

未校正カメラと未校正プロジェクタによる 3次元復元と仮想楽器への応用

西 江 桂 亮[†] 佐 藤 淳[†]

近年、ビデオプロジェクタとカメラを用いたプロジェクタカメラシステムによりマンマシンインタラクションを実現する方法が研究されつつある。プロジェクタは単焦点カメラと見なすことができるため、近年明らかになった多視点幾何を応用できるという大きなメリットがある。本稿では、未校正のプロジェクタと未校正のカメラを射影的に校正する新たなプロジェクタカメラシステムの校正法を示す。特に、プロジェクタ光によってできる対象物の影をカメラで観測することにより、基準儀などを用いずにプロジェクタカメラシステムが射影的に校正できることを示す。さらに、射影的に校正されたプロジェクタカメラシステムを用いて3次元情報が復元でき、この結果仮想楽器の一種である仮想キーボード（仮想ピアノ）が実現可能であることを示す。

3D Reconstruction from Uncalibrated Cameras and Projectors and Its Application to Virtual Music Instruments

KEISUKE NISHIE[†] and JUN SATO[†]

Recently, projector camera systems have been studied extensively for designing man machine interaction systems. Since projectors can be modeled by using the projective projection, we can apply the multiple view geometry to the projector camera systems. In this paper, we propose a new method for calibrating uncalibrated projector camera systems. In particular, we show that it is possible to calibrate camera projector systems projectively and reconstruct 3D information reliably by using shadows of objects made by projector lights. We also show that by using the projectively calibrated projector camera systems we can achieve virtual music instruments, such as virtual keyboard systems.

1. はじめに

近年、カメラとビデオプロジェクタを用いたマンマシンインタラクション¹⁾や複合現実感^{2),3)}が研究されつつある。プロジェクタは通常平面のスクリーンに投影して用いるため、カメラ画像平面、プロジェクタ平面、スクリーン平面の3つの平面間の関係を用いてプロジェクタカメラシステムを構築することが可能である。すなわち、カメラ画像平面、プロジェクタ平面、スクリーン平面の3つの平面間の平面射影変換(homography)が求まればプロジェクタカメラシステムを構築することができる。このようなことから、3つの平面間の関係を求めるプロジェクタカメラシステムの校正法が研究されている。

Sukthankerら⁴⁾は、スクリーンが矩形であることを利用してカメラ、プロジェクタ、スクリーン間の平

面射影変換を求める方法を提案した。また、Okataniら⁵⁾はプロジェクタの内部パラメータが既知であると仮定することにより、スクリーンの矩形輪郭がカメラによって観測できなくてもカメラ、プロジェクタ、スクリーン間の平面射影変換が求まることを示した。これらの研究ではスクリーンに投影する画像を制御することを目的としているため、各平面間の射影変換が求まればよい。

一方、プロジェクタ画像を非平面スクリーンへ投影したり、プロジェクタカメラシステムにより3次元計測を行ったりする場合には、カメラとプロジェクタの3次元位置および姿勢を校正する必要がある。このような校正は、カメラとプロジェクタ各々の 3×4 の投影行列を求めることでもある。東城ら⁶⁾は3次元座標が既知な基準儀を用いてカメラとプロジェクタの各々の投影行列を計算する方法を示している。しかし、このような基準儀を用いる方法では、3次元座標が既知な基準儀を用意する必要がある。

そこで本稿では、このような既知の基準儀を用いず

[†] 名古屋工業大学大学院情報工学専攻
Department of Computer Science and Engineering,
Nagoya Institute of Technology

にカメラとプロジェクタを校正し投影行列を計算する方法を提案する．多視点幾何の研究より，複数のカメラを校正することは，これらのカメラ間の multifocal tensor を計算することに等しいことが示された⁷⁾．プロジェクタは単焦点カメラと同じ幾何構造を持っていることから，プロジェクタカメラシステムにおいては多視点幾何を拡張して応用することが可能である．しかしプロジェクタは投光器であるため，カメラとまったく同様に扱うことはできない．本稿ではプロジェクタが投光器であるという性質を積極的に利用し，プロジェクタによってできるスクリーン上の影の情報を用いることにより，プロジェクタカメラシステムが射影的に校正可能であることを示す．また，このようにして校正されたプロジェクタカメラシステムから空間中の点の 3 次元情報が復元できることを示す．物体の影を使って物体の 3 次元情報を復元する研究として，郭ら⁸⁾ の線状物体の 3 次元位置計算法がある．また，影情報を用いて物体の表面の 3 次元復元を行う手法として，Savarese ら^{10),11)} の Shadow Carving や，Yu ら⁹⁾ の Shadow Graphs などが研究されている．しかし，これらは物体の影を作り出す点光源の位置が 3 次元空間中で既知であり，かつカメラの内部パラメータも校正済みであることを前提としている．これに対して本研究では，光源であるプロジェクタとカメラがともに未校正であることを前提に，これらの関係を射影的に校正し射影復元を行う手法を提案するものであり，あらかじめプロジェクタとカメラを校正する必要がないことが特徴である．

さらに本稿では，未校正プロジェクタカメラシステムの射影的校正の一応用例として，仮想楽器をプロジェクタカメラシステムにより実現する方法を示す．仮想楽器¹²⁾とは複合現実感技術を用いて仮想的に作り出した楽器であり，ユーザにあたかも現実中存在する楽器を演奏しているかのような感覚を与えるものである．本研究では，プロジェクタカメラシステムを用いることにより，仮想キーボード（仮想ピアノ）を実現する．類似システムとして，コンピュータキーボードを投影し，これに対するキー入力を可能にした例¹⁵⁾などがあるが，これらは単純に 2 次元座標を求めて入力キーを判定しているのみであり，指の 3 次元位置情報を獲得していないため，本稿で示す仮想キーボードのように音の強弱を実現することはできず，またカメラと投影装置の位置関係や投影する平面との位置関係を既知としている点で大きく異なる．

