

# 嗜好分類に基づく個人適応型サッカー映像の自動生成技術

窪田 進太郎<sup>†</sup> 有木 康雄<sup>†</sup> 塚田 清志<sup>‡</sup>

<sup>†</sup>神戸大学自然科学研究科 〒657-8501 神戸市灘区六甲台町 1-1

<sup>‡</sup>大阪株式会社開発部 〒530-0013 大阪市北区茶屋町 17-1

E-mail: <sup>†</sup> kuvota@me.cs.scitec.kobe-u.ac.jp, ariki@kobe-u.ac.jp, <sup>‡</sup> tsukada@mbs.co.jp

**あらまし** サッカーの固定映像からデジタルカメラワークを用いて、わかりやすく面白い映像を自動生成する研究を行っている。デジタルカメラワークとは高解像度の映像に対し、ある一部分をクリッピングし、そのフレームサイズと位置を制御することで仮想的にパンやズームを実現する技術である。これまで、パンとズームを中心としたデジタルカメラワークを研究し、個人に適応する映像を生成してきたが、今回は映像視聴の嗜好を分類することによって、簡易に映像の個人適応を行う手法を研究した。生成された映像を、個人の主観を反映させることができる AHP 法によって比較し、良好な結果を得た。

**キーワード** デジタルカメラワーク, サッカー, 自動映像生成, 個人適応, 嗜好分類, AHP 法

## Automatic Production Method Based on Preference Classification For Soccer-Game Videos

Shintaro KUBOTA<sup>†</sup> Yasuo ARIKI<sup>‡</sup> and Kiyoshi TSUKADA<sup>‡</sup>

<sup>†</sup> Science and Technology, Kobe University 1-1 Rokkodai-cho, Nada-ku, Kobe-shi, Hyogo, 657-8501 Japan

<sup>‡</sup> MBS 17-1 Chaya-machi, Kita-ku, Osaka-shi, Osaka, 530-0013 Japan

E-mail: <sup>†</sup> kuvota@me.cs.scitec.kobe-u.ac.jp, ariki@kobe-u.ac.jp, <sup>‡</sup> tsukada@mbs.co.jp

**Abstract** We are studying about automatic production of soccer sports videos for easy understanding by using digital camerawork on camera fixed videos. The digital camerawork is a movie technique which uses virtual panning and zooming by clipping frames from hi-resolution images and controlling the frame size and position. We have studied so far digital panning and zooming. In this paper, we propose a method of preference classification for personal adaptation. We compared the proposed technique with a conventional technique by AHP method that can reflect an individual subjectivity.

**Keyword** Digital Camerawork, Soccer, Automatic production of videos, Personal Adaptation, Preference Classification

### 1. はじめに

近年のデジタル技術の発達により放送インフラが整い、様々な新しいサービスが始まっている。BS, CS, 地上デジタル放送などの技術により、チャンネル数の増大に伴うコンテンツ数の増大、データ通信を使ったメタ情報付与などが可能になるが、それはすなわち放送局の作業量や、人件費、制作費などのコストの増大につながる。コスト面のほかにも、映像資源、編集者、カメラマンの数の有限性から、無尽蔵にチャンネル数を増やすことは不可能である。この問題を解決できるような映像生成技術が、今求められていると考えられる。

これを解決できる映像の種類として、アマチュアスポーツ映像がある。従来は、チャンネル数の制限や採算の面から放送局が積極的にアマチュアスポーツ映像を放送していくことはなかったが、チャンネル数の増

大と近年のスポーツブームにより需要が高まってきていると言える。

これまで我々は、この問題を解決するための技術として、デジタルカメラワークを用いたアマチュアスポーツ映像の自動撮影という観点から研究を進めてきた。デジタルカメラワークとは、高解像度の固定カメラにより映像を撮影し、各フレームに対してデジタル処理を用いたクリッピングを行い、擬似的なカメラワークを実現するものである。この技術の利点は、時間の遅延を許すならば、突発的なイベントにより映像の取りこぼしなどが起こらないことや、高解像度映像さえあればカメラワークを何度でも作り直せる点である。

現在は、放送インフラの発達から、対話型テレビという概念にも注目が集まっている。従来のような一方

が重視されるようになって来ている。

これを実現するための技術として、我々は個人適応型サッカー映像の自動生成を提案した。それまでの映像はプロのカメラマンの映像技術を模倣するように、すなわち、誰にでも受け入れられ易いような映像を自動的に生成していたが、視聴者一人ひとりが映像生成のパラメータを変更することにより、より個人に適応した映像を生成する技術である。しかし、複数のパラメータを個人が設定することは非常に煩雑であり、思ふようなパラメータの値を選択することは困難である。そこで、本研究では視聴嗜好をあらかじめ分類しておき、そこから選択して視聴する技術について提案する。

