移動ロボットのための SIFT 特徴量を用いたシーン認識

當眞 康史^{a,*} 萩原 勇人^a 浅海 賢一^a 小森 望充^a

Scene Recognition using SIFT Image Features for Mobile Robots

Yasufumi Touma^{a,*}, Hayato Hagiwara^a, Kenichi Asami^a, Mochimitsu Komori^a

(Received July 8, 2013; revised September 11, 2013; accepted September 12, 2013)

Abstract

This paper presents the scene recognition method using SIFT (Scale Invariant Features Transform) image features descriptor for the purpose of position revision of the mobile robot. The SIFT features descriptor has robustness for scale, rotation, and illumination in the scene image. Therefore, the mobile robot can benefit from such a characteristic in the time/space changing environment. In this proposal, the horizontal shifting distance between two scene images is calculated from the correspondence of SIFT features.

キーワード : SIFT, 特徴マッチング, シーン認識, 移動ロボット

Keywords : SIFT, feature matching, scene recognition, mobile robot.

1. はじめに

1.1 背景

科学技術の向上に伴い自動車工場をはじめ多くの工場オートメーション化が進み、ロボットアームなどの産業用ロボットが活躍し現代社会の産業を支えている。しかし多くの先進国で少子高齢化が進み、数年前から日本は超高齢社会となり、今後も高齢化が進むことが予想されている。それに伴いロボットアームなどの産業用ロボットだけでなくサービスロボットの市場も拡大が見込まれており、セキュリティや見守りなどの用途においてもサービスロボットが活躍すると予想されている⁽¹⁾。今後、サービスロボットの利用が予想される用途の一つに会社のオフィスや学校などの巡回・監視サービスがある。廊下をはじめとした屋内の環境下でカメラを搭載した移動ロボットが自動で巡回し、警備・監視を行うといったセキュリティ目的での利用である。

ロボットが自動で移動を行う場合、障害物等の情報を含んだ環境の地図が必要であり、画像情報をもとに環境地図を作成する手法の研究も数多くなされている⁽²⁻⁴⁾。しかし、廊下をロボットが移動・巡回する際に、タイヤの

スリップなどの様々な外乱的要素によって偶発的に正規の移動ルートから外れ、ロボットの実際の位置と通るべき位置でズレが生じてしまうことがある。このような状況に陥った場合、ロボットは正しい行動を継続することが困難になり、最悪の場合事故が発生する危険がある。したがって、ロボットが自動で巡回行動を続けるためには移動ロボットが自己の位置を認識・判断し、位置のズレが生じた際には正しい位置・ルートに復帰する機能が必要不可欠である。

1.2 研究目的

前述のように巡回・移動を行うロボットには自己の位置を補正する機能が必要である。自己位置補正機能を単純に実現する方法としてセンサーを多数搭載する方法などが考えられるが、省電力性という観点からセンサーはできる限り少ないほうが望ましい。また、監視を行う移動ロボットの場合、巡回中の映像を保存する機能が求められる。そこで、他のセンサーに比べ得られる情報量の多いカメラのみを使用し、カメラから得られる画像を処理することで自己位置補正を行うことを考える。

本研究では、物体認識によく用いられる SIFT 特徴量⁽⁵⁻⁶⁾を用いてシーン認識を行い、巡回する移動ロボットの自己位置補正に利用できる処理をソフトウェア上で構築・実現することを目指す。SIFT 特徴量は画像のスケール変化や回転、照明変化の影響が少なく、安定した特徴点検出とマッチングが行えるため、環境地図の作成に利

* Corresponding author. E-mail: thoma@es.ise.kyutech.ac.jp

^a 九州工業大学

〒804-8550 福岡県北九州市戸畑区仙水町 1-1

Kyushu Institute of Technology

1-1, Sensui, Tobata, Kitakyushu, Fukuoka, Japan 804-8550

用する研究も行われている⁽⁷⁻⁸⁾。したがって SIFT 特徴量は自己位置補正に有用であると考える。

2. SIFT 特徴量

Scale Invariant Feature Transform(SIFT)は David G. Lowe が 1999 年に”Object Recognition from Local Scale-Invariant Features”⁽⁵⁾で発表した手法で、画像のスケール変化と回転に不変な特徴量を記述することができる⁽⁵⁾。SIFT 特徴量はスケール変化などに不変なため、画像マッチングや物体認識に利用されている。本章では SIFT 特徴量を算出する手順について述べる。

