

## 未校正多視点可動カメラを用いた高精細な自由視点画像生成

伊藤 洋輔<sup>†</sup>      齋藤 英雄<sup>†</sup>

<sup>†</sup> 慶應義塾大学理工学部情報工学科

本研究では、広い空間で移動する物体を対象とした自由視点画像生成を目的とする。近年、複数台の固定カメラによって撮影された画像からの3次元モデリング・レンダリングに関する研究は盛んに行われているが、固定カメラで撮影された画像を用いる場合、移動する対象物体を画角に収めるために対象からズームアウトして撮影する必要があり、高精細な自由視点画像を生成することはできないという問題点があった。そこで本手法では、移動する対象物体を複数の可動カメラにより追跡し、ズームした状態で撮影された画像を用いることにより、高精細な自由視点画像を生成する。更に、未校正多視点可動カメラからの3次元モデリング・レンダリングを可能にするために、可動カメラの他に2台の固定カメラを設置し、これらにより決定される射影グリッド空間において3次元形状の復元を行うシステムを提案する。

## Synthesis of High Resolution Free-Viewpoint Images from Uncalibrated Multiple Moving Cameras

Yosuke Ito<sup>†</sup>, Hideo Saito<sup>†</sup>

<sup>†</sup> Department of Information and Computer Science, Keio University

We propose a novel method to synthesize free-viewpoint images for a moving object, which is captured by uncalibrated multiple moving cameras. If multiple fixed cameras are used for capturing a moving object, we must zoom out in order to capture the moving object within FOV of the cameras. Such zooming-out limits the resolution for the moving object in the captured images. In the proposed method, we use multiple moving cameras that capture the moving object in the center of the images with high resolution. For shape reconstruction of the object from the uncalibrated multiple moving camera images, two fixed cameras are employed for determining Projective Grid Space which defines a projective 3D coordinate in the object space. The coordinate in PGS can be related to every moving camera by fundamental matrices between the moving camera and the fixed cameras. In the experiment that is performed for demonstrating the efficacy of the proposed method, high resolution free-viewpoint images can be successfully synthesized by the proposed method.

### 1 はじめに

我々は、大規模な空間で行われる動きのあるイベントを対象として、多視点カメラによって撮影された映像をもとに自由に視点を变化させた映像を提示する研究に取り組んでいる。自由視点映像の生成システムとして実際に使用され、テレビ放送が行われたシステムとしては米国カーネギーメロン大学とCBSの共同開

発による Eye Vision[1] がある。スタジアムを取り囲むように数十台のカメラを配置し、そのうちの1台のカメラをプロのカメラマンが動かすと残りのカメラもフィールド上の同一視点を追隨するシステムである。また、Eye Vision の他には映画製作の現場において、100台程度の固定カメラを用いた Bullet-Time[2] と呼ばれるシステムも実現されている。

しかしこういったシステムは、単純に多数のカメラ

の切り替えにより自由視点を実現する手法をとっているため、視点間をシームレスにつなぐためには、非常に多くのカメラ台数が必要となってくる。カメラ台数が多くなるとキャリブレーション、カメラの制御などに手間がかかり、またシステムを構築するためのコストも多くなってしまふ。また、どれだけ多くのカメラを設置したとしても、その視点の数には限りがあるため、視聴者や映像製作者の望む厳密な意味での自由視点は実現できない。近年、コンピュータ処理能力の増大も伴い、コンピュータビジョンの技術を用いることで、限られた数の視点より得られた画像から仮想的な視点を合成する研究が盛んに行われており、大規模空間にも適用できる手法 [3] も提案されている。そういった研究では、高精細な画質、リアルタイムの画像生成処理などが目標となっており、本研究でも高精細な自由視点画像を一つの目標としている。

一般に自由視点画像を合成する際、入力画像の画質が出力画像の画質に大きく影響する。そのため高精細な画像を撮影するために、従来から頻繁に使用されている固定カメラではなく、可動カメラを使用することも考えられる。固定カメラと比較すると、可動カメラは広い撮影空間が対象であっても、少ない台数で高精細な画像を撮影することができる。しかし、可動カメラはその位置や姿勢が動的に変化するために、正確なカメラキャリブレーションを実現することは困難である。可動カメラを十分に高い精度でキャリブレーションする方法として、可動カメラの位置や姿勢を計測するセンサを利用する手法 [4] や、撮影環境中に人工的なマーカを設置し、撮影された画像を用いる手法などが提案されている。しかし、前者は特別な装置が必要であり、また後者は、特に本研究のように撮影空間が広い場合、キャリブレーション用の点が非常に多く必要となるという欠点がある。そこで、キャリブレーションを必要とせず、撮影された画像のみで対象の 3 次元モデリング・レンダリングが可能であれば、より制約の少ないシステムが可能になると期待できる。

