

座標変換にロバストな高速画像照合法と その移動ロボットナビゲーションへの応用の改良

九州工業大学 ○平塚規友 黒木秀一

Improvement of mobile robot navigation using fast-image matching robust to coordinate transformation

Noritomo Hiratsuka and Shuichi Kurogi, Kyushu Institute of Technology

Abstract: This paper presents a fast image matching method robust to coordinate transformation. And it aims at the navigation of the mobile robot for the guide post recognition with the camera installed on the mobile robot.

1. まえがき

本論文ではノイズ、明度変化、座標変換などを含む対象画像から任意の探索画像の位置を高速に探索する画像照合法を用いる。また、この手法を用いてカメラを移動ロボットに搭載し、自己位置推定を行うための人口ランドマークと左右折などの走行指令を示すバーコードを貼ったガイドポストを用いることで、これまで屋内用移動ロボットで用いていたこの手法を新たに導入した屋外用の移動ロボットのナビゲーションに応用することを目的とする。

2. 座標変換にロバストな高速画像照合法

2.1 高速画像照合

対象画像の中から、記号など特定のパターンを探索する手法としてテンプレートマッチングが広く用いられている。最も基本的な方法として、正規化相関があり、処理が簡単なことから産業分野などで用いられているが、明度変化などにより対象画像が変化し、照合性能が低下することがある。明度変化が結果に及ぼす影響を小さくするために増加分布相関といった画像の明度値を直接用いるのではなく画像の符号化を行って照合を行う手法、また、ブロック投票処理といったテンプレートが画像を複数の小領域ごとに分割し、各小領域に照合処理を行うとともに遮蔽等の例外部分を除くように統合処理をすることで遮蔽や明度変化に対するロバスト性を持たせている。しかし、拡大縮小、回転等の座標変化に対するロバスト性があまり高くないことを示す結果が報告されている。一方、回転等の座標変

換が施された多数のテンプレート画像を効率的に探索する手法として、高次元の探索画像をKL変換により低次元の固有空間に投影した後、FFT (Fast Fourier Transform) により高速画像探索する手法が提案されている。そこで本論文では、回転、拡大縮小、射影変換等を伴う探索画像に適応でき、ノイズや明度変化にロバストな高速画像照合法としてLOG符号とVQネットを用いる。

2.2 画像の相関とFFT

まず、2次元画像 p を $p = p(x, y)$ または $p = (\mathbf{x})$ と表す。また、 $\mathbf{x} = (x, y)$ は画像中の位置を表すベクトルである。2つの整数 N と M に対して $N > M$ とし、画素数がそれぞれ N_2 と M_2 の正方形画像 $p(\mathbf{x})$ と $q(\mathbf{x})$ の相関関数は

$$R(\mathbf{x}) \equiv p * q \equiv \sum_{i \in I_q} p(\mathbf{x} + \mathbf{i})q(\mathbf{i})$$

で定義されるがこれはフーリエ変換を用いて

$$p * q = F^{-1}(F^*(p)F(q))$$

とも計算できる。また、実際の画像 $\tilde{p}(\mathbf{x})$ と $\tilde{q}(\mathbf{x})$ が矩形画像であり各画素数が $\tilde{N}_1 * \tilde{N}_2$ と $\tilde{M}_1 * \tilde{M}_2$ のとき、 $\tilde{N}_1, \tilde{N}_2 (i=1,2)$ 以上で最小の2のべき乗の整数 N とし、 $N * N$ の画像 $p(\mathbf{x})$ と $q(\mathbf{x})$ を次のように生成するとする。まず2つの窓領域を $W_p = \{0,1, \dots, \tilde{N}_1 - 1\} \times \{0,1, \dots, \tilde{N}_2 - 1\}$ および、 $W_q = \{0,1, \dots, \tilde{M}_1 - 1\} \times \{0,1, \dots, \tilde{M}_2 - 1\}$ とし、窓関数 $w_p(\mathbf{x})$ と $w_q(\mathbf{x})$ を

$$w_p(\mathbf{x}) \equiv \begin{cases} 1 & \text{if } \mathbf{x} \in W_p \\ 0 & \text{otherwise} \end{cases}$$

$$w_q(\mathbf{x}) \equiv \begin{cases} 1 & \text{if } \mathbf{x} \in W_q \\ 0 & \text{otherwise} \end{cases}$$

で定義し, $p = \tilde{p}w_p$ と $q = \tilde{q}w_q$ を生成する. ここで各窓領域内での各画素値は正であるとする.

