

メディア間の同期情報を含んだ マルチメディア文書

米田 健 市川 敬史 佐藤 龍雄
林 正薫 屋代 智之 橋本 康江 松下 温
慶応大学理工学部

今日、CPUやメモリーが飛躍的に進歩した結果、コンピュータは数値や文字、グラフィックだけでなく音声や映像も扱えるようになった。

本論文では、時間的要因を含んだマルチメディア文書を、メディア間の時間関係、メディア間の空間関係に着目して構造化した。構造化された文書は、コンピュータの知的処理の対象となる新しいメディアと言える。

文書の構造化においてはODA (Office Document Architecture) の概念を利用した。そして、ASN. 1 (Abstract Syntax Notation One) によって文書を記述し、実際に文書交換システムを試作し、文書を構造化することの意義を確認した。

A Method for Specifying Structure
of Time Dependent Multimedia Document

Takeshi Yoneda Takafumi Ichikawa Tatsuo Sato
Tyonfun Lim Tomoyuki Yashiro Yasue Hashimoto Yutaka Matsushita
Faculty of Science and Technology
Keio University

In this paper, it is showed that multimedia information is characterized by the temporal and spatial relationships between media and a method for formally specifying time dependent multimedia document structure is proposed. The method is based on ODA(Office Document Architecture).

By means of the method, temporal relationships between media, as in the simultaneous representation of voice and video or voice and image, can be described. ASN.1(Abstract Syntax Notation One) is used for describing the time dependent document in order to interchange the document between computers.

1. はじめに

古くから、映像や音声はテレビ、ラジオによって扱われてきた。今では、音声はカセットテープに録音され、映像はビデオテープに録画されるのが当たり前となっている。記録された情報は、編集されたり、繰り返し再生されたりするが、それらの操作において情報の内容理解は人間が行い、機械自体は情報の内容を理解しない。

一方、コンピュータにおいては、CPUやメモリーが飛躍的に進歩し、数値や文字、グラフィックだけでなく音声や映像も扱えるようになった^[1]。コンピュータはデータ構造とアルゴリズムを用いて、知的な処理ができる。つまり、データ構造を理解しながら構造中の内容を取り出したり、新しい内容に更新したり、という処理ができる。

したがって、音声、映像、数値、文字、グラフィックというマルチメディア情報をコンピュータが扱えるデータ構造で表現することができれば、既存のビデオテープ、カセットテープを用いた処理とは異なったより知的で複雑な処理ができるのではないかと考えた。

本研究では、マルチメディア情報を構造化しコンピュータで処理するという形態を考えることにより、時間的要因を含むマルチメディア文書という新しいメディアを作ることができることを示す。具体的には、文書の中で、時間と共に静止画が変化し、静止画の変化に対応して音声やテキストが表示されたり、動画の内容に合わせて脚注が静止画で示されたり、ということが可能となる。

文書の構造化においてはODA (Office Document Architecture)^[2]の概念を利用した。また、文書をネットワークで接続されたコンピュータ間で交換するためにASN.1 (Abstract Syntax Notation One)^{[3][4]}によって文書を記述した。

そして、時間的要因を含む文書を交換するシステムを試作し、文書を構造化することの意義を確認した。

2. マルチメディア情報の捉え方

マルチメディア情報をコンピュータが理解できるデータ構造にするためには、マルチメディア情報がどのような要素によって構成されるかを理解しなくてはならない。マルチメディア情報の例として以下のようなものを考える。

- (1) 紙芝居のように、静止画が変わると共にそれに対応するテキストも変わる。
- (2) 図鑑のように、映像と音声による解説があり、それを補足するために静止画やテキストが必要に応じて表示される。

このようなマルチメディア情報と、テキストのみ映像のみといった単一メディアの情報の大きく異なる点は、個々のメディアの内容情報以外にメディア間の関係を表わす情報が必要になることである^[5]。 (1), (2)の例の場合、各メディアがお互いに関連して提示されるので、情報の中に各メディアがいつどこにどのような関連を持って表示されるのかという情報が含まれなくてはならない。つまり、メディア間の空間関係の情報とメディア間の時間関係の情報が必要となる。以下それらについて説明する。

