

# 多様な遠隔コラボレーションを実現する音声伝送システム

岸田 崇志<sup>†1</sup> 前田 香織<sup>†2</sup> 河野 英太郎<sup>†2</sup>  
近堂 徹<sup>†3</sup> 相原 玲二<sup>†4</sup>

広帯域ネットワークや QoS 保証のあるネットワークなど多様なネットワークが利用できるようになり、狭帯域ネットワークでは難しかったリアルタイム性の高い、また信頼性の高い音声伝送の実現が可能になってきた。そして、より品質の高い音声伝送への要求も高まっている。こうした変化にあわせ、音声伝送の利用場面も多岐にわたり、各場面で音声伝送システムに要求される条件も様々に異なる。本論文では、広帯域ネットワークで利用される様々な音声伝送場面をリアルタイム性、ロバスト性などの観点から分類した。これに基づき、各場面で要求度の高いパラメータを調整することにより、あらゆる利用場面に対応できる多目的な音声伝送システム “MRAT ( Multipurpose RAT )” の設計と開発を行った。MRAT は音声会議システムの 1 つである RAT ( Robust Audio Tool ) を広帯域ネットワーク上で使用することを想定して拡張し、低遅延音声伝送機能や Reed-Solomon 符号を用いた FEC によるエラー回復機構などを実装している。MRAT の評価としてはエンド-エンド間の低遅延化とロバスト性向上の検証を行い、その有用性を示す。また、MRAT の実証実験として 1 年間にわたって行った多地点での遠隔合唱や遠隔講義などについて記述し、多様な場面での実用性を示す。

## Development of an Audio Transmission System to Realize Various Scenes of Distance Collaboration

TAKASHI KISHIDA,<sup>†1</sup> KAORI MAEDA,<sup>†2</sup> EITARO KOHNO,<sup>†2</sup>  
TOHRU KONDO<sup>†3</sup> and REIJI AIBARA<sup>†4</sup>

It is becoming easy to use various networks such as broadband networks and QoS guaranteed networks. Also, it is possible to realize high robustness and high real-time audio transmission. Moreover, high quality audio transmission is required. The most required condition for audio transmission differs depending on various scenes. In our study, we classify conditions of audio communications required in various situations from the viewpoints like real-time and robustness and design a multipurpose audio transmission system called as “MRAT (Multipurpose RAT)”. We enhance RAT (Robust Audio Tool) that is one of audio communication tools, and add two new options. One is to keep shorter delays between end-nodes. Another is an error recovery using a Forward Error Correction with Reed-Solomon code to recover packet losses. By using these options, MRAT has three modes; a Chorus mode, a Conversation mode, and a Broadcast mode. Therefore, MRAT can correspond to various scenes. We show some evaluations of MRAT like measurements of end-to-end delays and effects of an error recovery. We also show practicality of MRAT by practical experiments such as distance chorus at multi-locations and distance lectures.

### 1. はじめに

従来、狭帯域ネットワーク上での音声伝送システムにはいかに狭帯域で品質の良い音声を送るかが議論の焦点になり、符号化や圧縮の効率化が求められてきた<sup>1)</sup>。しかし、近年は ADSL や光ファイバなどにより急速に広帯域ネットワークの普及が進んでおり、さらに QoS 保証のあるネットワークの利用も可能である。このような通信環境を用いて、遠隔会議、遠隔講義や遠隔講演など音声や映像を用いたマルチメディア通信

<sup>†1</sup> 広島市立大学大学院情報科学研究科  
Graduate School of Information Sciences, Hiroshima City University

<sup>†2</sup> 広島市立大学情報処理センター  
Information Processing Center, Hiroshima City University

<sup>†3</sup> 広島大学大学院工学研究科  
Graduate School of Engineering, Hiroshima University

<sup>†4</sup> 広島大学情報メディア教育研究センター  
Informatin Media Center, Hiroshima University

表 1 音声伝送利用場面の分類

Table 1 Classification of audio communication scenes.

利用場面	遠隔合唱	遠隔会議	遠隔講演
方向性	双方向	双方向	1 方向
許容遅延時間	100 ms 以下	400 ms 以下	任意
リアルタイム性	高	中	低
ロバスト性	低	中	高
主たる要求条件	拠点間での同期	会話に支障がないこと	伝送内容の信頼性

も一般的になってきている<sup>2),3)</sup>。こうしたネットワーク仕様の変化にとともに、単に音声を送るだけではなく、高いリアルタイム性や高信頼性を持つ音声伝送システムの開発が可能となってきた。また、従来実現が難しかった遠隔地間での合唱などリアルタイム性の高い遠隔コラボレーションも想定される。

こうした背景から、多様化する利用場面の中でそれぞれの音声伝送場面の要求条件に関する考察が重要と考えた。たとえば、遠隔合唱の場合、高品質かつ低遅延なリアルタイム性の高い音声伝送が必要となり、遠隔講演の場合は、伝送内容の信頼性が要求される。本論文では、様々なネットワーク仕様を想定し、音声伝送の利用場面をリアルタイム性やロバスト性(パケット損失に対する音声の劣化抑制)などの観点から新たに分類した。

