

映像散策のためのビデオハイパーメディア クライアントサーバ型 VHM のアーキテクチャ

坂田 哲夫 柴垣 斉

NTT 情報通信研究所

{sakata, shiba}@dq.isl.ntt.co.jp

〒 239 横須賀市光の丘 1-1

佐藤 哲司

NTT 技術部

satoh@es.hqs.ntt.co.jp

〒 163-19 東京都西新宿 3-19-2

あらまし 映像散策とは、主として動画によって提示される情報を受けとりながら、ユーザがインタラクティブに所望の情報を検索していくマルチメディアアプリケーションである。筆者らは映像散策のための枠組として VHM (ビデオハイパーメディア) を提案している。VHM は、映像・音声等のメディアデータを断片化してノードとし、シナリオと呼ぶデータ構造に格納することでインタラクティブなハイパーメディアを実現する。本稿では、VHM を用いて情報共有を図るため、マルチメディアコンテンツを提供する VHM サーバ・コンテンツとユーザの入力を解釈するリゾルバ・データの入出力を行なうブラウザの 3 つの要素からなるクライアントサーバ型のアーキテクチャを提案し、試作した実験システムの概要を述べる。

キーワード ハイパーメディア, マルチメディアデータベース

Video Hypermedia for Walk-through Applications; Client-server Architecture for VHM systems

Tetsuo SAKATA Hitoshi SHIBAGAKI

NTT Information and Communication Systems

Laboratories

{sakata, shiba}@dq.isl.ntt.co.jp

1-1 Hikarinooka Yokosuka-Shi Kanagawa 239 Japan

Tetsuji SATOH

NTT Technology Department

satoh@es.hqs.ntt.co.jp

19-2 Nishi-Shinjuku 3-chome

Shinjuku-ku, Tokyo 163-19 Japan

Abstract Video Walk-through is a multimedia application in which a user getting audio-visual service can search information by interacting with a system. We have proposed VHM (Video HyperMedia) for Walk-through applications. Using VHM system, an author writes multimedia contents from cuts which are fragments of media data such as Video and Audio data, and makes them into a scenario which is a data structure for interactive multimedia applications. In this paper, we propose a client-server architecture for VHM. It consists of three sub-system; a VHM-server serves multimedia contents, a resolver interprets the contents and user-inputs, and a browser treats with inputs from a user and output media-data for him. We also report a prototyped client-server VHM system.

Key words Hypermedia, Multimedia Database

1 はじめに

近年、ハードウェアの低廉化・通信の高速化を背景に、映像・音声など複数の種類のメディアをつかって情報を表現するマルチメディア情報システムの研究開発が盛んに行なわれている。

筆者らの提案するビデオハイパーメディア (Video HyperMedia: VHM) では、映像や音声のデータを断片(カット)化して蓄積・管理しておき、それらが再生される際の順序と条件をシナリオとして記述することで、マルチメディア情報(コンテンツ)を記述する。再生時にコンテンツを再構成することで、ユーザはマルチメディア情報をインタラクティブに検索できる。この、動画を中心としたハイパーメディアを用いたインタラクティブな検索を映像散策と呼ぶ [2, 3]。

映像散策が、不特定多数の一般ユーザに受け入れられるためには、次の2つの特徴が必要だと考える。(1) インタラクティブであること。(2) ユーザの種別ごとの要求に応えること。また、(3) 情報の提供者がデータの追加が容易に行なえることも必要である。本稿では、これら3条件を満たしつつ情報を共有するためのVHMシステムのサーバクライアント型のアーキテクチャを提案し、試作した実験システムの概要を紹介する。

2 要求条件

映像散策を用いて効果的に情報を提供するには、以下の要件を満たすべきである。

インタラクティブ性 ユーザインタラクションに際しては、提示されるさまざまな情報を手がかりにして所望の情報を引き出せるべきである。例えば、動画中に出現する被写体をクリックすることで検索できるべきである。これを被写体検索と呼ぶ。また、ストレスのない操作性を実現するためには、操作後短い時間で適切な情報を提供することが必要である。

