

セマンティックセグメンテーションを用いた三次元点群の位置合わせ

山田阜平[†], 山下隆義[†], 藤吉弘亘[†]

[†]: 中部大学, {satsuki@mprg.cs, takayoshi@isc, fujiyoshi@isc}.chubu.ac.jp

概要: 自己位置推定で用いられるIterative Closest Point(ICP)は, 対応付けと剛体変換推定を繰り返すことで, 2つの点群の位置合わせを行う. 一方の点群の各点に対し, 他方の点群における最近傍点を対応付ける. そして, 対応点間の距離を最小化する剛体変換を推定する. ICPは, 対応点同士が同一物体に属するか判定していないため, 点群の初期位置が大きく離れると, 位置合わせが失敗する. そこで, 位置合わせをする点群にセマンティックセグメンテーションを行い, 各点が属する物体の種類の情報を利用することで, 対応点同士が同一物体に属するか判定できる. 本稿では, セマンティックセグメンテーションラベルを対応付けに活用した位置合わせ手法を提案する. これにより, 誤った対応付けを抑制し, 位置合わせの更新回数を削減できる. また, 実験ではICPと比較し, 高精度かつ高速に位置合わせできることを確認した.

1. はじめに

自己位置推定は, ロボットが自律移動するために重要な機能の1つである. Global Positioning System(GPS)を用いて, 衛星などからの電波を利用する手法や, ジャイロセンサを用いて, ロボットの加速度を検出してロボットの姿勢を推定する手法により, 自己位置推定することができる. しかし, 屋内環境やロボットの車輪が滑りやすい環境など, GPSやジャイロセンサが使用できない環境が存在する. このような環境において, ロボットが記憶しているマップとレンジセンサなどで計測した三次元点群を対応付けることで, ロボットの正確な自己位置を推定することができる. そのため, 自己位置推定には, 異なる視点における三次元点群の位置合わせが重要となる.

代表的な点群の位置合わせ手法として, Iterative Closest Point(ICP)[1][2]がある. ICPは, 一方の点群を構成する各点に対し, 他方の点群における最近傍点を探索して対応付けを行い, 対応点間の距離を最小化する剛体変換行列を推定する. そして, 対応付けと剛体変換推定を繰り返し行うことで, 2つの点群の位置合わせを行う.

しかし, 対応点同士が同じ物体から取得したものかどうかを判定していないため, 点群の初期位置が大きく離れると, 位置合わせが失敗する. そこで, 位置合わせする点群の各点が属する物体の種類の情報

報を利用することで, 対応点同士が同じ物体に属するか判定ができる.

本稿では, PointNet[4]を利用して点群にセマンティックセグメンテーションを行い, セマンティックセグメンテーションラベルを対応付けに用いる手法を提案する. これにより, 誤った対応付けを抑制し, 位置合わせの更新回数を削減できる. また, セマンティックセグメンテーションを用いた対応付けにより, ICPよりも高精度かつ高速に位置合わせできることを示す.

2. 関連研究

位置合わせ手法とセマンティックセグメンテーションの関連研究について述べる.

2.1. Iterative Closest Point

ICPは, 位置合わせの基準となるモデル点群 P と異なる視点で計測されたデータ点群 Q の位置合わせを行う. データ点群 Q を構成する各点 q_j に対して, モデル点群 P における最近傍点 p_i を探索する. 最近傍点 p_i は, 対応付けのペア (p_i, q_j) として, 式(1)を用いて求める.

$$p_i = \operatorname{argmin}_{p \in P} \|p - (R^{icp} q_j + t^{icp})\| \quad (1)$$

対応付けのペアから対応点間の距離を最小化する回転行列と並進ベクトルを推定する. 回転行列 R^{icp} と並進ベクトル t^{icp} の更新式は式(2)のように表す.

