

## 多地点遠隔講義で自然なコミュニケーションを実現する DVリアルタイム処理機構の開発

小 峯 隆 宏<sup>†,††</sup> 勝 本 道 哲<sup>††</sup> 丹 康 雄<sup>†</sup>

双方向リアルタイム多地点遠隔講義には、講師と受講生との間の自然なコミュニケーションの実現が非常に重要である。しかし、これまでの遠隔講義システムは音声・映像等の伝送遅延時間が大きく自然な会話に困難である場合や、自然なコミュニケーションは可能であるが動作環境が2地点間接続に限定されている場合等、いくつかの技術的課題が残っている。我々は、双方向リアルタイム多地点遠隔講義で機能するデジタルビデオを用いたリアルタイム処理機構を提案している。本機構は、従来アナログ信号に変換してから行っていた音声ミキシング等の加工処理を、デジタルビデオ転送システムで伝送する音声・映像情報をデジタル信号のまま直接加工処理することで伝送遅延時間の短縮を実現するものである。本論文では本機構の比較性能評価を行った結果、自然な会話が可能で伝送遅延時間短縮を達成できたのでこれを報告する。

### Implementation of the Real-time Processing Mechanism with Digital Video for Advanced Interactive Distance Lecture among Multi-sites

TAKAHIRO KOMINE,<sup>†,††</sup> MICHIAKI KATSUMOTO<sup>††</sup> and YASUO TAN<sup>†</sup>

The realization of natural communications by the teacher and students is very important for successful interactive distance lectures among multi-sites. The conventional distance lecture systems have some technical difficulties such as the failure of natural communications with long transmission time-delay of audio/video information, and the system that only establishes natural communications within the limits of the point-to-point connection. We have developed the real-time processing mechanism with digital video for advanced interactive distance lecture among multi-sites. Using the digital video transport system, it directly edits audio/video digital signals, enabling the elimination of some A/D and D/A conversion steps. Evaluation of this mechanism showed that it maintains audio/video signal quality while greatly reducing the transmission time delay and achieves natural conversations among multi-sites, compared with conventional distance lecture systems with analog audio/video signal processing. Implementation of this mechanism as an intermediate node will enable participants to establish smooth conversation.

#### 1. ま え が き

講師から受講生への知識の伝授を行う講義には、講師が各受講生の理解度を探りながら講義内容の変更や受講生へ質問するといった行為や、受講生が把握できなかった講義内容について講師や他の受講生に質問するといった行為等、自然なコミュニケーションを実現できる環境が必要である。このため、講義というものは講師と受講生が同じ時間帯に同じ講義室内にいるこ

とが必須であった。しかし、インターネットに代表されるコンピュータネットワークの発達にともない、何らかの理由で講師と複数の受講生が講義室に揃わず多地点に分散している場合でも講義を実現できる遠隔講義システムの実現の可能性が高まってきた。そこで近年、受講生が都合の良い時間帯に蓄積されている講義コンテンツのオンデマンド型配信が受けられる遠隔講義システムや、講師の講義を複数の地点に分散している受講生がリアルタイムで受講する遠隔講義システムに関する研究がさかんである。前者の代表として SOI (School Of Internet) プロジェクト<sup>1),2)</sup>でのオンデマンド型講義ビデオ配信授業があり、複数の大学で提供されている講義を時間的、地理的な制約なしに受講できるという優れた利便性を持っているが、講義中の講

† 北陸先端科学技術大学院大学

Japan Advanced Institute of Science and Technology

†† 独立行政法人情報通信研究機構

National Institute of Information and Communications  
Technology

師と受講生による質疑応答といった行為がないため講義というより自己学習システムといえる。一方、リアルタイム遠隔講義システムも、実写映像を使ったシステムの研究<sup>3)</sup>、バーチャルリアリティ技術を活用したシステムの研究<sup>4)</sup>等が存在している。しかし、実写映像と衛星回線を活用したシステムは多数の遠隔受講会場の確保が容易になり受講生の地理的制約もなくなるという利点もあるが、衛星経由の音声・映像等の伝送遅延時間が大きい自然な質疑応答が困難という問題点がある。一方、バーチャルリアリティ技術を活用した場合は時間と空間を共有することで講師と受講生の一体感が増強されるという利点もあるが、講師が各受講生の理解度を探る等の細かい表情の把握は困難という問題点がある。また、実写映像と高速ネットワークを活用した場合の遠隔講義の研究開発<sup>5)-11)</sup>もこれまで数多く実現されており、質疑応答支援システムやプレゼンテーション手法の改善等による通常講義環境に近づくための提案がなされているが、自然なコミュニケーションを実現するには不十分な音声・映像品質である場合や2地点間構成という限定的な環境での自然なコミュニケーション実現という場合が多い。

本論文では、多地点に分散している講師と複数の受講生が伝送遅延時間の少ないインターネットを活用し音声と実写映像を中心とするストリーミング情報をリアルタイムに送受信することで講義を成立させる双方向リアルタイム遠隔講義の形態に焦点をあて、自然なコミュニケーションの実現手法<sup>12)</sup>に基づいたシステムに関して議論する。特に実空間内での自然会話の環境実現に関しては、TV 電話等の遠隔コミュニケーションでの自由会話で許容される音声・映像の伝送遅延時間が往復 400 ms であるという研究報告<sup>13)</sup>があり、多地点遠隔講義の場合には音声・映像等のストリーミング情報の符号化処理時間の短縮とともに音声ミキシング等の加工処理時間の短縮も考慮したシステム構成を検討する必要がある。一方、伝送遅延時間が小さくても伝送される音声・映像等のストリーミング情報の品質が低ければ、相手の話し声が聞き取りにくい、相手の細かな表情や動作が把握しにくい、といった状況が発生し、結果として自然なコミュニケーションの実現は困難になるため、音声・映像等のストリーミング情報の符号化処理時間が低くかつ品質が高い符号化方式を採用する必要がある。たとえば、現状の市販テレビ会議システムの双方向リアルタイム型遠隔講義への適用は、自然会話を実現させるためには音声・映像等のストリーミング情報の符号化処理時間が短くなるように伝送する同情報の品質を低く設定する必要があり、

