

## 手ぶれを考慮した MPEG2 からのカメラワーク検出の検討

土橋 健太郎<sup>†</sup> 小舘 亮之<sup>††</sup> 西塔 隆二<sup>††</sup> 富永 英義<sup>††</sup>

<sup>†</sup> 早稲田大学 理工学部 電子・情報通信学科  
〒169-8555 東京都新宿区大久保 3-4-1

<sup>††</sup> 早稲田大学 国際情報通信研究センター  
〒169-0051 東京都西早稲田 1-3-10

doba@tom.comm.waseda.ac.jp

動画像編集のためのサポートとして本研究では、撮影者の意図が含まれるカメラワークを検出する。従来の圧縮コンテンツからカメラワークを検出する研究ではニュース、映画、ドラマなどの安定したカメラワークで撮影された映像を対象としている。本研究では、デジタルビデオカメラ (DVC) で撮影された手ぶれを含んだ映像も対象とする。カメラワークを高精度に検出するには、カメラワークと手ぶれをうまく分離する必要がある。まず、手ぶれも含めたカメラの動いているシーンを検出する。その後、各フレームの動きベクトルの平均と、頻度分布の分散値を算出する事によりカメラワークと手ぶれを分離する。カメラワークの区間を検出した後に各フレームでの動きベクトルの大きさを調べる事によりカメラワークの速さも編集のサポートに利用できる。

キーワード 動画像編集, シーン検出, カメラワーク, 手ぶれ, MPEG

## Camera working parameter Extraction Methods using Motion vector information in MPEG2 video

Kentaro DOBASHI<sup>†</sup> Akihisa KODATE<sup>††</sup> Ryuji SAITO<sup>††</sup> Hideyoshi TOMINAGA<sup>††</sup>

<sup>†</sup> Dept. of Electronics, Information and Communication Engineering, WASEDA University  
3-4-1 Ohkubo Shinjuku-ku, Tokyo, 169-8555 JAPAN

<sup>††</sup> Global Information and Telecommunication Institute, WASEDA University

doba@tom.comm.waseda.ac.jp

Recently digital video cameras are getting popular, and video editing has become one of attractive applications even on home-use PCs with IEEE 1394. However video contents editing is still not very easy and some sorts of intelligent supporting system are requested, which realize semi-automatic editing or show contents information to editors. Therefore we focus on camera works that contain the intention of the users. As the first step, we have investigated a camera work scene (zooming and panning) extracting system including camera shakes. Additionally, scene distinction between camera work and camera shake has become possible, by considering motion vector characteristics.

keywords Video Editing, Scene Distinction, Camera work, Camera shake, MPEG-2

## 1. はじめに

高性能デジタルビデオカメラ、動画編集機能の搭載されたPCの普及により、一般ユーザでも手軽にデジタル映像コンテンツを撮影・所有できるようになった。しかし、現在のシステムにおいては録画時間以上の編集時間がかかり、必ずしも一般ユーザが使いやすい環境が整っているとは言えない。そこで、映像コンテンツを内容に基づいて意味的に構造化するVDA (Video Document Architecture) モデル<sup>(1)</sup>により、ユーザに内容を分かりやすく提示するシステムが有用だと考えられる。

これまでビデオインデキシングに関する研究が盛んに行なわれてきた。例えば、カット点を検出した大辻らの研究やカメラワークを構造化のパラメータとした検討<sup>(2)(3)(4)</sup>もされている。しかし、これらの研究は非圧縮コンテンツに対して行なわれている。これから、圧縮データとして流通する事が考えられる。そこで、MPEG符号化データを用いる方式について検討<sup>(5)</sup>する必要がある。また、Yang<sup>(6)</sup>やSmith<sup>(7)</sup>らの研究にみられるようにカメラ操作部分に撮影者の意図が反映されており、有用なシーンと考えられる。

従来の研究で対象としている動画コンテンツは、ニュース・ドラマ・映画などの安定したカメラワークである。本稿では、DVC (デジタルビデオカメラ) から撮影された手ぶれの含まれる動画コンテンツを対象としている。そこで、手ぶれを考慮にいれたカメラワークの検出を検討した結果について報告する。

