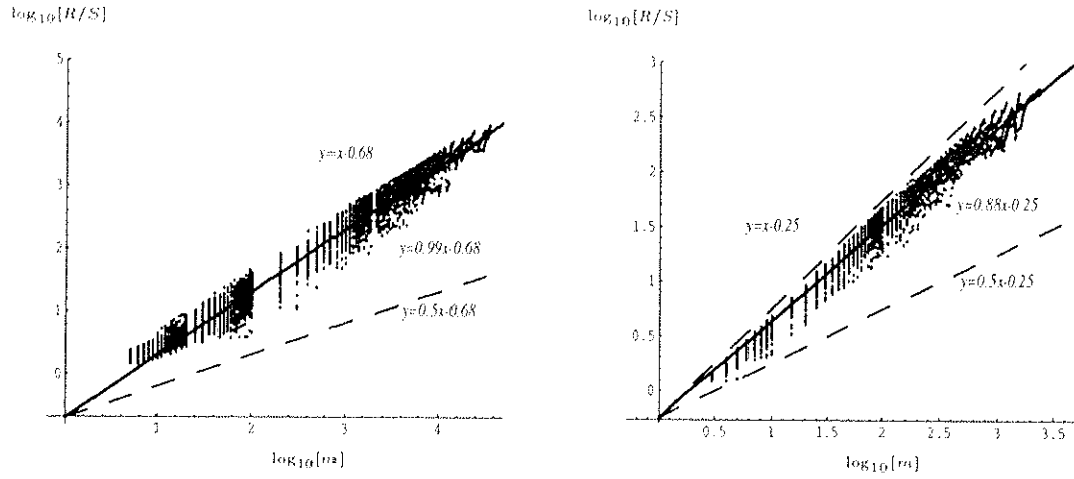


図 3.18: 混合データの R/S 解析. (左): フレーム単位, (右): GOP 単位.

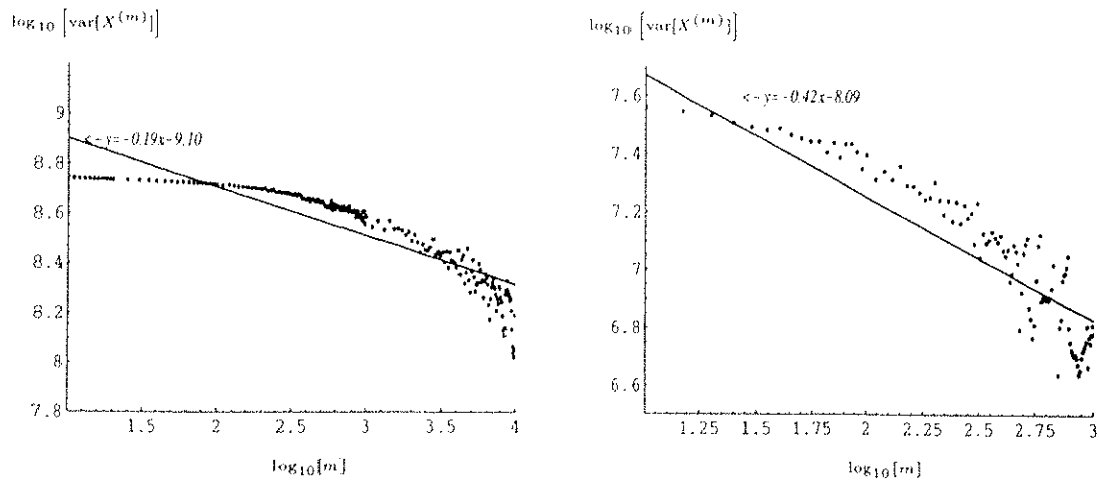


図 3.19: 混合データの分散解析. (左): フレーム単位, (右): GOP 単位.

