

## 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ーヤの映像表示方式 —

平野泰宏, 岩渕明, 花籠靖, 寺中勝美

NTT情報通信研究所 データベース研究部

筆者らは、映像を中心としたマルチメディア情報をユーザがインタラクティブな操作で取り出せるビデオハイパーメディア (VHM) 技術を提案してきた。本論文では、VHMの構成要素の一つであるインタラクティブプレーヤの機能と制御方式について述べる。

インタラクティブプレーヤでは、静止画再生ボード一枚を用いて、静止画の再生の他、可変速・反復再生等のトリック再生を含む動画再生、複数の動画の同時・連続再生、動画と音声の同期再生等の様々な再生を行なえる。

インタラクティブプレーヤでは、映像ストリーム毎に作成されたプロセスが静止画再生ボードをタイムシェアリングで使用し、時刻に応じたフレームを表示することで複数動画の同時再生を可能にしている。また、音声を再生するプレーヤと同期情報をやりとりすることで、動画と音声の同期再生を実現している。最後に、観光案内システム TakeMe での使用例を紹介する。

### 1 はじめに

コンピュータの高機能化により、映像、音声などのマルチメディア情報を容易に扱う事が可能になってきている。今や、単に映像を表示したり音声を流すだけでは様々なニーズを満たすことができなくなりつつある。ユーザの関心は、より自由で高度なプレゼンテーションに寄せられるようになってきている。こうした要求に答えるためには、ユーザの望む形態でメディア間の関連を記述し組み立てる技術や、組み立てに沿ってメディア情報を提供する技術が重要である。筆者らは、これらを総合的に実現するビデオハイパーメディア (Video Hypermedia: VHM) 技術を提案してきた [1] [2] [3]。

VHM 技術を使えば、美術館や街の中を順路に沿って複数の映像を撮影しておき、ユーザの操作にしたがってこれらを切れ目なく順次再生する事により、ユーザがインタラクティブにそ

の中を散策することができる。また、映像中の魚を指示して、魚の名前や特徴などの情報を表示させる事も可能になる。

本論文では、ビデオハイパーメディアの構成要素のうち、ユーザの作成したシナリオの中から選択したストーリーに沿ってメディア情報を提示するインタラクティブプレーヤの構成、機能および処理方式と、典型的なアプリケーションである観光案内システム「TakeMe」への適用について述べる。

### 2 ビデオハイパーメディア (VHM) システム

#### 2.1 概要

複数のメディアを柔軟に組合せ、インタラクティブなユーザの操作によってストーリー展開を変化させられるアプリケーションを構築できるように、オーサリングから実際のプレゼンテーションまでを支援する環境がビデオハイパーメディア (以下、VHM と称する) システムである。

Video Hypermedia for Video Walk through - Video Playback Method of Interactive Players -, Yasuhiro Hirano, Akira Iwabuchi, Yasushi Hanakago, Katsumi Teranaka, NTT Information and Communication Systems Laboratories

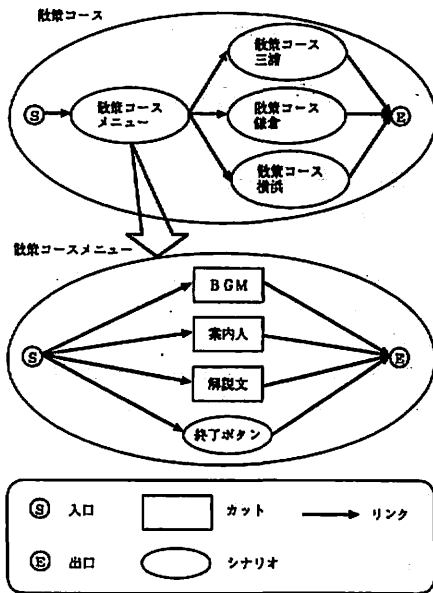


図 1: 構造化シナリオの例

## 2.2 構造化シナリオ

VHM システムでは、ストーリーの展開を有向グラフ構造を基本とした「シナリオ」で表現する。ノードにはプレゼンテーションの対象となる映像・音声・テキストなどのメディア（以後カットと呼ぶ）を配置する。カット間の関連である連続・同期・選択は、ノード間を結ぶリンクで表される。

複雑なストーリーや大きなシステムを1つのグラフ構造で記述すると、リンクが交差して視認性が悪く記述が難しい。また、シナリオの再利用やモジュール化が困難になる。そこで、シナリオは階層的な記述を可能とする。すなわち、ノードにはカットのみならず、それらを組み合わせたシナリオを置くこともできる。