## 2. 従来手法と問題点

従来のMPEGから動きベクトルを利用してカメラワークを検出する方法は、次のようである。

- (a) 各フレームごとに空間のグローバルな動きを検出する。
- (b) まとまったフレームごとに時間的なグローバルなカメラの動きを決定する。

空間的なグローバルな動きは、いくつかの方法で検出されている。<sup>(6)</sup>らは、9つのブロックに分割し角の4ブロックの動きベクトルの方向の多数ベクトルから求めている。他にも、<sup>(8)</sup>らは動きベクトルのヒストグラムからグローバルな動きを検出している。

しかし、これらの方法だけでは次のような問題点がある。

- 撮影者の意図しない手ぶれをカメラワークと判定してしまう。
- カメラワーク中に手ぶれが含まれると異なるカメラワークの判定をしてしまう事がある。

手ぶれによる問題は、図1を見ると明らかである。図1は、従来のグローバルな動きを検出する事によってカメラワークを検出した時の結果である。それぞれ、各フレームでの動き

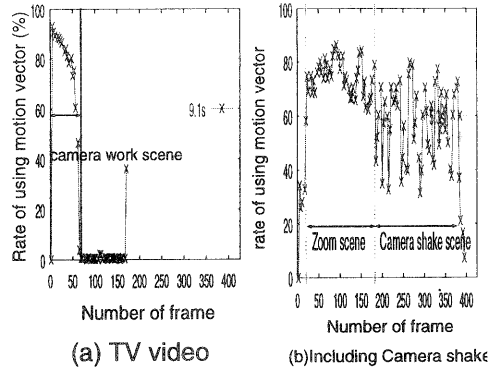


図1: 安定したカメラワークのコンテンツと手ぶれを含んだコンテンツでの検出

ベクトルの使用率が高い時がカメラワークと判定している。本研究でも、始めの処理としてこの方法を用いているが、この処理だけでは図1の手ぶれを含むシーケンスに対して正確に検出出来ない事が分かる。

そこで本研究ではカメラの動いているシーンを検出した後に、カメラワークと手ぶれを分離する方法を提案する。

## 3. カメラワークと手ぶれの検出方法

### 3.1 定義

本稿で用いる式を以下に定義する。

動きベクトル使用率:

$$RV_i = \frac{1}{N} \sum_m f_{m,i} \quad f_{m,i} = \begin{cases} 0 & v_{i,m}(x) = v_{i,m}(y) = 0 \\ 1 & \text{otherwise} \end{cases} \quad (1)$$

$v_{i,m}(x), v_{i,m}(y)$  は  $i$  番目のフレームの  $m$  番目のマクロブロック (MB) の動きベクトル (MV) の  $x$  方向,  $y$  方向を示す。  $N$  は、1 フレーム内の MB の数を表す。

平均動きベクトル変化量:

$$\theta_i = \tan^{-1} \left( \frac{AV_i(y)}{AV_i(x)} \right), \quad \Delta\theta_i = \theta_i - \theta_{i-1} \quad (2)$$

$AV_i(x), AV_i(y)$  は、各フレームの平均動きベクトル。このとき、横方向を  $0^\circ$  とすると、ベクトルの向き  $\theta_i$ 、 $\Delta\theta_i$  はベクトルの向きの変化である。

ベクトル向きの頻度分布の分散値:

$$\sigma_i = \frac{1}{N'} \sum_{j=1}^8 (n_{i,j} - \frac{N'}{8})^2 \quad (3)$$

ベクトルの頻度を図2のように8通りにクラスタリングし  $j$  番目の向きの頻度  $n_{i,j}$  と平均頻度  $\frac{N'}{8}$  を利用した分

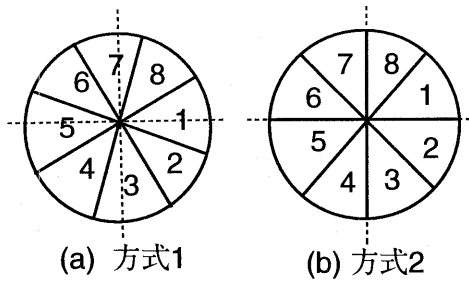


図 2: ベクトルの向きのカテゴリ

散値である。但し、 $N'$  は 1 フレームで発生した動きベクトルの数である。

動きベクトル量:

$$QV_i = \sum_m \sqrt{v_{i,m}(x)^2 + v_{i,m}(y)^2} \quad (4)$$

各フレームのベクトルの大きさの和を動きベクトル量と定義する。

### 3.2 カメラワークの分類

本研究では、カメラワークのパラメータを次の 5 つに分類した。

1. *fix* 区間 (開始, 終了フレーム)
2. 手ぶれ区間と程度
3. *Zoom* 区間とその速さ
4. *Pan* 区間とその速さ
5. 電源 ON/OFF 操作 (操作時フレーム)

