コンピュータビジョンと
イメージメディア 133−16
（2002．5．9）

多視点画像を用いたサッカー選手の追跡

岩瀬 幸子† 斎藤 英雄†*

† 慶應義塾大学理工学部情報工学科
* 科学技術振興事業団，さきがけ研究２１,「情報と知」領域

現在，サッカーは世界で最も盛んなスポーツであり，その映像データをもとに選手追跡を基本としたさまざまなシーン解析が報告されている．しかし，サッカーはプレーヤーが多いため，選手が重なって見えるといったオクルージョンにより追跡が失敗してしまうことが多かった．特に，移動物体が複数あるサッカーシーンの場合は，オクルージョン発生後の対象物体の対応付けが非常に難しい．そこで本研究では，サッカーのゲームシーンにおいて，複数の固定カメラから撮影された多視点画像を用いることでオクルージョンの問題を解決し，全カメラにおいて選手 1 人のロバストな追跡を行う手法を提案する．

Tracking Soccer Player Using Multiple Cameras

Sachiko Iwase†, Hideo Saito†*

† Department of Information and Computer Science, Keio University
* "Information and Human Activity", PRESTO, JST

Soccer is one of the most popular sports around the world, and various researches have done for soccer scene analysis by tracking soccer players. However, tracking is often failed when occlusion occurs, then we can not get accurate location of the player. In this presentation, we will show a method of tracking a soccer player using multiple cameras. By using multiple view images, we can get over the occlusion problem, and obtain robustness in player tracking.

1 はじめに

サッカーは，いまや世界で最も多くの人々に愛好されるスポーツであり，テレビでも頻繁に放送されている．現在サッカーに関しては，シーンの解析，復元，カメラワークの分析など様々な研究が報告されている．

サッカーシーンの解析についての研究は，数多くなされてきた．これらはテレビ画像のダイジェストの作成や戦術の解析などを目的としている．例えば，テレビ中継画像を利用して状況理解やダイジェストの作成を行う研究 [1] や，カルマンフィルタによる動き予測や，テクスチャ，色，形状といった複数の観測情報を利用して選手を追跡する研究 [2] などがある．このように，シーンの解析は戦術面の理解を目的としており，選手追跡が基本となっている．

また，シーン解析による，映像コンテンツの検索を目的とした，サッカー映像における特定映像イベント抽出システムの研究 [3] などもある．サッカー映像では，映像の中の選手やボールをオブジェクトとして取り出し，それぞれの動きや関係の変化に基づいて選手の動作を分類し，映像内容を把握する必要がある．サッカーシーンの解析は，他にもサッカーに関するいろいろな研究に応用されている．

さらに，サッカーシーンの 3 次元解析により，状況理解や臨場感のある映像生成を目的とした研究も行われている．例えば，多視点映像から中間映像を生成する研究 [5] などがある．これにより，たとえユーザーの希望する視点に実際にはカメラがなくても，その視点からの映像を得ることができる．

カメラワークの分析に関する研究としては，番組製作現場における知的ロボットカメラの実現を目的とし

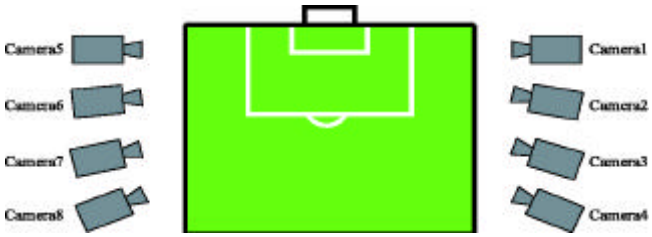


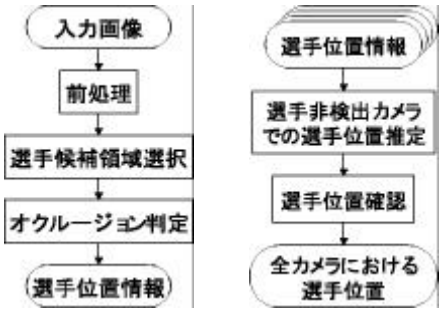
図 1: サッカーフィールドとカメラ位置の関係

ている [4]．カメラがあらかじめ決められた動きをすることは可能だが，実際に動画像処理を用いて対象物体を追跡し，視点を決めるまでには至っていない．しかし，サッカーの試合を複数の固定カメラで撮影し，ボールを追跡することにより，シーンを解析して最適なカメラ視点を決定する手法 [6] などが提案されており，知的自動撮影システムへの手がかりとなっている．