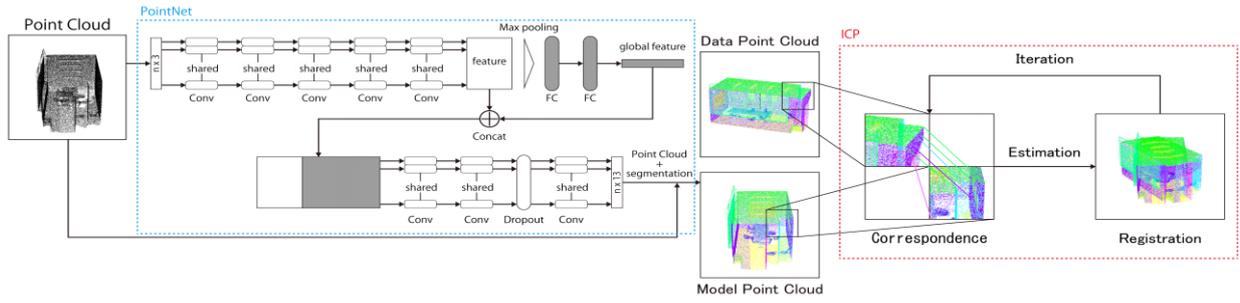


図1 提案手法のアルゴリズム

$$R^{icp}, t^{icp} = \operatorname{argmin}_{R, t} \sum_j^{|Q|} \|p_i - (Rq_j + t)\| \quad (2)$$

対応付けと回転行列および並進ベクトルの推定を繰り返すことで、位置合わせが可能である。しかし、ICPは点が属する物体の種類が一致しているかを考慮せず、最近傍点に対して対応付けを行うため、点群の初期位置が大きく離れると、誤った対応付けがされ、更新回数の増加や位置合わせ失敗などの問題が発生する。

そこで、特徴点を利用した手法が提案されている。エッジ上の点のみを用いて対応付ける手法[3]や座標だけでなく、法線方向・主曲率・モーメント等の不変特徴量の二乗誤差も加えた距離を用いる手法[6][7]、テクスチャや色の情報も利用する手法[8]は、位置合わせのロバスト性を向上させている。しかし、これらの手法は、特徴量と組み合わせることで、より離れた初期値からの位置合わせを可能にしているが、適切に組み合わせられない場合、位置合わせに失敗する。また、正しい対応付けの満たすべき幾何学的な性質を導き、これを用いて誤対応を除去する手法[9]が提案されている。しかし、この手法は、初期位置が離れすぎると、正確な位置合わせができない。

2.2. PointNet

三次元点群は、各点の順番によらず同形状を形成することができるデータ構造をしている。そのため、従来の画像を対象としたセマンティックセグメンテーション手法を用いることは難しい。PointNetは、点群を直接扱うことができるネットワーク構造として提案されている。これは、点群の各点に対して、個別に畳み込み処理を行い、局所特徴を抽出する。そして、Max poolingにより局所特徴を集約する。この処理では、各特徴に対して有効な点を選択する。これにより、点数や順番に対して不変な局所特徴となる。その後、集約した局所特徴を全結合層へ入力することで全体特徴を得る。全体特徴と局所特徴を結合して最終的

な各点の特徴とし、点群に対するセマンティックセグメンテーションを行う。

3. 提案手法

誤った点群の対応付けを抑制するため、セマンティックセグメンテーションのラベルにより同一物体の点群に対する対応付けを行う。図1に提案手法のアルゴリズムを示す。はじめに、点群をPointNetに入力し、点群に対してセマンティックセグメンテーションを行う。そして、セマンティックセグメンテーションのラベルを用いてICPにより対応付けを行う。ここで、位置合わせの際に基準となる点群をモデル点群、変換される点群をデータ点群とする。そして、モデル点群を $|P|$ 個の点の集合 P 、データ点群を $|Q|$ 個の点の集合 Q とする。また、 P, Q に含まれる各点を p, q とする。

3.1. 点群のセマンティックセグメンテーション

点群のセマンティックセグメンテーションにはPointNetを用いる。点群の点数を n とし、各点に対して畳み込み処理(Conv層)を複数層行い、局所特徴を抽出する。そして、Max poolingにより集約し、全結合層(FC層)から全体特徴を得る。全体特徴と局所特徴を利用して、各点のセマンティックセグメンテーションを行う。

3.2. セグメンテーションを考慮した対応付け

点 p のセグメンテーションラベルを $S(p)$ 、点 q のセグメンテーションラベルを $S(q)$ とする。データ点群 Q に含まれる点 q に対して、モデル点群 P における $S(q)$ と一致する $S(p)$ をもつ最近傍点 p を探索し、対応付けのペア (p, q) を求める。セグメンテーションラベルを用いることで、同一物体の点群の対応付けを行い、誤った対応付けを削減する効果が期待できる。

