

拡張多視点幾何に基づく任意視点・任意時刻映像生成

中村 直行 坂上 文彦 佐藤 淳

名古屋工業大学 情報工学専攻

〒466-8555 名古屋市昭和区御器所町

`nakamura@hilbert.elcom.nitech.ac.jp`, `{Sakaue,JunSato}@nitech.ac.jp`

あらまし: 近年、4次元以上の高次元空間における多視点幾何の性質が明らかにされつつある。高次元空間における多視点幾何では、従来の多視点幾何において不可能であったカメラと対象物が共に運動する動的なシーンにおける幾何学的な関係を扱うことが可能となる。本研究ではこの新たな多視点幾何理論を用いることで、空間中の運動点を最低1点画像中において追跡するのみで、任意の並進運動カメラ間の位置、姿勢、相対速度と、運動する点の情報が4次元時空間において復元でき、この結果、任意時刻における空間情報を任意視点カメラに投影した画像が仮想生成できることを示す。提案法を用いることで、空中に描画された図形や軌跡を、運動するユーザの視点位置に応じて適切に提示する空間描画システムが、最低1点の特徴点追跡のみで実現可能となる。

Generating Virtual Images Viewed from Arbitrary Viewpoint at Arbitrary Time by using Multiple View Geometry in Space-Time

Naoyuki NAKAMURA and Fumihiko SAKAUE and Jun SATO

Department of Computer Science and Engineering, Nagoya Institute of Technology
Gokiso-cho Showa-ku Nagoya 466-8555, Japan

`nakamura@hilbert.elcom.nitech.ac.jp`, `{Sakaue,JunSato}@nitech.ac.jp`

Abstract: Recently, multiple view geometry in higher dimensional space has been proposed and studied. In this multiple view geometry, we can derive geometric relations under situations in which multiple cameras and objects move. In this paper, we propose a method for reconstructing the 3D velocity, position, and orientation of cameras, as well as the 3D position of moving points, by using the multiple view geometry in higher dimensional space. We also show that we can generate virtual images of moving objects at arbitrary time, which are viewed from arbitrary viewpoints. By using the proposed method, we can construct “Virtual 3D Drawing System” just from single point tracking.

1 はじめに

近年、4次元以上の高次元空間における複数のカメラ間の幾何を記述した拡張多視点幾何が提案された [2, 4, 6, 8, 9]. この拡張多視点幾何は、従来の多視点幾何 [1, 7] では不可能であった空間中の点

とカメラが共に運動するような動的なシーンにおける幾何を扱うことができるなど、多くの可能性を秘めている。しかしこの拡張多視点幾何には未解明な点が多く、これを用いた復元理論なども未だに整備されていない。このために、現状では拡張多視点幾何理論を実際にアプリケーション等に

応用することが難しい。そこで本研究では、4次元空間から2次元画像への投影に基づく多視点幾何を例に、拡張多視点幾何において未解明となっていたユークリッド復元を行うための手法を提案する。さらに、この復元法を応用することで、任意の時刻における空間情報を任意視点のカメラに投影した画像が仮想生成できることを示す。これにより、空中に描画された図形や軌跡を、運動するユーザの視点位置に応じて適切に提示する空間描画システム [5] などが構築可能となる。

2 拡張多視点幾何

2.1 4次元空間における多視点幾何

本節では、従来の多視点幾何を高次元に拡張した拡張多視点幾何に関して、4次元空間から2次元画像への投影における多視点幾何を例に説明する。まず、4次元空間から2次元画像への投影について説明する。今、4次元空間中の点 $\mathbf{X} = [X_1, X_2, X_3, X_4, X_5]^T$ が射影カメラに $\mathbf{x} = [x_1, x_2, x_3]^T$ として投影されているとする。するとこの投影は 3×5 の投影行列 \mathbf{P} を用いて以下のように表すことができる。

$$\lambda \mathbf{x} = \mathbf{P} \mathbf{X} \quad (1)$$

ここで λ は定数倍の不定性を表すスカラーである。

次に、このような4次元から2次元への投影の元で成り立つ多視点幾何について説明する。従来の多視点幾何とは異なり、拡張多視点幾何では3~5視点の幾何拘束が存在する [4, 8]。これらのうち、本稿では第3節、第4節において3視点幾何を用いるため、ここでは3視点幾何拘束について詳しく述べることにする。4次元から2次元への投影に基づく多視点幾何において、拡張 trilinear 拘束は以下の式によって表される。

$$x^i x^j x^k \epsilon_{qjv} T_{ir}^j = 0_v \quad (2)$$

この式における全ての添え字はそれぞれ1から3までの値を取る。 x^i, x^j, x^k はそれぞれ第1~第3カメラ画像中における対応点である。また ϵ_{qjv} は $\{q, j, v\}$ から $\{1, 2, 3\}$ への置換が偶置換であれば1, 奇置換であれば-1, それ以外であれば0の値を取るテンソルである。 T_{ir}^j は4次元空間にお