2.3 LOGフィルタ

LOGフィルタは,

$$\nabla^2 G(\mathbf{x}) \equiv \frac{-1}{\pi\sigma^4} \left(1 - \frac{\|\mathbf{x}\|^2}{2\sigma^2} \right) e^{-\frac{\|\mathbf{x}\|^2}{2\sigma^2}}$$

で定義される. ($\mathbf{x} = (x, y)^T$, $\|\mathbf{x}\|^2 = x^2 + y^2$) である. これはガウス関数にラプラシアンを施したものである. LOG フィルタをフーリエ変換すると,

$$F(\nabla^2 G) = \frac{-\|\mathbf{u}\|^2}{2\pi} e^{-\frac{\sigma^2\|\mathbf{u}\|^2}{2}}$$

となる. ここで $\mathbf{u} = (u, v)^T$ であり, u と v はそれぞれ x と y 方向の周波数を表す. LOG フィルタは帯域通過フィルタであり, 低周波, 高周波のノイズを除去を行う.

2.4 LOG符号を用いた任意形状画像の高速照合

本手法はテンプレート画像 $q(\mathbf{x})$ 内の任意形状の窓領域 $W_r \subseteq W_q$ 内の部分画像

$$r(\mathbf{x}) \equiv q(\mathbf{x})w_r(\mathbf{x})$$

$$w_r \equiv \begin{cases} 1 & \mathbf{x} \in W_r \\ 0 & \text{otherwise} \end{cases}$$

で定義される $r(\mathbf{x})$ が正となる領域の画像を検索画像といい, この検索画像に最も照合する位置を対象画像 $p(\mathbf{x})$ の中から検出することを目的とする. 本手法ではまず $p(\mathbf{x})$ と $r(\mathbf{x})$ の LOG 符号画像

$$p_{SLOG} \equiv \text{sgn}_{w_p}(\mathbf{x})(p_{LOG}(\mathbf{x}))$$

$$r_{SLOG} \equiv \text{sgn}_{w_r}(\mathbf{x})(q_{LOG}(\mathbf{x}))$$

を生成する. ここで

$$\text{sgn}_{w_r}(\mathbf{x})(f(\mathbf{x})) \equiv \begin{cases} 1 & f(\mathbf{x}) \geq 0 \text{ and } w_r(\mathbf{x}) = 1 \\ -1 & f(\mathbf{x}) < 0 \text{ and } w_r(\mathbf{x}) = 1 \\ 0 & w_r(\mathbf{x}) = 0 \end{cases}$$

は画像 $f(\mathbf{x})$ の正負を符号化し窓関数を施した $w_r(\mathbf{x})$ 画像を表す.

次にこれらの相関を検索画像の画素数 $|W_r|$ で正規化した評価関数

$$\begin{aligned} R_{NCSLOG}(\mathbf{x}) &\equiv \frac{1}{|W_r|} (p_{SLOG} * r_{SLOG}) \\ &= \frac{1}{|W_r|} \sum p_{SLOG}(\mathbf{x} + \mathbf{i}) r_{SLOG}(\mathbf{i}) \end{aligned}$$

を正規化 LOG 符号相関 (NCSLOG; Normalized Correlation of Sign of LOG) と呼ぶ. この値を最大にする位置 $\mathbf{x} = \mathbf{x}_d$ を検出する. この式よりこの評価関数が 1 回の相関演算で高速に計算できることが分かる. ここで窓領域 W_q と W_r について説明を行う. まず, テンプレート画像 q の窓領域 W_q は検索したい任意形状画像よりも十分に大きい矩形領域であるとする. 一方, 検索画像 r の窓領域 W_r は検索したい任意形状画像の少し外側を縁取ったものとする. さらに W_r を表す窓関数 w_r は上述のように LOG フィルタを施した後の $q_{LOG} = (\nabla^2 G) * q$ に適用するが, これは LOG フィルタを施す前に W_r を適用すると, 検索画像の縁に W_r の外の値 $w_r(\mathbf{x}) = 0$ の影響が現れるので, その影響を除くためである.

