

J-012

## 個人適応型サッカー映像の自動生成技術

Automatic production Method with Personal Adaptation for soccer-game videos

窪田 進太郎  
Shintaro Kubota有木 康雄†  
Yasuo Ariki塚田 清志‡  
Kiyoshi Tsukada

## 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, 地上波デジタル放送などの技術により、チャンネル数の増大による放送コンテンツ数の増大、データ通信による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、制作費の増大につながる。コスト面のほかにも、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。

これまで我々は、この状態を解決するための技術として、アマチュアスポーツのデジタルカメラワークを用いた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、各フレームをデジタル処理して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る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。従来のような一方の放送ではなく、視聴者の意思や好みを反映した放送が重視されるようになってきている。

これを実現するための技術として、個人適応型サッカー映像の自動生成技術を提案する。これまでに我々が研究してきたデジタルカメラワークによる映像の自動生成技術では、プロのカメラマンの撮影技術を模倣するようにして、すなわち、誰にでも受け入れられ易いような映像を目指していたものであったが、本稿ではそれに加え、映像生成時に視聴者がいくつかの項目を設定することにより、視聴者の好みに合った映像も生成するという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ている。高解像度の映像を解析し、SD画質の映像を切り出す。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、本手法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解

† 神戸大学, Kobe Univ.

‡ 毎日放送, MBS

像度の高いカメラが開発される可能性があることと、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に実際のデジタルカメラワークの様子を示す。以後に述べる3種類のクリッピングサイズを用いている。

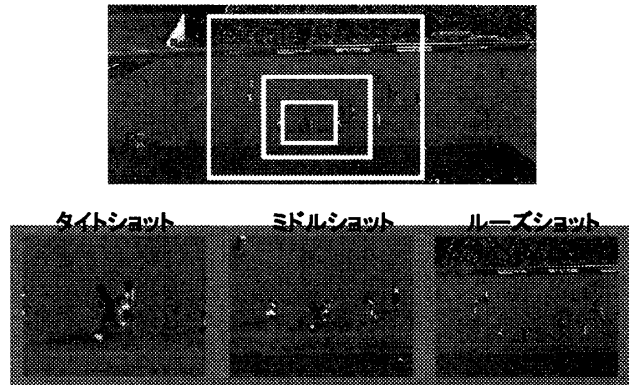


図1 デジタルカメラワークによるクリッピング

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(スローイン・フリーキックなど)の検出を行う。最後にカメラワーク部で視聴者から入力されたパラメータを考慮し、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やすさを考え、3章でカメラワーク、4章で状況認識について述べる。

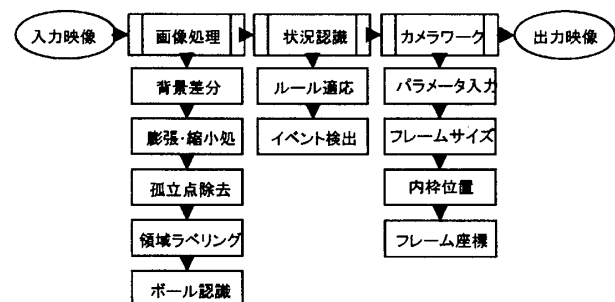


図2 提案手法における処理の流れ

## 3. カメラワーク

状況認識部から渡された各フレームごとのパラメータと、入力された視聴者のパラメータを用いて、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解像度と位置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 パラメータ入力

予め質問項目をいくつか用意しておき、それに対して視聴者が答えるという形式を用いた。質問項目としては、ズーム速度、ズーム処理、画面サイズ、イベント評価の四つの項目がある。

各項目の説明内容に関しては順次述べる。

ズーム速度は、ズームアウト、ズームイン時におけるズームの変化速度を五段階で評価するようになっている。一段階ごとにズーム処理の時間が20%変化している。

ズーム処理は、表1にしめすどのクリッピングサイズの使用頻度を高くするかということを決める項目で、質問項目は5段階で答えるようになっており、そのほかに三つのチェックボックスがついている。三つのクリッピングサイズそれぞれで固定することができるようにするためである。各段階で、クリッピングサイズ変化の閾値が20%ずつ変化する。クリッピングサイズの頻度が変われば、ボールをとらえる時間も変わるために、パンの速度も変わる。

画面サイズは、表XXで示している初期設定の3種類の解像度を変化させる評価で、5段階評価となっている。一段階ごとに各解像度が10%ずつ変化していく。

イベント評価は、イベント検出時にタイトショットを挿入し、イベントを強調するが、それを行うか行わないかということと、イベント検出時にイベント名を文字として出すか、出さないかというチェックがある。

### 3.2 クリッピングサイズ

TV映像を解析し結果、3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。本稿では表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のように遷移させる。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。

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

表1 ズーム処理における各クリッピングサイズ

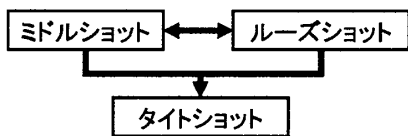


図3 クリッピングサイズの変化

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パラメータ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、あまり短すぎると見にくい映像になってしまうために、最低2秒以上は継続するように設定している。

