

Morphing と Visual Hull を用いた動的環境での実時間自由視点画像生成

石川 智也[†] 山澤 一誠[†] 横矢 直和[†]

[†] 奈良先端科学技術大学院大学 〒630-0192 奈良県生駒市高山町 8916-5

E-mail: [†] {tomoya-i, yamazawa, yokoya}@is.naist.jp

あらまし 近年、複数台のカメラからの画像を統合し、自由な視点での画像を生成する研究が盛んに行われている。しかし、従来の手法では撮影環境が静的でなければならないことや、撮影対象がカメラを配置した内側に限られること等の問題点があった。本研究では、全方位カメラで撮影される画像を静的領域と動敵領域に分離し、静的領域に対しては **Morphing** を用い、動敵領域に対しては **Visual Hull** を用いてそれぞれ実時間で自由視点画像生成を行い、それらの画像を統合しユーザに提示する手法を提案する。さらに実験としてライブ映像を入力とした自由視点画像生成を実時間で行い、本手法が動的環境においても有効であることを示す。

キーワード 自由視点画像生成, **Morphing**, **Visual Hull**, 実時間, 動的環境, 全方位カメラ

Real-Time Generation of Novel View of a Dynamic Scene Using Morphing and Visual Hull

Tomoya ISHIKAWA[†] Kazumasa YAMAZAWA[†] and Naokazu YOKOYA[†]

[†] Graduate School of Information Science, Nara Institute of Science and Technology

Takayama 8916-5, Ikoma, Nara, 630-0192 Japan

E-mail: [†] {tomoya-i, yamazawa, yokoya}@is.naist.jp

Abstract Recently, generation of novel view images from images which were acquired by multiple cameras has been investigated. It can be applied to telepresence effectively. The conventional methods need assumptions such as the static scene and the limited position of the object within the positions of cameras. In this paper, we propose the method of generation of novel view images of a dynamic scene. The acquired images from omni-directional cameras are divided into static regions and dynamic regions. The novel view images are generated by morphing of static regions and by visual hull of dynamic regions in real-time. In the experiment, we show that the implemented system can generate novel view images in real-time from live video streams.

Keyword Novel View, Morphing, Visual Hull, Real-Time, Dynamic Scene, Omni-directional Camera

1. はじめに

近年、複数台のカメラからの画像を統合し、自由な視点での画像を生成する技術が盛んに研究されている [1,2,3,4,5]. この技術は遠隔地の情景等の現実環境を仮想環境に取り込み、ユーザに提示するテレプレゼンスにおいて、ユーザが視点位置を自由に変更できることから、より臨場感の高いテレプレゼンスを可能にする。我々はこれまで、単眼全方位カメラや全方位マルチカメラシステムを環境中に配置し、その視点位置における情景を遠隔地で 360 度の周囲を見回す事が可能なシステムを研究してきた [6]. しかし、これらのシステムにおいて視点移動を可能とするためには予め環境中のあらゆる経路において画像を取得しておくか、あるいはロボットや車両にカメラを搭載し遠隔操作により視点位置を変更する必要があった。本研究は、従来のシステムに自由視点画像生成の技術を導入し、自由な視

点でのテレプレゼンスが可能なように拡張することを目標とする。

従来、自由視点における画像を生成する研究には、斉藤らによる室内に内向きに多数のカメラを配置した環境で自由視点画像を生成する手法 [1], 稲本らや古山らによるサッカーシーンにおける自由視点画像生成手法 [2,3], Seitz らによる射影幾何を **Morphing** に用いた自由視点画像生成手法 [4], 冨手らによる全方位カメラを用いた **Morphing** による自由視点画像生成手法 [5] 等がある。斉藤らの手法は配置されたカメラの内側の物体であれば高画質な自由視点画像が生成可能である。しかし、周囲の環境については自由視点の画像を生成できない。稲本らの手法は、サッカーフィールドやゴールポストは平面で構成されていることや遠景は平面と見做せること等のサッカーに特化した仮定を利用して高画質な自由視点画像生成を行っている。しかし、