次章以降では，まずよく知られたカメラとプロジェクタ間の平面射影変換の関係について述べる．次にカ

メラに投影される像とプロジェクタが作る影との間のステレオ視の関係について述べ，このステレオ視におけるエピポーラ幾何が自己エピポーラと呼ばれる強い拘束によって安定に求まることを示す．次に，このようにして求めたエピポーラ幾何からプロジェクタカメラシステムにより 3 次元復元が可能であることを示す．最後に，このようにして射影的に校正したプロジェクタカメラシステムにより，仮想キーボードが実現可能であることを示す．

2. カメラとプロジェクタ間の平面射影変換

今，図 1 に示すように未校正カメラと未校正プロジェクタと平面スクリーン Π を考え，プロジェクタの画像平面 π_p からスクリーン Π に投影された映像がカメラ画像平面 π_c に投影されるものとする．一般にプロジェクタは中心投影カメラと同一の幾何構造を持っているので， π_p と Π との間には平面射影変換 H_{ps} の関係が成り立つ．また Π と π_c との間にも平面射影変換 H_{sc} の関係が存在するため， π_c 上の点 x と π_p 上の点 x' との間には次式で示す平面射影変換 H_{pc} の関係が成り立つ．

$$\tilde{x} \sim H_{pc} \tilde{x}' \quad (1)$$

ここで， $(\sim)$ は斉次座標を表し， $(\sim)$ は定数倍の不定性を除いて等しいことを表す．また H_{pc} は以下に示すように H_{sc} と H_{ps} によって表される．

$$H_{pc} = H_{sc} H_{ps} \quad (2)$$

今，図 2 に示すようにプロジェクタ平面 π_p 上に N 個の点 x'_i ($i = 1, \dots, N$) を考え，これらの点がスクリーン上に投影された点をカメラに投影したものを，カメラ画像平面 π_c 上の点 x_i とする．すると，平面 π_p から平面 π_c への平面射影変換行列 H_{pc} は，よく知られているように最低 4 組以上の対応点 $\{x_i, x'_i\}$ から線形あるいは非線形に求めることができる⁷⁾．

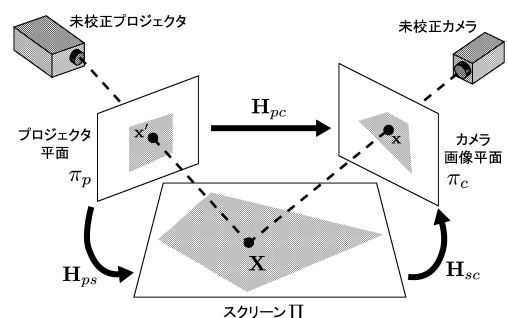


図 1 未校正カメラと未校正プロジェクタ
Fig. 1 Uncalibrated camera and projector.

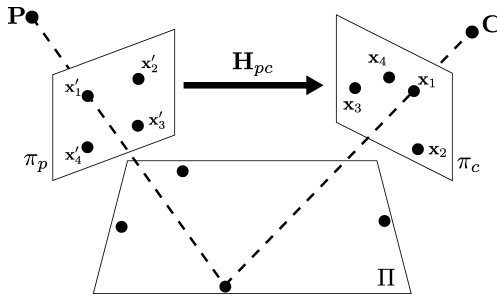


図 2 平面射影変換

Fig. 2 Planar projective transformation.

3. カメラとプロジェクタによるステレオ視

次に、カメラ画像とプロジェクタによる影の像との間に成り立つステレオ視の関係について考える。

図 3 に示すように 3 次元空間中のある点 X がカメラに点 x として投影されているとする。このとき、プロジェクタには点 X は投影されないで、カメラとプロジェクタの間には平面射影変換の関係以上の関係は何も分からないように見える。しかし、プロジェクタは一般に光を出す装置ともいえるので、その光学中心を点光源と見なすと、3 次元空間中の点 X は、プロジェクタの光学中心 P と点 X を結んだ直線がスクリーン Π と交差する点 S に影として投影される。この点 S がカメラに投影された点を s とすると、 s は平面射影変換以上の情報を我々に与える。

今、プロジェクタからカメラへの平面射影変換行列 H_{pc} が求まっているので、カメラ画像平面 π_c 上の点 s より、プロジェクタ平面 π_p 上の点 x' が以下のように求まる。

$$\tilde{x}' \sim H_{pc}^{-1} \tilde{s} \quad (3)$$

この点をスクリーンに投影すると図 3 から明らかに点 S に投影されることが分かる。しかし、プロジェクタをカメラとして考えると、点 x' は 3 次元空間中の点 X がプロジェクタ平面 π_p に投影されたものと見なすことができる。したがって、カメラとプロジェクタがともに 1 台ずつ存在する場合、3 次元空間中の点 X がカメラに投影された点 x と、プロジェクタが出す光によってスクリーン上に作られる影 S がカメラに投影された点 s との間にはステレオ視の関係があることが分かる。さらに、式 (3) より s から x' が求まることから、 x と x' の間のステレオ視を考えることもできる。したがってこれらはカメラ 2 台のステレオ視と同等に扱える。

また、 S' は C と X を結んだ直線がスクリーン Π と交差する点で、 s' は S' をプロジェクタに投影した

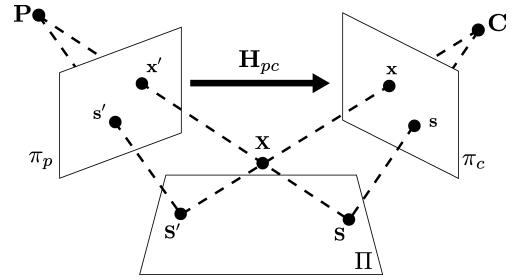


図 3 影を使ったステレオ視

Fig. 3 Stereo from shadows.

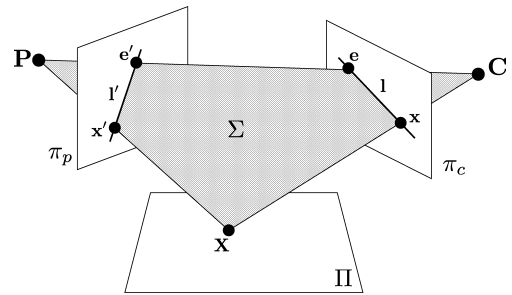


図 4 エピポーラ幾何

Fig. 4 Epipolar geometry.