本手法では可動カメラの他に 2 台の座標系構築用の固定カメラを設置し、射影グリッド空間 (Projective Grid Space) [5] を構築する。この射影グリッド空間は、従来困難であった未校正の多視点カメラの画像からの 3 次元復元の枠組みを提供するものであり、キャリブレーションが困難な可動カメラに非常に適した枠組みだと言える。この射影グリッド空間において視体

積交差法により 3 次元モデルを復元し、3 次元モデルより得られる入力視点間の対応関係からその中間視点を Image-Based Rendering により合成する。

本稿では、未校正の可動カメラと固定カメラを組み合わせた自由視点映像生成システムを提案する。また実画像を入力とした実験を通して、可動カメラを用いることで高精細な自由視点画像を生成でき、本手法が有用であることを示す。

## 2 理論

ここでは本手法で使用する理論について述べる。

### 2.1 射影グリッド空間

射影グリッド空間 (Projective Grid Space) は、基準となる 2 台のカメラの中心射影によって構成される 3 次元空間である。3 次元空間を離散的に扱う場合、2 次元画像における単位面積である画素 (Pixel) と同様に 3 次元の単位体積は Voxel と呼ばれる。またここでは Voxel を区切る直線をグリッドと呼ぶことにする。

図 1 は一般的に 3 次元空間を離散的に扱う場合の Euclidian Grid Space と、Projective Grid Space の関係を示したものである。Euclidian Grid Space はカメラとは関係なく、ある 1 点を原点として定義された直交 3 次元空間であり各 Voxel の大きさも等しい。これに対し、射影グリッド空間は複数台あるカメラのうち 2 台を基底カメラとし、そのカメラ 2 台のカメラによって中心射影的に決定される 3 次元空間である。この図のように射影グリッド空間は一般のユークリッド空間から見ると各 3 軸及び各グリッドは直交するとは限らない。また各 Voxel の大きさも一定ではなく、基底カメラに近いほど小さく、遠いほど大きくなるという性質を持つ。

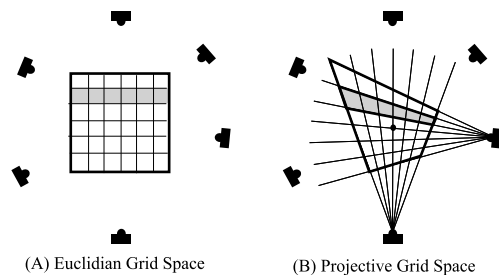


図 1: Euclidian Grid Space と Projective Grid Space

次に実際の射影グリッド空間の定義について述べる。まず複数台のカメラシステムのうちの2台を選択し、基底カメラ1、基底カメラ2とする。基底カメラ1から得られる画像のX軸、Y軸をカメラの視点から中心射影的に空間に投影したものを射影グリッド空間のP軸、Q軸とする。同様に基底カメラ2から得られる画像のXから射影グリッド空間のR軸を決定する。これらP、Q、R軸を3軸として定義される3次元空間が射影グリッド空間である。また射影グリッド空間における座標は各画像の画素とその視点とを結ぶ直線によって定義されるものとする。つまり射影グリッド空間のP座標、Q座標がそれぞれ基底カメラ1の画像のX座標、Y座標に対応し、R座標が基底カメラ2のX座標に対応する。基底カメラ以外の他の全てのカメラもこの座標系を用いるものとする。