2.5 座標変換画像の照合

画像をカメラで撮影する場合, 拡大縮小, 回転, 射影変換などを伴うことが多い. そのため, 座標変換画像を探索するための探索法を用いる. 画像 $p(\mathbf{x})$ に座標変換 $h(\mathbf{x}) = (x_h(\mathbf{x}), y_h(\mathbf{x}))^T$ を施した画像 $p(h(\mathbf{x}))$ について考える ($x_h(\mathbf{x})$ および $y_h(\mathbf{x})$ は $\mathbf{x} = (x, y)^T$)

$$p(h(\mathbf{x})) = \iint \delta(x_h(x, y) - u) \delta(y_h(x, y) - v) p(u, v) du dv$$

ここで δ はデルタ関数である. また本手法では主に次式で定義される射影変換の検討を行う.

$$h(\mathbf{x}) = \frac{\mathbf{A}\mathbf{x}}{\mathbf{b}^T \mathbf{x} + 1}$$

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$$

であり, $a_{11}, a_{12}, a_{21}, a_{22}, b_1, b_2 \in R$ である.

2.6 座標変換に対する照合関数の相似性

検索画像 $r(\mathbf{x})$ が座標変換 $h(\mathbf{x})$ を伴って対象画像 $p(\mathbf{x})$ に含まれる場合を考える. これを検討するためのひとつの尺度として次の関係を考える. すなわち $p(\mathbf{x})$ と $r(\mathbf{x})$ の相関を $R_{NCSLOG}(\mathbf{x})$ とし, $p'(\mathbf{x}) = p(h(\mathbf{x}))$ と $r'(\mathbf{x}) = r(h(\mathbf{x}))$ の相関を $R'_{NCSLOG}(\mathbf{x})$ とするとき, それらに相似な関係

$$R'_{NCSLOG}(\mathbf{x}) = t_{h(\mathbf{x})}(R_{NCSLOG}(\mathbf{x}))$$

が成立するときとこの照合結果は座標変換に影響されないといえる. さらにこの相似関係が成立すれば座標変換に不変な検出が行える. 線形変換 $h(\mathbf{x}) = \mathbf{A}\mathbf{x}$ を考える. 対象画像 $p(\mathbf{A}\mathbf{x})$ に LOG フィルタを施した LOG 画像を連続系で表すと

$$p_{LOG}(\mathbf{A}\mathbf{x}) \Big|_{\nabla^2 G(\mathbf{A}\mathbf{x})} = \int p(\mathbf{A}\mathbf{x} + \mathbf{A}\mathbf{u}) \nabla^2 G(\mathbf{A}\mathbf{u}) d\mathbf{u} \\ = |\mathbf{A}| \left[\int p(\mathbf{x}) \nabla^2 G(\mathbf{x}) d\mathbf{x} \right]$$

となる. ここで $|\mathbf{A}|$ は \mathbf{A} の行列式であり, $|\mathbf{A}| > 0$ である. よって対象画像の LOG 符号画像は

$$p_{LOG}(\mathbf{A}\mathbf{x}) \Big|_{\nabla^2 G(\mathbf{A}\mathbf{x})} = |\mathbf{A}| \left[p_{SLOG}(\mathbf{x}) \Big|_{\nabla^2 G(\mathbf{x})} \right]$$

となる. この式は $p(\mathbf{A}\mathbf{x})$ の LOG 符号画像(左辺)と $p(\mathbf{x})$ の LOG 符号画像を座標変換した画像(右辺)が等しいことを表す. 同様の関係は検索画像についても成立し, 従って相似関係式が成立する. 次に射影変換 $h(\mathbf{x}) = \mathbf{A}\mathbf{x}/(\mathbf{b}^T \mathbf{x} + 1)$ を考え, $p(h(\mathbf{x}))$ に $\nabla^2 G(h(\mathbf{x}))$ を施した LOG 画像は

$$p_{LOG}(h(\mathbf{x})) \Big|_{\nabla^2 G(h(\mathbf{x}))} = \int p \left(\frac{\mathbf{A}\mathbf{x} + \mathbf{A}\mathbf{u}}{\mathbf{b}^T \mathbf{x} + \mathbf{b}^T \mathbf{u} + 1} \right) \nabla^2 G \left(\frac{\mathbf{A}\mathbf{u}}{\mathbf{b}^T \mathbf{u} + 1} \right) d\mathbf{u}$$