像であり、以下の式より間接的に求まる。

$$\tilde{s}' \sim H_{pc}^{-1} \tilde{x} \quad (4)$$

4. カメラとプロジェクタ間のエピポーラ幾何

2 台のカメラにおける相対的なカメラの位置や姿勢の情報はエピポーラ幾何によって記述することができる^{7),13)}。これはカメラとプロジェクタが 1 台ずつ存在する場合においても同様に成り立つ。

今、図 4 のように空間中のある点 X がカメラ画像平面 π_c 上に点 x として投影され、また 3 章の方法によりプロジェクタ平面 π_p 上に x' として計算されているとする。すると、対応点 x と x' の間には次式で示すエピポーラ方程式が成り立つ。

$$\tilde{x}'^T F \tilde{x} = 0 \quad (5)$$

ここで F は fundamental 行列である。

また、 C, P を結ぶ直線が 2 つの画像平面 π_c, π_p と交わる点 e, e' はエピポールである。これらのエピポールは、 $F\tilde{e} = 0$ および $F^T\tilde{e}' = 0$ の関係を用いて F 行列より求まる。

カメラ 2 台の場合、3 次元空間中の点 X から得られる情報は 2 枚の画像平面に投影される点 x, x' のみである。しかし、プロジェクタカメラシステムでは、図 3 に示すように、カメラ 2 台の場合に比べてより

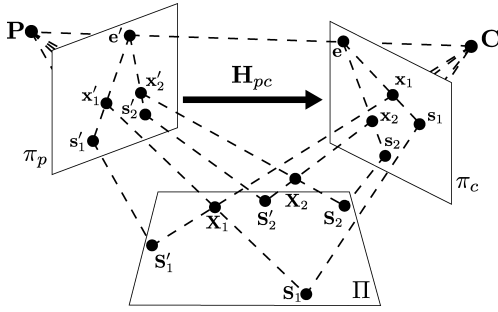


図5 カメラとプロジェクタの自己エピボラ

Fig. 5 Auto epipolar on cameras and projectors.

多くの投影情報 s, s' が得られる。これらの情報を用いることによって、対応点 2 点から以下に示すように F 行列が計算できる。

図 5 に示すように 3 次元空間中のある 2 点 X_1, X_2 がプロジェクタの光によってスクリーン Π 上に影 S_1, S_2 として投影され、それらがカメラ画像平面 π_c 上にそれぞれ x_1, x_2, s_1, s_2 として投影されているとする。このとき、図 5 から明らかなように点 x_1, s_1 を結んだ直線と、 x_2, s_2 を結んだ直線の交点は、カメラ画像平面 π_c におけるプロジェクタのエピポール e となる。

また、図 5 のプロジェクタ平面 π_p 上の 4 点 s'_1, x'_1, s'_2, x'_2 は、カメラ画像平面 π_c 上の 4 点 x_1, s_1, x_2, s_2 と平面射影変換行列 H_{pc}^{-1} を用いて以下のように求められる。

$$\begin{cases} \tilde{s}'_i \sim H_{pc}^{-1} \tilde{x}_i \\ \tilde{x}'_i \sim H_{pc}^{-1} \tilde{s}_i \end{cases} \quad (i = 1, 2) \quad (6)$$

カメラ画像平面 π_c と同様に、プロジェクタ平面 π_p 上の点 s'_1, x'_1 を結んだ直線と s'_2, x'_2 を結んだ直線との交点は、平面 π_p におけるカメラのエピポール e' となる。

以上より、プロジェクタ平面 π_p 上の 4 点には以下に示すようなエピボラ拘束が成り立つ。

$$\tilde{x}_i^T F_p \tilde{s}'_i = 0 \quad (i = 1, 2) \quad (7)$$

このように 1 枚の画像のみにエピボラ拘束が存在する場合を自己エピボラと呼ぶ。このとき、式 (7) の fundamental 行列 F_p は、次に示すような歪対称行列となる。

$$F_p = \begin{bmatrix} 0 & -f_1 & f_2 \\ f_1 & 0 & -f_3 \\ -f_2 & f_3 & 0 \end{bmatrix} \quad (8)$$

F_p は式 (8) に示すように 3 つの要素を持つが、定

数倍の不定性があるのでその自由度は 2 である。式 (7) より対応点 1 点から F_p に関して 1 つの拘束が得られる。したがって、2 点以上の N 点の対応点 $\tilde{s}'_i = (s_i^1, s_i^2, s_i^3)^T$, $\tilde{x}'_i = (x_i^1, x_i^2, x_i^3)^T$ ($i = 1, \dots, N$) より、以下の線形方程式を解くことにより F_p を計算することができる。

$$M f_p = 0 \quad (9)$$

ここで、 M は以下に示す $N \times 3$ 行列であり、また f_p は以下に示す 3 次元ベクトルである。

$$M = \begin{bmatrix} s_1^1 x_1^2 - s_1^2 x_1^1 & s_1^3 x_1^1 - s_1^1 x_1^3 & s_1^2 x_1^3 - s_1^3 x_1^2 \\ \vdots & \vdots & \vdots \\ s_N^1 x_N^2 - s_N^2 x_N^1 & s_N^3 x_N^1 - s_N^1 x_N^3 & s_N^2 x_N^3 - s_N^3 x_N^2 \end{bmatrix}$$

$$f_p = [f_1 \ f_2 \ f_3]^T$$

式 (9) の $\|f_p\| = 1$ なる条件のもとでの最小二乗解 f_p は行列 $M^T M$ の最小固有値に対応する固有ベクトルとして求まり、 F_p が求まる。同様にカメラ画像でもその平面における自己エピボラを表す fundamental 行列 F_c が求められる。

