

M-stage Vライブ 映像ストリーミング ゲートウェイの最新技術 —映像品質およびユーザビリ ティ確保の技術について—

インターネットから配信される映像と音声を、TV電話の仕組みでリアルタイムにライブストリーミング配信するシステムでは、映像品質およびユーザビリティの確保が重要な課題となる。この課題に対し、M-stage Vライブでの映像ストリーミングゲートウェイを用いて開発した技術により解決を図った。

あおやま はるみ みやた あつと
青山 春巳 宮田 篤人
おおぐし りょうへい いしまる ひろし
大串 亮平 石丸 浩

1. まえがき

ドコモでは、映像と音声をリアルタイムにライブストリーミング配信するサービス、M-stage Vライブを2003年5月より提供している[1]。このサービスでは、インターネットのストリーミング転送プロトコルであるRTP (Real-time Transport Protocol)[2]により、FOMA (Freedom Of Mobile multimedia Access) 端末など3G移動端末ビジュアルタイプヘオオーディオビジュアル通信規格“3G-324M”を使用し、TV電話の仕組みで映像と音声を配信している。このようなシステムでは、ライブ性の観点からリアルタイム性を確保することが重要であり、FOMA端末で再生される映像と音声をソースの映像と音声の実時間に比べ、ほとんど差がないように再生することが必要である。一方、リアルタイム性とは相反するが、インターネット側からの映像と音声の到達遅延の揺らぎやRTPペイロードサイズの変動をバッファリングし吸収することにより、再生映像品質を確保することも重要である。また、FOMA端末へ映像を表示するまでの時間、アーカイブコンテンツ再生終了時の切断時間を短縮することにより、課金対象となる余計な時間を排除することで、接続におけるユーザビリティをH.245などのTV電話との通信規格、RTSP (Real Time Streaming Protocol) などのリアルタイムストリーミングの規格に準拠しつつ向上することも重要な課題となる。

本稿では、M-stage Vライブにおける映像ストリーミングゲートウェイで採用した新技術のうち、映像品質の確保

という課題に対する解決策として、インターネット側からの映像と音声の到達遅延の揺らぎ、RTPペイロードサイズの変動を吸収する技術、エラー耐性を考慮しRTPペイロードから3G-324Mストリーミングデータを構成する技術について述べる。また、ユーザビリティの確保という課題に対する解決策として、FOMA端末へ映像を表示するまでの時間を短縮する技術、アーカイブコンテンツ再生終了時の切断時間を短縮する技術について述べる。

2. システム構成

M-stage Vライブにおける映像ストリーミングゲートウェイのシステム構成について図1に示し、詳細を以下に述べる。

映像ストリーミングゲートウェイは、モバイルの世界とインターネットの世界を結合するゲートウェイとして位置づけられ、M-stage Vライブにおいてはドコモコアネットワークとmopera (Mobile Operation Radio Assistant)[3]サーバ群の間に位置する。映像ストリーミングゲートウェイは、インタフェース変換装置 (ICE: Interface Converter Equipment) とゲートウェイ本体装置から構成される。

ICEはドコモのコアネットワークとゲートウェイ本体装置の間に位置し、ドコモのコアネットワーク側とのインタ

フェースであるPRI (Primary Rate Interface) 回線の終端、3G-324Mで使用されるマルチメディア通信制御プロトコルであるH.245[4]の制御、マルチメディア通信用多重化プロトコルH.223[5]上で多重化される音声データ・映像データ・H.245信号の多重・分離を行い、ドコモのコアネットワークのインタフェースに特化した装置としている。

ゲートウェイ本体装置は、ICEとmoperaプラットフォーム上のサーバ群との間に位置し、呼制御やmoperaサーバ群への接続認証、ストリーミングコンテンツ選択に関する制御、RTPプロトコルの終端など、ストリーミングクライアントとしての動作、音声・映像フレームを時系列に配置する同期制御、システム全体の運転制御などを行う。ゲートウェイ本体装置は同期制御など処理量が比較的多く、将来的にコストパフォーマンスの向上が見込まれる汎用サーバとすることで設備コストの低減に留意している。関連システムとして、FOMA端末ビジュアルタイプ、ドコモのコアネットワーク、インタワーク装置 (IWE: Inter Work Equipment)、moperaサーバ群 (認証サーバ、映像配信サーバ、ストリーミングサーバ、ドメインネームシステム (DNS: Domain Name System) サーバ、アクセス集計サーバ)、ライブカメラエンコーダがある。