一方、ロバスト性向上を目的とした音声伝送システムの1つにRAT(Robust Audio Tool)<sup>4)</sup>がある。RATは他の音声会議システムで装備していない多くのコーデックを持ち、狭帯域ネットワークでも使用できるようなエラー回復機構を持つなど、インターネット上の音声伝送システムとして有用である。しかし、狭帯域ネットワークでの利用を想定しているため、エラー回復機能が十分ではない。また、低遅延化も1997年頃のコンピュータのCPU性能を想定した実装で、さらなる改善が必要である。

そこで広帯域ネットワーク下で使用することを想定した利用場面の分類に基づいて、あらゆる利用場面に対応できる多目的音声伝送システムMRAT(Multipurpose RAT)<sup>5)</sup>の開発を行った。本論文の目的は、広帯域化するネットワークで必要とされる様々な利用場面について考察し分類すること、また、多様な利用場面を実現する音声伝送アプリケーションの必要性を主張するとともに、開発した音声伝送システムの有効性を示すことである。文献5)では、MRATの設計方針と実装の一部について記述した。本論文では、文献5)での設計に基づき実装を完成させ、評価を行った。さらにMRATを用いた数々の実証実験について論述する。

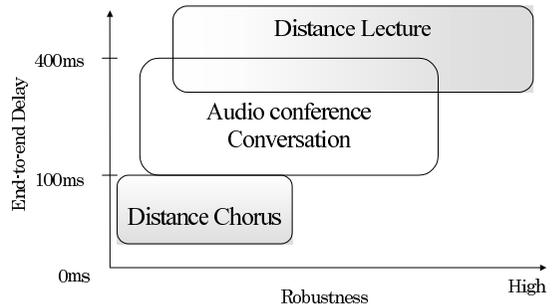


図 1 音声伝送利用場面とその要求条件

Fig. 1 Audio communication scenes and those requirements.

本論文では、まず2章で様々な利用場面における音声伝送に対する要求の分類について述べ、次に3章で多目的音声伝送システムの実装について詳述する。さらに、4章で評価と実証実験からMRATの有用性を示し、5・6章で今後の課題とまとめを述べる。

## 2. 音声伝送の利用場面とその要求条件

多様な音声伝送の利用場面に対応するために、我々は音声伝送の利用場面をリアルタイム性とロバスト性に着目し、典型的な利用場面として遠隔合唱、遠隔会議、遠隔講演という3つに分類した。表1は典型的な音声伝送システムの利用場面とその特徴を示し、図1は遅延時間とロバスト性の観点から図示したものである。図1のx軸はロバスト性の高さを示し、y軸は遅延時間を示す。片方向の遅延時間はITU-T G.114<sup>6)</sup>で1方向の伝送時間のガイドラインが定められているためそれを参考にした。

遠隔会議やゼミ形式の講義では双方向の会話(議論)が多く、主な要求条件は会話に支障のないことである。円滑なコミュニケーションを行うためには、遅延時間が重要な要因となる。そのため、この場面ではITU-T G.114を参考に遅延時間を400ms以下に定めた。ロバスト性は、その遅延時間を満たす範囲でより高いものが望まれる。

遠隔講演の場合、一方的に音声を流す時間が多く、双方向での会話がほとんどない。そのため、多少の遅

表 2 MRAT の特徴  
Table 2 The feature of MRAT.

	MRAT		RAT	VAT
	Chorus モード	Broadcast モード		
低遅延音声伝送		×	×	×
パケット損失耐性	×			×
バースト損失耐性	×		×	×

有利, やや有利, ×不利

延はあっても話し手の音声がかえれないということは避けたいので、主にロバスト性の高さが要求される。

遠隔合唱の場合、通信の方向性は双方向で、合唱の特性上拠点間で音声のタイミングがずれることは許されない。このため、最も重要となる条件はリアルタイム性である。実際に小学校の音楽の教師がネットワーク上で合唱を行う実験を幾度か繰り返した結果、100 ms を超える遅延では、互いの声のずれを感じ、合唱として成立しなかったため、遠隔合唱に必要な遅延時間を 100 ms 以下と定めた。ここで 100 ms 以下という遅延時間が妥当かどうか別の観点から考察する。摂氏 20°C の場合の音速を秒速 343.5 m とするとき、100 ms で音は約 34.3 m 進むので、34 m の舞台上で合唱を行う場合、ステージの両端では 100 ms の遅延時間が生じていることになる。そこで 100 ms という遅延時間は遠隔合唱の許容遅延時間の目安として妥当であると考えた。

### 3. Multipurpose RAT ( MRAT )

#### 3.1 特 徴

MRAT は様々な利用場面に対応するために開発した多目的音声伝送システムである。一般的にインターネットはベストエフォート型の通信ネットワークであるが、MRAT では利用場面にあわせ、低遅延化とロバスト性の高い音声伝送を実現することができる。MRAT は、遠隔合唱や遠隔合奏に用いる“Chorus モード”、遠隔会議や会話に用いる“Conversation モード”、遠隔講義や遠隔講演に用いる“Broadcast モード”の 3 つのモードを持っており、表 1 に示した様々な音声伝送の利用場面に対応することができる。表 2 にマルチキャストで音声伝送可能な 3 つのシステム、VAT ( Visual Audio Tool <sup>7)</sup>、RAT、MRAT の特徴を比較した。MRAT の特徴の 1 つである低遅延伝送モードの実現にはノンブロッキングリードを用いたバッファリング時間の制御技術を用い、Broadcast モードでの高いロバスト性の実現には Reed-Solomon 符号を用いた FEC ( Forward Error Correction ) を実装している。

図 2 に、MRAT のユーザインタフェースを示す。こ



図 2 MRAT の操作画面

Fig. 2 Overview of a control window of MRAT.

