

動画圧縮とマルチメディア・システム



動画を含んだマルチメディア・システムを構築するには、動画データの圧縮技術と、スムーズな動画表示のためのシステム構成技術を考慮しなければならない。各種の画像圧縮方式が CD-ROM タイトルやビデオ・オン・デマンドなどの用途に応じて使い分けられる。システム構成の点でもディスクへのアクセス方法などに工夫が必要だ。

山内 長承, 椎尾 一郎

日本アイ・ビー・エム 東京基礎研究所

広帯域ネットワーク・アプリケーション担当, 同主任研究部員

マルチメディア技術は実用段階に入っている。企業の情報システムにおいても顧客への情報提供や従業員の教育用などで、動画を含んだマルチメディア・システムが構築され始めた。ビデオ・オン・デマンド (VOD) システムも米国で次々と実験が始まっている。

動画を含んだシステムを構築するには、膨大な動画データを保存・転送に耐えうる程度の量に圧縮する技術や、システムにかかる負荷を制御してスムーズな動画表示を実現する構成技術などを考慮しなければならない。用途と技術の特性を知れば、最適なコストで動画システムを構築することができる。

動画を表現するには大量のデータが必要になる。例えば通常のテレビ画像をデジタル・データにすると、720×480 画素を毎秒 30 コマ (フレーム) 表示しなければならず、約 160 Mbit/秒のデータ量になる。わずか 1 分の動画に約 1.2 GB の記憶容量が必要になる。このままでは伝送路も約 150 Mbit/秒と高速な ATM (非同期転送モード) の実現を待たなければならない。現時点で圧縮技術は不可欠である。

現在一般的に使用されている動画圧縮方式には、国際標準の MPEG (Moving Picture Experts Group)-1 や MPEG-2, H.261, H.263 などのほか、商品化されて普及した Cinepak

や Intel Indeo, 規格化されていないが方式の総称として定着している Motion JPEG (M-JPEG) などがある。

動画圧縮の原理

まず動画データの圧縮の原理を説明しよう。ここでは MPEG を例に説明するが、MPEG 以外の動画圧縮方式も基本的には同じような手法を使っている。

一般にデータの圧縮は、内容の冗長な部分に着目してこれを取り除く技術である。MPEG では動画を、連続する静止画像 (フレームと呼ぶ) の列として処理する。「1 つのフレームの中で隣接する画素の濃淡・色調は似ている」という空間的な冗長度と、「近いフレーム同士の画像は似ている」という時間的な冗長度の 2 つに着目して、データ量を減らしていく。

[フレーム内の圧縮]

空間的な冗長度を際立たせるために、MPEG では画像を空間領域のデータ (各画素の明暗で示す) から、周波数領域のデータ (空間内の画素の変化を表現する) へ変換する。冗長性の表現力と実現のしやすさから、MPEG では「8×8 画素ごとの離散的コサイン変換 (DCT)」という手法を使う (図 1)。空間領

域で隣接する画素が似ている画像は、周波数領域に変換すると低周波成分（図1の変換後の行列データで左上の部分）に情報が集中し、右下の高周波成分は少なくなる。

人間の目は動画の高周波成分には鈍感である。この性質を利用して、DCT変換したデータを、低周波成分には多くのビット数を割り当ててきめ細かく表現し、高周波成分のためのビット数は省いて粗く表現することで、1枚のフレームを効率的に圧縮できる。

[フレーム間の圧縮]

時間的冗長度を取り除いて圧縮するために、フレーム間の差分を利用する。動きの小さい動画では、差分画像は0に近いデータばかりになる。この差分画像に上述したフレーム内での圧縮処理を行うと大幅に圧縮できる。

動きのある部分には「動き補償」という技術を使う。圧縮しようとするフレームを16×16画素に切り分けて、元のフレームの中から類似している部分を探し出し、この部分の差分を利用する。これで動きの大きい動画でも高い圧縮率を実現できる。

