

平成 12 年度

修士論文

超高速ネットワークにおける複数
AV ファイルの同期利用技術に関する研究

Study for the applications synchronizing multi-
audio-visual files through very high-speed networks

高知工科大学大学院 工学研究科 基盤工学専攻 博士前期課程

作成者 中平 拓司

指導教員 島村 和典 教授

2001 年 2 月 9 日

要 旨

超高速ネットワークにおける複数 AV ファイルの同期利用技術に関する研究

中平 拓司

情報ネットワークの高速化と、映像の符号化技術の発展により、ネットワークを介した映像利用の機会が増えている。特に、ネットワークの高速化は急激に進んでおり、近い将来、ギガビット級さらにはテラビット級の超高速ネットワーク時代が到来しようとしている。

本論文は、そうした超高速ネットワークにおける新しい映像利用の仕方のひとつとして、複数の同期関係にある符号化された映像コンテンツを、その同期関係を保持するように AV(Audio Visual) データを転送し、利用する技術に関する研究の成果をまとめたものである。特に、同期関係にある複数の AV ストリームを、同期性を保持して再生するための同期配送方式と、複数 AV ファイルの持つ同期関係を把握し、さまざまな要求に応えるための同期管理方式を重要な研究課題とした。

複数 AV ストリームの同期配送方式については、パケット到着の相対遅延を、実ネットワークを用いて検証した。その結果、送信元が同じであれば、ネットワークの使用率が 100%を超える過負荷の状態であっても、パケット到着の相対遅延は、ほとんどが 30[ms] 以内となり、NTSC 方式の映像であれば、同期して再生されることを実証した。

一方、複数 AV ファイルの同期管理方式については、同期関係にある複数 MPEG ファイル中にある各 I ピクチャの位置を、同期参照ファイルとして別途控えておき、必要に応じてそれを参照するという方式を提案した。実験により、提案方式が、1 秒から 3 秒の応答時間で正確に同期参照できることがわかり、その有効性が示された。

Abstract

Study for the applications synchronizing multi- audio-visual files through very high-speed networks

NAKAHIRA TAKUJI

Many Audio-Visual (AV) contents have been widely used through networks because of being high-speed and progress of encoding technologies. Especially, the speed of network has been rapidly fast and it is expected that the speed may reach Gb/s or Tb/s in the near future.

This thesis shows a new application using multi-AV files as a single content in the next generation very high-speed networks. The point is that they have the synchronism. The author describes two important research issues: synchronised delivering of multi-AV streams and managing the synchronism among multi-AV files.

In the former issue, relative delay of arriving packets was investigated on the real networks. Experiment shows that the most of relative delay distributes less than 30[ms] although the utilization of network exceeds 100 percent if the streams have the same sender. Also, it is demonstrated that MPEG files encoded from NTSC format videos can play back keeping their synchronism.

The author proposed a new method to refer the synchronism among MPEG files in the later issue. The position of I-pictures in the MPEG files is written in the another file called "Synchronization reference file". It is valid that the proposal method can refer the synchronized frames within one to three seconds by experiment.

目次

第 1 章	序論	1
1.1	研究の目的	1
1.2	研究の背景	1
1.2.1	ネットワークの高速化	2
1.2.2	動画像圧縮技術の発展	9
1.2.3	ネットワークを介した映像の利用	17
1.3	まとめ	20
第 2 章	複数 AV ファイルの同期利用システム	21
2.1	複数 AV ファイルの同期利用	21
2.1.1	システムの概要	21
2.1.2	コンテンツの例	22
2.1.3	関連研究との比較	23
2.2	システム参照モデル	24
2.2.1	蓄積系	25
2.2.2	配送系	26
2.2.3	再生系	27
2.3	要求条件	28
2.4	研究課題	29
2.4.1	複数 AV ファイルの同期配送	29
2.4.2	複数 AV ファイルの同期管理	29
2.5	まとめ	30
第 3 章	複数 AV ファイルの同期配送方式の検討	31

目次

3.1	IP 網における映像利用上の問題点	31
3.1.1	パケットの損失	31
3.1.2	パケットの遅延揺らぎ	33
3.1.3	対応策	34
3.2	複数 AV ストリーム配送における同時性	35
3.3	2 つの AV ストリームの同時性の検証	36
3.3.1	検証方法	36
3.3.2	実験条件	37
3.3.3	結果と考察	38
3.3.4	LAN における同時性	40
3.4	映像を再生した検証	41
3.4.1	検証方法	41
3.4.2	実験条件	41
3.4.3	結果と考察	43
3.5	まとめ	46
第 4 章	複数 AV ファイルの同期管理方式の検討	49
4.1	エンコードによる同期管理の問題点	49
4.1.1	可変長符号化 (VLC)	49
4.1.2	VLC による問題点	50
4.2	複数 MPEG ファイルの同期参照法	52
4.2.1	提案方式の概念	52
4.2.2	I ピクチャ検出アルゴリズム	53
4.2.3	他方式との比較	54
4.3	提案方式の検証	55
4.3.1	検証方法	55

目次

4.3.2	結果と考察	55
4.4	提案方式の応用例	58
4.4.1	異なる GOP フレーム数の MPEG ファイル間の同期参照	58
4.4.2	エンコード時に同期が取れていなかったときの補正	58
4.4.3	意味単位による同期参照	58
4.5	まとめ	59
第 5 章	結論	60
5.1	複数 AV ファイルの同期配送	60
5.2	複数 AV ファイルの同期管理	61
5.3	今後の課題	62
5.3.1	複数 AV ファイルの同期配送に関する課題	62
5.3.2	複数 AV ファイルの同期管理に関する課題	63
5.4	将来展望	63
謝辞		65
参考文献		66

目次

1.1	パケットによるデータ転送	2
1.2	OSI 参照モデル	3
1.3	TCP/IP プロトコルファミリと OSI 参照モデルとの関係	4
1.4	ATM の構成	6
1.5	動画像の構成	10
1.6	動画像圧縮処理のブロック図	12
1.7	MPEG ビデオ符号化の階層構造	15
1.8	MPEG システムの多重化モデル	16
1.9	MPEG-2 PS と TS	16
1.10	MPEG のデコーダモデル	17
1.11	VoD のシステム構成	18
1.12	ダウンロードとストリーミング	19
2.1	複数 AV ファイル同期利用のシステム参照モデル	25
2.2	蓄積系の構成	26
2.3	配送系の構成	27
2.4	再生系の構成	28
3.1	パケットの損失	32
3.2	パケットの遅延揺らぎ	34
3.3	階層符号化	35
3.4	複数 AV ストリーム配送の同時性	36
3.5	2 つの AV ストリームの同時性実験のネットワーク構成	38
3.6	パケット到着率に関する特性	39

目次

3.7	パケットの遅延に関する特性	40
3.8	相対遅延の分布図	40
3.9	実映像再生実験のネットワーク構成	42
3.10	実験に用いた映像	43
3.11	実映像の再生例 (符号化レート:3Mbps)	44
3.12	実映像の再生例 (符号化レート:6Mbps)	45
3.13	実映像の再生例 (符号化レート:12Mbps)	46
3.14	実映像の再生例 (符号化レート:15Mbps)	47
4.1	ハフマン符号化の例	50
4.2	MPEG におけるフレームサイズの例	51
4.3	VLC による同期管理の問題	52
4.4	複数 MPEG ファイルの同期参照法	53
4.5	同期参照実験のネットワーク構成	56
4.6	同期参照ファイルにより指定した位置から再生したときの出力例	57

表目次

1.1	Fast Ethernet と Gigabit Ethernet	7
1.2	DSL の種類と伝送速度の例	8
1.3	MPEG 標準の位置付け	13
1.4	MPEG のプロファイル	14
1.5	MPEG のレベル	14
1.6	ネットワークを介した映像利用の例	17
3.1	2 つの AV ストリームの同時性の実験条件	38
3.2	実映像再生の実験条件	42
3.3	実映像の再生結果	43
4.1	同じ長さの符号を割り当てる例	50
4.2	GOP ヘッダフォーマット	54
4.3	同期参照検証用ファイルのエンコード条件	56
4.4	提案方式の検証結果	57

第 1 章

序論

本章では、まず研究の目的を述べ、それに至るまでの技術的背景を振り返る。ネットワークの高速化と動画像圧縮技術の進展によって、ネットワークを介した映像利用が盛んに行われるようになり、将来、複数の同期した映像利用が進むことを述べている。

1.1 研究の目的

ネットワークの高速化は急速に進んでおり、現在のキロビット (kbps) 級、メガビット (Mbps) 級の速度から、近い将来、その速度はギガビット (Gbps) 級さらにはテラビット (Tbps) 級まで到達すると予測されている。本研究は、そのような次世代の超高速ネットワーク環境における新しいアプリケーションのひとつとして、複数の AV(Audio Visual) ファイルを同期させて利用するための諸技術を確立することを目的としている。

1.2 研究の背景

本研究の背景にある技術としては、大きく分けて二つの要素がある。一つは、伝送路や交換機、端末などから成る情報ネットワーク（以下ネットワーク）の高速化であり、もう一つは、MPEG などの動画像圧縮技術の発展である。ここで、これらの技術の発展の経緯を振り返っておく。

1.2 研究の背景

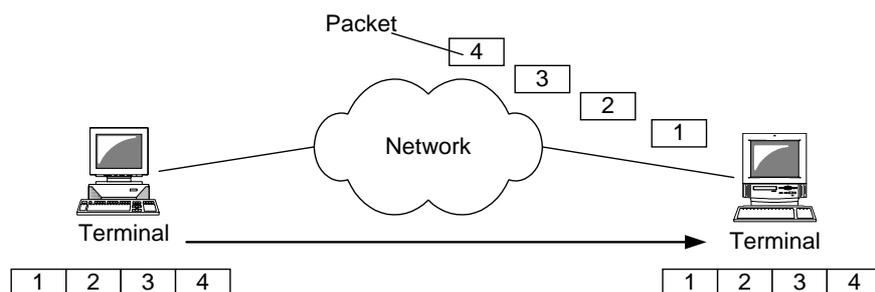


図 1.1 パケットによるデータ転送

1.2.1 ネットワークの高速化

(1) ネットワークのプロトコル

コンピュータネットワークでは，図 1.1 に示すように，送りたいデータを，「パケット (Packet)」と呼ばれる小さな塊に分解し，それを相手先まで届ける．相手先では，送られてきたパケットを集めて，再び元のデータに戻す．

このパケットをどのように作り，転送し，受信するのかというアーキテクチャとプロトコルは，当初ベンダーによってまちまちであったが，それらを統一した国際標準を作る機運が高まり，1983 年，ISO(International Standard Organization) が，OSI 参照モデル (Reference Model for Open System Interconnection) を発表した．OSI 参照モデルは，図 1.2 に示すように，通信モデルを 7 階層に分けた階層化プロトコルとなっている [1]．

ある端末から，ある端末までデータを送るとき，一番上の階層 (Layer) であるアプリケーション層のデータに対して，すぐ下の階層であるプレゼンテーション層が，それにヘッダと呼ばれる情報をデータの先頭に付加し，さらに下の階層に送る．以下，順に上の階層から送られたデータに対してヘッダを付加していき，最下層の物理層を介して，相手の端末までデータが送られる．

受信側の端末では，物理層から順に上の層に向かって，各層で処理を行い，送信側のその層で付加されたヘッダを取り除いてから上の層に残りのデータを送る．そして最終的に，受信側のアプリケーション層で，送信側が作った元のデータが取り出される．

OSI 参照モデルは，階層が多すぎて処理が複雑になることなどから，広く実装されること

1.2 研究の背景

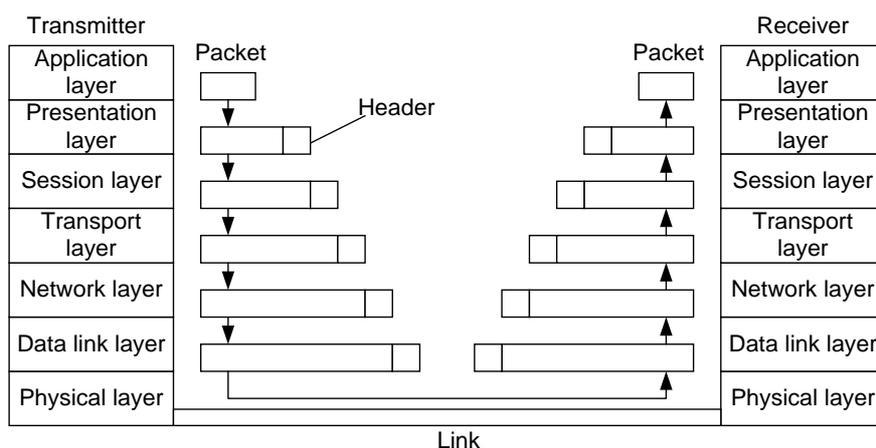


図 1.2 OSI 参照モデル

はなかったが、階層化という概念は、現在のネットワークプロトコルの事実上の標準として受け入れられている。

(2) TCP/IP とインターネット

コンピュータネットワークは、当初、その組織内だけで運用されており、様々なネットワークが独立している状態であったが、1970 年代に、米国国防総省が、軍事上の目的で、ARPA(Advanced Research Project Agency) ネットワークという、全米に広がる広域ネットワークを運用し、ネットワークの相互接続に関する研究を始めた。その成果は、現在のネットワーク技術の基盤となるものが多かったが、相互接続するためのプロトコルとして TCP/IP が確立され、現在のインターネット発展を支えている [2]。

TCP/IP とは、OSI 参照モデルの第 3 層と第 4 層である、ネットワーク層とトランスポート層に相当するプロトコル群のことを指し、TCP/IP プロトコルファミリを形成しているネットワーク層の代表的プロトコルとして、IP(Internet Protocol) があり、トランスポート層のプロトコルとして、TCP(Transmission Control Protocol) と UDP (User Datagram Protocol) の 2 種類がある。TCP/IP プロトコルファミリと OSI 参照モデルを対比させたものを図 1.3 に示す [3]。

IP 情報が、到着パケットに対して、そのパケットを次にどこに送り出せばよいかを決定、

1.2 研究の背景

OSI Reference model		TCP/IP Protocol family
Application layer		Application layer HTTP, TELNET, FTP, DNS, NFS, etc.
Presentation layer		
Session layer		
Transport layer		Transport layer TCP, UDP
Network layer		Internet layer IP, ICMP, ARP
Data link layer		Network Interface layer
Physical layer		Hardware

図 1.3 TCP/IP プロトコルファミリと OSI 参照モデルとの関係

転送する（ルーティング）。IP で扱われるパケットは、IP データグラムと呼ばれ、IP では、ルーティングした IP データグラムが正しく相手先に到達するという信頼性は提供されていない。IP の現在のバージョン 4 体系では、IP データグラムがどのアプリケーションのものを区別せずに、すべて同じ規則でルーティングする。これは Best Effort（最善努力）と呼ばれており、後の節で議論する。

通信の信頼性の提供は、TCP によって行われる。TCP では、通信する地点間に全二重の仮想回線を確立し、送信したパケットの到達確認・再送・フロー制御などを行う。ファイル転送・電子メール・WWW などは、TCP を用いている。

一方、UDP は、TCP のような再送やフロー制御などはせず、上位層から来たデータに対して最小限のヘッダを付加してすぐに IP にデータを送る。UDP で処理するパケットは UDP データグラムと呼ばれる。信頼性よりも速度を優先する場合には、一般に UDP が用いられる。後述するネットワーク上の映像配信にも、UDP を用いることが多い。

この TCP/IP により、世界中のネットワークが相互接続できるようになり、現在のインターネットが形成されるようになった。インターネット上のトラフィックは、当初はファイル転送 (FTP:File Transfer Protocol) が主であったが、現在は WWW の普及により、HTTP(Hyper Text Transfer Protocol) が大半を占めている。

1.2 研究の背景

(3) 高速ネットワーク技術

TCP/IP が汎用 PC にも広く実装されるようになると、インターネットの規模も利用者も、世界的に急増した。利用者が増加すれば、そのバックボーンやアクセス網の帯域が不足するため、帯域を上げる技術、例えば、ルータやスイッチの高速処理技術や、伝送路の高速伝送技術などが盛んに研究され、実用化が進んでいる。

この項では、そういった高速ネットワーク技術のうち、主にバックボーンとアクセス網における高速伝送技術を俯瞰する。

1. バックボーンにおける高速伝送技術

バックボーンにおいては、高速交換技術のひとつである ATM が導入されていることが多いため、ここでは、ATM について述べる。

- ATM(Asynchronous Transfer Mode) 非同期転送モード

ATM は、OSI 参照モデルのデータリンク層および物理層に相当するデータ交換技術であり [4]、図 1.4 にその構成を示す。通常の packets 交換では、packet サイズは可変であるが、ATM では、送信するすべてのデータを、「セル (Cell)」と呼ばれる 53 バイト固定長 (うちヘッダ 5 バイト) の塊に分解し、ATM スイッチにおいて、非同期に、到着したセルから順にハードウェアによりスイッチングされる。ATM では、仮想パス (Virtual Path) と、仮想回線 (Virtual Circuit) の 2 つの仮想コネクションにより相手先が特定され、セルのヘッダには、それらの識別子である、VPI(Virtual Path Identifier) と VCI(Virtual Circuit Identifier) が記述されている、ATM スイッチは、VPI と VCI によってスイッチング先のポートを一意に決定する。このように、データを固定長とし、処理を単純にすることで、スイッチング速度を向上させている。

近年のインターネットの発展により、IP データグラムを ATM 網でも転送できるようにすることが望まれ、そのための IP データグラムのカプセル化の手法が IETF(Internet Engineering Task Force) により規定され [5] [6]、IP over ATM と呼ばれている。

