

全方位画像を用いた記憶に基づく位置推定

粟飯原述宏[†] 岩佐 英彦^{††} 横矢 直和^{††} 竹村 治雄^{††}

Memory-Based Self-Localization Using Omnidirectional Images

Nobuhiro AIHARA[†], Hidehiko IWASA^{††}, Naokazu YOKOYA^{††}, and Haruo TAKEMURA^{††}

あらまし 実環境において自己位置を認識することは、ロボットの視覚誘導等において重要な課題である。本論文では、全方位画像を用いた記憶に基づく自己位置推定法を提案する。提案手法では、まず、位置推定に有効な大局的な情報を含む全方位画像から自己関連画像を生成することによって、センサの向きに依らない位置に固有な情報を抽出する。さらに、生成された自己関連画像から固有空間を構成し、固有空間内で最も近接した学習画像を検索することによって位置推定を行なう。また、大局的および局所的という2段階の位置推定処理を施すことによるロバストな位置推定手法も提案する。実際に屋内および屋外環境で撮影された時系列の全方位画像を用いて実験を行ない、提案手法の有効性を確認した。

キーワード 自己位置推定, 全方位画像, 自己関連, 固有空間

1. まえがき

ロボットが自律的に行動するための最も重要なタスクの1つとして、自己位置の認識がある。位置の認識を正確に行なうことができなければ、ロボットを正しくかつ安全に目的地まで誘導することはできない。ロボットの位置を認識するためには周囲の環境情報を獲得することが不可欠であり、超音波センサ、レンジファインダ、視覚センサ等の外界センサが主に使用されている。中でも視覚センサは、1度に多くの情報を獲得でき、また獲得された情報を位置推定以外にも様々な目的に利用できるため有望視されており、近年、視覚情報を用いた様々な位置推定法が提案されている。

例えば、位置情報の付加されたランドマークを予め環境内に人為的に配置しておき、獲得される画像からランドマークを抽出することによって位置を認識するという手法がある[1]。この手法は、視覚情報を用いて位置推定を行なうための最も簡単な解決法であるが、ランドマークが常に観測可能でなければならず、障害物が多く存在する様な環境で使用することは困難であ

る。また、予め環境に手を加えておかなければならないという問題もある。

環境に手を加えない方法としては、画像情報から環境の3次元構造を復元し、その距離情報を基に環境内での自己位置を認識するという手法がある。環境の3次元構造を復元するためには、全方位の動画からオプティカルフローを検出し、ロボットの自己運動パラメータと環境の3次元構造を同時に推定するもの[2]、2枚の全方位画像からステレオ視を用いて環境の奥行き情報を抽出するもの[3]等が提案されている。この手法は、未知の環境においても自己位置の認識が可能であるが、計算コストが高く、またノイズの影響を受けやすい。

環境が既知の場合には、画像情報から局所的な環境の特徴モデルを抽出し、既知の環境全体のモデルと対応させることによって環境内の自己位置の認識を行なうという手法がある。特徴モデルの全ての組合せに対してこの対応問題を解くことは計算量の点から困難であるため、実際には、同時に観測することができない特徴モデルの組合せを予め削除しておく等の、何らかの幾何的な制約を設けることによって、解の探索空間を削減する手法が用いられる[4]。この手法は、モデルの対応づけさえ正確に行なうことができれば位置推定の強力な手段になり得る。しかし、2次元画像から3次元の特徴モデルを復元する際に不良設定問題を解く

[†] ミノルタ株式会社, 大阪市

Minolta Corp., Osaka-shi, 541-0052 Japan

^{††} 奈良先端科学技術大学院大学情報科学研究科, 生駒市

Graduate School of Information Science, Nara Institute of Science and Technology, Ikoma-shi, 630-0101 Japan

必要があるため誤差が発生しやすく、また計算量も非常に大きくなるという問題がある。

以上の手法に対して、記憶に基づく認識法が注目を集めている [5] ~ [11]。記憶に基づく認識法とは、様々な場所で撮影された画像、あるいはそれらの画像情報に何らかの加工を施したものをそのまま記憶しておき、ロボットによって獲得された画像（あるいはそれを加工したもの）と記憶した情報とを直接照合することによって、位置の推定を行なうというものである。この手法は、画像の記憶容量をいかにして削減するかという問題があるが、2次元画像から3次元情報を抽出する必要が無い場合位置推定処理を高速化することが可能であり、また複雑な環境においても記憶する画像枚数を単純に増やすことで認識性能の向上が期待できる。

一般に、記憶に基づく認識法による自己位置推定においては、(1) どのような画像情報を記憶するか、(2) 画像の記憶容量や照合コストをいかにして削減するか、(3) 画像の類似性と位置の近傍性をいかにして保証するか、の3つの点が問題になると考えられる。記憶に基づく認識法を用いた位置推定法の従来手法としては、例えば、入力される視野画像列においてある閾値以上の変化がある画像のみを記憶し、テンプレートをを用いた相関演算により照合を行なう手法 [6] や、画素値の平均を特徴量とすることで大幅に画像情報を削減した全方位画像を記憶し、動的計画法を用いて画像を照合する手法 [8] がある。これらの手法は、記憶する情報量の少なさを画像の時系列性で補っているため、記憶した行動パターンと基本的に同じ（または類似した）行動をとる場合しか位置推定を行なうことができない。ロボットの行動パターンに制限を設けない方法としては、全方位画像をフーリエ変換することで位置に固有な特徴を抽出し、その低周波成分を記憶および照合に用いる手法 [5] がある。しかし、1枚につき 64×32 画素という記憶容量を必要とし、大量に記憶することは困難である。画像の記憶容量を削減するために固有空間法を用いる手法 [7] もある。この手法では、複数方向の観察結果を統合することによってロボットの位置を推定しているため、局所的に似通った景色を持つ場合でも正しく位置を推定することができるが、推定に必要な計算量が非常に多い。

本研究では、位置推定を行なうための画像として、周囲の大局的な視覚情報を含む全方位画像 [15] を使用する。ただし、ロボットの進行方向やセンサの向きによって、得られる画像が見た目上異なるため、ロ

