

サッカー映像の自動ゲーム分析方法の提案と評価

Automation of the Soccer Game Analysis

中 川 靖 士

要 約 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編集や解説を加えることが難しいため、視聴者が理解しにくい内容となっている。本稿は、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。

Abstract 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, I use such image processing techniques as binarization, compactness, Hough transform, affin 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. The experiment shows this system is useful enough to analyze the soccer game.

1. はじめに

インターネットのブロードバンド化にともない映像配信が行われ、衛星放送やケーブルテレビによるチャンネル数の増加により、生中継や録画された映像が増加している。映像の視聴時間には限界があるため、映像を効率よく検索できることが重要である。そのためには、構造化されたデータに関するデータとしてメタデータが必要となる。映像のメタデータは、ハイライトシーンだけを検索したダイジェスト番組、長い映像を短く見せる映像要約、ニュース映像とコマーシャルを分離する映像分割、見たい映像を取り出す番組表に利用される。

映像のメタデータを作成するためには、映像の内容を理解し構造化する必要がある。その方法は、手作業と画像解析の二つの方法がある。手作業は、映像の内容を理解し必要なデータを取り出すことができるが、予算や人手が必要となり、分類があいまいになる。画像解析は、ビデオエフェクト、テロップ、シーンチェンジ、画像差分を利用することで自動化できるが、映像の内容を理解するには至っていない。

数ある映像の中で、対象とする映像をサッカー映像に限定する。その理由は、ワールドカップにおける視聴者数においてサッカー映像が世界最大のスポーツ映像であるためである。また、ボールが比較的大きく、グラウンドが一定の大きさなので画像解析の対象として扱いやすいためである。

本研究は、メタデータを作成するためにサッカー映像の自動ゲーム分析方法を提案する。これは、画像解析に加え、サッカー解説者の知識を活用し、プレーを認識させスコアブックに記

録しチーム戦略を把握することである。画像解析と解説者の知識を用いたシステムを実装し、実際のサッカー映像に対して実験を行い、サッカー解説者が手作業で作成したスコアブックとの比較において本提案の有効性を評価した。

映像のメタデータ作成を主題にした動機は、過去に二つのインターネット中継を行ったが、限られた予算や人手で映像の実況、解説、演出、編集を十分に行うことができなかった経験による。一つ目は、米国ハワイ島の国立天文台すばる望遠鏡からの遠隔授業である。二つ目は、約 600 人が参加した奈良県の生駒市民駅伝の中継である。ハワイからの中継であることを説明する難しさやランナーに対する実況や解説が必要であるという課題が残った。また、人の場所を特定するために GPS や RF ID の利用を考えたが、電源、電波、予算の面から現段階では実用的でない判断し、映像情報だけを対象に絞り込んだ。

本稿理解の前提条件として画像解析の技法について理解している必要がある。しかし、画像解析について知識がなくても、画像解析の技術に解説者の知識を加えることにより、視聴者、選手、編集者などに有用な情報が取り出せることは理解できると思う。

本研究の新規性は、メタデータを作成する方法として、手作業と画像解析は古くから多く行われているが、画像解析で映像の内容を理解し構造化する実用的な技法は見つかっておらず、いかにそれを実現したかという点にある。画像解析にサッカー解説者の知識をいかに加えるかによってどこまで実用的なシステムを構築できるかがかかっている。

本研究の有用性は画像解析の分野だけでなく、対象映像をサッカーとすることによりロボカップと呼ばれるロボットによるサッカー大会への戦略のフィードバックが可能だと考える。また、人間の場所と行動を把握できるので、監視カメラによるセキュリティ確保や人間のネットワーク上へのエージェント化などに役立つと考える。

2. 目 標

インターネット中継の経験から、視聴者に映像内容を説明するための解説や演出が重要であることが判った。しかし、限られた予算と人手の中で製作されるインターネット中継では、生の映像がそのまま配信されることがある。

ニュース、ドラマ、歌番組はセリフや歌声によって視聴者に内容の理解を助けているので、解説の必要性の高いスポーツ中継を 1 例として選び出した。さらにスポーツ中継の中から、サッカー映像を選んだ。その理由は、オリンピックを凌ぐ、世界最大規模のスポーツイベントはサッカーワールドカップと言われているためである。例えば、1994 年第 15 回ワールドカップアメリカ大会では、359 万人の観客がスタジアムを訪れ、テレビ中継を観戦した人は延べ 321 億人に達した。さらに、センターライン上の延長上に配置されるカメラの映像を対象にすることによって、予算と人手が限られる現場に対応することができる。

本研究の目標は、映像の内容を理解し構造化するために、解説者の知識を活用することである。対象となる映像をスポーツ中継として、その映像の中から選手の場所やプレーと画像解析に解説者の知識を加えたシステムを構築する。

本システムの利用者は、インターネット中継やテレビ中継を家庭で見る視聴者に加えて、サッカーの選手自身や指導者、テレビ放送局や番組製作者とする。放送業界のように十分な予算と人手があるところだけでなく、一般家庭や草サッカーチームのような少ない予算と人手の利用者を想定している。

例えば、家庭の視聴者には、ゴールシーンだけを集めたダイジェスト番組、複数の試合中継から盛り上がったベストゲームを取り出すための番組表を提供する。ワールドカップを目指す草サッカーチームの選手やコーチには、戦略や戦術の確認、コーチングやスコアブックの作成を行う。インターネット中継を含む放送業界には、少ない予算と人手でも解説や演出を可能にする。

