オブジェクト連動データ放送システムのメタデータの伝送手法とその評価実験

A Transmission Method of Meta-data for the Object-linked Broadcasting System and Its Evaluation Tests

正会員 金次保明^{†*}, 正会員 高橋正樹^{††}, 正会員 三須俊彦^{††},
正会員 武智秀^{††}, 正会員 加井謙二郎^{††**}

Yasuaki Kanatsugu^{†*}, Masaki Takahashi^{††}, Toshihiko Misu^{††}, Masaru Takechi^{††} and Kenjiro Kai^{††**}

Abstract We propose a new up-to-the-minute data broadcasting service that uses the latest digital technologies. We produced prototype and carried out experiments to evaluate the performance of the system. The proposed service — the object-linked broadcasting service — provides various information regarding an object chosen by a viewer from a TV picture. We encountered some technical issues, such as observing broadcasting regulations, making dynamic descriptions for the objects, and working with one way transmission, in perfecting the system. To resolve these issues, we employed BML for browsing, ECMA Script for the dynamic description, Binary Table for the data structure, and DSM-CC for the transmission method. The experimental results show the system performance is acceptable. We found that an 80-second data carousel can carry related data for up to 20 objects.

キーワード：データ放送，メタデータ，オブジェクト，連動放送

1. ま え が き

テレビのデジタル化が進んでいる。BS デジタル放送の実施に続いて、地上デジタル放送も 2003 年 12 月から東名阪で開始され、2006 年 12 月までには、全都道府県の県庁所在地に拡大される予定である。そのような状況の中、デジタルの特長を生かした魅力ある新しいデータ放送サービスの開発が期待されている。その要求に応えるために、われわれは「オブジェクト連動データ放送サービス」を提案し、そのサービスを実現するための伝送手法を開発し、そのシステムを試作して、動作実験を行った¹⁾²⁾。オブジェクト連動データ放送とは、映像中に映っている物体（オブジェクト）を視聴者が選択することで、その映像オブジェクト

に関連した詳細な情報を視聴できるようにする放送サービスである。

これに似た放送サービスとして、「番組連動データ放送」がある。これは、データ放送を利用して番組に関連した情報を提供する放送サービスで、番組情報の補完、クイズ番組、ショッピング番組などに利用されている³⁾⁴⁾。また、現行のデータ放送の枠を超えた放送システムとして、TV-Anytime が研究開発されている⁵⁾。これはコンテンツ流通および消費のための規格であり、ARIB においてサーバ型放送方式の規格として採用されている⁶⁾。

受信機側の対応としては、EPG(Electronic Program Guide)やデータ放送を利用して、視聴者の嗜好や視聴環境に適應する受信機の開発も行われている。例えば、視聴者の好みの番組を選択する方式として、受信機に学習機能をつけて、簡単な操作でテレビを楽しむことのできる受信ナビゲーションシステムがある⁶⁾。また、興味のあるシーンのダイジェスト版を自動生成する方法についても研究されている⁷⁾。

以上、述べた放送サービスは、映像内のオブジェクトにはリンク情報が添付されていないため、映像内のオブジェクトを選択しても何も起こらない。視聴者は映像内のオブジェクトを直接選択できないので、オブジェクト毎の情報提供を受けられない。

2005 年 7 月 7 日受付、2006 年 1 月 10 日再受付、2006 年 2 月 8 日採録

[†]NHK エンジニアリングサービス

(〒157-8501 世田谷区砧 1-10-11, TEL 03-5494-2400)

^{††}NHK 放送技術研究所

(〒157-8501 東京都世田谷区砧 1-10-11, TEL 03-5494-3128)

* 現在、NHK 放送技術研究所

** 現在、NHK 放送技術局

[†]NHK Engineering Services, INC.

(1-10-11, Kinuta, Setagaya, Tokyo, 157-8510, Japan)

^{††}NHK Science and Technical Research Laboratories

(1-10-11, Kinuta, Setagaya, Tokyo, 157-8510, Japan)

* NHK Science and Technical Research Laboratories

** Broadcast Engineering Department, NHK

インターネットを利用した放送番組の補完サービスも増加している。例えば、デジタルテレビ専用のコンテンツを提供するネットワークサービス「Tナビ」⁹⁾、番組、ドラマ、タレントなどに登場する商品に関する情報を紹介、販売する「テレビト.com」¹⁰⁾、「Amazon.comのNBCサイト」¹¹⁾などがある。しかし、これらはテレビ番組とインターネットで独立して情報を提供しており、現時点では番組内のオブジェクトがインターネットに直接リンクされているわけではない。