2.1 メディア間の空間関係情報

メディア間の空間関係情報とは、各メディアが互いに関連を持ちながらどこに提示されるかを示す情報である。例としてFig. 1のような文書を考える。この文書は、テキスト1、2、グラフ、静止画から構成されており、各々は紙面の（あるいはディスプレイ上の）指示された位置にFig. 2のように割付けられている。Fig. 3の木構造はFig. 2の割付け情報をもとにFig. 1の文書を表現したもので、文書は一つのページからなり、そのページには4つのブロックが含まれていて、それぞれのブロックの内容がテキスト1、2、グラフ、静止画であるということを表示している。このようにメディア間の空間関係情報を含む文書はFig. 3のような木構造によって表現できる。

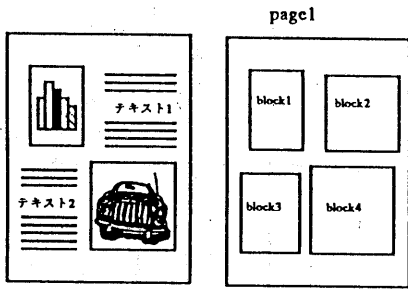


Fig.1 空間関係 Fig.2 割付け

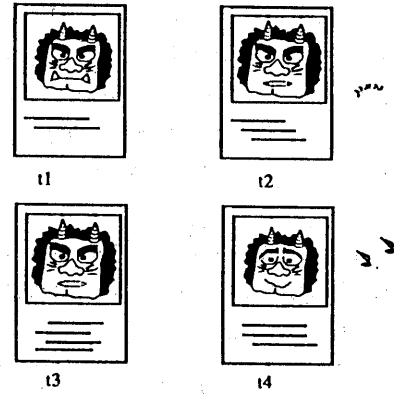


Fig.4 時間と共に変化する文書

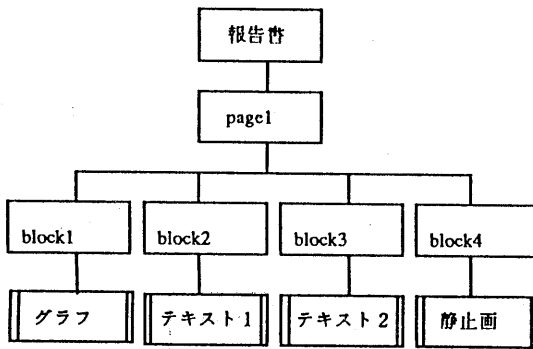


Fig.3 空間関係を含む文書表現する木構造

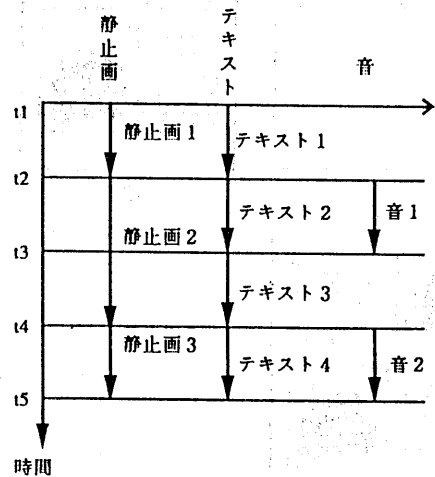


Fig.5 時間関係の表現

2.2 メディア間の時間関係情報

メディア間の時間関係情報とは、各メディアがそれぞれどのようなタイミングで提示されるかを表す情報である。Fig. 4に示した文書例では、時間とともに提示されている静止画、テキストが変化し、その変化に合わせて音楽も変わっている。Fig. 5はFig. 4の文書中のどのメディアがどのぐらいの時間提示されているかが表現されている。メディア間の時間関係はFig. 5のように表現できる。