我々の目指す遠隔講義実現のためには適切ではないといえる。

本論文では、多地点遠隔講義の場合での自然会話の実現手法として音声・映像等のストリーミング情報の符号化方式の符号化処理時間が小さく品質も高い方式を採用し、符号化されたストリーミング情報を IP パケットに搭載・転送し、音声ミキシング等の複数のストリーミング情報に対するリアルタイム加工処理もネットワーク中の IP パケットに搭載されたデータ形式のまま行い、再度 IP パケットに搭載・転送する、という手法を提案する。我々は、音声・映像等のストリーミング情報の符号化方式として符号化処理時間が短くかつ符号化後の音声・映像等のストリーミング情報の扱いが容易なデジタルビデオ (Digital Video<sup>14),15)</sup>: 以下 DV と記述) が上記の手法を適用可能な音声・映像等のストリーミング情報の符号化方式の1つであると考え、インターネット環境において複数の遠隔受講生が参加し自然コミュニケーションを実現する DV 遠隔講義システムを構築し、性能評価を行った。

以下、2章で、想定する多地点遠隔講義システム環境と、音声・映像等のストリーミング情報が生成されてから遠隔地で同情報が提供されるまでの所要時間に関する技術的課題について説明する。3章では、本論文で採用した DV 情報をインターネット上でリアルタイム伝送するアプリケーションシステム (Digital Video Transport System<sup>16)-18)</sup>: 以下 DVTS と記述) を多地点遠隔講義に利用した場合の技術的課題を述べる。4章では DV の特性を活用した DV リアルタイム処理機構について説明し、5章において実証実験による性能評価結果について述べる。

## 2. 多地点遠隔講義システム

まず本論文で想定する多地点遠隔講義システム環境を図1に示す。3地点以上の地理的に離れた会場間がインターネットで接続されている環境で、各会場にある音声・映像等のストリーミング情報の符号化・送信を行うエンコーダと受信・復号化を行うデコーダにより双方向のリアルタイム遠隔講義を行うと想定する。また、このような多地点遠隔講義では講師や遠隔会場から参加する受講生が講義に集中できる環境を整える必要があるため、各会場に配信する映像の選択、発言者の音量に合わせて行う音量レベル調整等の操作を講師や受講生以外の第三者が音声ミキサ等のリアルタイム加工処理装置で行うと想定する。図1では同機能をインターネット上にある中間ノード内に設置する場合を示しているが、同機能がどこか1つの会場内に存在

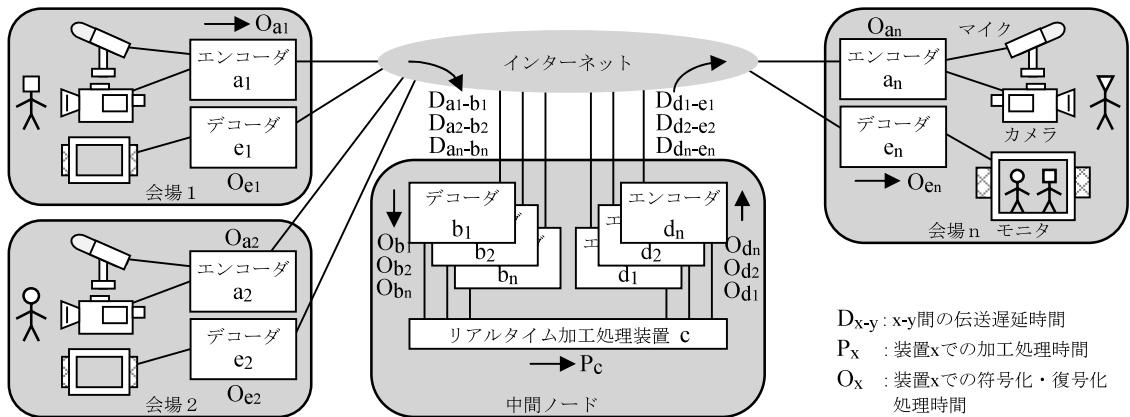


図 1 想定する多地点遠隔講義システム概念図

Fig. 1 Concept of advanced interactive distance lecture among multi-sites.

している場合も基本的な情報の流れは変化しない。

このような想定環境において、我々の目指すより自然なコミュニケーションを可能とする双方向リアルタイム遠隔講義の実現のためには多くの技術的課題があるが、本論文では、相手の話し声が明確に聞き取れ相手の細かな表情や動作が把握できる品質を持った音声・映像符号化方式の選定と、自然会話実現のために重要な音声・映像等の伝送遅延時間の短縮実現に絞って議論する。以下では、音声・映像等の伝送遅延時間に関する技術的課題について、図 1 を参照しながら伝送遅延時間短縮の可能性を項目別に説明する。図 1 内の英字記号はそれぞれ装置  $x$  から装置  $y$  間のネットワーク伝送遅延時間  $D_{x-y}$ 、装置  $x$  における音声・映像等のリアルタイム加工処理時間  $P_x$  および符号化・復号化処理時間  $O_x$  を表現しており、会場 1 でエンコーダ  $a_1$  に入力された音声・映像等の情報が会場  $n$  のデコーダ  $e_n$  から出力されるまでの所要時間（片方向伝送遅延時間） $T_{a_1-e_n}$  は

$$T_{a_1-e_n} = O_{a_1} + D_{a_1-b_1} + O_{b_1} + P_c + O_{d_n} + D_{d_n-e_n} + O_{e_n} \quad (1)$$