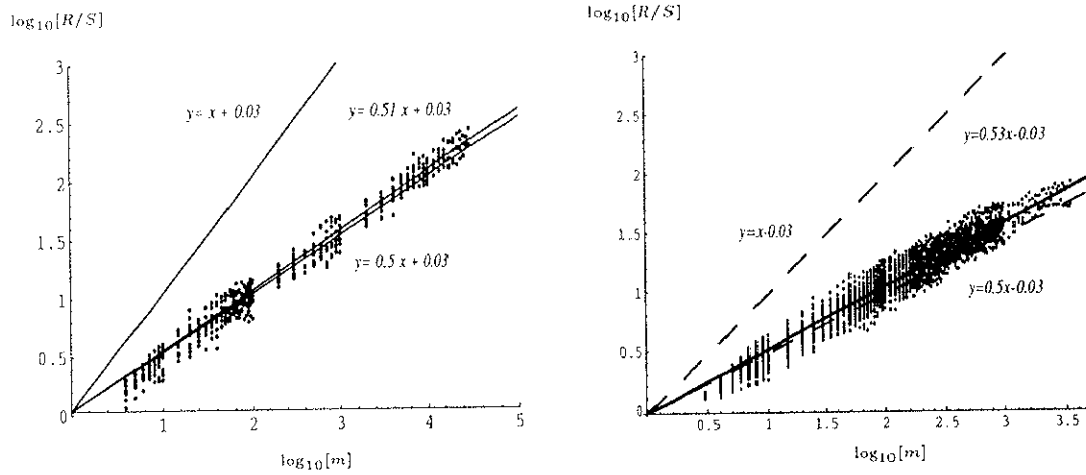


図 3.20: Poisson 乱数の R/S 解析. (左): フレーム単位, (右): GOP 単位.

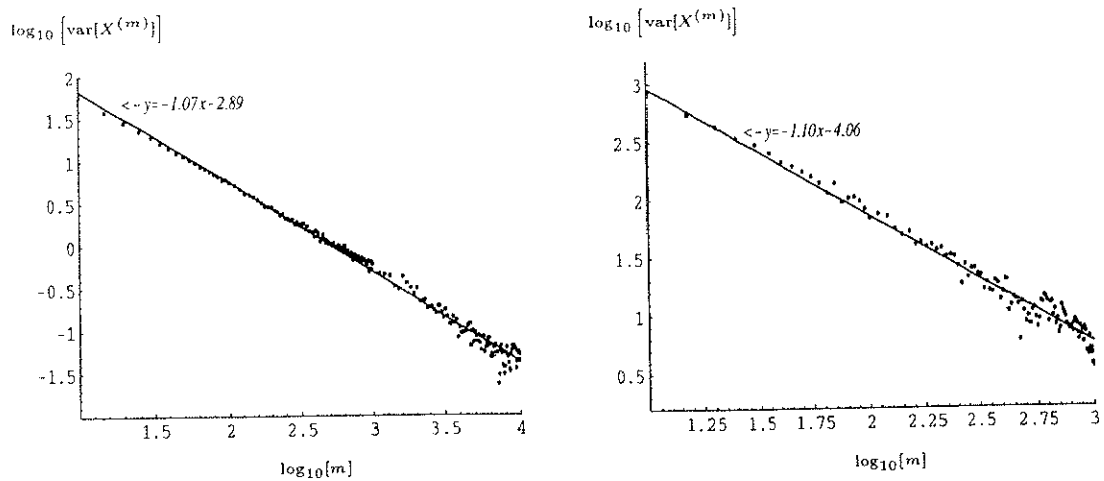


図 3.21: Poisson 乱数の分散解析. (左): フレーム単位, (右): GOP 単位.

