

マルチメディア会議システムにおけるバンド幅の 動的割り当てと動画像の画質制御

篠崎 雅英

日本アイ・ビー・エム (株) 東京基礎研究所

PCを利用して遠隔地を結ぶマルチメディア会議システムRTPは、INS Net 64を使って動画・音声・データを伝送することができる。通信コストの削減のため、1回線(2B、128kbps)のみを使用し、各メディアデータのバンド幅の動的割り当てを行うことによって効率的なデータ伝送を実現する。本システムでは、ISDNインタフェースを持ち、テレビ会議システム標準CCITT-H. 320準拠の機能を持つPC内蔵アダプタにより動画像伝送を実現した。PCにCODECを内蔵することにより、CODECと密接に連携した制御が可能になる。動画像のバンド幅に応じてフレームレートの変更および、拡大縮小を含めた画質の制御を行い、最適な画像を提供することができる。

A Dynamic Bandwidth Allocation and a Motion Picture Quality Control for a Multimedia Conferencing System

Masahide SHINOZAKI

IBM Japan, Ltd. Tokyo Research Laboratory
1623-14, Shimotsuruma, Yamato, Kanagawa, 242 JAPAN

In a multimedia remote conferencing system called RTP (Realtime Presentation System), motion pictures and voices and data are transferred by using INS Net64. In this system, we realize a effective data transmission on one ISDN line (2B, 128kbps) by dynamic bandwidth allocation for each media. We transfer motion picture by a CODEC adapter card which is based on CCITT H.320. When the bandwidth is changed, PC can control the CODEC adapter for motion picture quality enhancement such as frame rate control, and resizing.

1 はじめに

本研究所では、PC を使ってデスクトップで遠隔地との会議を実現する RTP(Realtime Presentation System) の研究を行なっている [1]。RTP では、TV64 と呼ぶ動画像圧縮伝送アダプタ・カードによって相手の表情などを動画像で伝えることができる。現在のテレビ会議システムでは、動画像の圧縮伝送方式として CCITT H.261 が主流となってきた。TV64 は、H.261 を含めたテレビ会議システム標準 H.320 に準拠した機能を実現している。これによって、RTP で動画像・音声・データといったマルチメディア・データの伝送が可能となっている。

本論文では、まず最初に RTP および TV64 について紹介する。そして、ISDN Net64 を使って動画像・音声・データを効率良く送るためのバンド幅の動的割り当て方式について述べる。また、バンド幅の変化に応じた動画像の画質制御方法について述べる。

2 リアルタイムプレゼンテーションシステム (RTP)

RTP は、ワークステーションを使った遠隔地間の会議システムである。ネットワークとしては、ISDN(2B+D, NTT INS ネット 64) または、Token-Ring LAN を使用する。RTP は、共有黒板、ビデオウィンドウを持つ (図 1)。共有黒板は、グラフィック・エディタの機能を持ち、ワープロやグラフィック・エディタなどのアプリケーションで作成したデータを使って、会議資料を作ることができる。ページめくり、拡大縮小、データ読み込みなど共有黒板上で行なう操作は、遠隔地にリアルタイムで反映される。ユーザは、共有黒板を使って遠隔地の相手にプレゼンテーションを行なう感覚で会議を進めていくこ

とができる。現在、RTP は、OS/2 2.0 上にインプリメントされている。

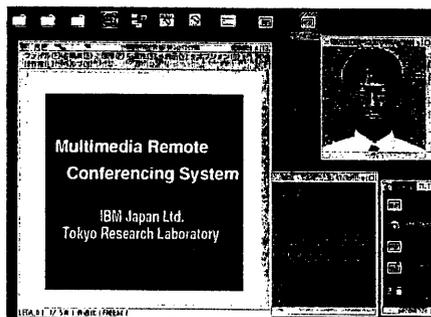


図 1: RTP の画面イメージ

3 動画像通信アダプタ TV64

TV64 は、国際標準 CCITT H.320 に準拠した画像通信アダプタである (図 2)。動画像・音声コーデックおよび ISDN インタフェースを内蔵している。アダプタ自身に NTSC 入出力機能と音声入出力、ハンドセットコネクタを持つ。動画像エンコーダは、NTSC 信号を入力とし、CCITT H.261 準拠の圧縮を行なう。圧縮結果はマルチメディア多重化部で音声・データと多重化し、ISDN インタフェースへ出力する。マルチメディア多重化は、CCITT H.221 に準拠している。受信側では、動画・音声・データを分離、動画はデコーダを通して NTSC 信号に変換し出力する。出力は、NTSC オーバーレイ・アダプタを使って PC 画面に合成する。また、通常の TV モニタにも出力できる。

