

自由視点映像生成のための簡易な操作による背景生成手法

野村 尚史[†] 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: [†] nomura@image.esys.tsukuba.ac.jp, [‡] {kitahara, kameda, ohta}@iit.tsukuba.ac.jp

あらまし スポーツシーンにおける自由視点映像生成の一環として、簡易な操作により背景領域の自由視点映像を生成する手法を提案する。本手法ではスポーツフィールドのような大規模空間においては、背景が十分遠くに存在するという特性を利用する。背景領域上の数点を指定するだけという簡易な操作により背景の3次元形状モデルを適切に構築し、さらに仮想的に背景領域を正面から撮影したように映像テクスチャを変形し、3次元形状モデルに貼り付けることにより、つなぎ目のズレのない、自然な見え方の自由視点映像生成を実現する。スタジアムや体育館などの大規模空間において撮影した多視点映像を用いて背景の自由視点映像の生成実験を行い、提案手法の有用性を確認した。

キーワード 自由視点映像, 3次元モデリング, テクスチャマッピング, スポーツシーン, 平面近似

3D Background Modeling Method with Simple Operations

Naobumi Nomura[†] Itaru Kitahara[‡] Yoshinari Kameda[‡] and Yuichi Ohta[‡]

[†] Graduate School of Systems and Information Engineering, University of Tsukuba

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

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

Abstract This paper proposes a novel 3D background modeling method that only needs simple manual operations. To generate 3D background model, background area is divided into “distant region” and “ground region”. The 3D model of distant region is represented as textured vertical rectangular planes. The ground region is modeled as a simple horizontal plane. These two regions are segmented by a few points given by manual operation. The segmentation is conducted based on the knowledge of the camera location and the target world. In order to show the validity of our method, we generated a 3D background video from multiple images recorded in the National Kasumigaoka Stadium and the Gymnasium of Japan Institute of Sports Science.

Keyword Free Viewpoint Video, 3D modeling, Texture Mapping, Sports Scene, Plane Fitting

1. はじめに

近年、コンピュータ技術の発展にともない、複数台のカメラを用いてさまざまな角度から映像を獲得し、それらの映像を計算機内部で融合することにより、自由な視点から映像を閲覧する自由視点映像メディアに関する研究が活発に行われている[1][2]。自由視点映像メディアは、娯楽、研究、教育、文化財保護などのさまざまな分野において応用が期待されているが、我々は特に、サッカースタジアムなどの大規模空間を撮影対象とし、多視点映像の撮影から3次元映像処理、ネットワークを介したデータ伝送、自由視点映像提示までを実時間で行う、自由視点映像中継方式に注目し、研究開発を行っている[3]。

従来の自由視点映像生成に関する研究は、物体の3

次元形状を陽に復元しない手法と復元する手法とに大別することができる。前者の例として、Zitnickらは、8台のカメラにより撮影された映像を Segmentation based stereo algorithm を用いて解析することにより、カメラ間の中間視点映像を生成する手法を提案している[4]。このシステムでは、入力映像を複数の layer に分け、各 layer において適切な処理を施し、合成することで人物の正確な3次元形状復元を行うことなく、仮想視点からの高精細な映像の鑑賞を可能としている。しかし3次元形状を陽に復元しない手法では、多視点画像間での対応点探索という計算コストの高い処理が必要となるため実時間化が困難であるという問題や、多視点カメラを設置した位置を内挿した箇所には仮想カメラを設置できないという問題が存在する。

これに対し3次元を陽に復元する手法は、視点設定の自由度が高く、また、近年急速に発展したコンピュータグラフィックスの技術を活用することにより、高速な映像生成処理が可能であるという特長を有する。Kanadeらは、半径5mの半球内に設置された51台のカメラにより撮影された多視点映像をもとに人物の正確な3次元形状を復元し、自由視点映像生成手法を提案している[5]。しかし、この手法によって自由視点映像を生成するためには膨大な計算コスト、データ量が必要となるため、大規模空間への対応やリアルタイムでの映像生成への適応が困難である。これに対して、Koyamaらは被写体となる人物(前景領域)の3次元形状を1枚の平面形状(ビルボード)で近似することにより、撮影から提示までの処理を実時間で実行可能とする手法を提案し、サッカースタジアムのような大規模空間における自由視点映像生成・伝送・提示を実現している[3]。