れは RAT とほぼ同様で、この操作画面上で同じセッションに参加している参加者を確認できる。MRAT は RTP 上で UDP/IP をトランスポートプロトコルとして用いており、さらにマルチキャストによる通信も可能である。Conversation モードは従来の RAT の機能を用いるので、本研究で追加した Chorus モードと Broadcast モードの特徴と技術を以降で詳述する。

#### 3.2 Chorus モード

Chorus モードはリアルタイム性に重点を置いたモードで遅延を最小限にすることを優先している。遅延は音声のサンプリングやエンコード処理、また、伝送によっても生じる。遠隔合唱の場合、低遅延化が必要で、2 章で述べたように Chorus モードでは許容遅延時間を約 100 ms 以内に定義した。遅延時間はオーディオデバイスバッファに關係する。

RAT にはオーディオデバイスのバッファリングによる遅延時間を小さくするためにそれを制御するアルゴリズムが実装されている<sup>8)</sup>。このアルゴリズムは *cushion* という変数を持ち、*cushion\_max* と *cushion\_min* という上限値と下限値の間に入っている場合は、その *cushion* 値をそのときに必要なバッファリング時間として処理する。上下限値に入っていない場合は上限値または、下限値がバッファリング時間と設定される。そのため、これらの上限値や下限値の設定が不適切な場合は不要な長さのバッファリング時間を設定して遅延時間が大きくなったり、処理に間にあわないくらい短いバッファリング時間を設定して音が途切れることになる。RAT では、ノンブロッキングリードを用いて

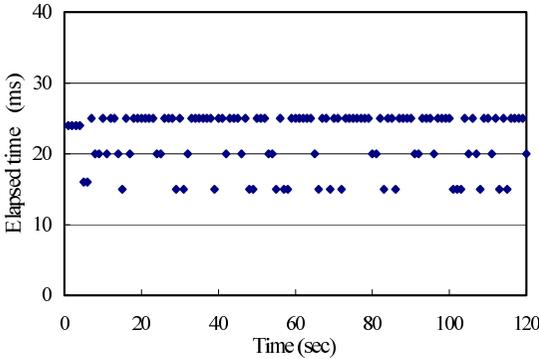


図 3 音声データ読み込み時間の推移  
Fig. 3 Read length variation.

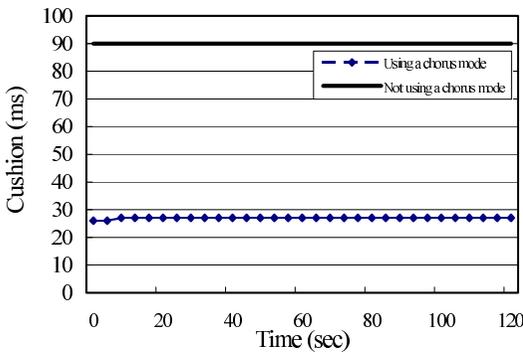


図 4 cushion の時間的推移  
Fig. 4 Evaluation of cushion.

オーディオデバイスバッファに蓄積された音声データの量を計ることで音声データが取得されてからどれだけ経過したかという時間を見積もる。この経過時間が長いほど未処理のデータがバッファに残る時間が長く、他の処理で CPU が使われていることになる。この経過時間が *cushion\_min*, *cushion\_max* の設定に反映されている。

現行の RAT では、1997 年当時の SUN Sparc 10 ワークステーションで測定された経過時間の結果を基に *cushion\_max* は 250 ms に、*cushion\_min* は 40 ms に設定されている。図 3 は CPU が Pentium III 1 GHz、メモリが 512 MB のコンピュータ上で前述の経過時間を 2 分間測定したものである。経過時間は 15 ms から 26 ms の範囲であることが分かる。近年の CPU で再測定した図 3 の結果より、MRAT では *cushion\_max* と *cushion\_min* の値をそれぞれ 40 ms と 0 ms に修正した。図 4 は Pentium III 1 GHz のコンピュータ上で 2 分間 *cushion* の測定を行ったもので、*cushion* 値は、90 ms であったものが約 26 ms となっていることが分かる。MRAT の Chorus モードでは、バッファリ

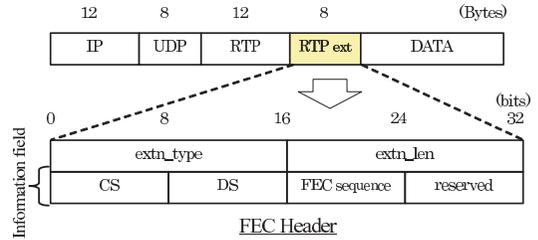


図 5 MRAT ( Broadcast モード ) のパケットフォーマット  
Fig. 5 Packet format of MRAT (Broadcast mode).

