

# SVMを用いた屋外移動ロボットの自己位置推定

## View-Based Localization in Outdoor Environments Based on Support Vector Learning

森田英夫\*<sup>1</sup>  
Hideo MORITA

ヒルド・ミヒヤエル\*<sup>2</sup>  
Michael HILD

三浦純\*<sup>1</sup>  
Jun MIURA

白井良明\*<sup>3</sup>  
Yoshiaki SHIRAI

\*<sup>1</sup>大阪大学 大学院 工学研究科 機械工学専攻

Department of Computer-Controlled Mechanical Systems, Graduate School of Engineering, Osaka University

\*<sup>2</sup>大阪電気通信大学 情報通信工学部 情報工学科

Faculty of Information and Communication Engineering, Osaka Electro-Communication University

\*<sup>3</sup>立命館大学 情報理工学部 知能情報学科

Department of Human and Computer Intelligence, College of Information Science and Engineering, Ritsumeikan University

This paper describes a view-based localization method using support vector machines (SVM) in outdoor environments. In the training phase, the robot acquires image sequences along the desired route and automatically learns the route visually. In the subsequent autonomous navigation phase, the robot moves by localizing itself based on the comparison between input images and the learned route representation. The SVM algorithm is applied to both recognizing general scene objects and discriminating locations based on the recognition results. This two-stage SVM-based localization approach exhibits a satisfactory performance for real outdoor image data without any manual adjustment of parameters and threshold values.

### 1. はじめに

屋外環境で走る自律走行車のナビゲーションは解決すべき重要な問題になってきており、特に車両の位置推定はキーテクノロジーの一つである。これまでに多くの手法が提案されてきた。これらの手法は、環境地図を用いるかどうか(いわゆる地図なしの方法)、もしくは、GPSのような地球規模のセンサ・システムが用いられているかどうか [Thrapp 01] (即ち、非ビジョン・ベースの方法)により分類化できる。GPSに基づく手法は、ある環境状況においては信頼性が低いことが知られており、また、地図に基づく手法では、しばしば地図の作成や維持のため相当の努力を要する。以上の理由から、視覚認識に基づく手法が必要であると考えられる。

本論文で提案する手法は、完全に視覚に基づく方法である。自律走行車は、カラービデオカメラが搭載されており、それを用いて学習走行の時に望みのルートに沿った連続画像を取得し、視覚的にそのルートを自動学習し、この学習データを後の自律走行を行うために貯めておく。この手法では、地図や座標系が不要である。我々は以前に、天候や季節の変化にロバストな物体認識法を用いる位置推定法を開発した [Katsura 03]。それと似た手法も、屋内 [Kidono 02] や屋外 [Takeuchi 98] の環境で応用させるために提案されている。これらの方法の一番困難な点は、(特徴量の選択を含めて) その環境の最も適切な内部的表現を見付け、そして、その内部的表現を学習により自動的に生成してくれる学習アルゴリズムを設計するところである。これと異なる視覚に基づく方法としては、道路境界のような局所的な特徴のみに依存する手法 [Ishiguro 92] があるが、実際の屋外環境においては、これらの特徴量が必ず抽出可能であるとは言えない。

最近、サポートベクターマシンのアルゴリズム (SVM) [Vapnik 98] は、様々な物体認識問題にうまく適用されてきた。それらの応用は、3次元の物体認識、顔認識

連絡先: 森田英夫, 大阪大学大学院 工学研究科 機械工学専攻, 大阪府吹田市山田丘 1, 06-6879-7333, morita@cv.mech.eng.osaka-u.ac.jp

、パターン照合ベースの物体追跡などを含めている。それらの方法は、主成分解析ベースの手法のような従来の方法よりよい識別結果を示している。視覚に基づく位置推定の問題は、学習の時に取得された環境画像集合の中から、現在与えられている入力画像に最も類似する画像を見つけ出す問題として解釈できるため、この問題は SVM ベースの認識に適した応用問題であると考えられる。

