

## マルチメディア文書の同期再生方法

斎藤 崇弘 中村 修 清水 剛

富士ゼロックス システムコミュニケーション研究所

マルチメディア文書は音声や動画像などの時間軸を持つ情報要素から構成されている。この文書を出力する際にはそれぞれの要素を同期させて再生する必要がある。本稿ではマルチメディアシステムの中で情報要素を同期再生するための一手法について述べる。本手法は情報要素の同期関係を表現するインタフェース記述を解釈して同期再生を行うことを目的とする。情報要素の再生時間を予測可能とするために、負荷変更を行って再生時間が変動しない再生部を実現した。その上で個々の再生部に対する先行制御時間を算出してスケジューリングを行い、再生を試みた。この結果、並列的および逐次的な要素の再生において0.27秒以内の誤差で同期を取ることができた。これは厳密な同期を要求しない用途に対して十分な誤差範囲と言えよう。

## A SYNCHRONIZATION CONTROL METHOD IN A MULTIMEDIA DOCUMENT

Takahiro Saito Osamu Nakamura Takeshi Shimizu

Fuji Xerox Co.,Ltd. Systems and Communications Laboratory

9-14 Naka-cho 4-Chome, Atsugi-shi, Kanagawa 243, Japan

Multimedia documents, having time axis, consist of elements such as sound and video data. These elements should be synchronized with each other while playing operation. This reports a method for the synchronization control in our multimedia system. The method interprets relations of each element, described in the interface script. First, several playing parts are realized without fluctuation by means of "load control", so that the system can predict how long time to play each element. Then we tried to schedule the elements by "timing control in advance". As the result, parallel and serial elements can be synchronized within 0.27 second. These errors are sufficient for documents which don't need to synchronize strictly.

## 1 はじめに

筆者らは、マルチメディア文書の作成支援を目的としたマルチメディア統合環境 MediaPreview [1]~[4]の研究を行っている。MediaPreviewはカードとリンクにより構成されるハイパーテキストモデルを基礎としている。様々な情報要素をカードとして実現しているため、情報要素の種別や保存形態に依存しない操作環境を提供している。また、時間的、空間的、論理的構造を表現するモデルを実現している。

マルチメディア文書を出力する上で最も重要な課題の一つとして、複数の情報要素間での同期の実現が挙げられる。特にデジタルデータとAV機器が混在したときには、実際のデータ保持形態や表現形式が異なるので同期を取ることが難しい。