3. 時間的要因を含むマルチメディア情報の構造化

以上の考察から、時間的要因を含むマルチメディア文書は、メディア間の時間関係情報、空間関係情報、各メディアの内容情報という3つの本質的な構成要素から成り立つことがわかった。本研究ではこの3つの構成要素をすべて取り込んだ構造として、Fig. 3の木構造のblockの上位に時間の要素を入れた構造を新しく提案する。具体的に、次のように時間と共に変化する文書を考えてみる。

(1)時刻t1において、block1,block2にそれぞれグラフ1、テキスト1が提示される。
 (2)時刻t2において、さらに、block3,block4にそれぞれテキスト2、静止画1が提示される。
 (3)時刻t3において、テキスト2がテキスト3に変わる。

これを図示したのがFig.6で、空間関係はFig.7、時間関係はFig.8のようになる。この文書の構造はFig.9のように表現できる。

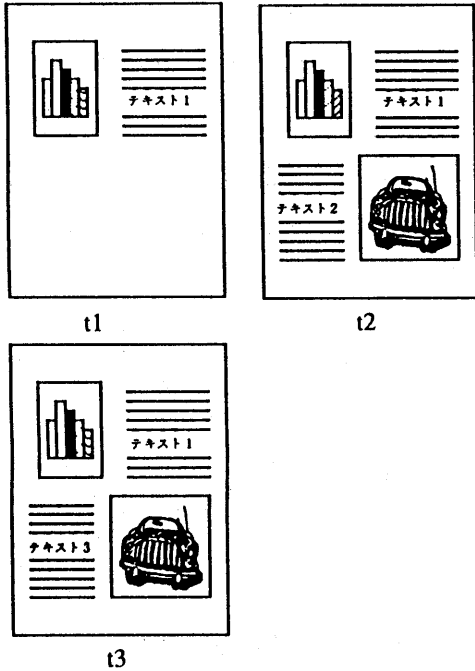


Fig.6 時間的要因を含む文書の例

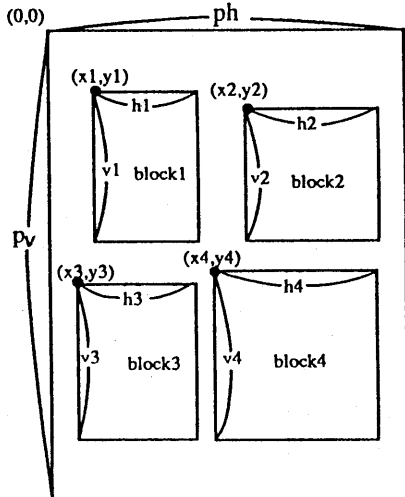


Fig.7 空間関係

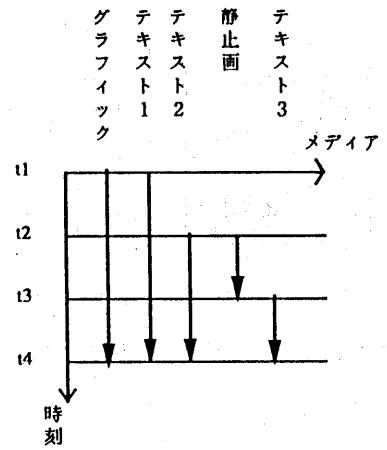


Fig.8 時間関係

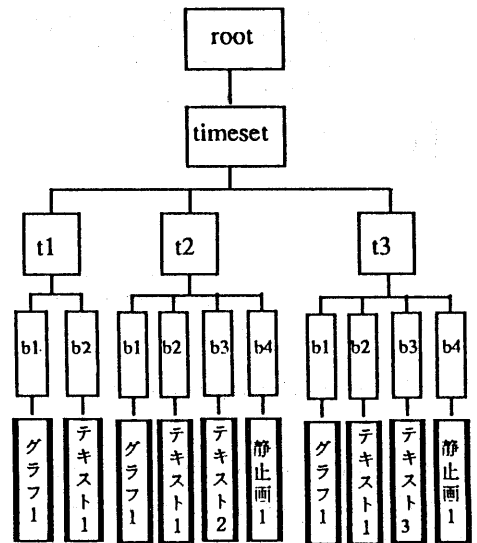


Fig.9 時間的要因を含む文書の構造

Fig. 9に示された構造はどの時間に、どこに何が表示されるか表現している。

木構造の一つ一つの構成要素を対象体と呼ぶ。この木構造では"root", "timeset", "time", "block", "content"の5種類の対象体が用いられている。"root"は木構造のルートを示す。"time"においては、下位に従属する内容を提示する時刻が記述される(提示開始時刻を0とする)。“block”においては、下位に従属する内容が提示される位置、寸法が記述される。例えば、Fig. 7より、block1は左上の位置が(x1, y1)、寸法が縦v1、横h1であるので、x1, y1, v1, h1がblock1の位置と寸法の属性の値として記述される。“content”においては実際の内容情報が記述される。“timeset”は、下位に従属する“time”に記述される時刻が同じ時間軸上にあることを示す。

したがって、Fig. 9の木構造は、この文書の提示状態が時刻t1, t2, t3と変わり、t1においては、block1, block2の場所にそれぞれグラフ1、テキスト1が提示され、t2においては、さらにblock3, block4にテキスト2、静止画1が提示され、t3においてはblock3だけテキスト2からテキスト3に変わるということを示している。

