2005－CVIM－149（14）
2005／5／12

未校正マルチカメラにおける3次元シーンフロー推定

水野 宏基　　藤吉 弘亘　　岩堀 祐之

中部大学 工学部 情報工学科
E-mail:hiroki@vision.cs.chubu.ac.jp, hf@cs.chubu.ac.jp, iwahori@cs.chubu.ac.jp

あらまし　従来，シーン内の物体の3次元運動 (3次元シーンフロー) を推定するには強カメラ校正情報をもとに，予め求めておいたシーンの Depth map や物体の表面の情報から推定を行っていた．しかし，カメラの台数が多い場合，強カメラ校正を行うには大変な作業を要する．そこで本稿では，未校正マルチカメラを用いて3次元シーンフローを推定する．提案する手法は，3次元空間を Projective Grid Space により定義し，その空間における全ての格子点を全てのカメラに投影し，その画像平面上のオプティカルフローの一致性をエピポーラ幾何による拘束を用いて調べることで3次元シーンフロー推定を行う．シミュレーションと実画像を用いた実験より推定したシーンフローを評価し，提案手法の有効性を示す．

A Method for Estimating Three Dimensional Scene Flow Using Non Calibrated Multiple Cameras

Hiroki Mizuno　　Hironobu Fujiyoshi　　Yuji Iwahori

Dept. of Computer Science, Chubu University
E-mail:hiroki@vision.cs.chubu.ac.jp, hf@cs.chubu.ac.jp, iwahori@cs.chubu.ac.jp

Abstract　Three dimensional scene flow can be estimated from information of the object's surface and depth map of the scene based on camera parameters obtained by strong calibration. However, increasing the number of cameras increases labor of the calibration. In this paper, we present a method for estimating three dimensional scene flow using non calibrated multiple cameras. Our approach uses the Projective Grid Space to define virtual three dimensional space, and then the three dimensional scene flow can be estimated by checking consistency of the optical flows computed for each cameras using constaint by the epipolar geometry. Effects of the proposed method are shown by experimental results of simulation and real images.

1　はじめに

オプティカルフローは，2次元画像間での同一対象の対応付けを行い，その移動量をベクトルで表現したものであり，実空間中のシーンの動きを画像平面に投影したものである．一方，3次元シーンフローは，3次元空間における同一対象の移動量を3次元のベクトルで表現したものである．このような3次元シーンフローは，人間の動きを理解するジェスチャー認識や，3次元形状復元の高速化と精度の向上，また3次元形状の時間的内挿等への応用が期待されている．

従来3次元シーンフロー推定には大きく分けて以

下の二つの手法が提案されている．
(1) 時系列のシーン内の 3 次元構造のマッチング [1]
(2) シーン内の 3 次元構造と，画像上のオプティカルフローとの統合 [2]
前者は時刻 t と $t-1$ でのシーン内の 3 次元構造を予め復元しておき，物体の表面形状とテクスチャを特徴とし，フレーム間での対応点を探索しシーンフローを推定する手法である．復元されたシーン内の 3 次元構造が正確で密である場合，サーフェイスマッチングにより正確な 3 次元シーンフローを推定することが可能である．また，復元された 3 次元構造が密でない場合でも，物体表面の局所的な凹凸情報と物体の色情報からフレーム間の対応を取る手法が提案されている [1]．後者は時系列のシーン内の 3 次元構造と画像平面上のオプティカルフローを統合する手法である．シーン内の物体の表面情報が予め正確で密に求められている場合には，その表面に画像平面上のオプティカルフローを逆投影するすることで 3 次元シーンフローを推定することができる．また，表面情報が密でない場合でも，2 台以上のカメラ画像から，シーンとオプティカルフローとの間の関係を計算することが可能であるため，シーン内の物体の表面上の各点において，各カメラにおけるオプティカルフローを投影することでシーンフローを推定することが可能となる．

しかし，これらの手法ではシーンの 3 次元構造を予め復元しておく必要がある．この問題に対して，Vedula 等はシーンの 3 次元構造を必要としないシーンフロー推定手法を提案している [2]．複数のカメラから得られたオプティカルフローを 3 次元空間に投影し，その一致性を利用してシーンフローを推定する．しかし，この手法では，正確な強カメラ校正を行い，3 次元空間への投影のために必要なカメラパラメータが既知である必要がある．一般に，このような強校正には大変な作業を要するという問題がある．

