

PHSによるライブメディア伝送時のメディア同期性能評価

加藤 正美[†] 臼井 憲義^{††} 田坂 修二^{†††}

Performance Evaluation of Live Media Synchronization in PHS

Masami KATO[†], Noriyoshi USUI^{††}, and Shuji TASAKA^{†††}

あらまし 本論文は、PHSによるライブメディアの伝送において、ビデオフレームサイズ^(注1)や再送による遅延の変動の影響を受けるメディア同期を検討の対象とする。特に、筆者らが以前に提案した蓄積メディア同期制御方式と整合性のあるライブメディア同期制御方式を考える。そのため、蓄積メディアの伝送に適用したスライド制御をライブメディアの伝送にも適用する。伝送遅延を小さくする必要があるライブメディアの特性を考慮し、前方へのスライド制御の追加、後方へのスライド制御の制限等を含むスライド制御の拡張方式を提案する。更に、送信側において空間的・時間的解像度制御を適用する。なお本研究でも蓄積メディアの場合と同様に、PHSの二つの通信チャンネルを使用し、H.263圧縮ビデオ情報をARQで、ADPCM音声情報を無手順で、リアルタイムに伝送しながら再生する場合を想定している。そして、ランダム誤りとバースト誤りの各環境下で、実際のビデオ情報と音声情報を使用したシミュレーションを行い、メディア同期や伝送遅延の性能評価、および主観評価により本提案方式の有効性を示す。

キーワード PHS, ビデオ伝送, ライブメディア, メディア同期, 伝送遅延

1. ま え が き

簡易型携帯電話 PHS (Personal Handy Phone System) [1]の登場により、モバイルコミュニケーションが急速に進展している。更に、データ伝送制御手順 PIAFS (PHS Internet Access Forum Standard)[2]は、高速・高品質な無線データ通信を可能にし、PHSを利用したモバイルコンピューティングも可能になった。しかしながら、PHSによるモバイルマルチメディアの実現には、データだけでなく、ビデオと音声も含めた統合伝送が不可欠である。

PHSでビデオ伝送を行うには、H.263[3]等の低ビットレート対応の動画像圧縮符号化方式を利用する。H.263はフレーム間相関を高度に利用した方式であり、一つのビデオフレームの伝送誤りが、連続したビデオフレームに影響を及ぼす可能性が高い。従って、フェージングの影響を受ける PHS で、H.263 圧縮ビ

デオ情報を伝送するには、誤り訂正能力に限界がある FEC (Forward Error Correction) よりも、確実に訂正できる ARQ (Automatic Repeat reQuest) の適用が重要になる。しかしながら、再送による伝送遅延は、送信側でのビデオの発生間隔と受信側でのビデオの表示間隔との同期 (メディア内同期) や、対応するビデオと音声の時間的關係 (メディア間同期) を乱してしまう。

これまでに筆者らは、PHSによる蓄積メディア伝送時のメディア同期性能について検討を行った[4]。特に、メディア同期制御方式としてスライド制御[4]を提案し、その有効性を確認した。しかしながら、マルチメディアコミュニケーションにおいては、蓄積メディアに加えて、ライブメディアの伝送も重要な構成要素である。特に、テレビ会議等では、両者を切り換えて使用することも考えられる。従って、一つのメディア同期制御方式でライブメディアと蓄積メディアを同時に扱えるようにする必要がある。

そこで、蓄積メディアに適用したスライド制御をライブメディアの伝送にも適用することを考える。このとき、次のような課題が発生する。蓄積メディアの伝送では、スライド制御により目標出力時刻 $t_n^{(1)}$ を後方に変更することで、伝送遅延が増加するものの同期外れの発生頻度を

(注1)：ビデオフレームサイズは圧縮符号化された各ビデオフレームの符号量 (ビット数) を示すものであり、空間的な画像サイズとは異なる。

[†] 三洋電機株式会社ハイパーメディア研究所, 岐阜県
Hypermedia Research Center, SANYO Electric Co., Ltd., Gifu-ken,
503-0195 Japan

^{††} 三洋電機株式会社情報通信開発研究所, 岐阜県
Information & Telecommunications Development Center, SANYO
Electric Co., Ltd., Gifu-ken, 503-0195 Japan

^{†††} 名古屋工業大学電気情報工学科, 名古屋市
Department of Electrical and Computer Engineering, Nagoya Institute of
Technology, Nagoya-shi, 466-8555 Japan

少なくできた。しかしながら、ライブメディアの伝送では、できるだけ伝送遅延を小さくする必要がある。これが、蓄積メディアとライブメディアとで本質的に異なる点である。従って、メディア同期を良好に維持しつつ、どのように伝送遅延を削減するかが大きな課題となる。

更に、PHS で H.263 圧縮ビデオ情報を伝送する場合、PHS の伝送容量とはほぼ等しくなるように H.263 の符号発生量を制御することが多いため、ビデオフレームサイズの大小が、伝送遅延と共にメディア同期にも影響を及ぼす。蓄積メディアの伝送では、ネットワークの状態に応じて符号発生量を制御することが困難であった。しかしながら、ライブメディアの伝送では、ネットワークの状態に応じて符号発生量を容易に制御できる。従って、どのようにビデオフレームサイズや符号発生量を制御して、伝送遅延を小さくおさえるかも課題となる。

一方、PHS によるライブメディアの伝送に関する他の研究として、文献[5]は、受信した NAK パケット数に応じた時間的・空間的解像度制御方式を提案し、伝送遅延に対する考察を行っている。しかし、メディア同期に対する考察はなされていない。また、文献[6]は、ビデオのメディア内同期性能を報告しているが、ビデオと音声のメディア間同期性能は明らかにしていない。

なお文献[7]でも、目標出力時刻を変更するメディア同期制御方式を提案し、ATM ネットワークにおけるライブメディアの伝送実験による評価を行っている。しかし、音声への影響を考慮したきめ細やかな制御はなされていない。

そこで本研究では、リアルタイムに発生する H.263 圧縮ビデオ情報と 32 kbit/s ADPCM 音声情報を、それぞれ PHS の一つの通信チャネルを使用し、PHS 基地局を介し

て ISDN 端末と PHS 移動局間で伝送する場合を考える。特に、ビデオは ARQ で、音声は無手順で伝送する (図 1)。

このとき、ビデオや音声のメディア内同期や、両者のメディア間同期を維持するために、受信側においてスライド制御を適用し、その拡張を試みる。特に、スライド制御により増加した伝送遅延を削減するために、目標出力時刻 $t_n^{(0)}$ を前方に変更できるようにスライド制御を拡張する。また、伝送遅延が大きくなりすぎないように、後方へのスライド制御を制限する。更に、送信側において、符号発生量やビデオフレームサイズを制御するために、空間的・時間的解像度制御を適用する。

そして、ランダム誤りとバースト誤りの各環境において、実際の H.263 圧縮ビデオ情報と ADPCM 音声情報を使用したシミュレーションを行い、メディア同期と伝送遅延の性能評価、および主観評価を行い、本提案方式の有効性を示す。