4. 可能となる知的処理

時間的要因を含むマルチメディア文書を構造化しコンピュータで処理することにより、既存のビデオデッキ、カセットテープレコーダではできない、知的な処理が可能となる。以下、検索、編集、提示という3つの処理について述べる。

4.1 検索

(1)内容に着目した検索

一度情報を見た後で、ある内容をもう一度見たいというときに、内容に着目した検索が必要となる。

構造化されたマルチメディア情報にはその内容部に内容が何であることを示すインデックスを付けることができる。したがって、内容に着目した検索が容易にできる。

(2)時刻に着目した検索

ある時刻に何が提示されているかを知りたい場合に、時刻に着目した検索が必要となる。その場合、着目した時刻を属性に持つ対象体“time”をまず見つけ、その下位に従属している内容を提示すればよい。

(3)位置に着目した検索

一度情報を見た後で、ある特定の位置に表示されていた情報をもう一度見たいというときに、位置に着目した検索が必要となる。その場合、着目した位置を含む対象体“block”をまず見つけ、その下位に属する内容を提示すればよい。

4.2 編集

(1)メディア別の内容編集

コンピュータは構造化されたマルチメディア情報をメディア別に処理しているので、あるメディアの内容を変更したい場合、そのメディアの内容だけを変更すればよい。これは、字幕スーパーのように、テキストが映像に埋め込まれてしまうビデオではできないことである。

(2)メディア別の提示場所に関する編集

あるメディアの提示場所を指定したり、変更したりすることは対象体“block”の中の位置の属性を指定、変更するだけで容易に行える。

(3)メディア別の提示タイミングに関する編集

あるメディアの提示タイミングを指定したり、変更したりすることはどの対象体“time”の下位に属させるかを指定、変更する、または“time”に記述される時刻そのものを変更することにより容易に行える。

4.3 マルチメディア情報の提示

“time”で指定された時刻にその“time”の下位に従属している内容を“block”で指定された位置に提示するという処理を行う。この提示処理はビデオやカセットのシーケンシャルな情報の再生と違い、コンピ

ュータが、どの時間にどの位置にどのメディアを提示すればよいかを把握して行われる。

5. ASN.1による記述

構造化された時間的要因を含む文書は、標準的なデータ記法で記述されることにより、あらゆるコンピュータ間で交換できることになる。そこで、Fig. 9で表現される構造を持った文書をASN.1(Abstract Syntax Notation One)で記述することにした。

