

# 差分画像を利用した アニメーション映像からのオブジェクト抽出

中神 央二<sup>†</sup> 渡辺 裕<sup>†</sup> 富永 英義<sup>†</sup>

<sup>†</sup> 早稲田大学 大学院 国際情報通信研究科  
〒 169-0051 東京都新宿区西早稲田 1-3-10

E-mail: †{ouji,tominaga}@tom.comm.waseda.ac.jp, †hiroshi@giti.waseda.ac.jp

あらまし 本稿では、アニメーション映像における動オブジェクト抽出の検討を行う。具体的には、セル画により表現される動き情報を利用する。アニメーション映像中の動きには、その制作手法から、スライド、リピート、ローリングなどの表現手法が用いられる。これら動作を検出するためのアルゴリズムとして、フレーム間の差分画像に基づいた方法を提案する。特に、映像中に大局的な動きがある場合について、その検出、補正アルゴリズムを結果とともに示す。実験の結果、大局的な動きを補正して差分画像を生成することで、精度良くオブジェクトを抽出できることを確認した。

キーワード オブジェクト抽出, 差分画像, アニメーション映像

## Object extraction using differential images from animation images

Ouji NAKAGAMI<sup>†</sup>, Hiroshi WATANABE<sup>†</sup>, and Hideyoshi TOMINAGA<sup>†</sup>

<sup>†</sup> Graduate School of Global Information and Telecommunication Studies,  
Waseda University  
Nishiwaseda 1-3-10, Shinjuku-ku, Tokyo, 169-0051 Japan

E-mail: †{ouji,tominaga}@tom.comm.waseda.ac.jp, †hiroshi@giti.waseda.ac.jp

**Abstract** In this paper, we propose an object extraction method from cartoon films. This method utilizes motions represented by celluloid images. In animation, some sheets of celluloid represent frames of actions. They are called “slide,” “repeat,” “rolling,” and so on. Proposed algorithm detects such actions by using differential images between continuous frames. When background images are moving, motion compensation is necessary to extract objects. Therefore, algorithms which detects and corrects background motion are also proposed. The experimental result shows effectiveness of this method.

**Key words** object extraction, differential image, animation movie

## 1 はじめに

我々は、対象となるコンテンツに最適な符号化方式を適用するという、コンテンツオリエンテッド符号化の概念を提唱している [1], [2]。なかでも、アニメーション画像の性質に着目し、静止画像に対する符号化の検討を行ってきた。これは、画像の形状情報に着目し、領域単位で符号化を行うものである。今回は、この符号化方式を動画画像へ応用するために、アニメーション映像中に含まれる、セル画で描かれた領域の抽出手法について考察する。

ここで対象とするのは、いわゆるセルアニメーションと呼ばれるものである。これらは、セル画と呼ばれる透明なシートの上に、主人公などのキャラクターを描いて制作される。背景部は、背景画として1枚の大きな絵を用意しておき、キャラクターの描かれているセル画を、背景画の上に重ねて撮影する。

一般に、彩色行程の煩雑さから、同一のセル画が何度も使用されることが多い。TV 放映されるアニメーション番組において、実際に作成されるセル画の枚数は、すべて異なるコマとして計算した場合の10分の1程度と言われている。したがって、映像中のどの部分が同一セル画によるものかを判定することができれば、符号化の際の情報量を減らすことが可能になる。

本稿では、その前段階として、セル画により表現されたオブジェクトを抽出する手法について、提案アルゴリズムと実験結果を述べる。

## 2 アニメーション映像における動きの分類

セル画の部分は、オブジェクトの動きを表現するために、1コマずつ僅かに変化させて、フィルムに撮影される。この際、同一、あるいは数枚のセル画の繰り返しによりキャラクターの移動等を表現することが多い。これには、様々な手法が用いられるが、最も基本的なものとして、次の3つが挙げられる。

スライド キャラクターの平行移動のこと。隣合うフレームにおいて、セル画を上下左右にわずかにずらして撮影することで表現される。この時、拡大・縮小、回転を加えることもあり得る。

リピート 限られた複数枚のセル画を繰り返し用いることで、動きを表現する。瞬きのシーンなど。

ローリング セル画のスライド等で一定の範囲を揺れるように反復移動させる方法。馬車の車輪など。

符号化への応用を考えた際、情報量の縮体のために最も効果が高いと考えられるのは、スライドとリピートの動きである。そこで、これらの動きに特化した検出、判定アルゴリズムが必要となる。なお、ローリングに関しては、スライドとリピートの組み合わせと考えることができる。

