

再投影誤差最小化による建物画像列からの 全周形状の一括復元

椋木 雅之†, 右田 剛史‡, 天野 晃#, 青山正人†, 浅田 尚紀†
† 広島市立大学 情報科学部 知能情報システム工学科
‡ 岡山大学 工学部 情報工学科
京都大学大学院 情報学研究科 システム科学専攻

あらまし 本論文では、屋外建物の全周形状復元問題に対して、簡便な初期値設定を行い、非線形最適化により直接解を求めることを提案する。画像列からの形状復元問題は、一般に非線形最適化問題となるが、非線形最適化では、初期値設定により局所解におちいり、正解が得られないことが問題とされていた。本論文で利用する初期値設定は、全周復元の問題設定を反映したものであり、実場面において比較的広く適用できる。また、局所解におちいった場合に、解を修正することにより正しい形状復元を行う手法も提案する。最後に、画像数数百枚、特徴点数数百点の規模のシミュレーションおよび実画像を用いた実験によって、本手法の有効性を示す。

Batch Structure from Motion for Entire 3D Shape Recovery from a Building Image Sequence Based on Reprojection Error Minimization

M. Mukunoki†, T. Migita‡, A. Amano#, M. Aoyama† and N. Asada†
† Department of Intelligent Systems, Hiroshima City Univ.
‡ Faculty of Information Science, Okayama Univ.
Graduate School of Informatics, Kyoto Univ.

Abstract We propose a procedure for recovering entire 3D shape from a building image sequence with non-linear optimization method. In this case, the initial values of the unknown variables in the optimization process is crucial. Thus the direct application of the non-linear optimization process has not been fully used in shape recovery. In this paper, first, we investigate an initialization method, which is very simple and relatively generic to the building shape recovery problem. Then, we propose a method for restoring the optimization process from a local minimum. Experimental results using real and synthetic image have shown that our algorithm stably recover 3D shapes.

1 はじめに

物体の三次元形状を取得することは、CG や VR のモデル生成や物体測定、3次元デジタルアーカイブなど多くの分野で必要とされている。小型の物体に対しては、レーザレンジセンサやステレオ画像計測器など、比較的簡便な装置が開発されており、形状取得が容易になりつつある。しかし、屋外の建物の全周形状のように、計測環境を制御できない場所にある大型物体に対しては、多くの人手をかけたり、大がかりな装置を用いた測量が必要であった。このような問題に対し、物体を撮影した画像列からその物体の3次元形状を復元する、画像列からの形状復元の手法は、大がかりな装置なしに形状復元が行えるため、盛んに研究されている。

画像列からの形状復元では、物体の形状は、物体上に設定した特徴点群によって表現される。形状復元問題は、各画像上の特徴点の2次元座標から、その特徴点の3次元座標と撮影したカメラの位置・姿勢を同時に求める問題として定式化できる。その解法の基本的な原理は、「画像上の特徴点と、その特徴

点の推定した3次元座標を画像上に再投影した点との誤差が最小となるようパラメータ推定を行う」という再投影誤差最小化の考えに基づいている。

再投影誤差最小化は、正規分布の誤差を仮定した下での、最尤推定の考えに従った定式化である [3]。カメラの投影モデルとして透視投影を仮定すると、この再投影誤差最小化は、非線形最適化問題となる [1]。この非線形最適化問題を直接解けば、最尤推定の意味で最適性が保証されるため、高精度な解が得られる。また、最適化の目的関数は、再投影誤差最小化の考えをそのまま利用したものとなる。その意味で、非線形最適化による解法は、単純かつ有効な方法といえる。

このような非線形解法は、従来バンドル調整 (bundle adjustment) と呼ばれ、他の手法で得られた復元結果の精度を向上するための「補正処理」として利用されてきた [3, 4]。これは、この非線形解法には高精度な形状復元を行えるという利点がある一方、以下の問題があるためと考えられる。

- 解析的な解法がなく、繰り返し処理による解の反復改良が必要なため、処理速度が遅い。

- 反復改良の初期値の与え方により、求めたい正解以外の局所解が得られる場合がある。

このうち、処理速度については、高速化を目指した研究がなされており [2, 12], 計算機性能の向上と相まって、現状ではそれほど問題にならなくなっている。局所解については、対象とする問題のクラスを限定し、適切な初期値設定を行うことで回避可能であると考えられる。