個人適応性 不特定多数のユーザを想定するシステムでは、個々のユーザの要求に合致した情報を提供することが望ましい。そのためには、ユーザの属性(使

用言語・提供される情報への嗜好など)によって、適切な情報を選択して提供するための枠組が必要である。

インクリメンタル性 システムが提供する情報が適宜に更新されることで、ユーザはそのシステムから、常に新しい情報を得ることが出来る。そのためには、新たなデータが容易に追加でき、速やかに利用可能とする機構が必要である。

3 VHMのデータ構造

先の節の要求条件を満たすためのデータ構造として、筆者らが提案しているVHMのデータ構造 [3] を概説する。

VHMの提供するマルチメディアコンテンツは、動画データ・音声データに代表される、ユーザに提示される個別のメディアの生のデータ(メディアデータ)と、それらを組み合わせる際の順序・条件を記述するためのデータ(シナリオ)とから構成される。

メディアデータを提供するには、データフォーマットの異種性を隠蔽でき、編集時にはできる限り複製しないことが望ましい。VHMのカットは、メディアデータへの参照情報と、そのデータのうち提示される区間の位置と長さを記録することで、この問題を解決している。

VHMでは、情報の提供者は複数のカットを必要に応じて組み合わせて提示することでコンテンツの構成法を記述する。これがシナリオである。シナリオは、メディアデータを参照するカットをノードとし、カットが再生される条件を付与したリンクによってそれらを結合した、ノードリンク型のデータ構造を有する(図1)。

シナリオでは、カットのみならず、他のシナリオをもノードとして参照できる。この時、シナリオ間の参照関係は循環しないように、つまり階層的になるように制限される。これによって、シナリオの段階的詳細化と部分的な再利用が可能になり、生産性と保守性とを改善している。

ノードを結合するリンクには条件式が付与される。あるノードの再生が終了した時点で、そのノードから

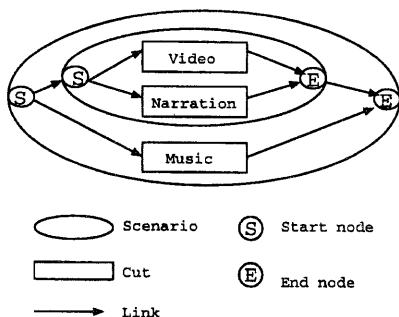


図 1: シナリオの例

出ていく全てのリンク上の条件式が評価され、充足された条件式をもつリンクに後続するノードが全て選択されて再生される。図 2 a では、スタートノード (図中の S) から 3 つのリンクが出ており、このうち Ex.1 と Ex.2 をもったリンクが有効になっている。この時には、同図 b のように、選択された video と narration(japanese) とが並列に再生される。

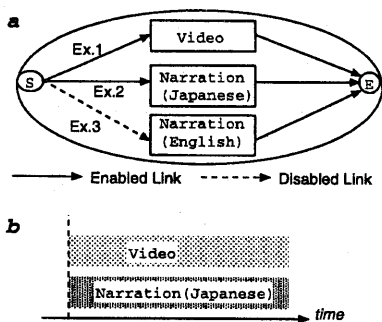


図 2: リンクの解釈

4 VHM システムのアーキテクチャ

VHM システムを実現するためには、4 つの機能が必要である。

1. 動画・音声などのメディアデータやシナリオを格納・管理する機能
2. メディアデータを表示デバイスに出力する機能

3. ユーザからの指示を受けとる機能

4. ユーザの指示に沿ってシナリオを解釈し、必要なメディアデータを決定する機能

第 1 の機能は、複数の端末で情報を共有する必要から、端末からは分離して実装すべきである。これを VHM サーバと呼ぶ。第 2・第 3 の機能は、入出力デバイスの制御を含んでいることとそれらのデバイスの異種性を隠蔽するために、VHM サーバから切り離して端末側の実装すべきである。この両機能を実装する部分をブラウザと呼ぶ。第 4 の機能は、ユーザとのインタラクションを実現するためにセッションを維持する必要がある、サーバとは独立して実装すべきである。しかもこの機能は端末の種別には依存しないため、ブラウザからも独立して実装すべきである。このサブシステムをリゾルバと呼ぶ。これら 3 つのサブシステムの関係を図 3 に示す。

