

論文

サッカーシーンにおける選手視点映像提示のための リアルタイム選手軌跡獲得手法

糟 谷 望 北 原 格 亀 田 能 成
 大 田 友 一

サッカーシーンにおける選手視点映像提示のための リアルタイム選手軌跡獲得手法

糟谷 望[†]北原 格^{†,††}(正会員)亀田能成^{†,††}(正会員)大田友一^{†,††}(正会員)[†] 筑波大学 大学院システム情報研究科^{††} 筑波大学 計算科学研究センター

〈あらまし〉 サッカーシーンにおける自由視点映像の提示法として、操作が容易で臨場感の高い「選手視点映像」を提案し、その生成に必要な、選手の移動軌跡を推定する手法について述べる。高所に設置した複数台の固定カメラで撮影した映像から抽出した影を含む前景領域をフィールドに射影し、影と選手の写像領域を用いて選手の位置を推定することにより、サッカーシーンで頻発する選手同士のオクルージョンの問題を軽減する。ユニフォームの色情報を用いて選手領域をチームごとにクラス分けすることにより、映像中にクロスプレーのような複雑な見え方が発生した場合でも、追跡の継続を可能とする。実際の競技場においてサッカーの試合を撮影した映像を用いて提案手法の有効性を検証し、選手視点映像を生成した。

キーワード：自由視点映像，選手視点映像，物体追跡

〈Summary〉 This paper proposes Player's View as a new presentation method of Free Viewpoint Video, and a method to estimate trajectories of soccer players for generating Player's View. We estimate the trajectories of soccer players by using multiple cameras set in high place to reduce occlusion. Shadow and Player regions extracted from the images are used for player's position estimation. Furthermore, our method utilizes color information of the players' uniforms to improve the accuracy of object tracking. We have applied our proposed method to a real soccer game and demonstrated its effectiveness. Player's View was generated from the trajectories successfully.

Key words: Free Viewpoint Video, Player's View, Object Tracking

1. はじめに

近年、複数の視点から撮影した映像をコンピュータで解析することにより、任意の視点からの見え方を生成する自由視点映像技術の研究^{1)~5)} が盛んに行われている。特に、スポーツ中継は、ユーザごとに注目している対象が異なるため、自由視点映像技術を用いた観戦方式の適

用に注目が集まっている^{6),7)}。しかし、従来の自由視点映像提示方式では、2自由度の入力装置であるマウスを用いて、自由度が6の自由視点映像を撮影する仮想カメラを操作するため、見たい視点を即座に指定することが困難である。また、カメラ操作をしながらでは映像鑑賞に集中できないといった問題が存在する。我々は、これらの問題の解決を目的として、選手の視点から試合を眺めたような“選手視点映像”の生成・提示に関する研究に取り組んでいる。選手視点映像では、見たい選手を選択するだけという簡易な入力操作により、臨場感のある映像を楽しむことができる。

“Online Trajectory Estimation of Soccer Players for Generating Player's View”

Nozomu KASUYA[†], Itaru KITAHARA^{†,††} (Member), Yoshinari KAMEDA^{†,††} (Member), Yuichi OHTA^{†,††} (Member)

[†] Graduate School of Systems and Information Engineering, University of Tsukuba

^{††} Center of Computational Sciences, University of Tsukuba

2. 選手視点映像

本稿で提案する選手視点映像とは、図1に示すように、選手の視点位置における映像である。選手視点映像では、観察視点位置は選手の存在するような地点に限定されるが、ユーザに要求される入力処理は、見たい選手を選択するといった簡易な操作であり、操作に気をとられることなく試合観戦に集中できる。また、競技が行われている空間に入り込んだ位置からの見え方が提示されるため、従来の映像にはない臨場感を得ることが、可能になると考えられる。

通常の映像撮影装置を用いて選手視点映像を撮影するためには、選手自身にカメラを取り付ける必要があるが、選手に負担をかけプレーの妨げになるため現実的ではない。本研究では、実際にはカメラが存在しない視点からの映像を仮想的に再現可能であるという、自由視点映像⁷⁾の特長を利用することにより、選手視点映像を生成する。

図2に選手視点映像生成のためのシステム構成を示す。まず、複数台の選手位置推定カメラを用いて選手のフィールド上での位置を推定し、同一選手の位置情報を時系列に対応付けることにより、選手軌跡情報を獲得する。次に、推定した選手位置を用いて三次元モデルを生成する。推定した選手位置に選手の大きさを想定したビルボードを配置し、同時に、テクスチャ獲得カメラで撮影した画

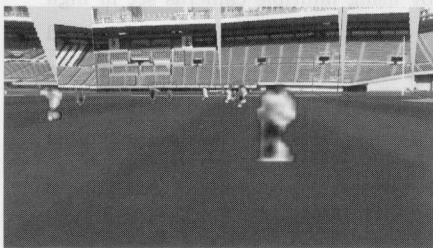


図1 選手視点映像の例
Fig. 1 Example of player's view

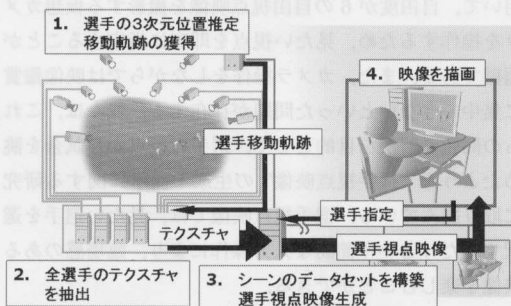


図2 選手視点映像生成システム
Fig. 2 Player's view generation system

像群から選手テクスチャ領域を獲得することにより、選手の三次元モデルを生成する。すべての選手の三次元モデルをフレームごとに束ねて、シーンのデータセットを生成する。映像を観察するユーザが、観察を希望する選手を指定すると、その選手の視点位置に仮想カメラを移動させ、その視点から見た映像の生成・提示を行う。以降のフレームでは、本稿で提案する手法によって獲得される選手の軌跡情報を用いて仮想視点を移動させ、選手視点映像の生成・提示を継続する。