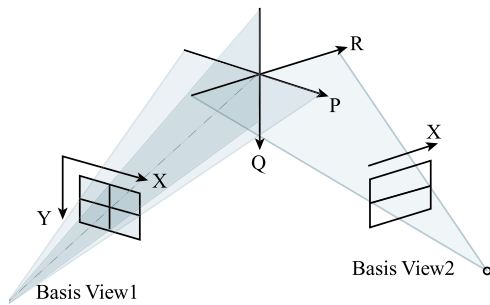


図 2: 射影グリッド空間の定義

次に射影グリッド空間における視点位置の3次元座標の定義を示す。基底カメラ1、2の各視点を  $C_1, C_2$  としその他カメラの視点を  $C_i$  とする。射影グリッド空間の定義より基底カメラ1から得られる画像のX軸、Y軸によってカメラ1の視点  $C_1$  から中心射影的にP軸、Q軸が定められているから、 $C_1$  のP座標、Q座標は一意には定まらずあらゆる値を取りうるが、ここでは処理の汎用性を考えて基底カメラ1の画像中央のX座標  $X1_c$ 、Y座標  $Y1_c$  をそれぞれ  $C_1$  のP座標、Q座標とする。射影グリッド空間のR軸は画像2のX軸として定義されることから、画像2に対するエピポールとなる  $C_1$  のR座標は  $e_{21}$  のX座標 ( $e_{21_x}$ ) となる。つまり、 $C_1$  の座標は  $C_1(X1_c, X2_c, e_{21_x})$  になる。同様に、 $C_2$  の射影グリッド空間におけるR座標も一意には定まらないが、基底カメラ2の画像中央のX座標  $X2_c$  を用いて定義されるものとする。 $C_2$  は画像1に対するエピポールとなるので、そのP座標、Q座標はそれぞれ  $e_{12_x}, e_{12_y}$  となる。つまり、 $C_2$  の座標は

( $e_{12_x}, e_{12_y}, X2_c$ ) と表すことができる。また、基底カメラ1、2以外のカメラの視点  $C_i$  の座標は、基底カメラ1、2のエピポールの座標を用いて、 $C_i(e_{1i_x}, e_{1i_y}, e_{2i_x})$  であることがわかる。

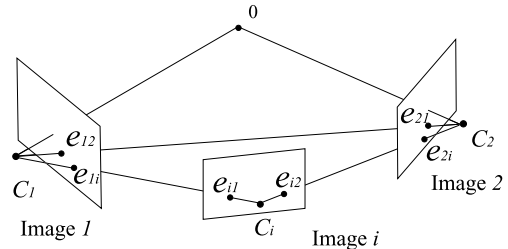


図 3: 射影グリッド空間における視点座標

### 3 提案する手法

#### 3.1 撮影環境

可動カメラはその位置・姿勢を変化させ、対象物体を追跡することができるため、広い撮影空間においても高解像度な画像が得られるという長所がある。しかし、位置・姿勢が動的に変化するためにカメラキャリブレーションを正確に行うことは困難である。また、可動カメラは対象物体に対してズームアップして撮影することが目的のため、もしカメラキャリブレーションを行うにしても、狭い画角に十分な数の特徴点を写り込ませる必要があるために撮影空間全体に多量の特徴点を配置する必要性が生じる。その多量の特徴点の3次元位置の計測は非常に手間を要し、また計測誤差も大きくなってしまふと考えられる。

そこで本手法では、可動カメラの他に2台の固定カメラを設置し、この2台の固定カメラを基底カメラとした射影グリッド空間 (Projective Grid Space) を座標系として構築する。この2台の基底カメラは座標空間設定用のカメラとして使用することにし、撮影空間全体を写すために撮影対象から離れた位置に設置する。

多視点による撮影システムとしてまず考えられるのは図4(a)のように、対象から見て扇状にほぼ同じ高さに各カメラを並べるような撮影システムである。ここで、射影グリッド空間の座標軸を Euclidian Grid Space の座標軸に近づけるために2台の基底カメラの視線方向のなす角度は  $90^\circ$  になるように設置を行う。

しかし図4(a)のように、2台の基底カメラの中心軸によって作られる平面上に可動カメラが存在するよう

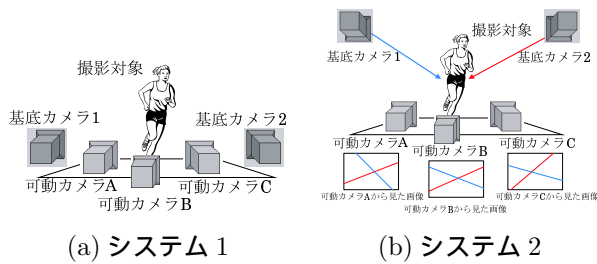


図 4: 撮影システム

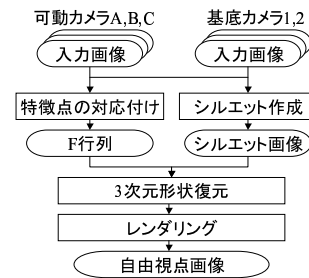


図 5: 本手法の流れ

なカメラ配置では、可動カメラにおける 2 台の基底カメラからのエピポーラ線の交差角が非常に狭く、平行に近い状態となる。詳細は 3.4 節で述べるが、本手法ではエピポーラ線の交差する点を正確に求める必要があり、エピポーラ線の交差角が狭い場合、 $F$  行列の推定誤差が及ぼす影響が大きくなってしまふ。