となる。

ネットワーク伝送遅延時間は、伝送距離や伝送装置の構成等に依存するものであり、本論文で扱う遠隔講義システムの範囲外である。なお、インターネット伝送遅延時間に関しては、研究開発用ギガビットネットワークを活用したギガビットイーサネットによる高速 IP 実験網の性能評価実験が実施されており、IPv4 の ping コマンドによる往復伝送遅延時間（Round Trip Time: 以下 RTT と記述）が東京・大阪間で 13.2 ms、東京・沖縄間で 46.8 ms という報告<sup>19)</sup>がされている。本論文ではこれらの値を 1 つの基準として議論する。

加工処理時間は、デコーダから出力される複数のストリーミング情報に対して音声ミキシング等の加工処理を行うための所用時間であり、音声・映像情報に関しては現在のところデコーダから出力されるアナログ情報をリアルタイム加工処理するためには一定の所要時間となっており、単独での処理時間短縮は困難である。しかし、ネットワーク上を流れているデジタル情報をデジタル状態のままリアルタイム加工可能な機構または装置が存在する場合は、中間ノード内のデコーダでのデジタル信号からアナログ信号へ変換する D/A 変換処理およびエンコーダでのアナログ信号からデジタル信号へ変換する A/D 変換処理を省略することができるためこの加工処理時間の短縮は可能となる。

符号化処理時間については、音声・映像等の符号化に要する処理時間が小さいエンコーダおよびデコーダを採用することで遅延時間短縮が可能である。ただし、1 章でも述べたように自然なコミュニケーションの実現にはある程度の音声・映像品質が必要であり、符号化処理時間と品質はトレードオフの関係になる傾向が強いため、符号化方式の選定から行う必要がある。また、符号化後のデジタル情報に対する加工処理が容易な符号化方式を採用した場合には、前述のとおり加工処理時間の短縮が可能であることも考慮する。

本論文では、音声・映像の符号化方式として符号化後のデジタル情報に対しても編集・加工処理が容易な DV を採用し、同情報をインターネット上でリアルタイム伝送する DVTS をエンコーダおよびデコーダとした双方向リアルタイム遠隔講義システムおよび複数の DV ストリーミング情報をデジタル情報のままリアルタイム加工処理する機構を開発することで、自然なコミュニケーションの実現を試みる。

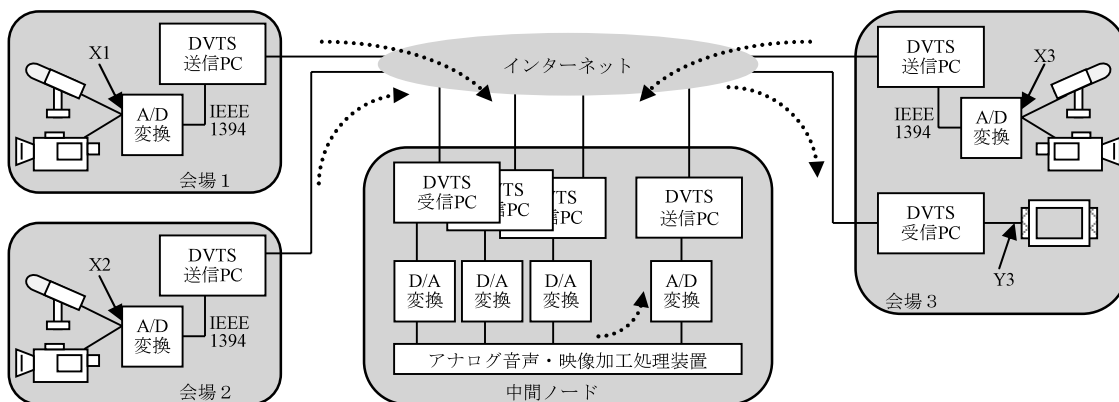
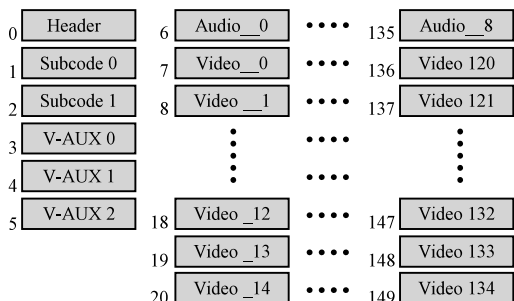


図 3 3 地点 DV 遠隔講義システムと会場 3 に向かうストリーミング情報の流れ

Fig. 3 Flowchart of streaming data on the DV interactive distance lecture among three sites.



Subcode : サブコードDIFブロック  
 V-AUX : 映像属性DIFブロック  
 Audio\_m : 音声DIFブロック (mはシーケンス番号)  
 Video\_n : 映像DIFブロック (nはシーケンス番号)  
 0~149 : DIFブロックシーケンス番号

図 2 DIF シーケンス内の DIF ブロック配置順序

Fig. 2 Transmission sequence of DIF blocks within a single DIF sequence.

### 3. 多地点 DV 遠隔講義システム

DV は、NTSC テレビ放送クラスの映像と 48-kHz サンプリング 16 bits リニア PCM の音声を提供するとともに、フレーム間符号化を行わずフレーム内符号化のみを行い、かつその符号化の圧縮率も低いため、MPEG-2<sup>20)</sup> や MPEG-4<sup>21)</sup> と比較して符号化処理時間が小さく、かつ提供される音声・映像品質が高くなっている。DV 情報は 80 バイト長サイズの DIF ブロックを最小構成単位として、DIF ブロック 150 個のまとまりが DIF シーケンス、DIF シーケンス 10 個のまとまりがフレーム (30 フレームで 1 秒に相当) という構成になっている。図 2 は 1DIF シーケンス内での DIF ブロックの配置構成を示しており、音声情報が格納される音声 DIF ブロック、映像情報が格納される映像 DIF ブロック、タイムコード等が格納されるサブコード DIF ブロック、映像の属性情報が格納される

映像属性 DIF ブロック等から構成され、その格納位置は固定となっている。また、DV の符号化処理は 1 フレーム単位で行われており、DIF シーケンス 10 個 (シーケンス番号 0~9) 中に 1 フレーム分の音声情報および映像情報が格納される構成になっている。このため複数の DV 情報のデジタル状態でのリアルタイム加工処理が容易と考えられ、従来のアナログ状態への変換をともなったアナログ情報のリアルタイム加工処理を省略可能である。

