

視点に依存した微小面群モデルによる スポーツシーンの自由視点映像生成

古谷 仁志[†] 北原 格[‡] 亀田 能成[‡] 大田 友一[‡]

[†] 筑波大学 大学院システム情報工学研究科

[‡] 〒305-8573 茨城県つくば市天王台 1-1-1

E-mail: [†] {furuya, kitahara, kameda, ohta}@image.esys.tsukuba.ac.jp

あらまし スポーツシーンを対象として、ユーザの視点位置にあわせて適切な自由視点映像を生成する手法について述べる。本手法では、ユーザが要求する視点移動に追従し、映像品質を保持しつつ、できる限り少ない計算・データ量で映像を生成できる。我々はボクセル空間において被写体の大まかな形状を復元し、各ボクセルに微小ビルボードを配置することにより、ボクセルとビルボードの長所を併せ持つ3次元形状表現を実現する。観察距離にあわせてボクセル空間の分解能を変化させることにより、生成映像の品質に対して適切な計算量を設定することが可能である。実際に代々木第2体育館において撮影した多視点映像を用いた自由視点映像生成実験により、提案手法の有効性を示す。

キーワード 自由視点映像, ボクセル, ビルボード, 視点依存, 形状表現

A User's Viewpoint-Dependent 3D Modeling and Rendering Method of Sports Scene by Using Microfacet Billboarding

Hitoshi FURUYA[†] Itaru KITAHARA[‡] Yoshinari KAMEDA[‡] and Yuichi OHTA[‡]

[†] Department of Intelligent Interaction Technologies, University of Tsukuba

1-1-1 Tennoudai, Tsukuba-shi, Ibaraki, 305-8573 Japan

E-mail: [†] {furuya, kitahara, kameda, ohta}@image.esys.tsukuba.ac.jp

Abstract This paper proposes a novel 3D modeling and rendering method that generates viewpoint-dependent optimized 3D video of sports scene. The generated 3D video fulfills the demand of image quality for observers. It is generated with the minimum amount of computation cost. In order to show the effectiveness of our method, we generate a 3D viewpoint-free videos from multiple videos recorded in the National Yoyogi gymnasium.

Keyword 3D Free Viewpoint Video, Voxel, Billboard, Viewpoint-Dependent, 3D Modeling

1. はじめに

近年、多視点から撮影された映像をコンピュータ内部で融合し、ユーザが希望する視点位置からの見え方を再現する、自由視点映像技術に関する研究が盛んに行われている[1]-[5]。この技術によって、視点の位置や方向を思い通りに操作しながら、実際のイベントを鑑賞するという、従来の映像メディアでは不可能だった視聴スタイルの実現が期待されている。特に、オリンピックやワールドカップのように世界中が注目する、いろいろな見方をする人たちがたくさんいるようなスポーツイベントでは、自由視点映像技術が有効に機能することが考えられる。

本研究では、屋内競技程度の広さにおいて、様々なスポーツシーンを対象に自由視点映像を生成すること

を目的とする。体育館では、参加人数、使用器具、競技範囲などが異なる多種多様な種目が開催されるため、それを観察する時の視点位置も多様に変化するであろう。また、ユーザの好みによっても、コート全体を観察するような離れた位置や、選手と同じ目線の高さの視点など、視点位置は変化することが考えられる。このように、スポーツシーンを対象とした自由視点映像の生成・提示を行う場合には、個別に変化するユーザの視点に対して、適切な自由視点映像を生成することが重要な課題となる。ここでいう“適切な映像”とは、ユーザが要求する視点移動に追従し、映像品質を保持しつつ、できる限り少ない計算・データ量で生成された映像のことを意味する。特に、ユーザ（観察視点）から選手（被写体）までの観察距離が大きく変わる場

合、維持すべき提示映像の画質にあわせて計算を効率的に削減する点が本手法の特徴である。

本論文では、スポーツシーンを対象に、ユーザの視点位置に追従して、適切な自由視点映像を生成する手法を提案する。本提案手法によれば、観察者視点や被写体であるスポーツ選手の位置関係が様々に変化する場合でも、必要最小限におさえた処理によって、ユーザの要求を満たす画質の自由視点映像を生成することが可能である。

2. 自由視点映像の生成手法

映像情報を入力として自由視点映像を生成する手法は、IBR (Image-Based Rendering) と呼ばれる。IBRには、多数のカメラで撮影した画像を用いて3次元モデルの生成を行わずに直接、自由視点映像を求める手法[1][2] (以後、狭義のIBRとする)と、対象物体の3次元形状をコンピュータビジョン技術により生成し、その3次元モデルにCG技術を適用することで自由視点映像の生成を行うIBMR (Image-Based Modeling and Rendering)がある。3次元モデルを復元しないIBRでは、カメラを密に配置する必要があるため、大規模空間で行われるスポーツイベントに適用することは現実的ではない。そのため本研究では、IBMRを用いて自由視点映像の生成を行う。