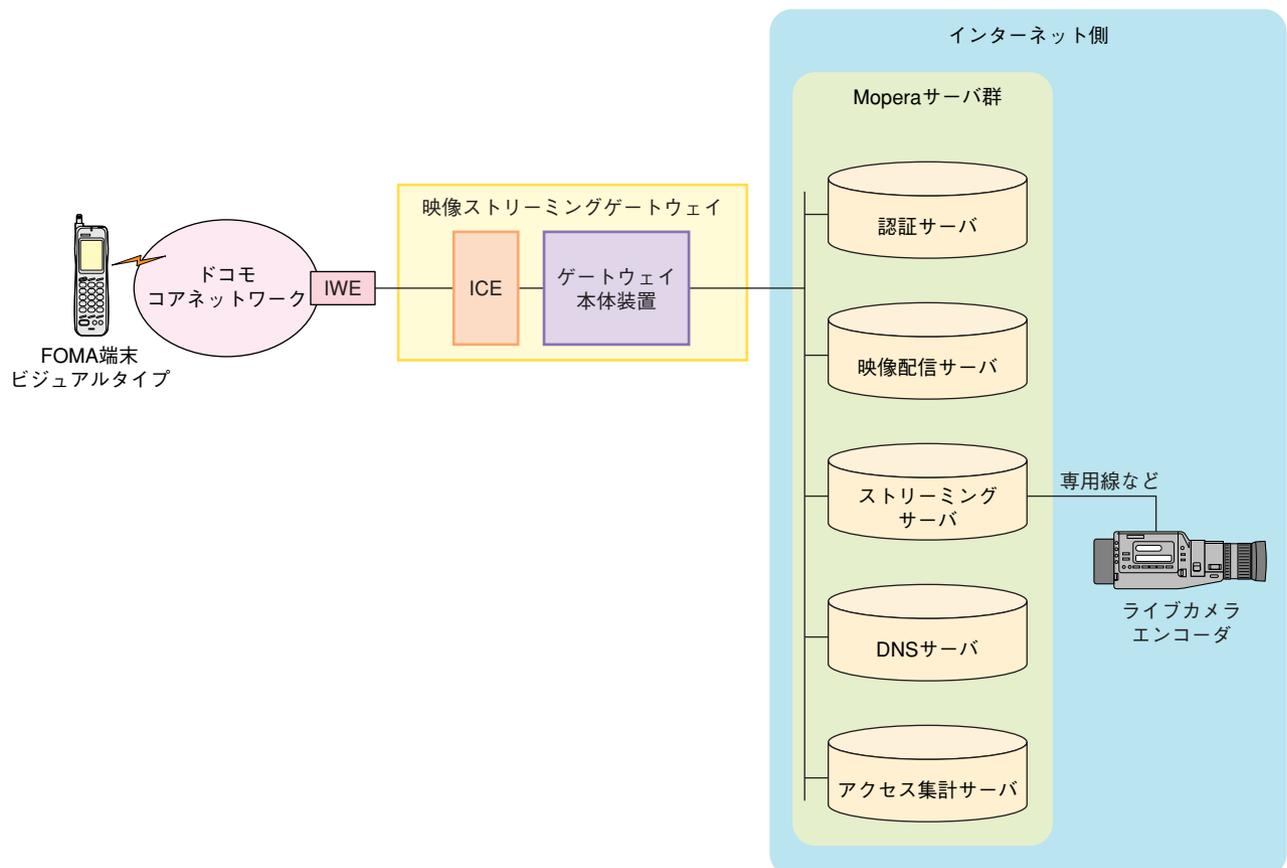


図1 映像ストリーミングゲートウェイを使ったシステム構成

3. 映像品質を確保する技術

映像品質を確保する技術として、M-stage Vライブにおける映像ストリーミングゲートウェイでは、“RTPプロトコルによりリアルタイムに転送される映像と音声の到達遅延の揺らぎとRTPペイロードサイズの変動を吸収する技術”と、“ストリーミング再生時のエラー耐性を考慮してRTPペイロードから3G-324Mストリーミングデータを構成する技術”を採用した。本章では、これらの技術について述べる。また、図2に技術概要を示す。

