

シーン状況と多視点映像の解像度を考慮した 自由視点映像生成支援手法

A Supporting Method to Control Virtual Viewpoint Capturing 3D Free-Viewpoint Video

大川原友樹¹⁾, 北原格¹⁾, 亀田能成¹⁾, 大田友一¹⁾
Tomoki OKAWARA, Itaru KITAHARA, Yoshinari KAMEDA and Yuichi OHTA

1) 筑波大学 システム情報工学研究科 知能機能システム専攻
(〒305-8573 茨城県つくば市天王台 1-1-1, s1220758@u.tsukuba.ac.jp)

Abstract: This paper proposes a method to support virtual viewpoint control capturing 3D free-viewpoint soccer video. Our proposed method estimates recommendation area to put a virtual camera by referring the captured texture resolution of a target player so that users can easily get better quality 3D video. When occlusion occurs in the sight from the inputted viewpoint, our method calculates more suitable viewpoint in where there is no occlusion. It is also possible to refine the appearance by applying cinematography. We conduct on subjective evaluations to confirm the effectiveness of our proposed method by using our developed pilot system.

Key Words: free-viewpoint, virtual camera control, cinematography

1. はじめに

自由視点映像とは、ユーザが自由に指定した位置からの見え方を観察することができる映像生成技術である[1]-[5]。スポーツを対象とする自由視点映像の場合、複数の選手間を移動するような映像や、フィールドを俯瞰した映像の再現が可能であるため、新たな映像メディア技術として期待されている。一方で、サッカーのように複数の移動物体が短時間で大きく動くような動的なシーンを対象とする場合、マウスなどの従来型インタフェースでは注視物体を追従しながらの観察が困難であるといった問題が存在する。

この問題を解決するために Watanabe ら[6]は 3次元位置センサとフィールド俯瞰映像を用いた自由視点映像撮影システムを提案したが、我々は、Watanabe らの手法には、以下に示す三つの課題が残っていると考えている。

- (1) ユーザは、選手の 3次元モデルにマッピングされているテクスチャ解像度に関する情報を持たないため、生成映像の画質を考慮した視点操作が困難である。
- (2) ユーザは、オクルージョンの発生状況を把握していないため、生成映像中でオクルージョンが発生しやすい。
- (3) 映像撮影技術に関する知識のないユーザは、注目選手を画面中心に配置しがちであり、その結果、その先の展開を予測させるフレーミングの映像を撮影するのが困難である。

本稿では、図 1 に示すように、上記に挙げる三点の問題をシーン状況と多視点映像の解像度を考慮することにより解決し、より適切な位置・姿勢に仮想カメラを設置することを目的とした手法を提案する。



図 1 提案する映像生成支援手法

2. 関連研究

本研究では、仮想カメラの視点を操作するインタフェースとして Watanabe ら[6]のシステムを用いる。ユーザは、両手に 3次元位置センサのマーカを持ち、それらを用いて

仮想カメラの視点位置と注視点位置を指定する．3次元位置センサで計測した各マーカの3次元位置から仮想カメラの位置・姿勢を算出し，自由視点映像生成技術を用いて対応した見え方を生成し，ディスプレイに提示する．このシステムは，フィールド俯瞰映像を提示することにより，選手やボールの位置関係を容易にユーザが把握できる機能を有する．しかし，前述したように，（1）テクスチャ解像度の問題（2）注視物体のオクルージョンの問題（3）注視物体のフレーミングの問題の三つの問題が存在する．

中嶋ら[7]は，多角形のポリゴンに対してオクルージョン判定を行う陰面消去アルゴリズムを説明している．非常に精度の良い結果を得ることができるが，サッカーを対象とする自由視点映像では，最大23枚のビルボードを扱うため，実時間処理が困難である．