これによって、意味的にまとまりを持った単位で階層的にシナリオを記述することで、可読性、再利用性を向上し、複雑なストーリーを簡易に記述できる。構造化シナリオの例を図1に示す。

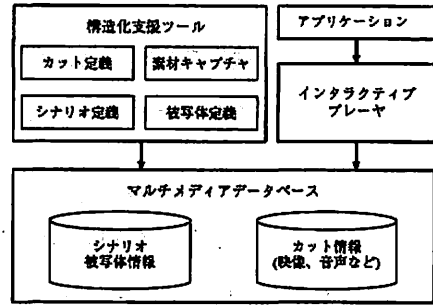


図 2: VHM システムの構成

## 2.3 被写体定義

構造化シナリオでは、ノード間のリンクの部分で静的にストーリー展開の変化を記述する。一方、ユーザが興味を持った映像中の物体に触れることでその物体の情報を表示したりストーリーを変化させることができると、より動的な展開が可能になる。この機能を実現するのが被写体検索である。

VHM システムでは、映像中の被写体が存在する時刻・位置情報を登録しておき、実行時にユーザがポイントした時刻・位置情報を用いて登録しておいた被写体の情報を検索し、得られた情報によって次のストーリー展開を決定する。

映像中の被写体の情報を登録することは一般には多くの作業量を必要とするが、VHM システムでは、動画中の被写体をユーザがポイントする際の精度はさほど高くないことに着目し、被写体の形状は矩形領域（またはその組合せ）で、動きは折れ線で近似し、数点のデータを入力すれば線形補間によって半自動的に登録できるようにした。

## 2.4 VHM システムの構成

VHM システムは、構造化支援ツール、インタラクティブプレーヤ、マルチメディアデータベースとで構成される(図2)。

構造化支援ツールは、映像、音声などの素材のキャプチャから被写体情報の設定、構造化シナリオの記述までを行なうオーサリングツール

である。

インタラクティブプレーヤは、ユーザの操作に応じたストーリー展開で、映像や音声を再生する。また、被写体検索の際にユーザがポイントした時刻や位置情報の取得も行なう。インタラクティブプレーヤの詳細は次章以降で述べる。

マルチメディアデータベースは、カットの情報、構造化シナリオおよび被写体情報を格納する。

### 3 インタラクティブプレーヤの機能

インタラクティブプレーヤは、静止画・動画および音声などのメディアを簡易にかつ自由に操作する機能を提供する。任意速度再生や正逆方向再生のような個々のメディアの操作の他に、構造化シナリオのメディア間の関連のうちの連続・同期のような複数のメディア間の制御も行なう。インタラクティブプレーヤの機能を表1に示す。

表1: インタラクティブプレーヤの機能

映像	正逆方向再生 複数映像再生 縮小再生 任意速度再生 静止・コマ送り スクロール・ズーム 論理座標指定
音声	音量調整 バランス調整 出力先指定 ステレオ・モノラル指定 BGM 再生
共通	任意区間指定再生 映像・音声同期再生 反復再生 往復再生 連続再生 論理時刻指定 イベントハンドラ

#### 3.1 基本機能

インタラクティブプレーヤでは、再生する素材と各種オプション（映像の位置やサイズ、音声の音量や出力先等）を設定し、スタート、ストップ等を指示するだけで、各種メディアを制御できる。再生方向、再生速度の変更や静止・コマ送りのような操作も自由に行なえる。

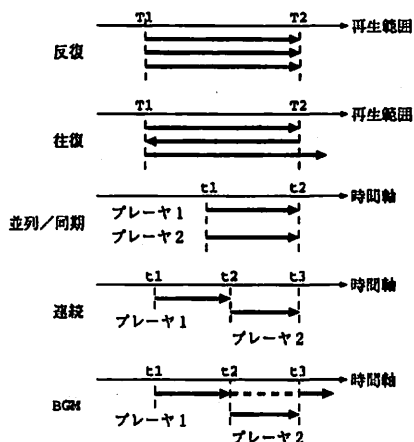


図3: 特殊再生

#### 3.2 特殊再生

トリックプレイを滑らかに行なえるように、また、アプリケーションによる制御を最小限にとどめられるように、図3のような特殊再生機能を持つ。

反復再生は、指定した範囲（図3のT1からT2）の再生を繰り返すものである。往復再生は、再生範囲の最後（図3のT2）まで来ると逆方向に再生し、再生範囲の最初（T1）まで来ると順方向に再生するという動作を繰り返すものである。往復再生を抜ける方向（図3では順方向）も指示できる。