3.1 映像と音声の到達遅延の揺らぎ、RTPペイロードサイズ変動の吸収

従来の映像ストリーミングシステムでは、映像と音声の到達遅延の揺らぎ、RTPペイロードサイズの変動を吸収する方法として映像と音声を再生クライアント側でバッファリング後に再生する方式、クライアント側の受信状況に応じ送信側で映像レートを下げるなどのレート制御を行う方

式が一般的である。しかし、ライブコンテンツのストリーミングのようにリアルタイム性が重視され、性能条件の制約でクライアント側でレート制御に対応していない場合は、先に述べたレート制御方式をそのまま適用することが困難である。TV電話での伝送レートである64kbit/sなど帯域の狭い伝送路を使ったライブコンテンツのストリーミングにおいては、エンコーダ側では長時間平均では64kbit/sとなっても、各々の時点を見ると、RTPパケット到達の揺らぎ、RTPペイロードサイズの変動により、64kbit/sがコンスタントに保障されているわけではない。FOMA端末ビジュアルタイプ向けストリーミングでは、FOMA端末向けの伝送レートは64kbit/sが保障されており、これに合わせ、RTP送信側の映像と音声の到達遅延の揺らぎ、RTPペイロードサイズの変動に依存しないストリーミングデータのレート制御が必要になる。M-stage Vライブにおける映像ストリーミングゲートウェイでは、揺らぎのあるRTPパケットを受け取って処理するゲートウェイ本体装置と常に64kbit/sの一定ビットレートで映像と音声のデータを多

