

## 多視点サッカー映像からの自由視点映像生成

林 邦彦<sup>†</sup>      齋藤 英雄<sup>†</sup>

<sup>†</sup> 慶應義塾大学理工学部情報工学科

近年、スポーツ映像などで自由な視点からの映像の要求が高まっている。従来手法 [5] では、サッカースタジアムのような大規模な空間で、カメラの中間視点画像しか合成できなかった。この従来手法は、空間の3次元モデルを復元をせずに、入力画像間の対応点を算出することで自由視点画像を生成している。そこで、本稿では、サッカーシーンを幾何的な特徴に応じて、動的領域と静的領域に分割し、各領域を平面近似して、それぞれ3次元モデルを簡易的に復元することによって、中間視点画像に限らない自由視点画像を生成する手法を提案する。本手法では、対象の3次元モデルを復元するために、まず、カメラキャリブレーションを行う。その際、サッカーグラウンドにマーカーを置く事はできないので、グラウンド面の平面パターンを利用してカメラキャリブレーションを行う。分割された各領域のうち、選手やボールである動的領域がグラウンドに垂直に立つ1枚の板(ビルボード)で表されるものとして3次元モデルを復元する。一方、静的領域に対しては、複数の平面領域に分割し、各平面の3次元モデルをあらかじめ復元する。本手法をサッカーシーンに適用し、従来生成できなかった中間視点以外の自由視点画像を生成することで有用性を示す。

## Synthesizing free-viewpoint images from multiple view videos in soccer stadium

Kunihiko Hayashi<sup>†</sup>, Hideo Saito<sup>†</sup>

<sup>†</sup> Department of Information and Computer Science, Keio University

We propose a new method for synthesizing free-viewpoint images from multiple view videos in soccer stadium. The previous method[5] can synthesize virtual viewpoint images at intermediate viewpoint between real cameras. The previous method estimates only pixel-wise correspondences between input images for synthesizing virtual viewpoint images without a 3D reconstruction. In this paper, we propose a new method that can synthesize virtual viewpoint images at free-viewpoints which are not limited at intermediate viewpoint by reconstructing simple 3D models after dividing a soccer scene into static regions and dynamic regions. To reconstruct the object soccer scene, we perform camera calibration without marker objects because we can't put such objects on the soccer ground. For reconstructing of dynamic regions, we employ a billboard representation of dynamic regions. The billboard representation can realize a simple 3D reconstruction and transform textures into the appropriate appearance of the aspect of dynamic regions at virtual viewpoint. For reconstructing of static regions, we divide two planes that are a ground region and a background region. We reconstruct 3D models in each region beforehand. By applying our method to a soccer scene, we can synthesize free-viewpoint images which are not limited at intermediate viewpoint by employing only pixel values of real images.

## 1 はじめに

近年、画像処理技術の発展に伴って、臨場感のあるスポーツ中継が行われるようになってきた。例えば、サッカーのオフサイドライン表示 [1] やフリーキック時のゴールまでの距離表示、競泳時に世界記録を示すラインの表示、などがある。しかし、現在のテレビ中継では、実カメラで撮影した映像のみが提供されていて、視聴者が視点を選択することはできない。このような背景から、コンピュータビジョンの分野において、ダイナミックなイベントを対象として、多視点カメラによって撮影された映像をもとに自由に視点を变化させた映像を提示する研究が盛んに行われている [2, 3, 4, 5, 6, 7, 9]。

金出らは、ステレオ視を用いてシーンの3次元モデルを正確に復元し、自由視点画像を生成する手法を提案した [9]。しかし、大規模な空間に対して、正確な3次元モデルを復元するためには、多くのカメラが必要であり、計算量がかかるという問題を抱えている。

一方で、正確な3次元モデルを復元せずに、複数の入力画像間における対応点を移動して自由視点画像を生成する手法 [5, 7] や、計算量を抑えるために簡易的な3次元モデルを復元し、大規模な空間で自由視点画像を生成する手法 [6] が提案されている。

