# 画像の標本化過程を考慮した高精度かつ頑健な平面の追跡

伊藤 栄介<sup>†</sup> 岡谷 貴之<sup>†</sup> 出口光一郎<sup>†</sup>

† 東北大学大学院情報科学研究科 仙台市青葉区荒巻字青葉

E-mail: *†*{ito,okatani,kodeg}@fractal.is.tohoku.ac.jp

あらまし カメラに対して相対的に運動する平面を,画像上で安定して高精度に追跡することはコンピュータビジョ ンの中心的課題である.これまで,照明変化,オクルージョンやモーションブラーなど,追跡の精度や安定性を損な う要因がいくつか指摘され,それらへの対策が提案されてきた.本稿では,追跡対象の平面が視線に対して大きく傾 いたり,遠くに離れる際に生じる平面の見えの実効的な解像度低下もまた,追跡性能に悪影響を与えることを指摘す る.その後,画像の標本化過程をモデル化し,平面追跡のアルゴリズムにこれを組み込んで解決する方法を示す.実 験により,従来の方法に比べてより高精度かつロバストに平面の追跡が行えることを示す. キーワード 平面追跡,画像生成,標本化

1. はじめに

本稿では,平面追跡の問題について考える.これは, カメラに対して相対的に運動する平面が,画像上のどこ にどのような形で写っているかを時々刻々推定するもの で,コンピュータビジョンの分野で古くから考えられて きた基本的な問題である.この問題のための方法として, 例えば,物体表面の小さな平面パッチを画像系列上で追 跡するLucas-Kanade法([1],[2])がよく知られ,SfM 等で使用する特徴点の軌跡を得るために使われてきた. 平面が画像上に占める位置・姿勢を推定できれば,そこ から平面の3次元空間における位置・姿勢を知ることも できる.このことから,ビジュアルサーボや複合現実感 (AR)への応用も盛んである.

この平面追跡の問題に対し,これまでに多数の方法が 提案されてきた.それらは2つのカテゴリーに分類でき る.一つは,画像から点や線分,閉曲線といった原始的 な特徴を取り出し,これらに基づいて平面を追跡するも のである(代表的なものに[3]~[5]).もう一つは,その ような特徴抽出を行わず,平面の見えを直接用いて最適 化計算により目的を達成する方法である([2],[6]~[12]). こちらはLucas-Kanade 法に見られるように,SSD など の見えの近さの尺度を,主としてニュートン法を用いて 最適化し,平面の画像上の位置・形状を定める.いずれ の方法にも長所と短所があり,前者の長所は,遮蔽への 耐性などの頑健さや,必ずしも画像の連続性を仮定しな くてもよいことであり,後者の長所は,計算量の少なさ や推定精度の高さにある.

本稿では,この後者の方法について考える.これに関 する研究は数多くあり([2],[6]~[12]),それぞれ異な る問題意識や動機に基づく.まず,最適化の方法に関す る研究がある. Baker らは,最適化の計算方法を分類 し,撮影画像の微分に基づいて反復計算を行う forward 法に比べて,テンプレートの微分に基づく inverse 法の 方が,より効率的であると述べている[2].Malisらは, この反復計算で用いるヘッセ行列を,一般的な方法より も高精度かつ小さな計算量で計算し得る ESM(Efficient Second-order Method)法を提案している[6],[8].これら の他にも,最適化の目的関数となるテンプレートと撮影 画像の近さの尺度を改善しようとする研究もある.例え ば近年,SSD 等従来の尺度より優れているとして,相互 情報量 (Mutual Information)に基づく尺度の導入が提案 されている[10]~[12].

一方,最適化に関するこのような研究とは別に,平面 追跡を難しくするいくつかの要因への対処方法も研究 されている.まず,平面の運動が結果的に平面への照明 の当たり方を変化させることがよくある.この場合,平 面の見えは,幾何学的のみならず濃淡値そのものが変化 することになり,そのことを考慮しなければ,追跡性能 は低下する.Silveira らは, ESM 法の枠組みで照明変動 を扱う方法を示している[13]. Dame らも,相互情報量 を用いる方法の研究[12]において照明変動への対処を議 論している.また,カメラのシャッター速度に比して対 象平面が高速に動く場合には,モーションブラーが発生 し, これも追跡性能を低下させる. これへの対策が Jin ら [14], そして Mei ら [9] によって示されている. Park らは,モーションブラーが直線的な平面の運動によって 生成される場合, ESM 法のわずかな修正でモーション ブラーが扱えると述べている[15].

本研究では,平面追跡を難しくする要因に関するこ れらの研究と類似の立場から,照明変動やモーションブ ラーと同様に,画像の実効的な解像度低下もまた,平面 追跡の性能を低下させる要因になり得ることを指摘する. 追跡の対象とする平面が,視線方向に対して大きく傾い たり,カメラから相対的に遠ざかるとき,その平面の像 の解像度は実効的に低下する.後に示すように,この場 合にテンプレートと画像とを直接比較すると,仮に正し い姿勢パラメータを与えても,両者は一致しなくなる. このことが考慮されなければ,追跡性能が低下してしま う.典型的には,途中で追跡に失敗する頻度が増加した り,追跡が行える場合でも姿勢パラメータの推定精度が 大きく低下する現象が見られる.われわれは,画像撮影 時の標本化過程を考察し,解像度低下をモデル化するこ とで,この問題を解決あるいは緩和する方法を示す.