ただしその他のカメラの動き、チルトやトラック、ブーム、ドリーは大きくパン・ズームに分類した。

### 3.3 カメラワーク検出の流れ

カメラワーク区間の検出に、P ピクチャの動きベクトルを利用した。まず、図 3 に全体の流れ図を示す。

- (a) パン、ズームをしているシーンは画像全体が動くので、1 フレーム内の多くのマクロブロックがベクトルを有すると考えられる。そこでカメラワークシーンでは、式 (1) の値が大きくなると想定される。但し、動きの激しいシーンで多いイントラ予測の場合も数えている。
- (b) パンとズームの判定をする。1 フレームの平均動きベクトルを計算する。ズームシーンは各 MB の動きベクトルの向きが均一に分布しているので平均ベクトルが 0 に近い値になる。手ぶれシーンでもどちらかに分類される。
- (c) パンと手ぶれの区別をする。手ぶれは細かいカメラワークと考える事ができる。安定したカメラワークに比べて、手ぶれはカメラの向きが周期を持って変化する。

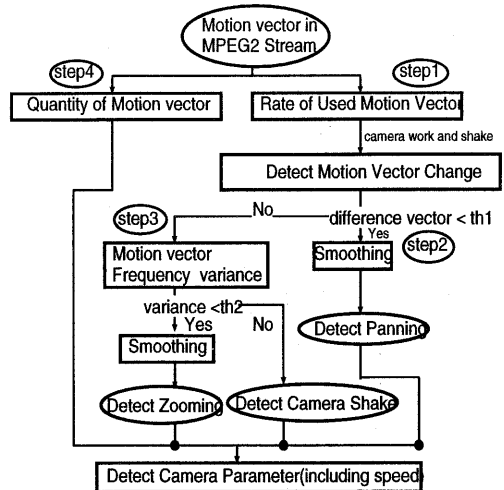


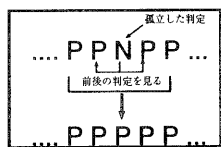
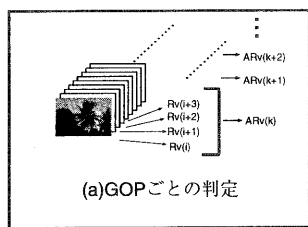
図 3: カメラパラメータ検出流れ図

る。そこで、式 (2) を考えると、手ぶれシーンで、頻繁に  $\Delta\theta_i$  が変化するので、カメラワーク区間と区別する事ができると考えられる。

- (d) ズームと手ぶれの区別をする。2 のステップでは、1 フレームの平均ベクトルを利用しているのでマクロブロックごとの動きベクトルが互いに相殺しあうズームには適していない。そこで、ズームと手ぶれの検出をするためにズームの特徴を利用する。ズーム区間では、全体的に見てマクロブロックの持つベクトルが均等に分布している。つまり式 (3)、フレーム内のベクトル頻度の分散値がズーム区間では小さな値になると想定できる。この動きベクトルの頻度分散値を調べる事によりズームと手ぶれを分離する事ができる。
- (e) 以上のステップをフレーム単位で行なう。次に時系列にロバストな判定を行なうために、まとまったフレームで判定をする。1GOP ごとに平均をしそれぞれのステップを判定する。また、ステップ 2 の変化は、手ぶれ区間ではベクトルがある周期を持って変化するので、1GOP ごとの最大値を採用し平滑化する。
- (f) カメラワークの速さを評価する。速いカメラワークをしているシーンは各マクロブロックの有するベクトルの大きさが大きくなる。この性質を利用して、式 (4) を計算すると、カメラワークの速さが動きベクトル量で評価する事ができると考えられる。

### 3.4 時間的な階層判定

カメラの動きをフレーム単位で判定を行なうと、例えば次のような場合にシーンの過分割が起きてしまう。



(b) 孤立した判定の回避

図 4: GOP ごとの判定

- カメラワーク中で手ぶれが起きた場合.
- カメラの目の前を一瞬オブジェクトが横切った場合.
- 撮影条件の瞬時的な照明条件の変化 (ビデオカメラの近くでカメラのフラッシュや雷が起きた場合)
- シーンチェンジが起きた場合 (動きベクトル使用率が大きくなる.)