そこで、本稿では、建物の全周形状を画像列から復元するというクラスの問題に対し、全周復元を想定した簡便な初期値設定方法を用い、非線形最適化による解法を直接適用することを提案する。また、局所解に陥った場合に、解の修正操作により正解を得るための手法についても提案する。初期値設定においては、その適用可能範囲が問題となる。本稿では、設定した初期値の適用範囲を評価するために、全特徴点のうち、画像列中の画像1枚に現れる点の割合を表す平均出現率に着目する。この値が一定以上であれば、ほとんどの場合、設定した初期値から正解に収束することを実験により示す。これにより、従来のバンドル調整で前提としていた逐次的な初期値推定が、全周復元の問題に対してはほとんど不要であることを示す。

以下、2. では、本稿で扱う全周形状の復元問題と形状復元に関する従来手法について述べる。3. では、非線形最適化による解法について説明する。4. では本稿で提案する非線形最適化の簡便な初期値設定法と、局所解に陥った場合の解の修正法について述べる。5. では、提案した初期値設定法の適用可能範囲について、実験を行う。6. では結論と今後の課題について述べる。

2 画像列からの全周形状復元

2.1 全周復元問題

本稿では特に、対象物体である建物の近くから周囲 360 度を撮影した画像列を用いて、形状を一括復元する問題を扱う。図 1 に示すように、撮影モデルとして、「円周撮影」と「疑似広角撮影」の 2 つを導入する。図 1(a) の「円周撮影」は、各カメラ位置で 1 枚の画像を撮影しながら円周上を一方向に移動するものである。一方、図 1(b) の「疑似広角撮影」は、円周撮影と同様の移動をしながら、各カメラ位置でカメラ方向を変化させて複数の撮影を行うものであり、より広い画角範囲の情報を得ることができる。疑似広角撮影の有効性は、実験の節で示す。

実際の撮影では、正確な円周上を移動するのではなく、建物の方向を向きながら全周を撮影するが、

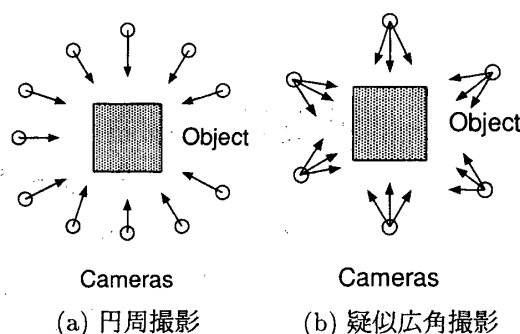


図 1: 撮影モデル

そのような場合も含めて「円周撮影」「疑似広角撮影」と呼ぶものとする。

実在する建物の撮影では、建物自身による隠蔽のため、1 枚の画像からは最大でも全周の半分程度の情報しか得られない。更に、建物周辺に木立や他の建造物が存在するため、遠方から撮影できず、建物に近接した位置から撮影した画像列を使うことが多い。そのため、全特徴点のうち 1 枚の画像内に現れる点の数は少なくなる。画像内に現れない特徴点のことを隠れ点とよぶ。本論文で扱う全周復元問題は、隠れ点が多い問題設定である。このような場合、復元すべき情報に対して、利用できる情報が減るため、一般に復元が困難となる。

2.2 従来研究

画像列からの形状復元については、従来から多数の研究がなされている。

多数の画像を一括して扱い、形状復元する手法として、因子分解法 [5] がある。因子分解法は、観測した特徴点をならべた計測行列が、特徴点の 3 次元位置成分とカメラの位置・姿勢成分に分離可能なことを利用している。以下の 2 つの制約が満たされている場合、特異値分解を用いて、一意の解を高速に求めることができる。

(a) 全特徴点が全画像で観測されている。

(b) カメラの投影がアフィン投影で近似できる。

再投影誤差最小化の観点からみると、この手法は、上記の制約のもとでの近似解法とみなすことができる。カメラと物体間の距離が十分離れている場合、この制約は満たされるが、全周復元では一般にこれらの制約は満たされない。

これらの制約を緩和する事によって、より広範な画像列を対象に、安定な形状復元を目指す研究も多数行われている。射影歪みに関しては、文献 [6] では、画像列を 2 画像ずつの組合せに分解し、それぞれからエピポーラ幾何を用いて得られる奥行き情報

(射影的奥行き)を統合することにより、対処した。文献 [7]では、数枚の画像を同時に扱い、射影的奥行きを非線形最適化により別途求めて、因子分解法に導入する手法を提案した。隠れ点に関しても、隠れ点を許容する特異値分解の手法を用いるもの [5, 9]、隠れ点のない要素のみからなる計測行列を構成して解を求めるもの [8]、特徴点に重みづけすることにより、隠れ点の影響を低減するもの [10] など様々な手法が提案されている。このような手法は、それぞれの問題について、対処法を提供しているが、隠れ点の多さや射影歪みの強さにより十分な精度が得られない、隠れ点と射影歪みに同時に対処することが難しいなど、全周復元への適用には不向きである。