そこで，本研究では複数の未校正カメラにおける，エピポーラ幾何を拘束条件とした 3 次元シーンフローの推定手法を提案する．本提案手法では，Projective Grid Space により 3 次元空間を定義し，その空間内の各点を投影する全てのカメラ画像中の点についてオプティカルフローの一致性を計算し，3 次元シーンフローを推定する．

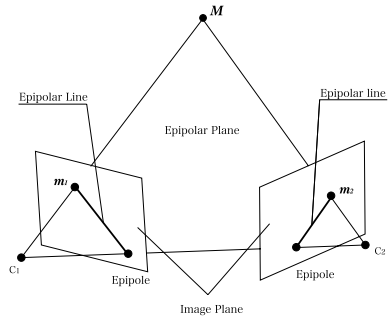


図 1: エピポーラ幾何における諸概念

2 未校正マルチカメラによる 3 次元空間の定義

複数のカメラを用いてシーンの 3 次元構造を得るには，画像間の対応が重要な課題となる．正確な対応を求めるには，強カメラ校正によって得られるカメラの内部･外部パラメータの精度が重要となる．しかし，使用するカメラ数が多い場合，カメラの強校正に費やす時間は膨大なものとなる．この問題に対して Saito らは，複数台の弱校正カメラの関係から定義される Projective Grid Space を提案している [4]．Projective Grid Space は 2 つのカメラを Base View とし，Base View1 の u 軸，v 軸，Base View2 の u 軸を用い，3 次元空間を定義する手法である．これにより強カメラ校正情報を必要とせず擬似的な 3 次元空間を扱うことができる．

2.1 エピポーラ幾何と F 行列

2 台のカメラ間には，エピポーラ幾何と呼ばれる画像間の幾何関係が存在する [5][6]．2 台のカメラで 3 次元空間中の同一点を撮影した場合，図 1 に示すように，その 3 次元空間中の点，各カメラのレンズ中心，画像上の点の全てを通るエピポーラ平面が存在する．エピポーラ平面とそれぞれの画像との交線であるエピポーラ直線は，カメラ間の関係から決定される．画像 1 における点 $\tilde{\boldsymbol{m}}_i = [u_i, v_i, 1]^T$ と画像 2 における対応点 $\tilde{\boldsymbol{m}}'_i = [u'_i, v'_i, 1]^T$ は式 (1) を満たす．

$$\tilde{\boldsymbol{m}}_i^T \boldsymbol{F} \tilde{\boldsymbol{m}}'_i = 0 \tag{1}$$

ここで $\boldsymbol{F}$ を基礎行列，式をエピポーラ方程式と言う．

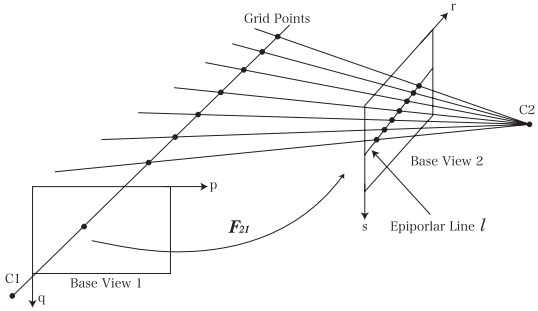


図 2: Projective Grid Space による 3 次元空間の定義

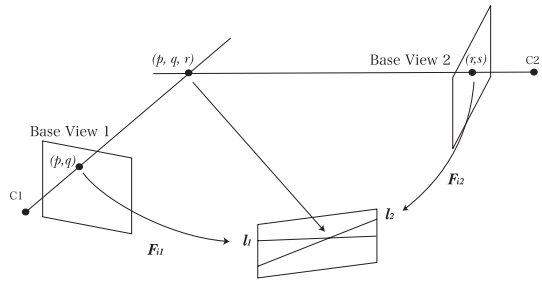


図 3: 任意位置に設置されたカメラへの Projective Grid Space 内の格子点の投影

2.2 Projective Grid Space

2 つの基本画像 (Base View) と呼ばれるカメラを選択し，2 つのカメラ画像から図 2 のように空間中に格子点を定義する．Base View 1 の座標 (p, q) は空間に 1 本の格子線を定義し，この格子線の節は Base View 2 の r か s のどちらかにより定義される．Base View 1 と Base View 2 の間の基礎行列 $\boldsymbol{F}_{21}$ はエピポーラ直線 $\boldsymbol{l}$ を Base View 2 上に定義し，空間中の格子点は $\boldsymbol{l}$ 上の点として表される．この空間を Projective Grid Space と呼ぶ．Projective Grid Space は 2 つの基本画像から定義され，格子上の節は (p, q, r) として表現される．格子上の点 (p, q, r) が基本画像上の座標 (p, q) と (r, s) に投影されたとすると，s は式 (2) として表される．

