

IP 網における DV 動画像伝送装置開発のための基礎検討

光枝 慶一[†] 都築 伸二[†] 平屋 光教[†] 小池 泰樹^{††} 山田 芳郎[†]

[†] 愛媛大学工学部電気電子工学科 〒790-8577 愛媛県松山市文京町 3

^{††} (現在, 富士通関西中部ネットテック(株) 中部事業所

〒464-0075 愛知県名古屋千種区内山 3-26-2 名古屋董友ビル)

E-mail: [†]{k-mitu,tsuzuki,mitsu,yamada}@sarah.ee.ehime-u.ac.jp, ^{††}koike.yasuki@kcn.fujitsu.com

あらまし 本稿では, デジタルビデオ (DV) カメラの高精細動画像を, IP 網を介してリアルタイムに伝送することを目的としたプログラム開発のための基礎検討結果を述べる. 輻輳によるパケットロス対策として DV データの伝送レート変換アルゴリズムを, また無音検出によるコマ落としによる伝送遅延の短縮法を提案している.

キーワード ビデオ会議システム, DV over IP, IEEE1394, 伝送レート変換, 無音検出, 遅延時間短縮

Basic Study for DV Transmission System over IP Network

Keiichi MITSUEDA[†], Shinji TSUZUKI[†], Mitsunori HIRAYA[†], Yasuki KOIKE^{††}, and Yoshio YAMADA[†]

[†] Faculty of Engineering, Ehime University Bunkyou-tyou 3, Matsuyama-shi, Ehime, 790-8577 Japan

^{††} Fujitsu Kansai-chubu Net-tech Limited Uchiyama 3-26-2, Chigusa-ku, Nagoya-shi, 540-0001 Japan

E-mail: [†]{k-mitu,tsuzuki,mitsu,yamada}@sarah.ee.ehime-u.ac.jp, ^{††}koike.yasuki@kcn.fujitsu.com

Abstract In this paper, our basic study results of a program, that was developed to transmit high quality video data from a digital video (DV) camera over IP Network in real time, are shown. We propose an algorithm to convert the DV data rate to be slow. It can be used when a packet-loss occurs due to the network congestion. Another algorithm, which can shorten the transmission delay-time, is also proposed. It is realized by dropping the DV frame when the voice is silent in the frame period.

Key words Video conference system, DV over IP, IEEE1394, Conversion of Transmission-Rate, Silent Period Detection, Shortening Delay Time

1. ま え が き

工場やプラント等の FA 分野, 遠隔診断等の医療分野, 遠隔教育等の教育分野などで, IP 網を使って高品質動画像をストリーム配信できる伝送装置が求められている.

NTSC クラスの動画像品質の符号化方式としては MPEG2 あるいは DV(Digital Video) がある. 前者は網の伝送容量に応じて, 1M~15Mbps といった画像圧縮の程度を変更できる特長があるものの, 高効率な符号化を行うため画像のフレーム間予測等を行っており, 画像のコマ送り等オーサリングを行う場合は DV に劣る.

一方, DV はフレーム間予測を行っていないためオーサリング用途に適するものの, 30Mbps 程度のストリームデータとなるため, インターネットでの配信は現状では帯域的に困難であ

り, JGN(Japan Gigabit Network) 等の専用回線や高速 LAN 内で用いられることが多い. 汎用のパソコン (PC) で動作する DV の伝送ソフトウェアとしては, DVTS [1] が有名であり, ギガビットネットワーク・シンポジウム 2001(沖縄) の映像配信等で用いられている^(注1).

筆者らは, (株)FA システムエンジニアリング社(松山市)と共同で, 以下のようなコンセプトに基づき, 独自の DV 動画像の伝送装置を製作している. なお, IEEE1394 インターフェースボードとそのドライバ周辺および受信データの描画 (Windows の direct show を使用) ソフトウェアは FA システムエンジニアリング社が, IP 網での伝送ソフトウェア (socket を使用) は筆者らが担当している.