そこで、本手法では上述のような問題を解決し、設置の容易さも考慮に入れた図 4(b) に示すようなカメラ配置による撮影システムを提案する。この撮影システムでは基底カメラから対象物を見下ろすように設置を行う。この配置は、いずれの可動カメラにおいても 2 本のエピポーラ線が十分に大きな角度で交差するので、 $F$  行列の推定誤差の影響を軽減できるカメラ配置であるといえる。

### 3.2 処理の概要

ここでは、本手法の処理の概要について述べる。まず大まかな処理の流れを図 5 に示す。

本手法は、可動カメラ-基底カメラ間の  $F$  行列の推定、射影グリッド空間内における対象物体の 3 次元復元、自由視点画像のレンダリングの 3 段階に大きく分けることができ、以後、順を追って解説する。この 3 つの処理を毎フレーム行い、仮想視点における画像を作成する。なお、この 3 つの処理の前処理として、2 つの基底カメラ間の  $F$  行列を推定する必要がある。

### 3.3 $F$ 行列の算出

射影グリッド空間を構築するためには、2 つの基底カメラ間の  $F$  行列を、あらかじめ求めておく必要がある。2 つの基底カメラは固定カメラなので、初期フレームで一度、2 視点間の点の対応付けを行い  $F$  行列を求

めてしまえば、次フレーム以降は更新を行う必要はない。本手法ではコーナ等の自然特徴点を用いて、2 つの基底カメラ間の点の対応付けを手動で行い、 $F$  行列を推定する。

次に、自由に可動するカメラを統一的に取り扱うためには、フレーム毎に 2 つの基底カメラに対する  $F$  行列を更新すればよい。初期フレームでは可動カメラ、基底カメラ間の特徴点の対応付けを手動で行い、 $F$  行列を推定する。次フレーム以降は初期フレームで対応付けた特徴点の追跡を行うことにより対応付けを行い、 $F$  行列を更新する。特徴点の追跡を行う際に、誤追跡が発生する可能性があるので RANSAC(Random Sample Consensus)[6] による誤追跡の除去を行う。

### 3.4 対象物体の 3 次元復元

本手法では視体積交差法 (SS 法) により 3 次元復元を行う。従来の視体積交差法 [7][8] では各カメラごとの射影行列を用いて、一定の空間中のボクセルを各シルエット画像上に投影していた。本手法では、射影グリッド空間中のボクセルを画像間の  $F$  行列を用いて各カメラのシルエット画像へ投影することにより SS 法を実行する。実際の 3 次元モデル作成の流れは次のようになる。

まず射影グリッド空間にある一定の空間を定義し、その空間に含まれるすべてのボクセルを各シルエット画像上に投影し、シルエットの内外判定を行うことにより、存在するボクセルの集合として 3 次元ボクセルモデルを復元する。

次に射影グリッド空間上のボクセルを各シルエット画像に投影する方法について述べる。投影するボクセルを  $A(p, q, r)$ 、基底カメラ 1, 2 から得られる各シルエット画像を画像 1, 画像 2, 可動カメラから得られ

るシルエット画像を画像  $i$  とする．またそれぞれの画像上の  $A$  の投影点を  $a_1, a_2, a_i$  , 画像  $h$  の画像  $k$  に対する  $F$  行列を  $F_{hk}$  と表すものとする．

ボクセル  $A(p, q, r)$  は射影グリッド空間の定義より, 画像 1 では  $a_1(p, q)$  に投影される．また画像 2 の画像 1 に対する  $F$  行列を用いて  $a_1$  を画像 2 に直線  $l$  として投影すると直線  $l$  は式 1 のように定義される．

$$l = F_{21} \begin{bmatrix} p \\ q \\ 1 \end{bmatrix} \quad (1)$$

射影グリッド空間の定義よりボクセル  $A(p, q, r)$  の画像 2 における投影点  $a_2$  の  $x$  座標は  $r$  であるから,  $a_2$  はこの直線  $l$  上の  $x$  座標が  $r$  の点として定めることができる．