本論文の構成は次のとおりである。2. では提案方式について、3. ではシミュレーションの仮定について述べる。4. では数値結果と性能評価を報告する。

2. メディア同期制御方式

本研究では、蓄積メディア同期制御方式と整合性のあるライブメディア同期制御方式を考える。そのため、蓄積メディアの伝送に適用したスライド制御を、ライブメディアの伝送にも適用し、更にライブメディアの特性を考慮した改良を加える。

ここで文献[4]に従い、各メディアは、その出力単位であるメディアユニット(MU)から構成されているものと仮定する。特に、ビデオは 1 ビデオフレームを 1 ビデオ MU とし、音声は一定の情報量をもって 1 音声 MU を構成すると仮定する。そして、メディア発生源 i における n 番目の MU の発生時刻を $T_n^{(i)}$ 、 n 番目と $n+1$ 番目の MU の発生間隔を $\sigma_{n,n+1}^{(i)} (= T_{n+1}^{(i)} - T_n^{(i)})$ と定義する (図 2)。なお、以後 $i=1$ はビデオ、 $i=2$ は音声を示す。

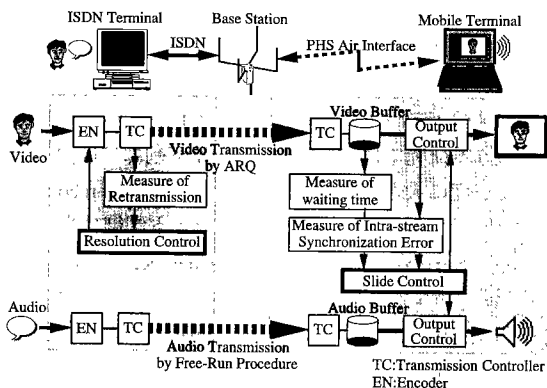


図 1 PHS ビデオ・音声伝送システム構成図

Fig. 1 Block diagram of video & audio transmission system for PHS.

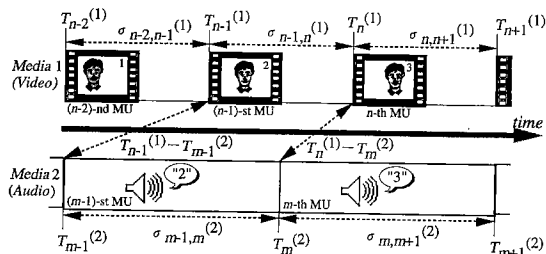


図 2 ビデオ MU と音声 MU の時間関係

Fig. 2 Temporal relationship between video and audio MUs.

このとき、メディア内同期は、単一メディア*i*を構成するMU間の発生間隔 $\sigma_{n,n+1}^{(i)}$ を、メディアの出力先においても出力間隔として維持することである。更に、メディア間同期は、メディア1とメディア2の間で対応するMUの発生間隔 $T_n^{(1)} - T_m^{(2)}$ を、メディアの出力先において同様に維持することである。

ここで、メディアの出力先におけるメディア*i*の*n*番目のMUの目標出力時刻を $t_n^{(i)}$ 、遅延ジッタがない理想の場合の目標出力時刻を $x_n^{(i)}$ 、到着時刻を $A_n^{(i)}$ 、出力時刻を $D_n^{(i)}$ 、到着してから出力されるまでの出力待ち時間を $\tau_n^{(i)} (= D_n^{(i)} - A_n^{(i)})$ と定義する。更に、 $D_n^{(i)}$ を $t_n^{(i)}$ から意図的に変化させる時間量としてのスライド時間を $\Delta S_n^{(i)}$ と定義し、総スライド時間を $S_n^{(i)} (= S_{n-1}^{(i)} + \Delta S_n^{(i)})$ と定義する。このとき、 $t_n^{(i)}$ 、 $D_n^{(i)}$ 、メディア内同期ずれ時間 $\Delta_n^{(i)}$ 、およびメディア間同期ずれ時間 $\Delta_n^{(1-2)}$ は、次式で与えられる。なお、*m*は、メディア1の*n*番目のMUの直前に発生したメディア2の番号を示す。

(i) 目標出力時刻

$$t_1^{(i)} = A_1^{(i)} + \tau_1^{(i)}, \quad x_1^{(i)} = t_1^{(i)} \quad (n=1)$$

$$t_n^{(i)} = x_n^{(i)} + S_{n-1}^{(i)}, \quad x_n^{(i)} = x_{n-1}^{(i)} + \sigma_{n-1,n}^{(i)} \quad (n \geq 2)$$

(ii) 出力時刻

$$D_n^{(i)} = t_n^{(i)} + \Delta S_n^{(i)} \quad \text{if } A_n^{(i)} \leq t_n^{(i)} + \Delta S_n^{(i)} \quad (n \geq 1)$$

$$D_n^{(i)} = A_n^{(i)} \quad \text{if } t_n^{(i)} + \Delta S_n^{(i)} < A_n^{(i)} \quad (n \geq 1)$$

(iii) メディア内同期ずれ時間

$$\Delta_n^{(i)} = D_n^{(i)} - t_n^{(i)} \quad (n \geq 1)$$

(iv) メディア間同期ずれ時間

$$\Delta_n^{(1-2)} = (D_n^{(1)} - D_m^{(2)}) - (T_n^{(1)} - T_m^{(2)}) \quad (n \geq 1, m \geq 1)$$

2.1 スライド制御の拡張

蓄積メディアの伝送では、ビデオと音声のメディア間同期を制御するために、スライド制御を適用した。これは、ビデオのメディア内同期ずれ時間 $\Delta_n^{(i)}$ を計測し、

その値に応じて $t_n^{(i)}$ を $\Delta S_n^{(i)}$ だけ遅らせるものである。ここではこれを、後方へのスライド制御と呼ぶ(図3(a))。この方式は、伝送遅延が増加するものの、同期外れの発生頻度を少なくできた。しかしながら、ライブメディアの伝送においては、できるだけ伝送遅延を小さくする必要がある。

そこで、後方へのスライド制御により増加した伝送遅延を削減することを考える。ここで、各ビデオMUの $t_n^{(i)}$ と $A_n^{(i)}$ の差を受信バッファでの目標出力待ち時間 $\tau_n^{(i)}$ と定義する。この値が大ききことは、伝送遅延が大きな状態でメディア同期を維持していると考えられる。従って、 $\tau_n^{(i)}$ を計測し、その値に応じて $t_n^{(i)}$ を $\Delta S_n^{(i)}$ だけ前方に変更できるようにスライド制御方式を拡張する。ここではこれを前方へのスライド制御と呼ぶ(図3(b))。これにより、例えば、一時的な通信環境の悪化に伴い、後方へのスライド制御が実行され、伝送遅延が増加したとしても、通信環境が回復した後に、伝送遅延を再び削減できる。

つぎに、後方へのスライド制御により大きな伝送遅延が発生しないように、後方へのスライド制御を制限することを考える。ここで、後方へのスライド許容時間を κ と定義する。そして、総スライド時間 $S_n^{(i)}$ が κ を超える場合には、後方へのスライド制御の実行を禁止する。