(注1) : <http://www.jgn.tao.go.jp/sympo2001/>

- (1) 汎用の Windows PC で動作
- (2) DV カメラ等の IEEE1394 インターフェースの DV ストリームデータを IP 網に送信
- (3) 受信データは PC の CRT に描画
- (4) 双方向で使う場合は、パーソナルなデスクトップビデオ会議を実現
- (5) 単方向で使う場合は、DV データをネットワーク越しにランダムアクセスする VOD(Video On Demand) システムや、データを遠隔編集するシステムを実現

つまり、DVTS のように DV データの配信に特化せず、パーソナルな DV データのネットワーク・アクセスツールの実現を目指しており、現在上記 (4) のための基礎検討を行っている。

本稿では、筆者らが担当する IP 網での伝送ソフトウェアについて述べる。ここでは、ベストエフォート型の IP 網を想定しており、輻輳によるパケット欠落 (以後、パケットロスと呼ぶ) や、パケットの到着時間の揺ぎによる遅延時間の累積が問題となる。通常、ストリームデータは UDP で送信するため、パケットロスが生じても再送しない。従って画像の一部が欠落してブロックノイズとなったり、音声途切れたりクリックノイズが発生する。また、遅延時間が累積していけば、ビデオ会議のような対話型通信が困難になる。通常、違和感なく対話するには 150m~200m 秒以下の遅延に押える必要がある。これらの問題に対して、伝送レートの変換や無音検出による遅延時間の短縮方法を提案し、それらの効果を述べる。

2. 使用機器とネットワークの構成

2.1 ネットワーク構成

本稿で使用したネットワークの構成を図 1 に示す。筆者らの研究室の PC1 から送信されたデータは、JGN の ATM 回線に入り、北陸 AP (アクセスポイント) で折り返されて、キャンパス LAN (ルーターのホップ数: 3) を経由して PC2 に帰って来る。双方向で実験するときは、PC2 から逆の経路で PC1 に送信する。なお、図中のルーター c7204(1) と ATM-SW (スイッチ)(1) 間では、10M~40Mbps の範囲でトラフィックをシェーピングして、仮想的に通信容量の小さいネットワークを模擬できるようにした。

2.2 使用機器の構成

図 1 に示した PC 周辺の構成を述べる。送信側の PC1 には IEEE1394 を介して DV カメラが接続してあり、2.1 節で述べたように受信側の PC2 に接続してある。表 1 にこれら PC のスペックを示す。ただし、ネットワークインターフェースは両者とも 100base-TX であり、表中の CPU 使用率およびメモリ使用量は、Windows のタスクマネージャで測定した値である。

DV カメラは SONY 製 DCR-TRV30 を使用した。

図 1 中のルーター c7204(1)(2) は、シスコ社製 7204VXR(NPE-200, IOS version 12.2) である。

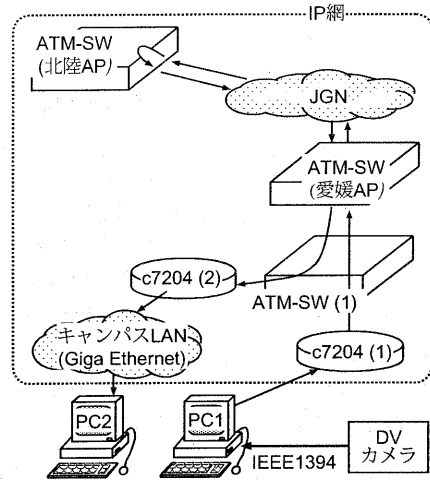


図 1 ネットワーク構成
Fig. 1 Network Configuration

表 1 使用した PC の仕様
Table 1 Specification of PCs

	PC1(送信側)	PC2(受信側)
メーカー	日本 IBM	日本 Gateway
CPU	PentiumIII 600MHz	Pentium IV 1.4GHz
メモリ	128MB	256MB
グラフィック	S3 Inc. Savage4 Xtreme	Matorox Millennium G550
IEEE1394 コントローラ	FASE 社 FWLink- 4001s	同左 (PC2 から送信、つまり 双方向通信のときに必要)
OS	Windows2000 Server	Windows2000 Professional
CPU 使用率	63%	62%
メモリ使用量	131MB	140MB