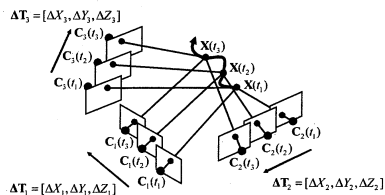


図 1: 並進運動カメラにおける多視点幾何

ける拡張 Trifocal tensor であり、 $3 \times 3 \times 3$ の3階テンソルで27要素をもつ。この T_{ir}^j は、従来の3次元空間における Trifocal tensor T_i^{rj} とは全く異なるものであり、カメラ間の位置や姿勢の情報だけでなくカメラの運動情報も含まれている。式(2)に示す拡張 trilinear 拘束では1組の対応点から3本の式が得られるが、そのうち線形独立な式は2本のみである。したがって、画像中から対応点が最低13点得られれば、拡張 Trifocal tensor T_{ir}^j を線形に求めることが可能となる。

2.2 並進運動カメラにおける多視点幾何

本節では、2.1節において説明した4次元空間における拡張多視点幾何を用いた例として、等速並進運動を行う複数台のカメラと、空間中で任意の運動を行う点との間に成り立つ幾何について説明する [2, 8]。この並進運動カメラによる拡張多視点幾何は、本稿第4節、第5節において用いる。

今、図1のように任意の運動を行う点 \mathbf{X} と等速並進運動する複数台のカメラが存在し、点 \mathbf{X} がこれらのカメラに毎時刻投影されているとする。ここで、カメラの X, Y, Z 軸方向の並進運動の速度を $[\Delta X, \Delta Y, \Delta Z]$ とする。この並進運動はそれぞれのカメラ毎に異なっていてよい。また時刻 t における点 \mathbf{X} の3次元座標を $\mathbf{X}_{3D}(t) = [X(t), Y(t), Z(t), 1]^T$ とする。この点 $\mathbf{X}_{3D}(t)$ の投影点を $\mathbf{x} = [x(t), y(t), 1]^T$ とすると、3次元点 $\mathbf{X}_{3D}(t)$ の並進運動カメラへの投影は以下のように表すことができる。

$$\lambda \begin{bmatrix} x(t) \\ y(t) \\ 1 \end{bmatrix} = \begin{bmatrix} P_{11} & P_{12} & P_{13} & P_{14} \\ P_{21} & P_{22} & P_{23} & P_{24} \\ P_{31} & P_{32} & P_{33} & P_{34} \end{bmatrix} \begin{bmatrix} X(t) - t\Delta X \\ Y(t) - t\Delta Y \\ Z(t) - t\Delta Z \\ 1 \end{bmatrix} \quad (3)$$

ここで、時刻 t における3次元点 \mathbf{X} を、時間要素 t を加えた4次元点 $\mathbf{X}_{4D}(t) = [X(t), Y(t), Z(t), t, 1]^T$ として表現する。すると

$\mathbf{X}_{4D}(t)$ の並進運動カメラへの投影は、式 (3) を変形して以下のように表すことができる。

$$\lambda \begin{bmatrix} x(t) \\ y(t) \\ z(t) \\ 1 \end{bmatrix} = \begin{bmatrix} p_{11} & p_{12} & p_{13} & -p_{11}\Delta X - p_{12}\Delta Y - p_{13}\Delta Z & p_{14} \\ p_{21} & p_{22} & p_{23} & -p_{21}\Delta X - p_{22}\Delta Y - p_{23}\Delta Z & p_{24} \\ p_{31} & p_{32} & p_{33} & -p_{31}\Delta X - p_{32}\Delta Y - p_{33}\Delta Z & p_{34} \end{bmatrix} \begin{bmatrix} x(t) \\ y(t) \\ z(t) \\ t \end{bmatrix} \quad (4)$$

ここで、各カメラの運動は等速並進運動であるため $\Delta X, \Delta Y, \Delta Z$ は定数である。したがって、式 (4) 右辺の 3×5 行列の要素は全て定数となる。よって、時間要素 t を加えた 4 次元空間から 2 次元画像への投影を考えたとき、等速並進運動をするカメラは 4 次元空間における静止カメラとみなすことができ、このようなカメラ複数台と運動点との間に成り立つ幾何学的な関係を、2.1 節で述べた 4 次元空間における多視点幾何によって表現することが可能となる。さらに、このような拡張された多視点幾何における multifocal tensor を計算することで、並進運動するカメラの校正、即ちカメラの位置・姿勢と運動を求めることができ、さらに運動点 \mathbf{X} も復元することが可能となる。

3 4次元空間におけるユークリッド復元

本節では 2 節で述べた 4 次元空間における拡張 Trifocal tensor を元に、4 次元空間をユークリッド復元する方法を示す。尚、本稿では各カメラの内部パラメータは既知であるとする。

従来の 3 次元空間における多視点幾何では、2 台のカメラ画像から計算した Bifocal tensor (Fundamental 行列) を元にカメラ行列をスケール付 3 次元ユークリッド変換の不定性を残して校正することで、ユークリッド復元を行うことができた [7]。ところが 4 次元空間における多視点幾何の場合、2 視点幾何が存在しないため Bifocal tensor を定義することができず、従来と同様の方法でユークリッド復元を行うことはできない。そこで本節では、Bifocal tensor の導出を経由することなく 4 次元空間においてユークリッド復元を行う方法を提案する。

本稿で提案する 4 次元空間でのユークリッド復元法では、まず画像からカメラ行列を射影不定性を含む状態で校正する。次にこのカメラ行列に 4 次元射影変換を施すことで、スケール付ユークリッド変換の不定性をもつカメラ行列を求める。最後に得られたカメラ行列を用いてユークリッド復元を行う。