## 3 差分画像を利用したオブジェクト抽出手法

先に述べたアニメ特有の表現手法から、同一のセル画によるオブジェクトが、シーケンス内の別のフレームにおいても登場する可能性は高い。そこで、検索のためのテンプレートとして、オブジェクトを抽出する必要があ

る。これには、オブジェクトの外形を精度良く抽出できる手法が必要になる。

まず、提案アルゴリズムのフローチャートを、図1に示す。前提として、動きを検出する際には、予めショットの単位へ分割されたシーケンスを対象とする。これは、シーン分割を行う自然画の手法が、精度良くアニメーション映像に適用できるという [13] の文献に基づいている。

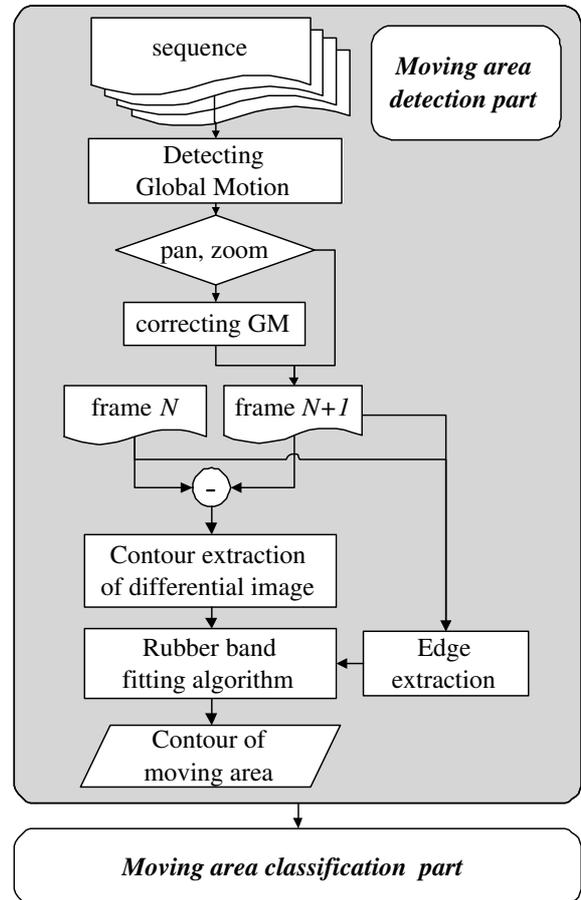


図1 動領域の検出・追跡アルゴリズム

### 3.1 動領域の抽出

アニメーション映像における動領域を検出するには、フレーム間の差分画像を利用する。

自然動画画像を対象としたオブジェクト抽出には、動きベクトルなどを利用する手法 [7] が提案されている。しかし、アニメーション映像の場合、自然画像と異なり、平坦な色遣いの領域が多く含まれる。そのため、正確なオブジェクト抽出に用いるのは難しい。また、予め背景画像を作成する手法も提案されているが [12]、画面中に占めるオブジェクトの割合が小さく、かつ大きな移動がある場合にしか適さない。これではアニメの場合、先に挙げたスライド以外の動きの検出に用いることが難しい。

一方で、アニメーション映像の特徴として、同一の背景領域を連続するフレームで用いることが多い。そのため、差分画像を生成することで完全に背景を消去することが可能である。隣接フレーム間で、差分を絶対値として扱い画像を生成した場合、変化した部分の領域が検出

される(図2参照)。その輪郭線は、前後フレームにおける各領域の輪郭線を併合せたものとなる。この粗い輪郭線情報を利用して、各フレームにおいて対応する領域の、真の輪郭線を抽出する。これには、元画像のエッジ情報と、輪ゴムのアルゴリズム [10] を用いる。

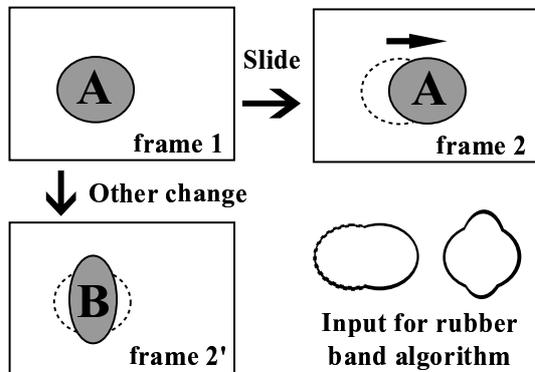


図2 動領域の粗い輪郭線検出

輪ゴムのアルゴリズムは、視覚的な凹凸情報を抽出してオブジェクト形状を段階的に表現する手法である。境界線構造の解析方式として、オブジェクトの回転・サイズ・微小ノイズに高い耐性を有する。

