

映像散策のためのビデオハイパーモデルの提案

坂田 哲夫, 木原 民雄, 小島 明, 佐藤 哲司

NTT 情報通信研究所

本論文では、3つの条件(メディアの構造化・インタラクション・自律性)を備えたマルチメディア・アプリケーションを映像散策と呼ぶ。更に、コンテンツ作成上の要請として順応性・モジュール性・構造独立性をも考慮に入れて、映像散策に適したデータモデルとしてビデオハイパーモデルを提案する。

本モデルの特徴は(1)画像・音声・テキストの組合せをシナリオという単位で構造化でき、(2)それらをリンクで結んで上位のシナリオを合成でき、(3)動画像中の被写体からリンクを張ることでシナリオを連結して新たなシナリオにできることである。

Video Hyper Model: A Data Model for Video Walk Through

Tetsuo SAKATA, Tamio KIHARA, Akira KOJIMA and Tetsuji SATOH

NTT Information and Communication Laboratories

In this paper, we propose multimedia application "Video Walk Through" which has three characters; activity, structure and autonomy, we think them important for multimedia applications. After taking adaptivity, modularity and data-independency into account to make contents efficiently, we propose "Video Hyper Model" which is for describing video walk through.

In this model, (1) one can take images, sounds and texts into "scenario" or a unit of structured multimedia, (2) can link them together to make them into an upper level scenario and (3) can link a subject in a moving image to a scenario.

1 はじめに

近年、映像や音声を扱うためのハードウェアの低廉化、通信のデジタル化を背景に、映像や音声を統合して扱うための、いわゆるマルチメディアシステムの提案がなされている。

これらのシステムの中で、実用化されもつとも広く使われているのが、WWW に代表される、ハイパーテキストである。

ハイパーテキストは、テキストを基本とし、そこからハイパーリンクで各種メディアの情報を引き出す仕組みであるが、次のような欠点がある。

- マルチメディアシステムとしてみた場合、複数メディアを組み合わせて提示するための方法がない。
- 情報を提示されているユーザが、何らかの働きかけをしない限り、新たな情報の提示の展開が生じない。

この二つが解決されれば、ユーザが動画を眺め、音声を聞くことで情報を受けとるときに、あらかじめ定められた情報を受動的に受けとるだけでなく、自分の興味や必要に応じて、次々と情報を選択することが可能となる。

こうしたマルチメディア応用を総称して我々は映像散策と名付ける。ビデオハイパーモデルはこれを実現するためのモデルである。

その実現にあたっては、マルチメディアは不特定多数のユーザへ情報を与える有用な手段する立場にたつて、佐藤等の主張する情報システムの3要件 [1] に言う、interactive かつ individual な情報を提供できるモデルとした。

以下、第2章ではこの映像散策を実現するための条件を考察する。それ以降、第3章で映像散策を実現するためのビデオハイパーモデルへの要請を考察し、第4章でこのデータモデルの定義を与え、第5章ではこれによる記述例を紹介する。

2 映像散策の提案

先の章で指摘したハイパーテキストの欠点は、例えば、以下のような応用例において顕著になる。

- 電子化マニュアル — ユーザは動画による解説画面を見、音声ガイダンスによる説明を受けつつ、場合によってはテキストを参照して、情報を得、自分が知りたい箇所を選択することで、より詳細な情報を提供する場面を視聴することが出来る。これには複数メディアの組合せを要する。
- ウォークスルー — 街路のような、広がりのある場所の動画像を切り替えながら再生することで、あたかもその街路を歩いているかのような動画像をユーザに与える。街路の分岐点でユーザが次に進みたい方向を選べるだけならば、ユーザが何もしないなら、ずっとそこで待たざるを得ない。自律的な再生が出来れば、「あらかじめ決めておいた」道を提示することも可能となり、単なるプレゼンテーションとユーザの意図に基づくナビゲーションとを統合して実現するメディアが可能となる。

映像散策は、上記のようなアプリケーションを含んでいると考えられる。この例の考察

から、映像散策を可能にするマルチメディアの条件は以下の3つであると考える。

構造化 映像を含む複数のメディア（動画・音声・テキストなど）を組み合わせてユーザに提示する方法を記述することが出来る。

自律性 ユーザが特別に何かの操作をしなくとも、時間の経過に沿って次々と情報が提示される。

インタラクション 提示される情報のある部分にユーザが興味を持った時、その意図をシステムに伝えることで、提示される情報の系列を変えることができる。¹

これらの他、映像散策を本質的に含んでいる応用としては、上記の他、マルチメディア化されたCAI・ゲームなどがあり、映像散策になっている応用分野はかなり広いと考えている。