時間情報に関する部分(対象体"timeset"、対象体"time"、時間属性"time"及び"interval")、時間の単位(1/60sec)、カラーの表現(Red, Green, Blueの256階調)、内容部の参照(reference-to-text-unit)といった拡張部分以外はODAにおける文書のASN.1による記述(データ型およびデータ値の記述)をそのまま用いた。

(1) ルート対象体の記述

```
layout-object {
  object-type 0,
  descriptor-body {
    object-identifier "1",
    subordinates {"0"},
  }
},
```

ルート対象体はlayout-objectである。object-typeは対象体の種類を示す。この場合は"root"であるので0である。object-identifierは木構造における対象体の位置を一意に識別するための対象体のIDである。この場合は1である。subordinates(従属体)は自分に直属する対象体の数を示す。1つならば"0"、2つならば"0","1"のように記述する。この場合は1つなので"0"である。

(2) "timeset"の記述

```
layout-object {
  object-type 1,
  descriptor-body {
    object-identifier "1 0",
    subordinates {"0","1","2"},
    positions {
      horizontal 0,
      vertical 0 },
    dimensions {
      horizontal ph,
      vertical pv }
  }
},
```

"timeset"は、layout-objectであり、共通の時間軸を持った"time"を従属体を持つ。また、画面上での文書の外枠の寸法もこの対象体の中で指定される。

object-typeは"time"であるので1。object-identifierは"1 0"。3つの従属体を持っている。Fig. 7より、文書の外枠の寸法は横(horizontal)がph、縦(vertical)がpvである。

ここで位置および寸法の単位はBMUを用い、
1BMU = 25.4mm * 1/1200
である。

(3) "time"の記述 ("t2"の記述例)

```
layout-object {
  object-type 2,
  descriptor-body {
    object-identifier "1 0 1",
    subordinates {"0","1","2","3"},
    user-visible-name "t2",
    time-attributes { interval 300 }
  }
},
```

"time"はlayout-objectである。object-typeは、"time"であるので2である。time-attributesは時間属性であり、timeとintervalがある。timeは下位に

従属する内容の表示時刻を示す。また、intervalは表示継続時間を示す。つまり、interval 300とは次の"time"の下位に従属する内容が表示される時刻までの時間間隔が300であることを示している。時間の単位は1/60 secである。ここではt2とt3の間隔を5secとしたいので intervalを300としている。また、user-visible-nameは対象体に付けられた名前である。

(4)"block"の記述 ("t2"に属する"b4"の記述例)

```
layout-object {
  object-type 4,
  descriptor-body {
    object-identifier "1 0 1 3",
    content-portions ("0"),
    user-visible-name "SEISHIGA",
    positions {
      horizontal x4,
      vertical y4 },
    dimensions {
      horizontal h4,
      vertical v4 }
    presentation-attributes {
      content-type 1 }
  }
},
```

"block"はlayout-objectである。object-typeは"block"なので4である。Fig. 7より、block4の位置は、左上の位置が(x4, y4)、寸法が横縦(h4, v4)である。content-typeは下位に従属するメディアの種類を表す。テキスト (formatted-character) であれば0、ラスタ図形 (raster-graphics) であれば1、アニメーション (animation) であれば2である。この場合静止画であるので1である。もしテキストであるなら、さらに、文字の色 (colour)、サイズ (font-size)、スタイル (font-style) を定める。

また、reference-to-text-unitの値として、前に表示される内容部のIDを記述することで、下位に従属する内容部を省略することができる。

(5)"content"の記述 (上記blockに属する静止画1の記述例)

```
text-unit {
  content-portion-attributes {
    content-identifier-layout "1 0 1 3 0",
    raster-gr-coding-attributes {
      number-of-pels-per-line 256,
      number-of-lines 245 }
  },
  content-information <car>
},
```

内容部はlayout-objectではなく、text-unitである。"内容部情報 (content-information) は内容の実体である。内容部属性 (content-portion-attributes) は内容部固有の属性であり、この場合 raster-gr-coding-attributes (ラスタ図形属性) として画像の縦横の画素数 (縦256、横245) を記述する。

