

## サッカーのインターネット中継が可能な自由視点映像方式

古山 孝好<sup>†</sup> 古谷 仁志<sup>†</sup> 向川 康博<sup>†\*</sup> 亀田 能成<sup>†</sup> 吉川 文人<sup>‡</sup> 大田 友一<sup>†</sup>

<sup>†</sup>筑波大学 〒305-8577 茨城県つくば市天王台 1-1-1

<sup>‡</sup>国立スポーツ科学センター 〒115-0056 東京都北区西が丘 3-15-1

E-mail: <sup>†</sup> {furuya,koyama,mukaigawa,kameda,ohta}@image.esys.tsukuba.ac.jp,

<sup>‡</sup> yoshikawa.fumito@jiss.naash.go.jp

**あらまし** 大規模空間におけるイベントであるサッカーシーンを対象に、自由視点映像のインターネット中継を実現する手法について述べる。自由視点映像の生成には3次元復元を用いる手法が一般的である。しかし、インターネット中継を行うには、自由視点映像の生成から、遠隔地にいる複数ユーザへ配信されるまでの処理をリアルタイムに行う必要がある。我々は人物ビルボードというCG表現で自由視点映像を生成し、これをあらかじめ作成した背景のCGモデルにリアルタイムではめることで、自由視点映像のインターネット中継を可能にする手法を提案する。実際にサッカースタジアムを対象とした実験を行い、遠隔地・複数ユーザへのインターネット中継ができることを示した。

**キーワード** 自由視点, 3Dビデオ, リアルタイム, 人物ビルボード, ストリーミング, インターネット配信

## Real-Time Transmission of 3D Video to Multiple Users via the Internet

Takayoshi KOYAMA<sup>†</sup> Hitoshi FURUYA<sup>†</sup> Yasuhiro MUKAIGAWA<sup>†\*</sup> Yoshinari KAMEDA<sup>†</sup>

Fumito YOSHIKAWA<sup>‡</sup> Yuichi OHTA<sup>†</sup>

<sup>†</sup>Department of Intelligent Interaction Technologies, University of Tsukuba 1-1-1 Tennoudai, Tsukuba-shi, Ibaraki,305-8577 Japan

<sup>‡</sup>Japan Institute of Sports Sciences 3-15-1 nishigaoka,kita-ku,Tokyo,115-0056 Japan

E-mail: <sup>†</sup> {furuya,koyama,mukaigawa,kameda,ohta}@image.esys.tsukuba.ac.jp,

<sup>‡</sup> yoshikawa.fumito@jiss.naash.go.jp

**Abstract** This paper presents a visualization method that can transmit live 3D video of a soccer stadium to remote users' PCs. To generate and transmit 3D video via the Internet in real time, each player is represented by "player-billboard". It is composed of one rectangle and one live texture. 3D video is visualized by rendering both player-billboard and background 3D model. In order to show the performance of our method, We captured videos with multiple cameras in a real soccer stadium and We transmitted 3D video to remote multiple users simultaneously.

**Keyword** Free Viewpoint, 3D Video, Real-Time, Player-Billboard, Streaming, Internet, Broadcasting

### 1.はじめに

映像メディアの発達により、我々は、遠隔地で行われるイベントをリアルタイムに試聴することが可能となった。特にスポーツイベントに関しては、ライブ中継の普及により、スポーツを見て楽しむというスタイルが広く浸透してきた。これに伴い、スポーツイベントを自分が望む様な様々な視点から自由に見たいという欲求が高まってきている。

このような要求を解決するために、1つのスポーツイベントにおいて、複数チャンネルの選択権を与える試みがなされている。しかし、この方式においても、視聴者が選択できる視点は、与えられたうちの限られた映像視点の1つにすぎず、自由な視点から映像を見ることができているとは言えない。

そこで、近年、対象空間を多数の視点から撮影し、この多視点映像をコンピュータで解析することにより、自由視点映像を生成する研究が盛んに行われている。

\*現在 大阪大学産業科学研究所

Kanade らが提案した Virtualized Reality は、半径 5m の半球内を対象とし、半球上に設置した 51 台のカメラが半球内の動的なイベントを同期して撮影し、そのデータを Multi Baseline Stereo という手法で解析することにより、対象物体を 3 次元メッシュモデルとして構築し、仮想視点からの映像を生成した[1][2]。しかし、この方法では、映像を生成するための計算コストやデータ量が非常に大きいため、大規模空間への対応やリアルタイムでの映像生成・提示は困難である。

