

単眼動画画像からの可展面物体の 3 次元変形とその展開テクスチャの復元

柴山 裕樹[†] 小室 孝^{††} 渡辺 義浩[†] 石川 正俊[†]

[†] 東京大学大学院情報理工学系研究科 〒113-0033 東京都文京区本郷 7-3-1

^{††} 埼玉大学大学院理工学研究科 〒338-8570 埼玉県さいたま市桜区下大久保 255

E-mail: †{Hiroki_Shibayama, Yoshihiro_Watanabe, Masatoshi_Ishikawa}@ipc.i.u-tokyo.ac.jp,

††komuro@mail.saitama-u.ac.jp

あらまし 単眼動画画像から、非剛体物体の 3 次元変形とともに同変形を平面に展開した際のテクスチャも同時に推定する手法を提案する．このような手法は、モデリング、検査、書籍電子化等の応用展開において強力な技術となる可能性があると考えられる．本手法のキーポイントは、物体の可展面特性に着目して平面展開のプロセスを復元手法に導入している点である．このようなアプローチによって、複数の撮像画像から得られる情報を統合的に整合する枠組みを構築することが可能となっている．この枠組みの下、物体の変形の動きをカメラ上で観測することで、連続曲面の変形と高解像度な展開テクスチャの両者が復元されることを示す．

キーワード 画像復元, 形状復元, レジストレーション, 可展面

1. はじめに

複数枚の画像から形状を復元する技術は、模型のモデリングから建築物の復元まで多くの魅力あるアプリケーションを生み出し続けている [8, 1]．しかし、その多くの技術は、対象物体が剛体であることを前提として様々な課題を解決してきた．対象が非剛体である場合には、形状が変形するという大きな前提の違いにより、新たな手法による問題の解決が求められている．対象が非剛体である場合の形状復元として、近年、いくつかの手法が提案されているが [13, 12, 16]、剛体の場合ほど標準的な手法が確立されていない．そのため、コンピュータビジョンの分野においてチャレンジングなトピックとなっている．

非剛体の変形推定は、技術的なチャレンジング性とともにもその応用においても、産業検査からエンターテインメントの分野まで幅広く革新的な貢献をすることができると考えられる．これまでの同様の手法に関する報告では決定的な応用展開を示唆するものはなかったが、有用性の高いアプローチが提案されている．例えば、紙、缶、T シャツ、ヨットの帆などの変形を推定する評価実験が行われている．特に、材質の非伸縮性を拘束条件として利用し、変形のみを復元に焦点が置かれてきた．

ここで、従来の変形モデルで導入されている対象の非伸縮性は、対象が元々平面から変形したものであれば平らな状態に展開することができることを意味していると考えられる．この展開する操作は、復元問題において重要な意味を持つ可能性がある．具体的には、全ての観測画像に対して展開されたテクスチャ画像は 1 種類しか存在しないという条件を利用することができる．このように、観測された画像を利用して対象表面のテクスチャの

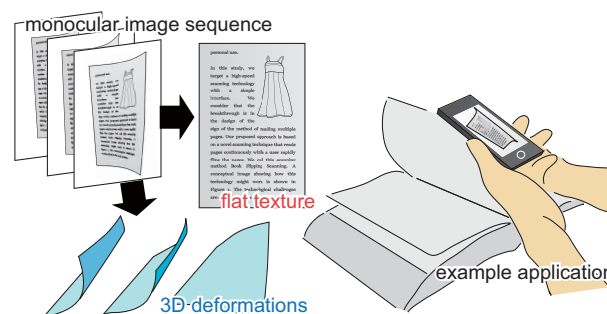


図 1 本手法では、単眼動画画像を元に対象の 3 次元形状とその展開テクスチャを復元する．新しいスタイルの書籍電子化等の応用が期待できる．

復元を合わせて実施する事には、応用とテクニカルな 2 つの側面で大きなアドバンテージがあると考えている．

例えば、我々は書籍電子化の分野への貢献を 1 つの応用として目論んでいる．近年、書籍電子化へのニーズが全世界で高まって来ている．書籍電子化には高速性と高精細が必須であるが、その両者を兼ね備えたシステムは未だに実現されていない．これまで、書籍を 1 ページずつ固定して比較的平坦な状態にされた紙面を順に電子化するスタイルが一般的であり、時間を要するという問題があった．

この問題を解決するために、書籍のページめくり動作中に動きを止めることなく電子化を行う *Book Flipping Scanning* が提唱されている [11]．この新しい方法によって、業務用途から個人用途までに応じて、いくつかのシステム構成が生み出されると考えられる．例えば、3 次元センサを用いることで形状をリアルタイムに取得し、カメラによって得られた歪んだ書籍画像を補正するタイプのシステムが提案されている [11]．しかし、3 次元セ

ンサを必要とするシステムの構成は特殊な装置を必要とする欠点がある。

本論文のキーテーマとなる、単眼カメラを用いた非剛体の変形とテクスチャの同時復元は、パーソナルな書籍電子化に新しいスタイルを提供することができると考えられる。図 1 に示すように、いつでもどこでもカメラをかざし、その前で書籍をめくればすべての情報を一瞬で高品質に取得することができるはずである。