SIFT 処理は、特徴点(キーポイント)の検出と特徴量記述の大きく 2 つの段階に分けられる⁽⁶⁾。SIFT のキーポイントは周辺領域に濃淡変化を多く持つ点が選ばれる。SIFT 処理の大まかな流れを以下に示す。

(a) detector

- ・スケールと極値の検出
- ・キーポイントのローカライズ

(b) descriptor

- ・オリエンテーションの算出
- ・特徴量の記述

2.1 スケールと極値の検出

スケール探索を行う方法として、Scale-normalized Laplacian-of-Gaussian(以下、LoG)というガウシアンフィルタを用いたスケールスペースが提案されている。LoG オペレータは、ガウシアンフィルタで平滑化した画像に 2 階微分であるラプラシアンフィルタを作用させる。LoG は、ガウシアン⁽⁷⁾のスケール σ を変化させた画像に対し LoG オペレータを適用し、その出力値が極大となるスケールを検出するという手法である。しかし、LoG は計算コストが大きいという欠点があるため、SIFT では LoG の近似である Difference-of-Gaussian(以下、DoG)を用いる。

(a) Difference-of-Gaussian

DoG 画像は入力画像を異なるスケールで平滑化した画像の差分で、ガウス関数 $G(x, y, \sigma)$ 、入力画像を $I(u, v)$ 、平滑化画像を $L(u, v, \sigma)$ とすると DoG 画像 $D(x, y, \sigma)$ は以下のように表される。

Fig. 1. Example of DoG computation.

$$G(x, y, \sigma) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right) \quad (1)$$

$$L(u, v, \sigma) = G(x, y, \sigma) * I(u, v) \quad (2)$$

$$D(u, v, \sigma) = (G(x, y, k\sigma) - G(x, y, \sigma)) * I(u, v) \\ = L(u, v, k\sigma) - L(u, v, \sigma) \quad (3)$$

この処理を k 倍ずつ増加させた異なるスケール間で行う。DoG 画像作成の例を Fig. 1 に示す。 $L(u, v, k\sigma_0)$ と $L(u, v, \sigma_0)$ の差分から DoG 画像が 1 枚生成され、ほかのスケールも同様にして DoG 画像を複数枚作成する。

ガウシアンフィルタのウィンドウサイズが大きくなると計算コストの増加などの問題が発生する。そこで SIFT では画像のダウンサンプリングを行う。Fig. 2 にダウンサンプリング時のイメージ図を示す。

画像 I をスケール $2\sigma_0$ で平滑化した画像 $L_1(2\sigma_0)$ と、画像 I を $1/2$ のサイズにダウンサンプリングしスケール σ_0 で平滑化した画像 $L_2(\sigma_0)$ は式(4)の近似が成り立つ。⁽⁹⁾

$$L_1(2\sigma_0) \approx L_2(\sigma_0) \quad (4)$$

この近似を利用し計算コストの増加を抑えている。

(b) 極値検出

スケールと極値検出のイメージ図を Fig. 3 に示す。SIFT のキーポイント候補である極値検出は、隣り合うスケールの DoG 画像 2 枚と注目画素のある DoG 画像の 3 枚の DoG 画像を 1 セットにして行う。注目画素と周囲の 26 画素(同スケールの DoG 画像の 8 画素 + 異なるスケールの DoG 画像の 18 画素)を比較し、極値(極大値もしくは極小値)であればキーポイント候補とする。異なるスケールの DoG 画像と比較することで空間だけでなくスケールも考慮した特徴点の算出ができる。

2.2 キーポイントのローカライズ

2.1 で検出されたキーポイント候補の中にはノイズや開口問題の影響を受けやすい点も含まれているため、低コントラストの点やエッジ上の点を除外することで安定したキーポイントを得る。加えて、サブピクセルを推定することでキーポイントのより正確な場所とスケールの特定を行うことができる。

(a) エッジ上にあるキーポイント候補の除外

エッジ上の点は直線線上の対応点を見つけないことができない開口問題などの影響を受けやすい。そこで、へ

Fig. 2. Example of down sampling.

Fig. 3. Example of detecting the local extrema.