以上のように，現在のサッカーシーンの解析は選手追跡が基本であり，そしてその確実性が求められている．しかし，サッカーというスポーツ上画面ではオクルージョンが多く発生し，またプレーヤーも多いため，選手の追跡を確実に行うのは非常に難しい．よって，サッカーに関する研究では，いかにロバストな選手追跡ができるかという点が重要となってくる．

そこで本研究では，サッカーのゲームシーンにおいて，多視点画像を用いることでオクルージョン問題を解決し，用いる全てのカメラにおいてロバストに選手を追跡する手法を提案する．ここで，オクルージョンとはカメラの位置によって複数の選手が重なっているように見えることをいい，実際に選手同士が接触している場合は含まない．

また本手法では，カメラの校正について，非常に手間のかかる多数のカメラの強校正を使わずに弱校正のみを行うだけでよいので，労力を格段に削減することができる．このように多視点を協調して用いるときには，カメラ間の幾何学的関係を知ることが大切である．しかし，屋外の場合のように広い撮影空間を対象とする場合には，校正用のポイントを広範囲にわたって正確に設置することが困難であるため，カメラキャリブレーションを行うことは難しい．しかし本研究では，各カメラに共通に見えるマーカ点から求められる，カメラ間の対応関係を示す Fundamental Matrix(F-Matrix) を用い，容易に複数のカメラの位置関係を知ることができる．



(a) カメラ内処理　(b) カメラ間処理

図 2: 処理の流れ

2　手法

はじめに撮影環境についてであるが，本手法ではゴール前の空間を囲むようなあらゆる方向からの視点を必要とする．そこで，設置可能な場所であるメインスタンドとバックスタンドの最上部に各 4 台ずつ図 1 のように配置し，計 8 台のカメラを固定して同時に試合の撮影を行う．そして，スタンド最上部から斜め下に見下ろすようにして選手を撮影する．

このような通常のスタジアムでの撮影では，3 次元位置が既知の点を用いた強校正を行うことは殆ど不可能であり，画像間の対応点を自然特徴点を用いて検出することによる弱校正しか事実上行うことができない．本手法では，20 点程度の自然特徴点を手入力により検出することにより各カメラ間の F-Matrix を推定した．

本手法ではまず，初期フレームにおいて，用いる全てのカメラにおいて追跡する選手を手入力で与える．そして，毎フレーム全カメラにおいて追跡する選手の位置を求めることで，選手の追跡を行っていく．

本手法の流れを図 2 に示す．まず，各カメラについてカメラ内処理を独立で行うことにより選手追跡を試みる．オクルージョンがないと判定されたカメラでは，カメラ内処理のみで選手が追跡できたとし，その位置を保存する．選手非検出カメラ，すなわちオクルージョンが発生している，または選手が画角外にいるカメラについては，F-Matrix によってあらわされるカメラ間のエピポーラ幾何情報を利用するカメラ間処理を行い，8 台のカメラの情報を統合することで選手の位置を求めていく．本手法では単眼でオクルージョン判定ができていることを前提とする．