ボットの位置に固有な特徴量を全方位画像から抽出する。抽出される特徴量は非常に大きなサイズになるので、情報圧縮の手法として広く普及している固有空間法 [12], [13] を用いて、記憶容量と照合コストを削減する。さらに、照合すなわち位置推定の段階においては、大局的および局所的と言う2段階の推定処理を行なうことにより、ロバストな位置認識を実現する。

本論文では、2章において、全方位画像の取得方法を示すと同時に、自己相関画像の作成による、位置に固有な特徴量の抽出法を提案する。続いて3章では、全方位画像から作成した自己相関画像の容量を、固有空間法を用いて削減する手法について述べる。画像の類似性を評価することで位置推定を行なう場合、画像が類似していても撮影位置が必ずしも近接しているとは限らない、という問題がある。そこで、記憶する画像集合をいくつかのクラスに分割し、入力画像がどのクラスに属するかを先に決定することによって、誤った画像検索候補を削除する。本論文では、4章で、クラス決定のために部分空間法を用いる手法について述べ、さらに、ロボットの移動の時間的空間的連続性を利用することにより、クラス決定のロバスト性を向上する手法を提案する。そして、5章において、屋内および屋外で撮影された全方位画像を用いて行なった位置推定の実験結果を通して、提案手法の有効性を検証する。最後に、6章で結論および今後の展望について述べる。

2. 位置に固有な特徴量の抽出

2.1 獲得する画像情報

位置推定を行なうための画像情報としては、通常のカメラ画像 [6], [7]、垂直スリット画像を時間軸に沿って並べたパノラマ画像 [9] 全方位画像 [5], [8] 等が使用されている。通常のカメラ画像を用いた場合、環境内のある方位に対する局所的な情報しか一度に獲得できないため、カメラ方向が異なれば、たとえ同じ位置で撮影されたものであっても画像は変化してしまう（同一通路内での行きと帰りでは撮影される画像は異なる）。従って、一方向のみを観測するだけでは、カメラ方向のずれ等により位置推定が困難になることが予想される。これを解決するためには、同じ位置で複数の方向の画像を撮影し、その観測結果を統合すればよい [7] が、位置推定に時間がかかるという問題がある。

時系列のパノラマ画像を用いた場合は、この画像自体が時系列性を含む表現になっているため、学習した

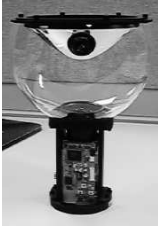


図1 視覚センサの外観
Fig. 1 An outlook of the image sensor



図2 全方位画像の例
Fig. 2 A sample of omnidirectional images

経路を再び辿るのであれば、記憶したパターンと入力されたパターンの出現順序が変化しないという拘束条件の下で安定した照合を行なうことができる。しかし、途中で停止や逆走等の動作が入ると照合は困難になる。

一方、全方位画像は周囲 360 度の視野情報を有するため、(1) 同一位置で撮影された画像は同一の情報を含む、(2) 1 枚の画像内に位置推定に必要とされる大局的な情報を含む、(3) 環境の局所的な変動による画像変化の影響を、変化していない部分の画像情報を用いて吸収できる可能性が高い、という性質を持ち、位置推定に適していると考えられる。

そこで本研究では、全方位視覚センサ HyperOmni Vision [16] を使用して取得した全方位画像を用いて位置推定を行なう。HyperOmni Vision は鉛直下向きの双曲面ミラーと鉛直上向きのカメラで構成された視覚センサで、周囲 360 度の大局的な視野領域が実時間で得られるという特徴がある。HyperOmni Vision の外観を図 1 に、実際に取得された全方位画像の例を図 2 に示す。

2.2 自己相関画像の生成

図 2 に示す全方位画像は、移動ロボットの向きが異なるような場合、すなわちセンサ軸を中心としてカメラが回転している場合には、たとえ同一位置で撮影されていても見かけ上異なる画像になってしまう。この問題に対する最も単純な解決策として、撮影位置毎に少しずつ回転させた画像を作成し、それらを全て記憶する方法が考えられるが、記憶容量や照合のための計算コストの点から望ましくない。そこで本研究では、全方位画像からセンサ位置に固有な自己相関画像を生成することを考える。自己相関画像は、撮影された 1 枚の全方位画像を少しずつ回転させ、元の画像と回転させた画像との相関値を計算することにより求めるものであり、同じ場所で撮影された全方位画像であれば、



(a) ある位置で撮像された全周パノラマ画像



(b) カメラを 120 度回転させて撮像された全周パノラマ画像

図3 全周パノラマ画像の例

Fig. 3 Samples of omnidirectional panoramic images

センサ軸まわりの回転に対して不変であるという性質を持つ。このことから、全方位画像から生成される自己相関画像の類似性を判定することにより、センサ位置の推定を行なうことができる。自己相関画像は、具体的には以下の手順により求める。

まず、撮像された全方位画像から $M \times N$ の全周パノラマ画像を生成する (図 3)。生成されるパノラマ画像は、全方位視覚センサのミラーの焦点から見た円筒面への透視投影画像である。そして自己相関画像を求めるために、図 4 に示す様に、 P 本のラインベクトルで構成されるテンプレートを全周パノラマ画像の上から下まで逐次更新していきながら、各テンプレートの水平方向への自己相関関数を計算する。具体的には、 n 番目のテンプレートを $T_n(x, y)$ ($x=1, \dots, M$, $y=1, \dots, P$) とすると、 $T_n(x, y)$ に対する自己相関関数 $r_n(k)$ は次式を用いて求める。

$$r_n(k) = \frac{\sum_{x=1}^M \sum_{y=1}^P (T_n(x, y) - \bar{T}_n)(T_n(x+k, y) - \bar{T}_n)}{\sum_{m=1}^M \sum_{y=1}^P (T_n(x, y) - \bar{T}_n)^2} \quad (1)$$

ここで、 \bar{T}_n はテンプレート $T_n(x, y)$ の全ての画素値の平均である。式 (1) は、生成される自己相関画像の n 行目の第 k 成分が、テンプレート $T_n(x, y)$ とその各成分を k 画素シフトしたものと正規化相関値であることを示している。 $(N-P+1)$ 個のテンプレ