ッセル行列 H を用いてエッジ上にあるキーポイント候補点の除外を行う。キーポイント候補点のヘッセ行列 H は次式で定義される。

$$H = \begin{bmatrix} D_{xx} & D_{xy} \\ D_{xy} & D_{yy} \end{bmatrix} \quad (5)$$

D はキーポイント候補点の DoG 画像画素値の 2 階微分を表す。

ここでヘッセ行列 H の固有値を α, β ($\alpha < \beta$) とすると、対角成分の和 $\text{Tr}(H)$ と行列式 $\text{Det}(H)$ は以下ようになる。

$$\text{Tr}(H) = D_{xx} + D_{yy} = \alpha + \beta \quad (6)$$

$$\text{Det}(H) = D_{xx}D_{yy} - (D_{xy})^2 = \alpha\beta \quad (7)$$

γ を 2 つの固有値の比とし、 $\gamma = \alpha/\beta$ とすると主曲率は次式で表せる。

$$\frac{\text{Tr}(H)^2}{\text{Det}(H)} = \frac{(\alpha + \beta)^2}{\alpha\beta} = \frac{(\gamma + 1)^2}{\gamma} \quad (8)$$

エッジ上の点では主曲率が大きくなるため、閾値処理を行い主曲率が閾値以下の点を厳選し、閾値より大きい点 (エッジ点) は除外する。

(b) キーポイント候補のサブピクセル推定

3 変数 (x, y, σ) の 2 次関数でフィッティングを行うことで、キーポイント候補点の位置とスケールをより正確に特定することができる。ある候補点 $x = (x, y, \sigma)^T$ の DoG 関数 $D(x)$ をテイラー展開すると次式のようになる。

$$D(x) = D + \frac{\partial D}{\partial x} x + \frac{1}{2} x^T \frac{\partial^2 D}{\partial x^2} x \quad (9)$$

上式の x に関する偏導関数を求め、それを 0 とし解くと次式を得る。 $\hat{x}$ はサブピクセルの位置を表す。

$$\hat{x} = -\frac{\partial^2 D}{\partial x^2}^{-1} \frac{\partial D}{\partial x} \quad (10)$$

(c) 低コントラスト点の除外

周りの画素とのコントラストが小さいキーポイント候補はノイズなどの影響を受けやすい。そのため、算出したサブピクセルを用いて低コントラスト点を除外する。式(10)を式(9)に代入し式(11)を得る。

$$D(\hat{x}) = D + \frac{1}{2} \frac{\partial^2 D}{\partial x^2} \hat{x} \quad (11)$$

Fig. 4. Illustration of orientation assignment.

式(11)はサブピクセル $\hat{x}$ における DoG 出力値を表す。このサブピクセルの DoG 出力値の絶対値が閾値より小さい場合低コントラスト点として除外する。

エッジ点の除外と低コントラスト点の除外処理後に残ったキーポイント候補点を特徴点 (キーポイント) とし、以降の処理を行う。

2.3 オリエンテーションの算出

特徴量の記述を行う前に、検出した各キーポイントにオリエンテーションを割り当てる。オリエンテーションとは特徴点の固有の方向であり、このオリエンテーションの向きに合わせ特徴量の記述を行うことで画像の回転に対する不変性を得る。オリエンテーション算出の例を Fig. 4 に示す。

オリエンテーションはそのキーポイントが検出された平滑化画像 $L(u, v)$ から算出される勾配強度 $m(u, v)$ と勾配方向 $\theta(u, v)$ を用いて決定される。勾配強度 $m(u, v)$ と勾配方向 $\theta(u, v)$ はそれぞれ次式で定義される。

$$m(u, v) = \sqrt{f_u(u, v)^2 + f_v(u, v)^2} \quad (12)$$

$$\theta(u, v) = \tan^{-1} \frac{f_u(u, v)}{f_v(u, v)} \quad (13)$$

$$f_u(u, v) = L(u + 1, v) - L(u - 1, v) \quad (14)$$

$$f_v(u, v) = L(u, v + 1) - L(u, v - 1) \quad (15)$$

上式を用いてキーポイント周辺領域の勾配方向のヒストグラムを作る。勾配方向ヒストグラムは 360° を 10° ごとの 36 方向に量子化し、キーポイントの持つスケール σ でガウシアンをかけ重み付けすることでキーポイントに近い勾配がより強い影響を与える。この 36 方向の勾配ヒストグラムの最大値の 80% 以上のピークをそのキーポイントのオリエンテーションとして割り当てる。

2.4 特徴量の記述

2.3 で割り当てたオリエンテーションを基準に特徴点 (キーポイント) に 128 次元の特徴量の記述を行う。特徴量記述の例を Fig. 5 に示す。

まずオリエンテーションの方向に合わせてキーポイントを中心にキーポイントのスケール σ の円が内接するような矩形領域を用いる。その矩形領域内部を 4×4 の 16 ブロックに分け、ブロックごとに勾配方向ヒストグラムを作る。この勾配方向ヒストグラムは 8 方向 (45° ずつ) で量子化されるがオリエンテーション割り当ての際と同

Fig. 5. Illustration of the local image descriptor.

