

サッカー映像の自動ゲーム分析

中川靖士[†] 羽田久一[†]
今井正和^{††} 砂原秀樹[†]

インターネットのブロードバンド化により映像配信が盛んに行われるようになりつつある。しかし、限られた人手や予算の中で映像内容を理解し、整理や編集することが難しいため、視聴者にとって理解しにくい内容となっている。本研究では、解説者が手作業で行っているサッカー映像のゲーム分析を自動化する。映像に対して2値化、円形度、ハフ変換、アフィン変換、オプティカルフローを用いて画像解析を行い、選手、ボール、ラインを認識し、視点移動、カメラワーク認識、プレー認識を行う。その結果を時間、場所、チーム、プレーについて数値化したスコアブックで管理する。さらにスコアブックからボールキープ時間やボールの移動軌跡に展開する。実際の映像に対して実験を行い、手作業との比較において本手法の有効性を確認した。

Automation of the Soccer Game Analysis

Yasushi Nakagawa,[†] Hisakazu Hada,[†] Masakazu Imai^{††}
and Hideki Sunahara[†]

Recently it has become popular to deliver the video data on the Internet, and become necessary to add commentary information to the video stream at low cost. In this paper, we use such image processing techniques as binarization, compactness, Hough transform, affine transformation, and optical flow, so that the players, the ball, and the lines on the field can be recognized automatically. The numerical values from the recognition are managed in the scorebook. Furthermore the ball holding time and the locus of the ball are derived from this scorebook. The experiment shows this system is useful enough to analyze the soccer game.

1. はじめに

一般家庭におけるインターネットのブロードバンド化により映像配信が盛んに行われるようになりつつある。従来の地上波テレビ、衛星放送、ケーブルテレビに加え新たなチャンネルとして情報発信が行われ、生中継のみならず蓄積される録画映像も増加している。

しかし、チャンネル数の増加に伴い編集作業が十分なコンテンツが発信されることがある。これは限られた人手や予算の中で既存のテレビ放送のような編集作業を行ったコンテンツを作成することが難しいためである。また、インターネット放送の映像データは現場の様子を映しているだけの生データで、スポーツ中継における実況を行うアナウンサーやゲーム分析を行

う解説者が存在しないため、視聴者にとっては理解しにくい内容となっている。

本研究ではインターネット放送のコンテンツ充実のためにスポーツ中継における解説者の代わりにゲーム分析を自動的に行うシステムを提案する。対象とするスポーツは世界最大の視聴者数を誇り、各国で盛んに行われているサッカーとする。

ゲーム全体を把握するためには、サッカー映像からプレーを認識し、スコアブックを作成し、チームの特徴や戦術を理解することが求められる。実際の映像に対して実験を行い、解説者が行っていた手作業との比較において本手法の有効性を確認した。

2. 関連研究

丸尾ら¹⁾は、映像コンテンツの検索を目的としてサッカー映像における特定映像イベント抽出システムを提案している。カメラの状態推定を行い、静的オブジェクトとしてゴールとライン、動的オブジェクトとしてボールと選手を検出し、最終的にコーナーキックを特定映像イベントとして抽出する。この研究は映像

[†] 奈良先端科学技術大学院大学 情報科学研究科
Graduate School of Information Science, Nara Institute
of Science and Technology

^{††} 鳥取環境大学 情報システム学科
Department of Information System, Tottori University
of Environmental Studies

コンテンツの検索が目的で、視聴者の理解を助ける情報を付加するものではない。また、ゲームを戦略的観点から分析するものではない。

掛水ら²⁾はサッカーを科学分析に基づいて確率の高いシュートやトレーニング方法を解説している。サッカーゲームのデータ化は「いつ」「どこで」「だれが」「なにを」に集約でき、ゲーム中に起こる全ての情報を数値データに置き換えることができる。この研究のデータ作成は人手による作業が多く自動化されていない。また、得られた戦略的な分析情報は監督が練習計画、選手選考、作戦変更、選手交代に使用しているものの、視聴者には公開されていない。

ORAD 社³⁾は Tactical Replay と Quick Mode を販売している。Tactical Replay はテレビ放送でのスポーツ分析向けにアニメーション化されたプレイブックを作成し解説者のゲーム分析を助ける。Quick Mode はハイライトのアニメーションやサッカープレイの統計グラフを作成する。実行処理速度は高速であるがテレビ放送局向け専用を用意されているので、一般視聴者が利用することはできない。