ここで、プロジェクタ平面 π_p からカメラ画像平面 π_c への平面射影変換 H_{pc} が求まっているので、

$$\tilde{s}'_i \sim H_{pc}^{-1} \tilde{x}_i \quad (10)$$

が成り立つ。これを式 (7) に代入すると、

$$\tilde{x}_i^T F_p H_{pc}^{-1} \tilde{x}_i = 0 \quad (11)$$

となる。ここで、

$$F = F_p H_{pc}^{-1} \quad (12)$$

とおくと、

$$\tilde{x}_i^T F \tilde{x}_i = 0 \quad (13)$$

となり、これは式 (5) に示したようにカメラとプロジェクタのエピボラ拘束式となる。したがって、カメラとプロジェクタの間の fundamental 行列 F は対応点が 2 点以上あれば式 (12) により求まることが分かる。

5. カメラとプロジェクタによる 3 次元復元

4 章においてカメラとプロジェクタ間のエピボラ幾何が求まったので、これを用いることにより影情報を用いたカメラとプロジェクタによる 3 次元復元が可能となる。

近年の多視点幾何の研究より、カメラ間のエピボラ幾何を求めることとこれらのカメラを射影的に校正することはまったく等しいことが明らかになっている。すなわち、前節の方法で F 行列が求まると、カメラの投影行列 A とプロジェクタの投影行列 A' はそれ

ぞれ以下のように求まる．

$$A = \begin{bmatrix} I & 0 \\ [e']_x F & e' \end{bmatrix}$$

$$A' = \begin{bmatrix} [e']_x F & e' \end{bmatrix}$$

ここで $[\cdot]_x$ はベクトル積を行列の積で表すための歪対称行列である．このようにしてカメラとプロジェクタの投影行列 A, A' が求まると、これを 2 台のカメラの投影行列と同等に扱い、カメラ画像上の対象物の像 x と対象物の影の像 s から、ステレオ視により 3 次元復元を行うことができる．ただし、投影行列 A, A' には射影的な不定性が存在するため、これらの投影行列を用いて 3 次元復元を行った結果は、射影的な不定性が残る射影復元となる．

ここでもしもカメラとプロジェクタの内部パラメータが既知であったり自校正などにより得ることができたりすれば、この射影的な不定性は取り除くことができ、ユークリッド復元を得ることができる．一方、完全にユークリッド的な 3 次元情報が得られなくても、射影復元のままで、あるいはアフィン復元に持つていくことにより、様々なアプリケーションに応用することもできる．次章では射影復元結果に対して情報を付加することによりアフィン復元が得られ、この結果、未校正プロジェクタカメラシステムにより仮想キーボードが実現できることを示す．

6. 仮想キーボードの実現

本章では、以上のようにして射影的に校正したプロジェクタカメラシステムを用いて、仮想楽器の 1 つである仮想キーボードを実現する方法を示す．

6.1 仮想キーボードの仮想配置

まず初めに仮想キーボードをスクリーン Π 上の任意の位置に仮想的に配置する．そのためにはプロジェクタ平面 π_p とスクリーン Π との関係が分からなければならないが、今分かっているのはカメラとプロジェクタ間の関係のみである．そこで、実際に仮想キーボードを配置する位置を指で指し示すことによりプロジェクタとスクリーンとの関係を求める．

図 6 に示すようにスクリーン上で指示した 4 点 K_i ($i = 1, \dots, 4$) を仮想キーボードの角 4 点とし、これらがカメラ画像平面 π_c 上に k_i ($i = 1, \dots, 4$) として投影されているとする．このとき、プロジェクタ平面 π_p 上の 4 点 k'_i ($i = 1, \dots, 4$) が以下のように求まる．

$$\tilde{k}'_i \sim H_{pc}^{-1} \tilde{k}_i \quad (i = 1, \dots, 4) \quad (14)$$

これらの点は、3 章で説明したように、プロジェク

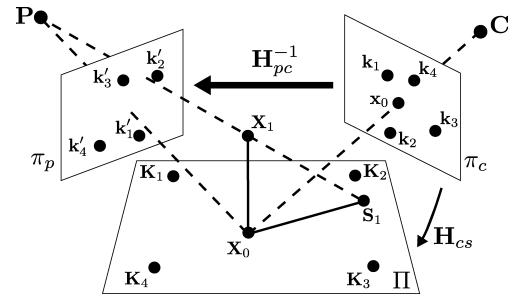


図 6 仮想キーボードの仮想配置と打鍵の検出
Fig. 6 Virtual setting of virtual keyboard.

タをカメラと見なしたときにスクリーン Π 上の 4 点 K_i ($i = 1, \dots, 4$) をプロジェクタ平面 π_p に投影した点であると考えられる．したがって、プロジェクタ平面 π_p 上の 4 点 k'_i ($i = 1, \dots, 4$) がキーボードの角となるよう平面 π_p 上にキーボードを描画すれば、スクリーン上において指で指し示した位置にキーボードが仮想配置される．

ただし、本稿ではスクリーン上の 4 点 K_i の指示を行う場合に、スクリーンからある一定の間隔だけ離れたスクリーンに平行な平面 Φ 上において指先を 4 カ所 R_i ($i = 1, \dots, 4$) に移動させ、この指先の 4 点 R_i のスクリーン上における影 4 点をキーボードの角 K_i とする．このときの平行平面 Φ を基準平面と呼ぶことにする．基準平面の実際の高さは知る必要はない．このようにスクリーン上ではなくスクリーンから一定の間隔だけ離れた平面上において 4 点を指示する理由は、スクリーン上のキーボードの角の 4 点を指示しつつ、次節で述べる射影復元からアフィン復元への不定性の除去を行うためである．

6.2 アフィン復元と打鍵の強さ

仮想キーボードを実現するためには、指先で押したキーを判別すると同時に、どのような強さでこのキーを押したかを計算し、これを発生する音の強弱として表現する必要がある．押されたキーの判別は指先がスクリーン平面と接触した瞬間のキーボードに対する指の 2 次元位置によって判別することができる．一方、キーを押した強さは指の打ち下ろし速度に比例すると考えられるので、ある微小な時間間隔 Δt における指の 3 次元位置の変化から求める必要がある．指の各時刻における 3 次元位置は、前節に述べた方法により、画像における指先の像とプロジェクタによる指先の影の像から射影復元できる．しかし射影復元には 3 次元射影変換の不定性が残っていることから、射影復元結果から音の強さを求めることはできない．そこで射影復元された結果に対して情報を付加することによりア

