

選択的マルチメディア通信方式 SMAP における 自動優先度付けシステムの提案と実装

音嶋 肇 † 橋本 豊大 ‡ 石原 進 * 水野 忠則 *

選択的マルチメディア通信方式は、マルチメディアコンテンツの意味的に重要な情報に高い優先度を付けて、その時の帯域幅にあわせてその情報を優先的に送信する。これにより、スループットが低下しても重要なシーンを確実に転送することができる。しかしながら、その優先度付け処理は、映像であればフレームの1枚1枚に人が優先度を付けなくてはならず、非常に煩わしい作業である。また、当然リアルタイムで生成されるマルチメディアデータに対して手動で優先度を付けることは不可能である。そこで、本研究ではこの優先度付けを自動的に行う自動優先度付けシステムを提案する。このシステムは、動画データのシーン毎にテンプレートマッチングを行い、シーン判定を行うことで自動的に優先度を付けるものである。

Automated priority assigning system for Selective Multimedia Access Protocol

Hajime Otojima †, Toyohiro Hashimoto ‡, Susumu Ishihara *, Tadanori Mizuno *

Selective Multimedia Access Protocol (SMAP) assigns high meaning priority to important information of multimedia content and transmit the information in accordance with the bandwidth and the priority. Even if the throughput of the network deteriorates, SMAP can transmit the important scenes surely. However, the priorities of each piece of a frame must be assigned by hand. Because of this, the priority assigning process is very troublesome work. Furthermore it is impossible to assign meaning priority by manual operation for the multimedia data which naturally is formed by realtime. In this paper, we propose the automated priority assigning system to assign this meaning priority automatically. This system assign the priority automatically to the scenes of the data by matching with template and determining the scene.

1 はじめに

最近ではモバイルコンピューティングという言葉も広く普及し、それに伴いつつでもどこでもデータ通信を行う環境が広まってきている。しかしながら、ワイヤレス通信環境の中で扱われるマルチメディア情報は、十分なネットワークサービスを提供できないため、低解像度の画質、低品質の音声に限られる。

この理由には、1) マルチメディア情報を扱うのに十分な帯域幅を提供できない 2) パースト的な誤りや、移動時のハンドオフにより転送が途切れるということが挙げられる。このため、ビデオやオーディオなどの連続的メディアを通信を用いて扱う場合には、通信回線は比較的大容

量の帯域幅、およびスムーズな転送手段を提供することが重要となってくる。

このような問題を解決するためにこれまで開発されてきた技術は、MPEG などのようにマルチメディア情報の空間的、時間的解像度を調節して、利用可能な帯域幅に合わせてマルチメディア情報を圧縮する方法 [1] などがある。しかしながらこれら従来の技術は、マルチメディア情報の内容の意味的な重要度を考慮せずに単純に品質を落とす処理を行う。このため、データの中の重要な情報が欠落してしまう恐れがある。

この問題に対して、我々はコンテンツ内の意味的に重要な情報を優先して提供する選択的マルチメディア通信方式 SMAP (Selective Multimedia Access Protocol) [2] を提案している。この方式は、マルチメディアデータの各映像フレーム、音声ブロックに対し、その意味的な重要度に基づいてあらかじめ優先度を付け、高優先度な情報を優先して受信者に提供することができる。

しかしながら、SMAP では扱えるデータは蓄積型データに限られ、さらにあらかじめそれらのデータに優先度

† 静岡大学大学院情報学研究所

Graduate School of Information, Shizuoka University

‡ 静岡大学大学院理工学研究所

Graduate School of Science and Engineering,
Shizuoka University

* 静岡大学情報学部

Faculty of Information, Shizuoka University

付けツールを使って、人の手によって優先度付けを行わなければならない。優先度付けはフレーム単位で行うため、その作業に多くの手間と時間がかかってしまう。これに対し、筆者らは自動インデックス入力機能や自動最適化機能などが新たに加えられた改良型の優先度付けツール [3] を導入したが、これにより人による優先度付け作業の負担は軽減されたものの、やはりその作業は煩わしく、依然として扱えるデータは蓄積型データに限られる。

本稿では、この選択的マルチメディア通信方式において優先度付け作業を自動的に行い、リアルタイムで生成されるデータも扱えるような自動優先度付けシステムを提案する。以下、2章で選択的マルチメディア通信方式について述べ、3章で自動優先度付けシステムについて述べる。4章では、自動優先度付けシステムのサブシステムの評価について述べる。