加藤ら[8]は，サッカーを撮影する際のカメラワークを分析することで，卓越したカメラマンは注目する選手（以降，注目選手と呼ぶ）の移動方向に対して，水平方向にカメラをパンさせる撮影技法を用いていることを紹介している．一方，北原ら[9]や西井ら[10]が提案する3分割法は，同様の操作を垂直方向にも適用させるため，前空きよりも効果的に余白を活用できる．提案手法では，プロのカメラマンが撮影するような，サッカーの状況の先の展開を予測させる映像を生成するために，3分割法を用いて注目選手を安定的にフレーム内に収める処理を実現する．

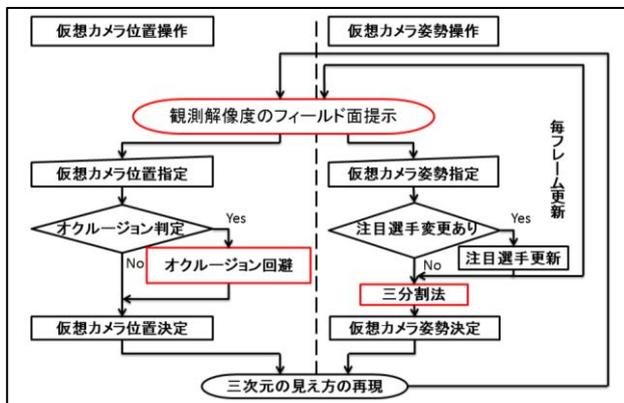


図 2 映像生成支援手法の処理の流れ

3. シーン状況と多視点映像の解像度を考慮した映像生成支援手法

本稿で提案する手法は，図 2 に示すように，ユーザが仮想視点位置を決定する際に参考となる情報（注視物体の観測解像度）を提示する操作支援処理と，ユーザによって入力された仮想視点位置をシーンの状況（選手の位置情報）を用いて仮想カメラの位置・姿勢補正処理の二つの部分で構成される．操作支援処理では，注視物体の観測解像度をフィールド面に提示することにより，選手テクスチャの状態が良好な位置に仮想カメラを誘導する．仮想カメラの位置・姿勢補正処理では，ユーザが入力した仮想カメラの位置・姿勢によって生成される映像に対して，注目選手と他

選手とのオクルージョン判定を行い，よりオクルージョンが少ない位置に仮想カメラを移動させる．また，入力された映像に対して，3分割法と呼ばれる映像撮影技法を適用し，仮想カメラの姿勢操作によってフレーミングを補正することにより，撮影知識を持ち合わせていないユーザでも，その先の展開を予想させるような映像が撮影を実現する．

3.1 多視点映像の解像度を考慮した操作支援処理

図 3 に示すように，同一選手のテクスチャの解像度は，カメラ毎に異なっている．これは，カメラの焦点距離（ズーム）を固定した撮影では，多視点カメラと被写体の距離に反比例して観測解像度が減少するためである．提案手法では，多視点カメラと注目選手の距離から注視物体の観測解像度を算出し，その情報を仮想俯瞰映像上に重畳提示することにより，選手テクスチャの状態が良好な位置に仮想カメラを誘導する．

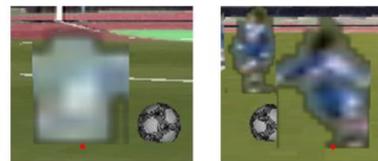


図 3 異なるカメラから撮影された同一選手のテクスチャ

人物ビルボードの矩形サイズを獲得するために，まず Kasuya ら[11]の手法を用いてフィールド上の選手位置を推定する．獲得された選手位置を基に，注目選手を囲む矩形を生成し．適切な射影変換を介して，仮想カメラが撮影する画像上での観測サイズを計算する．図 4 に示すように，ある多視点映像上での注目選手の3次元モデル（人物ビルボード）の観測サイズが横 w ，縦 h [pixel]であるとすると，観測解像度 S は，式 1 により算出される．計算された S を用いて，各時刻における最大値を RGB 表色系における青(0,0,255)，最小値を赤(255,0,0)とし，フィールド面に提示することでユーザは一目で画質と観測解像度の関係を理解することができる．