フィン復元を得る方法を考える．アフィン復元された結果には 3 次元アフィン変換の不定性が残るが，距離の比は不定性なく一意に求まるため，音の強さを制御するためにはユークリッド復元が行えなくてもアフィン復元が行えれば十分である．

射影復元結果からアフィン復元結果を得るには，もともと実空間において無限遠に存在する点が，復元結果中においても無限遠点となるよう射影復元結果に対して 3 次元射影変換を行えばよい．これは，実空間において平行な平面上に存在する点が復元結果中においても平行な平面上に存在するように変換することでもある．

そこで，まず前節において指で指示した 4 点 R_i がスクリーンと平行な平面上に存在することを用いて，これら 4 点 R_i の射影復元結果が，これら 4 点のスクリーン上における影 K_i の復元結果と平行となるよう射影復元結果を射影変換する．この変換により，スクリーン平面上の無限遠点が復元結果においても正しく無限遠の点となる．しかしこの状態ではスクリーン平面上にはない無限遠点が，復元結果中においてまだ正しく無限遠の点とはなっていない．そこで次にこれらの点が復元結果中において正しく無限遠点となるよう射影変換する．このために，仮想キーボードを打鍵する場合には，スクリーン近傍において指がほぼ等速直線運動していることを利用する．直線上において等間隔に並ぶ 3 点 X_0, X_1, X_2 と無限遠点 X_∞ の計 4 点がなす複比は射影変換のもとで不変であり，その値はつねに -1 となる．この性質を用いることにより，射影的に至んだ直線上等間隔の 3 点からその直線上の無限遠点が求まる．この点が実際に無限遠の点となるよう復元結果を射影変換することにより，アフィン復元を得ることができる．

このようにして得られたアフィン復元結果においては比が不変である．したがって，アフィン復元結果においてスクリーンに指が接する瞬間から Δt 時刻前の指の高さが H であり基準平面 Φ の高さが H' であるとき，打鍵の強さ h は以下に示すように H' に対する指の高さ H として求まる．

$$h = \frac{H}{H'} \quad (15)$$

7. 実験

本章では，まず 4 章で提案した影情報を用いたカメラとプロジェクタのエピポーラ幾何計算法を評価する．次に 6 章で述べた仮想キーボードを実装し評価を行った結果を示す．

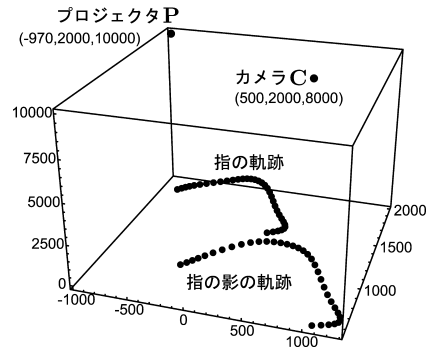


図 7 シミュレーションの配置図

Fig. 7 Simulation experiment.

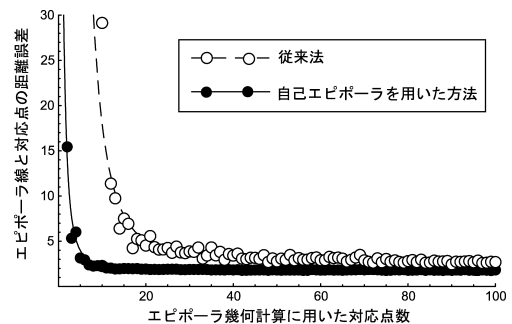


図 8 エピポーラ幾何計算評価

Fig. 8 Evaluation of computed epipolar geometry.

7.1 影情報によるエピポーラ幾何計算の安定性

まず，カメラとプロジェクタのエピポーラ幾何計算および射影復元の安定性をシミュレーション実験により評価する．

7.1.1 エピポーラ幾何計算のシミュレーション

図 7 に示すようにプロジェクタ P とカメラ C をそれぞれ $(-970, 8000, 10000)$, $(500, 2000, 8000)$ の位置に配置し，カメラに投影された指と影の軌跡よりランダムに 2 組から 100 組の対応点を取り，fundamental 行列を計算した．このとき，8 点以上の対応点から線形に fundamental 行列を求める正規化 8 点法と，自己エピポーラを用いて 2 点以上の対応点から線形に fundamental 行列を求める提案法を比較評価した．評価は計算に用いた点と異なる点をランダムに 50 点とり，これらの点に関する画像中のエピポーラ線と対応点との距離により行った．この結果を図 8 に示す．なお，画像ノイズとして標準偏差 1 ピクセルのガウスノイズをカメラ画像に印加した．

図 8 の横軸はエピポーラ幾何計算に用いた対応点数であり，縦軸は fundamental 行列から計算したエピポーラ線と対応点の距離を表す．図 8 より，自己エピポーラを用いた提案法の方が，従来法より少ない対応

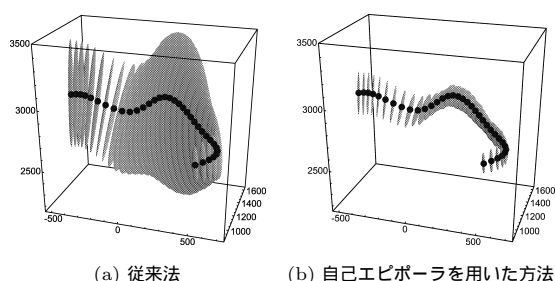


図 9 射影復元結果

Fig. 9 Results from projective reconstruction.