我々の提案する「オブジェクト連動データ放送サービス」は、映像内のオブジェクトに関連する情報がオブジェクトに直接リンクされており、視聴者が画面内のオブジェクトを選択すれば、そのオブジェクトに関連する情報を入手することができる。このサービスを実現するために、オブジェクトに関連する情報の伝送手法と受信機の機能を検討し、システムの試作を行った。さらに、システムの動作実験を行い、その有効性を確かめた。

本文では、まず、提案する放送サービスとシステム構成について述べ、次に、本システムの特徴であるメタデータ記述法について述べる。最後にシステムの動作実験を行いパフォーマンスを確認したので、これについて報告する。

2. 提案する放送サービスとシステム構成

2.1 オブジェクト連動データ放送サービス

図1に本サービスの概要を示す。本サービスは、視聴者がテレビを見ている時に、その画面の中に興味のあるオブジェクトがあれば、それに向けて指をさせばそのオブジェクトの詳細なデータが表示されるサービスのことである。インターネットのWebページのハイパーリンクと同様な機能であるが、リンクの対象が映像内のオブジェクトであることが異なる。現行のデジタル放送の枠内での開発をめざすため、一般の受信機でも映像と音声再生できるようにしなければならない。そのため、映像ストリームと音声ストリームには手を加えず、データ放送用のストリームを利用することを提案する。オブジェクトに関連するデータと、表示方法などを記述したデータをメタデータと呼び、これをデータストリームで伝送する。

2.2 システム構成

提案した放送サービスを実現するために検討したシステム構成を図2に示す。

システムは大きく放送局側と受信側に分かれる。

放送局側では、まず、映像の中のオブジェクトを分離・追跡してオブジェクトの形状と位置を求める。この技術が「オブジェクト分離・追跡技術」である¹²⁾。「メタデータ記述」では、新たに定めたメタデータのデータ構造に基づき、関連データとオブジェクトのリンク情報、および、受信機での関連データの表示方法などの処理を記述している。また、これらのデータを映像と音声信号とに多重化する。「メタデータ・データベース」では、取材時などに得られた情報

図1 オブジェクト連動データ放送サービス
Object-linked broadcasting service.

図2 システムの全体構成
System Configuration.

を保存しておき、オブジェクトのリンクの際に、オブジェクトに関連するデータを検索するために利用される。また、「オーサリングツール」はこれらの処理を扱いやすくするためのユーザインターフェースである。

受信側では、受信した信号を映像、音声、データに分離した後、映像を表示するとともに、視聴者がオブジェクトを選択したときに関連データを表示する。視聴者がオブジェクトを選択したかどうかの判断、および、関連データの表示制御は「メタデータ表示処理」で行われる。「ポインティングデバイス」は視聴者が指さしている場所を取得するためのユーザインターフェースである¹³⁾。

メタデータ記述は、メタデータの構造および伝送方式に関する技術であり、本システムの特徴となる技術であるため、次章で詳細を述べる。

3. メタデータ記述法

3.1 システム設計の課題

システムの開発にあたり、メタデータをどのように伝送するかを検討した。検討にあたって、次のような課題があった。

- (1) 放送規格の枠内で利用しなければならない
- (2) オブジェクトは動画であり、動的なメタデータもある
- (3) 伝送路が放送波であり、一方通行である

(1) に関しては、データ放送の伝送規格に基づいてデータを伝送する必要がある、伝送できるデータの種類の伝送方式が制限される。また、従来システムとの互換性を確保する必要がある、従来の受信機 STB(Set Top Box) でコンテンツが受信できなくなることは避けなければならない。

(2) に関しては、オブジェクトの位置および形状データは時々刻々変化するため、静的なメタデータと区別し、オブジェクトとデータとの同期をとる必要がある。

(3) に関しては、伝送路が基本的に下りのみであり、視聴者の要求をサーバに送ることはできない。したがって、インタラクティブ性を確保するために、すべてのデータを受け取った後、受信側ですべての処理を行う必要がある。ただし、一部のデータに関しては、ネットワークを利用する方法も検討した。

3.2 メタデータの種類

まず、オブジェクト連動データ放送サービスを実現するために必要なメタデータの種類の検討を行った。受信側でオブジェクトを特定するために必要なオブジェクトの「属性データ」と、表示によって視聴者にオブジェクトを説明するための「関連データ」とに分類した。