ング時間を近年の CPU に適応させ過剰なバッファ時間を抑えることで低遅延化が達成できた。RAT のノンブロッキングリードを使用する限り、CPU の性能が向上すれば同様の修正をすることで遅延をさらに短縮できる可能性がある。しかし、サウンドカードなどのハードウェア内での遅延時間もあるため遅延時間を完全に 0 にすることはできない。

### 3.3 Broadcast モード

Broadcast モードは高いロバスト性を実現するためのモードである。RAT には Lecture モードというオプションが指定でき、このオプションを指定することで、通常より少し多めにジッタ緩衝バッファなどのバッファ時間を見積もる。また、Media-Specific FEC (Forward Error Correction<sup>9)</sup> のオプションを指定することでエラー回復が図られる。しかし、現行の RAT では 3 つ以上の連続したパケット損失の回復ができないため、インターネットで生じるバースト損失(連続してパケットが損失すること)には対応できない。そこでよりロバスト性を高めるために MRAT の Broadcast モードはバースト損失に強い Reed-Solomon 符号を用いた FEC を実装した。

Reed-Solomon 符号はブロック符号の 1 つであり、送信すべき情報を  $M$  個の連続した複数ビット単位(シンボル)に分割し、このシンボルを単位として符号化、複合化を行う誤り訂正符号である。1 ブロックは  $N(N = 2^M - 1)$  個のシンボルによって構成される。そのうち  $K$  個がデータシンボルであり、 $N - K$  個が冗長シンボルである場合、 $(N, K)$  の冗長度を持つといい、 $(N, K)$  の冗長度を持つ Reed-Solomon 符号を  $RS(N, K)$  符号と呼ぶ。MRAT では、冗長度を増やすと処理遅延も増すため、 $N$  を 15 として実装している。

実装には RTP 拡張ヘッダを FEC ヘッダとして定義した。図 5 に MRAT ( Broadcast モード ) のパケットフォーマットを示す。extn\_type, extn\_len フィールドは RTP 拡張で定義された固定フィールドで、extn\_type には 1 を使用している。これに CS ( code

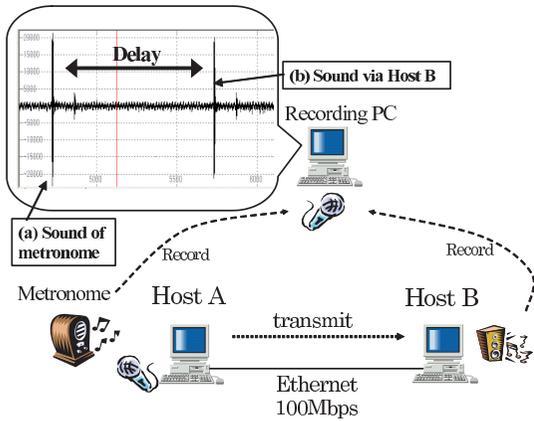


図 6 遅延時間測定の実験環境

Fig. 6 Experimental environment on delay measurements.

symbol size), DS (number of data symbols), FEC シーケンスフィールドを追加して FEC ヘッダとしている. CS フィールドはコードシンボル数を示し, DS フィールドはデータシンボル数を表す. FEC シーケンスフィールドは FEC の 1 ブロック中のシーケンス番号を表すフィールドである.

4. 評価

4.1 遅延時間

4.1.1 処理遅延時間の測定

各モードでの遅延時間を測定した. 実験環境を図 6 に示す. Host A の仕様は CPU が Pentium III 600 MHz でメモリが 384 MB, OS は Vine Linux 2.5, Host B の仕様は CPU が Pentium III 1 GHz でメモリが 512 MB, OS は Vine Linux 2.1 である. Host A で 1 秒間に 1 回のテンポでメトロノームの音を入力し, その音を Host B で受信しスピーカから出力する. 録音用 PC では, メトロノームの音とネットワークを経て Host B 経由で戻ってきた音を録音し波形の時間軸のずれを比較し測定した. 測定結果を表 3 に示す. Broadcast モードは冗長度 (15,11), (15,12), (15,13) それぞれに対して遅延時間を測定した. 表 3 は 10 回測定した遅延時間の平均を表示している. この測定は同一 LAN 内で行っており同一 LAN 内のパケット転送遅延は 0.1 ms 程度であることを確認しているため, 表 3 の遅延時間は処理遅延と見なせる.

表 3 より, Conversation モードは片側の遅延時間が 132 ms, Chorus モードは 72 ms となり, 表 1 で定義した条件を満たしている. Broadcast モードでは最も冗長度が高い場合でも表 1 で定義された値より十分に小さくなっており, 遠隔会議で定義した遅延時間よ

表 3 遅延時間の測定結果  
Table 3 Measurement of delays.

モード	遅延時間 (ms)
Broadcast (15,11)	143
Broadcast (15,12)	138
Broadcast (15,13)	138
Conversation	132
Chorus	72

りも小さく収まっている. ゆえに, Broadcast モードは遠隔会議においてもよりロバスト性の高いモードとして使うことが可能である.