一般的に画像解析は高価な専用ワークステーションを使用することが多いが、本システムでは利用者を考慮し、一般家庭に多く普及しているパソコンで解析を行う。

3. 関連研究

掛水ら^[1]はサッカーを科学分析に基づいて確率の高いシュート方法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、「だれが」を複数、「どこで」を3次元空間と考えるとゲーム中で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と述べている。データの使い方は長期的な練習計画や選手選考、中期的な対戦相手の戦略や自チームの戦術確認、即時的なハーフタイムでの作戦変更や選手交代に役立てている。

ORAD社^[2]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スコアブックを作成し、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ー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に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^[3]は、サッカーのテレビ中継を対象として、選手認識、カメラ認識、ボール認識を行った結果を3次元化した映像へ展開している。選手認識はユニフォームのシャツとパンツの二つの組み合わせを使用しゴール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三つまで絞り込んでいる。ボール認識は重力や空気抵抗も考慮している。課題としてユニフォーム色の自動検出、画面切り替わり対応、計算処理時間の短縮が挙げられている。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、映像を変更することに主眼が向けられている。そのため、プレー認識やゲーム全体を把握して戦略的観点から分析するものではない。

丸尾ら^[4]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もない。

4. 提案手法

本研究の目的は画像解析を用いて映像解説のためのプレー認識を自動化し、個人利用が可能な環境でスコアブックを用いて戦略分析を行うことである。この目的を達成するために、サッカー映像の自動ゲーム分析方法として、画像解析にサッカー解説者の知識を活用したシステムを提案する。その概要を図1に示す。

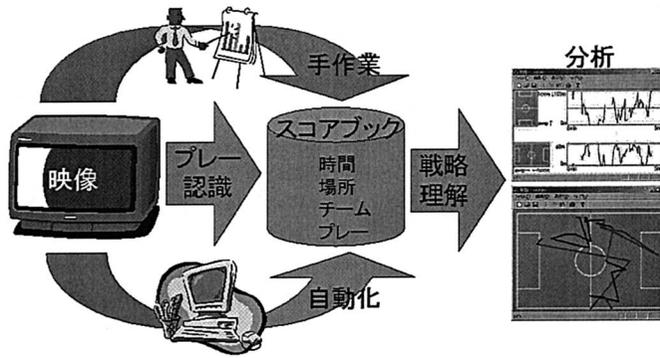


図 1 提案手法

提案手法は、プレーの認識、スコアブックの作成、チーム戦略の分析で構成される。プレー認識を行うために画像解析とサッカー解説者の知識を用いて、選手認識、ボール認識、ライン認識、視点移動、カメラワーク認識を行う。スコアブックは時間、場所、チーム、プレーを数値化したデータで管理する。ボールキープ時間やボール移動軌跡からチーム戦略の分析を行う。

関連研究と本研究を比較する。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。関連研究は、サッカー解説者のスコアブックは手作業で作成され、視聴者には公開されていない。そのため、個人的にスコアブックを利用することができず、異なる試合やチームの戦略分析ができない。また、ダイジェスト番組のためのメタデータ作成では映像解説や戦略分析を行っていない。本研究では、関連研究の問題点をカバーするようにバランスをとったシステムを構築する。

4.1 プレー認識

掛水ら^[1]の研究によると、サッカーのプレーは15種類に分類することができる。映像からプレーを認識しスコアブックを作成するために、映像の中からサッカーを構成するオブジェクトとして、選手、ボール、ラインを認識し抽出する。さらに、視点移動やカメラワークを認識することでオブジェクトの場所を特定する。そして、選手とボールの距離や位置によってプレー認識を行う。

4.2 スコアブック

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。idはデータの通し番号である。時間(time)は映像開始からの時刻(秒)で表す。場所(x, y)はグラウンドの左手前を原点とするボールの位置をx座標、y座標(m)で表す。チーム(team)はユニフォームの色でホーム(home)とアウェー(away)に分類する。どちらとも区別がつかない場合はunknownとした。チーム全体を把握するため選手個人は対象としない。プレー(play)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。これは、サッカーの基本プレーはパスであるので、パスを中心に記録するという知識によるものである。また、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、スローインなど15種類に分類したプレーを類推することも可能である。fileは画像解析の対象ファイルである。

表1のデータはFIFA コンフェデレーションズカップ2001 カメルーン vs 日本(新潟スタジアム)の5分間の結果である。このデータは日本の先取点のシーンで、解説者が行っているように手作業で抽出した。その作業時間は5時間であった。

表1 スコアブック

Id	Time(秒)	x(m)	y(m)	team	play	file
1	18.6	103.0	0.0	home	1	0558
2	20.3	80.4	7.5	away	1	0608
3	22.8	100.3	7.8	away	1	0685
4	23.7	105.0	7.5	away	0	0712
5	46.0	100.0	30.2	home	1	1380

4.3 チーム戦略分析

スコアブックからチーム戦略の分析をするために、ボールキープ時間、ボールの移動軌跡、シュートチャンスの模式図を作成した。元データは手作業で作成した表1のスコアブックである。

図2にチーム別の毎分ボールキープ時間を示す。チーム別の毎分ボールキープ時間は、上側がカメルーンで、下側が日本の1分間ごとのボールキープ時間を表している。2分までの日本はあまりボールキープできていないが、3~4分は30秒以上キープしカメルーンを圧倒している。4~5分は日本がゴールを決めたため、カメルーンはボールに触ることができなかった。サッカーの試合を相手より有利に進めるためには、できるだけ自分のチームがボールの支配力を高め、ボールをキープすることが大切である。図2はボールの支配力を知ることができる。