IBMRの代表的な3次元形状表現手法としてボクセルを用いる方法[3][4]と、ビルボードを用いる方法[5]がある。ボクセルを用いる手法は、ボクセルのサンプリング間隔を調整することにより、あらゆる分解能の物体形状表現が可能であるという特長がある。しかし、スポーツシーン全体を俯瞰しているような、ユーザから被写体までの距離が比較的遠い場合には、それほど詳細な形状表現が必要とされないため、高解像度のボクセルモデルによる形状復元処理は、計算効率が悪くなる。一方、ビルボードを用いる手法では、多視点カメラで撮影された画像を、ほぼそのままの状態ビルボードにマッピングして提示するため、生成映像の写実性が高いという特長がある。しかし、選手を1枚のビルボードで表現しているため、正確な3次元形状表現は不可能であり、ユーザの視点と被写体である選手までの距離が比較的近い場合には、形状の歪みが顕在化する。これは広い意味での生成映像の画質劣化といえる。

本研究では、上記の問題を解決するために、ボクセルモデルにより大まかな形状を復元し、その各ボクセル上に微小なビルボードを配置する、マイクロファセット・ビルボーディング法[6]を用いて自由視点映像の生成を行う。各ボクセルに配置されるビルボードには、カメラのテクスチャをそのまま貼り付けるため、生成

される映像の画質が大きく損なわれることがない。さらに、ボクセル空間の分解能を視点に依存して変化させることにより、計算量・データを適切に設定することが可能である。この3次元形状表現手法を微小面群モデルと呼ぶ。

3. 微小面群モデル

微小面群モデルとは、対象物体を微小なビルボードの集合により表現する手法である。微小面群モデルを生成するために必要な処理は、主に以下の3つである。

- 視点に依存したボクセル空間の構築
- 選手の3次元形状復元
- 微小ビルボードの配置と実写画像のマッピング

この3つの処理について具体的に述べる。

3.1. 視点に依存したボクセル空間の構築

観察距離に依存してボクセル空間の分解能を変化させることにより適切な計算量を設定する。形状復元にかかる計算量は、ボクセル空間におけるボクセル数を k 、射影変換にかかる計算量を p 、カメラ台数を c とすると $O(cpk^3)$ となる。従って、ボクセルの数を $1/n$ にすると $O(n^3)$ で計算量を減らすことができる。

ボクセルの分解能は、生成映像の品質を落とさない範囲で行うべきである。そこで本研究では、ユーザの観察距離が変化した場合に、ビルボードの観測サイズ(生成画像上での面積)を一定とするボクセルの解像度で復元処理を行うことにより、映像画質を落とさずに適切な計算量の設定を実現する。ユーザの視点が選手に近い場合は、ボクセルのサイズは小さくなり、逆に遠い場合には大きくなるように、視点と選手との距離に応じてボクセルのサイズは動的に変化するが、図1に示すように、生成映像中における1枚のビルボードの観測サイズは一定であり、映像品質は維持されているといえる。ボクセルの解像度は、ボクセル空間を構築する世界座標系とユーザが観察する画面上のスクリーン座標系の、二つの座標系において、画面上のビルボードの情報から透視投影変換により求める。