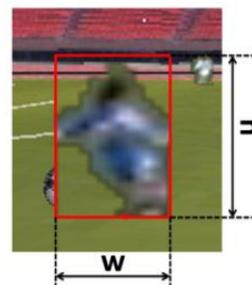


図 4 観測解像度の算出方法

$$S = wh \quad (1)$$

注視物体の観測解像度をフィールド面に提示した様子を図 5 に示す。注目選手を中心に、観測されるテクスチャの解像度に応じて、多視点カメラに対応する各エリアが色分け提示されていることが確認できる。この場合、青い部分に仮想カメラの位置を指定すると観測解像度の高いテクスチャが得られる。



図 5 選手の観測解像度をフィールド面に提示した結果

3.2 注視物体のオクルージョン状態を考慮した補正処理

ユーザが指定した仮想カメラの視点位置から注視物体を観察した際に、他の選手など手前の物体によってオクルージョンが発生する場合、オクルージョンが発生しない位置に仮想カメラの位置を移動することで、注視物体が観察しやすい映像を生成する。渡邊[12]らの考察を参考にして、位置補正の範囲は、仮想カメラの注視点位置を中心とした半球面上で、ユーザが指定した視点位置から上下左右 45 度ずつとする。図 6 に示すように、オクルージョン判定は、人物ビルボードの 4 隅と仮想カメラの視点位置を結ぶ 4 本の線と他選手のビルボード面との衝突判定によって行う。

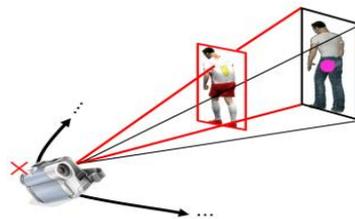


図 6 注目選手と他選手との衝突判定の様子

上述した視点位置の補正範囲内で、ユーザが入力した仮想カメラ視点位置から、水平方向、垂直方向に 5 度刻みで視点を移動させながらオクルージョン判定を行う。図 7 に出力結果を示すように、注目選手が観察しやすい映像が生成されていることが確認できる。

4. 3 分割法を用いたフレーミング補正処理

映像撮影知識の乏しいユーザが撮影した映像によく見られる画面中央に注視物体を撮影した映像に対して、3 分割法を適用することで、その先の展開を予測させる効果を有する映像を生成する。画面中央に捉えた注視物体を適切な位置に移動させるための仮想カメラの姿勢補正量を、パン角 $d\phi$ 、チルト角 $d\omega$ とする。仮想カメラの水平画角 θ_x 、垂直画角 θ_y とすると、補正量は式 2, 3 により求めることができる。



図 7 ユーザの入力画像（上）と補正処理後画像（下）

$$d\phi = \theta_x \times \frac{1}{2} \times \frac{1}{3} = \frac{1}{6} \theta_x \quad (2)$$

$$d\omega = \theta_y \times \frac{1}{2} \times \frac{1}{3} = \frac{1}{6} \theta_y \quad (3)$$

3 分割法を用いて注視物体を移動する位置は、図 8 に示すように、左右二つの候補があるが、注目選手の運動方向により広い余白ができるような位置を選択する。運動方向が頻繁に切り替わることを避けるために、過去 8 フレームを平滑化した値を選手の運動方向とする。

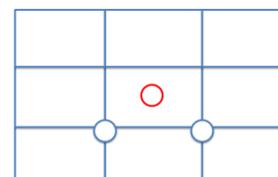


図 8 3 分割法における画面のグリッド線

3 分割法の適用例を図 9、図 10 に示す。二つの図は、同時刻における画面中央に選手を配置した画像であり、図 10 は図 9 に対して 3 分割法を適用した結果である。図 10 では 3 分割の基準フレームに沿うように注目選手の移動方向を考慮した配置をしており、図 9 には写っていない選手が図 10 では確認できるため、ユーザにその先の展開を予測させるフレーミングの映像提示が実現できていることが確認できる。