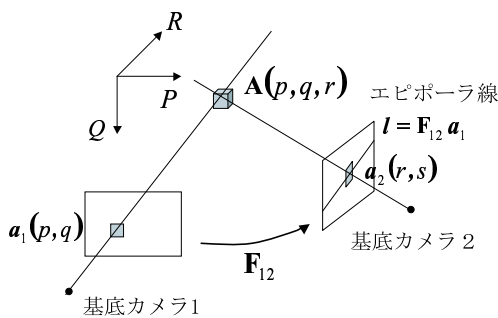


図 6: 基底カメラ画像への投影法

次に, 画像  $i$  の投影点  $a_i$  の決定法を述べる．先ほどと同様に  $F_{1i}$  を用いて  $a_1$  を画像  $i$  に直線  $l_1$  として投影する．また  $a_2$  を  $F_{2i}$  で画像  $i$  上に直線  $l_2$  として投影する．この 2 直線  $l_1$  と  $l_2$  の交点が画像  $i$  の投影点  $a_i$  となる．

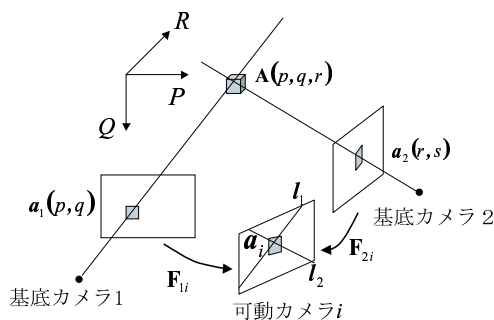


図 7: 可動カメラ画像への投影法

最後にボクセルモデルから 3 次元表面形状を復元する．その際, Marching Cubes 法 [9] と比較して位相問題に対してロバスト, 実装が容易という観点から Deformed Cubes 法 [10] を使用し, 3 次元表面形状を復元する．

### 3.5 レンダリング

作成された 3 次元モデルをもとに自由視点画像を合成する．自由視点を合成する手法は様々なものが提案されているが, そのひとつに, 合成しようとする付近の実視点画像を複数枚選び, 画像間の対応関係を 3 次元モデルから求め, この対応関係に基づいて見かけ上の中間視点画像を合成する手法が提案されている [11] . このレンダリング手法は 3 次元モデル復元誤差の影響を受けにくく, 画質劣化の少ない画像を合成できるという長所があるが, 自由視点は現実のカメラ視点の中間にしか設定することはできないという短所がある．

本研究の目的として高精細な自由視点画像を作成することがあげられるため, このレンダリング手法は適していると言える．また, 本手法ではカメラはイベントに対し良い方向から撮影できており, 各カメラの中間に仮想視点が設定できれば必要とされる視点をほぼ合成することができるので, 本手法ではこのレンダリング手法を用いることにした．

#### 3.5.1 Z-Buffer 法

各入力画像に対して, 3 次元モデル表面のどの三角パッチがオクルージョンとなっているかを判定するため, 入力視点において Z-Buffer を作成する．シルエット画像にボクセルを投影するときと同じ要領 (3.4 節) で 3 次元モデル表面の各三角パッチの 3 頂点を入力画像へ投影する．入力画像に Z-Buffer を用意し, 各画素の Z-Buffer には, その画素に投影された三角パッチの重心の 3 次元位置とその画像の視点の 3 次元位置との距離を格納する．同画素に投影される三角パッチが複数ある場合, 視点に一番近い三角パッチに対する距離を格納する．また投影されるボクセルがない画素の画素値は初期化しておく．このようにして Z-Buffer に距離情報を持った入力画像を作成する．

ここでの距離は, 射影グリッド空間において式 2 の

ように定義される．

$$D = \sqrt{(p_1 - p_2)^2 + (q_1 - q_2)^2 + (r_1 - r_2)^2} \quad (2)$$

ただし， $(p_1, q_1, r_1), (p_2, q_2, r_2)$  は射影グリッド空間における任意の 2 点とする．ここで三角パッチと視点との距離を求めるために，射影グリッド空間における各視点の 3 次元位置が必要になるが，射影グリッド空間の定義 (2.1 節) によりエピポラ幾何を用いて求めることができる．