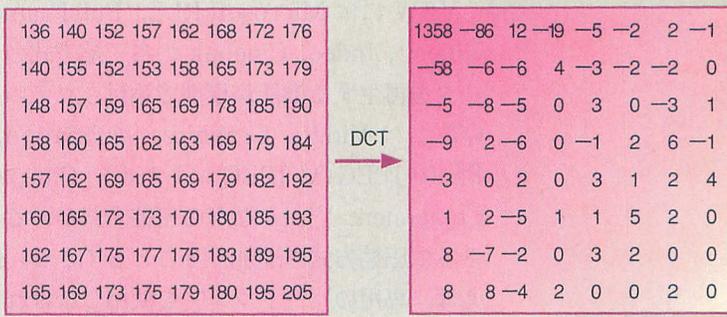
MPEGはCD-ROMタイトルのような蓄積型動画のために開発された。そこで時間的に過去のフレームだけでなく、未来のフレームも参照して圧縮する方法を使っている。

MPEGでは「I (Intra-coded) ピクチャ」、「P (Predictive coded) ピクチャ」、「B (Bi-directionally-predictive coded) ピクチャ」の3種類のフレームで動画を構成する。

Iピクチャはフレーム内のデータ圧縮だけを行ったフレームである。圧縮効率は低いが、差分を取り続けることで蓄積される誤差を解消するため、早送りなどのトリック・プレイのために必要である。

Pピクチャはフレーム内の圧縮と過去のフレームを参照した圧縮を併用したフレーム。Bピクチャはさらに未来のフレームも参照

図1 ●MPEGでは画像（フレーム）ごとに離散コサイン変換（DCT）という手法を使って圧縮する。8×8画素ごとの濃淡の階調を示したデータ（左）を右側のようなデータに変換し、情報のまばらな右下部分のデータを切り捨ててデータ量を減らせる



して圧縮したフレームだ（図2）。ただしいづれも参照するのはIまたはPピクチャであり、Bピクチャは参照には使わない。

I、P、Bの各ピクチャは任意に組み合わせることができるが、蓄積用動画の場合はIとPの間に数枚のBピクチャを挿入するのが一般的である。通信に利用する場合は、未来の参照フレームを待つための遅延を避けるため、Bピクチャを使わないことが多い。

動画圧縮方式の特性

MPEGをはじめとする各種の圧縮方式には、それぞれに適した用途がある。主な用途別に適性を見ていこう。

[マルチメディア・ソフトの動画]

パソコン用の一般消費者向けマルチメディア・ソフトでは、再生のために特別なハードウェアが必要な圧縮方式は歓迎されない。ソフトウェアだけで再生できる圧縮方式が普及している。この種の方式は再生だけでなく、比較的短時間で専用ハードなしで圧縮もできるので、エンドユーザーがプレゼンテーションに動画を取り込む用途などにも使われる。

パソコンで動画を再生するソフトとして、Video for Windows (VFW) や QuickTime (QT) が普及している。この2つは動画を

扱うための仕組みを提供するソフトウェアで、圧縮方式の名称ではないが、各種の圧縮方式に対応している。

VFW には Microsoft RLE, Full Frames, Video 1, Indeo, Cinepak の各圧縮方式のデータを再生する機能が標準で添付されている。QT には Video, Graphics, Animation, Photo-JPEG (M-JPEG にあたる), Cinepak, Component Video などの圧縮方式がある。さらに圧縮方式を追加することもできる。例えば、専用のハードウェアを使用しなければならない M-JPEG や MPEG をオプションでサポートする。

コンピュータ・グラフィックスで作った動画像には RLE や Animation が適している。実写やアニメ映画では Video 1, Indeo, Cinepak, Video が使われる。Indeo や Cinepak は圧縮率が高く、プロセサの性能に対する再生可能なコマ数や画面サイズも大きい。しかしアニメのようにベタ塗り部分の多い画像に対しては、Indeo や Cinepak ではひげ状のノイズが目立つので、Video 1 や Video が適している。