2 選択的マルチメディア通信方式 SMAP

2.1 SMAP の概要

選択的マルチメディア通信方式 SMAP [2] とは、ワイヤレス通信環境のような利用可能帯域幅が狭い通信環境において、効率的に蓄積型のマルチメディアデータにアクセスすることを目的として我々が提案しているマルチメディアアクセスプロトコルである。SMAP ではマルチメディアデータに対し、意味的重要度に基づいて優先度を付与することによって、コンテンツ内の重要な情報を優先して転送し、高品質な再生を提供することができる。

2.1.1 優先度付け

マルチメディアデータの作成者、配布者などは、各映像フレーム、音声ブロックに対して、その意味的な重要度に基づいて優先度を4段階(0~3)で与える。優先度の与え方については優先度を付ける者の主観に委ねられるが、例としてゴルフのレッスンビデオの場合、スイングフォームを表示するフレーム、商品を表示するフレームに高優先度を与えるといったことが考えられる。

SMAP はマルチメディアデータに与えられた優先度に基づいて、利用可能帯域幅に適応する選択的転送を行う。

2.1.2 選択的転送

前述の優先度付け作業によって付与された優先度に基づいて、利用可能帯域幅の中で映像フレーム・音声ブロックを選択的に転送する。マルチメディアデータの各映像フレーム・音声ブロックに付けられた優先度に基づき、利用可能帯域が狭い場合は、低優先度のはスキップし、消費帯域幅を少なくしつつ、高優先度の情報を優先的に転送する。

2.2 優先度付けツール SMAPED

SMAP において優先度付け作業を行うツールとして、SMAPED というアプリケーションが実装されている [2]。

このツールを使うことによって、マルチメディアデータの作成者、配布者などのユーザは、連続した映像フレーム、音声ブロック群を選択して、その選択範囲の映像フレーム、音声ブロックに対して4段階の優先度を与えることができる。優先度付け作業が終わると優先度ファイルが生成されるので、サーバのストレージにマルチメディアデータとともに配置し、サーバの選択的転送スケジューリングで利用される。

2.3 改良版優先度付けツール EDplus

前述した優先度付けツール SMAPED において、ユーザによる優先度付け作業をさらに効率的に行えるように改良したのが改良版優先度付けツール EDplus [3] である。このツールは、インデックス入力機能と自動最適化機能の2つを新しい機能として導入している。

インデックス入力機能は、映像からシーンチェンジを検出して各シーンのインデックスを作成し、ユーザがそのインデックスに対してフレーム単位ではなく、シーン単位での優先度付けを行う。

自動最適化機能は、従来の SMAPED やインデックス入力機能などを用いて作成した優先度ファイルとマルチメディアデータを参照し、各フレームの役割や同優先度フレームの密度やサイズなどを指標として最適化を行い、SMAP をより効率的に機能させるものである。

2.4 課題

SMAP の優先度付け作業における課題として、次の2点が挙げられる。

- 優先度付け作業の負担

優先度付け作業は人手によるものなので、その作業には多くの時間と手間を要する。

- 生のデータが扱えない

同様に、優先度付け作業は人手によるものなので、扱えるデータは蓄積型メディアに限られる。

以上のような課題から、優先度付け作業を自動で行うシステムを使用することで、人による優先度付け作業の負担をなくし、かつリアルタイムで生成されるデータに対しても SMAP で送信できると思われる。本研究では自動優先度付けシステムを提案し、そのサブシステムの評価を行った。

3 自動優先度付けシステム Auto-SMAP

3.1 システムの概要

提案するシステムは、従来人手によって行ってきたコンテンツの意味的な重要度に基づく優先度付け作業を自動で行うものである。あらゆる内容のデータに対しての完全な自動化は困難と考えられる。しかしながら、野球中継などのスポーツ中継のように生放送されることが多く、似たようなシーンが周期的に続くといった特徴のあるコンテンツではあらかじめ特徴のあるシーンの情報の優先度テンプレートを用意することで優先度付けの自動化が実現できると考える。

3.2 優先度付けの方針

図1に自動優先度付けシステムの構成を示す。自動優先度付けは、主にテンプレートマッチングの手法を用いる。しかしながら、リアルタイムで生成される画像フレームのすべてにテンプレートマッチングを行い、画像フレームの1枚1枚に優先度を与えるのは比較的計算量が多くなってしまふと考えられる。よって本システムでは、シーンチェンジフレームで区切られる同一シーン内のフレームには同一の優先度を与えるものとする。

