

第 4 章

フラットマルチスケラブル高能率高信頼度映像伝送方式

4.1 まえがき

従来より、オリンピック中継など、厳格な品質の要求される映像の素材伝送や一次分配においては、伝送の高信頼性確保のため、現用系と予備系からなる2系統の回線を用意することが多い(2重化構成)。通常時は現用系が使用され、障害発生時には予備系に切替えて伝送が続行される。しかし、そのような障害は稀にしか起こらず、通常、予備系の伝送容量は無駄になっている。この無駄を軽減し伝送の高能率化をはかることは重要な課題である。

また近年、デジタル圧縮を用いた素材映像伝送が盛んになるにつれ、デジタル伝送画質の監視が運用上重要になりつつあり、画質監視の自動化に対する期待は大きい。

一方、圧縮符号化・復号技術はMPEG-2により一応の確立を見ており、すでに既存の装置が広く使用されている。これらを置換する場合には多大なコストの発生が予想される。

そこで本章では、以上の点に鑑み、映像素材・番組の2重化伝送における高能率化と伝送画質監視を、既存の符号化装置などの構成をなるべく変更せずに実現する方式を求めるところとする。

以下、4.2で複数系統による伝送法の現状と問題点について述べた後、4.3で、2系統の符号化・復号系での符号化雑音の差異について、提案方式で活用する点を中心に説明する。4.4で上記目的を達成するフラットマルチスケラブル映像伝送方式を提案する。本方式は、単一伝送路による伝送よりも高信頼性を有する2重化伝送に対し、高能率性を付与する。すなわち、同一伝送速度のもとでの高画質化あるいは同レベルの画質をより少ない伝送速度で達成することができる。4.5で計算機シミュレーション実験によりその有効性を示す。4.6でまとめるとともに素材伝送以外への応用についても言及する。

4.2 複数系統による伝送法の現状と問題点

2重化伝送は、複数のストリームに分けて信号を伝送する方式の1形態ということができる。そこでまず以下に、本検討の観点から、複数系統を用いた伝送法の現状と問題点をまとめる。

(1) 完全冗長2重化方式

完全冗長2重化方式(図4.1)では、前述の通り、現用系/予備系どちらの系統からでも、映像を受信することができる。これは、専用線をベースとした素材伝送/一次分配においてしばしば使用される。しかし前述の通り、両系統とも正常の時間が大半であるため、通常は片系統分の伝送容量が全く無駄となっているという問題がある。

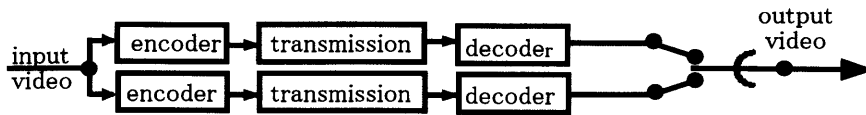


図 4.1: 完全冗長 2 重化系

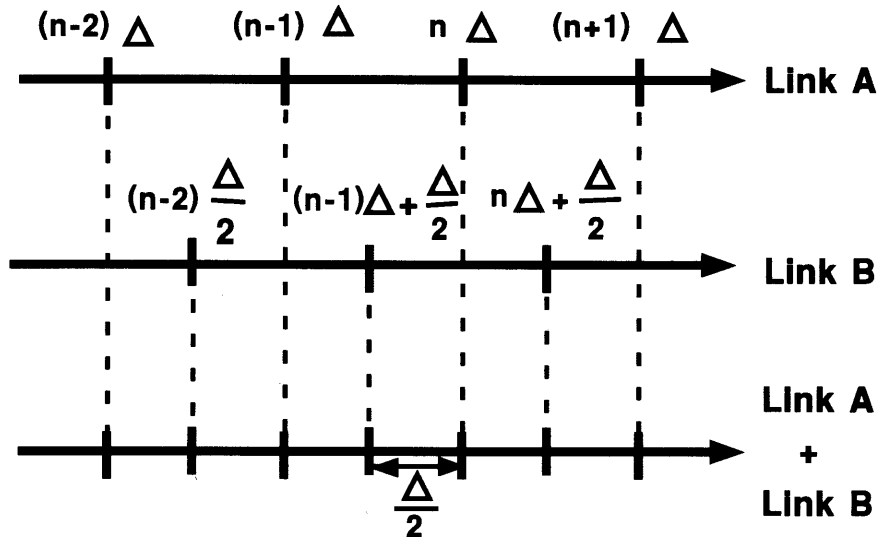


図 4.2: 相補的量子化

(2) MPEG-2 などのスケーラブルプロファイル [1]

近年多く提案されている，マルチキャストなどの応用を想定したスケーラブル符号化は，上位／下位階層といういわば主・従の関係に基づく符号化法であるといえる．主である低階層ストリームから，低解像度の画像を復元できるものの，従である補強ストリームだけでは，画像復元ができないという制限がある．従って，仮に低階層ストリームリンクに障害が発生した場合は，送信側に通知して低階層ストリームと補強ストリームのリンクを入れ換えさせるなど，双方向通信を前提とした追加の処理が必要となる．

従って，映像伝送の二重化に応用しようとする時，既存の伝送構成の大幅な変更を必要とする．

(3) 相補的量子化による複数系統伝送

量子化代表値を，複数系統（ N 系統あるとする）で相補的な入れ子型にすることにより，任意の 1 系統から復号できると同時に， N 系統揃えばステップサイズを等価的に $1/N$ にできる，例えば，2 系統で相補的な量子化器を使用することにより，両系統揃った場合のステップサイズは等価的に半分となる（図 4.2）．すなわち 6 dB の SN 向上が見込める．

しかしながらこの方法は，局所復号画像を必要とする予測符号化に適用すると，受信側で片系統しか利用できない場合，予測のミスマッチが生じ，劣化が大きくなる．従って，動き補償予測を行なう MPEG-2 などの動画像符号化には適さないといえる．

(4) Multiple description coding

この方式は，(3)と同様な考えに基づき，符号化方式自体の工夫により，片系統だけからの復号を可能とするとともに，両系統揃えばより高画質な画像が復号できるという方式である．これには，(3)と同様，量子化器の工夫により行なう方法 [2] や，2 系統での DCT 係数の組を直交変換することにより行なう方法 [3] などがある．しかし，前者は (3)と同様，フレーム間予測符号化には使用できない．また後者は，[3]において動き補償予測符号化に適

用されているが、既存の符号化／復号装置を大きく変更する必要が生じ、本検討の目的にそぐわなくなる。

一方、受信画像の品質自動測定（客観品質評価）に関しては、従来、原画との比較に基づく場合と、受信画像のみから行なう場合がある。原画との比較に基づく方式は、既にいくつか提案されている [4]。しかしながら、実際の伝送中に受信側で客観品質評価をするためには、原画なしで実行する必要がある、いまだ性能的に十分なものは提案されていない。

以上より、本検討の目的、すなわち、

- 複数系統それぞれから映像を復元可能
- 複数系統の合成により受信映像の高画質化が可能
- 伝送画質監視の実現
- なるべく既存の装置構成を変更しない

を満足するためには、上記の方式では、どれも不足であるといえる。

4.3 各系統の符号化ノイズ

ここで、各系統では独立に符号化・復号がなされるという前提に立つ。これは、コーデックを含めた2重化系では、通常成立する条件である。この時の画質劣化（符号化ノイズ）について考察する。