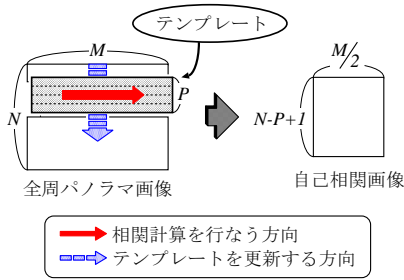


図4 自己相関画像の生成

Fig. 4 The generation of autocorrelation images

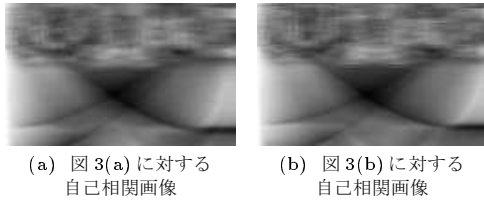


図5 自己相関画像の例

Fig. 5 The autocorrelation images generated from omnidirectional panoramic images in Fig. 3

レート全てについて式 (1) の処理を行なうことにより、 $M \times (N - P + 1)$ の自己相関画像が得られる。しかし、生成される自己相関画像の対称性のため、実際に考慮するのはその半分の領域だけでよい。この手法を用いて作成された自己相関画像を、以下では $AutoCor^{(P)}$ と表記する。

実際に図 3 の全周パノラマ画像に対して求めた自己相関画像 $AutoCor^{(1)}$ を、相関値が大きいほど輝度値が高くなるように表示したものを図 5 に示す。ここで、図 3 の 2 つの全周パノラマ画像は、同一位置で撮像されたものであるが、センサ軸回りの回転によって見かけ上異なる画像となっている。しかし、生成される自己相関画像は非常に類似していることが分かる。

3. 固有空間法による画像記憶

前節の方法により、1 枚の全方位画像から 1 枚の自己相関画像が生成される。ただし、自己相関画像を直接類似度の判定に用いると計算量が膨大になってしまう。そこで、このような自己相関画像の集合から識別に有用な特徴軸を抽出し、より簡潔に表現する。

本研究では、実環境中で空間的に連続した位置で撮影された全方位画像を、記憶する学習画像集合として使用する。このため、これらの全方位画像から生成さ

れる自己相関画像の集合は、互いに相関の強いものになっていると考えられる。こうした相関の強い画像集合から効率良く特徴を抽出する手法として、固有空間法が知られている。固有空間法は、Karhunen-Loève 展開 (KL 展開) に基づいて求められた特徴軸を用いて固有空間 (特徴空間) を構成し、その固有空間に画像を投影することによって画像データを圧縮する、線形な特徴抽出法である。固有空間法によって抽出されたこれらの特徴は、固有空間の作成に用いられた個々の画像を識別するのに有効な特徴であり、顔画像の認識 [12] やオブジェクトの姿勢推定 [13] などで高い推定性能を示している。また、実際に位置推定に固有空間法を用いた研究 [7] も報告されており、その有効性が示されている。以下では、具体的に固有空間法による画像からの特徴抽出法を示す。

L 枚の画像からなる画像集合 S を考える。各画像を、画素の輝度値を要素に持つ画像ベクトル Φ_j ($j = 1, \dots, L$) とする。画像内の画素数を n とすれば、画像ベクトル Φ_j は n 次元空間内の点 $\Phi_j = (x_1, \dots, x_n)$ として表すことができる。このとき、画像集合 S に対する共分散行列 C は、

$$C = \frac{1}{L} \sum_{j=1}^L (\Phi_j - \bar{\Phi})(\Phi_j - \bar{\Phi})^T \quad (2)$$

によって求められる。ここで、 $\bar{\Phi}$ は Φ_j ($j = 1, \dots, L$) の平均ベクトルである。画像ベクトル集合 $\{\Phi_j\}$ に対する固有空間 E は、 C に関する固有方程式

$$C \mathbf{u}_k = \lambda_k \mathbf{u}_k \quad (k = 1, 2, \dots, K, \dots, n) \quad (3)$$

$$\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_K \geq \dots \geq \lambda_n$$

を解くことによって求められる固有ベクトル \mathbf{u}_k を直交基底ベクトルとして構成される。そして、 $\nu_{nk} = (\Phi_j - \bar{\Phi})^T \mathbf{u}_k$ とするとき、画像ベクトル Φ_j は固有値の大きい上位 K 個の固有ベクトルを用いて

$$\Phi_j \approx \bar{\Phi} + \sum_{k=1}^K \nu_{jk} \mathbf{u}_k \quad (4)$$

と近似できる。

一般にこの K 次元の近似は、元の n 次元のデータとの 2 乗誤差を最小にするという意味で最適であることが知られている。このとき式 (4) は、平均ベクトル $\bar{\Phi}$ を原点とする K 次元の固有空間へ画像ベクトル Φ_j を射影することを意味している。すなわち各画像は、

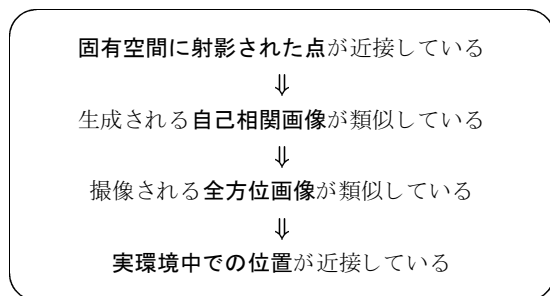
本来の n 次元の空間中の点から、次元数の低い K 次元の固有空間内の点へと変換される。この結果、 n 次元空間での類似度計算を K 次元空間での類似度計算に置き換えることができるため、 K の値を小さく設定することによって画像サイズと計算量の大幅な削減が可能となる。