DVTS は、IEEE1394 ケーブル上を流れてくる DV 情報を適当な数の DIF ブロックを一塊として RTP/UDP/IP パケットに搭載しインターネット上にリアルタイム伝送する DVTS 送信機能と、インターネット上を流れてくる IP パケットから DV 情報を再構築し IEEE1394 ケーブル上へリアルタイム伝送する DVTS 受信機能で構成される。また、DVTS は様々な OS 環境の通常仕様の PC で稼働でき、特に Windows OS で稼働する DVTS 受信 PC では DV 情報の同 PC のモニター画面出力およびスピーカ出力になり DV 情報の IEEE1394 規格への格納作業が必要なくなるため、FreeBSD OS、Linux OS で稼働する DVTS 受信 PC と比較して最終出力までに要するまでの時間が約 1 フレーム分短縮され結果往復伝送遅延時間の短縮が実現される。

図 3 は、DVTS を活用した 3 地点 DV 遠隔講義での各会場から会場 3 へ向けた片方向の音声・映像等のストリーミング情報の流れを示している。アナログ音声・映像情報が各会場のポイント X1, X2, X3 から入力され、A/D 変換後に DVTS 送信 PC により伝送され、中間ノードにおいて DVTS 受信 PC、D/A 変換装置、アナログ音声・映像加工処理装置、A/D 変換装置、DVTS 送信 PC を経由し、会場 3 の DVTS

受信 PC によりポイント Y3 での情報出力が実現される。なお、会場 1 および 2 での DVTS 受信 PC においても同様の音声・映像情報の流れがある。この多地点 DV 遠隔講義システムは、MPEG-2 や MPEG-4 を活用した場合と比較して符号化処理時間が小さいためある程度の往復伝送遅延時間の短縮は期待できる。しかし、この遠隔講義システム構成には、中間ノードでの A/D および D/A 変換をともなうアナログ音声・映像加工処理装置が伝送遅延時間の増大と音声・映像等の品質劣化を生じさせるという技術的課題がまだ残っており、自然なコミュニケーション実現には不十分であり、これを改善することでさらなる往復伝送遅延時間の短縮を図る必要がある。

#### 4. DV リアルタイム処理機構

本章では、中間ノードでの A/D および D/A 変換をともなうアナログ音声・映像加工処理による伝送遅延時間の増大と音声・映像等の品質劣化に関する技術的課題を解決するために、我々が開発した DV リアルタイム処理機構について説明する。3 章で述べたように DV の符号化処理は 1 フレーム単位で行われている。特に 48-kHz サンプリング 16bits リニア PCM の音声情報に関しては、図 4 に示すように 2 チャネル(図中の ch-0, ch-1)の PCM 音声情報各 1 フレーム分が 8 バイト長のヘッダ情報と 16bits サンプリング音声データ群である 72 バイト長のデータ領域から構成される音声 DIF ブロック 90 個に分割格納され、1,500 個の DIF ブロックで構成される 1 フレーム分の DV 情報の中の固定格納位置に配置されている。これは、複数の会場から伝送されてくる DV 情報各 1 フレーム分に関して、同一の DIF シーケンスのシーケンス番号 (0~9) でかつ同一の DIF ブロックのシーケンス番号 (0~149) を持つ複数の音声 DIF ブロックに対するデジタル加工処理が可能であることを示している。ここで、複数のリニア PCM 音声情報に対する音声ミキシングはサンプリング音声データ単位の加算演算で実現されるため、図 4 の構成である複数の音声情報に対する音声ミキシングは、同一の格納位置に相当する音声 DIF ブロック内のサンプリングデータ単位の加算演算で実現される。一方、DV の映像情報に関しては 1 フレーム内で完結しているフレーム内符号化処理であるため、複数の会場から伝送されてくる DV 情報各 1 フレーム分に関して、1 フレーム分の映像 DIF ブロック 1,350 個すべてを配信したい映像情報の DIF ブロックで上書きすることで配信映像の選択も実現できる。

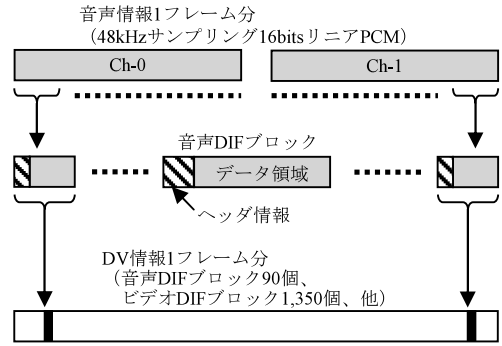


図 4 音声情報の DV 情報への格納手法  
Fig. 4 Relationship between audio data and DV data.

我々はこれら DV 情報の特性を活用することで、多地点から到着する DV 情報のデジタル情報での音声ミキシング処理および全地点への配信映像の選択等を行う DV リアルタイム処理機構を実現した。図 5 は、図 3 の構成に DV リアルタイム処理機構を導入した場合の 3 地点遠隔講義システム構成と各会場から会場 3 へ向けた片方向の音声・映像等のストリーミング情報の流れを示している。DV リアルタイム処理機構は、各会場から伝送される DV 情報を受信する DV 受信機能 R1, R2, R3, 1 フレーム分の DV 情報を一時格納する共有メモリ、共有メモリ内の DV 情報に対する加工処理を行うリアルタイム DV 処理機能、各会場に向けて DV 情報を送信する DV 送信機能 S1, S2, S3 から構成される。なお、会場 1 および 2 での DVTS 受信 PC においても同様の音声・映像情報の流れがあり、その場合は図 5 内の DV 送信機能 S3 がそれぞれ DV 送信機能 S1 および S2 に置き換わる。本機構は、図 3 の中間ノードの機器構成と比較して A/D および D/A 変換が省略され、複数の PC で実行していた DV 送受信機能とアナログ音声・映像加工処理装置で実施していたリアルタイム加工処理機能を 1 台の PC 上のアプリケーションプログラムとして稼動する。