$$s = \frac{-(l_x r + l_z)}{l_y} \tag{2}$$

ベクトル $\boldsymbol{l} = (l_x, l_y, l_z)$ は点 (p, q) における Base View 2 上のエピポーラ直線で，式 (3) として表される．

$$\boldsymbol{l} = \boldsymbol{F}_{21} \begin{bmatrix} p \\ q \\ 1 \end{bmatrix} \tag{3}$$

Projective Grid Space と任意位置に設置されたカメラの画像との関係は，任意の画像 i と 2 つの基本画像との間の基礎行列 $\boldsymbol{F}_{i1}$ と $\boldsymbol{F}_{i2}$ により決定され，格子上の点 (p, q, r) から画像 i へ投影された点は次の手順から導かれる．(p, q, r) が Base View 1 の (p, q) 上に投影されている時，i 番目の画像へ投影される格子点は $\boldsymbol{F}_{i1}$ と (p, q) から与えられるエピポーラ直線 $\boldsymbol{l}_1$ 上に存在する．同じように，Base View 2 の (s, r) 上に投影されている時，i 番目の画像へ投影される格子点は Base View2 の (s, r) と $\boldsymbol{F}_{i2}$ から与えられる $\boldsymbol{l}_2$ 上に存在する．これより，図 3 に示すように，i 番目の画像上のエピポーラ直線 $\boldsymbol{l}_1$ と $\boldsymbol{l}_2$ の交点が格子点 (p, q, r) を投影した点となる．

3 シーンフローの推定手法

従来の 3 次元シーンフロー推定手法では，シーン内の物体の形状の情報や，Depth map，またカメラの内部・外部パラメータ情報を必要とする．我々は，未校正マルチカメラを対象としているため，これらの情報は未知である．そこで，Projective Grid Space においてエピポーラ幾何による拘束条件を用いて 3 次元シーンフローを推定する．

3.1 画像間のオプティカルフローの一致性

Base View 間でのフローの一致性

図 4 に示すように，Projective Grid Space における 2 台の Base View 間では，3 次元空間中の点 $\boldsymbol{M}$ にシーンフローが存在する場合，その点を投影した 2 つの画像上にそれぞれオプティカルフローは存在する．点 $\boldsymbol{M}$ を Base View 1 に投影した点を $\boldsymbol{m}_1$ とし，その点のフローの終点を $\boldsymbol{m}'_1$ とする．また，点 $\boldsymbol{M}$ を Base View 2 に投影した点 $\boldsymbol{m}_2$ にもフローが存在する．この場合，Base View 2 における点 $\boldsymbol{m}_2$ のフローの終点 $\boldsymbol{m}'_2$ は，点 $\boldsymbol{m}'_1$ における Base View 2 上のエピポーラ直線 $\boldsymbol{l}'$ 上に存在する．これらのことから，点 $\boldsymbol{m}'_2$ と $\boldsymbol{l}'$ の間には次式が成り立つ．

$$\boldsymbol{m}'^T_2 \boldsymbol{l}' = 0 \tag{4}$$

よって，式 (4) を満たす 2 つの Base View 間の点 $\boldsymbol{m}_1$，$\boldsymbol{m}_2$ の組ではオプティカルフローが一致する．