連続再生は、一つのプレーヤの再生が終了すると、指定したプレーヤの再生をスタートさせるものである。複数のメディアを滑らかに接続できる。

BGM再生は、音声再生中に別の音声の再生を指示すると、前の音声プレーヤの情報をスタックに積み、再生が終了するとスタックから取り出して再生を再開するものである。普段はBGMを再生し、ナレーションが入ると自動的にBGMを中断し、ナレーションの終了とともに自動的にBGMが再開するようなことを簡単に実現できる。

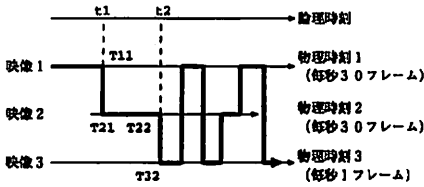


図 4: 論理時刻

### 3.3 論理時刻

図4は、サッカーの試合を複数のカメラで撮影した映像のように、互いに時間的関連のある複数の映像の例である。この例では、映像2の開始時刻は他の映像と異なっている。また、映像3のフレームレートは他の映像と異なっている。これらの映像からユーザーが見たいものを随時選択して表示するアプリケーションを考える。

例えば、映像1の時刻 T11 を再生中に映像2への切替を指示すると、映像1と映像2の開始時刻の差を考慮して映像2の時刻 T21 (≠ T11) から再生する必要がある。また、映像2の時刻 T22 を再生中に映像3への切替を指示すると、映像2と映像3のフレームレートの違いも考慮して映像3の時刻 T32 から再生する必要がある。このように、時間的関連を有する複数のメディアから選択して再生する際には、各メディア間の時間軸のずれを補正する必要がある。

そこで、インタラクティブプレーヤでは、ユーザーから見た時刻(論理時刻)とメディアから見た時刻(物理時刻)をマッピングする機能を設けた。各メディアが論理時刻上のどこに位置するかを登録しておき、再生位置の指定は論理時刻で行なう。再生中にメディアを切替える場合、切替たいメディアを指定するだけで対応する時刻から再生される。

また、どの時刻にどの映像を表示するかを登録することにより、自動的に映像を切替えながら表示するようにした。例えば図4では、論理時刻  $t_1$  までは映像1、論理時刻  $t_1$  から  $t_2$  までは映像2、論理時刻  $t_2$  からは映像3・・・のよう

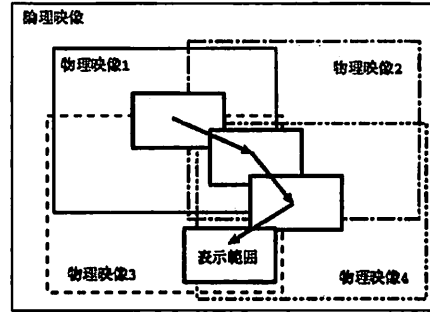


図 5: 論理座標

に指定すると、自動的に太線のように映像を切替えながら再生する。

### 3.4 論理座標

図5は水族館の水槽を複数のカメラで同時に撮影した映像のように[4]、互いに位置的に関連のある映像の例である。これらの映像からユーザーが見たい範囲を随時指定して表示するアプリケーションを考える。

ここでは、全ての映像でカバーされる範囲(この例では水槽全体)を仮想的な映像空間とし、その仮想的な空間の座標(論理座標)を設定し、論理座標で指示された表示範囲から、表示する映像を選択し、表示する範囲を映像内の座標(物理座標)で表現する必要がある。

そこで、インタラクティブプレーヤでは、各映像の論理座標内での位置情報を登録し、論理座標で表示範囲を指示すると適切な映像を選択して表示するようにした。

また、論理座標上の表示範囲の時間変化の情報にしたがって自動的に表示する映像と表示範囲を変えながら表示するようにした。例えば、特定の魚を中央に含むような表示範囲のデータを与えると、その魚の動きを追跡できる。

ここで、表示範囲の時間変化の情報は、特に再生時間が長くなると、膨大なものになることが予想される。そこで、表示範囲の情報をデータベースに格納しておいて少しずつ読み込みながら再生するなどの自由度を得られるように、