図 9 選手を中心に配置した画面



図 10 選手を3分割法によって配置した画面

5. システムの構築

提案手法を用いたパイロットシステムを、以下の機材を用いて構築した。映像提示及び仮想視点制御用 PC として Mouse Computer Lm-i530B(Intel Core i5 3.2GHz 2.0GB RAM)を使用した。自由視点映像提示ディスプレイとして、BenQ E2400HD(解像度 1920×1080 画素)を使用し、フィールド俯瞰映像提示ディスプレイとして、三菱 LDT322V(解像度 1366×768 画素)を使用した。また、仮想カメラの位置・姿勢を決定するために、MicronTrackerSx60(サンプリング周波数 48Hz)を使用して、マーカの 3次元位置を計測する。また、Koyama らの人物ビルボード方式[5]を用いて、フレームレートは 30[fps]で自由視点映像の生成・提示を行う。

6. まとめ

本稿では、Watanabe らのシステムに存在した三つの問題をシーン状況と多視点映像の解像度を考慮することにより解決し、より適切な位置・姿勢に仮想カメラを設置するための手法について説明した。提案手法では、注視物体の観測解像度をフィールド面に提示することにより、注視物体のテクスチャの状態が良好な位置に仮想カメラを誘導した。また、ユーザが入力した映像に対して、注目選手と他選手とのオクルージョン判定を行い、よりオクルージョンが少ない位置に仮想カメラを移動させる。さらには入力された映像に対して、3分割法と呼ばれる撮影技法を適用し、フレーミングを補正することにより、撮影知識を持ち合わせていないユーザでも、その先の展開を予想させるような映像を生成した。

参考文献

[1] T.kanade, P.Rander,P.J.Narayanan,"Virtual-ized Reality:Constructing Virtual Worlds rom Real Scenes," IEEE Multimedia Vol.4, No.1, pp.34-47, 1997

[2] W.Matusik,C.Buehler, R.Rasker, S.J.Gortler, L.McMillan, "Image-Based Visual Hulls," ACM SIGGRAPH, pp.369-374, 2000

[3] Joel Carranza,Christain Theobalt, Marcus A. Magnor, Hans-Peter Seidel,"Free-Viewpoint Video of Human Actors," ACM Transaction on Graphics, vol.22, No.3,pp.569-577, 2003

[4] 谷本正幸, "自由視点テレビ", 3次元映像フォーラム, Vol.15 No.3, pp.17-22, 2001

[5] Takayoshi Koyama,Itaru Kitahara,Yuichi Ohta,"Live Mixed-Reality 3D Video in Soccer Stadium"ISMAR ,pp.178-187, 2003

[6] Tetsuya Watanabe, Itaru Kitahara, Yoshinari Kameda, Yuichi Ohta,"3D FREE-VIEWPINT VIDEO CAPTUREING INTERFACE BY USING BIMANUAL OPERATION",IEEE, 2010

[7] 中嶋 正之, 安居院 猛, 高木 志朗, "二物体の優先関係を用いた隠面消去法", テレビジョン学会技術報告 Vol.71, No.3 ,pp13-18, 1984

[8] 加藤 大一郎, 山田 光穂, 阿部 一雄, 石川 秋男, 尾幡 昌芳, "被写体を追尾撮影時の放送カメラマンのカメラワーク特性分析",テレビジョン学会誌 Vol.50, No12, pp.1941-1948, 1996

[9] 北原 浩, "画面構成のテクニック", 玄光社, pp.76-77, 2006

[10] 西井 育生, "CINEMATOPGRAPHY & DIRECTING", 株式会社ボーンデジタル, pp.127-128, 2003

[11] Nozomu Kasuya, Itaru Kitahara, Yoshinari Kameda, Yuichi Ohta, "Real-time soccer player tracking method by utilizing shadow regions", ACM Multimedia, 1319-1322, 2010

[12] 渡邊哲哉, 北原格, "正確で直感的なカメラ操作を可能とする両手を用いた自由視点映像撮影インタフェース", 筑波大学大学院博士課程 システム情報工学研究科修士論文, 2009