この背景には,次のような動機がある.平面追跡には そもそも限界があり,対象とする平面がカメラの視線方 向に対して完全に垂直を向いたり,平面がカメラから無 限遠方に遠ざかる場合,追跡は(当然ながら)不可能で ある.われわれの関心は,追跡が理論的に不可能なこの ような条件にどれだけ近いところで,追跡を行えるか にある.つまり,一定の追跡性能が維持できる限界を, 従来方法よりも広げたい.著者らの知る限り,同じ動機 を持ち,上述の問題を解決しようとした研究はこれま でにない.例えば,平面追跡の手法の評価用にテスト データ[16]が公開されており,そこでは対象平面のテク スチャの多寡や,照明変動やモーションブラーの影響は 十分意識されているものの,このような限界性能を試す データは含まれていない.

本研究のこのような視点は,ほとんどの応用で大事だ と思われるが,特にARやヒューマンインタフェースへ の応用で重要である.多様な条件下での様々なユーザに よる利用を想定すれば,平面の運動の仕方は,追跡アル ゴリズムにとって都合良く制御されたものではあり得な いからである.上述のような限界性能を向上できれば, この種の応用のユーザビリティの向上にも直結すると思 われる.

# 2. 問題の定式化と関連研究

#### 2.1 基本的な問題

現在の画像を *I*,追跡対象となるテンプレートを *I*\* と書く.テンプレート *I*\* は,*I*より小サイズの画像で, 追跡対象とする平面の矩形領域を真正面からカメラで撮 影したものに相当する.カメラと平面の相対位置姿勢が 変化し,それにともなって *I* が変化するとき,*I*上で平 面を追跡したい.

テンプレート  $\mathcal{I}^*$  とそれを撮影した画像  $\mathcal{I}$  に対し,  $\mathcal{I}^*$ の点  $\mathbf{p}^*$ を  $\mathcal{I}$ の点  $\mathbf{p}$  に写す変換(warp)を

$$\mathbf{p} = w(\mathbf{p}^*) \tag{1}$$

と表す.この変換は理想的な透視カメラを仮定すると,次の2次元射影変換で与えられる.

$$w(\mathbf{p}^*; \mathbf{H}) \propto \mathbf{H}\mathbf{p}^* \tag{2}$$

Hは2次元射影変換を表す3×3行列で, pおよび p\*

はそれぞれ  $\mathcal{I}$  および  $\mathcal{I}^*$  上の画像座標の同次表現(3次 元ベクトル)であり,  $\propto$  は両辺のベクトルが長さを除い て等しいことを表す.なお,これ以降  $\mathbf{p} \ge \mathbf{p}^*$  それぞれ, 文脈に応じて同じ画像の点についての同次座標と非同次 座標のどちらかを表すものとする.

今,追跡対象の平面の同一点の画像上の明るさが,平 面姿勢によらず変化しないとすれば,平面を撮影した画 像 $\mathcal{I}(\mathbf{p})$ とテンプレート $\mathcal{I}^*(\mathbf{p}^*)$ は次のような関係を持つ ことになる:

$$\mathcal{I}(w(\mathbf{p}^*; \mathbf{H})) = \mathcal{I}^*(\mathbf{p}^*). \tag{3}$$

このことから,式の両辺の差が最小となるように H を 定めることを考える.この最小化はニュートン法に基づ く反復計算によって行う.具体的には,現在の推定値Ĥ が与えられたとき,

$$\mathbf{H}_{o} = \underset{\mathbf{H}}{\operatorname{argmin}} \sum_{i} \left[ \mathcal{I}(w(\mathbf{p}_{i}^{*}; \hat{\mathbf{H}}\mathbf{H})) - \mathcal{I}^{*}(\mathbf{p}_{i}^{*}) \right]^{2} \quad (4)$$

なる増分 H<sub>o</sub>を求める(なおこの式で H を再定義してお リ, ĤH が (3) 式の H に相当することに注意する).そ の後, Ĥ  $\leftarrow$  ĤH<sub>o</sub> と更新し,以降これを繰り返す(Baker らの分類[17]によれば,これは forward compositional 法にあたる.)新しい画像フレームを処理するときは,前 フレームでの推定値 Ĥ をそのまま使って(つまり初期値 として),この反復計算を収束するまで繰り返す.