山田ら⁴⁾は、サッカーの TV 中継を対象とし選手認識、カメラ認識、ボール認識を行い 3D 化した映像へ展開している。選手認識はユニホームのシャツとパンツの 2 つの組み合わせを使用しキーパーも判別している。カメラ認識はアフィン変換の特殊式を使用し係数を 3 つまで絞り込んでいる。ボール認識は重力や空気抵抗までも考慮している。この研究の出力結果は複数画面によるワイドスクリーンやグラウンド真上からの図であり、プレー認識やゲーム全体の把握して戦略的観点から分析するものではない。

3. 設 計

本研究における提案手法は、解説者が手作業で行っているスコアブックの作成を自動化することである。その概要を図 1 に示す。

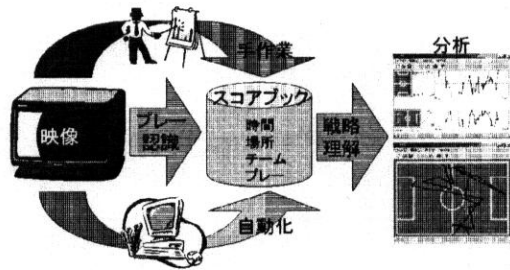


図 1 提案手法

具体的には、画像解析を用いて選手認識、ボール認識、ライン認識、視点移動、カメラワーク認識、プレー認識を自動的に行う。

スコアブックは選手がボールにタッチするごとに時間、場所、チーム、プレーを 1 ラインとして記録した。時間は映像開始からの時刻 (秒) で表す。場所はグラウンドの左手前を原点とする x 座標、y 座標 (m) で表す。チームはユニホームの色でホームとアウェーに分類する。プレーは同一選手のボールタッチ数として扱う。1 タッチはダイレクトパス、2 タッチはトラップとパス、3 タッチ以上はドリブルとパスに分類する。場所とプレーの情報と合わせることでコーナーキック、ゴールキック、スローインを類推することも可能である。

スコアブックからチーム別ボールキープ時間、XY 成分に分割した時間ごとのボール位置、グラウンドを 12 分割した地域別プレー回数、ボールの移動軌跡などの模式図を作成する。これにより、チームの特徴や戦略を理解することが可能となる。

前提として、視聴者の個人利用を目指し特別なカメラや機材を使用しない一般的なサッカー映像を対象とする。また、特定選手の活躍やゴールシーンにとらわれることなくゲーム全体を把握することである。

関連研究と本研究を比較した結果を表 1 に示す。本研究の特徴は、画像解析を用いて映像解説のためプレー認識を自動化し、個人利用が可能な環境で戦略分析を行うことである。

表 1 関連研究と本研究の比較

	映像解説	自動化	個人利用	戦略分析
丸尾ら	×	○	○	×
掛水ら	×	×	×	○
ORAD 社	○	○	×	○
山田ら	○	○	○	×
本研究	○	○	○	○

4. 実 装

本章では、画像解析を用いて選手認識、ボール認識、ライン認識、視点移動、カメラワーク認識、プレー認識を自動的に行う方法について述べる。また、実装における問題点や手作業の部分も記述する。

4.1 選手認識

選手を認識するために芝生の緑色と選手のユニホームの色を使用する。

まず、動画像 (AVI ファイル) を 1 フレームごとの画像サイズ 720x480 ピクセル、色数は 1677 万色 (32 ビット) の静止画像 (BMP ファイル) に切り出す。この画像の色データは赤 (R)、緑 (G)、青 (B) で与えら

れているが、RGBの間には相関関係があり、Rの値が高い場合は、GやBの値も高くなる場合が多い。

そこで、RGBを色の明るさを表す輝度(Y)、色の種類を表す色相(H)、色の濃さを表す彩度(S)に変換して使用する。輝度は相関関係もなく人の目の感覚に近い白黒画像となる。

YHSのヒストグラムを作成し、そのピーク部分を取り除くことによりグラウンドを除外する。グラウンドの輝度よりも暗い部分、明るい部分に分ける閾値を設けて2値化処理を行う。ユニホームの色に対応する色相によって選手の敵味方のチームを判別する。

選手の画素の上にテンプレートマッチングを行い、72x36ピクセルの四角を使って選手を認識させた。その結果を図2に示す。この図のように、ゴール前の混戦にならない限り選手同士のマークは厳しくないのので、この方法は選手認識において許容範囲内であると言える。