3. サッカーシーンにおける選手軌跡獲得

サッカーの試合における選手追跡処理の特徴としては、フィールド全体をカバーした広角レンズによる撮影を行うため、各サッカー選手の観測サイズが小さいこと、コンタクトスポーツであるため、選手が重なったり交差したりするシーンが頻繁に発生すること、サッカーの試合は屋外で開催されることが多いため、日照変化が発生することなどがある。また、自由視点・選手視点映像の生成・提示をサッカー中継に適用することを考えた場合、実時間での処理が必要となる。本章では、これらの問題に対する、我々の取り組みについて述べる。

3.1 関連研究

これまでにも、映像中の動物体を追跡し軌跡情報を獲得する研究が数多く行われている。

テンプレートマッチング^{8)~10)}を用いてサッカーシーンのように多数の物体を同時に追跡する場合、計算コストが増大するため実時間追跡が困難となる。また、物体の見え方に強く依存するため、選手の接触や交錯によって発生するオクルージョンに弱いといった問題が存在する。

オクルージョンに対しロバストで、かつ実時間追跡を実現した手法が提案されている^{11),12)}が、これらの手法は、十分高い解像度で選手領域が撮影されている必要があり、太陽光によって強く影の出るシーンは対象としていない。我々が撮影対象とするサッカーの試合の映像では、サッカー選手の観測サイズは小さく、大きな日照変化も不可避であるため、適用が困難である。

サッカー選手の追跡に関する研究も多数行われている^{13)~16)}。しかし、「選手の影が強くでるシーンに対応していない」、「高解像度の映像が必要である」、「処理に時間がかかる」など、上述した問題をすべて解決している手法は提案されていない。

3.2 提案する選手追跡手法

本研究では、上述のような比較的難しい条件において撮影せざるを得ないサッカーシーンの映像から、リアルタイムで選手軌跡の獲得を行う手法を提案する。処理の大まかな流れを図3に示す。撮影には2台の固定カメラ

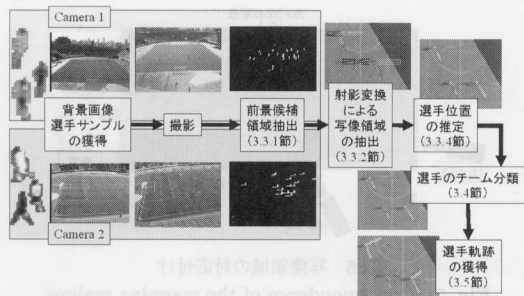


図3 選手軌跡獲得の流れ

Fig. 3 Players' trajectories estimation

を用いる。各カメラにおいてあらかじめ背景画像と選手の見え方のサンプルデータを取得しておく。撮影された映像に対して背景差分法を適用し、選手候補領域を抽出する。抽出した選手候補領域を二次元射影変換によってサッカーフィールド面上に投影する。複数台のカメラ映像からの画像を用いて選手位置を推定することにより、ある選手についてオクルージョンが発生した場合でも、他のカメラで撮影した映像を用いて、対象選手を見失う問題を軽減することができる。更に、ユニフォーム色を用いて検出領域をチーム分類することにより、選手が交錯するような場合に発生しやすい誤追跡を削減する。最後に、選手の位置とチーム分類結果を用いてフレーム間の対応をとることにより、選手の軌跡を獲得する。

3.3 選手位置推定

本手法の特徴の一つは、従来の物体追跡では誤差の原因とされてきた選手の影領域について、これを除去するのではなく、逆に選手位置推定のための情報として活用することにより、選手位置推定処理の安定化を実現することである。

図4に、2台のカメラを用いた場合を例にとりて、選手位置推定の流れを示す。前述した撮影環境のため、抽出された前景候補領域には選手領域とその選手の影領域が含まれていると考えられる。これらの領域をフィールド面に射影し、選手領域と影領域からなる複数の画像を得る（以降、この画像を投影画像と呼ぶ）。これらの画像の主軸の交点を最小2乗法で求め、選手がフィールド上に立っている位置を推定する。

3.3.1 前景候補領域の抽出

本研究で用いるカメラは、撮影環境に固定されているため、背景差分法によって前景候補領域を抽出することができる。選手は、通常フィールド上にのみ存在するので、フィールド以外の領域にマスクをかけることにより、観客席領域などにおける誤検出を除去する。日照変動の影響は主としてY成分のみに表れるため、YUV色空間