以下に優先度付けの概略を記す。

1. あらかじめ扱うマルチメディアデータと同種類のデータから意味的に重要なシーンのフレームを取り出してテンプレートを作成
 2. 生成されるデータをバッファリング
 3. 各フレームの輝度値ヒストグラムを作成
 4. ヒストグラム差分に基づいてシーンチェンジを検出
- 5a. シーンチェンジが検出された場合

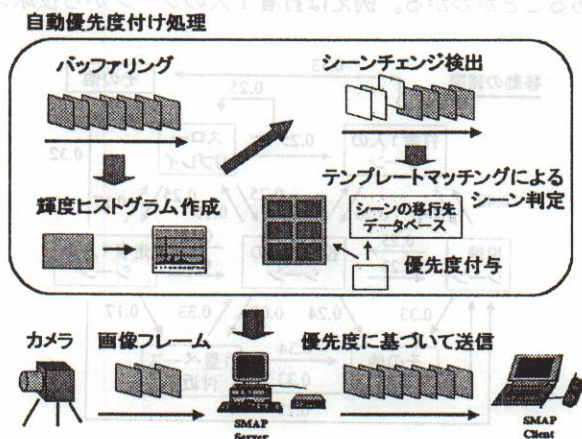


図1: システムの構成図

テンプレートマッチングおよび、シーンの移行先データベースに基づき、シーンを判定、優先度を付与

- 5b. シーンチェンジが検出されなかった場合
 - 4に戻って処理を続ける。
6. 優先度に基づいてSMAP Senderで送信

上記の手順に基づいてリアルタイムに生成されるマルチメディアデータに対して自動的に優先度付けを行う。以下それぞれの処理について具体的に述べる。

3.2.1 テンプレートの作成

自動優先度付けシステムにおいて、優先度付け処理を開始する前にあらかじめテンプレートを作成しておく必要がある。テンプレートは扱うマルチメディアデータと同種類のデータから重要と思われるシーンの代表画像を数枚抽出して作成する。例えば、野球中継に対して自動優先度付け処理を行ってSMAPで送信する場合、あらかじめピッチャーが投球するシーンや、バッターが1人で大きく映っているシーンなどの代表的な画像を抽出してテンプレートを作成する。テンプレート画像にそれぞれ優先度を付けておくことで、マッチングしたシーンの優先度を決定する。

重要シーンのテンプレートは、そのシーンの中の1フレームの輝度ヒストグラムとして表し、そのヒストグラムを用いてマッチング処理を行う。ヒストグラムを用いるのは、画像内の物体が移動したとしても、ヒストグラムの変化がないという性質を利用するためである。また、シーン判定を比較的寛容に行い、マッチング処理の計算量を抑えるために、ヒストグラムの濃度値を256段階から、32段階へと単純化して用いる。ヒストグラムの例を図2に示す。

3.2.2 シーンチェンジ検出

シーンチェンジの検出は、連続する2フレーム間におけるヒストグラムの各輝度値の度数の差分に基づいて行う。しかしながら、すべての連続する2フレーム間でシーンチェンジ検出を毎回行うと計算コストが非常に大きくなる。そこで図3に示すように、まず最初にあるインターバルでのシーンチェンジ検出を行い、そこでシーンチェ



図2: ヒストグラムの例

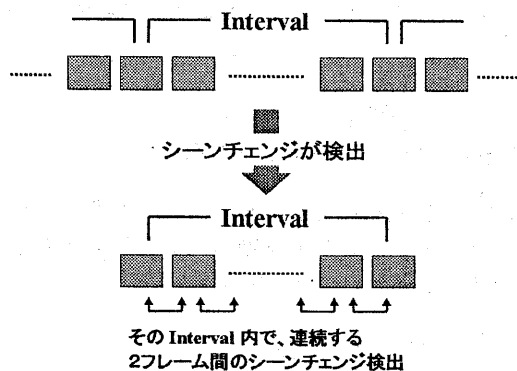


図 3: シーンチェンジ検出の方法

ンジが検出されればそのインターバル内で連続する 2 フレーム間のシーンチェンジ検出を行って、シーンチェンジフレームを検出する。

3.2.3 テンプレートマッチングによるシーン判定

シーンチェンジ検出により抽出されたシーンの先頭フレームに対してテンプレートマッチングを行い、シーン判定を行う。テンプレートマッチングの際に行う処理は、以下のような順序で行われる。

