

## 次世代放送向け多チャンネル映像のデータベース化

中家 路也 徐浩源 富井尚志 有澤博

横浜国立大学 工学部 電子情報工学科  
〒240-8501 横浜市保土ヶ谷区常盤台 79-5  
E-mail:{nakaie, tommy, arisawa}@arislab.dnj.ynu.ac.jp  
横浜国立大学 総合情報処理センター  
〒240-8501 横浜市保土ヶ谷区常盤台 79-5  
E-mail:haoyuan@ynu.ac.jp

### あらまし

本稿では、昨今盛んに行なわれている次世代放送の研究において新たな提案を試みている<sup>1</sup>。一般的にいわれる次世代放送の特徴としては、インタラクティブ性、ディジタル放送性、高画質性(ハイビジョンなど)などがあげられる。しかし、近年急激に普及しているインターネットの基盤である高速ネットワークの技術や、映像などの大容量データを蓄積することを可能にしたストレージ技術を用いることにより、次世代放送にさらなる可能性があるのではないかと考えられる。そこで、本稿では次世代放送における多視点カメラによる新しい撮影環境を提案し、その環境によって取得された映像などのデータを高速ネットワークを用いてデータベースに蓄積するシステム、MCVDBを提案しその設計を行なう。また、このシステムにおけるプリミティブである高速ネットワークの検証をおこない、スキーマに基づくデータベースへの蓄積システムについて述べる。

## Multi Channel Video Database for Next-Generation Broadcasting

Michiya Nakaie Haoyuan Xu Takashi Tomii Hiroshi Arisawa

Division of Electrical and Computer Engineering, Faculty of Engineering,  
Yokohama National University  
79-5, Tokiwadai, Hodogaya-ku, Yokohama 240-8501 JAPAN  
E-mail:{nakaie, tommy, arisawa}@arislab.dnj.ynu.ac.jp  
Information Processing Center  
Yokohama National University  
79-5, Tokiwadai, Hodogaya-ku, Yokohama 240-8501 JAPAN  
E-mail:haoyuan@ynu.ac.jp

### Abstract

This paper presents a new methodology in researching next-generation broadcasting. Generally a feature of next-generation broadcasting is interactivity, digital broadcasting, high definition. As for interactivity, audiences may request to a broadcasting station and using the results, he edits programs. But in recent years, a high speed network technology on a base of rapidly spreading Internet, it is possible to accumulate very large multimedia data like a video data by a technology of storage, so we suppose a new possibility of next-generation broadcasting. So this paper presents a new filming environment for a multi channel camera, and a system of transferring the obtained video data in it on high speed network, then accumulating those data in database. We design the system(MCVDB), and verify high speed network in this system, describe accumulating system on the database schema.

## 1 はじめに

本稿では、昨今盛んに行なわれている次世代放送の研究において新たな提案を試みている。一般的にいわゆる次世代放送の特徴として、インタラクティブ性やデジタル放送、高画質性などがある。しかし、近年急激に普及しているインターネットの基盤である高速ネットワーク(図1)の技術や、映像といった大容量データを蓄積することを可能にしたストレージ技術を用いることにより、次世代放送にさらなる可能性があるのではないかと考えた。

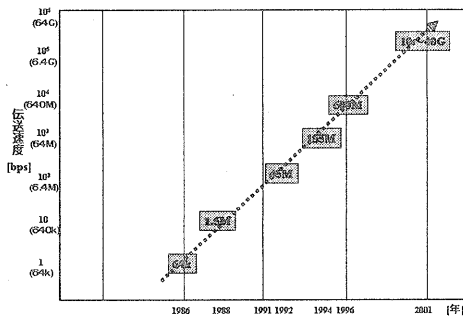


図1: コンピュータ・ネットワークの高速化

そこで次世代放送においては、現在の野球中継の放送ように1チャンネル映像を放送するだけでなく、同時刻における様々な視点から撮影している映像やステレオ映像を用いることにより、球場の中で様々な視点から見ている感覚を味わえるといった、臨場感豊かな放送が求められると予測される。また、多視点カメラ映像を用いて解析処理を行なうことにより、被写体の空間位置を抽出するといった研究も近年盛んに行なわれていて、多視点カメラ映像の取得やデータベース化が求められると考えられる。