QT のファイルや VFW のファイルを再生できるソフトは、Macintosh と Windows の両方に提供されている。ただし VFW フ

ファイルを Macintosh で再生するには、一部を加工する必要がある。最近では Windows と Macintosh の両方で動作するハイブリット CD-ROM という形式でタイトルを作ることが多く、QT ファイルの方が都合がよい。

[ビデオ品質の動画]

VideoCD や VOD システムは、ビデオ・テープやテレビ放送と同等の画質の動画を供給することを狙っている。これらを再生するには現状では専用ハードウェアが必要だが、家庭市場向けには安価に再生用ハードを製造できる圧縮方式が必要である。一方、圧縮処理はサービス提供企業が行うので、ある程度時間や費用がかかってもかまわない。この用途に適しているのは MPEG 方式である。

MPEG には MPEG-1 と MPEG-2 の 2 種類の規格がある。MPEG-2 はインタレース (飛び越し走査) 方式の動画への対応などの機能が追加されたが、両者の圧縮方法は基本的に同じものだ。MPEG-1 は家庭用ビデオ・テープ並みの画質 (352×240 画素×30 フレーム/秒) で、MPEG-2 はテレビ放送並みの画質 (720×480 画素×30 フレーム/秒) で使われることが多い。圧縮後のデータ量は上記の画質の場合で、動画と音声を併せてそれぞれ約 1.5 Mbit/秒、約 6 Mbit/秒になる。

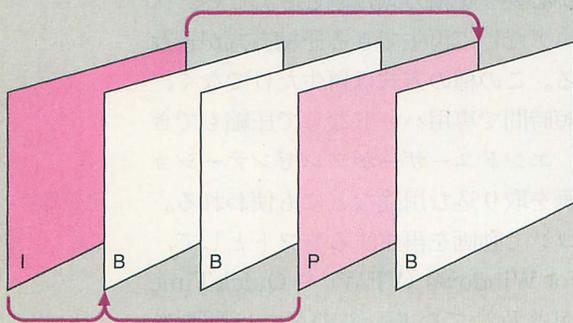
[編集段階での動画]

デジタル・ビデオ素材の編集に使う圧縮方式には、画質がよいことと、編集が容易であることが要求される。

画質を維持して圧縮する手法の 1 つに、人間の眼が鈍感な色の情報を間引く方式がある。QT の Component Video はこの方式で、編集前の素材を扱うのに適している。

コマ単位で自由に編集するには、MPEG で採用しているような他のフレームを参照する方式は適さない。静止画像の圧縮技術であ

図2●MPEGでは連続する画像(フレーム)が類似していることを使い、変化部分だけを抽出することで、さらにデータ量の圧縮を行う。一定の周期ごとの画像(Iピクチャ)は、実際の画像をDCTを使って圧縮する。数コマ先の画像(Pピクチャ)はこのIピクチャからの変化分を、さらにその間の画像(Bピクチャ)は前後のIピクチャとPピクチャとの変化分を、それぞれDCTで圧縮する



る JPEG (Joint Photographic Coding Experts Group) 方式により各フレームを独立に圧縮する M-JPEG 方式などが使われる。

[テレビ電話の動画]

テレビ電話やテレビ会議システムでは、電話回線のような低速の通信網で使用でき、対話の遅延が小さいことが要求される。この用途向けには国際標準の ITU-T H.261, H.263 規格がある。原理的には MPEG に似た圧縮方式だが、激しい動きの画像には伝送するフレーム数を落として対応できる点が特徴だ。

ビデオ・システムの構成技術

VOD などを実現するのに不可欠なデジタル・ビデオ・システムの技術は、研究から実用化段階へ進みつつある。システム全体の一般的な構成を図3に示す。実時間で途切れない動画転送を行うためには、従来のコンピュータ・システムと異なる要件を満たさなければならない。