1.2 研究の背景

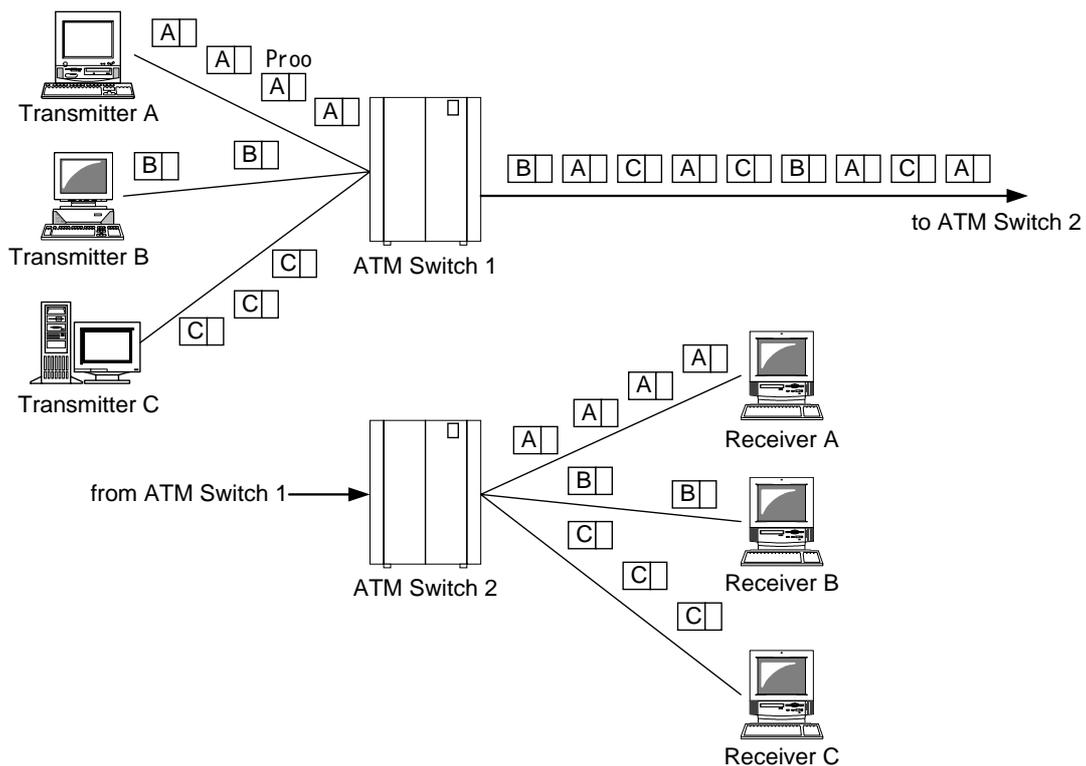


図 1.4 ATM の構成

ATM の伝送媒体には光ファイバが多く用いられ、代表的な NNI(Network Network Interface) 伝送速度は、156Mbps,622Mbps, および 2.4Gbps である。ATM は、長距離網のバックボーンや、キャンパスネットワークなど比較的広範囲の LAN(Local Area Network) のバックボーンとして導入される事が多い。

2. アクセス網における高速伝送技術

バックボーンだけでなく、アクセス網においても、様々な高速伝送技術が登場している。ここではその例として、Fast Ethernet/Gigabit Ethernet と DSL を取り上げる。

- Fast Ethernet/Gigabit Ethernet

LAN では、1976 年に発表された Ethernet が、IEEE802 委員会で標準化されて以降、現在に至るまで最も広く利用されているデータリンク層技術である。当初、Ethernet 規格の伝送速度は 10Mbps であったが、1995 年、IEEE802.3u において、伝送速度を 100Mbps に

1.2 研究の背景

表 1.1 Fast Ethernet と Gigabit Ethernet

	Fast Ethernet	Gigabit Ethernet
伝送速度	100Mbps	1000Mbps
全二重通信の可否	可能	可能
フロー制御	対称フロー制御	非対称フロー制御
主な伝送媒体	UTP	光ファイバ
主な用途	PC,WS とスイッチとの接続	スイッチ - サーバ間 スイッチ同士の接続

した Fast Ethernet が標準化された。現在は，Fast Ethernet のネットワークアダプタやリピータも安価となり，ほとんどの LAN 環境に普及している。

元々 Ethernet は，伝送媒体をスイッチ制御によらず共有していたが，Fast Ethernet では，全二重通信やフロー制御もサポートするようになり，スイッチ型のネットワークとして利用できるようになった。

さらに 1998 年には，IEEE802.3z において，1Gbps(1000Mbps) の伝送速度を目指した Gigabit Ethernet が標準化され，LAN においては，他のアクセス網技術よりも一足早く，ギガビット化が実現された [7]。

Fast Ethernet と Gigabit Ethernet の特徴について，表 1.1 に示す。Fast Ethernet は，物理媒体に非シールド撚り対線 (UTP:Unshelded Twist Pair Cable) が多く使われており，Gigabit Ethernet では，光ファイバが使われている。Gigabit Ethernet は，アクセス網技術としてのみならず，ATM に代わる LAN のバックボーンとなる可能性を秘めている。

- DSL(Digital Subscriber Line) デジタル加入者線

企業や大学などの LAN からは，10Mbps 以上と比較的高速にインターネット接続ができていたが，家庭からインターネットに接続するとなると，数年前まではアナログの電話回線で 56kbps，ISDN でも，銅線の場合，最大 128kbps と低速であった。

1.2 研究の背景

表 1.2 DSL の種類と伝送速度の例

方式	上りの速度	下りの速度
ADSL	16-640kbps	1.5-9Mbps
HDSL	1.5Mbps	1.5Mbps
VDSL	1.6-2.3Mbps	52Mbps

アナログの電話回線では、これまで音声しか伝送する必要がなかったため、人間の可聴周波数に合わせて、4kHz の信号を変調伝送していたが、専用のモデム (MODEM:変復調器) によって、より高い標本化周波数でデジタル信号を変復調することで、これまでの加入者線のままであっても、数 Mbps 程度の伝送速度を得ることが可能となった。この原理による伝送技術を総して、DSL と呼ばれている [8]。

DSL には、何種類もの技術があるが、代表的なものを表 1.2 に示す。上りは相対的に狭帯域だが、下りの伝送速度が高速な ADSL (Asymmetric Digital Subscriber Line)、1.5Mbps を目指した高速の HDSL (High bit-rate Digital Subscriber Line)、さらに下りで 50Mbps 相当の速度を持つ、VDSL (Very High bit-rate Digital Subscriber Line) などがある。これらはまとめて、xDSL と表記される。

xDSL を用いれば、銅線媒体の電話回線も、中高速のデジタル AV 情報の転送路となる。

- その他の技術

さらに、加入者宅まで光ファイバを敷設し、すべての伝送路を光ファイバにする FTTH (Fiber To The Home)、オフィスや集合住宅の近くまで光ファイバを敷設し、そこから先は無線などを使う FTTC (Fiber To The Curb)、FTTB (Fiber To The Building) など、アクセス網の主要伝送設備を光化する技術もあり、一部でその試験サービスも始まっている [9]。

伝送路の高速化に対して、ルータやスイッチなどの処理速度がボトルネックとならないよ

1.2 研究の背景

う，その高速化の技術についても研究が進んでおり，ギガビットルータやギガビットスイッチといわれる，Gbps 相当の転送能力を持ったルーティング機能製品が登場している [10]．

1.2.2 動画像圧縮技術の発展

この節では，本研究のもう一つの背景である，動画像圧縮技術について述べる．

(1) 動画像圧縮の必要性

● 画像の表現方法

コンピュータ内のデータ処理はすべて 2 進数のデジタル値で行われており，文字・音声・静止画などのメディアもデジタル化されて，コンピュータ内で処理される．動画像も同様にデジタル化される．

一般に動画像は，図 1.5 に示すように，ある瞬間を捉えた静止画像（フレーム）を，時間と共に連続して素早く切り替えて見せることで，画像内の人物や背景などの動きを表現する [11]．1 秒間に見せるフレーム数のことを，フレーム周波数と呼ぶ．日本や米国で採用されているテレビジョン信号規格の NTSC では，フレーム周波数は 29.97Hz，欧州で採用されているの PAL では，25Hz となっている．

デジタル動画像では，フレームあるいはフレームのサブセットであるフィールドという画面を，同図左側に示す画素 (Pixel) と呼ばれる小さな濃度表現点を縦横に配置し，各画素の輝度および色情報をデジタル値で表現する．

コンピュータのモニタなどでは，赤 (R)，緑 (G)，青 (B) の 3 原色の和で各画素の輝度と色を表し，通常各色は 8bit のデジタル値である．真っ黒な画素は，(0,0,0)，真っ白な画素は，(255,255,255) となる．

これに対し，テレビジョン信号では，各画素の輝度成分 (Y) と，赤との色差成分 (U)，青との色差成分 (V) の 3 成分で各画素の輝度と色を表している．これらも，8bit のデジタル値を取るが，Y は，0 から 255 まで，U と V は，-128 から 127 までの値を取る．YUV

1.2 研究の背景

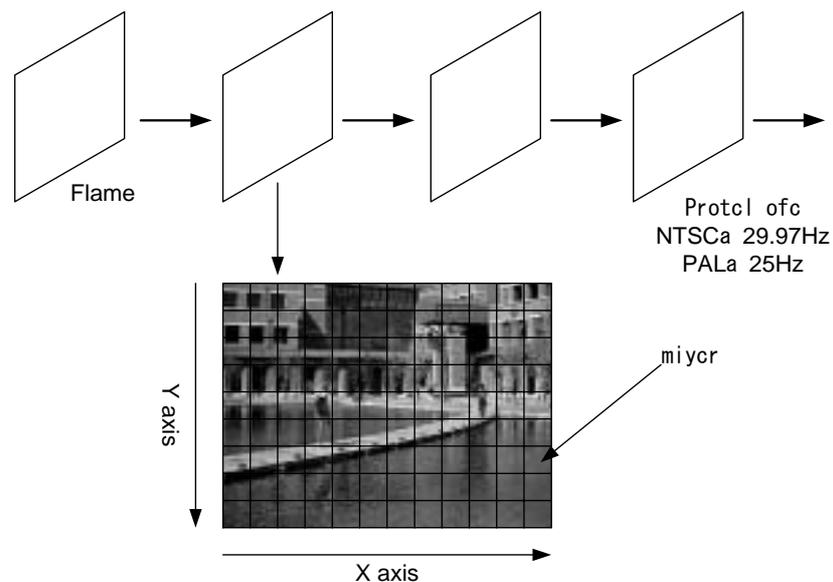


図 1.5 動画像の構成

の各成分と RGB の各成分との間には，次の変換式が成り立つ [12] .

$$Y = 0.299R + 0.587G + 0.114B \quad (1.1)$$

$$U = -0.169R - 0.3316G + 0.500B \quad (1.2)$$

$$V = 0.500R - 0.4186G - 0.0813B \quad (1.3)$$

● デジタル動画像のデータ量

デジタルの動画像は，どの程度のデータ量となるのか試算する．例えば， 720×480 画素のフレームで， 30Hz のフレームレートで動く 1 時間の動画の場合，そのデータ量 S は，次式のようになる．

$$S = (8 \times 3) \times 720 \times 480 \times 30 \times 3600 \simeq 896\text{Gbit} = 112\text{Gbyte} \quad (1.4)$$

このような大容量となれば，動画像をコンピュータで扱うことに大きな支障をきたす．1 秒間に発生するデータ量で言うと， 250Mbps 近くとなり，これをそのままネットワークを介して利用するのもまだ一般的でない．

1.2 研究の背景

ただ幸いなことに、画像は、人間の視覚特性を利用することによって、見た目の劣化は少なく、かつその情報量を大きく削減することが可能である。これが画像圧縮である。続いて、動画像圧縮における基本的な処理について述べる。

(2) 動画像圧縮の基本処理

図 1.6 に、動画像圧縮の基本処理ブロック図を示す [13]。圧縮により実際に行われるのは、元の動画像データを、別の形式のある符号に変換することである。これを符号化、またはエンコード (Encode) と言い、逆に、エンコードされた符号からもとのデータに復元することを復号、またはデコード (Decode) と言う。エンコードおよびデコードをする処理系のことを、それぞれエンコーダ (Encoder)・デコーダ (Decoder) と言い、合わせてコーデック (Codec) とも呼ばれる。

一般に符号化は、元の情報を損失せずに復号できるロスレス符号化と、元の情報の一部を失うロッキー符号化に分類できるが、動画像では、ある程度の情報が失われても、人間の視覚ではそれほど劣化を感じないため、ロッキー符号化が使われる。

動画像の符号化および復号の典型的な処理手順を、図 1.6 に示す。まず、入力信号は、A/D 変換された後 YUV 値に変換される。RGB 値を YUV 値に変換するのは、人間の視覚は、Y の変化に対して、U と V の変化には鈍く、UV を削減しても画像の劣化が少なくてすむためである。

次に、画像圧縮の重要な汎用技術となっている DCT(Describe Cosine Transform:離散コサイン変換) が施される。これにより、画像データを周波数領域に変換した後、人間の目が鈍感となる高周波成分を削除することで、情報量を減らす。またこのとき、DCT により変換された周波数成分は、低周波数成分は細かく、高周波成分は荒く量子化される。

次のエントロピー符号化は、各シンボルにそのエントロピーに近い長さの符号を割り当てる符号化であり、ハフマン符号化がその代表例である。ハフマン符号化は、出現頻度が高いデータ系列には短い符号を割り当て、出現頻度が低いデータ系列には長い符号を割り当てることで情報量を減らす可変長符号化である。

1.2 研究の背景

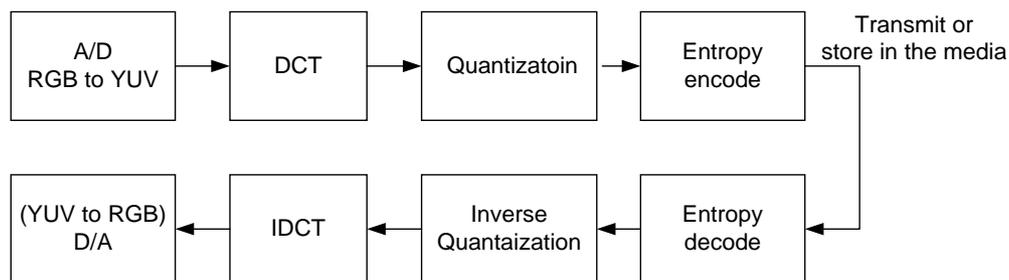


図 1.6 動画圧縮処理のブロック図

(3) MPEG(Moving Picture Experts Group)

動画圧縮には様々な種類があるが、この項では、動画圧縮の国際標準であり、ネットワーク転送対象の動画に多用される MPEG について概説する。

(1) MPEG の経緯

MPEG は、当初、74 分程度の動画を 1 枚の CD-ROM に記録することを目指して 1988 年に始まった ISO の標準化活動のグループ名で、そこで制定された標準自身を指すことが多い [14]。

MPEG 標準として、1992 年、当初の目標である、CD-ROM などの媒体への蓄積を想定した MPEG-1 が制定された。続いて 1994 年、通信・放送への応用を想定した MPEG-2 が制定された。さらに無線など低ビットレート環境への対応や、オブジェクトベース符号化などを盛り込んだ MPEG-4 が、1998 年に標準化されている [15]。

さらに 1996 年に活動が始まった MPEG-7 では、それまでとは目的が異なり、AV コンテンツを有効に検索するための記述子の標準化を目指している。MPEG-7 の正式名称は「マルチメディアコンテンツの記述インタフェース」(Multimedia Content Description Interface) である。

これらの各標準の位置付けを、表 1.3 に示す。

1.2 研究の背景

表 1.3 MPEG 標準の位置付け

標準	位置付け
MPEG-1	蓄積用, 1.5Mbps 程度
MPEG-2	通信・放送用, 6-15Mbps 程度
MPEG-4	オブジェクトベース符号化, 低ビットレート用など
MPEG-7	シーン記述インタフェースの策定

(2) プロファイルとレベル

MPEG では, アプリケーションに応じて柔軟に符号化を適用できるようにするため, プロファイル (Profile) とレベル (Level) という概念により, 仕様を分割している. プロファイルは, 符号化ツールを適当な機能単位で区切ったものであり, レベルは, 画像サイズなどのパラメータを大きさで区切ったものである. MPEG のプロファイルとレベルの種類を, 表 1.4 と表 1.5 に示す.

シンプルプロファイル (SP) は, 後述する順方向予測のみを用いたもので, メインプロファイル (MP) は, 双方向予測を用いたものである. SNR スケーラブルプロファイルは, 画質に関して階層的な構造を持ち, 空間スケーラブルプロファイルは, 画像サイズに関して階層的な構造を持つ. ハイプロファイル (HP) は, これらすべてを含むプロファイルである.

レベルは, 画像サイズにより分類され, SIF を扱うローレベル (LL), NTSC/PAL 程度の画像を扱うメインレベル (ML), HDTV を扱うハイレベル (HL) などがある. これらプロファイルとレベルのうち, 機能面でもコスト面でも, 多くのアプリケーションに対応できるように, メインプロファイルとメインレベルが定められている. メインプロファイル, メインレベルの符号化は, "MP@ML" と表記され, 他の組み合わせも同様に書かれる.