そこで、フレーム単位で求めた各値を MPEG における GOP ごとにまとめて判定する方法を検討した。図 4(a) に示すように、1GOP 毎の動きベクトル使用率、向きの変化、頻度分布分散値の平均を算出して一定の閾値によって判定する。GOP ごとの判定の後に、シーンチェンジやカメラのフラッシュなど瞬時の変化に対しては、図 4(b) のように参照 GOP の判定結果が前後 GOP の判定と異なる場合は、前後の GOP の判定を参照し判定結果の統合を行なう。この GOP は本研究で用いた符号化器では約 0.5[s] である。

#### 4. 実験と結果

デジタルビデオカメラ (SONY 製, DCR-PC1000) で静止画オブジェクトをズーム、パンを交え撮影したコンテンツを入力して実験を行なった。コンテンツの種類は表 1 に、MPEG 符号化器の性能を表 2 に示す。MPEG2 符号化には、Canopus 社の MVR-D2000 を使用した。

##### 4.1 動きベクトル使用率による評価

図 5 に、ズーム・パンの場合を示す。グラフは、横軸がフレーム番号、縦軸が使用率である。縦の破線は、主観評価によるカメラワーク区間である。図 5(a) のズームは徐々に使用率が大きくなっているのが分かる。図 5(b) のパンは高い使用率である事がわかる。本稿では、動きベクトル使用頻度判定法と定義する。

表 1: 実験で使用したシーケンス

シーケンス名前	カメラワークの速さ [s]	手ぶれ
zoom1	6.9s	有り
zoom2	0.9,1.4,5.3,5.8s,7.2,9.1	有り
pan1	2.3s,3.5s,5.5s,14s	無し
pan2	11.5s	有り

表 2: MPEG2 符号化器の使用

MPEG2 シーケンス要素	パラメータ
frame rate	30 frames/s
frame in GOP	15 frames
I/P frame distant	3 frames
picture format	frame and field picture coding
horizontal size	720 pixels
vertical size	480 pixels
bit rate	4M bits/sec

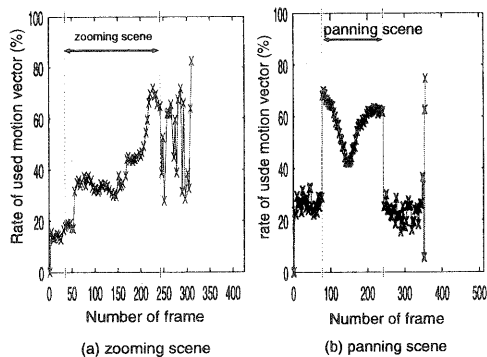


図 5: 動きベクトル使用率

##### 4.2 平均動きベクトルの向きの変化による手ぶれの検出 (パン)

本実験のシーケンスは、3.1 の実験に利用したものと異なり、パンの前後で手ぶれが起きているシーケンスを入力データとして用いた。図 6(a) に平均ベクトルの向きの変化を示したものである。さらに、図 6(b) は 1GOP ごとに平滑化して、上の波形が使用率からカメラワーク区間と判定したもの、下の波形が手ぶれと判定したものである。これにより使用率で検出したカメラワーク区間から手ぶれシーンを分離することができる。