3. DV 動画伝送プログラム

3.1 ユーザインターフェース

DV 動画の受信を行っている時の、PC の画面例を図 2 に示す。図の右半分が、受信した DV データの描画用ユーザインターフェースである。プレビューウィンドウのサイズは、720 × 480 [pixel] または 360 × 240 [pixel] を選択できるが、図 2 では、360 × 240 [pixel] を選択した場合である。第 1 章で述べたように、将来の拡張性を考えコマ送り等のボタンも用意している。図の左下は、受信用の (コマンドプロンプト) ウィンドウである。将来は GUI 化する予定である。図の左上は Windows タスクマネージャで CPU 使用率とメモリ使用量をモニタしている様子である。

3.2 DV フォーマット

後述のレート変換や、無音検出アルゴリズムでは、DV のデータの性質を用いている。そこで本節では、DV のデータフォー

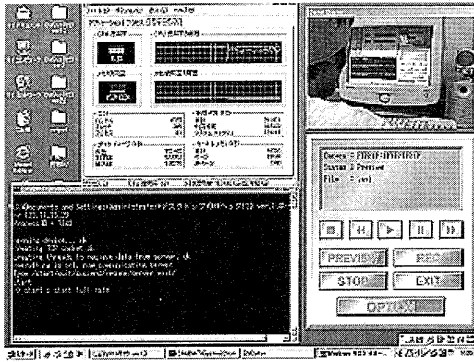


図2 ユーザインターフェイス
Fig. 2 User Interface

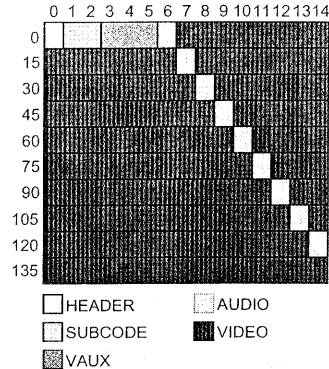


図4 DIFシーケンス
Fig. 4 DIF Sequence

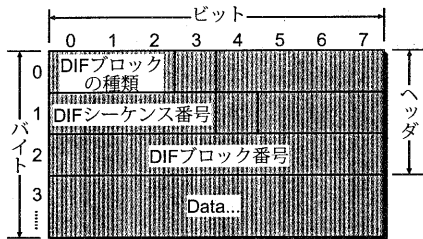


図3 DIFブロック
Fig. 3 DIF Block

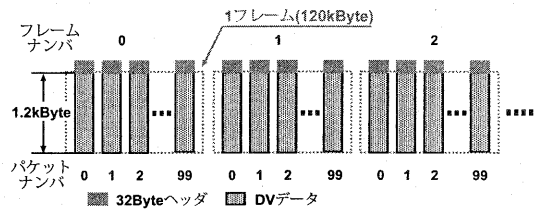


図5 DVデータのフォーマット
Fig. 5 Packet format of DV data

マットについて説明する [2], [3].

DV データは、80Byte の DIF ブロックが複数集まって構成される。150 個の連続した DIF ブロックの集合を DIF シーケンスと呼び、さらに 10 個の DIF シーケンスで画像 1 フレームが構成される。各 DIF ブロックは、ヘッダ 3Byte が ID と呼ばれ、残りの 77Byte にデータが格納されている。図 3 に DIF ブロックのデータ構造を示し、以下に DIF ブロックの ID の詳細を示す。

第 0 バイト

0-2 ビット: DIF ブロックの種類
(000:HEADER, 001:SUBCODE, 010:VAUX,
011:AUDIO, 100:VIDEO)

3 ビット: デフォルト 1

4-7 ビット: 任意 (1,0 どちらでもよい)

第 1 バイト

0-3 ビット: DIF シーケンス番号 (0-9 の 2 進表示)

4 ビット: ビットレート (0:25Mbps, 1:50Mbps)

5-7 ビット: デフォルト 1

第 2 バイト

0-7 ビット: DIF ブロック番号 (0-134 の 2 進表示)

また、DIF シーケンスは、HEADER:1ヶ、SUBCODE:2ヶ、VAUX:3ヶ、AUDIO:9ヶ、VIDEO:135ヶの、計 150 ヶの DIF ブロックから構成されており、図 4 に示すように並んでいる。

3.3 伝送ソフトウェア

本ソフトウェアでは、図 5 に示すように、画像 1 フレーム分の DV データ (120kByte) を 100 個にフラグメントして送信する。このフラグメント数は、フレームサイズ 120kByte を割り切ることができ、かつ データサイズが Ethernet の MTU 以下で最大となるように求めた。