屋内など別の環境には適用できないという問題がある。また、古山らの手法はサッカープレイヤーと自由視点の位置は離れているという仮定を利用し、平面に画像を投影することで近似的に自由視点画像を生成する。このため処理が高速であるが、稲本らの手法同様別の環境に適用することは困難である。Seitzらの手法はMorphingを用いるため自由視点画像を高速に生成できるが画像間の対応を与えておく必要があり動的環境に適用するのは困難である。冨手らの手法は全方位カメラを用いて自由視点の全方位画像を生成するため、360度の周囲を見回すこと可能である。しかし、Seitzらの手法同様画像間の対応が必要であり、動的な環境に適用困難である。

そこで本研究では、環境中に全方位カメラを複数台配置し、その環境中の静的領域についてはMorphingを用い、動的領域についてはVisual Hullを用いて自由視点画像生成を行う手法を提案する。さらに本手法をライブ映像に適用し実時間での自由視点画像生成を行う。

2. 自由視点画像生成

2.1. 処理の流れ

図1に提案手法の処理の流れを示す。本手法では環境中に全方位カメラを複数台配置し、それらのカメラから得られる画像を用いて自由視点画像生成を行う。まず、それぞれのカメラからの画像に対し、背景差分処理により画像を静的領域と動的領域に分離する。静的領域に対してはMorphing、動的領域に対してはVisual Hullを用いてそれぞれ自由視点での画像を生成し、それらを合成した画像を最終的な自由視点画像として提示する。また、本手法ではカメラは固定して設置されており、位置・姿勢は既知であるものとする。

2.2. 静的領域と動的領域の分離

ここでは、複数の全方位カメラで取得された画像それぞれに対して静的領域と動的領域を分離する処理を行い、動的領域のみが含まれる全方位画像を生成する。この処理を行うための基本的な戦略としては、予め動的領域の存在しない背景画像を撮影しておき、背景画像と入力された画像との差によって動的領域を検出し分離することが考えられる。ただし、この戦略では照明変化やノイズに弱く長時間に渡っての自由視点画像生成は困難である。そこで本手法では、蛍光灯やCRTディスプレイのフリッカー、緩やかな照明変化、カメラに依存するノイズ等を考慮した背景差分処理の手法[7]を用いる。

この手法は、時系列に入力される画像の各画素につ

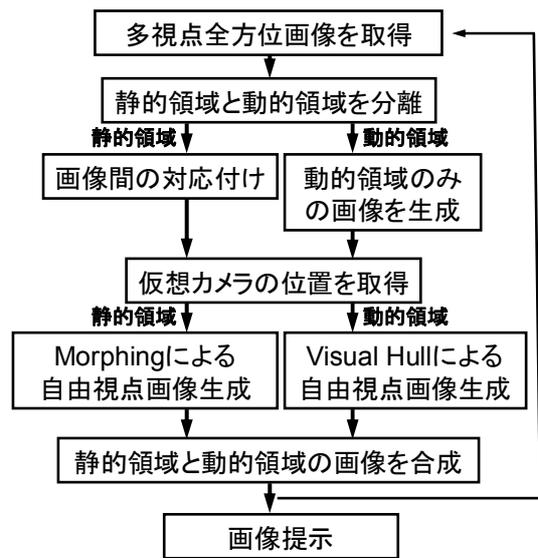


図 1. 提案手法の処理の流れ

いて平均輝度値や分散値を保持し、現フレームの輝度値が与えられた背景画素の輝度値のモデルの範囲であれば背景とみなし、それ以外の場合は動的領域とみなす。さらに、動的領域と判断された領域に対してその領域が新たに環境に加わった静的物体である場合を考慮して、ある程度時間が経つと背景と判断するよう背景画像の輝度値を変化させる。画像中の各背景画素の輝度 I は以下の式でモデル化する。

$$I = \bar{I} + \sigma \sin(2\pi\omega t) + k\zeta \quad (1)$$

\bar{I} は輝度の時間平均、 σ は輝度の振幅、 ω は輝度の周波数、 k は-1~+1までの値をとる係数、 ζ はカメラのみに依存した雑音を表す。 $\sigma \sin(2\pi\omega t)$ の項は蛍光灯やCRTディスプレイなどのフリッカー、窓の外の樹木の揺らぎなどの影響を表し、 $k\zeta$ の項はカメラに依存するノイズを表す。このとき背景画素の輝度値 I は以下の範囲をとり得る。

$$\bar{I} - \sigma - \zeta \leq I \leq \bar{I} + \sigma + \zeta \quad (2)$$

入力画素の輝度が式(2)の範囲内であればその画素は背景画素とみなし、そうでなければ動的領域とみなす。以下に処理手順を示す。 ζ は予め動的領域の存在しない環境で測定しておく必要があるが、環境によって変更する必要はない。以下に時系列で入力される画像に対する処理手順を示す。