##### 4.3 フレーム内のベクトル頻度分布の分散による手ぶれ検出 (ズーム)

ズーム区間と、手ぶれシーンの分離を目指した結果を図 7 に示す。ベクトル向き頻度の分散による実験の結果は図

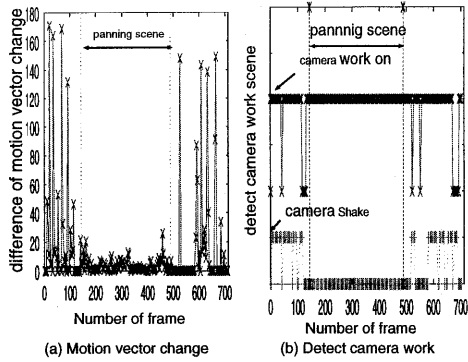


図 6: 平均動きベクトル向きの変化と手ぶれ・カメラワーク検出

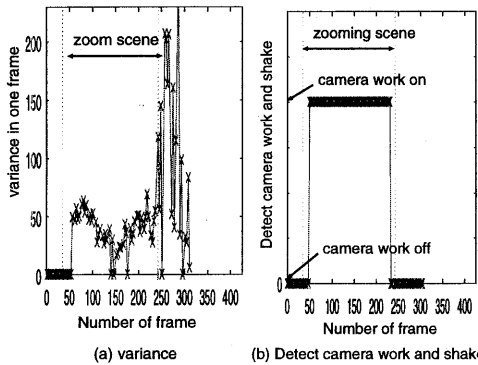


図 7: フレーム内のベクトル向き頻度分布の分散値と検出結果

7(a), その結果手ぶれシーンと検出した結果を図 7(b) に示す。

#### 4.4 結果

手ぶれとカメラワークの検出結果の例を表 3 に示す。

従来法とは、動きベクトル使用率のみからカメラワーク区間を検出した場合、提案手法では、さらに平均動きベクトル変化やベクトル頻度分布の分散値を用いてカメラワーク区間を検出した場合である。シーケンス pan1 のように手ぶれの含まれない場合は、従来手法においても抽出が可能である。しかし、その他の手ぶれの含まれるシーケンスで従来法を適用すると手ぶれ部分も過剰検出してしまっているのが分かる。手ぶれを考慮した提案手法では、手ぶれシーンを排除できているので主観評価に近い結果が出ているのが分かる。その他のカメラワークの速さでも同様の結果が出た。しかし、別に撮影した暗い廊下でのカメラを固定して撮影したシーンで動きベクトル使用率が高くなり誤検出してしまっている場合があった。

表 3: カメラワークと手ぶれの検出結果の例

シーケンス	主観評価	従来法	提案
zoom1	1.1s-8.0s	1.6s-9.2s	1.6s-7.7s
zoom2	0.7s-6.0s	0.3s-12.3s	0.3s-5.8s
pan1	2.6s-4.7s	2.1s-4.9s	2.7s-4.3s
pan2	4.8s-16.2s	0s-22.2s	4.6s-17.0s

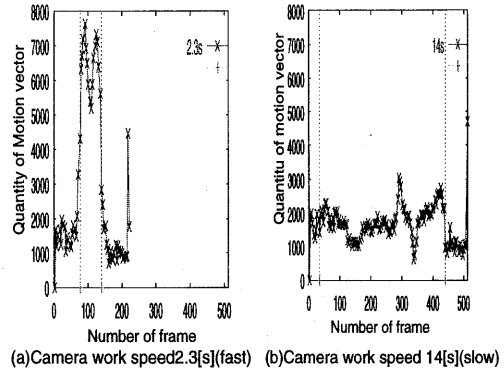


図 8: カメラワークの速さと動きベクトル量

#### 4.5 カメラワークの速さ

カメラワークの速さは、式 (4) で定義した動きベクトル量を調べることにより判定する。速いカメラワークでは、シーンの移り変わりが速いので MB ごとの参照するフレームが速くなる。つまり、速いカメラワークはフレーム内の動きベクトル量の和が大きく、遅いカメラワークでは小さくなる。同じシーンを速いパンとゆっくりなパンで撮影したコンテンツでの動きベクトル量の違いを図 8 に示す。

#### 5. 外枠判定

これまで、動きベクトルの探索を 1 フレーム内に存在する全てを対象としてきた。しかし、ズームの開始点などで 1 フレームの中心部分がほとんど変化しないので動きベクトル使用率が低いという問題があった。カメラワークの種類ごとの動きベクトル使用率を図 9 に示す。

図 9 で分かるように、ズーム区間でフレームの中心部分の動きベクトル使用率が低くなっている事がわかる。そこで、図 10 に示すようにフレームの外枠部分での動きベクトル使用率に着目する。

Yang<sup>(6)</sup> にも、フレームを 9 つのブロックに分割し角の 4 つを利用しているがより外枠を利用する事により検出率を上げる事が可能になる。同じシーケンスを用いて、全体に動きベクトル使用頻度判定法の場合と外枠判定法を図 11 に示す。全体で動きベクトル使用率を見ている方は、ズームの始めで低い値になってしまっている。つまり、カメラ固定の時