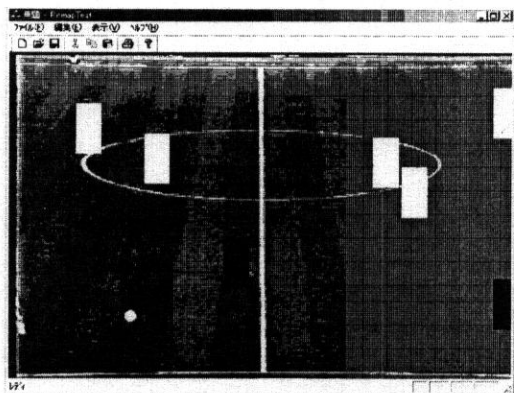


図2 選手認識

問題点は4つある。選手同士の競り合いによって前後に重なったときに敵味方のチームを判別できない。ゴールキーパーや審判のユニホームを判別できない。ラインや観客席を選手として誤認識する。選手の画素に対してテンプレートの大きさが一定のため誤差が発生する。

2値化閾値(輝度)の算出とユニホームの色(色相)の決定は手作業で行っている。

4.2 ボール認識

ボールを認識するためには画像上で円に近いというボールの形状を使用する。

選手認識の2値化処理で取り出された全ての画素に対して、連結成分ごとに同じラベル番号を振る。このラベルごとに面積S、周囲長l、円形度eを求める。

面積は物体の画素数を加算し、周囲長は物体の輪郭線の画素数を加算する。円形度は、

$$e = \frac{4\pi S}{l^2} \quad (1)$$

で与えられ、形状の複雑さを測る特徴量であり、円に近いほど1に近づく性質も持つ。

この円形度が最大となるラベルの画素をボールとして認識させた。その結果を図2に合わせて示す。対象画素の色が白く、大きさが30から100ピクセルとし、半径8ピクセルの円として認識させた。

ボールが選手やラインと重なる場合があるのでボールの移動可能範囲を計算し、フレーム間予測を使うことで精度向上と処理速度向上を行っている。1フレーム目は全画面に対して検索し、2フレーム目以降は移動可能範囲のみを検索対象としている。

問題点が3つある。色と大きさが似ているため選手が白いソックスを履いているとボールと間違える。ボールの移動速度が速いと見かけの形状が楕円に近くなり円形度が小さくなる。1フレーム目から選手やラインと重なっている場合は認識不可能である。

これら問題点は、1フレーム目さえ認識できればフレーム間予測を用いて補完できるので、この方法はボール認識に対して効果的である。

4.3 ライン認識

ライン認識はグラウンド上の選手やボールを誤認識しないため、選手やボールの位置を決定するパラメータとしてライン同士の交点やサークルの頂点を求めるために必要である。ライン認識のためにハフ変換を用いる。

ハフ変換はある直線 $y = ax + b$ は、原点から下ろした垂線の長さ ρ 、 x 軸となす角 θ を用いると

$$\rho = x \sin \theta + y \cos \theta \quad (2)$$

で表せる。デジタル画像の中に n 個の点で構成された直線があれば、それらは ρ - θ 平面上では n 個の三角関数が交わったところを表す。さらに ρ - θ 平面で交わる点の数が多いほうから取り出していくことで、長い直線を順番に検出する方法である。

また、ハフ変換の一般式を使えば円や楕円を取り出すことも可能である。つまり、タッチラインやエンドラインなどの直線だけでなくセンターサークルやペナルティーアークの楕円や円弧も検出できる。

問題点が3つある。量子化誤差によって厳密な線分を取り出せない。交点の数が多いところから取り出すので短い線分を取り出せない。楕円を検出するには5つのパラメータが必要となり、5次元空間でのピーク検索が必要となり計算量が増大する。