本論文では、変形中の非剛体を 1 台のカメラによって撮像した動画像から、その変形形状と展開テクスチャを復元するアルゴリズムを提案する。提案する手法は、可展面モデルを対象物体の条件として組み込むことで、展開のユニーク性を保持する。これによって、その変形を連続曲面として復元すると同時に展開テクスチャも復元する。

本論文の構成を述べる。2. 節では、関連研究について述べる。3. 節では、今回の問題を定式化し、手法の全体構成を述べる。4. 節では、手法の構築に置いて必要となる可展面モデルを説明する。5. 節では提案する手法の具体的な計算方法を説明する。6. 節では実験結果について報告する。7. 節で論文をまとめる。

2. 関連研究

本章では、画像を用いた形状復元に関する関連研究について述べる。まず、対象物体が剛体である場合、対象やカメラを動かす事で得られる複数視点画像を用いて推定する枠組みが広く用いられている [8]。近年では、対象物体の構造に応じた先験情報を拘束条件として与えることで精度が向上する手法も提案されている。例えば、室内環境においては Manhattan-world 仮説が効果的である事が示されている [6]。このように、対象の形状に関する特徴情報を予め手法に組み込む方法が有用であると考えられる。

一方、対象が非剛体である場合は変形が発生するため、全ての撮像画像において対象の形状が同じであるという仮定は成り立たない。このため、画像の枚数を増やすことによって形状復元における不確定性を除去することが困難である。そこで重要となるのが、いかに効果的かつ最小の制約を組み込んで真の形状を復元するかという点である。

非剛体を対象とした関連研究では、画像間での変形を許容するレジストレーション手法が提案されている [5]。特に、医療画像のレジストレーションが主要な応用となっている。また、紙などの材質でレジストレーションを行う手法も報告されている [2, 7]。これらは、Radial Basis Function 等の基底関数を用いて画像内での 2 次元の自由な変形を記述している。しかし、3 次元形状が復元されない点や予め参照画像を必要とする点から、本論文の目標とは異なる。

非剛体曲面の 3 次元形状を復元するためには、その変

形に関して仮定が必要となる。そのひとつが、学習を用いた手法である [14, 3]。この種の手法は、予め大量の学習データを用意し、撮像画像とそれらのデータを比較することで正解の変形を導こうとするものである。学習データを用意することが可能である場合は効果的であると考えられる。しかし、多くのアプリケーションにおいては導入が厳しい条件となると考えられる。

この他のアプローチとして、対象の変形モデルに非伸縮性を導入する事で形状の復元を可能にした手法がある [13, 12]。ただし、この手法では、既知の変形時における観測画像が参照として予め取得されていることが前提となっている。加えて、曲面の変形はメッシュ構造によってモデル化されており、メッシュによって描かれる各パッチ内での形状は平面であると仮定されている。メッシュの密度に応じて逆問題の複雑さが増大するという問題があり、精度の高い推定は望めないと考えられる。

このように、参照画像を必要とする制約は、アプリケーションの可能性を制限すると考えられる。この問題を解決するために、参照画像を必要としない手法が提案されている [16]。この手法でも、上述の手法と同様に非伸縮性のメッシュ構造を用いた変形モデルが採用されている。同手法では、大まかな初期形状を推定した後に、最適化プロセスによって精度を上げている。しかし、メッシュの問題とともに、推定の安定化のために変形のモードを学習によって用意しておくことを前提とする制限がある。

以上のように、単眼の動画像から非剛体物体の 3 次元形状を復元する既存の手法は、強い制約や先験的信息を必要としたり、精度が十分でない問題を抱えている。また、これらの手法では変形の復元のみ焦点が当てられている。そのため、展開するために必要な推定精度が達成されているかどうかについては考察されていない。

3. 問題設定

本節では、問題設定とその定式化について述べる。入力は、 N 枚の画像群 $\{I_k\}_{k=1}^N$ である。出力は、 N 個の形状 $\{\Theta_k\}_{k=1}^N$ と 1 枚の展開テクスチャ J である。各形状 Θ_k は、各入力画像 I_k を撮像した時の対象の形状である。

これらのデータについて以下のような前提を与える。まず、入力画像は内部パラメータが既知である単眼のカメラで撮像されたものとする。また、展開テクスチャ J は対象を平らに展開した際の表面のテクスチャであるとする。そして、対象物体は伸縮や断裂を起こす事無く平面に展開することが可能であるとする。これは曲面全体でガウス曲率が常にゼロである事を意味する [4]。

次に、観測モデルについて述べる。各入力画像 I_k は、 J が表面に貼られた物体が変形した様子をカメラで撮像することで取得されたものである。この撮像されたテクスチャの変形情報が、形状 Θ_k と展開テクスチャ J の復