さらに、フラグメントされた DV データには、フレーム番号やパケット番号、タイムスタンプ等の情報を付加するために 32Byte のヘッダを付けており、これらの情報を元にパケットロスや遅延時間の測定をした。

フラグメントした DV データに付加するヘッダには、先頭の 4Byte にフレームナンバーを、次の 4Byte にパケットナンバー (0~99) を、その次の 4Byte にタイムスタンプを押しした。

本ソフトウェアでは、プログラムの実行時に引数を指定することにより、これらの情報を用いてパケットロスや遅延時間の表示を行うことが出来るように実装している。

3.3.1 パケットロスの測定方法

フレームナンバーは 0 から unsigned long の最大値までを、パケットナンバーは、1 フレームを 100 分割しているため、0~99 まで順にカウントしており、プログラムの実行時に引数によってパケットロスの測定を行うように指定すると、受信側でこれらの値を検出して、パケットロスが起こるたびにロス・カウンタを増やして行く。

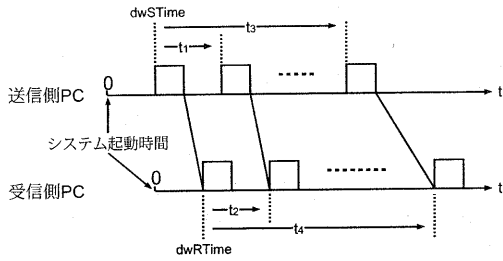


図6 遅延時間の測定

Fig. 6 Measurement of delay time

3.4 遅延時間の測定方法

C言語の標準ライブラリで提供されている時間関数の精度は低く、本研究における遅延時間の測定には不十分である。そこで、Windowsのマルチメディアコンポーネントが用意しているtimeGetTime()関数を用いることにより、システム起動からの時間を1msの精度で測定できるようにした[4]。

この関数を利用して、パケット・ヘッダにタイムスタンプを押し、その値から遅延時間を測定する。まず、図6に示すように最初のパケットのヘッダに付いているタイムスタンプの時刻と、このパケットを受信した時刻をそれぞれdwSTimeとdwRTIMEというの2つの変数に保持しておく。

そして、以後はこれらの時間より、受信したパケットは送信側PCが最初のパケットを送ってから何ms後に送られたか(図6中の t_1 や t_3)と、受信側PCが最初のパケットを受信してから何ms後に届いたか(図6中の t_2 や t_4)を計算し、それぞれの差をとる($t_2 - t_1$, $t_4 - t_3$)と伝送遅延時間が測定できる。

本ソフトウェアでは、実行時に引数で指定すると、こうして算出した遅延時間を1秒ごとに表示する仕組みになっている。

4. レート変換によるパケットロス軽減

4.1 レート変換の原理

伝送容量の細い回線やトラフィックの多い回線で、DV動画を伝送するためには、データ量を減らさなければならない。データ量を減らす手段として、MPEGに代表されるフレーム間予測圧縮技術が挙げられるが、本研究で扱っているDVの符号化はフレーム内圧縮を行っているものの、フレーム間の圧縮は行っていない。そこで音声DIFはそのまま、画像DIFを間引き、その量を変更することで、データの伝送レートを変更するようにした。つまり、 m フレーム中1フレームだけは間引かずフレーム内の全てのデータを送信し、それ以外のフレームは、音声DIFのみを送信するというものである。この m を間引き係数と呼ぶことにする。

ここで、間引き係数と伝送レートの関係を式(1)に表す。式(1)において、 $29.97[\text{frame/s}]$ はフルレート($m=1$)時のDV動画の、1秒間に送信されるフレーム数である。 $120,000[\text{Byte}]$ は1フレームのデータサイズ、 $32[\text{Byte}]$ はヘッダのサイズ、100は映像フレームのフラグメント数、 $7,200[\text{Byte}]$ は1フレーム内の音声データのサイズ、6は音声データのフラグメント数で