しかし、多視点から撮影している映像をスイッチングして、1チャンネル映像を放送するといった現在の放送方法では、同時に複数のチャンネルの映像を放送することが困難であるため、多視点カメラの映像をマルチチャンネルTVとして見る事ができない。更に、これまでの放送方式では多視点カメラの映像を放送と同時に、データベースに蓄積する必

要性がなかったため、このようなデータベースは実現されていない。

そこで著者らは、多視点カメラによって撮影されている映像を、高速ネットワークと大容量ストレージを用いてデータベース化することにより、マルチチャンネルTVを実現することができると考えた。このシステムを次世代放送向け多チャンネル映像のデータベース化システム(Multi Channel Video DataBase:MCVDB)と呼ぶ。また、MCVDBにおいては映像データをデータデータベースに蓄積後、そのデータが画像解析などにも用いられると想定しているため、元の情報の欠落がない非圧縮映像データを用いることも必要である。しかし、図1のように、ネットワークの高速化が進んでいるとはいえ、ネットワークの帯域は有限であり、非圧縮映像データ(ここではD1映像)をリアルタイムに伝送するためには最低160Mbpsの帯域が必要となる。このような帯域を多チャンネル分確保することは困難であり、実用的ではない。そこで、限られたネットワークリソースを有効に利用して、大規模な多チャンネル非圧縮映像データを情報の欠落なくデータベースに蓄積する方法が必要となる。また、多視点カメラの映像を蓄積する上では、各カメラのフレームレートが異なることも考えられ、このような場合に撮影時間とのカメラ固有の時間との対応づけをどのようにしておこなうかが問題となる。更に、どのようにして多チャンネル映像の放送を実現するかが課題となる。

論文の構成を述べると、2章では多チャンネル映像のデータベース化及び、マルチチャンネルTVを実現するために必要とされる機能について述べ、MCVDBのシステム設計を行なう。3章において多視点カメラから取得した映像やカメラパラメータといったデータの特徴に触れ、多視点カメラ映像のデータベース実現するためのスキーマを設計する。そして、4章では本システムを実現する上で重要となる高速ネットワークの評価を行ない、最後にまとめを述べる。

## 2 次世代放送向け MCVDB システム

著者らが提案する、次世代放送向け多チャンネル映像データベース化システムは、多視点カメラの非

圧縮映像データを、高速ネットワークを用いてリモートにあるデータベースに蓄積することによって、マルチチャンネルTVと多視点カメラの映像データベースを実現しようというものである。このシステムを実現するための要件として映像のバッファリング、フレーム補完、ネットワークによる高速伝送、インデックスとの対応づけ蓄積があげられる。著者らはMCVDBシステムを設計する上で、上記の要件をプリミティブとしてかかげ、この要件を解決することにより本システムを実現することを目標にしている。

1. 映像のバッファリングと再送機能:多チャンネル映像においては高速なネットワークを用いて伝送した場合にも、フレームドロップが生じる。そこで、ドロップしたフレームをバッファリングすることが可能であること、後に再送することが可能であること。ただし、バッファは撮影対象を十分に蓄積するだけの容量があるものとする。
  2. フレーム補完機能:映像がフレームドロップした場合には、バッファリングしているフレームを後でデータベースの所定の位置に埋め込む。すなわち、フレームドロップした時点がどこであるかを記録しておき、データベースに蓄積する上で時間軸にそって蓄積を可能にする機能。
  3. 高速ネットワーク伝送機能:映像を伝送するための帯域としては、最低でも1チャンネルの映像をフレームドロップなしに伝送できることが可能であること。
  4. インデックスとの対応づけ蓄積機能:撮影に用いられるカメラは様々な機種と考えるため固有のフレームレートを持つ。そのため、各カメラが持つカメラ時間が存在しその時間と撮影時間との対応が記述できその対応に基づき蓄積する機能。
- 1は各カメラに設置されるディスクが行なう処理である。2、4はデータベース側で行なわれる処理である。3は高速ネットワークに要求される要件である。上記の機能が本システムの上でどのような流れで処理を行なうかを具体的に説明する。