図 6 は、3 地点間双方向リアルタイム遠隔講義システムでの DV リアルタイム処理機構全体の流れを示している。会場 1, 2, 3 から伝送されてくる DV 情報はそれぞれ DV 受信機能 R1, R2, R3 において RTP/UDP/IP パケットから取り出され、このうち配信映像として選択されている DV 情報のみが配信 DV 情報用共有メモリに格納されるとともに、格納された順番で随時共有メモリから DIF シーケンスのシーケンス番号 p, DIF ブロックのシーケンス番号 q の DIF ブロックが読み出され、DV 送信機能 S1, S2, S3 から各会場の DVTS 受信 PC を目的地とした RTP/UDP/IP パケットに搭載され転送される。ただし共有メモリか

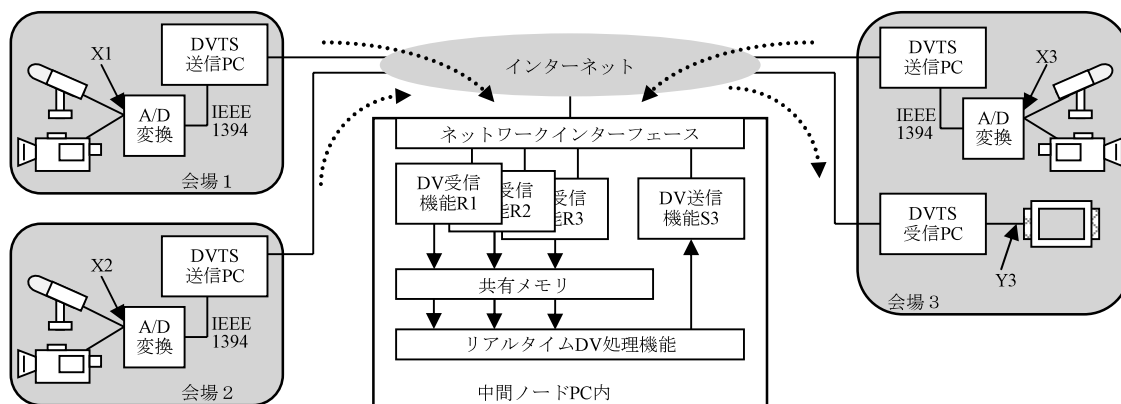


図 5 DV リアルタイム処理機構を導入時の会場 3 に向かうストリーミング情報の流れ  
 Fig. 5 Flowchart of streaming data in the case of the real-time processing with DV at an intermediate node.

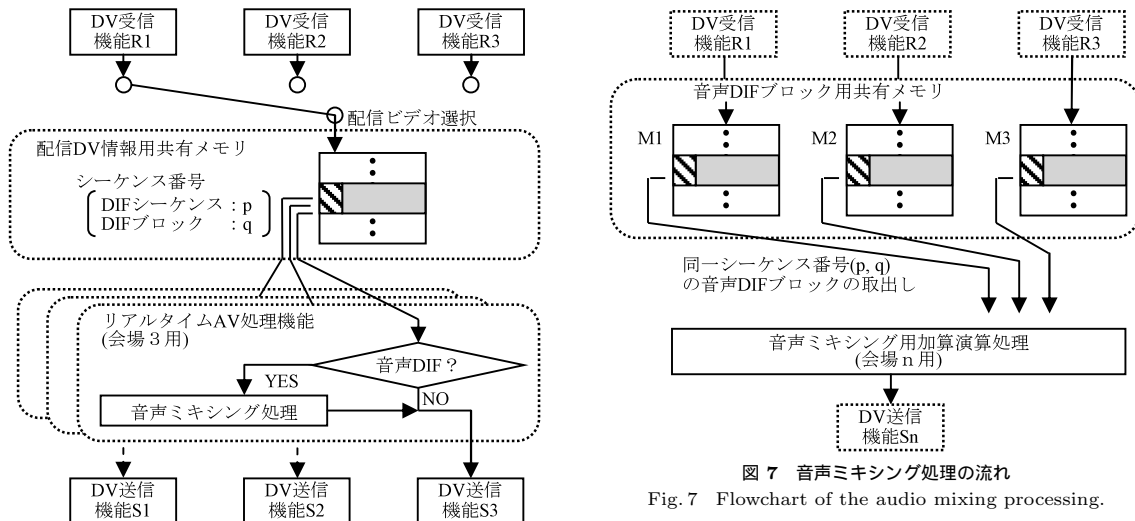


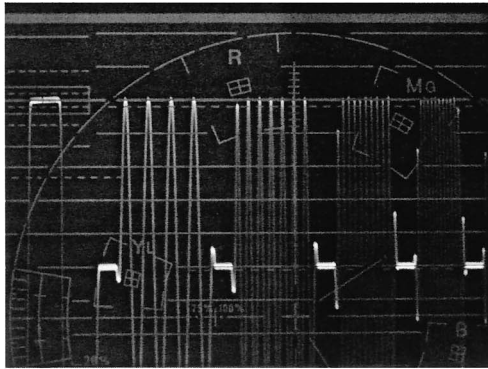
図 6 DV リアルタイム処理機構全体の流れ  
 Fig. 6 Flowchart of the DV real-time processing.