### 3.5.2 自由視点画像の合成

先ほど作成された距離情報を持った入力画像を 2 枚用いて仮想視点画像を作成する．ここでは物体上の各点と視点との距離の違いによって生じる画像上の視差を利用して，その 2 視点の間における各点の画像上の座標を決定する．

いま，投影すべき三角パッチの頂点のうち一つを  $V$  とし，その距離情報をもった実視点画像 1 上での投影点を  $v_1(x_1, y_1)$ ，実視点画像 2 上での投影点を  $v_2(x_2, y_2)$  とし，画像間の距離を調節するウェイトを  $w$  とすると仮想視点の画像における  $V$  の投影点  $v_3$  の座標は

$$v_3 = w \begin{bmatrix} x_1 \\ y_1 \end{bmatrix} + (1 - w) \begin{bmatrix} x_2 \\ y_2 \end{bmatrix} \quad (3)$$

と表される．投影すべき三角パッチの頂点のうち，残りの 2 頂点も同様に仮想視点に投影する．

仮想視点と 2 つの実視点における三角パッチの位置が決定したのち，実視点上の三角パッチから仮想視点上の三角パッチへのアフィン変換を用いることにより，三角パッチを構成する画素値を式 3 と同様に線形的に決定する．その際，三角パッチを実視点画像に投影したとき，三角パッチを構成する各画素について，射影グリッド空間中のその三角パッチと実視点間との距離と，実視点画像の Z-Buffer に格納されている距離の比較を行う．もし距離が異なる場合，その画素はオクルージョンとなっていることになる．この場合，その画像の座標は仮想視点画像における投影点の座標計算に用いるが画素値は用いない．画素値はもう一方の画像の画素値をそのまま用いる．2 つの実視点両方に対してオクルージョンの場合，この 2 つの視点では見えないものと判定して画素値は求めない．これを 3 次元

モデル表面上の全ての三角パッチに対して行い，仮想視点画像をレンダリングする．

## 4 実験と考察

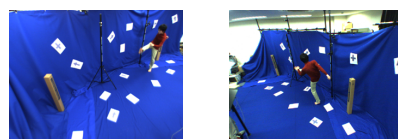
本手法を多視点画像に対し適用した．実験 1 では本システムを使用し，また，実験 2 では固定カメラのみによる撮影システムを使用し，本システムとの比較を行った．

実験に用いたカメラの仕様は，可動カメラ，固定カメラともに同様である．本手法ではカメラの姿勢が動的に変化し，また，運動する物体に対して自由視点を合成するので，各カメラの厳密な同期がとれていることが重要である．そのため，撮影に用いたカメラは各カメラの厳密な同期がとれるものを使用した．可動カメラは台数分の撮影者により手動で動かし，対象を追跡・撮影した．

撮影を行う前に，各カメラ間で特徴点の対応をとるためにマーカ，ポール，箱などを撮影環境に配置した．

### 4.1 可動カメラと固定カメラを用いた実験

撮影環境として可動カメラ 3 台，固定カメラ (基底カメラ) 2 台を用い，移動する人物の撮影を行った．各カメラ配置は図 4(b) に示したとおりである．撮影された  $640 \times 480$  の解像度をもつ画像を図 8 に示す．



(a) 基底カメラ 1 (b) 基底カメラ 2



(c) 可動カメラ A (d) 可動カメラ B (e) 可動カメラ C

図 8: 入力画像

これらをもとに実際に作成した 3 次元モデルに対して視点をを変えてレンダリングしたものを図 9 に示す．なお，この 3 次元モデルは射影グリッド空間の座標を

ユークリッド座標とみなして OpenGL によりレンダリングを行ったものである。

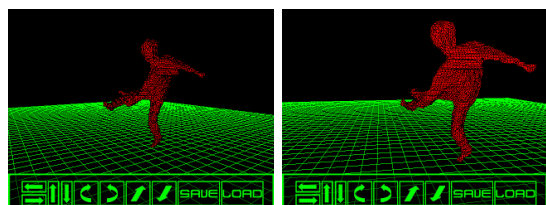


図 9: 3次元表面モデル

次に、作成した3次元モデルを用いて合成した自由視点画像の例を図10に示す。これは、可動カメラB-可動カメラC間の中間画像をウェイトを変化させて合成したものである。レンダリング画像のテクスチャの劣化や欠落が少ないことから、本手法では、姿勢が動的に変化するカメラでも、カメラ校正を行わずに自由視点が合成できたことが確認できる。