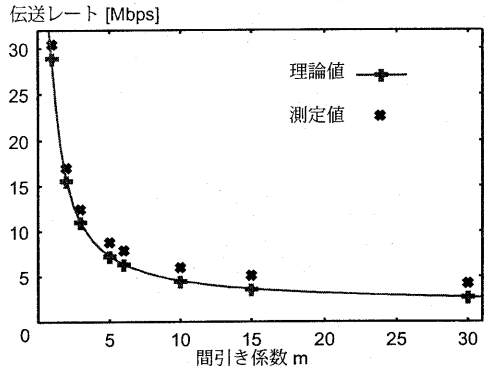


図7 間引き係数と伝送レートの関係

Fig. 7 Relationship between data-skipping coefficient and transmission-rate

ある。

伝送レート

$$= \left\{ \frac{29.97[\text{frame/s}]}{m} \times (120,000[\text{Byte}] + 32[\text{Byte}] \times 100) + (29.97[\text{frame/s}] - \frac{29.97[\text{frame/s}]}{m}) \times (7,200[\text{Byte}] + 32[\text{Byte}] \times 6) \right\} \times 8[\text{bit}] \text{ Mbps} \quad (1)$$

式(1)で求めた伝送レートの理論値と、実際の伝送時における測定値の比較を図7に示す。ただし、この測定は、図1中のルーター、c7204(1)、のインターフェースを5分間監視し、その平均を求めた値である。測定値が理論値よりもわずかに大きくなっているが、これは測定値にはUDP/IP等のプロトコルヘッダのサイズが加算されているからである。

本提案方式は単純なアルゴリズムであるため、レートを素早く変換できる。したがって、回線の輻輳状況に応じて適応的にレートを変換する用途に適している。その反面、 m が大きいたまには再生映像がコマ送りのようになるので、動きの速い映像には向きである。

4.2 音声のみを伝送する仕組み

伝送レートを下げる場合は、画像データを送らず、DVフレームから音声データのみを抜き出して送信する。以下ではその実現方法について述べる。

まず、送信側では、DVデータをフレーム毎にIEEE1394ドライバから受信し、バッファへ格納した後、図8に示すようにフレーム内に散在している合計7,200ByteのAUDIO DIFを抜き出し、送信する。

受信側では、ディスプレイに表示し終わったフレームに、受信した音声データを上書きして(フレームの再構成、と呼ぶ)再生する。これにより、描画される画像は前回と同じで音声のみ更新されて再生される。よって、ディスプレイ上ではコマ送り画像のように見えるが、実際には29.97fpsで表示することになる。

なお、送受信PC間の通信制御用にTCPコネクションを別途張り(データ自身はUDPで伝送)、受信側PCから送信側

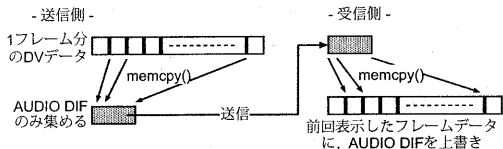


図8 音声データのみ伝送
Fig. 8 Audio data transmission without video data

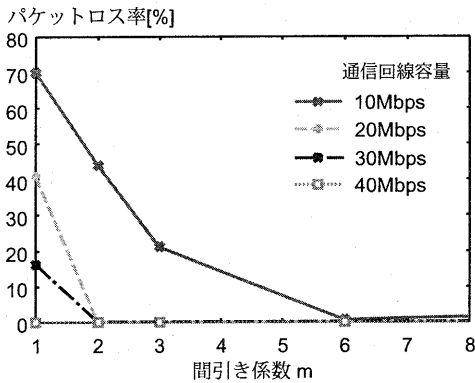


図9 間引き係数とパケットロス率の関係
Fig. 9 Data-skipping coefficient and packet-loss rate

PC に対して間引き係数を送信し、それに応じたレートで送信するように実装している。これにより、受信側 PC からの間引き係数変更要求を受信した送信側 PC は、即座に伝送レートを変更できる。

4.3 各帯域における間引き係数とパケットロス率の関係

図1中のルーター、c7204(1)、でのシェーピングの値を10M~40Mbpsの間で変えたときの、間引き係数 m とパケットロス率の関係を図9に示す。パケットロス率は伝送開始からの値を1秒おきに表示してあり、この値をそれぞれの間引き係数につき5分間ずつ測定し、それらの平均をとってプロットした。