図3にボールの毎分ごとの移動軌跡を示す。ボールの毎分ごとの移動軌跡は、グラウンドを真上から見た図で3~4分の1分間の軌跡である。日本チームは左サイドを起点にして攻撃を続けているものの、攻めあぐねて後方向や横方向のパスが目立っている。ボールキープ時間は増えるものの、相手ペナルティエリアに入ることができず決定力不足が見て取れる。図3はサッカーの攻撃を再現でき、ゴールに至ったパスはもちろん途中のパスも流れを再現できるのでチームの攻撃の形や戦術を知る上で役に立つ。

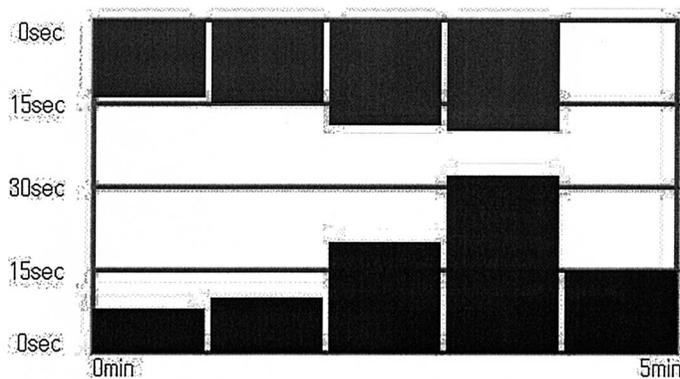


図2 ボールキープ時間

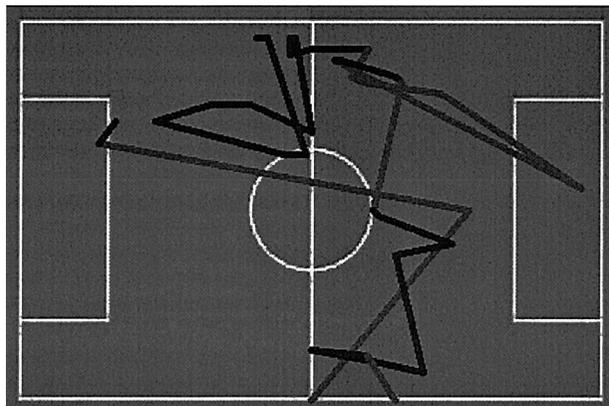


図 3 ボールの移動軌跡

図 4 にゴールする確率が高いシュートチャンスの模式図を示す．シュートチャンスは，視聴者にとってエキサイティングなシーンを抽出するために，ゴールする確率が高い場所とプレーの組み合わせをポイント化し，時間を軸に表示している．大橋ら^[5]やチャールズヒューズ^[6]の研究より以下の四つの傾向が与えられている．縦軸はそれぞれの傾向を 1 point として，最高 4 point を与えている．

- ・シュートの 80% はペナルティエリア内で行われる
- ・ゴールの 80% はダイレクトシュートあるいは 1 トラップシュートである
- ・ゴールの 80% は 2 本以下のパス回しでシュートまで行われる
- ・ゴールの 46% はセットプレーから発生する

このことから，カメルーンは 4 回あったシュートチャンスをことごとく外し，日本は 1 回しかなかったチャンスをものにして得点をあげたことになる．

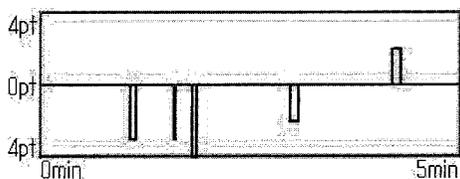


図 4 シュートチャンス

このように，これらの模式図から，チームの特徴や戦略を理解することが可能となる．また，シュートチャンスはサッカーのみどころそのものであるため，ダイジェスト番組や電子番組表に活用できる．例えば，ダイジェスト番組編集する場合にシュートチャンスのポイントに応じてシーンを組み合わせればよい．また，視聴者が複数のサッカー番組の選択に困った時にはシュートチャンスの多いサッカー試合を選べばおもしろいものを選べることになる．

これらの模式図は手作業で抽出したスコアブックからであるが，メタデータ作成のため映像内容を理解し構造化している．

5. 実 装

画像解析にサッカー解説者の知識を加えることにより、映像解説のためのプレー認識を自動化し、スコアブックを作成する部分を詳細に述べる。具体的には、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、プレー認識の方法について述べる。

5.1 選手認識

審判が悪天候の元でも敵味方の識別を可能にするために、サッカーのユニフォームはホームチームがカラー、アウェーチームが白色を基調にしたものを選ぶことになっている。また、発展途上国のテレビ事情を考慮し白黒画面においてもユニフォームからチームを認識可能にしている。この知識を利用して映像から選手認識を行う。

動画像を1フレームごとの静止画に切り出す。MPEG ファイルは画像圧縮により色情報が抜けることがあるため、動画像として画像圧縮されていない AVI ファイルを使用する。同様に静止画も画像圧縮されていない BMP ファイルを使用する。画像サイズ 720 x 480 ピクセル、色数は 1677 万色 (32 ビット) である。