ここで増分 H は  $3 \times 3$  行列で 9 成分あるが, その自由 度は 8 しかなく, パラメータの取り方が重要である.よ い方法は, H が特殊線形群 SL(3) に属するようにパラ メータをとることで, 具体的には次のような指数写像を 用いて, 8 次元ベクトル x によって H = H(x) を

$$\mathbf{H}(\mathbf{x}) = \exp\left(\sum_{j=1}^{8} x_j \mathbf{G}_j\right) \tag{5}$$

と表現することである([7],[8]).ここで  $G_j$ (j = 1, ..., 8) はリー代数の生成作用素となる  $3 \times 3$  行列である(中身 は[8]を参照). H をこのように表現することで det H = 1が約束され, H のスケール倍の不定性が除去される.こ の表現に基づき,反復の毎ステップで

$$\mathbf{x}_{o} = \operatorname*{argmin}_{\mathbf{x}} \sum_{i} \left[ \mathcal{I}(\hat{\mathbf{H}}\mathbf{H}(\mathbf{x})\mathbf{p}_{i}^{*}) - \mathcal{I}^{*}(\mathbf{p}_{i}^{*}) \right]^{2}$$
(6)

を求め, $\hat{\mathbf{H}} \leftarrow \hat{\mathbf{H}}\mathbf{H}(\mathbf{x}_o)$ と更新し,これを繰り返す.

#### 2.2 関連研究

本研究ともっとも関連が深いのが,モーションブラー への対策を述べた研究 [9], [14] である.Jin ら [14] (およ び Mei ら [9])のアイデアは,モーションブラーの影響を 取り除くのに,効果および計算効率の観点から,撮影画 像のブラーを除去する(deblurring)のではなく,テン プレートにブラーを反映させることである.



図 1 標本化による画像の実効的解像度低下の例.上段:平面がカメラに正対するとき.下段:傾いたとき.左列から右列へそれぞれ,追跡中の画像,画像から求めた *I*(*w*(**p**<sup>\*</sup>; **H**)),およびテンプレート*I*<sup>\*</sup>(**p**<sup>\*</sup>)(テンプレートは上下段で同一).画像およびテンプレートの画素数はそれぞれ640×480,192×192.

モーションブラーが,画像上で均一な $\mathbf{v}(=[v_x,v_y]^{\top})$ 方向への等速ブラーであるとすると,モーションブラーを含む画像は,瞬時画像(ブラーのない画像) $\mathcal{I}^u(\mathbf{p})$ を用いて, $\mathcal{I}(\mathbf{p}) = \int_0^1 \mathcal{I}^u(\mathbf{p} - \mathbf{v}t) dt$ によって与えられる. さらに Mei らは,任意の等速ブラーが

$$\mathcal{I}_{\hat{\mathcal{H}},\mathcal{H}(\mathbf{x})}(\mathbf{p}) = \int \mathcal{I}^{u}(\mathrm{e}^{-t\hat{\mathcal{H}}}\mathrm{e}^{-t\mathcal{H}(\mathbf{x})}\mathbf{p})dt$$
(7)

と表現できることを示した.ここで $\mathcal{H}$ は,それに対応す る射影変換 H を与える (5) 式の,指数関数の内部の行列 を表す.また $\hat{\mathcal{H}}$ は, $\mathcal{H}(\mathbf{x})$ の累積(現フレームでシャッ ターが開いてから今に至るまでの積分)を表す.モー ションブラーが完全に平面の運動によって決まるとする と,(6)式は

$$\mathbf{x}_{o} = \underset{\mathbf{x}}{\operatorname{argmin}} \sum_{i} \left[ \mathcal{I}(\hat{\mathbf{H}}\mathbf{H}(\mathbf{x})\mathbf{p}_{i}^{*}) - \mathcal{I}_{\hat{\mathcal{H}}_{b},\mathcal{H}(\mathbf{x})}^{*}(\mathbf{p}_{i}^{*}) \right]^{2}$$
(8)

に置き換えられる (ここでは式を簡素化するため,フレーム間隔中シャッターが開き続けるものとした.詳しくは[9]を参照.)

なお,(8)式にしたがって計算を行えば原理的にはモー ションブラーを扱えることになる.しかしながらわれわ れの実験では,[9]で報告されているような効果が一定程 度認められるものの,モーションブラーをモデル化しな い場合に比べて反復計算の収束性が低下し,追跡の安定 性が損なわれてしまうことを経験している.このことは, (8)式において,差分をとる2つの成分がともに未知数 xに依存し,関数がより複雑な構造を持ち最適化が安定 して行えなくなるためであると思われる.

# 3. 標本化過程を考慮した平面追跡の方法

前節で述べたように,平面追跡では,真の姿勢 $H_0$ を

推定するため、その推定値  $H_1$  で撮影画像を変形した  $\mathcal{I}(H_1p^*)$  と、テンプレート  $\mathcal{I}^*(p^*)$  を比較することを行 う、図1のように、例えば平面が大きく傾いている場合 など、 $\mathcal{I}(H_1p^*)$  はその変形に応じて、実効的に解像度が 低下する、その結果、仮に  $H_1 = H_0$  であったとしても、  $\mathcal{I}(H_1p^*)$  は  $\mathcal{I}^*(p^*)$  と一致しなくなる、実際、このよう な場合には頻繁に追跡性能の低下が見られる、

ここでは,モーションプラーを扱った Jin ら [14] や Mei ら [9] らの方法同様に,このような解像度低下に合わせ てテンプレートを適切に修正することで,追跡性能の低 下を抑えることを考える.