なお、文献[8]で提案した受信バッファ占有率を利用した後方へのスライド制御は、伝送遅延の増加をまねくため、ライブメディアの伝送では適用しないこととする。

ところで、蓄積メディアの伝送におけるスライド制御では、出力時刻の変更が音声に及ぼす影響を考慮し、Gradual recovery^(注2)とQuick recoveryの2種類の変更方法を採用した[8]。Gradual recoveryは、音声への影響を小さくするために、 $t_n^{(i)}$ を少しずつ変更して徐々にメディア間同期を修復する。またQuick recoveryは、音声への影響は大きいけれども、 $t_n^{(i)}$ を一度に大きく変更し、早急にメディア間同期を修復する。ライブメディアの伝送を扱う本研究では、後方へのスライド制御には、伝送遅延が一度に大きく増加しないようにGradual recoveryのみを適用する。また、前方へのスライド制御にも、音声への影響を小さくするためにGradual recoveryを適用する。

ここで文献[4]に従い、Gradual recoveryによる後方へのスライド制御を実行するための $\Delta_n^{(i)}$ のしきい値を T_{h1} 、

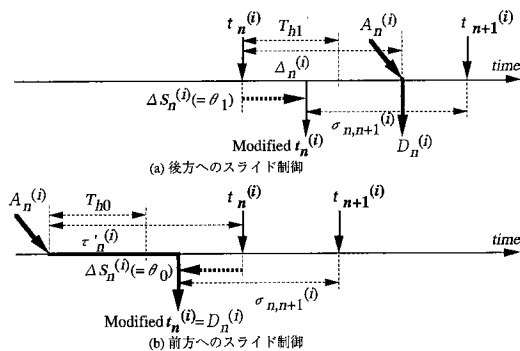


図3 スライド制御

Fig. 3 Slide control scheme.

(注2) : $\Delta_n^{(i)}$ や受信バッファの占有率に応じて実行されるという点で、文献[9]で提案されたGraceful Recoveryとは異なる。

そのスライド時間を $\theta_1 (>0)$, スライド制御を実行できる最小許容時間間隔を ω_1 と定義する. また, 前方へのスライド制御を実行するための $\tau_n^{(1)}$ のしきい値を T_{n0} , そのスライド時間を $\theta_0 (<0)$, スライド制御を実行できる最小許容時間間隔を ω_0 と定義する. このとき, 以下の条件に応じて, スライド制御を実行する.

$$(i) \quad T_{n1} \leq \Delta S_n^{(1)}, \text{ かつ } S_{n-1}^{(1)} + \Delta S_n^{(1)} \leq \kappa$$

後方にスライド制御を実行する. なお, 1ビデオMUに対し1回のスライド制御しか実行されず, $\Delta S_n^{(1)} = \Delta S_m^{(2)} = \theta_1$ となる (図3(a)に対応).

$$(ii) \quad \Delta S_n^{(1)} < T_{n1}, \text{ かつ } T_{n0} \leq \tau_n^{(1)}$$

前方にスライド制御を実行する. なお, (i) と同様に $\Delta S_n^{(1)} = \Delta S_m^{(2)} = \theta_0$ となる (図3(b)に対応).

(iii) 上記 (i), (ii) 以外の場合
スライド制御を実行せず.

一方, メディア同期の基準時刻の決定方法にも注意が必要である. 蓄積メディアの伝送では, 最初のビデオMUの $A_1^{(1)}$ に $\tau_1^{(1)}$ を加えて, 基準時刻を決めていた. しかし, 伝送遅延に制約があるライブメディアの伝送に同じ方法を適用すると, 最初のビデオMUサイズ (注3) が, 基準時刻の決定に大きな影響を与えてしまう. 特に, H.263 ビデオ情報の最初のビデオMUは, MUサイズが大きなI(Intra)フレームである. そのため, 蓄積メディアと同様に基準時刻を決めると, 通信開始時に伝送遅延が大きくなってしまう. また, 最初のビデオMUを受信しなければ, 音声MUを出力できないため, BERが悪い環境では, 音声のMU遅延が大きくなってしまう. そこで, ライブメディアの伝送では, 最初の音声MUの $A_1^{(2)}$ に $\tau_1^{(2)}$ を加えて, 基準時刻を決めることとする. なお, 本研究では $\tau_1^{(2)}$ を0と仮定する.

2.2 空間的・時間的解像度制御

PHSでビデオ伝送を行う場合, H.263の符号発生量はPHSの伝送容量とほぼ等しくなるように制御されることが多い. 例えば, 送信側において, 空間的・時間的解像度制御を適用し, 符号発生量を制御することが一般的に行われている. このとき, MUサイズの大小が伝送遅延やメディア同期に影響を及ぼすため, どのようにMUサイズや符号発生量を制御するかが課題となる.

文献[10]は, ネットワークの伝送速度を固定値と仮定した場合における空間的・時間的解像度制御方式を規定している. まず, 空間的解像度制御では, 目標MUレートを定め, 各ビデオMUのMUサイズが, 伝送速度と目標MUレートから定まる目標MUサイズに近くなるように, 量子化特性を制御している. 特に, 直前

のビデオMUの量子化特性やMUサイズに加え, マクロブロックごとの符号発生量を計算し, マクロブロック単位で量子化特性を制御している. 一方, 時間的解像度制御では, 送信バッファのビデオ情報量を, 各ビデオソースのサンプル時刻ごとに計算で求め, この値があるしきい値より小さくなるサンプル時刻までフレームスキップを行っている. しかしながら, 送信バッファのビデオ情報量は, ネットワークの伝送速度が固定値であると仮定して計算されており, 再送等による通信状態の変化には対応していない.

文献[5]は文献[10]を拡張し, 再送回数に応じて, 時間的解像度と空間的解像度の両方を制御する方式を提案している. しかしながら, 再送回数に応じて量子化特性を直接変更する空間的解像度制御は, 画質の大幅な劣化を引き起こす可能性がある. また, 動きがあるビデオソース等では, 量子化特性を変更しても, 符号発生量が削減されるとは限らない.

そこで本研究では, 再送等による通信状態の変化を考慮した時間的解像度制御と, これを考慮しない空間的解像度制御を適用することを考える.

まず, 時間的解像度制御では, 文献[10]の方式に再送による通信状態の変化を考慮した拡張を行う. すなわち, ビデオソースをサンプルするごとに, 送信バッファのビデオ情報量を実際に計測し, その値があるしきい値より小さければ, そのビデオフレームを符号化する. このとき, 実際に計測した送信バッファのビデオ情報量は, 再送の影響を直接反映しており, 通信状態に十分よく対応した符号発生量の削減を確実に期待できる.

一方, 空間的解像度制御として, 再送を考慮しない文献[10]と同じ方式を採用する. すなわち, 各ビデオMUのサイズが, 目標MUサイズに近くなるように, 量子化特性を制御する. これにより, 通信状態の悪化により, 画質までもが大幅に劣化してしまうことを回避できる.

3. シミュレーションの仮定

3.1 伝送制御方式の仮定

各メディアの伝送に使用する伝送制御方式に関し, 以下の仮定を置く.