稲本らは、カメラ間の射影幾何関係を求める弱校正を行い、複数の入力画像間における対応点を移動することによって入力画像間の中間視点画像を生成する手法を提案した [5]。この手法は、サッカーシーンをグラウンド領域、背景領域、選手とボールの動的領域という3つの領域に分割し、各領域の幾何学的な性質を利用している点が特徴的である。しかし、この手法は、視点の範囲が入力画像間に制限されていた。その理由は、3次元モデルを復元していないことである。また、3次元モデルを復元するために、カメラの強校正をしていないことが挙げられる。

亀田らは、簡易的な3次元モデルを復元することにより、視点の範囲が入力画像間に限られるという問題を解決し、リアルタイムでユーザが好きな視点からサッカーの試合を鑑賞する手法を提案した [6]。しかし、選手やボールなどの動的な領域を除く静的領域は、CGを用いて生成されているため、リアリティが若干欠けるといえる問題がある。

木村らは、テニスシーンに対して、実画像の明るさのみを用いて、選手視点画像を生成する手法を提案した

[7]。しかし、この手法では、仮想視点内にテニスコート上の特徴点数点が常に見えている必要があり、広い空間での自由視点画像生成は難しい。また、適用されたテニスシーンには、選手が2人とボールが1つであり、サッカーのように多くの動的領域が存在する場合には対応していない。

そこで、本稿では、多視点サッカー映像から自由視点映像を生成する手法を提案する。本手法は、入力画像間に制限されず、CGを用いず、かつ、テニスコートより広いサッカー競技場という空間内で、自由視点映像を生成する。本手法は、従来手法 [5] に基づいて、サッカーシーンをグラウンド領域、背景領域、選手とボールの動的領域という3つの領域に分割し、各領域の幾何学的な性質を利用して、それぞれの領域の簡易的な3次元モデルを復元することにより、中間視点に限らない自由視点画像を生成する。また、3次元モデルを復元するために従来手法 [5] では行われなかったカメラの強校正を行う。カメラの強校正を行う際、サッカーグラウンドにマーカーを置く事はできないので、グラウンド面のパターンを利用したカメラキャリブレーションを行う。

分割された各領域のうち、選手やボールである動的領域はグラウンドに垂直に立つ1枚の板（ビルボード）で表されるものとして3次元モデルを復元する [6]。なお、選手を表すビルボードの平面は、サッカーグラウンドを真上から見た場合に、ビルボードと仮想視点を結ぶ直線に対して垂直であるものとする。このビルボード表現によって、選手やボールの3次元モデルを簡易的に復元することができる。しかし、実際の選手は平面形状ではなく、3次元形状をしている。そのため、3次元形状を2次元平面に近似したことに起因するひずみをできるだけ小さくするために、仮想視点が各ビルボードを見込む方向にできるだけ近い実カメラを選択し、ここから動的領域のテクスチャをビルボード上にレンダリングする。

グラウンド領域、背景領域に対しては、ホモグラフィを利用して、自由視点画像を生成する。

本手法をサッカーシーンに適用し、従来生成できなかった中間視点以外の自由視点画像を生成することで有用性を示す。また、従来生成できなかったサッカー選手から見た選手視点画像を生成することで有用性を示す。

## 2 提案手法

### 2.1 撮影環境

撮影環境は、図1のように4台の固定カメラを配置した。この撮影環境の下、サッカーの試合を撮影し、この動画像を入力画像とした。入力画像はすべて720x480画素、24-bit-RGBのカラー画像である。入力画像の例を図2に示す。