# 3.1 解像度低下のモデル

テンプレートをどのように修正すべきかを考えるため に,撮影画像  $\mathcal{I}(\mathbf{p})$  がどのように生成され,さらにテン プレートと比較するためにいかに変形されるかを調べる. 対象とする平面パタン (=テンプレート)上の点  $\mathbf{p}^*$  が 写像  $\mathbf{p} \propto \mathbf{H}_0 \mathbf{p}^*$  によって画像の点  $\mathbf{p}$  に写るとする.ここ で  $\mathbf{H}_0$  は真の射影変換である.今,画像撮影時の空間方 向の標本化を無視したときの撮影画像を  $\mathcal{I}'(\mathbf{p})$  と書くと, これは単に

$$\mathcal{I}'(\mathbf{p}) = \mathcal{I}^*(\mathbf{H}_0^{-1}\mathbf{p}) \tag{9}$$

と与えられる.標本化を前提とすると,撮影画像は,プ レフィルタ  $f(\mathbf{p})$  をこの  $\mathcal{I}'(\mathbf{p})$  に適用した連続関数

$$\mathcal{I}''(\mathbf{p}) = \mathcal{I}'(\mathbf{p}) * f(\mathbf{p}) = \mathcal{I}^*(\mathbf{H}_0^{-1}\mathbf{p}) * f(\mathbf{p})$$
(10)

を,標本化したものとしてモデル化できる.その標本値は,撮影画像の各画素  $\mathbf{p}_j (j = 1,...)$ での値  $\mathcal{I}_j'' \equiv \mathcal{I}''(\mathbf{p}_j)$ である.これを連続領域に再構成したものが撮影画像  $\mathcal{I}(\mathbf{p})$ であると考え,再構成フィルタ $h(\mathbf{p})$ を用いて

$$\mathcal{I}(\mathbf{p}) = \sum_{j} \mathcal{I}_{j}^{\prime\prime} h(\mathbf{p} - \mathbf{p}_{j})$$
(11)

と表す.

撮影画像  $\mathcal{I}(\mathbf{p})$  は, テンプレート  $\mathcal{I}^*(\mathbf{p}^*)$  と比較すべく 変形される.この変形を与える写像を  $\mathbf{p} \propto \mathbf{H}_1 \mathbf{p}^*$  とする. (なお,追跡が高精度に行われていれば,  $\mathbf{H}_1$  は  $\mathbf{H}_0$  とな るはずである.) 変形後の画像を  $\tilde{\mathcal{I}}(\mathbf{p}^*)$  とすると,これは  $\tilde{\mathcal{I}}(\mathbf{p}^*) \equiv \mathcal{I}(\mathbf{H}_1 \mathbf{p}^*)$ によって与えられ,(11) 式より

$$\tilde{\mathcal{I}}(\mathbf{p}^*) = \mathcal{I}(\mathbf{H}_1 \mathbf{p}^*) = \sum_j \mathcal{I}_j'' h(\mathbf{H}_1 \mathbf{p}^* - \mathbf{p}_j)$$
(12)

と与えられる.これに(10)式を代入すると

$$\tilde{\mathcal{I}}(\mathbf{p}^*) = \sum_{j} \left[ \mathcal{I}^*(\mathbf{H}_0^{-1}\mathbf{p}_j) * f(\mathbf{p}_j) \right] h(\mathbf{H}_1\mathbf{p}^* - \mathbf{p}_j) \\
= \sum_{j} \left[ \mathcal{I}^*(\mathbf{p}_j^*) * f(\mathbf{H}_0\mathbf{p}_j^*) \right] h(\mathbf{H}_1\mathbf{p}^* - \mathbf{H}_0\mathbf{p}_j^*)$$
(13)

を得る.ただし2番目の等式は $\mathbf{p}_j^* \ge \mathbf{p}_j^* \equiv \mathbf{H}_0^{-1} \mathbf{p}_j$ と定義し, $\mathbf{p}_j$ を置換して得られたものである.

(13) 式は, H<sub>0</sub>の姿勢をとる平面を, H<sub>1</sub>で変形して得られる画像のモデルである.従来手法のテンプレートをこれで置き換えれば目的は果たされる.(13) 式中の H<sub>0</sub> は不明なので,まずこれを H<sub>1</sub> で近似し,その後 H<sub>1</sub>を  $\hat{H}H(x)$ で置き換えたものを $\tilde{I}_{\hat{H}H(x)}(p^*)$ と記すことにすれば,更新量 x を

$$\mathbf{x}_{o} = \operatorname*{argmin}_{\mathbf{x}} \sum_{i} \left[ \mathcal{I}(\hat{\mathbf{H}}\mathbf{H}(\mathbf{x})\mathbf{p}_{i}^{*}) - \tilde{\mathcal{I}}_{\hat{\mathbf{H}}\mathbf{H}(\mathbf{x})}(\mathbf{p}_{i}^{*}) \right]^{2}$$
(14)