4. 画像の類似性に基づく自己位置の推定

本研究では、床面あるいは地面、すなわち 2 次元平面上でのロボットの移動を考える。このため、空間的に連続した位置で撮影された学習画像集合は、互いに類似したものになっていると考えられる。一般に、類似した画像は固有空間に射影しても近接している [13] ことから、連続的に類似した学習画像集合は固有空間内において、実環境中での位置を表す x 軸方向と y 軸方向の 2 つのパラメータにより表現される学習多様体 (曲面) を構成する。このとき位置推定は、入力画像に最も近い学習多様体上の点を検出し、検出された点から実空間中の 2 つのパラメータを推定することによって行なう。

4.1 擬似画像問題

本研究では、全方位画像から生成される自己相関画像の類似性を固有空間内で判定する。このため、以下に示す $A \Rightarrow B$ (A ならば B) という論理関係のうちいずれかが成立しない場合、たとえ固有空間内で画像が近接していても実際の撮影位置は近接していないという問題 (以下、**擬似画像問題**と呼ぶ) が発生すると考えられる。



擬似画像問題が発生すると、構成される学習多様体はある一部分で交差したものになっているため、入力画像を固有空間に射影したとき、その近傍の複数の学習画像候補の中に誤った (すなわち位置推定の致命的な失敗に導く様な) 画像候補が存在し、位置を誤推定する可能性がある。(図 6(a))。しかし、入力画像に類似した画像を検出する際の検索候補から、予め誤った

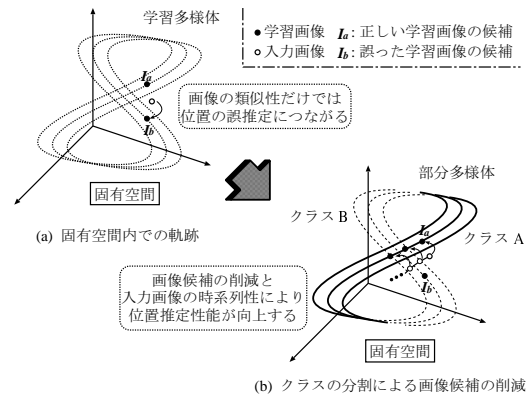


図 6 クラス分割による位置推定性能の向上

Fig. 6 Improvement of self-localization by classification

画像候補を削除することができれば位置推定の失敗を防ぐことが可能である。

このため、学習画像集合 (学習多様体) をいくつかのクラス (部分多様体) に分割しておき、入力画像の時系列性を用いて、正しい画像候補を含むクラスを予め画像検索候補から抽出することを考える (図 6(b)). ここで、入力画像として用いられる時系列画像はロボットが移動しながら撮影されるものであり、時間的・空間的に連続したものであるため、固有空間内では 1 つの連続した軌跡として表される。すなわち、過去数フレームの画像についても考慮した入力画像の軌跡を、固有空間内のある部分多様体 (部分曲面) 上の曲線として対応づけることによって、1 フレームの情報からでは得られない実環境の空間的な連続性も考慮した画像の対応づけが可能になる (図 6(b)). 従って、各クラスに何らかの特徴づけを施すことによって入力画像を正しいクラスに対応づけておき (大局的位置推定)、対応づけられたクラス内でさらに画像の類似性を判定する (局所的位置推定) ようにすれば、擬似画像問題を回避した位置認識が可能になるとと思われる。以下では、提案する 2 段階の推定法について説明する。

4.2 大局的な位置推定

本研究では、クラスの分割と対応づけに部分空間法を用いる。部分空間法とは、クラス毎にその分布を最適に近似するような部分空間を構成し、入力パターンとの距離が最小となる部分空間を検出することによって、その部分空間を構成するクラスに入力パターンを対応づけるという手法である [17]。本研究では、この

部分空間として固有空間を、また固有空間との距離の尺度として入力画像と固有空間との余弦を用いる。すなわち、入力画像ベクトルとそれを各固有空間に投影した投影ベクトルとの余弦を求め、これを最大にするようなクラスに入力画像を分類する。

まず、前節で述べた固有空間法を用いて、クラス毎に固有空間を構成しておく。一般に部分空間法は、画像をいずれかのクラスに分類するという処理を行なうだけであるので、画像集合全体の平均を基準として各固有空間を構成する [17]。しかしここでは、クラスに分類された後、そのクラスが構成する固有空間内でさらに類似画像を検索するため、固有空間を構成する際にクラス平均を基準とする方が、クラス内の画像を識別する際には望ましい。すなわち、このとき構成される各固有空間は、クラス平均からの分散が大きい軸を基底ベクトルとして構成されるため、クラスの特徴をよく表したのものとなっている。しかし、クラス毎にこれらの固有空間を構成するために各固有空間の原点が異なり、入力画像と異なる固有空間との類似度を単純に比較することはできず、各固有空間の原点を揃える必要がある。そこで、各固有空間にクラス平均の成分を含めた軸を新たに付け加える。具体的には、式 (5) で計算される基底ベクトルをクラス毎に作成し、各固有空間の軸として新たに付け加える [14]。

$$\mathbf{u}_0 = \frac{\bar{\Phi} - \sum_{k=1}^K \alpha_k \mathbf{u}_k}{\left\| \bar{\Phi} - \sum_{k=1}^K \alpha_k \mathbf{u}_k \right\|} \quad (5)$$

$$\alpha_k = \bar{\Phi}^T \mathbf{u}_k \quad (k = 1, 2, \dots, K)$$

このとき、固有空間は $(K+1)$ 個の基底ベクトルで構成されることになり、 $\rho_{jk} = \Phi_j^T \mathbf{u}_k$ とすれば式 (4) は次のように書き直すことができる。

$$\Phi_j \approx \sum_{k=0}^K \rho_{jk} \mathbf{u}_k \quad (6)$$

式 (6) により、クラス毎に構成される各固有空間の原点は 1 つに重なり、異なる固有空間との類似度の比較が可能になる。

類似度の比較の際には、類別の精度を向上させるために入力画像の時系列性も考慮する。すなわち、現在の入力画像と固有空間との距離のみを類別基準とするのではなく、過去の距離に対しても重みづけた累積類似度（累積余弦）を類別の基準とする。以下に、具体的な累積距離の計算方法を示す。