しかし、これら従来研究では、シーン中の物体や人物といった前景領域に注目したものがほとんどであり、それ以外の背景領域の見え方は、あらかじめ用意した3次元CADモデルを用いて生成するにとどまっている。近年、サッカースタジアムなどの大規模な施設を建設する場合、3次元CADモデルを用いた設計施工が一般的となっているが、そのモデル生成には依然多くの労力が必要とされる。本発表では、3次元CADモデルが存在しない運動施設においても、高品質な背景領域の提示を伴う自由視点映像を生成することを目的とし、簡易な操作により背景領域の3次元モデルを生成する手法について提案する。

2. 簡易操作による背景生成手法

大規模空間を撮影対象とした場合、背景領域は遠方に存在している遠景領域とそれ以外の地面領域に二分することができる。本節では、各領域に対して適切な処理を施すことにより、自由視点映像を生成する手法について述べる。

2.1. 背景領域の3次元モデル

本手法では、スタジアム、体育館などに設置された、カメラパラメータが既知の複数台のカメラで撮影された多視点映像を入力とする[6]。

我々は大規模空間におけるスポーツイベントでは、以下のような特性が背景領域に当てはまると考え、カメラにより撮影された背景画像を、コート、ピッチといった地面領域と、壁・観客席といった遠景領域に二分する[7]。

- 選手(前景領域)はコート、ピッチといった1枚の平面上(地面領域)で競技を行う。

- 背景領域は主に地面領域と壁・観客席のような遠景領域によって構成されている。
- 遠景領域は、カメラから十分に遠いため、地面領域に対して直立する平面とみなすことができる。
- 撮影に用いるカメラは地面領域に垂直に設置されている

一般に観察者は、ピッチ、コート上でプレーする選手に注目することが多く、壁や観客席などの遠景領域に注目することは少ない。また遠景領域はカメラから十分に遠方に存在しているため、3次元的な見え方の変化を正確に再現する必要は少ないと考えられる。我々は、これら遠景領域が持つ特性に着目し、複数枚の平面群として形状を近似し、それらに撮影映像のテクスチャを貼り付けることにより3次元モデルを生成する手法を提案する。映像上の数点を指定するという非常に簡易な処理だけで、背景領域を構成する平面群の分割処理を実現できる点が、本手法の大きな特長である。

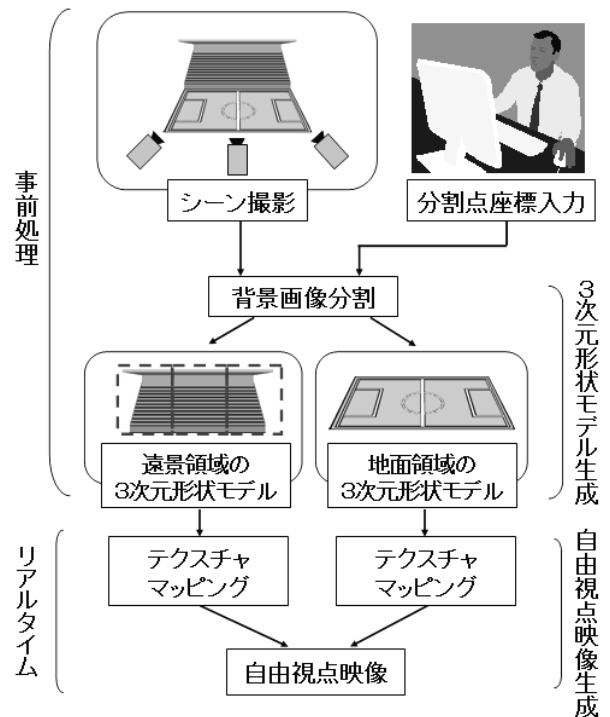


図1: 処理の概要

2.2. 背景領域生成手法の概要

図1に、背景領域の自由視点映像生成処理の概要を示す。まず、撮影された多視点画像を小領域に分割する。この小領域への分割では、多視点画像中の地面領域と背景領域との境界を示す特徴点を手動で指定する必要があるが、1台の撮影カメラに対して、数点のクリック処理を1度だけ行えばよく、それほど労力を要さない。分割された小領域を3次元空間中に射影する

ことにより、地面領域に直立する平面として、遠景領域の3次元形状が復元される。最後に、3次元形状モデルに多視点映像のテクスチャを貼り付け、背景領域の自由視点映像を生成する。