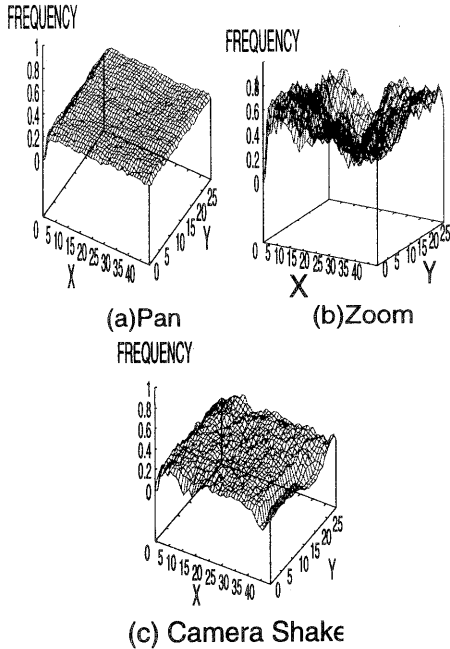


図 9: 各カメラワークでの動きベクトル使用率

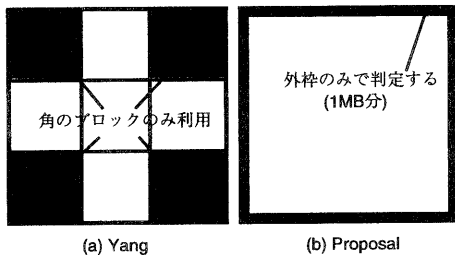


図 10: 外枠判定方法

と区別が付きにくい、外枠判定をしているほうは、ズーム区間で常に高い動きベクトル使用率になっている。このため、外枠判定をする事によりズームの始めの未検出を防げる事が分かった。

## 6. まとめ

本稿では、動画を編集・要約する際に役立つ有用な情報として、撮影者の意図が反映されたカメラワークを検出する手法を提案した。そこで、従来の研究が対象としている安定したカメラワークのコンテンツだけでなく、手ぶれを含んだ動画コンテンツに対してもカメラワークを検出する事を目

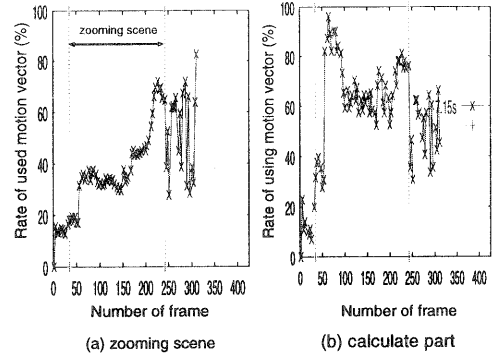


図 11: 外枠判定の結果

的とした。まず、動きベクトル使用率から、カメラの動いている区間を検出し、パンとズームの性質を持つ区間に分類する。その後、平均ベクトルの変化、ベクトル頻度分布の分散値からカメラワーク区間と手ぶれ区間の分離する手法を提案した。さらに、動きベクトル量を測る事によりカメラワークの速さもカメラパラメータの要素として加える。また、外枠判定を導入する事によって、シーン検出の精度を上げる事ができた。

今後は、1.動きオブジェクトを含むシーンへの対応、2.ズームとパンの両操作を同時に含むシーンへの対応を目指し手法の改良を行なっていく予定である。

## 参考文献

- (1) 宮森 恒: “構造特徴に基づく映像情報の生成および編集方式の研究”, 早稲田大学大学院 理工学研究科博士論文 (1997).
- (2) 阿久津, 外村, 橋本, 大庭: “動画像インデキシングを目的としたカメラ操作の規定法”, 信学論, **J75-DII**, pp. pp.22-30 (1992).
- (3) 谷口, 阿久津, 外村: “PanoramaExcerpts: パノラマ画像の自動生成・レイアウトによる映像”, 電子情報通信学会論文誌, **D-11-j82** (1999).
- (4) 上田, 宮武, 吉澤: “認識技術を応用した対話型映像編集方式の提案”, 信学論 (1992).
- (5) H.-C. H.Liu and G. L.Zick: “Scene Decomposition of MPEG Compressed Video”, **SPIE**, **2419**, .
- (6) Y. Yang: “A Study on Camera Work Extraction Technology from Compressed Information”, 映像情報メディア学会, **53**, (1999).
- (7) M. Smith and Kanade: “Video Skimming and Characterization through the Combination of Image and Language Understanding Techniques”, **IEEE CVPR** (1997).
- (8) 大野, 青木, 吉田, 酒井: “動きベクトルと色情報を用いた階層的 MPEG 画像検索”, 映像メディア処理シンポジウム, No.5, pp. pp.101-102 (2000).