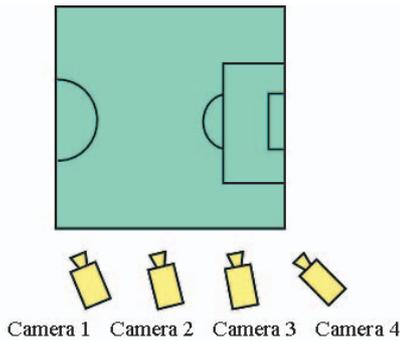


図 1: 撮影環境

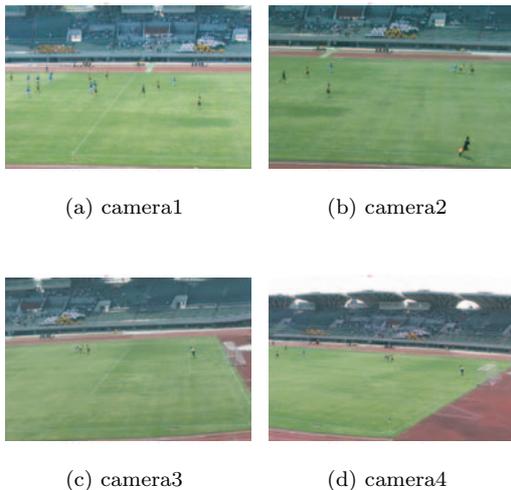


図 2: 入力画像

### 2.2 カメラキャリブレーション

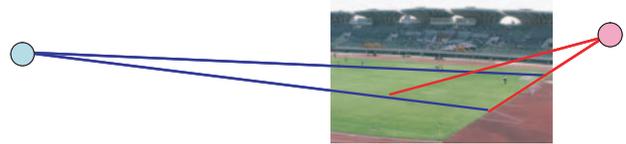
対象の3次元モデルを復元するために、カメラキャリブレーションを行う。一般的なカメラキャリブレーション手法では、3次元情報が既知な物体を置く必要がある。しかし、サッカーグラウンドにこのような物体を自由に置くことは困難である。そこで、本研究においては、このような物体を置かず、図3のように、グラウンド上に描かれた2組の平行線の交点から生成

される消失点を用いてカメラキャリブレーションを行う [10]。

カメラキャリブレーションとは、カメラの射影行列  $P$  を算出することである。カメラの射影行列は、カメラの内部パラメータ  $A$  と外部パラメータで構成されている。外部パラメータは、回転行列  $R$  と並進ベクトル  $t$  で構成されている。

まず、内部パラメータを算出する。2組の消失点に関して、画像上の座標を  $(x_1, y_1)$ ,  $(x_2, y_2)$  とすると、式 (1) が成り立つことが知られている。この式を用いて、内部パラメータの構成要素である焦点距離  $f$  を求め、内部パラメータを算出する。

$$x_1 y_1 + x_2 y_2 + f^2 = 0 \quad (1)$$



●, ● : Vanishing points

図 3: 消失点

次に、外部パラメータを算出する。ここでは、外部パラメータを算出するために、入力画像と上部仮想視点画像 (図7(a)) 間のホモグラフィ  $H$  を用いる。上部仮想視点画像とはサッカーグラウンドを真上から見た画像のことである。

(2) 式のように、ホモグラフィ  $H$  は、画像 (平面)  $I_1$  上の座標を別の画像 (平面)  $I_2$  上の座標に変換する平面射影行列である。ホモグラフィは、 $I_1, I_2$  間の4点以上の対応点から算出することができる。(2) 式の  $x_1, x_2$  はそれぞれ画像  $I_1, I_2$  上の座標である。

$$x_2 \simeq H x_1 \quad (2)$$

現実空間のサッカーグラウンドの広さや規格は競技場ごとに決まっているので、特徴点 (サッカーグラウンド上に引かれているラインの交点) のグラウンド平面における2次元座標を獲得することができる。一方、これらの特徴点の画像平面における2次元座標も手動で獲得することができる。このようにして得られるグラウンド平面と画像平面間の2次元座標の対応関係を4組以上用いて、入力画像内に撮影されているサッカーグラウンドと上部仮想視点間のホモグラフィを算出することができる。