(1) ビデオ情報の伝送

PHSの32 kbit/s非制限ディジタルベアラサービスにおいて, SR-ARQ (Selective Repeat - ARQ) を採用する[2]. ARQフレームの構造は640 bitを一つのARQフレームと

(注3): 本研究では1ビデオフレームを1ビデオMUと仮定しており, ビデオMUサイズはビデオフレームサイズに等しくなる.

し、ARQ 制御情報とユーザデータを常に組み込むビジーバッキング方式とする(図4)。なお、ARQ フレームの同期は通信開始時に確立するものとする。

(2) 音声情報の伝送

ビデオとは別の通信チャンネルにおいて、32 kbit/s 音声サービスと同様にADPCM音声情報を無手順で伝送する。

3.2 通信環境の仮定

ランダム誤り環境およびバースト誤り環境におけるシミュレーションを行うにあたり、以下の仮定を置く。

(A1) ビット誤りは、基地局と移動局間の無線部における情報チャンネル TCH のみに発生する。ISDN 端末と基地局間の有線部におけるビット誤りは無視する。

(A2) ランダム誤り環境におけるビット誤りの発生は、幾何分布に従う。

(A3) バースト誤り環境として2状態マルコフ連鎖モデル(ギルバートモデル)を適用する[11]。なお、バースト状態から良好状態への遷移確率 γ 、良好状態からバースト状態への遷移確率 σ は、フェージング周波数8 Hz(歩行時に相当)における各受信レベルとビット誤り率(BER)、平均バースト長の測定結果から算出した[6]。また、バースト状態と良好状態間の状態遷移の発生は、幾何分布に従う。

(A4) ISDN 端末と基地局間の伝搬遅延時間、および ARQ フレーミングによる処理遅延時間の合計を 20 ms と仮定する。また、基地局と移動局間の無線部において電波伝搬遅延は無視する。

(A5) ACK/NACK 等の ARQ 制御情報には誤りが発生しないものとし、また再送要求の廃棄は行わない。

3.3 ライブメディアの仮定

シミュレーションで使用するH.263圧縮ビデオ情報とADPCM 音声情報の作成に関し、以下の仮定を置く。

(A1) H.263で規定された四つのコーディング・オプションは、いずれも使用しない。

(A2) 1 ビデオフレームを1 ビデオ MU とする。

(A3) ビデオの目標符号発生レートを29.2 kbit/sとする。この値は、ビデオ伝送におけるスループットの最

大値に等しい。また、目標 MU レートを 15 MU/s とする。従って、目標 MU サイズは 1,947 bits となる。

(A4) n 番目のビデオ MU の発生時刻 $T_n^{(1)}$ は、符号化時にビデオフレームに設定された Temporal Reference (TR_n) により、次式で与える。但し、 TR_1 (最初のビデオフレームの TR) を 0 とする。

$$T_n^{(1)} = TR_n / 30$$

(A5) 音声の 1 MU サイズを 1,600 bit、MU 発生レートを 20 MU/s の固定値とし、 n 番目の MU の発生時刻 $T_n^{(2)}$ を次式で与える。

$$T_n^{(2)} = (n - 1) / 20$$

(A6) 使用したライブメディアの諸元を以下のようにする。

メディアソース：人の上半身の画像

(テレビ電話を想定)

画像サイズ：sub-QCIF (128 × 96 ピクセル)

ソース時間：120.767 s

3.4 メディア同期制御の仮定

メディア同期制御に関し、以下の仮定を置く。

(A1) 後方へのスライド制御において、しきい値 T_{h1} によるスライド時間 θ_1 は 50 ms、最小許容時間間隔 ω_1 は 1 s とする。この値は、スライド制御実行時に音声情報を巻き戻して出力する場合に、その音声が遅和感の少ないエコーとして聞こえるように選んだ[4]。特に θ_1 は、伝送遅延を小さくするために、蓄積メディアの場合 ($\theta_1 = 100$ ms) よりも小さい値を選択した。なおしきい値 T_{h1} は、蓄積メディアの場合と同じ値である 100 ms とする。また、スライド許容時間 κ は 250 ms とする。この値は静止衛星回線のラウンドトリップ伝送遅延(約 270 ms)を基準に定めた。

なお、Quick Recovery のしきい値 T_{h2} は ∞ 、ビデオの受信バッファ占有率による Gradual Recovery のしきい値 T_{h3} は ∞ とし、いずれも実行不可とする。

(A2) 前方へのスライド制御において、しきい値 T_{h0} によるスライド時間 θ_0 は -50 ms、最小許容時間間隔 ω_0 は 1 s、しきい値 T_{h0} は 100 ms とする。これらの値は、後方へのスライド制御の条件と対称になるように選定した。

(A3) 第 1 音声 MU の出力待ち時間は、 $r_1^{(2)} = 0$ s とする。

3.5 評価システム

図5に実際のH.263圧縮ビデオ情報とADPCM音声情報を使用する評価システムのソフトウェア構成を示す。H.263エンコーダは、シミュレーションプログラムか

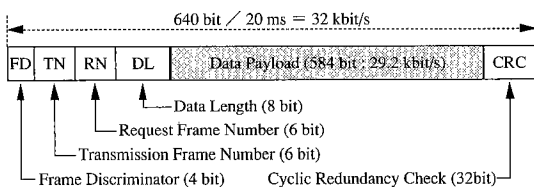


図4 ARQフレームの構成
Fig. 4 Structure of ARQ frame.

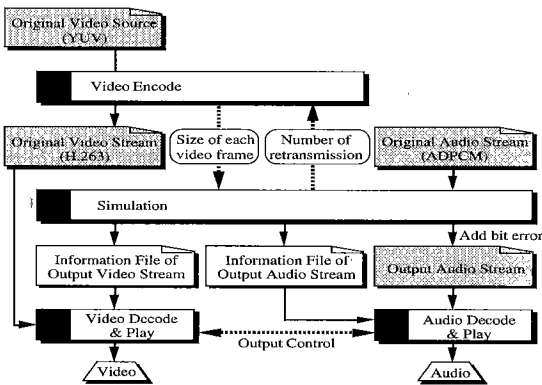


図5 評価システムの構成
Fig. 5 Block diagram of the evaluation system.