提案手法は、分割毎に画像上の2点を指定するだけ、という簡易な操作により、自由視点映像生成が可能であるため、カメラ台数が増加した場合でも比較的容易に対応可能である。また、その分割間隔はユーザが生成結果を見ながらインタラクティブに操作できるため、労力に対して良好な生成結果が得られるという利点を有する。

3. 地面領域の自由視点映像生成

本節では、映像から地面領域の3次元モデルを復元するために必要な地面と画像平面間の座標変換について述べる[8][9]。3次元空間中のある点 (x,y,z) と、画像上での観測点 (u,v) との間の幾何学的関係は、撮影に用いたカメラの射影変換行列を用いて、式(1)のように表すことができる。

$$\begin{pmatrix} \lambda u \\ \lambda v \\ \lambda \end{pmatrix} = \begin{pmatrix} C_{11} & C_{12} & C_{13} & C_{14} \\ C_{21} & C_{22} & C_{23} & C_{24} \\ C_{31} & C_{32} & C_{33} & C_{34} \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix} \quad (1)$$

ただし、モデル復元処理を簡易化するため、 z 軸は鉛直上向きにとるものとする。地面を高さ $z=h$ の平面とすると、式(1)は

$$\begin{pmatrix} \lambda u \\ \lambda v \\ \lambda \end{pmatrix} = \begin{pmatrix} C_{11}x + C_{12}y + C_{13}h + C_{14} \\ C_{21}x + C_{22}y + C_{23}h + C_{24} \\ C_{31}x + C_{32}y + C_{33}h + C_{34} \end{pmatrix} \quad (2)$$

となる。この式は z を定数と認めると、次式のように書き直すことができる。

$$\begin{pmatrix} \lambda u \\ \lambda v \\ \lambda \end{pmatrix} = \begin{pmatrix} C_{11} & C_{12} & hC_{13} + C_{14} \\ C_{21} & C_{22} & hC_{23} + C_{24} \\ C_{31} & C_{32} & hC_{33} + C_{34} \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = \mathbf{B} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} \quad (3)$$

式(3)右辺の 3×3 行列 \mathbf{B} は明らかに逆行列が存在するので、この逆行列 \mathbf{B}^{-1} を求めることにより、画像上の点 (u,v) を地面領域上の任意の点 (x,y,h) に投影することが可能となる。この2次元射影変換を用いて、撮影多視点画像を地面の3次元モデルに貼り付けたものを、自由な視点位置においてレンダリングすることにより、地面領域の自由視点映像が生成される。

4. 遠景領域の自由視点映像生成

4.1. 分割画像群による遠景領域モデルの生成

本手法では、遠景領域を実形状とは異なる複数の平面で表現するが、背景形状モデルを構成する平面の枚数が少な過ぎると、撮影映像中の背景領域を3次元形状モデルに貼り付ける際に見え方に大きな歪みが生じ、生成映像の品質劣化の原因となる。そこで、図2に示すように、遠景領域を小領域に分割し、その分割小領域ごとに3次元平面近似を行うことにより近似精度を高め、歪みの少ない背景自由視点映像を生成する。

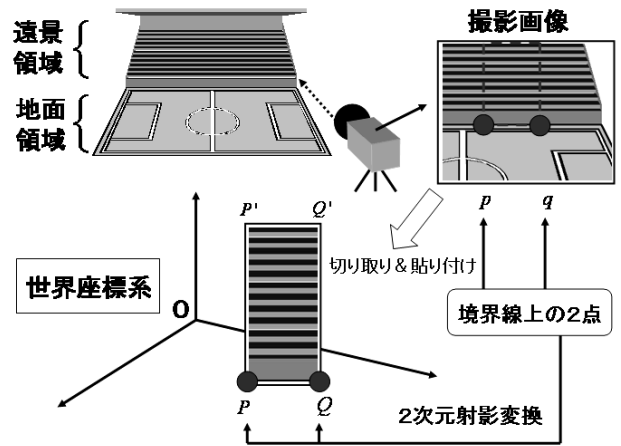


図2：遠景領域分割処理による形状復元

まず、入力画像中の遠景領域と地面領域の境界線上2点の座標 $p = (u_p, v_p)$ 、 $q = (u_q, v_q)$ を手動で指定し、その2点を下端とする矩形領域を分割小領域として切り取る。この2点を以降、分割点と呼ぶ。次に2つの分割点 p, q を式(3)を用いて地面上に射影し、3次元座標 $P = (x_p, y_p, h)$ 、 $Q = (x_q, y_q, h)$ を算出する。3次元空間中に投影した2点の座標間の距離 d と、分割画像のアスペクト比 α を式(4)に与えることにより、分割画像の上端隅の3次元空間中における対応点 $P' = (x_p', y_p', z_p')$ 、 $Q' = (x_q', y_q', z_q')$ を求めることができる(アスペクト比は分割小領域の高さを幅で割った値とする)。これら分割小領域の四隅の3次元座標から地面に直立する遠景領域の3次元形状が復元される。