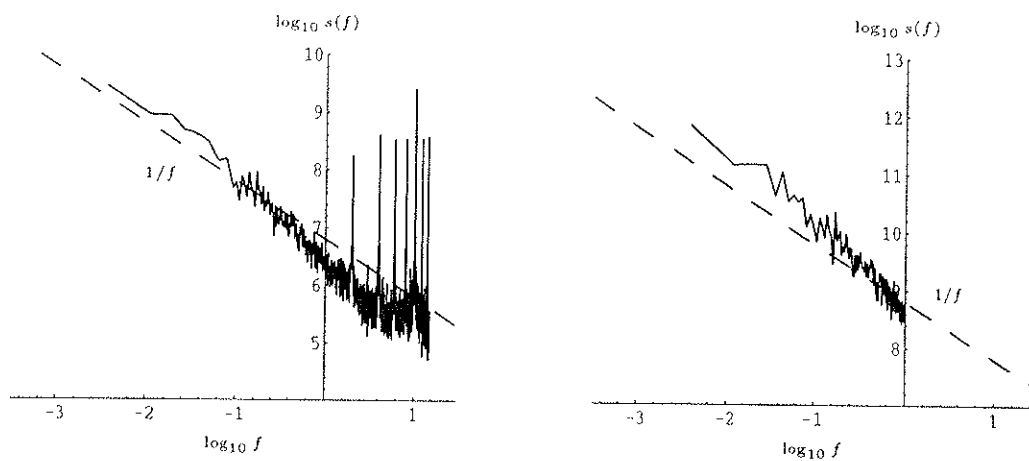


図 3.22: LD1 のスペクトル解析. (左): フレーム単位, (右): GOP 単位.

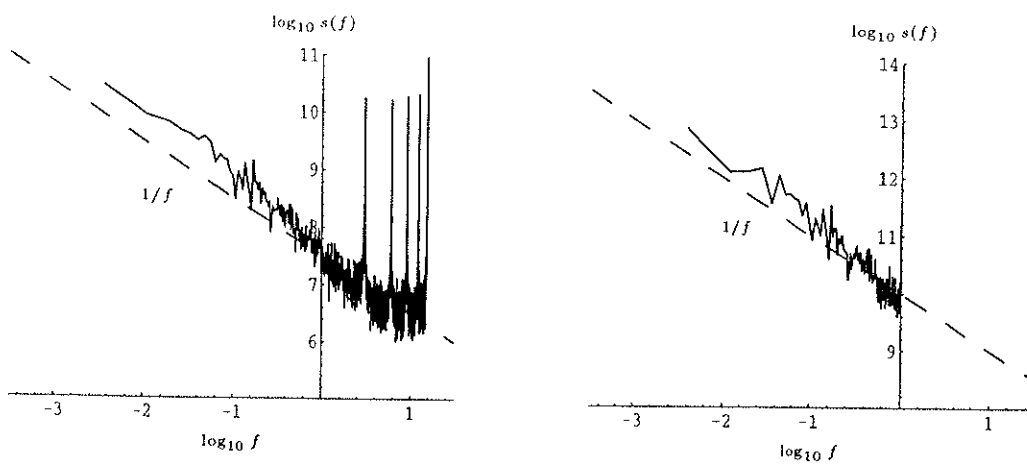


図 3.23: LD1S のスペクトル解析. (左): フレーム単位, (右): GOP 単位.

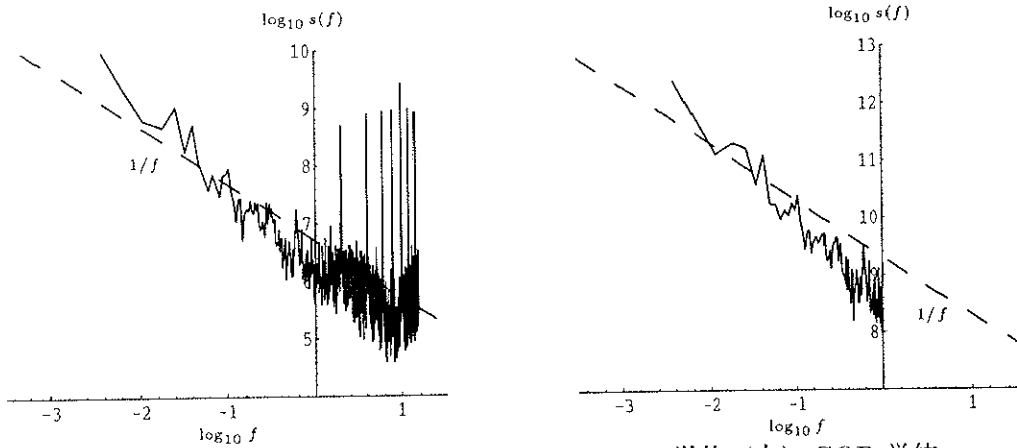


図 3.24: LD2 のスペクトル解析. (左): フレーム単位, (右): GOP 単位.