表示範囲情報の登録用のハンドラを設定できるようにした。ハンドラは周期的に呼び出したり、登録済の情報が残り少なくなったら呼び出したりできる。

### 3.5 イベントハンドラ

ビデオハイパーメディアでは、例えば美術館の中を歩く映像を再生中に彫刻の前を通った時にその解説を出す、英語の音声の流れた時には日本語の字幕を出すなど、特定の位置を再生した時に何かの処理をさせたい場合がある。そこで、メディアを再生中に指定した時刻に達した時点で、アプリケーションに通知したり、指定した関数を起動したりできるようなイベントハンドラが必要になる。

ここで、メディアをインタラクティブに操作する場合、指定した時刻に達したときの再生方向や、通常に再生してその時刻に達したのか途中を飛び越してその時刻を越えたのかによって処理が異なる場合がある。例えば、彫刻の解説の場合は、途中を飛び越して彫刻の前に来た場合でも解説を表示したいし、逆方向再生で彫刻の前から離れた場合には解説を消したい。そこでインタラクティブプレーヤでは、イベントハンドラにどのようにしてその時刻に達したか、どの方向に再生したか等を知らせる機能を持たせた。

## 4 インタラクティブプレーヤの実装

### 4.1 実装環境

インタラクティブプレーヤは、以下の環境で実装した。

ハードウェア： SUN SPARC Station  
Parallax XVideo  
or PowerVideo  
ソフトウェア： SunOS 4.1.3 or 4.1.4  
X11R5

上記環境において実現される主要諸元を、表2に示す。

表 2: インタラクティブプレーヤの主要諸元

映像	圧縮方式：	JPEG または JFIF
	映像サイズ：	最大横幅 1024 pixels
	再生レート：	640 × 480 pixels で 最大 30fps 程度
音声	符合化方式：	$\mu$ law または PCM
	チャンネル数：	2
	ビット幅：	8 または 16
	再生レート：	最大 48kHz

### 4.2 音声再生方式

今回実装した環境においては、音声デバイスへ書き込むことによって音声を指定した再生レートで再生できるが、インタラクティブな操作に対応できるようにしておく必要がある。

そこで、インタラクティブプレーヤでは、音声再生用のプロセスを作成し、読み込んだ音声ファイルを非ブロックモードで少しずつデバイスに書き込むようにした。

### 4.3 動画再生方式

#### 4.3.1 タイミング合わせ

今回の実装で使用した Parallax Graphics の XVideo および PowerVideo は静止画の圧縮伸長ボードであり 1 フレームごとのデコード処理時間にばらつきがあるため、動画の再生レートが一定とならないという問題がある。

そこでインタラクティブプレーヤでは、映像毎に再生用のプロセスを作成し、基本的には各映像プロセスがタイマーを用いて、処理が早過ぎる場合にはタイマー割り込みまで休止し、処理が遅れた場合にはその時刻のフレームまで飛ばして次の処理を行なうようにした。タイミング合わせの様子を図 6 に示す。

図 6 では、4 フレーム目の処理が遅れたため、次の 5 フレーム目の処理を省略して、6 フレーム目を処理する。このように、いつどのフレームを再生するかをチェックすることで、指定したレートで動画を再生できる。

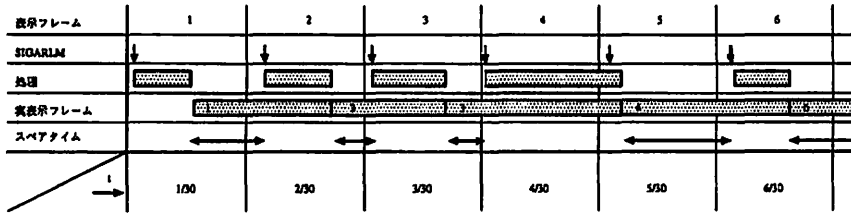


図 6: 動画再生方式

#### 4.3.2 複数動画の同時再生

Parallax Graphics のビデオボードは、同時に複数の映像を取り扱うことはできない。そのため、複数の動画を同時に再生するには、ビデオボードを切替えて使用する必要がある。

そこでインタラクティブプレーヤでは、映像毎にプレーヤプロセスを作成するようにした。4.3.1で述べたように各映像プレーヤは時刻を確認して再生するフレームを決定するため、OS のスケジューリングによりプロセッサを割り当てられた際に適切なフレームを表示するため、ビデオボードをタイムシェアリングして動作する。