本論文では、ルートに沿った景色を自動的に学習する視覚認識に基づくナビゲーション・システムについて述べる。ここでは、手動で設定する閾値を全く用いない。特徴抽出の段階の後、特徴ベクトルは SVM のアルゴリズムにより学習される。その後ナビゲーション・位置推定を行うとき、特徴は学習の時と同様の方法で抽出され、学習された SVM によりクラス分けされ、ルートに沿った自律走行車の位置推定を行う。

以下の節では、システム構成についてより詳細に述べ、サポートベクター学習と分類化について説明し、大学校舎内の屋外環境における実験の結果を示す。

### 2. 手法の概要

図1は SVM に基づく位置推定のプロセスを示している。このプロセスは2段階からなる。一つ目の段階では、画像中の対象物が認識される。色やエッジ密度などの特徴が入力画像から抽出され、各ピクセルごとに特徴ベクトルが生成され、特徴画像として保存される。次にこの特徴画像が複数の SVM に並列に入力される。それぞれの SVM は1つの特定の物体クラスを認識するために学習させたものである。各々の SVM の出力は画像中の認識された物体の位置を表す図となる。そして、最終の認識結果を得るために、出力されたすべての図におけるベクトルがつなぎ合わせられる。天候や季節の変化の、様々な条件のもとで取得されたそれぞれの対象クラスの画像を入力データにし、SVM を学習させることにより、天候や季節の変化による対象物の見えの変化の問題をこの段階において解決する。

第1段階の物体認識の結果を用い、第2段階において、自律走行車の位置推定を行う。指定されたルート内の1ヶ所だけを

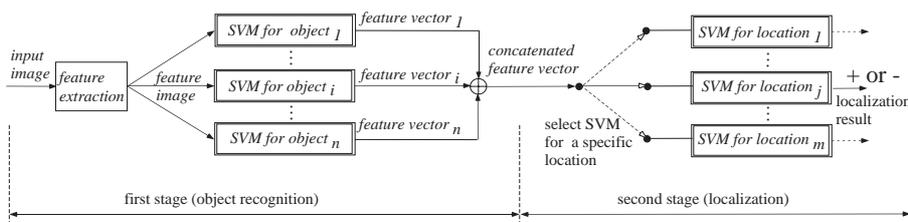


図 1: Two-stage localization using SVMs.

他のルート内の場所から識別できるように、それぞれの SVM を学習させる。この識別は、原画像ではなく、第 1 段階の認識結果（即ち、つなぎ合わせられた特徴ベクトル）に基づき行われるため、屋外環境における位置推定をロバストに行えるようになっていく。

自律走行車がある与えられている位置を通過しているかどうかを検証するとき、入力画像をその位置を推定するために学習させた SVM を応用し、その入力画像のデータを処理する。自律走行車が学習したルートに沿って移動しているとき、自律走行車は位置推定のための SVM を次々入れ換えて用いる。この場合には、一度に一つだけの SVM が使われる。



図 2: Example images.

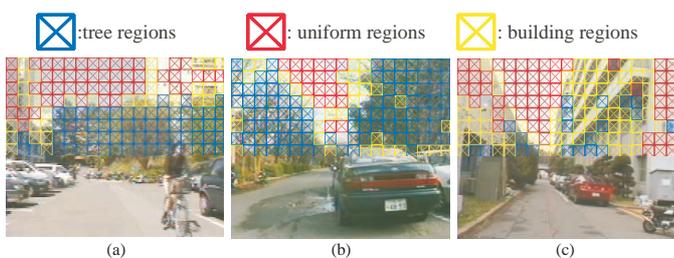


図 3: Recognition results.