の音声 DIF ブロックと入れ替える。これにより、各会場の音声のボリューム調整は DV リアルタイム処理機構において自由に変更可能になる。

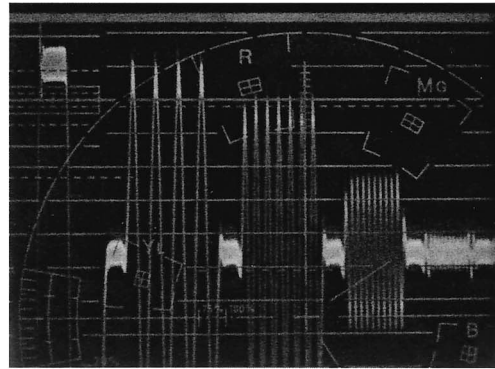
5. 評 価

ら読み出した DIF ブロックが音声 DIF ブロックである場合は、各会場から伝送されてくる音声の音量調整を行う音声ミキシング処理が追加される。図 7 は、DV リアルタイム処理機構内での音声ミキシング処理の流れを示している。まず、会場 1, 2, 3 から伝送されてくる DV 情報のうち音声 DIF ブロックのみが音声 DIF ブロック用共有メモリ M1, M2, M3 内の所定位置にそれぞれ格納される。図 6 の配信 DV 情報用共有メモリから読み出された DIF ブロックが音声 DIF ブロックの場合、その DIF シーケンスのシーケンス番号 p および DIF ブロックのシーケンス番号 q が同一である音声 DIF ブロックを音声 DIF ブロック用共有メモリ M1, M2, M3 からそれぞれ取り出し、各会場向けに設定された音声ミキシング処理を行い元

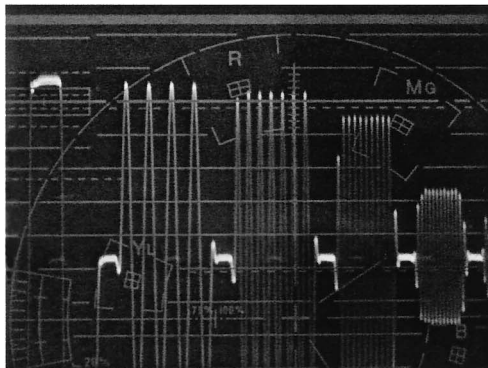
本章では、図 3 に示すようなアナログ音声・映像加工処理装置を用いた DV 遠隔講義システム構成（以後、構成 A と記述）、我々の提案する図 5 に示すような DV リアルタイム処理機構を用いた DV 遠隔講義システム構成（以後、構成 B と記述）、図 5 の構成で中間ノード PC を用いず DVTS 送信 PC から DVTS 受信 PC への直通伝送を行う 2 地点間 DV 遠隔講義システム構成（以後、構成 C と記述）の 3 つのシステム構成を対象にして比較性能評価を行う。評価項目は、伝送する音声・映像情報の品質劣化度の評価、伝送遅延時間の測定評価、妨害品質尺度による主観的評



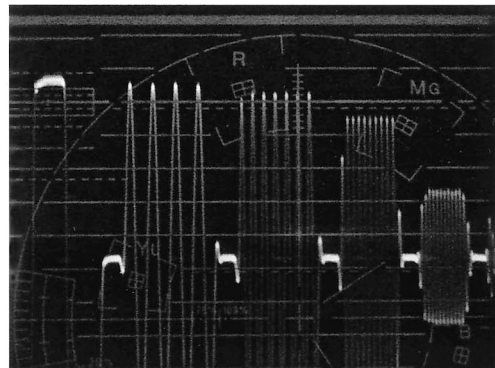
入力波形：100 RE マルチバースト信号



構成A (アナログ音声・映像編集加工処理装置を使用)



構成B (DVリアルタイム処理機構を使用)



構成C (2地点DV直通伝送)

図 8 入力波形とポイント Y3 での出力波形の比較

Fig. 8 Waveforms used for evaluating quality of DV streaming data.

表 1 主な機器の諸元

Table 1 Specifications of major components.

機器名称	諸元
DVTS送信PC	DVTS version 1.0b OS : FreeBSD 4.4 CPU : Pentium-3 1GHz Memory : 256MB
DVTS受信PC	DVTS for WinXP OS : Windows XP CPU : Pentium-M 1.6GHz Memory : 512MB
アナログAV 編集装置	FUTEC MXProDV (アナログAVポート使用)
A/D、D/A変換	SONY DVMC-MS1
中間ノード	OS : FreeBSD 4.4 CPU : Pentium-3 1GHz Memory : 256MB

価である．今回の比較性能評価に用いた主な機器の諸元を表 1 に示す．

### 5.1 品質劣化度の評価

品質劣化度の評価は，構成 A, B, C それぞれのポイント X1 に特定のテストパターン信号を入力し，ポイント Y3 からの出力信号を取り出し，その入力され

たテストパターン信号と出力信号をオシロスコープ表示画面で比較し，出力表示波形の差異から評価を行った．なお，本評価でのインターネット環境はスイッチングハブ 1 台のみの構成で実施している．図 8 は，テストパターン信号として NTSC ジェネレータで生成された 100 RE マルチバースト信号のポイント X1 での入力波形と，構成 A, B, C それぞれのポイント Y3 での出力波形を示している．この図から，入力波形に対する構成 B での出力波形の歪の度合いは，構成 A での出力波形の歪の度合いと比較して小さくなっており，構成 C の出力波形とほぼ一致していることが分かる．このため，我々の提案する DV リアルタイム処理機構を用いた DV 遠隔講義システム構成は，アナログ音声・映像加工処理装置を用いた場合と比較して品質劣化を低く抑えることができ，かつ DV リアルタイム処理機構での品質劣化はないといえる．