3.1 4次元射影復元

本節では、4 次元空間の拡張 Trifocal tensor を元に射影不定性をもつカメラ行列を求める方法について説明する [3, 9]。今、空間内に 3 台のカメラが存在するとし、これらのうち 1 台のカメラ行列を座標系の基準として $\mathbf{P}_{1p} = [\mathbf{I}|\mathbf{0}|\mathbf{0}]$ とおく。このとき、これに対する第 2, 第 3 カメラ行列は以下のように定義することができる。

$$\mathbf{P}_{2p} = [\mathbf{H}_{12}|\mathbf{u}_2|\mathbf{u}'_2] \quad (5)$$

$$\mathbf{P}_{3p} = [\mathbf{H}_{13}|\mathbf{u}_3|\mathbf{u}'_3] \quad (6)$$

これらの式において、 $\mathbf{H}_{1n}(n=2,3)$ は第 1 カメラ画像から第 n カメラ画像への平面射影変換行列を表す。また $\mathbf{u}_n, \mathbf{u}'_n$ は、第 n カメラ画像における第 1 カメラ視点の投影像であるエピポール \mathbf{e}_{n1} 上の任意の 2 点である (4 次元から 2 次元への投影におけるエピポールは点ではなく直線であることに注意)。

\mathbf{H}_{12} は拡張 trilinear 拘束式 (2) を元に導出した以下の式より求められる。

$$H_{12}^j_i = p''^r T_{ir}^j \quad (7)$$

ここで p''^r は第 3 カメラ画像中の任意の点を表す。またエピポール \mathbf{e}_{21} は以下の式より求められる。

$$(\mathbf{H}_{12}^T + \lambda \mathbf{H}'_{12}{}^T) \mathbf{e}_{21} = \mathbf{0} \quad (8)$$

ここで \mathbf{H}'_{12} は、式 (7) と同様の方法で、 p''^r とは異なる第 3 カメラ画像中の点より求めた平面射影変換行列である。また λ は 0 でないスカラーであり、行列 \mathbf{H}'_{12} についての行列 $-\mathbf{H}'_{12}{}^T$ の一般化固有値として求められる。こうして得られた式 (8) より \mathbf{e}_{21} が求められる。しかし、式 (8) における 3×3 行列 $(\mathbf{H}_{12}^T + \lambda \mathbf{H}'_{12}{}^T)$ の 3 行は従属しており、その rank は 1 である。 \mathbf{e}_{21} を計算するためには rank が 2 以上必要であるため、第 3 カメラ画像中からさらに他の点を 2 点選択して別の平面射影変換行列 $\mathbf{H}''_{12}, \mathbf{H}'''_{12}$ を求め、式 (8) と同様の式をもう 1 組用意することで、エピポール \mathbf{e}_{21} を導出することが可能となる。

また、式 (6) に示す第 3 カメラ行列は、拡張 trilinear 拘束式 (2) において第 2 カメラと第 3 カメラを入れ換えた以下の拡張 trilinear 拘束式を元に平面射影変換行列とエピポールを導出して校正する。

$$x^i x'^j x''^k \epsilon_{krv} T_{ij}^r = 0_v \quad (9)$$

ただし、以上の方法で第2, 第3カメラ行列をそれぞれ求めると、異なる拘束式(2), (9)を元に別の拡張 Trifocal tensor から導出を行っているため、第1, 第2カメラより成る射影空間と、第1, 第3カメラより成る射影空間が異なるものになる。このように属する射影空間が異なるカメラ行列からでは、3.2節にて説明するスケール付きユークリッド変換の不定性をもつカメラ行列の導出を行うことができない。そのため一方のカメラ行列をもう一方のカメラが属する射影空間に4次元射影変換することで、空間を合わせておく必要がある。これは、 \mathbf{P}_{1p} と \mathbf{P}_{2p} を用いて画像より復元した4次元点 \mathbf{X} と、 \mathbf{P}_{1p} と \mathbf{P}_{3p} を用いて復元した4次元点 \mathbf{X}' から4次元射影変換行列を求めて、これを一方のカメラ行列にかけることで実現できる。

3.2 4次元ユークリッド復元

本節では各カメラの内部パラメータを既知とした条件の元で、3.1節において導出した射影不定性をもつカメラ行列をスケール付ユークリッド変換の不定性をもつカメラ行列へと変換する方法について説明する。ここで4次元空間におけるスケール付ユークリッド変換には、カメラのワールド座標に対する絶対位置及び姿勢の不定性と系全体のスケールの不定性が含まれている。したがって、スケール付ユークリッド変換の不定性を含むカメラ行列を $\mathbf{P}_{1e}, \mathbf{P}_{2e}, \mathbf{P}_{3e}$ とすると、これらは第1カメラ \mathbf{P}_{1e} を基準として以下のように表される。

$$\begin{aligned} \mathbf{P}_{1e} &= [\mathbf{I}|\mathbf{0}|\mathbf{0}] \\ \mathbf{P}_{2e} &= [\mathbf{R}_{12}|\alpha\Delta\mathbf{T}_{12}|\alpha\mathbf{T}_{12}] \\ \mathbf{P}_{3e} &= [\mathbf{R}_{13}|\alpha\Delta\mathbf{T}_{13}|\alpha\mathbf{T}_{13}] \end{aligned} \quad (10)$$