原信号を x 、2系統の復号信号を x_1, x_2 、雑音を d_1, d_2 ($x_1 = x + d_1, x_2 = x + d_2$)、 n_1, n_2 をそれぞれ d_1, d_2 の信号分散とする。また $\rho_{d_1 d_2}$ を d_1 と d_2 の相関係数とする。受信2映像 x_1, x_2 を平均化すると、その信号の符号化ノイズ d_m は、

$$d_m = \frac{x_1 + x_2}{2} - x = \frac{d_1 + d_2}{2}$$

となり、そのパワー n_m は次のようになる。

$$n_m = \sigma^2 \left(\frac{d_1 + d_2}{2} \right) = \frac{n_1 + n_2 + 2\sqrt{n_1 n_2} \rho_{d_1 d_2}}{4} \quad (4.1)$$

従って、よく知られているように、2系統のノイズが独立かつ電力が等しければ、SNは3 dB向上する。逆に、相関が強まるにつれて、SNの向上度は低下する。また、2系統で雑音電力の差異が大きいと、もともと雑音の小さい系統にとっては平均化は逆効果となる。

ここで、どのような場合に平均化が有効かについて簡単に考察する。式(4.1)と $n_m < n_1, n_2$ より、

$$\frac{1}{3}(n_1 + 2\rho_{d_1 d_2} \sqrt{n_1 n_2}) \leq n_2 \leq 3n_1 - 2\rho_{d_1 d_2} \sqrt{n_1 n_2} \quad (4.2)$$

すなわち、例えば d_1, d_2 が全く無相関の場合は、 $\frac{1}{3}n_1 \leq n_2 \leq 3n_1$ となり、雑音電力が互いに3倍までの大きさの範囲内にはいってれば、平均化の効果があることになる。相関が強まるにつれ、この範囲が狭くなっていくこととなる。

通常の素材伝送の2重化のようにビットレートが等しい場合（すなわち雑音電力が2系統で等しい場合）には、次のようにして2重化系の高効率化が可能となるといえる。

- (1) 符号化装置に入力される前に、映像に何らかの処理を加えることにより、2系統の符号化雑音がなるべく独立に近づくようにする。
- (2) 2系統の復号後、上記前処理に対応する後処理を加える。

(3) 2 系統の映像を平均化することにより、片系統よりも高画質化を実現する。

以上に基づいた方式を 4.4.1 で提案する。

次に、伝送画質の監視に関し、次のように考える。

x の信号分散を s とする。受信信号 x_1 の分散と、2 受信信号の差分信号 $x_1 - x_2$ の分散はそれぞれ

$$\sigma^2(x_1) = \sigma^2(x + d_1) = s + n_1 + 2\sqrt{sn_1}\rho_{xd1} \quad (4.3)$$

$$\sigma^2(x_1 - x_2) = n_1 + n_2 - 2\sqrt{n_1n_2}\rho_{d1d2} \quad (4.4)$$

ここに ρ_{xd1} は x と d_1 の相関係数である。これらより、

$$\frac{\sigma^2(x_1)}{\sigma^2(x_1 - x_2)} = \frac{1 + \frac{n_1}{s} + 2\rho_{xd1}\sqrt{\frac{n_1}{s}}}{1 + \frac{n_1}{n_2} - 2\rho_{d1d2}\sqrt{\frac{n_1}{n_2}}} \cdot \frac{s}{n_2} \quad (4.5)$$

上式において、符号化雑音は原信号に対し十分小さく、それらの相関も十分小さい、すなわち $1 \gg$

$\frac{n_1}{s} + 2\rho_{xd1}\sqrt{\frac{n_1}{s}}$ と仮定し、 $\frac{n_1}{n_2} = c$ とおいて変形すると次式が得られる。

$$10\log \frac{s}{n_2} = \alpha + 10\log \frac{\sigma^2(x_1)}{\sigma^2(x_1 - x_2)} \quad (4.6)$$

ここに $\alpha = 10\log(1 + c - 2\sqrt{c}\rho_{d1d2})$ である。

式 (4.6) の α が、画像の種類によらず、2 重化系の構成毎にほぼ一定とみなせれば、受信側で観測可能な x_1, x_2 だけから、それらの SN が求まることとなる。

なお、仮に $x_1 = x$ (1 系が原画) とした場合には $c = 0$ となり、式 (4.6) は s/n_2 の計算式そのものとなる。すなわち、式 (4.6) は通常の SN 計算式の拡張となっている。

4.4 フラットマルチスケラブル映像伝送方式

以上の考察に基づき、本章ではフラットマルチスケラブル映像伝送方式を提案する。その概念は、次の通りである。すなわち、2 系統のストリームが上下の関係にある通常のスケラブル方式に対し、本方式では、これら複数ストリームをフラットに取り扱う。そして、任意の単一ストリームから元の画像が復元できるとともに、複数ストリームを組み合わせれば、それらの相互補完により、更に高品質な画像を得られるようにする。

本方式の大きな特長として、既存の符号化器/復号器に手を加えることなく、それらの前処理/後処理として実現できることがあげられる。

本方式は、元来高信頼性のある複数系統による映像伝送において、高能率化を達成する「相補的サンプリング法」(4.4.1) と、画質評価を達成する「符号化ノイズ残留相関法」(4.4.2) からなる。

4.4.1 相補的サンプリング法による高能率化

提案方式

本方式は、3 系統以上の複数系統にも適用可能であるが、ここでは簡単のため、2 系統の場合を考える。この 2 重化系の高能率化を達成するため、各系統の復号画像の平均化による符号化ノイズ平滑化効果を利用する。

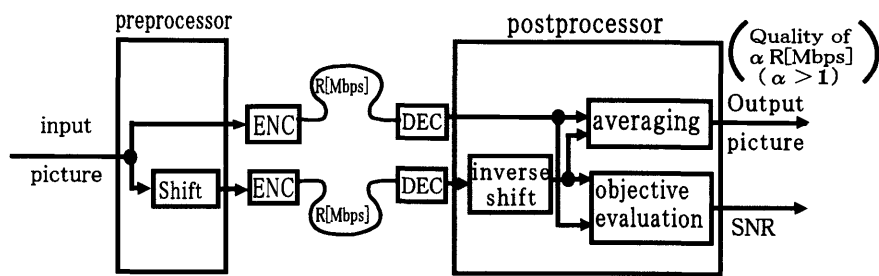


図 4.3: フラットマルチスケラブル映像伝送

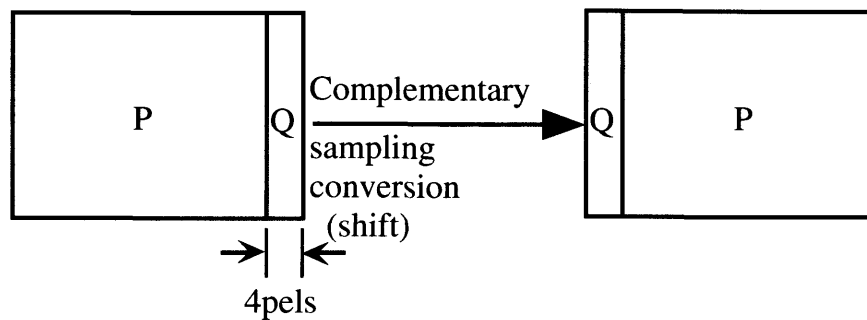


図 4.4: 相補的サンプリング変換の例

両系統のノイズを独立に近づけ、平滑化効果を最大限にするために、次のように考える。

MPEG-2 などのブロックベース符号化では、DCT を 8×8 のブロック毎に適用することが一般的である。そこで、DCT ブロック境界をずらすことができれば、2 系統の符号化ノイズの相関が小さくなることを期待できる。そのためには、片系統の画素位置をずらせば良い。