4.1.2 転送遅延

ここでは, 転送遅延が MRAT に与える影響について考察する. MRAT では送信側から音声データを RTP/UDP を用いて送信しており, 受信側から送信側に通信速度などに関する制御を行っていない. そのため, 転送遅延が大きくなると転送遅延時間が音声の再生遅延時間に上乗せされるだけである. ただし, MRAT では大きなジッタはパケット損失として処理される点に注意する必要がある. このことを示すために, 遅延が 15 ~ 20 ms 程度 (RTCP による統計情報) のインターネット接続環境を用いて同様な測定を行ったところ, 再生時間が約 90 ms となり, 転送遅延分が上乗せされていることを確認している (20 回の測定の平均値). また, 4.3.1 項で示す実証実験の環境では 10 Mbps の広域イーサネットを用い, 転送遅延は 2 ~ 4 ms で, 音声再生までの時間は 70 ~ 75 ms であった.

当然, 転送遅延が大きいくところでは, MRAT で想定しているすべての利用場面を実現することはできない. 本研究で遠隔合唱を行うためには 100 ms 以下を許容遅延時間と仮定しているため, 転送遅延 30 ms 以下を実現できるネットワークでなければならない.

4.2 Broadcast モードでの FEC の性能評価

4.2.1 Reed-Solomon FEC の性能評価

Reed-Solomon 符号を用いた FEC を使用した場合のパケット損失回復性能の評価を行った. 測定環境を図 7 に示す. 測定方法は MRAT (Broadcast モード) を使用し Host A からパケット損失発生役の PC (Loss generator)<sup>10)</sup> 経由で Host B へ伝送する. Loss generator は, 疑似乱数 (BSD4.3 準拠の random 関数を使用) を用いて任意の確率でパケット独立にパケット損失を発生させる. この測定に関して, パケット損失に偏りがなく収束するまで十分な時間 (1 時間程度) 約 10 万パケットを採取し測定を行った. その際, Loss generator では 1, 2, 4, 6, 8, 10% の確率でパケット損失を発生させた. 冗長度 (N, K) におい

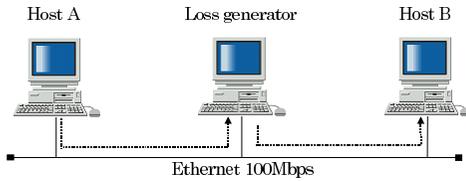


図 7 FEC 性能評価の実験構成図

Fig. 7 Experimental environment on FEC Performance.

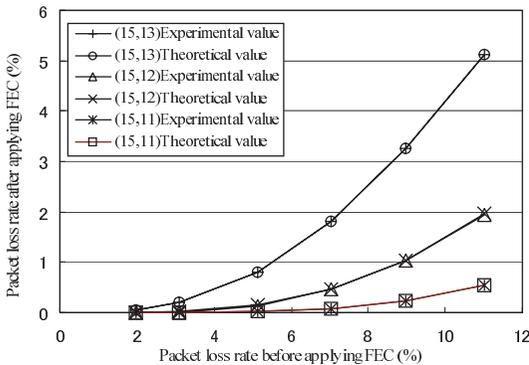


図 8 Broadcast モードでの FEC 性能評価

Fig. 8 FEC performance measurement of Broadcast mode.

て  $N$  は 15 固定とし、 $K$  を 11, 12, 13 と変化させ、それぞれの場合について FEC 復元後のパケット損失率の実測値を測定し、FEC 復元後のパケット損失率の理論値と比較した。FEC 復元後のパケット損失率  $P$  は式 (1) を用いた<sup>12)</sup>。式中の  $p$  は各パケットの損失率で、 $i$  はデータパケット  $K$  個中 FEC 復元後に損失しているパケット数である。パケット損失率の実測値の測定結果とともに式 (1) から算出したそれぞれの場合における理論値を図 8 に示す。 $x$  軸は FEC 復元前のパケット損失率、 $y$  軸は FEC 復元後のパケット損失率である。図 8 に示したように Reed-Solomon 符号を用いた FEC を用いることによってパケット損失率を少なく抑えることができた。また、理論値と実測値はほぼ一致しており理論値どおりの性能がでていることが分かる。

$$P = \frac{1}{K} \sum_{i=0}^k i_k C_i p^i (1-p)^{k-i} \quad (1)$$

$$\sum_{j=t}^{N-K} N-K C_j p^j (1-p)^{N-K-j}$$

$$t = \begin{cases} 0 & \text{if } N-K < i \leq K \\ N-K-i+1 & \text{if } 0 \leq i \leq N-K \end{cases}$$

現行の RAT では、ジッタに関しては 375 ms 以上の

ジッタが発生した場合はパケット損失と見なされるので、ジッタが大きいネットワークでも Broadcast モードの FEC は有効である。

#### 4.2.2 Reed-Solomon FEC と Media-Specific FEC との比較

今回、実装した Reed-Solomon FEC (以降、RS-FEC) と RAT に従来実装されていた Media-Specific FEC (以降、MS-FEC) のパケット損失回復性能を比較する。MS-FEC は RFC2198<sup>11)</sup> で提案されている FEC である。MS-FEC は、送信ホストでは送信前に音声データのペイロード部を複製し次のパケットのペイロードに追加しピギーバックする。その際、元の音声データと複製された音声データそれぞれに個々のエンコード方式を選択することができる。受信ホストでは、損失したパケットがあれば後方のパケットにピギーバックされていたものから復元する。MS-FEC では冗長データをピギーバックで送信するためバースト損失については対処できない。RS-FEC と違い冗長パケットを生成する際に複雑な演算を用いないシンプルな FEC であり、エンコードやデコード処理の負担が少ないという特徴がある。