ここで、モデル点群 P の i 番目の点 p_i とデータ点群 Q の j 番目の点 q_j を $|Q|$ 組ある対応付けのペア (p_i, q_j) とし、 q_j を p_i に近づける回転行列を R^{icp} 、並進ベクトルを t^{icp} とする。 R^{icp} と t^{icp} の初期値はそれぞれ単位行列と単位ベクトルとする。これにより、 q_j と対応した

\mathbf{p}_i は式(3)のように求められる。

$$\mathbf{p}_i = \underset{\{\mathbf{p} \in P | S(\mathbf{q}_j) = S(\mathbf{p})\}}{\operatorname{argmin}} \|\mathbf{p} - (\mathbf{R}^{icp} \mathbf{q}_j + \mathbf{t}^{icp})\| \quad (3)$$

式(3)の計算量は、 $S(\mathbf{p})$ を持つ点の数を $|S(\mathbf{p})|$ 、 $S(\mathbf{q}_j)$ を持つ点の数を $|S(\mathbf{q}_j)|$ とすると、 $O(|S(\mathbf{q}_j)||S(\mathbf{p})|)$ となる。また、従来のICPの対応付けの計算量は全探索であるため、 $O(|Q||P|)$ となる。それぞれ $|S(\mathbf{p})| \leq |P|$ 、 $|S(\mathbf{q}_j)| \leq |Q|$ であるため、提案手法は従来のICPより計算量を削減できる。

3.3. 対応付けを用いた剛体変換推定

3.2節で求めた対応付けのペア $(\mathbf{p}_i, \mathbf{q}_j)$ から、 \mathbf{q}_j を \mathbf{p}_i に近づけるような回転行列 \mathbf{R}^{icp} と並進ベクトル \mathbf{t}^{icp} を推定する。回転行列 \mathbf{R}^{icp} と並進ベクトル \mathbf{t}^{icp} の更新式は式(4)のように表す。

$$\mathbf{R}^{icp}, \mathbf{t}^{icp} = \underset{\mathbf{R}, \mathbf{t}}{\operatorname{argmin}} \sum_j^{|Q|} \|\mathbf{p}_i - (\mathbf{R} \mathbf{q}_j + \mathbf{t})\| \quad (4)$$

3.2節で求めた対応付けのペアの中には誤ったペアが含まれている可能性がある。そこで、RANSAC[10]により誤ったペアの影響を抑える。対応付けのペアからランダムに対応付けを選択して対応点間の距離の算出と、剛体変換推定を繰り返し行う。そして、距離が中央値となる剛体変換を正しい推定とし、点群を変換する。

3.4. 収束条件

回転行列 \mathbf{R}^{icp} と並進ベクトル \mathbf{t}^{icp} について、収束判定条件に満たしている場合は終了し、満たしていない場合は3.2節の処理から繰り返す。本研究では、更新回数が k 回であるとき、点群間の距離の総和 d_k は式(5)のように求められる。

$$d_k = \frac{1}{|Q|} \sum_j^{|Q|} \|\mathbf{p}_i - (\mathbf{R}^{icp} \mathbf{q}_j + \mathbf{t}^{icp})\| \quad (5)$$

距離の総和 d_k と収束の閾値 τ を用いて、式(6)のように判定条件を設定する。

$$d_k \leq \tau \quad (6)$$

4. 評価実験

提案手法の有効性を調査するため、評価実験により評価する。

4.1. 実験概要

PointNetによる点群のセマンティックセグメンテーションの精度と位置合わせにセマンティックセグメンテーションを導入した効果について評価実験を行う。

セマンティックセグメンテーションの評価では、識別精度を算出する。位置合わせの評価では、更新回数や誤差、処理時間を比較する。比較に用いる手法はICPとEdge point ICPである。

評価実験で使用するPointNetの構造を表1に示す。畳み込み層(Conv)が8層、全結合層(FC)が2層の計12層のPointNetを用いる。入力層は、XYZとRGB、XYZを部屋の大きさごとに正規化したnormalized XYZの9次元の情報を持つ4,096点の三次元点群を入力する。これより、入力ユニット数は36,864となる。また、各層の活性化関数にはReLUを用いる。PointNetの学習パラメータを表2に示す。使用するデータセットは、S3DIS Dataset[4]である。学習には、Area1-4とArea6の203部屋を使用する。評価には、