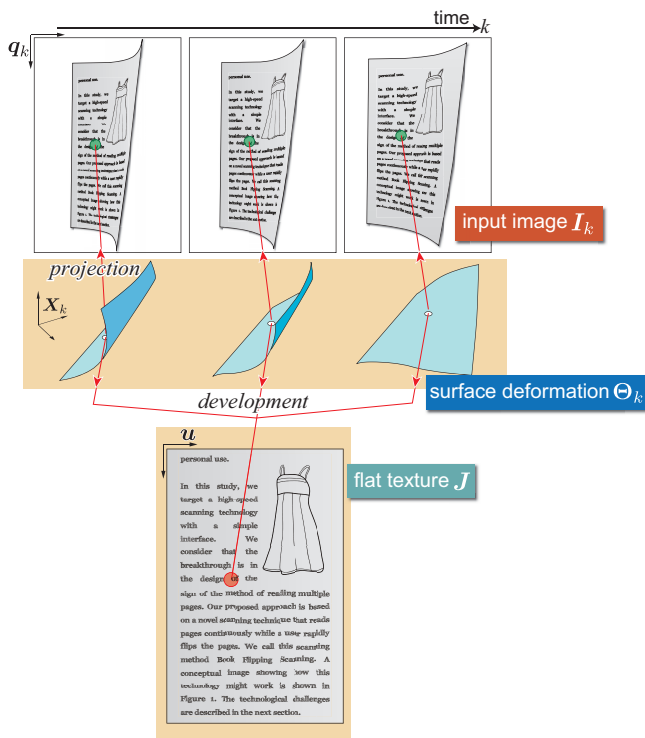


図2 問題構造と各座標系における点の対応関係．

元に用いられる．この観測の概要を図2に示した．ここで， k 番目の撮像画像と展開テクスチャの2つの座標系の変換を次式で表す．

$$u = g(q; \Theta_k). \quad (1)$$

この式は，画像上の点 q が変形と撮像を含むプロセス g を通して平坦な展開テクスチャ上の点に投影されることを意味する．今回の場合，このプロセスは変形を表すパラメータ Θ_k によってコントロールされる．

式 (1) に記述される観測モデルと展開テクスチャ J を用いて，各撮像画像 I_k は次式で表される．

$$I_k(q) = \tilde{I}_k(q; \Theta_k, J) = \sum_{\|q - \tilde{q}_i\| < \epsilon} B_i J(g(\tilde{q}_i; \Theta_k)). \quad (2)$$

ここで， $I_k(q)$ と $J(u)$ は各画素における輝度値を表している．この式は， k 番目の入力画像 I_k がレンズぼけとサブサンプリングのプロセスを介して，展開テクスチャ J から生成されることを表している．ここで， B_i はぼけ関数のカーネルである．

なお，対象物体は可展面の制約の下で自由に变形することができ，我々の手法ではこのような自由な変形を許容するが，全ての観測画像における対象物体の展開テクスチャは1種類である．このように，展開操作を導入することによって，撮像された複数枚の画像同士の整合性

を得ることができる．つまり，入力画像の枚数を増やすことによって条件数を増やし，変形を推定する逆問題を解ける環境を構築することができる．さらに，画像枚数の増加に従って展開テクスチャ1枚に用いられる情報量が増し，より高解像度な展開テクスチャが得られることが期待できる．加えて，この高解像度の展開テクスチャ J を再帰的に表面形状の推定にも用いる事で，形状推定の精度の向上にも寄与できると考えられる．

このアプローチに基づいて，単眼のカメラによる複数枚の撮像画像から形状と展開テクスチャを，以下の問題を解くことで復元する．

$$\min_{\{\Theta_k\}_{k=1}^N, J} \sum_k \sum_q \|I_k(q) - \tilde{I}_k(q; \Theta_k, J)\|^2 + f(J). \quad (3)$$

ここで，第2項は展開テクスチャ J に関する正則化項である．この項は必ずしも必要であるわけではなく，今回の実験においてはこの項を用いることなく高解像度な復元を実現することができた．

具体的な定式化では，展開可能な非剛体を少ないパラメータを用いて記述し，問題の規模を小さくすることが必要である．本論文では，微分幾何学分野で用いられる可展面の概念を導入する．このモデルを用いることで，対象の変形は少ないパラメータで記述できるだけでなく連続曲面として表現できるとともに，一意に展開できる性質を組み込むことが可能となる．変形をこの他のモデルで表した場合，展開テクスチャの一意性が保証されず，安定した推定精度を保つことが困難になると考えられる．可展面モデルの詳細については4.節で述べる．

また，式 (3) のように表される今回の問題では，変形 $\{\Theta_k\}_{k=1}^N$ と展開テクスチャ J の両者の同時推定が前提となっている．しかし，このような同時推定は複雑な逆問題となり，解を得ることが難しい．そこで本手法では，曲面形状と展開テクスチャを交互に推定する．また，非線形最適化によって解の探索を行うため，推定には初期解を与える必要がある．以上から，本手法は次のように構成される．

(1) 初期値推定 このステップでは，初期の推定パラメータを求める．本手法では，曲面を複数のパッチに分割し，隣接するパッチとの接続を補償しながらホモグラフィ推定を行うことで初期値を推定する．詳細は5.1節で述べる．

(2) 展開テクスチャ復元 全ての入力画像 $\{I_k\}_{k=1}^N$ を用いて展開テクスチャ J を復元する．このステップにおいて，変形 $\{\Theta_k\}_{k=1}^N$ は一定であるとする．詳細は5.2節で述べる．