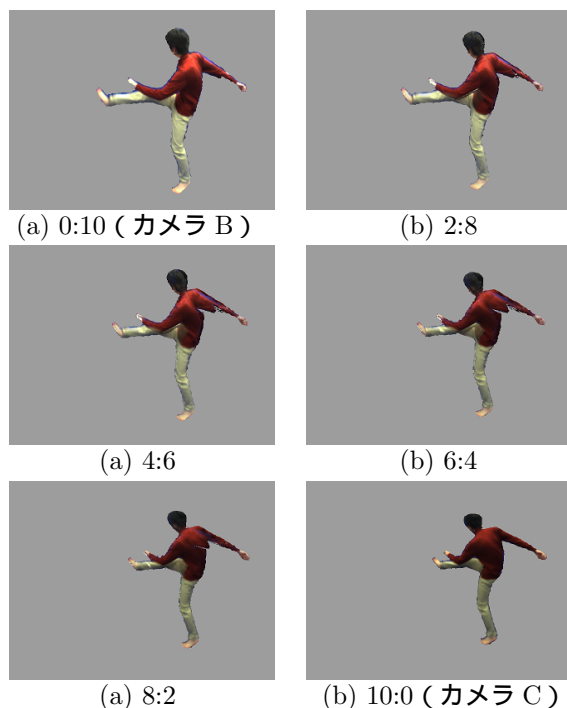
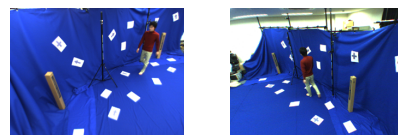


図 10: 可動カメラ B-可動カメラ C 間の中間視点画像

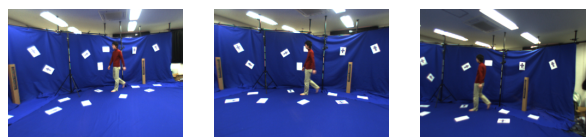
## 4.2 固定カメラのみを用いた実験

撮影環境として固定カメラ5台を用い、移動する人物の撮影を行った。各カメラ配置は図4(b)と同様の

ものであるが、可動カメラの代わりに固定カメラを設置した。この固定カメラは撮影空間全体を写すように、対象物体から離れた位置に設置を行った。なお、実験1と同様に両端の2台のカメラを基底カメラとする。撮影された  $640 \times 480$  の解像度をもつ画像を図11に示す。



(a) 基底カメラ 1 (b) 基底カメラ 2



(c) 固定カメラ A (d) 固定カメラ B (e) 固定カメラ C

図 11: 入力画像例

入力画像から3次元モデル作成し、それをもとに合成した自由視点画像の例を図12に示す。これは基底カメラ1-固定カメラA間の中間画像をウェイトを変化させて合成したものである。実験2の結果は、実験1の結果と比較すると、縦・横で約  $1/2$  の解像度しか得られていないことがわかる。また、レンダリング画像の欠落や劣化が実験1, 2ともに同程度なことから、本システムの有効性が示されたといえる。

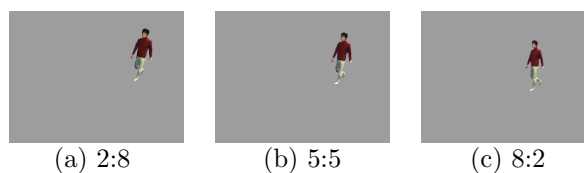


図 12: 基底カメラ 1-固定カメラ A 間の中間視点画像

## 4.3 考察と検討

本手法では、注目するフレームで新たに出現した特徴点の対応付けは行っていない。そのためフレームが進むにつれて、カメラの画角の外に出てしまう特徴点や撮影対象によってオクルージョンとなってしまう特徴点を追跡できなくなり、対応点が少なくなることにより F 行列の精度が落ちてしまう。本手法では、あらかじめ基底カメラ同士で多くの点の対応付けを行って