次に、このホモグラフィを用いて外部パラメータを算出する。式 (3) のように回転行列  $R$  は、3 つの列ベクトルとして表すことができるので、 $A, R, t, H$  の関係式は、式 (4) で表される。

$$R = [r_1 \ r_2 \ r_3] \quad (3)$$

$$A[r_1 \ r_2 \ t] = H \quad (4)$$

式 (4) で、 $A, H$  は既知なので、 $r_1, r_2, t$  が算出できる。ここで、回転行列  $R$  の外積の性質を (式 (5)) を利用して、 $r_3$  を求める。この行列演算により外部パラメータを算出する。

$$r_1 \times r_2 = r_3 \quad (5)$$

## 2.3 手法の概要

本手法の目的は、従来生成できなかった中間視点に限らない自由視点映像を生成することである。そのためには、対象空間の 3 次元モデルを復元する必要がある。しかし、サッカー競技場は大規模な空間なので、少数のカメラで撮影した入力画像から、一般的なステレオ法などで 3 次元モデルを復元するのは困難である。

そこで、本手法では、対象空間を静的な領域と動的な領域に分け、それぞれの領域を平面で近似することで、3 次元モデルを復元する。静的領域はグラウンド領域と背景領域に分割される。動的領域は選手とボールである。そして、このように近似された 3 次元モデルを用いて自由視点画像を生成する。

本手法のフローチャートを図 4 に示す。

まず、各カメラの入力画像を手動でグラウンド領域、背景領域、動的領域に分割する。

分割された各領域のうち、選手やボールである動的領域がグラウンドに垂直に立つ 1 枚の板 (ビルボード) で表されるものとして 3 次元モデルを復元する [6]。なお、選手を表すビルボードの平面は、サッカーグラウンドを真上から見た場合に、ビルボードと仮想視点を結ぶ直線に対して垂直であるものとする。このビルボード表現によって、選手やボールの 3 次元モデルを簡易的に復元することができる。しかし、実際の選手は平面形状ではなく 3 次元的な形状をしている。そのため、レンダリングの際に、3 次元形状を 2 次元平面に近似したことに起因するひずみをできるだけ小さくする必要がある。このひずみを最小にする処理が最適なカメラの選択である。

静的領域では、時間が変化しても静的領域は変化しないと仮定し、あるフレームにおける 3 次元モデルを生成し、これを自由視点映像のすべてのフレームにおける静止領域として利用するものとする。

グラウンド領域は、世界座標系の  $XY$  平面に存在するものとし、ホモグラフィを用いて生成した上部仮想視点画像 (図 7(a)) を用いて、自由視点画像を生成する。

背景領域は、三角測量の原理から入力画像の背景領域内の 4 点の 3 次元座標を復元する。さらに、ホモグラフィを用いて自由視点画像を生成する。

最後に 3 つそれぞれの領域で生成された画像を合成し、自由視点画像を生成し、これを各フレームごとに繰り返し替えることにより自由視点映像を生成する。

### 2.3.1 動的領域の自由視点画像生成

この処理では、入力画像内の選手とボールの座標が必要である。そのため、あらかじめ入力画像内の選手とボールの座標を含むデータベースを構築しておくものとする。本稿の目的は選手の追跡ではなく自由視点映像生成であるため、動的領域の追跡は手動で行った。しかし、[8] のような選手の自動追跡を適用することもできる。また、自由視点映像生成のために必要となる動的領域のテクスチャを抽出するために、あらかじめ背景差分を行い、データベースを構築しておくものとする。

前述のように、選手やボールである動的領域に対しては、図 6 に示したように、動的領域がグラウンドに垂直に立つ 1 枚の板 (ビルボード) であると考えられる。このビルボードの平面は、サッカーグラウンドを真上から見た場合に、ビルボードと仮想視点を結ぶ直線に対して垂直になるように配置するものとする。

しかし、実際の選手は平面形状ではなく、3 次元的な形状をしている。そのため、ビルボードを用いて自由視点をレンダリングすると、3 次元形状を 2 次元平面に近似したことに起因するひずみが発生する。このひずみをできるだけ小さくするために、仮想視点が各ビルボードを見込む方向にできるだけ近い実カメラを選択し、ここから動的領域のテクスチャをビルボード上にレンダリングする。

例えば、図 5 のように、仮想視点と選手やボールの位置を線で結び、この直線と最も近いカメラを選択する。図 5 では、ボールと仮想視点を結ぶ直線は、カメラ 1 と最も近いので、ボールのテクスチャはカメラ 1