### 3. 物体認識のためのサポートベクター学習

#### 3.1 認識される対象物

本論文では、大学の構内を対象環境とする。そこには、建物、木、自動車、自転車等の物体が存在する。図 2 に構内を撮影した画像例を示す。構内に存在する物体のうち、自動車、自転車などの見えが日時によって変化するため、これらの物体を位置推定のために用いず、代わりに、カメラが地面と水平に設置されていることを仮定し、画像の上半分（水平面より上）に存在するかなりの大きさを持つ建物、空、木を認識の対象とする。本論文では以下の 4 種の物体を認識する。

- 季節による葉の色の変化も考慮した葉のある木。
- 枝のみが観察される葉のない木。
- 空、建物の壁等の画像中の一様な領域。
- 建物の窓や境界線等の画像中の強い直線が観察される領域。

#### 3.2 物体認識に用いられる特徴

画像の上半分（即ち、 $304 \times 128$  ピクセル）を  $16 \times 16$  ピクセルの小ウィンドウの集合に分割し、各小ウィンドウ内の色やエッジ点という特徴量を調べ、各小ウィンドウを上記の対象物クラスの一つに属するように分類化する。木の領域に対しては、2 つの対象物のクラス（葉のある木と葉のない木）を設けているが、両方の種類の木領域に「木」というラベルを割り与える。用いられる特徴は下記のように処理される。

3.2.1  $(r, g, b)$ : これらは「正規化された色成分」である。各色成分は、 $r = R / (R + G + B)$  のように計算される。色特徴を生成するため、各ピクセルにおける正規化された色成分を小ウィンドウ全体で平均をとる。各成分は  $[0, 1]$  の範囲に正規化される。この特徴は上記の 4 つのすべての対象クラスのために使われる。

3.2.2  $f_{density}$ : 1 つの小ウィンドウ内のエッジの総数をウィンドウの面積で割った「エッジ密度」である。ここでのエッジとは、ソーベルオペレータにより計算された閾値 3 を越えるピクセルが対象となる。この閾値は、画像中のノイズの度合を観察し、決定した。この特徴も全対象クラスのために使われる。

3.2.3  $f_{distrib}$ : この特徴は「エッジ方向の分散の度合い」を循環統計学に基づき計測したものであり、葉のない木のクラスを認識するために有用である。1 つの小ウィンドウに  $n$  個のエッジがあると考え、エッジの方向や強度は  $\phi_1, \dots, \phi_n$  と  $a_1, \dots, a_n$  で表される。まず、各方向の値を単位円上の点に変換し、以下のようにエッジ強度で重みづけした重心位置  $(\bar{x}, \bar{y})$  として計算される。

$$(\bar{x}, \bar{y}) = \frac{1}{\sum_{i=1}^n a_i} \left( \sum_{i=1}^n a_i \cos 2\phi_i, \sum_{i=1}^n a_i \sin 2\phi_i \right). \quad (1)$$

$r = (\bar{x}^2 + \bar{y}^2)^{1/2}$  の値について調べると、 $r$  はエッジ方向がばらつく程小さくなり、その値は  $[0, 1]$  の範囲に収まる。そのため、 $S = 1 - r$  を方向付けされた分散として用いる。よって、 $f_{distrib} = S$  が成立する。

3.2.4  $f_{hough}$ : これは 1 つの小ウィンドウ内のエッジ点のハフ空間上での投票値の最大値である。この値は、建物の窓や境界線のような強い直線が存在する場所で大きくなる。この投票最大値を経験的に得られた 300 という値で割ることで正規化している。

上記の特徴からなる 6 次元特徴ベクトルは各小ウィンドウにおいて得られる。各入力画像は 6 次元特徴ベクトルを保存す

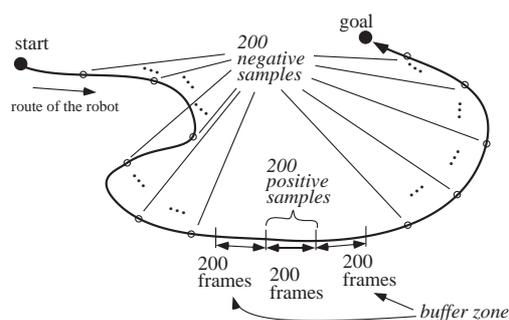


図 4: Making training data for localization SVM.