すなわち、次のような相補的サンプリング法を提案する (図 4.3)。

- (1) 符号化前処理部において、入力画像を 2 系統に分配する際、片系統の画像の画素位置を、サンプリング変換により一定量ずらす。
- (2) 符号化後処理部においては次の処理を行なう。
 - (a) 片系統の画素位置のずれを元に戻す。
 - (b) 両系統の復号画像を平均化する。

(1)の画素ずらしを式で表現すると、次のようになる。元の画像を、簡単のため次のように 1 次元信号で表す。ここに N は水平方向のサンプル数である。

$$x(0), x(1), \dots, x(i), \dots, x(N-2), x(N-1)$$

この時、 n 画素分だけ右にずらした信号とは、次のように表現される。

$$x(N-n), \dots, x(N-1), x(0), \dots, x(N-n-1)$$

すなわち、右端からはみ出す分は、左端に挿入される。

例として図 4.4に、水平方向に 4 画素ずらす場合を示す。この図では、右方向のずらしのため右端からはみ出る Q の部分が、左端に挿入されている。

なお、この画素ずらしにおいては、非整数画素のずらし量も、デジタルフィルタによるサンパ

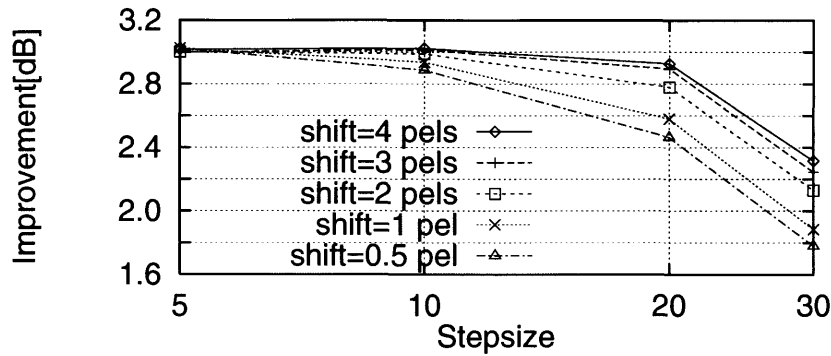


図 4.5: 平均化による SN 比改善. 入力信号の分散は 100.

リング変換を行なうことにより, 実現することができる.¹ その場合でも, 次のように, ずらし/戻しによる画質劣化は十分小さい. すなわち, 11 タップのサンプリング変換フィルタ (係数: [2, -6, 11, -24, 81, 81, -24, 11, -6, 2]/128) を用いた半画素ずらしの場合, 変換/逆変換により, 一般的な自然画像で約 46dB 以上の PSNR が得られた. これは, 符号化劣化に比べると無視できるほどの劣化量である.

確認実験

画素ずらしによるノイズの無相関化を確認するため, 予備実験を行なう. 次のような条件のもと, モンテカルロ法により, 2 系統の平均化信号の SN 向上度と量子化ステップサイズの関係調べた結果を図 4.5 に示す.

- 入力画像: ラプラス分布をなす 1 次元白色信号
- 相補的サンプリングの構成: 片系統を 0.5 ~ 4 画素ずらす
- 符号化: ブロックサイズ 8x1 の 1 次元 DCT, 一様線形量子化

図 4.5 より, 量子化ステップサイズ (DCT 係数を一様線形量子化する際の刻み幅) が十分小さいと, 画素ずらしにより平均化信号の SN が 3 dB 向上していることが確認できる.

また, ステップサイズが大きくなると, 概して SN 向上度は小さくなる. この理由は, 次の通りである. すなわち, 量子化ステップサイズが大きくなるにつれ, 切捨て状態になる係数 (量子化代表値 0) が増加する. その場合は, 量子化ノイズが入力信号に等しくなり, 2 系統でのノイズの相関が増大するためである.

さらに, 図 4.5 より, 4 画素ずらしの場合が最も SN 向上度が大きいといえる. 4 画素は DCT ブロックサイズの半分に相当しており, 最もずれ方が大きい場合である. そのため, ノイズの相関が最も小さくなるものと考えられる.

¹ サンプリング変換フィルタを $f(i)$ と表すと, 元の信号 $x(i)$ をずらした結果の信号 $y(i)$ は

$$y(i) = x(i) \otimes f(i) = \sum_{k=0}^{N-1} x(k) f(i-k)$$

と表される (\otimes は畳み込みを表す). 帯域保存型の内挿補間処理となる.

4.4.2 符号化ノイズ残留相関法による客観画質評価

本節では、2系統の符号化ノイズの差異を利用して、2系統の受信映像だけから、それらと原画とのSN比および視覚特性を考慮したWSNRを推定する方式を提案する。2系統で符号化ノイズ電力が異なる場合も適用可能である。

SN比の推定方式

式(4.6)において、 α が、画像の種類によらず、2重化系の構成毎にほぼ一定とみなせれば、次のようにして、2系統の復号画像 x_1, x_2 だけから、それらと原画とのSN比の推定が可能となる。下記に、 x_2 のSN比を求める方法を示す。 x_1 の場合も同様である。

- (1) 推定式係数の決定. あらかじめ定めた数種類のテスト画像を、当該2重化系の構成で符号化することにより、定数 α を決定する。
- (2) 実運用時には、受信画像から $\sigma^2(x_1)/\sigma^2(x_1 - x_2)$ を測定し、式(4.6)に当てはめることにより s/n_2 を得る。
本方式の有効性を4.5.2で示す。

WSNRの推定方式

原理 本節においては、WSNRとして、文献[5]にある定義を用いる。すなわち、原画と復号画に、空間周波数の振幅特性が式(4.7)で表される2次元フィルタ（視覚特性フィルタ）をかけ、その結果の画像どうして求めたSN比をWSNRとする。

$$H(r) = (0.2 + 0.45r)e^{-0.18r} \quad (4.7)$$

ここに r は半径方向の周波数（単位視角度当たりのサイクル数）である。

前節までの記号を使い、 x, x_1, x_2, d_1, d_2 に本フィルタをかけた信号をそれぞれ $\hat{x}, \hat{x}_1, \hat{x}_2, \hat{d}_1, \hat{d}_2$ とする。 \hat{s} を \hat{x} の信号分散、 \hat{n}_1, \hat{n}_2 を \hat{d}_1, \hat{d}_2 のパワーとする。 \hat{x} と \hat{d}_1 の相関係数を $\rho_{\hat{x}\hat{d}_1}$ 、 \hat{d}_1 と \hat{d}_2 の相関係数を $\rho_{\hat{d}_1\hat{d}_2}$ とする。式(4.3)、式(4.4)と同様に

$$\sigma^2(\hat{x}_1) = \hat{s} + \hat{n}_1 + 2\sqrt{\hat{s}\hat{n}_1}\rho_{\hat{x}\hat{d}_1} \quad (4.8)$$

$$\sigma^2(\hat{x}_1 - \hat{x}_2) = \hat{n}_1 + \hat{n}_2 - 2\sqrt{\hat{n}_1\hat{n}_2}\rho_{\hat{d}_1\hat{d}_2} \quad (4.9)$$