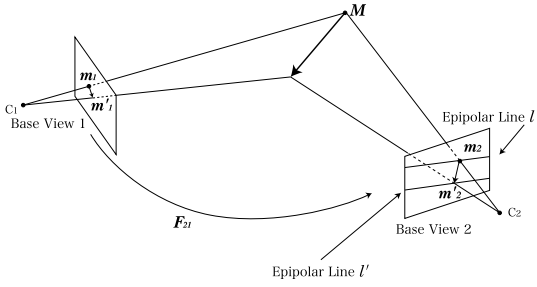


図 4: Base View 間のフローの一致性

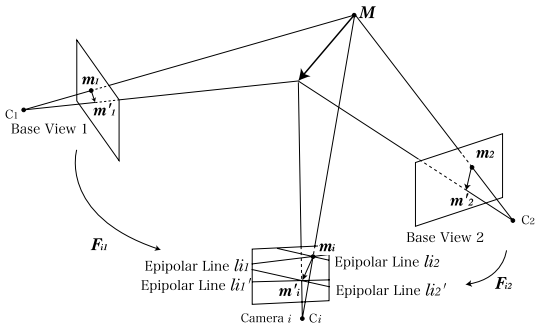


図 5: 任意の画像におけるフローの一致性

任意カメラにおけるフローの一致性

任意位置に設置されたカメラ画像においては，図 5 に示すように，2 つの Base View との間の基礎行列からフローの一致性を求めることができる．点 $\boldsymbol{M}$ にシーンフローが存在すると，i 番目のカメラに投影されるフローの終点の位置は，Base View 1 におけるフローの終点 $\boldsymbol{m'_1}$ と，Base View 1 とカメラ i との間の基礎行列 $\boldsymbol{F}_{i1}$ から求められるエピポーラ直線 $\boldsymbol{l'}_{i1}$ と，Base View 2 におけるフローの終点 $\boldsymbol{m'_2}$ と，Base View 2 とカメラ i との間の基礎行列 $\boldsymbol{F}_{i2}$ から求められるエピポーラ直線 $\boldsymbol{l'}_{i2}$ の交点となる．i 番目のカメラにおけるフローの終点の座標を $\boldsymbol{m'}_i$ とすると，次式が成り立つ．

$$\begin{cases} \boldsymbol{m'}_i^T \boldsymbol{l'}_{i1} = 0 \\ \boldsymbol{m'}_i^T \boldsymbol{l'}_{i2} = 0 \end{cases} \tag{5}$$

よって，式 (5) の両式を満たす点 $\boldsymbol{m'}_i$ は，オプティカルフローが一致することになる．さらに，カメラの台数が増えた場合も，同様に式 (5) を満たす点であれば，オプティカルフローは一致すると言える．

3.2 シーンフローの推定アルゴリズム

2 次元オプティカルフローから 3 次元シーンフローの推定アルゴリズムを以下に示す．

Step1 各入力カメラ画像に対し時刻 t と $t-1$ のオプティカルフローを計算

Step2 Projective Grid Space のある格子点 (初期探索では Base View1 から最も近い格子点を選択) について Base View1 と Base View2 においてオプティカルフローの一致性を計算

Step3 フローが一致する場合，任意のカメラ C_i においてもオプティカルフローの一致性を計算

Step4 全てのカメラにおいてオプティカルフローが一致した場合，2 台の Base View からシーンフローの大きさと方向を計算

Step5 次の格子点に移り **Step2**∼**Step4** を繰り返し，全ての格子点に対しオプティカルフローの一致性を計算

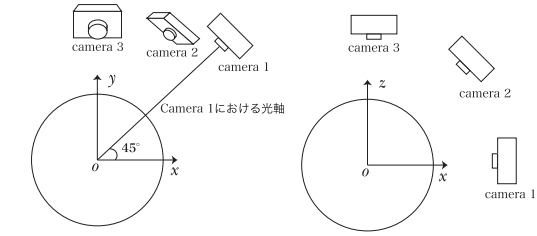
本アルゴリズムにおいてオプティカルフローの一致性を計算する際には，シーン内の物体のオクルージョンを考慮する必要がある．この問題に対し Seitz 等は Voxel Coloring と呼ばれる，1 ビットの Z バッファを用いた手法を提案している [7][8]．シーンフロー推定を行う前に，全てのカメラの各画素に対し，0 の値に初期化した Z バッファを与える．カメラにはシーン内の物体の表面だけが投影されることから，オプティカルフローが一致した点について Z バッファを 1 にする．次の格子点に移り，その格子点の Z バッファが 1 でる場合は，その点はシーンフローの計算対象外とする．

4 シーンフロー推定実験

提案した手法を用いた 3 次元シーンフローの推定実験を行う．シミュレーションと実画像を用いての実験を行い，提案手法の有効性を示す．

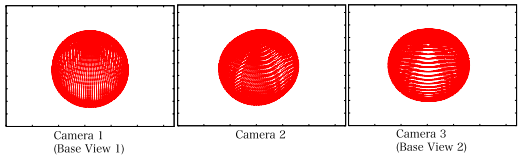
4.1 シミュレーションによる実験

シミュレーションでは，3 次元空間中に球を定義し，その表面上の点の移動の軌跡を画像平面に投影する．カメラは，図 6 に示すように，camera 1 を $x-y$ 平面で 45 度の角度で原点を見下ろす位置に配置し，camera 2 と camera 3 はそれぞれ camera 1 から $x-z$ 平面で 45 度ずつ回転させた位置とする．球の表面上の点 90×180 点を x 軸方向に平行移動さ

