GPU を用いた単眼時系列画像からのリアルタイムで密な 3 D復元

関　晃仁† Oliver Woodford†† 岡田　隆三† Bjorn Stenger†† Roberto Cipolla††

† 株式会社 東芝 研究開発センター

†† Toshiba Research Europe Limited Cambridge Research Laboratory

E-mail: †{akihito.seki,ryuzo.okada}@toshiba.co.jp,

††{oliver.woodford,bjorn.stenger,roberto.cipolla}@crl.toshiba.co.uk

あらまし　手持ちの単眼カメラの映像から SfM(Structure from Motion) 法により 3 D形状の復元を行うには，時々刻々と変化するカメラの位置・姿勢推定と，画像間の対応位置を正確に求めることが重要である．本稿では，単眼カメラの入力から 3 D形状の出力までをリアルタイムで行うシステムを提案する．提案法では，カメラの運動推定と，密な 3 D形状復元で 2 段階の処理とし，前段の複雑な処理を CPU，CPU では計算時間はかかるが，処理が単純で並列処理が可能な後段を GPU で計算を行うことで，全体としてリアルタイムな 3 D形状の取得を可能とする．

キーワード　3 D復元，GPU，リアルタイム，マーカーレス

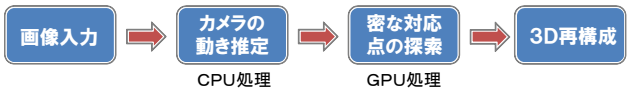

図 1　画像からの 3 次元再構成の流れ．

1. はじめに

プラントの建設・保守・解体，製品検査の他，自動車やロボットの周辺監視装置では，周囲や現場の状況把握のため 3 D復元技術が必要である．前者の社会インフラ応用ではリードタイムの短縮によるコスト削減や，車載応用ではリアルタイム性が機能上必須であることなど，高速に処理できる 3 D復元技術が求められている．

本稿では，近年の計算機の高速化や GPU を用いた超並列計算により，安価に高速な計算手段を手に入れることができようになってきていることを踏まえ，処理に応じて CPU と GPU を使い分けることで，単眼カメラで撮影した時系列画像から密な 3 D形状のリアルタイム復元手法を提案する．

2. 提案方法

図 1 は，提案方法の概略図である．提案法では，マーカーなどの特殊な目印を必要としない一般環境下の利用を想定する．そこで，画像が入力されると，まず画像から疎な特徴点を抽出し，時系列画像間で対応点の探索を行う．対応関係から基本行列を推定し，続く画像間で特徴点のオクルージョン判定やバンドル調整により安定かつ精度よくカメラの運動と疎な特徴点の 3 D形状を復元する．ここまでの処理は比較的複雑で並列処理がしにくいため，ＣＰＵで計算を行う．

続いて，求められた運動情報によって決まる幾何拘束を使って密な 3 D形状を文献 [1] の方法を用いて推定する．複数の画像間を用いて対応点の探索を行うことにより，安定して対応点が求められることが知られているが [2]，類似度分布を保持するためのメモリを大量に必要とする．そこで，図 2 に示すように，奥行に応じた類似度分布を従来のノンパラメトリックな方法から特定の関数の係数で表現するパラメトリックな方法とすることで，奥行の離散化の間隔に応じて増大していたメモリの使用量を削減し，関数の係数パラメタを更新することによって対応位置を決定する．この計算は各画素ごとに並列に計算でき，かつ処理が単純なため GPU で処理することができる．類似度分布をパラメトリックに表現しているために，類似度分布が急峻な対応位置を決定しやすい位置は少ないフレーム数でよく，逆に平坦な位置ではより多くのフレームを使って安定に対応位置を決定することができる．画像全体の 3 D座標を連続的に求めるには，同図で示すように画像をグリッド状に分け，ここでは例えば $(x, y) = (4, 3)$ のセルの対応位置が推定されたら，セル内のパラメータを初期化し，新しく対応点探索を始める．つまり，時刻が進むについて，各セルで異なる時刻からの対応点探索を行っていることになる．