点からより安定なエピポラ幾何計算が可能であることが分かる．これは提案法は従来法と比べて計算の空間が 8 自由度から 2 自由度に減少するために画像ノイズの影響を受けにくくなるためと考えられる．さらに，従来法ではそもそも幾何学的に 7 自由度の問題を線形 8 自由度の問題に置きかえて計算しているため，余計な自由度を加えているために計算が不安定になるのに対して，提案法では幾何学的に 2 自由度の問題を線形 2 自由度の問題として解いているために安定になると考えられる．

7.1.2 射影復元のシミュレーション

次に，図 7 に示す配置で，高さ 3000 付近で動かした指の軌跡の射影復元を行った．その結果を図 9 (a)，(b) に示す．なお，射影復元には，対応点 30 点から従来法および自己エピポラを用いた提案法より求めた fundamental 行列を用いた．また画像ノイズとして標準偏差 1 ピクセルのガウスノイズを印加し，100 回繰り返し計算を行って，復元結果の不確定領域を計算した．

図 9 (a)，(b) 中の楕円体は，手の軌跡の 3σ の不確定領域を表す．この不確定領域より，従来法より自己エピポラを用いた提案法の方が格段に安定に射影復元ができていたことが分かる．ただし，射影復元結果には射影的な不定性があるので，正しい射影基底を割り当てて評価を行った．

7.2 仮想キーボードシステム

本節では実際に仮想キーボードを実装し，打鍵の検出の精度や打鍵の位置や強さの計算の精度を評価する．

7.2.1 システム構成

本システムでは，図 10 に示すように，机の上をスクリーンとし，これに向かってカメラとプロジェクタを 1 台ずつ配置した．これらのカメラとプロジェクタの位置や姿勢，内部パラメータは未知であり，本システムではこれらを未校正のまま利用する．なお，本システムではキーボードの演奏を実現するために指先と指

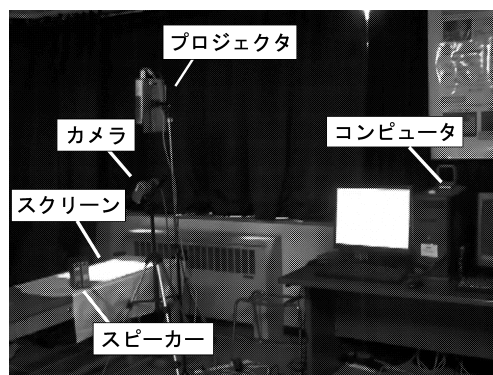


図 10 システム構成

Fig. 10 Experimental system.

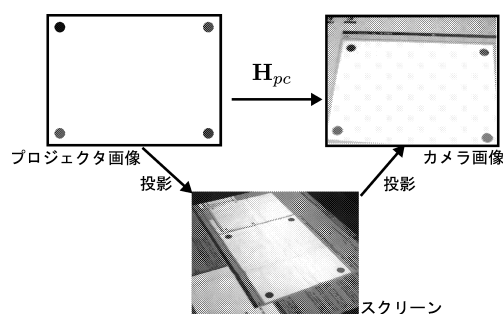


図 11 プロジェクタとカメラの間の平面射影変換

Fig. 11 Planar projective transformation between a projector and a camera.

先の影を抽出し対応付ける必要があるが，システムを簡単にするために，指を 1 本だけ伸ばした状態で演奏するという条件のもとで画像中における指先点の像と指先点の影の像との間の対応を自動的に求める．そのためにまず背景差分を用いて「手と手の影の領域」を抽出する．次に，色検出により「手の領域」と「手の影の領域」を分離する．このようにして得られた手の像と手の影の像のそれぞれの上端を指先と指先の影と考えて対応点とする．

以上のようなシステム構成のもとで，まず初めに 4 点の基底点をプロジェクタを通してスクリーンに投影する．このときスクリーンに投影された基底点をカメラにより観測し，プロジェクタとカメラの間の平面射影変換 H_{pc} を求める．この実験の様子を図 11 に示す．

H_{pc} が求まると，スクリーン上で指し示した位置がコンピュータ内の仮想空間のどこにあたるかが計算できる．そこで，次に仮想キーボードの配置を行う．ここで，後述する打鍵の強さの計算のために，スクリーンと基準平面の関係も一緒に求める．そこで，図 12 に示すようにスクリーンから基準となる高さにスクリーンと平行な平面を仮定し，その基準平面上の

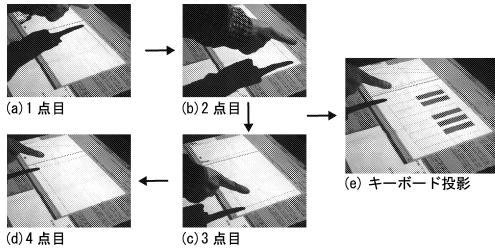


図 12 仮想キーボードの配置

Fig. 12 Setting of virtual keyboard.

4 点を指で指し示す．基準平面は実際には存在していないので，ユーザがスクリーンと並行となるよう指を動かすことによりこれら 4 点を得る．このとき，指の影がスクリーンに投影されているので，この影 4 点を仮想キーボードの角 4 点とし，そこへ 6.1 節に示す方法で仮想キーボードを配置する．

仮想キーボードの角 4 点を指し示すと同時に，カメラ画像面上の自己エピポーラを表す fundamental 行列 F_c を 4 章に示す方法で計算する． F_c は，カメラ画像に映る指と指の影の対応点 2 組から求めることができるが，図 8 に示すように対応点が多ければ多いほど安定に fundamental 行列を計算することができる．仮想キーボードシステムはリアルタイム処理のため指を動かせば動かすほど対応点の組が多く得られるため，仮想キーボードの角 4 点を指し示している間の対応点をできるだけ多く F_c の計算に用いる．ただし，指と影が重なって対応が不安定になるような状況が生じた場合，fundamental 行列の計算精度が向上するとは限らない．したがって，対応点の組において，指の影が指に隠れて得られないなど明らかに間違っている組を使わないように，RANSAC¹⁴⁾ を用いて outlier をはじく．このようにして取得した対応点は安定であるため，fundamental 行列 F_c の計算精度の向上が期待できる．また F_c の計算は安定な対応が得られるキーボード配置時においてのみ行い，次に述べる指の 3 次元復元時には行わない．以上から F_c が求まるので， F_c からカメラに映るプロジェクタのエピポール e を求める．このようにしてエピポーラ線を求めた結果を図 13 に示す．なお，図 13(a)，(b)，(c) はそれぞれ指の高さを変えたときの結果である．