### 5.2 伝送遅延時間の測定評価

伝送遅延時間の測定評価のために，オシロスコープ上に表示される信号波形が明らかに異なる 2 種類のテストパターン信号が効果的であるため，今回は画面す

すべてが赤色のテストパターン信号と画面すべてが青色のテストパターン信号を用いた。なお、DV での音声情報と映像情報は同期しているため音声信号と映像信号の伝送遅延時間は等しいので今回の測定結果は DV ストリーミング情報の伝送遅延時間といえる。測定は、構成 A, B, C それぞれのポイント X1 に入力するテストパターン信号を切り替え、X1 での入力波形と Y3 での出力波形をオシロスコープ上で並列表示することで、2 つの波形の表示位置のずれの量から X1 から Y3 までのストリーミング情報の伝送遅延時間を測定する。なお、本評価でのインターネット環境はスイッチングハブ 1 台のみの構成で実施しているため、今回の測定結果にはネットワーク伝送遅延時間は含まれない。各システム構成での X1 から Y3 までの片方向の伝送遅延時間の測定結果を表 2 に示す。この測定結果から、DV リアルタイム処理機構を導入した構成 B の伝送遅延時間は、構成 A の場合と比較して大幅に短縮されていることが分かる。また、この測定結果から推定されるストリーミング情報の往復伝送遅延時間は、東京・大阪間のインターネット想定環境 (RTT が 13.2 ms) の場合、構成 B で 395.2 ms となり、遠隔コミュニケーションでの自由会話で許容される音声の往復伝送遅延時間 400 ms の範囲に入っていることが分かる。

表 2 片方向伝送遅延時間の測定結果

Table 2 Result of measured values for one way transmission delay.

システム構成	遅延時間 (平均値) (標準偏差)	
アナログ音声・映像編集加工処理装置を使用 (構成 A)	447.8 ms	4.5 ms
DV リアルタイム処理機構を使用 (構成 B)	191.0 ms	2.5 ms
2地点 DV 直通伝送 (構成 C)	156.8 ms	1.9 ms

### 5.3 妨害品質尺度による主観的評価

文献 13) の評価手法を参考に構成 A, B, C の会場 1 および 3 に被験者 1 名ずつ配置し自由会話を行ってもらった形態での主観的評価を実施した。被験者は、筆者の所属する機関の事務系職員 20 名で、男性 12 名女性 8 名、遠隔会議の経験者 10 名未経験者 10 名という内訳である。評価手法は、会場 1 および 3 にいる被験者 1 名ずつによる 1 分間の自由会話後に表 3 に示す 5 段階妨害尺度によるオピニオン評点評価を行ってもらった試験を、評価基準となる往復伝送遅延時間がゼロの場合と評価対象の遠隔講義システム構成の場合を交互に行った。

表 4 は、評価した 5 種類の遠隔講義システム構成の内訳と評価結果を示している。本試験では遅延発生装置を用いることでネットワーク伝送遅延時間を模擬しており、特に我々の提案する DV リアルタイム処理機構を活用した構成 B に関しては東京・大阪間のインターネット環境を想定した片方向 7 ms の遅延のほかに、東京・沖縄間のインターネット環境を想定した片方向 24 ms の遅延と参考用の片方向 100 ms の遅延の場合も試験を行った。なお、今回の試験結果について t-検定<sup>22)</sup>を行ったところ、5 種類の試験の順番による差異、性別による差異、遠隔会議の経験有無による差異についてはどれも危険率 5% で差異がないと推定された。表 4 から、評点が 4 以上つまり「自由会話をしていて遅延が気にならない」と評価した被験者は、東

表 3 5 段階妨害品質尺度

Table 3 Degradation category rating.

評点	品質尺度
5	遅延が全く感じられない
4	遅延がわずかに感じられるが、気にならない
3	遅延がわずかに気になる
2	遅延が気になる
1	遅延が非常に気になる

表 4 妨害品質尺度による主観的評価結果

Table 4 Result of mean opinion score evaluation with degradation category rating.

システム構成	ネットワーク 模擬遅延時間	推定される 往復遅延時間	オピニオン評点の分布					オピニオン 評点平均値
			[5]	[4]	[3]	[2]	[1]	
アナログ音声・映像編集加工処理装置を使用 (構成 A)	7 ms	909.6 ms	0	2	7	11	0	2.6
DV リアルタイム処理機を使用 (構成 B)	100 ms	582.0 ms	0	4	14	2	0	3.1
同上	24 ms	430.0 ms	2	12	6	0	0	3.8
同上	7 ms	396.0 ms	2	16	2	0	0	4.0
2地点 DV 直通伝送 (構成 C)	7 ms	327.6 ms	12	6	2	0	0	4.5



京・大阪間を想定したインターネット環境での構成 A では 10%であるのに対し、同インターネット環境での構成 B および C では 90%であった。また、東京・沖縄間を想定したインターネット環境での構成 B でも 70%の被験者が「自由会話をしている遅延が気にならない」と評価している。このため、DV リアルタイム処理機構を用いた DV 遠隔講義システムを利用することで遠隔地間での自由会話が実現され、遠隔講義での自然なコミュニケーション実現に大きく近づいたといえる。

## 6. おわりに

本論文では、多地点に分散している講師と複数の受講生が地上のインターネットを活用し音声と実写映像を中心とするストリーミング情報をリアルタイムに送受信することで講義を行う双方向リアルタイム遠隔講義に焦点をあて、自然なコミュニケーションの実現手法に基づいたシステムに関して議論を行った。

我々は、多地点遠隔講義の場合での自然会話の実現手法として、符号化処理時間が短くかつ符号化後の音声・映像等のストリーミング情報の扱いが容易な DV を採用し、インターネット上を流れる IP パケットに搭載されているデジタル状態の複数の DV ストリーミング情報に対して中間ノードで直接加工処理を行う DV リアルタイム処理機構を提案した。評価の結果、DV リアルタイム処理機構を用いた遠隔講義システムは、従来の講義システムと比較して往復伝送遅延時間が短縮され自由会話をを行うことができるようになり、かつストリーミング情報の品質劣化抑制という面においても有効に機能していることが分かった。

なお、中間ノードでの複数ストリーミング情報に対するリアルタイム加工処理機能は、DV 情報に限定された機能ではなく、また音声ミキシングや配信映像の選択以外のデジタル加工処理にも適応可能であるため、今後別の形態で有効利用される可能性があると考えられる。

謝辞 本研究の一部は、文部科学省の研究開発委託事業である「IT プログラム」<sup>23)</sup> により行われた。ここに記して感謝する。

## 参 考 文 献

- 1) 大川恵子, 伊集院百合, 村井 純: School of Internet—インターネット上での「インターネット学科」の構築, 情報処理学会論文誌, Vol.40, No.10, pp.3801-3810 (1999).
- 2) WIDE 大学 School of Internet.