これには、まず概念的にオブジェクトに輪ゴムをかける。そして、輪ゴムとオブジェクトとの接触/非接触関係を、画素の反転処理により抽出する。この反転処理の操作を反復することで、オブジェクトの全境界線特徴の抽出を行うものである。

今回は、輪ゴムの始点を、差分画像より検出した粗い輪郭線情報とし、その対象を、元画像におけるエッジとする。アニメーション画像にはたくさんのエッジ成分が含まれるが、これによりそのフレームにおいて変化している領域の輪郭線のみを検出することができる。本方式の利点として、処理量が少ないこと、ピクセルベースでの抽出が可能になることが挙げられる。

ただし、元の領域の一部のみが変形する場合には、輪ゴムは内側からかける必要がある。すなわち、 $A$  が、 $A' \supset A$  または、 $A' \subset A$  になる場合である。このとき、差分の輪郭線から、外側にたどることにより、前者では後フレームの  $A'$ 、後者では前フレームの  $A$  の領域を検出することができる。

最後に、検出した輪郭線を用いて、元画像から後段のマッチングに用いるテンプレート画像を切り出す。

### 3.2 大局的な動きへの対応

アニメーション映像においても、背景画像を動かした場合、自然画におけるカメラモーションのような大局的な動きが発生する。前項で述べた動領域の抽出手法では、ショット中にこの大局的な動きが存在する場合、領域の抽出に失敗するという問題がある。単純な差分画像の生成では、背景領域を消去することが不可能になるためである。

そこで、ショットにおける大局的な動きを判別し、次にそれを補正することが必要となる。前者にはオプティカルフローを基盤とする手法、後者には全探索による手法を用いた。オプティカルフローのみでは、ピクセルレ

ベルで大局的な動きを決定することが難しいこと、逆に全探索のみではショットにおけるパンやズームといった判定を行うことができず、正確な対応点探索を行うことができないからである。

はじめに、ショットにおけるカメラモーションを決定する手法について述べる。前述の通り、一般にオプティカルフローや動きベクトルを人工的な画像に用いた場合、平坦な色遣いの部分で、対応点が正確にとれないという問題がある。そこで、木村 [5]、西原 [6] らの手法を参考に、フローの信頼性を高める処理を行った。

木村ら [5] の手法では、勾配法を基に、フローの分布に関する仮定を置いている。一つは画像の小領域では、オプティカルフローは一定であること、もう一つは、画像全面に渡り滑らかに変化するということである。この仮定を基に、抽出された各フローの信頼度関数を評価することで、サブピクセルレベルの小さな移動の検出精度に優れる結果が得られるとしている。

本研究では、アニメーション映像に精度良く適用するために、西原ら [6] の手法をもとに、さらにエッジの法線方向にあたるフローに重みを付加した。これにより、アニメーション映像においても、大局的な動きを反映したフローを抽出することが可能になった。

ショットから抽出されたフローは、土橋ら [11] の手法によりクラスタリングを行うことで、パン、ズームの判定を行う。全画面でオプティカルフローが近い向きを持つパン区間では、特定のクラスの頻度が高くなる。一方、ズームでは均等に各クラスに動きベクトルが分布する。検出の手順としては、次のようになる。

1. 8つのクラス(図3参照)の頻度の、全体に占める割合を出す。
2. ヒストグラムの中で最大の頻度のクラスを探す ( $c1$  とする)
3. その最大のクラスの両側のクラスの内に大きい方のクラスを選ぶ ( $c2$  とする)
4.  $c1$  と  $c2$  のクラスの割合を加える。
5. 閾値によりパン・ズームの分離を行なう。

ここで、最大頻度のクラスのみの割合を用いないのは、動き予測の誤差から2つのクラスに大多数が分布する場合があるからである。

次に、検出されたパン、ズームの情報をもとに、各フレーム間において背景部がどのような動きをしているかを探索する。これには、秦泉寺ら [9] の手法で提案される、差分画像におけるオブジェクトの面積を最小にする動きが、大局的な動きを表現している、という条件を用いる(図4参照)。パンのシーンであれば、最尤クラスタの方向に対して全探索を行い、差分画像において閾値以上の画素の数が最も少なくなる点を求める。