る  $19 \times 8$  配列に変換される．この配列を「特徴画像」と呼び、物体認識のための SVM の入力となる（図 1 参照）．各 SVM では、6 次元特徴ベクトルの中の 4 つの成分、あるいは、6 つの成分が使われる．

### 3.3 物体認識のための SVM の学習

各対象物クラスに対して 1 つの SVM を用いる．様々な天候条件での各季節において、大学のキャンパスで取得した画像を用い、1 つの小ウィンドウに対してただ 1 つの対象物だけが存在した小ウィンドウの約 300 個を手動で選択し、学習データとした．これらの小ウィンドウをそれぞれ 6 次元特徴ベクトルに変換した．結果としては、4 つの対象クラスに対応する、6 次元特徴ベクトルからなる集合を 4 個取得した．

1 つの与えられたクラスに対して 1 つの SVM を学習させるため、そのクラスに対応する集合に含まれているすべての 6 次元特徴ベクトルを「ポジティブ・サンプル」とし、それ以外のクラスの集合から同じ数の 6 次元特徴ベクトルをランダムに選び「ネガティブ・サンプル」とした．

各 SVM は 6 次元特徴ベクトルを入力し、もし出力が正なら 1 を、それ以外は 0 を返す．特徴画像のサイズは  $19 \times 8$  なので、SVM は 152 次元の 0,1 の 2 値をとる「特徴ベクトル」を生成する（図 1 参照）．

### 3.4 認識結果

図 3 に図 2 に示した画像の認識結果を示す．× マークのついた各ブロックは木、一様、人工物領域の認識結果を表す．これらの結果を著者らの以前の手法 [Katsura 03] と比べて、同様な結果が得られたが、本論文の新しいアプローチは手動で調節するパラメータや閾値を一切必要としない大きな利点がある．

## 4. 位置推定のためのサポートベクトル学習

第 1 段階での認識結果を用いた SVM により、第 2 段階で位置推定を行う（図 1 参照）．「木」、「一様」、「人工物」の 3 種類のラベルが存在するため、第 1 段階の出力は、3 つの特徴ベクトル（152 次元 0-1 のベクトル）となる．これら 3 つの特徴ベクトルを 1 つの 456 次元の特徴ベクトルにつなぎ合わせ、位置推定を行なう SVM の入力として用いる．

### 4.1 SVM の学習データの生成

自律走行車が移動するルートに沿って設定した各地点に対して、1 つの SVM を作成する．ある地点の近辺で取得したデータを「ポジティブ・サンプル」とし、その他の場所で取得したデータを「ネガティブ・サンプル」とし、位置推定の各 SVM を学習させる．学習データの詳細な取り方は以下のようなものである．

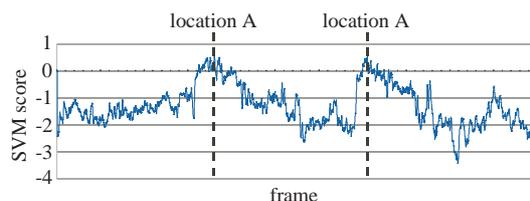
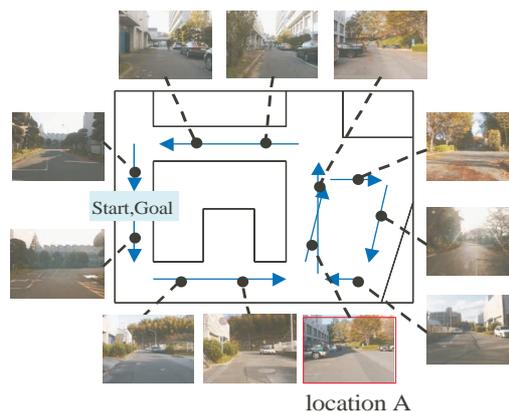


図 5: Result of a localization experiment in our campus.