- (1) 次のシーンの予測
- (2) 次のシーンの候補に対してテンプレートマッチングを行い、シーン判定
- (3) 優先度の付与

まず、後述するシーンの移行先データベースを使用して、次のシーンの予測をする。野球放送などの同じようなシーンが周期的に現われるデータの場合には、直前のシーンの種類から、次に現われる可能性の高いシーンの予測ができる。

次に、上記のシーン予測に基づいて移行する確率が高いシーンのテンプレートから順にテンプレートマッチングを行う。

テンプレートマッチングには、イベントアクティビティモデルに基づくシーン検索手法 [4] や、問い合わせ処理によるシーン検索手法 [5] などがある。これらの手法は、いずれも蓄積型マルチメディアデータに対するデータベース検索の手法である。したがって、複雑な処理を用いて、検索は正確に行われる。しかしながら、これらの手法はリアルタイムの画像通信には適さない。

本システムで行うテンプレートマッチングは、リアルタイムに生成される画像フレームに対して行われるものである。最終目的はフレームに優先度を付けて送信することであるので、多くの計算量を必要とする手法は好ま

しくない。したがって本システムでは、比較的単純な輝度ヒストグラムを用いたテンプレートマッチングを採用することで計算量の軽減をはかる。

マッチング方法は、テンプレート画像の輝度ヒストグラムと、ストレージ内の画像フレームの輝度ヒストグラムの各輝度値の差分を取る。シーンごとに求めているしきい値に対し、その差分値が任意のシーンのしきい値を下まわれば、そのシーンを該当するシーンとして認識する。しきい値は、テンプレート中のあるシーンの画像において、そのシーン以外の全てのシーンに対してヒストグラム差分を取り、その値の最小のものをそのシーンのしきい値として設定する。これにより、マッチング処理において差分値がしきい値を越えれば、すぐにそのシーンとのマッチング処理を中止したり、しきい値を下まわればすみやかにシーンの判定を行って次のマッチング処理に移ることができ、計算量を抑えることができる。

シーンが判定されると、テンプレートの優先度に基づいてそのシーン内のフレームすべてに同じ優先度を付与する。

3.2.4 シーンの移行先データベース

シーンの移行先データベースを以下に説明する。シーンチェンジの前のシーンが分かっている場合、次に起きるシーンの候補はシーンの移行先データベースを参照することにより、その移行の確率が高い順に得られるものとする。したがって、これにより得られる移行する可能性の高いシーンのテンプレートから順にテンプレートマッチングを行うことによって効率良くシーンの判定が可能となる。

ある野球放送の 10 分間を解析した結果、代表的シーン間の移行傾向が図 4 のような状態遷移図として得られた。図 4 より、シーン間の移行の傾向は均等ではなく、偏りがあることがわかる。例えば打者 1 人のシーンから投球シー

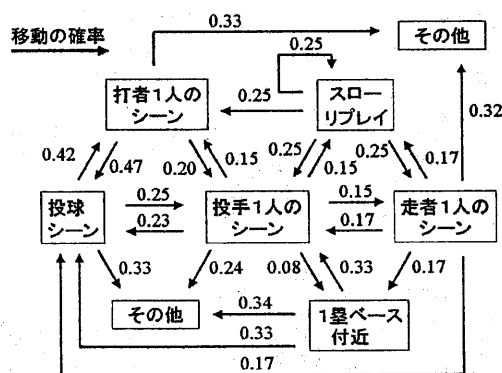


図 4: シーンの移行傾向例

ンへの移行確率は0.47であるから、打者1人のシーンの後のシーンチェンジの際には、まず最初に投球シーンへのテンプレートマッチングを行う。この後、移行確率の高い順に他のシーンのテンプレートと比較することによって効率的なテンプレートマッチングが行える。

本稿では以後、シーン予測の際に用いるデータベースをシーンの移行先データベースと呼ぶことにする。

4 サブシステムの評価

自動優先度付けシステムのサブシステムである、シーンチェンジ検出処理とテンプレートマッチング処理の評価を行った。

4.1 シーンチェンジ検出

シーンチェンジ検出の評価に使用するデータとして、1999年のプロ野球日本シリーズ第2戦、2回表2死からの8分間の実際の野球放送からデジタルデータ(avi形式、サイズ160×120 pixels、1008フレーム、15fps)を用意した。評価方法は、インターバルを5/10/15/20フレームと変化させて、シーンチェンジ検出を行う際の計算量を比較した。上記のそれぞれのインターバルでシーンチェンジ検出を行った結果が図5である。