生成した映像と従来手法の映像、TV映像などをAHP法による評価で比較し、良好な結果を得ることができた。

## 2. 提案手法の概要

本システムで用いる固定映像の撮影には、デジタルハイビジョンカメラを用いた。720p(1280x720)の高解像度映像を解析し、480p(640x480)の映像を切り出す。ズーム処理がなければこのまま現行の放送で用いられているレベルの画質を保つことが可能である。しかし、本研究ではパンに加えてズームも用いるため、デジタルズームによる画質劣化が起こる。この問題は、将来的な映像技術の発達により1080p(1920x1080)などの解像度の高いカメラが開発される可能性があると考えられることと、また、ソフトウェア的にも高解像度化が可能であることから、本研究では扱っていない。図1に実際のデジタル処理の様子を示す。

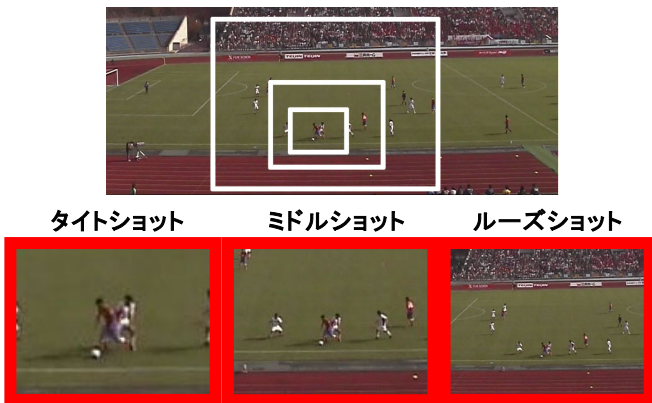


図1. デジタル処理によるクリッピング

提案手法における処理の流れを図2に示す。まず全景映像を取り込み、画像処理部で選手とボールを抽出し、特にボールについては座標の移動を追う。次に、状況認識部でボールや選手の座標、保持時間などから状況の認識、イベント(フリーキック、コーナーキックなど)の検出を行う。最後にカメラワーク部で視聴者

から選択された視聴タイプを考慮し、それぞれの状況に適切なフレームサイズ、フレーム座標を選び最終的なカメラワークを決定する。処理の流れは状況認識、カメラワークの順だが、本紙では内容的な解り易さを考え、3章でカメラワーク、4章で状況認識について述べる。

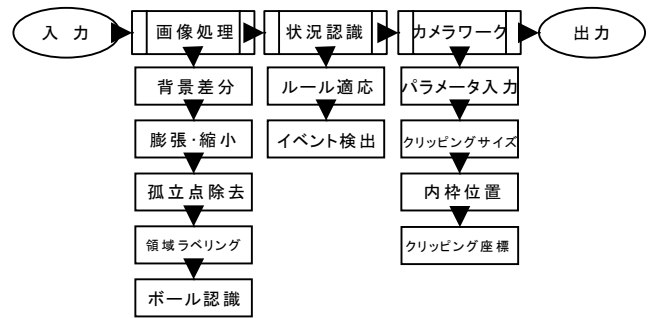


図2. 処理フローチャート

## 3. カメラワーク

状況認識部から渡された各フレームのパラメータと、入力された嗜好分類を用いて、クリッピング座標を連続的に変化させ、状況に応じてクリッピングサイズと座標を変化させることによって、固定映像からデジタルカメラワークを用いたサッカー映像を生成する。

### 3.1. クリッピングサイズ

TV映像を解析した結果、三種類程度のクリッピングサイズを用いるのが適切であると判断した。この結果、本研究では、表1に示すようなタイトショット、ミドルショット、ルーズショットを用いる。これを図3のように変化させる。ミドルショット、ルーズショット間は連続的に変化させ、随所にタイトショットを断続的に挿入する。

表1. 各クリッピングサイズ

タイトショット	ミドルショット	ルーズショット
120x90	240x180	480x360

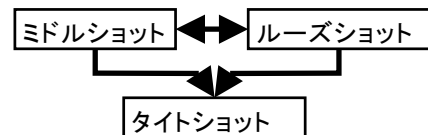


図3. クリッピングサイズの変化

状況認識部から渡されたパラメータと視聴者が入力した嗜好分類を用い、適切なクリッピングサイズを決定する。タイトショットは断続的に挿入するため、あまり短すぎると醜い映像になってしまうために、最低でも二秒以上は継続するように設定した。

