

テクスチャ付距離画像計測システムを用いた 実環境の 3D マッピングの高速化

Improving a processing time of a 3D mapping
using a textured range image sensing system

○ 内田 裕己 (中央大) 正 寺林 賢司 (中央大)
正 梅田 和昇 (中央大)

Yuki UCHIDA, Chuo University, uchida@sensor.mech.chuo-u.ac.jp
Kenji TERABAYASHI, Chuo University
Kazunori UMEMA, Chuo University

In this paper, we propose a method for improving a processing time of a 3D mapping. 3D map is generated from range images measured from multiple viewpoints by registration and integration of multiple range images. A range image is obtained from a 100Hz real-time textured range image sensing system. To improve the processing time of the previous method, Kanade-Lucas-Tomasi feature tracker (KLT) [1] method and Graphics Processing Unit (GPU) [2] for parallelizing some processes is used. The effectiveness of the proposed method is verified to compare with the previous method in experiments.

Key Words: 3D model, 3D mapping, GPU, KLT, Range Image Sensor

1. 序論

実世界で知的に行動するロボットを実現するためには環境を認識することが重要である。環境認識を行うための技術として三次元環境のマッピングの研究が盛んに行われており、近年では実環境で行動することを目的としてリアルタイム処理の研究が行われている [3]。我々は小型高速距離画像センサを構築しこのセンサを用いて三次元環境のマッピングを実現している [4] が、計算量が多く処理が遅い。特に計算量の多い処理は画像の特徴点トラッキングを行う際に用いる SIFT (Scale-Invariant Feature Transform) 特徴点の抽出 [5]、三次元対応点決定手法の二つである。そこで本研究では SIFT 特徴に比べて高速な手法である KLT トラッカーを用いることと高度な並列計算を得意とする GPU によって三次元対応点決定を並列処理することにより 3D マッピングを高速化することを目的とする。

2. 三次元環境地図生成

2.1 概要

本手法はセンサから得た複数シーンのテクスチャ付距離画像を位置合わせすることによって三次元環境のマップを作成する。その際に用いる位置合わせ方法は、最新のシーンと直前のシーンの二つの距離画像の三次元の対応点を取得する、対応点間の距離を最小とするように距離画像同士を重ね合わせるという二段階の処理になっており、この処理を繰り返すことによってマップを拡大していく。

2.2 テクスチャ付距離画像の取得

図 1 に本研究で使用するテクスチャ付距離画像計測システムを示す。レーザプロジェクタは 361 点のスポット光を投射する小型のものである。モノクロカメラのレンズには、レーザの波長のみを透過するように、前面にローパスフィルタが取り付けられている。このセンサはレーザプロジェクタとモノクロカメラを用いた投光法により距離画像を取得する。得られた距離画像に三角分割を行い、平面を定義しカラーカメラから取得したカラー画像を平面にテクスチャマッピングすることによってテクスチャ付距離画像を取得する。本センサは 100Hz でテクスチャ付距離画像を取得可能である。

2.3 三次元の対応点決定

距離画像同士をレジストレーションするためには、三次元空間での対応をとる必要がある。まず最新のシーンと直前のシーンのカラー画像間で特徴点を抽出し、対応点を決定する。得られた対応点とカメラの光学中心を結んだ直線とカラー画像に対応した三角分割を行った距離画像の平面とで交点計算を行い、得られた交点を三次元における対応点とする(図 2 参照)。