- ① Nフレームまでは背景画像の生成を行う。Nフレームまでの画像を用いて各画素の輝度の平均値と分散値を求め、背景画像とする。Nは経験的に設定する値である。また、Nフレーム目までは動的領域の検出処理は行わない。

- ② N+1 フレームからは動的領域の検出と背景画像の更新を行う。現フレームの各画素が式(2)の範囲内であれば背景画像の輝度の平均値と分散値を更新する。範囲外であればその画素は動的領域と判断し、現フレームの偏差に重みを加えて分散値のみを更新する。

以上の処理により、入力される画像から動的領域のみが含まれる画像を生成する。

2.3. Morphing

ここでは、静的領域に対して **Morphing** を行い自由視点の全方位画像を生成する富手らの手法[5]を用いる。この手法は環境中の複数の地点で全方位画像を取得し、画像間の対応を与えることにより自由視点における全方位画像を生成する。以下にその手順を示す。なお、全方位カメラの位置、姿勢および入力画像間の対応点は既知とする。

- ① 与えられた各対応点について、それらの点の3次元位置を計算する。
- ② ①で得た3次元点を自由視点画像上に投影する(図2参照)。
- ③ 投影された点群に対してデローニー三角形分割を行い、三角形パッチを生成する。
- ④ ③で得た三角形パッチに対応する三角形の画像を各入力画像から切り出し、自由視点位置に基づいて算出される重みを用いてブレンドし、自由視点における全方位画像とする。

2.4. Visual Hull

動的領域に対する自由視点画像生成に **Visual Hull** を用いる。**Visual Hull** とは、多視点から撮影された物体のシルエットとカメラの光学中心によって形成される錐体の積集合空間を指す。この **Visual Hull** を表現する手法として視体積交差法[8]がある。しかし、視体積交差法では空間をボクセルで表現するため、対象とする空間が広がるとそのデータ量が膨大となる。そこで、本研究では **Visual Hull** を表現する手法に **Image-based Visual Hull**[9]を用いる。この手法は自由視点における画像面の各画素に対して、仮想カメラの光学中心(自由視点)とその画素を結ぶ直線が **Visual Hull** を貫通するか否かを判定することにより自由視点画像を生成する。その手順を以下に示す。

- ① 仮想カメラの画像面のある画素について、仮想カメラの光学中心とその画素を結ぶ直線を各実カメラの画像面上に投影する。各実カメラの画像面上に投影された仮想カメラの光学中心