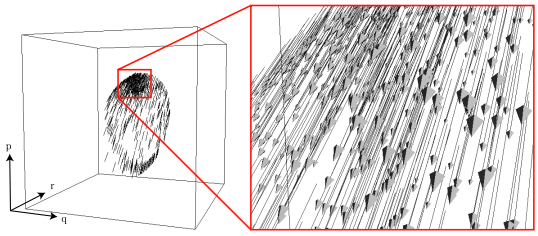


(a) x-y平面におけるカメラ配置　(b) x-z平面におけるカメラ配置

図 6: シミュレーションモデルにおけるカメラ配置

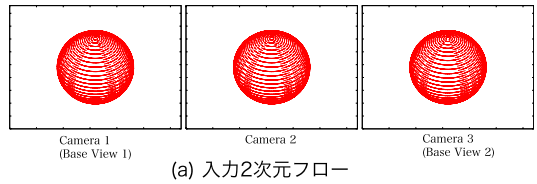


(a) 入力2次元フロー

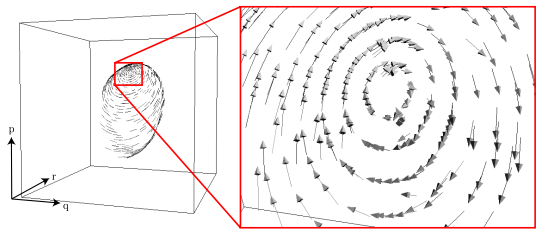


(b) 推定された3次元シーンフロー

図 7: シミュレーションにより作成した 2 次元フローと推定されたシーンフロー (平行移動)



(a) 入力2次元フロー



(b) 推定された3次元シーンフロー

図 8: シミュレーションにより作成した 2 次元フローと推定されたシーンフロー (回転運動)

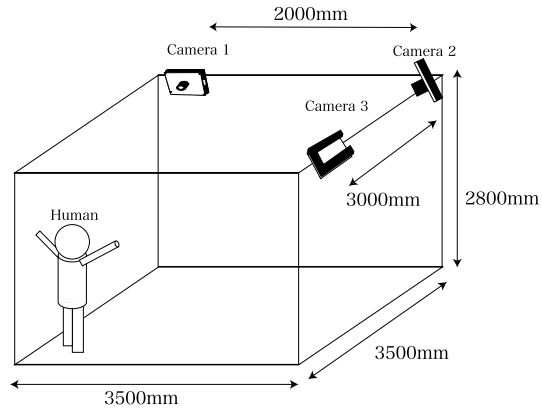


図 9: カメラ配置

せた場合と，y 軸を中心に回転させた場合の各カメラに投影した画像を入力とする．

平行移動に対する推定結果

図 7 に，球モデルが平行移動した際のオプティカルフローと，その入力から推定したシーンフローを示す．図 7(b) の右側の図は，推定されたシーンフローの一部を拡大したものである．この図から全てのフローが一定の方向を向いていることがわかる．作成した球モデルの運動は平行移動であるため，このことから，推定したシーンフローは正しいことがわかる．またシーンフローは物体表面上のみに存在する．

回転に対する推定結果

図 8 に球モデルが回転した際のオプティカルフローと，シーンフロー推定結果を示す．推定されたシーンフローの一部の拡大から，推定されたフローは球の中心を軸に回転していることがわかる．

4.2 実画像を用いた実験

実画像を用いた実験環境を，図 9 に示す．天井に設置した 3 台のカメラ間の基礎行列は，弱校正から予め求めておく．実験に使用したカメラは Point Gray Research 社の Dragonfly(IEEE-1394 カメラ) であり，各カメラ間のシャッタータイミングは同期している．実験映像は，人物が手を下から上へ上げているもので，撮影された映像の時刻 t と $t-1$ のオプティカルフローを計算し，3 次元シーンフローを推定した．オプティカルプローの計算にはブロックマッチング法を用いた．実際に推定に用いた映像の時刻 t と $t-1$，またそれに対し計算されたオプ