ここで、 $\mathbf{R}_{1n}, \Delta\mathbf{T}_{1n}, \mathbf{T}_{1n}(n=2,3)$ はそれぞれ、第1カメラのカメラ座標を基準としたときの第nカメラのもつ回転行列、速度ベクトル、並進ベクトルを表す。 α は系全体のスケールの不定性を表すスカラーである。本節では、 $\mathbf{P}_{1p}, \mathbf{P}_{2p}, \mathbf{P}_{3p}$ を同一の4次元射影変換行列 $\mathbf{H}_{5\times 5}$ を用いて変換し、式(10)に示すカメラ行列を求めることを目的とする。ここで、式(10)に示すカメラ行列 $\mathbf{P}_{ne}(n=1,2,3)$ の左3列を計算する行列を \mathbf{H}_O 、残りの2列を計算する行列を \mathbf{H}_T とすると、 $\mathbf{H}_{5\times 5}$ は以下のように表せる。

$$\mathbf{H}_{5\times 5} = \mathbf{H}_O\mathbf{H}_T \quad (11)$$

以降では \mathbf{H}_O と \mathbf{H}_T の導出法について説明する。

3.2.1 正規直交化による直交空間への変換

本節では、式(5), (6)における \mathbf{H}_{1n} を式(10)における \mathbf{R}_{1n} に変換する行列 \mathbf{H}_O の導出法について説明する。ここで \mathbf{R}_{1n} は基準の座標系に対するカメラの回転行列であるため、正規直交行列である。したがってこれを拘束とし、 \mathbf{H}_{1n} を正規直交化することで \mathbf{R}_{1n} へと変換する。ところが、Gram-Schmidtの正規直交化法などを用いてそれぞれの \mathbf{H}_{1n} を変換すると、変換後のカメラ行列は異なる空間におけるカメラ行列となってしまう。そこで、3台のカメラ行列を同一の行列 \mathbf{H}_O を用いて変換することで、カメラ3台を同一空間のまままで正規直交変換する方法を示す。まず3台のカメラ行列 $\mathbf{P}_{1p}, \mathbf{P}_{2p}, \mathbf{P}_{3p}$ に右から 5×5 の4次元射影変換行列 \mathbf{H}_O をかけて新たなカメラ行列を得る。このカメラ行列においてそれぞれの左から3列を抜き出した 3×3 の部分行列を $\mathbf{Q}_1, \mathbf{Q}_2, \mathbf{Q}_3$ とする。これらの行列には \mathbf{H}_O の25要素が未知数として含まれているが、これら3つの行列が正規直交行列となることを元に得られる拘束式を解くことで \mathbf{H}_O の25要素を求める。ただしカメラ行列には定数倍の不定性があるため、 $\mathbf{P}_{1p}, \mathbf{P}_{2p}, \mathbf{P}_{3p}$ を共通の変換行列 \mathbf{H}_O で変換したところで、 $\mathbf{Q}_1, \mathbf{Q}_2, \mathbf{Q}_3$ の各行ベクトルのノルムが3台のカメラで同一の値になるとは限らない。そこで、3台のカメラ行列毎に固有のスカラー s_1, s_2, s_3 を用いて $\mathbf{Q}_1, \mathbf{Q}_2, \mathbf{Q}_3$ の各行ベクトルのノルムが1となるようにする。したがって、 \mathbf{H}_O と s_1, s_2, s_3 を用いた $\mathbf{P}_{1p}, \mathbf{P}_{2p}, \mathbf{P}_{3p}$ の変換は以下の式で表される。ここで、この変換によって得られる3台のカメラ行列をそれぞれ $\mathbf{P}_{1o}, \mathbf{P}_{2o}, \mathbf{P}_{3o}$ とする。

$$\begin{aligned} \mathbf{P}_{1o} &= s_1\mathbf{P}_{1p}\mathbf{H}_O = [s_1\mathbf{Q}_1|\mathbf{0}|\mathbf{0}] \\ \mathbf{P}_{2o} &= s_2\mathbf{P}_{2p}\mathbf{H}_O = [s_2\mathbf{Q}_2|s_2\mathbf{u}_2|s_2\mathbf{u}'_2] \\ \mathbf{P}_{3o} &= s_3\mathbf{P}_{3p}\mathbf{H}_O = [s_3\mathbf{Q}_3|s_3\mathbf{u}_3|s_3\mathbf{u}'_3] \end{aligned} \quad (12)$$

ここで、 $\mathbf{P}_{1o}, \mathbf{P}_{2o}, \mathbf{P}_{3o}$ の左3列から成る部分行列は、式(10)に示す $\mathbf{P}_{1e}, \mathbf{P}_{2e}, \mathbf{P}_{3e}$ の左3列の部分行列と等しい。即ち、以下の拘束式が成り立つ。

$$s_1\mathbf{Q}_1 = \mathbf{I} \quad (13)$$

$$s_2\mathbf{Q}_2 = \mathbf{R}_{12} \quad (14)$$

$$s_3\mathbf{Q}_3 = \mathbf{R}_{13} \quad (15)$$

これらの式を用いて \mathbf{H}_O の 25 要素と s_1, s_2, s_3 を求める。ここで、式 (13) を満たすこと、3 台のカメラ行列の 4 列目・5 列目に変換を与えないことという 2 つの条件より $s_1 = 1$ となり、また \mathbf{H}_O の 25 要素のうち 19 要素が以下のように決定される。

$$\mathbf{H}_O = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ h_{41} & h_{42} & h_{43} & 1 & 0 \\ h_{51} & h_{52} & h_{53} & 0 & 1 \end{bmatrix} \quad (16)$$