図 13 中の直線は計算により求めたエピポーラ線である．自己エピポーラにおけるエピポーラ線上には，指の点とエピポールのほかに，対応する指の影の点に乗るが，図 13 では確かに指の影の点をエピポーラ線が通過している．このことからエピポーラ幾何がほぼ正しく求まっていることが分かる．

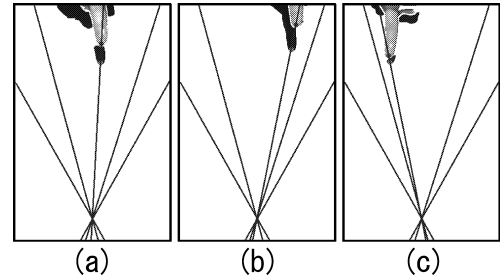


図 13 エピポーラ線の計算

Fig. 13 Computed epipolar geometry.

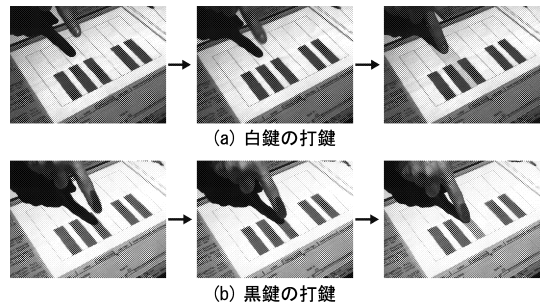


図 14 演奏の様子

Fig. 14 Playing the virtual keyboard.

以上のようにキーボードを配置したのち，求めた fundamental 行列を用いて指の 3 次元復元を行う．スクリーンに投影されたキーボードの鍵盤への打鍵の様子をリアルタイムに 3 次元復元することで，打鍵した鍵盤の識別と打鍵の強さ計算を行い，それに対応した音をスピーカから出す．なお，キーボードへの打鍵時，指がスクリーンに接地して指先の影が正しく取得できなくなる．しかし，指先の影の位置は指先の位置と同じになるため，指先の位置を用いて指先点の復元を行う．このとき，接地したかどうかの検出が必要あるが，本システムでは単純に指と指の影の対応が正しくとれている状態からとれなくなったときを，指で指の影が隠れたとき，すなわち接地したときとする．なお，指先点と指先点の影の対応が正しいかどうかの判断は，指の位置から fundamental 行列を用いてエピポーラ線を計算し，そのエピポーラ線上に取得した指先点の影が乗っており，なおかつ前フレームまでの動きから予測される位置より大きく外れていないとき，正しく対応がとれているとする．

以上のようにして演奏している様子を図 14 に示す．図 14(a) は白鍵の「ソ」を押した様子で，「ソ」を押したときにその鍵盤の色が変わって確かに「ソ」を押したということが識別できていることが分かる．また，図 14(b) は黒鍵の「ファ#」を押した様子であり，白



図 15 聖者の行進

Fig. 15 Score used for the experiment.

表 1 「聖者の行進」の演奏結果
Table 1 Evaluation of the virtual keyboard.

演奏回数	正しい音が出た	正しい音が判定できなかった	合計	成功率
1	32	0	32	1.0
2	32	0	32	1.0
3	31	1	32	0.97
4	31	1	32	0.97
5	30	2	32	0.94
合計	156	4	160	0.975

鍵よりも幅の小さい黒鍵への打鍵も正しく識別できていることが分かる。また、図 14 (a) の方が (b) よりも音が強く出ているが、図 14 (a) の時間幅と (b) の時間幅は同じなので、同じ時間でより大きく指が動いた図 14 (a) の方が、音が強くなるのは正しいということが分かる。

以上のとおり提案法に基づき、未校正カメラと未校正プロジェクタを用いて仮想キーボードを実現することができた。

7.2.2 打鍵の位置計算の妥当性

次に、仮想キーボードにおいてどれぐらいの精度で正しく打鍵が識別でき、正しい音が鳴るかを評価した。

仮想キーボードは楽器なので、ただ音を鳴らすだけでなく、実際に曲を演奏した方が自然である。そこで、図 15 に示す「聖者の行進」を 5 回ずつ演奏して全音符に対して正しく出た音の割合を調べた。

図 15 に示す「聖者の行進」を 5 回演奏した結果を表 1 に示す。ここでは、1) 正しい音が出たとき、2) 正しい音が判定できなかったとき、のそれぞれの発生回数を調べた。

表 1 において成功率が 9 割を超えていることから、ほぼ正しく音を出せているといえる。正しい音が出なかったのは、失敗した音が短い音符であったときである。このことから、仮想キーボードにおいては、指の動きが速い場合には正しく音を出すのが難しいことが分かる。これは、本実験で用いたカメラ系では、ブルアリングなどにより、高速運動時の画像位置が正しく得られないことに起因していると考えられる。

7.2.3 打鍵の強さ計算の妥当性

次に、6.2 節で述べた打鍵の強さ h が正しく計算できているかどうかを評価した。計算により求めた打鍵の

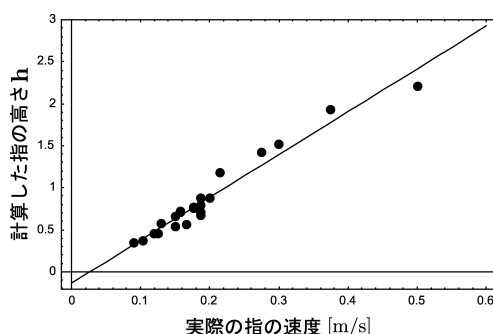


図 16 実際の指の速度と計算した打鍵の強さの関係

Fig. 16 Real motions and computed motions of a finger.