様に行う。16ブロックそれぞれで8方向の勾配ヒストグラムを作成するため、 $16 \times 8 = 128$ 次元の高次元の特徴量となる。オリエンテーションに座標軸を合わせて特徴量を記述することで、画像の回転に対する不変性を付加し、さらに128次元の特徴量ベクトルの長さを特徴量ベクトルの総和で正規化することにより照明変化の影響が軽減される。

3. シーン認識

本研究は、移動ロボットの位置の補正に SIFT 特徴量を利用することを目的としている。したがって、ロボットの移動により変化した画像から現在の画像がどんなシーンであるかを正しく認識し、それに基づきロボットの位置の補正を行うことが求められる。

シーン認識は数枚の画像に対し SIFT 処理を行い、画像ごとに特徴量を算出する。そのため画像間で特徴量のマッチングを行い正しく対応点を発見する必要がある。また、マッチングを行った対応点からカメラの移動距離等の情報を検出できなければ移動ロボットの位置を補正することはできない。本章では、前章で説明した SIFT 特徴量を移動ロボットの自己位置補正に適用するための方法を述べる。

3.1 画像マッチング手法

画像ごとに SIFT 特徴量を求め、それぞれの画像で算出された各特徴点の特徴量を比較し対応点探索を行う。対応点探索は異なる2枚の画像間で算出された特徴量間のユークリッド距離を用いて行う⁽⁹⁾。このユークリッド距離が最小となる点同士を対応点として検出することで、画像間のマッチングを行う。Fig. 6 に対応点マッチングの例を示す。赤い点が特徴点、緑の線が特徴点同士の対応を表す。

画像1内のある特徴点を $k1$ 、画像2内のある特徴点を $k2$ とし、 $k1$ の特徴量を v^{k1} 、 $k2$ の特徴量を v^{k2} 、特徴量間のユークリッド距離を d とすると、 d は次式のようになる。

$$d(v^{k1}, v^{k2}) = \sqrt{\sum_{i=1}^{128} (v_i^{k1} - v_i^{k2})^2} \quad (16)$$

ある特徴点間で最小のユークリッド距離を d_1 、2番目に

Fig. 6. Example of images matching.

Fig. 7. Example of image matching result by using different similarity factor.

小さいユークリッド距離を d_2 とする。この d_1 が d_2 の A 倍以下のとき、すなわち以下の関係式を満たすとき対応点として検出する。

$$d_1 < A \times d_2 \quad (17)$$

類似度 A は1以下の正の実数である。式(17)を用いることで、対応点が見つけれなかった場合にミスマッチの発生を抑制することができる。Fig. 6 のマッチングは $A=0.7$ で行っている。Fig. 7 に A の値を変えてマッチングを行った場合の例を示す。

類似度 A を小さく設定すると検出される対応点が減少し、 A を大きく設定するとミスマッチが増加することがわかる。 A が小さい場合、ユークリッド距離がそのほかの点より大幅に小さくしなければ対応点とならないため、対応点の検出がシビアに行われる。逆に A が大きい場合、本来対応付けされるべき点が発見できなくてもユークリッド距離が最も近い点が対応点と検出されてしまうため、ミスマッチを誘発してしまう。後の章で述べるシーン認識では、 $A=0.47$ に設定し実験を行っている。

3.2 自己位置推定の概要

3.1 で述べた手法を用いることで異なる画像の特徴点同士をマッチングすることができる。本節では、そのマッチング結果に基づき自己位置判断を行う方法を示す。自己位置の判断は、目標となるテンプレート画像と入力画像の特徴点をマッチングし、その変化を検出すること

Fig. 8. Input Images for scene recognition.

Fig. 9. Illustration of photography points.