これら問題点を解決する方法を実装できなかったため、この方法はライン認識に対して有効ではなかった。

そこで、対応策として先に求めた円形度が0.1以下の画素をラインと見なして選手認識やボール認識の際に除外した。

4.4 視点移動

グラウンド上の選手やボールの位置を特定するために、画像座標からグラウンド座標へ変換する視点移動を行う。視点移動にはアフィン変換を使用する。

アフィン変換は3次元方向の拡大縮小、移動、回転を一つの式で表す。アフィン変換のうち人の目線にもっとも近い透視投影法の一般式は画像座標 (x, y) 、グラウンド座標 (X, Y) とすると

$$x = \frac{AX + BY + C}{PX + QY + R} \quad (3)$$

$$y = \frac{DX + EY + F}{PX + QY + R} \quad (4)$$

として与えられる。このアフィン変換で変換した結果を図3に示す。グレーで上下に引き伸ばされた画像がグラウンド座標へ変換したものである。

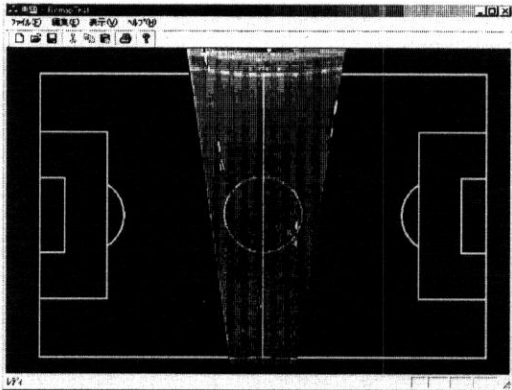


図3 視点移動

この係数 $A \sim F$ 、 $P \sim R$ の9つを一意に決定するためには基準点として位置が判っている4点が必要となる。基準点としてはライン同士の交点やサークルの頂点が有効である。

しかし、ハフ変換の結果からライン認識が思うようにできないため、基準点4点の決定を手作業で行っている。

問題点は3つある。カメラのレンズひずみのため厳密な変換できない。手動でも基準点4点を決定できないカメラアングルが存在する。基準点4点の与え方によっては係数が一意に決定できない。

これら問題点は、変換誤差は1m以内であることやカメラワーク認識で補完可能なので、この方法は視点移動に対して有効である。

4.5 カメラワーク認識

アフィン変換の基準点4点を決定できないカメラアングルに対応するため、オプティカルフローを使ったカメラワーク認識を行う。また、ゲーム分析の対象外となるカメラアングルを除外するためカメラアングルも認識する。

オプティカルフローは連続するフレーム間の情報を認識可能な動きベクトルに置き換え、ブロックマッチング法と勾配法の2つの方法がある。ブロックマッチング法は動きベクトルを推定したい画素ブロックと最も似ているブロックを検索範囲から探し出す。勾配法は画面の輝度の勾配とフレーム間差分の関係から動きベクトルを算出する。

ブロックマッチング法においては、全探索法とツリー探索法の2つの探索法がある。全探索法は検索範囲内の全てのブロックにおいて最適なものを探し出す。ツリー探索法は幾つかのブロックに分けて評価を行ってから、さらに詳細にブロックを分けて評価を行いツリー的に順次選択する。ここでは、処理時間がかかるが最適な動きベクトルを計算可能な全探索法を採用する。

このオプティカルフローの結果を図4に示す。ブロックサイズを 32×32 ピクセル、検索範囲を 64×64 ピクセルとした。左から右へ向かう矢印で表しているのは見かけのグラウンドの移動であり、実際にはカメラが右から左へ動いている。

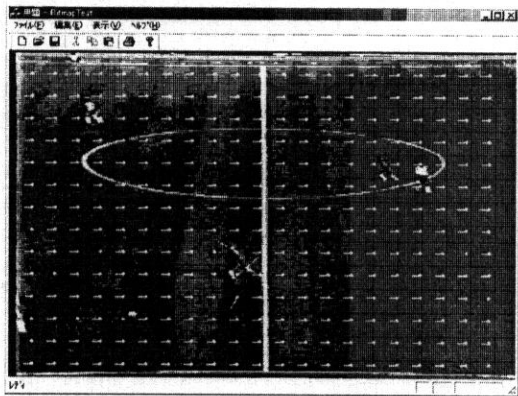


図4 カメラワーク認識

全画素に対するオプティカルフローの中央値を用いることにより、アフィン変換の基準点4点を決定できないカメラアングルに対応可能となった。さらに、ア

フィン変換の基準点4点の決定を1フレーム目だけ手作業で行えば、2フレーム目以降は自動的に算出することも可能になった。よって、この方法でカメラワーク認識を行えば視点移動を補完することができる。

問題点は2つある。ブロックマッチング法と全探索法の組み合わせは検索精度が高い代わりに検索速度が遅い。パーンやチルトは対応可能であるが中央値を用いるのでズームに対応できない。