これにより、残る未知数は \mathbf{H}_O の残りの 6 要素と s_2, s_3 の合計 8 つとなる。次にこの 8 変数を、拘束式 (14), (15) を用いて求める。式 (16) に示す \mathbf{H}_O を式 (12) に代入して得られた $s_2 \mathbf{Q}_2, s_3 \mathbf{Q}_3$ において、各行列内での異なる行との内積が 0 の式 (直交性の拘束式) を 6 本と、各行列の行ベクトルのノルムが 1 の式 (正規性の拘束式) を 6 本得る。こうして得られた 12 本の拘束式を 8 個の未知数に関して解くことで、 s_2, s_3 と \mathbf{H}_O の要素が全て求められる。そしてこれらの値を用いて式 (12) を計算することで、以下に示す 3 つのカメラ行列 $\mathbf{P}_{1o}, \mathbf{P}_{2o}, \mathbf{P}_{3o}$ を得ることができる。

$$\begin{aligned} \mathbf{P}_{1o} &= [\mathbf{I}|0|0] \\ \mathbf{P}_{2o} &= [\mathbf{R}_{12}|s_2 \mathbf{u}_2|s_2 \mathbf{u}'_2] \\ \mathbf{P}_{3o} &= [\mathbf{R}_{13}|s_3 \mathbf{u}_3|s_3 \mathbf{u}'_3] \end{aligned} \quad (17)$$

3.2.2 時間要素の拘束を用いたユークリッド空間への変換

次に、式 (17) のように求めた $\mathbf{P}_{1o}, \mathbf{P}_{2o}, \mathbf{P}_{3o}$ を、式 (10) に示すスケール付ユークリッド変換の不定性をもった状態のカメラ行列 $\mathbf{P}_{1e}, \mathbf{P}_{2e}, \mathbf{P}_{3e}$ に変換する行列 \mathbf{H}_T の導出法について説明する。ここで \mathbf{H}_T は、 $\mathbf{P}_{1o}, \mathbf{P}_{2o}, \mathbf{P}_{3o}$ の左 3 列を変換しないこと、また第 1 カメラに関しては $\mathbf{P}_{1o} = [\mathbf{I}|0|0]$ を $\mathbf{P}_{1e} = [\mathbf{I}|0|0]$ のまま変換しないことという 2 つの条件を満たさなければならないため、その 25 要素のうち 21 要素が以下のように決定される。

$$\mathbf{H}_T = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & h_{44} & h_{45} \\ 0 & 0 & 0 & h_{54} & h_{55} \end{bmatrix} \quad (18)$$

ここで、 \mathbf{H}_T の右下の 4 要素はカメラ行列の 4 列目、5 列目に関する 1 次元射影変換を表し、この

4 要素の間で定数倍の不定性をもつ。これにより 4 要素のうち独立な未知数は 3 つとなるため、このうち 1 つの要素を定数として残りの 3 要素を求めることとする。ここで、この 3 要素を求める拘束式を得るために、4 次元ユークリッド復元点のもつ不定性について考える。4 次元空間におけるスケール付ユークリッド変換の不定性をもつカメラ行列を式 (10) とした場合、等速並進運動する 3 台のカメラの各時刻毎における 3 次元位置及び姿勢の相対的な関係は、スケール付ユークリッド変換において不変である。これは、この 3 台のカメラが属する空間を 4 次元空間として見たとき、時間軸方向に関しては不定性が含まれないことを意味する。したがって、4 次元ユークリッド復元された点はその時間要素に関してのみ不定性をもたない。即ちこの時間要素は、カメラを運動させながら対象物を投影した際のカメラ画像のフレーム番号と一致する。よってこれを拘束条件として \mathbf{H}_T の残りの 3 要素を求める。まず式 (17) のように求めた $\mathbf{P}_{1o}, \mathbf{P}_{2o}, \mathbf{P}_{3o}$ のうち 2 台のカメラ行列を用いて、画像点より 4 次元点 \mathbf{X}_o を復元する。ここで、4 次元ユークリッド復元点 \mathbf{X}_e と \mathbf{X}_o の間の関係は $\mathbf{X}_e = \mathbf{H}_T^{-1} \mathbf{X}_o$ のように表される。このとき 4 次元ユークリッド復元点の第 4 要素はカメラ画像のフレーム番号と一致することより、以下の式を得る。

$$t = \{\mathbf{H}_T^{-1} \mathbf{X}_o\}[[4]] \quad (19)$$

この式において左辺の t は、 \mathbf{X}_o を復元するとき用いたカメラ画像中の投影点が撮影されたときのカメラ画像のフレーム番号である。また右辺は、 \mathbf{X}_o に左から \mathbf{H}_T^{-1} をかけて得られた座標の第 4 要素を表す。今、未知数の数は \mathbf{H}_T のもつ 3 つであるから、式 (19) が 3 本以上、即ち 3 フレーム分の式が得られていれば \mathbf{H}_T を求めることができる。この \mathbf{H}_T を用いることで、スケール付 4 次元ユークリッド変換の不定性をもった 3 台のカメラ行列 $\mathbf{P}_{1e}, \mathbf{P}_{2e}, \mathbf{P}_{3e}$ が求められる。