本稿では、まずマルチメディア文書の表現形式について述べる。次に時間的構造を持つマルチメディア文書を再生する上での同期における問題点とその解決方法について述べる。

## 2 MediaPreview システム

マルチメディア文書における同期を考えるために、時間的構造の表現形式と問題点を述べる。

### 2.1 マルチメディア文書モデル

一般的なマルチメディア文書の空間的構造と時間的構造の関係について図1に示す[5]。

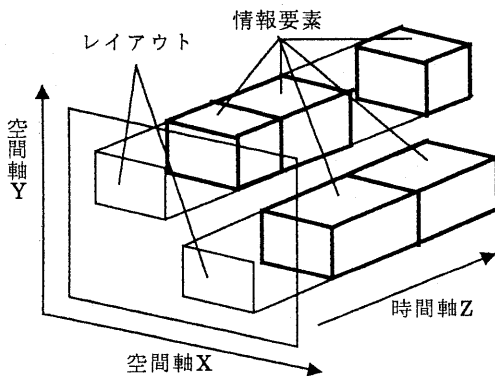


図1 空間的構造と時間的構造の関係

マルチメディア文書における時間的構造は、逐次的な時間関係と並列的な時間関係の組合せに

よって表現できる[6]。MediaPreviewでは、逐次的、並列的な時間関係をグループ化しハイパーテキストを用いてモデル化した(図2)。

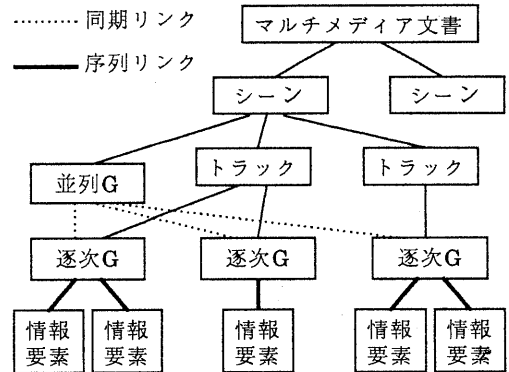


図2 マルチメディア文書モデル

- 逐次グループ

情報要素間の逐次的な時間関係を序列リンクにより管理する。これは情報要素を連続して再生することを意味する。

- 並列グループ

逐次グループ間の並列的な時間関係を同期リンクで管理する。これは情報要素を同時に再生、終了することを意味する。

- トラック

複数の逐次グループの空間的関係を管理する

- シーン

マルチメディア文書の段落を管理する。

### 2.2 インタフェース記述

図3にシーンの編集インタフェース例を示す。

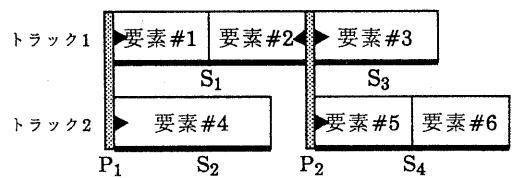


図3 シーン編集例

縦のシェードは並列グループ、▶は同期リンク、太線は逐次グループを意味する。例えばP<sub>2</sub>では要素#2の終了と同時に要素#3と要素#5が開始されることを意味する。また、S<sub>1</sub>では要素#1と要素#2が連続して再生されることを意味する。

MediaPreview ではこれをインタフェース記述に変換する。例えば、図3を変換すると以下のようになる。

```

Start Scene #1
  SeqGroup Seq#1 (element#1 element#2)
  SeqGroup Seq#2 (element#4)
  SeqGroup Seq#3 (element#3)
  SeqGroup Seq#4 (element#5 element#6)
  ParaGroup Para#1
    (Start:Seq#1 Start:Seq#2)
  ParaGroup Para#2
    (End:Seq#1 Start:Seq#3 Start:Seq#4)
EndScene #1
  