(3) 3次元形状復元 1枚の入力画像 I_k と展開テクスチャ J から，対応する3次元形状 Θ_k を復元する．このステップにおいて，展開テクスチャ J は固定されている．詳細については5.3節で述べる．

(4) 繰り返し (2) 及び (3) を解が収束するまで繰り返す。

4. 可展面モデル

可展面は，曲線 $\gamma(s)$ とベクトル $\xi(s)$ によって記述される．ここで γ は準線と呼ばれる．また， ξ は母線と呼ばれる．可展面では曲面全体においてガウス曲率は常にゼロとなっている [4]．この条件から，曲面を平面に展開することが可能となる．可展面の例を図 3 に示した．

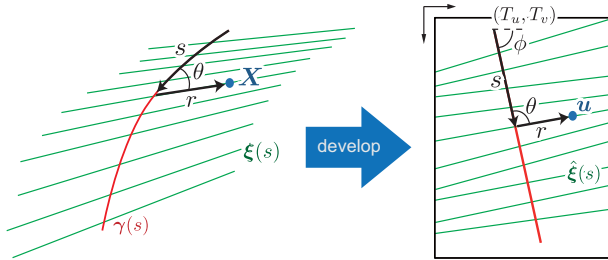


図 3 可展面とその平面展開．

曲面上の 3 次元点 X は 2 つのパラメータ (s, r) を用いて以下の式で表される．

$$X(s, r) = \gamma(s) + r \cdot \xi(s) \quad (4)$$

$$\xi(s) = \frac{\gamma''(s) \times \gamma'''(s)}{\|\gamma''(s) \times \gamma'''(s)\|} \quad (5)$$

パラメータ (s, r) はそれぞれ準線と母線上での長さを表している．曲面上の 3 次元点 X は，準線上を s 進み，その後同点を通る母線上を r 進んだ点であると言い換えることができる．

曲面がこのように定義されることから，曲線 γ を決定すれば曲面全体の形状を一意的に定めることができる．従って，曲面を復元するためには準線を規定するパラメータを推定すれば良い．

今回は曲線 γ として，次のような Bezier 曲線を用いた．

$$B(t) = \sum_{i=0}^{n-1} p_i \cdot \binom{n-1}{i} \cdot t^i (1-t)^{n-1-i} \quad (6)$$

ここで点群 $\{p_i\}_{i=0}^{n-1}$ は Bezier 曲線の制御点を表している．制御点を決定することで，可展面の変形は一意的に定まる．また，次の式によって Bezier 曲線のパラメータ t は弧長 s に変換される．

$$s(t) = \int_0^t \left\| \frac{dB(t')}{dt'} \right\| dt' \quad (7)$$

5. 3次元変形と展開テクスチャ復元

5.1 初期値推定

まず，曲面を記述する初期パラメータ $\{\Theta_k\}_{k=1}^N$ を推定する．ここでは，対象曲面を小領域に分割し，各領域の形状を平面と仮定して法線ベクトルの分布を求め，その分布から全体の形状を推定する手法を用いる．法線ベクトルは，各領域について推定されたホモグラフィを分解することで取得することができる．このような領域パッチベースの手法では，隣接するパッチ同士の接続を補償することが重要であると考えられる．

このような手法は，従来の単眼カメラによる非剛体曲面の変形復元でも導入されている [16]．同手法では，画像面上で領域に区切り，領域ごとにホモグラフィ推定を行い，得られたホモグラフィから領域の法線ベクトルを求めている．ここで，隣接する領域は，入力画像上で領域の端の部分において互いに重複している．これによって，重複領域に存在する画像点が同領域の全ての平面上の点であることを仮定し，隣接する領域の接続性を保証している．ただし，この仮定は近似であり，現実とは異なる．つまり，実際は画像上のオーバーラップした部分の点に対応する 3 次元空間上の点のすべてが 2 つのパッチ平面に属するとは限らない．

この他にも，曲面を小領域に分割し，ホモグラフィ推定をベースとしてその形状を推定する手法が報告されている [15]．この手法は，キャリブレーションされたステレオカメラによって撮像された 2 枚の画像から，曲面形状を推定することを目的としている．この手法では，2 枚の画像同士で対象の形状が同一であるので前提が異なるが，その定式化は今回の場合にも有効であると考えられる．同手法では，各パッチの法線ベクトルを直接求めるのではなく，パッチの各頂点の奥行きを用いて法線情報を記述している．この方法を用いる事で，パッチ同士の接続性を効率的に記述することができると考えられる．

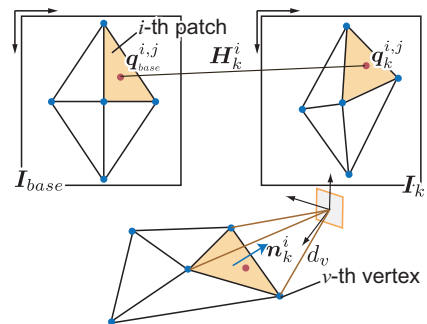


図 4 パッチベースの初期値推定手法．

本手法での初期値推定の概要を図 4 に示した．まず，1 枚の画像を基準画像 I_{base} として選び， I_{base} をパッチに分割する．そして，基準画像と他のもう 1 枚の画像 I_k との間で対応点の探索を行う．対応点の探索には，SIFT [9]