(3) MPEG ビデオ符号化

MPEG のビデオ符号化 [16] には, DCT とランレングスハフマン符号化が採用されており, フレーム内のデータのみを使って符号化するフレーム内符号化と, 手前のフレームから,

1.2 研究の背景

表 1.4 MPEG のプロファイル

プロファイルの種類	内容
シンプル	I ピクチャと P ピクチャから成る
メイン	I,P,B すべてのピクチャから成る
SNR スケーラブル	画質に関して階層構造を持つ
空間スケラブル	画像サイズに関して階層構造を持つ
4:2:2	YUV4:2:2 フォーマットを扱う
High	上記すべてを包含する

表 1.5 MPEG のレベル

レベルの種類	内容
Low	SIF を扱う
Main	ITU-R601 サイズを扱う
High1440	欧州の HDTV の画像サイズを扱う
High	米国の HDTV の画像サイズを扱う

次のフレームを符号化する順方向予測符号化，前後のフレームからその間のフレームを符号化する双方向予測符号化の 3 種類のフレームの符号化がある．これらの符号化されたフレームは，それぞれ，I ピクチャ，P ピクチャ，B ピクチャと呼ばれる．

符号化されたビットストリームは，図 1.7 に示すように，階層的に構成され，最終的には，複数のピクチャをひとかたまりにした GOP(Group Of Pictures) の繰り返しで構成される．

P ピクチャと B ピクチャに較べて，I ピクチャはデータ量が多くなるが，前後のフレームデータが欠落すると，P ピクチャと B ピクチャはデコードできなくなるため，適宜 I ピクチャを間欠的に挿入することでエラー耐性を上げ，また，ランダムアクセスにも対応できるようにしている．

1.2 研究の背景

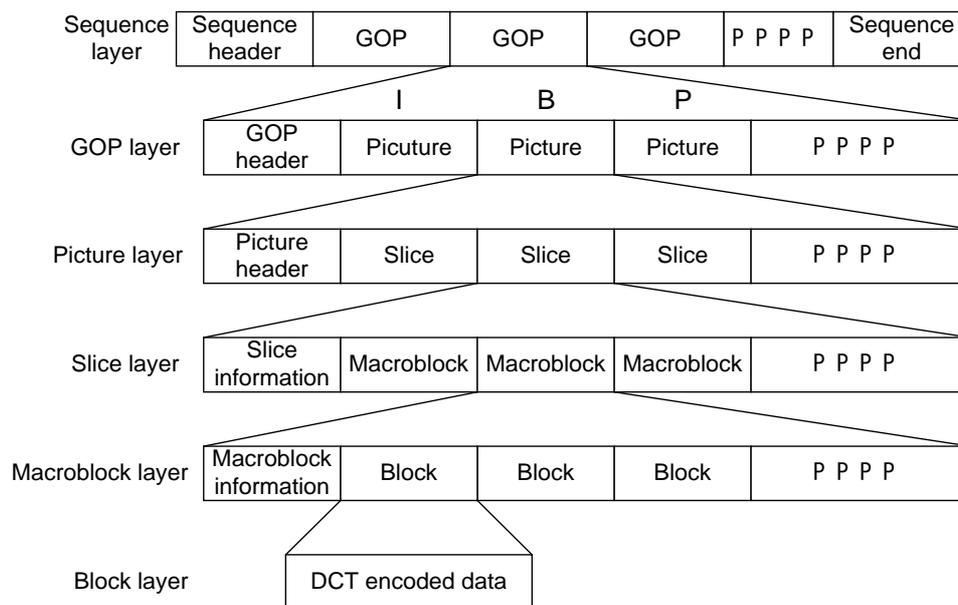


図 1.7 MPEG ビデオ符号化の階層構造

(4) MPEG システム

ビデオの符号化ビットストリームと、オーディオの符号化ビットストリームを多重化・分離するための規定が、MPEG システムである [17]。MPEG-2 におけるシステムモデルを、図 1.8 に示す。

MPEG におけるビデオとオーディオの符号化ビットストリームは、それぞれ、ビデオエレメンタリーストリーム (Video Elementary Stream)、オーディオエレメンタリーストリーム (Audio Elementary Stream) と呼ばれ、それぞれのストリームが適当な長さにパケット化される。パケット化されたストリームは、Packetized Elementary Stream (PES) と呼ばれる。その後、ビデオとオーディオの PES が 1 つのストリームに多重化されるが、MPEG-1 では、ほぼ同じ大きさになるように PES を集めて、パックと呼ばれる単位で多重化する。

MPEG-2 では、MPEG-1 と同様の多重化方式であるプログラムストリーム (Program Stream: PS) と、PES を 188byte の固定長パケットに分割するトランスポートストリーム (Transport Stream: TS) の 2 種類の多重化が規定されている。これらの関係を、図 1.9 に示す。TS において 188byte の固定長が採用されたのは、ATM のセルや、誤り訂正符号などの親和性が良く、高速通信に向いているためである。

1.2 研究の背景

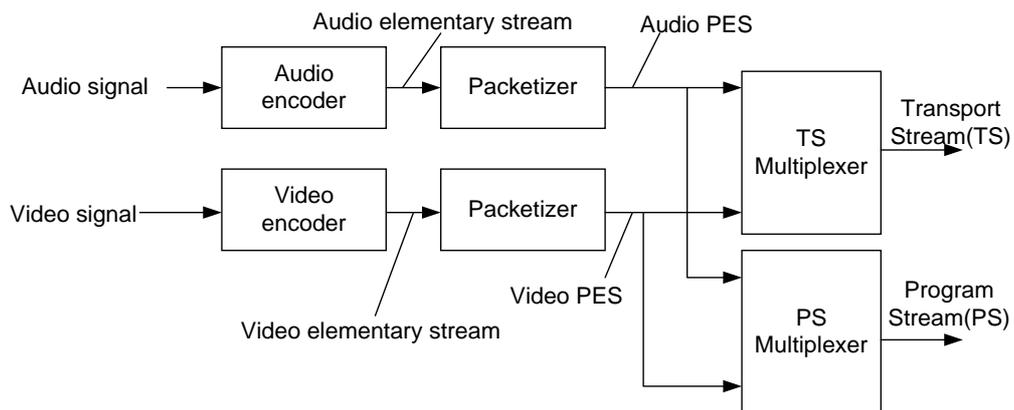


図 1.8 MPEG システムの多重化モデル

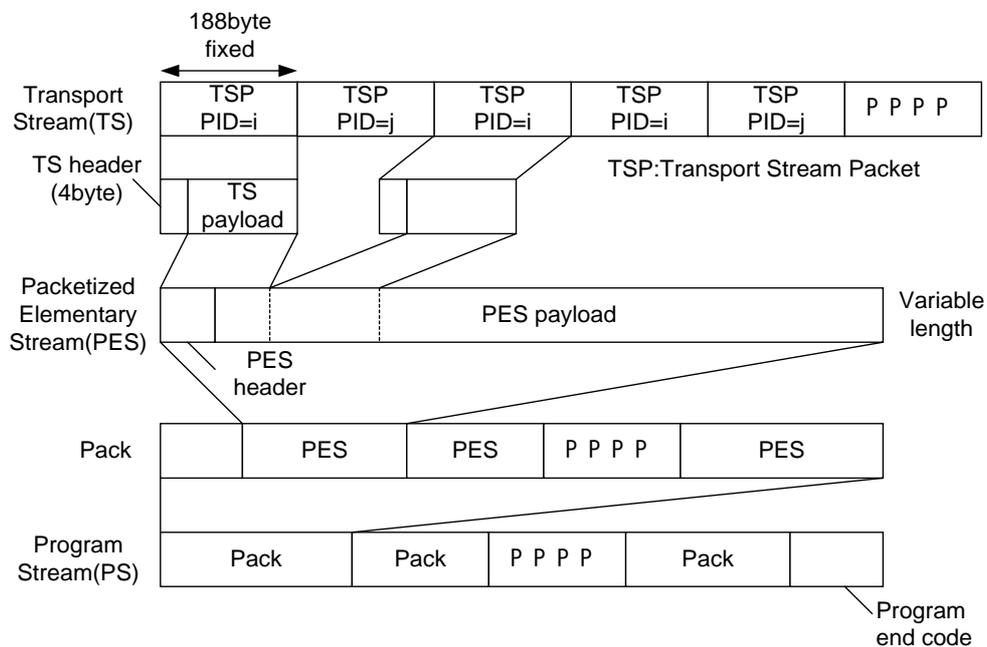


図 1.9 MPEG-2 PS と TS

一方、デコーダでは、図 1.10 に示すように、多重化されたストリームを分離し、それぞれのエレメンタリーストリームをビデオデコーダとオーディオデコーダに送り、デコードされる。ビデオとオーディオの再生の同期を取るため、MPEG では、タイムスタンプ (Timestamp) と呼ばれる時刻印をエンコード時に挿入するようになっており、デコーダでは、そのタイムスタンプを取り出して、デコードされたビデオ信号とオーディオ信号の出力のタイミングを調整する。

1.2 研究の背景

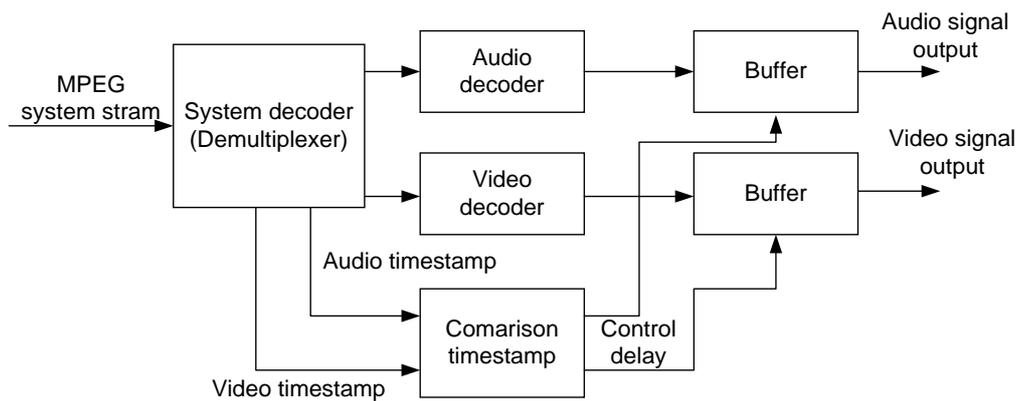


図 1.10 MPEG のデコーダモデル

表 1.6 ネットワークを介した映像利用の例

	人 対 人	人 対 システム
蓄積型	-	ビデオオンデマンドなど
リアルタイム型	ビデオ会議など	ライブ中継など

1.2.3 ネットワークを介した映像の利用

動画は、テキストや音声、静止画などの他のメディアと較べて非常にデータが大きくなるため、低速なネットワークを介して利用するのは困難であったが、前述したようにネットワークが高速になり、圧縮率の高い動画像符号化が可能になったことで、圧縮された映像をネットワークを介して利用する動きが活発になった。特にここ数年の間に、インターネットを介した映像配信のサービスが増え、低ビットレートではあるが、蓄積配信・リアルタイム配信共に実用化されている。

ネットワークを介した映像の利用例を、表 1.6 に示す [18]。この項では、本研究と関係が深い、ビデオオンデマンドと、インターネットにおけるストリーミングについて述べる。

(1) ビデオオンデマンド

ネットワークの高速化と動画像圧縮技術の進歩に加え、ハードディスクなど、その映像を蓄積する記憶装置も大容量化が進んだことで、1990 年代前半に、米国を中心として、ビデ

1.2 研究の背景

オンドemand (Video on Demand:VoD) と呼ばれるシステムが流行した。

VoD は、図 1.11 に示すようなシステム構成となっており、まず、ネットワーク上に、MPEG などで圧縮された映像ファイル (AV ファイル) を大量に蓄積する (蓄積される装置群は、ビデオサーバと呼ばれる)。ユーザは、ネットワークに接続された自分の端末から、ビデオサーバにアクセスし、見たい映像 (コンテンツ) を選択する。ビデオサーバは、ユーザ端末までその AV ファイルを送り、端末で再生される。このとき、ユーザは、要求した時点で好きな映像 (コンテンツ) を視聴できるため、”on Demand” という名前が付けられている [19]。

VoD のコンテンツとしては、映画が主流であり、米国だけでなく、日本や欧州でも様々な実験モデルが作られた。また、VoD を対象とした国際標準組織 DAVIC(Digital Audio Visual Council) もでき、標準化が行われたが [20]、当時はまだ AV ファイルをやりとりするのに十分なネットワーク速度ではなく、回線や記憶装置などのコストがまだ高かったため、結果的に広く普及していない。

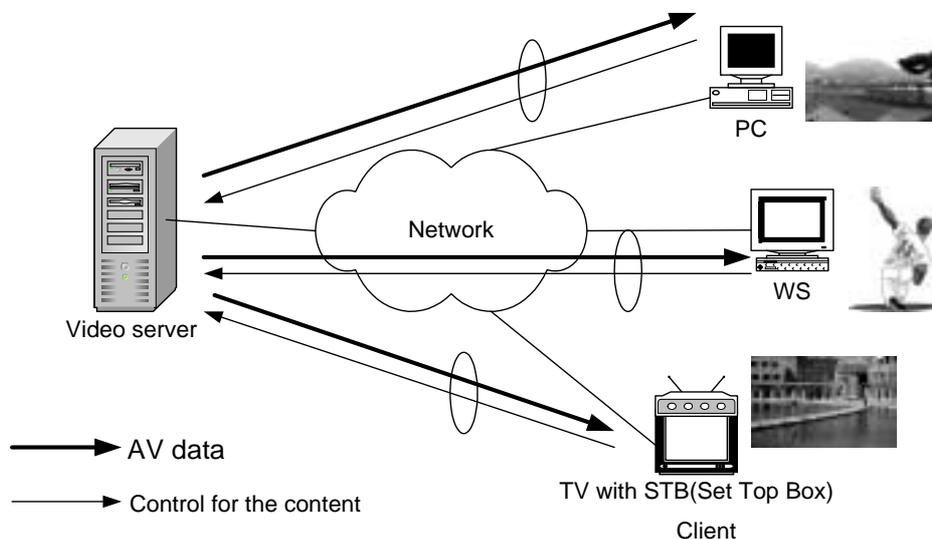


図 1.11 VoD のシステム構成

しかし、1990 年代後半になってから、ネットワークの高速化や記憶容量の大容量化が急激に進み、コストも下がってきたことから、近い将来、VoD が本格的に実施できる期待が再び高まっている。

1.2 研究の背景

(2) インターネットにおけるストリーミング

VoD が流行した当初は，ATM など，そのサービス専用のネットワークを整備していたが，1990 年代後半にインターネットが世界的に利用されるようになると，やがて，インターネットを介して音声を配信する技術が現れ始め，その後まもなく，映像を配信する技術が実用化された．

当時のインターネットでは，家庭からでは数 10kbps，LAN からでも数 Mbps 程度のスループットしか得られないのが一般的であったため，他のコンテンツのように，AV ファイルをダウンロードしてから再生するのは非現実的であった．

そこで，ダウンロードに代わる映像配信の手段として，ストリーミング (Streaming) という技術が登場した．これは，図 1.12 に示すように，AV ファイルのデータをパケットに分割してクライアントまで配信を始め，クライアントでバッファリングする．そして，バッファにある程度データがたまったら再生を開始し，並行してネットワークからパケットを受信し続けることで，連続して AV ファイルが再生できるようにしたものである．

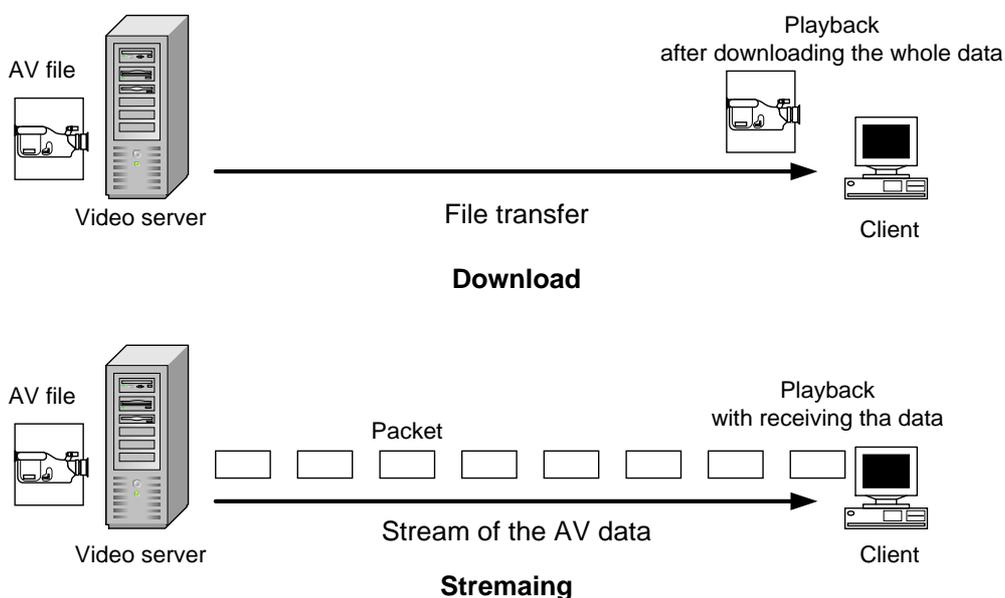


図 1.12 ダウンロードとストリーミング

これにより，低ビットレートでもそれに合わせたエンコードをすることで，リアルタイム

1.3 まとめ

に近い状態で映像を視聴することが可能となった。また、蓄積配信だけでなく、イベントの生中継などのリアルタイム配信も実用化されている [21]。

映像配信では、ある時刻までに目的地までパケットを到着させることが重要であるため、処理が単純で処理遅延の少ない UDP を用いるのが普通である。IETF では、映像配信など、リアルタイム性が要求される通信のトランスポートプロトコルとして、RTP(Realtime Transport Protocol)[22] を規定しており、実際にインターネットでも使われている。

RTP では、シーケンス番号やタイムスタンプなど、映像などの実時間通信向けのヘッダフォーマットを規定しているが、ペイロードフォーマットは、個々の符号化フォーマットごとに別途規定されている [23] [24]。

1.3 まとめ

本章では、研究の目的を述べた後、それに至る背景技術として、ネットワークの高速化と、動画像圧縮技術の進展を取り上げた。これらの要因から、ネットワークを介した映像利用が広く行われるようになり、その例として、ビデオオンデマンドと、インターネットにおけるストリーミングを紹介した。

特に、インターネットが急激に普及したことから、IP の上ですべてのアプリケーションを利用する動きが進み、映像配信はその最右翼として研究・開発が進められ、低ビットレートでの実用化もされた。今後もネットワーク技術の発展は急速に進むことが予想されており、本格的に映像を利用できる日もそう遠くはないと筆者は考えている。

第 2 章では、本研究の主題である、次世代の超高速ネットワーク環境によって実現可能となる、複数 AV ファイルの同期利用システムについて、議論を展開する。

第 2 章

複数 AV ファイルの同期利用システム

本章では，本研究の主題である，複数 AV ファイルの同期利用についてまず説明し，これまでの研究との関連性などを議論する．次に，システム参照モデルを規定し，要求条件を整理した後，そこから生じる研究課題を導く．