次に、MS-FEC の FEC 復元後のパケット損失率の理論値を導出する。各パケットの損失率を  $p$  とすると MS-FEC ではオリジナルとピギーバックの 2 つのパケットが両方損失した場合、1 パケットは回復できなくなる。そのため、復元後の損失率の理論値  $P$  は式 (2) のようになる。

$$P = p^2 \quad (2)$$

それぞれの FEC での FEC 復元後のパケット損失率の理論値を比較したものを図 9 に示す。ほぼ同帯域を使用する RS(15, 8) 符号と MS-FEC では、パケット損失が 10% 生じている場合、MS-FEC による復元後のパケット損失率に対して、RS-FEC では非常に小さなパケット損失率に抑えることができる。RS(15, 11) 符号では、使用帯域が MS-FEC より 26.7% 少なく、さらにパケット損失 10% 下では、約 2.7 倍パケット回復率が高まっている。また、MS-FEC では連続する 2 個のパケット損失しか回復できないが、RS(15, 11) 符号では連続する 4 個までのバースト損失の回復が可能で、より高いロバスト性が実現できていることが分かる。

一方、MS-FEC は冗長部 (ピギーバック部分) の圧縮形式によって、転送するデータ量が変化する。ピギーバック部分がオリジナルの圧縮形式と同じ場合は 2 倍になり、オリジナルより高い圧縮率の形式だとそれによって冗長データ量が決まる。たとえば、オリジ

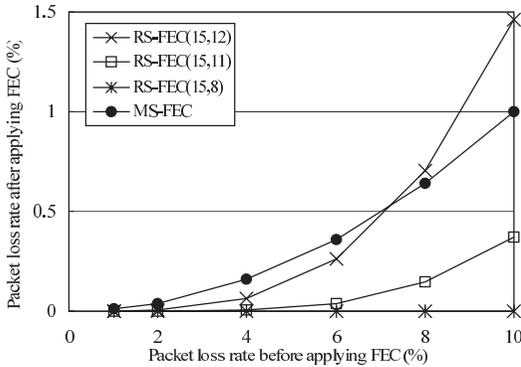


図 9 Reed-Solomon FEC と Media-Specific FEC の比較  
Fig. 9 Compare FEC performance.

ナルのペイロードの圧縮形式が 16 bit Linear PCM で冗長部の圧縮形式が G.726 ADPCM の場合は転送データ量は約 1.25 倍となる。ただし、この場合受信側でオリジナルと同じ品質が再生されない。RS-FEC では冗長度 (N,K) によって、転送データ量が変化する。たとえば、RS(15,11) の場合、FEC なしに比べて約 1.36 倍の転送量となる。RS-FEC の場合、再生品質は送信側と同じである。

4.3 実証実験

MRAT を用いて 1 年間にわたり様々な実証実験を行った。Chorus モードの実証実験として広島市内の学校で 2 地点間での遠隔合唱と遠隔合奏、3 地点間での遠隔合唱を行った。さらに、佐賀の小学校と広島市の小学校での 2 地点間での遠隔合唱も行っている。

Broadcast モードの実証実験としては、佐賀大学と広島大学、広島市立大学の 3 地点間で 28 回にわたる遠隔ゼミで使用し、さらに、広島市内の高大連携の遠隔授業で 4 回使用した。そのうちの 2 地点間での遠隔合唱、3 地点間による遠隔ゼミを紹介する。

4.3.1 遠隔合唱

Chorus モードを用いて「マメ de がんすプロジェクト」<sup>13)</sup> の実証実験ネットワークを用いていくつかの学校間で遠隔合唱や遠隔合奏を行った。そのうち 2 小学校間で行った遠隔合唱について詳述する。実験構成図を図 10 に示す。実験ネットワークは 10 Mbps の広域イーサネット、各拠点間の転送遅延は 2 ms 程度であった。各合唱地点 (白島小, 南観音小) は伴奏を聞きながらそれぞれアルトパートとソプラノパートに分かれ、それぞれ相手の歌声をネットワーク経由で聞きながら合唱した。伴奏は合唱をしない第 3 の地点 (広島市立大学) から送信した。音声のエンコーディング方式として 16 bit linear PCM を用い、サンプリ

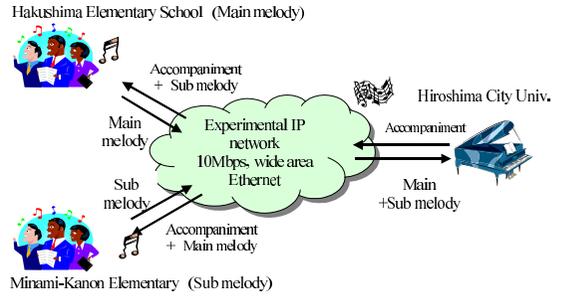


図 10 遠隔合唱の実験構成図  
Fig. 10 System configuration of distance chorus.