ただし, $|\mathbf{b}^T \mathbf{u}| \ll 1$ とし, $\mathbf{x}' = \mathbf{A}\mathbf{x}/(\mathbf{b}^T \mathbf{x} + 1)$, $\mathbf{u}' = \mathbf{A}\mathbf{u}/(\mathbf{b}^T \mathbf{x} + 1)$ とおく. $d\mathbf{x}' = |\mathbf{A}| d\mathbf{u}/(\mathbf{b}^T \mathbf{x} + 1)$ となるため,

$$p_{LOG}(h(\mathbf{x})) \Big|_{\nabla^2 G(h(\mathbf{x}))} \cong (\mathbf{b}^T \mathbf{x} + 1) |\mathbf{A}|^{-1} \left(p_{LOG}(\mathbf{x}) \Big|_{\nabla^2 G(\mathbf{x})} \right)$$

が得られる. ただし $\nabla^2 G_{\sigma}$ は $\nabla^2 G$ の広がりを表す σ として元の σ の代わりに位置 \mathbf{x} に依存する $\sigma' = \sigma/(\mathbf{b}^T \mathbf{x} + 1)$ を使用したものである. 通常, \mathbf{x} の範囲は原点を中心とした $\mathbf{b}^T \mathbf{x} + 1 > 0$ の範囲を扱うので $(\mathbf{b}^T \mathbf{x} + 1) |\mathbf{A}|^{-1} > 0$ となり,

$$p_{SLOG}(h(\mathbf{x})) \Big|_{\nabla^2 G(h(\mathbf{x}))} \cong^T h(\mathbf{x}) \left(p_{LOG}(\mathbf{x}) \Big|_{\nabla^2 G_{\sigma}} \right)$$

が得られる. 同様の関係は検索画像にも成立するので, 相似関係式が近似的に成立するが, この結果の導出には $\nabla^2 G_{\sigma}$ を使用する必要がある. 一方, 検索画像 $r(\mathbf{x})$ の窓領域の大きさが小さく, その領域で $|\mathbf{b}^T \mathbf{x}| \ll 1$ が成立するときは射影変換画像 $r(h(\mathbf{x}))$ は線形変換画像 $r(\mathbf{A}\mathbf{x})$ で近似できるので $\nabla^2 G_{\sigma}$ ではなく $\nabla^2 G$ を使用しても相似関係式が成立する. 以上の導出過程を見ると, 射影変換を施した $\nabla^2 G h(\mathbf{x})$ があれば対象画像 $p(h(\mathbf{x}))$ から $p_{SLOG} \Big|_{\nabla^2 G h(\mathbf{x})}$ が生成でき, さらに $p(h(\mathbf{x}))$ と $w(h(\mathbf{x}))$ があれば $r_{SLOG} \Big|_{\nabla^2 G h(\mathbf{x})}$ が生成でき, これらを用いて相似関係式を満足する位置 $h(\mathbf{x}_d)$ を検出できる.

2.7 VQ ネットを用いた座標変換画像の照合

対象画像 $p(\mathbf{x})$ の中に座標変換 $h(\mathbf{x}) \in H$ により変形された検索画像 $r(h(\mathbf{x}))$ があり, その位置を検出することを考える. ここで座標変換集合 $H = \{h_j(\mathbf{x}) | j = 0, 1, 2, \dots\}$ のパラメータの範囲はあらかじめ分っており, H の要素は一般には無限個あるものとする. 以上の仮定の下では検索すべき画像は理論的には連続無限個あるが, 実際にはそれらを量子化した有限個の検索画像を用いれば十分である. さらにある座標変換 $h_j(\mathbf{x})$ に対して前節で示した画像照合を行うにはテンプレート画像 $q_j = q(h_j(\mathbf{x}))$ 窓関数 $w_j = w_r(h_j(\mathbf{x}))$ および LOG フィルタ $v_j = \nabla^2 G(h_j(\mathbf{x}))$ があればよい. そこで本手法では次のような VQ ネットを用いる. まずこの VQ ネットは N_c 個のユニットから成り, 各ユニット $i (i = 1, 2, \dots, N_c)$ はテンプレート画像 q_i 検索窓 w_i および LOG フィルタ v_i を記憶ベクトルとして学習する. この学習は競合再初期化学習法 (CRL) を用いる. この学習後 VQ ネットの記憶 (q_i, w_i, v_i) を用いて $R_{NCSLOG}(\mathbf{x})$ を計算する. ここでテンプレート画像 q_j と対象画像 $p(\mathbf{x})$ に対して, 検索画像の窓関数と LOG フィルタはそれぞれ w_i と v_i を用いる. すべての $i = 1, 2, \dots, N_c$ とすべての位置 $\mathbf{x} \in W_p$ に対して最も大きな $R_{NCSLOG}(\mathbf{x})$ の値を示す位置 $\mathbf{x} = \mathbf{x}_d$ を検出位置とする.