検討の結果、オブジェクト属性データとしては、オブジェクトを特定するためのオブジェクト ID、オブジェクト名、オブジェクトの位置と形状のデータ、同期をとるための時間情報を採用することとした。オブジェクトを特定するためにはオブジェクト ID だけで充分であり、オブジェクト名は不要であるが、記述したファイルの視認性を考慮してオブジェクト名を採用した。

オブジェクト関連データとしては、テキスト形式の説明文が不可欠である。その他の関連データとして、3 次元画像、Web ページ、音楽データを採用した。

ショッピング番組や科学番組においては、オブジェクトの 3 次元の形状を知りたいという要求がある。そのため、3 次元画像をオブジェクト関連データの一つとして採用した。3 次元画像を表示するブラウザとしては、一般的に使われていて、無料で入手できる VRML(Virtual Reality Modeling Language) ブラウザと Quick Time プレーヤを採用した。Quick Time は、厳密な意味では 3 次元情報を持っていないが、データサイズが VRML に比べて小さく、簡易バージョンとして用いることができる。

オブジェクトを説明する他の手法として、インターネットの Web ページを利用することとした。これは Web ページを表示するための HTML ブラウザが普及しているからである。オブジェクトのリンク情報としては Web ページの URL とし、これを伝送することとした。

オブジェクトが楽器や楽譜である場合、その音を聞きたいという要求がある。また、音楽番組の補完的なサービスとして利用できるため、音楽データを関連データの一つとした。ただし、今回は音楽データとして MIDI データを採用した。これは音を符号化したデータに比べて MIDI デー

表 1 メタデータの種類の種類

Meta-data.	
オブジェクト属性データ	オブジェクト ID
	オブジェクト名
	オブジェクトの形状
	オブジェクトの位置 時間情報
オブジェクト関連データ	説明文 (テキスト)
	静止画
	3 次元画像 (VRML,QT)
	Web ページ (URL) 音楽 (MIDI データ)

タのサイズが小さいことと、MIDI データのプレーヤの入手が容易だからである。採用したメタデータを表 1 に示す。

3.3 放送規格に基づいたデータ構造

メタデータのデータ構造および伝送方式について検討を行った。放送を実施する場合、放送規格および運用規定に準拠する必要がある、少なくとも、既存受信機が問題なく動作する必要がある。そのため、ARIB 規格で規定されているデータ構造および伝送方式を採用する必要がある。ARIB 規格のサーバ型放送の中には、メタデータのデータ構造が規定されているものの、規定されているメタデータが、著作権情報などを中心とするもので、我々の提案するオブジェクトに関するデータが含まれていないこと、運用規定が現在審議中であることから、メタデータの構成には、現行の規格で規定されているバイナリテーブルを採用することとした。バイナリテーブルは、ARIB 規格に規定されており、ECMA Script で読むことができるからである。なお、ECMA Script もまた ARIB 規格で規定されている。

バイナリテーブルはその名の通り、2 次元配列をした構造である。メタデータのオブジェクト属性データの中で、オブジェクトの形状と位置のデータ (オブジェクト形状データと呼ぶ) は動的であり、時間とともに変化する。したがって、1 秒ごとのオブジェクト形状データを各レコードに格納することとした。このバイナリテーブルを「オブジェクト形状テーブル」と呼ぶ。オブジェクト形状テーブルはオブジェクト毎に存在し、そのレコード数はオブジェクトが出現している秒数となる。

オブジェクト形状データ以外の関連データは時間に依存しない静的なデータである。したがって、オブジェクト形状データ以外のデータを格納するバイナリテーブルは、1 オブジェクトを 1 レコードに割当て、レコード内にオブジェクトの静的なデータを記述する形式とした。テーブルのレコード数はオブジェクト数と同数になる。これを関連データ管理テーブルと呼ぶ。静的データのうち、オブジェクト ID、オブジェクト名、説明文を必須のデータとし、その他の関連データはオプションとして扱うこととした。そこで、説明文はオブジェクト ID と同様に直接記述することとしたが、その他の関連データはファイルとして伝送し、そのファイル名を「マルチメディアファイル名」として記述することとした。

	オブジェクト ID	オブジェクト名	説明文	参照オブジェクト ID	マルチメディアファイル名
サイズ(Byte)	1	V	V	1	V
型	U	S	S	U	S

(a) 関連データ管理テーブル(可変長)

	頂点	中心座標(水平)	中心座標(垂直)	中心からの距離(水平)	中心からの距離(垂直)
サイズ(Byte)	2	3	3	3	3
型	U	S	S	S	S

(b) オブジェクト形状テーブル

	時刻(秒)	オブジェクト数	シーン開始時間(秒)
サイズ(Byte)	2	1	4
型	U	U	S

(c) 時間管理テーブル

V:Variable
U:Unsigned Integer
S:String

図 3 バイナリテーブルの構造
Data allocation in binary table.