別の方法として、2または3枚の画像からの形状復元結果を逐次的に繋ぎ合わせて近似解を得て、それを初期値として非線形最適化により、再投影誤差を最小化する方法 [4] もある。2または3枚の画像からの形状復元では、カメラの位置・姿勢の相対関係を線形化可能であり、探索を要しない閉じた形式の解が与えられる。これらの結果を全周に渡って統合し、非線形最適化の初期値とすることにより、一意の解が得られる線形解法の利点と、高精度な解が得られる非線形手法の利点を兼ね備えた復元が行えることから、大規模な画像列を用いた形状復元を行う場合の1つの現実的方法といえる。線形化を行わない場合でも、画像を逐次的に増やしながらか非線形解法を適用することにより、初期値推定の問題を解決する試みがなされている [13]。しかし、初期値推定に用いる部分画像列からの復元は全情報を用いた復元に比べ、解の精度が低くなる傾向がある。特に2枚の画像を使う場合、射影幾何学に基づく代数的操作が誤差を増幅することも指摘されている。さらに、多段階の手法では、手順も複雑となり、誤差の蓄積により、適切な初期値推定が行えないことも考えられる。本論文では、全周復元を仮定した場合、このような複雑な初期値設定は必ずしも必要ないことを実験により示す。

本論文と同様に、再投影誤差の最小化を非線形最適化により直接解く方法もいくつか報告されている。文献 [1]は、近似解も要さず観測データのみから初期値を構成する方法によって非線形最適化により形状復元が行えることを示した。しかし、全周形状の一括復元は対象としておらず、また、具体的な問題での初期値の適用可能性についても、十分に検討されていない。

3 再投影誤差最小化による形状復元

3.1 定式化

物体は P 個の特徴点で表現される。各特徴点は、物体とカメラに共通のある適当な座標系 (世界座標系) 上にあり、固有の3次元座標 s_p ($p = 1 \dots P$) を持つ。これらの点は最大 F 枚の画像上に投影され観測される。第 f 画像 ($f = 1 \dots F$) 上で観測された、第 p 特徴点の座標を u_{fp} とする。特徴点 p が画像 f で観測されていない場合 u_{fp} は定義されない。

各画像を撮影したカメラの位置・姿勢は、世界座標系における3次元の回転と並進 (R_f, t_f) で表される。カメラ位置は直接のパラメータではなく、 $-R_f^T t_f$ の式で与えられる。投影モデルは透視投影とし、焦点距離は全画像を通じて共通の既知の値 l とする。則ち、ズームの異なる画像は扱わない。

これらのパラメータを用いて、形状復元は次のように定式化される [1]。

$$\arg \min_{\mathbf{x}} E(\mathbf{x}) \quad (1)$$

$$E(\mathbf{x}) := \sum_{(f,p)} |P[R_f s_p + t_f] - u_{fp}|^2$$

$$\mathbf{x} := (s_p, q_f, t_f \text{ を並べたベクトル})$$

$$\text{ただし、} P(x, y, z)^T := \frac{1}{z}(x, y)^T$$

ここで R_f は自由度3の回転行列であり、 q_f は R_f のパラメータを並べたベクトルである。また、特徴点の画像座標とその画像間での対応 u_{fp} は予め与えられているものとする。従って、画像間での対応問題は解決済みである。

この式は、推定した形状 s_p を推定したカメラパラメータ R_f, t_f で投影した場合の投影座標 $P[\cdot]$ と、観測値 u_{fp} との誤差 (これを再投影誤差と呼ぶ) の2次ノルムの総和が最小となることを要求するものであり、観測座標に等方性正規分布の誤差を仮定した場合の最尤推定を与える式である。投影座標と観測値との誤差を再投影誤差と呼ぶ。この式を最小化することにより、最尤推定の意味で最適な形状復元が行えることになる。

各特徴点は、2枚以上の画像で観測されていなければ、その位置を推定することができない。各カメラの位置・姿勢を求めるには、2画像以上で対応する特徴点が6点以上、または、3画像以上で対応する特徴点が4点以上観測されている必要がある。信頼できる推定のためには、これらの下限値よりも多くの画像で多くの特徴点が観測される必要がある。

3.2 最適化アルゴリズム

全周復元の問題に対して、式(1)を目的関数とする非線形最適化により、解を求める。一般に、非線形最適化では、繰り返し処理により解を逐次更新する反復法が用いられる。計算法としては、Levenberg-Marquardt法などの任意の汎用非線形最適化手法が利用できるが、問題の規模が大きい場合、より計算量の少ない、前処理付き共役勾配法 [2] が適している。