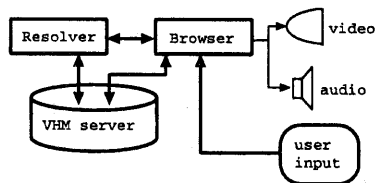


図 3: サーバクライアント型 VHM のアーキテクチャ

以下、VHM サーバ・ブラウザ・リゾルバに求められる機能と構成について述べる。具体的な実装例は次節に挙げる。

4.1 VHM サーバ

VHM サーバは、画像や音声のメディアデータと、それらを組み合わせてマルチメディア化するためのシナリオとを格納する。VHM サーバには格納・検索・更新管理の 3 種の機能が必要である。

格納 シナリオのデータはノードリンク構造を持っており、ナビゲーションを主体とした検索に適した格納構造が必要である。動画と音声のメディアデータは、等時的な (isochronous) 読み出し操作を前提として格納する必要がある。

検索 映像散策を実現するため、基本的にはナビゲーション型の検索を実現する必要がある。充実した映像散策を可能とするために、静止画やボタンに対するユーザの指示だけでなく、動画上の被写体もまた指示の対象とすべきである。すなわち、被写体検索機能が必要である。さらに、ユーザの属性によって、適切な情報を選択するための機能も必要である。

更新管理 VHM サーバは、全てのサブシステムに対してデータを供給するため、シナリオやメディアデータの一貫性を保持する必要がある。例えば、あるクライアントが編集作業をしている間、他のクライアントが同一のシナリオを編集できないように更新を管理する必要がある。

4.2 ブラウザ

ブラウザは、端末上でデータを再生し、ユーザとのインタラクションを行なう部分である。

再生 メディアデータを端末上に出力する再生機能には、等時的なデータを出力する機能が必要である。また、同一のシーンを示す動画データと音声データとを同期して再生する必要がある。

インタラクション ユーザインタラクションを可能にするため、端末からユーザの指示を受けとって、リゾルバの解釈できる形式で渡す機能が必要である。例えば、動画上の被写体検索では、どのフレーム上のどの座標でクリックが発生したかを通知する必要がある。

4.3 リゾルバ

リゾルバは、シナリオの意味とユーザの指示にしたがって、VHM サーバ中のどのメディアデータを取り出すべきかを判断する。シナリオからは一意には定まらない要素については、ユーザインタラクションのような、発生するたびに状況が異なる要素と、ユーザ属性のように一度決めれば明示的に変更されない限り変更されない要素とに分けて取り扱う必要がある。

表 1: 実験システムの構成

CPU	SUN SS-20 (主記憶: 96MB)
OS	SUN-OS 4.1.3
動画フォーマット	Motion JPEG
音声フォーマット	2Ch., 16 bit 直線, 48K sample/sec

5 実験システムの概要

前節で述べたアーキテクチャの概念的な仕様を実装する際の課題と解決法を示すため、我々が試作したVHM 実験システムを紹介する。

このシステムは、表 1 のような構成である。

VHM サーバ VHM サーバのシナリオを格納する部分は、排他制御と検索の高速化を実現するために、オブジェクト指向データベース管理システム (DBMS) 上のアプリケーションプログラム (AP) として実現している。シナリオを格納するためのスキーマの構造を図 4 に示す。ナビゲーションな検索と被写体検索は、共に、DBMS を通じて行なう。

シナリオがメディアデータを利用するために、メディアデータを参照するという手法をとっている。これによって、シナリオをどのように変更しても、メディアデータは影響を受けなくなり、シナリオ間でのメディアデータの共有と、シナリオ変更からの保護が実現している。