によって定めればよい.しかしながら実時間性を前提と すると,(14)式の直接計算はあまり現実的でない. $\tilde{\mathcal{I}}(\mathbf{p}^*)$ は撮影画像の標本グリッド  $\mathbf{p}_j^*$ についての和を計算する し,また未知数を含むフィルタを適用した後での標本化 を要するからである.そこで,(13)式の $\tilde{\mathcal{I}}(\mathbf{p}^*)$ を,テ ンプレート $\mathcal{I}^*(\mathbf{p}^*)$ に線形フィルタ $g(\mathbf{p}^*;\hat{\mathbf{H}}\mathbf{H}(\mathbf{x}))$ を畳込 んだ

$$\widetilde{\mathcal{I}}(\mathbf{p}^*) \approx \mathcal{I}^*(\mathbf{p}^*) * g(\mathbf{p}^*; \widehat{\mathbf{H}}\mathbf{H}(\mathbf{x}))$$
(15)

のように,近似的に表すことを考える.このようにすれば,従来手法に対し計算量の増加を最小限に抑えられる. そこで,そのような線形フィルタgを求めたい.テン

プレート  $\mathcal{I}^*$ も離散データであり, 各画素  $\mathbf{p}_i^*(i=1,\ldots)$ での標本値  $\mathcal{I}_i^*$ が与えられているとすると, 連続領域の  $\mathcal{I}^*(\mathbf{p}^*)$ は, 再構成フィルタ $h(\mathbf{p}^*)$ を用いて

$$\mathcal{I}^*(\mathbf{p}^*) = \sum_i \mathcal{I}_i^* h(\mathbf{p}^* - \mathbf{p}_i^*)$$
(16)

と表せる.この式と(13)式を(14)式に当てはめて, $\mathbf{p}_i^*$ を  $\mathbf{p}_i^*$  と近似的に同一視すると

$$h(\mathbf{p}^* - \mathbf{p}_i^*) * g(\mathbf{p}^*; \tilde{\mathbf{H}}) = h(\tilde{\mathbf{H}}(\mathbf{p}^* - \mathbf{p}_i^*))$$
(17)

の関係を得る.ただし, $\hat{\mathbf{H}} \equiv \hat{\mathbf{H}}\mathbf{H}(\mathbf{x})$ であり,また上で はプレフィルタ f の効果を無視した.この式を満たす g は,次のように計算される.再構成フィルタをガウス 関数

$$h(\mathbf{p}^*) \propto \exp\left(-\frac{1}{2\sigma^2} \mathbf{p}^{*\top} \mathbf{p}^*\right)$$
 (18)

とすると,(17)式の右辺は,

$$h(\tilde{\mathbf{H}}(\mathbf{p}^* - \mathbf{p}_i^*)) \propto \\ \exp\left(-\frac{1}{2\sigma^2}(\mathbf{p}^* - \mathbf{p}_i^*)^\top \tilde{\mathbf{H}}^\top \tilde{\mathbf{H}}(\mathbf{p}^* - \mathbf{p}_i^*)\right) \quad (19)$$

と書ける. (17) 式の関係とガウス関数の畳込みの性質 ( $N(\mathbf{a}, \mathbf{A}) * N(\mathbf{b}, \mathbf{B}) = N(\mathbf{a} + \mathbf{b}, \mathbf{A} + \mathbf{B})$ )から,

$$g(\mathbf{p}^*; \tilde{\mathbf{H}}) \propto \exp\left(-\frac{1}{2}\mathbf{p}^{*\top} (\tilde{\mathbf{H}}^{-1} \tilde{\mathbf{H}}^{-\top} - \sigma^2 \mathbf{I})^{-1} \mathbf{p}^*\right)$$

を得る .

以上をまとめると,更新量の計算式は

$$\mathbf{x}_{o} = \underset{\mathbf{x}}{\operatorname{argmin}} \sum_{i} \left[ \mathcal{I}(\hat{\mathbf{H}}\mathbf{H}(\mathbf{x})\mathbf{p}_{i}^{*}) - \mathcal{I}^{*}(\mathbf{p}_{i}^{*}) * g(\mathbf{p}_{i}^{*};\hat{\mathbf{H}}\mathbf{H}(\mathbf{x})) \right]^{2}$$
(21)

となる.この式は,前述したモーションブラーの場合の (8)式同様,差分をとる2項がともに変数xに依存してお り,収束性に懸念がある.今の場合,テンプレートの修 正は,現在の平面姿勢の瞬時値にしか依存しない.しか も,xは更新量で微小であるから,そのH(x)によるテ ンプレートの修正量はわずかである.そこでテンプレー トの修正量を,次のようにxに依存しないような形に近 似する.

$$\mathbf{x}_{o} = \underset{\mathbf{x}}{\operatorname{argmin}} \sum_{i} \left[ \mathcal{I}(\hat{\mathbf{H}}\mathbf{H}(\mathbf{x})\mathbf{p}_{i}^{*}) - \mathcal{I}^{*}(\mathbf{p}_{i}^{*}) * g(\mathbf{p}_{i}^{*}; \hat{\mathbf{H}}) \right]^{2}$$
(22)

なおモーションブラーの場合,テンプレートの修正は各 フレーム間隔内での平面の微小運動の積分値に依存する から,このような近似は不可能である.