や連続画像の特徴点追跡を用いる。次に、得られた対応点を用いて N_p 個のパッチのホモグラフィ $\{H_k^i\}_{i=1}^{N_p}$ を推定する。この時、ホモグラフィ H_k^i の推定には基準画像での i 番目のパッチ領域における対応点のみが用いられる。さらに、推定された各 H_k^i を、回転 R_k^i 、平行移動 t_k^i 、法線ベクトル n_k^i に分割する [10, 18]。

ここまで各領域の推定は個別に行われており、得られた法線ベクトル $\{n_k^i\}_{i=1}^{N_p}$ も隣接領域間の接続を考慮していない。そこで、接続性を補償した初期値を得るために、次の様な最適化プロセスを適用する。

$$\min_{\{d_k^l\}_{l=1}^{N_d}, \{R_k^i\}_{i=1}^{N_p}, \{t_k^i\}_{i=1}^{N_p}} \sum_i \sum_j \|q_{base}^{i,j} - H_k^i q_k^{i,j}\|^2 \quad (8)$$

ここで、2 画像 (I_{base}, I_k) の i 番目の領域での j 番目の対応点は $(q_{base}^{i,j}, q_k^{i,j})$ で表される。また、 d_k^l は l 番目の頂点の奥行きである。 i 番目のパッチに属する 3 つ以上の頂点によって多角形のパッチが形成され、その法線ベクトル n_k^i を求める。

この最小化問題を解く事で、曲面のおおまかな形状を表す 3 次元点群が得られる。取得された各入力画像に対する 3 次元点群を用いて、初期の曲面形状 $\{\Theta_k\}_{k=1}^N$ を推定することが出来る。ここで、曲面形状を記述するパラメータセット Θ_k は次のように書ける。

$$\Theta_k = \begin{bmatrix} P_k \\ \psi_k \end{bmatrix}. \quad (9)$$

ここで、 $P_k = [p_0, \dots, p_{n-1}]$ は式 (6) で用いた曲面形状を記述する制御点群である。 P_k は [17] で提案された手法を用いて推定することが出来る。この手法は、3 次元点群から可展面の曲面を推定するものである。また、 ψ_k は、展開テクスチャの座標系を合わせるためのレジストレーション変数である。詳細は式 (12) で定義する。

以上の手法によって推定されたパラメータ P_k を用いて、各観測画像 I_k からそれぞれ展開テクスチャを得ることが出来る。これらの展開テクスチャから 1 枚を基準画像として選び、他の画像の座標系が一致するように $\{\psi_k\}_{k=1}^N$ を推定する。この初期パラメータを用いて、次節で述べる展開テクスチャ推定を行う。

5.2 展開テクスチャ復元

展開テクスチャ復元では、変形パラメータ $\{\Theta_k\}_{k=1}^N$ を固定し、定数として扱う。従って、目的的最適化問題は次のように書ける。

$$\min_J \sum_k \sum_q \|I_k(q) - \tilde{I}_k(q; \Theta_k, J)\|^2 + f(J). \quad (10)$$

この問題を解くためには、式 (1) や式 (2) で定義され

た関数 g を定義する必要がある。この関数を定義することで、カメラ座標系と展開テクスチャの座標系との間の対応をとることができる。

この対応は 2 つの段階に分けて考えることができる。まず 1 段階目が、カメラ画像上の 2 次元点 q と曲面上の 3 次元点 X との対応である。これは次のように定式化することができる。

$$\min_{\eta} d(KX(\eta), q)^2. \quad (11)$$

ここで、 K はカメラの内部パラメータである。 $d(x, y)$ は 2 点間の距離を表している。曲面上の 3 次元点 X は、パラメータ $\eta = [t, r]$ によって表される。ここで、 t は Bezier 曲線のパラメータであり、 r は母線の長さである。弧長 s ではなく、Bezier 曲線のパラメータ t を直接用いた方が効率的であると考えられる。

2 段階目は、曲面上の 3 次元点 X と展開テクスチャ上の 2 次元点 u との間の対応である。可展面の場合、面上での距離は曲面上でも展開後の平面上でも一定であるので、次の式が成り立つ。

$$u = T + s\rho(\phi) + r\hat{\xi}(s). \quad (12)$$

可展面では、準線 γ は測地線となるため、展開された面上では直線に変換される。また、準線と母線のなす角は展開前後に変わらない。 $\psi = [T, \phi]^t$ は、展開テクスチャの座標系に合わせるためのレジストレーション変数である。展開テクスチャ復元においてはこれらのパラメータも一定とした。同変数は形状復元の際に推定される。