まず、時刻 t における入力画像と固有空間 E_i との距離を計算し、これを $d_i(t)$ とする。このとき、時刻 t までの累積距離を以下の漸化式を用いて更新する。

$$\begin{cases} D_i(0) = d_i(0) \\ D_i(t) = \alpha \cdot d_i(t) + (1 - \alpha) \cdot D_i(t-1) \end{cases} \quad (7)$$

ここで、 α は正規化係数 ($0 \leq \alpha \leq 1$) であり、いかなる t においても重みの和が 1 に正規化された累積距離が得られる。そして、この累積距離を基にクラス分類を行なう。

4.3 局所的な位置推定

式 (6) を用いて固有空間に射影された画像は、固有空間内では 1 つの点として表現される。そこで本研究では、固有空間内での画像の類似性の尺度として、固有空間における点間の距離を用いる。すなわち、固有空間内での 2 点間の距離が近ければ、それら 2 つの画像は類似しているとする。ここで、固有空間 $E^{(i)}$ 内の 2 点 $\psi_a^{(i)}$ と $\psi_b^{(i)}$ 間の距離にはユークリッド距離 $d_{a,b}^{(i)}$ を用いる。

$$d_{a,b}^{(i)} = \left\| \psi_a^{(i)} - \psi_b^{(i)} \right\| \quad (8)$$

局所的な位置推定は以下の様にして行なう。前節の大局的な位置推定により、入力画像はあるクラス i に分類されているとする。ここで、クラス i が構成する固有空間 $E^{(i)}$ において、 L 枚の学習画像からなる学習多様体を $\{\psi_j^{(i)} | j = 1, \dots, L\}$ とする。また、入力画像を固有空間 $E^{(i)}$ 内の点として表したものを、 $\psi^{(i)}$ とする。このとき、 $\psi^{(i)}$ と全ての $\{\psi_j^{(i)}\}$ との距離を式 (8) を用いて求め、距離が最小となる学習画像を検出する。そして、検出された学習画像に付加された位置情報により、現在の位置を推定する。

5. 屋内環境における実験

5.1 実験環境

本章では、位置推定の精度がどの程度のものか正確に調査するために、屋内環境において行なった実験について述べる。学習画像および入力画像の撮影位置を、図 7(a), (b) にそれぞれ示す。学習画像は 48 枚、入力画像は 63 枚を使用している。63 枚の入力画像のうち、32 枚は学習画像撮影位置（学習位置）と同一の位置で撮影されたもの、残りの 31 枚は 2 つの学習位置の中間点で撮影されたものを用いた。学習位置の間隔は 50cm、入力画像撮影位置（入力位置）の間隔は

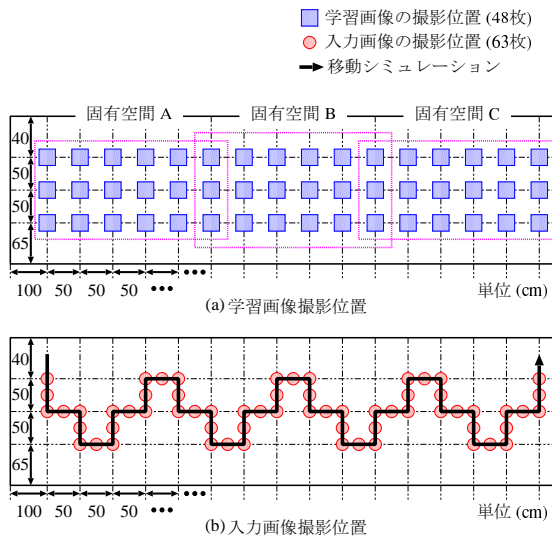


図7 全方位画像の撮影環境（建物内の通路）

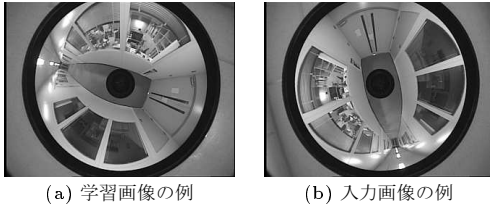


図8 実験に用いた全方位画像の例
Fig. 8 Samples of omnidirectional images for experiments

25cmであり、学習画像と入力画像の方位差（センサ軸まわりの回転）は約120度である。また、学習画像と入力画像は同時時間帯に撮影されたものであり、環境の変化（照明条件の変化等）は起きていない。実際に撮影された全方位画像の例を図8に示す。図8(a)では使用した学習画像の例を、また(b)では左の学習画像と同一位置で撮影された入力画像を示す。

5.2 実験結果および考察

まず、48枚の全方位画像（320×243、256階調）をそれぞれ256×64の全周パノラマ画像に変換した。そして、変換された全周パノラマ画像から自己相関画像を生成し、18枚ずつ3つの部分画像集合（クラス）に分けて固有空間A～Cを作成した（図7(a)）。次に、63枚の入力画像を用いて位置推定を試みた。

位置推定は、クラスの分類とクラス内での距離に基づく画像検索の2段階で行なわれる。まず、クラスの

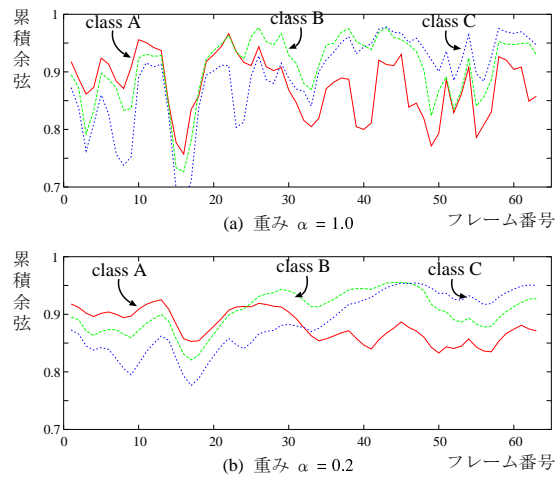


図9 入力画像と3つの固有空間との類似度（累積余弦）
Fig.9 The similarity between input images and eigenspaces