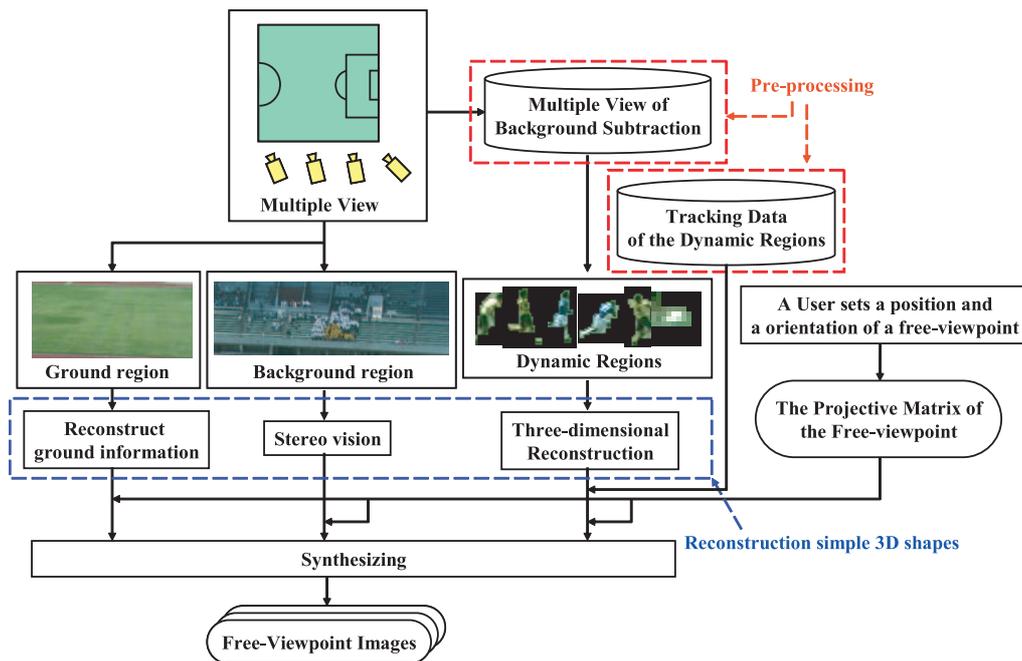


図 4: 手法の概要

から選択する．同様に，青色のユニフォームを着た選手に関して，この選手と仮想視点を結ぶ直線は，カメラ 2 と最も近いので，ボールのテクスチャはカメラ 2 から選択する．このようにして，各動的領域ごとにビルボードのレンダリングに用いるテクスチャが映っているカメラを選択する．選択したカメラに動的領域が映っていない時・オクルージョンが発生した時には，隣のカメラを選択するか，複数の選手を同一選手と見なす．

最後に，自由視点の射影行列と 3 次元モデルを用いて，動的領域の自由視点画像を生成する．

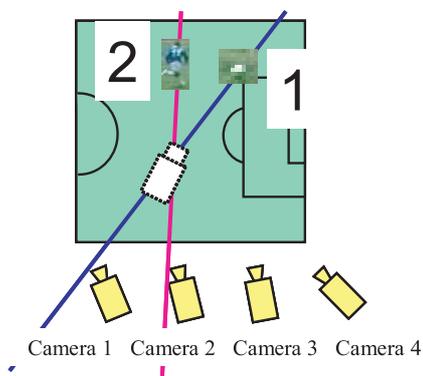


図 5: 最適なカメラの選択

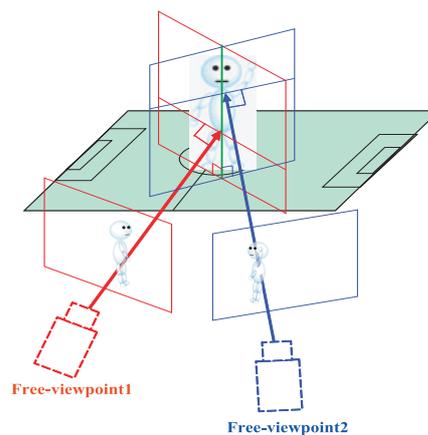


図 6: ビルボード法

### 2.3.2 グラウンド領域の自由視点画像生成

グラウンド領域に対しては，3 次元モデルを復元するために入力画像と上部仮想視点画像間のホモグラフィを用いる．グラウンド領域は世界座標系の XY 平面に存在するものとし，2.2 章で算出した入力画像と上部仮想視点画像（図 7(a)）間のホモグラフィを用いて図 7(a) のような 3 次元モデルと等価である上部仮想視点画像を生成する．その際，一枚の入力画像ではサッカーグラウンド全体をカバーできないので複数の入力画像を用い，それぞれ生成された上部仮想視点画像を切り