### 4. 実験結果

提案手法の有効性を調べるため,いくつかの実験を 行った.実験では,前節で述べた提案手法と従来手法を 比較した.従来手法とは,Malis らの ESM 法に照明変 化モデル[13] および Park らのモーションブラーのモデ ル[15] を加えたものである.提案手法は,さらに提案し た解像度低下のモデル((22) 式)をこれに加えたもので ある.

#### 4.1 実験方法

# 4.1.1 実験装置

追跡対象の平面が視線方向に対して大きく傾いた場合 の,(a)変形パラメータの推定精度,ならびに(b)追跡の 安定性・頑健性の2つが実用上重要である.これらを評 価するため,3枚の平面が互いに正確に90度の角度(角 度誤差は0.01度以下)をなすように機械加工した物体を 用いた(図6に全体像がある).垂直をなす2枚の平面 を選び,片方には追跡対象となる平面パタンを,もう一 方にはチェスボードパタン(11×8)を貼り付けた.こ の平面パタンを提案手法(および比較対象とする従来手 法)にて追跡し,その推定精度をチェスボードパタンを 使って求めたものと比較した.平面パタンとチェスボー ドパタンが互いに直交することから,平面パタンとが視線 に対して大きく傾くとき,チェスボードパタンが視線 に対し,その平面姿勢推定の精度はほぼ最良となる のでグランドトゥルースとするにふさわしい. その他の装置は次の通りである.3.2GHz の Core i7 を CPU に持ち, nVidia 社の GTX480 を GPU に持つ PC に,毎秒 60 フレームで 640 × 480 画素の画像を撮影する Point Grey Research 社の Grasshopper をカメラと接続 して用いた.主要な計算を GPU 上で行うことで,カメ ラの画像転送にかかる時間を除いた正味の計算時間を 3 から 4 ミリ秒程度に抑えることが出来ている.したがっ て実時間での実行に十分以上の余裕がある.

#### 4.1.2 精度の評価方法

精度の評価は,平面追跡の結果得られる2次元射影変換Hそのものではなく,それを元に計算される平面パタンの空間姿勢を用いて行った.これは,平面の姿勢の方が誤差をより直観的に理解でき,またそれが主にARを中心とする応用で最終的に必要とされるものであるからである.

平面の姿勢を求めるには,平面パタン,チェスボー ドパタンいずれの場合も,カメラの内部パラメータを あらかじめ校正しておく必要があり,ここでは Zhang の方法 [18] を利用した.平面パタンの場合,平面パタ ン上に定義した3次元座標系の点xから画像座標系の  $| (u,v) \land \mathbf{O}$ 投影を  $[u,v,1]^{\top} \propto \mathbf{K}(\mathbf{Rx}+\mathbf{t})$ と書くと,  $\mathbf{H} \propto \mathbf{K}[\mathbf{r}_1,\mathbf{r}_2,\mathbf{t}]$ を得る.ここで $\mathbf{K}$ は,校正したカメラ の内部パラメータからなる  $3 \times 3$  行列で,  $\mathbf{r}_1$  および  $\mathbf{r}_2$ は,回転行列Rの1,2列ベクトルである.この関係か ら, K<sup>-1</sup>H を計算し, その 3×3 行列の第3 列ベクトル を t の推定値とし, 第 1, 2 列ベクトル q1, q2 を用いて 新たに $3 \times 3$ 行列  $[q_1, q_2, q_1 \times q_2]$ を作成し,その特異値 分解における特異値をすべて1に置き換えて計算される 行列を, Rの推定値とした.この方法は内部パラメータ が既知であることを前提とすれば必ずしも最適なものと は言えない(注1)ものの、少なくとも2つの方法の性能を 公平に比較する目的では問題ないはずである.また,平 面追跡のアルゴリズムは一般に,Hを推定するように定 式化・最適化されていることも理由である.

一方, チェスボードパタンから平面の姿勢を推定する 方法は, チェスボードのコーナー点とその画像上の位置 の対応を使って,再投影誤差が最小になるようにチェス ボードの位置姿勢(上述のRとtにあたる)を推定し, 平面パタンがチェスボードと直交する事実を用いて,こ れを平面パタンの位置姿勢に換算した.

#### 4.2 実験結果

#### 4.2.1 精度評価実験

図3に示すような平面パタンを対象に追跡実験を行った.図2にその追跡結果を示す.テンプレートサイズは 160×160 画素である.図は,上述のように計算した平 面の姿勢Rおよびtのうち,特に提案手法と従来手法と の間で差が顕著に現れるRの時間変化を示す.図の左の 3 つのプロットは,各方法で推定した R を回転軸・角度 表現した 3 次元ベクトル  $[\omega_1, \omega_2, \omega_3]$ の各成分にあたる. 一番右のプロットは,チェスボードパタンで推定したも のとの R のずれを,回転角の大きさに直して表示したも のである.なお,すべてのプロットの単位は度(degree) である.図3は,図2中の3つの時点(第400,527,547 フレーム目)におけるスナップショットである.