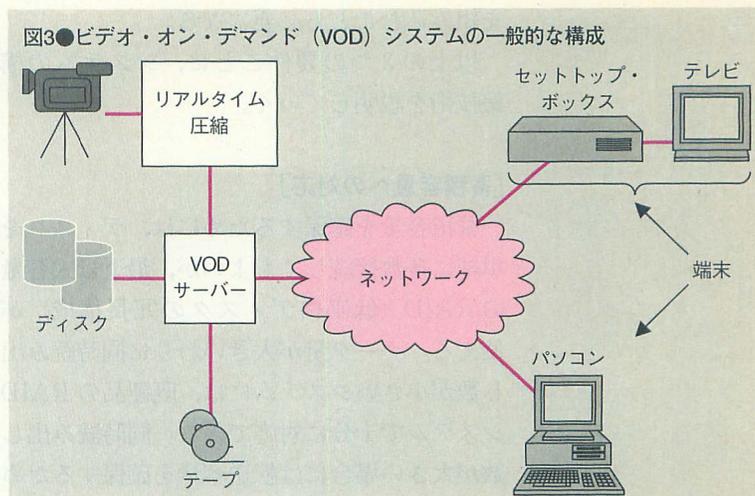
[要件1：蓄積容量]

デジタル動画は圧縮後も大量のデータである。通常 1.5 Mbit/秒程度の MPEG-1 データで1時間当たり 700 MB 弱、6 Mbit/秒程度の MPEG-2 ではその4倍の 2.7 GB に達する。1000 時間分の映画を MPEG-2 方式でディスク上に保管するには、2.7 TB もの容量が必要になる。

[要件2：転送速度]

ビデオ・データは1つの動画ストリームを転送するだけなら、あまり過大な負荷にはならない。米国で VOD の実験に使われた MPEG-1 は 1.5 Mbit/秒程度なので、標準的な LAN の通信速度より遅く、パソコンなどで実用的に使えるスピードである。

しかし多数のストリームを同時に送り出さ



なければならない VOD サーバーを構築する場合には、大きな転送容量が必要になる。同時にサポートするストリーム数は小規模な構内システムで 20~30 程度、既存の CATV システムを置き換えるようなシステムでは数千に及ぶものまで考えられる。今後画質向上のため MPEG-2 へ移行していくと、ますます大きな容量が求められる。

[要件3：転送のリアルタイム性]

動画データを途切れなく実時間で転送するには、経路の各所で発生する待ち時間を想定してリソースをスケジューリングする必要がある。例えばディスクのシーク時間、データ転送時間、プロセサの待ち時間、主記憶の空き待ち時間、共有ネットワークの空き待ち時間などを考慮しなければならない。

受信側にバッファを置いて遅れのばらつきを吸収するのは1つの解決策だ。ただしバッファを大きくすると、バッファにデータを蓄積する時間だけ遅れも大きくなる。対話型の操作 (VOD では再生や一時停止など) が多用される場合、あまり大きなバッファを持つことは望ましくない。

容量を越えたアクセス要求には接続拒否をするための管理機構 (Admission Control)

を組み込むことも必要になる。

以上の3つの要件ごとに、システムの構築技術を説明していく。

[蓄積容量への対応]

蓄積容量を確保するためには、ディスクを単純に多数接続してもよいが、最近は大容量の RAID (低価格ディスクの冗長構成) が使える。データ量が大きいわりに同時読み出し数が小さいシステムには、既製品の RAID システムで十分に対応できる。同時読み出し数が大きい場合には転送容量を確保するため、マルチプロセッサやクラスタ構成にして各プロセッサにディスクをつなぎ、システム全体で多数のディスクを収容する。

アクセスがまれなデータ、例えばめったに見られることのない映画のようなものを保管する場合、テープ・ライブラリや光ディスク・ライブラリを使った階層的な記憶装置構成でコスト削減を図る。映画などは書き換えの必要がないことも考慮してメディアを選ぶ。

階層記憶を使うとテープなどを自動装着するにも時間がかかるので、その時間を視聴者に感じさせない工夫が必要になる。

[転送速度への対応]