## 2.1 システム構成

本節では、MCVDBシステムのマルチチャンネル映像放送と、多チャンネル映像のデータベース化の実現方法について述べる。MCVDBのシステム構成は図2のようになっている。

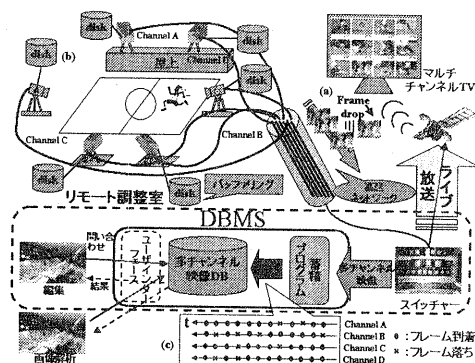


図2: システム構成図

本システムでは、多視点カメラによって映像を取得することを前提としている。しかし、この多視点カメラによって取得されるデータは映像以外にも、カメラアングルやカメラポジションといったデータも取得されるものとした。というのは、このようなデータはジャイロやGPS(Global Positioning System)を用いることで取得可能であり、本システムでは多視点カメラによって取得した映像データは、解析処理などにも利用されると考えている。そのことから、このようなデータを取得してデータベースに蓄積しておくことは後の解析処理にも有用であると考えた。

次に、多視点カメラによって取得された映像などのデータは、高速ネットワークを用いて伝送されるのであるが、ここで一つの仮定を行なう。

多視点のカメラ映像のどれを放送するかの決定はディレクターが行ない、撮影対象の状況を見て多視点カメラ映像、または単視点カメラ映像の選択をする。

例えば、野球の放送ではピッチャーが投げるかまえに入れば、投球するということがわかり、視聴者の興味がバッターとピッチャーの対決に集中する。このような場面には通常ピッチャーの後部からピッ

チャーをバッターを写している単視点映像による放送が行なわれると考えられる。つまり、このような場面ではディレクターは単視点放送を行なうと予想される。ところが、バッターがバッターボックスを外した場合には、視聴者のバッターやピッチャーに対する興味が薄れると考えられる。そこで、そのような場面では、ディレクターは多視点カメラ映像を用いて球場内を様々な角度から写している映像を選択すると仮定するのである。このようなことから、常に多視点映像が放送に用いられるわけではなく、状況に応じて単視点映像との切替が行なわれると考えられる。そこで著者らは、このような放送の性質を利用して、高速ネットワークの帯域の有限性に対応できると考えた。つまり、多チャンネルの放送ではネットワークリソースの有限性からフレームドロップが生じる(図 2-(a)) ため、そのような映像データをそのままデータベースに蓄積することが実用的とはいえない。そこで著者らは、フレームドロップが生じたものに関してはカメラに附属するディスク(図 2-(b)) を用いてバッファリングを行ない、後で再送することにより有用な映像データベースを実現しようと考えた。そのためには、実際にフレームドロップが生じた時点を記録しておく仕組み(図 2-(c)) が必要となる。

### 3 多チャンネル映像のスキーマ設計

前節でも述べたように、多チャンネルカメラによって取得した様々な情報を逐次データベースに蓄積し、様々な問い合わせに応じることのできるスキーマを設計する必要がある。まず、実世界から多チャンネルカメラによって取得されるデータについての特徴を具体的に述べ、次にスキーマ設計を行なう。

#### 3.1 蓄積する情報

高速ネットワークからデータベースに渡されるデータは、映像などのような原データや撮影された時刻などのインデックスデータであるが、具体的なデータは以下のようなものである。

- 映像

取得される映像は、様々な位置に設置された多チャンネルのカメラによって取得されたものである。また、場合によっては立体映像、つまり、2 台のビデオカメラをセットに何組かを用いて取得される。注意すべき点はこれらのカメラが全て同期されて(フレーム間の対応づけがなされて)データベースに蓄積される必要があるということである。