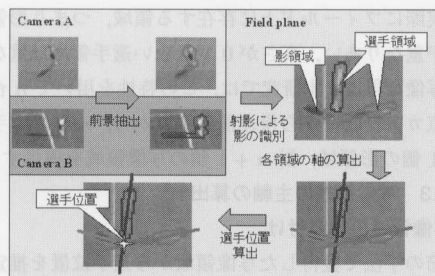


図4 選手位置推定処理

Fig. 4 Players' position estimation

を用いることにより、RGB色空間に比べ安定した前景抽出が可能となる。更に、画像中の選手領域の大きさは選手の身長とカメラパラメータを用いてある程度予測できることから、前景候補領域の面積に対してしきい値処理を行い、ノイズやボール領域を除去する。この際のしきい値は選手の一般的な体の大きさと撮影カメラの射影変換行列から算出する。

3.3.2 射影変換による画像領域の獲得

前節で抽出した前景領域には、選手領域のほかに、その選手の影領域も一緒に検出される。影領域を識別する手法としてHSV色空間を用いた識別法¹⁷⁾が提案されているが、彩度の低い黒いユニフォームを着た選手を影と識別してしまう問題や、解像度が低いと画素がつぶれ、周囲の色の影響により正しく識別できないといった問題が存在する。そこで本研究では、影は地面上に張り付いているという幾何的特性を利用し、従来の選手位置推定手法ではノイズとして取り扱われていた影領域を活用することにより、選手の三次元位置の推定精度を向上させる手法を提案する。なお、この推定選手位置は、地面上に拘束 ($Z=0$) して推定した選手の足元位置である。

空間中にある三次元点 (X_w, Y_w, Z_w) と、その点が画像上で観測される位置 (u, v) との関係は、3行4列の射影行列 P で表すことができる。ここで、Z軸の正方向を鉛直上向きとし、撮影空間がすべてフィールド面上に存在すると仮定すると、Z成分が0となり P の3列目を無視することができる。その結果、以下の式 (1) のように、フィールド面上の点 ($X_w, Y_w, 0$) とその画像上での観測点 (u, v) との関係は二次元射影変換を用いて表すことができる。ここで p_{ij} は射影行列 P の i 行 j 列の要素とする。

$$\begin{bmatrix} \lambda u \\ \lambda v \\ \lambda \end{bmatrix} = \begin{bmatrix} p_{11} & p_{12} & p_{14} \\ p_{21} & p_{22} & p_{24} \\ p_{31} & p_{32} & p_{34} \end{bmatrix} \begin{bmatrix} X_w \\ Y_w \\ 1 \end{bmatrix} \quad (1)$$

このようにして求めた二次元射影変換の逆変換を用いて複数視点で撮影された画像をフィールド面上に投影する

と、実際にフィールド上に存在する領域、つまり影領域のみが重なりあい、高さが 0 ではない選手領域は重ならない画像を得る。本研究では、この特性を用いて n 台の多視点カメラ映像から得た前景領域から、 n 個の選手領域と 1 個の影領域、計 $n + 1$ 個の画像領域を獲得する。

3.3.3 画像領域の主軸の算出

(i) 画像領域の対応付け

前節の手法で獲得した画像領域から選手位置を推定するためには、同一選手によって発生する画像領域を対応付ける必要がある。獲得した影領域に膨張処理を施し、射影誤差による影領域のずれを吸収する。その後、影領域、カメラごとの投影選手領域ごとにラベリング処理を行うことで連結領域を獲得する。この連結領域の位置関係を元に画像領域の対応付けを行う。

カメラとフィールド、光源とフィールドの位置関係は固定であると仮定できるため、画像領域の方向から各々の画像を判別できる。異なる種類の画像領域の組み合わせを求めることにより、図 5 の点線のように画像領域を対応付ける。同一選手に対応する画像領域は互いに隣接し、カメラ・光源の位置関係が図 5 に示すような場合、カメラ 1 の画像に対してカメラ 2 の画像は左上に、影領域は左に位置し、カメラ 2 の画像に対して影領域は下に位置するという情報を元に対応付けする。

図 6 のように一方のカメラで隠れが起きている場合にも、他方のカメラの画像と影領域の対応を、位置関係と隣接状況をもとにとることができる。画像 a に隣接している領域は、影領域である画像 c であり、これが画像 a の下にあることから選手 A のものと推定する。画像 a にはほかに隣接している領域がないため、カメラ 1 画像中で選手 A は隠れが発生していると推測できる。そのような場合には対応の得られた二つの領域から選手位置を推定する。また、画像 b に対しては、影領域である画像 c が下に、カメラ 1 の画像である画像 d が右下に隣接しており、これらが選手 B に対応するものとして対応付けることができる。

(ii) 画像領域の主軸の算出

選手領域と影領域の主軸方向を、投影画像上における画像処理により算出すると、領域の面積が小さい場合、推定精度が低下する可能性がある。本手法では、カメラとフィールドの幾何的な関係が固定であることを利用して、カメラの射影行列を用いてフィールド上の位置に対応する選手領域の主軸方向をあらかじめ計算しておく。また、影領域の主軸方向は、日時・緯度経度の情報を用いて計算することができる。