いるため、それらの点を可動カメラへ投影することで、注目フレームで新たに出現した特徴点を対応付けることが可能である。

まず、3.3 節で示した処理を行った後、基底カメラ同士で対応付けられている全ての点を、エピポーラ線の交差する点として可動カメラに投影する。可動カメラの画角内に投影された点のうち、まだ対応付けられていない点は注目フレームで新たに出現した特徴点、または誤追跡除去の段階で除かれた点である。そのような点があった場合、新たに出現した特徴点であるかどうかの判定が必要である。その判定には、基底カメラ上のその点と、可動カメラ上の対応する点での相関を取ることが考えられるが、基底カメラと可動カメラで視点が大きく違うために、その点の周辺の見え方も大きく違い、単純にウィンドウをかけただけでは正確な相関を取ることができない。近年、Tuytelaars らによりアフィン不変な領域を見つけ出し、その対応付けを行うという Wide baseline stereo matching の手法 [12] が提案されており、同様の手法を用いて、基底カメラ上の点の周辺におけるアフィン不変領域と、可動カメラ上の点の周辺におけるアフィン不変領域を見つけ、それらのモーメント不変量を比較することで相関を求めることができると考えられる。

基底カメラ同士で対応付けられている全ての点を可動カメラに投影する際、もし基底カメラと可動カメラの間における F 行列の精度が低い場合、投影される座標は本来投影されるべき座標からずれてしまう。そのため、投影された座標の周辺でアフィン不変領域のモーメント不変量が最も近いものを探索したのち、先ほどのような判定を行う必要があると考えられる。

## 5 おわりに

本稿では、移動する対象物体を複数の可動カメラにより追跡し、ズームした状態で撮影された画像を用いることで高精細な自由視点画像の生成を行った。撮影の際、可動カメラの他に固定カメラを 2 台設置し、その 2 台を基底カメラとした射影グリッド空間を構築することにより、手間のかかるキャリブレーションを行うことなく各カメラを取り扱うことができる新規なシステムを提案した。

実験では、実画像を用いた Image-Based Rendering

により自由視点画像の作成を行い、実際に高精細なレンダリング画像が得られることを確認した。また、固定カメラのみによる自由視点画像と比較し、高精細なレンダリング画像が得られていることから本手法の有用性を示されたと言える。

## 参考文献

- [1] <http://www.ri.cmu.edu/events/sb35/tksuperbowl.html>
- [2] [http://whatisthematrix.warnerbros.com/cmp/sfx-bullet\\_text.html](http://whatisthematrix.warnerbros.com/cmp/sfx-bullet_text.html)
- [3] 北原格, 大田友一, 齋藤英雄, 秋道慎志, 尾野徹, 金出武雄, “大規模空間における多視点映像の撮影と自由視点映像生成”, 映像メディア学会誌 Vol.56, No.8, pp.120–125, 2002
- [4] M.Takemura, Y.Ohta, “Diminishing Head-Mounted Display for Shared Mixed Reality”, Proc. 1st International Symposium on Mixed and Augmented Reality 2002 (ISMAR2002), pp.149–156, 2002
- [5] H.Saito, T.Kanade, “Shape reconstruction in projective grid space from large number of images”, Computer Vision and Pattern Recognition (CVPR-99), pp.49–54, 1999
- [6] M.A.Fischler, R.C.Bolles, “Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography”, CACM, 24(6), pp.381–395, 1981
- [7] M.Potmesil, “Generating Octree Models of 3D Objects from Their Silhouettes in a Sequence of Images”, Computer Vision, Graphics and Image Processing, Vol.40, pp.277–283, 1987
- [8] 源田大輔, 向川康博, 尺長健, “視点と法線を組み合わせた任意視点映像の生成法”, 情報処理学会研究報告 CVIM 2003-137-7, pp.53–60, 2003
- [9] William E. Lorensen, Harvey E. Cline, “Marching Cubes: a high resolution 3D surface construction algorithm”, Computer Graphics(Proceedings of SIGGRAPH '87), Vol. 21, No. 4, pp.163–169, 1987
- [10] T.Nagae, T.Agui, H.Nagahashi, “Object surface construction from volume data with appropriate topology”, EICE D-II, Vol. J76-D-II, No. 8, pp.1704–1711, 1993
- [11] 矢口悟志, 齋藤英雄, “Projective Grid Space における多視点 Silhouette 画像からの自由視点画像生成”, 電子情報通信学会技術報告 PRMU2000-26, pp.23–28, 2000
- [12] T.Tuytelaars, L.Van Gool, “Wide baseline stereo matching based on local, affinity invariant regions”, In Proc. BMVC, pages 412–425, 2000