

映像の重要度と地図を利用した多視点映像のブラウジング方式

秦 淑彦[†], 廣瀬 竜男^{††}, 田中 克己^{††}

プラントや道路の監視, スポーツやフェスティバルの TV 中継などでは, 広い地域に多数のカメラが設置され, 必要に応じてオペレータに切り替えて表示される。カメラ台数が数十から数百といった多視点映像の場合, 見たい映像を探すのに時間がかかる, どこを撮影しているのか分からない, 一つのカメラで撮影できない広い空間で発生する事象をマクロな視点で捉えられないなどの問題がある。これらを解決するために, 映像の重要度と地図を利用した多視点映像のブラウジング方式を提案する。まず, ユーザの注視点, アラームや通報等のイベント, 建物や道路等のオブジェクトに基づき各映像の重要度を計算して, その重要度に応じて表示すべき映像を選択し, 強調表示や画質の割り当てを行う。さらに, 地図を背景として複数の映像を撮影範囲の近傍に表示する。本稿では, 多視点映像のブラウジングにおける課題を説明した後, 提案方式について述べる。また提案方式の有効性を評価するためのプロトタイプについても述べる。

A Browsing Method for Multiple Perspective Video Using their Importance and a Map

Toshihiko Hata[†], Tatsuo Hirose^{††}, Katsumi Tanaka^{††}

Several pictures from a lot of video cameras in a broad area are selected and displayed for monitoring plant facilities and outside broadcasting. If more than several tens of cameras are set up, there are some problems to be solved. For example, it takes much time to find a necessary video scene and it is difficult to grasp a camera field of the image. It is also difficult to understand the entire state of a phenomenon that occurs in a broad area and is taken pictures by plural cameras. We propose a new browsing method using the importance of pictures and a map for those problems. The importance of pictures is calculated from some elements that are objects such as facilities, events such as an alarm, and a user's observation point. Based on the importance, the pictures to be displayed are selected and their display ways such as changing enhancement and image quality are decided. Finally those pictures are displayed on the map. We describe this browsing method after we explain the above problems in this paper. We also describe our prototype to evaluate the method.

1. はじめに

プラントや道路の監視, スポーツやフェスティバルの TV 中継などでは, 広い地域に多数のカメラが設置され, 必要に応じてオペレータに切り替えて表示される。カメラの低価格化に伴い設置されるカメラ台数も増加しており, 現状の数十から将来は数百というケースも考えられる。一方, 映像をデジタル圧縮してハードディスク等に連続的に記録し, ネットワーク経由で配信するデジタル映像システムの導入も始まっている。このようなシステムでは, 見たい映像に素早く

アクセスできる, 記録を止めずに繰り返し見直せる, 被写体に起こったイベントをトリガに映像を表示するなど, 容易な検索や視覚的效果のある表示等の機能が提供されている[1]。しかし, カメラ台数が数十から数百といった多視点映像の場合, 見たい映像を探すのに時間がかかる, どこを撮影しているのか分からない, 一つのカメラで撮影できない広い空間で発生する事象をマクロな視点で捉えられないなどの問題がある。

本研究では, これらの問題を解決するために, 映像の重要度と地図を利用した多視点映像のブラウジング方式を提案する。まず, ユーザの注視点, アラームや通報等のイベント, 建物や道路等のオブジェクトに基づき各映像の重要度を計算して, その重要度に応じて表示すべき映像を選択し, 強調表示や画質の割り当てを行う。さらに, 地図を背景として複数の映像を撮影範囲の近傍に表示する。

以下, 2 章で多視点映像とそのブラウジングにおける課題および関連研究, 3 章で提案方式, 4 章で方式

[†] 三菱電機 (株) 産業システム研究所
Industrial Electronics & Systems Laboratory,
Mitsubishi Electric Corporation.

^{††} 神戸大学大学院自然科学研究科情報メディア科学専攻
Division of Information and Media Sciences, Graduate
School of Science and Technology, Kobe University.