画像のフォーマットは CIF(352x288) と QCIF(176x144) に対応する。フレームレートは、最大 1 秒間あたり 15 フレームに設定することが可能である。

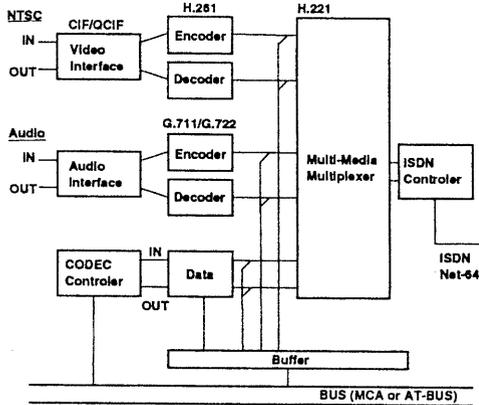


図 2: TV64 のシステム構成

TV64 は PC の拡張スロットに装着するアダプタであるため PC とのインタフェースは、システムバスを使う。マイクロチャンネル・アーキテクチャ・バスまたは AT バスを通じて、CODEC の制御とデータ入出力ができる。アプリケーション・プログラム作成のために、OS/2 でのデバイスドライバと API(Application Program Interface) を用意している。API としては、接続のためのダイアリング機能など ISDN のコントロールや動画像・音声・データのコントロールなどがある。

TV64 は、H.221 のフォーマットの状態でデータを PC に取り込んで記録・再生する機能を持つ。ISDN インタフェースから受信したデータを、バスを通して PC 側で読み込み、ハードディスクなどに保存する。画像と音声同期したままの状態での保存・再生することができる。また、通信中に保存したデータをリアルタイムで再生し、ISDN 回線に送ることも可能である。この機能を使って、動画像付き留守番電話の機能を容易に実現できる。

4 RTP におけるデータ通信

RTP においては、遠隔地間で様々なデータを送信しながら会議を行なう。

RTP では、リモートポイントと呼ぶ機能があり、共有黒板上でポイントを動かすとその動きが相手に伝わる。通常のプレゼンテーションで使用するポイントを遠隔地でもサポートする。リモートポイントを使っている時は、定期的にポイントの位置情報を遠隔地に送る必要がある。この際、ディレイができるだけ少ないことが要求される。ポイントの位置情報は 100byte 程度であり、大きなバンド幅を必要としない。

RTP の共有黒板では、プレゼンテーション資料に文字をフリーハンドで書き込む、アノテーション付ける、図形を作成する、スキャナや他のアプリケーションからビットマップを読み込む、といった操作が可能であり、これらの操作を行なうと通信が発生する。共有黒板上の操作は、そのまま相手側に伝えられ、リアルタイムで同一の内容を共有することができる。

アノテーション機能を使うと、プレゼンテーション資料に対して遠隔地の相手に注目して欲しい部分にラインマーカのように線を引くことができる。例えば、1 単語をフリーハンドで丸く囲んだ場合には、約 1Kbyte 程度の通信が発生する。

また、共有黒板に対して、図形オブジェクトを書き込むことができる。RTP の共有黒板では、線や円、四角形、矢印などの図形を書くことができる。図形を書いた場合の通信量は、約 300byte である。

共有黒板には、スキャナからイメージを取り込むこともできる。A4 1 ページの一般的なビジネス文書を 120dpi モノクロ階調なし (1 ビット) で読み込んだ場合のデータ発生量は約 180Kbyte である。これを 1/10 程度に圧縮してから送信する。

プレゼンテーションパッケージを使って会議を行なう場合には、あらかじめパッケージ全体を送信する。会議中は、ページめくりの情報のみを送ることによって情報量を減らしている。グラフィックスを使った5ページ程度のプレゼンテーションパッケージのサイズの目安は、50Kbyte から 100Kbyte 程度である。

5 バンド幅の動的割り当て

TV64 では、CCITT H.221 に準拠したマルチメディア多重化を実現している。チャンネル間同期機能によって、Net64 の 2 つの B チャンネル分の 128kbps を音声・画像・データの各チャンネルに分割して割り当てることができる。