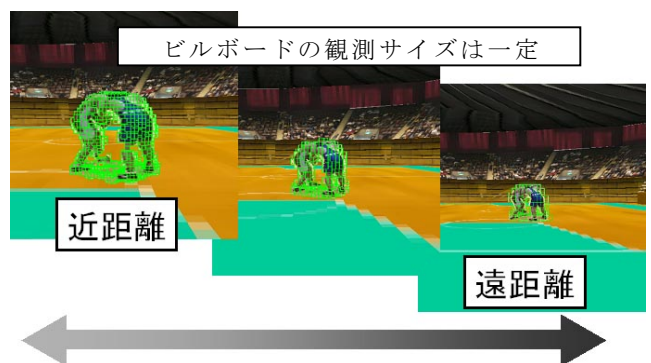


図 1: 観察距離に依存したビルボードサイズの設定

ここで、生成画像を観察したユーザに画質の劣化を感じさせないビルボードの観測サイズは主観評価実験により調査した。

先述したビルボードのサイズを設定するための、主観評価実験について述べる。この実験では、ビルボードのサイズを変えて生成した自由視点映像を被験者に観察してもらい、映像画質の評価実験を通じて、ビルボードサイズと生成自由視点映像の関係を調査する。その結果からユーザが満足に映像を観察できる解像度のビルボードを求め、本研究で用いるボクセル空間の分解能のパラメータとする。ビルボードの生成画像上で観測される領域の1辺の幅を5,10,15,20,25,30画素とした6パターンの映像について、映像品質の満足度評価を実施した。10名の被験者に対し、5つのシーンの映像において評価実験を行った結果を図2に示す。横軸がビルボードの観測サイズで、縦軸は評価値である。ビルボードの観測サイズが小さくなるにつれて品質に対する評価が向上する。一方、ビルボードのサイズが変化した時の形状復元に要する計算量の指標として、射影変換計算の回数を図中の曲線グラフで示す。計算量は、ビルボードサイズの3乗に反比例して減少していることから、高速な微小面群モデルの生成を実現するためには、ビルボードサイズをある程度大きく設定することが必要であるといえる。以上のことからビルボードは、5~10画素程度の幅で観測されるように設定するのが適当であると考えられる。

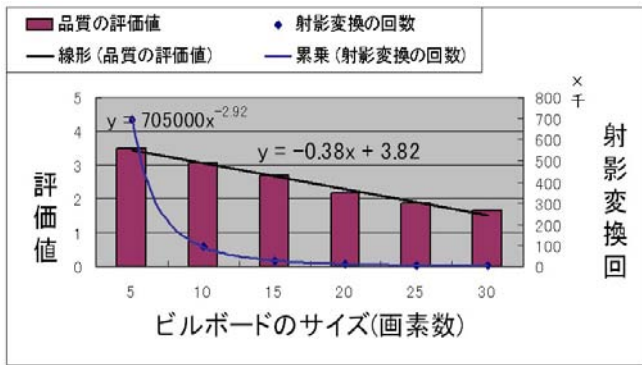


図 2: ビルボードのサイズと映像品質の評価

3.2. 視体積交差法による3次元形状復元

本研究では、視体積交差法を用いて選手の形状復元を行う。複数のカメラによって撮影された映像中のシルエット情報のみを用いて形状復元を行うため、照明条件の変化などによる多視点画像上での観測色の変化にロバストであるという長を有する。凹形状の復元が不可能であるという問題が存在するが、微小面群モデルでは、選手の形状は大まかな精度で復元できれば良いことから視体積交差法を選択した。

視体積交差法の手順を紹介する。まず、被写体のシ

ルエット（前景領域）を抽出し、撮影カメラの光学中心を頂点、シルエットを断面として3次元空間に被写体の視体積を形成する。次に、複数の画像から形成された視体積の積集合を求めることにより、被写体の3次元形状を推定する。

3.3. 微小なビルボードを用いたテクスチャマッピング

視体積交差法によって得られた選手の表面部分の3次元形状モデルに対してテクスチャ情報を与える。本研究では、各ボクセルの中央にボクセル幅程度の微小ビルボードを配置し、そのビルボードに映像のテクスチャをマッピングすることで、見え方を再現する。各ビルボードは、その法線がユーザの視線方向と常に一致するように回転させる。ビルボードを用いることにより、大まかに復元された3次元形状しか獲得できない場合でも、写実的な映像の生成が可能になる。また、マッピングする画像内の各画素には、 α 値を与えることで、前景と背景の領域を分け、前景領域のみを描画する。

テクスチャ情報として採用する多視点映像の選択には視点位置、ビルボードの位置、カメラの位置情報を用いる。図3のように、カメラからビルボードへのベクトルと視点からビルボードへのベクトルを求め、そのなす角がもっとも小さくなるカメラからの映像を用いるカメラとして選択する。

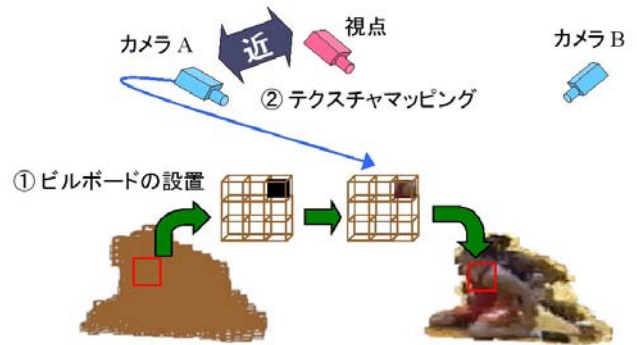


図 3: ビルボードへのテクスチャマッピング

通常、テクスチャのマッピングには視線と最も近いカメラが選択されるが、マッピングの対象となる選手が他の選手により遮へいされている場合には対象の選手に他の選手のテクスチャがマッピングされてしまう。本研究では、遮へいの生じていないカメラの中から視線と最も近いカメラを選択する方法により、この問題を解決する。