画像の色データは赤 (R)、緑 (G)、青 (B) で与えられている。しかし、 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ので、人間の感覚とずれが生じる。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。 RGB から YHS への変換式は、

$$Y = 0.3R + 0.59G + 0.11B \dots(1)$$

$$H = \tan^{-1} \frac{R - Y}{B - Y} \dots(2)$$

$$S = \sqrt{(R - Y)^2 + (B - Y)^2} \dots(3)$$

によって与えられる。輝度はその他の値に対して相関関係がなく人の目の感覚に近い白黒画像となる。

サッカー映像のほとんどの部分を占めるのがグラウンドの緑色である。そこで、 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の画素を除外する。そのヒストグラムの輝度は 80 から 170 の間がピークとなっていた。ヒストグラムの結果から得られた値を閾値として 2 値化処理を行う。それに加え、ユニフォームの色に対応する色相によって選手の敵味方を判別する。

2 値化の画素の上にテンプレートマッチングを行い、72 x 36 ピクセルの四角を使って選手を認識する。72 x 36 ピクセルの値は、映像上に映る選手の大きさが最大となるグラウンド手前の位置で、実際の画素数を数えて算出したものである。

テンプレートマッチングの方法は、テンプレートの 4 倍を 1 ブロックとして、1 ブロックの中で最も画素を多く含むテンプレートの位置をそのブロックの代表点とする。さらに、1 ブロック内で隣接する代表点どうしを比較して最も大きいものを残す。ノイズを除去するために大きさの制限 (100 ピクセル以上) を選手として判定する。

結果を図 5 に示す。この図のように、選手認識の正解がホーム 2 名とアウェー 2 名のところをホーム 1 名とアウェー 3 名に間違えて認識している。これは選手がセンターサークルと重な

っているために発生している．この方法は選手認識において選手の場所を特定することができるが，チームを特定するのが難しいことが判った．

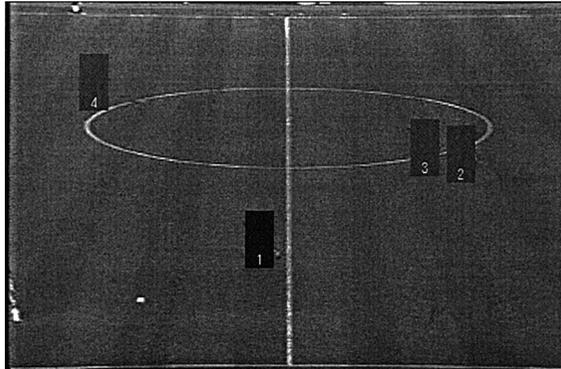


図 5 テンプレートマッチングの結果

5.2 ボール認識

ボールを認識するために，画素の特徴パラメータを使用する．ボールは円に近い形状と選手やラインと比較して面積が小さいという二つの特徴パラメータを持っている．

ラベリングは，つながっている画素（連結成分）に同じラベル（番号）をつけ，異なった連結成分には異なった番号をつけることである．ラベリングの結果からつながっている画素ごとの形状や大きさを知ることができる．具体的には，選手認識の2値化処理で取り出された画素に対して，端から走査してラベルが付けられていない画素を見つけ，その画素にラベル付けを行う．さらに，その隣接する画素に同じ番号をつけていき隣接するものがなくなるまで繰り返す．全ての画素についてこれを行う．

このラベルごとに特徴パラメータとして重心 $G(x, y)$ ，面積 S ，周囲長 l ，円形度 e を求める．重心は，

$$G(x, y) = \left(\frac{1}{n} \sum_{i=0}^{n-1} x_i, \frac{1}{n} \sum_{i=0}^{n-1} y_i \right) \quad \dots(4)$$

で与えられ，画素位置の平均値である．面積は物体の画素の数である．周囲長は物体の輪郭線の画素数であり，斜め方向は $\sqrt{2}$ 倍し積算する．円形度は，

$$e = \frac{4\pi S}{l^2} \quad \dots(5)$$

で与えられる形状の複雑さを測る特徴量であり，円に近いほど1に近づく性質を持つ．

実際のサッカー画像の2値化画素に対してラベリングを行い，円形度が最大で面積の大きさが30から100ピクセルのラベルをボールとして認識させる．

5.3 ライン認識

ラインは選手やボールに比べ連続した長い画素の集まりである．ラインは直線ばかりでなく，センターサークルやペナルティアークなどの曲線も存在し，映像によっては途切れて見える場

合もある．この知識を利用してライン認識を行う．

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試したが有効でなかったため、円形度を使用した．

ハフ変換とは、エッジ検出法の一つで、点の集合で表された画像中における線分を抽出する処理である．人工の物体の幾何的な規則性（直線や円など）を検出するために、物体の境目になっている不連続な線分を連続した境界線として抽出することができる．具体的には、与えられた画像上で多くの点がついている直線を見つけ出し、元の画像を直線だけで表される画像に変換する処理を行う．一般に雑音の多い直線や途切れた直線の検出に強いとされる．

$x - y$ 平面上のある直線は、

$$y = ax + b \dots (6)$$

として表現できる．また、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

$$\rho = x \sin \theta + y \cos \theta \dots (7)$$

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 $\rho - \theta$ 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 $\rho - \theta$ 平面上で交わる点の数が多い部分から

$$y = -\frac{x}{\tan \theta} + \frac{\rho}{\cos \theta} \dots (8)$$

によって取り出してくることで、長い直線を順番に検出する方法がハフ変換である．

実際にサッカー画像に対して直線に対するハフ変換を行った．輝度を用いた白黒画像を 2 値化した．この 2 値化で与えられた画素に対して式 7 を適応し $\rho - \theta$ 平面へ展開したものが図 6 である．三角関数の交点が 20 個以上の部分をグレーで表示し、120 個以上の部分を黒で表示する．さらに、閾値値を上げ 180 個以上の交点に対して式 8 を適応し $x - y$ 平面へ展開する．