さらに、式(4.5)と同様に、

$$\frac{\sigma^2(\hat{x}_1)}{\sigma^2(\hat{x}_1 - \hat{x}_2)} = \frac{1 + \hat{n}_1/\hat{s} + 2\rho_{\hat{x}\hat{d}_1}\sqrt{\hat{n}_1/\hat{s}}}{1 + \hat{n}_1/\hat{n}_2 - 2\rho_{\hat{d}_1\hat{d}_2}\sqrt{\hat{n}_1/\hat{n}_2}} \cdot \frac{\hat{s}}{\hat{n}_2} \quad (4.10)$$

また同様に符号化雑音は原信号に対し十分小さく、それらの相関も十分小さい、すなわち $1 \gg \hat{n}_1/\hat{s} + 2\rho_{\hat{x}\hat{d}_1}\sqrt{\hat{n}_1/\hat{s}}$ と仮定し、 $\hat{n}_1/\hat{n}_2 = \hat{c}$ とおいて変形すると次式が得られる。

$$10 \log \frac{\hat{s}}{\hat{n}_2} = \hat{\alpha} + 10 \log \frac{\sigma^2(\hat{x}_1)}{\sigma^2(\hat{x}_1 - \hat{x}_2)} \quad (4.11)$$

ここに $\hat{\alpha} = 10 \log(1 + \hat{c} - 2\sqrt{\hat{c}\rho_{\hat{d}_1\hat{d}_2}})$ である。

求めるWSNRは \hat{s}/\hat{n}_2 にほかならない。観測可能な $\frac{\sigma^2(\hat{x}_1)}{\sigma^2(\hat{x}_1 - \hat{x}_2)}$ と、WSNRの関係式が得られたこととなる。

なお、SN比の場合と同様、仮に $x_1 = x$ (1系が原画)とした場合には $\hat{c} = 0$ となり、式(4.11)は \hat{s}/\hat{n}_2 の計算式そのものとなる。すなわち、式(4.11)は通常のWSNR計算式の拡張となっている。

表 4.1: 各画像の 22.5Mbps での PSNR の改善 [dB]

	che.	fla.	gre.	mar.	mob.	soc.
PSNR[dB]	41.2	39.4	37.8	35.6	38.4	36.2
Δ PSNR	2.0	1.9	1.8	1.9	2.0	2.0

提案方式 式 (4.11) において, $\hat{\alpha}$ が, 画像の種類によらず, 2 重化系の構成毎にほぼ一定とみなせれば, 次のようにして, 2 系統の復号画像 x_1, x_2 だけから, それらと原画との WSNR の推定が可能となる. 下記に, x_2 の WSNR を求める方法を示す. x_1 の場合も同様である.

- (1) 推定式係数の決定. あらかじめ定めた数種類のテスト画像を, 当該 2 重化系の構成で符号化し, 式 (4.7) のフィルタを適用することにより, 式 (4.11) の定数 $\hat{\alpha}$ を決定する.
- (2) 実運用時には, 受信画像に式 (4.7) のフィルタをかけ, $\sigma^2(\hat{x}_1)/\sigma^2(\hat{x}_1 - \hat{x}_2)$ を測定し, 式 (4.11) に当てはめることにより \hat{s}/\hat{n}_2 を得る.
この方式が有効であることを 4.5.2 で示す.

4.5 計算機シミュレーション実験

提案するフラットマルチスケラブル映像伝送方式の要素である相補的サンプリング法と符号化ノイズ残留相関法の有効性を, 計算機シミュレーション実験により検証する.

4.5.1 相補的サンプリング法の検証

提案した相補的サンプリング方式を実画像に適用し, その SN 向上度を調べる. 次のような実験条件で行なった.

- 符号化: MPEG-2 TM5
- 片系統の画素ずらし量: 水平方向 4 画素, 垂直方向 0 ライン

また, テスト画像の種類としては, さまざまな絵柄を含むように, ITU-R 標準テスト動画像と ITE 標準 HDTV テスト動画像 (標準 TV サイズに切り出し) から, 計 6 種類 (cheerleaders, flamingoes, green leaves, marching in, mobile & calendar, soccer) を選択した.

図 4.6 に, 片系統の PSNR および 2 系統の平均化後の PSNR とビットレートとの関係を示す. PSNR は 6 画像の平均値を用いる. また, 表 4.1 に, 画像毎の, ビットレート 22.5Mbps での PSNR (平均化後) と平均化による PSNR の向上度を示す.

図 4.6 より, 提案方式により, 5-45Mbps の範囲で, 約 1.2-2.2dB の PSNR 向上が得られることがわかる. また表 4.1 より, この PSNR 向上度の画像による差異は小さく, 例えば 22.5 Mbps の符号化時, 約 2dB の向上となることがわかる.

各ビットレートにおける量子化スケール (量子化ステップサイズに対応) の平均値は, 45Mbps で 5.0, 22.5Mbps で 9.6, 11.2Mbps で 16.8, 5.6Mbps で 28.7 であった. すなわち, 数値実験の図 4.5 と比較すると, 同じ量子化ステップサイズにおける PSNR の向上度は, 実際の MPEG-2 符号化の場合 (図 4.6) の方が, 小さいということになる. この違いの原因としては, 実際の MPEG-2 符号化では, 一様線形量子化ではなくデッドゾーン付きの量子化器を使用していることや, 画素値の整数化による丸め誤差が発生することなどがあげられる.

以上の実験は, 垂直方向のずらしがない場合について行なった. この理由は, 次の実験で示すように, 水平方向に 4 画素ずらししている限り, 垂直方向のずらし量が PSNR 向上に及ぼす影響はほ

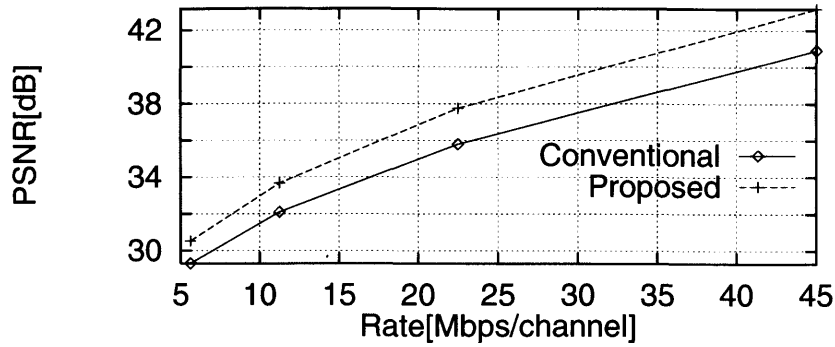


図 4.6: PSNR の改善とビットレート

表 4.2: 垂直方向のずらし量 [frame lines] と PSNR の改善 [dB]. 水平方向のずらし量は 4 画素. (22.5 Mbps)

shift	0	2	4	6	8
Δ PSNR	1.96	2.00	2.04	2.04	2.00

とんど無視できるからである。

すなわち, 水平方向 4 画素ずれのもと, 垂直方向のずらし量を変化させた時の PSNR 向上度 (6 画像平均) の変化を, 表 4.2 に示す. また, 水平方向にはずらさず, 垂直方向のみのずらし量を変化させた時の結果を表 4.3 に示す.