Fig. 1 Textured range image sensing system

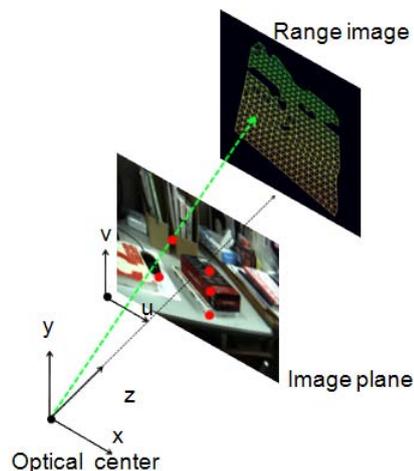


Fig. 2 Intersection calculation for 3D coordinates obtaining of 2D feature point

2.4 線形最小二乗法による回転並進行列の推定

2.3節で述べた方法により得られた最新のシーンにおける三次元点群 $X = \{x_i\}$ と直前のシーンにおける三次元点群 $P = \{p_i\}$ の対応点間の距離の2乗和の平均を最小にする R と t を求める。

$$f = \frac{1}{N_p} \sum_{i=1}^{N_p} |x_i - Rp_i - t|^2 \quad (1)$$

ここで、 R は 3×3 の回転行列、 t は並進ベクトルを表す。 N_p は対応の総数である。この最小化問題を四元数を用いた手法で解く [6]。このとき、特徴点の対応は誤対応を含むため、LMedS 推定 [7] により誤対応を除去して上記の処理を行う。

2.5 誤差楕円体による統合モデルの作成

回転行列 R と並進ベクトル t を用いて位置合わせし、距離画像を重ね合わせることで三次元マッピングを行うことができる。しかしながら、フレーム間の距離画像同士を逐次的にレジストレーションしカメラ運動を推定する方法は、誤差が累積してしまうため正しくマッピングを行うことができないという問題がある。さらに上記の方法でマッピングを行った場合、データ量が膨大になってしまう。

本研究ではデータ量の増加を抑えつつ、累積する誤差を低減する手法として、誤差楕円体による統合モデルを作成する。本手法で用いる誤差楕円体とは、距離画像センサより求めた計測点を中心とし、半径を各軸方向の計測誤差とする楕円体のことである。誤差楕円体は以下の式で表す。

$$\frac{X^2}{\sigma_X^2} + \frac{Y^2}{\sigma_Y^2} + \frac{Z^2}{\sigma_Z^2} = 1 \quad (2)$$

σ : 各軸の計測誤差

最初の二フレームにおける距離画像の統合において、新しく計測した二フレーム目の距離画像の計測点を中心として誤差楕円体を作成し、一フレーム目の距離画像点群に対してこの誤差楕円体の内外判定を行う。楕円体の内部にある一フレーム目の計測点と誤差楕円体の中心である二フレーム目の計測点をまとめて一つの計測点として統合する。このとき統合した計測点の位置は統合に用いた点の平均位置とする。

同様にして、二フレーム目以降は、統合したモデルを構成している点群と新しく計測した距離画像点群とで統合処理を行う。このように統合モデルを更新していくことで対象の3Dマッピングを行う。

3. 3D マッピングの高速化

3.1 KLT の利用

KLT とは、画像同士の特徴点トラッキングに広く用いられている手法である。画像のフレーム間で小さな変化をする特徴点を仮定しているため高速なトラッキングが可能となっている。一方、従来手法で用いている SIFT 特徴は回転やスケール変化にロバストであるが処理速度が遅いという問題がある。本研究で用いているセンサは高速なセンシングが可能であるためカラー画像のフレーム間の変化は微小とみなせる。そのため本センサと KLT を組み合わせることで十分な精度で特徴点トラッキングが可能である。以上より3Dマッピングを行う際の二次元の対応関係を求める処理では KLT を用いる。

3.2 GPU を用いた三次元対応点決定処理の並列化

GPU は、グラフィックス専用のプロセッサであり、その特徴として膨大な並列計算を処理できる点が挙げられる。そのため、同じ計算を複数回おこなう場合にその処理を並列化することにより処理速度を上げることができる。

本手法では 2.3 節で述べた交点計算を得られた特徴点と距離画像の平面で全探索的に処理していく。多い時では数千点の特徴点と 600 個以上の平面との処理になるため計算コストは非常に大きい。しかしこの処理は逐次的に処理する必要がなく並列計算に適していると考えられる。