分類に用いる累積類似度（累積余弦）をグラフに表したものを図9に示す。図9では、固有空間の次元数を10（寄与率：98～99%）とした場合の各フレームにおける入力画像と各固有空間との累積余弦を、式(7)の重み α が1.0（過去の履歴は考慮しない）、0.2の場合についてそれぞれグラフで示している。入力画像は、各フレームにおいて累積余弦が最大となるクラスに分類される。このグラフから、 α の値を小さくする（過去の履歴に対する重みを大きくする）ことで累積余弦が滑らかに遷移する様になり、ノイズ等に影響されることなく、より安定したクラスの分類が可能になることが分かる。

次に、1位選択されたクラスが構成する固有空間内での距離を基に、位置推定を行なった結果について示す。表1では、重み α を0.5とし、テンプレートの幅 P を1、2、4、8、16と変化させた時の、固有空間の次元数に応じた位置推定の正解率をそれぞれ示している。さらに、各次元数において最高の正解率を太字で示している。ここでの正解率は、学習時と同一位置で撮影された画像に対してはその学習位置が正しく推定された場合を正解とし、それ以外の画像に対しては最も近接した学習位置（入力位置の両隣の学習位置）が推定された場合を正解とすることによって求めた。この正解率にはクラス分類の際の誤りも含まれている。本実験では、自己相関画像として *AutoCor*⁽¹⁶⁾ を用いた場合に、98.4%の位置推定率が得られた。この結

表1 位置推定の正解率 ($\alpha = 0.5$)
Table 1 Accuracy rates of self-localization ($\alpha = 0.5$)

		P の値				
		1	2	4	8	16
次元数	5	77.8	77.8	84.1	84.1	87.3
	10	84.1	84.1	87.3	87.3	95.2
	15	90.5	93.7	96.8	95.2	98.4

果より、提案手法が屋内等の狭い環境内での位置推定に適していることが示された。また、P の値を大きくした方が位置推定率は向上している。これは、P の値を大きくすることによって、僅かな画像変化やノイズによる影響を吸収できるためである。

6. 屋外環境における実験

6.1 実験環境

本実験では、実環境中で撮影された時系列の全方位画像を用いて提案手法の有効性を調査した。なお本実験においては、学習画像、入力画像ともにこの時系列画像を用いることにする。すなわち、ロボットの行動を1次元的な移動のみに制限している。この制限は、屋外という広い環境において、あるルートに沿って移動するロボットの位置推定という目的のためには妥当なものであると考えられる。

実験画像は、全方位視覚センサを自動車に搭載し、本学敷地内の外周路に沿って移動しながらビデオ撮影することによって取得した。実際に通過した経路を、図10に矢印によって示す。外周路の1周は約1kmであり、自動車の速度を時速20kmでほぼ一定に保ちながら、外周路に沿って2周回分撮影した。撮影されたビデオ画像を電子画像ファイルに変換し、さらに毎秒7.5枚の割合でサンプリングすることによって、1周目のデータとして1328フレーム、2周目のデータとして1455フレームの全方位画像データ(320×243, 8bit濃淡階調)を取得した。撮影間隔は、自動車の速度とサンプリングレートを考慮すると、概ね50~100cmと推定される。実際に実験に用いた全方位画像の例を、図11に示す。

6.2 実験結果および考察

位置推定の実験は、1周目のデータを学習画像集合、2周目のデータを位置を推定すべき入力画像集合として行なった。まず、画像の記憶であるが、前章で述べた様に、1328枚の学習画像集合を予めいくつかのクラスに分割しておく。分割されたクラスにも空間的な連続性を持たせるため、クラスの分割には画像集合を単

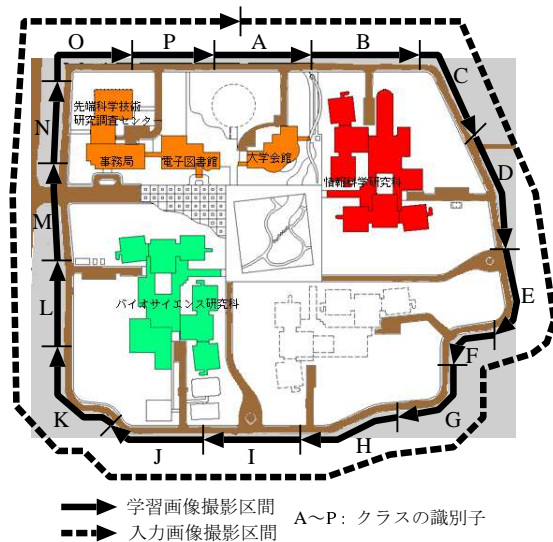


図10 全方位画像の撮影環境 (学内の外周路)

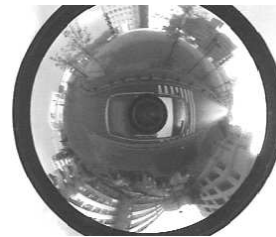


図11 実験に用いた全方位画像の例

Fig.11 A sample of omnidirectional images for experiments

純に等間隔に分割していく等分割法を用い、クラスの個数は16(クラスA~P)とした。実際に実環境領域がどのように分割されたかを、図10に実線の矢印で示す。このとき、各クラスに含まれる画像の枚数は一律83枚となるが、実際には、前後2フレームずつを隣接クラスと重複させた85枚の画像集合を用いて1つのクラスとした。1つのクラスに含まれる実環境中での自動車の移動距離は約60mと推定される。そして、これらの画像集合を256×64の全周パノラマ画像に変換した。さらに、全周パノラマ画像から自己相関画像を生成し、16個のクラス毎に固有空間を作成した。

次に、1455枚の入力画像集合を用いて位置推定実験を行なった。この入力画像が実際の経路上で対応している位置を、図10に破線で示す。また、予め目視により、入力画像に対して正解となる学習画像の対応