$$\begin{aligned} P' &= (x_p', y_p', z_p') = (x_p, y_p, h + d\alpha) \\ Q' &= (x_q', y_q', z_q') = (x_q, y_q, h + d\alpha) \end{aligned} \quad (4)$$

こうして事前に生成した遠景領域の3次元形状モデルに競技中に撮影した背景画像を貼り付けていくことにより、歪みの少ない自由視点映像を生成することができる。図3に1枚の画像から4枚の分割小領域を

設定し,生成した遠景領域の自由視点映像の例を示す.この手法は簡単でありながら背景領域を構築できるという点で有用である.しかし,この方法で正しく遠景領域をモデル化できるのは,撮影カメラが背景領域を構築する物体群に対して正対しているという制約が成り立つときである.

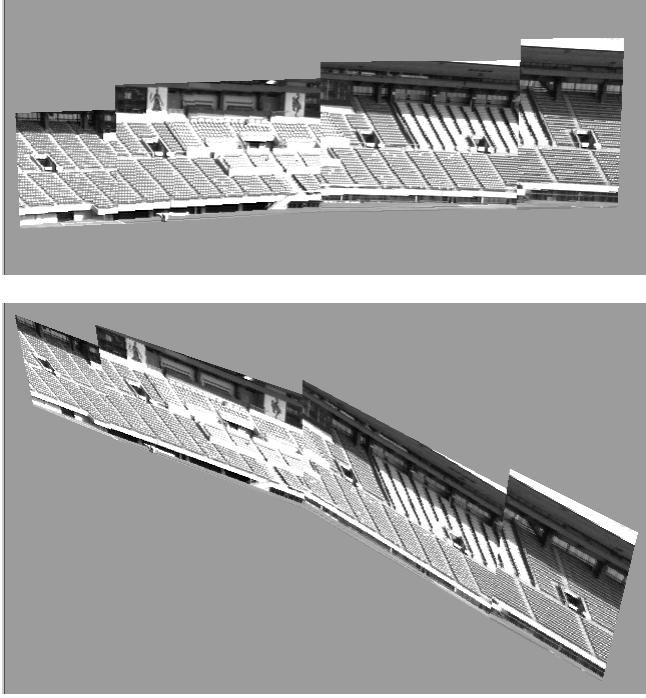


図 3 : 遠景領域の自由視点映像の例
(上 : 生成例、下 : 異なる視点での映像)

4.2. 分割小領域の 2 次元射影変換

複数の分割小領域を背景 3 次元形状モデルに貼り付けた場合,画像間の解像度の違い,画像中の壁の奥行きの違いなどが原因となり,図 5 (上)の入力に対して,図 5 (中)に示すように分割小領域のつなぎ目においてズレが発生する場合がある.このズレを解消するためには,遠景領域面を正面から撮影した映像を貼り付ける必要があるが,物理的制約から常にそのような撮影を行うことができるとは限らない.そこで本手法では,遠景領域の直立仮定を用いて,4 点の対応関係から遠景領域平面と撮影画像間の 2 次元射影行列を求め,この 2 次元射影変換を用いて遠景領域を仮想的に正面から撮影した映像を生成することにより,分割小領域のつなぎ目におけるズレを解消する.

4.2.1. 2 次元射影行列の算出

2 次元射影変換は,次式で表される平面間の射影関係である.

$$s\tilde{m} = H\tilde{m}' \quad (5)$$

ここで \tilde{m} , \tilde{m}' は 2 枚の平面間における対応点の同次座標である.4 点以上の対応関係がわかっているならば,2 次元射影行列 H を求めることができる[6].推定誤差を低減させるためには,4 点以上の対応関係から最小自乗法を用いて 2 次元射影行列を算出するのが一般的であるが,本研究では操作の簡易性を優先し,4 点の対応関係のみを用いて 2 次元射影行列を算出する.