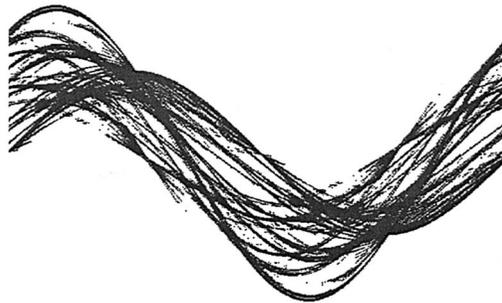


図 6 ハフ変換

タッチラインやペナルティエリアなど比較的長い線分は抽出可能であるが、ゴールエリアなどの短い線分は取り出すことができなかった．閾値を下げることにより短い線分も抽出可能となるが、線分だけでなく選手やボールも反応してしまい、抽出される線分が錯綜する結果となる．タッチラインやセンターラインを含む画像に対してハフ変換を用いた場合は比較的長い線

分で構成されるカメラアングルなので実用的であることが判った。

ハフ変換によるライン認識に対して有効ではなかったことの対応策として先に求めた周囲長や円形度の利用を考える。周囲長が 100 ピクセル以上をラインとする。上下のユニフォームの色が異なる場合は正しく認識できる。しかし、上下が同一色のユニフォームの選手を誤認識することがあったので、周囲長を用いた方法を採用しなかった。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する。周囲長や円形度を用いる方法は簡単な処理なため、直感的に判りやすく効果が高いと言える。

5.4 視線移動

センターラインの延長上にあるカメラの映像は手前の選手が大きく映り、奥の選手が小さく映る。この画像のままでは、位置や移動距離を特定することが難しくスコアブックに展開できない。そこで、映像を変換しグラウンドを真上から映した画像に変換すると把握しやすい。この知識のもとに視点移動を行う。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 次元方向の幾何学変換である拡大縮小、移動、回転を一つの式で表したものである。例えば、センターライン延長線上にある観客席から撮影するカメラアングルは斜め上方からの映像となるが、視点を移動することによりグラウンド真上からの映像に変換することができる。

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（逆変換）は、画像座標 (x, y) 、グラウンド座標 (X, Y) とすると

$$x = \frac{AX + BY + C}{PX + QY + R} \dots (9)$$

$$y = \frac{DX + EY + F}{PX + QY + R} \dots (10)$$

として与えられる。

この線形補間法と透視投影法を用いて、輝度を用いた映像を変換した結果を図 7 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。上下のタッチラインが少し湾曲しているが、センターラインやセンターサークルは綺麗に変換されている。

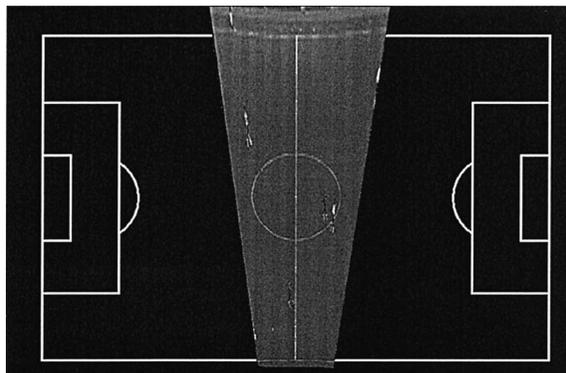


図 7 アフィン変換の結果

また、ボール位置を画像座標からグラウンド座標へ変換するためには、座標変換だけを行えばよいのでアフィン変換の順変換を用いる。逆変換、順変換のいずれもこの係数 $A \sim F$, $P \sim R$ の九つを一意に決定するためには基準点としてあらかじめ位置が判っている4点が必要となる。

その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。また、ライン同士の交点やサークルの頂点が映像に映らない場合があるので、手作業でも基準点4点を決定できないカメラアングルが存在する。そこでカメラワーク認識の結果を用いて補完する。

5.5 カメラワーク認識

センターカメラのサッカー映像は左右に変化（パン）している。逆に拡大縮小（ズーム）や上下に変化（チルト）はほとんど起こらない。左右の変化によって、アフィン変換の基準点4点を決定できない場合がある。また、中継映像を単調にさせないために、複数カメラ切り替え（スイッチング）によりセンターカメラ以外の映像が使われることがある。

これらの知識を利用し、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とは、時間的に連続する画像から画像中の移動物体を解析するために、画像の輝度情報から画素毎における動きベクトルを求めることにより、動きを認識する方法である。その動きには、画像中の物体が動く場合と撮影しているカメラが動く場合がある。ここではカメラの移動を対象とし、時間的に連続するフレーム間の動き情報を認識可能な情報に置き換える。

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方法である。具体的には、2フレームからなる連続した画像があり、2フレーム目をブロックの大きさを 8×8 とし、1フレームの検索範囲を 32×32 とし、注目したブロックの探索範囲とする。探索範囲内をブロック画素の差分の絶対値和を用いて比較し、最も小さいブロックを探し出す。同様に全てのブロックにおいて繰り返す。

実際のサッカー画像に対して、ブロックマッチング法と全探索法を用いたオプティカルフローの結果を図8に示す。

ブロックの大きさと検索範囲の組み合わせで、実際にテストを行った。最も安定した結果が得られたので、ブロックの大きさ 32×32 、検索範囲 64×64 を採用した。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。

多くの動きベクトルはカメラの移動に対応しているが、図8の中央の選手は右から左へ走っているため、その付近の動きベクトルは選手の移動を表している。そこで、カメラワーク認識の動きベクトルは、全画素に対するオプティカルフローの中央値（メディアン）を用いる。これは選手の動きベクトルはカメラの動きベクトルに比べ最大値が最小値にかたよる傾向があるため、平均値を採用するよりも精度が高くなる。