以上の関係によって、カメラ座標系と展開テクスチャの座標系との対応が構築される。この対応を用いることで、式 (10) で表される最適化問題を解くことが可能となる。

5.3 3次元形状復元

曲面の復元では、展開テクスチャ J は定数とする。ここでは、各入力画像 I_k に対応する各形状 Θ_k を個別に復元する。復元は次の式を解くことで行われる。

$$\min_{\Theta_k} \sum_q \|I_k(q) - \tilde{I}_k(q; \Theta_k, J)\|^2. \quad (13)$$

この式を解くためには、非線形最適化を伴うため、解の探索のための反復計算を行う必要がある。しかし、曲面形状のパラメータが変化すると式 (1) のマッピング関数 g も変化するため、反復の度に同関数を導出し直す必要がある。前節の展開テクスチャ復元では同関数の計算は一度で良かったが、今回の場合、これは非常に長い計算時間を必要としてしまうために現実的ではない。

そこで、輝度値を評価する代わりに、入力画像 I_k と推定された展開テクスチャ J との間で対応する特徴点同士の距離の誤差を評価することで、形状の復元を行う手法を提案する。展開テクスチャ復元の場合に対して評価関数の形が異なる問題があるが、式を持つ意味は本質的に等しいと考えられる。実際の実験における同時推定では解が収束したことから、この定式化による問題は無いと考えられる。

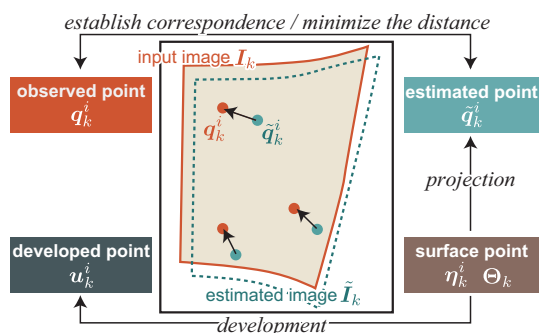


図5 3次元形状復元の概要。対応する特徴点同士の距離の誤差を最小化することで復元を行う。

図5に形状復元の概要を示す。まず、推定された変形パラメータ Θ_k を用いて推定撮像画像 \tilde{I}_k を作成する。画像の作成はこの時点での形状復元においてこの一度きりである。次に、入力画像 I_k と作成された推定撮像画像 \tilde{I}_k の間で特徴点の対応を取得する。今回の場合、SIFTが特徴点抽出と対応点取得に対して効果的であると考えられる。

このように画像間の対応点を求めることで、カメラ上で撮像された点が実際に展開テクスチャ上のどの点に対応するのかを求めることができる。展開テクスチャは形状復元では固定されているので、展開テクスチャ上の各特徴点の位置も変化することはない。以上の処理によって取得された、形状復元における観測データ V_k^i を次式に示す。

$$V_k^i = \begin{bmatrix} q_k^i \\ u_k^i \end{bmatrix}. \quad (14)$$

観測データ V_k^i は q_k^i と u_k^i の2点から構成される。 q_k^i は入力画像 I_k 上での i 番目の特徴点の座標を表し、 u_k^i は展開テクスチャ J 上の対応する点の座標を表している。

提案手法では、カメラ画像同士の対応点同士の距離の誤差を最小化することで、正しい曲面形状を復元する。ただし、同問題は、十分な量の入力画像が得られない場合に局所最適解に落ちる可能性があり、適切な解を得ることができないため、正則化項を加える必要がある。今回は展開画像上における紙面の4隅の位置を拘束条件として加えた。提案する推定手法を下式に示す。

$$\min_{\Theta_k, \{\eta_k^i\}_{i=1}^{N_c}} \sum_i \|V_k^i - \tilde{V}_k^i(\Theta_k, \eta_k^i)\|^2 + \sum_j \|C_k^j - \tilde{C}_k^j(\Theta_k, \eta_k^j)\|^2. \quad (15)$$

Θ_k は式(9)で定義された推定形状パラメータである。また、式(15)で定義されたもう1つの推定パラメータ $\{\eta_k^i\}_{i=1}^{N_c}$ は N_c 個の点の曲面上の位置を表している。この Θ_k と $\{\eta_k^i\}_{i=1}^{N_c}$ のパラメータセットを用いて、観測データの推定値 \tilde{V}_k^i を求めることができる。ここで、4隅の点に対応する観測データ V_k を $C_k^j, j=0, 1, 2, 3$ と書いた。観測データの内、撮像画像上の点 \tilde{q}_k^i は曲面上の点をカメラに投影することで求められる。また、同様に展開テクスチャ上の点 \tilde{u}_k^i は式(12)から求められる。

本手法では、式(15)の第一項はカメラ画像同士の対応点同士の距離の誤差と展開テクスチャ上での対応点同士の距離の誤差を表しており、これらを最小化することで、曲面形状を復元すると同時に展開テクスチャが一意である条件も満たすことができる。また、第二項の正則化項は、カメラ画像上の4隅の点に対応する展開画像上の点と展開画像の4隅の座標との間の距離を表している。

6. 実験

まず、提案手法の評価のためにシミュレーションを用いた実験を行った。この実験では、予め用意した展開テクスチャを曲面上にマッピングし、仮想的なカメラに投影する事で得られる仮想の観測画像を用いた。解像度は 300×300 とし、全部で6枚の入力画像を作成した。また、推定プロセスにおける形状パラメータの初期値には正解値にランダムに10%のノイズを与えたものを用いた。



