

## 画像間のマッチングに基づく屋外環境下における大域的な位置推定

桂 浩章 † Michael Hild ‡ 三浦 純 † 白井良明 †  
† 大阪大学大学院工学研究科電子制御機械工学専攻  
〒 565-0871 大阪府吹田市山田丘 2-1  
‡ 大阪電気通信大学総合情報学部情報工学科  
〒 572-8530 大阪府寝屋川市初町 18-8  
E-mail:katsura@cv.mech.eng.osaka-u.ac.jp

自己位置の推定は自律移動ロボットやナビゲーションシステムなどに必要な機能の一つである。本論文では、予め屋外環境の多数の地点で撮影した画像データベースと現在位置で撮影した画像をマッチングすることによって現在のおよその位置を推定する手法について述べる。対象とした環境が屋外であるため、天候や日照条件により、近い位置で撮影した画像であっても複数の画像間で物体の見え方が変化するので、単純な画像同士の比較ではうまくいかない。しかし、ほぼ同じ視点で撮影した画像では、建物や木といった物体の存在の有無や、それらの位置関係の変化が少ない。そこで画像からそれぞれの物体に関する知識を用いて建物と木を認識し、認識結果を用いてマッチングを行う手法を提案する。実際に、屋外で撮影した画像を使って位置の推定実験を行い、本手法の有効性を示す。

## Rough Localization Based on Image Matching in Outdoor Scene

Hiroaki Katsura †, Michael Hild ‡, Jun Miura † and Yoshiaki Shirai †

† Dept. of Computer-Controlled Mechanical Systems, Osaka University

2-1 Yamadaoka, Suita, Osaka 565-0871, Japan

‡ Dept. of Engineering Informatics, Osaka Electro-Communication University

18-8 Hatsu-Cho, Neyagawa, Osaka 572-8530, Japan

E-mail:katsura@cv.mech.eng.osaka-u.ac.jp

Localization is one of the necessary functions for autonomous mobile robots, navigation systems, and so on. This paper describes a localization method for outdoor scene, which is based on the comparison between the current image and a set of images taken in various places, times of the day, and seasons. Since appearance of objects may change according to the lighting condition, a simple image comparison does not work. However, the placement of objects in the image is almost the same if images are taken from similar viewpoints. Therefore we propose a strategy of first recognizing objects such as buildings and trees, and then comparing the results of recognition. Experimental results using actual outdoor images show the usefulness of the proposed method.

### 1 はじめに

自己位置の推定は自律移動ロボットやナビゲーションシステムなどに必要な機能の一つである。そこで、本論文では、あらかじめ屋外環境内の多数の地点で撮影した画像データベースと現在位置で撮影した画像を比較、対応づけ(マッチング)することによって、日照条件等、環境の見え方の変化にも関わらずその位置を推定する手法を提案する。

松本ら [1] は、経路に沿って撮影した環境の

白黒画像を平滑化によって圧縮したものをデータベースとして記憶し、現在の画像と記憶した画像の画素値を直接比較するブロックマッチングを行うことで位置を推定している。この研究では、屋内環境を想定しており見え方の変化が少ないため画像の直接比較を用いることができるが、屋外環境に適用することは難しい。Takeuchiら [5] は、画像を色の分布を表す行列で表し、2枚の画像から得られる行列の距離をSSDによって計算し、画像間の距離を定義して

いる．また Belongie ら [3] は，画像検索において画像を HSV 表色系で表した色とテクスチャの分布を用いて表している．しかし，これらの研究では，日照条件，季節の変化による物体の色と見かけの色の変化を考慮していない．

そこで本研究では，位置が変化しない対象物をできるだけ簡単な方法で認識し，この認識結果を用いてマッチングを行う手法を提案する．

屋外シーンの認識においてシーン中の物体に関する知識を用いる手法が提案されている [2][4]．本研究でも同様に知識に基づく認識を行う．認識の際に，ある物体である可能性はあるが，そうであると決定できない領域もある．このような領域に，“可能性のある領域”というラベルを付けて柔軟なマッチングを行うことにより認識率を高める．

マッチングは，画像内に存在するものがどのように写っているのかを比較することによって行う．具体的には比較する 2 枚の画像のある物体であると認識された領域をそれぞれ重ね合わせた時に，重なり具合を評価することで行う．

本論文では第 2 節で，物体の認識手法について述べる．第 3 節では，2 枚の画像がどれだけ似ているかを評価する手法について述べる．また第 4 節で位置推定実験を行い，本手法の有効性を示す．

## 2 建物と木の認識

本研究では，大学構内の屋外環境を対象とし，認識する物体との距離が少なくとも 10m 以上となるように撮影した 256×240 のカラー画像を用いる．現画像の例を図 3 に示す．

対象としている大学構内の屋外環境中には，建物，木，自動車，自転車等が存在する．このうち自動車，自転車等は日時によって変化するために認識対象として適さない．そこで画像の上部 (画像上部 135 画素) に存在する建物，木のみを認識の対象とする．本研究では以下の 3 種の領域を認識する．

1. 側壁のように色が一樣で大きな領域 (一樣領域と呼ぶ)