GPU は階層的なスレッド構造を持っている。各スレッドに1つの特徴点を割り当て、図 3 のように並列に交点計算を行うことによって高速に処理することができる。

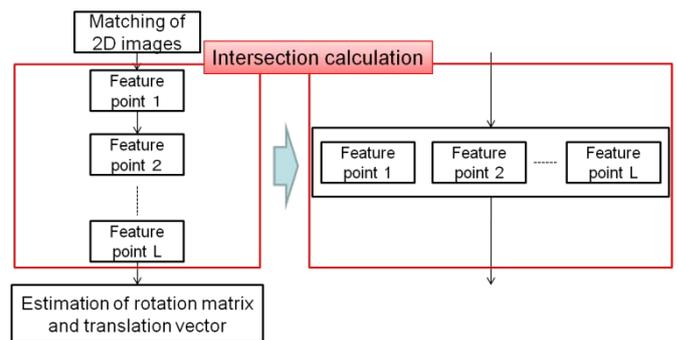


Fig. 3 GPU processing of intersection calculation

4. 実験

提案手法の有用性を検証するため、図 4 のシーンに対して従来手法と提案手法を使って 3D マッピングを行い、結果を比較した。PC は Intel Core2 Extreme(3GHz), 4GB RAM を搭載したものを使用した。またグラフィックスカードは NVIDIA 社製の GeForce GTX 285 を使用した。なお 3D マッピングに用いたデータは図 4 のシーンは距離画像とそれに対応するカラー画像 122 フレーム分を使用した。

従来手法と提案手法の 3D マッピングの結果を図 5 に示す。これらのマッピング結果から提案手法で従来手法と同様にマッピングできていることが確認できる。よって SIFT 特徴に代わり KLT を適用可能であることがわかる。またマップ作成にかかった合計時間の比較結果を表 1 に示す。これより従来手法に比べて 2 倍以上速くなったことがわかる。さらに 1 フレームあたりの合計時間、二次元対応、三次元対応の処理時間の比較結果を表 2 に示す。高速化を行った処理である二次元の対応付けと三次元への拡張処理を見るとそれぞれ約 56 倍、20 倍速くなっており、大幅に高速化できていることが確認できる。また 1 フレームあたりの合計時間は幅を持った表記になっているが、これは本手法の 3D マッピングはマップが拡大するに従い、処理時間が増加していくためである。処理時間が増加する問題は誤差楕円体による統合マップ作成の処理に原因がある。この処理は 2.5 節で述べたように 2 フレーム目以降は統合モデルの点群と新しく計測した点群とで統合処理を行う。そのため統合モデルが拡大すると参照するデータ量も増加するため処理時間の増加につながる。今後解決すべき点である。

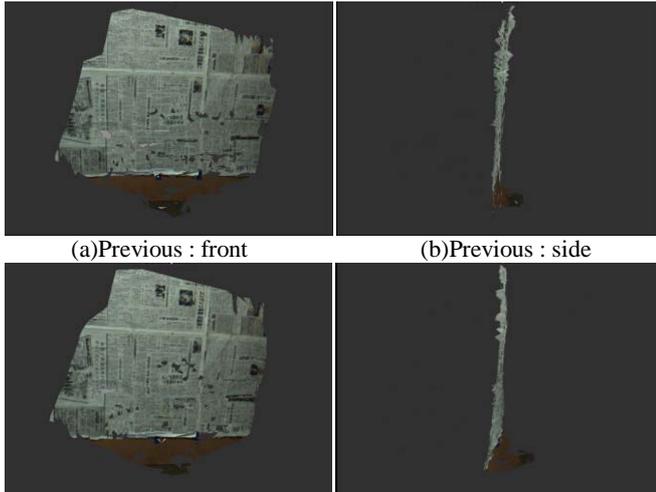
次に、より複雑なシーンに対して提案手法を用いた 3D マッピングが行えるのかを確認するため、図 6 のシーンに対して 3D マッピングを行った。3D マッピングに用いたデータは 156 フレーム分を使用した。実験結果を図 7、表 3、4 に示す。マッピング結果から複雑なシーンに対しても提案手法で 3D マップが作成可能であることがわかる。