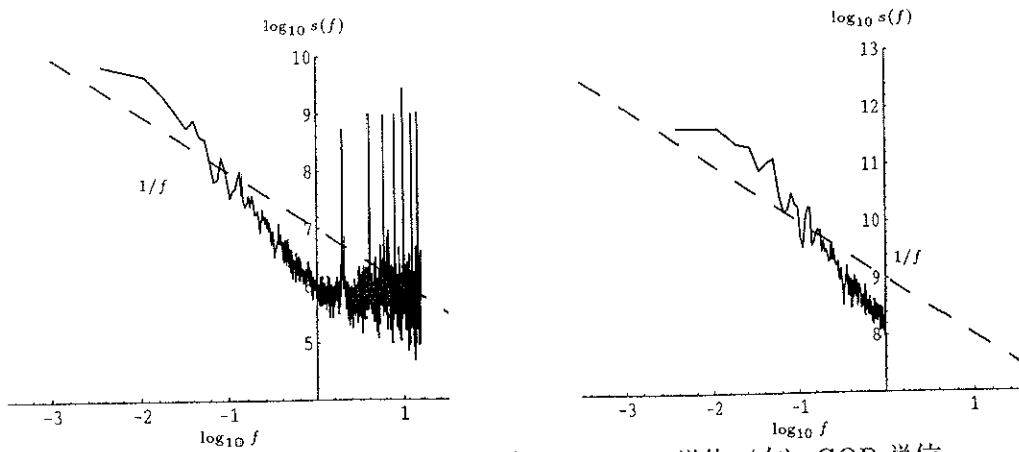


図 3.25: LD3 のスペクトル解析. (左): フレーム単位, (右): GOP 単位.

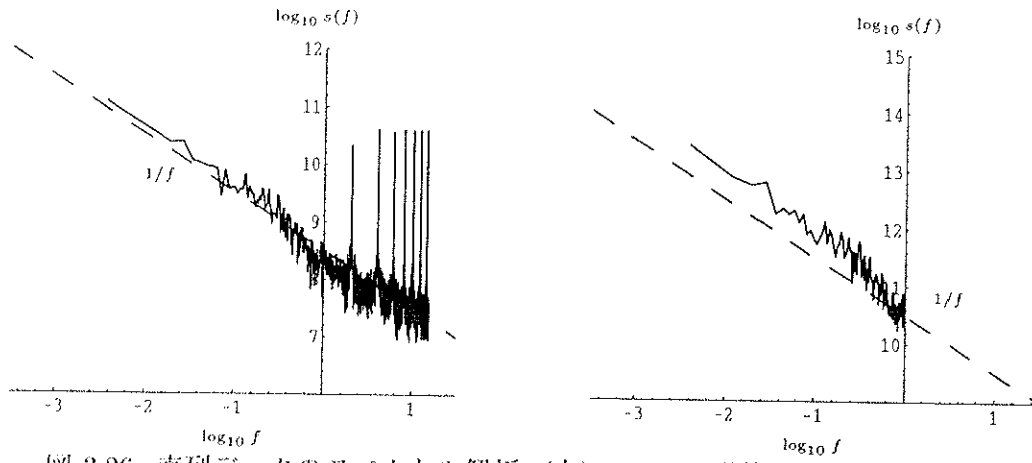


図 3.26: 直列データのスペクトル解析. (左): フレーム単位, (右): GOP 単位.