- カメラ情報

本システムにおいてはカメラによって取得される映像データだけでなく、カメラの情報を GPS やジャイロを用いて取得し、放送と同時にデータベースに蓄積される。このような情報の例としては、個々のカメラがある時点にもつ焦点距離などのパラメータや、撮影環境におけるカメラの位置やカメラのアンクルといったデータなどがあげられ、このような情報をもとにカメラの撮影領域などが算出されデータベースに蓄積される。

#### 3.2 映像データベースのスキーマ

多視点映像のデータをデータベースに蓄積するためのスキーマを設計するためには、エンティティベースのデータモデルが必要である。ここでは、データモデルとして、ER モデルなどの意味データモデルを拡張した AIS モデル [1], [2], [3] を用いて設計を行っている。図 3-(a) に、本システムのデータベースで用いるスキーマを示す。このスキーマにおいて、プリミティブとなっているのが撮影時刻とカメラ時間の対応を行なっている部分である。本来、カメラ時間と撮影時刻との対応づけを行なうのが困難である。なぜなら、各カメラのフレームレートの違いなどから、ある撮影時刻に対応するカメラ時間が必ず存在するわけではないからである。このような問題を、撮影時間の最小単位を時区間(3-(b)) とすることにより解決している。

#### 3.3 スキーマに基づく蓄積システム

次に、前節で設計したスキーマに沿ってリアルタイムにデータベースに蓄積するシステムについて述べる。MCVDB においては、ネットワークから転送される映像や時間情報などのデータを先に設計した

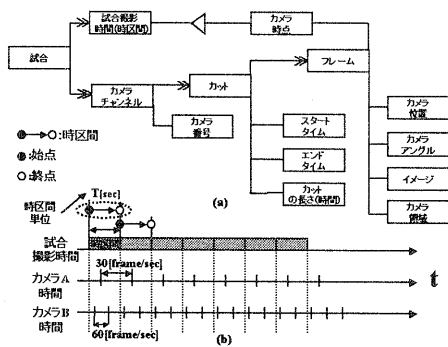


図 3: 多チャンネル映像データベースの例

スキーマに基づきながら蓄積する必要がある。つまり、映像をフレームの列とであると考えた時に、そのフレームと撮影された時間との対応(つながり)などを記述し、それをインデックス情報として蓄積するのである。このような対応づけを行なうことにより、様々な検索が可能になるのである。そこで、このような処理を次の図4に示すようなシステムによって、実現できるのではないかと考えた。

そこで、図4を例に、カメラのチャンネルが3チャンネルの場合においてこのシステムについて説明する。3チャンネルの各映像は撮影時間などのデータとともに伝送されてくる。このようなことから、まず映像の各チャンネルに対して1ポートずつ割り振る。また、撮影時間などのインデックスデータについても、各チャンネル毎に対して専用のポートを割り振る。この結果、図4においては全てで6つのポートが割り振られることになる。このようにして、割り振られたポートに送られてきたデータを受信プログラムで受取り、一つの値(データ)を受けとる毎にエンティティを順次生成していく。つまり、インデックスやフレームの数だけエンティティは生成される。また、受信プログラムではそのエンティティ間の繋がり(対応)も蓄積と同時に結び付けていく。例えば、FRAMEというエンティティ(ID:fl)と、IMAGEというタイプのエンティティ(ID:i1)の繋がりを結ぶのである。このように、データの到着時に順次エンティティIDを生成しながら、エンティティ間の繋がりを結ぶという処理を行なうのが受信プログラムで

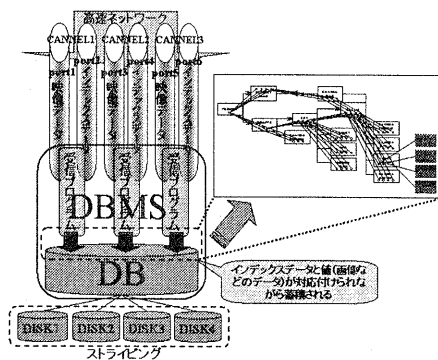


図 4: データ蓄積システム

ある。

以上の処理によって、データがデータベースに蓄積されるのである。

#### 4 高速ネットワークによる伝送と蓄積

MCVDBにおいては多視点カメラから撮影された映像などのデータは高速ネットワークを用いて伝送される。このようなことから、高速なネットワークによる高速伝送を実現する必要がある。