2.1 カメラ内処理

2.1.1 前処理

図 3 に前処理の流れを示す．選手領域を抽出するには，まず背景差分を行う．フレーム差分を用いると影の影響は除去されるが，動きの少ない選手が検出されにくい．今回用いる入力画像においては影がほとんど現れなかったので，背景差分を用いる．背景差分画像に対して，明度と面積に関して経験的に求められた閾値を用いて閾値処理を行い，二値画像を生成する．その際，膨張収縮をし，ノイズ除去も行う．ある入力画像とその背景画像をから生成した 2 値画像を図 4 に示す．このように生成された二値画像に対してラベリング処理を行い，特徴量として各ラベルの重心と面積を求めておく．本手法では，この 2 値画像において選手が検出されていることを前提とする．



図 3: 前処理

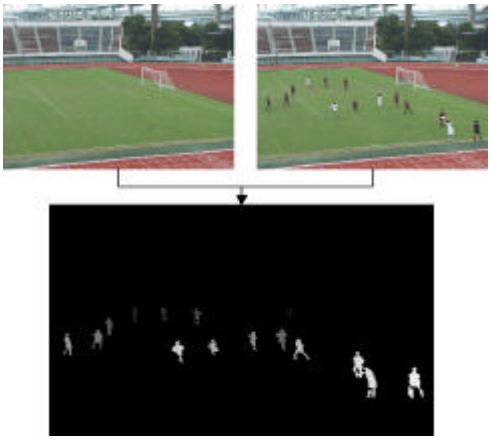


図 4: 抽出した選手領域

2.1.2 選手候補領域選択

前フレームにおける追跡する選手の位置をもとに，現フレームにおいて前処理で抽出された選手領域の中から，追跡する選手候補領域を求める．前フレームにおいて選手が画角内にいた場合については，前フレームで選択された領域との移動距離が最小となる選手領域を選択する．前フレームにおいて選手が画角外にいた場合については，選手候補領域の選択は行わず，カメラ間処理を用いて選手の位置を求める．

2.1.3 オクルージョン判定

求められた選手候補領域において，追跡する選手が他の選手と重なっていないかどうかの判定を行う．この判定は前フレームと現フレームで，選手候補領域の面積と，追跡する選手の周りにいる選手の人数を比較することで行う．前フレームにおける選手の状態によって判定の仕方が異なる．例えば，前フレームに比べて現フレームにおいて，選手候補領域の面積が増え，またその周りのラベル数が減った場合については，現フレームでオクルージョンが発生したと判定する．

ここでオクルージョンが発生していないと判定されたカメラについては，カメラ内処理のみで選手の追跡ができているとし，求められた選手の位置を保存する．オクルージョンが発生していると判定されたカメラ，または選手が画角外にいるカメラといった，選手非検出カメラについては次のカメラ間処理を行うことにより選手の位置を求めていく．

2.2 カメラ間処理

ここでは，カメラ内処理で求められた選手の位置情報を統合して，カメラ間処理を行う．選手の位置が求められたカメラを用いて，選手非検出カメラにおいても選手の位置を推定することができる．その理論と方法を以下に示す．

2.2.1 選手非検出カメラにおける選手位置推定

エピポーラ幾何　カメラを移動させたとき，あるいは複数のカメラがあるとき，複数の視点における相対的なカメラの位置や姿勢の情報は，このエピポーラ幾何と呼ばれる画像特徴の幾何によって記述できる．

図 5 に示すように，2 つのカメラで三次元空間の同じ点 $\mathbf{X}$ を見ているとすると，点 $\mathbf{X}$ と 2 つのカメラのレンズの中心 $\mathbf{C}$，$\mathbf{C}'$，そして両方の画像 π，π' における点 $\mathbf{X}$ の投影 $\mathbf{x}$，$\mathbf{x}'$ が 1 つの平面，エピポーラ平面 Σ 上に存在することになる．また，エピポーラ平面 Σ と画像面 π，π' との交線 $\mathbf{l}$，$\mathbf{l}'$ をエピポーラ線といい，視点 $\mathbf{C}$ と視点 $\mathbf{C}'$ を結ぶ直線が 2 つの画像面 π，π' と