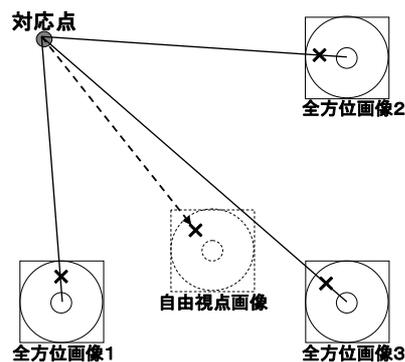


図 2. 対応点の自由視点画像への投影

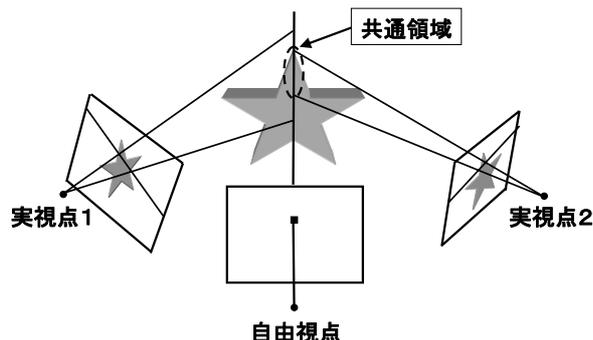


図 3. 共通領域の有無の判定

がエピポール、直線がエピポーラ線となる。

- ② 各実カメラに投影されたエピポーラ線上で動的領域と交差する領域を調べ、全てのエピポーラ線が動的領域と交差する領域(共通領域)があるならば、仮想カメラからの直線は **Visual Hull** を貫通する。共通領域が無い場合は、**Visual Hull** を貫通しないと判断できる(図3参照)。
- ③ 仮想カメラからの直線が **Visual Hull** を貫通する画素についてはその画素の色を決定するために、共通領域への各実カメラからのベクトルと自由視点からの直線のベクトルとを比較し、最もベクトルの向きが近い実カメラの画素の色を対象としている画素の色とする。
- ④ ①~③の処理を自由視点画像上の全ての画素に対して行う。

以上の処理により動的領域に対する全方位の自由視点画像を生成可能である。しかし、自由視点画像上の全画素について処理を行うと計算量が多く実時間での画像生成は困難となる。

そこで、計算量削減処理として予め動物体の環境中での位置を計算し、その位置を仮想カメラの画像面上に投影し動的領域を限定する。動物体の位置の計算には寺沢らの手法[10]を用いる。以下では、説明を分かりやすくするためにカメラを二つ使った場合の動物体の位置推定について述べる。

まず、各実カメラからの動的領域のみが含まれる全方位画像に対して直交座標から画像の中心を原点とする極座標 (r, θ) に変換し、 θ 軸方向に作成したヒストグラムから連続したある閾値より大きい値を持つ部分の画素領域の方位角範囲 $\theta_n (\theta_{n1}, \theta_{n2})$ を求める (図 4 参照)。

次に各実カメラから得られた方位角範囲から動物体の存在領域を推定する。図 5 より、動物体の存在領域は方位角範囲 $\theta_n (\theta_{n1}, \theta_{n2})$ の各ベクトルの交点から計算できる。各ベクトルの交点座標は、世界座標系でのカメラの位置ベクトルと方位角範囲の両端の角度に対応する単位ベクトル及び媒介変数を用いて以下の式で求めることができる。

(交点座標) =

(カメラの座標値) + (単位ベクトル) × (媒介変数)

よって 2 つの直線の交点の座標値は、直線を延ばしたカメラの位置ベクトルをそれぞれ $\vec{S}_1 (S_{1x}, S_{1y})$, $\vec{S}_2 (S_{2x}, S_{2y})$, 直線の単位ベクトルをそれぞれ $\vec{V}_1 (V_{1x}, V_{1y})$, $\vec{V}_2 (V_{2x}, V_{2y})$, 媒介変数をそれぞれ M_1 , M_2 とすると、以下の式で表せる。

$$\begin{aligned} (\text{交点の}x\text{座標}) &= S_{1x} + V_{1x} \times M_1 \\ &= S_{2x} + V_{2x} \times M_2 \end{aligned}$$

$$\begin{aligned} (\text{交点の}y\text{座標}) &= S_{1y} + V_{1y} \times M_1 \\ &= S_{2y} + V_{2y} \times M_2 \end{aligned}$$

以上の式より、各媒介変数は以下の式で求めることができる (図 6)。

$$M_1 = \frac{(S_{2x} - S_{1x}) \times V_{2y} - (S_{2y} - S_{1y}) \times V_{2x}}{V_{1x} \times V_{2y} - V_{1y} \times V_{2x}}$$

$$M_2 = \frac{(S_{1x} - S_{2x}) \times V_{1y} - (S_{1y} - S_{2y}) \times V_{1x}}{V_{2x} \times V_{1y} - V_{2y} \times V_{1x}}$$

各カメラからのばした全ての 2 直線について以上の処理を行い、動物体の存在領域を推定する。動物体の存在領域が分かればその領域を仮想カメラの画像面上に投影し、投影された範囲の画素に対してのみ Visual Hull の処理を行う。

3. 実験

提案手法により自由視点画像を生成するシステムを実装し、ライブ映像を入力とした実時間自由視点画像生成を行った。図 7 に示すような室内環境に、IEEE1394 で接続可能な全方位カメラを 3 台設置し、それらをすべて 1 台の画像取得・生成用 PC に接続した。使用した全方位カメラは、640×480 の解像度の画像を最高 15fps で取得可能であり、垂直方向には仰角 12 度、