音声部分は、CCITT G.711 および G.722 をサポートし、データレートは 48,56,64kbps から選択することができる。

データ部分は、H.221 の LSD(Low Speed Data) チャンネルを使い、300, 1200, 4800, 6400, 9600, 14400bps の中から選択することができる。画像部分は、128kbps から音声と LSD のバンド幅を引いた残りが自動的に割り当てられる。将来的には、HSD(High Speed Data) チャンネルにも対応する予定であるが、現在はデータの記録再生機能を使って B2 チャンネル全体を 64kbps のデータ用チャンネルとして利用するモードをシミュレートしている。スキャナから読み込んだイメージを転送する場合などには、この HSD チャンネルを使って高速にデータを転送する。

PC とのデータ通信はバスを通して行なわれる。300bps から 14400bps の LSD では、データは、カード内部のバッファに蓄積し、カード内部のコントロール・プログラムが送受信を行なう。PC からカード内部のバッファにデータを読み書きするための API を用意

している。一方、B1 と B2 の全てのデータを記録するモードでは、カード内部のバッファを PC のメモリ空間にマップすることにより高速なデータ転送を行なっている。

バンド幅の割り当て変更は、CCITT H.242 に準拠したプロトコルによって行なう(図 3)。送信側からバンド幅割り当て変更の BAS(Bitrate Alignment Signal) コマンドを H.221 のフォーマットで送信する。その次のデータ・フレームから新しい割り当て方式で、データを送信する。受信側では、BAS コマンドを受信すると受信モードを変更し、次のフレームから新しいフレームを受信できるようにする。受信モード変更の要求は PC 側にも通知され、画質制御のための情報として利用することができる。TV64 に内蔵したコントロールプログラムにより、PC からバンド幅変更のための BAS を受け取ると動的に割当を変更する。

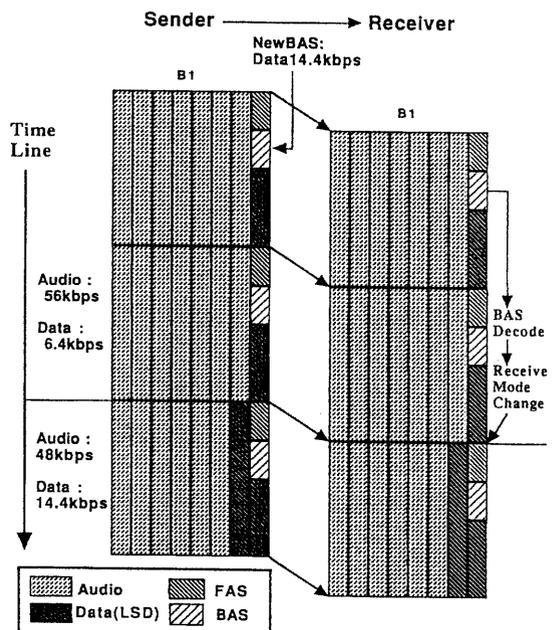


図 3: バンド幅の割り当て変更

通信中に動的にバンド幅変更を行なう場合に送信側と受信側のバンド幅割当をタイミング良く変更しないと、切替時にデータのロスが生じてしまう。BASを受信しデコードしてバンド幅割当を変更するタイミングを合わせることによって、切替時のデータロスをなくしている。

PCのデータ通信モジュールは、アプリケーションからのデータ転送要求に応じて転送遅延ができるだけ小さくなるように割当を決める。割当を変更する場合には、送信側でTV64に対して変更要求を出す。受信側では、TV64に内蔵したコントロールプログラムが、H.242プロトコルに従ったBASをデコードし、自動的に受信モードを変更する。H.221フォーマットのフレームのデコードは、TV64内部で行なうためPC側では切替に供なう処理の負荷は増加しない。また、切替時にデータロスやチャンネル間の同期はずれが起こらないようにインプリメントしているため、動画像や音声途切れたりすることはない。したがって、動画・音声を常時流したまま、共有黒板やリモートポインタのデータを効率良く転送することができる。

RTPでは、前述したように異なる性質を持ったデータを転送する。基本的には、比較的データサイズの小さいリモートポインタやグラフィックスのデータと、スキャナから読み込んだイメージやプレゼンテーション資料などサイズの大きなデータに分かれる。

通常は、LSDを利用し、データ送信リクエストの量に応じてデータ通信モジュールがデータのバンド幅を変更する。サイズの大きなデータ転送が発生した場合は、PC側のデータ通信モジュールがTV64にバンド幅変更要求を出し、HSDを利用可能にする。更に、サイズが大きなスキャナイメージを送る場合には、一時的に画像・音声を止めて128kbpsで高速転送するモードも可能である。例えば、会議の進行上必要なイメージで