2.1 複数 AV ファイルの同期利用

第 1 章で述べたように，ネットワークの高速化に伴い，新しい映像の利用技術に関する研究が活発に行われ，また，実用化もされるようになった．高速ネットワーク技術の発展は目覚しく，近い将来，ユーザが，End-to-End で 100Mbps オーダ，さらには 1Gbps オーダの超高速ネットワーク環境に常時接続できる，ブロードバンド時代の到来が現実味を帯びてきている．

本研究では，そのような超高速ネットワークにおいて，その性質を活かした，従来にない新しい映像利用のシステムとして，「複数 AV ファイルの同期利用」を目指した．

2.1.1 システムの概要

ネットワークを介した映像利用のシステムとしては，先に述べたように大きく分けて蓄積配信型と，リアルタイム配信型とに分かれるが，本研究では，主に蓄積配信型の新しいシステムとして，複数 AV ファイルの同期利用技術を研究の対象とした．

2.1 複数 AV ファイルの同期利用

複数 AV ファイルの同期利用とは、ある事象、すなわち後にユーザに提供するコンテンツを、エンコードされた AV ファイルとしてネットワーク上のサーバに蓄積するとき、そのコンテンツを、同期関係にある複数のエンコードされた AV ファイルの集合から成るようにし、ユーザが、任意の時刻にそれら複数の AV ファイルから成るコンテンツを利用するというものである [25]。

従来のビデオオンデマンドなどのコンテンツ配信では、1 ユーザが同時に 1 つの映像を利用していたが、このシステムでは、同時に複数の映像を、それらの同期関係を保持しながら利用する点が特徴である。

ここでは、このシステムによって提供・利用可能となる新しいコンテンツの例と、他の研究との関連性について述べる。システムの具体的なモデルについては、次節で述べる。

2.1.2 コンテンツの例

この項では、そういった複数の AV ファイルを同期利用するコンテンツの例を幾つか取り上げる。これらは、お互い組み合わせることも可能である。

(1) 多視点映像

例えば、現在のスポーツ中継などでは、広い競技場内の様々な箇所を写すカメラや、注目選手をずっと追跡するカメラなどを置いて、番組の進行に合わせて放送するカメラショットを切り替えている。そのような多視点から捉えた複数の映像は、互いに同期関係にあるため、これらの映像を同期関係にある AV ファイル群とすれば、エンドユーザは、好きな視点からの映像を、好きなタイミングで変えながら再生できる。

(2) 多地点同時進行映像

複数の地点で同時進行であるイベントが行われたり、各地点間のコミュニケーションが図られているような場合、各地点の映像をそれぞれ AV ファイルとして蓄積すると、それらは同期関係にある複数の AV ファイル群となる。例えば、日本と米国を結んだ国際会議や、日

2.1 複数 AV ファイルの同期利用

中韓 3 大学を結んだ合同授業など，多地点事象を同期して視聴させ得る．

(3) 広角映像

ある風景を，複数のカメラで別々の角度で撮影しておき，後でそれらを繋ぎ合わせてパノラマ画像を生成するという研究例がある．これを，動画にも拡張すれば，同期関係にある複数の AV ファイル群として利用できることになる．史跡や名所などの紹介に用いると，その美しさや迫力がいっそう増して見えることになる．

また，類似の研究として，魚眼レンズを 2 つ用意して全方位の映像を生成するものもあり，これを動画にも適用できれば，180 度パノラマ映像だけでなく，360 度全方位の映像も利用できるようになる．

(4) 複数ビデオオブジェクト

これまでの映像は，フレームという矩形領域に入っている物（オブジェクト）をひとつの画としていたが，その中にも複数の動く物，ビデオオブジェクトを配置し，それらを同期させて提示する．例えば，PC を併用するある講義を，講師・講師の PC 画面・白板などをそれぞれ AV ファイルとして蓄積しておき，それらを同期させて提示すれば，好きなタイミングで好きな箇所を拡大するなどして，その講義に実際に出席したときと同様の感覚で見ることができる．

MPEG-4 では，このようなオブジェクトベースの符号化が特徴となっているが，ひとつの動画に，複数の静止画やテキストや音声が入るといったモデルであるため，この例は，MPEG-4 の発展形とも言える．

2.1.3 関連研究との比較

本研究に限らず，ネットワークを介した映像利用に関しては，蓄積配信型，リアルタイム配信型，コミュニケーション型を問わず，既に多くの研究例があり，実用化も進んでいる．

2.2 システム参照モデル

特に、[26],[27],[28],[29] などでは、本研究と同様に、複数の AV データを 1 ユーザが利用するという観点に立って研究が行われている。

本研究と同様の映像コンテンツ配信型の研究動向を見ると、帯域の向上に伴ってより高精細・高画質の映像を送ろうとするのが主流となっている。例えば、最近の民生用カメラなどにも導入されている DV(Digital Video) フォーマットの映像を、広帯域のネットワークで伝送した例などがある [30]。

DV は、MPEG のような順方向予測符号化や双方向予測符号化はせず、すべてのフレームに、フレーム内符号化が使われているが、さらに広帯域のネットワークを用いて、D1 と呼ばれる 270Mbps 相当の非圧縮の映像を、TCP/IP によって伝送することに成功したという報告もされている [31]。

こういった研究に対する本研究の意義としては、

- ギガビット級の超高速ネットワークの利用の仕方として、高精細・高画質以外の新しい映像利用の方向性を示している。
- ただ配信するだけでなく、同期関係を保持させることを目指している。
- 映像コンテンツの在り方そのものに新しい観点を持ちこんでおり、情報ネットワーク上だけでなく、既存の放送番組などにも変革をもたらす可能性を持っている。
- 蓄積配信だけでなく、リアルタイム配信型やコミュニケーション型のアプリケーションへの応用が期待できる。

などが挙げられる。

2.2 システム参照モデル

研究の第一段階として、対象とするシステムの参照モデルを、図 2.1 のように規定した。モデルは、(1)AV ファイルを生成してサーバに蓄積しておく蓄積系、(2) 蓄積されたコンテンツをサーバからユーザ端末まで届ける配送系、(3) 届けられた AV ファイルデータをデコード、再生する再生系の 3 つのサブシステムを持つ。これら 3 つのサブシステムについ

2.2 システム参照モデル

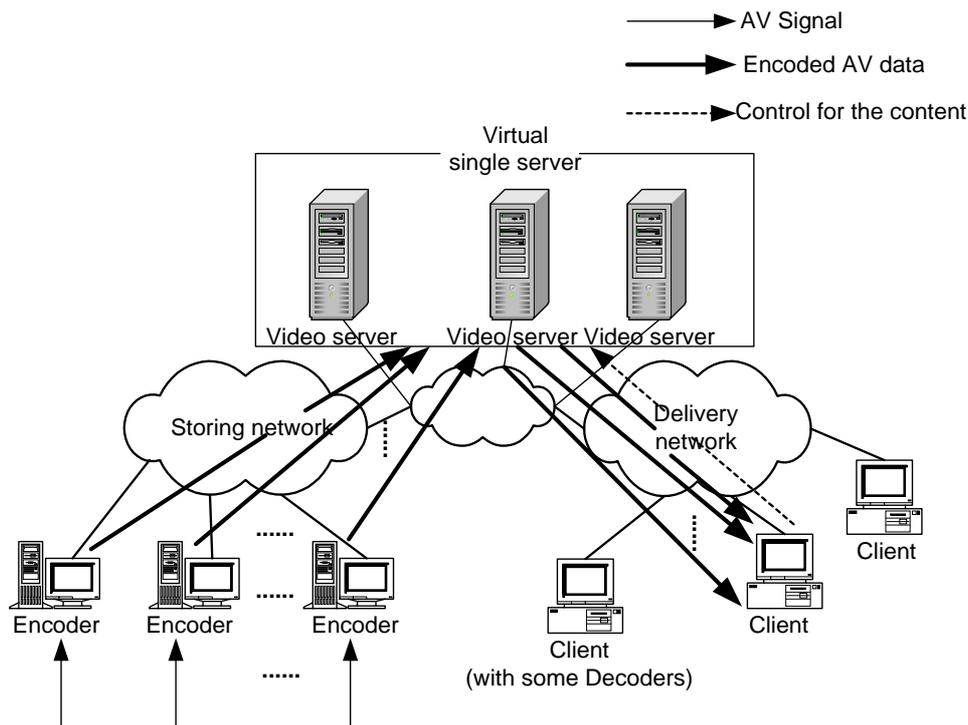


図 2.1 複数 AV ファイル同期利用のシステム参照モデル

て、順に構成を示す。

2.2.1 蓄積系

蓄積系は、エンドユーザに提供されるコンテンツを生成・蓄積するサブシステムである。図 2.2 に示すように、コンテンツを複数のカメラで撮像する撮像系、あるいはテープなどの媒体に記録する記録系と、それから出力される AV 信号をエンコードして複数の AV ファイルを生成するエンコーダ系、AV ファイルを蓄積・管理するビデオサーバと、そこまで転送する蓄積転送ネットワークによって構成される。エンコーダ系は、生成する AV ファイルの数だけのエンコーダを持つ。

AV ファイルをビデオサーバまで転送する方法には、すべてのエンコードが終了してファイルが出来上がってから、ビデオサーバまでファイル転送する方法と、リアルタイムにエンコードされたデータをビデオサーバに転送し、エンコード終了時に、AV ファイルをビデオサーバ内に生成する方法がある。

2.2 システム参照モデル

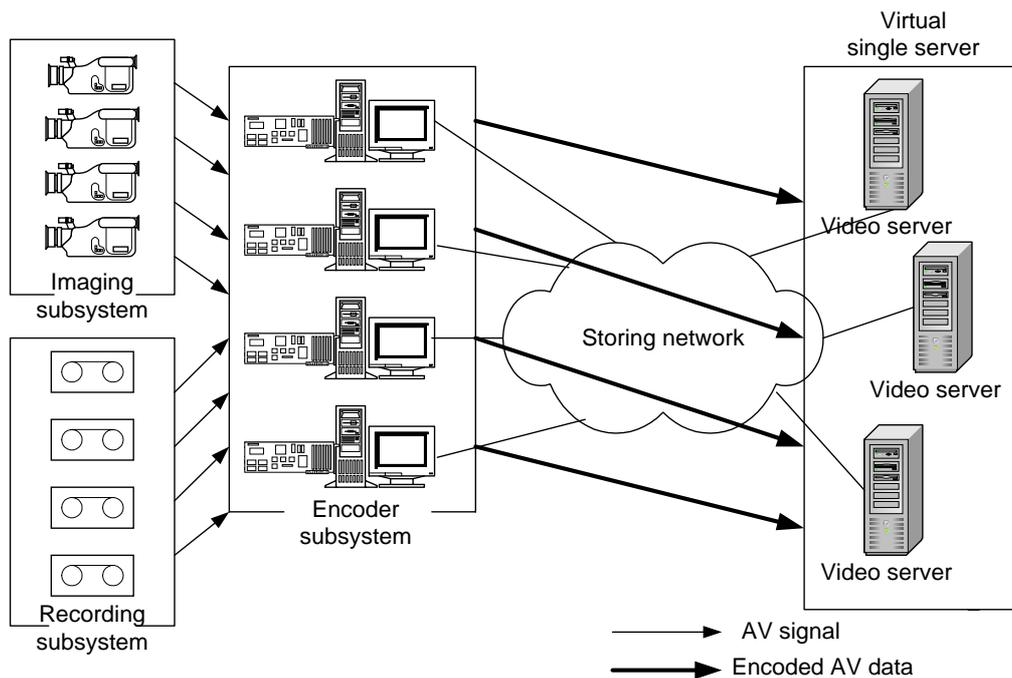


図 2.2 蓄積系の構成

2.2.2 配送系

配送系は、蓄積されたコンテンツを、エンドユーザからの要求に応じて、後述の再生系まで配送するサブシステムであり、図 2.3 に示すように、ビデオサーバとエンドユーザ所有の端末、それを結ぶネットワークから成る。

ビデオサーバは、単一で構成されることもあるが、負荷分散、コンテンツ発生場所の分散などの理由からより一般的な状況を考え、複数に分散されると規定した。ただし、エンドユーザからは、単一のサーバからコンテンツを提供してもらうように見せ、分散していることは意識させない。このことを、サーバの仮想単一化と呼ぶ。

コンテンツの配送は、エンドユーザの端末からの要求をビデオサーバが受け取ってから行われる。端末には、複数のデコーダが導入されており、同時に再生できる最大の AV データ数は、デコーダの数と等しい。

ビデオサーバと端末を結ぶネットワークは、長距離で、伝送速度は、バックボーンで数 Gbps 以上、アクセス網で 100Mbps 以上の高速ネットワークとなっており、ネットワーク

2.2 システム参照モデル

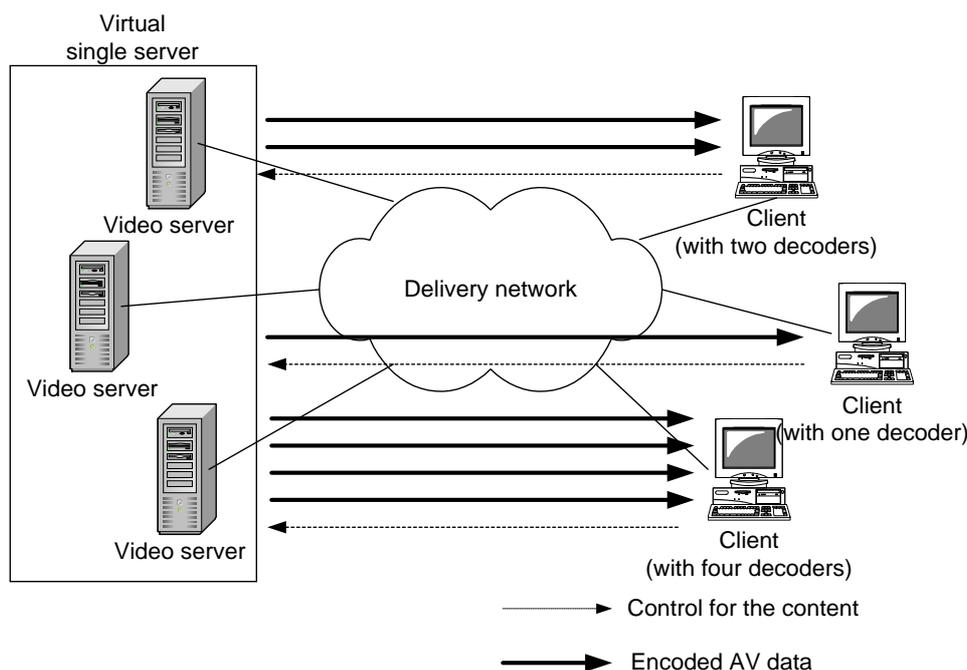


図 2.3 配送系の構成

層は IP で統一されているものとする。ゆえに、高速ではあるが、ネットワークの QoS は必ずしも保証されていない。コンテンツの配送時には、AV データを端末まで届ける下りチャンネルと、端末からビデオサーバに様々な制御命令を送る上りチャンネルが割り当てられる。

2.2.3 再生系

再生系は、エンドユーザが所有する端末とその周辺装置、ソフトウェアから成り、図 2.4 に示すように構成される。端末は、複数のデコーダを持ち、ビデオサーバから配送された AV データをデコードする。デコードされた映像は、ビデオディスプレイおよびスピーカなどに出力される。

ビデオサーバへの接続とコンテンツの要求、デコードされた信号の出力は、端末にインストールされたソフトウェアによって実行される。再生中のコンテンツに対しては、ビデオサーバに対して様々な制御を on-Demand に要求でき、そのためのユーザインタフェースも備えている。

2.3 要求条件

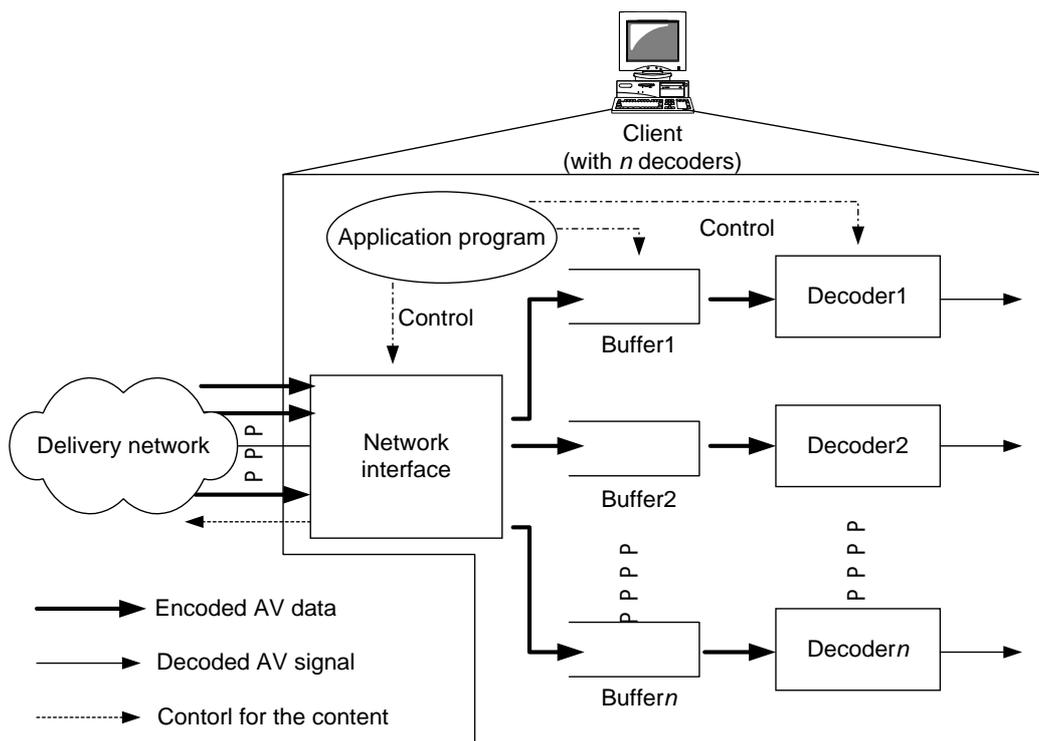


図 2.4 再生系の構成

2.3 要求条件

システム参照モデルを規定した次の段階として、このシステムへの要求条件を以下のように定義した。

- エンコーダは、標準に準拠したものを使用し、生成された AV ファイルは、標準に準拠した状態で同期利用できること。
- AV ファイルは、他の AV ファイルと同期利用できると共に、単独のファイルとしてもデコード・再生できること。
- 一旦蓄積されたコンテンツ (AV ファイル群) は、任意の時刻のエンドユーザからの配送要求を受け付け、配送を開始すること。
- エンドユーザ所有の端末まで届けられるコンテンツは、その同期関係を保持しながら再生すること。
- デコーダは、標準に準拠したものを使用し、同時に再生できる AV ファイル数は、端末