強さ h は打鍵直前の指の速度に比例しなければならない。そこで何十回か打鍵を行い、実際の指の速度と計算した打鍵の強さ h の関係を調べた。

図 16 に実際の指の速度と計算した打鍵の強さ h の関係を示す。なお、打鍵の回数は 22 回である。横軸は実際の指の速度を表し、縦軸は計算した打鍵の強さ h を表す。また、図中の直線は全 22 点を用いて最小二乗法により求めた直線である。

図 16 より実際の指の速度と計算した打鍵の強さはほぼ比例の関係にあることが分かる。また、最小二乗法で求めた直線はほぼ原点近くを通っていることから、ほぼ正しく打鍵の強さが計算できていると考えられる。

また図 16 に示すように、実験で現れた指の速度は $0.1 \sim 0.2$ [m/s] に集中している。これは演奏時に指を動かす速度がこの近辺に集中していたことを表している。このように実際の演奏時にはよく現れる指の速度帯域が存在すると考えられる。指の高さ計算は、一般に運動速度が遅いほど画像ノイズの影響を大きく受け、速度が速くなるほど画像ノイズの影響が少なくなると考えられる。しかし一方で、指を動かす速度が速いほど画像に大きなブルアリングが生じ計算結果が悪くなる。今後は演奏でよく現れる速度帯域を調査し、その速度帯域においていかにして計算精度を上げるかを調べる必要がある。

8. ま と め

本研究では、未校正なプロジェクタカメラシステムを射影的に校正する方法を示した。特に、プロジェクタによってできる対象物の影の情報を利用することにより、基準儀などを用いずにカメラとプロジェクタの投影行列が求まることを示した。また、このようにして校正したプロジェクタカメラシステムを用いて、対象物の影の情報より対象物の 3 次元情報を計算する方法を示した。さらに、このようにして射影的に校正さ

れたプロジェクタカメラシステムを用いて仮想楽器の一種である仮想キーボードを実現する方法を示した。仮想キーボードでは、スクリーン上の2次元的な指の位置のみでなく、スクリーンに接触するときの指の速度を計算するために指の3次元的な位置を計算する必要がある。提案法では、プロジェクタカメラシステムの投影行列が射影的に求まるため、このような速度計算が可能となることを示した。

実際に仮想キーボードを構築し評価実験を行うことにより、プロジェクタカメラシステムの射影的校正が正しく行われていることを示した。

今後の課題としては、2台以上の複数台のカメラとプロジェクタの関係性を効率良く校正する方法への拡張が考えられる。また、仮想キーボードの今後の課題としては、複数の指が扱えるシステムへの拡張など、より実際のキーボードの機能に近づけることがあげられる。

参 考 文 献

- 1) Oka, K., Sato, I., Nakanishi, Y., Sato, Y. and Koike, H.: Interaction for entertainment contents based on direct manipulation with bare hands, *Proc. IFIP International Workshop on Entertainment Computing*, pp.391-398 (2002).
- 2) Milgram, P. and Kishino, F.: A taxonomy of mixed reality visual display, *IEICE Trans. Inf. & Sys.*, Vol.E77-D, No.12, pp.1321-1329 (1994).
- 3) 田村秀行, 大田友一: 複合現実感. 映像情報メディア学会誌, Vol.52, No.3, pp.266-272 (1997).
- 4) Sukthankar, R., Stockton, R.G. and Mullin, M.D.: Smarter Presentations: Exploiting Homography in Camera-Projector Systems, *Proc. International Conference on Computer Vision*, Vol.1, pp.247-253 (2001).
- 5) Okatani, T. and Deguchi, K.: Autocalibration of a Projector-Camera System, *IEEE Trans. PAMI*, Vol.27, No.12, pp.1845-1855 (2005).
- 6) 東城賢司, 日浦慎作, 井口征士: プロジェクタを用いた3次元遠隔指示インタフェースの構築, 日本バーチャルリアリティ学会論文誌, Vol.7, No.2, pp.169-176 (2002).
- 7) Hartley, R. and Zisserman, A.: *Multiple View Geometry*, Cambridge University Press (2000).
- 8) 郭, 谷内田, 辻: 影情報を利用した線状物体の端点の三次元位置計測法, 日本ロボット学会誌, Vol.2, No.2, pp.151-157 (1983).
- 9) Yu, Y. and Chang, J.T.: Shadow Graphs and 3D Texture Reconstruction, *IJCV*, Vol.62,

No.1-2, pp.35-60 (2005).

- 10) Savarese, S., Rushmeier, H., Bernardini, F. and Perona, P.: Shadow Carving, *Proc. ICCV2001* (2001).
- 11) Savarese, S., Rushmeier, H., Bernardini, F. and Perona, P.: Implementation of a Shadow Carving System for Shape Capture, *Proc. 1st International Symposium on 3D Data Processing Visualization and Transmission (3DPVT02)* (2002).
- 12) 江川宗一郎, 佐藤 淳: 複合現実感による仮想楽器の実現, 第10回画像センシングシンポジウム, pp.161-166 (2004).
- 13) 佐藤 淳: コンピュータビジョン—視覚の幾何学, コロナ社 (1999).
- 14) Fischler, M.A. and Bolles, R.C: Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography, *Comm. ACM*, Vol.24, No.6, pp.381-395 (1981).
- 15) <http://www.tanomi.com/limited/html/00034.html>

(平成 17 年 9 月 20 日受付)

(平成 18 年 3 月 20 日採録)

(担当編集委員 杉本 晃宏)



西江 桂亮

平成 17 年名古屋工業大学工学部電気情報工学科卒業。現在同大学大学院情報工学専攻修士課程在学中。コンピュータビジョンに関する研究に従事。



佐藤 淳

昭和 59 年名古屋工業大学工学部機械工学科卒業。平成 8 年ケンブリッジ大学大学院博士課程修了。同年ケンブリッジ大学工学部助手。平成 10 年名古屋工業大学工学部助教授。平成 16 年より同大学大学院情報工学専攻教授。この間に、ATR 人間情報通信研究所客員研究員等、コンピュータビジョン、視覚誘導、視覚的ユーザインタフェースの研究に従事。博士 (Ph.D.)。著書に『コンピュータビジョン—視覚の幾何学』(コロナ社) 等。BMVC'94 最優秀科学論文賞, BMVC'97 最優秀科学論文賞, 第 5 回画像センシングシンポジウム論文賞受賞。電子情報通信学会, IEEE 各会員。