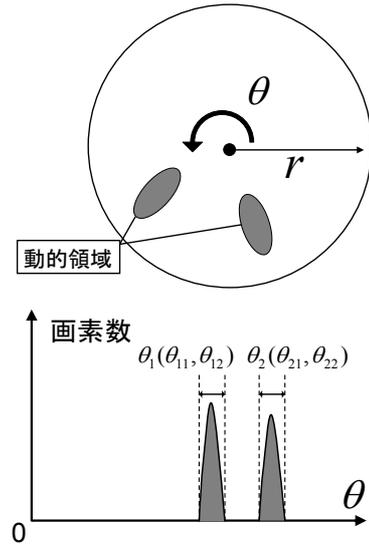


図 4. 極座標変換によるヒストグラムの作成

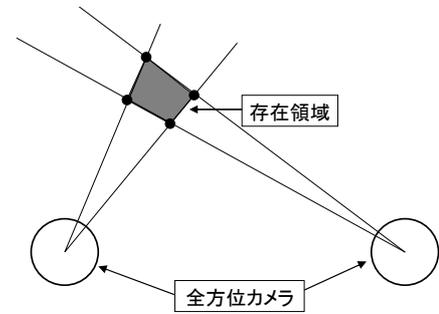


図 5. 推定される存在領域

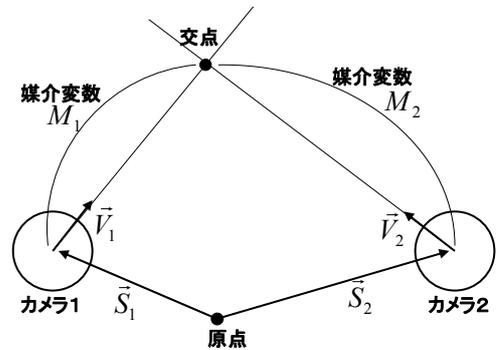


図 6. 直線の媒介変数表示

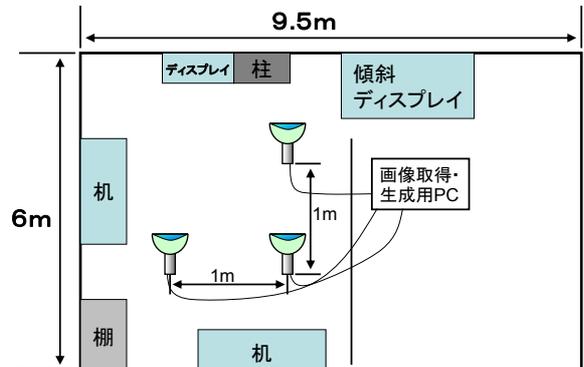


図 7. 実験環境

俯角 50 度, 水平方向には 360 度の視野を持つ. カメラ間の距離は 1m とした.

今回, 実験を簡略化するために **Morphing** に用いる画像は予め静止環境を撮影したものを使用し, 入力される映像を使用して画像を更新する処理は行わない. そして, 静的環境の対応点は予め手動で与えるものとし, 全方位カメラの位置・姿勢は既知とした. また, **Morphing** による制限から自由視点画像を生成できる範囲は各実カメラを結ぶ平面上に限定した. 前節において **Image-based Visual Hull** の計算量削減について述べたが, その計算量削減処理に加えて, ユーザが見ている範囲の画像のみを処理することとしている. 全方位カメラを使用する利点として, 同じ全方位画像を用いて複数のユーザがそれぞれ別の方向を見ることができるといことが上げられるが, 自由視点でのテレプレゼンスにおいては, それぞれのユーザがそれぞれ違う視点位置から見ることが想定されるため, ユーザが見ている範囲のみの自由視点画像を生成することは妥当だと考える. また, 動的領域の自由視点画像については穴の除去のためにメディアンフィルタを適用している.