2. 窓枠や建物の端等ある程度の長さのある直線が存在する部分 (人工物領域と呼ぶ)

3. 木領域

以下各々の領域の認識手法について述べる．

### 2.1 一樣領域の認識

原画像からエッジを抽出しエッジによって囲まれた領域を初期領域とし，この初期領域を Ohlander の方法 [6] を用いて，9 種 (R,G,B), (T, $\theta$ ,S), (Y,I,Q) の色情報から似た色の領域に分割する．分割結果から大きな領域を抜き出した後，空領域を認識し除くことによって一樣領域の認識を行う．空領域は，一般に画像上端に接している明るくて面積が大きい領域である．ただし，建物の上部が画像上端に接している場合に空領域と判断される可能性があるため，領域の境界線の上部に建物の境界に相当する明確な直線があるかどうかを調べ直線がなければ空領域とする．また，面積があまり大きくない領域の場合，この領域を空かどうか判断するのは困難である．そこでこのような領域は，一樣領域の可能性のある領域とする．

他に，例えば青空の中に雲が浮かんでいて，青空と雲がそれぞれ別領域に分割され，さらに雲の領域が画像上端に接していない場合にも雲の領域を空と認識しない．そのために，空領域と隣あった画素数を調べ，この画素数が多ければこの領域を空領域とする．

図 4 に図 3 から一樣領域を検出した結果を示す．

### 2.2 人工物領域の認識

人工物領域には，直線部分に多くのエッジが検出される．そこで画像を 16×15 のウィンドウに分割し，エッジの割合が大きいウィンドウを対象に認識を行う．

対象となるウィンドウから明確な境界を直線として抽出することによって確認し，直線が抽出できるウィンドウを人工物領域として認識する．直線の抽出には，ハフ変換を用いている．

図 5 に，図 3 から人工物領域を認識した結果を示す．

### 2.3 木領域の認識

木領域には、葉や枝によって多くのエッジが検出される．そこで画像を  $16 \times 15$  のウインドウに分割し、エッジの割合が大きく、あらかじめ作成した葉の色のモデルに当てはまるウインドウを木領域と認識する．

エッジを検出する際、閾値を一意に決めて検出すると、日照条件により影に入っている木領域内にエッジを多く検出できず、認識の対象とならない．そこで影の部分の木の領域を認識の対象とするために、ウインドウの平均明度を計算し、平均明度が小さいウインドウに関して平均明度に応じた閾値でエッジを検出して、エッジの割合を計算する．

葉の色は、季節や日照条件の変化によって変わってくるので、本研究では緑の木の領域(ひなた、影)、色づいた木の領域(ひなた、影)の4個のモデルを作成する．そこでこれらの状況下にある木にウインドウを設定して、明度  $T$  と色相  $\theta$  の関係を調べる．図1に、夏の晴れた日と秋の晴れた日に撮影した画像内の木領域について調べた例を示す．図1を見ると、ひなたの木の領域では、明度  $T$  が変化しても色相  $\theta$  は、あまり変化しない．また影になっている木の領域では、明度  $T$  が小さい部分では、空の青色成分の影響を受けるために色相  $\theta$  が緑から青の方に広く分布していることがわかる．秋の色づいたひなたの木の領域では、色相  $\theta$  が少し赤の方に移

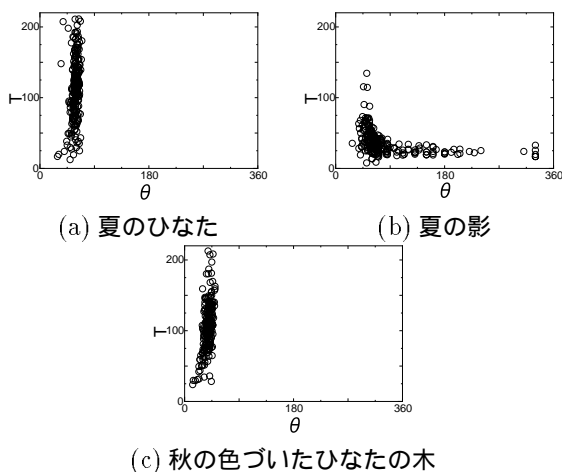


図1: 木の領域の  $T$ - $\theta$  の関係

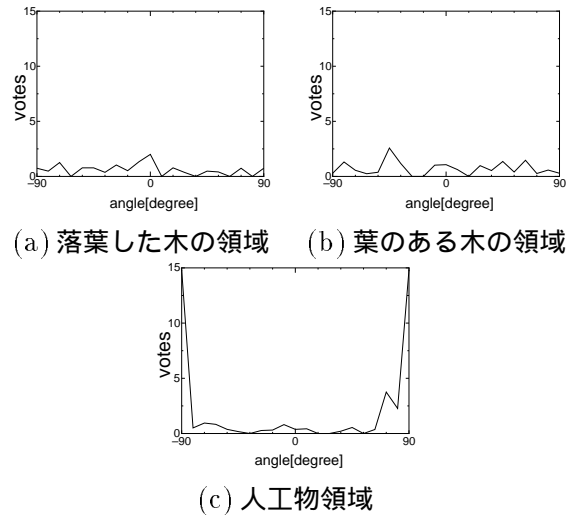


図2: エッジの勾配方向のヒストグラム

動している．これより葉の色のモデルを作成し認識を行う．本研究ではひなたの木の領域で作成したモデルにあてはまるウインドウを木領域と、影の木の領域で作成したモデルにあてはまるウインドウを木の可能性のある領域とする．