評価を目的としたプロトタイプについて、各々述べる。

2. 多視点映像のブラウジング

2.1 多視点映像

本稿で扱う多視点映像とは、ビルや競技場等の撮影対象となる空間を数十から数百という多数のカメラで撮影したライブ映像および蓄積映像である。多視点映像は以下に示すメタデータを有する。これらメタデータは何時、何処で、何を、何故撮影したかを記述するものであり、映像検索に有効に利用できる。また、これらメタデータは各種センサ装置等より生成されるので、映像データ入力と同時に自動的に映像と関連付けて記録・配信することができる。

- 撮影時刻：撮影時刻を示すタイムスタンプで、何時という属性に対応する。
- カメラと撮影範囲：撮影したカメラの識別子とカメラの位置や方向等の撮影範囲データで、何処でという属性に対応する。地図データ等と組み合わせれば、何をという情報も得られる。撮影範囲は外部制御可能な首振りカメラや GPS (Global Positioning System)、ジャイロ等のセンサから得られる。
- イベント：イベントとは人や設備などのオブジェクトに発生した事象であり、例えば、夜間出入り口から人が侵入した、交差点で2台の車が衝突した、3番バッターがホームランを打ったなどである。イベントデータはイベント ID、イベントタイプ、発生時刻等から成り、何故という属性に対応する。各種センサ装置から発生されるだけでなく、電話通報や実況中継の音声認識して自動的にメタデータとして記録することもできる。
- 解説や注釈等：被写体の注釈や事象の分析結果など、ユーザが加えたデータである。テキスト入力や音声認識により、メタデータとして記録される。

さらに、多視点映像では、全空間を隙間なく撮影しているわけではなく、一部の空間を覗き見るように疎に撮影しているという特徴がある。また、蓄積映像については、例えば記録時間が数日に及ぶ長期間の場合、全てを残しているわけではなく、最新の映像やイベント発生前後の映像のみ記録するといった、時間的に疎な記録をしている。(図1)

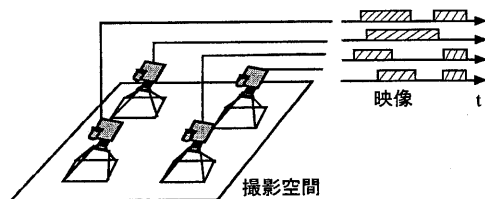


図1 時空間的に疎な性質を有する多視点映像

2.2 ブラウジングにおける課題

多視点映像から必要な映像クリップを探し出し観察するための映像ブラウジングでは、以下の課題がある。

① 多数のカメラへの対応

複数カメラの映像を効率的に観察するために一つのモニタに複数の映像を同時に表示する方法がよく用いられる。しかし、興味あるものをざっと探すような目的でも同時に数個、注視する時はせいぜい2個の映像を見るのが限界である。従って、数十から数百ものカメラがある多視点映像では、何らかの規則に従い重要な映像を自動選択して表示する必要がある。

② 撮影範囲の直感的把握

実空間を撮影した映像の場合、映像だけ見せられても何処を撮影したのか分からないという撮影範囲の直感的把握の問題がある。カメラ台数が少ない場合も問題であるが、カメラ台数が多いとより深刻になる。

③ 広域で発生する事象の把握

カメラの死角は多視点映像によらず一般的な問題であるが、疎に撮影した多視点映像を用いて、広域で発生した事象をマクロな視点でうまく捉えられるかという固有の課題がある。例えば、不審者を追跡したり、プラント事故での事象伝播を観察する場合に、空間の一部を覗き見している複数の映像をどのように提示すれば、観察者は全体の様子を頭の中にうまくイメージできるかという課題である。

④ 長時間にわたる事象の要約

短時間に概要を把握するための映像の要約生成は、多視点映像に限らず重要な課題である。しかし、多視点映像の場合、撮影時刻、写っているオブジェクトの位置や意味的關係などにより、複数の映像が相互に関連しているため、複数ストーリーを対象とした要約生成が必要である。また、要約を提示する際、時間的・空間的な把握ができるよう、撮影時刻や位置などの時空間情報が要約に含まれていなければならない。

2.3 関連研究

映像データの検索については多くの取り組みがなされている[2]。しかし、その大半は映画やニュース番組等を対象としており、本稿で取上げる多視点映像に対する研究はほとんどない。

撮影範囲が既知である複数のカメラで撮影した映像を蓄積するとともに、背景オブジェクトと映像より抽出した歩行者等のオブジェクトから成る3次元モデルを使って、指定した歩行者が写っている映像を検索するシステムが提案されている[3]。オブジェクトの位置と時刻を指定した問い合わせや、3次元モデルに映像をマッピングして表示するなど興味ある方式を提案しているが、対象は狭い空間を複数の異なる視点から撮影した映像であり、本稿で扱う多視点映像とは異なる。