この画像系列では,平面を,視線に対する角度が70度 付近から90度付近まで動かし,戻している(図2の $\omega_1$ がほぼこの角度に相当する).従来手法は,550フレー ム目手前で大きな誤差を生じた後,追跡に失敗した. 方,提案手法は全系列で追跡に成功している.そこに至 る過程で推定精度をチェスボードのそれと比較すると, 平面の傾きが小さい場合,従来手法と提案手法は似たよ うな精度を示すが,傾きが大きくなったときの振る舞い が大きく異なり,提案手法が明確に勝る.

これは図 3 を見ても確かめられる.同図 (a)-(d) には, 推定された平面姿勢を使って平面上にとった座標フレー ムを視覚化したものを表示してある.同図 (a), (b) (従 来手法)と(c), (d) (提案手法)を比べると,上段 (400 フレーム),中段 (527 フレーム)では両者にほとんど違 いはないが,下段 (547 フレーム)の結果において,従 来手法による推定に大きな誤差が含まれることが見て取 れる.また図 3 には,画像を推定パラメータで変換した ものを (f) に,解像度低下のモデルに基づいて修正した テンプレートを (g) にそれぞれ示している.両者は各フ レームである程度近く,提案した解像度低下のモデルの 妥当性を裏付ける.(h) はテンプレート修正に用いた線 形フィルタ g であるが,平面追跡の結果に応じて変化し ていることが分かる.

別な平面パタンに対して撮影した画像系列の結果を図 4に示す.最初の平面パタン(テンプレート)は比較的 滑らかな濃淡変化を示したのに対し,こちらのそれ(図 5(e))はシャープなエッジのみで構成されている点で違 いがある.図5に特定の3フレームでの画像および推定 姿勢を可視化したものを示す.結果は若干の違いは認め られるものの,図2の系列とほぼ同様である.従来手法 は400フレーム手前で追跡に失敗し、その直前での精度 が大きく低下していることが認められる.一方,提案手 法は全フレームで追跡を行えている.ただし図2の系列 とは異なり,平面の視線に対する傾きが90度に近い領域 で,姿勢の推定精度のばらつきが大きくなっている.こ れは平面パタンの濃淡構造によるものと考えられる.し かし,推定のバイアスは大きくなっていないこと,追跡 自体は継続できているため,必要ならば時間方向に平滑 化することで実行的な精度を向上させることも可能だろ う.実世界の平面に仮想物体を重ね込む AR では,追跡 対象の平面が視線に対し大きく傾いたり遠くに配置され る状況が想定される.提案手法はそのような場合に最も 効果を発揮する.例を図6に示す.従来手法で精度良く

<sup>(</sup>注1): 平面追跡の定式化の際, H の代わりに R, t を未知パラメー タとする方がより高精度だろう.



図 2 各方法 (chessboard: チェスボード, proposed:提案手法, previous: 従来手法) で推定した平面姿勢の回転成分の時間変化. 左の3つは回転を回転軸・角度表 現したベクトルの3成分.一番右はチェスボードによる推定とのずれを角度の 大きさで表したもの.従来手法 (previous) が途中で途切れているのは,追跡に 失敗したため.



(a)

図 3 図 2 の画像系列の追跡結果のスナップショット.上段,中段,下段はそれぞれ 第 400, 527, 547 の 3 フレーム . 各列 (a) から (h) は次の通り . (a) 従来手法の 追跡結果(赤い四角形)と、そこから復元した平面上の座標フレーム(赤青緑 3 色の線分で座標軸を表示,右側の大きな方),およびチェスボードパタンから 推定した平面上の座標フレーム(同様に3色で座標軸を表示,左側の小さな方. ただし見やすいように平面上で並行移動した).(b)これを拡大表示したもの. (c) 提案手法の追跡結果と復元した座標フレーム(表示は(a)と同じ).(d) その 拡大図.(e)追跡したテンプレート(上中下段で同一).(f)収束後の I(Ĥp\*). (g) 提案手法による修正後のテンプレート.(h) 修正に用いたフィルタg.

追跡できないような場合でも,提案手法は安定して高精 度に追跡でき,ARの表示品質を高く保つことができる と分かる.

# 4.2.2 追跡安定性

平面を手で自由に動かして約1000フレームからなる 画像系列を得,従来手法と提案手法をこれに適用した. 結果を図7に示す.従来手法はこのうち6フレームで追 跡に失敗したが,提案手法は全フレームで安定して追跡 を行えた.また,図8に別の画像系列に対する結果を示 す.この系列では,対象平面はカメラから徐々に遠ざか る.従来手法は,ある距離以上のところで追跡に失敗し た.一方で提案手法は,それよりはるかに遠い距離でも 正しく追跡できた.なお,この系列では,追跡対象の平 面パタンに[16]に収録されている画像を用いた.その他 の平面パタンに対する結果とともに,以上の結果は添付 資料のビデオにも収録されている.

#### 5. ま と め

本稿では、平面追跡の問題において、追跡対象となる 平面が視線に対して大きく傾いたり,遠ざかるなどする とき,撮影画像上での実効的な解像度が低下し,このこ とが追跡性能を低下させることを指摘した.このような 性能低下を防ぐため,画像撮影時の空間方向の標本化過 程,および平面追跡の最適化計算で行われる画像の変形 をモデル化し、これを考慮した最適化計算を行う方法を 示した.核となるアイデアは,追跡する平面パタン(テ ンプレート)を,追跡中の平面姿勢から作成した線形 フィルタを適用することで修正し,これを最適化に用い ることである.従来手法に対する計算量の増加は小さく, 追跡は実時間で十分行える.いくつかの実験を通じて, 提案手法は従来手法よりも明確に優れることを示した.