選手はフィールドに直立していると仮定すると、選手の画像領域の主軸は、選手の三次元位置に垂直に立てた

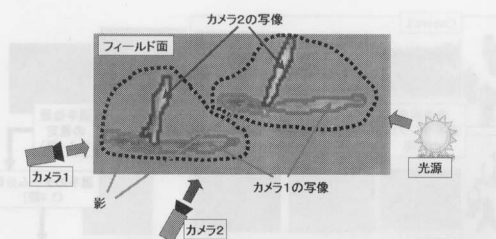


図 5 画像領域の対応付け

Fig. 5 Correspondence of the mapping regions

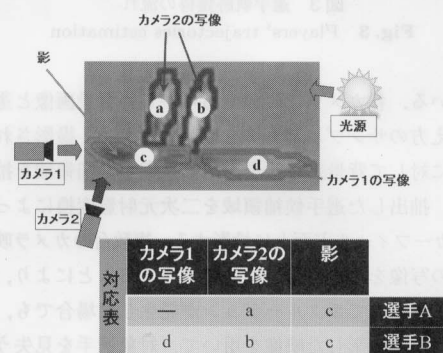


図 6 隠れが存在する場合の画像領域の対応付け

Fig. 6 Correspondence of the mapping regions in the case of occurring occlusions

直線を投影画像に投影したものと一致する。このようにして、投影画像上のすべての位置において、選手領域の主軸方向を事前に求めておくことにより、算出処理の高速化を図る。

太陽は、無限遠に存在する点光源と考えることができるため、地面に直立している物体の影はすべて同じ方向を向くと仮定する。また、短時間のシーンでは影領域の主軸の方向の変化は微小であるため、本手法では、影の方向を一定と仮定することとする。具体的には、1 フレーム目における影領域の主軸角度の平均を影領域の主軸方向として利用する。

この時点では選手の位置は未知であるため、各投影領域の重心位置において主軸の傾きを求め、その値を各領域の傾きとして採用する。正確な位置推定を行うためには、暫定傾き値を用いて推定した選手位置において主軸の傾き値を再度算出する処理を再帰的に実行する必要があるが、投影画像中の選手領域がそれ程大きくないため、再起処理によって求まる選手位置と重心位置が近接していること、近隣領域では主軸角の変化が小さいことから、実時間処理の実現を優先し再帰的处理を省略している。

3.3.4 投影画像を用いた選手位置の推定

カメラを 2 台以上使用すれば、投影画像において同一選手によって発生する複数の画像領域を獲得できるため、

それらの領域の主軸の交点から選手位置を算出することができる。本手法では、更に投影画像上に写像された影領域の情報を利用することにより、より頑健な位置推定を実現する。

前節で求めた同一選手に対応する $n+1$ 個の領域の主軸の交点を算出し、選手位置とする。各領域の傾きを θ_i 、重心位置を (x_{0i}, y_{0i}) とし、各領域について

$$a_i = \sin \theta_i, b_i = \cos \theta_i, a_i x_{0i} + b_i y_{0i} + c_i = 0 \quad (2)$$

となるような直線をあてはめ、直線の交点を

$$\sum_i^{n+1} (a_i x + b_i y + c_i)^2 \quad (3)$$

が最小となるような点 (x, y) を求め、この点を選手の足元位置とする。最小 2 乗法を用いて交点の算出を行うことにより、複数の領域がほぼ平行となるような場合や、曇天で影領域が発生しない場合、オクルージョンによって抽出される領域数が減った場合にも、同じ計算で足元位置の推定が可能になる。なお、ナイターの試合のように人工照明が複数あるような場合には、太陽光のように人物の影が強くでることはなく、前景領域抽出の際に影領域は削除されるため、曇天時と同様の処理で対応可能である。

3.4 選手のチーム分類

追跡物体同士が交錯したり、重なったりするような場合、誤った物体を追跡してしまうことがある。また、一度重なった状態で一緒に追跡物体が移動し、その後、物体が離れたときに、どちらの物体を追跡していたのかを判断することは困難である。我々は、サッカーでは、プレー中に選手同士が接触することがよくあるが、そのように近接した状態でプレーするのは、異なるチームの選手である場合が多いことに注目し、追跡領域をチームごとに分類しておくことで、物体同士の距離が近い場合においても、安定した追跡を実現する。具体的には、サッカーの試合では、選手はチームごとに決まった色のユニフォームを着るというルールを利用して、クラスタリング処理によって映像中の選手領域のチーム分類を行う。なお、追跡処理時間の高速化を実現するため、計算コストの大きい学習プロセスを 1 度だけ実行することとし、それ以降では、初期学習に基づいたチーム分類を行う。

3.4.1 学習プロセス

サッカーではユニフォームの色 (クラス数) が数種類に限定されるため、学習には非階層的手法の k-means 法を用いる¹⁸⁾。k-means 法は、初期種子点によって大きく結果が異なってしまうことがあるが、本システムでは、事前に取得したユニフォーム色を初期種子点として与える

ことにより適当な結果を得る。学習データは初期の 30 フレームから抽出した全選手領域のユニフォーム色である。ユニフォームの色は抽出された選手領域のうち、ユニフォームの領域であることが予想される上部 20~50% の領域の平均色を、その選手のユニフォーム色とした。色空間としては、色の違いを比較するのに適している $L^*a^*b^*$ 均等色空間を用いることにより、クラスタリング精度の向上を図る¹⁸⁾。

3.4.2 クラスタリング

1 フレームごとに選手の色情報データを追加し、k-means 法によって代表点を更新する処理を実行するためには、膨大な処理コストが必要となる。そこで、学習プロセス終了後は代表点の更新を行わず、学習プロセスで得た各クラスの代表点からの距離が最も近いクラスへの分類処理のみを行う。この際、映像中で観測されるユニフォーム色は、各々分布が異なるため、距離の判定にはマハラノビス距離を用いる。オクルージョンなどによって選手のユニフォーム色が適切に抽出できない場合には、間違ったクラスに分類される可能性が高い。チーム分類に誤りがあると追跡の際に別の選手であると判断してしまうという問題が生じるため、マハラノビス距離が大きい場合には、クラス分類をせず、未確定クラスとして判別処理を行う。