前処理付き共役勾配法は、以下の式で表される。

$$\begin{aligned} x_0 &= \text{初期値} \\ d_0 &= C^{-1}g_0 \\ x_{i+1} &= x_i - \alpha_i d_i \\ d_{i+1} &= C^{-1}g_{i+1} + \beta_i d_i \\ \text{ただし } \alpha_i &= \arg \min_{\alpha} E(x_i - \alpha d_i) \\ \beta_i &= \frac{(C^{-1}g_{i+1})^T (g_{i+1} - g_i)}{g_i^T C^{-1}g_i} \\ g_i &= \nabla E(x_i) \end{aligned}$$

ここで、 C は前処理行列であり、ヘッセ行列 $(\partial^2 E / \partial x_i \partial x_j)$ の対角ブロック成分のみを残した帯行列を用い、16反復毎に再計算を行う。本稿で扱った問題では、画像数 $F = 30 \sim 180$ 、特徴点数 $P = 100 \sim 1200$ 、未知数が $1000 \sim 5000$ 程度となるが、この場合、計算時間はPentiumIII 1GHzのPCを用いて0.1~10秒程度となる。

この非線形解法では、式(1)を目的関数として直接最適化を行っており、透視投影に関する近似は含まれていない。また、式(1)では、隠れ点の存在は問題とならない。さらに、入力画像列の利用できる特徴点すべての情報を一括して用いていることから、局所解に陥らなければ、高精度な解が得られる。

4 初期値と収束可能性

4.1 注視回転による初期値

建物画像列からの全周復元では、カメラはほぼ水平な平面上に位置し、建物方向を向いて移動しながら撮影が行われると考えられる。また、カメラは建物に対して、原則として1方向に移動しており、画像列は撮影順序に従って与えられるものとする。この撮影状況を物体上に設定した注目点を中心にカメラが等速円運動をしながら物体を観測する注視回転によりモデル化する(図2)。ここで、カメラ位置と光軸は全て同一平面(水平面)上にあるものとする。

初期値として必須の成分には、カメラの位置、姿勢情報と、復元対象の形状情報の3種類がある。こ

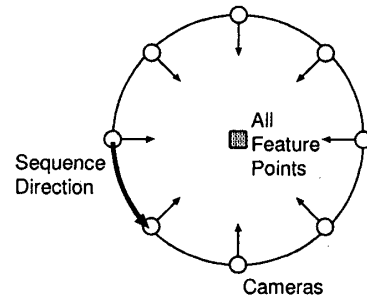


図2: 注視回転による初期値

のうち、形状情報は一般に未知であり、初期値として何らかの仮定を導入することは現実的でない。そこで、形状の初期値としては全特徴点が原点に位置するという縮退した状態を用いる。カメラの位置、姿勢情報については、注視回転を仮定して設定する。この初期値(以下、注視回転による初期値と呼ぶ)を数式で表すと次式ようになる。

$$\begin{cases} R_f = R_{\omega_f}^z, t_f = (0, 0, r)^T & \text{for all } f \\ s_p = (0, 0, 0) & \text{for all } p \end{cases} \quad (2)$$

ここで、 r はカメラの円軌道の半径であるが、計算空間の3次元座標系には物理単位が与えられていないため、 $r = 1$ として良い。従って、撮影状況に応じて与えるべきパラメータは、全回転量 ωF (または、画像間の平均回転移動量 ω)のみとなる。全周復元では、撮影時のカメラ軌跡の方向(ωF の符号)を考慮して、 $\omega F = 360$ 度または $\omega F = -360$ 度のいずれかを選ぶ。

このような初期値設定により、どの範囲の問題を解くことができるかが重要であるが、非線形性のため、解析的な検討は困難である。経験的に得られた傾向を以下に列挙する。初期値設定の有効性については、5.の実験で評価する。

- 対象物体の形状にはほとんど依存せず、1点に縮退した初期値から正解を得ることができる。
- 全回転量の大きさにもそれほど影響を受けない。実際の回転量が270度であっても360度を初期値として正解がえられる。
- カメラ軌跡の方向は重要で、撮影時と異なる設定をすると、局所解に陥る可能性が高い。
- 図1(d)の疑似広角撮影では、各撮影点で仰角や方位角が大幅に変動するが、これに対しても、式(2)を初期値として正解が得られる。また、疑似広角撮影の方が円周撮影よりも安定して解が得られる。
- 対象物体に近接して撮影するほど局所解に陥りやすく、復元が失敗する可能性が高くなる。

なお、全周より狭い範囲の部分形状復元に対して、撮影状況にあわせて ωF を定めることにより、同じ枠組で復元が行える。この際、全周復元の場合と同様に、カメラ軌跡の方向は重要であるが、全回転量の大きさにはそれほど影響を受けない傾向がある。