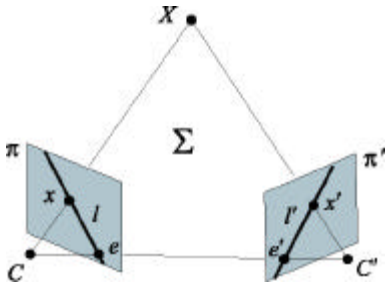


図 5: エピポーラ幾何における諸概念

交差してできる点 $\mathbf{e}$，$\mathbf{e}'$ をエピポールと呼ぶ．

カメラが固定である場合，エピポールは固定なので，片方の画像上の 1 点を取るとその 3 点で平面ができ，もう片方の画像上では直線に対応することになる．この点と直線の関係を表した行列が F-Matrix である．

カメラ座標系 (x, y) と他のカメラでの対応直線 (エピポーラ線) との関係式を式 1 に示す．

$$\begin{pmatrix} a \\ b \\ c \end{pmatrix} = \mathbf{F} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} \tag{1}$$

式 1 において $\mathbf{F}$ が F-Matrix であり，a, b, c は対応直線:$ax + by + c = 0$ のそれぞれの係数に対応する．

選手位置推定方法　エピポーラ幾何により，あるカメラから撮影された画像中の画素から，他のカメラで撮影された画像上に，対応する直線を算出することができる．よって，2 つのカメラにおいて選手の位置が求められていれば，オクルージョンで選手の位置が求められない画像においても，求められた 2 つのそれぞれの選手の位置に対応する直線の交点を算出することで，その点を選手の位置と考えることができる．図 6 に，オクルージョン発生時の選手の位置推定方法を示す．選手が画角から外れている場合についても，同様に選手の位置を求めることができる．

今回は全部で 8 つのカメラを用いて実験を行っているが，先に述べたの選手の位置推定方法では，選手の位置が求められた 2 つ以上のカメラを必要とする．0，あるいは 1 つのカメラでしか選手の位置が求められなかった場合は，位置推定を行うことができない．2 つ以上のカメラで選手の位置が求められた場合は，その中でカメラ間の距離が最大となる 2 つのカメラを選択する．本手法では図 1 のようにサッカーフィールドとカメラの位置関係の概略が分かっているため，カメラ

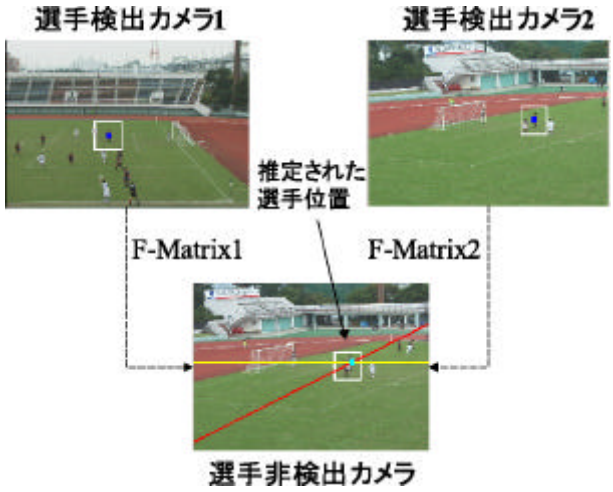


図 6: オクルージョン発生時の選手位置推定

間距離の相対関係を得ることができる．カメラ間距離が最大となるカメラを選択するのは，カメラ間の距離が小さいと，その 2 つのカメラから得られるエピポーラ線の交角が小さくなり，交点にずれが生じてしまうからである．

2.2.2　選手位置確認

以上の処理により，用いた全てのカメラの画像上で追跡する選手の位置が求められる．しかし，オクルージョンが発生した後の追跡や，途中のフレームから選手が現れた場合の追跡では，追跡したい選手とは違う選手を追ってしまうことがある．そこで，より安定して追跡を行っていくために，複数のカメラの情報を用いて選手の位置確認をする．

位置確認の方法は，まず，位置確認を行うカメラの求められた選手の位置座標をもとに，カメラ内処理のみで選手の位置が求められたカメラに対して，図 7 のようにそれぞれエピポーラ線を引いていく．