通信回線容量に応じて m の値を調整することで、パケットロス率を軽減できることが分かる。回線容量が10Mbpsの場合は完全にロス無くすることは出来ないが、 m を6以上にすれば、時々再生画像や音声にノイズが生じる程度であり、実用上は問題と思われる(ただし、5[frame/s]以下の動画像になるため、動画像の種類によって、実用的かどうかの判断は異なる)。

5. 遅延時間の短縮法

5.1 無音検出によるコマ落とし

本稿で試作したDVデータ伝送システムでは、送信側PCに比べて、受信側PCではフレームの再構成や映像の再生などの処理があるためCPU負荷が高くなっている。また、リアルタイムの双方向伝送(ビデオ会議等)において、遅延時間の最小化は不可欠である。そこで、この受信側PCでのCPU負荷を軽減し、遅延時間を最小化するために、DVフレーム中の音声データを解析し、無音の場合(あるいは背景雑音しかない場合)、

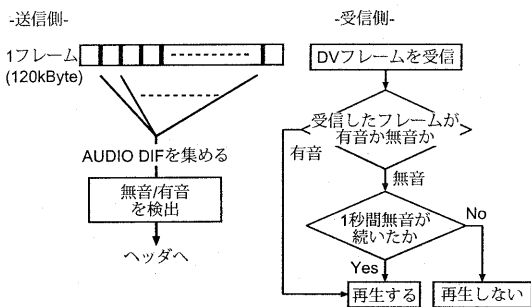


図10 無音検出アルゴリズム
Fig. 10 Silent Detection algorithm

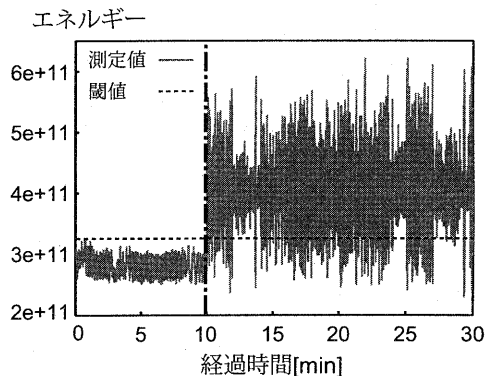


図11 音声のエネルギー例
Fig. 11 Example of energy of audio data

当該フレームは一切再生しないようにする(コマ落とし、と呼ぶ)実装を行った。

この実装の方法は、まず、送信側でDVの1フレーム内の音声DIF(各DIFの最後の2Byteは除く)から片チャンネルの音声のエネルギーの二乗和を求め、この値が閾値よりも多い/少ないで、そのフレームは有音であるか無音であるかを検出する[5]。そして、この検出結果をパケット・ヘッダに記して送信する。

受信側PCでは、受信したパケットのヘッダを見て、そのフレームが有音であればそのまま再生、無音であれば映像音声ともに再生しない。ただし、無音でも時々画像を更新するように、1秒間に1フレームは必ず表示するようにした。以上のアルゴリズムを、図10に示した。

5.2 有音/無音の閾値

測定に用いたサンプル音声のエネルギーの測定結果を図11に示す。図11中の一点鎖線より左側では音声は入力せず、映像のみを入力し、右側では両方を入力した。図11に示したように、有音/無音の閾値は、音声を入力していないときのエネルギーの最大値付近で設定した。以降の測定では、この閾値を用いて測定を行っている。

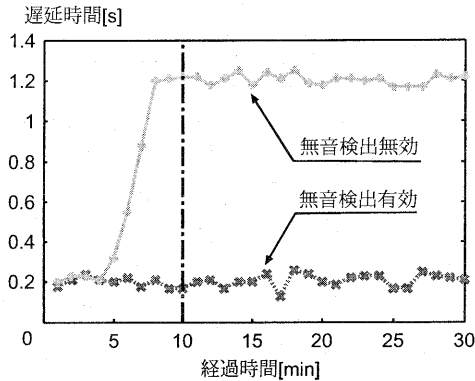


図 12 遅延時間の比較

Fig. 12 Comparison of Delay Time

5.3 遅延時間の測定

図 11 で使用した音声サンプルを用い、無音検出によるコマ落しを行う場合の遅延時間短縮の効果を調べるために、コマ落しを行わない場合との比較を行った。ここでの遅延時間は、DV カメラの入力から受信 PC の CRT に描画されるまでの時間である。