4.2 平均出現率

ある与えられた形状復元問題が最適化アルゴリズムにより解けるかどうかは、初期値の設定法だけでなく、カメラ-物体間距離や、物体形状の凹凸、利用する画像数、特徴点数など様々な要因に依存する。これらを総合的に表す指標として、特徴点の平均出現率を用いる。平均出現率は、各特徴点が全画像のうち何%で観測されるかを表す量で、文献 [5] で用いられている「観測行列の fill-ratio」と等価なものである。各特徴点が平均して全画像の $r\%$ で観測される場合、平均出現率は $r\%$ とする。これは、各画像に平均して全特徴点のうち $r\%$ が観測されると考えることもできる。すなわち、平均出現率は、画像間での特徴点の対応、あるいは観測された点の割合を表している。

カメラの画角が一定のままカメラ-物体間距離が縮まれば、画像内の特徴点数は減少し、平均出現率は低下する。同様に、物体形状の凹凸が多ければ、隠れ点を多く生じ、平均出現率は低下する。このように、平均出現率は、与えられた2次元情報からの形状復元の困難さを反映すると考えられる。

4.3 解の修正

注視回転による初期値は、多くの場合で正解を得ることができる。しかし、隠れ点が非常に多い画像列の場合には、局所解に陥ることがある。このような場合にも、式 (1) の枠組の中で、特徴点の削除と再追加を繰り返すことにより、解を修正して、局所解から抜け出すことができる場合がある。そのため以下のような手順をとる。

- (i) 局所解の原因となった特徴点を除去する。
- (ii) 残った特徴点のみを用いて、形状と運動を再推定する。
- (iii) 除去した特徴点を再追加する。

以下、この手順について説明する。

式 (1) の目的関数において総和の範囲を限定した次式を導入する。

$$E(\mathbf{x}, L) = \sum_{(f,p) \in S(L)} |P[R_f \mathbf{s}_p + \mathbf{t}_f] - \mathbf{u}_{fp}|^2$$

ここで、 $S(L)$ は考慮すべき添字の集合で、 L は制御パラメータである。特徴点 p が L 枚以上の画像に現れる場合のみ $S(L)$ に (f, p) を含めるものとする。

理論上は、特徴点は2画像以上で観測されていれば($L=2$)復元可能であるが、少数の画像でしか現れない特徴点が誤った位置に推定されることによって、カメラ位置や姿勢が誤って推定され、局所解となる場合が多い。手順 (i) では、 L を増加させることによって、局所解の原因となっている特徴点(群)を一旦除去する。手順 (ii) では、残った点を用いて、形状復元を行うことにより、カメラ位置、姿勢を推定しなおす。多数の画像で観測されている特徴点のみを用いると、カメラ位置、姿勢や形状を正しく推定できる傾向がある。手順 (iii) では、手順 (ii) で推定された正しい値をもとにして、除去した特徴点群を含めた全ての特徴点を用いて式 (1) を最小化する。これによって、局所解から脱出する。一回で解を修正しきれない場合、 L の増加、減少を複数回繰り返すことで修正を行う。

一般に、局所解では、 $E(\mathbf{x}, L)$ の値が大きいのに対して、局所解の原因となっている特徴点が除去できると、 $E(\mathbf{x}, L)$ の値が急激に減少する。この性質を使って、 $E(\mathbf{x}, L)$ の閾値処理によって、除去する特徴点を選択するパラメータ L を自動的に決定する。自動処理により局所解から抜け出せない場合は、手動によりパラメータ L を調整する。

この手順では、局所解に陥った際に、 L を減らすことによって平均出現率の高い特徴点のみからなる部分問題を解いている。この部分問題は元の問題よりも容易な問題となっているため、正解が得られる可能性が高い。その結果を利用するため、局所解から抜け出すことができる解釈できる。この手順のみを用いて必ずしも局所解から抜け出せるわけではないが、パラメータ L を増加させても十分な数の特徴点が残る画像列では、正解が得られると期待できる。

5 実験

本節では、全周復元問題に対して、提案した初期値設定がどの程度有効であるか確かめるため、数値的に生成した疑似建造物を用いた模擬実験、並びに、実画像での実験結果を示す。