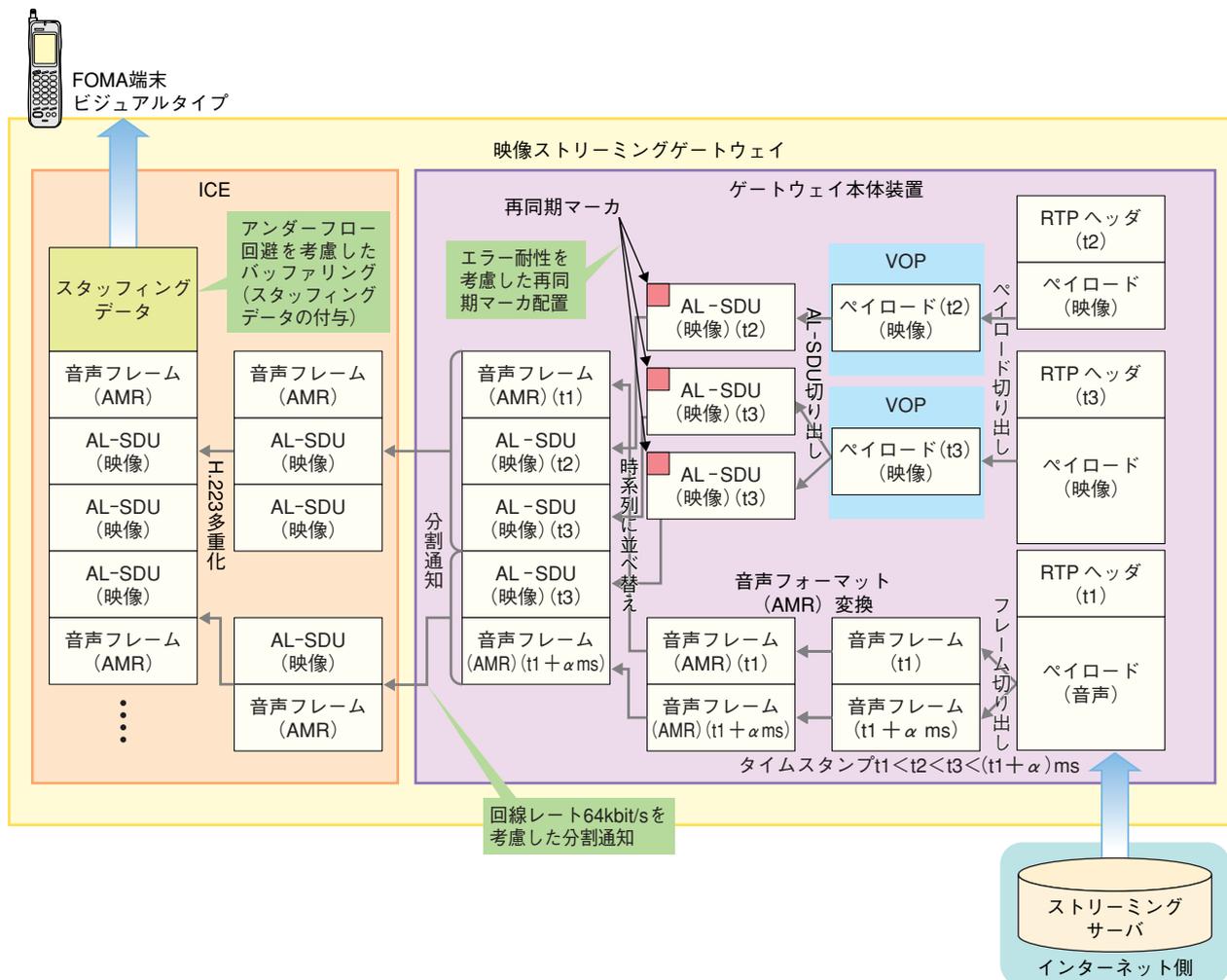


図2 映像品質を確保する技術の概要

重化してFOMA端末へ送るICEという特徴的な分離構成において、ゲートウェイ本体装置とICEでの多段のバッファリングと、ゲートウェイ本体装置でのフロー制御を行う。これにより、映像と音声のリアルタイム性を極力損なわないように、映像と音声の到達遅延の揺らぎ、RTPペイロードサイズの変動を吸収する手法を採用している。ゲートウェイ本体装置は、インターネット側からタイムスタンプ情報 (t_1, t_2, t_3) のついた音声と映像のRTPパケットを受信する。音声については、RTP上の音声ペイロードをフレームタイプに応じて複数に切り出し、その各々をFOMA端末向けフォーマットの適応マルチレート (AMR: Adaptive Multi Rate) フレームへ変換する。変換した各フレームが時系列上等間隔 (α ms) になるよう各フレームのタイムスタンプに α を付加する。映像については、RTP上の映像ペイロードを1VOP (Video Object Plane) 化する。VOPはFOMA端末側で指定する最大サイズでH.223上でストリーミングデータをFOMA端末に送信する際のデータ単位であるAL-SDU (Adaptation Layer Service Data Unit) に分割する。分割したVOPおよび音声フレームには、RTPヘッダに存在したタイムスタンプ情報を対応させており、各々に対応する絶対時刻から時系列に配列しICEへ送信する。その際、ICE側にてオーバーフローおよびアンダーフローが起きないように、64kbit/sに見合ったレートになるよう一定時間のバッファリングおよび送出伝送レートの平準化を行う。ICEでは映像と音声のアンダーフローによるFOMA端末側での映像の乱れをさらに回避するために一定時間のバッファリングを行う。この一定時間は、映像と音声のリアルタイム性を極力損なわないように、必要最低限の値となるようなチューニング設定を可能とする。

3.2 エラー耐性を考慮した3G-324M ストリーミングデータの構成技術

ゲートウェイ本体装置は、インターネット側からの映像のRTPパケットを受け、RTP上の映像ペイロードを1VOP化し、FOMA端末に送信する際のデータ単位であるAL-SDUに分割する。映像の符号化データに誤りが混入し、それ以降の符号の同期が取れず復号不可能な場合、誤りの含まれる符号化データをスキップし次に復号すべき符号化データの開始位置を示すための再同期マークが用いられる。AL-SDUへの分割にあたっては、VOP中の各再同期マークがAL-SDUの先頭にくるように配置する。この再同期化により、無線区間伝送時にビット誤りが発生し映像が乱れた場合でも、FOMA端末側にて即時に映像を乱れから回復することを可能とし、画像劣化の最小化を図っている。

4. ユーザビリティを確保する技術

M-stage Vライブにおける映像ストリーミングゲートウェイにて採用したユーザビリティを確保する技術として、“FOMA端末へ映像を表示するまでの時間を短縮する技術”と、“アーカイブコンテンツの再生終了時の切断時間を短縮する技術”について述べる。また、図3にそれらの技術の概要を示す。