図6 入力画像(シミュレーション)。

入力画像を図6に示す。推定を4回繰り返した後の推定曲面形状を図7に示す。図の赤線と青線はそれぞれ推定曲面上での準線と母線を表している。また、推定された展開テクスチャを図8に示す。展開テクスチャの解像度は 300×300 とした。

さらに、推定結果と真値を比較し、数値誤差を評価した。ここでは図6の右下の形状について、各反復時の曲面の平均誤差を図9に示す。同図で緑の形状が初期の曲面を表している。同様に、青の形状が4回反復後の曲面

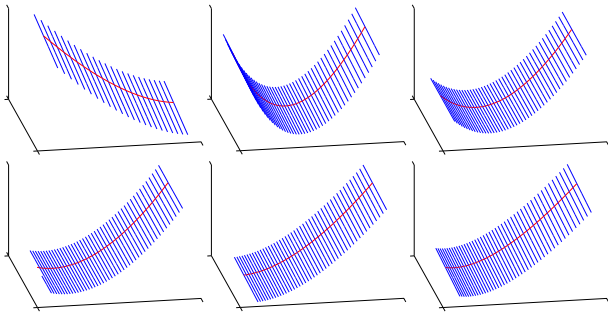


図7 推定形状．正解パラメータにランダムに10%のノイズを加えたものを初期値として用いた．

高精度な推定を達成しており，復元結果は実際のアプリケーションに対しても実用性が高いことが期待される．



図11 入力画像（変形する紙面を撮像）．



図8 右：推定展開テクスチャ．左：入力画像例．

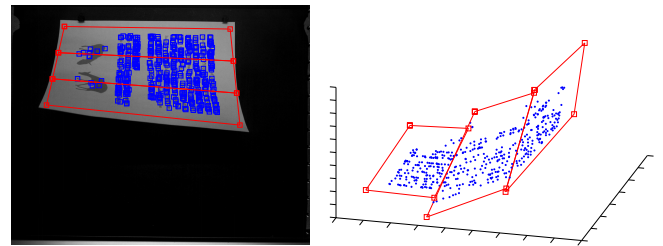


図12 初期値推定結果．

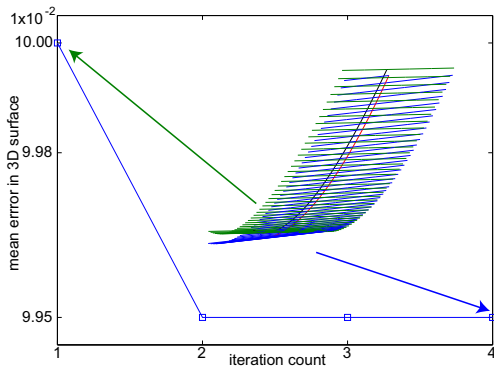


図9 形状復元の誤差．

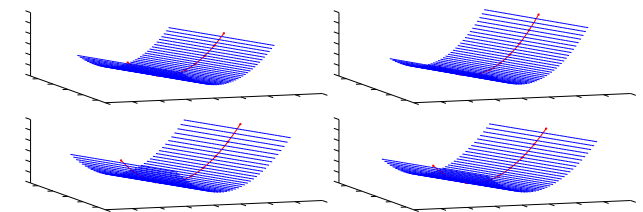


図13 推定形状．

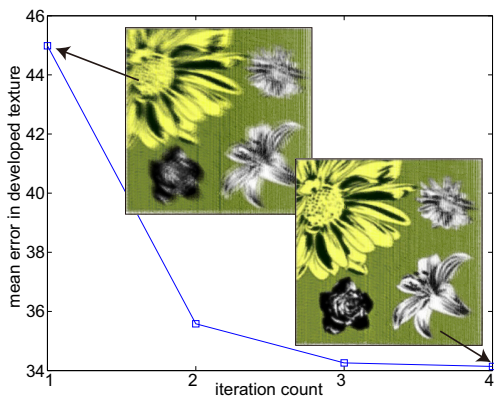


図10 展開テクスチャ復元の誤差．

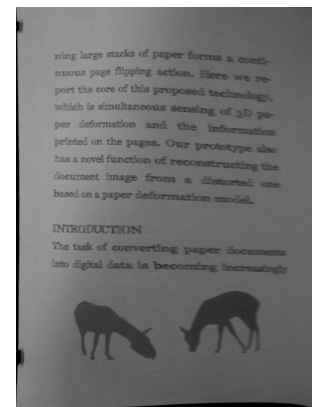


図14 推定展開テクスチャ．

である．図9に示した誤差値に対して，紙面はおよそ $20 \times 20 \times 20$ の立方体に収まる程度の大きさであり，十分小さな誤差に収まっていることが分かる．また，展開テクスチャ上のそれぞれの画素における輝度値の平均誤差を図10に示す．これらの結果に示されているように，反復とともに精度が向上していることが分かる．本手法は