らフィードバックされる再送回数を利用して、空間的・時間的解像度制御を適用しつつ、ビデオソースからH.263ビデオ情報を生成する。一方、シミュレーションプログラムは、通信プロトコルやエラー環境等のシミュレーションを行い、メディア内、およびメディア間同期ずれ時間、再送回数等を求め、各メディアの出力情報ファイルを生成する。特に、ADPCM音声情報にはエラー環境に応じたビット誤りを付加する。そして各メディアは、これらの出力情報ファイルをもとに出力制御されて再生される。なお、H.263圧縮符号化と情報ファイルの生成、および復号化と再生には、H.263ソフトウェアエンコーダ/デコーダ[12]を一部改造して使用した。

4. 数値結果

ランダム誤り環境とバースト誤り環境の両方の通信環境におけるメディア同期性能を評価する。

4.1 ランダム誤り環境

ランダム誤り環境における評価結果を以下に示す。まず、SR-ARQによるビデオ通信チャネルのスループット特性は図6のとおりである。横軸はBERである。なおスループットは、単位時間当りに送信が完了した情報量をPHSの伝送容量(32 kbit/s)で正規化したものとする。

(1) メディア同期性能の評価

図7にメディア間同期ずれ時間 $\Delta_n^{(1-2)}$ の2乗平均値の平方根(RMS: Root Mean Square)、図8にビデオの平均MU出力レート、図9にビデオのメディア内同期ずれ時間 $\Delta_n^{(1)}$ のRMS、図10に音声のメディア内同期ずれ時間 $\Delta_n^{(2)}$ のRMSを、BERとの関係で示す。平均MU出力レートは、受信側で単位時間当りに出力したMU数の

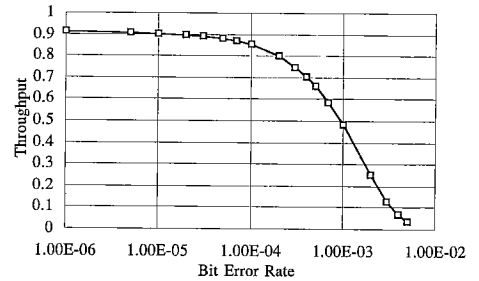


図6 スループット(ランダム誤り環境)
Fig. 6 Throughput in random error environment.

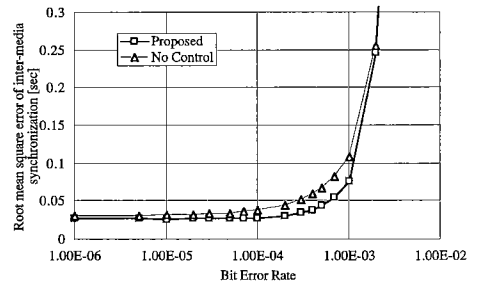


図7 $\Delta_n^{(1-2)}$ のRMS(ランダム誤り環境)
Fig. 7 Root mean square of $\Delta_n^{(1-2)}$.

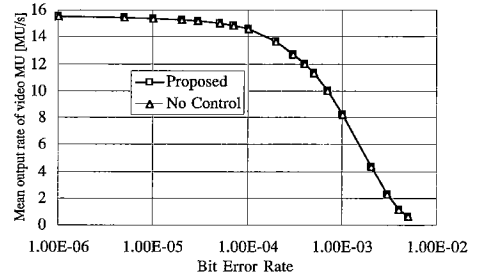


図8 ビデオの平均MU出力レート(ランダム誤り環境)
Fig. 8 Mean output rate of video MU.

平均値と定義する。また各図において、□印はスライド制御を適用した場合(Proposed)、△印は適用しない場合(No Control)を示す。但し、No Controlの場合でも、ビデオには目標出力時刻 $t_n^{(1)}$ と到着時刻 $A_n^{(1)}$ の関係に応じたメディア内同期制御を行うものとする。このとき $S_n^{(1)}$ は常に0である。また、ビデオにもメディア内同期制御を行わない方式も考えられるが、この方式は本論文のNo Controlよりも更に性能が悪くなる。

まず、ライブメディアの伝送においても、スライド制御がメディア同期に有効であるかどうかを調べる。図7によれば、スライド制御の適用によりメディア間同期が改善できることがわかる。

ここでメディア間同期性能を評価するために、文献[13]の結果を引用する。これは、メディア間同期ずれ時間が80 ms以内であればメディア間同期は良好な状態であり、160 ms 以上であれば同期外れの状態であると報告している。例えば、BERが 1.0×10^{-3} 以下において、スライド制御を適用すれば、メディア間同期ずれ時間のRMSを76 ms以下（No Controlでは109 ms以下）におさえられ、良好なメディア間同期を維持できる。このとき、ビデオの平均MU出力レート（図8）は8.3 MU/s以上を確保でき、十分動きのある画像を得ることができる。一方、BERが 2.0×10^{-3} 以上では、スライド制御を適用してもメディア間同期ずれ時間のRMSが246 ms以上となり、同期外れの状態となる。このとき、ビデオの平均MU出力レートは4.39 MU/s以下である。従って、BERが 2.0×10^{-3} 以上では、もはや同期性能を議論できる状況にないと言える。

なお、スライド制御を適用しない場合でも、BERが 5.0×10^{-4} という比較的悪い通信環境において、メディア間同期ずれ時間のRMSは68 ms以下であり、良好なメディア間同期を維持できる。これは、再送回数を考慮したビデオの時間的解像度制御により、あるビデオMUの伝送遅延の影響が次のビデオMUでは少なくなるよ

うにビデオフレームをスキップしたためである。

つぎに、メディア内同期の性能を調べる。図9によれば、スライド制御の適用によりビデオのメディア内同期も改善される。例えば、BER = 1.0×10^{-3} 以下において、ビデオのメディア内同期ずれ時間のRMSは78 ms以下（No Controlでは106 ms以下）である。

一方、音声のメディア内同期はスライド制御の適用により大きくなる（図10）。しかしながら、音声のメディア内同期ずれ時間のRMSは最大でも9 msであり、ビデオのメディア内同期ずれとと比較すれば十分小さな値である。なお、BERが 1.0×10^{-3} 以上において、音声のメディア内同期ずれ時間のRMSが減少するのは、後方へのスライド許容時間 κ によりスライド制御の発生が抑制されたためである。

(2) 平均MU遅延時間

つぎに、スライド制御が伝送遅延にどのような影響を及ぼすかを調べる。図11に音声の平均MU遅延時間、図12にビデオの平均MU遅延時間を、BERとの関係で示す。なお、平均MU遅延時間は、出力時刻 D_n^0 と発生時刻 T_n^0 の差の平均値と定義する。

図11によれば、音声の平均MU遅延時間は、スライド制御を適用しなければ、BERによらず常に一定の値

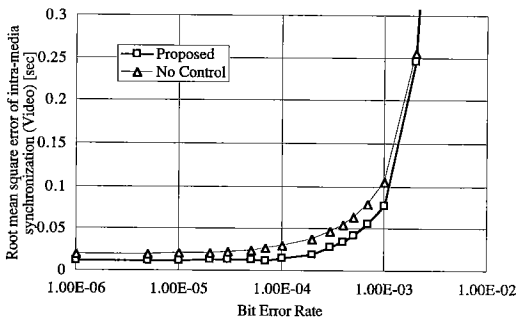


図9 ビデオの $\Delta_n^{(1)}$ のRMS(ランダム誤り環境)
Fig. 9 Root mean square of $\Delta_n^{(1)}$.

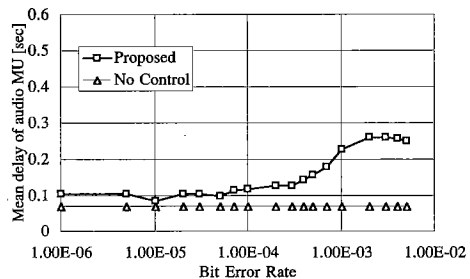


図11 音声の平均MU遅延時間(ランダム誤り環境)
Fig. 11 Mean delay of audio MU.