計算量の評価のための指標として、動画像内のすべてのシーンチェンジの検出に用いた2フレーム間の輝度値の差を計算した回数を用いた。

図5より、インターバルを大きくすれば単純に計算量が減るわけではなく、適したインターバルを用いることによって計算量を最小にできることがわかる。今回の場合は、インターバルが10フレームの時が一番計算量が少なく、インターバルを取らないで行うノーマルのシーンチェンジ検出と比べて、計算量がおよそ4分の1になった。

しかしながらインターバルの最適値の取り方は、シーンチェンジの正確さや、インターバルを大きく設定して

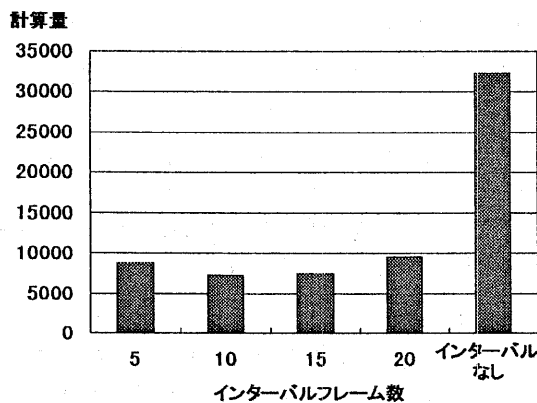


図5: インターバルと計算量

表1: マッチングの成功率と計算量と計算量比率

	成功率 (%)	計算量	計算量比率
データベース未使用	81.8	1648	1.58
データベース使用	63.6	1041	1.00

も送信に間に合うかといった問題とのトレードオフに委ねられる。また、映像の種類によっても大きく違ってくるのが予想される。

4.2 テンプレートマッチングによるシーン判定

本節では、自動優先度付けシステムにおけるテンプレートマッチングによるシーン判定の評価について述べる。連続するフレームからシーンチェンジ検出を行い、その結果シーンチェンジが検出されたフレームに対してのみテンプレートマッチングによるシーン判定を行う。

評価の対象とするのは、シーンの移行先データベースを用いない場合および、シーンの移行先データベースを用いる場合の2つのテンプレートマッチング手法である。

シーンの移行先データベースを用いない場合のテンプレートマッチングは、シーンチェンジを検出したフレームと、テンプレート中のすべてのフレームとのマッチング処理を行って、その中から一番近いシーンを選び出すものである。一方、シーンの移行先データベースを用いた場合のテンプレートマッチングは、シーンチェンジを検出した時、テンプレート中の移行確率の高いシーンから順番にマッチング処理を行い、しきい値を越えなければ、即シーンを決定してマッチング処理を終えるものである。したがって、シーン判定には間違った判定が伴う。

比較の対象とするのは、以下の通りである。

- マッチングの成功率
- マッチング処理の計算量

マッチング処理の計算量の指標として、2フレーム間の各ヒストグラム情報の輝度値の差を求めた回数とした。今回評価に用いたデータは、4.1と同じデータである。また、この1008枚のフレームから、10枚をテンプレートとして作成した。上記の2種類のテンプレートマッチングを行った結果が表1である。シーン判定の正否は、人間の主観に基づいて判断した。

表1より、シーンの移行先データベースを用いる場合のテンプレートマッチングの成功率は63.6%、用いない場合の成功率は81.8%となった。一方、シーンの移行先データベースを用いない場合の計算量は、用いた場合の約1.6倍となった。これより、シーンの移行先データベースを用いる場合は、用いない場合と比べて成功率はやや劣るも