自律走行車が約  $0.9 [m/s]$  のスピードで移動しながら、 $30 [frame/s]$  というレートで画像を取得する．与えられた地点に対して、ある地点の近辺で取得した連続した 200 枚の画像を選択する（これは約  $6 [m]$  の移動量に対応する）．これらの画像を「ポジティブ・サンプル」として用い、200 個の特徴ベクトルを作成する．「ポジティブ・サンプル」の前後に中性の間隔 (buffer zones) を設置し、200 個の「ネガティブ・サンプル」を作成するため、残りのフレームから規則的な間隔を置いて 200 枚の画像を選ぶ（図 4 参照）．これらのポジティブ、ネガティブなサンプルは位置推定を行なう SVM を学習させるために用いる．他の地点に対しても同様の方法で学習サンプルの選択を行い、別の SVM を学習させる．

### 4.2 SVM による位置推定

もし入力が「ポジティブ・サンプル」と判断されれば、SVM は正の値を出力する．自律走行車が与えられた地点を通過しているかどうか知るために、現在の入力画像から得られる特徴ベクトルを位置推定を行なう SVM に入力し、その出力値が正かどうか調べる．しかし、位置推定を行なう SVM からの出力にはノイズが多く含まれており、しばしば誤った正の出力が返される経験がある．このため、この出力値（超平面からの距離）を、現在のフレームを中心として 21 フレームで平均した値を用いる．これを位置推定におけるスコアと呼ぶことにする．

### 4.3 位置推定の評価基準

SVM を用いた位置推定には 2 つのケースがある．学習したルートに沿って移動しているため、自律走行車が現在、その大まかな位置を知っていれば、その予測される位置に対応した SVM（もしくは、予測した位置の近辺の 2, 3 個の SVM）に現在の画像を入力し、その出力スコアが「正」であるかを確認すればよい．逆に、自律走行車が現在の位置に関する情報を全く持っていなければ、現在の入力画像をすべての SVM に入力し、どの SVM からの出力が「正」になるかを調べる必要がある．もし複数の SVM から「正」の値が出力されれば、最も高いスコアを出力するものが現在の位置を示している可能性が最

も高い。

上記の2つのケースを考慮し、SVMに基づく位置推定の手法のために、以下の2つの統計的な指標を用いる。

1. 認識率：設定した地点の総数に対する、SVMにより正しく認識された地点の数の比率。これは前者のケースに適用する。
2. 最大得点率：設定した地点の総数に対する、SVMにより得られた出力が「正」であり、かつ最も高いスコアを持っている地点の数の比率。

#### 4.4 位置推定の結果

図5に示す約350[m]のルートで実験を行なう。自律走行車はスタートからゴール地点まで約0.9[m/s]のスピードで移動し、30[frames/s]というレートで画像を取得していく。1回の走行では、約12,000枚の画像が取得される。学習用とテスト用の2組の画像を取得する。学習画像列は、2004年の11月13日の4時頃(晴れ)にロボットを手押しでルートに沿って移動させながら取得した。テスト画像列は、2004年12月20日の2時頃(曇り)に車両制御用のジョイスティックでロボットを操縦しながら取得した。この2組の画像列の相違点は以下のようなものであった。

1. テスト画像列では、多くの木が葉を落としており、学習画像列ではそのような木は存在しなかった。
2. 天候が違いため、対象物の見え方が異なった。
3. 学習画像列においては、自律走行車を手押しで移動させたが、テスト画像列では、ジョイスティックで移動させたため、画像のブレが大きくなっていった。