オブジェクト形状テーブルのサイズはオブジェクトが出現する時間に依存し、関連データ管理テーブルについても、そのサイズは出現するオブジェクト数に依存する。オブジェクト数が多く、出現する時間が長いと各テーブルのサイズが大きくなる。テーブルのサイズが大きいと、すべてのデータを送るのに時間がかかる。また STB に大きな容量のメモリが必要となる。そのため、番組を適当な時間で区切り、複数のテーブルに分割して伝送することとした。テーブルのサイズは、時間だけでなくオブジェクトの数にも依存するので、一定時間で分割するのは合理的でない。一つのシーンの中で出現するオブジェクトが大きく変化することは少ないが、シーンが変わると出現するオブジェクトも変化することが多い。そのため、テーブルをシーンで分割すると効率的なことが予測される。テーブルの分割とシーンの関係については 3.4 節で述べる。

この二つのテーブル以外に、映像とメタデータとの同期をとるために、時間管理テーブルを採用した。映像ストリームのタイムコードと時間管理テーブル内のシーン開始時刻との比較を行い、シーンの開始のタイミングを合わせることで、映像とデータとの同期をとる。また、オブジェクト数から参照すべきオブジェクト形状テーブルの数を取得することができる。

以上の検討から、メタデータのデータを以下の 3 つのバイナリテーブルで構成することとした。

- (1) 関連データ管理テーブル
- (2) オブジェクト形状テーブル
- (3) 時間管理テーブル

各テーブルの構造を図 3 に示す。オブジェクト形状テーブルの図はオブジェクト形状が矩形の場合であり、この場合に限ったフォーマットである。より複雑な形状を表現する場合は、頂点数に従って欄を増やすことができる。

3.4 伝送方式

オブジェクト形状データは動的であるため、ストリーム型データとして扱う方が望ましい。映像と音声以外のストリームとしては、字幕用データを伝送している独立 PES が現在のデジタル放送で採用されている。しかし、独立 PES は、字幕以外の使用が規定されていないこと、エラー処理が

図 4 ECMA Script による動的処理
Procedure by ECMA Script.

脆弱であることからオブジェクト形状テーブルの伝送には、ファイル型データを伝送するために規定されている DSM-CC (Digital Storage Media Command Control), すなわち、データカールセル伝送形式を採用した。また、関連データ管理テーブルと時間管理テーブルについても、データカールセル伝送を採用した。

バイナリテーブルは基本的にシーン毎に分割され、モジュールに格納されて伝送される。すなわち、バイナリテーブルが更新されると、DSM-CC のモジュールが更新される。番組によっては、シーンが異なっても出現するオブジェクトが変化しない場合や、逆に、同じシーンでも異なるオブジェクトが次々に出現する場合もある。このような場合、コンテンツの内容や情景で決められるシーンとは異なる単位で、バイナリテーブルを分割することがある。ここでは、バイナリテーブルを分割した単位、つまり、DSM-CC のモジュールが更新されてから次に更新されるまでの期間をシーンと呼び、その時間をシーン時間と呼ぶこととする。

3.5 動的処理

伝送路が放送波であることから、上り回線が使用できない。このため、STB は送られてきたデータのみを使用し、カーソル位置とオブジェクトとの位置が一致しているかどうかの判断などの処理を STB で行う必要がある。このような動的処理を STB 内で行うため、ECMA Script を採用した。また、メタデータの表示には BML を採用した。現行の STB には、ECMA Script インタプリタと BML ブラウザが設定されている。

図 4 にその処理アルゴリズムの一例を示す。ポインティングデバイスからカーソルの位置を取得した後、カーソルの位置がオブジェクトの範囲に入っているかどうかの判断を行う。範囲に入っていれば関連データを表示する。範囲に入っていない場合は次のオブジェクトの範囲に入っているかどうかの判断を行う。オブジェクト数だけ繰り返し、オブジェクト数を超えると、再度、カーソルの位置を取得する。オブジェクトの位置、関連データ、オブジェクト数はデータストリームから取得する。カーソルの位置は指しポインターから取得する。

図 5 マルチメディアデータの処理
Process of Multi-Media Data.