このようにして、得られたカメラ行列を元に画像点を復元することで、4 次元空間におけるユークリッド復元を行うことができる。

4 拡張多視点幾何を用いた任意時刻映像生成

本節では、3節において説明した拡張多視点幾何における復元法をアプリケーションへ応用した例として、空間描画システムを提案する。まず提案する空間描画システムについて説明する。今、カメラ付HMDを装着したユーザが空間内に複数存在しているとする。ここで、ユーザが指などによって空中に文字や図形などを描画し、この描画を行う指先を、それぞれのユーザが装着したカメラ画像中においてトラッキングする。そして現在時刻において、その時刻までに得られている指先の投影点全てを現在時刻の視点位置・姿勢のカメラに投影した映像を仮想生成する。こうして得られた映像を各ユーザのHMD上に表示することで、ユーザに空間描画を体感させることができる。

このとき、ユーザがシステム動作中静止しており、カメラの視点位置・姿勢が変化しないのであれば、単純に過去の投影点全てをHMD上に表示するだけで空間描画システムを実現できる。しかし、ユーザが運動しカメラの視点位置・姿勢が毎時刻変化する場合、ユーザの運動によって、過去に描かれた描画軌跡の形や大きさなどの見え方は変化するため、現在のカメラに過去の投影点をそのまま投影しても正しい映像を得ることはできない。

そこで、拡張多視点幾何を用いてこの問題を解決する。対象の空間に時間の要素を加えて4次元時空間における多視点幾何の理論を用いることで、任意の時刻・任意の視点位置・姿勢のカメラにおいて適切な映像を生成することができ、空間描画システムの構築が可能となる。

4.1 任意時刻映像生成の基礎理論

本節では、拡張多視点幾何における復元法を用いて任意時刻・任意視点映像を生成するための基礎となる理論を説明する。まず、全てのカメラの運動はそれぞれ任意の等速並進運動であると仮定する。そして描画を行う指先を特徴点としてこれらのカメラ画像中でトラッキングし、これらのカメラのうち2台を用いてこの特徴点を毎時刻4次元空間中において復元する。このとき復元を行う2台のカメラの内部パラメータ・外部パラメータがともに既知であれば、この復元は古典的復元となる。ある時刻にお

けるカメラ画像のフレーム番号を t_1 とし、このときの古典的復元された点を $\mathbf{X}(t_1) = [X_1, Y_1, Z_1, t_1, 1]^T$ とする。次に、別の時刻におけるカメラ画像のフレーム番号を t_2 とし、この時刻 t_2 における視点位置・姿勢のカメラ画像に、過去の時刻 t_1 において復元された点 $\mathbf{X}(t_1)$ を正しく投影する方法について考える。ここで、4次元復元点 $\mathbf{X}(t_1)$ の第4要素 t_1 は、その点を復元したときまでにカメラが初期位置から何フレーム分等速並進運動したかを表す値となっている。したがって、この $\mathbf{X}(t_1)$ の第4要素 t_1 を、投影したい時刻のフレーム番号 t_2 に置き換えた点 $\mathbf{X}(t_2) = [X_1, Y_1, Z_1, t_2, 1]^T$ を作成し、この点をカメラ行列を用いて投影することで、時刻 t_1 に撮影された点を時刻 t_2 のカメラ視点位置から見た映像として正しく生成することが可能となる。

4.2 ユークリッド復元を用いた任意時刻映像生成

4.1節において説明した方法は、4次元空間中の点を古典的復元した場合のものである。この場合、古典的復元を行うために各カメラの内部パラメータ・外部パラメータの情報が必要となる。内部パラメータは事前に各カメラを校正しておくことで得ることができるが、外部パラメータに関してはユーザの位置・姿勢や運動速度の情報が必要となるため、これを得ることは難しい。そこで本節では、内部パラメータが既知、外部パラメータが未知であり、復元が4次元ユークリッド復元となる場合について、任意時刻映像の生成が可能であることを示す。今、3台のカメラ画像から計算した拡張Trifocal tensorを元に3.2節において説明した方法でカメラ行列を式(10)のように求め、これを用いて4次元ユークリッド復元した点を $\mathbf{X}_e(t_1)$ とする。このとき $\mathbf{X}_e(t_1)$ は、この点の真の座標 $\mathbf{X}(t_1)$ に対してスケール付4次元ユークリッド変換行列 \mathbf{W}_e をかけたベクトルとなる。またこの \mathbf{W}_e は、ワールド座標系に対する第1カメラの回転 \mathbf{R}_1 、速度 $\Delta\mathbf{T}_1$ 、並進 \mathbf{T}_1 、スケール α の不定性より成る。これにより、 $\mathbf{X}_e(t_1)$ と $\mathbf{X}(t_1)$ の関係は以下の式によ

て表される.