ルート上に50ヶ所の地点を選び、上記で説明したように200個の「ポジティブ・サンプル」と200個の「ネガティブ・サンプル」を用い、SVMを学習させた。50組のSVMに対して約12,000のテスト画像を処理させた。各フレームに対してのスコアは、与えられたフレームを中心とした21枚の連続した画像に対するSVMの出力値の平均値とする。図5に位置推定の結果を示す。地点Aで学習させたSVMに対して、すべてのフレームを処理し、スコアを計算した。自律走行車は地点Aを二度通過したため、グラフでは2つの顕著な「正」のスコアの領域が現れている。この結果から自律走行車は地点Aを正しく認識したことがわかる。

各SVMに対して学習データと同様のデータを処理させ、100%の認識率を得たことでSVMの性能を立証した。テストデータに関しては、認識率は88%で最高得点率は78%であった。同じルートにおいて、以前発表した位置推定の手法[Katsura 03]により得られた評価と、今回の結果を比較した。以前の手法では、各対象物の領域が学習時と入力画像との間でどれほどマッチしているかを計算し、閾値を用いることでマッチングの成功率を表現していた。認識率は95%であったが、最高得点率は57%しかなかった。これは、以前の手法では最高得点率を犠牲にすることで高い認識率を得るため、比較的低い閾値を用いていたことを意味している。しかし、今回提案しているSVMによる手法ではパラメータや閾値を一切用いることなく、以前に匹敵する性能を示している。これが新しい手法の最も重要な点である。

別の約1[km]の長さのルートにおいても実験を行った。この実験では、学習用、テスト用の両方のデータを2005年1月13日(晴れ)の3時頃に取得した。このルートはほとんど景色

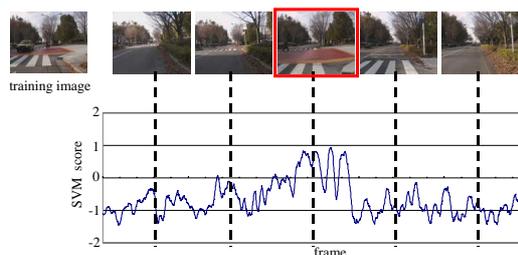


図6: Localization result when views are similar.

の変化しない、いくつかの長い一本道を含んでいる。50ヶ所の選ばれた地点に対して、認識率は100%であり、最高スコア率は92%であった。このような高い性能が得られたのは、学習データとテストデータを同じ日に取得したためであると考えられる。

## 5. まとめ

本論文では、サポートベクター学習に基づく屋外環境での新しい位置推定の手法について述べた。ここでは、物体認識のためのSVMとその結果を用いた位置推定のためのSVMによる、2段階のプロセスをとっている。この手法を用い、実際の屋外環境で実験を行ない、我々の以前の手法と比較した。以前の手法では、手動で設定した多くのパラメータや閾値が存在したが、本手法においては、一切用いていない。

## 参考文献

- [Ishiguro 92] H. Ishiguro, K. Nishikawa, and H. Mori: Mobile robot navigation by visual sign patterns existing in outdoor environments. *In Proceedings of IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*, pp. 636-641, 1992.
- [Katsura 03] H. Katsura, J. Miura, M. Hild, and Y. Shirai: A view-based outdoor navigation using object recognition robust to changes of weather and seasons. *In Proceedings of 2003 IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*, pp. 636-641, 1992.
- [Kidono 02] K. Kidono, J. Miura, Y. Shirai: Autonomous visual navigation of a mobile robot using a human-guided experience. *Robotics and Autonomous Systems*, Vol. 40, No. 2-3, pp. 121-130, 2002.
- [Takeuchi 98] Y. Takeuchi and M. Herbert: Evaluation of image-based landmark recognition technique. Technical Report CMU-CS-95-116, The Robotics Institute, Carnegie Mellon University, July 1998.
- [Thrapp 01] R. Thrapp, C. Westbrook, and D. Subramanian: Robust localization algorithms for an autonomous campus tour guide. *In Proceedings of 2001 IEEE Int. Conf. on Robotics and Automation*, pp. 2065-2071, 2001.
- [Vapnik 98] V.N Vapnik: *Statistical Learning Theory*. John Wiley and Sons, New York, 1998.