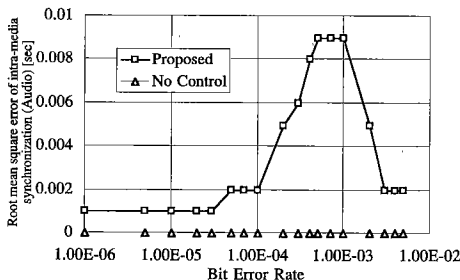


図10 音声の $\Delta_m^{(2)}$ のRMS(ランダム誤り環境)
Fig. 10 Root mean square of $\Delta_m^{(2)}$.

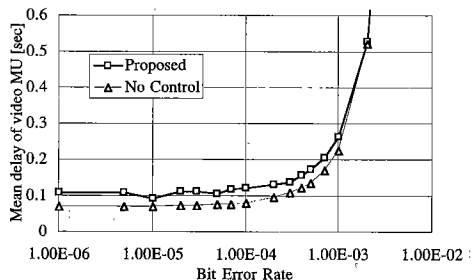


図12 ビデオの平均MU遅延時間(ランダム誤り環境)
Fig. 12 Mean delay of video MU.

(70 ms) である。しかし、スライド制御を適用すると、BERが大きくなるにつれて平均MU遅延時間も大きくなる。これは、BERの増加に伴い、再送によるビデオのメディア内同期ずれが頻繁に発生し、後方へのスライド制御も多く発生するためである。しかしながら、例えばBERが極端に悪い状況になったとしても、音声の平均MU遅延時間は、静止衛星回線のラウンドトリップ伝送遅延(約270 ms)にはほぼ等しい値である262 ms以下に抑制される。これは、後方へのスライド許容時間 κ (=250 ms)により、スライド制御の発生を制限したためである。

また、図12によれば、ビデオの平均MU遅延時間はスライド制御の適用により増加する。しかしながら、BERが 1.0×10^{-3} 以下において265 ms以下であり、音声の平均MU遅延時間と同程度におさえられる。一方、BERが 2.0×10^{-3} 以上では、スライド制御の適用の有無にかかわらずビデオの平均MU遅延時間は増大する。

(3) 主観評価

つぎに、スライド制御が音声とビデオに及ぼす影響を主観的に評価する。

まず、音声の品質について主観評価を行う。後方へのスライド制御時の音声は、スライド時間($\theta_1 = 50$ ms)に相当する量だけの音声情報を巻き戻して出力することで、違和感の少ないエコーとして聞こえる[4]。しかし、前方へのスライド制御時の音声は、スライド時間が $\theta_0 = 50$ msと小さいにもかかわらず、音飛びとして認識される。しかしながら、音声は無手順で伝送されるため、ビット誤りにより音声にノイズが付加される。このノイズと音飛びは主観的には区別が付きにくい。従って、前方へのスライド制御による音声品質の低下は、顕著には感じられない。なお、音飛びの影響を少なくするために、無音状態を検出し、無音時に前方へのスライド制御を行うことも考えられる。

つぎに、ビデオの品質について主観評価を行う。後方へのスライド制御によりビデオの出力はポーズする。また、前方へのスライド制御によりビデオの出力は早送りとなる。しかし、スライド時間が50 msと小さいため、主観的にはほとんど影響がない。

(4) 空間的解像度制御の画質への影響

更に、空間的解像度制御が画質に及ぼす影響を調べる。図13に量子化器情報(PQUANT: Quantizer Information for P-picture)の平均値、およびビデオMUサイズの平均値とBERの関係を示す。なお、PQUANTは、符号化時に各ビデオフレームの中に設定されており、量子化ステップサイズの半分の値(1から31)である。

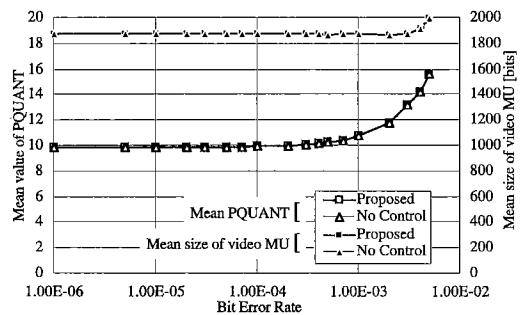


図13 量子化器情報の平均値とビデオMUサイズの平均値
Fig. 13 Mean value of PQUANT & mean size of video MU.

これによれば、BERが 1.0×10^{-3} 以下では、量子化器情報は10~11程度とほぼ一定であり、また、ビデオのMUサイズも目標MUサイズ(1,947 bits)によく制御されている。このとき、実際の画像を観察してみても、通信環境の悪化に伴う画質の劣化は少ない。しかしながら、BERが 2.0×10^{-3} 以上では、BERの悪化に伴い量子化器情報が急激に大きくなる。これは、平均MU出力レートの低下(4.4 MU/s以下: 図8)に伴いフレーム間の相関が少なくなり、目標MUサイズに近くなるように空間的解像度制御を行うには、量子化ステップサイズを大きくせざるをえなかったためである。

(5) 後方へのスライド許容時間 κ の影響

図14に音声の平均MU遅延時間、図15にメディア間同期ずれ時間 $\Delta_n^{(1-2)}$ のRMSを、後方へのスライド許容時間 κ との関係で示す。ここで、○, △, □印は、それぞれBERが 1.0×10^{-4} , 5.0×10^{-4} , 1.0×10^{-3} の場合を示す。なお、 $\kappa = 0$ msはNo Controlと等価である。

これによれば、メディア間同期と音声の平均MU遅延は、 κ をパラメータにしてトレードオフの関係にあることがわかる。すなわち、 κ を大きくすればメディア間同期ずれ時間のRMSは小さくなるが、音声の平均MU遅延が大きくなる。逆に、 κ を小さくすれば音声の平均MU遅延は小さくなるが、メディア間同期ずれ時間のRMSが大きくなる。この傾向は、BERが大きいほど顕著である。

ところで、メディア間同期を保証するか、音声のMU遅延を保証するかは、アプリケーションに応じて異なることが多い。例えば、リアルタイムの英会話レッスンのように、音声と口の動きを同期させる必要がある場合には、メディア間同期を保証する必要がある。また、活発なディスカッションを行う場合には、音声のMU遅延を重視する必要がある。また、同じアプリケーションであっても、状況によっては保証するものが変わることも

考えられる。

そこで、メディア間同期と音声のMU遅延の双方をどの程度保証するかを調整するために、後方へのスライド許容時間 κ を使用することを考える。例えば、BERが 1.0×10^{-3} 以下の通信環境において、良好なメディア間同期（80 ms 以下）を実現したければ、 κ を 250 ms 以上に設定すればよい。但し、音声の平均MU遅延は228 ms以上になる。また、音声の平均MU遅延を小さくしたければ、 κ を更に小さく、若しくは0に設定すればよい。但し、メディア間同期性能は低下する。このように、後方へのスライド許容時間 κ を使用すれば、アプリケー

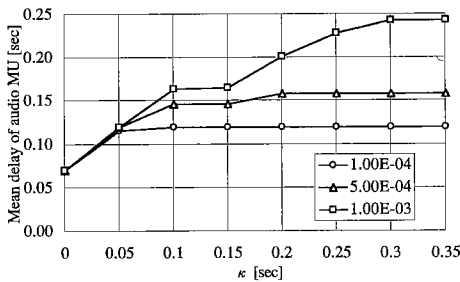


図 14 音声の平均 MU 遅延時間と κ の関係

Fig. 14 Mean delay of audio MU versus κ .