3.6 マルチメディアデータの呈示

テキスト以外のオブジェクト関連データのデータ構造について検討を行った。これらのデータをマルチメディアデータと呼び、そのファイル名を関連データ管理テーブルに記述して伝送する。マルチメディアデータ本体の伝送は、DSM-CCを利用する方法とネットワークを利用する方法の両者に対応できるようにした。

関連データ管理テーブルに記載されているマルチメディアデータ名がファイル名の場合、STB はファイル名で参照したマルチメディアファイルをデータストリームから取得し、そのデータを PC に転送する。PC では STB から転送されたファイル名の拡張子によって、適切なプレーヤまたはブラウザを起動してデータを呈示する。

関連データ管理テーブルに記述されているマルチメディアデータ名がURLの場合は、そのURLをPCに転送する。PCでは、転送されたデータがURLの場合は、HTMLブラウザを起動し、ネットワークを通して関連するサーバにアクセスする。ダウンロードしたファイルがHTMLの場合は、HTMLブラウザで表示し、MIDI, VRML, QuickTimeの場合は、適合するプレーヤーを起動して、マルチメディアデータを再生する。図5はマルチメディアデータの処理を示す。

4. システムの動作実験

システムの動作確認とそのパフォーマンスを評価するための実験を行った.

4.1 実験に用いた TS コンテンツ

システムの動作を確認するために、バイナリーテーブル、BML ファイル、モノメディアをモジュール化し、映像と音声との多重化を行い TS(Transport Stream) を作成した。より現実的な状況での動作を確認するために、複数のシーンと番組とから構成される TS とした。番組数は 10 番組で、各番組の長さは 3 分である。番組の切り替えは PMT を更新することで実現し、ローカルコンテンツの PID を番組毎に変更した。さらに、1 番組を六つのシーンに分割し、各シーンに対応したデータカールセルを多重化することとした。実験に用いた TS コンテンツの構成を図 6 に。TS

図 6 実験に用いた TS の構成
Transport stream for the experiment.

表 2 TS 素材データの属性
Attribute of transport stream.

TS の長さ	30 分
PMT の数 (番組に対応)	10
カルセルの数 (シーンに対応)	60
オブジェクトの数	506
TS のビットレート	約 21 Mbps
DSM-CC のビットレート	1.58 Mbps
関連データの種類	テキスト MIDI URL

図 7 実験システム
Experimental system.

コンテンツの属性を表 2 に示す。オブジェクト数は 506 であり、すべてのオブジェクトにオブジェクトを説明するテキストを関連データとして記述した。テキスト以外のオブジェクト関連データとしては、MIDI データを用いた音楽データと URL を用いた Web ページを採用した。

4.2 実験システム

実験システムを図 7 に示す。作成した TS を ASI (Asynchronous Serial Interface) として伝送し、STB で受信した後に、映像、音声、データを分離し、開発した BML ブラウザを用いて、映像およびデータを表示した。BML ブラウザは基本的には ARIB の規格に準拠しているが、ポインティングデバイスのインターフェース、カーソル表示処理、TCP/IP インターフェースの各機能を STB に追加した。

4.3 番組切り替え時のメタデータ取得時間

TS の受信, BML ブラウザの動作, ECMA Script の動

作, ポインティングデバイスとのインターフェースなど, システムの基本的な機能については問題なく動作していることを確認した。

番組の切り替えについては, ローカルコンテンツのコンテンツ ID が更新され, スタートアップ用の BML 文書が読み込まれ, データ放送を開始する準備が行われ, 問題なく切り替わることを確認した。ただし, 番組が切り替わった直後に, データ放送の準備が整うまでに約 5 秒かかることがわかった。これは, 番組が切り替わってからメタデータを表示できるようになるまでに, 約 5 秒かかることを意味するが, 市販されている STB でも現状は, チャンネルを切り替えた直後はデータ放送が表示できないことを考えると, 実用上問題のない値と考えられる。

4.4 シーン切り替え時のメタデータ取得時間

先に定義したように, シーンとはモジュールが更新されてから次に更新されるまでの期間のことを言い, 関連データ管理テーブルとオブジェクト形状テーブルはモジュール化されているため, シーン毎に更新される。モジュールの更新中はメタデータの表示ができないため, シーンの切り替え直後はメタデータが表示できない。この時間が長いと実用的でないため, この時間を調査する必要がある。しかし, この時間を直接測定することは困難であるため, モジュールが更新されてからメタデータが表示可能となるまでの時間(メタデータ取得時間)を測定した。この時間は, モジュールのサイズに依存し, モジュールのサイズは, シーンに出現するオブジェクト数とシーン時間に依存する。なぜなら, オブジェクト数は, 関連データ管理テーブルのレコード数とオブジェクト形状テーブルの数に比例し, シーン時間はオブジェクト形状テーブルと時間管理テーブルのレコード数に比例するからである。このため, オブジェクト数とシーン時間に対するメタデータの取得時間を測定した。また, この測定により, 実用的に使用できるオブジェクト数とシーン時間を推定することができる。