出し、つなぎ合わせることによって、上部仮想視点画像を生成する。

最後に、自由視点カメラの射影行列を用いてこの上部仮想視点画像の明度値を自由視点画像に挿入し、図7(b)に示すようなグラウンド領域の自由視点画像を生成する。

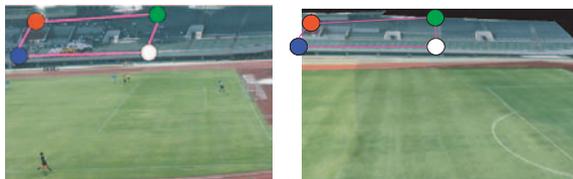


(a) 上部仮想視点画像 (b) グラウンド領域の自由視点画像

図 7: グラウンド領域の自由視点画像生成

### 2.3.3 背景領域の自由視点画像生成

背景領域に対しては、入力画像の背景領域(図8(a))と自由視点画像の背景領域(図8(b))間のホモグラフィを用いて、自由視点画像をレンダリングする。そのため、入力画像の背景領域と自由視点画像の背景領域間で4点以上の対応点が必要となる。4点の3次元座標が既知ならば、自由視点の射影行列を用いて、自由視点画像に4点を投影し、4点の対応点が算出できる。また、この4点の3次元座標を算出することは背景領域の3次元モデルを簡易的に復元したことに等価である。本手法は、カメラキャリブレーションをしているので、三角測量の原理により4点の3次元座標をあらかじめ算出する。三角測量の際、2枚の入力画像間の対応点4点は手動で対応付ける。最後に、このホモグラフィを用いて、入力画像の背景領域を自由視点画像に挿入し、背景領域の自由視点画像を生成する。



(a) camera2 (b) 背景領域の自由視点画像

図 8: 背景領域の自由視点画像生成

## 3 実験と結果

サッカーシーンに対して本手法を適用し、中間視点以外の自由視点画像と様々な位置からの自由視点画像を生成することで有用性を示す。また、従来実現できなかった選手視点画像を生成することで有用性を示す。

### 3.1 自由視点画像生成

本手法では、ユーザが自由視点の位置と方向を決めることができる。実験では、図9のように自由視点カメラを配置し、またすべての自由視点カメラが同じ1点を見ているように方向を固定して、各自由視点カメラにおける自由視点画像を生成した。図11に結果画像を示す。図11のキャプションの座標は、世界座標系を図10のように設定した場合の自由視点カメラの位置の座標である。図11(a)~図11(d)と入力画像(図2)を比較して、中間視点と同じようなカメラの動きに対しても自由視点画像が生成できていることがわかる。図11(e)~図11(h)と入力画像(図2)を比較して、サッカーフィールド内部からの自由視点画像を生成することで、中間視点以外の自由視点画像を生成できていることがわかる。図11(i)~図11(l)と入力画像(図2)を比較して、明らかに高い位置からの自由視点画像を生成することで、中間視点以外の自由視点画像を生成できていることがわかる。このように、本手法を用いることにより、中間視点画像だけでなく、従来手法[5]では生成できなかった中間視点以外の自由視点画像を生成することができる。また、ユーザが自由視点の位置と向きを設定することで従来手法[5]よりも自由度の高い自由視点画像を生成することができる。