3 ビデオハイパーモデルへの要請

我々は映像散策には、3つの条件が必要であることを先に論じた。本章では、実用上の状況を考慮に入れて、映像散策を可能にするためのマルチメディアのデータモデルであるビデオハイパーモデルへの要請を考察する。

先に論じたのは、情報の受け手から見た時の条件であったが、情報の提供側のことも考慮するなら、次の条件をも満たすべきである。

順応性 ユーザの状況やメディアを再生する機器など、再生時に定まる条件によって、ユーザに与える情報を特化できること。

こうすれば、さまざまなタイプのユーザ

¹これは佐藤等[1]のいう interactive の要件をハイパーメディアにおいて実現する一手法である。

に与えるべきコンテンツ・さまざまな端末でそれに応じて再生されるコンテンツを一元的に記述できるようになる。これは、佐藤等[1]の言う individual の概念を、あらかじめ格納されている情報によるハイパーメディアによって実現する一手法である。

モジュール性 コンテンツをモジュール化・階層化できること。こうすれば、コンテンツの再利用が容易になり、また、部分的に開発することが可能になるため、複雑なコンテンツの作成も見通し良く行なえる。

構造独立性 実際のメディアデータと、ビデオハイパーモデルのデータとは独立性を持たせ、動画像や音声などのメディア素材データの入れ換えを容易にする。

先に挙げた3条件と上記の3条件の全てを満たす、映像散策のためのデータモデルを、ビデオハイパーモデルと呼ぶ。次の章では、これらの要請を満たすデータモデルを設計する。

4 ビデオハイパーモデル

先の章で見た通り、ビデオハイパーモデルには6つの基本的要請が課せられている。これを大きく二群に分けると、情報の構造を規定する静的な側面（シナリオ）と、ユーザとのインタラクションを可能にする動的な側面（被写体検索）とがある。そこで、ビデオハイパーモデルを構築するため、静的・動的の両側面から、このデータモデルを設計する。

4.1 シナリオ — ビデオハイパーモデルの静的側面

ビデオハイパーモデルへの6つの要請のうち、5つ—構造化・自律性・モジュール性・順応性・構造独立性—はデータ構造に依存する静的な要請と言える。これらを満たすデータ構造をシナリオと呼ぶ。

構造独立性を除く4つの要請はいずれも、一つの全体的なコンテンツがいくつかの小部分から構成され、必要に応じて組み合わせられる構造を持つことを要請している。

すなわち、構造化からは、音声・動画の素材を別個に作っておいて、ユーザへの与え方を、組合せの形で記述できることが要請される。

次に自律性の要請から、モデル上では、あるコンテンツの小部分の次に自律的に再生される小部分を記述できる必要がある。このため、先行関係を記述できることが必要である。

順応性からは、コンテンツには条件によって再生される小部分とそうでない部分とが含まれている必要がある。これを時間的に見れば、コンテンツの分岐が記述できることである。

この三つの条件を同時に満たすデータ構造としては、有向グラフがあるが、単にコンテンツの小部分を組み合わせ、有向グラフにただけでは構造化とモジュール化の要請には十分応えられない。この両者はともに、有向グラフをまとめてあたかも1つのノードとして扱うことを求めている。そこで、有向グラフ内の部分グラフをもノードとして扱えるようにしてこの点を解決した²。

² 藤川等の Harmony では、同様の有効グラフ構造を用いて複数のメディアを同期して表示する方法を実現している。[2]

このデータ構造をシナリオとして提案する。(図1を参照)