で行われる。Fig. 8 はそのイメージである。移動ロボットの自己位置補正に利用する場合、目的のルート上の画像や目的地の画像をテンプレート画像、カメラからの入力画像とし、それぞれの画像から算出した特徴点のマッチングを行うことで目的の位置と現在の位置とのズレを検出し補正を行うことができる。

本論文における実験では、テンプレート画像を 3 枚用意し、テンプレート画像を入力画像とマッチングし、最も近いと判断したテンプレート画像とどんなズレが発生したかを検出する。実験で用いた画像の撮影位置を Fig. 9 に示す。Fig. 9 は実験に使用する画像の撮影を行った廊下を真上から眺めた図で、以降で行う説明の利便性のため、各位置で撮影した画像に図のように、テンプレート画像の 3 枚を No.1~No.3, 入力画像を IM1~IM6 と番号を割り当てた。

テンプレート画像は、ある廊下の壁面から 1.7m, 3.2m, 4.7m の位置で撮影した。テンプレート画像とマッチングを行う入力画像は、テンプレートと同様に壁面から 1.7m, 3.2m, 4.7m の位置で左右方向に 75cm だけ平行移動した位置で撮影を行い実験に使用した。実験ではある 1 枚の入力画像とテンプレート 3 枚とマッチングさせ、対応点が最も多く一致したテンプレート画像を最も近い画像として選択し、2 枚の画像間の対応点の座標変化を算出す

Fig. 10. Illustration of photography area.

る。求めた対応点の座標変化量から横方向への移動距離を算出する際、テンプレート画像の撮影した位置までの距離(1.7m, 3.2m, 4.7m)は既知の情報であるとして処理を行った。

3.3 画像変化の検出手法

3.2 でも述べたように自己位置を判断し正しい位置へ補正するためにはテンプレート画像と入力画像の変化を検出する必要がある。本節では、2 枚の画像から画像間の変化を検出する手法について述べる。

(a) 横移動の検出手法

左右方向の平行移動を検出するために Fig. 10 のような三角形を考える。Fig. 10 はカメラの撮影領域を示している。画角が θ のカメラで撮影を行う場合、距離 Z だけ離れた位置で幅 L の範囲が撮影されるとき次式が成り立つ。

$$L = 2Z \tan\left(\frac{\theta}{2}\right) \quad (18)$$

カメラが右に W だけ移動した場合、相対的に物体が左に W だけ移動したように見える。この時画像内の物体が左にピクセル x 分移動して写っていたとすると実際の移動距離 W は、 U を画像の横方向の画素数とすると次式で求めることができる。

$$W = \frac{x}{U} \times 2Z \tan\left(\frac{\theta}{2}\right) \quad (19)$$

x には、マッチングした対応点同士の横座標の差の平均値を用いる。また、 Z に代入する値はテンプレート画像までの距離であり、これは初期条件として既知の値であるからカメラの横方向の移動距離 W を計算することができる。

4. 実験結果

第 2 章で述べた SIFT 特徴量と第 3 章で述べた手法を用いてテンプレート画像 3 枚とテンプレートの位置から平行移動した画像 6 枚をマッチングさせ画像間の変化を検出する実験を行った。本章では、各入力画像 IM1~IM6 と最も一致度の高いテンプレート、及びその際算出した

Table 1. Table of scene recognition result.

Image number	Selected template
Average of difference in corresponding points (pixel)	
Average of horizontal shifting distance (cm)	

	Distance 1.7m		Distance 3.2m		Distance 4.7m	
Right 75cm	IM1	No.3	IM3	No.2	IM5	No.3
	314.0		-113.8		-68.1	
	Left 326.4		Right 80.9		Right 70.8	
Left 75cm	IM2	No.3	IM4	No.2	IM6	No.3
	92.3		106.1		96.3	
	Left 96.0		Left 75.5		Left 100.0

Fig. 11. Images of scene recognition result.
(left : IM2, right : IM1)

対応点の座標変化を示す。

IM1～IM6 の各画像がどのテンプレートに最もマッチングし、算出した対応点の座標変化を本章で示す。本論文中的実験は、市販のデジタルカメラ (SONY Cyber-shot DSC-W610) を使用し、VGA サイズ (640×480 ピクセル) の画像を用いて行った。