シナリオの更新の際には、シナリオ相互の一貫性を保つため、編集システムがオープンしているシナリオは、他の編集システムではオープンできないよう排他制御を実施している。

メディアデータは、再生時の高速アクセスを可能とするため通常のファイルに格納している。メディアデータの更新や削除処理に伴う潜在的なシナリオの不整合を防ぐために、VHM サーバだけがメディアデータの更新と削除を実行できるようにカプセル化した上で、VHM サーバはメディアデータの処理と不可分な操作としてシナリオをも変更している。

等時性が要求される音声データは、メディアデー

タ自体を格納するストリームファイルとサイズやフレーム数を格納する属性ファイルとに分けて格納される。同じく等時性が要求される動画ファイルでは、ストリームファイル、属性ファイルに加え、動画ファイル上の個々のフレームの位置の情報を格納するインデックスファイルの3つのファイルに格納される。

このような格納・管理の手法を用いることで、高速なメディアデータの読み出し、メディアデータの保護と共有、シナリオとメディアデータとの整合性の確保を実現している。

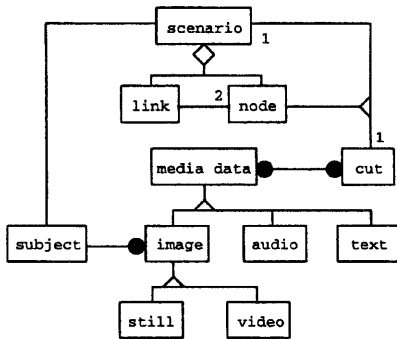


図 4: VHM サーバのスキーマ構造

ブラウザ ブラウザは、後述するリゾルバから、再生すべきメディアデータの指示を受けとって、表示デバイス上に出力する(図7)。動画データと音声データとは等時的な処理が必要なため、平野らが開発したインタラクティブプレイヤーを用いた[1]。

ブラウザのユーザインタフェース機能は、リゾルバからの指示にしたがって、ユーザに対してメッセージを出して入力を促す機能をもつ。そのため、リゾルバから送られてきた変数情報を元に、入力要求画面を作り出す(図6)。

被写体検索のためのインタフェースでは、画面上の物理座標とクリックした時刻から、メディアデータ上の論理座標とフレーム番号へ変換する機能を備える。

リゾルバ リゾルバは、シナリオを展開してどのメディアデータを再生すべきかをブラウザに指示する。

実験システムでは、リンク上の条件式は、実装を簡略化するために <変数名> = <定数> の構文に限定した。

変数には、ユーザの属性などの一度設定したら変更しない値を格納する環境変数と、ユーザインタラクションに用いられ、読み出す毎に値が変更されるインタラクション変数との2つのクラスを設けた。リゾルバは、式中に値の不定な変数があると、その値をユーザに入力させるようにリゾルバに指示する。インタラクション変数はリンクから読み出した時点では常に値が不定になるよう決めてある。このため、インタラクション変数を持つリンクが検索されるたびに必ずユーザへのプロンプトが生じる。環境変数は、一度設定されたらその値を保持し続ける。また、ユーザとのセッションの起動時に値を設定することも可能である。

このような変数と値によるプロトコルを採用したことで、リゾルバ-ブラウザ間のインタフェースを簡略化することができた。

同一のノードから複数のリンクの条件が充足される場合、シナリオの最初のノードに限って、並列に再生を進める。また、シナリオの出口のノード(図中の㊸)では、同シナリオ中で再生していた全てのノードの再生が終了するのを待ち合わせる。シナリオの意味をこのように制限することで、ユーザが応答する限り無限に待ち合わせが生じない解釈が可能となった。

VHM エディタ シナリオを編集し、メディアデータを投入するために、VHM エディタを用意している。これは GUI を用いて直観的にシナリオを編集する機能を持つ(図5)。