測定結果を図 12 に示す。ただし、図 12 中の一点鎖線は図 12 の一点鎖線の時刻と同一である。

図から分かるように無音の検出を行った場合は、行わなかった場合に比べて遅延時間の累積増加は見られなかった。これは、無音のフレームを廃棄(コマ落し)したことによって、バッファのキューに累積しているデータが無音の度に無くなるためである。また、遅延時間は約 200msec 程度に押えることができおり、ビデオ会議等の双方向通信に使用できる程度に短縮できた。

6. おわりに

本稿では、IP 網での DV データ伝送ソフトウェア開発のための基礎検討結果について述べた。ベストエフォート型の IP 網で想定される問題、輻輳によるパケットロス、およびパケットの到着時間の揺ぎによる遅延時間の累積に対する対策法を提案し、それらの有用性を述べた。

輻輳によりパケットロスが生じる場合は、第 4 章で述べたレート変換法を用いて、伝送データ量を減らすことによって、ロス率を軽減できることを示した。実験した範囲では 10Mbps までレートを下げても実用上は問題ないと思われた。伝送遅延については、第 5 章で述べたように、受信側 PC で無音のフレームを再生しないことによって解消出来ることを示した。

以上の結果より、今後の実装としては、レート変換と遅延短縮法を組合せ、図 13 に示すようなアルゴリズムで動作するプログラムに改良する予定である。

また、本稿では、1 対 1 で片方向の伝送のみについて述べたが、双方向の伝送を行う場合には、それぞれの PC には大きな CPU 負荷がかかる。この負荷を軽くするためのソフトウェアの改善や、PC のハードウェアスペックなどについても検討す

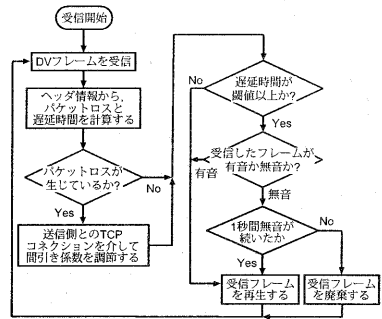


図 13 今後の実装について

Fig. 13 Future implementation

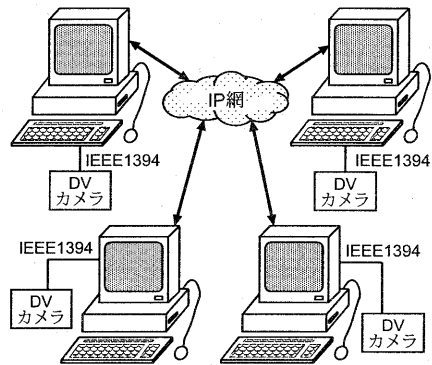


図 14 多地点ビデオ会議システム

Fig. 14 Multi-points video conference system

る必要がある。

さらには、図 14 に示すような多地点ビデオ会議システムの開発や、現在普及しつつある IPv6 への対応も検討して行く予定である。

謝 辞

本研究を進めるにあたり、共同研究を行っている(株)FAシステムエンジニアリング社殿に深く感謝します。また JGN 回線を使用した評価実験にご協力して頂いている、北陸先端科学技術大学院大学 丹康雄助教授に感謝致します。

本研究は、JGN プロジェクト No.JGN-R13024 に基づいて実施している。

文 献

- [1] 小川 晃 通, "DV 転送システム," <http://www.sfc.wide.ad.jp/DVTS/index-j.html>.
- [2] 小池 泰樹, "IP 網におけるリアルタイム動画像伝送方式に関する研究," 愛媛大学大学院理工学研究科博士前期課程電気電子工学専攻学位論文, 2002.
- [3] 久保田 幸雄 編著, "図解 デジタルビデオ読本," 株式会社 オーム社, 東京, 1996.
- [4] W. リチャード・スティーヴンス 著, 篠田 陽一 訳, "UNIX ネットワークプログラミング 第 2 版 Vol.1," 株式会社ピアソン・エデュケーション, 東京, 2001.
- [5] 都築, 山田, 田崎, 吉田, "非同期ネットワークでの音声同期再生と VSAT 衛星回線中継器への実装," 電気学会論文誌, 117-C, 11, pp.1557 - 1565, Nov.1997.