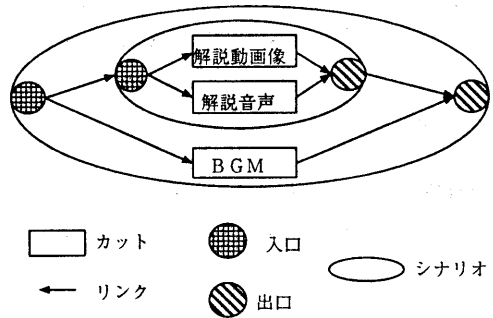


図1: シナリオの例

この図1のシナリオは、ある内容を解説する動画と音声とが同時に流され、それにBGMが同時にながれることを表現している。

シナリオは基本的には、リンクおよび他のシナリオから再帰的に構成される有向グラフ構造をしており、その要素は次のように定義される。

- シナリオは、シナリオどうしがリンクで結ばれたものである。
- それ以上構造を持たないシナリオをカットと呼ぶ。カットには動画・音声のほか静止画・テキストなどの種類がある。

上記のカットにはメディアデータそのものを格納するのではなく、それらへの参照情報を格納する。こうすることで、カットの内容を大幅に変更せずにメディアデータを入れ換えることが可能となり、ビデオハイパーモデルの構造独立性が保たれる。

ここに、シナリオどうしを結合するリンクは次のような属性を持つ。

元 シナリオ上のリンクは、遷移元のシナリオのIDを持つ。

先 遷移先のシナリオの ID を持つ。

条件 一つのシナリオから複数のリンクが出ている場合、そのうちのどれが選択されるべきかを定めるための条件を与える。

先・元は自律性のために、条件は順応性のために必要である。

これに加えて、モジュール化の要請から、シナリオは他のシナリオから参照されるために入口・出口と名付ける特別な節点を持つ。入口からメディアの再生が始まってシナリオを辿ってゆき、出口に到達して、かつその先に他のシナリオが接続されていないならば、再生を終了するものと解釈する。

4.2 被写体検索 — ビデオハイパーモデルの動的側面

先のシナリオは 6 条件のうちの 5 つを解決している。ここでは、コンテンツの再生時の問題であるインタラクションを実現するための動的側面について述べる。

映像散策では、メディアの自律性があるため、ユーザが特に操作しない場合でも一定の内容が展開してゆく。しかし、これのみではユーザに満足を与えることは難しく、ユーザインタラクションを可能にすることが必要である。そこで、シナリオの流れを切り替える機能として、被写体検索を提案する。

これを可能にするためには、ユーザが興味を抱いた動画中の部位をポインティングデバイスで指示することで、システムに「興味のあるか」を伝え、それに従って、関連するシナリオを提供するように、再生の流れを切り替える機能を具備すればよいと考える。

こうすることによって、既に与えられたシナリオではあるが、ユーザはシナリオの進行

に影響を及ぼし、それによって提示される情報の系列が変わるので、ユーザへも影響が及ぶというインタラクションが成立する。

4.3 ビデオハイパーモデル — 静と動とが結合されたモデル

これまで、ビデオハイパーモデルの静的ならびに動的な側面についてそれぞれ説明してきた。しかし、これらがバラバラのままでは不都合である。特に被写体検索の結果として、動画像や音声による情報をユーザに与えようとする時、それらの情報を構造化する手段が必要になる。そこで我々は、シナリオと被写体検索の 2 つの要素が統合されたビデオハイパーモデルが必要であると考えている。

4.3.1 統合のしかた

シナリオと被写体検索を統合するために、リンクの機能を拡張する。シナリオの内部でシナリオどうしを結び付けているリンクに、被写体を指示する機能を付加すれば、被写体検索によるシナリオの切り替えと、シナリオによる情報の構造化とが統一されたモデルの上で実現できるはずである。

そこで、先に提案したリンクを次のように拡張する。リンクの元に遷移元のシナリオの ID がある場合に加え、被写体検索の場合には被写体として識別すべき対象を記述する。例えば、高野等が実施しているように [3]、動画像の上の被写体をボタン (矩形の領域) の集合を割り当てる場合は、
(画面のフレーム番号, ボタンの座標)
の集まりを付加すれば良い。(図 2 を参照)

4.3.2 統合による利点

1. 被写体検索とシナリオ構造の設計が同一

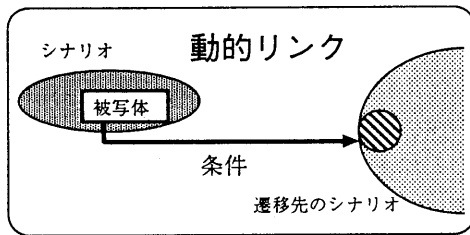
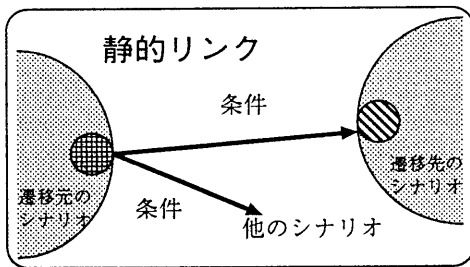