実行例 我々は、この実験システム上に、小規模な横浜名所案内のコンテンツを記述した。メディアデータのファイルは40、カット数は20、下位の階層まで含めたシナリオの総数は10である。この例では、メディアデータを準備する時間を除いて、1時間未満で作成できた。

このコンテンツのシナリオは、先の図5に示したものである。リゾルバは、同図中左上の㊸ノードから解釈を始め、リンクの矢印に沿って解釈を進める。中

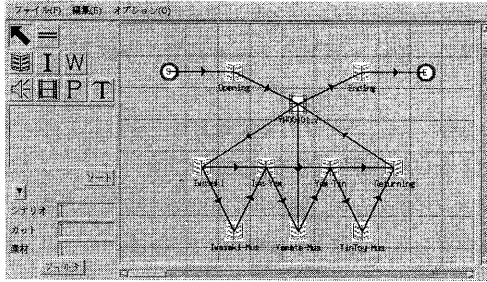


図 5: VHM エディタ画面

中央や上のフィルムのアイコンのノードからは二つのリンクが伸びている。このリンクの条件式にはインタラクション変数が使われており、これによって、図 6 に示すようにユーザへのプロンプトが提示される。ユーザからの入力を受けると、対応するリンクをたどって再生を進め、図 7 のように情報を提示する。このようにして、ユーザはシナリオで記述されるマルチメディアに対して、映像散策をすることができる。

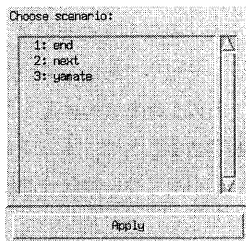


図 6: ブラウザへの入力

6 まとめ

本稿では、映像散策というマルチメディアアプリケーションを用いて、情報を共有化するために、クライアント-サーバ型の VHM のアーキテクチャを提案した。このアーキテクチャは、情報の再利用性とデータフォーマットの隠蔽とを考慮し、VHMサーバ・ブラウザ・リゾルバの 3 つのサブシステムから構成される。

今後の課題としては、実験システムでは実装され

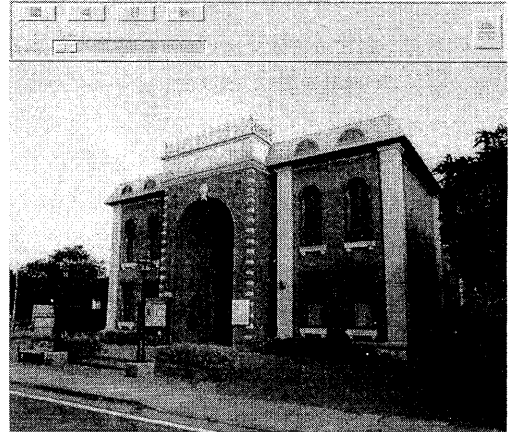


図 7: ブラウザ画面

なかった、画面レイアウト機能と、より詳細なインタラクションを実現するための機能とを実現することが重要と考えられる。

謝辞 本論文の図 7 に掲載したデータは、京浜急行と日本電信電話とが実施中のマルチメディア共同利用実験システムより引用しました。

参考文献

- [1] 平野 泰宏, 岩渕 明, 花籠 靖, 寺中 勝美: “映像散策のためのビデオハイパーメディア — インタラクティブプレイヤーの映像表示方式, 情処ワークショップ論文集 Vol.96, No.1, pp.221-228 (Oct., 1996).
- [2] 坂田 哲夫, 木原 民雄, 小島 明, 佐藤 哲司: “映像散策のためのビデオハイパーモデルの提案”, 情報処理学会研究報告 [データベースシステム], DBS-104-27, pp.209-216 (Jul. 1995).
- [3] Sakata, T., Kojima, A., and Satoh, T.: “Hyper Video: Video Based Hypermedia for Walk-Through Applications”, Proc. Multimedia Japan 96, pp. 148-154 (Mar. 1996).