```

このインタフェース記述を解釈して情報要素を再生する。

### 2.3 システム構成

図4にシステム構成を示す。

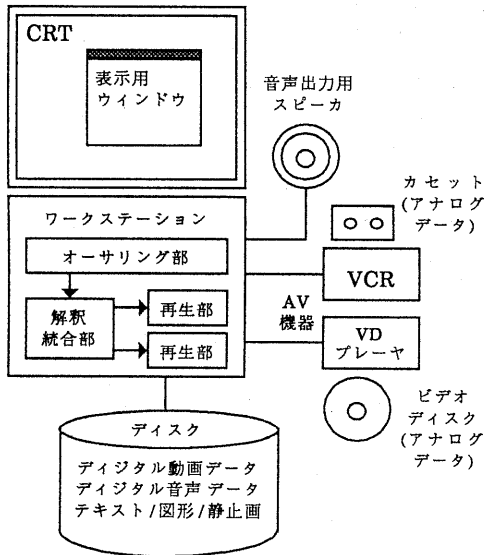


図4 システム構成

マルチメディア文書を構成する情報要素は以下のように再生される。

#### ① テキスト/図形/静止画データ

コード化したデータがディスクに蓄積され、展開して表示用ウィンドウに表示される。

#### ② デジタル動画データ

圧縮してディスクに蓄積され、再生部で伸張しながら表示用ウィンドウに表示される。

#### ③ デジタル音声データ

PCMのように圧縮してディスクに蓄積され再生部で伸張しながらスピーカへ出力される。

#### ④ AV機器(アナログデータ)

ワークステーションを介してVCRやビデオディスクプレーヤのようなAV機器を操作し、アナログ動画や音声の再生が行われる。

### 2.4 問題点

インタフェース記述を単純に解釈して各情報要素の再生部に要求を出すシステムでは、例えば図5のように所望値と実現値に差が生じてしまう。

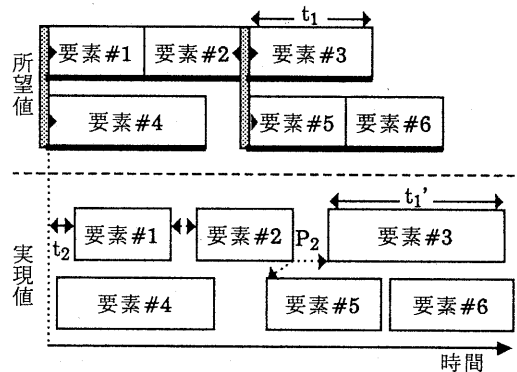


図5 再生例

#### (1) 再生時間が延びる

デジタル動画/音声データにおける再生において、例えば、要素#3の再生時間が本来 $t_1$ であるべきものが、 $t_1'$ に延びることがある。これはデータを伸長するための時間がかかるためである。

#### (2) 開始と終了時間が遅れる

例えば、要素#1の再生開始時間が、 $t_2$ だけ遅れてしまう。これは内部の処理時間やAV機器の動作時間などが必要なためである。

#### (3) 同期関係が崩れる

例えば、 $P_2$ で指定した並列グループの同期関係で、要素#2の終了と同時に要素#3と要素#5が開始していない。これは、(1)(2)の要因の他、インタフェース記述の同期表現を反映した制御を行っていないためである。

### 3 同期再生の方針

同期再生を実現するために、まず再生時間を予測することを試みる。再生時間を予測可能とするために次のような方針で進めることにする。図6はその方針を示す。

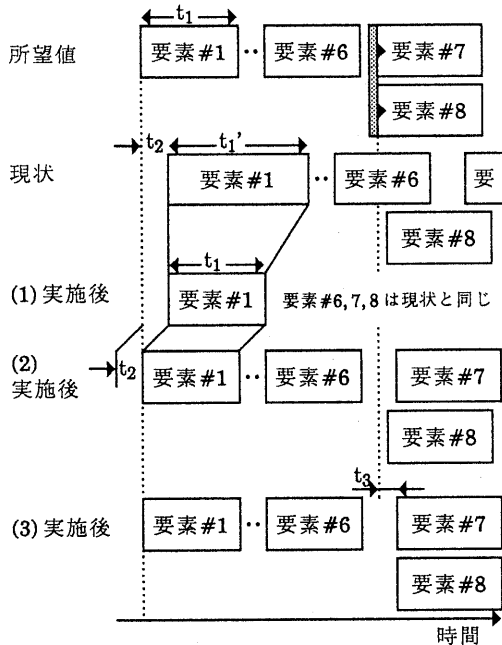


図6 解決方針

#### (1) 再生時間を固定的に再生する

デジタルデータ再生において、情報要素を時間的に短縮、伸長して同期させスケジューリングする方法も提案されている[7][8]。ここでは再生時において情報要素を固定時間で再生させる ( $t_1$  は  $t_1'$  のように延びないようにする)。

#### (2) 開始と終了時間の遅れを先行制御で回避する

情報要素の再生部での遅れを計算し、 $t_2$  時間前にコマンドを発行するような先行制御を行う

#### (3) ユーザー指定の同期関係を重視する

例えば、要素#6と要素#7が同一のAV機器で、かつシーク動作が入った場合は時間が必要である。このときは、並列グループで示される要素#7と要素#8の同期関係を重視し、 $t_3$  に合わせるよう要素#8を移動する。

### 4 同期再生の手法

#### 4.1 固定長の再生方法

再生する情報要素には表1のような特徴がある

表1 情報要素の特徴

情報要素	特徴	CPU 負荷	CPU 負荷 変更	出力 バッファ	出力 時間 遅れ
テキスト/図形/静止画		中	不可	なし	中
デジタル動画		大	可	なし	中
デジタル音声		小	不可	あり	小
AV機器		小	不可	なし	大