図 2: 統合されたリンクによるシナリオの記述

のレベルで行なえる。

2. 被写体検索の場合の参照先が複数になっても、シナリオの分岐と同様の手法で処理できる。同一の被写体から複数の情報を選択的に引き出す機能は、CAI やゲームのようにユーザの事情によって与えて良い情報が違う場合には必須の機能である。
3. 被写体検索の指示する先をシナリオにできるので、被写体検索による遷移とシナリオによる遷移とが同一のものとして扱える。このため、次の利点がある。

- 被写体検索の先にくるシナリオを取り替えることで、全体的なシナリオを部分的に更新することが出来る。これはインクリメンタルなコンテンツの制作をより一層容易にする。

ると同時に、シナリオの再利用方法が増えるため、既存のシナリオを元に新たなシナリオをたやすく制作できる。

- ユーザから見ると、シナリオの切り替え方法に関係なく、ずっと動画(映像)を見たままで場面の切替が行なわれるので、臨場感が向上する。

5 ビデオハイパーモデルの記述例

本章ではビデオハイパーモデルを使った映像散策の例を示す。ここではウォークスルーを題材とし、説明音声の特化・部分シナリオの差し替え・交差点の記述法とを示す。

説明音声の順応 シナリオの順応性の例として、ある一つの動画に対応する説明音声の言語をユーザの要求によって、切り替える方法を示す。この例では動画像 (*video*) はつねに選択されるが、音声の方は *Lang* という変数の値によって *Japanese* か *English* かが決まる。

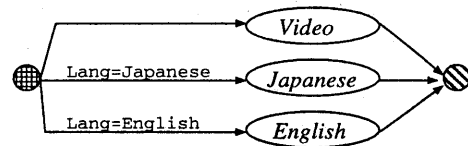


図 3: シナリオの順応の例

部分シナリオの差し替え 仮定の町を題材とするウォークスルーのシナリオにおいては、街路のイメージの一部を差し替えたいことが比較的頻繁に生じると予想している。例えば、

ある建物を増設する・建物内の特定のフロアを模様があえるなどである。

今、図4のような街路をモデルとしたウォークスルーを作ったとしよう。ここで、図中のAの部分に新たに建物1を追加して、その中もまたウォークスルーが出来るように拡張することを考える。そこで、右図中のBで表されるような、この建物1の内部のウォークスルーシナリオを作成して追加する。

また、この建物1をまた別の建物と差し替える場合には、この部分のシナリオだけを入れ換えれば良い(右図のB'のように)。このように、シナリオを利用したウォークスルーはコンテンツを段階的に開発すること・部分ごとに保守することを可能にする。

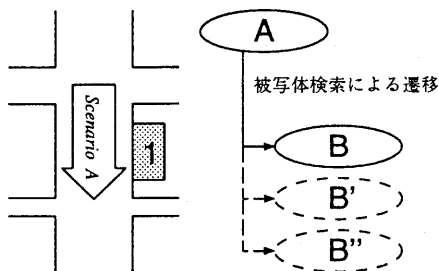


図4: ウォークスルーでの建物の追加

交差点 ウォークスルーでの交差点は、ノードとリンクが複雑に絡まり合う箇所である。これに対して、シナリオの構造化・モジュール化の利点を発揮する方法を示す。

十字路の交差点は図5のように一般化できる。³ここで、今矢印Aに沿って交差点に進入する場合を考える。この時、ユーザは右左折と直進の中から次の動作を選ぶことが出来るの