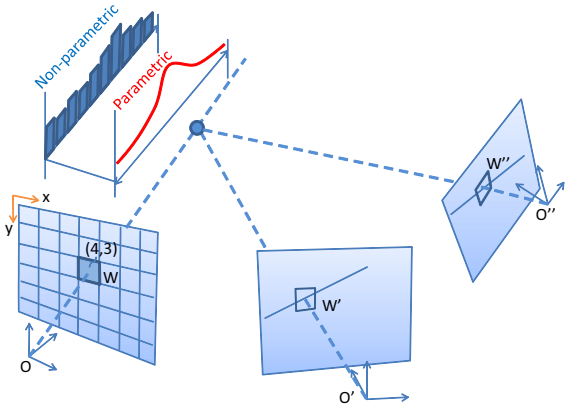

図 2　密な対応点探索の概要

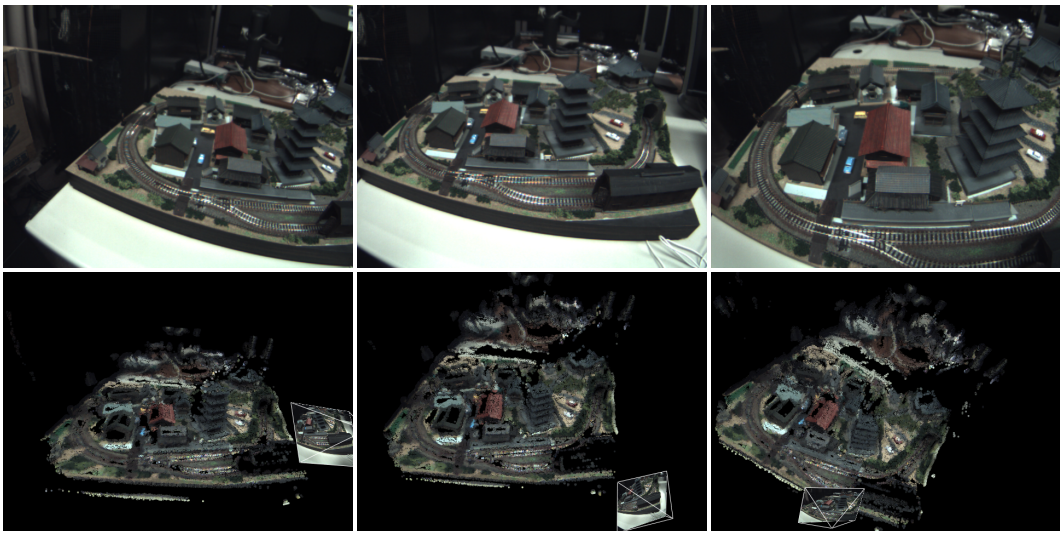

図 3　シーン 1：入力画像の一部 (上段)．左から右に時刻が進む．3 D復元結果 (下段)

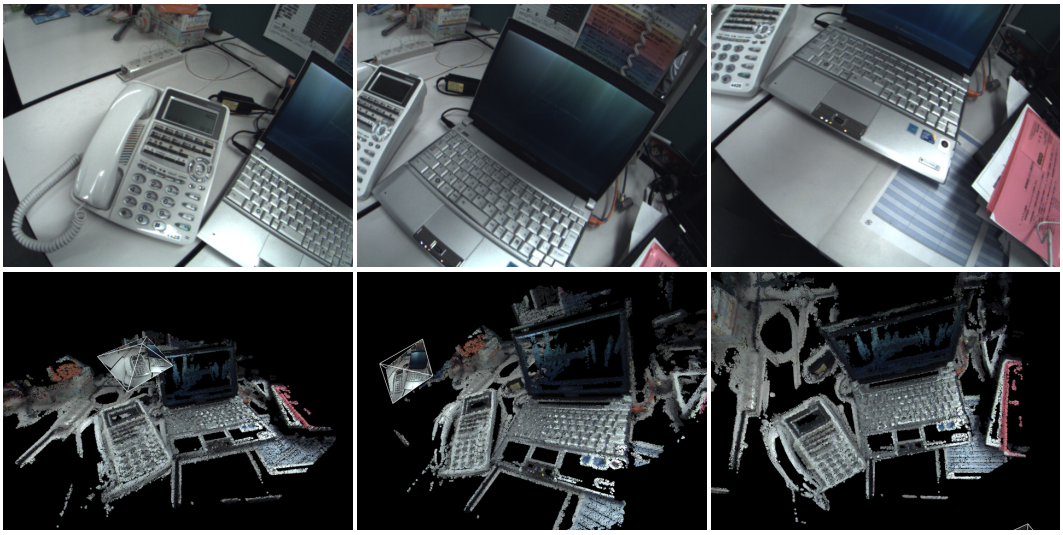

図 4　シーン 2：入力画像の一部 (上段)．左から右に時刻が進む．3 D復元結果 (下段)

3. 実 験 結 果

GPU を搭載したラップトップＰＣを用いたデモシステムを作成した．カメラは PointGray 社製 FireflyMV を，ＰＣは CPU が Intel Core-i7 2630Q，メモリは 20GB，GPU は Nvidia Quadro 2000 を用いてデモシステムを作成した．画像の入力から 3 D復元までの計算は 37ms で行われ，毎秒 58 万点の対応点探索を行っている．図 3 は，模型の左側から右側に向かってカメラを動かして撮影したシーンの画像の一部を上段に，下段に推定された結果を角度を変えて観察した様子である．建物などの 3 D形状が正しく復元されていることがわかる．図 4 は，机上のラップトップＰＣと電話機を復元した結果である．テクスチャの無い部分は復元されていないが，模様のある部分は正しく形状が復元されている．

4. お わ り に

本稿では，手持ちの単眼カメラからの時系列画像から，リアルタイムで密な 3 D復元を可能とする手法を提案した．提案法では，撮影環境にマーカーなどの特殊なマークを必要としないため，さまざまな環境下に適用可能である．汎用計算機を用いてリアルタイムに処理するために，カメラの動き推定は CPU，動き情報から得られる拘束条件を使った密な対応点探索は GPU で処理を行った．実験では，これらの処理を実装したラップトップＰＣを用いて 37ms でリアルタイムに処理ができることを確認した．

文　　献

[1] G. Vogiatzis and C. Hernandez: “Video-based, real-time multi-view stereo”, Image and Vision Computing, Vol. 29, pp. 434–441 (2011).

[2] M. Okutomi and T. Kanade: “A multiple-baseline stereo”, IEEE Trans. on PAMI, **15**, 4, pp. 353–363 (1993).

本論文に掲載の商品，機能等の名称は，それぞれ各社が商標として使用している場合があります．