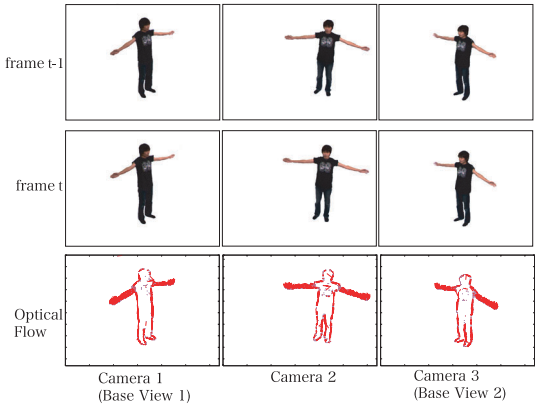


図 10: 実験に用いた映像とオプティカルフロー

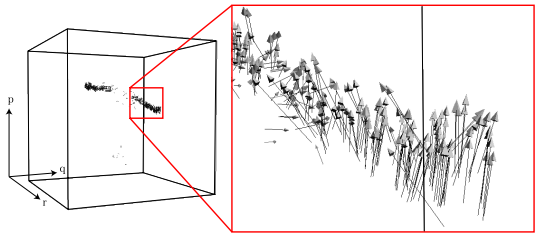


図 11: 推定された 3 次元シーンフロー

ティカルフローを図 10 に示し，推定された 3 次元シーンフローを図 11 に示す．

図 11 より，腕部のフローは上方向に向いていることがわかる．また腕の先の方が長いフローが存在していることから，3 次元シーンフローは正しく求められていることがわかる．

定量的な評価として，推定された全てのフローに対する方向分布を表 1 に示す．表中の p，q，r は Projective Grid Space における軸を，$+$，$-$ はそれぞれの軸に対する正負の方向を表している．表 1 より，カメラ 3 台の結果では q 軸 $+$ 方向 (上方向) のフローが最も多く求められていることがわかる．また，カメラの台数が 2 台の場合より，3 台の方が方向分布が小さく，より正しいフローが推定されていることがわかる．これは，カメラの台数を増やすことにより，2 次元オプティカルフロー算出のノイズ等の影響を減少させることができたためである．

表 1: シーンフローの方向の分布 [%]

	$p+$	$p-$	$q+$	$q-$	$r+$	$r-$
2 台	17.1	7.5	42.9	9.5	10.4	12.6
3 台	13.3	4.5	62.1	0.5	12.3	7.2

5 おわりに

本稿では，エピポーラ幾何を利用することにより，未校正マルチカメラにおける 3 次元シーンフロー推定手法を提案した．本提案手法では 3 次元シーンフローは，Projective Grid Space 内の各点での相対的な長さ，方向を推定することが可能であり，Projective Grid Space により定義された空間において有効であることを確認した．

今後は，実空間に対応した座標軸について 3 次元シーンフローを推定する予定である．また，推定された 3 次元シーンフローを用いてジェスチャー認識に取り組む予定である．

参考文献

[1] 青木公也, 輿水大和 ”時系列距離画像を用いた 3D フロー検出の検討” 動的画像処理実利用化ワークショップ 2005 pp.140-145, 2005

[2] Sundar Vedula, Simon Baker, Peter Rander, Robert Collins, Takeo Kanade ”Three-Dimensional Scene Flow” Proceedings of the 7th International Conference on Computer Vision, Vol. 2, September, 1999, pp. 722 - 729

[3] Sundar Vedula, Simon Baker, Takeo Kanade ”Spatio-Temporal View Interpolation” Proceedings of the 13th ACM Eurographics Workshop on Rendering, June, 2002.

[4] Hideo Saito, Takeo Kanade ”Shape Reconstruction in Projective Grid Space from Large Number of Images” IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR ’99), June, 1999.

[5] 佐藤淳 ”コンピュータビジョン -視覚の幾何学-” コロナ社

[6] 徐剛 ”写真から作る 3 次元 CG” 近代科学社

[7] S. M. Seitz, C.M. Dyer ”Photorealistic scene reconstruction by voxel Coloring” Proc. Computer Vision and Pattern Recognition 97, pages 1067-1073, 1999.

[8] S. M. Seitz and K. N. kutulakos ”Plenoptic Image Editing” Proc. 6th Int. Conf. Computer Vision, 1998.