$$\begin{aligned}
 \mathbf{X}_e(t_1) &= \mathbf{W}_e \mathbf{X}(t_1) \\
 &= \begin{bmatrix} \alpha \mathbf{R}_1 & \alpha \Delta \mathbf{T}_1 & \alpha \mathbf{T}_1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ t_1 \end{bmatrix} \\
 &= \begin{bmatrix} \alpha r_{11} & \alpha r_{12} & \alpha r_{13} & \alpha \Delta T_1^1 & \alpha T_1^1 \\ \alpha r_{21} & \alpha r_{22} & \alpha r_{23} & \alpha \Delta T_1^2 & \alpha T_1^2 \\ \alpha r_{31} & \alpha r_{32} & \alpha r_{33} & \alpha \Delta T_1^3 & \alpha T_1^3 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ t_1 \\ 1 \end{bmatrix} \\
 &= \begin{bmatrix} \alpha r_{11} X + \alpha r_{12} Y + \alpha r_{13} Z + \alpha \Delta T_1^1 t_1 + \alpha T_1^1 \\ \alpha r_{21} X + \alpha r_{22} Y + \alpha r_{23} Z + \alpha \Delta T_1^2 t_1 + \alpha T_1^2 \\ \alpha r_{31} X + \alpha r_{32} Y + \alpha r_{33} Z + \alpha \Delta T_1^3 t_1 + \alpha T_1^3 \\ t_1 \\ t_1 \end{bmatrix} \quad (20)
 \end{aligned}$$

この式から分かるように, 変換行列 \mathbf{W}_e によって $\mathbf{X}(t_1)$ の第 4 要素 t_1 は, 速度ベクトル $\Delta \mathbf{T}_1$ の要素とスケール α との積として, 4 次元ベクトルの上から 3 行にも含まれるようになる. このとき基準カメラが固定カメラであるとする, $\Delta \mathbf{T}_1 = \mathbf{0}$ となり, $\mathbf{X}_e(t_1)$ に含まれる t_1 は第 4 要素のみとなるため, 第 4 要素の t_1 を任意のフレーム番号 t_2 に置き換えることで任意時刻の 4 次元点に変換できる. 以上のようにフレーム番号を置き換えた 4 次元点をカメラ行列を用いて投影することで, 任意時刻・任意視点映像の生成が可能となる.

5 任意時刻映像生成実験

本節では, 提案した拡張多視点幾何における 4 次元ユークリッド復元を用いて, 任意視点・任意時刻映像の生成が可能であることをシミュレーション実験によって示す. まず始めに, 提案法が理論的に正しいことを確認するために, 画像ノイズを考慮せずに実験を行った. 実験環境として, 図 2 に示すように空間中で点 \mathbf{X} が運動し, 3 次元的に軌跡を描く. そしてこの点 \mathbf{X} を, 図 2 のように等速並進運動を行う 3 台のカメラ画像中に毎時刻投影する. ただし第 4 節において説明したように 4 次元ユークリッド復元を用いて任意時刻映像生成を可能にするために, 第 1 カメラのみ固定カメラとした. またこれらのカメラの画角は, 並進運動を長時間続けても点 \mathbf{X} がフレームアウトしないように十分大きなものと仮定する. 図 2 において, 黒い曲線は点 \mathbf{X} による描画軌跡を表し, 青色, 緑色, 桃色の点はそれぞれ第 1, 第 2, 第 3 カメラの視点 $\mathbf{C}_1, \mathbf{C}_2, \mathbf{C}_3$ を表す. 次に 3 台のカメラから得られた画像より拡張 Trifocal tensor, 平面射影変換行列, エピボールを計算し, これらを元に第 2, 第 3 カメラ行列を式 (5), (6) のように校正する. そし

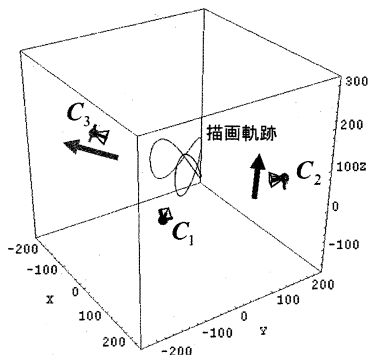


図 2: 実験環境

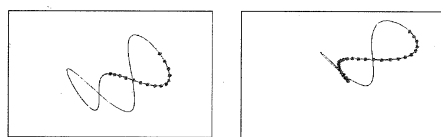
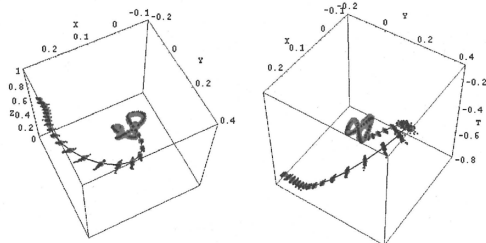


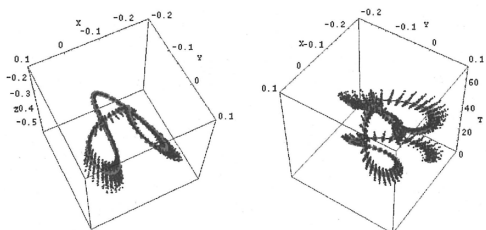
図 3: 任意時刻映像生成結果 (第 3 カメラ画像)

てこれらのカメラ行列をスケール付 4 次元ユークリッド変換の不定性をもつカメラ行列へと変換し, これによって画像点を 4 次元ユークリッド復元して任意時刻映像生成を行った. その結果のうち, 第 3 カメラ画像のものを 2 時刻分, 図 3 に示す. 図 3 における緑色の点は現在時刻の \mathbf{X} の投影点, オレンジ色の点は提案法によって過去の \mathbf{X} を現在の第 3 カメラに投影した点であり, 黒色の曲線はその真値である. 図より, 視点の変化に応じて過去の映像が正しく生成できており, 提案法による任意時刻映像生成が可能であることが確認できる.