2.4 研究課題

が持つデコーダの数と等しくなること。

- 再生中のコンテンツに対するエンドユーザからの制御として、再生位置の変更・再生する AV ファイルの変更・再生の中断・再開ができ、これらの制御を行った場合にも、同期関係を保持して再生を続けること。

2.4 研究課題

三つ目の段階として、現状のネットワーク技術および映像の符号化技術をそのまま適用したとき、これらの要求条件を満足するかどうか検討した。その結果、(1) 配送系における、複数 AV ファイルの同期配送、(2) 蓄積系における、複数 AV ファイルの同期管理、の 2 つを、最初の重要な研究課題とした。

2.4.1 複数 AV ファイルの同期配送

IP など、Best Effort 型のプロトコルで映像配信を行うとき、エンドユーザに対しては、ネットワークの品質が保証されないため、一つの AV ストリームに対して、そのパケットの損失や遅延揺らぎを抑えて、滑らかな再生をする技術が盛んに研究されている。本研究では、同期関係にある複数の AV ストリームを配送しなければならず、さらに複数ストリーム間の同期配送を保証するという、新しい技術を考える必要がある。

2.4.2 複数 AV ファイルの同期管理

要求条件にあるような再生中のコンテンツに対する制御を行ったり、また、再生される映像の同期関係が崩れたときの再同期を行うためには、あらかじめ、ある AV ファイル中のあるデータに同期している、あるいは同期させるべき、他の AV ファイル中のデータがわかるような、複数ファイル間の同期管理を保證する方式が必要となる。

2.5 まとめ

本章では、まず最初に、第 1 章で述べた技術的背景から出てくる新たな映像利用技術のひとつとして、複数 AV ファイルの同期利用を研究の主題に選んだことを述べた。続いて、そのシステム像と応用例、他の研究との関係について整理し、システム参照モデルを定義した。

定義されたシステム参照モデルは、AV ファイルを生成・蓄積する蓄積系、ユーザ端末まで届ける配送系、再生とその制御を行う再生系の 3 つのサブシステムから成り立つ。各サブシステムごとに条件を仮定し、システムへの要求条件を整理した。

最後に、要求条件に対して、配送系において、蓄積された複数 AV ファイルデータの同期配送技術が必要となること、蓄積系において、それら AV ファイル群の同期管理技術が必要となることを述べた。続く第 3 章と第 4 章では、前者の同期配送技術、後者の同期管理技術について議論する。

第 3 章

複数 AV ファイルの同期配送方式の 検討

本章では、最初に、IP 網における映像利用上の問題点について整理し、その後、配送系における複数 AV ファイルの同期配送方式について、シミュレーション実験と実装による評価を行う。

3.1 IP 網における映像利用上の問題点

インターネットが広く普及したことで、あらゆるネットワークを IP ベースのものにして運用・管理しようとする動きが進んでいる。例えば、従来の電話網に対して、IP 網で音声を転送し、会話をする Voice over IP (VoIP) という技術も出現している。

先に述べたように、現在の IP (バージョン 4) は、信頼性を提供しないプロトコルであり、アプリケーションの区別なく、すべての IP データグラムを同じ規則に従ってルーティングする、Best Effort 型のプロトコルである。このため、映像を利用、特に遠隔地の端末で映像を再生させるときに、以下の問題が生じる [32]。

3.1.1 パケットの損失

一般にルータの性能は、そのルータが、単位時間あたりに何パケットのルーティング処理を実行できるかで論じられることが多い。理想的には、入力ポートに到着するパケットは瞬時に出力ポートに転送されるが、実際には、転送処理時間の分だけ待たされるため、この値

3.1 IP 網における映像利用上の問題点

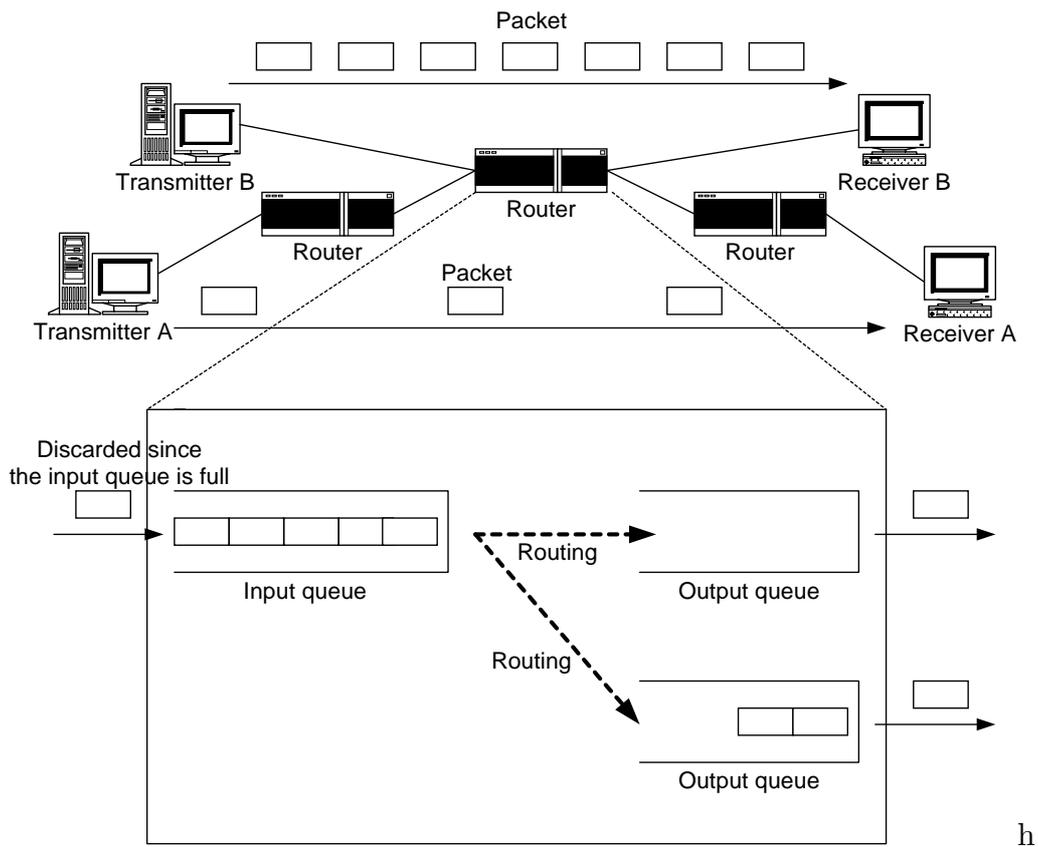


図 3.1 パケットの損失

は有限となる。

パケット転送のキュー管理は、図 3.1 に示すように、各ポートには、通常キュー (Queue) が用意されており、ある時間帯に集中して同じポートにパケットが到着しても、それらをキューイング (Queuing) するようになっている。IP では、IP データグラムを区別をせずに、同じポートに到着したデータグラムは、同じキューに入れるため、キューの容量を越えるパケットが到着した場合は廃棄される問題がある。

このため、ネットワークの利用率が上がってくると、ストリーミングをしている間に、受信端末に至る途中のルータでパケットが損失することになる。その結果、再生時の映像にノイズが入ったり、MPEG であれば、I ピクチャが損失し、続く P ピクチャや B ピクチャの再生ができないという事態が生じる。

3.1 IP 網における映像利用上の問題点

3.1.2 パケットの遅延揺らぎ

ネットワークを經由してパケットを送信する場合，送信端末におけるパケットの送信時刻を t_s ，受信端末における受信時刻を t_r とすると， t_r は，次式で表される．

$$t_r = t_s + t_t + t_p \quad (3.1)$$

ここで， t_t は，パケットの送信所要時間， t_p は，パケットの伝搬遅延時間であり，それぞれ，

$$t_t = S/B \quad (3.2)$$

$$t_p = D \times L \quad (3.3)$$

と表される．ただし， S は，パケットサイズ (bit)， B は，伝送速度 (bit/s)， D は，伝送路 1 km 当たりの伝搬遅延 (s/km)， L は，伝送路の距離 (km) である．

ゆえに，パケットの到着時刻はネットワーク条件で一意に決まり，連続してパケットを送信した場合，受信側での到着間隔は，送信側での送信間隔と等しくなる．

しかしながら，ネットワークの利用率が上がり，あるルータのキューが混雑すると，図 3.2 に示すように，損失はしなくても，その混雑具合によってパケット到着までの遅延が大きくなる．逆に，混雑が解消されると，到着までの遅延は小さくなる．このように，ネットワークの状態によって，パケットの到着遅延が揺らぐことになる．この揺らぎのことを，遅延揺らぎ (Delay Jitter)，または単にジッタと言う．

映像は，決まったフレームレートで再生されるため，このようにパケットの遅延揺らぎが生じると，再生時の映像がぎこちなく動いたり，コマ落ちしたように見えてしまう．例えば NTSC の映像であれば，毎秒約 30 フレームで再生するため，約 33ms 以内に次のフレームデータを到着させる必要がある．

3.1 IP 網における映像利用上の問題点

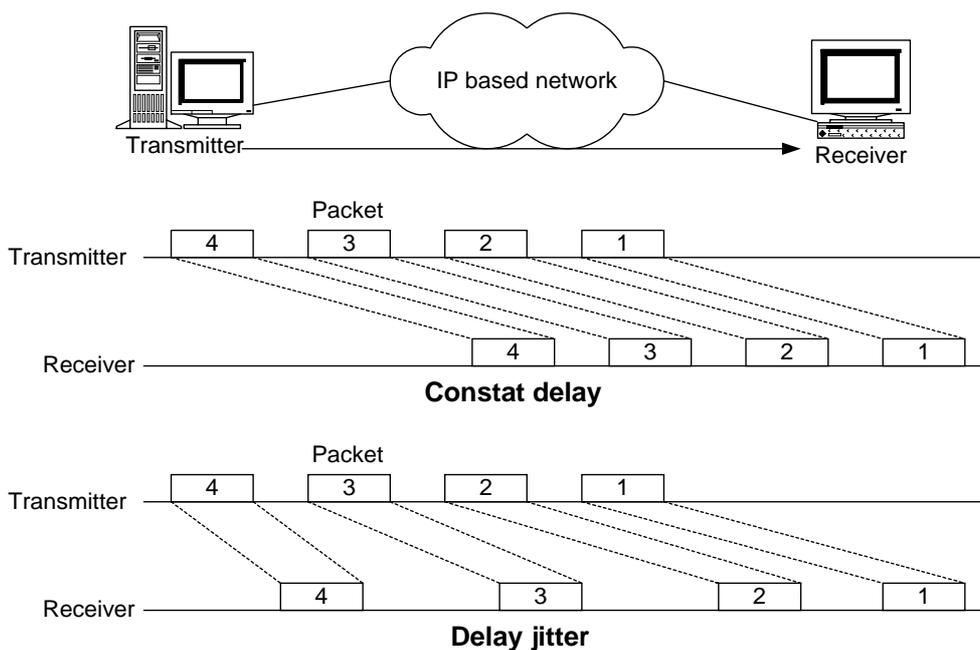


図 3.2 パケットの遅延揺らぎ

3.1.3 対応策

これらネットワークのサービス品質 (Quality of Service:QoS) が保証されていないときにも、映像の再生品質を保つための対応策が数多く検討されているが、典型的な例としては、階層符号化により、ネットワークの QoS の変化に適応させる方式がある。

階層符号化とは、映像をエンコードする際に、ファクシミリのように最終的に表示したい画像を一気に符号化せずに、図 3.3 に示すように、最低限表示したい画質が得られるようにまず符号化し、続いて、より高画質となるような符号化を階層的に施すものである。デコードされる際には、まず最下層の最低限の画質に相当する符号化データがデコードされ、その後順番に高画質に相当する符号化データのデコードが行われる。

配信するときには、ネットワークの QoS が良好なときはすべての階層の符号化データを送信し、最高画質で再生するが、QoS が少し低下すると、上の階層の符号化データは送信せず、さらに QoS が悪化した場合には、最下層の最低限の符号化データのみ送信するようにする。こうすることで、ネットワークの QoS の変化に適応的に対処でき、ある画質以上の再生は保証できるようになる [33]。

3.2 複数 AV ストリーム配送における同時性

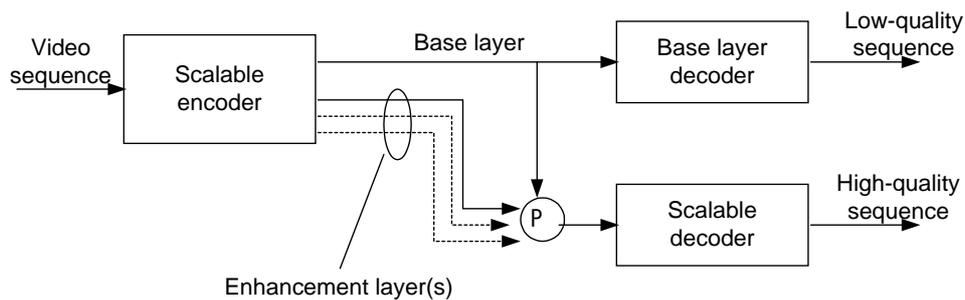


図 3.3 階層符号化

3.2 複数 AV ストリーム配送における同時性

前節を要約すると、IP 網など、QoS が保証されないネットワークを経由して、AV ストリームを配送するときには、「パケットが失われず、フレームレートによって決まる再生時刻の前に到着するよう制御すること」が課題となる。

本研究では、同期関係にある複数の AV ストリームを配送するため、これに加えて、図 3.4 に示すように、AV ストリーム間の同時性を保つ必要がある。つまり、同期関係にある各 AV ストリームのパケット群のうち、あるパケットの到着が他のパケットの到着よりも遅れると、再生時に各映像の同期が取れない原因となる。

ゆえに、この同時性を保証する配送方式が要求されるが、一つの AV ストリームを配送するときの再生品質を保つ方式は、先に述べたような階層符号化など、多くの研究・方向がされているのに対し、複数の AV ストリームを配送したときの同時性に関する研究・報告はほとんどない。

そこで本研究では、まず、現在インターネットにおける映像配信の主流となっているストリーミングによって、複数 AV ストリームを配送したときの同時性に関する特性を把握し、その妥当性について検証することにした [34]。

3.3 2つのAVストリームの同時性の検証

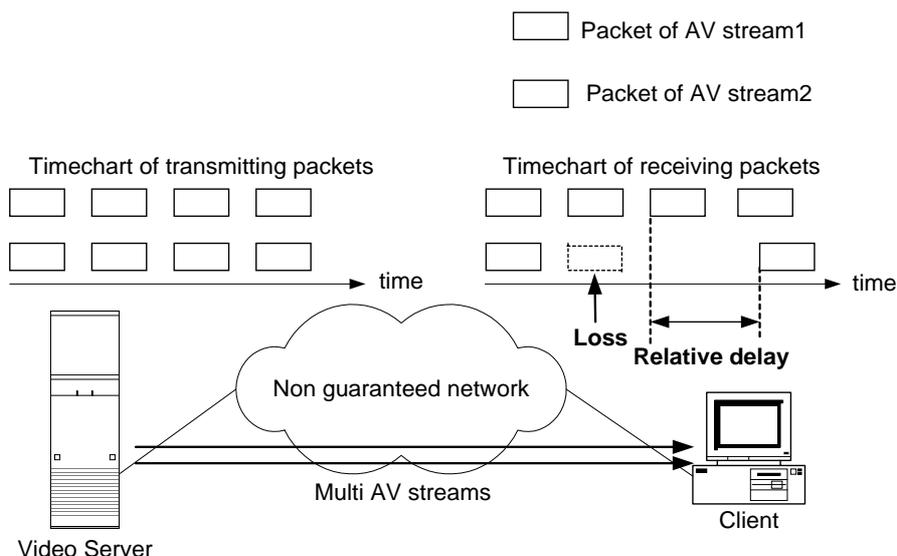


図 3.4 複数 AV ストリーム配送の同時性

3.3 2つのAVストリームの同時性の検証

3.3.1 検証方法

1つの送信端末（ビデオサーバ）にある2つの同期関係にあるAVファイルが蓄積されており、これらは同じ再生時間で、最初から最後まで同期して再生したいものとする。AVファイルのエンコード条件は2つとも同じとする。

送信方法は、2つのAVファイルデータを、同じサイズの packets に分割して、同じ送信レートでストリーミングする。このとき、packet のシーケンス番号を、1byte のヘッダとして付与する。

2つのAVファイルは、同じ時間軸で捉えた映像をエンコードされたものとして、同じシーケンス番号のpacket を同時に送信し、受信端末において、packet 到着の有無と、到着したpacket の到着時刻を測定し、前述の同時性について検証した。

ここでは、従来の1ストリームに着目した項目と、2つのストリームの同時性に関する項目があるが、まず、個々のストリームの特性については、次の項目を検証した。

- packet 到着率 (Packet arrival rate) : 送信したpacket のうち、受信端末に到着した

3.3 2つのAVストリームの同時性の検証

パケットの割合

- パケット到着間隔 (Packet interval) : パケットを受信してから、次のパケットを受信するまでの時間

2つのストリームの同時性に関しては、次に定義する項目を検証した。

- 完全パケット到着率 (Complete packet arrival rate) :
同時に送信された2つのパケットが、両方とも到着した割合
- 不完全パケット到着率 (Uncomplete packet arrival rate) :
同時に送信された2つのパケットのうち、どちらか片方が到着した割合
- 完全パケット損失率 (Complete packet loss rate) :
同時に送信された2つのパケットが、共に損失した割合
- 到着相対遅延 (Relative delay) :
同時に送信された2つのパケットのそれぞれの到着時刻の絶対差