表1 PointNetの構造

layer		
Conv1	フィルタサイズ	1 × 9
	ユニットサイズ	4,096 × 64
Conv2	フィルタサイズ	1 × 1
	ユニットサイズ	4,096 × 64
Conv3	フィルタサイズ	1 × 1
	ユニットサイズ	4,096 × 64
Conv4	フィルタサイズ	1 × 1
	ユニットサイズ	4,096 × 128
Conv5	フィルタサイズ	1 × 1
	ユニットサイズ	4,096 × 1024
Max pooling	フィルタサイズ	4,096 × 1
	ユニットサイズ	1,024
FC1	ユニットサイズ	256
FC2	ユニットサイズ	128
Concat	ユニットサイズ	4,096 × 1052
Conv6	フィルタサイズ	1 × 1
	ユニットサイズ	4,096 × 512
Conv7	フィルタサイズ	1 × 1
	ユニットサイズ	4,096 × 256
Conv8	フィルタサイズ	1 × 1
	ユニットサイズ	4,096 × 13

表2 PointNetの学習パラメータ

学習係数	0.001
バッチサイズ	24
Epoch 数	50
学習誤差関数	ソフトマックスクロスエントロピー関数

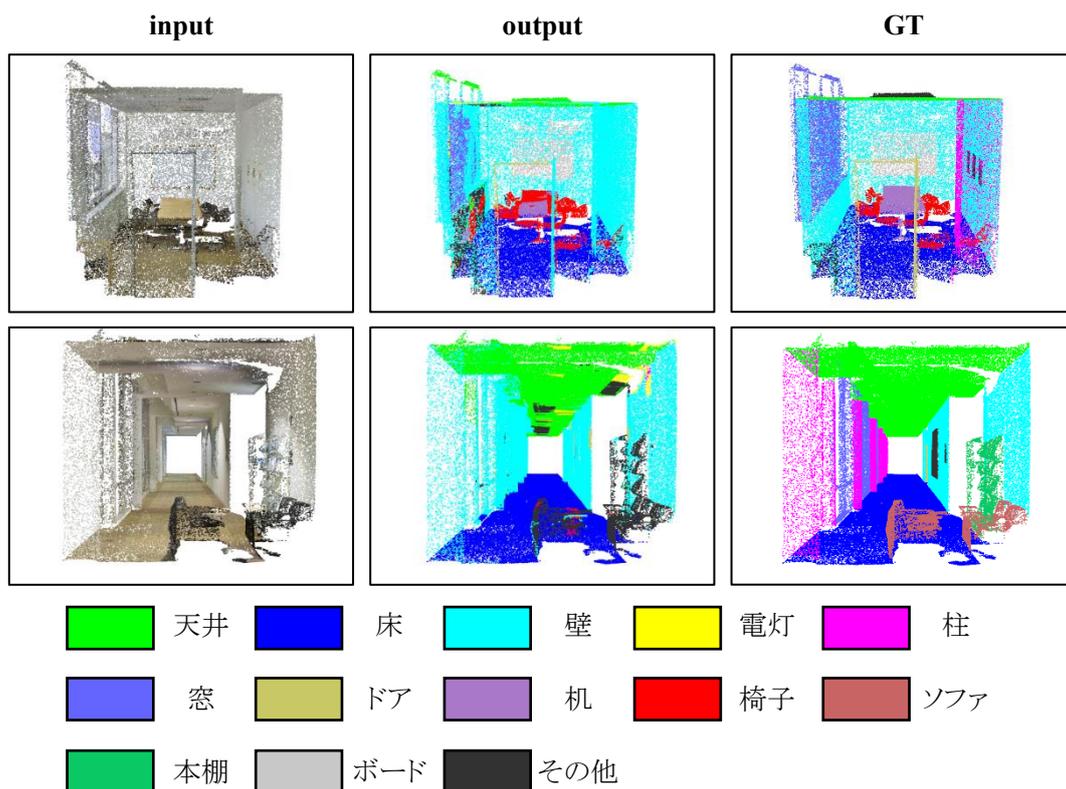


図2 セマンティックセグメンテーション例

表4 平均更新回数[回]

	回転[度]									
	0	10	20	30	40	50	60	70	80	90
ICP	28	26	27	29	31	34	36	40	48	74
Edge point ICP	26	24	26	28	30	33	35	39	47	73
提案手法	22	20	20	21	21	21	21	22	23	23
削減率[%]	21.4	23.1	25.9	27.6	32.3	38.2	41.7	45.0	52.1	68.9