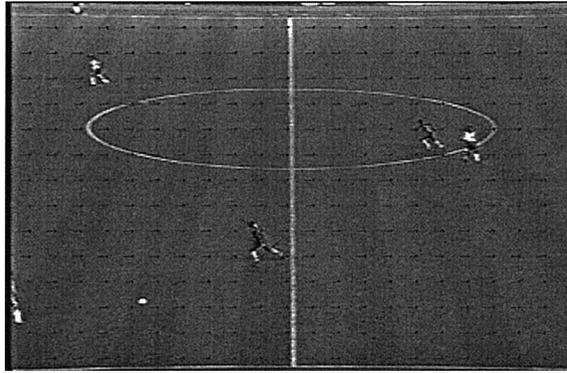


図 8 カメラワーク認識

このオプティカルフローの中央値を使って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か検証する。1 フレーム目の手作業で決定した基準点に対してオプティカルフローの中央値を加算して 10 フレーム分移動させる。10 フレーム目の基準点を手作業で決定する。オプティカルフローの補完によって算出した位置と本来の 10 フレームの基準点の位置を比較した。その位置の誤差はおおよそボール 1 個分であり、この誤差は 1 m 以内であるので許容範囲内である。

この結果より、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オプティカルフローを使って補完することが可能となることが判った。

サッカー中継映像の中には複数のカメラアングルが存在する。このカメラアングルの種類とサッカーのプレーの関係について調べる。サッカーはボールがフィールド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しプレーが止まる。この時間をアウトオブプレーと呼ぶ。その逆でボールがフィールド内にありプレーが継続している時間をインプレーと呼ぶ。

実際のサッカー映像がどのカメラアングルから撮影されたかを調べた。インプレーが 62% でアウトオブプレーが 38% である。さらに、インプレーのうちでセンターカメラが映す映像の割合は 94% になる。このことから、ゲーム分析するにあたって十分な情報が得られると判断し、本研究における実装では、画像解析対象をセンターカメラに限定する。

センターカメラ以外のカメラとして、スタンドから選手の全身をズームで映すカメラがある。この映像を特定する方法を考える。スタンドから選手の全身をズームで映すカメラの場合は、映像の中心に選手認識で用いたテンプレートよりも大きな画素が存在する。ベンチ内や交代選手を映すハンディカメラの場合は、グラウンドを示す緑色の画素が少ないことが挙げられる。他のカメラアングルにおいても同様の特徴をもつことが判った。

そこで、センターカメラの映像のみを画像解析対象とするため、画面中央 (x 方向, y 方向とも $1/3 \sim 2/3$ の範囲) に選手のテンプレートより大きな画素があった場合とグラウンドを示す緑色が映像の半分以下だった場合は除外することにした。この方法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る。

5.6 プレー認識

サッカーは手を使わずにボールを相手ゴールに入れるスポーツである。これは、選手とボー

ルが接触することによりボールを移動させていることである。その接触の仕方や回数によってプレーが認識される。この知識を用いて、本研究において画像解析によるサッカープレーの認識方法を定義する。選手とボールの接触、重なり、すり抜けのパターンを使用する。このパターンより、パス、トラップ、ドリブル、スルーを認識させた。

プレーは基本的に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。

プレーは同一選手のボールタッチ回数として表現するため、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。そのために選手とボールの移動可能範囲を設定する。この移動可能範囲にある選手やボールを連続するフレーム間で追跡を続け、次にボールタッチが起きたときに、前回と同一選手であるかを判定する。

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。このタッチ回数をスコアブックに記録する。

選手とボールが重なって移動する場合はドリブルとして扱う。3タッチとの違いは1タッチ後に選手の周りからボールを認識できなくなり、その後、その同一選手の周りから改めてボールが認識できた場合である。このドリブル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箇所をスコアブックに記録する。

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にかかる例外で、選手の手前や裏側を選手がボールに触れることなく通過する場合を指す。1タッチとの違いは、ダイレクトパスが選手に重なる前後のボールの方向やスピードが変化するのに対して、スルーの場合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ことで判定する。

ボールの方向やスピードを保持しておくために、移動予想範囲を設定する。これは、現フレームと前フレームにおけるボール座標から次フレームのボール座標を予測し、その座標を中心とする半径8ピクセルの円を移動予想範囲とする。ボールと選手が重なる前から移動予想範囲を算出し、重なった後も次フレームの予想位置を現フレームのボール位置と仮定して移動予想範囲の算出を継続する。選手からボールが離れる時に、画像から得られるボールの中心座標が移動予想範囲にあった場合をスルーと判定する。

6. 実 験

実装したツールを用いてサッカー映像からスコアブックを作成する実験を行った。その方法、結果、評価、考察について述べる。

6.1 方 法

本研究の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の作成を行った。FIFAコンフェデレーションズカップ2001のカメルーン vs 日本とブラジル vs カナダの試合である。それぞれのシーンは以下のとおりである。

- ・選手がボールを受けてパスを出す。
- ・ゴール前の混戦で敵味方がボールを奪い合う。

- ・選手がボールを受けてドリブル後パスを出す .
- ・選手が二つのパスをつなぐ .
- ・選手がシュートしキーパーがボールを弾く .
- ・選手がドリブル突破をする .