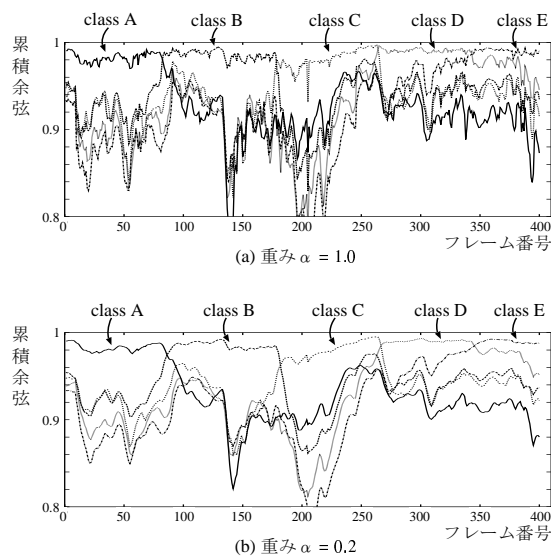


図 12 入力画像と 5 つの固有空間 (A~E) との類似度 (累積余弦)
 Fig. 12 The similarity between input images and eigenspaces(A-E)

づけを行なっておく。位置推定は、屋内環境における実験と同様に、クラスの分類とクラス内での距離に基づく画像検索の 2 段階で行なわれる。まず、クラスの分類のために必要な、入力画像と各固有空間との累積余弦をグラフに表したものを図 12 に示す。図 12 では、固有空間の次元数を 20 (寄与率: 91 ~ 96%) とした場合の、各フレームにおける入力画像と各固有空間との累積余弦を、式 (7) の重み α が 1.0, 0.2 の場合についてそれぞれグラフに示している。ただし、グラフ表示の都合上、ここでは最初の 400 フレーム (約 300m) の結果についてのみ示し、さらに 16 個のクラスのうち、この区間中のいずれかのフレームで 1 位に選択された 5 つのクラス (A~E) に対する累積余弦のみ表示している。図 12 のグラフより、過去のフレームを考慮した方がより滑らかに累積余弦が遷移しており、安定したクラス分類が可能になることが分かる。

本手法では、記憶していない画像が得られた場合、すなわち未知の環境に入り込んだ場合の処理については定義していない。しかし、図 12(b) を見ると、正しいクラスに分類されているときの累積余弦は殆どこのところ高い値 (0.98 以上) となっており、このあたりに閾値を設定することによって未知環境かどうかの判断が可能になると考えられる。

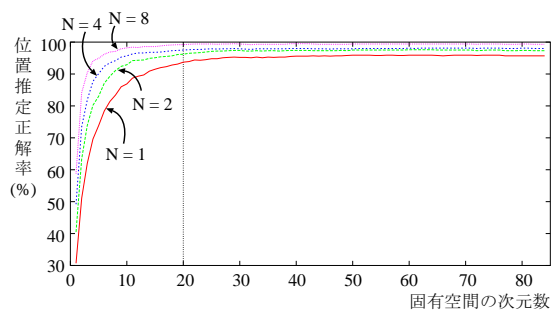


図 13 固有空間の次元数に対する位置推定の正解率 ($\alpha = 0.5$)

Fig. 13 Accuracy rates of self-localization ($\alpha = 0.5$)

さらに、上述した様に、予め目視により入力画像と学習画像 (実環境中での位置) の対応づけを行なっておき、推定された位置がその対応位置の $\pm N$ フレーム以内に入っている時を正解とした場合の正解率を以下に示す。図 13 では、自己相関画像として *AutoCor*⁽¹⁾ を用い、重み α を 0.5 とした時の固有空間の次元数に応じた位置推定正解率を、 N が 1, 2, 4, 8 の場合についてそれぞれグラフで示す。図 13 のグラフより、固有空間の次元数を 20 とした場合でも、入力に用いた 1455 枚の画像集合の約 96% が 2 フレーム (約 1.5m) 以内の誤差の範囲で、また約 99% が 8 フレーム (約 6m) 以内の誤差の範囲でそれぞれ推定されていることが分かる。また、30 次元以上では正解率の変化は殆ど生じていないため、正解率を維持したまま固有空間の次元数を大幅に削減することが可能である。以上の結果より、提案手法は、自己相関画像の 8192 次元に対して位置推定に利用する画像情報の次元数を数十次元にまで極端に削減した場合においても、非常に高い推定性能を有することが確認できた。

表 2 では、重み α を 0.5, N を 2 とし、 P の値を 1, 2, 4, 8, 16 とした時の位置推定正解率を、固有空間の次元数が 5 から 30 までそれぞれ示している。さらに、各次元数において最大の正解率を太字で示している。この結果から分かる様に、 P の値を増加させることによって正解率が低下しており、屋内環境における実験結果と比べて全く逆の結果になっている。これは、屋外で撮影された全方位画像は撮影対象までの距離が一様に長く、僅かな位置ずれでは画像の変化が殆ど生じないため、 P の値を増加させて高さ方向の画素変化による影響を吸収しようとする、逆に位置判別

表2 位置推定の正解率 ($\alpha = 0.5, N = 2$)
Table 2 Accuracy rates of self-localization
($\alpha = 0.5, N = 2$)

		P の値				
		1	2	4	8	16
次元数	5	83.0	82.9	82.5	81.1	78.0
	10	92.9	92.6	91.9	90.4	89.8
	15	95.1	94.6	94.2	93.9	92.1
	20	96.2	96.1	95.5	95.1	93.5
	25	97.1	96.6	96.0	95.4	93.2
	30	97.4	96.8	96.4	95.6	93.4

表3 クラスに含まれる画像枚数
Table 3 The number of images included in a class

クラス数	1	2	4	8	16
画像枚数	1328	666	334	168	85

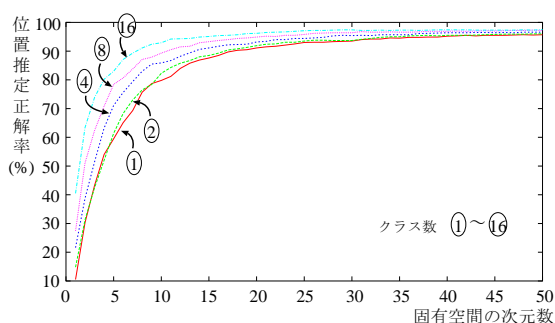


図14 クラス分割数に対する位置推定の正解率
($P = 1, \alpha = 0.5, N = 2$)

Fig. 14 Accuracy rates of self-localization for the number of class partition ($P = 1, \alpha = 0.5, N = 2$)