### 3.2. クリッピング座標

図4に示すように、クリッピング映像の内部に枠を設ける。クリッピング座標は、基本的にはボールを追従するように変化させていく。しかし、完全に追従してしまうと、ボールの微小な動きにとらわれたり、ノイズを拾ったりし、予測のできない動きをしてしまう。それを防ぎ、より視聴のしやすい映像を生成するために、内枠の中に存在するときはクリッピング座標を変化させず、内枠からはみ出した時は、収まるようにクリッピング座標を変化させるという方法をとる。ボール座標の変化量が小さいときはスムーズに変化し、大きいときは、クリッピングサイズを大きくすることにより対処している。また、この処理は一定フレームまとめて行い、その際に最小二乗法を用いて滑らかな座標変化ができるように処理を行っている。

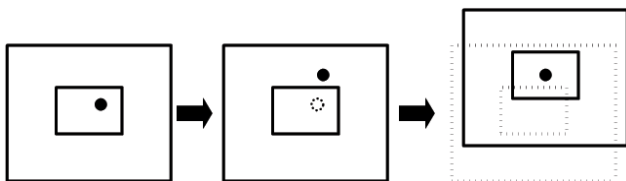


図4. ボールの細かい動きに追従しない内枠

### 3.3. 内枠位置

内枠は基本的には画面中央に設けているが、シーンによっては内枠の位置をシフトさせることにより、適切な映像を生成している。例えば、ドリブルで前にすばやく進んでいくシーンなどでは、進行方向のフォーメーションがどのように展開していくのを見るために、内枠を進行方向と逆に移動させる。

### 3.4. 嗜好分類

これまでの研究では、プロのカメラマンを模倣したカメラワーク決定の手法に加え、視聴者それぞれがカメラワーク決定のパラメータを変更できる手法を示した。パラメータとは、クリッピングサイズの変化速度、クリッピングサイズ、イベント処理などである。しかし、複数のパラメータを数値的に変化させるために、視聴者の思い描いている値に決定することは難しい。

パラメータの組み合わせにより、カメラワークは無数に決定することができるが、実際に選択されたパラメータの分布には偏りがあり、それを嗜好として分類したものを選択できるようにした。嗜好を六つに分類して、それぞれ望遠、個人、イベント、瞬間、オフENS、ディフェンスと名付けた。

望遠とは、ルーズショットを多く使い、全体を見やすくする嗜好である。フォーメーションなどが確認しやすく、また、画角に対してのボール追従速度が小さいために、パン動作が小さく見やすい。

個人とは、個人をより細かく観察するためにミドルショットからタイトショットを多く用いる嗜好である。ボールの画角に対してのボール追従速度が上がるために、パン動作が大きくなるが、細かいボールの流れが確認できる。

イベントとは、フリーキック、スローイン、ゴールなどのイベントを重視した嗜好である。イベント時の選手のポジションの取り方、ボールを動かすタイミングなどを確認しやすい。

瞬間とは、望遠と個人の両方を嗜好である。基本的には大きくとったクリッピングサイズを用いるが、細かい動作、たとえばボールの取り合いなどの時にだけ小さいクリッピングサイズを用いるものである。一般的な嗜好のズーム処理を、より極端にした状態といえる。

オフENS、ディフェンスとは、クリッピング座標をオフENSとディフェンスが均等に入るように決定するのではなく、どちらか片方に偏らせる嗜好である。オフENSやディフェンスの全体の動きを確認することができる。

これに加えて、これまでの研究で用いたプロのカメラマンを模倣したものの中から七つを選択可能にしている。

## 4. 状況認識

画像処理部で処理した画像データを用いてサッカーの試合の状況認識を行う。本研究ではデジタルカメラワークを用いた映像生成技術の有用性を示すことが目的であるため、画像処理部のボール認識に失敗した部分は手動により座標を入力した。状況認識とは、映像中のイベント検出や、データを解析することによって、その場面は視聴者にとってどのようなカメラワークで見たい場面なのかを判別することである。カメラワーク部では嗜好分類により状況認識は変化するが、ここでは一般的な嗜好での認識を行っている。ここでいう一般的とは、プロのカメラマンが撮影してTVで放送されるような、という意味である。TV放送から、撮影技術を以下に述べるようにルール化した。