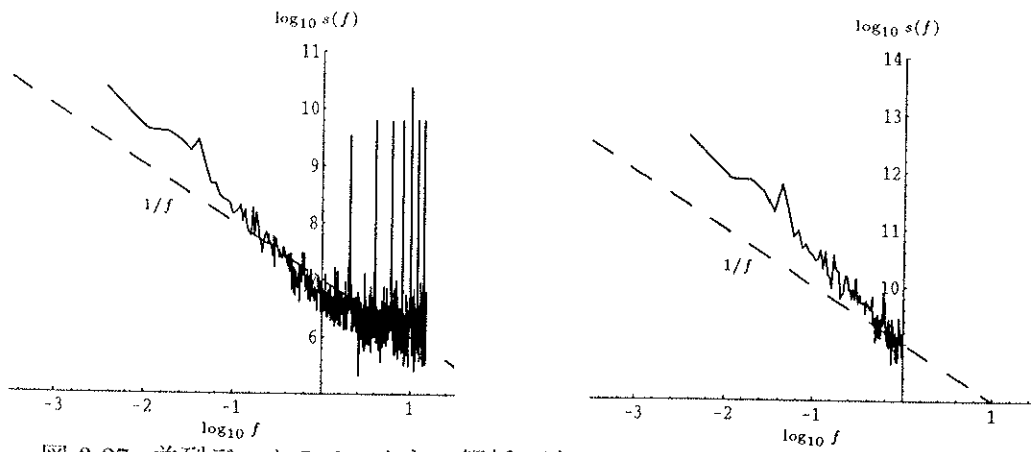


図 3.27: 並列データのスペクトル解析. (左): フレーム単位, (右): GOP 単位.

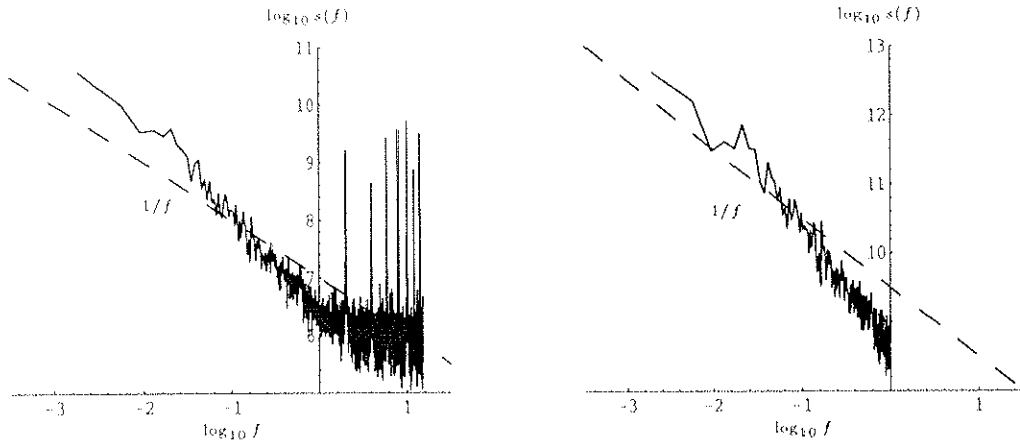


図 3.28: 混合データのスペクトル解析. (左): フレーム単位, (右): GOP 単位.