5.1 シミュレーション画像による復元

凹凸のある建物を想定し、球を変形させた次式の曲面を対象物体とする。ただし、 $x-z$ 平面を地面と考える。

$$\begin{bmatrix} [1 - \nu + \mu \cos(\rho(\varphi - \pi/4))] \cdot \cos \theta \cos \varphi \\ [1 - \tau] \cdot \sin \theta \\ [1 - \nu + \mu \cos(\rho(\varphi - \pi/4))] \cdot \cos \theta \sin \varphi \end{bmatrix}$$

$$|\theta| \leq \pi/2, \quad |\varphi| \leq \pi$$

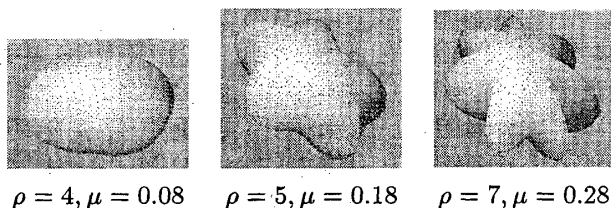


図 3: 対象物体 ($\nu = \tau = 0.2$)

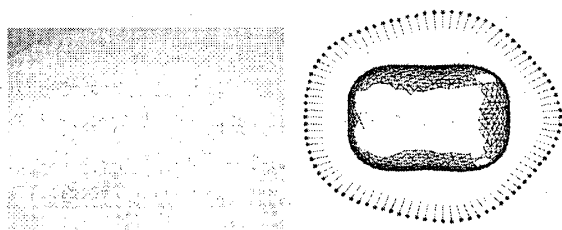


図 4: 4角形を撮影した例 (最も近い距離)

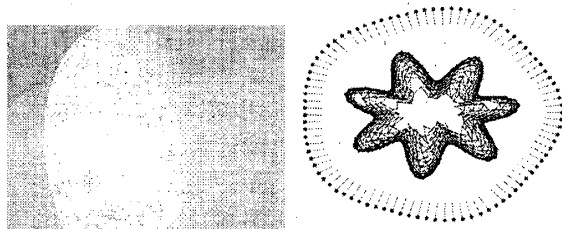


図 5: 7角形を撮影した例 (最も遠い距離)

この曲面は θ - ϕ 表現を変形したもので、 ν, τ は縦・横・奥行の比を決めるパラメータ、 ρ は凹凸の数、 μ は凹凸の深さを表す。 $\nu = \tau = \mu = 0$ のときこの曲面は球となる。球面上で点密度がほぼ一定となるように特徴点を設定し、上の式に従って変形する。更に、近接点を結んだ 3 角形群で曲面を構成し、特徴点の隠蔽判定に用いる。

真円からある程度離れたカメラ軌跡を生成するために、視点の決定にも上の曲面の式を用いる。対象物体とは異なるパラメータを用いて、対象物体よりも大きな曲面を作り、 $\theta = 0$ で切った断面の曲線を上下/左右に平行移動させたものに沿って視点を設定する。光軸方向は、この曲線の接線によって決めるものとする。このため、光軸方向は必ずしも物体中心を向いていない。

形状、画像数等を以下のように変化させ、合計 135 種類のデータについて復元の成功率を調べた。

形状: 図 3 に示す 4 角形、5 角形、7 角形を用いた。角数の増大と共に物体の凹凸も増加させたため、4 角形の場合ほとんど凹凸はないが、7 角形の場合、物体は深く凹んでおり隠蔽が起こりやすい。隠蔽が多

いと特徴点对応は少なくなり、復元困難となる。

特徴点数: 各形状につき 300, 600, 1200 と変化させた。特徴点数が少ないほど隣接画像での対応数が減り、復元困難となる。

画像数: 物体周囲を一回転する 5 角形 ($\rho = 5, \mu = 0.03, \nu = 0.1$) の軌跡に沿って 30, 60, 90ヶ所から、22.5 度見上げる角度で画像を作成した。各画像の大きさは 600x450pixel で、焦点距離は画像の横幅と同じ (横画角が約 53°) とした。画像数が少ないほど隣接画像間の対応が少なく復元困難となる。

カメラ-物体間距離: 距離が遠い場合復元は容易で、カメラが物体に近づくにつれ隣接画像間での対応が減り復元困難となる。ここでは、復元成功率が大きく変化するような 5 段階の距離を用いた。図 4, 5 に画像および物体とカメラ軌跡の上面図の例を示す。5 段階のうち最も遠い距離の場合でも、各画像には物体の一部しか入らないことがわかる。上面図において周辺部の線付の点はカメラ位置と光軸を表す。その内側の点の集合は物体の概形を表す特徴点である。なお、画像列の 2 画像以上に現れない特徴点は表示されていないため、特徴点は主に物体辺縁部に集中している。

復元には、全回転量が 360 度の注視回転を仮定した初期値を用いた。観測に誤差は加えず、復元結果の再投影誤差 RMS が 0.1pixel 未満の場合を復元成功と判定した。