Area5の68部屋を使用する。学習に使用するセマンティックセグメンテーションラベルは、天井、床、壁、電灯、柱、窓、ドア、机、椅子、ソファ、本棚、ボード、その他の13クラスである。また、位置合わせには、動かない天井、床、壁、柱、窓、ボードの6クラスを使用する。使用するモデル点群は、計測した点群で、使用するデータ点群は、モデル点群に対して回転と並進の変換を加えて、ランダムにノイズを付与した点群である。回転はz軸を中心に0~90度まで10度間隔で、並進はy軸において0~3.0mまで0.5m間隔で変動させ、評価する。

4.2. セマンティックセグメンテーションの評価

PointNetによる点群のセマンティックセグメンテーションの精度を評価する。表3にセマンティックセグメンテーションの精度を示す。また、セマンティックセグメンテーション例を図2に示す。図2により、柱を壁と誤

表3 セマンティックセグメンテーションの精度[%]

	mean IoU	overall accuracy
PointNet	42.38	79.86

識別している。これは、柱は壁の一部であり、特徴が似ているためだと考えられる。一方で、天井、床、壁など面積が大きい物は識別できていることから、よりよい特徴を抽出に成功していることがわかる。

4.3. 更新回数の比較

ICPとEdge point ICP, および提案手法について、距離の総和 d_k が0.01以下となる時の更新回数を比較する。また、更新回数の上限を100回として、収束していない場合でも終了する。ノイズとして、データ点群の点をランダムに10%減らしている。各回転角における平均更新回数と、ICPと比較した提案手法の更新回数の削減率を表4に示す。表4より、提案手法は初期位置に関係なくICPよりも少ない更新回数で位

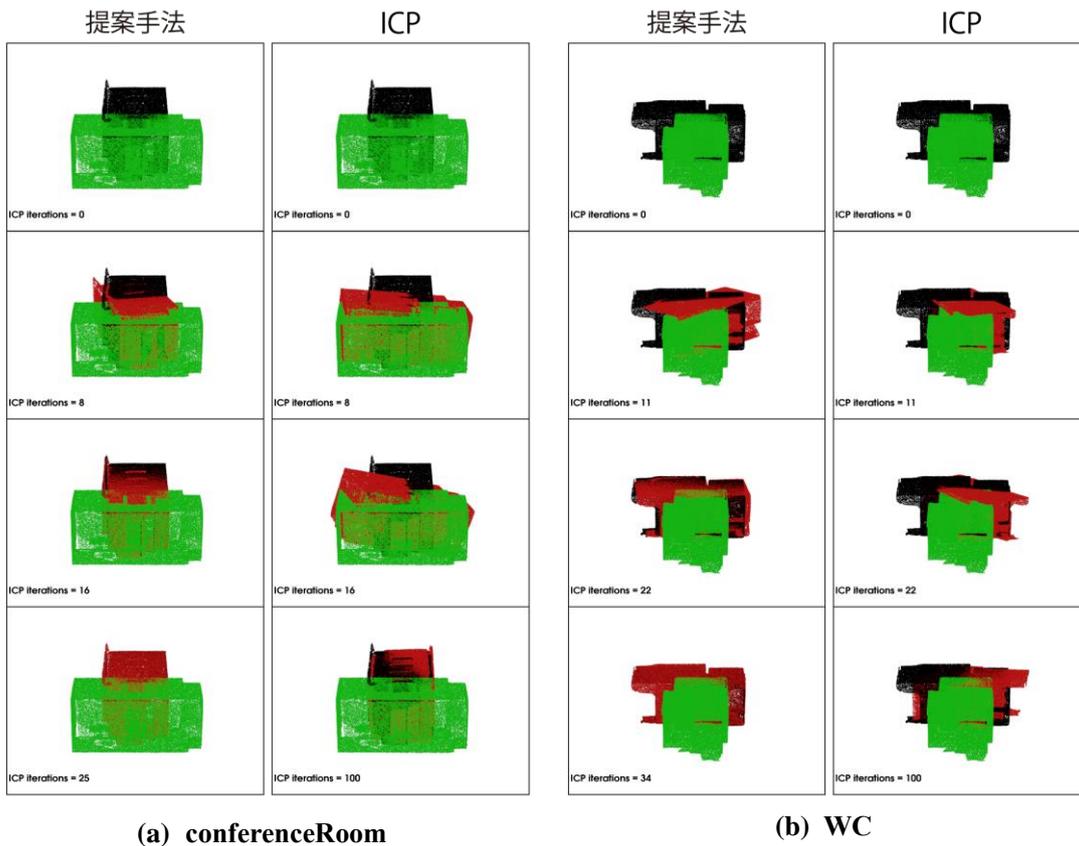


図3 位置合わせ例

位置合わせが可能である。また、提案手法は回転角が大きいほど、削減率が向上し、少ない更新回数で位置合わせできる。特に、90度回転した際の位置合わせは、ICPと比べて平均更新回数を約68.9%削減できている。これより、初期位置が離れるほどセマンティックセグメンテーションの有効性が顕著に表れている。