表 4.2 より, すでに水平方向にずらしてある場合, 垂直方向のずらし量を変化させても, PSNR 向上にはほとんど影響がないことがわかる. 一方, 表 4.3 より, 垂直方向のみのずらしによっても, 水平方向のみのずらしにほぼ匹敵する PSNR 向上 (約 1.9dB) が見られる. すなわち, 水平方向あるいは垂直方向のずらしのみで, すでに可能な相関低下は実現されているといえる. なお, 垂直方向の 0 以外のずれ量による変化が少ない一因としては, 次のことが考えられる. すなわち, 特に MPEG-2 の場合, DCT モードとして field DCT と frame DCT があり, 適応的に選択されている. このため, 水平方向と違い, 垂直方向のブロック境界が一定していないことがあげられる.

このように, 水平方向にずらしていれば, 垂直方向のずらし量による影響がほとんど無視できるので, 本提案方式においては, 処理の簡単化のため, 垂直方向のずらしなしで行なうことにする.

次に, 図 4.7 には, 各種レートに相当する PSNR を得るために 2 重化伝送系で必要とされる, 2 系統の合計レートを示す. 本図のデータは, 図 4.6 のデータより, 完全冗長 2 重化系と提案方式によるフラットマルチスケラブル伝送系について算出した. 従来の完全冗長 2 重化系では, あるレート相当の画質を実現するために, その 2 倍のレートを要する. 一方, 提案方式では, それに対し約 25% のレート削減を実現している.

表 4.3: 垂直方向のずらし量 [frame lines] と PSNR の改善 [dB]. 水平方向のずらしはなし. (22.5 Mbps)

shift	0	2	4	6	8
Δ PSNR	0.00	1.76	1.85	1.88	1.76

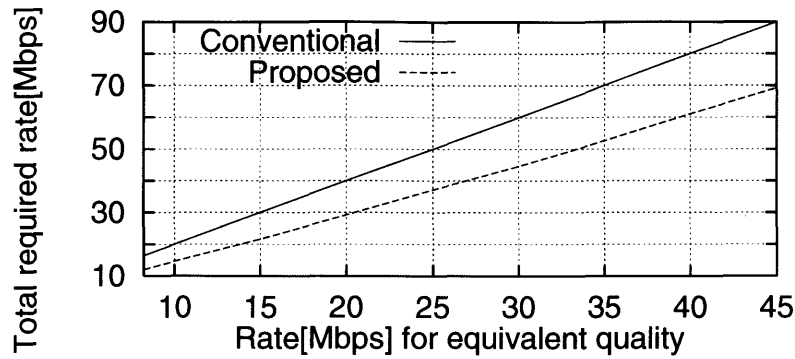


図 4.7: 同じ画質を達成するために要する各 2 重化システムでのビットレート

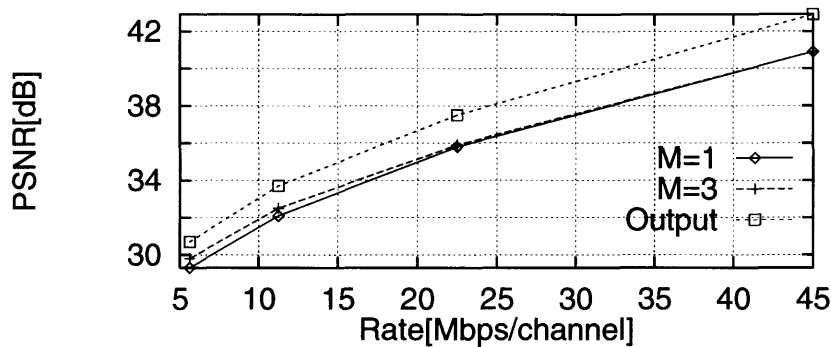


図 4.8: PSNR の改善とビットレート (2 系統で符号化構造が異なる場合)

以上は、通常の映像伝送 2 重化系を想定し、同一のコーデックが 2 系統で使用されていることを前提としている。そこで次に、2 系統で符号化構造が異なる場合について実験する。片系統を B ピクチャなし ($M = 1$)、もう片系統を $M = 3$ で符号化した場合の各系統の PSNR 及び平均化後の PSNR を図 4.8 に示す。この図を見ればわかるように、IBP 構造が異なる 2 重化系の場合でも、合成により平均 0.9-2.0dB の SN 向上が得られることが確認できる。

さらに、参考のため、SNR スケーラビリティとの比較を表 4.4 に示す。前述の通り、SNR スケーラブル方式においては、エンハンスメントレイヤのみからの復号は不可能であるが、提案方式であるフラットマルチスケーラブル方式では、どちらの系統からでも復号可能である。提案方式はそのような長所があるとともに、11Mbps においては、スケーラブル符号化の中では一般にオーバーヘッドが小さいとされる SNR スケーラビリティとほぼ同等の性能を有する。同じ SN の差でも高レートではあまり目立たず、低レートほど目立つことを勘案すると、本方式は、この面からも優れた性能を有するといえる。

表 4.4: 2 系統合成時の PSNR の向上 [dB](a) 提案方式 (b)SNR スケーラブル方式 (ベースレイヤとエンハンスメントレイヤは同一ビットレート)

Bitrate [Mbps/channel]	22.5	11.2
(a)	2.0	1.6
(b)	3.2	1.8

表 4.5: PSNR の改善 [dB](a) 相補的サンプリング変換なし (I ピクチャの時間的位置ずれのみ)
(b) 相補的サンプリング変換あり (I ピクチャの時間的位置ずれあり)

Bitrate[Mbps]	45	22.5	11.2	5.6
(a)	1.4	1.1	0.9	0.7
(b)	2.3	2.0	1.6	1.3

以上により、相補的サンプリング法の有効性を実証できた。

なお、MPEG-2 符号化などにおいては、I B P 構造があるため、相補的サンプリングすなわち画面ずらしを行なわなくても、例えば I ピクチャの時間的位置が 2 系統間でずれていれば、符号化ノイズも異なることとなり、平均化による SN 向上が見込まれる。そこで、相補的サンプリングを用いず、I ピクチャの位置を互いに 7 フレームずらして符号化・復号 (N=15, M=1) した 2 系統を平均化したときの SN 向上度を表 4.5 に示す。表には、比較のため、相補的サンプリング変換 (4 画素ずらし) を行った場合の SN 向上度 (図 4.6 に対応、ただし I ピクチャのずらしも含む) を含めてある。この表より、相補的サンプリング変換を行わない場合の PSNR の向上度は、行う場合に比べ 0.6-0.9dB ほど低下するものの、約 1dB あることがわかる。