中継映像の中には幾つかのカメラアングルが存在する。本研究では対象映像をセンターライン延長線上にあるカメラに限定する。これは、センターライン延長線上にあるカメラはボールがフィールド内にありプレーが継続しているインプレーを映す割合が60%を越え、他のカメラアングルは5%以下であるためである。残りの約35%はボールがフィールドの外に出た場合やファールが行われた場合にゲームが中断するアウトオブプレーである。これは、掛水ら²⁾の研究の結果と合致し、ゲーム分析するにあたって十分な情報が得られるのでセンターライン延長線上にあるカメラに限定する。

センターライン延長線上にあるカメラ以外のカメラアングルを対象としないために、グラウンドを示す緑色が映像の半分以上だった場合と選手のテンプレートより大きな画素があった場合を除外している。これにより選手や監督の表情を捕らえたカメラアングルを省くことができたので、この方法がカメラアングル認識に効果がある。

4.6 プレー認識

画像解析でサッカーのプレーを認識する方法は、選手とボールの接点や重なりを認識することとした。

プレーは同一選手のボールタッチ回数として表現する。そのために、フレーム間でのボールに触った選手とボールの両方を追跡する。選手とボールが重なったところでタッチ回数を1にする。その後、タッチした選手とボールが分かれてから再び接触すれば、タッチ回数に1を加える。タッチした選手以外が触れば、タッチ回数を1に戻す。

また、ドリブルの場合に選手とボールが重なって移動することがある。その場合はタッチした時(ドリブル開始)とパスを出した時(ドリブル終了)の2箇所を記録する。

しかし、選手とボールが重なってもプレーに関係ない場合があるこれをスルーと呼ぶ。スルーは、3次元上で行われているサッカーを2次元の映像データから理解するとき起こる問題で、選手の手前や裏側をボールに触れることなく通過する場合を指す。これは

選手に重なる前後のボールの方向やスピードが保たれているかどうかで判定する。

問題点は、選手認識において選手の画素に対してテンプレートの大きさが一定のため誤差を含むので、プレー認識もその誤差の影響を受ける。選手がグラウンドの手前や奥に位置すると見かけの大きさが変化し、人間のシルエットは四角で近似することが難しいことが原因である。このため、プレー認識の認識率低下を招いた。

5. 実験

実験方法はサッカー映像の1シーンを100フレームとし、2試合6種類のシーンに対して実装したツールを用いて自動的にスコアブック作成を行った。それぞれのシーンは以下のとおりである。

- (1) 選手がボールを受けてパスを出す。
- (2) ゴール前の混戦で敵味方がボールを奪い合う。
- (3) 選手がボールを受けてドリブル後パスを出す。
- (4) 選手が2つのパスをつなぐ。
- (5) 選手がシュートしキーパーがボールを弾く。
- (6) 選手がドリブル突破をする。

評価方法は手作業で作業して作成したスコアブックと比較し認識率を求める。プレーに関しては、再現率(R)検索されるべきデータのうち検索されたデータの割合と適合率(P)検索されたデータのうち望ましいデータの割合を用いる。その結果を表2に示す。

表2 実験結果

	time	place	team	play(R)	play(P)
scene 1	100%	100%	100%	100%	100%
scene 2	100%	100%	0%	86%	60%
scene 3	100%	100%	40%	67%	100%
scene 4	100%	100%	52%	80%	67%
scene 5	100%	100%	33%	83%	71%
scene 6	100%	100%	88%	75%	82%

時間と場所の認識率は100%であり、プレーの再現率、適合率とも平均80%を越える結果となった。しかし、チームの認識率は平均52%しかなく満足な値ではない。

実験の考察として、チームの認識率が低い原因は4つある。scene 2のようにゴール前の混戦の場合は選手同士の重なり合う。ユニホームの背番号やソックスの色や靴の色などが微妙に影響する。ユニホームの色が緑でグラウンドの色に近い。プレーがスルーだった場合にチームを変更しない必要がある。