移動カメラで撮影した映像に対し、GPS やジャイ

口等の位置センサから得られた撮影範囲情報を利用し、映像のフレームやカット単位に地図上にその撮影範囲を示し、映像がどこを撮影しているのか直感的に把握させるブラウジング方式が提案されている[4], [5]. これらは実空間を異なる視点から撮影した映像のブラウジングとして、撮影範囲情報や地図を用いるという点で本提案方式と同様のアプローチをとっているが、多視点映像の効果的なブラウジング方式に焦点をあてたものではない。

3. 提案方式

3.1 課題解決のアプローチ

提案するブラウジング方式は 2.1 で述べた①～③の課題解決を狙っており、以下のアプローチをとる。

① 多数のカメラへの対応

各映像の重要度を求め、重要度の高い映像を優先的に表示する。映像の重要度を決める要素としては、映像に写る人や設備といった“オブジェクト”，オブジェクトに発生する事象である“イベント”，ユーザの注目する地点やオブジェクトの位置を示す“注視点”を考える。これら要素の重要度は観察する対象の状態やユーザの意図により動的に変更され、それに対応して映像の重要度が変化し、目的に合った映像が自動的に選択・表示される。例えば、プラントの設備監視であれば、駆動系や冷却系など系統の種類に応じて重要度を変更すれば、系統の種類毎にそれを写す映像が優先的に表示される。

② 撮影範囲の直感的把握

対象空間の地図やレイアウト図を背景として表示し、その上に映像を撮影範囲近傍に表示する。撮影範囲をより明確にするために、地図上で撮影領域をハッチング図形として描画する等を行う。当然、地図のスクロールや縮尺変更などの地図操作に連動し、表示される映像も更新される。

③ 広域で発生する事象の把握

②と同様に地図上に複数映像を表示し、撮影時刻に基づき同期再生する。広域事象に対応したイベントデータが存在する場合は、そのイベントの重要度を高くして、関連する映像を優先的に表示する。これにより、複数のカメラで捉えた物体の動く様子や、煙の広がっていく様子などを位置と時間を意識しながら観察することができる。また、移動体等の観察では、移動体が通過した後も、移動体が写る映像フレームを消さずにそのまま表示しておくなどして、地図上で軌跡がわかるようにする(図2)。

処理手順は、要素の重要度から映像の重要度を計算するステップと、映像の重要度に基づき表示すべき映像の選択や強調表示の度合い等を決めるステップからなる。以下に各ステップにおける処理内容を説明する。

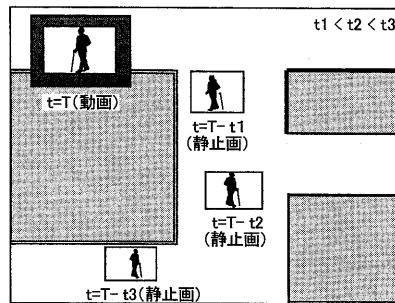


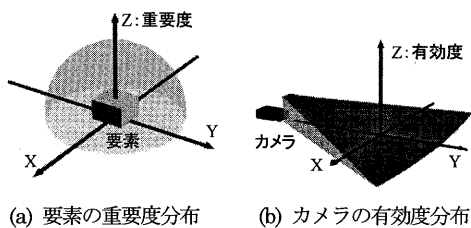
図2 移動体の軌跡が把握できる表示

3.2 重要度の計算

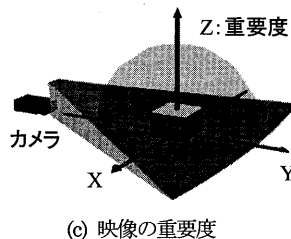
重要度の計算は、要素が存在する空間での計算と時間軸上での計算がある。

(1) 空間での計算

各映像に対し、重要度の高いオブジェクトやイベント、またはユーザが指定する注視点位置が写っているか否か、さらにどの程度写っているかにより映像の重要度を計算する。例えば図3(a)に示すように、要素は存在する位置を原点とし、2次元平面上に分布する重要度の関数を有する。ここで、XY平面は撮影空間の地図であり、Z軸はある位置における要素の重要度の値である。



(a) 要素の重要度分布 (b) カメラの有効度分布



(c) 映像の重要度

図3 空間での重要度の計算

分布は要素自身や観察の仕方により適切なものを選択する。例えば、オブジェクトのサイズが大きい、オブジェクトの周囲も観察したい場合には原点を中心に広がった分布になるし、オブジェクトを見る方向で重要