Fig. 4 Measured object 1



Fig. 6 Measured object 2



(a)Previous : front

(b)Previous : side

(c)Proposed : front

(d)Proposed : side

Fig. 5 Mapping results

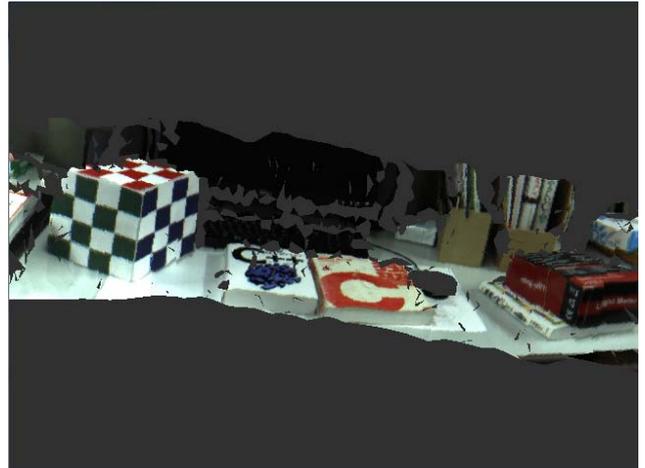


Fig. 7 Mapping result

Table 1 Processing time of 3D mapping

	Previous	Proposed
Time [s]	635	266

Table 3 Processing time of 3D mapping

	Previous	Proposed
Time [s]	552	215

Table 2 Processing time per 1 frame

	Previous	Proposed
Total [ms]	4062~7285	1150~2537
2D Matching [ms]	2911	52
3D Matching [ms]	243	12

Table 4 Processing time per 1 frame

	Previous	Proposed
Total [ms]	3072~5350	803~2178
2D Matching [ms]	2204	38
3D Matching [ms]	169	15

5. 結論

テクスチャ付距離画像計測システムを用いた 3D マッピング手法を示し、KLT 及び GPU を用いて提案手法の高速化を行った。また実環境の 3D マッピングを行い、処理時間を解析し、提案手法の有用性を示した。

今後の課題としてさらなる高速化、高精度化、構築した 3D マップを用いた物体認識が挙げられる。統合マップ作成の処理はマップ拡大による膨大なデータの参照が原因である。そのため GPU を用いて並列処理を行うことである程度は解決できる問題である。またカラーカメラとモノクロカメラを同軸にすることによりテクスチャマッピングの高精度化が期待できる。さらに作成したマップをオブジェクト別に切り出し、モデルベースマッチングに利用することができる。

謝辞 本研究は科研費(20500164) の助成を受けたものである。

文 献

- [1] C. Tomasi and T. Kanade, "Detection and Tracking of Point Features," CMU Tech. Rep. CMU-CS-91-132, 1991; <http://vision.stanford.edu/~birch/klt/>.
- [2] 青木尊之, 額田彰, "はじめての CUDA プログラミング", 工学社, 2009.
- [3] 加賀美聡, "ステレオカメラからのオンライン 3 次元 SLAM", ロボット学会誌, vol.26, no.4, pp.310-313, 2008.
- [4] H. Ishiyama, K. Terabayashi, K. Umeda, "A 100Hz Real-time Sensing System of Textured Range Images," Proc. of 2010 International Symposium on Optomechatronic Technologies, SP-SH-7, 2010.
- [5] D. G. Lowe, "Distinctive image features from scale-invariant keypoints", International Journal of Computer Vision, vol.60, no.2, pp.91-110, 2004.
- [6] P. Besl, and D. McKay, "A method for registration of 3-d shapes," in Proc. of IEEE Transactions on Pattern Analysis and Machine Intelligence, vol.14, no.2, pp.239-256, 1992.
- [7] P. J. Rousseeuw, "Least Median of Squares Regression", in J. American Stat. Assoc., vol.79, no.388, pp.871-880, 1984.