4. 自由視点映像生成システム

システムの概要を図4に示す。本研究で用いる多視点カメラは、コートを上から見下ろすような天井カメ

ラが1台、コート周辺を囲うような外周カメラ数台で構成される。天井カメラは選手の位置を推定するために用いる。外周カメラは微小面群モデルを生成するために用いる。ここで、微小面群モデルにより生成されるのは、シーンに対して動的に変化する選手などの前景領域のみである。本研究ではシーンに対して静的な背景領域に関してはあらかじめ作成した3Dモデルを用いる。ユーザはこれらをあわせて描画することで自由視点映像を生成する。

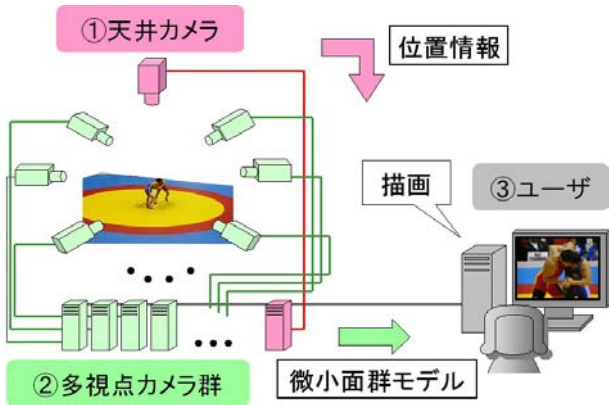


図4：自由視点映像生成システムの概要

全体の処理の流れを示す。まず、多視点同期撮影システムを用いて天井カメラと周辺カメラの画像を取得する。次に、これらの画像から選手のシルエット領域を抽出し、天井カメラ画像については抽出したシルエット領域の位置から選手の3次元位置を推定する。そして、ユーザの視点情報を基に設定されたボクセルサイズで、推定された選手の3次元位置にボクセル空間を構築する。形状復元とビルボードへのテクスチャマッピングにより微小面群モデルを生成し、最後に、微小面群モデルと背景の3次元モデルを組み合わせて描画を行い、自由視点映像を生成する。

5. 実験

本手法の有効性を示すために、実際に代々木第2体育館において9台のカメラを設置し、レスリングを対象として実験を行った。9台のうち1台は天井からコートを見下ろすように設置した天井カメラであり、残りの8台のカメラはコート周辺を囲うように設置してある。本実験では、3CCDビデオカメラ(SONY,BRC300)を用いて640×480の画像を30fpsで取得した。

実際に生成した自由視点映像の例を図5に示す。このとき1枚のビルボードの幅は8画素とした。このように、ユーザは様々な位置からの映像を見ることが出来る。本研究で用いた微小面群モデルでは、ビルボードを用いることにより、写実的な映像の生成が可能である。また、選手はビルボードにより平面群で近似しているが、ユーザからの見え方は十分に再現できてい

ることが確認できる。視点の移動に伴って選手の復元の精度を変化させることにより、ユーザの視点が選手に近い場合でも、それに伴った自由視点映像を生成することができている。また、視点が選手から離れた場合には計算量をおさえるために形状復元が粗くなるが、生成される映像に影響がないことがわかる。



図5：生成された自由視点映像の例

6. 結論

ユーザの視点に依存して適切な自由視点映像生成を得る手法を提案した。本論文では、選手の復元精度をユーザの視点情報を基に変化させ、微小面のビルボード集合により選手の3次元モデルの生成を行い、映像の品質を落とさずに計算コストを軽減することを実現した。今後の課題は、映像の品質を向上させるために、カメラ間の色合いを統一することなどがあげられる。

文献

- [1] L.Marc,P.Hanrahan, "Light Field Rendering", Proc. SIGGRAPH, pp.31-42(1996)
- [2] S.J.Gortler,R.Grzeszczuk,R.Szeliski,M.F.Cohen, "The Lumigraph", Proc.SIGGRAPH, pp.43-54(1996)
- [3] F.Caillette and T.Howard, "Real-Time Markerless Human Body Tracking with Multi-View 3-D Voxrl Reconstruction", Proc.BMVC, Vol.2, pp.597-606(04)
- [4] 上田恵, 有田大作, 谷口倫一郎, "3次元ビデオ映像のオンライン生成", 画像の認識・理解シンポジウム, Vol.2, pp.283-288(04)
- [5] T.Koyama,I.Kitahara,Y.Ohta, "Live 3D Video in Soccer Stadium", SIGGRAPH, 2003
- [6] S.Yamazaki, R.Sagawa, H.Kawasaki, K.Ikeuchi and M.Sakauchi, "Microfacet Billboarding", Proc. the 13th Eurographics Workshop on Rendering, pp.175-186(02)