また、Wurmlin らは、人物 1 人を対象として、前方 160° 程度の範囲に配置した 6 台のカメラを用いて 3 次元形状の復元を行った[3]。このシステムでは、物体の 3 次元形状を point-base のデータ形式で表現し、PRk-tree という構造体に格納することにより、データ量の削減と、逐次的な映像描画が可能である。しかし、複数ユーザへは対応しておらず、また、オフラインでの処理を含んでいるため、リアルタイムでの映像生成は困難である。

本研究では、大規模空間で行われるスポーツイベント、特に、サッカーを対象として、複数のユーザが遠隔地にいながらにして、各々が望む自由な視点からの映像を、インターネットを介してライブで中継することができる自由視点映像方式を提案する。

このために、我々は、人物ビルボード表現による自由視点映像生成手法を提案する。本提案手法では、選手を 1 枚の平面で近似し、この平面に多視点映像から獲得した選手映像を適切に貼り付けることにより自由視点映像の生成・提示を行う。そのため、従来研究されてきた多くの手法に比べ、少ない計算コストで自由視点映像を生成することができる。また、少ないデータ量であるため、インターネットを介して自由視点映像を高いフレームレートで配信することが可能である。

## 2.人物ビルボード

本研究で提案する自由視点映像手法では、人物ビルボードを用いることが計算コストとデータ量の削減に貢献している。人物ビルボードとは、各選手を 1 枚の平面で近似し、多視点映像から獲得した選手のテクスチャをこの平面にそのまま貼り付けることにより、各選手を表現する手法である。図 1 に概要を示す。下の赤枠が描画された人物ビルボードを示す。本手法では、選手の 3 次元形状の復元を行わないため、視点の移動による選手の見え方の変化を正確に再現することは難しい。しかし、対象空間を取り囲むように設置した複数のカメラによる多視点映像の中から、ユーザが望む仮想視点位置に最適なテクスチャを抽出することで、視点の移動による物体の見え方の変化を近似的に再現することが可能である。このことは図 2 のように、あ

る程度離れた視点からでは人物ビルボードと 3 次元モデルとの違いはほとんど気にならないことからわか

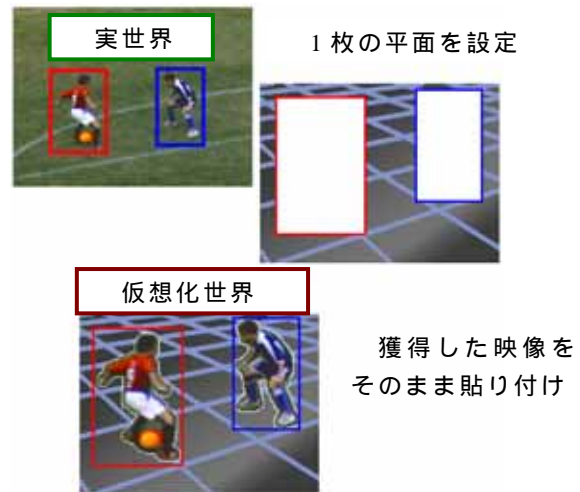


図 1 人物ビルボード

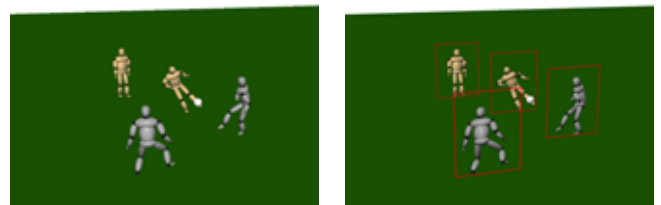


図 2 3 次元モデル(右)と人物ビルボード(左)の見え方の比較

る。人物ビルボードは、選手領域のみを対象とし、サッカースタジアム構造物などの変化のない背景領域に関しては、事前に 3 次元モデルとして作成しておく。この 3 次元モデルの中に人物ビルボードを配置して描画することにより、自由視点映像を生成する。

本手法では、選手の 3 次元形状を復元する処理が不要であるため、自由視点映像を生成するための計算コストを、他の自由視点映像生成手法と比べ、大幅に抑えることが可能である。また、各選手を表すのに必要なデータは、平面の位置、大きさと、テクスチャ画像のみであるため、データ量を小さくすることが出来る。

## 3.インターネット中継システム

本研究で提案する手法に基づいて構築したシステムの概要を図 3 に示す。本システムは、人物ビルボードを用いて、大規模空間における自由視点映像をリアルタイムに生成し、インターネットを介して遠隔地の複数ユーザへ同時配信を行う。システムはシーン解析部、多視点映像撮影部、サーバ部、自由視点映像提示部の 4 つからなる。