モジュールが更新された瞬間から, 関連データが表示可能になるまでの時間, すなわち, メタデータの取得時間を STB の内部時計を用いて測定した。測定に要する処理の負荷は無視できるほど小さい。オブジェクト数とシーン時間以外の TS の条件は 4.1 節と同様である。実験に用いた TS のオブジェクト数は, 10, 20, 30, 40, 50 の 5 種類である。シーン時間は, 10 秒から 100 秒までの 10 秒間隔とした。各条件でのメタデータの取得時間を測定した結果を図 8 に示す。取得時間はシーン時間よりもオブジェクト数に依存することがわかる。

4.5 メタデータ取得時間に関する考察

測定の結果, メタデータ取得時間の実測値にはばらつきがあることがわかった。実用的に伝送できるオブジェクト数とシーン時間を求めるためには, メタデータを確実に取得できる時間, すなわち, メタデータ取得時間の最大値を求めなければならない。そこで, メタデータの取得時間を以

図 8 メタデータの取得時間 (実測値)
Capturing time for meta-data.

下の処理に分けてメタデータ取得時間の最大値を推定した。

- (1) セクション番号 0 の待ち時間 (T_1)
- (2) セクションが揃うまでの時間 (T_2)
- (3) セクションを表示可能な形式にする時間 (T_3)

3 つに分割した取得時間について考察する。

- (1) セクション番号 0 の待ち時間

データカルーセル伝送では, メタデータは複数のセクションに分割されて順番に伝送される。すべてのセクションが伝送されたら, 再びはじめのセクションから伝送される。このため, モジュールが更新された直後に STB に到達するセクションは不定である。しかし, STB はモジュールのセクション番号 0 のセクションから取得を開始するため, STB は, モジュールが更新されてから, セクション番号 0 のセクションを受信するまで, データの取得を待たなければならない。この待ち時間の最大値はセクション番号 0 の伝送間隔となり, 測定値はゼロから最大値まで一様に分布すると予想される。セクション番号 0 の待ち時間の最大値 T_1 は, セクション番号 0 の読み込み時間を含めるとカルーセルの周期であり, データの総数とビットレートから式 (1) で表すことができる。

$$T_1 = \frac{(S_b + S_k) \times 4066 \times 8}{B} \quad (1)$$

ここで, S_b は常用データモジュールのセクション数, S_k はシーンデータモジュールのセクション数, B は DSM-CC のビットレートを表す。4066 は 1 セクションのバイト数である。常用データモジュールとは, BML 文書など番組を通して同一のデータが使用され, シーンによって変化しないデータのモジュールのことである。本実験ではこれに 2 セクション使用しており $S_b = 2$ である。

それに対して, シーン毎に異なるデータをシーンデータモジュールと呼ぶ。このモジュールは, メタデータを格納したバイナリテーブルと, メタデータの表示枠などの画像ファイルから構成される。画像ファイルのサイズはシーンに依存しないで固定であり, 実験では 2 セクションを使用している。バイナリテーブルは, 関連データ管理テーブル, オブジェクト形状テーブル, 時間管理テーブルの 3 種類からなり, それぞれのバイト数を L_{mt} , L_{ot} , L_{tt} とすると, セクションは 4066 バイトバウンダリとなるから, シーンデータモジュールのセクション数 S_k は

$$S_k = \lceil ((L_{mt} + L_{ot} + L_{tt})/4066) \rceil + 2 \quad (2)$$

となる．ここで， $\lceil \cdot \rceil$ は天井関数であり，与えられた実数を下回らない最小の整数を示す．

次に，各テーブルのサイズを以下に示す．関連データ管理テーブルは可変長であるが，今回は最大値を求める．オブジェクト名の最大サイズが 20 バイト，関連データの最大サイズが 200 バイトである．本実験では，マルチメディアファイルは使用していないので，マルチメディアファイル名の欄は空である．オブジェクト ID と参照オブジェクト ID に，各 1 バイトを割当てているので，関連データ管理テーブルの 1 レコードは 222 バイトとなる．したがって，関連データ管理テーブルのサイズ L_{mt} は，オブジェクト数を N_{obj} とすると，