提案手法は,従来手法と比べて平面追跡の限界のより 近くに迫れたと考えられる. すなわち, 平面と視線が 90 度をなす場合や,平面が無限遠方にある場合を境界とす



図 4 別の画像系列に対し,各方法で推定した平面姿勢の回転成分の時間変化.図の 配置は図 2 と同じ.



図 5 図 4 の画像系列の追跡結果のスナップショット.上段,中段,下段はそれぞれ 図 4 中の第 297,360,395 フレーム.(a) 従来手法の追跡結果,(b) その拡大 図,(c) 提案手法の追跡結果,(d) その拡大図,(e) テンプレート,(f) 収束後の *I*(Ĥ**p**\*),(g) 修正後テンプレート,および(h) フィルタg.

る極限条件に,従来よりも近いところで追跡を行えるようになった.ただし提案手法では,テンプレートの修正 を線形フィルタで表すこととし,また標本化を含む画像 の生成過程をモデル化する際いくつかの近似を行うなど により,追跡性能と計算量のバランスを図っている.し たがって追跡性能のみを重視すれば,提案手法を上回る 性能を達成することは可能かもしれない.現実には実時 間性の制約からくる計算量の小ささも重要であるので, 両者をより高い次元でバランスする方法を実現できるか どうかが将来の課題となる.

### 文 献

- B. Lucas and T. Kanade: "An iterative image registration technique with an application to stereo vision", Proc. the International Joint Conference on Artificial Intelligence, pp. 674–679 (1981).
- [2] S. Baker, R. Gross and I. Matthews: "Lucas-kanade 20 years on: A unifying framework: Part 4", International Journal of Computer Vision, 56, pp. 221–255 (2004).
- [3] V. Lepetit and P. Fua: "Keypoint recognition using randomized trees", IEEE Trans. PAMI, 28, 9, pp. 1465–1479 (2006).
- [4] M. Özuysal, P. Fua and V. Lepetit: "Fast keypoint recognition in ten lines of code", Proc. CVPR (2007).
- [5] S. Holzer, S. Hinterstoisser, S. Ilic and N. Navab: "Distance transform templates for object detection and pose estimation", Proc. CVPR (2009).
- [6] E. Malis: "Improving vision-based control using ef-

ficient second-order minimization techniques", Proc. ICRA (2004).

- [7] S. Benhimane and E. Malis: "Real-time image-based tracking of planes using efficient second-order minimization", Proc. IEEE/RSJ International Conference on Intelligent Robots Systems (2004).
- [8] S. Benhimane and E. Malis: "Homography-based 2d visual tracking and servoing", Int. J. Rob. Res., 26, pp. 661–676 (2007).
- [9] C. Mei and I. D. Reid: "Modeling and generating complex motion blur for real-time tracking.", CVPR (2008).
- [10] N. Dowson and R. Bowden: "A unifying framework for mutual information methods for use in non-linear optimisation", Proc. Ninth European Conf. Computer Vision, pp. 365–378 (2006).
- [11] N. Dowson, R. Bowden and S. Member: "Mutual information for lucas-kanade tracking (milk): An inverse compositional formulation", IEEE Trans. PAMI, **30**, 1, pp. 180–185 (2008).
- [12] A. Dame and E. Marchand: "Accurate real-time tracking using mutual information", Proc. ISMAR (2010).
- [13] G. Silveira and E. Malis: "Real-time visual tracking under arbitrary illumination changes", Proc. CVPR (2007).
- [14] H. Jin, P. Favaro and R. Cipolla: "Visual tracking in the presence of motion blur", Proc. CVPR (2005).
- [15] Y. Park, V. Lepetit and W. Woo: "ESM-Blur: Handling and rendering blur in 3D tracking and augmentation", Proc. ISMAR (2009).
- [16] S. Lieberknecht, S. Benhimane, P. G. Meier and



図 6 平面の追跡結果に基づいて立方体を重畳したもの.上段:従来手法.下段:提 案手法.



図 7 ある画像系列に対して,従来手法で追跡を継続できなかった画像(上段).提 案手法は系列を通して追跡を行えた(下段).



図 8 対象平面が徐々にカメラから遠ざかる場合の結果.最上段から,従来手法の追 跡結果,提案手法の結果,比較対象となる修正後のテンプレート,および撮影 画像を変形したもの.各列は画像系列内の1フレームに対応する.

N. Navab: "A dataset and evaluation methodology for template-based tracking algorithms", Proc. ISMAR (2009).

[17] S. Baker and T. Kanade: "Limits on super-resolution and how to break them", Proceedings of Computer Vision and Pattern Recognition (2000).

[18] Z. Zhang: "A flexible new technique for camera calibration", IEEE Trans. PAMI, 22, 11, pp. 1330–1334 (2000).