3.5 選手軌跡の獲得

前述した手法で得た選手位置のフレーム間における対応関係を求めることにより、選手軌跡情報を獲得する。その際、チーム分類の結果を用いることにより誤追跡を防ぐ。また、途切れてしまった同一選手の軌跡を軌跡間の対応を求めることにより結合する手法について述べる。

3.5.1 フレーム間の対応関係の取得

チーム分類結果を用いて各フレームの選手位置のフレーム間での対応関係を求めることにより、選手軌跡情報を獲得する。図 7 にフレーム間の対応の取り方の例を示す。今回使用した映像は毎秒 30 枚のフレームレートで撮影されているため、フレーム間で選手がそれほど大きく移動

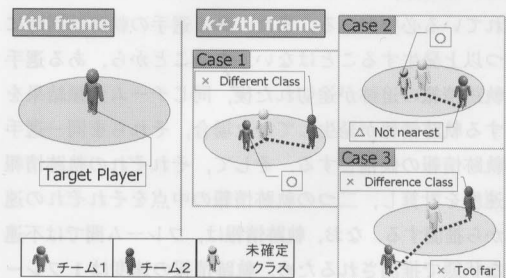


図 7 フレーム間の対応関係

Fig. 7 Correspondence between frames

しないと仮定し、図 7 の Case 1, Case 2 のように、追跡している選手とチーム分類結果が同じ、または未確定となっている選手のうち、前のフレームで観測された位置から最も近く、かつ移動距離がしきい値以下の位置にいる選手を、対応する選手として選択する。ただし、追跡中の選手のチーム分類結果が未確定となっている場合には、クラスに関係なく最も近い選手を選択する。そして、新たなフレームでクラスが確定している選手が選択された場合には、その時点で追跡中の選手のクラスを確定する。図 7 の Case 3 のように、最も近い同一クラスの選手までの距離がしきい値を超える場合には、背景差分処理のエラーにより選手検出漏れが発生したと判断し、そのフレームでは選手を見失ったものとし、次のフレームで追跡選手を探索する。その際、追跡選手は見失っている間にも移動していることが予想されるので、しきい値を広げて探索する。これを数フレーム繰り返すことにより、見失った選手を再捕捉する。再捕捉できた時点で、それ以前の見失っていたフレームでの位置を、再捕捉した位置と見失った位置の線形補間によって求める。

サッカーは広いフィールドを一杯に活用する競技であるため、同一チームの選手同士が、複数台のカメラ映像を用いても分割できないほどの近距離で長時間プレイするケースが多発しないこと、選手軌跡推定処理用の映像は、サッカースタジアムの電光掲示板やスタンド屋上のような比較的高所から撮影されているため、映像中における選手同士の隠れが発生しにくいことから、3.3 節の手法で十分に対応できる。しかし、コーナーキックなどのセットプレイ時には同一チームの選手同士による隠れが問題となることが予想されるため、今後は Ohno¹⁹⁾らのように隠れの前後関係や移動速度から予測を行うことでこれを解決する必要がある。

3.5.2 同一選手の軌跡の結合

フレーム間の対応をとるだけでは、対応取得に失敗して同一選手の軌跡が分離してしまうことがある。そこで、分離した同一選手の軌跡を、移動速度情報を用いて結合する。その際、同一選手の軌跡は同じチームとして分類されている必要がある。また、ある選手の軌跡は同時に二つ以上発生することはない。このことから、ある選手の軌跡情報の追跡が途切れた後、同じチーム分類結果を有する軌跡情報が発生していた場合、それらを同一選手の軌跡情報の候補とする。そして、それぞれの軌跡情報の速度を計算し、二つの軌跡情報の中点をそれぞれの速度から推測する。なお、軌跡情報は、フレーム間では不連続な状態で推測されるため、軌跡情報の速度は 1 フレームごとではなく、10 フレーム分の平均速度とした。推測した 2 点が十分に近い場合には、それらは同一選手の軌

跡情報であると判断し、結合する。今回このしきい値は 0.2m 刻みで増やしていき、誤りなく結合できた最大の値である 1m とした。しかし、他のシーンに本手法を適応させた場合にはこのしきい値は必ずしも適切であるとはいえない。そのため今後はしきい値の適切な値を調査する必要がある。

4. 実験

4.1 実験環境

本手法の有効性を検証するために、国立霞ヶ丘競技場において実証実験を行った。実験では、ピッチの半面を撮影対象とし、2 台のカメラ (カメラ 1, カメラ 2) を図 8 に示すように、電光掲示板とメインスタンド屋根上の比較的高い位置に配置した。カメラはプログレッシブ撮影可能な SONY 製の DXC-9000 を使用し、GPS からの信号を用いて同期撮影を行っている。撮影画像のサイズは VGA (640 画素 × 480 画素)、フレームレートは 30 fps である。撮影画像中の各画素と三次元空間の座標を対応させるため、事前に三次元測量機を用いてカメラキャリブレーションを行った。