ング周波数は 32 kHz、チャンネルはモノラルで送信を行った。そのため使用する帯域は 1 拠点 512 kbps であった。2 拠点間は双方で TV 品質の動画を送信することができ (mpeg2ts<sup>10),12)</sup> 使用), 児童はスクリーンに投影されたパートナの姿を見ながら合唱を行った。RAT では、片側の遅延が大きく合唱を行えなかったが、MRAT を用いることで遅延は片側 70 ~ 75 ms 程度となり、わずかなエコーを感じる程度に収まった。このときのネットワークのジッタは約 7 ms (RTCP の統計情報による表示) であった。

4.3.2 遠隔ゼミ

Broadcast モードの実証実験として 2002 年 4 月 ~ 12 月までの期間、佐賀大学、広島大学、広島市立大学間で週に 1 回の遠隔ゼミを行った。遠隔ゼミでは、実験ネットワークとして JGN<sup>14)</sup> (ATM45 Mbps) を用いた。映像伝送には mpeg2ts を用い、音声伝送には MRAT (Broadcast モード) を用いた。エンコード方式は G.726 であり冗長度は (15,11) を用いたので送信帯域は各拠点 160 kbps となる。システム構成図を図 11 に示す。当初は Conversation モードを使用していたが、Conversation モードではパケット損失によりノイズが頻繁に生じたため、より良い音質を確保するため 9 月からは Broadcast モードを使用した。

図 12 にゼミで使用したネットワークで 1 時間計測を行った際 (広島市立大学—広島大学間) の FEC 適用前と FEC 適用後のパケット損失率を示す。FEC 適用前のものは従来の RAT、FEC 適用後のものは Broadcast モードを使用したものである。x 軸は時間を示し、y 軸はパケット損失率を示しており、パケット損失率は 30 秒ごとに集計した平均値である。図 12 より、FEC 適用前では、1 時間の平均のパケット損失率は 0.120% であったが、FEC を適用することでパケット損失率は 0.019% と非常に小さく抑えられたことが分かる。

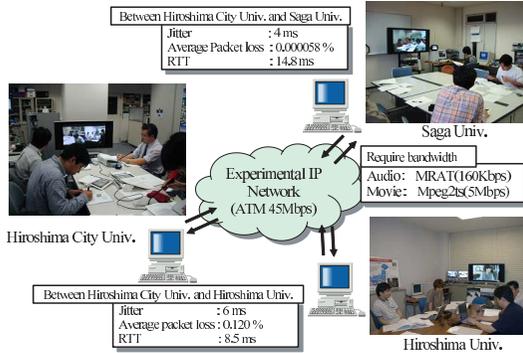


図 11 遠隔ゼミの実験構成図

Fig. 11 System configuration of distance seminar.

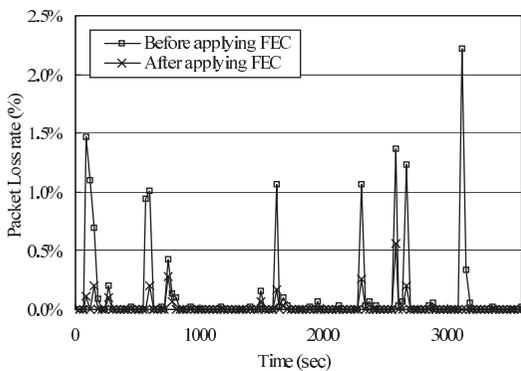


図 12 パケット損失率の時間的推移

Fig. 12 Error recovery of packet losses.

## 5. 考 察

ここでは関連技術との比較をまとめ、開発したシステムの問題点について考察する。

音声伝送システムに関する技術との比較は、3.1 節の VAT, RAT, MRAT の特徴比較のとおりである。音楽情報を送信する同様な技術として RMCP (Remote Music Control Protocol)<sup>3)</sup> が提案されている。RMCP は、MIDI (Musical Instruments Digital Interface) と LAN を融合することにより、LAN に接続されたコンピュータを通じ MIDI 楽器をコントロールし、遠隔地間で合奏や演奏などを行うために提案されたプロトコルである。このプロトコルは音楽データを符号化することが前提で、MRAT とはアプローチが異なる。RMCP では符号化されない人間の声は送ることができず合唱は行えない。

次に、MRAT のパケット損失耐性で用いた技術に関しては 4.2.2 項で述べたように MS-FEC との比較を行った。これにより、MRAT はバースト損失への耐性が高いことを示した。しかし、RS-FEC を用いる

MRAT でも図 12 のように大きなバースト損失が生じる場合はそれらをすべて回復できるわけではない。今後、使用するネットワークの特性によって、冗長度を動的に変更できる FEC などを検討していきたい。

その他の問題点として、RS-FEC は計算量も多く処理負荷も多いため、ロバスト性を高めると遅延時間が増加することがあげられる。遅延時間短縮とロバスト性強化にはトレードオフの関係があるため MRAT では Chorus モードはロバスト性は低く Broadcast モードでは Chorus モードほどの低遅延は望めない。また、Chorus モードを転送遅延の大きなネットワークで利用することは可能だが、4.1.2 項で述べたように、遠隔合唱を行うには転送遅延の上限がある。

## 6. おわりに

我々は音声伝送の利用場面の要求条件について考察を行い、リアルタイム性とロバスト性の観点から大きく 3 つに分類した。そして、それらをもとにあらゆる利用場面に対応できる多目的な音声伝送システム MRAT の開発を行った。また、MRAT の評価を行い、多様な利用場面での有用性を示した。さらに、遠隔合唱や遠隔合奏、そして遠隔ゼミを行い広帯域ネットワーク上で多様な遠隔コラボレーションが実現できることを示した。