システムはまず、シーン解析部に接続されたカメラにより獲得される映像から、選手の 3 次元位置を推定

する。サーバ部は、シーン解析部にアクセスして選手の3次元位置情報を取得し、多視点映像撮影部の各

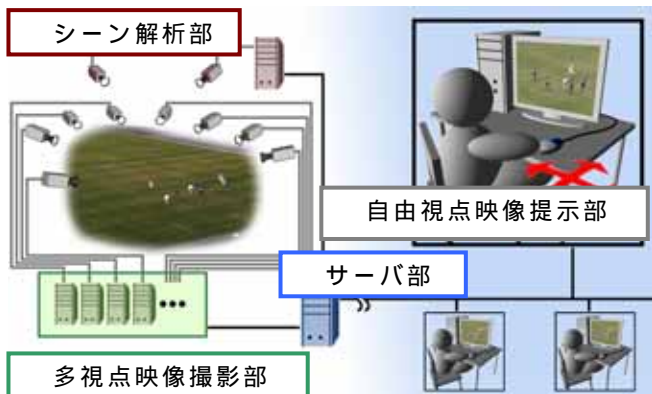


図3 インターネット中継システムの概要

PCへ転送する。多視点映像撮影部の各PCは、得られた3次元位置を基に、それぞれが撮影している映像内の必要な領域に対してのみ、選手のテクスチャ抽出を行い、それをサーバ部へ転送する。サーバ部はこれら全ての選手のテクスチャをまとめ、自由視点映像を生成するためのデータセットの構築を行う。一方、遠隔地にいる各ユーザは自身が望む仮想視点位置情報と時間情報を、インターネットを介してサーバ部に送信する。サーバ部はこの情報を基に、各ユーザへの自由視点映像を生成するために必要なデータを選択し、配信を行う。この繰り返しにより、自由視点映像の提示を行う。このとき、ユーザはマウスなどを用いて任意の時点で自由に視点を変えてよい。

#### 4.自由視点映像生成手法

人物ビルボードを用いた自由視点映像の生成には、以下の3段階の処理が必要である。

- 選手領域の抽出
- 選手の3次元位置の推定
- 選手のテクスチャの抽出

##### 4.1 選手領域の抽出

サッカースタジアムのような屋外環境下では、撮影される映像に日照が与える影響は大きい。そこで、本研究では、屋外での日照変化に対応できる、選手領域の抽出手法を用いる。サッカーシーンにおいて、選手領域はフィールド上のみ存在するとし、撮影画像中のフィールド部分のみについて領域の抽出を行う。

まず、背景モデルを、フィールドを映している画素ごとに1つのガウス分布により表現する。そして、2つのガウス分布で表される背景分布モデルを毎フレーム得られる新たな画素値を用いて逐次更新していくことにより日照変化に対応する。背景であると判断された画素は、新たなデータを基に、学習を行いながらモ

デルの更新を行う。

本手法では、ある画素が長時間選手領域であり背景モデルが更新できない場合を考慮するために、フィールド上の画素について画素値を基にクラスタリングを行う。同一クラスタの背景モデルの変動を考慮することにより、選手領域である画素の背景モデルについても更新を補う。

##### 4.2 選手の3次元位置の推定

各選手の3次元位置を推定するために、対象空間全体を捉え比較的高い位置に、光軸がほぼ直交するように2台のカメラを設置する。その2台のカメラ上で抽出した2つの選手領域の幾何対応をとることにより選手の3次元位置を推定する。

幾何的対応をとるために、各々のカメラにおいて選手領域の外接長方形をフィールドに投影し、その重心とカメラを結ぶ直線を求める。この2つの直線の好転が求める3次元位置である。

##### 4.3 選手テクスチャの抽出

多視点映像からの選手テクスチャの抽出には、推定された各選手の3次元位置の情報を用いる。まず、多視点映像の各カメラに対して、選手の足元位置から人物の大きさに合わせた一定サイズの平面をフィールドとカメラ光軸の両方に垂直になるように設定する。次に、この平面をカメラ画像上に投影することで、矩形領域を求め、この矩形領域のみに対して選手抽出処理を行う。この方法では、選手同士の重なりが生じた場合に不必要な領域を含む可能性がある。そこで、重なりが生じる可能性がある場合には、手前と奥の選手について次のようにしてテクスチャの抽出を行う。手前の選手についてはCollinsらが提案したSpace-Sweep[4]を用いて手前だけの選手領域を求め、テクスチャを抽出する。後ろの選手については手前の選手によって隠された領域のテクスチャを獲得することは原理的に不可能で、他の視点から抽出されたテクスチャを用いることで対応する。