### 3.2 選手視点画像生成

試合中、選手が常にボールを見ているという仮定を用いると、本手法を適用することにより選手視点画像を生成することができる。選手視点の位置の座標は世界座標系における選手の座標であり、選手視点の方向の座標は、世界座標系におけるボールの座標である。したがって、動的領域の追跡位置が保存されているデータベースから選手視点の射影行列を算出できる。この選手視点の射影行列を用いて選手視点画像を生成する。生成された選手視点画像を図12に示す。図12(a)は、注目する選手を示している。この選手からの選手視点画像が、図12(b)~図12(d)である。また、図12(b)~

図 12(d) のキャプションは、ユーザが設定した焦点距離である。このように、内部パラメータの構成要素である焦点距離を変えることでユーザの好みに合わせた選手視点画像を生成することができる。

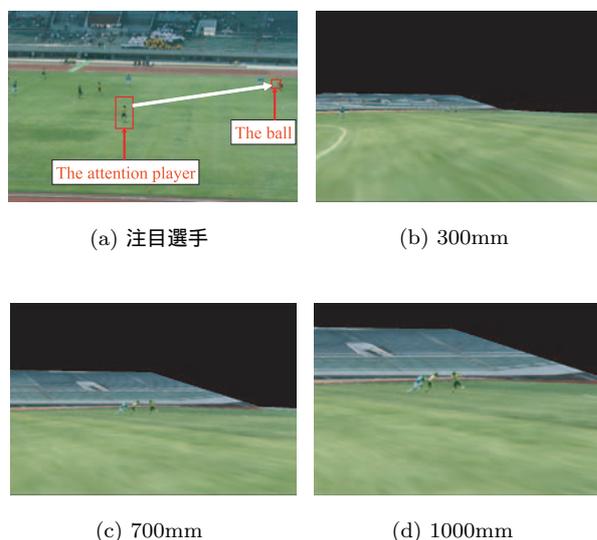


図 12: 生成された選手視点画像

## 4 結論

多視点サッカー映像から自由視点映像を生成する手法を提案した。本手法は、サッカーシーンを幾何的な特徴に応じて動的領域と静的領域に分割し、それぞれ 3 次元モデルを簡易的に復元することによって、中間視点画像に限らない自由視点映像を生成した。本手法では、対象の 3 次元モデルを復元するために、まず、カメラの強校正を行う。その際、サッカーグラウンドにマーカーを置く事はできないので、グラウンド面の平面パターンを利用してカメラの強校正を行う。分割された各領域のうち、選手やボールである動的領域がグラウンドに垂直に立つ 1 枚の板（ビルボード）で表されるものとして 3 次元モデルを復元する。静的領域はさらに、グラウンド領域と背景領域に分割される。グラウンド領域は、ホモグラフィを用いて、3 次元モデルをあらかじめ復元する。背景領域は、三角測量の原理を用いて、3 次元モデルをあらかじめ復元する。

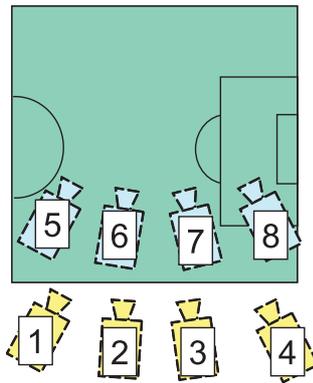
実際にサッカースタジアムで撮影したサッカーの多視点映像に対して本手法を適用し、中間視点画像だけでなく、従来手法 [5] では生成できなかった中間視点

以外の自由視点画像を生成することで本手法の有効性を示した。

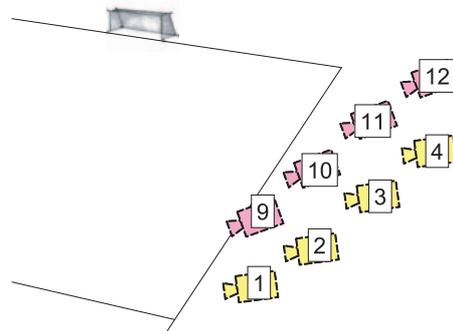
最適なカメラの選択に対する誤差減少、テクスチャの滑らかな変形、動的領域の自動追跡、背景差分の精度向上、テクスチャの高解像度化、詳細な 3 次元モデルの構築、キャリブレーションの精度向上、リアルタイムのレンダリングなどが今後の課題として挙げられる。