実験対象としたシーンの抽出理由は、プレーに幅を持たせ、パス、ドリブル、スルー、シュートを含むように考慮した。また、ゴール前のシーンを入れることにより、選手が混戦状態の認識率を算出した。実験に使用した機材を表 2 に示す。OS はいずれも Microsoft Windows 2000 Professional + Service Pack 1 である。

表 2 実験機材と処理時間

HW	CPU	Memory	Disk	処理時間
SONY VAIO RX	Intel Pentium III 1GHz	384MB	120GB	90分
IBM ThinkPad X20	Intel Pentium III 600MHz	192MB	20GB	120分
IBM ThinkPad X20	Intel Celeron 500MHz	192MB	20GB	170分

6 2 結 果

評価方法としては手作業で作成したスコアブックと比較し認識率を求めた。具体的には手作業で作成したスコアブックと自動化ツールのスコアブックを比較し、その一致率を認識率としている。

時間、場所、チームは手作業のスコアブックに対して一致している割合を認識率として表す。プレーは、検索もれや検索ミスを考慮して、手作業のスコアブックに対して再現率と適合率を用いる。再現率 (R) は検索されるべきデータのうち検索されたデータの割合 (検索もれの少なさ) を表し、適合率 (P) は検索されたデータのうち望ましいデータの割合 (検索ミスの少なさ) を表す。その結果を表 3 に示す。

表 3 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%
平均	100%	100%	52%	82%	80%

6 3 評 価

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% しかなかった。再現率、適合率とも 95% 程度を示せば充分であると言われているが、チームの認識率とプレーの再現率、適合率は満足な値でなく、改善が必要である。

チームとプレーの認識率が低いとチームの戦略分析を行う上での障害となる。特にチーム

認識率が 52% というのは影響が大きい。この点について考察を行う。

6.4 考 察

チームとプレーの認識率の低下原因を調べる。また、このツールが複数の試合会場や放送局に対応可能なことが判ったが、対応できない試合会場やシーンについても調査した。時間と場所の認識率が高い理由は二つある。

- ・時間はフレーム枚数から算出されるので誤差が発生しない
- ・場所はフレーム間予測を用いて補完しているので誤差が少ない

選手認識の問題点として、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、グラウンドの奥でプレーする時に誤差が発生することが上げられる。チームの認識率とプレー認識率が低い原因は六つある。

- ・ゴール前の混戦の場合は選手同士が重なり合う。特にヘディングで競り合う場合が難しい
- ・ユニフォ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
- ・ユニフォームの色が緑でグラウンドの色に近い
- ・スルーだった場合もチームを変更しているため
- ・直線的なドリブルはスルーとして誤認識する
- ・選手認識で実際の画素とテンプレートの大きさの違いにより細かいプレーを認識できない

Scene 2 のチーム認識率が 0% なのは、ゴール前の混戦で敵味方がヘディングでボールを奪い合うシーンである。これは人間の解説者でも間違える可能性があるため、敵味方を判別しなくても戦略分析を行う知識を導入する必要がある。

6.5 模 式 図

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 9 に示す。これは scene 1 の結果で、連続したプレーを模式化することができた。

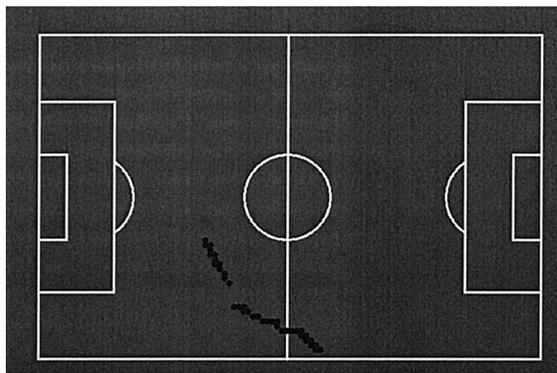


図 9 ボールの移動軌跡

ツールの処理速度が遅いため、ゲーム全体でなく数秒間のプレーを模式化している。それで

も、選手がボールを受けてパスを出している様子を把握できた。処理速度が向上されてゲーム全体を処理できるようになれば、サッカーの戦略を分析することは可能であると考えている。

実験結果と模式図より、解説者の代わりにプレー認識を自動化し、スコアブックを用いてサッカーの戦略分析することが可能なことが判った。

7. 課 題

本章は今後の課題として、ツールの処理時間、手作業の自動化について述べる。

7.1 処 理 時 間

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。これは、認識率向上を主目的としたため、1ピクセル単位で全て解析した結果である。

具体的にはCPUの処理速度に依存し、表2に示す処理時間が必要となった。今回の実験で特別なハードウェアを用意しなかった理由は、一般視聴者の家庭にあると思われる仕様のパソコンで処理することにより利用者の拡大を図るためである。

処理時間が長いことの対策として、認識率低下を招くことなく、画像処理の簡素化が必要となる。つまり、画像処理において手間をかけることなく認識率を保てれば処理時間の高速化を図ることができると考えている。目標は、手作業で5分間のシーンを5時間で処理できるので、この程度にしたいと思う。現状は、5分間のシーンを処理するのに136時間必要で、手作業の27.2倍かかる。

また、処理を分割することにより並列処理を行うことで処理時間を短くすることが可能だと考える。分割可能なシーンはアフィン変換の基準点4点の決定可能な画像から始まるシーンであれば良く、まったく別のシステムで処理を行うことが可能である。今後のCPU処理速度の向上や分散並列処理を行うグリッドコンピューティングの発達などで本システムの処理速度は改善されると思われる。