そこで、本稿では高速ネットワークを実現する手法として、マルチメディアデータの伝送に適しているといわれる ATM(Asynchronous Transfer Mode)を用いた。そして、伝送速度の測定実験と多チャンネル映像の伝送実験を行なった。実験においては次に示す点について注目し実験を行なった。

- ATMの通信方式として IPOA(IP over ATM)と LANE(LAN emulation) という二つの方式を用い、この二つの方式を比較する実験を行なった。
- 実験においてはソケットを用いたプログラムを使用し、その通信の際には通信プロトコルとして TCP/IP を用いた。また、実験においてウィンドウサイズをパラメータの一つとして用い、ウィンドウサイズと伝送速度の関係に注目した。

- 。リアルタイムに伝送される映像データ(非圧縮データ)を送信し、映像にフレームドロップが生じずに蓄積できるかを確認した。

まず最初にあげた ATM の通信方式であるが、IPOA では IP パケットを ATM セルにのせるため、TCP/IP または UDP による通信しか行なえない。また、LANE では従来の LAN(Ethernet、FDDI など)の通信方式が行なえるが、オーバーヘッドによる伝送速度の低下などの問題がある。

次に、ソケットのプログラムを用いた伝送を行なう際に、あらかじめ送受信側でウィンドウサイズを決めて伝送速度の測定を行なった。このウィンドウサイズとは TCP/IP プロトコルにおいて一度に送っても良いフレームの個数を示すもので、この値が大きいほどフレームの再送などが起きなかった場合に伝送効率が良くなる。しかし、ネットワークのトラヒックやネットワークインターフェースの性能などによりこのサイズに限界がある。

最後に、このような伝送速度測定実験を行なうことで最適なウィンドウサイズがわかることから、この結果を反映した最適な伝送設定を行ない、多チャンネル映像を伝送してチャンネル数とフレームドロップの関係を測定した。この結果はマルチチャンネル TV を実現する上で重要なサンプルとなると考えられる。というのは、伝送帯域は有限でありその帯域に応じたチャンネル数の決定を行なう必要があると考えられるからである。

また、このような ATM-LAN を構築する上でネットワークで問題となる点をいくつか明らかにした。そこで、この実験を行なった環境およびその伝送速度実験の結果について次に述べる。

#### 4.1 実験

最初に、前述の伝送速度実験におけるネットワーク構成図 5(a) を示す。また、この実験は表 1 にあがる 3 台のワークステーションを用いて行なった。この伝送速度実験に関係のあるパラメータとして MTU(Maximum Trasfermission Unit) サイズが、伝送速度に大きく影響することがわかった。この MTU サイズは、フレームの最大長を表すものでこのサイズによっても、伝送速度に大きな変化があった。というのは、今回の実験に先駆けネットワークベンダの協力により、LANE を用いた通信方式では MTU サ

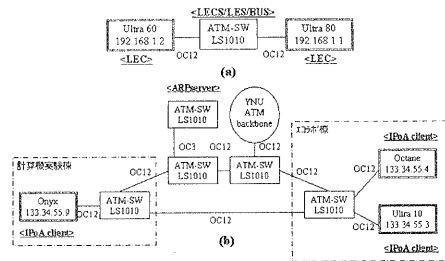


図 5: ネットワーク構成図

イズに大きく依存することが明らかになった。これまで、LANE による通信方式では上記の環境で MTU サイズ 1500[Kbyte] において、伝送速度の最大値は 270[Mbps] であったのに対し、MTU サイズ(ATM スイッチ、ATM-NIC 両者とも)を 9234[Kbyte] に設定し、ATM スイッチを一つはさんだ通信では、最大 500[Mbps] の値を示した。この実験環境のネットワーク構成図 5(b) 示す。このようなことから、MTU サイズを 9234[Kbyte] に設定し上記の環境で LANE 通信方式に適用しようとしたところ、ATM スイッチの OS のバージョンの違いにより相互接続を確立できなかった。