4.2.2. 仮想的に正面から撮影した見え方の生成

既知である 2 つの分割点を点 p , q とする.次に式 (3)を用いて点 p , q を 3 次元空間中の地面平面に射影し,この点をそれぞれ点 P , Q とする.点 P , Q 間の距離 d を求め,この 2 点の鉛直上方,距離 d のところに点 P'' , Q'' を定める.図 4 に示すように点 P , Q , P'' , Q'' は 3 次元空間中で正方形を構成する.式 (1)により点 P'' , Q'' を撮影画像上に射影し,点 p'' , q'' を求める.4 点 P , Q , P'' , Q'' が正方形に見えるように仮想正面カメラを用意し,このカメラから見た 4 点をそれぞれ \hat{p} , \hat{q} , \hat{p}'' , \hat{p}'' とする.前節で述べたように,点 p , q , p'' , q'' と点 \hat{p} , \hat{q} , \hat{p}'' , \hat{p}'' の 4 組の対応関係から 2 次元射影行列 H を算出し,この射影変換により,背景領域を正面から撮影したような見え方に撮影画像を変換し,3 次元形状モデルに貼り付ける.各点の関係を以下に示す.

$$\begin{cases} p = (u_p, v_p) \\ q = (u_q, v_q) \\ p'' = (u_{p''}, v_{p''}) \\ q'' = (u_{q''}, v_{q''}) \end{cases} \quad (6)$$

$$\begin{cases} P = (x_p, y_p, h) \\ Q = (x_q, y_q, h) \\ P'' = (x_{p''}, y_{p''}, h + d) \\ Q'' = (x_{q''}, y_{q''}, h + d) \end{cases}$$

$$\begin{cases} \hat{p} = (u_{\hat{p}}, v_{\hat{p}}) \\ \hat{q} = (u_{\hat{q}}, v_{\hat{q}}) \\ \hat{p}'' = (u_{\hat{p}''}, v_{\hat{p}''} + (u_q - u_p)) \\ \hat{q}'' = (u_{\hat{q}''}, v_{\hat{q}''} + (u_q - u_p)) \end{cases}$$

図 5 (上)に示した映像と分割点を与えた場合における 2 次元射影変換後の自由視点映像を図 5 (下)に示す.このように正面から背景領域を撮影していない映像でも,2 次元射影変換を行うことで,分割小領域のつなぎ目におけるズレが解消されていることがわかる.