<http://www.sfc.wide.ad.jp/soi/>

- 3) 田中健二, 近藤喜美夫: 大学間衛星ネットワーク(スペース・コラボレーション・システム)の構成, 電子情報通信学会論文誌 D-I, Vol.J82-D-I, No.4, pp.581-588 (1999).
- 4) 白戸仁博, 佐々木整, 竹谷 誠: パーチャルリアリティ技術を用いた遠隔教育システムの開発と適用, 電子情報通信学会論文誌 D-I, Vol.J83-D-I, No.6, pp.619-626 (2000).
- 5) 吉野孝ほか: インターネットを介したパーソナルコンピュータによる遠隔授業支援システムの開発と適用, 情報処理学会論文誌, Vol.39, No.10, pp.2788-2801 (1998).
- 6) 重野 寛, 間下直晃, 植原常宣, 松下 温: 講義イベントに着目した XML ベース遠隔教育システム, 情報処理学会論文誌, Vol.42, No.9, pp.2319-2327 (2001).
- 7) He, A. ほか: Tele-Seminar Room: ギガビットネットワークを利用した多地点リアルタイム双方向通信に基づく遠隔ゼミ支援システム, 情報処理学会論文誌, Vol.43, No.2, pp.647-660 (2002).
- 8) 若原俊彦: ATM-PVC 網を利用した遠隔講義システムの構成と特性, 電子情報通信学会論文誌 B-I, Vol.J81-B-I, No.8, pp.494-506 (1998).
- 9) 香川修見, 片山 薫, 上林弥彦: 遠隔教育システムにおける回答支援のための質問選択機構, 電子情報通信学会論文誌 D-II, Vol.J80-D-II, No.7, pp.1878-1886 (1997).
- 10) 前田香織, 相原玲二, 大槻説乎: 遠隔講義のためのマルチメディア教材提示, 情報処理学会論文誌, Vol.40, No.1, pp.161-167 (1999).
- 11) He, A. ほか: RIDEE-SPS: リアルタイム双方向遠隔教育環境のプレゼンテーションシステム, 情報処理学会論文誌, Vol.44, No.3, pp.700-708 (2003).
- 12) Komine, T., Tsuchiya, T., Katsumoto, M. and Tan, Y.: Development of the real-time processing proxy node for multimedia applications on interactive telecommunications, *Proc. International Symposium on Towards Peta-Bit Ultra-networks*, pp.156-161 (2003).
- 13) 栗田孝昭, 井合 知, 北脇信彦: オーディオビジュアル通信における伝搬遅延の影響, 電子情報通信学会論文誌 B-I, Vol.J76-B-I, No.4, pp.331-339 (1993).
- 14) IEC 61834: Helical-scan digital video cassette recording system using 6,35 mm magnetic tape for consumer use (525-60, 625-50, 1125-60 and 1250-50 systems), International Electrotechnical Commission.
- 15) 久保田幸雄: 図解デジタルビデオ読本, オーム社 (1995).
- 16) Ogawa, A., Kobayashi, K., Sugiura, K., Nakamura, O. and Murai, J.: Design and imple-

mentation of DV based video over RTP, *Packet Video 2000*, pp.140-146 (2000).

- 17) Sugiura, K., Sakurada, T. and Ogawa, A.: Demonstration of high quality media transport system using Internet, *IPSJ Journal*, Vol.41, No.12, pp.1321-1326 (2000).
- 18) Digital Video Transport System.  
<http://www.sfc.wide.ad.jp/DVTS/>
- 19) 町澤朗彦ほか：GbEによるネイティブIPv6広域実験網の設計と構築，マルチメディア，分散，協調とモバイルシンポジウム，pp.85-90 (2001).
- 20) 藤原 洋：ポイント図解式最新 MPEG 教科書，アスキー (1994).
- 21) 大久保榮，川島正久：要点チェック式 H.323/MPEG-4 教科書，IE インスティテュート (2001).
- 22) 市原清志：バイオサイエンスの統計学，南江堂 (1990).
- 23) VizGrid Project WWW page.  
<http://www.vizgrid.org/>

(平成 16 年 5 月 7 日受付)

(平成 16 年 9 月 3 日採録)



小峯 隆宏 (正会員)

1965 年生．1989 年電気通信大学電気通信学部物理工学科学士課程修了．同年郵政省通信総合研究所 (現・独立行政法人情報通信研究機構) 入所．以降，遠隔会議・遠隔コミュニ

ケーションの高度化に関する研究に従事する．電子情報通信学会通信方式研究会専門委員．



勝本 道哲 (正会員)

1997 年東洋大学大学院博士課程修了．同年郵政省通信総合研究所 (現・独立行政法人情報通信研究機構) 入所．1997～1999 年まで感性情報処理を用いたインターネットによる情報検索手法の研究開発に従事．1999～2001 年まで次世代インターネット技術である超高品質コンテンツ配信技術の研究開発に従事．2002 年よりインターネットアプリケーション技術の研究開発に従事する．現在，インターネットアプリケーショングループリーダー．情報処理学会 DPS 研究会運営委員．電子情報通信学会 IA 研究会専門委員．平成 11 年度情報処理学会山下記念賞受賞．IEEE Computer Society，ACM 各会員．工学博士．



丹 康雄 (正会員)

1965 年生．1993 年東京工業大学大学院理工学研究科博士後期課程修了．博士 (工学)．同年より北陸先端科学技術大学院大学情報科学研究科助手．同情報科学センター助教授を経て 2001 年情報科学研究科助教授．2004 年より国立情報学研究所助教授併任．計算機ネットワーク，ユビキタスコンピューティング，情報家電等に興味を持つ．情報処理学会ユビキタスコンピューティングシステム研究会運営委員，宅内情報通信・放送高度化フォーラム特別会員．電子情報通信学会，IEEE，ACM 各会員．