図 6 は円周撮影法において、横軸を平均出現率、縦軸をカメラ-物体間距離としたときの復元成功/修正による成功/失敗の分布を示す。図 6 より平均出現率が 15% 以上の全ての場合で注視回転の初期値から正解が得られた。カメラ-物体間距離が近くなるにつれて、平均出現率も低下してくるが、8% 程度でも、解の修正により正解が得られる場合があった。また平均出現率が下がると復元成功率も下がる傾向があり、与えられた問題が解けるかどうかを表す指標として平均出現率が適切であることも確認できる。単純な比較はできないが、Tomasi and Kanade は、彼らの手法を 16% の問題に適用している [5]。平均出現率の観点から本手法は、全周復元の比較的広い範囲の問題に対して、設定した初期値から非線形最適化により正解を得られることが示された。

同様に図 7 に疑似広角撮影法における復元成功/修正による成功/失敗の分布を示す。画像数は 6×10 (10ヶ所から各 6 枚)、 6×20 , 6×30 の 3 通り、他のパラメータは前記実験と同じである。同一位置から複数の画像を撮影しているが、定式化上は、この情報は全く利用しておらず、初期値も円周撮影と同じものを用いた。復元結果の上面図の一例を図 8 に示す。平均出現率が 11% 以上の全ての場合に注視回

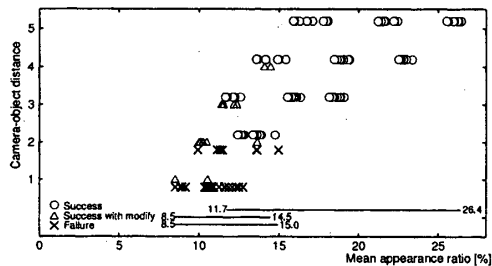


図 6: 復元成功/修正による成功/失敗の分布 (円周撮影法)

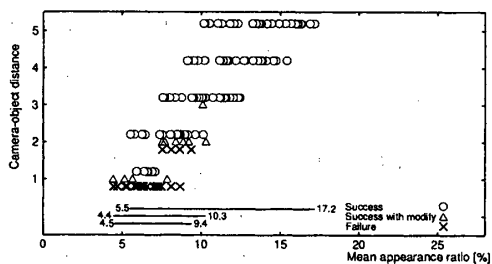


図 7: 復元成功/修正による成功/失敗の分布 (疑似広角撮影法)

転の初期値から正解が得られた。また、平均出現率が5%程度でも、解の修正により正解が得られる場合があった。円周撮影よりも疑似広角撮影の方が、設定した初期値からより離れた問題設定となっているが、結果的には、平均出現率が低く、カメラ-物体間距離が近い場合でも、円周撮影の場合より成功率が高い。これは、同じ画像対応数でも、疑似広角撮影の方がカメラ位置の空間的分布が広がるために生じる利点と考えられる。

なお、ここでは用いなかったが、疑似広角撮影では、同一位置から撮影した複数の画像のカメラ位置に対して、カメラ位置が同一であることを利用して定式化することにより、推定すべきパラメータが減少し、更に容易かつ高精度に解を求めることが期待できる。このように、全周復元に対しては、円周撮影よりも疑似広角撮影の方が、より適切であると考えられる。

カメラの移動速度や距離に偏りのある場合(図9)に対しても良好な結果が得られることを確認しており、設定した初期値は汎用性の高いものと言える。

5.2 実画像を用いた復元

ここでは、実画像列からの復元結果を2例示す。

図10に、原爆ドームを撮影した画像29枚中の2枚と復元結果の上面図、を示す。この実験例では、建物から離れた位置から撮影を行っており、平均出現率 r は17%と高いため、注視回転の初期値から

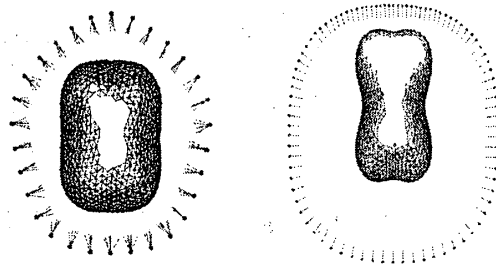


図 8: 疑似広角撮影法による復元結果
図 9: 速度や距離に偏りがある場合の復元結果

局所解に陥ることなく正解が得られ、解の修正は不要であった。

図11の例では、建物に近い位置から撮影しており、建物のごく一部しか撮影されていない画像を含む。そのため、平均出現率は6%と低い。また、近距離から広範囲を撮影するため、光軸を上下左右に大きく動かしながら撮影しており、疑似広角撮影となっている。この例では注視回転の初期値からでは部分的に正しくない結果が得られた。その後、解の修正を行ったところ、図11の復元結果が得られた。図右上方の建物の垂直な角が良好に復元されているなど、建物全体が良好に復元できていることが確認できる。