ネットワークの広帯域化によって、音声伝送アプリケーションの需要は多様化するとともに、より対面コミュニケーションに近いものへの要求が高くなる。本論文はそうした要求を解決する 1 つのアプローチを示すことができたと考えている。

謝辞 本研究に際し、御協力いただいたマメ de がんすプロジェクトのメンバに感謝します。佐賀大学理工学部渡辺健次助教授と広島大学情報メディア教育研究センター西村浩二助手には有益なご助言をいただき、感謝いたします。本研究の一部は広島市立大学特定研究費 (平成 13 年度 1803, 平成 14 年度 2206, 平成 15 年度 3207), 通信・放送機構 JGN プロジェクト (P3410005, G13013) の支援を受けて実施されている。ここに記して感謝の意を示す。

## 参 考 文 献

- 1) 広井 順, 徳田恵一, 益子貴史, 小林隆夫, 北村 正: HMM に基づいた極低ビットレート音声符号化, 電子情報通信学会, Vol. J82-D2, No. 11, pp. 1857-1864 (1999).
- 2) Nishihori, Y., Okabe, S. and Yamamoto, Y.: Creating Cross-cultural Learning Communities on the Internet and the Japan Gigabit Network

— Integration of Media Tools into Collaborative Learning, *Proc. International Conference on Computers In Education*, Vol.1, pp.323–327 (2002).

- 3) Goto, M., Neyama, R. and Muraoka, Y.: RMCP: Remote Music Control Protocol — Design and Applications, *Proc. 1997 International Computer Music Conference*, pp.446–449 (1997).
- 4) Robust Audio Tool. <http://www-mice.cs.ucl.ac.uk/multimedia/software/rat/>
- 5) 岸田崇志, 河野英太郎, 前田香織: 多目的な音声伝送システムの設計, 情報処理学会研究報告 DSM-26-3, pp.13–18 (2002).
- 6) ITU-T G Series. <http://www.itu.int/>
- 7) Visual Audio Tool. <http://www-nrg.ee.lbl.gov/vat/>
- 8) Kouvelas, I. and Hardman, V.: Overcoming Workstation Scheduling Problems in a Real-Time Audio Tool, *Proc. Usenix Annual Technical Conference*, pp.235–242 (1997).
- 9) Hardman, V., Sasse, A., Handley, M. and Watson, A.: Reliable Audio for Use over the Internet, *Proc. INET'95, Honolulu, HI*, pp.171–178 (1995).
- 10) 近堂 徹, 大塚玉記, 西村浩二, 相原玲二: MPEG2 over IP 伝送システム mpeg2ts の開発と性能評価, DICOMO 2002 シンポジウム論文集, pp.157–160 (2002).
- 11) Perkins, C., Kouvelas, I., Hodson, O., et al.: RFC2198 — RTP Payload for Redundant Audio Data.
- 12) 大塚玉記, 西村浩二, 相原玲二, 前田香織: FEC を用いた MPEG2 over IP システムの開発と評価, 情報処理学会研究報告 DSM-24-8, pp.43–48 (2001).
- 13) マメ de がんす Project. <http://www.csi.ad.jp/activity/MAMEdeGansu/>
- 14) Japan Gigabit Network. <http://www.jgn.tao.go.jp>

(平成 15 年 5 月 16 日受付)

(平成 15 年 12 月 2 日採録)



岸田 崇志 (学生会員)

2001 年広島市立大学情報科学部情報工学科卒業。2003 年同大学大学院情報科学研究科博士課程前期修了。現在, 同大学院情報科学研究科博士課程後期在学中。IP ネットワーク上の音声伝送技術やマルチメディア応用に関する研究に従事。電子情報通信学会学生会員。

研究に従事。電子情報通信学会学生会員。



前田 香織 (正会員)

1982 年広島大学総合科学部卒業。広島大学工学部助手 (財) 放射線影響研究所技術員, 広島市立大学情報科学部助手を経て, 1999 年より広島市立大学情報処理センター助教授。博士 (情報工学)。コンピュータネットワーク, マルチメディア情報通信に関する研究に従事。電子情報通信学会, 教育システム情報学会各会員。



河野英太郎 (正会員)

1990 年広島工業大学工学部電気科卒業。1992 年東京工業大学大学院修士課程修了。1996 年同博士課程単位取得退学。同年広島市立大学情報処理センター助手。コンピュータネットワークの研究に従事。電子情報通信学会会員。



近堂 徹 (学生会員)

2001 年広島大学工学部第二類 (電気系) 卒業。2003 年同大学大学院工学研究科博士課程前期修了。現在, 同大学院工学研究科博士課程後期在学中。IP ネットワーク上の高品質動画像伝送に関する研究に従事。



相原 玲二 (正会員)

1981 年広島大学工学部第二類 (電気系) 卒業。1986 年同大学大学院博士課程修了。同大学同学部助手, 同大学集積化システム研究センター助教授を経て, 現在, 同大学情報メディア教育研究センター教授。工学博士。マルチプロセッサシステムの設計, コンピュータネットワークの研究に従事。電子情報通信学会, IEEE Computer Society 各会員。