あれば、そのイメージが相手に届くまで会議が進まないため、画像や音声を止めてもデータ転送を優先するモードも必要であると思われる。

画像や音声を止めている時は、動画表示用ウィンドウにメッセージを表示するなどしてユーザの混乱を防ぐことができる。また、送信と受信を非対称のバンド幅割り当てにすることも可能である。この場合、イメージの受信側では、受信状況を動画と音声を使って送信側に伝えることもできる。

6 データ通信時の画質の制御

フレームレートの制御動画・音声以外のデータの比率が大きくなると動画に割り当てるデータのバンド幅は、小さくなる。そのため画像圧縮のパラメータをデータのバンド幅に応じて変更しないと、画質の劣化が激しくなってしまう。本システムでは、画像CODECの各種のパラメータをPCから制御するためのインタフェースをインプリメントしている。画質制御のポイントは、フレームレートと画像サイズの2点である。

画像のデータバンド幅が減少した場合に、フレームレートを下げることによって画質の劣化を防ぐことができる。顔画像を伝送して、遠隔地に表情や動作を伝えている場合には、フレーム数があまりにも減少すると動画の効用があまりなくなってしまう。そこで画像サイズを変更することによって、できるだけフレーム数を落さずに画像を伝送するようなコントロールを行なうことができる。

TV64は、国際標準のH.261に準拠しているため、CODEC部分の画像サイズは、CIF(352x288)とQCIF(176x144)の2種類である。本システムは、動画像をPCのモニタにオーバーレイし、1つのウィンドウとして表示する。この際、ウィンドウのサイズは可

変である。オーバーレイボードの機能を使ってディレイなしに受信した動画を拡大、縮小して表示することができる。

画像のバンド幅が減少した時に、フレーム数の減少を抑えようとブロック化が激しくなり画質の劣化が起きる。この時に画像フォーマットを QCIF にし、動画ウインドウのサイズを小さくすることによって、ユーザの目から見た画質の劣化を抑えながらフレーム数を多くすることができる。バンド幅が変化すると TV64 から受信側 PC にメッセージが送られる。受信側 PC では、バンド幅変更のメッセージを受けると、画像のバンド幅に合わせて動画ウインドウのサイズを変更する。

また、静止画像の質が重視される文書や写真などを伝送している場合には、ソース画像を一時フリーズさせることによって、画像のバンド幅が減少した時にも、画質を保つことができる。

このようにバンド幅が減少した場合に、フレーム数を重視した画質制御や、静止画像を重視した画質制御が可能である。今後、動き情報などを TV64 から PC に通知して、自動的にモードを選択するような機能も考えていきたい。

7 まとめ

RTP において、データ転送の状態に応じて、バンド幅を変化させることにより、効率的なデータ転送を実現する方式について述べてきた。TV64 のマルチメディア多重化機能を使うことによって、PC 側の処理に負担をかけることなく、バンド幅の動的変更が実現できる。また、データ送信リクエストの状況に応じてデータ通信モジュールが自動的に TV64 にバンド幅変更要求を出すのでアプリケーションは、通信バンド幅を意識せずにデータ転送を行なうことができる。

現在の TV64 のプロトタイプでは、音声バンド幅として最低でも 48kbps を必要とする。そのため HSD 利用時には、動画像のバンド幅はかなり小さくなってしまう。Net64 の一回線のみを使ってデータ転送と動画像転送を両立させるには、音声のバンド幅をより小さくする必要がある。将来、CCITT G.728 の機能を付加すれば、音声のバンド幅を 16kbps まで削減することができる。その場合、HSD 利用時にも画質と音声のバンド幅の組合せの選択範囲が広がり、RTP の利用状況に合わせた画質の制御が可能になる。

今後、RTP でカラーの高品質イメージなどを扱うようになれば、ますますデータのバンド幅が必要となる。その場合、Net64 一回線で常時、動画像を伝送するのは難しくなると思われる。その場合には、音声の有無や動きの激しさなどの情報を CODEC から PC に通知し画像・音声の ON/OFF するようなコントロールも必要になると思われる。RTP のようなマルチメディアを使った会議システムを限られた転送容量で実現するためには、それぞれの会議の場面で必要な情報を選択して、自動的に伝送するようなくみが重要である。

参考文献

- [1] 中島, 安藤, フィン, 村上, 篠崎, 黒澤, “共有ウインドウと動画を用いた遠隔マルチメディアプレゼンテーションシステム”, 情報処理学会論文誌, Vol34-6, 1993 掲載予定