この結果を用いて差分画像を生成することで、大局的な動きがある場合にも、前項で述べた、無い場合と同様の手法でオブジェクトを抽出することが可能になる。

#### 3.2.1 実験結果

大局的な動きの検出、補正に関して、実験を行った。いずれもパンやズームを含むショットに対して、オプティ

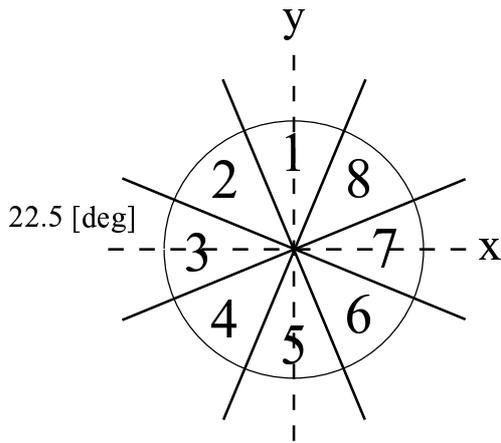


図 3 8 方向へのクラスタリング

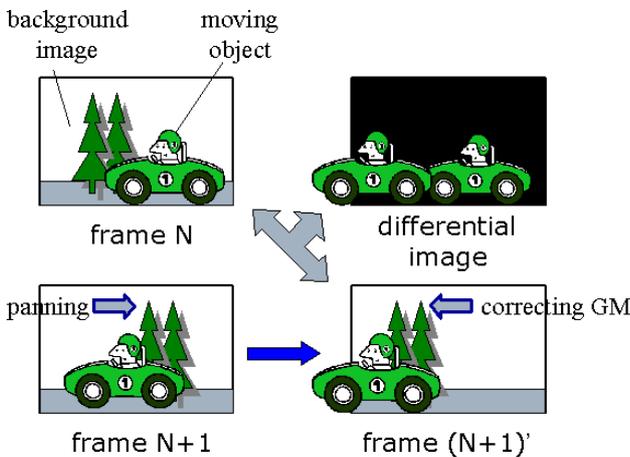


図 4 大局的な動きの補正

カルフローを用いることでその検出を行った。そして、パンを含むショットに関しては、全探索法により求めた補正值と、手で求めた実測値を比較した。

実験条件を表 1 に示す。図 5 に、パン (“pan1” ~ “pan3”) およびズーム (“zoom1”) を含むショットにおける各フレーム間の全フローを同一平面上に展開したものと、それをクラスタリングした結果を示す。なお、“pan1”, “pan3” はオブジェクトの動きを含み、“pan2” および “zoom1” はオブジェクトに動きは無い。

表 1 実験条件

テスト画像	“pan1”, “zoom1”	“pan2”, “pan3”
画像解像度	352[pixel] × 240[line]	640[pixel] × 480[line]
色空間	YCbCr	YCbCr

フローの信頼度を高める前置処理の結果、平坦な色遣いを含む映像においても、パンやズームの検出を行うことが可能であることが確認された。特に、“pan3” ではオブジェクトの動きが比較的多いが、クラスタリングアルゴリズムと併せることで、正確な判定が可能になっている。