理想的には、完全パケット到着率が1、到着相対遅延が0となる。

3.3.2 実験条件

実験条件を、表 3.1 に、実験に用いたネットワークの構成を図 3.5 にそれぞれ示す。一般に配送系は、長距離のネットワークとなることから、高知にあるビデオサーバから、JGN(Japan Gigabit Network) [35] を経由して千葉県幕張にある受信端末まで、2つのストリームを送ることにした。高知 - 幕張間の JGN 回線は、20Mbps の ATM CBR(Constant Bit Rate) となっている。

実験は、図 3.5 に示すように、AV ストリームに対する負荷ストリームを発生させ、負荷ストリームの送信レートをパラメータに取って、前節の項目について測定を行い、結果を整理・分析した。

3.3 2つの AV ストリームの同時性の検証

表 3.1 2つの AV ストリームの同時性の実験条件

送信ストリーム数	2
パケットサイズ	12500byte
送信レート	10Mbps
負荷パケットサイズ	50000byte
負荷送信レート	0,5,10,20Mbps
高知 - 幕張間回線容量	20Mbps ATM CBR
使用プロトコル	UDP/IP

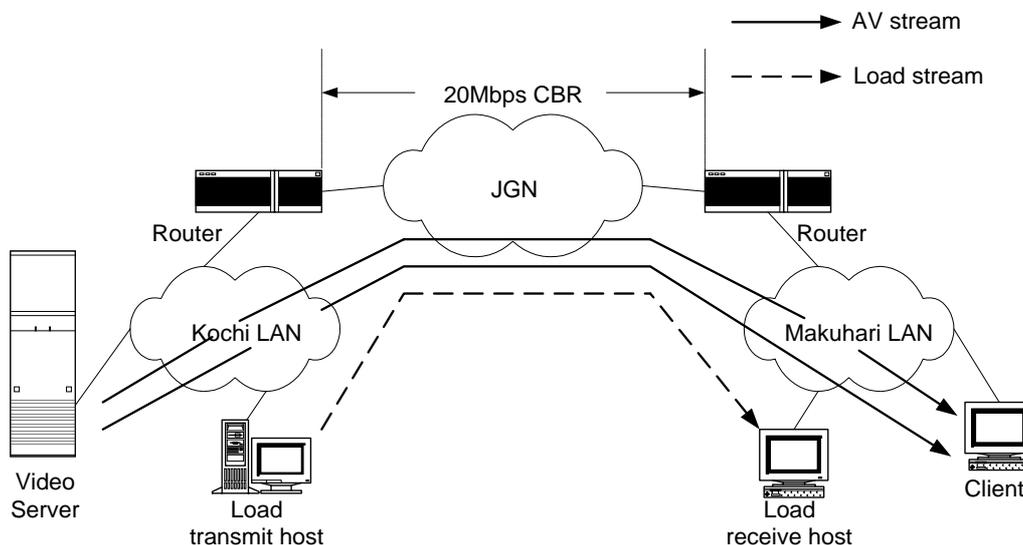


図 3.5 2つの AV ストリームの同時性実験のネットワーク構成

3.3.3 結果と考察

測定値から上記の項目について整理したところ、パケットの到着率に関する特性は、図 3.6 に、パケットの遅延に関する特性は、図 3.7 に示すとおりとなった。

図 3.6 左に示す結果より、負荷が大きくなるにつれて両方のストリームのパケット到着率は減るが、必ずしも同じ特性を示すとは限らないことがわかる。また、同図右に示す結果を見ると、負荷に反比例して完全パケット到着率が下がり、その分、負荷に比例して完全パ

3.3 2つの AV ストリームの同時性の検証

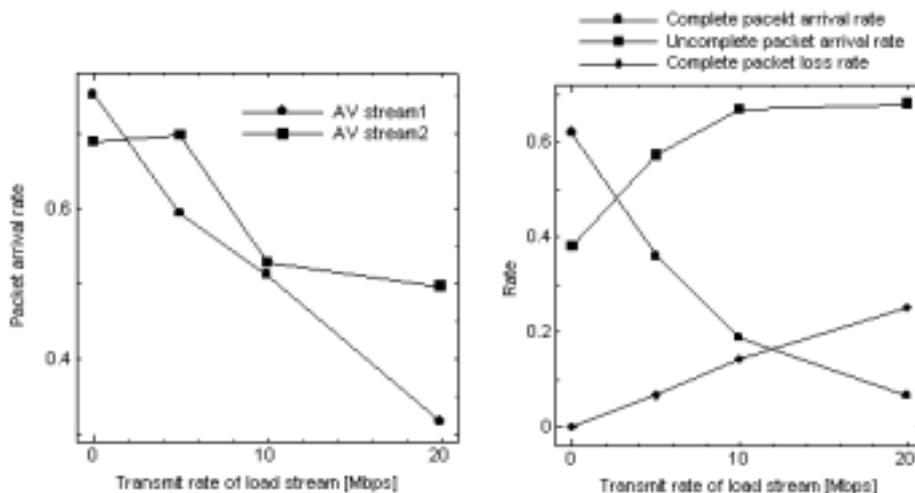


図 3.6 パケット到着率に関する特性

ケット損失率が上がっている．このことから，複数 AV ストリームの同時性を考えたときには，パケット損失への対処が必要となる．

図 3.7 を見ると，パケットの到着間隔は，負荷に比例して長くなっており，2つのストリームが同様の特性を示している．これに対し，相対遅延の方は，負荷が大きくなってても，その平均値は 10ms 以下を保っている．負荷が上がれば，遅延量は増えるのが一般的であるため，この特性は注目に値する．さらに詳しい特性を得るために，今度は相対遅延の分布を調べることにした．

結果は，図 3.8 に示すとおりとなった．負荷が大きくなるにつれて計数されるパケット数が減っているのは，完全パケット到着率が下がっているためである．同図より，負荷が 20Mbps のときでも，完全到着したパケットのほとんどは，30ms 以下の相対遅延となっていることがわかる．

ゆえに，実際に映像を再生させる場合，NTSC もしくは PAL のフレーム周波数であれば，同期して再生できることが予測される．

3.3 2つの AV ストリームの同時性の検証

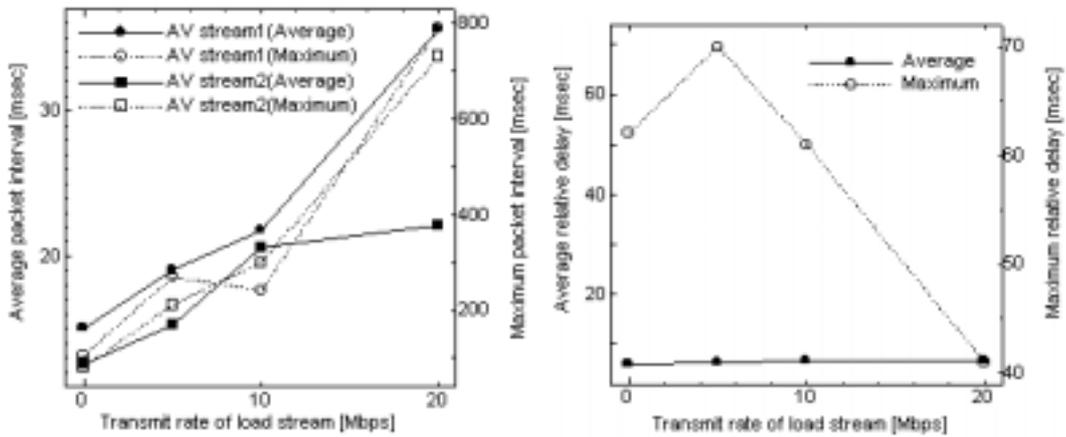


図 3.7 パケットの遅延に関する特性

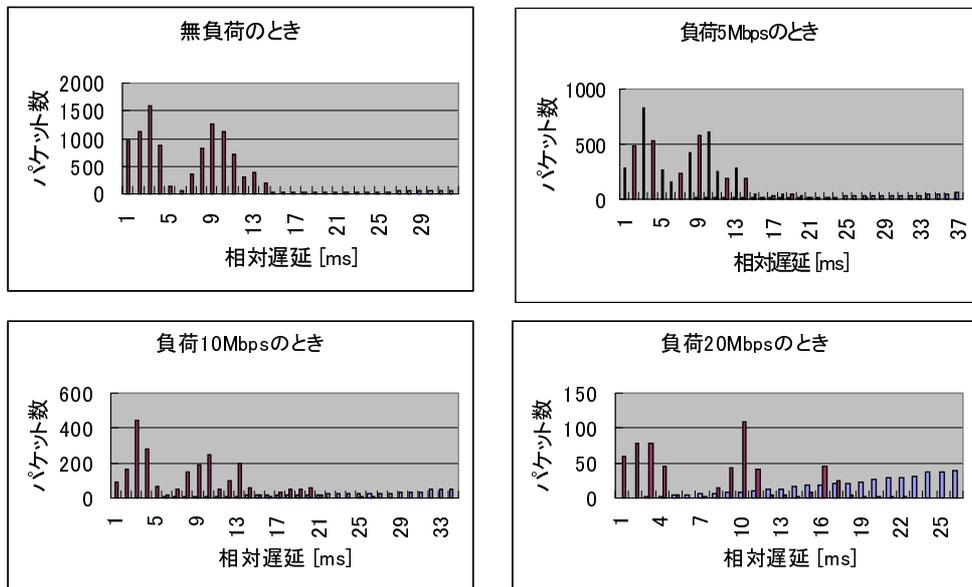


図 3.8 相対遅延の分布図

3.3.4 LAN における同時性

JGN などの長距離網における検証と並行して、LAN における同時性についても合わせて実験を行った。その結果、負荷が大きくなっても、到着したパケットの相対遅延は、無負荷のときのそれとほぼ同じになることがわかった [36]。

ゆえに、ネットワークのトポロジーに関わらず、同じ場所から送信されるストリームであれば、その到着相対遅延は負荷に依らずに、途中のルータやスイッチなどの処理時間に支配

3.4 映像を再生した検証

されるものと考えられる。

3.4 映像を再生した検証

前節の実験結果から，同一サーバからであれば，ストリーミングによって，NTSC のような秒間 30 フレーム程度の映像を同期関係を保持しながら再生できることが可能と考えられる。

そこで今度は，実際に同期関係にある 2 つの映像を MPEG ファイルにエンコードしてビデオサーバに置き，2 枚の MPEG デコーダボードを持つ端末（デコーダ端末）にストリーミングして，映像をデコード・再生し，同期関係を保持しながら再生されるかどうか，検証した。

3.4.1 検証方法

擬似ストリームのときと同様に，まず，同期関係にある 2 つの映像を同じ条件でエンコードし，同じビデオサーバに保管しておき，デコーダ端末に向かって，同じ条件でストリーミングを開始する。ただし，デコーダ端末における処理を簡単にするため，シーケンス番号などのヘッダは付与せず，再送やフロー制御もしない。

こうして，再生される映像の同期関係および映像の品質を，目視により確認した。

3.4.2 実験条件

実験条件を，表 3.2 に，実験に用いたネットワークの構成を図 3.9 にそれぞれ示す。実験に用いた映像は，図 3.10 に示すように，バレーボールをしている 2 つのチームのそれぞれのコートからカメラにより撮影したものであり，同期関係を保持しながら再生することが要求される。

この映像を，MPEG-2/TS(Transport Stream) ファイルにエンコードし，符号化レートを，3Mbps, 6Mbps, 12Mbps, 15Mbps の 4 種類用意し，同じ条件でエンコードされた 2 つ

3.4 映像を再生した検証

表 3.2 実映像再生の実験条件

映像フォーマット	NTSC
ファイルフォーマット	MPEG-2/TS
符号化レート	3,6,12,15Mbps
GOP フレーム数	15
再生時間	約 5 分
送信パケットサイズ	16384byte(16kbyte)
送信レート	符号化レートより 1,2Mbps 高く設定
使用プロトコル	UDP/IP

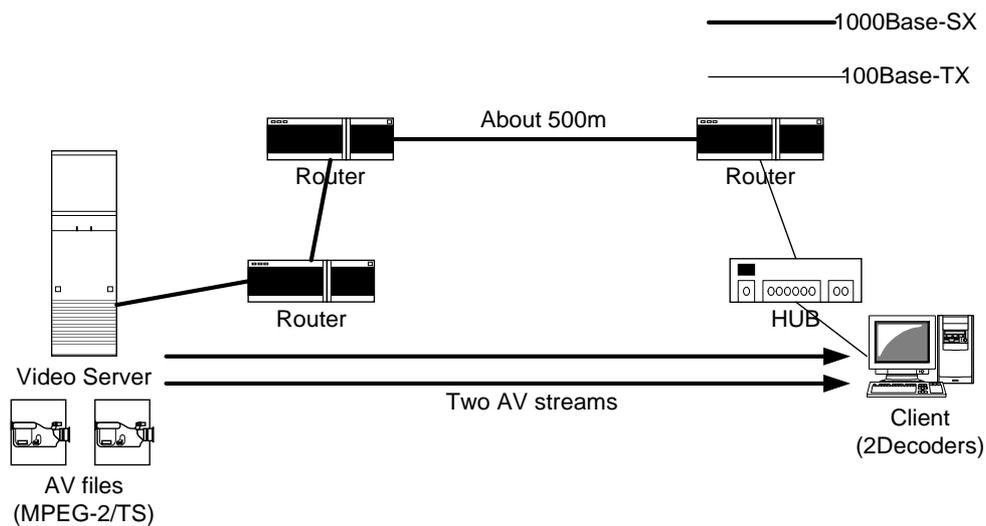


図 3.9 実映像再生実験のネットワーク構成

のファイルをストリーミングした [?][?]。ストリーミングのパケット転送レートは、符号化レートと等しいか、1Mbps から 2Mbps 高く設定し、再生される映像の同期性と、映像の品質を確認した。

3.4 映像を再生した検証



図 3.10 実験に用いた映像

表 3.3 実映像の再生結果

符号化レート	同期性	映像の品質
3Mbps	同期して再生	数フレーム欠落し，画が飛んで見える．
6Mbps	同期して再生	時々動きが鈍くなり，ノイズが混じる．
12Mbps	同期して再生	時々動きが鈍くなり，ノイズが混じる．
15Mbps	同期して再生	時々動きが鈍くなり，ノイズが混じる．

3.4.3 結果と考察

再生した結果は，表 3.3 に示すとおりとなった．また，再生中のある時刻における映像の出力例を，図 3.11, 図 3.12 図 3.13, 図 3.14 にそれぞれ示す．

結果を見ると，符号化レートが低いと，動きの激しいところでノイズが出たり，数フレーム欠落し，動いている人物や物体が飛んで見えたりしたが，符号化レートが高いと，映像の品質は改善され，そのように画が飛ぶことはなかった．ただし，最初は滑らかに動いていて

3.4 映像を再生した検証



図 3.11 実映像の再生例 (符号化レート:3Mbps)

も、しばらく経過すると動きが鈍くなり、その後また元に戻るという状態が繰り返された。また、動きの激しい所では、ブロック上のノイズが出たり、あるいは帯状のノイズが出るがあった。

しかしながら 2 つの映像の同期性に注目すると、いずれの場合でも、図 3.11 から図 3.14 に示すように、終始同期関係を保持しながら再生された。

3.4 映像を再生した検証



図 3.12 実映像の再生例 (符号化レート:6Mbps)

再生される映像と合わせて、ビデオサーバにおけるパケット送信のタイミングと、クライアントにおけるパケット受信のタイミングも調べてみたところ、意図したレートで送信され、また、遅延揺らぎもなかった。また、1つのファイルのみストリーミングした場合には、ノイズもなく、動きも滑らかであった。

このことから、2つの映像を再生したときに品質が落ちるのは、クライアントにおけるバッファリングがうまくできていない・処理能力の限界・PCIバスなどのボトルネックの存

3.5 まとめ



図 3.13 実映像の再生例 (符号化レート:12Mbps)

在などが原因と考えられるが，原因を特定し，性能を上げるのは，今後の課題である．

3.5 まとめ

本章では，複数 AV ファイルの同期利用における研究課題のひとつとして取り上げた，複数 AV ストリームの同期配送方式について検討した．まず，現在の IP 網における映像配信の問題点を整理し，パケットの損失と遅延揺らぎが，再生される映像の品質に関わる重要な

3.5 まとめ



図 3.14 実映像の再生例 (符号化レート:15Mbps)

ネットワーク品質であることを述べた。

IP 網の集合体であるインターネットでは、現在、ストリーミングによる映像配信が主流となっているが、複数の同期関係にある AV ファイルのストリーミングについては、これまで研究例がないため、本研究では、複数の AV ファイルをストリーミングしたときの、パケット到着の同時性を検証した。その結果、同じ場所から送信されるストリームであれば、パケットの損失がなければ、フレーム周波数 30Hz 程度の映像は同期して再生できることが推測されたため、続いて、実際の映像を用いて検証を行った。

3.5 まとめ

2つの同期関係にある映像を MPEG-2/TS にエンコードし、LAN においてストリーミングし、1つのクライアント上で2つのストリームをデコード・再生したところ、動きが鈍くなったり、ノイズが入ることはあったが、同期自体は取れたまま、終始再生することを確認した。

第 4 章

複数 AV ファイルの同期管理方式の 検討

本章では、蓄積系における複数 AV ファイルの同期管理方式について検討する。まず、エンコードによって生じる、同期管理上の問題点を述べ、それに対応する同期管理方式として、複数 AV ファイルの同期参照法を提案する。続いて、提案方式の有効性について検証する。

4.1 エンコードによる同期管理の問題点

4.1.1 可変長符号化 (VLC)

一般にある情報源を符号化する場合、高い確率で出現するシンボルには短い符号を割り当て、出現する確率が低いシンボルには、長い符号を割り当てることで、平均符号語長を短くすることができ、圧縮率を上げることができる。このように、シンボルによって違う長さの符号を割り当てる符号化のことを、可変長符号化 (Variable Length Coding: VLC) という。