4.1 FOMA 端末へ映像を表示するまでの 時間の短縮

インターネット側からの映像と音声は、RTPやRTSPなどのリアルタイムストリーミングの規格に準拠して配信され、これらはH.245などTV電話との通信規格に準拠してFOMA端末ビジュアルタイプへストリーミングされる。ストリーミング開始前に、FOMA端末側からの接続要求に基づき、発信者番号によるRADIUS (Remote Authentication Dial In User Service) 認証、発信時に指定するサブアドレス番号に対応するコンテンツURL (Uniform Resource Locator) の解決など、FOMA端末側と複数サーバとの間でネゴシエーションを実施する。FOMA端末側との論理チャネル開設などの処理と各サーバとの通信処理において、通信エラー発生時を考慮し、各通信を無駄のないようにする観点からは、FOMA端末側との処理と各サーバとの処理をシリアル化することも考えられるが、ストリーミング開始までに時間を要する点でユーザビリティ上問題がある。そこで、最低限順序性を保障する必要があるシーケンス以外はパラレルに処理することで、ストリーミング開始までの時間を短縮する方式としている。パラレル処理するシーケンスとして、M-stage Vライブ映像ストリーミングゲートウェイでは、FOMA端末側の復号処理で必要となるインターネット側のコンテンツ情報DCI (Decoder Configuration Information) の授受、インターネット側の音声と映像をFOMA端末向けにストリーミングデータ化する際に必要となるAL-SDUの最大サイズの授受などが挙げられる。パラレル化にあたっては、インターネット側からの音声と映像を早く取得しすぎること、映像と音声のリアルタイム性が損なわれないよう、FOMA端末側との論理チャネル開設完了時に最初の映像と音声のRTPパケットの取得を完了できるタイミングを計ってRTSP[6]のコンテンツ再生開始を指示するPLAYメソッドを発行し、タイムリーにインターネット側からの映像と音声のデータを取得する方式としている。