プレー認識率が低い原因は2つある。選手認識で実際の画素とテンプレートの大きさの違いのため、細か

いプレーを認識できない。直線的なドリブルはスルーとして誤認識してしまう。共通する原因はスルーの処理が弱いことである。

また、実験対象にできない試合やシーンが4つ存在する。オプティカルフローを使ってアフィン変換の補完を行っているので、グラウンドに芝目がない試合会場だと場所の認識ができない。昼間の試合で観客席や選手自身の影がグラウンドにはっきりと映る映像は、2値化の認識ができない。ボール認識でフレーム間予測を使用しているが、1フレーム目やパス出で誤認識が起こるとそれ以降すべてのデータが間違いになってしまう。ゴールキックなどの放物線を描くボールは、映像から飛び出すこと、地面で大きくバウンドすることがあるので、ボール認識の追跡が難しい。

作成されたスコアブックのデータからボールの移動軌跡を図5に示す。これは scene 1 の結果であり、連続したプレーをモード化することができた。

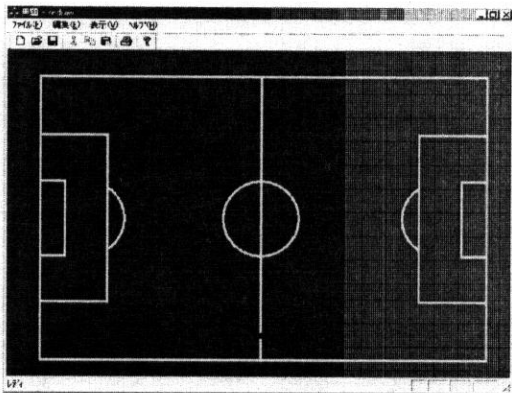


図5 実験結果

6. 課題

画像処理の手法を複数組み合わせ使用しているので各画素に対して多くの繰り返し演算が発生し、映像約3秒に対して約90分の処理時間が必要となった。対策として、オプティカルフローにおいて検索方法を変更して、勾配法やツリー探索法を採用したい。これは検索精度が低下するが検索速度の向上が見込める。また、マルチスレッド対応のプログラムに書き換え、マルチCPUのマシンで実行したいと考えている。

手作業で決定しているパラメータの自動計算は次の4つの方法で対応できると考えている。2値化の閾値はヒストグラムのクラス間分散を利用する。ユニホームの色はフレーム間予測より動的オブジェクトとして

抽出する。アフィン変換の基準点4点はハフ変換で求める直線の交点や円弧の頂点を利用する。ライン抽出はハフ逆変換の時に一定の幅を持たせる。

認識率の向上が見込めない部分や実験対象にできない試合やシーンが存在する。その原因は3つある。3次元のグラウンドで展開されるサッカーのゲームを1台のカメラから得られる2次元情報で解析するので情報の欠落が発生する。ゴール前の混戦の場合は選手同士の重なり合う場合は人間が手作業で行ってもチームやプレーの判断が難しい。カメラアングルにより1フレーム目に誤認識が発生するとフレーム間予測の全てを間違えてしまい訂正するロジックを導入できない。

7. おわりに

インターネット放送において限られた人手や予算の中で既存のテレビ放送のようなコンテンツを提供するために、解説者の代わりサッカー映像の自動ゲーム分析する方法について提案した。提案手法はプレーを認識し、スコアブックを作成し、チームの戦略理解することを画像解析によって自動化することである。選手認識、ボール認識、ライン認識、視点移動、カメラワーク認識、プレー認識のために、2値化、円弧度、ハフ変換、アフィン変換、オプティカルフローなどの画像解析を用いて実装した。スコアブックとして時間、場所、チーム、プレーを数値化して記録した。課題として画像解析に時間がかかり実行速度が十分でないためゲーム全体の分析はできないこと、1フレーム目のパラメータ決定は手動で行うことがある。しかし、2フレーム目以降の数秒間の連続したプレーを自動分析することは可能である。認識率はシーンによってばらつきがあり、場所やプレーの認識率は十分に利用価値があるが、チーム認識率はスルーの処理強化により向上が望まれる。今後は手作業で行っている1フレーム目のパラメータ決定を自動化したいと考えている。

参考文献

- 1) 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, 電子情報通信学会, IE99-17, 1999.
- 2) 掛水 隆, 大橋 二郎: サッカーおもしろ科学, 東京電機大学出版局, 1996.
- 3) ORAD 社: <http://www.orad.co.jp/>
- 4) 山田: サッカー中継画像の認識, 大阪大学工学部電子制御機械工学専攻白井研究室, 2001, <http://www-cv.mech.eng.osaka-u.ac.jp/~yamada/research/research.html>