図 8 に実験環境の様子と自由視点位置における生成画像を示す. 図 8 は左から仮想視点位置と視線方向, 実験環境の様子 (男女が歩行), 自由視点全方位画像, 自由視点全方位画像を平面透視投影変換した画像である. 画像取得・生成用 PC は Pentium4-3.2GHz を搭載した PC を用いて各自由視点画像を約 200ms で生成することができた. 内訳は, 同期した画像の取り込みと背景差分処理に約 50ms, **Morphing** に約 50ms, **Image-based Visual Hull** とメディアンフィルタに約 100ms となっていた. 生成した自由視点画像について, 動的領域に対してはユーザが見ている範囲のみ計算するため, 生成した全方位画像には男性は映っていない. 現実環境と生成した自由視点画像を比較すると, 動物体の位置が約 50cm 程度ずれていることが分かる. これは実カメラを設置した際の位置・姿勢のずれや静的環境の対応点数が少ないために **Morphing** によって生成した画像に歪みを生じていることが原因として考えられる. 同様に, オクルージョンの影響により対応点を与えられなかった領域については画像の歪みが大きかった. また, 今回の実験環境ではカメラ間の距離が狭く, 相互のカメラによりがオクルージョンができてしまい, 動的領域を正しく検出できない場合があった.

4. まとめと今後の課題

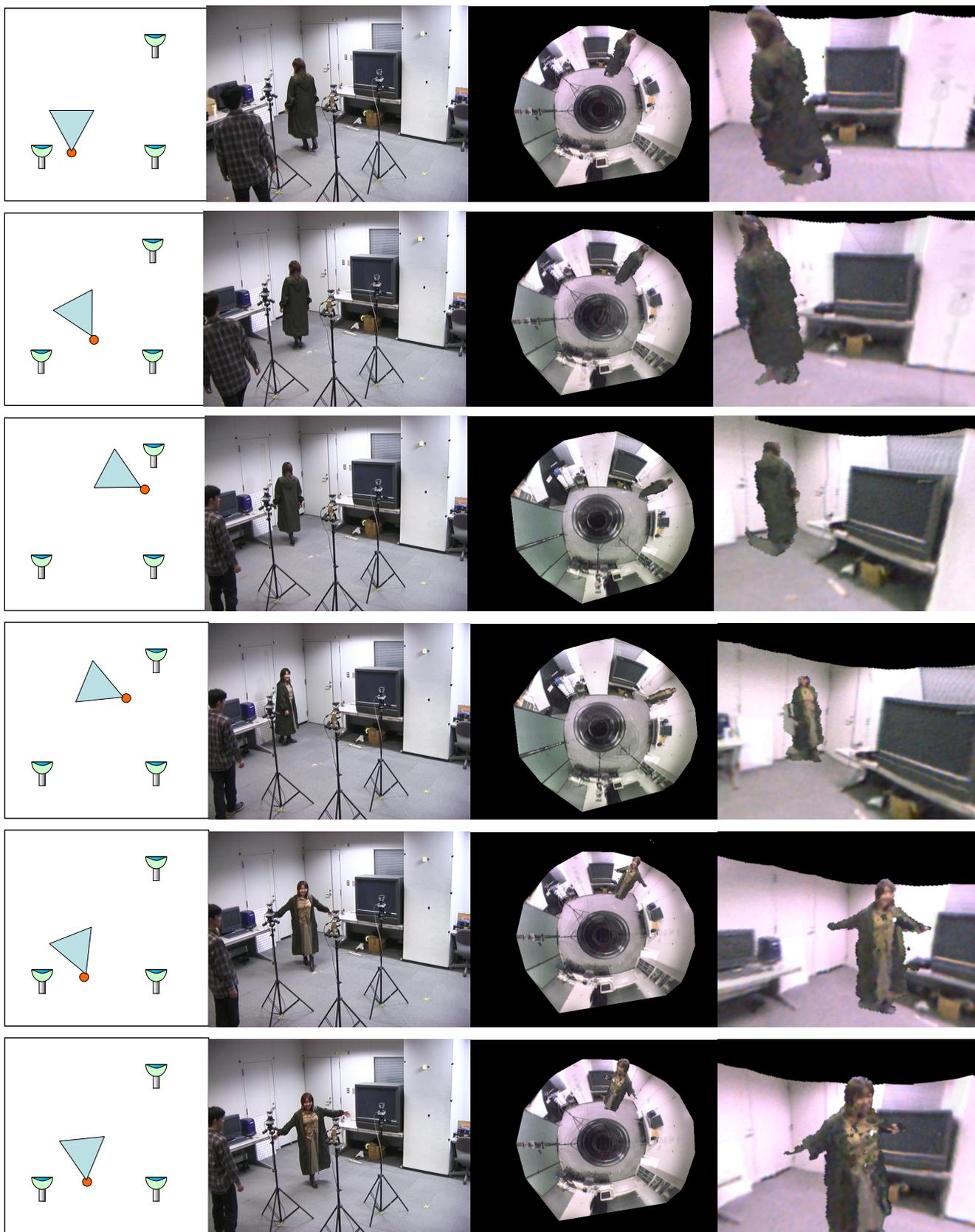
本稿では, **Morphing** と **Visual Hull** を用いることでシーンに依存せず動的環境に適用可能な自由視点画像生成手法を提案した. さらに, 自由視点画像生成の計算