#### 4.4 動画と音声の同期再生

音声を伴った映像を再生する場合、映像と音声の同期が必要になる。

音声に関しては、デバイスへ書き込むことによって指定した再生レートで再生可能であり、映像に関しては時刻を確認して適切なフレームを表示するが、両者の再生開始時刻を一致させる必要がある。例えば、単に音声プロセスと映像プロセスに同時に再生開始を指示する方法では、特にシステムの負荷が高い場合に、両者の再生開始時刻がずれることがある。

そこでインタラクティブプレーヤでは、再生開始時に音声プロセスが再生開始時刻を画像プロセスに送り、映像プロセスが時刻を音声に合わせるようにした。

ここで、音声の再生開始時刻に正確に合わせ

ようとすると、映像ファイルの読み込みから表示までの所要時間が問題になる。さらに、この所要時間は映像の大きさ、ファイルの読み込み速度、システムの負荷に依存して変動する。

そこで、映像プロセスが1フレームの表示に要する時間を測定し、それらの移動平均を計算して、表示すべきフレームを決定する際に使用するようにした。

#### 4.5 構成

インタラクティブプレーヤの構成を図7に示す。既に述べたように、インタラクティブプレーヤは、映像/音声といった各メディアごとにプレーヤプロセスを生成する。そうして、それらのメディア間の同期再生、連続再生のような制御を容易に、かつアプリケーションが制御を意識しなくてもよいように、アプリケーションとは独立の制御プロセスを生成する。

アプリケーションはマルチメディアライブラリ (MM ライブラリ) を用いて制御プロセスと通信する。制御プロセスは、映像/音声プレーヤの生成、再生、消滅等の制御を行なう。

MM ライブラリはC言語または Tcl/Tk から使用可能で、簡単なプログラム記述でアプリケーションを構築することができる。

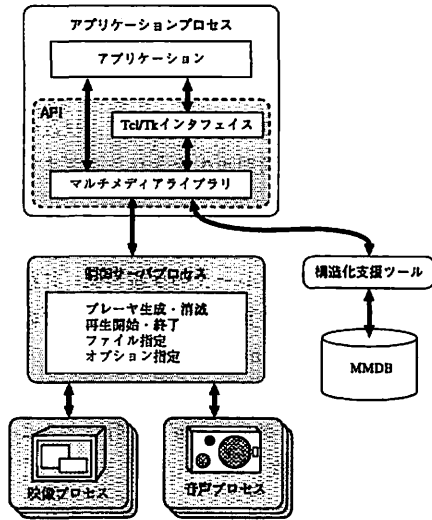


図 7: インタラクティブプレーヤの構成

## 5 観光案内システム「TakeMe」への応用

観光案内システム「TakeMe」[5]をVHMシステムを用いて構築した。「TakeMe」は、NTTが行っている「マルチメディア共同利用実験」の一環として京浜急行電鉄駅で運用されているシステムであり、複数の駅のホームに設置されている。沿線の観光地域の多岐にわたる情報を蓄積している。

ここでは、「TakeMe」の約300のシーンから、インタラクティブプレーヤの特徴的な機能を使用しているものをいくつか紹介する。

図8は、ユーザが指示したように観光地を歩き回る散策コースの交差点のシーンである。ユーザから進む方向を指示されるまで交差点で停止している際に、1フレームを固定して表示すると映像が固まった感じになり違和感が生じる。そこで、立ち止まった状態で撮影された数フレームを往復再生することによって、例えば木の葉が風に揺れている状態が違和感なく再現される。

ユーザから進む方向を指示されると、往復再

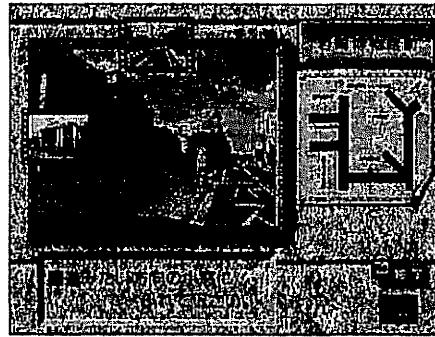


図 8: 交差点のシーン

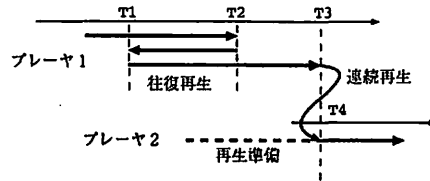


図 9: 交差点のシーンでのプレーヤの動作

生から抜け出て、次のシーンへと移る。ここで、映像ファイルは分岐点ごとに分割されているため、交差点に差しかかる前の映像と後の映像を滑らかに繋げて再生しなければならない。そこで、ユーザから進む方向を指示されてから交差点に差しかかる前の映像が終了するまでの間に、指示された方向に進む映像を設定したプレーヤを準備し、連続再生機能を用いて無隣断で再生する。このシーンでのプレーヤの動作を図9に示す。