#### 5.実験

国立霞ヶ丘競技場において、サッカーフィールドの半面を対象に撮影実験を行った。

シーン解析部にはカメラを2台使用した。多視点映像撮影部のために観客席の全周にわたって8台のカメラを設置した。各カメラの近くにPCを設置し、シーン解析部、多視点映像部、サーバ部のすべてのPCをLANで接続した。また、各PCはNTPにより時刻同期をとった。

実際に作成した自由視点映像の例を図4示す。このように、ユーザはピッチに降り立ったような視点や、



図 4 自由視点映像生成の例

スタジアム上空から見下ろしている視点など、自由な視点からの映像を見ることが出来る。多視点映像から抽出した映像を適切に貼り付けることにより、写実性の高い映像を得ることができていることがわかる。また、選手同士の重なりが生じた場合にも、これらのテクスチャに対して他の選手の体の一部が誤って表示されることはほとんどない。

本実験において、登場するプレイヤー全ての人物ビルボードの生成は 30fps で動作することを確認した。本手法による人物ビルボードの生成速度は、多視点映像撮影部のスループットで決定されている。一方、シーン解析部については、背景更新処理を行わない場合には 30fps で動作するが、選手の 3 次元位置推定の精度を高めるために、背景を更新しながら選手領域の抽出を行わせた場合のスループットは現在のところ 11fps 程度である。なお、本システムでは、テクスチャ抽出を行う自由視点映像撮影部と 3 次元位置推定を行うシーン解析部は並行して実行されるため、シーン解析部のスループットが自由視点映像生成の全体の処理速度に影響を与えることはない。

実際にインターネットを介して、遠隔地にいるユーザへ自由視点映像を配信する実験を行った。帯域の異なる 4 種類の回線を用いて、単独配信と同時配信を行った場合にユーザが受け取る映像のスループットとレイテンシを測定した結果を表 1 に示す。個別配信の実験結果から、一般的なブロードバンド回線においても、本システムによりフレームレートの高い自由視点映像を配信することができていることがわかる。また、LAN 上で 14 人を対象とした同時配信において、各ユーザに対するスループットは個別配信の場合とほぼ同程度となった。これは、各ユーザに対するテクスチャの選択にかかる処理が小さいため、ユーザ数が増加したことによるサーバ部の処理量の増加が、スループットにほとんど影響を与えていないことを示している。

なお、各選手あたりの平均のデータ量は 960[byte]であった。従って、30fps で映像を配信した場合には、約 4.0[Mbps]の帯域が必要になる。

## 6.おわりに

大規模空間におけるスポーツイベントである、サッカーシーンを対象として、複数のカメラから撮影した映像を解析することにより、遠隔地にいる多数のユーザへインターネットを介して、各々の要求する自由視点映像を配信する方式を提案した。

本研究では、各選手を人物ビルボードによって表現することにより、少ない計算コストとデータ量で自由視点映像を生成することに成功した。

また、自由視点映像を生成するために、動的な背景モデルに基づく選手領域の抽出や、2 台のカメラ間の幾何関係を用いた 3 次元位置の獲得、選手同士の重なりを考慮した選手テクスチャの抽出を行った。

実際に国立霞ヶ丘競技場において、10 台のカメラを用いて撮影を行い、リアルタイムでの自由視点映像の生成、インターネットを介した遠隔地ユーザへの自由視点映像の配信、複数ユーザへの同時配信実験を行った。本実験において、全ての登場プレイヤーの、人物ビルボードの生成を 30fps で行うことができた。また、インターネットを介した遠隔地への配信を 15fps 以上の高いフレームレートで実行可能であることを確認した。

表 1 遠隔地への配信実験の結果

	回線	Throughput	Latency
個別配信	ADSL 8[Mbps]	24[fps]	128[ms]
	CATV 6	15	213
	ADSL12	21	145
	CATV12	17	183
	LAN100	29	104
同時配信 14 人	LAN100	29	106

## 文 献

- [1] T. Kanade, P. Pander, and P.J. Narayanan: "Virtualized Reality: Constructing Virtual Worlds from Real Scenes", IEEE Multimedia, Vol.4, No.1, pp.34-47 1997
- [2] P.Narayanan, P.Rander and T.Kanade, "Constructing Virtual Worlds Using Dense Stereo", Proc.of the Int.Conf.on Computer Vision (ICCV)'98, pp.3-10, 1998
- [3] S.Wurmlin, E.Lamboray, O.G.Staadt, M.H.Gross, "3D Video Recorder", Proc.of Pacific Graphics'02, pp.325-334, 2002
- [4] R.T.Collins, "A Space-Sweep Approach to True Multi-Image Matching", Proc. of Computer Vision and Pattern Recognition (CVPR)'96, pp.358-363, 1996