### 4.1. ズーム処理のためのルール

クリッピングサイズの変化のきっかけになるルールは、大部分がボールの座標の変化によるものである。

基本的には、ボールの座標がゴールに近づいたときと、ボールがあまり動かないときにズームインしてミドルショットとし、その逆のときにズームアウトしてルーズショットとする。ゴール前でボールがすばやく動く場面や、ピッチ中央でゴールがあまり動かない場面もあるが、それぞれに優先順位をつけているために、適切なクリッピングサイズが得られる。ボール一定時間停止した状態を、フリーキックなどのイベントが起

こる前の状態と判断し、選手を一時的にピックアップするためにタイトショットを挿入する。イベント認識に関しては後述する。

#### 4.2. 内枠の位置に関するルール

前節で述べたように、本研究ではボールの微少運動に反応しないカメラワークを生成するための内枠を設けている。通常は画面の中央部に設けているが、すばやい攻撃展開によりゲームの関心が移動していくときに、ボールを中心に捕らえるのではなく進行方向に余裕を持って撮影することにより、次のプレイがどのように起こるかを判断することができる。そのためのルールとして総オブジェクト（選手）の平均座標が同じ方向に動いているときに、内枠を左右にシフトさせるというルールを設ける。

#### 4.3. イベント認識のためのルール

本研究におけるイベントは、現時点ではボールや選手の座標などを用いて簡易的に表現できる六種類のものである。それはフリーキック、ゴールキック、スローイン、コーナーキック、ペナルティキック、キックオフ（ゴール）である。キックオフに関しては、ゴールそのものを検出することは難しいが、その後のキックオフを検出することにより間接的に検出することができる。これらのイベントはボールが一定時間以上静止する種類のものであり、ズームインへの変化と同時に起こる。サッカー映像の解析から、平均的にこの状態が6秒以上続いたときにイベントが発生したと認識し、その全体の状況を確認するために、一度ズームアウトし、その後イベントの進行方向に内枠位置をずらすことにより、イベントを適切に表示することができる。

各イベントを同定するために、ボール座標、選手座標の重心とボール座標の距離を基準にしている。それぞれのルールを以下の表2に示す。

表2. 状況認識ルール

イベント名	ボール座標	集団からのボール座標	クリッピングサイズ
フリーキック	フィールドエリア	遠い	ルーズショット
ゴールキック	ゴールエリア	遠い	ルーズショット
スローイン	タッチライン外	中間	ミドルショット
コーナーキック	コーナースポット	中間	ミドルショット
ペナルティキック	ペナルティスポット	中間	タイトショット
キックオフ	センタースポット	中間	ルーズショット

### 5. 結果と考察

#### 5.1. 実験環境

実験に用いた映像は第83回全国高校サッカー選手権大会京都府大会決勝 京都朝鮮高級学校 対 桂高等学校の試合を用いた。

この映像から、デジタル処理により、デジタルカメラワークを用いたサッカー映像を自動生成した。デジタルカメラワークを用いたサッカー映像は、一般的な嗜好で生成されたものと、嗜好を選んで生成したものの二種類を評価した。この映像に加えて、プロカメラマンが撮影した地上波放送による中継映像と、デジタル処理をする前のコート全景の高解像度固定映も比較の対象とした。

#### 5.2. 評価基準

スポーツ映像の評価基準としては表3のようなものが考えられる。全景固定映像に関しては1と2が評価不可能だが、それも映像の価値として判断する。

表3. AHP法の評価基準

基準	評価
1. ズーム処理	良い <-> 悪い
2. パン処理	良い <-> 悪い
3. ショットサイズ	良い <-> 悪い
4. ショット持続時間	適当 <-> 不適當
5. 画質	きれい <-> 粗い
6. 試合の流れ	良い <-> 悪い

#### 5.3. AHP法を用いた主観比較法

映像の評価というものには絶対的な基準が存在しないために、視聴者の主観評価が必要となる。このような事象の比較は、単純に各判断基準のスコアを合計するだけでは行うことはできず、各項目の重要度は視聴者の嗜好によって異なることが多い。そのために、個人の嗜好を反映した一対比較法として有効であるAHP(Analytic Hierarchy Process)法が知られている。この手法では評価項目に重みをつけることが可能なため、個人の嗜好が反映され、正確な評価が可能となる。また、デジタルカメラワークを用いた技術として、HDDレコーダーなどに録画された映像をAHP法を用いて評価し、デジタル処理して再生成するということも考えられるため、この手法は有効であるといえる。AHP法についての詳しい情報はそれを解説する論文を参照されたい。また、AHP法の各値の計算を表計算ソフト上で自動的に行うマクロが「公開されており、計算にはそれを用いた。