図10はスロットマシンを用いてユーザが行きたい場所を直接指定するシーンである。ここでは、3つのスロットがバラバラに回転している感じを出すために、複数映像の同時再生機能を用いて、3つの動画をそれぞれ異なる速度で再生している。また、3つのスロットに反復再生を指定することによって、一回りした時に継目のところが滑らかにつながって再生される。

「TakeMe」では、例えば交差点や道路、ス

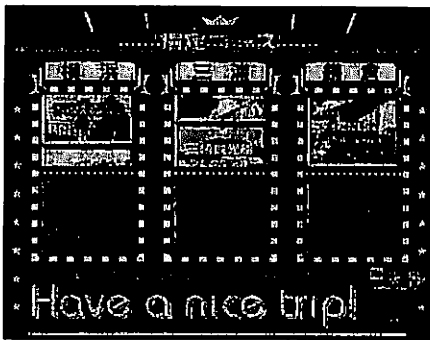


図 10: スロットマシン

ロットマシンのシーンではBGMを再生し、名所のシーンでは解説の音声再生し、道路やスロットマシンに戻ってくるとBGMを再開する。インタラクティブプレーヤーのBGM機能を利用すると、最初にBGMの再生を開始し、名所のシーンでは解説音声付の動画の再生を開始すれば、BGMはスタックに積まれて解説の音声再生され、道路やスロットマシンに戻って来た時には音声無しの動画の再生を開始すれば、自動的にスタックからBGMを取り出されて再生される。インタラクティブプレーヤーでは、音声付の動画と音声無しの動画を区別せずに扱えるので、アプリケーションは再生したいメインの動画を順次指示するだけでよく、動画が音声付かどうかやBGMのOn/Offを気にしなくてもよい。

## 6 おわりに

映像を中心としたマルチメディア情報をユーザがインタラクティブな操作で再生できるビデオハイパーメディア(VHM)システムの構成要素の一つであるインタラクティブプレーヤーの機能と制御方式について述べた。

インタラクティブプレーヤーは、静止画・動画および音声などのメディアの任意速度再生や正逆方向再生のような個々のメディアの操作を簡単に行なえる。

さらに、映像や音声毎に再生用のプロセスを作成し、各映像プロセスがタイマーを用いて再生すべきフレームを決定し、制御プロセスを介して同期情報をやりとりすることで、複数の動画や音声の同時・同期再生を含むさまざまな再生を行なえるようにした。

最後に、観光案内システムTakeMeにおいて、インタラクティブプレーヤーの反復再生、往復再生、連続再生、同時再生、BGM機能がどのように生かされているかを示した。

今後は、画面レイアウトも含めたプレゼンテーション構築全体をグラフィカルに行なえるように拡張を進めていく。

## 参考文献

- [1] 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”, 信学技報(データ工学), DE95-35, pp. 65-72, July 1995.
- [2] 花籠靖, 小島明, 佐藤哲司, “ビデオハイパーメディアによるビジュアル電子図書館の構築”, 情処研報(データベースシステム), 96-DBS-107, pp. 17-24, Mar 1996.
- [3] 小島明, 花籠靖, 坂田哲夫, 佐藤哲司, “実写ウォークスルー型電子図書館の構築”, 信学技報(オフィスシステム), OFS95-50, pp. 31-36, Mar 1996.
- [4] 片岡良治, 紺谷精一, 三浦史光, 佐藤哲司, 井上潮, “実写映像を用いたマルチメディア情報システム「ビデオリアリティ」の構築”, 信学技報, DE94-78, pp. 1-8, 1995.
- [5] 岸田義勝, 木原民雄, 平野泰宏, 岩淵明, 寺中勝英, “ビデオハイパーメディアの観光案内システム「TakeMe」への応用”, 情処研報(マルチメディア通信と分散処理), 96-DPS-76, pp. 55-60, May 1996.