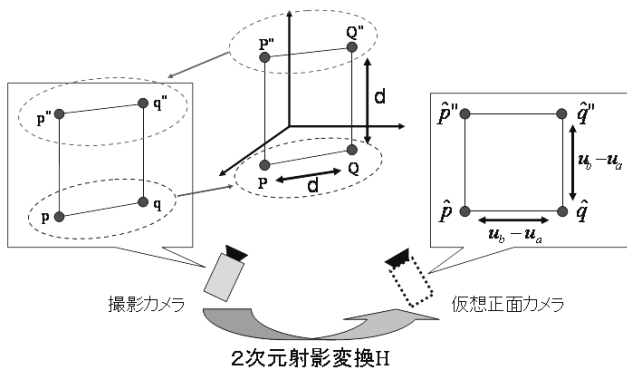


図 4 : 撮影画像と形状平面間の
2次元射影変換の推定

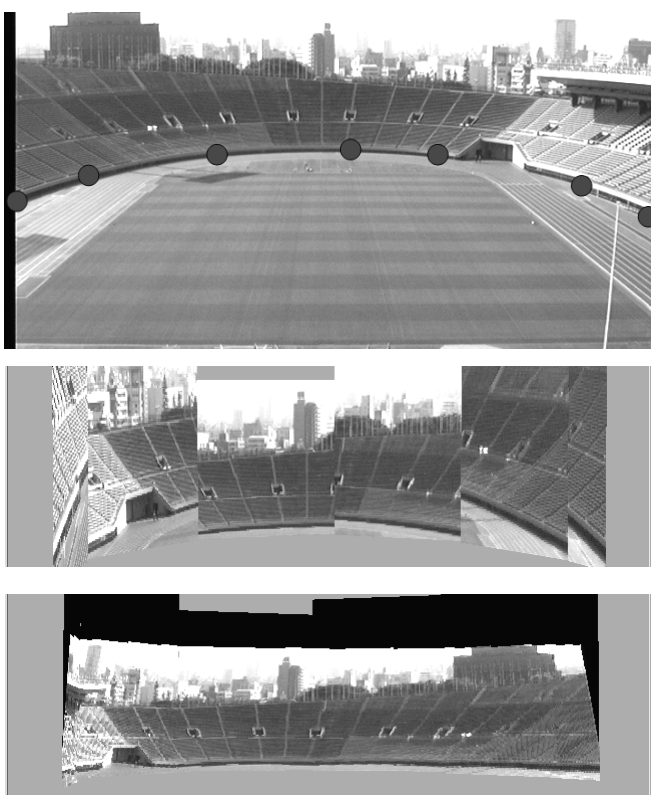


図 5 : 背景領域の自由視点映像生成の例
(上:入力画像と分割点の位置, 中:撮影テクスチャをそのまま貼り付けた例, 下:正面から撮影したテクスチャに変換後貼り付けた例)

4.3. 背景画像の更新

サッカーの試合では、観客の入退場や、応援の盛り上がりによって背景が動的に変化する場合がある。そこで、本研究では、遠景領域にマッピングする背景画像を随時更新し、より臨場感のある自由視点映像を生成する。

5. 実験

図 6 に示すように、国立霞ヶ丘競技場に 3CCD ビデオカメラ(SONY, DXC9000)10 台を設置して多視点映像の撮影実験を行った。撮影映像の解像度は 640×480 画像、フレームレートは毎秒 30 枚である。

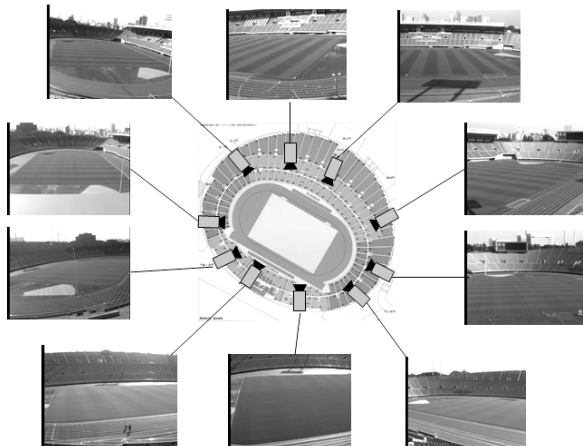


図 6 : 霞ヶ丘競技場実験でのカメラ配置

入力多視点画像に対して、計 32 個の分割点を指定し 16 枚の分割小領域を取得し、背景領域の 3 次元モデルを生成した。入力分割点 32 点を平面図上に示したものが図 7 である。画像の分割に要した入力手作業の時間は 10 分程度であり、十分簡易であると言える。実際に生成された自由視点映像を図 8 に示す。地面と同じ高さにて視点を設定した場合の見え方や、競技場上空から全体を俯瞰したような見え方など、自由な視点からの映像を観察することができる。また、分割小領域のテクスチャを 2 次元射影変換し 3 次元形状モデルに貼り付けることで、つなぎ目のズレが目立たない自然な見え方が再現されていることがわかる。

国立スポーツ科学センターにおいても、図 6 と同様にコートの周囲にカメラ(SONY,BRC300)を 10 台配置した撮影実験を行った。この実験で撮影した映像を用いて、4.3 節で述べた背景画像の更新処理を行った。図 9 に、人物が壁際を歩く様子を撮影した映像を用いて、背景領域の自由視点映像を生成した例を示す。背景の時系列変化に合わせて画像を更新することで、撮影空間の雰囲気により把握しやすい自由視点映像が生成・提示されていることがわかる。