また、図 4.6 と表 4.5 を比較すると、相補的サンプリング変換を行った上に I ピクチャのずらしを加えても、それを加えない場合と有意な差異はない。相補的サンプリング変換のみによって、すでに可能なまでの符号化ノイズの相関低下が実現されているものといえる。

相補的サンプリングを含む場合の PSNR の向上度が I ピクチャずらしのみの場合よりも大きいこと理由は、次のように考えられる。すなわち、I ピクチャずらしのみの場合、極端な例として、2 系統で、予測画像が全く同一で量子化パラメータ (MQUANT) も同一のピクチャでは、符号化ノイズも同一となる。したがって平均化による PSNR 向上は望めない。一方、相補的サンプリング (ブロックの境界ずらし) があれば、たとえそのような極端な場合でさえも、DCT のブロック境界が異なる以上、符号化ノイズが異なる可能性が強い。このように、2 系統の符号化ノイズをなるべく確実に相異ならせられることが、相補的サンプリングがある場合に I ピクチャずらしのみの場合に比べて PSNR 向上度が大きくなる理由であると考えられる。

4.5.2 符号化ノイズ残留相関法による客観画質評価の検証

SN 比の推定法の検証

前節と同じ 6 種類の動画像を 30 フレームずつ、5-45Mbps の 4 種類のレートで符号化することにより、その復号画 x_2 と原画との実効 SN (RSNR= $10 \log s/n_2$) を測定しておく。原信号電力 s としては符号化単位である 16x16 のマクロブロック毎に平均値分離した後の信号分散を使用する。

一方、相補的サンプリング法により 4 画素ずらして符号化したもう 1 系統の復号画 x_1 と上記の復号画 x_2 から、 $\frac{\sigma^2(x_1)}{\sigma^2(x_1 - x_2)}$ を測定する。

これらをプロットした図 4.9 では、図中の点は、式 (4.6) に当てはめた場合、 $\alpha = 2.0$ で平均誤差は ± 0.6 dB におさまることがわかる (破線で示した回帰直線)。

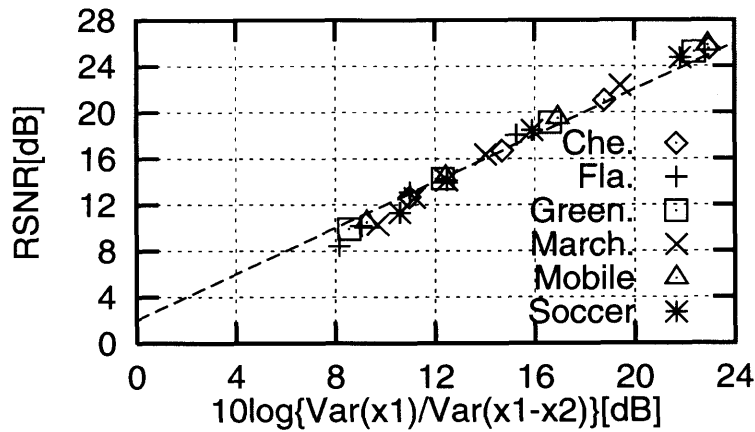


図 4.9: 2 受信画像からの SN 比推定

その理由は、式 (4.5) から式 (4.6) を導く際の仮定、すなわち $1 \gg \frac{n_1}{s} + 2\rho_{xd1}\sqrt{\frac{n_1}{s}}$ が成立しているからと考えられるが、次にそれを確かめる。まず $\frac{n_1}{s}$ は、図 4.9 における $RSNR$ から逆算できる。例えば $RSNR$ が 8dB から 24dB として、 $\frac{n_1}{s}$ は最大で 0.16 (低ビットレート)、最小で 0.004 (高ビットレート) である。また ρ_{xd1} を測定したところ、表 4.6 のようになった。これらより $\frac{n_1}{s} + 2\rho_{xd1}\sqrt{\frac{n_1}{s}}$ を計算すると、表 4.7 のようになり、約 -0.07 から +0.01 の範囲に収まっている。すなわち、前記仮定が成立していることが確認できた。

表 4.6: 原画像信号 x と符号化雑音信号 $d1$ の相関係数 ρ_{xd1}

Bitrate [Mbps]	45	22.5	11.2	5.6
Che.	-0.009	-0.034	-0.071	-0.127
Fla.	-0.018	-0.081	-0.142	-0.194
Green.	-0.016	-0.065	-0.150	-0.268
March.	-0.018	-0.090	-0.175	-0.252
Mobile	-0.015	-0.054	-0.119	-0.196
Soccer	-0.026	-0.082	-0.159	-0.240

表 4.7: $n_1/s + 2\rho_{\hat{x}d1}\sqrt{n_1/s}$ の値

Bitrate[Mbps]	45	22.5	11.2	5.6
Che.	0.0019	0.0018	0.0007	-0.0048
Fla.	0.0112	0.0131	0.0083	-0.0032
Green.	0.0013	-0.0022	-0.0206	-0.0692
March.	0.0031	-0.0043	-0.0271	-0.0606
Mobile	0.0011	-0.0002	-0.0090	-0.0279
Soccer	0.0003	-0.0053	-0.0239	-0.0566

以上より、2 系統の復号画像から $\frac{\sigma^2(x_1)}{\sigma^2(x_1 - x_2)}$ を測定し、図 4.9 の直線に当てはめれば、その復号画像の SN 比が推定できることになる。

WSNR の推定法の検証

まず、4.5.2 において符号化・復号した 6 種類の画像の WSNR を測定しておく。方法としては、4.4.2 に示した方法を用いる。ただし、単位視角度当たりの画素数は、文献 [6] と同様、64pel/degree とした。また、2 系統の復号画 x_1, x_2 に対する式 (4.7) のフィルタリングにより、 \hat{x}_1, \hat{x}_2 を求め、これらから $\frac{\sigma^2(\hat{x}_1)}{\sigma^2(\hat{x}_1 - \hat{x}_2)}$ を測定する。

これらをプロットした図 4.10 では、図中の点は、式 (4.11) に当てはめた場合、 $\hat{\alpha} = 1.1$ で平均誤差は $\pm 0.6\text{dB}$ におさまることがわかる (破線で示した回帰直線)。

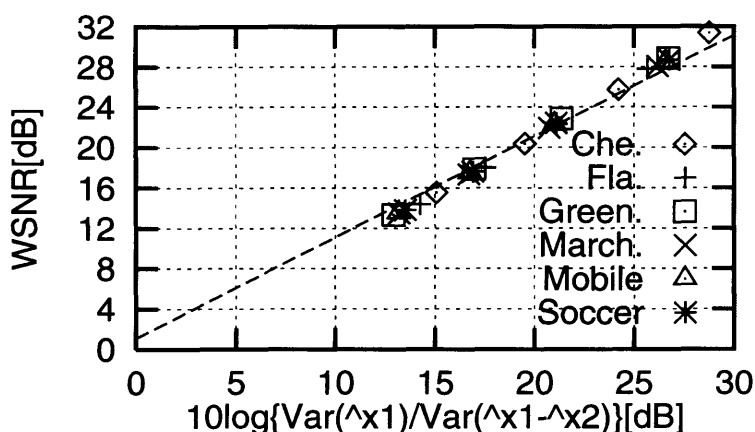


図 4.10: 2 受信画像からの重みつき SN 比の推定

その理由は、式 (4.10) から式 (4.11) を導く際の仮定、すなわち $1 \gg \hat{n}_1/\hat{s} + 2\rho_{\hat{x}d1}\sqrt{\hat{n}_1/\hat{s}}$ が成立しているからと考えられるが、次にそれを確かめる。まず \hat{n}_1/\hat{s} は、図 4.10 における WSNR から逆算できる。例えば WSNR が 12dB から 32dB とし、 \hat{n}_1/\hat{s} は最大で 0.06 (低ビットレート)、最小で 0.0006 (高ビットレート) である。また $\rho_{\hat{x}d1}$ を測定したところ、表 4.8 のようになった。