他の対象体も同様に記述される。

6. 試作システムの構成

時間的要因を含んだ文書を交換するシステムを試作した。システムは、SUN-Workstation上でC言語を用いて実装された。データの流れと処理はFig. 10のようになっている。

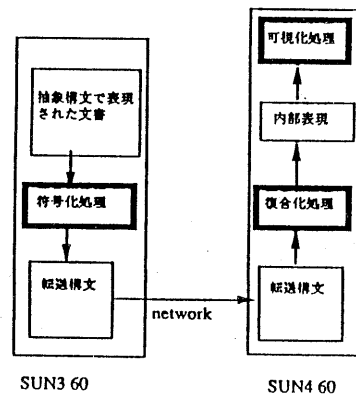


Fig.10 データの流れと処理

(1) 符号化処理

抽象構文は、符号化処理により、オクテット単位のバイナリコードに変換される。変換規則は標準で規定されている^[4]。

具体例を示す。

抽象構文	転送構文
layout-object {	A221
object-type 2,	020102
descriptor-body {	311C
object-identifier "1 0 0",	41053120302030
user-visible-name "t0",	8E027430
subordinates {"0", "1", "2"},	A009120130 12013110132
time-attributes {	BA04
interval 300 }},	8102012C

(2) 復号化及び可視化処理

バイナリで送られてきたデータは復号化部により内容が解釈されてFig. 12のように、それぞれポイントで結ばれた構造体に変換される。

この構造体はそれぞれFig. 11の対象体に対応し、対象体の属性の情報以外に可視化に必要なウィンドウの情報などを持つ。

可視化はX-window環境でC言語を用いて行った。時間的な制御はこの構造体を左から順にたどって行くことにより実現した。

まとめ

マルチメディア情報を、各メディアの内容、メディアの空間情報、時間情報という点から構造化した。構造化した情報をコンピュータで処理することにより、時間的要因を含む文書という新しい表現メディアを作った。

実際にコンピュータ間で時間的要因を含む文書の交換を行い、有効性を確かめた。

参考文献

- [1] 稲葉 則夫、浅見 直樹、"21世紀へ、始動するマルチメディア"、日経エレクトロニクス、no. 471、pp. 124-126, Apr. 17, 1989
- [2] ISO 8613:1989, Information processing -Text and office system- Office Document Architecture(ODA) and Interchange Format
- [3] JIS X 5603 "開放型システム間相互接続の抽象構文記法1 (ASN.1) 仕様
- [4] JIS X 5604 "開放型システム相互接続の抽象構文記法1 (ASN.1) の基本符号化規則仕様
- [5] Joseph S. Sventek "AN ARCHITECTURE SUPPORTING MULTI-MEDIA INTEGRATION", Proc. IEEE Office Automation Symposium, pp. 46-55, 1987
- [6] 梶浦 正規、菅野 正孝、水谷 憲司 "マルチメディア通信に関する概念モデルの一検討" 情報処理 マルチメディア通信と分散処理研究会, pp17-24, 1990.

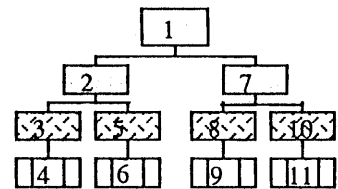


Fig.11 文書構造

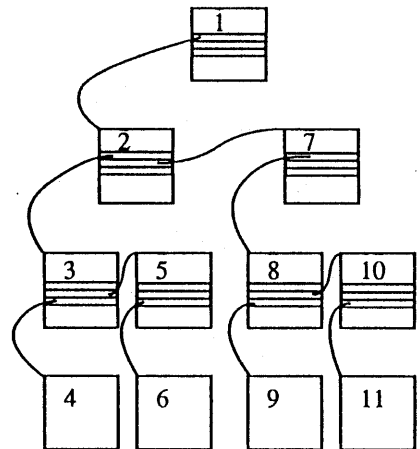


Fig.12 機種に依存した構造