## 参考文献

- [1] CyberPlay url: <http://www.orad.co.il>.
- [2] I.Kitahara, Y.Ohta, H.Saito, S.Akimichi, T.Ono, T.Kanade. "Recording Multiple Videos in a Large-scale Space for Largescale Virtualized Reality" Proc. Of International Display Workshops (AD/IDW'01), pp.1377-1380 (2001)
- [3] D.Snow, O.Ozier, P.A.Viola, W.E.L.Grimson. "Variable Viewpoint Reality" NTT R&D, Vol.49, No.7, pp.383-388 (2000)
- [4] S.Yaguchi and H.Saito. "Arbitrary View Image Generation from Multiple Silhouette Images in Projective Grid Space" Proc. of SPIE Vol.4309 (Video-metrics and Optical Methods for 3D Shape Measurement), pp.294-304 (2001)
- [5] N.Inamoto and H.Saito. "Intermediate View Generation of Soccer Scene from Multiple Views" ICPR2002, Vol.2, pp.713-716 (2002)
- [6] Y.Kameda, T.Koyama, Y.Mukaigawa, F.Yoshikawa and Y.Ohta. "Free Viewpoint Browsing of Live Soccer Games" ICME, pp.747-750 Vol.1 (2004)
- [7] K.Kimura and H.Saito. "Video synthesis at tennis player-viewpoint from multiple view videos" IEEE VR2005, pp.281-282 (2005)
- [8] S. Iwase and H.Saito. "Parallel tracking of all soccer players by integrating detected positions in multiple view images" ICPR2004, 23-26, pp.751 - 754 Vol.4 (2004)
- [9] T.Kanade, P.J.Narayanan, P.W.Rander. "Virtualized reality: concepts and early results" Proc. of IEEE Workshop on Representation of Visual Scenes, pp.69-76 (1995)
- [10] Gilles Simon, Andrew W.Fitzgibbon, Andrew Zisserman. "Markerless tracking using planar structures in the scene." Proc. of the International Symposium on Augmented Reality, pp.120-128 (2000)



(a) 自由視点カメラの配置 1



(b) 自由視点カメラの配置 2

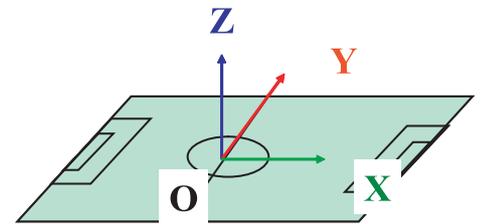


図 10: 世界座標系

図 9: 自由視点カメラの位置



(a) 自由視点カメラ 1  
(X=0m, Y=-50m, Z=10m)



(b) 自由視点カメラ 2  
(X=18m, Y=-50m, Z=10m)



(c) 自由視点カメラ 3  
(X=35m, Y=-50m, Z=10m)



(d) 自由視点カメラ 4  
(X=50m, Y=-50m, Z=10m)



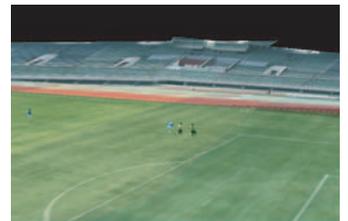
(e) 自由視点カメラ 5  
(X=0m, Y=-25m, Z=10m)



(f) 自由視点カメラ 6  
(X=18m, Y=-25m, Z=10m)



(g) 自由視点カメラ 7  
(X=35m, Y=-25m, Z=10m)



(h) 自由視点カメラ 8  
(X=50m, Y=-25m, Z=10m)



(i) 自由視点カメラ 9  
(X=0m, Y=-50m, Z=25m)



(j) 自由視点カメラ 10  
(X=18m, Y=-50m, Z=25m)



(k) 自由視点カメラ 11  
(X=35m, Y=-25m, Z=10m)



(l) 自由視点カメラ 12  
(X=50m, Y=-50m, Z=25m)

図 11: 生成された自由視点画像