次に、パンのショットにおいて、全探索を行った結果を図 6 ~ 図 8 に示す。各フレームにおける大局的な動き

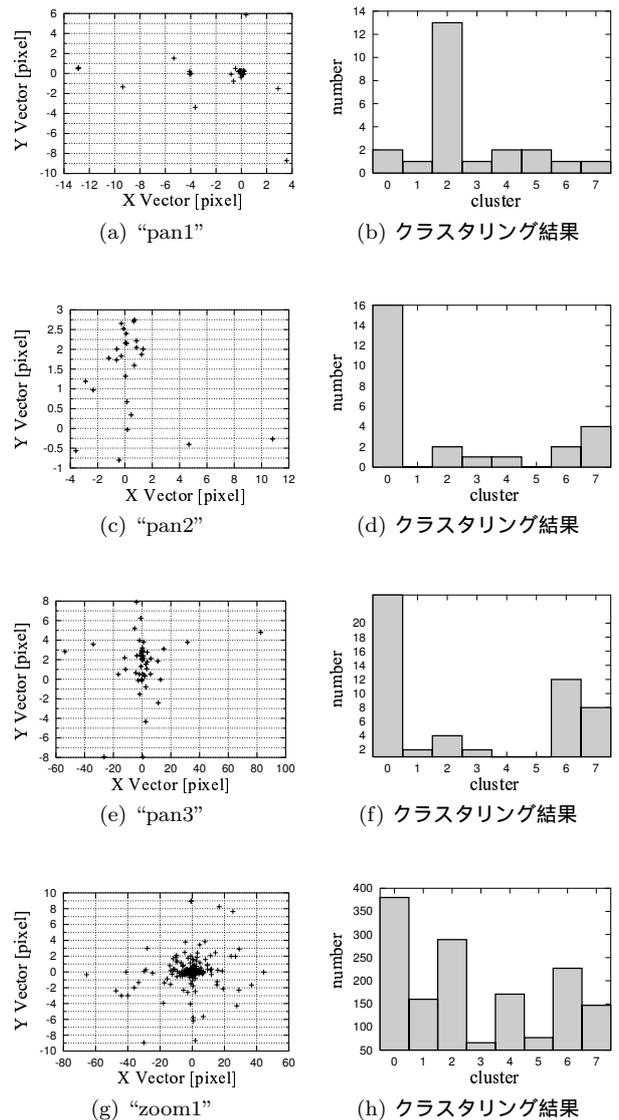


図 5 大局的な動きがあるショットでのオプティカルフローを時系列上にグラフにしたもので、横軸はフレーム番号、縦軸は画素数である。

単純な全探索法では、誤探索の問題がある。画面中に占める割合の大きいオブジェクトが移動する場合、そちらを大局的な動きと判定することがある。ショットのシーン判定を行わない場合、いくつかのフレームにおいて誤判定した。しかし、前段階で抽出されたパン方向へ探索範囲を絞ることにより、これを回避できた。すなわち、オブジェクトがカメラモーションと別の方向に動く場合は、その探索結果を除外することが可能になる。結果を見ると、検出結果はほぼ実測値と等しく、正確に抽出することが可能であるといえる。

### 3.3 オブジェクト抽出結果

提案手法によるオブジェクト抽出は、エッジ情報に基づいているため、輪郭線法の入力として理想的な初期輪郭線が得られれば、正確なオブジェクトの抽出が可能になる。ただし、差分画像のエッジが、背景領域における輪郭線と交差している場合などは、初期輪郭線が得られ

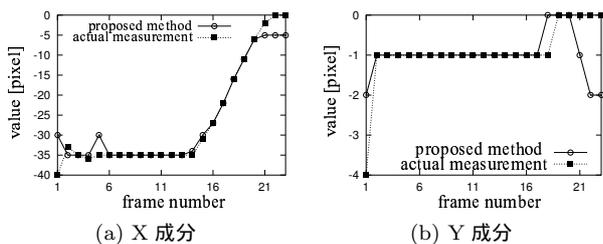


図 6 “pan1” における検出結果

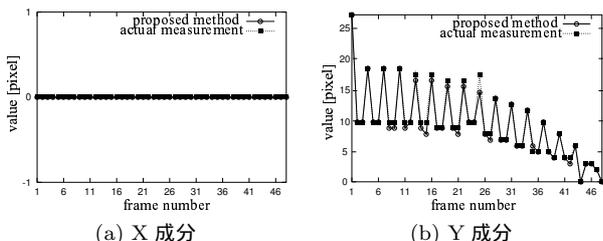


図 7 “pan2” における検出結果

ず、オブジェクトの抽出そのものに失敗する。

表 2 にオブジェクトの抽出結果を示す。フレーム間でオブジェクトに移動等の変化がある場合のうち、差分画像から初期輪郭線が得られた場合を成功と判断し、その割合を示す。テスト画像は、先ほどのシーケンス (“pan1” および “pan3”) に加え、大局的な動きが無い場合のショット (“fix1” および “fix2”) について示す。これらの実験条件は、“pan1” と同じである。

表 2 オブジェクト抽出結果

テスト画像	オブジェクトに 変化がある場合	初期輪郭線獲得に 成功した場合	割合 %
pan1	7	5	71
pan3	20	14	70
fix1	8	8	100
fix2	18	17	94

大局的な動きがある場合は、1 ピクセルでもその補正がずれると、背景領域による偽の輪郭線が発生し、オブジェクトの輪郭線と交差するため輪郭線が得られないことがある。そのため、全探索において、オブジェクトの面積最小となる点という条件に加え、さらに精度を上げるための拘束条件が必要になると考えられる。一方で、大局的な動きが無い場合には、オブジェクトの正確な抽出が可能となった。