性が変わる場合は非対称の分布となる。またカメラも図 3(b)に示すように、ある位置に存在する被写体をどの程度写しているかを示す有効度の分布を有する。有効度分布はカメラの撮影方向や焦点距離等の撮影範囲により決まる。これら要素の重要度分布とカメラの有効度分布の積を一つの要素に対する映像の重要度と定義する(図 3(c))。そして全ての要素に対する重要度の総和を映像の重要度とする。これを式で表現すれば、要素 j の重要度 w_j 、カメラ i の有効度 e_i とした時、カメラ i で撮影した映像の重要度 W_i は、

$$W_i = e_i \times \sum_j w_j \quad (1)$$

となる。

(2) 時間軸上での計算

イベントは空間での重要度分布の他に、発生時刻を原点とし、時間軸上に広がりを持つ重要度分布を有する。通常はイベント発生時刻の前後の映像が重要となるので、例えば図 4(a)に示す正規分布等の関数可以利用できる。また、センサ装置の感度が鈍い場合にはイベント発生時刻より前にピークを持つ関数が適している(図 4(b))、イベント発生後の状況が重要である場合には図 4(c)の関数が適しているなど、イベントの性質や観察の目的に応じて使い分けられることができる。これを式で表現すれば、時刻 t におけるイベント k の重要度 $w_k(t)$ 、カメラ i の有効度 $e_i(t)$ とした時、カメラ i で撮影した時刻 t の映像の重要度 $W_i(t)$ は、

$$W_i(t) = e_i(t) \times \sum_k w_k(t) \quad (2)$$

となる。ここで、カメラの有効度を時間の関数としているのは首振りカメラや移動カメラへの対応である。

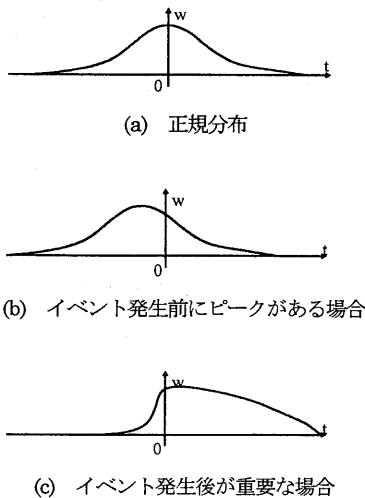


図 4 イベントの重要度関数

このイベントの重要度計算は、映像再生時に周期的

に行う。これにより、撮影時刻の経過に伴い映像の表示が自動的に更新される。

3.3 表示方法

重要度に基づいた表示方法について、以下に述べる。

(1) 代表選択

重要度の高い映像を選択して表示する(図 9)。同時に表示する個数は、背景の地図を覆い隠さない等の画面レイアウトや表示端末等の処理能力により上限の個数が決まり、注視か概要把握かといった観察目的に応じて上限の個数内で変更できるようにする。また、閾値以上の重要度を有する映像を選択したり、重要度の高い N 個の映像を選択するなど、目的に応じて使い分ける。

(2) 強調表示

重要度に従い、表示サイズや映像ウィンドウの枠の色などを変更し、重要なものほど目立つように表示する(図 9)。代表選択と同様、観察目的に応じてパラメータを変更可能とする。また、強調の度合いは重要度の絶対値または順番から決定する。

(3) 映像品質

重要度の高い映像ほど画像の解像度を高くしたり、フレームレートを上げるなど、重要度に従い映像品質を変更する。これは、重要なものほどより詳細かつ鮮明に表示して注意を引くという視覚的效果と、表示端末やネットワーク等における処理能力の有効利用を狙いとする。与えられた処理リソースと重要度から各映像の品質を決める QoS (Quality of Service) 割り当てのアルゴリズムが必要であり、今後の課題である。

(4) 軌跡表示

移動体や煙の伝播など、各場所の状況を確認するだけでなく、事象が空間上でどのように変化したかを併せて見たい場合がある。これには、図 2 のように、変化の軌跡を地図上で明示する。図 2 であれば、イベントの重要度が閾値以下になったときに、その映像の表示を中止し、代わりにピークの重要度を有する映像フレームを静止画として表示することにより実現できる。

(5) 地図連動

スクロールや縮尺変更等の地図操作により、表示する地図の領域が変化する。これに合わせて、新たな地図表示領域を撮影する映像を選択し重畳表示する。映像の重要度は地図の表示領域に依存しないが、代表選択等の表示方法を決定するには、地図表示領域を撮影した映像だけを対象とする方法と、地図表示領域とは無関係に全ての映像を対象する方法がある。