そこで、IPOA を用いた通信方式に変更して実験を行なうことにした。IPOA においても MTU サイズに大きく依存することがわかったため、MTU サイズを ATM-NIC では 9234[Kbyte]、ATM スイッチにおいては 4470[Kbyte] に設定し伝送速度の測定をおこなった。IPOA においては ATM スイッチの OS のバージョンによる問題もなく、相互接続が可能であったため伝送速度の測定を達成することができた。その伝送速度測定による実験結果を図 6,7,8,9,10,11 に示す。この伝送速度の実験結果から、1 チャンネルの映像データ(D1 映像)などのマルチメディアデータを十分に伝送可能な帯域を確保することができると考えられる。

次に、この伝送帯域を用いてリアルタイムに映像の伝送を行ない、フレームドロップが生じずに蓄積することができるかを確かめる実験を行なった。この実験には映像として D1 カメラ(表 2)で撮影している映像データを使用し、先ほどの実験結果をもとに IPOA 方式を用いて伝送し高速大容量ストレージ

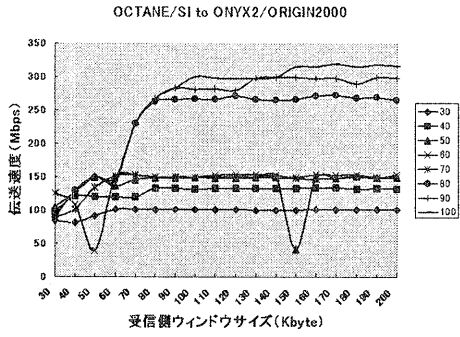


図 6: 測定結果 1

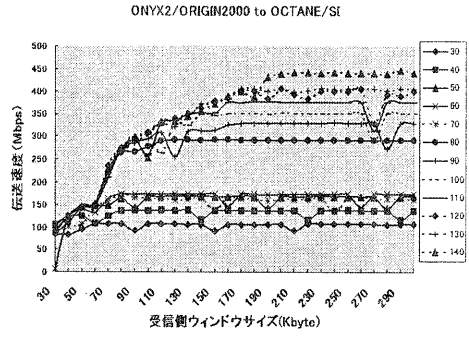


図 9: 測定結果 4

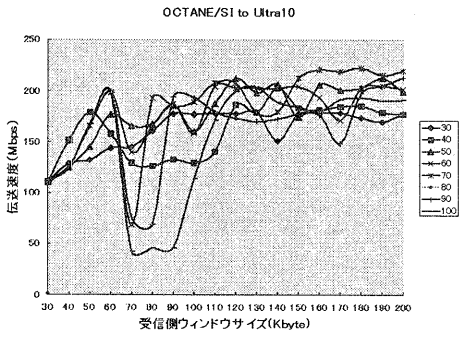


図 7: 測定結果 2

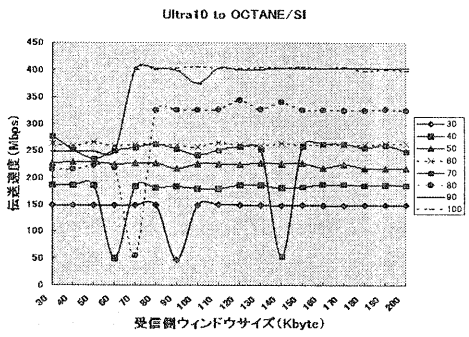


図 10: 測定結果 5

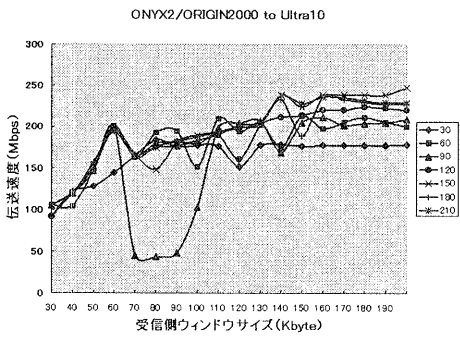


図 8: 測定結果 3

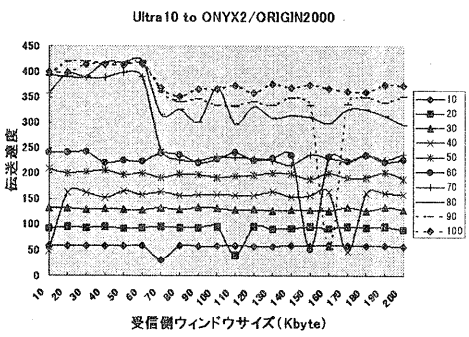


図 11: 測定結果 6

ワークステーション環境	
Machine name	SGI ONYX2/ORIGIN2000
CPU	260[MHz]MIPS RISC R10000(×8)/ 190[MHz]MIPS RISC R10000(×48)
Memory size	4[Gbyte]/16[Gbyte]
Disk size	162[Gbyte]/117[Gbyte]
ATM-NIC	FORERunner 200E
Machine name	SGI OCTANE/SI
CPU	195[MHz] MIPS RISC R10000
Memory size	128[MByte]
Disk size	9[Gbyte]
ATM-NIC	FORERunner 200E
Machine name	Sun Ultra 10
CPU	300[MHz]UltraSPARC-III
Memory size	256[Mbyte]
Disk size	2.5[Gbyte]
ATM-NIC	FORERunner 200E