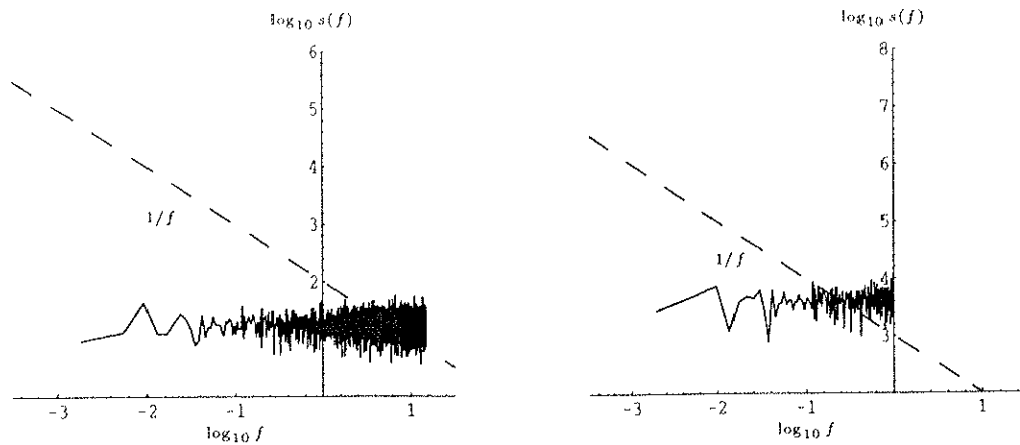


図 3.29: Poisson 乱数のスペクトル解析. (左): フレーム単位, (右): GOP 単位.

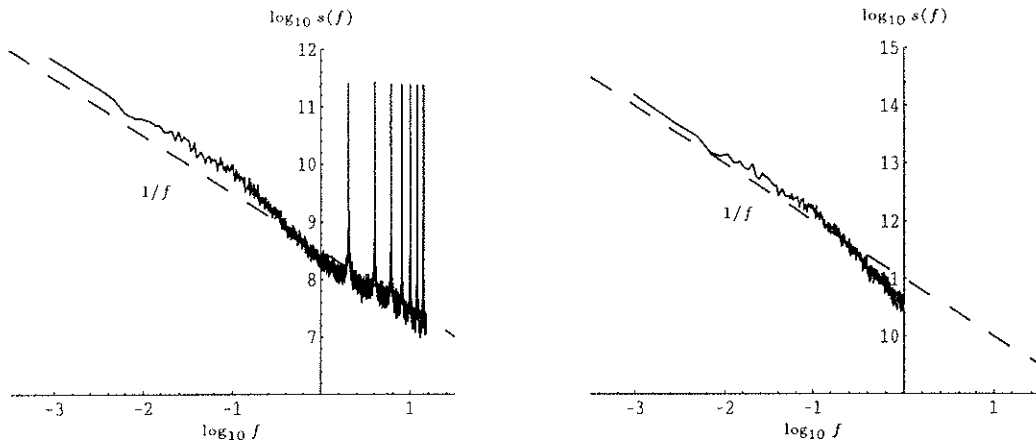


図 3.30: B4 のスペクトル解析. (左): フレーム単位, (右): GOP 単位.

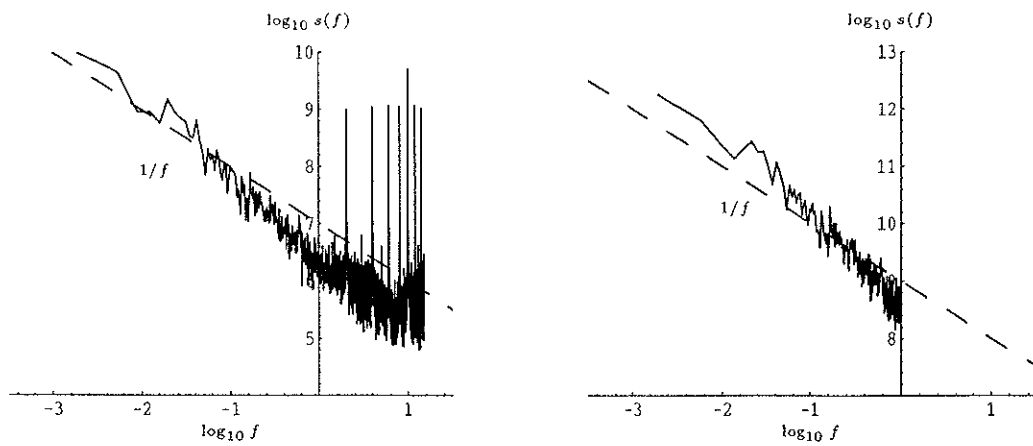


図 3.31: B4S のスペクトル解析. (左): フレーム単位, (右): GOP 単位.