³この他にも実装方法はいろいろあり得るが、カットをなるべく再利用する方針に沿ってこの図のようにした。

が自然であるが、この図にある全ての矢印に相当する全部で28個の動画カットを24個のリンクで結ぶ必要がある。

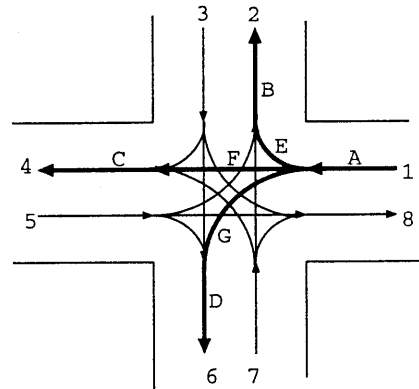


図5: ウォークスルーでの交差点の記述

通常の有向グラフだけを用いるシステムでは、交差点ごとにこうしたシナリオを記述する必要があるが、我々のモデルでは、一度これをシナリオとして登録してから、交差点が出現するごとにそれを複製して、実動画像を入れ換えれば良く、作業の効率を大きく改善できる。

さらにこの交差点の例は、4重の対称性を持っているので更に効率化が可能である。まず、図6の右図のように、進路1に沿って進入する場合のシナリオのみを記述する。交差点全体ではこれと同じ形のシナリオが4つあるので、一度このシナリオを作ってから、シナリオを複製して、実画像入れ換えを行えば(同図左)、交差点自身もさらに少ない少ない手間で作成できる。

このように、シナリオの階層化機能は、小さなシナリオを繰り返し再利用して複雑なシナリオを少ない手間で作ることを可能とする。ただし、シナリオの再利用には、建物の入れ換

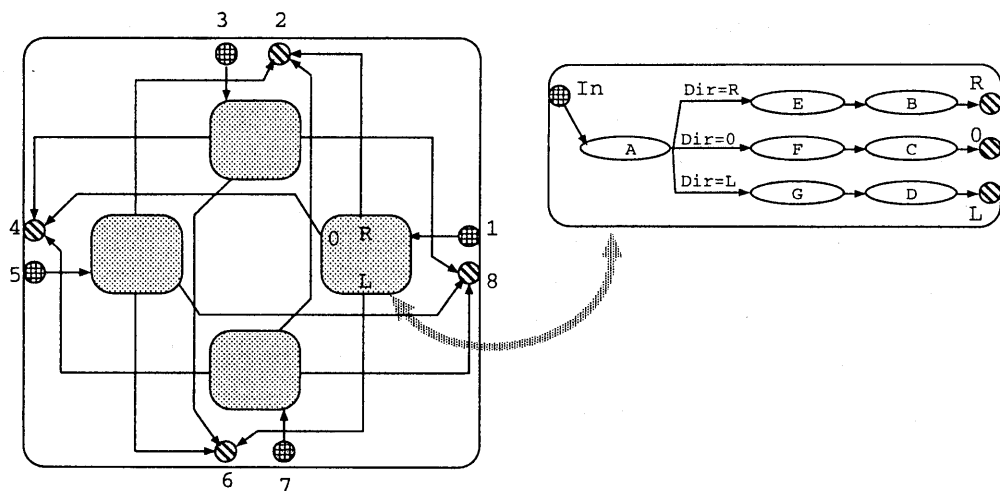


図 6: 交差点の階層的記述

えの例のようなデータまで含めた再利用と、交差点の例のような構造のみの再利用という二つの形態があることがわかる。後者の場合、既に作ったシナリオの構造だけしか利用しないのに、シナリオ全体の複製を一旦作っており無駄が多い。この解決は今後の課題である。

6 まとめ

我々は、映像散策を実現するためのマルチメディアとして、ビデオハイパーモデルを提案した。このモデルはメディアの構造化・自律性・インタラクション・構造独立性・モジュール性・順応性の6つの条件を満たしている。ハイパーテキストとの大きな差異は、この6つの条件を満たすことで、マルチメディアによるプレゼンテーションとナビゲーションとが一つに統合できる点にある。

参考文献

- [1] 佐藤, 片岡, 井上: “ビデオリアリティ: 映像を用いた情報検索手法の高度化”, 情報処理学会データベースシステム研究会 DBS-99-38, July 1994.
- [2] 藤川, 下條, 松浦, 西尾, 宮原: “オブジェクト指向に基づくハイパーメディアシステム Harmony の構築”, 信学論 (D-I), J-57-D-I, 11, pp. 1015-1024, November 1992.
- [3] 高野, 的場, 原: “ビデオ・ハイパーメディアのナビゲーション方式”, Symposium on Human Interface(8th), pp. 607-612, October 1992.