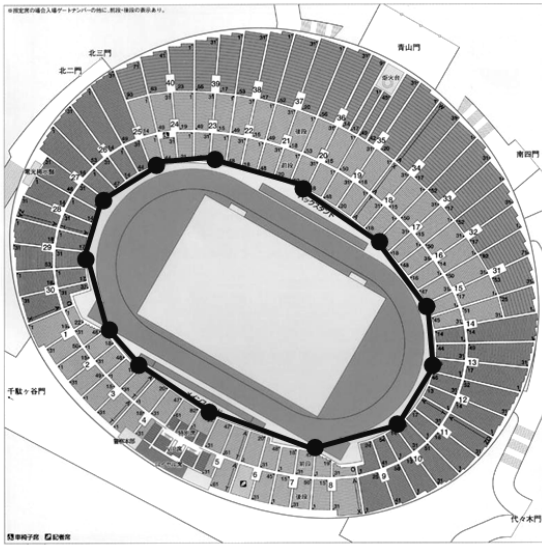


図7：霞ヶ丘競技場での分割点の空間的位置

6. まとめ

スポーツシーンを対象として、簡易な操作により背景領域における自由視点映像を生成する手法を提案した。本研究では、カメラから撮影された背景画像を地面領域と遠景領域に二分し、領域毎に自由視点映像を生成している。遠景領域においては、背景画像上の数点を指定するだけという簡易な操作により、写り込んだ遠景領域の3次元形状を適切に表現した3次元形状モデルを生成することにより、視点移動に伴う自然な見え方の変化を実現した。また、背景領域を仮想的に正面から撮影したように撮影画像を変換し貼り付け処理を行うことにより、分割小領域のつなぎ目のズレを解消し、自然な見え方の自由視点映像生成を実現した。実際の大規模空間において撮影した多視点映像を用いて自由視点映像を生成することにより、提案手法の有用性を確認した。今後は、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課題として挙げられる。本研究の一部は、総務省戦略的情報通信研究開発推進制度“ネットワークによる自由視点映像のライブ配信とインタラクティブ提示”により実施したものである。

文 献

- [1] F.Caillette and T.Howard,“Real-time Markerless Human Body Tracking with Multi-view 3-D Voxel Reconstruction”,Proc.SIGGRAPH, pp.43-54, 1996.
- [2] S.Yamazaki,R.Sagawa,H.Kawasaki,K.Ikeuchi and M.Sakauchi, “Microfacet Billboarding”,Proc. the 13th Eurographics Workshop on Rendering,pp.175-186, 2002.
- [3] T.Koyama, I.Kitahara and Y.Ohta, “Live Mixed-Reality 3D Video in Soccer Stadium”, Proc. of the 2nd IEEE and ACM International Symposium on Mixed and Augmented Reality (ISMAR2003), pp.167-178, 2003.
- [4] C.L.Zitnick,S.B.Kang,M.Uyttendaele,S.Winder and R.Szeliski, “High-quality Video View Interpolation Using a Layered Representation”, ACM SIGGRAPH and ACM Trans. on Graphics, Vol23, Issue 3, pp.600-608, 2004.
- [5] T.Kanade, P.Rander and P.J.Narayanan, “Virtualized Reality: Constructing Virtual Worlds from Real Scenes”, IEEE Multimedia, Vol.4, No1, pp.34-37, 1997.
- [6] 徐剛, 辻三郎, “3次元ビジョン”, 共立出版, 1998.
- [7] N.Inamoto,H.Saito, “Free Viewpoint Video Synthesis and Presentation of Sporting Events for Mixed Reality Entertainment”, International Conference on Advances in Computer Entertainment Technology (ACE 2004), pp.42-50, 2004.
- [8] R.Szeliski,“Video Mosaics for Virtual Environments”,IEEE Computer Graphics and Applications(c), Vol. 16, No.2, pp.22-30, 1996.
- [9] NaviView:仮想車載カメラ映像による運転者の視覚支援, 電子情報通信学会論文誌, Vol.J82-D-II ,No.10,pp.1816-1825, 1999.

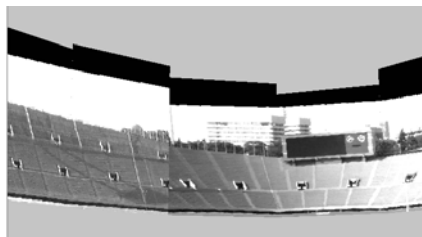


図8：生成した自由視点映像の例

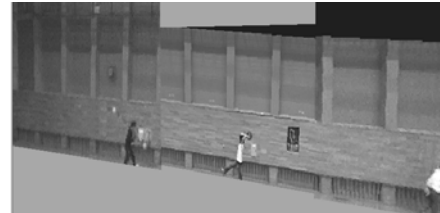
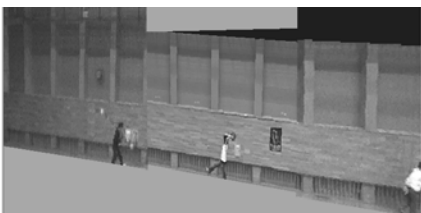


図9：背景画像を更新した例