カメラ内処理のみで選手の位置が求められたそれぞれのカメラにおいて，その選手の位置座標とエピポーラ線との距離を算出する．その距離があらかじめ設定した閾値 (後述の実験では閾値を 10 とした) 内であれば，確認を行うカメラにおいてそのままその選手の位置座標を保存する．その距離が適当でない場合は，位置確認を行うカメラにおいて，最初に確認を行った選手に近い選手から同様なチェックを行っていく．そして各選手とエピポーラ線との距離を算出し，その距離が適当となる選手の位置座標を保存する．

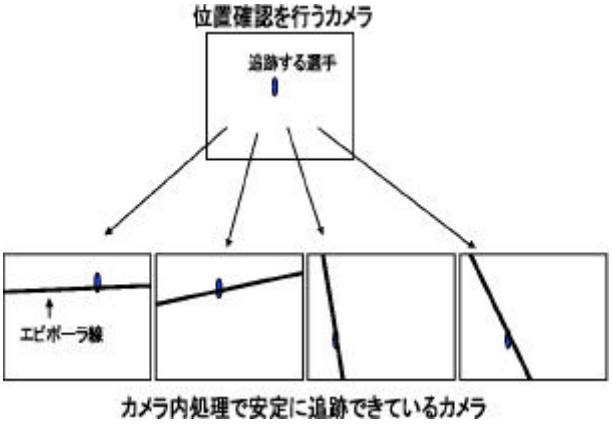


図 7: 選手の位置確認



(a) 本手法による選手軌跡



(b) 実際の選手軌跡



(c) 1 つのカメラによる選手軌跡

図 8: シーン 1，カメラ 4 における選手軌跡

3　実験

2000 年秋の関東大学サッカーリーグ戦の数試合について，8 台のカメラを用いて実際にスタジアムで撮影した多視点ビデオ画像を用いて，本手法の有効性を確認する実験を行った．入力画像には，RGB24bit，画像サイズ 720×480pixel，フレーム間隔 15frame/sec の多視点時系列画像を用いた．実験は，シーン 1(350 フレーム) とシーン 2(190 フレーム) に対して行った．

4　結果と検討

4.1　追跡結果

図 8 において，シーン 1 のカメラ 4 における，本手法による選手軌跡を図 (a)，実際の選手軌跡を図 (b) に示す．軌跡を比較すると，本手法による選手軌跡は実際の選手軌跡と同じような軌跡をしていることが分かる．ところどころで選手軌跡が途切れているが，これは，本手法ではエピポーラ幾何が正確に求まっていることを前提としているが，実際用いる F-Matrix には多少の誤差があるため，エピポーラ線の交点を求める際にずれが出てきてしまったものと考えられる．しかし，全体としては正しい選手追跡が行われている．

表 1 にシーン 1 のいくつかのカメラにおける選手追跡成功率を示す．成功率を見ると，何回かオクルージョンが発生していても，どのカメラにおいても高い確率で追跡が成功していることがわかる．この結果により，他のカメラ，または他のシーンにおいても良好な結果が得られることが予測できる．

表 1: 追跡率

カメラ番号	オクルージョン発生回数	追跡失敗 [フレーム数]	成功率
2	4	0	100%
4	6	8	97.7%
6	6	18	94.9%
8	6	3	99.1%

4.2　1 つのカメラによる追跡との比較

本手法は複数のカメラを用いたが，本手法と 1 つのカメラによる追跡とではどの程度追跡結果に差が出るのか，比較を行った．1 つのカメラによる追跡では，他のカメラからの情報を得ることができないので，オクルージョンが発生した時には正確な選手の位置を求めることができない．また，一旦選手の追跡に失敗してしまい，他の選手を追っていってしまうと，再び正しい選手に追跡を戻すことができない．図 8 において，その軌跡を比べてみると，違いは明らかである．