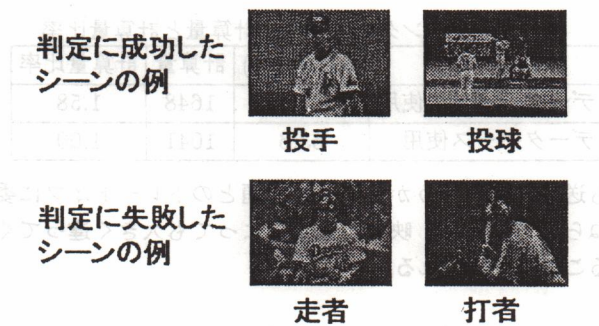


図 6: シーン判定の成功例と失敗例

の、計算量は少なく抑えられることが確かめられた。シーン判定の正否の傾向は図 6 に示す通り、判定に成功したシーンの多くは、シーンチェンジを検出した時の画像内の人物の位置や背景がいつも同じであることが多いのがわかる。これに対して、判定に失敗したシーンは、画面内に大きく人物が映っていてそれを中心に画面が動き、その都度背景が大きく変わるためシーン判定に失敗する。例えば、走者が映っているシーンでは走者の位置によって背景が変わり、それによってシーン判定に失敗を持たらしたと考えられる。

4.3 考察

テンプレートマッチングによるシーン判定の評価を行った結果、シーンの移行先データベースを用いることで、マッチングの成功率をある程度維持しながら、マッチング処理の計算量を抑えることができた。しかしながら、輝度ヒストグラム情報を用いてマッチング処理を行うので、意味的に同じシーンであっても、背景の色の違いなどからマッチング処理を失敗するケースも見られた。したがって、テンプレートを作成する時には、同じシーンのフレームであっても背景の色などが違うフレームを数種類用意することで解決できるものと思われる。また、選手のユニフォームの色の違いや、扱っている野球放送がデイゲームなのか、ナイトゲームなのかによって、それぞれに応じたテンプレートを数種類用意しておくことで、これらの問題に対処できると思われる。

優先度付けの根本的問題として、テンプレート作成者の感じる意味的重要さと、見る側、すなわちユーザの感じる意味的重要さにはズレがあるという点が挙げられる。これは、最終的にテンプレートをユーザ側が選択できる機能を提供することで、解決できるものと思われる。例えば野球放送の場合、自分がひいきにしているチームが映っている画面を映像中に多く含ませたいのならば、そのチームが多く含まれるテンプレートをユーザ側に選択してもらおうといった具合である。このようにして、この

問題に対して解決法を考える必要がある。

5 おわりに

本稿では、SMAP の人手による優先度付け作業の省略と、SMAP を生放送へ適用することに向けて、自動優先度付けシステムの提案とその実装、および評価を行った。さらに、シーンの移行先データベースを用いることによって、テンプレートマッチングをより効率的に行えることがわかった。

今後の課題は、優先度付けをさらに詳細に行う方法の検討である。今回実装したシステムは、シーン単位で優先度付けを行うため、シーン内のイベントの意味的な重要性を考慮していない。したがって、シーン内でも優先度の差をつける方法を考える必要がある。その方法として、シーン内の物体の形状を計算量を抑えつつ判定する手法や、シーン内の物体の意味的に重要な動きの検出、長いシーン内において優先度の高いイベントの検出などが挙げられる。これらの手法を今後検討する。

さらに、人が付けた優先度と、自動で付けた優先度を比較することで、本システムの評価を行う。また、今回は野球放送を用いて本システムを評価したが、野球放送以外の一般的な動画像へ本システムを適用することも今後の課題である。

参考文献

- [1] 和田 正裕, ビデオ圧縮符号化技術の昨日 * 今日 * 明日, 情報処理学会誌, Vol. 41, No. 1, pp. 59-63, Jan. 2000.
- [2] 太田 賢, 渡辺 尚, 水野 忠則, ワイヤレス通信環境における選択的マルチメディア通信方式の実装, 情報研報, Vol. 97, No. 35, pp. 141-146, April, 1997.
- [3] 橋本 豊大, 増田 彰久, 太田 賢, 水野 忠則, 選択的マルチメディア通信方式 SMAP における優先度付けツール, マルチメディア、分散、協調とモバイル (DI-COMO '99) シンポジウム論文集, pp. 103-108, June, 1999.
- [4] 牛尼 剛聡, 広部 一弥, 渡辺 豊英, 利用者の視点に基づくシーン検出のためのイベント-アクティビティモデル, 電子情報通信学会論文誌, Vol. J82-D- I, No. 1, pp. 256-267, January, 1999.
- [5] 田淵 仁浩, 村岡 洋一, 動画像データベース中の系列データを指定する条件の不完全さを許容できる問い合わせ処理と MeSOD モデル, 電子情報通信学会論文誌, Vol. J76-D- I, No. 6, pp. 288-299, June, 1993.