$$L_{mt} = 222N_{obj} \quad (3)$$

となる．

オブジェクト形状テーブルは，1 レコードが 14 バイトであり，シーン時間分のレコードが必要である．また，オブジェクト数分のテーブルが必要であることから，オブジェクト形状テーブルのサイズ L_{ot} は，シーン時間を T_s とすると

$$L_{ot} = 14T_sN_{obj} \quad (4)$$

となる．

時間管理テーブルは，1 レコードが 7 バイトであり，シーン時間分のレコードが必要であるから，時間管理テーブルのサイズ L_{tt} は

$$L_{tt} = 7T_s \quad (5)$$

となる．

これにより，オブジェクト数とシーン時間からセッション番号 0 の待ち時間 T_1 を算出することができる．オブジェクト数が 30 の時の算出された最大値 T_1 と実測値の比較を図 9 に示す．セッション 0 の待ち時間の計測は STB のタイマーを利用した．ただし，タイマーの精度が 1/60 秒であるため，実測値は約 0.017 秒の誤差を生じる．この比較から実測値は推定値以下であることが分かる．図中の実線は式 (1) から求めた推定値で点が実測値である．推定値は待ち時間の最大値であり，測定値はその範囲内で一様に分布すると考えられ，図においても測定値が推定値以下で分布していることがわかる．

(2) セクションが揃うまでの時間

STB はインターリーブ方式でセッションを取得する．本実験では 2:1 インターリーブ方式を採用した．2:1 インターリーブ方式では 1 回目に偶数番号のセッションを読込んだ場合，2 回目は奇数番号のセッションが読込まれる．すなわち，読込みに必要なシーンデータモジュールのセッション数は 2 倍となる．厳密には，セッション番号 0 のセッションはす

図 9 セクション番号 0 の待ち時間 (推定値と実測値)
Waiting time for Section 0.

図 10 セクションが揃うまでの時間 (推定値と実測値)
Capturing time for Modules.

でに読込まれているから，セッション数の 2 倍マイナス 1 となる．常用モジュールに関しては，シーンが変化しても読込む必要はないが，カーセルに含まれるので，その間，読み飛ばす必要がある．したがって，セッション番号 0 の取得した後にモジュールに必要なセッションが揃うまでの時間 T_2 は以下の式で表すことができる．

$$T_2 = \frac{(S_b + S_k \times 2 - 1) \times 4066 \times 8}{B} \quad (6)$$

セッションが揃うまでの時間とその実測値を図 10 に示す．実測値がほぼ計算通りの値になっていることがわかる．このことにより，データの取得に関しては，STB はオーバーヘッドなく動作しており，推定通りの動作をしていることがわかる．

(3) セクションを表示可能な形式にする時間

モジュールに必要なセッションが揃うと，BML のスクリプトに従ってモジュールのバイナリテーブルが読出され，表示可能な形式にして STB のメモリに記録される．この処理は，処理中のタスク数や負荷によって変化するため算出できない．このため，実測値を使うこととした．実測値はばらつきがあるため，実測値の最大値 T_3 を用いることとした．

(4) メタデータの取得時間の最大値

以上，算出した最大値 T_1, T_2, T_3 を合計し，実測値と比較を行った．図 11 にその結果を示す．実線は推定値を点の実測値を示す．すべての実測値は推定値以下であることから，最大取得時間の推定は適切と考えられる．この結果から，メタデータの取得時間が最大 1 秒となる条件は，「オ

図 11 メタデータの取得時間 (推定値と実測値)
Capturing time.

プロジェクト数が 20 個以下でシーン時間が 80 秒以下」であることがわかる。実際の運用を考えた場合、シーンが切り替わった直後に視聴者がオブジェクトを指定することは困難なため、シーンの切り替え後 1 秒間、操作ができないとしても許容されると思われる。

5. む す び

放送の本格的なデジタル化を迎えて、デジタルの特徴を生かした放送サービスの一つとして、我々は「オブジェクト連動データ放送サービス」の提案を行い、このサービスを実現するためのシステムを試作し、動作確認を行った。

このサービスを実現するための技術的課題は、できるだけ現行規格に準拠させることであった。その結果、

- (1) ブラウザに BML を使用する
- (2) 動的処理に ECMA Script を使用する
- (3) データ構造にバイナリテーブルを使用する
- (4) 伝送方式に DSM-CC を使用する