これらより $\hat{n}_1/\hat{s} + 2\rho_{\hat{x}\hat{d}1}\sqrt{\hat{n}_1/\hat{s}}$ を計算すると、表 4.9 のようになり、約 -0.0003 から -0.07 の範囲に収まっている。すなわち、前記仮定が成立していることが確認できた。

以上より、2 系統の復号画像に、式 (4.7) のフィルタをかけ、 $\frac{\sigma^2(\hat{x}_1)}{\sigma^2(\hat{x}_1 - \hat{x}_2)}$ を測定し、図 4.10 の直線に当てはめれば、復号画像の WSNR を推定できることとなる。

表 4.8: \hat{x} と $\hat{d}1$ の相関係数 $\rho_{\hat{x}\hat{d}1}$

Bitrate[Mbps]	45	22.5	11.2	5.6
Che.	-0.019	-0.034	-0.070	-0.130
Fla.	-0.045	-0.076	-0.123	-0.183
Green.	-0.026	-0.051	-0.106	-0.197
March.	-0.029	-0.073	-0.134	-0.193
Mobile	-0.024	-0.050	-0.108	-0.188
Soccer	-0.040	-0.095	-0.171	-0.263

表 4.9: $\hat{n}_1/\hat{s} + 2\rho_{\hat{x}\hat{d}1}\sqrt{\hat{n}_1/\hat{s}}$ の値

Bitrate[Mbps]	45	22.5	11.2	5.6
Che.	-0.0003	-0.0008	-0.0042	-0.0155
Fla.	-0.0020	-0.0058	-0.0151	-0.0334
Green.	-0.0006	-0.0022	-0.0108	-0.0385
March.	-0.0007	-0.0053	-0.0180	-0.0369
Mobile	-0.0004	-0.0018	-0.0110	-0.0348
Soccer	-0.0016	-0.0087	-0.0277	-0.0657

2 系統の符号化構成が異なる場合

次に、本推定方式が、上記のように 2 系統で符号化構成が同じ場合のみならず、異なる場合でも適用可能であることを検証する。

前の実験と同様、6 種類の動画像を 30 フレームずつ、5-45Mbps の 4 種類のレートで符号化することにより、 (s/n_2) と $\{\sigma^2(x_1)/\sigma^2(x_1 - x_2)\}$ を得、式 (4.6) に当てはめる。

符号化は MPEG-2 TM5 とし、その構成としては、次の 6 種類をとった。

- (a) N=15, M=3
- (b) N=15, M=1
- (c) N=2, M=0 (IBIB...)
- (d) N=15, M=1, bitrate をもう片系統に比べ 5Mbps 増加
- (e) N=15, M=1, bitrate をもう片系統に比べ 10Mbps 増加
- (f) N=15, M=3, 2 段タンデム接続

本実験では、画面ずらしは行わず、I ピクチャの時間的位置を 2 系統で異ならせた。すなわち、符号化前処理は行なわない構成（2 系統は独立に符号化）で、受信側だけで処理を行なうことに相当する。

各符号化構成を種々に組み合わせた 2 重化構成のそれぞれについて、図 4.9 に対応するグラフを 図 4.12 と 図 4.13 に示す。回帰直線の切片値 α とそれに対する各点の平均誤差も記した。

図 4.12 と 図 4.13 より、符号化構成の組合せ方により程度の差はあるものの、全ての場合において式 (4.6) の直線性が良く成立している。すなわち、 α は、画像による区別をせず、符号化構成の組合せ毎に 1 個の値で代表させても問題ないとみなせる。

すなわち、タンDEM接続 (図 4.11) や 2 系統でレートが異なる場合も含め、上記回帰直線をテーブルとして用いることにより、受信画像 x_1, x_2 から SN を平均 1dB 未満の誤差で高精度に推定可能である。

以上により提案方式の有効性が検証できた。

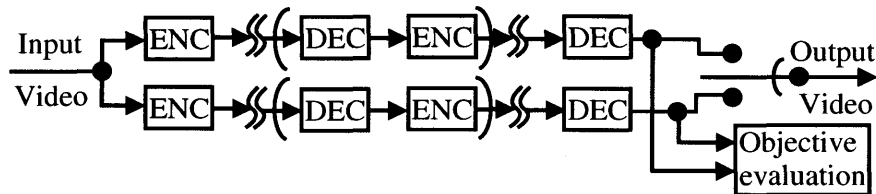


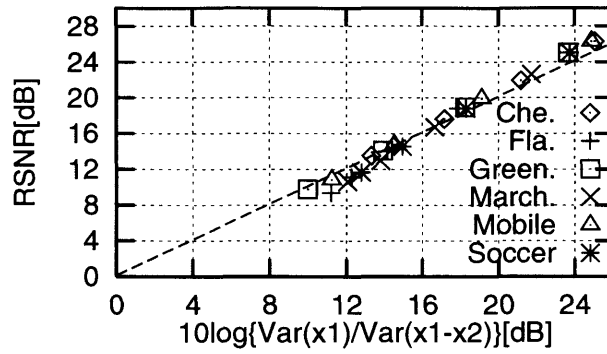
図 4.11: 2 系統の復号画像からの客観画質評価 (タンDEM接続)

4.6 本章のまとめ

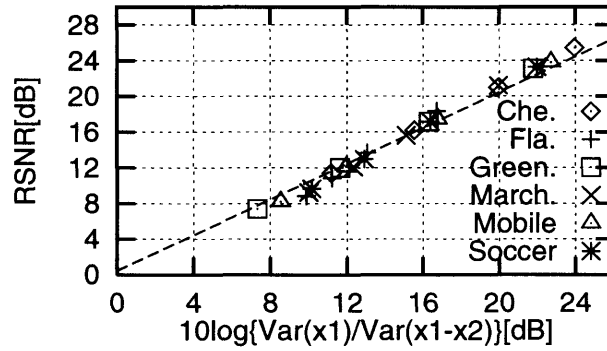
本章では、高能率・高信頼度伝送と伝送画質の自動監視を実現するフラットマルチスケラブル映像伝送方式を提案した。

本方式は、従来の映像素材伝送に利用される完全冗長 2 重化系を高能率化・高機能化する方式として、既存の符号化/復号装置構成に前・後処理を加えるだけで実現可能である。2 系統の符号化雑音の差異を積極的に活用することにより、画質向上と画質評価を達成する。また、この画質評価法は、タンDEM接続や符号化構造・レートが 2 系統で異なる場合など、様々な構成においても適用可能である。