#### 4 まとめと今後の課題

セルアニメーション映像を対象に、ショット中からセル画によるオブジェクトを抽出するための手法を提案した。実験により、カメラ固定の状態では、高い確率でオブジェクトの抽出に成功することを確認した。

また、ショット中に大局的な動きを含む場合への対応を考察した。実験の結果、カメラワークの判定は、オブジェクトの動きによらず成功することがわかった。大局的な動きを補正して差分画像を生成することで、オブジェクトの抽出率が向上した。しかし、補正に 1 画素で

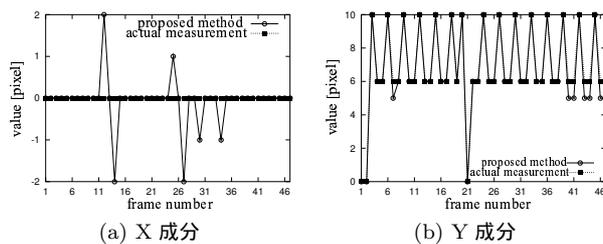


図 8 “pan3” における検出結果

もずれが生じた場合、抽出に失敗することがあることも判明した。

今後の課題としては、ズームや、パンとズームを同時に含むような場合でも適用できるかどうかを確認することが挙げられる。

#### 参考文献

- [1] O. Nakagami, T. Miyazawa, H. Watanabe, and H. Tominaga, “Study on two-layer coding for animation images,” In Proc. *IEEE Int'l Conf. on Multimedia and Expo (ICME)*, Aug. 2002
- [2] N. Shimizu, T. Miyazawa, W. Kameyama, H. Watanabe, H. Tominaga, “A novel decoder downloadable system for content-oriented coding,” In Proc. *IEEE Globecom 2002*, Nov. 2002
- [3] 岩田, 安居院, “セルアニメーション画像の間の閉領域同定方法に関する研究,” 信学技法, PRL80-9, pp. 9-16, Sep. 1980
- [4] 塩原, 後藤, 吉田, “3次元構造に基づくアニメーション画像のフレーム間対応,” 信学論誌 (D-II), Vol. J73-D-II, No. 10, pp. 1687-1695, Oct. 1990
- [5] 木村, 太田, 金谷, “精密なノイズモデルによるオプティカルフローの検出,” 情処研報, 96-CV-99-6, pp. 37-42, Mar. 1996
- [6] 西原, 吉田, 酒井, “信頼度に基づくブロック統合を用いた動画像の高精度動きベクトル推定とその応用,” 映像情報メディア学会誌, Vol. 53, No. 1, pp. 148-156, Jan. 1999
- [7] P. Anandan, M. Irani, R. Kumar, and J. Bergen, “Video as an image data source: efficient representations and applications,” In Proc. *IEEE Int'l Conf. Image Processing (ICIP)*, Oct. 1995
- [8] 田中, 佐野, 大原, 奥平, “パラメトリックテンプレート法とその高精度ロバストマッチングへの応用,” 信学論誌 (D-II), Vol. J-83-D-II, No. 4, pp. 1119-1130, Apr. 2000
- [9] 秦泉寺, 渡辺, 岡田, 小林, “MPEG-4 スプライト符号化を用いた超低レート動画像圧縮,” 信学論誌 (D-II), Vol. J84-D-II, No. 5, pp. 758-768, May 2001
- [10] 小館, 朱, 金, 浦野, “輪ゴムかけの概念に基づく境界線の構造記述と未定義文字処理システムへの応用,” 信学論誌 (D-II), Vol. J84-D-II, No. 1, pp. 1-11, Jan. 2001
- [11] K. Dobashi, A. Kodate, H. Tominaga, “Camera working parameter extraction for constructing video considering camera shake,” In Proc. *IEEE Int'l Conf. on Image Processing (ICIP)*, Oct. 2001
- [12] 境田, 苗村, 金次, “背景差分法と時空間 watershed による領域成長法を併用した動画像オブジェクトの抽出,” 信学論誌 (D-II), Vol. J84-D-II, No. 12, pp. 2541-2555, Dec. 2001
- [13] Bae, Y. J., 高橋, 中嶋, “セルアニメーション制作補助のためのキャラクタ動作データベース構築 第 1 報: Shot 変化の自動検出,” 通信全大, D-12-80, Mar. 2002
- [14] 角, 中嶋, “2D アニメーション用標準動画像データベースシステム,” 通信全大, D-12-81, Mar. 2002