4. プロトタイプ

提案方式の評価を目的としたプロトタイプを開発中であり、完成した機能を使って実験を開始した段階である。構成、実装機能、実験について説明する。

4.1 システム概要

(1) システム構成と実験手順

図5にシステム構成を示す。複数のビデオカメラで撮影した映像を一度ビデオテープに記録した後、パソコンのキャプチャボードで取り込み圧縮してハードディスクに格納する。映像のメタデータは全てテキストエディタを用いて人手で作成し、テキストファイル形式で格納する。表示プログラムを起動すると、地図データ、メタデータ、映像データを読み出し、画面に表示する。

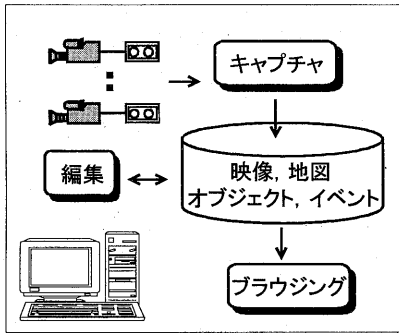


図5 プロトタイプ構成

(2) 実装機能

映像の重要度計算は、テキストファイルに定義されたオブジェクトとイベントの重要度、マウスで指定した注視点に対して行う。オブジェクトおよび注視点に対する重要度計算は、各カメラに対して要素が撮影範囲に存在するか否かを判定し、存在するものに対して以下の計算式で重要度を求める。(図6)

$$W_i = \sum_j (w_j * F_{ij}) / (l_{ij} * S_i) \quad (3)$$

ここで、 w_j は要素の重要度、 S_i 、 F_{ij} はカメラ i の撮像素子の大きさと焦点距離、 l_{ij} はカメラ i から要素までの距離である。この式は、その要素が撮像面上で占める面積の割合を近似的に表す。イベントに対しては空間での重要度計算は行わず、対応するカメラに対してのみ時間軸上の重要度計算を行う。重要度分布は正規分布を用いている。

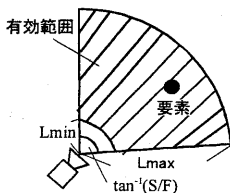


図6 重要度計算における範囲チェック

各要素を個別に、または組み合わせて計算するよう設定できる。また、イベント重要度の時間変化に対応するよう、再生中は周期的に重要度を再計算して画面を更新する。さらに、マウスで指定される注視点の移動に対応するよう、マウスイベントの発生毎に重要度を再計算して画面を更新する。

上記重要度に基づき、重要度の高い N 個の映像を選択し、映像ウィンドウの大きさや色を変化させて地図に重畳表示する。全ての映像は撮影時刻に基づき同期して再生する。代表選択の数や重要度と強調表示の関係はテキストファイルで定義できるほか、表示プログラム実行中に対話的に変更できる。スクロールや縮尺変更への対応は、現在実装中である。

(3) S/W 構成と動作

図7に表示プログラムの S/W 構成を示す。対話制御スレッドはユーザによる映像再生や注視点変更等の対話操作を受け付け、他のスレッドにメッセージを送る。重要度計算&表示スレッドは、対話制御スレッドやイベント監視スレッドからのメッセージを受け、重要度を計算し各映像の表示方法を決定して、表示定義テーブルを更新する。表示制御スレッドは対話制御スレッドや重要度計算&表示スレッドからのメッセージを受け、表示定義テーブルに従い画面を更新する。イベント監視スレッドは再生中の撮影時刻とイベントの発生時刻を比較し、一定時間以内にイベントが存在すれば重要度計算&表示スレッドにメッセージを送る。