その他、100枚以上からなる全周実画像列5例においても良好な復元結果が得られた。

6 おわりに

本稿では、注視回転という簡単なモデルを用いて撮影位置と姿勢を設定し、建物画像列からの全周形状復元問題を再投影誤差最小化の考えに基づき、非線形最適化により直接解く方法を提案した。さらに、この方法だけでは正解が得られない場合を考慮して、解の修正により局所解から抜け出し正解を得る方法について述べた。

実験の結果、提案した初期値は、その仮定となる注視回転から離れた画像列に対しても、正解を得ることが可能であった。このとき、従来のバンドル調整法を適用する際に行われていた特徴点位置の初期推定処理は全く必要なかった。また、シミュレーション実験から、全周復元のような隠れ点の多い問題に対しても、全周の10~15%程度が撮影された画像からなる画像列から形状復元が行えることが確認できた。このことは、簡便な初期値設定法が、全周復元の広い範囲の問題に適用可能であり、従来のバンドル調整法が前提としていた複雑な初期値推定は、大幅に簡略化できる可能性があることを示している。



図 10: 原画像の例と復元結果原爆ドーム ($P = 122, F = 29, r = 17\%$)



図 11: 原画像の一部と復元結果講堂 ($P = 189, F = 280, r = 6\%$)

形状復元の代表的解法として知られる因子分解法との比較実験は行っていないが、全周復元問題は、射影歪みが強く隠れ点が多いことから、本来の因子分解法の定式化では、精度の高い解が得られないことは明らかである。因子分解法を改良したより複雑な手法の中には、同じクラスの問題に適用可能なものもあると考えられるが、本手法は、一貫した処理により、近似誤差なしに最尤推定の意味で最適な解が得られる点で、利点がある。

復元においては、円周撮影より、疑似広角撮影の方が、平均出現率が低くても形状復元が成功する傾向があり、全周復元に適していた。これは、後者の撮影法の方が、各特徴点を広範囲から撮影し、拘束が強いためである。本論文では、平均出現率を問題の難しさを表す指標として用いたが、カメラ間の位置関係など他の指標も導入することにより、形状復元が正しく行える条件を明確にし、より良い撮影方法を検討していくことが今後の課題である。

参考文献

- [1] R. Szeliski and S. B. Kang: "Recovering 3D Shape and Motion from Image Streams using Non-Linear Least Squares," CVPR, pp.752-753, (1993).
- [2] 右田剛史, 天野晃, 浅田尚紀: "3次元形状・運動復元のための高速非線形最適化計算法," 情報処理学会論文誌, vol.44, no.11, pp.2864-2872, (2003).
- [3] B. Triggs, P.F. McLauchlan, R.I.Hartley, A. W. Fitzgibbon: "Bundle Adjustment - A Modern Synthesis," Vision Algorithms: Theory and Practice, LNCS 1883, Springer, pp.298-375 (2000).
- [4] M. Pollefeys, F. Verbiest, and L. V. Gool: "Surviving Dominant Planes in Uncalibrated Structure and Motion Recovery," 7th ECCV, pp. II-837-851, (2002).
- [5] C. Tomasi, T. Kanade: "Shape and Motion from Image Streams under Orthography: a Factorization Method," IJCV 9(2), pp. 137-154, (1992).
- [6] P. Strum, and B. Triggs: "A Factorization Based Algorithm for Multi-Frame Projective Structure and Motion," ECCV, pp. 709-720, (1996).
- [7] 植芝 俊夫, 富田 文明: "奥行きパラメータの逐次推定による多視点透視投影画像のための因子分解法," 信学論, vol.J81-D-II, no.8, pp. 1718-1726, (1998).
- [8] 藤木 淳, 蔵田 武志: "計量アフィン射影モデルの逐次型因子分解法" 信学論, vol.J84-D-II, no.8, pp. 1663-1673, (2001).
- [9] D. Jacobs: "Linear fitting with missing data: Application to structure-from-motion and to characterizing intensity images," CVPR, pp. 206-212, (1997).
- [10] H. Aanas, R. Fisker, K. Aström and J. Carstensen: "Robust Factorization," PAMI 24-9, pp. 1215-1225, (2002).
- [11] A. R. Dick, P. H. S. Torr, and R. Cipolla: "A Bayesian Estimation of Building Shape Using MCMC," 7th ECCV, pp. II-852-866, (2002).
- [12] P. F. McLauchlan: "A Batch/Recursive Algorithm for 3D Scene Reconstruction," CVPR, pp.738-743, (2000).
- [13] 佐藤 智和, 神原 誠之, 横矢 直和, 竹村 治雄: "マルチベースラインステレオ法を利用した動画からの屋外環境の三次元モデル化," パーチャルリアリティ学会論文誌, Vol.7, No.2, pp.275-282, (2002).