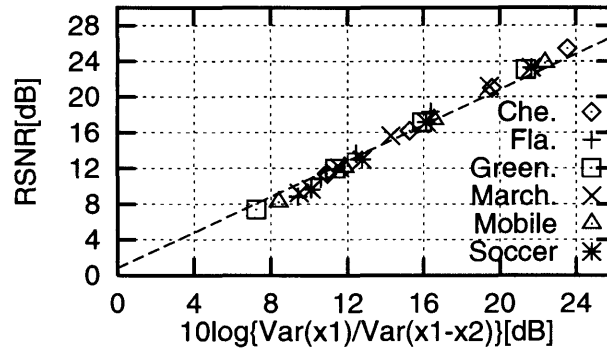
さらに、近年盛んなインターネットなどの公衆網を介した画像伝送では、パケットロスに起因する画像乱れがしばしば発生する。提案方式を利用すれば、静止画/動画及びコーデックの種類によらず、既存の部分を変更せず単なる機能追加により、その高信頼度化が簡易に達成されることになる。例えば、UDP によるインターネット動画像伝送を 2 重化し、通常時は両系統平均化による高画質画像を出力する。そして、IP パケットのロスの発生が片方のデコーダで検知された場合には、正常な方のストリームのデコード画像のみを出力するようにすればよい。



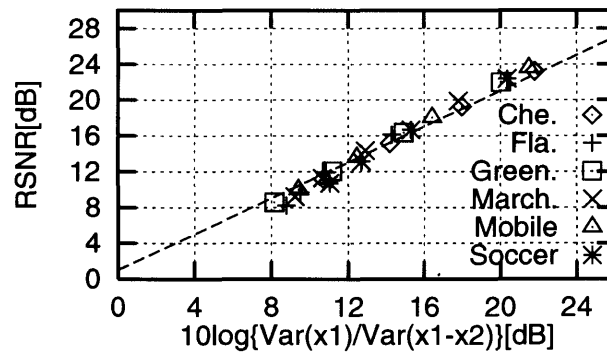
符号化構成：(b) / (b). 回帰直線 ($\alpha=0.2$) からの平均誤差は 0.74dB.



符号化構成：(b) / (c). 回帰直線 ($\alpha=0.5$) からの平均誤差は 0.61dB.

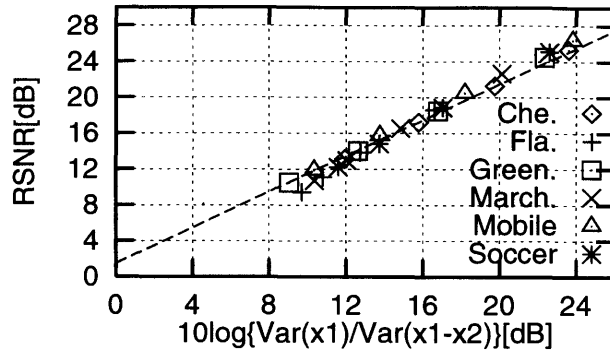


符号化構成：(a) / (c). 回帰直線 ($\alpha=0.8$) からの平均誤差は 0.66dB.

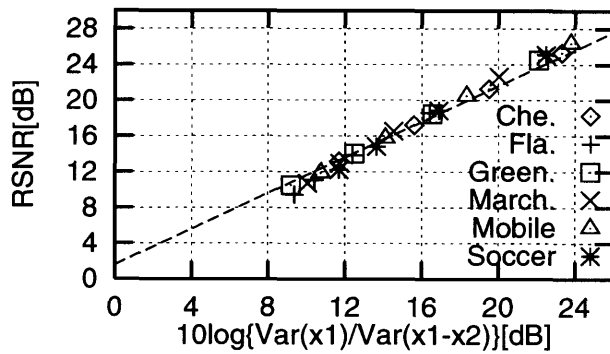


符号化構成：(f) / (f). 回帰直線 ($\alpha=1.0$) からの平均誤差は 0.62dB.

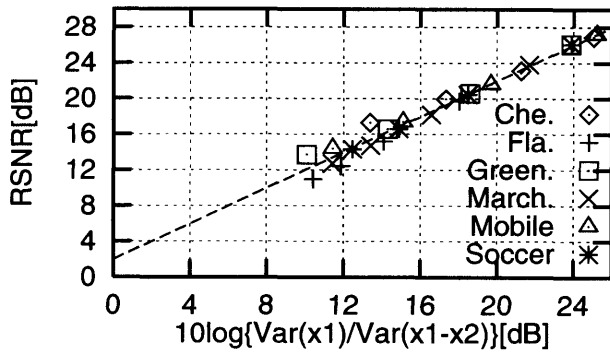
図 4.12: 符号化構成が異なる 2 系統の画像からの SN 比推定 (1)



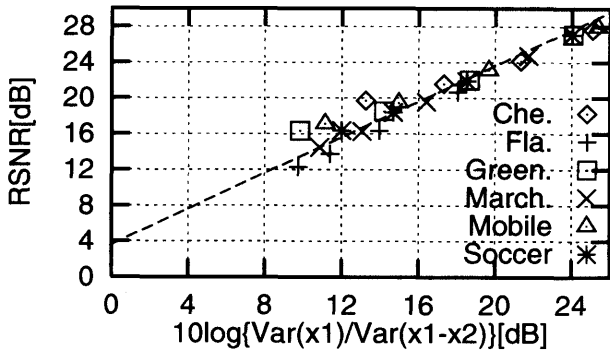
符号化構成：(b) / (a). 回帰直線 ($\alpha=1.5$) からの平均誤差は 0.56dB.



符号化構成：(a) / (a). 回帰直線 ($\alpha=1.6$) からの平均誤差は 0.55dB.



符号化構成：(b) / (d). 回帰直線 ($\alpha=2.0$) からの平均誤差は 0.54dB.



符号化構成：(b) / (e). 回帰直線 ($\alpha=3.6$) からの平均誤差は 0.92dB.

図 4.13: 符号化構成が異なる 2 系統の画像からの SN 比推定 (2)

本章の参考文献

- [1] ISO/IEC 13818-2: “Information technology – Generic coding of moving pictures and associated audio information: Video” (2000).
- [2] V.A.Vaishampayan: “Design of multiple description scalar quantizers”, IEEE Trans., **IT-39**, 3, pp. 821–834 (1993).
- [3] A.R.Reibman, Y.Wang, M.T.Orchard and R.Puri: “Multiple description coding for video using motion compensated prediction”, Proc. IEEE International Conf. on Image Processing, Vol. 3, Kobe, Japan, pp. 837–841 (1999).
- [4] T. Hamada, S. Miyaji and S. Matsumoto: “Picture quality assessment system by three-layered bottom-up noise weighting considering human visual perception”, SMPTE Journal, **108**, 1, pp. 20–26 (1999).
- [5] N.B.Nill: “A visual model weighted cosine transform for image compression and quality assesment”, IEEE Trans., **COM-33**, 6, pp. 551–557 (1985).
- [6] B.Chitprasert and K.R.Rao: “Human visual weighted progressive image transmission”, IEEE Trans., **COM-38**, 7, pp. 1040–1044 (1990).