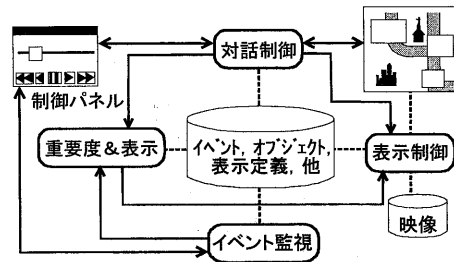


図7 S/W モジュール構成

4.2 実験

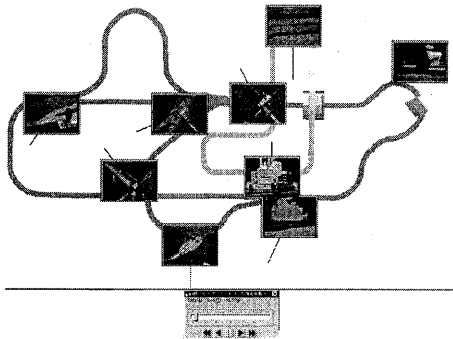
プロトタイプを用いて最初の実験を行った。移動体が地図上に表示される複数の映像ウィンドウを通してどのように見えるのかを評価するため、玩具の電車2台と車2台をレール上で走らせ、9台の固定カメラで撮影した(図8)。移動体の通過や移動体の衝突をイベントとして定義した。現在評価途中であるが、以下のようなことが分かった。

- 重要度を付けずに9個の映像を同時に表示した場合(図9(a))、変化のある映像に次々視線が移動する。衝突と通過が離れた個所で発生した場合、衝突の方

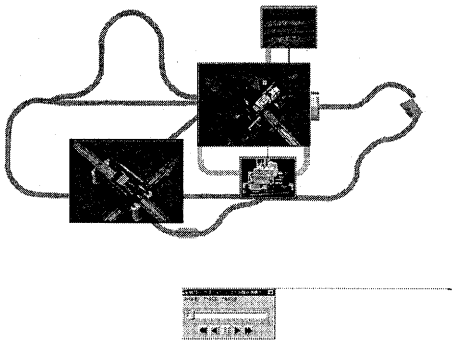
を見落としてしまう場合がある。変化を見落とさずに観察できるのは、近傍 (21" モニタで 100cm² 程度) に表示された 4 個の映像までである。



図8 撮影風景



(a) 重要度を付けずに表示



(b) 代表選択と強調表示

図9 実験におけるブラウジング画面例

- 衝突イベントを通過イベントの重要度より高くして強調表示することにより (図 9(b)), 通過に気をとられて衝突を見落とすことがなくなる。この表示方法は重要度の高いイベントを先に観察させるが、同時に重要度の低いイベントも発生していることをユ

ーザに認識させる効果がある。

- 特定の電車に対する通過イベントの重要度を上げて強調表示すると、その電車の移動の様子が分かりやすくなる。ただ、撮影方向により映像の表示の向きが地図の方向と合わず、地図上で異なる方向に移動するように見える場合、移動方向を勘違いすることがある。
- 強調表示の場合、頻繁に急激なサイズ変更が起こると目が疲れるため、緩やかな変化が好ましい。

5. まとめ

本稿では、オブジェクト、イベント、注視点の重要度に基づき映像の重要度を求め、重要な映像を優先して地図上に表示する新たな多視点映像のブラウジング方式を提案した。これにより、見たい映像を探すのに時間がかかる、どこを撮影しているのか分からない、一つのカメラで撮影できない広い空間で発生する事象をマクロな視点で捉えにくいといった課題を解決する。現在、評価用のプロトタイプを開発し実験を始めたところであり、今後その評価結果や考察について別途報告する予定である。

謝辞 本研究の一部は、文部省科学研究費重点領域研究「高度データベース」(領域番号 275 (08244103)) の援助を受けています。また、本研究の一部は、日本学術振興会未来開拓学術研究推進事業における研究プロジェクト「マルチメディア・コンテンツの高次処理の研究」(プロジェクト番号 JSPS-RFTF97P00501) によっています。ここに記して謝意を表すものとします。

参考文献

- [1] 秦 淑彦, 塚田 晶宇, 尾崎 稔, 坊 覚, “分散型履歴映像データの効果的検索・再生方式と実装,” 電子情報通信学会論文誌 D-I, Vol. J82-D-I, No. 1, pp.234-246 1999.
- [2] A. K. Elmagarmid, H. Jiang, A. A. Helal, A. Joshi, M. Ahmed, “Video Database Systems Issues, Products, and Applications,” Kluwer Academic Publishers, 1997.
- [3] A. Katkere, J. Schlenzig, A. Gupta, R. Jain, “Interactive Video on WWW: Beyond VCR-like Interfaces,” Fifth International World Wide Web Conference, Computer Networks and ISDN Systems, Vol.28, pp.1559-1572, 1996.
- [4] ART+COM, <http://www.artcom.de/>.
- [5] 有川 正俊, “カメラの時空間記述情報を利用したビデオデータの空間ブラウジング,” 情報処理学会データベースシステム研究会研究報告, Vol.98, No.58, pp.225-232, 1998.