図 9 は，シーン 1 のカメラ 4 において，1 つのカメラによる追跡と本手法による追跡が，実際の軌跡とどれくらい距離があるかということをグラフで示したものである．1 つのカメラによる追跡では，最初の 40 フレームくらいは正しく追跡が行われているが，それ以降のフレームでは異なる選手を追跡している．そして，その後再び正しい選手に追跡が戻ることがない．しかし本手法による追跡では，フレームによっては選出追

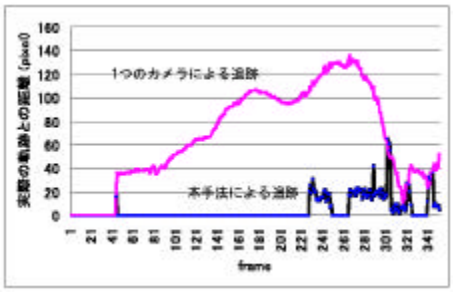


図 9: 実際の選手軌跡との距離による比較

跡に失敗していることがあるが，その後再び正しい選手に追跡が戻り，全体としては追跡に成功していることがわかる．このように，複数のカメラを用いる追跡では，1 つのカメラによる追跡に比べてロバストな選手追跡を行うことができる．

4.3 検討

以上のように，本手法では基本的には良好な追跡結果が得られている．シーン 2 では，カメラによっては選手が始め画角から外れており画像上に現れていなかったが，他のカメラには現れておりそのカメラで選手の位置を求めることができたため，途中から画像上に現れてもそこから正しい追跡を行うことができた．本来ならば，初期フレームで追跡する選手を決定することができないため追跡を行うことができないが，本手法では他のカメラを用いての選手位置推定が可能なので，途中から追跡を開始することができる．

本手法を用いても例外的な場合には失敗する．本手法では，カメラ内処理においてオクルージョンの有無がわかっていることを前提としていたが，そのオクルージョン判定に失敗した場合には追跡も失敗してしまうことが多い．また，カメラ間処理の選手位置確認についても，大勢の選手で混みあったシーンの場合は他の選手を追跡し始めてしまうことがある．しかし，基本的には高い確率で追跡に成功している．

5 おわりに

8 つのカメラで撮影されたサッカーシーンにおいて本手法を用いて選手追跡を行ったところ，オクルージョンがある場合にも安定した追跡をすることができ，弱校正のみを行うことにより得られるカメラ間の射影的関係を表す F-Matrix を用いて複数のカメラから得られる情報を統合した，本手法の有用性が示された．

サッカー選手の追跡に関する基礎研究は近年盛んに行われているが，オクルージョンの問題により，1 人の選手でさえ安定に追跡を行うことは困難である．そこで，従来の方法による選手追跡に共通なオクルージョン問題を克服するため，本手法を提案した．今後課題としては，まずはより安定した追跡を目指し，そして将来的には 1 試合通して選手を追跡し，その情報を用いて様々なシーン解析に応用したいと考えている．例えば，今回は 1 人の追跡を行ったが，本手法を全ての選手の追跡に拡張し，その選手情報をもとに，戦術解析，シーンの復元，番組製作現場におけるカメラ制御などを行うこともできる．用いる全てのカメラにおける選手位置が毎フレーム求められるため，本手法の応用分野は広いと考えられる．

参考文献

[1] 上原 恵次，吉田 貴夫，小沢 慎治："サッカー中継のシーン解析"，電気学会情報処理・産業システム情報化合同研究会，IP-96-42，IIS-96-41(1996)

[2] 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

[3] 丸尾 二郎，岩井 儀雄，谷内田 正彦，越後 富夫，飯作 俊一："サッカー映像からの特定映像イベントの抽出"，信学技報，Vol.99，No.179，pp.31-38(1999)

[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] 稲本奈穂，斎藤英雄："サッカーシーンにおける多視点映像からの中間映像生成"，信学 PRMU 研究会，PRMU2001-151(2001)

[6] 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 PRMU 研究会，PRMU23-34(2000)