可変長符号化の典型的な例としては、ハフマン符号がある [37]。例えば、4 つのシンボル、 a_1, a_2, a_3, a_4 があり、それぞれの出現確率が、0.5, 0.3, 0.15, 0.05 とする。シンボルの数は 4 つなので、例えば、表 4.1 のように符号化できる。

ここでは、2bit ですべてのシンボルを符号化できるため、平均符号語長も同じく 2bit である。これらのシンボルを、図 4.1 に示す 2 分木によってハフマン符号化を施すと、平均符号語長は 1.7bit となり、情報量の圧縮ができたことになる。

4.1 エンコードによる同期管理の問題点

表 4.1 同じ長さの符号を割り当てる例

シンボル	符号
a_1	00
a_2	01
a_3	10
a_4	11

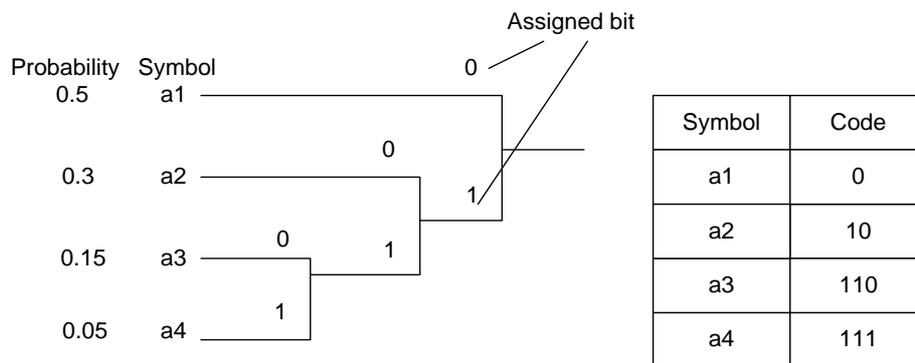


図 4.1 ハフマン符号化の例

4.1.2 VLC による問題点

動画像について、そのフレームサイズを考えて見ると、非圧縮のデジタル動画像の場合、フレームサイズ $S(\text{bit})$ は、次式により決まる一定の値となる。

$$S = a \times X \times Y \quad (4.1)$$

ただし、 a は、1 画素を表すのに使われるビット数 (例えば、RGB を 8bit ずつで表すと、 $8 \times 3 = 24\text{bit}$ となる)、 X は、フレームの横方向の画素数、 Y は、フレームの縦方向の画素数である。

ところが、一般的に動画像の符号化は、元々の情報量が多いため、その圧縮率を高める必要があることもあり、VLC となっている。例えば、MPEG では、DCT の後にハフマン符号化を施している。また、ピクチャの符号化に予測符号化を採用しているため、各フレーム

4.1 エンコードによる同期管理の問題点

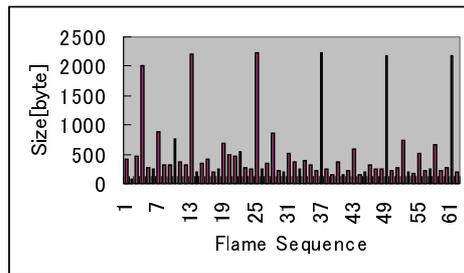


図 4.2 MPEG におけるフレームサイズの例

のサイズは，図 4.2 に示すように可変となる [38]。I ピクチャは，自身のデータのみで符号化するため，予測符号化をする P ピクチャや B ピクチャに対して，サイズが大きくなっている。

ここで，本研究のように，複数の AV ファイルを同期させて利用することを考える。図 4.3 に示すように，フレームサイズがすべて同じであれば，ある AV ファイル中のあるフレームに同期させたい他の AV ファイルのフレームデータが，そのファイル中のどの位置にあるのか，機械的に特定することができる。しかし，VLC では，フレームサイズが異なるため，機械的に特定することができない。

これまでの動画像符号化では，一つの動画像のオーディオとビデオの同期を取るためのメカニズムがタイムスタンプなどによって提供されている。それに合わせる形で，一つの AV ストリームにおける再生時の同期品質について，これまで研究がされてきた [39]。しかしながら，従来の符号化技術や，映像技術に関する研究では，複数の動画像の同期に関しては考慮されていなかったため，何らかの新しい同期管理法が必要となる。

そこで，本研究では，複数の同期関係にある AV ファイルの同期管理法のひとつとして，第 2 章で示した要求条件を満たす，複数の同期関係にある MPEG ファイルの同期参照法を提案し，その有効性を検証した。

4.2 複数 MPEG ファイルの同期参照法

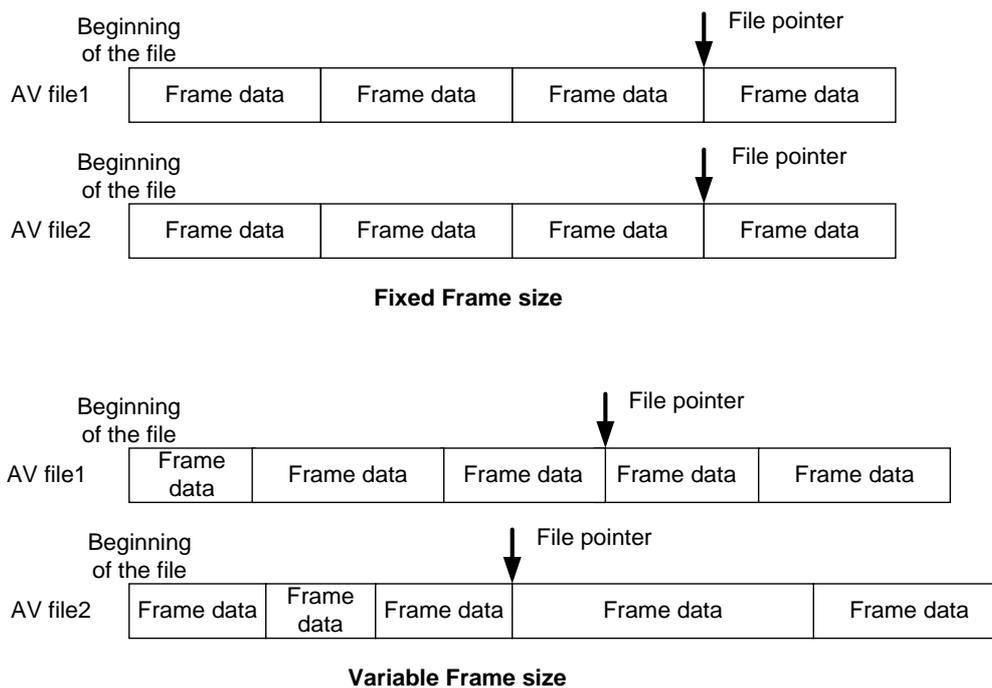


図 4.3 VLC による同期管理の問題

4.2 複数 MPEG ファイルの同期参照法

4.2.1 提案方式の概念

図 4.4 に、提案する複数の MPEG ファイルの同期参照法を示す。対象となるファイルは、次の条件でエンコードされているものとする。

- 対象とするファイルは、同じ時間軸で動いている映像をエンコードしている。
- 各ファイルのエンコード開始の同期が取れている [40]。
- 各ファイルの GOP フレーム数とピクチャの並び方、再生時のフレームレートは同じである。

MPEG では、I ピクチャのみが、自らのデータだけでデコードできることから、各ファイル中の I ピクチャが、ファイルのどの位置（何 byte 目）にあるのかをあらかじめ調べておく。それを、同図右側に示すようなファイル（以下同期参照ファイル）の形で、MPEG ファイルと共に保管しておく。同期参照ファイルは、同じ行に並んである値が、各ファイルの同

4.2 複数 MPEG ファイルの同期参照法

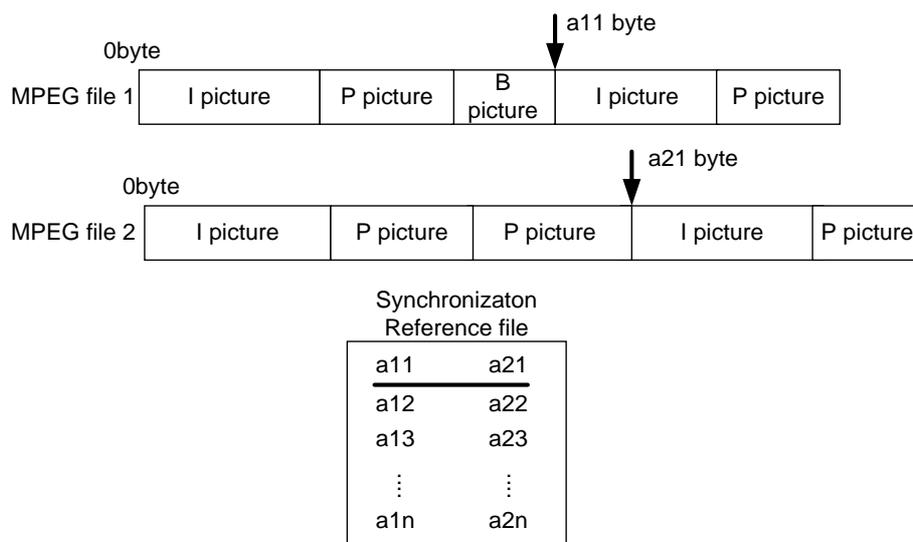


図 4.4 複数 MPEG ファイルの同期参照法

期させたい I ピクチャの位置を示すように生成する。

これにより、1つの MPEG ファイル中のある I ピクチャに同期した他のファイル中の I ピクチャの位置をそのファイルから読み出し、そこからデータを転送することで、要求条件にあったような、同期の取れたランダムアクセスや、同期が崩れたときの再同期などの制御が可能となる。

4.2.2 I ピクチャ検出アルゴリズム

提案方式を実現するには、各 MPEG ファイル中の I ピクチャがある位置を検出する必要がある。本研究で用いた I ピクチャ検出アルゴリズムは、次の手順となっている。

1. 対象となる MPEG ファイルをオープンする。
2. ファイルの先頭から検出を開始する。MPEG ビデオ符号化では、GOP 層におけるヘッダが、表 4.2 のように規定されているため、そのスタートコードである、“000001B8”(16 進数)を探す。
3. GOP の最初のピクチャは、I ピクチャとなっているので、スタートコードが見つかった TSP(Transport Stream Packet) の、ファイルの先頭からの位置(先頭から何バイト目

4.2 複数 MPEG ファイルの同期参照法

表 4.2 GOP ヘッダフォーマット

コード名	ビット数	内容
group start code	32	GOP スタートコード (0x000001B8)
time code	25	タイムコード (時, 分, 秒, ピクチャ)
closed gop	1	GOP の独立性を示すフラグ
broken link	1	GOP 内 I ピクチャ前の B ピクチャの正当性フラグ

か) を, テキストファイルに出力する .

4. 以下, ファイルの最後まで同じ処理を繰り返し, ファイルをクローズする .

すべての MPEG ファイルの I ピクチャ検出処理が終わったら, その出力テキストファイルを, 図 4.4 に示したフォーマットの同期参照ファイルに結合する . この処理には, C 言語などのコンパイラを使わずとも, awk や perl などのテキスト処理ツールを使った方が, 簡単に処理が記述できる .

4.2.3 他方式との比較

提案方式以外に, 複数 AV ファイルの同期参照を行おうとすると, 次のような方法が考えられる .

(1) AV ファイル間タイムスタンプの挿入

MPEG では, ファイル間の同期に関する規定がないため, 各ファイルの同期させたい箇所, 独自のタイムスタンプを挿入し, デコーダでは, そのタイムスタンプに従ってデコードを行うことで, ファイル間の同期を取る方法である .

しかしこの方法では, デコーダに対し, 独自のタイムスタンプを認識させる処理を追加する必要があり, それに対応していないデコーダを使うと, 元々標準にはないデータが混じることになるため, エラーとなってしまう .

4.3 提案方式の検証

(2) 既存のタイムスタンプ値の利用

提案方式と同様の同期参照ファイルを，PTS(Presentation Time Stamp) など，既存のタイムスタンプ値を並べることで同期を取る方法である．

(1) のように，ファイルのデータ自体に手を加えることがないため，デコーダも従来のままでよいが，提案方式と較べると，参照したタイムスタンプ値を持つ TSP などを探索してから，データの転送を開始することになるため，処理が遅くなってしまう．

4.3 提案方式の検証

提案した同期参照法が，実際の MPEG-2/TS ファイルにおいて，どの程度の精度で同期参照ができるかどうか，検証した．

4.3.1 検証方法

まず，表 4.3 に示す同じ条件で，同期関係にある 2 つの MPEG ファイルを生成し，前節に示した処理アルゴリズムによって，同期参照ファイルを作成する．検証用のネットワークには，図 4.5 に示す LAN を使った．同図左側ビデオサーバのある建物と，同図右側クライアントのある建物との間は，約 500m の距離がある．

検証は，同期参照ファイルの中から，無作為にある行のデータを取り出し，そこに書かれてある位置にファイルポインタを移動させ，クライアントへのデータ転送を開始し，クライアントにおいて同期の取れた再生ができているかどうか，目視により確認した．また，転送されてから再生が始まるまでの時間も，time コマンドにより計測した．

4.3.2 結果と考察

検証した結果を，表 4.4 に示す．また，実際の出力例を，図 4.6 に示す．同期参照ファイルからは，20 行を無作為に選んだ．まず，同期参照の精度については，同じ GOP フレーム

4.3 提案方式の検証

表 4.3 同期参照検証用ファイルのエンコード条件

映像フォーマット	NTSC
ファイルフォーマット	MPEG-2/TS
ファイル数	2
符号化レート	3,6,12,15Mbps
GOP フレーム数	15
再生時間	約 5 分間

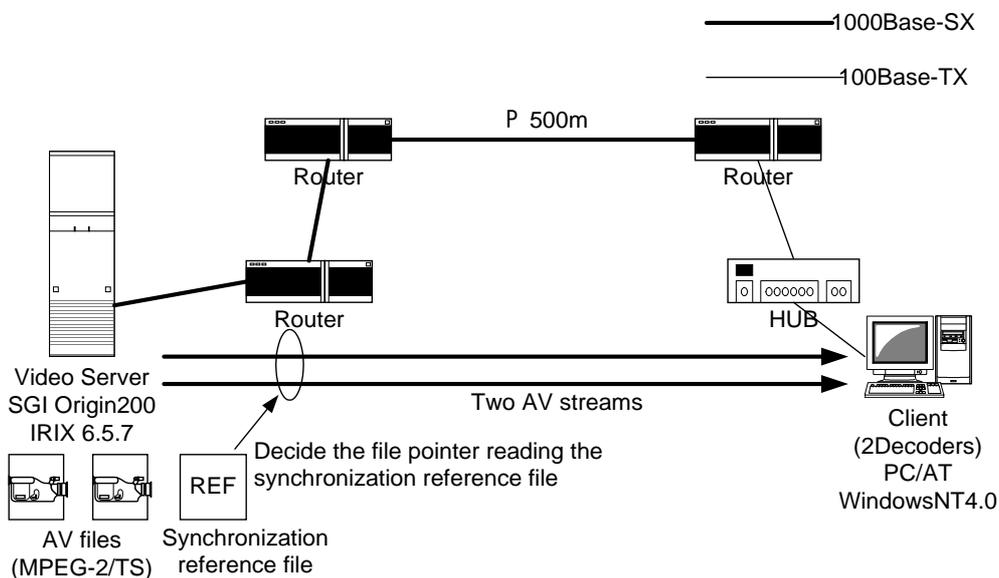


図 4.5 同期参照実験のネットワーク構成

数のファイル同士であれば，図 4.6 のように，作成した同期参照ファイルのどこを読んでも，2 つの映像は同期して表示され，その後も同期して再生が続いた．少なくとも，目視による同期のずれは起きなかったことから，提案方式の精度は実用的な範囲にあると言える．

次に，データを転送開始してから，最初のフレームが表示されるまでの時間を見ると，早いときで約 1 秒，遅いときでも約 3 秒以内に表示された．遅延の原因としては，データは，16kbyte のパケットにして転送され，クライアントのバッファに何個か溜まってからデコードを開始するようにしていたことが考えられるが，このような実装上のパラメータを適当に

4.3 提案方式の検証

表 4.4 提案方式の検証結果

符号化レート	GOP フレーム数	同期参照の精度	再生までの待ち時間
3Mbps	15	同期して再生	1 秒から 3 秒
6Mbps	15	同期して再生	1 秒から 3 秒
12Mbps	15	同期して再生	1 秒から 3 秒
15Mbps	15	同期して再生	1 秒から 3 秒

調整することで、遅延をさらに抑えることが可能と思われる。



図 4.6 同期参照ファイルにより指定した位置から再生したときの出力例

4.4 提案方式の応用例

4.4 提案方式の応用例

検証により，同期してエンコードされ，同じ GOP フレーム数であれば，それらファイルの同期参照が可能であることがわかった．この節では，提案方式の応用例について述べる．

4.4.1 異なる GOP フレーム数の MPEG ファイル間の同期参照

検証では，同じ GOP フレーム数を持つファイルを対象としたが，各々の GOP フレーム数が異なる場合には，次のようにして同期参照が可能となる．

今，同期関係にある n 個の MPEG ファイル f_1, f_2, \dots, f_n があり，それぞれの GOP フレーム数が g_1, g_2, \dots, g_n であるとする．このとき，提案方式によって同期参照ファイルを作成すると， g_1, g_2, \dots, g_n の最小公倍数， $LCM(g)$ の i 倍 ($i = 1, 2, \dots$) 行目に書かれてある位置から f_1, f_2, \dots, f_n を読み出せば同期していることになる．ゆえに， $LCM(g)$ フレーム単位での同期参照が可能となる．

4.4.2 エンコード時に同期が取れていなかったときの補正

現在，MPEG など複数の AV ファイルを生成するのに，自動的に同期を取ってエンコードを開始する手段はまだなく，手動で複数のエンコードを開始するため，同期が取れないこともあり得る．そのときは，同期参照ファイルに書かれた位置から I ピクチャを順次デコード・表示できるようにすれば，目視により，同期参照ファイルを補正することができる．

4.4.3 意味単位による同期参照

MPEG-7 などに代表されるように，近年，映像のシーン記述やインデキシングなど，映像の意味単位による管理が，映像分野における重要な研究テーマのひとつとなっている．