次に，実際の画像を用いた実験について述べる．この実験では，1枚の紙が変形している様子を撮像した動画画像を入力として用いた．撮像画像の解像度は 1280×600 であった．また，4枚の画像を入力として用いた．入力画像例を図11に示す．初期値推定の結果の例を図12に示す．推定された形状例を図13に示す．また，推定され

た展開テクスチャを図 14 に示す。なお，展開テクスチャの解像度は 1129×1500 とした。実験の結果，実際の画像においても本手法が有効であることがわかった。

7. ま と め

本論文では，単眼カメラで撮像された連続画像を用いて，変形する物体の 3 次元形状と 1 枚の展開テクスチャを推定する新しい手法を提案した。提案手法は，曲面を平面に展開する操作の導入によって可能となった観測画像間の整合性を軸として，復元を達成している。このような，平面展開を導入した新しいアプローチは，理論的にも実用的にも有用性が高い。提案手法では，曲面は少ないパラメータセットで記述するとともに効率的な解法が構築されており，ロバストな推定が可能となっている。実験の結果，本手法によって曲面形状と展開テクスチャを復元できることを示した。

次の課題として，複数の非剛体曲面が存在する様な場合においても，今回のアプローチを用いて形状や展開テクスチャを復元することを考えている。また，自己遮蔽などが発生した状況に対する解決策も考える必要がある。計算時間については，並列処理モジュールを用いることで処理速度の向上が期待できる。例えば，形状復元に関しては形状ごとに個別に推定を行えるため，並列化が可能である。また，展開テクスチャの復元においても並列化が可能であると考えられる。これらを解決していくことで，今回提案した新しい復元手法は新しい検査システムや書籍電子化などの具体的な様々な応用に大きな貢献を果たすことができると期待される。

謝 辞

本研究の一部は，戦略情報通信研究開発推進制度 (SCOPE) の下で行われた。

文 献

- [1] S. Agarwal, N. Snavely, I. Simon, S. M. Seitz, and R. Szeliski. Building rome in a day. In *Proceedings of the International Conference on Computer Vision*, pages 72–79, 2009.
- [2] A. Bartoli and A. Zisserman. Direct estimation of non-rigid registrations. In *Proceedings of the British Machine Vision*, 2004.
- [3] V. Blanz and T. Vetter. A morphable model for the synthesis of 3D faces. In *Proceedings of the International Conference and Exhibition on Computer Graphics and Interactive Techniques*, pages 187–194, 1999.
- [4] M. P. D. Carmo. *Differential Geometry of Curves and Surfaces*. Prentice Hall, 1976.
- [5] W. R. Crum, T. Hartkens, and D. L. G. Hill. Non-rigid image registration: theory and practice. *The British Journal of Radiology*, 77:140–153, 2004.
- [6] Y. Furukawa, B. Curless, S. M. Seitz, and R. Szeliski. Reconstructing building interiors from images. In *Proceedings of the International Conference on Computer Vision*, pages 80–87, 2009.
- [7] V. Gay-Bellile, A. Bartoli, and P. Sayd. Direct estimation of

nonrigid registrations with image-based self-occlusion reasoning. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 32(1):87–104, 2010.

- [8] R. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, 2000.
- [9] D. G. Lowe. Object recognition from local scale-invariant features. In *Proceedings of the International Conference on Computer Vision*, pages 1150–1157, 2007.
- [10] E. Malis and V. Manuel. Deeper understanding of the homography decomposition for vision-based control. In *Research Report INRIA Sophia-Antipolis*, 2005.
- [11] T. Nakashima, Y. Watanabe, T. Komuro, and M. Ishikawa. Book flipping scanning. In *Adjunct Proceedings of the Symposium on User Interface Software and Technology*, pages 79–80, 2009.
- [12] M. Perriollat, R. Hartley, and A. Bartoli. Monocular template-based reconstruction of smooth and inextensible surfaces. In *Proceedings of the Asian Conference on Computer Vision*, 2010.
- [13] M. Salzmann, J. Pilet, S. Ilic, and P. Fua. Surface deformation models for non-rigid 3D shape recovery. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 29(8):1481–1487, August 2007.
- [14] M. Salzmann, R. Urtaun, and P. Fua. Local deformation models for monocular 3D shape recovery. In *Proceedings of the Computer Vision and Pattern Recognition*, 2008.
- [15] S. Sugimoto and M. Okutomi. A direct and efficient method for piecewise-planar surface reconstruction from stereo images. In *Proceedings of the International Conference on Computer Vision and Pattern Recognition*, 2007.
- [16] A. Varol, M. Salzmann, E. Tola, and P. Fua. Template-free monocular reconstruction of deformable surfaces. In *Proceedings of the International Conference on Computer Vision*, pages 1811–1818, 2009.
- [17] Y. Watanabe, T. Nakashima, T. Komuro, and M. Ishikawa. Estimation of non-rigid surface deformation using developable surface model. In *Proceedings of the International Conference on Pattern Recognition*, pages 197–200, 2010.
- [18] Z. Zhang and A. Hanson. 3D reconstruction based on homography mapping. In *Proceedings of the ARPA Image Understanding Workshop*, pages 1007–1012, 1996.