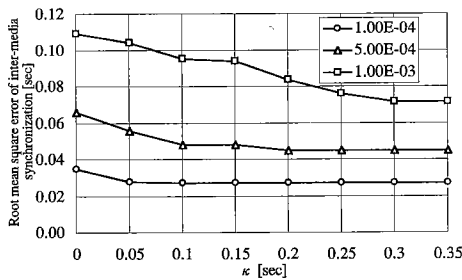


図 15 $\Delta_n^{(1,2)}$ の RMS と κ の関係

Fig. 15 Root mean square of $\Delta_n^{(1,2)}$ versus κ .

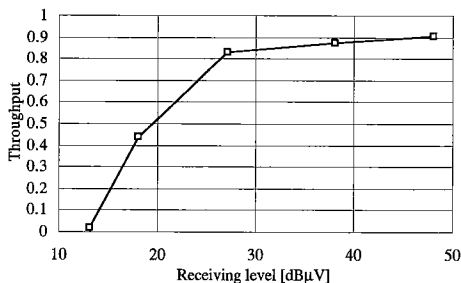


図 16 スループット(バースト誤り環境)

Fig. 16 Throughput in burst error environment.

ションや状況に応じて、メディア間同期と音声のMU遅延の性能をどの程度保証するかを調整できる。

4.2 バースト誤り環境

つぎにバースト誤り環境における評価結果を示す。まず、SR-ARQによるビデオ通信チャネルのスループット特性は図 16 のとおりである。横軸は受信レベルである。

(1) メディア同期と MU 遅延時間

図 17 にメディア間同期ずれ時間 $\Delta_n^{(1,2)}$ の RMS, 図 18 にビデオのメディア内同期ずれ時間 $\Delta_n^{(1)}$ の RMS, 図 19

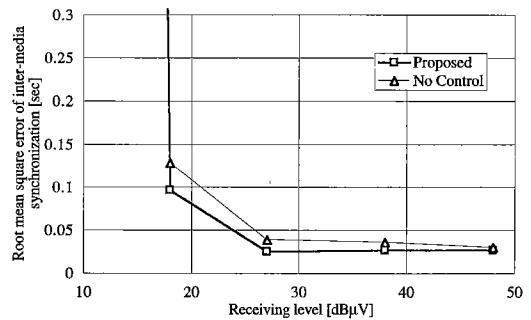


図 17 $\Delta_n^{(1,2)}$ の RMS (バースト誤り環境)

Fig. 17 Root mean square of $\Delta_n^{(1,2)}$.

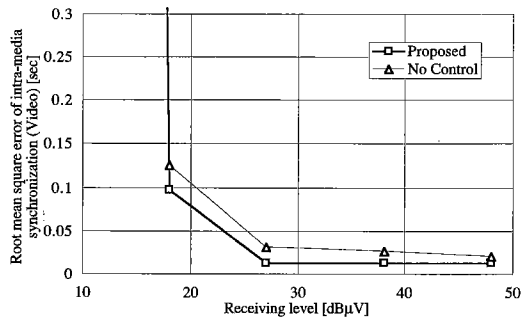


図 18 ビデオの $\Delta_n^{(1)}$ の RMS (バースト誤り環境)

Fig. 18 Root mean square of $\Delta_n^{(1)}$.

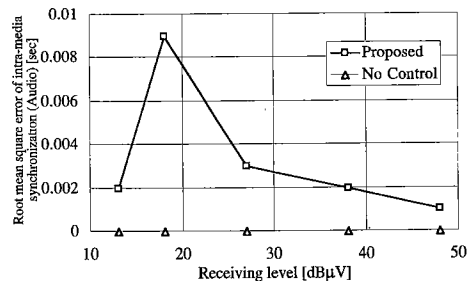


図 19 音声の $\Delta_m^{(2)}$ の RMS (バースト誤り環境)

Fig. 19 Root mean square of $\Delta_m^{(2)}$.

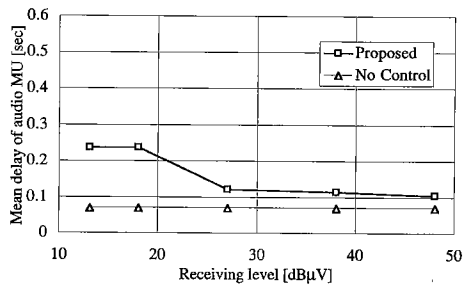


図 20 音声の平均 MU 遅延時間(バースト誤り環境)
Fig. 20 Mean delay of audio MU.

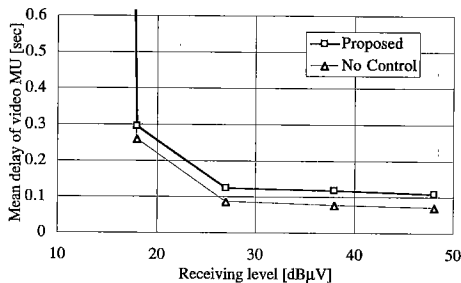


図 21 ビデオの平均 MU 遅延時間(バースト誤り環境)
Fig. 21 Mean delay of video MU.

に音声のメディア内同期ずれ時間 $\Delta_m^{(2)}$ の RMS, 図 20 に音声の平均 MU 遅延時間, 図 21 にビデオの平均 MU 遅延時間を, 受信レベルとの関係で示す。

これによれば, スライド制御はバースト誤り環境においてもメディア間同期 (図 17), およびビデオのメディア内同期 (図 18) に対して効果的であると言える。また, スライド制御は音声のメディア内同期を乱すものの, その値は十分小さい (図 19)。更に, ビデオや音声の平均 MU 遅延時間は増加する (図 20, 図 21)。しかしながら, 低受信レベルにおける音声の平均 MU 遅延時間の増加は, 後方へのスライド許容時間 κ により抑制できる。

(2) スライド制御の時間的経過

総スライド時間 $S_n^{(2)}$ は, 出力側での基準時刻の変化量を表しており, 音声の MU 遅延の状態を示す評価尺度として利用できる。そこで図 22 に, 受信状態が比較的悪い受信レベル 18 dB μ V のバースト誤り環境における総スライド時間 $S_n^{(2)}$ とビデオの MU サイズの時間的変化を示す。

通信開始直後, $S_n^{(2)}$ が 200 ms になるまで後方へのスライド制御が連続して発生している。これは, バースト誤りによる再送が頻繁に発生するだけでなく, 最初のビデオ MU である I フレームの MU サイズ (8,720 bits) が大きいためである。

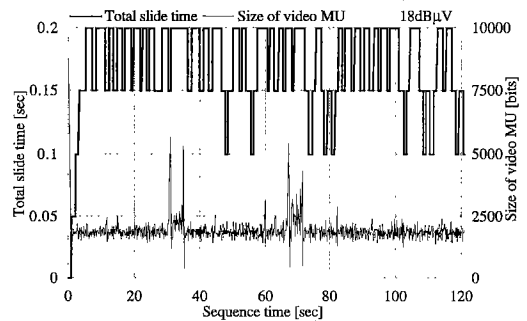


図 22 総スライド時間とビデオ MU サイズの時間的変化
Fig. 22 Progress of total slide time & size of video MU.