Table 1 上部にシーン認識結果の表の構成を示し、下部には各入力画像に対する結果を示す。Table 1 の selected template とはその画像と最もマッチしたテンプレート画像であり、そのテンプレート画像とのマッチング結果をもとに自己位置の推定を行う。2 段目は、2 つの画像間における対応点同士の座標の差を取りそれを平均したもので、3.3 で示した式 (19) を用いて横方向の移動距離計算を行うための値である。

(a) 横移動の検出

テンプレートの位置から横方向へ 75cm 平行移動した IM1～IM6 の画像に対し処理を行った結果を Table 1 に示す。また、マッチングを行ったそれぞれの画像を Fig. 11 ～ Fig. 13 に示す。

Fig. 12. Images of scene recognition result.
(left : IM4, right : IM3)

Fig. 13. Images of scene recognition result.
(left : IM6, right : IM5)

壁から距離 1.7m 地点の画像である IM1, IM2 に距離的に最も近いテンプレートは No.1 であるが、対応点を発見できず、テンプレート No.3 とマッチしてしまった。その上、マッチングできた No.3 上の対応点は全く異なる特徴点であった為、移動距離を正しく算出することができなかった。しかし、壁から距離 3.2m 地点の画像 IM3, IM4 は距離が最も近いテンプレートである No.2 とマッチングでき、3.2 節で述べた手法を用いて計算した移動距離が右に約 80.9cm(IM3)と左に約 75.5cm(IM4)となり、実測値の 75cm に近い値となった。また、壁から距離 4.7m 地点の画像 IM5, IM6 では移動距離が右に約 70.8cm(IM5)と左に約 100.0cm(IM6)と算出され、IM3, IM4 の結果に比べ実際の移動距離と大きな差が出てしまった。

5. 考察

第 4 章で示した横方向移動距離の実測値に対する計算結果の相対誤差を求めた。その結果を Table 2 に示す。壁からの距離が 3.2m の IM3, IM4 では相対誤差がそれぞれ

Table 2. Table of relative error.

IM3	IM4	IM5	IM6
7.9 %	0.6%	-5.6%	33.4 %

約 7.9%,0.6%と 10%以下で計算することができた。しかし、IM6 では約 33.4%の大きな誤差が生じてしまった。この原因として IM5, IM6 と一致したテンプレート画像である No.3 の画像がわずかに右方向に移動した場所で撮影されたという点がある。本来 IM5 と IM6 の位置の中央で撮影されるべき画像が IM5 寄りに撮影されたため、IM5 の移動距離は小さく、IM6 の移動距離は大きく算出されたと考えられる。また、IM5, IM6 の画像は壁からの距離が 4.7m であり、IM3, IM4 に比べ距離が遠いためカメラのオートフォーカス機能によりカメラの画角に変化があった恐れがある。

今回実験で使用した入力画像とテンプレート画像は異なる日時に撮影されたものであるため照明状態が異なっている。しかし、SIFT 特徴量を用いたことで、多少照明が変化しても特徴点をマッチングすることが可能である。したがって、廊下などの一般的な環境で 사용할ことができる。加えて、壁面から 3.2m 地点の IM3, IM4 での相対誤差は 10%未満であり、実距離にすると誤差は 7.5cm 未満である。大きさが数十 cm の移動ロボットが幅 1.5m の廊下を走行すると想定した場合、この誤差は許容できる程度だと考える。一方で、4.7m 地点での誤差は大きく、改善する必要がある。

今回の実験では同じ視線方向での前後移動と横方向移動に対してテンプレートマッチングを行うことができた。本提案手法では、同様のシーン画像に対して撮影位置の変化を判断することができるため移動ロボットの自己位置の修正(補正)に適していると考えられる。

6. おわりに

自律移動ロボットナビゲーションにおいて、実用的かつ効果的な巡回行動を実現するためには、ロボットに予め行動エリアの地図を持たせておく方法⁽¹⁾がとられることが多い。その一方で、柔軟な巡回行動を達成するためには、ロボットの環境認識能力の向上が求められる。本研究で行った SIFT 特徴量を用いたシーン認識機能はより自由度の高いロボットナビゲーションにとって有効であるといえる。

また、SIFT 特徴量を用いたロボットナビゲーションとして、特徴点の座標を主に利用する手法^(2-3,7)が提案されているが、本研究の様に視覚的理解の能力を実現するものではなかった。したがって、本研究で構築したシーン認識機能を移動ロボットに持たせることはサービスロボットの適用範囲の拡大につながると考えられる。