撮影対象は第 56 回全日本大学サッカー選手権大会決勝、法政大学対早稲田大学の試合中の、日中で影の強く出ている 30 秒間程度のシーンである。今回使用したシーンではキーパーのユニフォームと白いユニフォームとの区別がつかなかったため、チーム分類は白のユニフォームのチーム、オレンジのユニフォームのチーム、黒のユニフォームの審判の三つのチームに分類した。

4.2 選手軌跡獲得実験

4.2.1 選手検出・チーム分類結果

表 1 に影領域を用いた場合と、用いなかった場合の選手検出結果を示す。影領域を用いなかった場合の位置推定は投影選手領域の二つの主軸の交点を選手位置として求めた。全選手数は二つのカメラで重なる撮影範囲にいる選手数の全 600 フレームでの合計、検出選手数は、投

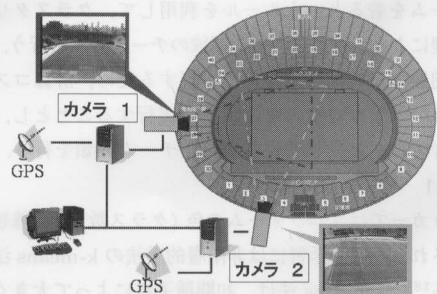


図 8 実験環境

Fig. 8 Experimental environment

表 1 選手検出精度

Table 1 Accuracy of player detection

	全 選手数	検出 選手数	誤検出 選手数	FN 率	FP 率
提案手法	9859	8321	294	15.60%	2.98%
選手領域のみ	9859	7231	213	26.66%	2.16%

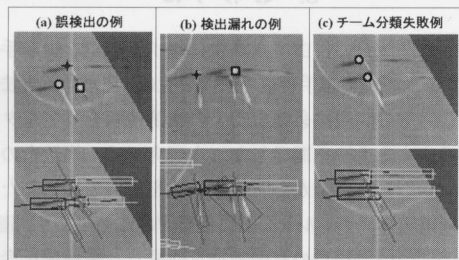


図 9 検出・分類失敗例

Fig. 9 Example of failure

影画像上で 2 ピクセル円が重なっていた場合、つまり 3 ピクセル以内の誤差で位置推定が行われたものを検出選手としたときの数である。誤検出選手数は存在しないはずの位置に選手を検出してしまった数である。また、FN (False Negative) 率は存在するはずの選手を検出できなかった割合を示し、FP (False Positive) 率は誤検出率で、それぞれ以下の式で算出した。

$$\text{FN 率} = \frac{\text{全選手数} - \text{検出選手数}}{\text{全選手数}} \times 100 \quad (4)$$

$$\text{FP 率} = \frac{\text{誤検出選手数}}{\text{全選手数}} \times 100 \quad (5)$$

影領域を利用しなかった場合には、選手が画像中で重なってしまった場合などに選手が検出されないという見失いが多く発生した。このことから提案手法はオクルージョンに対して頑健であることがいえる。提案手法で誤検出された選手のうち多くは図 9 (a) に示すように、ほかの選手の射影領域によって選手領域が分割されたために生じている。これらの領域は、軌跡獲得時における統合処理により、除去することが可能である。また、本システムでは領域内の選手の数 considerando していないため、図 9 (b) に示すように一つの領域に複数の選手が入り込むケースが発生する。その結果、フィールド全体で検出される選手の数在实际よりも少ない False Negative が発生する。この誤差は、領域内に含まれる選手数をフレーム間の対応によって推測することにより減少させることができると考えられる。

表 2 にチーム分類結果を示す。正解の確定クラスへの分類確率 (分類成功率) は 91.89% で、間違っただけの確定クラスへの分類確率 (分類失敗率) は 3.17% であった。未確定クラスに分類される選手が多いが、フレーム間の対応

表 2 チーム分類結果

Table 2 Results of team classification

	分類結果				合計
	白 クラス	橙 クラス	黒 クラス	未確定 クラス	
白クラス	3309	74	1	128	3512
橙クラス	155	3374	0	195	3724
黒クラス	14	1	422	59	496
複数人抽出	67	115	4	403	589
合計	3545	3564	427	785	8321

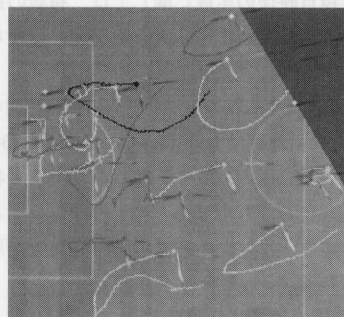


図 10 獲得選手軌跡

Fig. 10 Estimated trajectories of players

表 3 軌跡獲得精度

Table 3 Accuracy of trajectories estimation

		追跡率 [%]	重複 切断数	重複 追跡数	誤差平均 [m]
例	選手 A	100.0	0	0	0.23
	選手 B	91.9	3	0	0.43
	選手 C	88.8	1	17	0.44
平均		87.29	1.79	3.47	0.35
標準偏差		11.36	1.78	4.41	0.08

をとることによりクラスを確定させることができる。一方、間違っただけのクラスに分類された場合、統合処理に悪影響を与えるため、できる限り少なく抑える必要がある。なお、間違っただけの分類が行われるのは、図 9 (c) に示すように、一方のカメラでは選手が隠れているなど、同一選手の領域であることの判断を誤っていることが主な原因である。