#### (1) CPU 負荷

ソフトウェアでの処理の重さであり、デジタル動画の処理はCPUに負荷がかかる。

#### (2) CPU 負荷変更

処理方法によりCPU負荷を変えられるかどうかで、デジタル動画の場合には再生フレーム数を少なくして処理を軽くすることができる

#### (3) 出力バッファ

出力のためのキューを備えているかどうかであり、デジタル音声の場合にはPCM用のバッファがある。

#### (4) 出力時間遅れ

再生部に開始命令が発行されてから実際に再生が開始されるまでの時間であり、特にAV機器の再生では出力時間遅れが大きい。

以上の特徴を鑑み、デジタルデータの再生部では以下のような方法を取ることににより再生時間を固定的なものとする。

#### ① デジタル動画データ再生部

情報要素を連続的に再生するために負荷に応じて再生フレームを変更する[9]。これは、タイマを監視しながら時間に応じたフレームを表示する。

#### ② デジタル音声データ再生部

CPUの余裕があるときにハードウェアで用意されているバッファを満たす制御を行う。

## 4.2 先行制御による同期制御方法

同期再生を行うための構成を図7に示す。

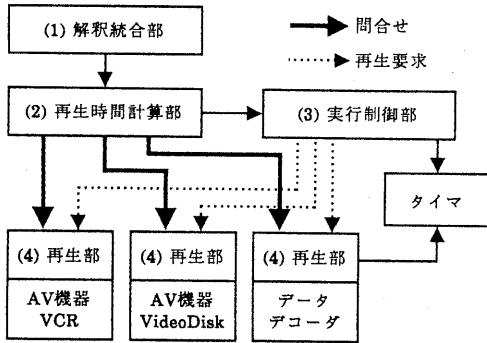


図7 同期再生のシステム構成例

各部分での処理内容を以下に示す。

### (1) 解釈統合部

インタフェース記述を解釈し、個々の情報要素の再生開始および終了時間に展開し、再生時間計算部に通知する。

### (2) 再生時間計算部

それぞれの遅延時間を各再生部に問合せ、情報要素の再生開始、終了時間から制御のための指定時間を計算する。この制御テーブルを作成し実行制御部に送る。テーブルの例を表2に示す。

表2 制御テーブルの例

再生部名	オペレーション	指定時間
VCR	PLAY	12:29:55.50
デジタル動画	PLAY	12:29:59.30
VCR	STOP	12:30:25.50
ビデオディスク	PLAY	12:30:28.85
ビデオディスク	STOP	12:30:59.50
デジタル動画	STOP	12:30:59.80

### (3) 実行制御部

表2のようなテーブルにしたがってタイマを監視しながら、対応する再生部に要求を出す。

### (4) 再生部

先行制御時間のための問合せに対して、自分の再生における遅延を返す。AV機器再生での遅れ ( $t_{AVdelay}$ ) とデジタルデータ再生、テキ

スト等の再生での遅れ ( $t_{DDdelay}$ ) は次式で表される。

$$t_{AVdelay} = t_{cal1} + t_{cal2} + t_{com} + t_{device}$$

$t_{cal1}$ : 実行制御部での処理時間

$t_{cal2}$ : 再生部での処理時間

$t_{com}$ : AV機器とのコマンド通信時間

$t_{device}$ : AV機器がコマンドを受信してから実際に動作するまでの時間。状態により変わる。

$$t_{DDdelay} = t_{cal1} + t_{cal2} + t_{read}$$

$t_{read}$ : デジタルデータ読み出し時間

それぞれの再生部では、 $t_{cal2} + t_{com} + t_{device}$ 、または、 $t_{cal2} + t_{read}$  を返す。この値は、実測値の平均から算出した値を用いる。

## 4.3 並列、逐次同期の制御方法

並列、逐次同期を維持するために図7の再生時間計算部での処理内容を以下のごとく追加する。

表2を作成するとき、同一再生部に対してのオペレーションの時間的な重なりがあるかチェックする。オペレーションの重なりとは、例えば、PLAY動作中にSEEK動作(データ格納媒体中の位置検索)が来ていないかなどである。もし重なりがある場合には、並列グループ、逐次グループの関係を崩さないように重なりの方を移動する。