90度回転、3.0m並進した場合のICPおよび提案手法による位置合わせ例を図3に示す。図3における黒色の点群はモデル点群、緑色の点群は初期位置のデータ点群、赤色の点群は変換されるデータ点群を示している。提案手法はセマンティックセグメンテーションを用いることで、正しい対応付けができており、図3(a)では25回の更新、図3(b)では34回の更新で位置合わせに成功していることがわかる。しかし、ICPは初期位置が大きく離れると、誤った対応付けを行い、位置合わせを失敗する問題がある。そのため、図3(a)、図3(b)のどちらも更新回数100回目のとき、反対向きに位置合わせしている。

4.4. 誤差の比較

ICPとEdge point ICPおよび提案手法について、20回更新したときの誤差を比較する。また、ノイズとしてデータ点群の点をランダムに10%減らしている。各回転角における平均誤差を表5に示す。表5より、提案

表3 平均誤差[$\times 10^{-2}\%$]

	回転[度]									
	0	10	20	30	40	50	60	70	80	90
ICP	3.93	3.74	4.27	6.16	9.76	14.7	19.1	27.2	48.2	61.0
Edge point ICP	3.69	3.5	3.95	5.81	9.66	14.8	19.4	27.9	49.3	61.4
提案手法	1.78	1.66	1.62	1.6	1.66	1.74	1.9	2.06	2.24	2.44

表4 平均処理時間[分]

	回転[度]									
	0	10	20	30	40	50	60	70	80	90
ICP	35.6	35.9	29.3	28.7	29.7	31.6	30.6	29.1	37.9	33.9
Edge point ICP	35.8	35.2	29.8	29.5	30.0	31.9	31.6	30.0	35.5	34.9
提案手法	14.1	15.9	17.7	18.5	17.1	16.9	17.1	17.0	16.9	16.9

手法は初期位置に関係なく、ICPよりも速く誤差が収束していることがわかる。これは、セマンティックセグメンテーションラベルにより、同一物体の点同士の対応付けを行うことで、誤った対応付けをすることなく、正しい対応付けができているためである。

4.5. 処理時間の比較

4.4節と同じ条件で、20回更新したときの処理時間を比較する。各回転角における更新1回あたりの平均処理時間を表6に示す。表6より、提案手法の処理時間はICPの半分程度に抑えられている。ICPの対応付

表5 各ノイズ付与率における平均更新回数[回]

	ノイズ付与率[%]								
	10	20	30	40	50	60	70	80	90
ICP	74	73	73	74	74	74	73	73	74
Edge point ICP	73	73	73	73	73	73	73	74	75
提案手法	23	24							

表6 各ノイズ付与率における平均誤差 $\times 10^{-2}\%$

	ノイズ付与率[%]								
	10	20	30	40	50	60	70	80	90
ICP	61.0	60.4	61.7	60.7	60.9	62.2	61.3	61.2	62.5
Edge point ICP	61.4	60.7	61.8	60.3	61.0	62.5	61.4	60.7	63.3
提案手法	2.44	2.46	2.51	2.64	2.61	2.67	2.55	2.61	2.48

表7 各ノイズ付与率における平均処理時間[分]

	ノイズ付与率[%]								
	10	20	30	40	50	60	70	80	90
ICP	33.9	31.8	27.2	11.9	10.9	8.89	5.16	2.06	1.26
Edge point ICP	34.9	32.5	27.1	23.7	14.3	9.41	6.85	4.02	1.06
提案手法	16.9	13.9	9.95	4.83	4.05	3.25	2.79	1.42	0.40

けの計算量は全探索であるのに対して、提案手法の対応付けの計算量は同じセマンティックセグメンテーションのラベルを持つ点のみの探索であるため、計算量は削減され、処理時間を削減した。