ネットワーク環境	
ATM-Switch	LS1010(LightStream1010)
OS	IOS ver.11.2.8(16MDRAM)

表 1: 実験に用いた環境

frame rate	30[frame/sec]
frame size	699840[byte]

表 2: D 1 映像

である Onyx2/Origin2000 を用いて蓄積した。また、この実験を行なう際には事前に伝送前の映像を用意し、その映像との比較によってフレームドロップが生じているかどうかを調べた。次に、チャンネル数を増加させた場合の実験を行なった。この測定方法としては、ある時間内に届くはずのフレーム数で実際に届いたフレーム数で割った時の値(×100)を測定値とする。(図6の測定で用いたワークステーション)間で行ない、チャンネル数による影響が映像伝送に及ぼすかどうかを確かめた。その結果を図12に示す。この図からわかるように、チャンネル数が増えることによってフレームドロップの割合が増加していることがわかることから、ネットワークの帯域に応じてチャンネル数を決定する必要があることがわかる。

## 5 まとめ

本稿では、次世代放送において求められるマルチチャンネルTVを実現する MCVDB システムの設計を行なった。また、この MCVDB では多チャンネル映像のレコーダとしてデータベースを用いることを

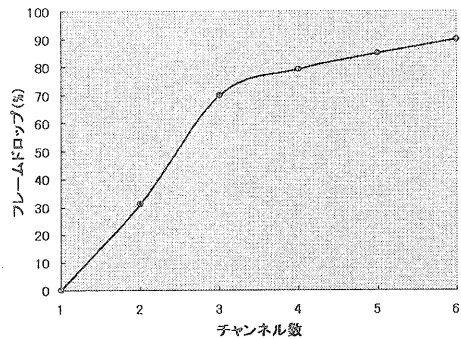


図 12: フレームドロップ

提案した。このシステムにおけるプリミティブの一つである、高速ネットワークの伝送速度測定実験を行なった。さらに、同一ネットワークにおける多チャンネル映像の伝送実験を行ない、多チャンネル映像伝送のフレームドロップ発生率を測定した。これらの実験結果により、非圧縮映像によるマルチチャンネルTVにおいて、要求されるクオリティに対する必要伝送帯域を割り出すことができた。

## 参考文献

- [1] H.Arisawa,H.Nagae,Y.Mochizuki:“Representation of Complex Objects in Semantic Data Model“AIS” and Implementation of Set Operators”,IEICE TRANSACTIONS,vol.E74,No.1,1991.
- [2] H.Arisawa,T.Tomii,H.Yui,and H.Ishikawa: “Data Model and Architecture of Multimedia Database for Engineering Applications”, IEICE TRANS.INF. & SYST.,vol E78-D,No.11,1995.
- [3] 石川 英彦,有澤 博:“データベーススキーマの効果的な図示手法”,情報処理学会データベースシステム研究会.1996