(a) ディスク読み出しの転送速度

同時にアクセスするストリーム数が 30 程度のシステムでは、1.5 Mbit/秒の MPEG-1 の動画データを想定すると、合計で 5.6 MB/秒の速度が必要になる。これは SCSI インタフェースで接続される市販のディスクでも十分に対応できる速度だ。

動画データをディスクに格納する場合、1度にアクセスできるセクターの長さを大きく取ると転送効率が向上する。動画再生のためのデータ読み出しは必ずセクターの先頭から最後まで順に進んでいくので、長いセクターを使えば再生時間に比べてシーク (磁気ヘッドの移動) 時間の比率が相対的に下がる。

同時に要求されるストリーム数が多く、単一ディスクの転送能力を越える場合には、複数のディスクからの同時読み出しや RAID で対応することになる。複数ディスクから同時に読み出す場合、ディスクが並列に動作し、かつデータが同時に転送できなければならない。例えば 1 台のパソコンやワークステーションをサーバーとするなら、図 4 のように複数の SCSI アダプタを装着して、複数の転送経路を確保する。

人気映画のように 1 つのストリームにアクセスが集中する場合、図 5 のように 1 本のストリームを異なるディスクに分割して格納する「ストライピング」が有効である。1 台のディスクに連続してデータを格納するとアクセス待ちが発生してしまうが、ストライピングを使うと 1 つのストリームへの同時アクセスが可能になる。

(b) システム・バスの性能

図 4 の例で SCSI の本数を増やしていくと、次にサーバーの入出力バスの転送容量がネックになる。特殊なハードウェアを使うシステム、例えばヒューレット・パッカードの