4.2.2 軌跡獲得結果

提案手法によって獲得した選手軌跡を図 10 に示す。各線の色はその選手のチーム分類結果を示す。また、手動で与えた軌跡の正解データと獲得軌跡を比較した結果を表 3 に示す。追跡率は、獲得した軌跡のフレーム数を該当選手が正解データ付近にいたフレーム数で割ったものである。切断数は軌跡が途中で切れている数、重複追跡数は、同じ選手を同時に複数の選手として追跡してしまった数、誤差平均は正解データと適合がとれた軌跡との誤

差の平均値である。選手 A のようにほぼ全フレームで追跡できた選手は 18 人中 5 人であった。選手 B のように切断の多くは他の選手とのオクルージョンの影響である。切断がおこると映像生成ができず、ユーザに再度入力を要求してしまうため、更に切断を減らすことが望まれる。切断の中には短い区間の切断が多々あり、推定した選手位置にフィルタリングを施すことで精度が向上すると思われる。また、選手 C のように一人の選手を複数の選手として追跡してしまったものが多くみられた。しかし、選手視点映像の生成に関しては、複数の選手として多重に検出されていても、大きな問題にはならない。

4.2.3 処理時間

選手軌跡を獲得する際の 1 フレームの処理時間を表 4 に示す。使用した計算機は Intel(R) Core(TM)2 Duo CPU E6550 @ 2.33 GHz, 2.00 GB RAM である。各々の処理がほぼ 30 ms 以下に収まるため、パイプライン処理を行うことにより、オンラインでの実時間処理が可能であると考えられる。

4.3 選手視点映像生成実験

前述した手法で獲得した選手軌跡を用いて選手視点映像を生成する。生成に用いた自由視点映像生成・提示システムは Koyama ら⁷⁾ のシステムを改良したものをを用いた。

生成した選手視点映像を図 11 に示す。サッカーシーンに没入した臨場感の高い映像が生成されていることがわかる。また、仮想視点の操作は周囲を見回す自由度だけであるため、初心者でも容易に扱える。しかし、従来の自由視点映像ではフィールド全体を俯瞰するような視点

点からの映像が生成されることが多かったのに対して、フィールドに入り込んだ近い位置からの映像となる選手視点映像では、選手テクスチャの粗が目立ってしまう。高解像度の選手テクスチャを獲得し、より違和感の少ない選手視点映像の生成が今後の課題である。

5. おわりに

サッカー試合の自由視点映像に対する新たな提示手法として、仮想的に選手視点の映像を生成・提示する方法を提案し、その実現に必要な選手の軌跡情報をリアルタイムで推定する手法を開発した。本研究では、複数台のカメラで撮影した画像から背景差分法を用いて抽出した前景候補領域をフィールド面上に投影することにより、個々の選手について複数の写像領域を作成し、影領域を含め複数の領域の主軸を統合して選手の位置推定を行うことで、安定に位置推定を行うことを可能にした。また、ユニフォームの色情報を用いることにより、選手同士が交錯する場合における誤追跡の問題を改善した。国立霞ヶ丘競技場において、実際のサッカーの試合を撮影した映像を用いて選手追跡実験を行い、提案手法の有効性を検証した。そして、獲得した選手軌跡をもとに選手視点映像を生成した。撮影データの解像度が低く映像から選手の視線方向を推定するのは困難なため、ユーザが視線方向を自由に選択できるように設計した。従来の自由視点映像に比べ、操作が簡易で臨場感のある映像が生成できることを確認した。

今度の課題としては、試合全体をととしての選手の追跡が考えられる。その際に必要となる切断軌跡の結合処理において、グラフカットによる一貫性を利用して軌跡を取得する手法²⁰⁾ などを実時間処理の可能性を検討しつつ導入することにより、撮影シーンの変化に対応した結合処理の実現を考えている。また、各段の処理をパイプラインで処理するシステムを実装し、撮影から追跡までの処理時間を実時間で行うこと、選手視点映像提示インタフェースを構築し、映像提示方式としての操作性に関する評価実験の実施、選手視点映像の画質の向上なども検討すべき課題である。

謝 辞 国立霞ヶ丘競技場で実験をするに当たり協力してくださった、国立スポーツ科学センターの白井克佳氏、産業技術総合研究所の吉川文人氏に深く感謝いたします。

参考文献

- 1) T. Kanade, P. Rander, and P.J. Narayanan: "Virtualized Reality: Constructing Virtual Worlds from Real Scenes", IEEE Multimedia 1997, Vol.4, No.1, pp.34-47 (1997).
- 2) W. Matusik, C. Buehler, R. Rasker, S.J. Gortler,

表 4 選手軌跡獲得処理時間

Table 4 Processing time of trajectory estimation

処理内容	処理時間 [ms]
前景候補抽出	11~12
射影変換	29~30
写像領域抽出	10~12
選手位置推定	0.1~2
選手色抽出とクラスタリング	10~11