4.6. ノイズ付与率ごとの比較

ICPとEdge point ICPおよび提案手法について、回転を90度に固定し、ノイズ付与率を10~90%まで10%間隔で変更した場合の更新回数と誤差、処理時間を比較する。各ノイズ付与率における平均更新回数を表7に、平均誤差の比較を表8に、更新1回あたりの平均処理時間の比較を表9に示す。表7と表8より、ノイズ付与率を変動させても、平均更新回数と平均誤差に変化は確認できない。表9より、ノイズ付与率が高くなるほど処理時間が削減されていることが確認できる。これは、単純に点の総数が減ったためであるが、提案手法はノイズ付与率が高くなり、対応付けで使用できる点が減ったとしても、ICPよりも少ない更新回数で位置合わせに成功している。

5. おわりに

本稿では、セマンティックセグメンテーションを用いた位置合わせ手法を提案した。提案手法により、更新回数を削減した。また、同じ更新回数するとき、初期位置に関係なく精度と処理速度を向上させた。今後の課題として、大域的な位置合わせと組み合わせた2段階の位置合わせによる処理の効率化やセマンティックセグメンテーションの精度が良いラベルの利用が

挙げられる。

参考文献

- [1] Y. Chen, and G. Medioni: Object modelling by registration of multiple range images, Image and vision computing, 10(3):145-155, 1992.
- [2] S. Rusinkiewicz, and M. Levoy: Efficient Variants of the ICP Algorithm, 3DIM, 2001.
- [3] 友納正裕: エッジ点追跡に基づくステレオカメラを用いた三次元SLAM, 日本ロボット学会誌, 27(7):759-767, 2009.
- [4] C. R. Qi, et al.: Pointnet: Deep learning on point sets for 3d classification and segmentation, CVPR, 2017.
- [5] I. Armeni, et al.: 3D Semantic Parsing of Large-Scale Indoor Spaces, CVPR, 2016.
- [6] J. Feldmar, and N. Ayache: Rigid, affine and locally affine registration of free-form surfaces, International journal of computer vision, 18(2):99-119, 1996.
- [7] G. C. Sharp, et al.: ICP registration using invariant features, PAMI, 24(1):90-102, 2002.
- [8] G. Godin, et al.: A method for the registration of attributed range images, 3DIM, 2001.
- [9] Y. Liu, and M. A. Rodrigues: Geometrical analysis of two sets of 3D correspondence data patterns for the registration of free-form shapes, IROS, 33(4):409-436, 2002.
- [10] M. A. Fischler, and R. C. Bolles: Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography, ACM, 24(6):381-395, 1981.

山田卓平: 2017年中部大学工学部情報工学科卒, 現在同大学大学院工学研究科情報工学専攻博士前期課程在学中, 三次元点群の研究に従事。

山下隆義: 2002年奈良先端科学技術大学院大学博士前期課程修了, 2002年オムロン株式会社入社, 2011年中部大学大学院博士後期課程修了(社会人ドクター), 2014年中部大学講師, 2017年中部大学准教授。人の理解に向けた動画像処理, パターン認識・機械学習の研究に従事。2009年画像センシングシンポジウム高木賞, 2011・2014・2015年画像センシングシンポジウム オーディエンス賞, 2013年電子情報通信学会 情報・システムソサイエティ論文賞, 2013年電子情報通信学会 PRMU研究会研究奨励賞, 2016年電子情報通信学会 査読功労賞, 2016年画像センシングシンポジウム 優秀学術論文賞, 2016年画像センシングシンポジウム 最優秀学術論文賞, 2017年IWAIT2017 Best Paper Award。

藤吉弘亘: 1997年中部大学大学院博士後期課程修了, 1997年米カーネギーメロン大学ロボット工学研究所Postdoctoral Fellow, 2000年中部大学工学部情報工学科講師, 2004年中部大学准教授, 2005年米カーネギーメロン大学ロボット工学研究所客員研究員(~2006年), 2010年中部大学教授, 2014年名古屋大学客員教授。計算機視覚, 動画像処理, パターン認識・理解の研究に従事。2005年ロボカップ研究賞, 2009年情報処理学会論文誌 コンピュータビジョンとイメージメディア 優秀論文賞, 2009年情報処理学会山下記念研究賞, 2010・2013・2014年画像センシングシンポジウム優秀学術賞, 2013年電子情報通信学会 情報・システムソサイエティ論文賞, 2015年永井科学技術財団 財団学術賞, 2015年電子情報通信学会 情報・システムソサイエティ査読功労賞他。