図4●ビデオ・サーバーではディスクからの読み出し経路を複数にし、同時に転送できるようにして、大量データの転送に対応する

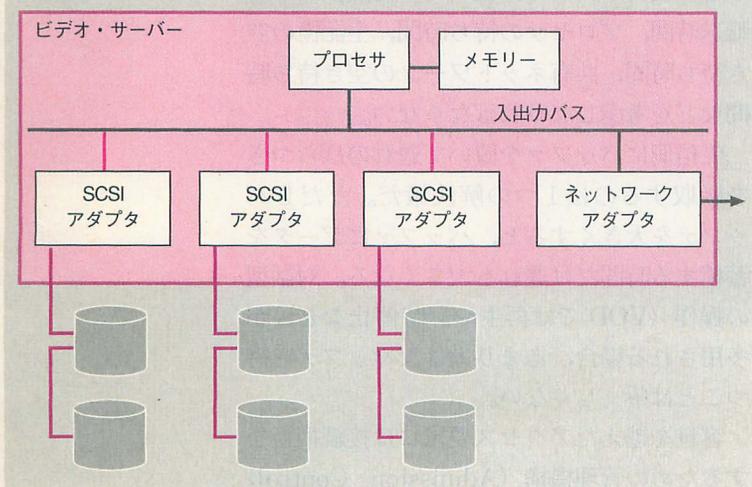
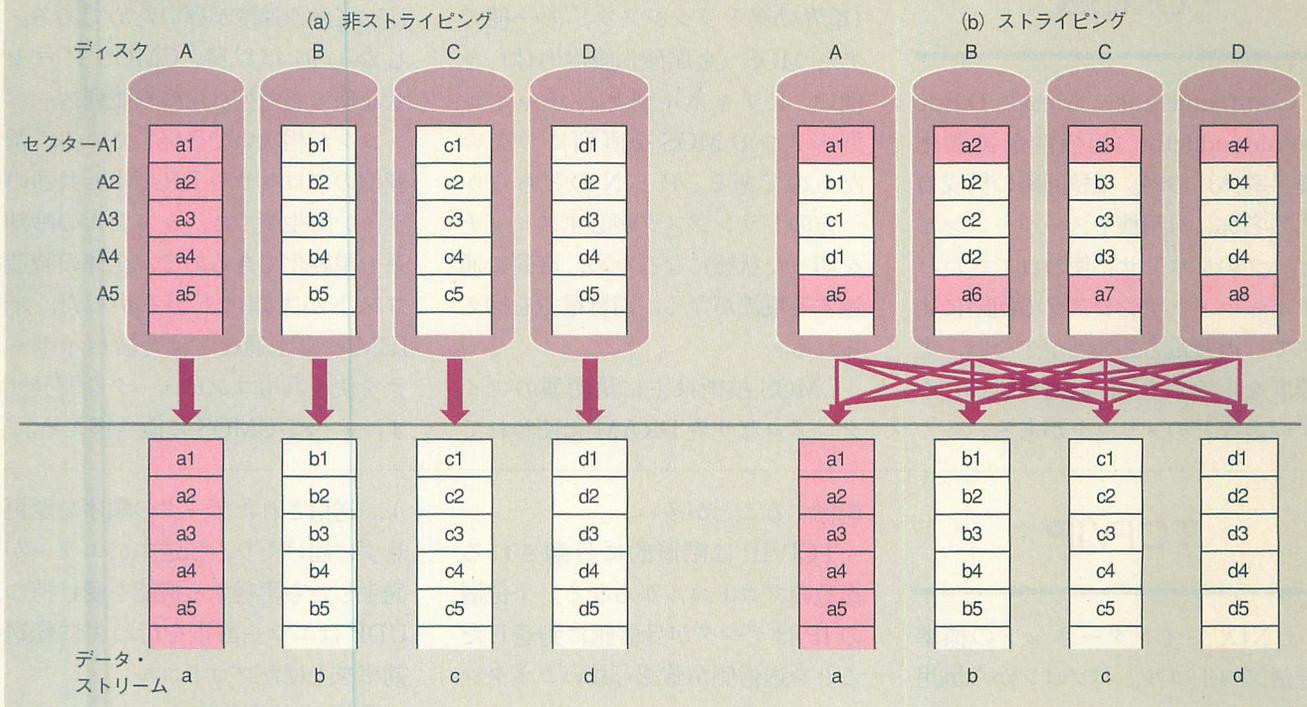


図5●1つの動画データにアクセスが集中する可能性がある場合、データをディスクにストライピング方式で書き込むと、同時アクセスの多重度を上げることができる



ビデオ・サーバーは内部バスを高速にして対応している。

汎用的なハードでシステムを構築する場合、複数のパソコンやワークステーションをつないでクラスタ構成にして、システム・バスの能力不足をカバーする。マイクロソフトのビデオ・サーバー・システム Tiger はこの手法を採用している。

(c) プロセッサ能力、主記憶容量

ビデオ・サーバーのプロセッサの仕事は、ディスクからの転送作業とネットワークへの送出程度で、負荷はあまり重くない。ただしネットワーク側がTCP/IPのLANのように複雑なプロトコル処理を必要とする場合、その処理がプロセッサの負荷になることがある。

主記憶容量への要求はその使い方に依存する。ディスクから読み出したデータを主記憶に格納（キャッシュ）しておき、複数のスト

リームにそのデータを繰り返し供給するようなシステムでは、大きな主記憶を必要とする。格納せずに単に転送バッファとして使う設計であれば、極端に大きな主記憶は必要ない。

[転送のリアルタイム性への対応]

リアルタイム性を確保するため、サーバー上ではディスク、プロセッサ、主記憶などの資源のスケジューリングと管理が必要になる。スケジューリングの技法には、制御機器などのリアルタイム処理の分野で使われてきたEarliest Deadline FirstやRate Monotonicなどの技術を適用できる。

現実のVODサーバーは既存の商用システムを使って構築することが多い。その場合、完ぺきなリアルタイム性を要求すると、OSなど基本ソフトの改造が必要になり、コストが膨大なものになる。リアルタイム性は実用レベルで妥協することが多いようである。NC