処理を削減することによりライブ映像を用いた実時間自由視点画像生成が可能であることを実験により示した.

今後は, 実環境と生成した自由視点画像との位置ずれの軽減や生成する画像の画質向上のためにカメラのキャリブレーションなどについて検討を行う. また, 今回は予め手動で行った実カメラの背景画像間の対応付けの自動化を行う. さらに, より多くのカメラを環境中に配置し, それらのカメラからの画像を用いて環境中をウォークスルーできるようにシステムを拡張する. その際, 自由視点位置に依存せず並列に処理可能な背景差分処理や動的領域の位置推定などと, 各ユーザの自由視点位置に依存する画像生成処理などを別々の PC で行うことにより, 複数人が同時に別々の自由視点画像を見ることが可能なシステム構成にする予定である.

文 献

- [1] H.Saito, S.Baba, and T.Kanade, "Appearance-Based Virtual View Generation from Multicamera Videos Captured in the 3-D Room", *IEEE Trans. On Multimedia*, Vol.5, No.3, pp.303-316, September 2003.
- [2] 稲本奈穂, 齊藤英雄, "視点位置の内挿に基づく 3 次元サッカー映像の自由視点鑑賞システム", *映像情報メディア学会誌*, Vol.58, No.4, pp.529-539, 2004.
- [3] 古山孝好, 北原格, 大田友一, "仮想化現実技術を用いた大規模空間自由視点映像のライブ配信と提示", *電子情報通信学会 技術研究報告 PRMU2003-22*, pp.61-66, May 2003.
- [4] S.M.Seitz and C.R.Dyer, "View Morphing", *ACM SIGGRAPH'96*, pp.21-30, 1996.
- [5] 富手要, 山澤一誠, 横矢直和, "複数の全方位画像を用いた広範囲なウォークスルーの実現", *画像の認識・理解シンポジウム(MIRU2002)講演論文集*, Vol. II, pp.353-358, August 2002.
- [6] 石川智也, 山澤一誠, 横矢直和, 佐藤智一, 池田聖, 中村豊, 藤川和利, 砂原秀樹, 横矢直和, "Web ブラウザと全方位動画画像を用いたテレプレゼンスシステム—車載全方位カメラや全方位マルチカメラシステムを用いた実装—", *画像の認識・理解シンポジウム(MIRU2004)講演論文集*, Vol. I, pp. 127-132, July 2004.
- [7] 森田真司, 山澤一誠, 横矢直和, "全方位画像センサによるネットワークを介したテレプレゼンスシステムと遠隔監視システム", *電子情報通信学会 技術研究報告 PRMU2002-83*, September 2002.
- [8] 松山隆司, 高井勇志, ウ小軍, 延原章平, "3次元ビデオ映像の撮影・編集・表示", *日本バーチャルリアリティ学会論文誌*, Vol.7, No.4, pp.521-532, 2002.
- [9] W.Matusik, C.Buehler, R.Raskar, S.Gortler, and L.McMillan, "Image Based Visual Hulls", *ACM SIGGRAPH2000*, pp.369-374, 2000.
- [10] 寺沢征彦, 山澤一誠, 竹村治雄, 横矢直和, "複数の全方位画像センサを用いた遠隔監視システムにおける複数移動物体の存在領域推定", *電子情報通信学会 技術研究報告*, PRMU2000-195, February 20001.



自由視点位置

実験環境の様子

自由視点全方位画像

平面透視投影変換画像

図 8. 実験環境の様子と生成された自由視点画像