#### 5.4. 結果と考察

AHP法を用いて各評価基準の重要度を求めた結果は、表3の評価基準番号順に、(0.058, 0.182, 0.087, 0.084, 0.105, 0.483)となった。試合進行が最も重要度が高いと判断され、その次にパンが続き、他はあまり差のない重要度となった。

重要度を加味して求めた各映像の総合評価値を以下の図5に示す。TV映像の評価が最も高く、次に提案手法映像、次に高解像度固定映像(HD映像)とな



っている。最も重要視される評価値である試合の流れが、本手法では高評価を得ており、全体における各評価の値の割合が TV 映像と類似していることから、全体としても高評価を得ている。ズームインを行っているため、画質の評価がやや下がっているが、この問題は将来的に、より高解像度のカメラを用いたり、また、低解像度画像を高解像度画像に変換する手法を研究したりすることにより改善されると考えられる。提案手法が全景固定映像よりも高評価を得ていることから、単なる固定映像に提案手法の処理を行うことにより映像評価を上げることは可能であるといえる。このことから、放送局がコストの問題から従来放送コンテンツとして扱わなかった対象を、提案手法を用いることによって放送コンテンツにできる可能性は十分ある。

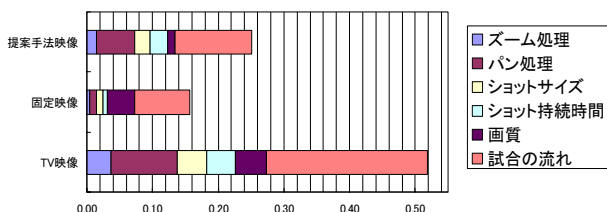


図 5. AHP 法による評価結果

次に、一般的な嗜好で生成された映像と、各被験者が嗜好を選択した後に生成された映像を比べて評価を行った。変更前と変更後を比べて、どちらがよかったかという評価を行った結果、15人中10人が変更後、5人が変更前というものだった。この5人に関しては、選手の同定や、フォーメーションの解析、個人のテクニクの撮影、複数カメラでの撮影など、本研究の現段階では到達していない要求を求めるものであった。これらの項目に関しては、次章に記す。現段階の範囲では一般的な嗜好に比べて高い評価を得られることが確かめられたと言える。

また、イベント検出に関しては、本研究ではコート全体ではなくコート半分を撮影した映像を用いたために、不完全なイベントが多数存在したが、認識できる範囲ではスローイン以外の約90%のイベントを検出することができた。スローインは50%程度となった。これは、イベント検出の手法としてボールが静止している時間を用いているために、ボールを置いてすぐにプレイを始める場合などには対応できないからである。また、スローインと同定できるボールの座標範囲が狭いことにも理由として考えられる。

## 6. まとめ

本研究では、デジタルカメラワークを用いることにより、サッカー映像を仮想的に自動撮影する方法として、ボールを中心に、映像生成のルールを用いて状

況認識を行う方法に加え、視聴者があらかじめ分類された嗜好を選択することによって、それに応えて高い評価を得られるという手法を提案した。

画質の評価が低かった点に関しても、現在ソフトウェア的に画質劣化を防ぐ技術を研究している。

現段階で行えていない選手の同定に関して、ボールと同様に、高い追跡性を持った技術を適用することにより、今回の結果では満足できない視聴嗜好にも応えられると考えられる。またこれらが実現することにより、フォーメーション解析なども可能となるために、サッカー教育等にも用いることが可能となる。

また、現在複数カメラを用いたスイッチング技術の研究も行っている。現在は一つの視点からの撮影を行っているが、TV映像をより模倣するためには個別の選手や、ゴール前のズーム映像などは欠かせない。これが実現すれば視聴者の評価はより向上すると考えられるため、継続して研究を行う予定である。