もし, 後方へのスライド制御だけを適用したならば, 音声の MU 遅延はその後大きな値を維持してしまう。しかし, 前方へのスライド制御の適用により, ネットワークの状態に応じて音声の MU 遅延を再び減少できる。例えば, 通信開始直後に 200 ms まで達した $S_n^{(2)}$ は, 100 ms まで減少することも数回発生している。

また, 通信開始から 30 ~ 35 秒経過時, および 65 ~ 70 秒経過時, ビデオの MU サイズが一時的に大きくなっている。これは, ビデオのシーンが大きく変化したためである。このとき, MU サイズの増加よりやや遅れて, $S_n^{(2)}$ も増加している。これは, ビデオの MU サイズの増大が, ビデオの MU 遅延の増加, およびメディア内同期ずれを引き起こし, そのビデオ MU を受信した後に後方へのスライド制御を実行したためである。

更に, ビデオの MU サイズが大きくないにもかかわらず, $S_n^{(2)}$ が増加している。これは, バースト誤りによるビデオの伝送遅延がメディア内同期を乱し, 後方へのスライド制御を実行したためである。すなわち, 通信環境が良好状態からバースト状態に移移すると, 後方へのスライド制御が実行され, メディア間同期を維持しようとする。逆に, 良好状態に移移すると, 前方へのスライド制御が実行され, メディア間同期を維持したまま MU 遅延を削減しようとする。

このように後方へのスライド制御は, ビデオの MU サイズの一時的な増加や再送が引き起こすメディア同期の乱れを吸収するのに有効である。更に, 前方へのスライド制御は, 後方へのスライド制御により増加した MU 遅延を削減するのに有効である。

5. む す び

本研究は, PHS の二つの通信チャネルを使用し, ライブの H.263 圧縮ビデオ情報を ARQ で, ADPCM 音声情報

を無手順で、リアルタイムに伝送しながら再生することを考えた。そして、ライブメディア同期制御方式と蓄積メディア同期制御方式の整合性を実現するために、蓄積メディアに適用したスライド制御を採用し、これにライブメディアの特性を考慮した拡張を行った。すなわち、前方へのスライド制御の追加、後方へのスライド制御の制限等の拡張を行った。更に、送信側において空間的・時間的解像度制御を適用した。そしてランダム誤りとバースト誤りの各環境下において、実際のビデオ/音声情報を使用したシミュレーションを行い、メディア同期や伝送遅延の性能評価、および主観評価を行った。その結果、次のような結論を得た。

拡張したスライド制御は、ランダム誤りとバースト誤りの両環境において、音声のメディア内同期を故意に乱すものの、ビデオのメディア内同期やメディア間同期の維持に有効である。また、ライブメディアの伝送で重要な評価基準である伝送遅延は、スライド制御の適用により大きくなる。しかしながら、後方へのスライド許容時間 κ を適切に設定すれば、音声の平均MU遅延時間の増加を抑制できる。更に、メディア間同期と音声のMU遅延はトレードオフの関係にあり、後方へのスライド許容時間 κ を使用すれば、どちらを保証するかを調整できる。

また、時間的解像度制御のみにおいて再送による影響を考慮すれば、通信環境の悪化に伴い、MU出力レートは低下するものの、量子化ステップの増加、すなわち画質の劣化を防ぐことができる。また、スライド制御が音声やビデオに及ぼす影響も、ほぼ良好であることを主観評価によって確認した。今後、提案方式を実装し、その性能を実験により確認する予定である。

謝辞 本研究は、一部、通信・放送機構 平成9年度「創造的情報通信技術研究開発推進制度」によって行われた。関係各位に感謝する。

文 献

- [1] 社団法人電波産業会, 第二世代コードレス電話システム標準規格 第三版 RCR STD-28, Nov. 1997.
- [2] PHSインターネット・アクセス・フォーラム, PHS Internet Access Forum Standard (PIAFS)仕様書, March 1997.
- [3] ITU-T Recommendation H.263, Video coding for low bit rate communication, March 1996.
- [4] 加藤正美, 白井憲義, 田坂修二, “PHSによる蓄積メディア伝送時のメディア同期性能評価,” 信学論(B-II), vol.J80-B-II, no.9, pp.749-759, Sept. 1997.
- [5] 的場直人, 近藤 靖, 山階正樹, “ARQを用いた無線動画通信の伝送遅延の改善法の検討,” 信学技報, RCS96-80, pp.95-101, Aug. 1996.
- [6] 加藤正美, 白井憲義, 田坂修二, “PHSにおけるビデオ伝

送方式の性能評価,” 信学論(B-II), vol.J79-B-II, no.10, pp.646-656, Oct. 1996.

- [7] Y.Ishibashi, S.Tasaka, and A.Tsuji, “Measured performance of live media synchronization mechanism in an ATM network,” Conf. Rec. IEEE ICC '96, pp.1348-1354, June 1996.
- [8] M.Kato, N.Utsui, and S.Tasaka, “Stored Media Synchronization Based on Buffer Occupancy in PHS,” Proc. IEEE PIMRC '97, pp.1049-1053, Sept. 1997.
- [9] Y.Ishibashi, E.Minami, and S.Tasaka, “Performance Measurement of a Stored Media Synchronization Mechanism: Graceful Recovery Scheme,” IEICE Trans. Commun., vol.E79-B, no.3, pp.399-411, March 1996.
- [10] ITU-T SG-15 Experts Group on Very Low Rate Visual Telephony, Video Codec Test Model, TMN5, Jan. 1995.
- [11] R.W.Lucky, J.Salz, and E.J.Weldon, Principles of data communication, McGraw-Hill, Inc., 1968.
- [12] Source program code of TMN H.263 Encoder ver.2.0 & Decoder ver.2.0, Telenor R&D, June 1996.
- [13] R.Steinmetz, “Human perception of jitter and media synchronization,” IEEE J. Sel. Areas in Commun., vol.14, no.1, pp.61-72, Jan. 1996.

(平成9年10月8日受付, 10年3月18日再受付)



加藤 正美 (正員)

昭59名古屋大・理・物理卒。同年三洋電機(株)入社。現在、同社 ハイパーメディア研究所 主任研究員。移动通信システム、無線・有線統合型マルチメディアネットワークの研究に従事。



白井 憲義 (正員)

昭46岐阜大・工・電気卒。同年三洋電機(株)入社。現在、同社 情報通信事業本部 情報通信開発研究所 部長。PBXをはじめ、移动通信システム、特にPHS公衆基地局の研究・開発に従事。



田坂 修二 (正員)

昭46名古屋大・工・電気卒。昭51東大大学院博士課程了。同年名古屋大・情報勤務、現在、同大・電気情報・教授。昭59～60 UCLA 客員研究員。工博。マルチメディア通信、無線ネットワークの研究に従事。IEEE, ACM, 情報処理学会各会員。