性能が低下してしまうことが原因であると考えられる。

6.3 クラス分割の効果

学習画像集合に対するクラス分割の効果調べるために、クラス数を1, 2, 4, 8, 16と変化させて位置推定実験を行なった。クラスの分割には等分割法を用い、各クラスに含まれる画像枚数は表3の通りである。ただし、1328枚の学習画像集合に対して等分割を行なった後、前後2フレームずつを隣接クラスと重複させることによって得られる画像集合を用いて、1つのクラスを構成している。このときの固有空間の次元数に応じた位置推定正解率を図14に示す。ここで、自己相関画像としては *AutoCor*⁽¹⁾ を用い、重み α は0.5, N は2としている。

図14の結果から分かる様に、本研究で提案する部分空間法を用いたクラス分割を行なうことによって位置推定正解率は改善している。これは、1つの固有空間

間を作成する画像枚数が増えれば増えるほど、固有空間の次元数を削減することによって発生する誤った画像候補の数が増加するためと思われる。位置推定正解率の改善は、固有空間の次元数が低いほど顕著であり、次元数が高くなると正解率にそれほど差は見られない。しかし、記憶容量を小さくするために固有空間の次元数は低くすることが望ましく、低次元の固有空間においても高い位置推定性能を有している提案手法は、この目的に適していると考えられる。

7. むすび

本論文では、全方位画像からカメラ位置に固有な自己相関画像を求め、その類似性を固有空間内で判定することによって自己位置の推定を行なう手法を提案した。本手法は、部分空間法を用いて探索空間を削減した上で、入力された画像に最も類似したものを記憶した画像の中から検索することによって位置の推定を行っており、クラス分類処理と画像検索処理を統合した位置推定法であると言える。実際に屋内および屋外において撮影された全方位画像を用いて位置推定の実験を行なった結果、どちらの場合においても提案手法が高い位置推定性能を有することを確認した。

本手法の利点として、以下の点が挙げられる。まず、記憶に基づく手法を用いているため、あらゆる環境の画像を記憶することによって環境条件に依存しない位置認識が可能である。また、場所毎に学習画像を撮影し固有空間を作成するので、新たな画像の学習が容易である。さらに、特徴量の計算に用いる自己相関関数はフーリエ変換を用いることで高速に計算でき、さらに類似画像の検索には射影等の比較的小さな計算量だけで済むため、位置推定処理の高速化が期待できる。

今後は、未知の物体が出現した場合の処理や、位置推定の高速化について検討する予定である。

文 献

- [1] C. Becker, J. Salas, K. Tokusei and J. C. Latombe: "Reliable Navigation Using Landmarks," *Proc. IEEE Int. Conf. on Robotics and Automation*, vol.1, pp.401-406, 1995.
- [2] 榎本雅幸, 三橋 渉: "全方位画像からの自己運動と3次元構造の復元," 信学技報, PRU91-126, pp.1-8, 1991.
- [3] 茶園 篤, 山澤一誠, 横矢直和, 竹村治雄: "全方位ステレオ視による3次元情報の獲得," 映像情報メディア学会誌, vol.51, no.10, pp.1745-1753, 1997.
- [4] K. T. Simsarian, T. J. Olson and N. Nandhakumar: "View-Invariant Regions and Mobile Robot Self-Localization," *IEEE Trans. Robotics Automat.*,

- vol.12, no.5, pp.810–816, 1996
- [5] 前田武志, 石黒 浩, 辻 三郎: “全方位画像を用いた記憶に基づく未知環境の探索,” 情処研報, CV92-10, pp.73–80, 1995.
- [6] 松本吉央, 稲葉雅幸, 井上博允: “視野画像列の記憶に基づく移動ロボットのナビゲーション,” 第13回日本ロボット学会学術講演会予稿集, no.1, pp.313–314, 1995.
- [7] 前田佐嘉志, 久野義徳, 白井良明: “固有空間解析に基づく移動ロボットの位置認識,” 信学論 (D-II), vol.J80-D-II, no.6, pp.1502–1511, 1997.
- [8] 小島 浩, 伊藤慶明, 岡 隆一: “Reference Interval-Free 連続 DP を利用した移動ロボットの時系列画像による位置同定システム,” 信学論 (D-II), vol.J80-D-II, no.3, pp.724–733, 1997.
- [9] J. Y. Zheng and S. Tsuji: “Panoramic Representation for Route Recognition by a Mobile Robot,” *Int. J. Computer Vision*, vol.9, no.1, pp.55–76, 1992.
- [10] 栗飯原述宏, 岩佐英彦, 横矢直和, 竹村治雄: “全方位画像を用いたイメージ記憶に基づく位置の推定,” 画像電子学会研究会予稿, 97-04-06, pp.21–24, 1997.
- [11] 栗飯原述宏, 岩佐英彦, 横矢直和, 竹村治雄: “時系列全方位画像を用いた記憶に基づく位置推定,” 信学技報, IE97-158, pp.101–108, 1998.
- [12] M. A. Turk and A. P. Pentland: “Face Recognition Using Eigenfaces,” *Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR '91)*, pp.586–591, 1991.
- [13] H. Murase and S. K. Nayar: “Visual Learning and Recognition of 3-D Objects from Appearance,” *Int. J. Computer Vision*, vol.14, no.1, pp.5–24, 1995.
- [14] M. Uenohara and T. Kanade: “Vision-Based Object Registration for Real-Time Image Overlay,” *Proc. Computer Vision, Virtual Reality and Robotics in Medicine (CVRMed '95)*, pp.13–22, 1995.
- [15] 八木康史: “全方位画像の取得とその応用,” 信学技報, PRMU96-140, pp.89–96, 1997.
- [16] 山澤一誠, 八木康史, 谷内田正彦: “移動ロボットのナビゲーションのための全方位視覚センサ HyperOmni Vision の提案,” 信学論 (D-II), vol.J79-D-II, no.5, pp.698–707, 1996.
- [17] 石川則之, 有木康雄: “正規化複数部分空間法による顔認識と話者認識の統合,” 信学技報, PRMU96-41, pp.31–38, 1996.