7.2 手作業の自動化

選手認識のため、2値化閾値（輝度）の算出とユニフォームの色（色相）の決定は手作業で行っている。選手認識で作成されるYHSのヒストグラムを活用し、グラウンドの色の明るさとユニフォームの色の種類を決定する。

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、アフィン変換の基準点4点の決定を手作業で行っている。2フレーム目以降はオプティカルフローで補完するが、1フレーム目は手作業で確認する必要がある。サッカー画像ファイルを開き、基準点となる4点を決定し、画面座標 (x, y) を取り出す。それに対応するグラウンド座標 (X, Y) を模式化された 640×480 ピクセルのグラウンドから取り出す。合計八つの座標を手作業で抽出する。

1フレーム目の手作業で決定しているパラメータの自動計算は次の方法で対応できると考えている。選手認識の閾値の決定と視点移動の基準点4点の決定については、以下の六つの方法である。これらの方法は関連研究で有効に活用されているので、本研究にも取り入れる必要がある。

- ・ヒストグラムのクラス間分散を利用する。

- ・ フレーム間予測より動的オブジェクトとして抽出する．
- ・ 判別基準に基づく画像の2値化を利用する．
- ・ 画素間の差分（微分）とエッジ抽出技術であるゼロクロッシングを用いる
- ・ ハフ変換で求める直線の交点や円弧の頂点を利用する．
- ・ ライン抽出は逆ハフ変換の時に一定の幅を持たせる．

8. おわりに

映像のメタデータ作成するために映像の内容を理解し構造化する必要がある．しかし、画像解析だけを用いた方法は映像の内容を理解することができないので、手作業に頼る方法が取られている．今後、増え続ける映像に限られた予算や人手で対応しきれないと思われる．

そこで、サッカー映像からその内容を取り出すための知識を最も持っているのはサッカー解説者なので、サッカー映像の自動ゲーム分析方法として、画像解析にサッカー解説者の知識を活用したシステムを提案した．

提案手法は、プレーの認識、スコアブックの作成、チーム戦略の分析で構成される．プレー認識を行うために画像解析とサッカー解説者の知識を用いて、選手認識、ボール認識、ライン認識、視点移動、カメラワーク認識を行った．スコアブックは時間、場所、チーム、プレーを数値化したデータで管理する．ボールキープ時間やボール移動軌跡からチーム戦略の分析を行った．プレー認識の画像解析は、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た．

本研究の結論として、認識率の向上、処理時間の高速化、1フレーム目に対するパラメータの自動計算を行う必要があるが、サッカー映像の自動ゲーム分析方法として、画像解析にサッカー解説者の知識を活用したシステムが有効である．

その理由は、画像解析だけの場合は選手とボールを追跡して位置を認識するだけであったが、画像解析にサッカー解説者の知識を加えた場合は選手とボールの位置に加えてプレーを認識することが可能になった．提案手法は場所とプレーを認識できるので、グラウンド俯瞰図に画像変換するだけでなく、手作業で行われていたプレー解説やチーム戦略の分析を行うことができる．さらに、シュートチャンスを把握することにより、ダイジェスト番組や番組表に役立つデータを抽出可能である．これはメタデータ作成のため映像内容を理解し構造化する一例を示すことができたとと言える．

また、手作業の場合1シーン平均6.3プレーの作業が必要であったが、本システムは1フレーム目だけの処理を行えば良いので、84%の作業量削減が可能である．さらに、センターカメラのアンクルのみを使用するので38%のアウトオブプレーを除外することができた．

1フレーム目の作業である色と基準点の決定はサッカー解説者のような専門知識を必要とせず、だれでも行うことができるため、最小限の機材と人手で作業をまかなうことが可能となった．

今後の発展としては、解説者のように模式図を解説して一言に要約する機能、より多くのサッカーの知識を取り込むこと、画像処理だけでなく音声処理を追加し歓声の大きさの変化を使用することが考えられる．

本研究を行うにあたり、ご指導いただいた奈良先端科学技術大学院大学情報科学センター砂原秀樹教授、鳥取環境大学情報システム学科今井正和教授、奈良先端科学技術大学院大学付属

図書館研究開発室羽田久一助手に感謝する。また、本論文の執筆にあたり、ご指導いただいた日本ユニシス株式会社第一ソフトウェアサービスセンター鹿毛公氏、古澤純一氏、インテグレーションサービス部羽田昭裕氏に感謝する。

-
- 参考文献**
- [1] 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
 - [2] Orad 社 <http://www.orad.co.jp/>
 - [3] 山田明人, 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
 - [4] 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イベントの抽出, 電子情報処理学会, IE 99 17, 1999
 - [5] 大橋二郎, 田嶋幸三, 掛水隆, サッカーゴールへの科学, 東京電機大学出版局, 1997
 - [6] Charles Hughes, サッカー勝利への技術戦略, 大修館書店, 1996
 - [7] 中川靖士, 羽田久一, 今井正和, 砂原秀樹サッカー映像の自動ゲーム分析, 情報処理学会研究報告マルチメディア通信と分散処理, 2002 DPS 106:193-198, 2002
 - [8] 中川靖士, サッカー映像の自動ゲーム分析方式の提案と評価, Master's these, 奈良先端科学技術大学院大学, 2002

執筆者紹介 中川靖士 (Yasushi Nakagawa)
1967年生。2002年3月奈良先端科学技術大学院大学情報科学研究科修了。1991年4月日本ユニシス(株)入社。UNIXシステムの受入評価とシステム検査に従事後, Windowsシステムの技術情報提供とアプリケーション開発業務に従事。現在, ビジネスアグリゲーション事業部マーケティング室に所属。