## 5 結果

以上のシステムを試作し、誤差を測定した。測定方法は、表示画面をVCRで撮影し再生してフレーム数をカウントしたものである。そのため、測定の精度は1/30秒となる。

図8に測定箇所を示す。

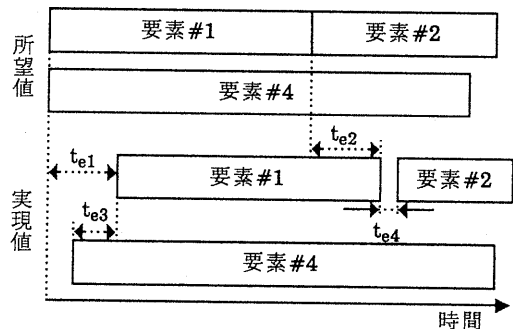


図8 測定箇所

$t_{e1}$ : 指定時間に対する再生開始誤差

$t_{e2}$ : 指定時間に対する再生終了誤差

$t_{e3}$ : 並列グループの再生誤差

$t_{e4}$ : 逐次グループの再生誤差

表3に測定結果を示す。

表3 測定結果(単位:秒)

	従来の 平均値	本方式での最悪値 (0内は標準偏差)
再生開始誤差(VCR)	1.1	0.33 (0.12)
再生開始誤差(VD)	0.4	0.30 (0.08)
再生終了誤差(VCR)	0.5	0.30 (0.09)
再生終了誤差(VD)	0.4	0.33 (0.11)
並列グループの誤差	1.3	0.27 (0.16)
逐次グループの誤差	1.3	0.23 (0.08)

(1) 指定時間に対する再生開始および終了誤差

個々の情報要素を再生するときの開始および終了の誤差は0.33秒以下に収まっている。他の情報要素が再生されない場合には十分な精度である。この誤差の原因としては、AV機器本体の誤差とオペレーティングシステム内で生じる誤差(例えば、タイマの誤差、プロセス切り替えに伴う誤差)が考えられる。

(2) 情報要素間の同期

従来、秒のオーダーでの同期しか取れなかったものが、並列再生、逐次再生とも誤差が0.27秒以下に収まっている。これは目的としていた情報要素間の並列関係と逐次関係における同期ずれがかなり改善された。ただし、人の口の動きと声を一致させることが必要な場合などは誤差の減少が必要となる。

## 6 おわりに

以上のように、情報要素の再生時間を予測可能とするために、負荷変更を行って再生時間が変動しない再生部を実現した。その上で個々の再生部に対する先行制御時間を算出してスケジューリングを行い、再生を試みた。この結果、並列的および逐次的な要素の再生において0.27秒以内の誤差で同期を取ることができた。これは厳密な同期を

要求しない用途に対して十分な誤差範囲と言えよう。

今後の課題としては、さらなる誤差の減少や高精細な動画、高品質な音声に対応可能な再生部の実現などがある。

## 7 参考文献

- [1] 清水: MediaPreview: マルチメディア情報操作における階層的インターフェイスモデル, Proceedings of Advanced Database Symposium '91, pp.117-125 (1991)
- [2] 清水他: MediaPreview: マルチメディア情報を利用した番組企画/シナリオ作成システム, 情報処研報, Vol.92, No.89 pp.29-36 (1992)
- [3] 中村他: MediaPreview: マルチメディアシステムにおけるデータ操作記述フォーマットの設計, 情報全大秋 (1993)
- [4] 清水他: MediaPreview+: データベース・インタフェースを備えたマルチメディア・ドキュメント生成支援システム, 情報処コンピュータシステム・シンポジウム, pp.9-16 (1993)
- [5] R. Ogawa et al.: A Scenario-based hypermedia: a model and a system, Proceedings of the European Conference on Hypertext, INRIA, France, Nov, pp.38-51 (1990)
- [6] A. Poggio et al.: CCWS: A Computer-Based Multimedia Information System, IEEE Computer, Oct. (1985)
- [7] M.C. Buchanan et al.: Specifying Temporal Behavior in Hypermedia Documents, Proceedings of the European Conference on Hypertext '92, Dec. (1992)
- [8] R. Hamakawa et al.: Object Composition and Playback Models for Handling Multimedia Data, ACM Multimedia (1993)
- [9] Inside Macintosh: QuickTime, Apple Computer, Inc. (1992)