## 文 献

- [1] 矢島史, 角谷和俊, 田中克己, “映像上での動きの直接描画によるサッカー映像検索,” 情処学会研究報告-情報学基礎, Vol.2002, No.041, May.2002.
- [2] 新田直子, 馬場口登, “放送型スポーツ映像の意味内容獲得のためのストーリー分割法,” 信学論(D-II), vol.J86-D-II, no.8, pp.1222-1233, Aug.2003.
- [3] 宮内進吾, 馬場口登, 北橋忠宏, “テキスト・音声・画像の協調的処理による放送型スポーツ映像におけるハイライト検出とインデクシング,” 信学会(D-II), vol.J85-D-II, no.11, pp.1692-1700, Nov.2002.
- [4] 加藤大一郎, “新しい番組制作支援技術知的ロボットカメラと放送番組への応用,” NHK技研R & D, No.48, pp.34-47, 1998.
- [5] 加藤, 山田他, “被写体を追尾撮影時の放送カメラマンのカメラワーク特性分析,” テレビジョン学会誌, Vol.50 Num.12, pp.1941-1948(1996.12).
- [6] 越後富夫, 宮森恒, 飯作俊一, “GMRFとオブジェクトフローを利用したサッカー映像のオブジェクト抽出,” 画像の認識・理解シンポジウム(MIRU'98), TP3-05, 1998.
- [7] 大西正, 輝泉正夫, 福永邦雄, “デジタルカメラワークを用いた自動映像生成,” 画像の認識・理解シンポジウム(MIRU2000), pp.I-331-I-336, Jul.2000.
- [8] 熊野雅仁, 岩本健, 有木康雄, 塚田清志, “ボールと選手に着目したデジタルカメラワークの実現法,” 画像の認識・理解シンポジウム(MIRU2004), pp.II-341-II-346, Jul.2004.
- [9] 藤村光, 杉原厚吉, “優勢領域に基づいたスポーツチームワークの定量的評価,” 信学論(D-II), vol.J87-D-II, no.3, pp.818-828, Mar.2004.
- [10] 宮森恒, 越後富夫, 飯作俊一, “短時間動作記述によるスポーツ映像の表現と検索方式の検討,” 知能情報メディアシンポジウム, S-9, 1999.
- [11] 瀧剛志, 長谷川純一, “チームスポーツにおける集団行動解析のための特徴量とその応用,” 信学論(D-II), vol.J81-D-II, no.8, pp.1802-1811, Aug.1998.

- [12] 三須俊彦, 苗村昌秀, 境田慎一, 鄭文涛, 金次保明, “ 複数情報の融合によるサッカー選手のロバストな追跡法, ” 信学技報, PRMU2001-67, 2001.
- [13] 大野義典, 三浦純, 白井良明, “ サッカーゲームにおける選手とボールの追跡, ” CVIM 114-7, 1999.
- [14] 高橋正樹, 三須俊彦, 合志清一, “ オブジェクト抽出技術のスポーツ番組への応用, ” SLDM 111-1, 2003. [15] 前田英作, 高橋裕子, 石井健一郎, “ 環境変動にロバストな物体有無判定法, ” 信学会 (D-II), vol.J74-D-II, no.12, pp.1731-1740, Dec.1991.
- [15] 田中健二, 鈴木健治, 佐藤正人, 荒川佳樹, “ 高精度映像(WHD :Wide/Double HD) 伝送システム, ” 信学会 (D-II), vol.J84-D-II, no.6, pp.1094-1101, Jun.2001.
- [16] 松山隆司, 和田俊和, 波部齊, 棚橋和也, “ 照明に頑健な背景差分, ” 信学会(D-II), Vol.J84-D-II, no.10, pp.2201-2211, Oct.2001.
- [17] 境田慎一, 苗村昌秀, 金次保明, “ 背景差分法と字空間 watershed による領域性長方を併用した動画画像オブジェクトの抽出, ” 信学会 (D-II), Vol.J84-D-II, no.12, pp.2541-2555, Dec.2001.
- [18] 土田勝, 川西隆仁, 村瀬洋, 高木茂, “ 背景差分法による物体検出を目的とした逐次モンテカル口法による背景推定, ” 信学会(D-II), Vol.J87-D-II, no.5, pp.1062-1070, May.2004.
- [19] 松本圭介, 須藤智, 斎藤英雄, 小沢慎治, “ サッカーシーンにおけるボール追跡に基づく最適視点決定システム, ” 信学技報, PRMU2000-06, Jun.2000.
- [20] <http://vision.kuee.kyoto-u.ac.jp/lecture/dsp/kadai/2003/report3.html>
- [21] 片山 禎昭, “ AHP によるシステム評価, ” 情報研報告, IPSJ, Oct.1990.
- [22] <http://www.isc.senshu-u.ac.jp/thc0456/AHP/>