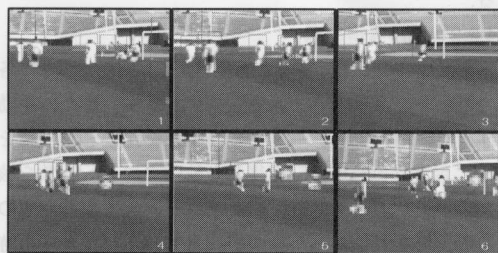


図 11 生成した選手視点映像

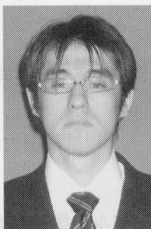
Fig. 11 Generated player's view

- and L. McMillan: "Image-Based Visual Hulls", ACM SIGGRAPH, pp.369-374 (2000).
- 3) J. Carranza, C. Theobalt, M.A. Magnor, and H.-P. Seidel: "Free-Viewpoint Video of Human Actors", ACM Trans. on Graphics, Vol.22, No.3, pp.569-577 (July 2003).
- 4) 谷本正幸: "自由視点テレビ", 三次元映像フォーラム, Vol.15, No.3, pp.17-22 (2001).
- 5) 石川彰夫, 酒澤茂之, 小池 淳: "分割された局所領域での内挿補間による自由視点映像の生成", 信学会技術研究報告, IE, 画像工学, Vol.106, No.155, pp.37-42 (July 2006).
- 6) N. Inamoto and H. Saito: "Immersive Observation of Virtualized Soccer Match at Real Stadium Model", ISMAR, pp.188-197 (2003).
- 7) T. Koyama, I. Kitahara, and Y. Ohta: "Live Mixed-Reality 3D Video in Soccer Stadium", ISMAR, pp.178-187 (2003).
- 8) 安居院猛, 長尾智晴, 斉藤智明, 中嶋正之: "ブロック特徴を用いた歩行人物の抽出と追跡", テレビジョン学誌, Vol.45, No.10, pp.1213-1220 (Oct. 1991).
- 9) 内田恭輔, 三浦 純, 白井良明: "雑踏における歩行者の追跡", PRMU99-209, pp.39-44 (2000).
- 10) 福島健介, 西原明法: "テンプレートマッチングを用いたリアルタイム選手追跡システムの開発とバドミントンシングルのゲーム分析", CAS2003-36, pp.91-96 (2003).
- 11) 岩元浩太, 麻田知嗣, 小松尚久: "信頼度パラメータを用いた広域動画における人物追跡", IN2001-137, pp.13-18 (2001).
- 12) 佐竹純二, 尺 長健: "階層的注視点制御による動画画像上の複数人物追跡", 信学論誌, Vol.J86-D-II, No.8, pp.1212-1221 (2003).
- 13) 坂本和則, 横田健一郎, 大野啓司, 鈴木康平, 有澤 博: "多視点テレビ放送を用いた映像内のオブジェクト位置の認識と追跡", 信学論誌, Vol.J87-D1, No.2, pp.175-187 (Feb. 2004).
- 14) B.M. Junior and R.O. Anido: "Distributed Real-Time Soccer Tracking", VSN'04, October 15.
- 15) 三須俊彦, 苗村昌秀, 高橋正樹, 和泉吉則: "オブジェクト追跡と背番号認識の連携による動画画像用スポーツ選手同定手法", FIT 2003, LI-012.
- 16) J. Sullivan and S. Carlsson: "Tracking and Labelling of Interacting Multiple Targets", In Proc. 9th European Conf. on Computer Vision (2006).
- 17) R. Cucchiara, C. Grana, M. Piccardi, A. Prati, and S. Sirotti: "Improving shadow suppression in moving objects detection with hsv color information", In IEEE Intelligent Transportation Systems Conference Proc., pp.334-339 (2001).
- 18) "デジタル画像処理", CG-ARTS 協会 (2006).
- 19) Y. Ohno, J. Miura, and Y. Shirai: "Tracking Players and Estimation of the 3D Position of a Ball in Soccer Games", Proc. 15th Int. Conf. on Pattern Recognition, pp.145-148, Barcelona, Spain (Sep. 2000).

- 20) X. Wang, K. Tieu, and E. Grimson: "Learning Semantic Scene Models by Trajectory Analysis", Proc. of ECCV2006, Vol.3, pp.110-123, Graz, Austria (2006).

(2009年2月9日受付)

(2009年6月22日再受付)



谷本 正幸

2007年, 筑波大学第三学群工学システム学類卒. 2009年, 同大学院システム情報工学システム情報工学研究科知能機能システム専攻博士前期課程修了. 現在, 同博士後期課程在学中. 自由視点映像に関する研究に従事.



北原 格 (正会員)

1996年, 筑波大学大学院理工学研究科修了. 同年シャープ(株)入社. 2000年, 筑波大学先端学際領域研究センタ助手. 2003年, ATR 研究員. 2005年, 筑波大学大学院システム情報工学研究科講師, 2008年, 同准教授. コンピュータビジョン, 複合現実感の研究に従事. 2006年度画像電子技術賞受賞. 博士(工学).



亀田 能成 (正会員)

1991年京大・工・情報卒. 1996年同大学院博士後期課程認定退学. 同年同大学助手. 2001~2002年米国MIT客員研究員(併任). 2003年筑波大学講師. 2004年同大学院助教授, 現, 准教授. 博士(工学). 複合現実感やマルチメディア処理などの研究に従事. 電子情報通信学会シニア会員, 画像電子学会理事(2007~2008).



大田 友一 (正会員)

1978年, 京都大学工学部情報工科学科助手. 1981年, 筑波大学電子・情報工学系講師, 1992年, 同教授. 2004年, 筑波大学大学院システム情報工学研究科教授, 2009年, 同研究科長. IAPR, IEICE, IPSJ 各フェロー. 電子情報通信学会理事(情報・システムソサイエティ会長).