### 3.3 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捕らわれたり、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、人間による撮影に近づけるために、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取る。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、後述するが、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をまとめて行い、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している。

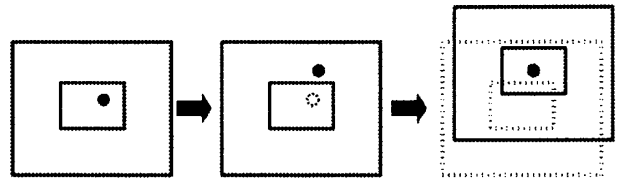


図4 ボールの細かい動きに追従しない内枠

### 3.4 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、ドリブルで前に素早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かを見るために、内枠を進行方向と逆に移動させる。

## 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はどのようなシーン、すなわちどのようなカメラワークでみたい場面なのかを判別することである。本紙では個人の嗜好を反映させることができるが、この段階では一般的な嗜好での認識を行う。ここでいう一般的とは、TVで放送されるような、という意味である。テレビ映像から、撮影技術をルール化した。

### 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となるルールは、基本的にボールの座標を中心に考える。

基本的には、ボールの座標がゴールに近づいた時と、ボールがあまり動かないときをズームイン、すなわちミドルショット、その逆のときにズームアウト、すなわちルーズショットとする。ゴール前でボールが素早く動いたときや、ピッチ中央でゴールがあまり動かないときなどの、条件に優先順位をつけているために、それぞれ適切なクリッピングサイズが得られる。ボールが極端に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、本手法ではボールの微小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、素早い攻撃展開によりゲームの関心が移動していくときに、ボールを中心にと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の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5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、ペナルティキックである。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、ズームインへの変遷と同時に起こる。サッカー映像の解析から、平均的にこの状態が6秒以上続いた時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のルールを以下の表XXに示す。

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

表2 イベント認識ルール

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。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固定映像の合計3種類を比較の対象とした。

そして、それに加えて、一般的な嗜好で生成された映像と、個人嗜好を反映した映像を見比べて評価を行った。

5.2 評価基準

映像の評価としては表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が評価できないが、それも映像の評価として判断する。

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をするためには、単純に各判断基準のスコアを合計するだけで計ることはできず、視聴者の嗜好により

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

表3 サッカー映像の評価基準

項目の重要度が異なることが多い。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、正確な比較が可能となる。また、デジタルカメラワークの生成において、将来的には各家庭の嗜好を反映した映像を、HDDレコーダーなどによりデジタル処理して生成するという考えられるため、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の各値の計算を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。

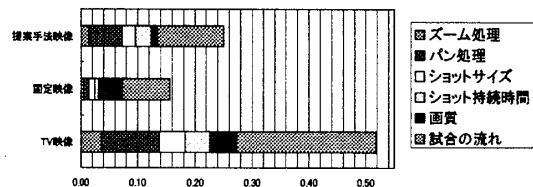


図5 AHP法による評価結果

次に、一般的な評価基準で生成された映像と、各被験者がパラメータ入力を行っ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ところ、7人の被験者のうち、6人が変更後、1人が変更前という結果になった。しかし、この1人も、パラメータ入力を複数回行うことによって、

自分の好みの映像を生成することができた。一般的な嗜好に比べて、高い評価を得ることが確かめられた。

また、本紙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ーを始める場合などには対応できないからである。このことに関しても更なる研究が必要である。

## 6. まとめ

本紙では、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、視聴者がパラメータを入力することによってより視聴者の嗜好という点に関して高い評価が得られるという手法を提案した。現段階では選手などの同定を行うことはできていないが、それができれば、パラメータ入力の際に、特定の選手を多く撮影するということが可能になる。また、パラメータ入力の項目も現段階では4つと少ないので、より詳細な設定を目標とする。

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。

カメラ一つでの固定映像撮影を行っているが、これを複数にしてスイッチングを行うという考えもある。選手とボールを完全に追跡することができれば、デジタルカメラワークは、短時間の遅延さえ許せば人間の撮影表現を遙かに上回ることが可能であると考えられる。2006年のワールドカップでは、ボールにチップを埋め込み、ラインを割ったかどうかの判断を行う技術が使われる。これを応用すれば選手とボールの追跡をすることも可能であると考えられるため、本研究を継続していきたい。

### 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 とオプティカル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.1-331-1-336, 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.11-341-11-346, 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.

[16] 田中健二, 鈴木健治, 佐藤正人, 荒川佳樹, “高精細度映像(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.

[17] 松山隆司, 和田俊和, 波部斉, 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.

[18] 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.

[19] 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ロ法による背景推定,” 信学会(D-II), Vol.J87-D-II, no.5, pp.1062-1070, May.2004.

[20] 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.

[21] <http://vision.kuee.kyoto-u.ac.jp/lecture/dsp/kadai/2003/report3.html>

[22] 片山 禎昭, “AHP によるシステム評価,” 情処研報告, IPSJ, Oct.1990.

[23] <http://www.isc.senshu-u.ac.jp/thc0456/AHP/>