提案方式は，機械的に複数の MPEG ファイルの同期管理を行うが，意味単位での同期管理，例えば，エンドユーザが「一郎が登場するシーンを見たい」と要求すれば，それに対応した各ファイルのフレームが同期して読み出されるということも可能である．まず，先に挙

4.5 まとめ

げた補正法と同様に，同期参照ファイルに書かれた位置から順次デコード・表示し，任意の箇所で，そのシーンを表すキーワードを入力する．そして，入力されたキーワードを，同期参照ファイルのその行の右側に書きこむ．

これにより，エンドユーザにコンテンツを提供する際，キーワードの一覧を一緒に提示し，好きなキーワードを選択してもらえば，それが書きこまれてある行からデータを転送することで，そのシーンから，各ファイルが同期して再生される．

4.5 まとめ

本章では，複数 AV ファイル間が持つ同期関係の管理方式について検討し，従来の符号化法では，複数 AV ファイル間の同期管理について考慮されておらず，新たな管理方式を構築する必要があることを述べた．

本研究では，その管理方式のひとつとして，MPEG-2/TS ファイルの同期参照ファイルの作成による同期管理方式を提案した．提案方式は，各 MPEG-2/TS ファイルにおける I ピクチャの位置を検出し，それらをひとつのファイルに結合するという単純な処理により，ファイルを参照し，指定されてある位置からデータを転送することで，同期が取れるようになっている．

提案方式の有効性を検証するため，実際に同期関係にある 2 つの MPEG-2/TS ファイルを，エンコード条件を変えて複数用意し，同期参照ファイルを作成した後，同期参照ファイルの中から無作為に選んだ位置からファイルの再生を開始したところ，エンコード条件に関わらず，同じ GOP フレーム数であれば，1 秒から 3 秒以内に同期して再生が始まることを確認した．

また，提案方式の応用例として，異なる GOP フレーム数のファイル同士の同期参照，エンコード時に同期がとれていなかったときの補正，意味単位による同期参照を取り上げた．このように，提案方式は，処理は単純でかつ柔軟な運用ができる優れた方式であると言える．

第 5 章

結論

本章では、結論として、本研究で取り上げた 2 つの研究課題、複数 AV ファイルの同期配送と複数 AV ファイルの同期管理について得られた成果を整理し、今後の課題と将来展望について述べる。

5.1 複数 AV ファイルの同期配送

複数 AV ファイルを同期利用するとき、IP 網など、QoS が保証されていないネットワークでは、パケットの損失や遅延揺らぎによって、同期して映像が再生されない問題があるため、一つ目の研究課題として、複数 AV ファイルを、エンドユーザの端末で同期して再生させるための同期配送方式を挙げた。

課題への取り組み方としては、複数の AV ファイルから成るコンテンツをネットワーク経由で取得して再生するという発想およびそれを実装した系がなかったため、まず、現在の映像コンテンツ配信の主流であるストリーミングを複数 AV ファイルの配送にも擬似的に適用し、その特性を把握することから始めた。

その結果、同一サーバからの配送であれば、ネットワーク帯域の 2 倍という過負荷のパケット転送が行われていても、クライアントに到着したストリームの相対遅延は、ほとんどが 20ms 以下であった。

このことから、フレーム間隔が約 33ms となる NTSC の映像であれば、パケットの損失率が低ければ、複数 AV ファイルをストリーミングにより配送することで、それらが同期して再生されることが推測されるので、続いて、複数 AV ストリームを受信し、デコード・再

5.2 複数 AV ファイルの同期管理

生ずる再生系を作り、実際に同期関係にある映像を用いて、同期して再生されるか検証することにした。

サーバとクライアント間が 500m 程の距離である LAN において、2 つの同期関係にある NTSC 映像を MPEG-2/TS にエンコードし、それらをストリーミングした。その結果、映像の品質は、1 つのファイルをストリーミングしたときよりも落ちるが、終始同期がとれた状態で再生した。

ゆえに、同一のサーバに同期させたい AV ファイル群を蓄積しておけば、パケット損失率を低く抑えることで、ストリーミングによる同期再生が可能であることが言える。

5.2 複数 AV ファイルの同期管理

第 2 章で述べたように、複数の AV ファイルが同期して利用できるようになると、様々な応用が考えられる。ところが、MPEG を始めとした現在の符号化標準では、複数の AV ファイルの同期性を管理するという事は想定されていない。ゆえに、新たに複数 AV ファイルの同期管理技術を確立する必要があり、もう一つの研究課題とした。

対象の AV ファイルとして、国際標準であること・通信および放送の分野で利用されていること・今後のネットワークの高速化や端末の処理能力向上により、インターネット上での利用が見込まれるものとして、MPEG-2/TS ファイルを選んだ。

同期関係にある複数の MPEG-2/TS ファイルにおいて、それらの同期情報を、I ピクチャを参照点とした同期参照ファイルとして作成し、それに書かれた位置からファイルを読み出せば同期していることを保証するという、同期管理法を提案し、同期関係にある映像を用いて、3 秒以内に同期参照ファイルに書かれた位置から同期して再生開始することを実証した。

提案方式は、符号化されたビットストリームを直接操作することがないため、標準準拠のエンコーダやデコーダがそのまま使える。また、デコーダを 2 つ持っているユーザに対しては、2 つの同期した AV ファイルを提供でき、4 つ持っているユーザには、最大で 4 つの同期した AV ファイルを同時に提供できるといった、スケーラビリティも持っている。

5.3 今後の課題

また，提案方式は，GOP フレーム数が異なるファイル同士であっても，I ピクチャを参照点とすることから，GOP フレーム数の最小公倍数のフレーム単位での同期管理が可能となる．例えば，GOP フレーム数 6 の MPEG-2/TS ファイルと，GOP フレーム数 15 の MPEG-2/TS ファイルであれば，30 フレーム単位での同期管理が可能である．

他にも，同期参照ファイルに書かれた位置の I ピクチャをデコードして表示するようになれば，エンコード時に同期がとれていなくても補正できるため，同期している事象であっても，エンコードは非同期に行える．さらに，デコードしたシーンに対応したキーワードを付与することで，現在の映像技術の重要な研究テーマのひとつである，意味単位での映像管理も可能となる．

5.3 今後の課題

今後の課題としては，以下に示すものが挙げられる．

5.3.1 複数 AV ファイルの同期配送に関する課題

同一サーバに蓄積されている AV ファイル群であれば，個々のファイルをストリーミングすることで，その同期性が保てることが言えたが．第 2 章で述べたように，一般にファイルは，複数のサーバに分散されると仮定している．ゆえに，分散している複数 AV ファイルの同期配送方式を考える必要がある．

また，元々ストリーミングは，低速なネットワークでも，リアルタイムに映像を見たいという要求から生まれた技術であり，ギガビット級の超高速ネットワーク時代においてもそれが最善かどうか，吟味する余地がある．例えば最近の論文では，622Mbps の ATM ネットワークを想定し，AV ファイルをいくつかのセグメントに分けて，バースト的に転送する方式などが提案されている [41]．

本稿では，IP 網などの QoS が保証されていないネットワークを想定していたが，近年，DiffServ など，QoS を保証する技術に関する研究が活発に行われている [42]．IP も，次の

5.4 将来展望

バージョン 6 では、QoS を保証できるように設計がされている [43][44]。このため、ネットワークの QoS がどの程度保証されるかによって、その同期配送方式がどれだけ有効となるのかという点についても考慮することが望まれる。

5.3.2 複数 AV ファイルの同期管理に関する課題

本稿では、MPEG-2/TS ファイルを対象に、I ピクチャを同期参照点とできることを述べたが、シーンによっては、意図する同期参照点が、I ピクチャではなく、P ピクチャや B ピクチャになることが考えられる。

P ピクチャや B ピクチャは、手前のフレームもしくは手前と次のフレームをデコードしてからでないとデコードできないため、この点について考慮した同期管理方式が必要となる。

検証に用いた映像は、終始同じ場面を捉えてあったが、シーンの切り替わりなど、前のフレームとの動きが激しい場合にも同期参照ができるかどうか、引き続き検証したい。

また、MPEG-4 などのオブジェクトベース符号化では、シーンを構成する各オブジェクトの符号化ビットストリームは多重化されず、個々にデコードされた後、オブジェクトの配置や時間関係を記したシーン記述部を参照してシーンを構成する。このため、提案方式のように、多重化されたファイル中のビットストリームから、目的のフレームデータがある位置を記しておくという手法が使えないため、オブジェクトベース符号化に対応した同期管理技術も必要となる。

5.4 将来展望

最後に、本研究の将来展望について述べる。ネットワークを介した複数の AV ファイルの同期利用という技術が確立されると、第 2 章のシステム参照モデル以外にも、ライブ配信システムや、ビデオサーバに蓄積された映像を、ネットワークで繋がった遠距離にいる仲間と一緒に視聴するなど、蓄積配信型とリアルタイム配信型とが混じったシステムへの応用が考えられる。

5.4 将来展望

本研究では、複数 AV ファイルの素材として、同期関係にある実写映像を用いたが、最近のテレビ番組や映画では、CG(Computer Graphics) が随所に用いられている。CG も、Visual のひとつであり、かつデータ量も大きいことから、超高速ネットワークにおいて配送される AV コンテンツとして期待できる。本研究の応用として、実写と CG の組み合わせからなる同期 AV コンテンツの生成とその利用についても、視野に入れている。

謝辞

本研究を進めるにあたり，入学当初から，指導教員として終始御指導，御助言していただいた島村 和典教授に深く感謝致します．また，副査として有益な御助言をしていただいた，情報通信ネットワークコースの清水 明宏助教授ならびに情報システムコースの岩田 誠助教授にも深く感謝致します．同じく，有益な議論をしていただき，実験設備も利用させて下さった，通信・放送機構 高知通信トラヒックリサーチセンター ギガビットネットワーク 研究開発プロジェクトの加藤 寛治研究員，高松 希匠研究員，神田 敏克研究員の皆様にも感謝致します．

参考文献

- [1] 宮原秀夫, 尾家祐二 共著, “基礎情報工学シリーズ 9 コンピュータネットワーク”, 森北出版, 1992.
- [2] Paul Izzo, “Gigabit Networks”, John Wiley & Sons, INC., 2000.
- [3] 竹下隆史, 村山公保, 荒井透, 苅田幸雄 共著, “マスタリング TCP/IP 入門編第 2 版”, オーム社, 1998.
- [4] 加納貞彦, 栗林伸一 編著, “やさしい ATM ネットワーク信号方式 マルチメディア通信のための新しい信号方式 ”, 電気通信協会, 1996.
- [5] Juha Heinanen, “Multiprotocol Encapsulation over ATM Adaptation Layer 5”, IETF Network Working Group, RFC1483, Jul., 1993.
- [6] M. Laubach, “Classical IP and ARP over ATM”, IETF Network Working Group, RFC1577, Jan., 1994.
- [7] スティーブン サンダース 編集, Stephen Saunders 著, 井早 優子, 林田 朋之, 米沢 寿員 翻訳, “ギガビットイーサネット徹底解説”, 日経 BP 社, 1999.
- [8] 山野誠一, 堺和則, 三好清司, 松本一也, “多様化とブロードバンド時代を迎えたアクセスネットワーク技術小特集/2.xDSL アクセス技術”, 電子情報通信学会誌, Vol.84, No.2, pp.84-91, 2001.
- [9] 渡辺隆市, 山口一雄, “多様化とブロードバンド時代を迎えたアクセスネットワーク技術小特集/3. 光アクセス技術”, 電子情報通信学会誌, Vol.84, No.2, pp.92-97, 2001.
- [10] 米原明史, 船江 英章, 河内 康郎, “IP インフラストラクチャを提供する高速ルータ”, IPSJ Magazine, Vol.42, No.1, pp.76-81, Jan., 2001.
- [11] 越智宏, 黒田英夫 著, “JPEG&MPEG 図解でわかる画像圧縮技術”, 日本実業出版社, 1999.
- [12] 小野定康, 鈴木純司 共著, “わかりやすい JPEG・MPEG2 の実現法”, オーム社, 1995.

参考文献

- [13] 小野文孝, 渡辺裕 共著, “高度映像技術シリーズ 1 国際標準画像符号化の基礎技術”, コロナ社, 1998.
- [14] 映像情報メディア学会 編, “総合マルチメディア選書 MPEG”, オーム社, 1996.
- [15] 三木弼一 編著, “多彩な映像、音声を自在に符号化する MPEG-4 のすべて”, 工業調査会, 1998.
- [16] INTERNATIONAL STANDARD, “Information technology Generic coding of moving pictures and associated audio information: Video”, ISO/IEC 13818-2, 1996.
- [17] INTERNATIONAL STANDARD, “Information technology Generic coding of moving pictures and associated audio information: Systems”, ISO/IEC 13818-1, 1996.
- [18] Francois Fluckiger, “Understanding Networked Multimedia Applications and Technology”, Prentice Hall, 1995.
- [19] 映像情報メディア学会 編, 笠原久嗣 著, “マルチメディアビデオオンデマンド”, 昭晃堂, 1999.
- [20] The Digital Audio-Visual Council(DAVIC), “DAVIC 1.3 Specifications”, Geneva, Switzerland, Feb, 1997.
- [21] 一之瀬進 他, “小特集 ここまで来た！ インターネットにおける映像技術”, 映像情報メディア学会誌, Vol.52, No.12, pp.1756-1778, 1998.
- [22] H. Schulzrinne, S. Casner, R. Frederick, V. Jacobson, “RTP: A Transport Protocol for Real-Time Applications”, IETF Audio-Video Transport Working Group, RFC1889, Jan., 1996.
- [23] D. Hoffman, G. Fernando, V. Goyal, M. Civanlar, “RTP Payload Format for MPEG1/MPEG2 Video”, IETF Audio-Video Transport Working Group, RFC2250, Jan., 1998.
- [24] Y. Kikuchi, T. Nomura, S. Fukunaga, Y. Matsui, H. Kimata, “RTP Payload Format for MPEG-4 Audio/Visual Streams”, IETF Audio-Video Transport Working

参考文献

- Group, RFC3016, Nov., 2000.
- [25] 中平拓司, 島村和典, “超高速ネットワークにおけるマルチビデオストリームアプリケーション”, 2000 年電子情報通信学会総合大会講演集, B-7-143.
- [26] 尾野徹, 大田友一 他, “仮想化現実技術による自由視点三次元映像スタジアム通信の実証研究”, 通信・放送機構 平成 12 年度研究発表会予稿集, 資料発表編, pp.277-278, 2000.
- [27] 横矢直和, “小特集 3 次元映像メディアのコンピュータ合成技術 3-2 3D パノラマ画像合成技術”, 映像情報メディア学会誌, Vol.54, No.3, pp.338-342, 2000.
- [28] 齋藤隆弘, 市川忠嗣 他, “高度三次元動画像情報の通信技術に関する研究開発 (統合化情報) 空間共有コミュニケーションの実現に向けて ”, 通信・放送機構 平成 12 年度研究発表会予稿集, 席上発表編, pp.299-306, 2000.
- [29] 宮崎英明, 亀田能成, 美濃導彦, “複数のカメラを用いた複数ユーザに対する講義の実時間映像化法”, 電子情報通信学会論文誌, D-II, Vol. J82-D-II, No.10, pp.1598-1605, 1999 年 10 月.
- [30] 杉浦一徳, 櫻田武嗣, 小川晃通, “高品質メディア IP 転送技術の実証実験報告”, IPSJ Magazine, Vol.41, No.12, pp.1321-1326, Dec., 2000.
- [31] 勝本 道哲, “超高品質メディアの高速 IP 転送技術報告”, IPSJ Magazine, Vol.41, No.12, pp.1327-1331, Dec., 2000.
- [32] Martyn J. Riley, Lain E.G. Richardson, “Digital Video Communications”, Artech House, 1997.
- [33] Xin WANG, Henning SCHULZRINNE, “Comparison of Adaptive Internet Multimedia Applications”, IEICE Transactions on communications, Vol.E82-B, No.6, pp.806-818, Jun., 1999.
- [34] 中平拓司, 島村和典, “WAN 環境におけるマルチ AV ストリーム配送の同時性に関する一検討”, 電子情報通信学会 技術研究報告, SSE2000-105, 2000.
- [35] Japan Gigabit Network, URL <http://www.jgn.tao.go.jp/>

参考文献

- [36] 中平拓司, 坂田青児, 島村和典, “LAN 環境におけるマルチ AV ストリーム配送時の相対遅延特性”, 平成 12 年電気関係学会四国支部連合大会講演論文集, 12-23, 2000 年 10 月.
- [37] 嵩 忠雄, “情報工学入門選書 (6) 情報と符号の理論入門 ”, 昭晃堂, 1989.
- [38] MPEG-4 and H.263 Video Traces for Network Performance Evaluation,
URL <http://www-tkn.ee.tu-berlin.de/research/trace/trace.html>
- [39] 布目 敏郎, 田坂 修二, 石橋 豊, “蓄積メディアストリーミング技術の性能比較”, 電子情報通信学会論文誌, Vol.J83-B, No.6, pp.824-835, 2000 年 6 月.
- [40] 野中健史, 中平拓司, 島村和典, “マルチストリーム AV ファイル同期初期化法の一検討”, 平成 12 年電気関係学会四国支部連合大会講演論文集, 12-22, 2000 年 10 月.
- [41] Satoshi UNO, Hideaki TODE, Koso MURAKAMI, “Simple and Efficient Video-on-Demand Scheme with Segment Transmission over High Speed Network”, IEICE Transactions on communications, Vol.E84-B, No.1, Jan., 2001.
- [42] Chris Metz, “Differentiated Services”, IEEE Multimedia, July-September 2000, pp.84-90.
- [43] S. Deering, R. Hinden, “Internet Protocol, Version 6 (IPv6) Specification”, IETF Network Working Group, RFC1883, Dec., 1995.
- [44] クリスチャン・ウイテマ 著, 村井 純, WIDE プロジェクト IPv6 分科会, 松島 栄樹 翻訳, “IPv6 次世代インターネット・プロトコル”, ピアソン・エデュケーション, 1997.