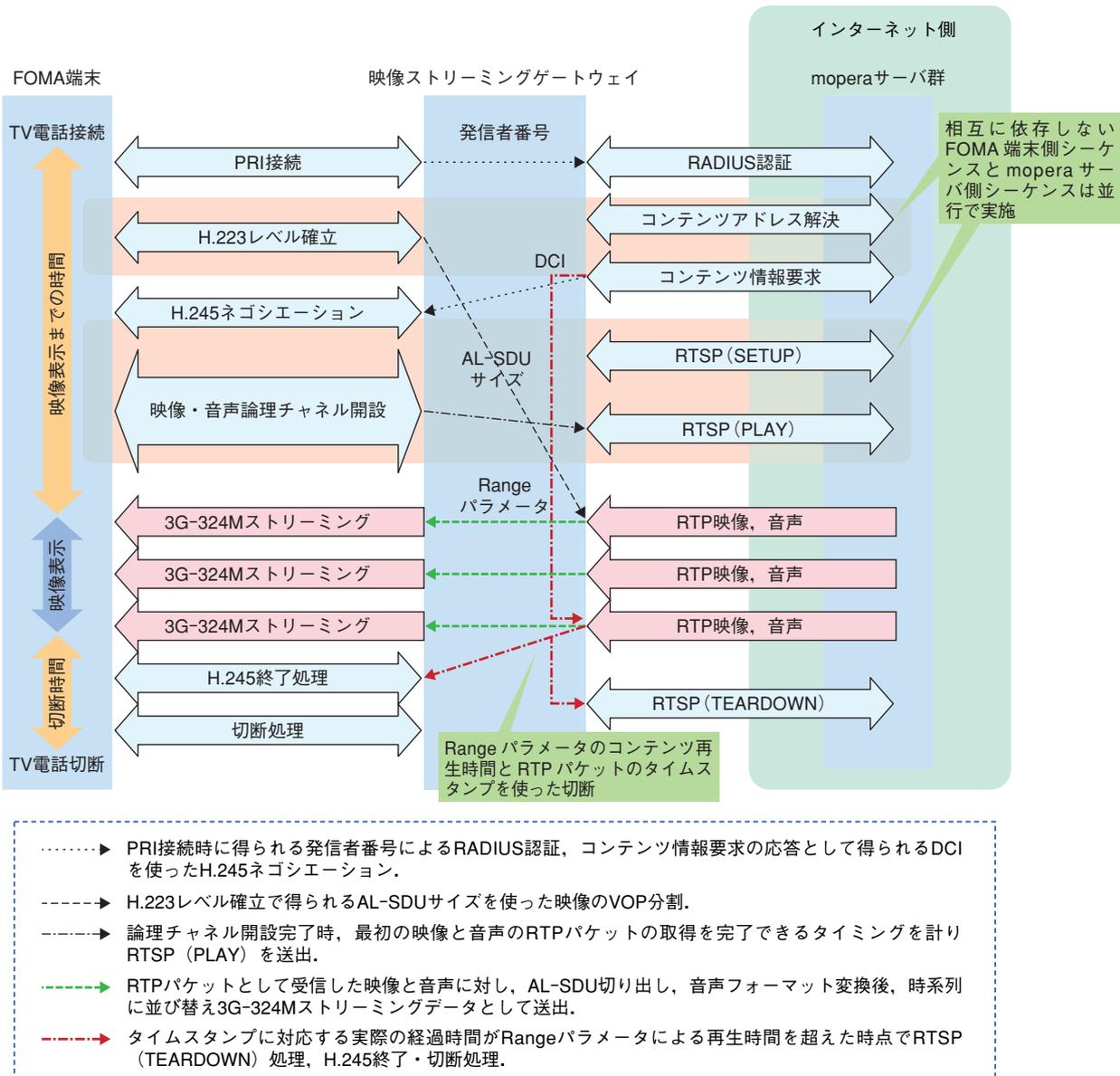


図3 ユーザビリティを確保する技術の概要

4.2 アーカイブコンテンツ再生終了時の切断時間の短縮

インターネット側からは RTP や RTSP などのリアルタイムストリーミングの規格に準拠して映像と音声の取得を行うが、VOD (Video On Demand) のように、アーカイブ的なコンテンツのストリーミングにおいて、インターネット側からコンテンツの終結時にセッションの終結を指示する RTSP の TEARDOWN 処理がされない場合がある。この場合、ユーザの意思によって切断するか、インターネット側からの映像と音声の RTP パケットが到達しなくなる契機で映像ストリーミングゲートウェイにて切断を行うしかなく、VOD コンテンツ終了から切断までに時間を要する場合がある。これに対し、映像ストリーミングゲートウェイでは、コンテンツ情報要求時に通知されたコンテンツの再生

時間長を示す Range パラメータから分かる再生時間と、逐次受信する RTP パケットのタイムスタンプに対応する実時間から分かる経過時間を常に比較している。タイムスタンプに対応する実際の経過時間が、Range パラメータによる再生時間を超えた時点で RTSP の TEARDOWN 処理、H.245 の終了処理および切断処理を行う。これにより、短時間でコンテンツ終了を検出して切断し、課金対象となる通信時間を最小限に抑える独自の方式としている。

5. あとがき

本稿では、M-stage V ライブにおける映像ストリーミングゲートウェイにて採用した技術として、映像と音声の RTP パケットの到達遅延の揺らぎ、RTP ペイロードサイズ変動の吸収など映像品質を確保する技術と、映像表示まで

の時間の短縮など接続におけるユーザビリティを確保する技術について述べた。今後は、高品質な音声コーデックへの対応、コンテンツ提供を容易化するコンテンツ閲覧制御への対応など、サービス性をさらに拡大できるような技術の検討と評価を行っていく予定である。

文 献

- [1] 大久保, ほか: “FOMA Vライブサービス向け映像配信ゲートウェイの開発,” 本誌, Vol.11, No.3, pp.42-47, Oct.2003.
- [2] Schulzrinne H. et al: RTP: “A Transport Protocol for Real-Time Applications,” RFC1889, Jan.1996.
- [3] 久保, ほか: “mopera SETP2と今後の展開,” 本誌, Vol.7, No.4, pp.32-39, Jan.2000.
- [4] ITU-T Recommendation H.245: “Control Protocol for Multimedia Communication,” Jul.2001.
- [5] ITU-T Recommendation H.223: “Multiplexing Protocol for Low Bit

Rate Mobile Multimedia Communication,” Jul.2001.

- [6] Schulzrinne H. Rao A. and Lanphier R: “Real Time Streaming Protocol (RTSP),” RFC2326, Apr.1998.

用 語 一 覧

AL-SDU : Adaptation Layer Service Data Unit
 AMR : Adaptive Multi Rate (適応マルチレート)
 DCI : Decoder Configuration Information
 DNS : Domain Name System (ドメインネームシステム)
 FOMA : Freedom Of Mobile multimedia Access
 ICE : Interface Converter Equipment (インタフェース変換装置)
 IWE : Inter Work Equipment (インタワーク装置)
 mopera : Mobile Operation Radio Assistant
 PRI : Primary Rate Interface
 RADIUS : Remote Authentication Dial In User Service
 RTP : Real-time Transport Protocol
 RTSP : Real Time Streaming Protocol
 URL : Uniform Resource Locator
 VOD : Video On Demand
 VOP : Video Object Plane