今後の課題として、精度の向上とステレオカメラ化、そして処理のリアルタイム化がある。今回用いた手法では、壁からの距離が大きくなると移動距離の計算値に含まれる誤差も大きくなる。したがって、目的地からより離れた位置での自己位置補正を行う場合精度の改善が必要となる。また、本論文では単眼で実験を行ったが、ステレオカメラに拡張することで画像から物体までの距離や障害物の判断を行う⁽¹⁰⁾。加えて、今回は SIFT 特徴量を用いたが、SURF 等のより高速な手法へ置き換え、処理の高速化を目指す。さらに、高速化した処理を FPGA 実装することにより特徴量検出とシーン認識機能のリアルタイム化を行い、移動ロボットへ搭載することを目指す。

文 献

- (1) T. Yukawa, O. Hosoya, N. Saito, H. Okano : "Navigation of a care and welfare robot using cell map information", Proc. of the IASTED Int. Conf. Intelligent Systems and Control (ISC 2005), pp. 43-48, 2005.
- (2) Stephen Se , Ho-kong Ng , Piotr Jasiobedzki , Tai-jing Moyung : "Vision based modeling and localization for planetary exploration rovers", 55th International Astronautical Congress, pp. 1-11, 2004
- (3) Stephen Se, David G. Lowe, James J. Little : "Vision-Based Global Localization and Mapping for Mobile Robots", IEEE Transactions on Robotics, pp. 364-375, 2005
- (4) 城殿清澄, 三浦純, 白井良明 : 「誘導による移動経験に基づく視覚移動ロボットの自律走行」, 日本ロボット学会誌, Vol.19, No. 8, pp. 1003-1009, 2001
- (5) David G. Lowe : "Object Recognition from Local Scale-Invariant Features.", Proceedings of the International Conference on Computer Vision, Volume 2, pp. 1150-1157, 1999
- (6) David G. Lowe : "Distinctive Image Features from Scale-Invariant Keypoints..", International Journal of Computer Vision, Volume 60, Issue 2, pp. 91-110, 2004
- (7) Stephen Se, David G. Lowe, Jim Little : "Vision-based Mobile Robot Localization And Mapping using Scale-Invariant Features", Proceedings of the IEEE International Conference on Robotics and Automation, Volume 2, pp. 2051-2058, 2001
- (8) David Schleicher, Luis M. Bergasa, Manuel Ocaña, Rafael Barea, Elena López : "Real-Time Stereo Visual SLAM in Large-Scale Environments based on SIFT Fingerprints", IEEE International Symposium on Intelligent Signal Processing, pp. 1-6, 2007
- (9) 八木康史, 斎藤英雄 編, 藤吉弘亘, 山下隆義, 岡田和典, 前田英作, ノジクヴァンソン, 石川尋代, ドゥソルビエフランソワ 著 : 「- CVM チュートリアルシリーズ - コンピュータビジョン最先端ガイド」, 2 巻, pp. 1-21, 2010
- (10) Don Murray, Jim Little : "Using real-time stereo vision for mobile robot navigation", Journal of Autonomous Robots, Volume 8, Issue 2,

當眞 康史

2013年3月九州工業大学工学部総合システム工学科卒業。同年4月同大学大学院工学府先端機能システム工学専攻博士前期課程入学，現在に至る。画像計測，ロボットビジョンの研究に従事。

萩原 勇人

2013年3月九州工業大学工学府先端機能システム工学専攻博士前期課程修了。同年4月同大学大学院工学府先端機能システム工学専攻博士後期課程入学，現在に至る。リアルタイムロボットビジョンの研究に従事。

浅海 賢一

1992年3月九州工業大学情報工学部知能情報工学科卒業。1994年3月九州工業大学大学院情報工学研究科情報科学専攻修士課程修了。1997年3月九州工業大学大学院情報工学研究科情報科学専攻博士後期課程単位取得退学。博士（情報工学）。現在，九州工業大学大学院工学研究院基礎科学研究系准教授。組込みシステム，ロボットビジョンの研究に従事。

小森 望充

1982年3月大阪大学基礎工学部制御工学科卒業。1984年3月大阪大学大学院基礎工学研究科修了。博士（工学）。現在，九州工業大学大学院工学研究院先端機能システム工学研究系教授。超電導メカトロニクス，磁気浮上メカトロニクスの研究に従事。