ことで、システムを構築し、提案するサービスが実現可能なことを示した。

また、試作したシステムを用いて実験を行った結果、80 秒以内のシーンに 20 個までのオブジェクトのメタデータを伝送できることがわかった。実験に用いた STB は開発用ではあるが、現行の STB と同程度のスペックであり、また、STB に入力した信号は現行の BS デジタル放送と同程度のビットレートであることから、デジタル放送の信号形式の妥当性、および、STB の実現性の観点から、提案した「オブジェクト連動データ放送サービス」の実現性の可能性が得られたと考える。ただし、オブジェクトの抽出追跡、メタデータの制作、および、これらをリンクさせるオーサリングシステムについては、研究開発を進めているもののサービスを開始できる技術レベルには至っていない。実用化にはこれらの研究開発が必要であり、課題として今後も取り組んでいく。

なお、本研究は独立行政法人情報通信研究機構の助成を受けて行われた。

〔文 献〕

- 1) 金次保明, 苗村昌秀, 鄭文寿, 三須俊彦, オブジェクト連動データ放送システムの検討, 2001 年映像学年次大, pp.176-177(2001)
- 2) 金次保明, 苗村昌秀, 境田慎一, 鄭文寿, 三須俊彦, オブジェクト連動データ放送システムの検討, 2001 年信学総大, A-16-28, p.351 (2001)
- 3) 勘解由哲, 藤井達也, 鈴木鎮男, データ放送番組の制作と送出, 映像学誌, 56, 10, pp.1546-1548 (2002)
- 4) 岡村和男, 角谷和俊, 宮部義幸, IplusTV: 時系列ハイパーメディアを用いたデータ放送方式とその実装, 情報学論データベース, Vol.40, No.SIG5(TOD2), pp.91-102 (1999)
- 5) ARIB STD-B38:サーバ型放送における符号化、伝送及び蓄積制御方式標準規格, 社団法人 電波産業会
- 6) 磯部忠, 藤原正雄, 金田裕之, 浦谷則好, 森田寿哉, テレビ受信ナビゲーションシステムの開発と評価結果, 映像学技法, 27, 42, pp.5-8 (2003)
- 7) 両角聡, 向井信彦, 小杉信, 野球中継ダイジェスト版の自動生成システム, 映像学技法, 27, 66, pp.33-36 (2003)
- 8) <http://www.tv-anytime.org/>
- 9) <http://tnavi.net/>
- 10) <http://www.televito.com/>
- 11) <http://www.tv-marker.com/>
- 12) 鄭文寿, 三須俊彦, 苗村昌秀, 和泉吉則, 金次保明, 複数カメラの協調動作による動オブジェクトの抽出, 信学論, 86-D-II, 3, pp.397-408 (2003)
- 13) 金次保明, 浦谷則好, 小峯一見, 指さしポインターの開発とその性能評価実験, 信学技報, 104, 329, pp.9-12, (2004)

かなつぐ やすあき
金次 保明 1980 年, 関西大学大学院修士課程修了。同年, NHK に入局。大分放送局を経て, 1983 年より NHK 放送技術研究所に勤務。1995 年より 1998 年まで ATR 人間情報通信研究所, 2000 年より 2005 年まで NHK エンジニアリングサービスに転出, 現在に至る。その間, デジタル映像処理, 放送方式, 視覚情報処理の研究に従事。正会員。

たかはし まさき
高橋 正樹 1997 年, 慶應義塾大学理工学部管理工学科卒業。1999 年, 同大学大学院修士課程修了。同年 NHK 入局。岡山放送局技術を経て, 2002 年より同放送技術研究所に勤務。画像認識の研究に従事。正会員。

みつひろ としひろ
三須 俊彦 1994 年, 東京大学工学部電子工学科卒業。1996 年, 同大学院修士課程修了。1999 年, 同大学院博士課程修了。1996 年より 1999 年まで, 日本学術振興会特別研究員。1999 年, NHK に入局, 同放送技術研究所に勤務し, 画像オブジェクト抽出・追跡, パターン認識, 画像計測, データ放送の研究に従事。博士 (工学)。正会員。

たけち まさる
武智 秀 1990 年, 東北大学大学院工学研究科 (電子工学専攻) 修士課程修了。同年, NHK 入局。以来, 放送技術研究所において衛星放送システム, デジタル伝送方式, マルチメディアサービス, デジタル受信機アーキテクチャの研究に従事。現在, 同研究所ネットワークシステム専任研究員。正会員。

か いけんじろう
加井 謙二郎 1981 年, 早稲田大学理工学部電子通信学科卒業。同年, NHK に入局。富山放送局, 技術本部を経て, 1989 年より放送技術研究所にて, ハイビジョン文字表示, データ放送, ISDB の研究に従事。現在, 放送技術局チーフエンジニア。正会員。