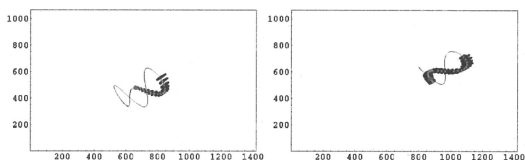
次に各カメラ画像から得られた点に標準偏差 1pixel のガウシアンノイズを印加して同様の計算をすることにより, 提案法の精度評価実験を行った. ところがこの計算過程において, 式 (8) を元にしたエピボールの導出が特にノイズの影響を受けて不安定なものとなった. そのため本実験では, エピボールのみ真値を用いて以降の任意時刻映像生成までを行った. この計算をノイズを変えて 100 回繰り返し行った結果のうち, 第 1, 第 3 カメラによる \mathbf{X} の 4 次元射影復元結果を図 4 に示す. また提案法により射影復元結果から計算した 4 次元ユークリッド復元結果を図 5 に示す. これらの復元結果は 4 次元空間の点となっているため, これを視覚的に分かりやすく表示するために, $[X, Y, Z, t]$



(a)[X, Y, Z] 空間 (b)[X, Y, t] 空間
図 4: 4 次元射影復元の精度評価



(a)[X, Y, Z] 空間 (b)[X, Y, t] 空間
図 5: 4 次元ユークリッド復元の精度評価



(a) 時刻 t_{15} (b) 時刻 t_{30}
図 6: 任意時刻映像生成の精度評価

の 4 要素から 3 要素ずつを切り出した結果を 3 次元表示した。また、第 3 カメラ画像の任意時刻映像生成結果の例を 2 時刻分、図 6 に示す。図 4、図 5 における青色の点は提案法による復元結果であり、赤色の点は真値である。青色の点は赤色の点周辺に分布しており、概ね正しく 4 次元射影復元できていることが分かる。また射影復元結果より計算した 4 次元ユークリッド復元の結果も概ね正しく計算できていることが分かる。また図 6 において、緑色の点は現在時刻の \mathbf{X} の投影点である。また提案法によって過去の \mathbf{X} を現在の視点位置の第 3 カメラに投影した結果を青色の点で示す。またこれに対する真値をオレンジ色の点で示す。図より、青色の点はオレンジ色の点の周辺に分布しており、概ね正しく映像生成が行えていることが分かる。ただし、この結果は真値のエピポールを用いて計算したものである。

以上の実験結果より、ノイズが存在する状況において精度良くエピポールを計算する方法が開発できれば、提案法により任意時刻映像生成が安定に行えることが分かる。

6 結論

本稿では、拡張多視点幾何におけるユークリッド復元法を提案した。またその応用例として、任意時刻映像生成法について述べた。今後の課題として、ノイズの影響下において精度良くエピポールを導出する計算法を開発する必要がある。この問題を解決することで、提案した任意時刻映像生成法を実環境において適用することが可能となる。

参考文献

- [1] Hartley, R. and Zisserman, A.: *Multiple View Geometry in Computer Vision*, Cambridge University Press (2000).
- [2] 早川和孝, 佐藤淳: 時空間における多視点幾何, 電子情報通信学会論文誌, Vol. J90-D, No. 8, pp. 1878-1887 (2007).
- [3] Kozuka, K. and Sato, J.: Rectification of 3D Data Obtained from Moving Range Sensors by using Multiple View Geometry, *Proc. International Conference on Image Analysis and Processing*, pp. 117-122 (2007).
- [4] Kozuka, K. and Sato, J.: Multiple View Geometry for Mixed Dimensional Cameras, *Proc. International Conference on Computer Vision Theory and Applications*, Vol. 1, pp. 5-12 (2008).
- [5] 中村直行, 坂上文彦, 佐藤淳: 多視点幾何に基づく 3 次元空間描画の実現, 画像の認識・理解シンポジウム, pp. 57-64 (2007).
- [6] 佐藤淳: テンソルと多視点幾何, CVIM チュートリアルシリーズ コンピュータビジョン最先端ガイド 1, Tokyo, 八木康史 and 斎藤英雄, pp. 33-42 (2008).
- [7] 佐藤淳: コンピュータビジョン—視覚の幾何学—, コロナ社 (1999).
- [8] Wan, C. and Sato, J.: Multiple View Geometry under Projective Projections in Space-Time, 画像の認識・理解シンポジウム, pp. 205-211 (2007).
- [9] Wolf, L. and Shashua, A.: On Projection Matrices $P^k \rightarrow P^2$, $k = 3, \dots, 6$, and their Applications in Computer Vision, *Proc. 8th International Conference on Computer Vision*, Vol. 1, pp. 412-419 (2001).