2.8 LOG 符号画像の座標変換に対するロバスト性

ここで LOG 符号画像が微小な座標変換に対してロバスト性を持つことを示す. まず W_p 内で座標変換が $h(\mathbf{x}) \cong \mathbf{A}\mathbf{x}$ で線形近似できると仮定し,

LOG 画像 $q_{LOG}(\mathbf{Ax}) = q(\mathbf{Ax}) * \nabla^2 G(\mathbf{Ax})$ と $q_{LOG}((\mathbf{A} + \delta\mathbf{A})\mathbf{x}) = q((\mathbf{A} + \delta\mathbf{A})\mathbf{x}) * \nabla^2 G((\mathbf{A} + \delta\mathbf{A})\mathbf{x})$ の違いを検討する. 画像 q が位置 \mathbf{Ax} の近傍の影響範囲で2次近似できるとすると,

$$\begin{aligned} q_{LOG}((\mathbf{A} + \delta\mathbf{A})\mathbf{x}) &\cong \frac{1}{|\mathbf{A} + \delta\mathbf{A}|} \int (z^T [q^{(2)}(\mathbf{Ax})] z)^2 G(z) dz \\ &\cong \frac{|\mathbf{A}|}{|\mathbf{A} + \delta\mathbf{A}|} q_{LOG}(\mathbf{Ax}) \end{aligned}$$

が得られる. よって両辺を2値に符号化すると

$$r_{LOG}((\mathbf{A} + \delta\mathbf{A})\mathbf{x}) \cong r_{SLOG}(\mathbf{Ax})$$

が得られる. 従って画素値が画像上の位置の変化に対して1次や2次で緩やかに変化する部分が多い画像では原画像よりもLOG符号画像 $r_{SLOG}((\mathbf{A} + \delta\mathbf{A})\mathbf{x})$ の方が $\delta\mathbf{A}$ の影響を受けにくい. つまりLOG符号画像を用いる方がVQネットが必要とするユニット数を少なくできる. 実際に2.で示す手法を用いることで正規化相関, 選択的正規化相関よりも少ないユニットで高い照合率を示している.[1]

3. 用いる移動ロボットとその誘導方法について

これまで2.で示した手法を用いて自己位置推定を行うためにガイドポスト(Fig. 3)に貼られた4つの特徴点をもつ人工ランドマークと右左折などの走行指令を示す2bitのバーコードを貼ったガイドポストを複数使用し, 屋内用移動ロボット(Fig. 1)に搭載したカメラでこれらを読み込むことでナビゲーションを行っていた. そこでこの手法をカメラを搭載した屋外用移動ロボット(Fig. 2)に応用し, さらにロボットのオドメトリ情報を用いて屋外でのロボットの誘導を行うシステムを開発している.

4. 結論

現在このカメラシステムおよび屋外用移動ロボットシステムは開発中であるがこれらのシステムが実際に用いることができればランドマークの特徴点を探索することで高速な自己位置推定を行うことができると考えられる.



Fig.1. iRobot ATRV-mini



Fig.2. MobileRobots P3-AT



Fig.3 . Guidepost involving artificial landmark (4 marks on the corners) for localization and a barcord at the center for indicating the direction to go.

参考文献

- [1] 渕川, 黒木, 西田: ラプラシアン・ガウシアン符号とベクトル量子化ネットを用いる座標変換にロバストな高速画像照合法, 電子情報通信学会論文誌. D-II, 情報・システム, II-パターン処理 J87-D-II(10), 1940-1950, 2004-10-01
- [2] 渕川, 黒木, 松尾, 宮本, 西田: ガイドポストと一台のカメラを用いる移動ロボットナビゲーション, 計測自動制御学会論文集 42(1), 62-69, 2006-01-31