落葉した木の領域の場合には、枝の間に背後のものが写ってしまい葉の色のモデルとはあわないことがある．一般に葉や枝によって得られるエッジの勾配方向はばらついていて、強度が小さい．例として図2に、落葉した木の領域、葉のある木の領域、他の領域(例として人工物領域)について、エッジの勾配方向のヒストグラムを示す．図2は、エッジの勾配方向を  $20^\circ$  ごとに分け、エッジの強度を重みとしてかけて頻度を計算し、エッジの数で正規化をしている．図2より木の領域では他の領域と比べヒストグラムの最大値が小さいことがわかる．葉のモデルに合わないウインドウを対象に図2のヒストグラムを作成し、その最大値が小さいウインドウを木の可能性のある領域とする．

図6に、図3から木領域を検出した結果を示す．

### 3 マッチング

複数の地点で撮影した画像から前節の処理を行い、領域の認識結果を記憶しデータベースを作成する．ある地点で撮影した画像を入力と

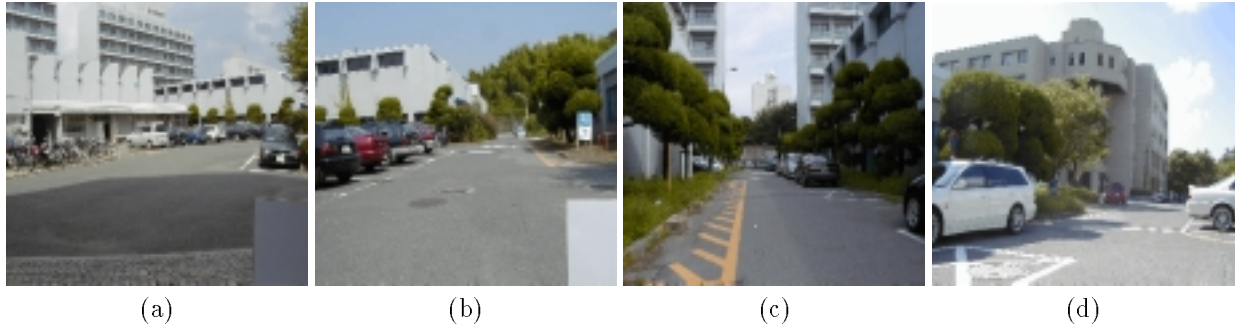


図 3: 原画像の例



図 4: 一様領域の認識結果



図 5: 人工物領域の認識結果



図 6: 木領域の認識結果

■ 一様領域    ▨ 一様領域の可能性のある領域    ▤ 空領域    ▧ 人工物領域  
 ▩ 人工物の可能性のある領域    ▦ 木領域    ▪ 木の可能性のある領域

して同様の認識を行い、この画像の撮影位置をデータベース内の学習画像の認識結果とマッチングすることによって位置を推定する。

### 3.1 評価手法

入力画像とデータベース内に記憶したある画像(学習画像)がどれだけ似ているかを評価する。撮影位置、方向が大体同じである入力画像と学習画像には、同じ物体が写っていて、それらの相対的な位置関係がほぼ等しい。そこでマッチングの評価を以下のように行う。

1. 撮影位置が多少変化しても、物体との距離が大きいため画像内の物体の位置はあまり変化しないが、撮影方向の微小な変化によって、物体の画像中の位置が水平方向に

ほぼ平行移動する。また、撮影位置の地面の起伏により光軸が上下に変化し、垂直方向にほぼ平行移動する。そこで、この平行移動をキャンセルするために、入力画像に対して学習画像をある程度の平行移動を許し、ある移動量  $(dx, dy)$  で重ねた時、画像が重なりあった部分で両方の認識結果の重なり程度  $(s(dx, dy))$  と表す) で評価する。そして、移動量を許容される範囲内で変化させながら評価を行い、一番良い評価  $(s_{max})$  と表す) をマッチングの度合(照合度)とする。

2. 入力画像と学習画像から得られる3種の領域それぞれについて重なり程度(重複度)

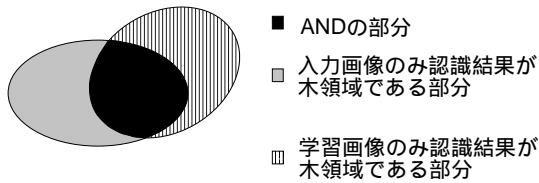


図 7: 木領域の認識結果を重ね合わせた場合

を計算し、すべての重複度が一定値以上であれば  $s(dx, dy)$  を計算する。ただし、画像内で 3 種のうち、ある領域の面積が入力画像、学習画像ともに小さいときには信頼性が低いので、その領域での重複度の計算を行わない。

それぞれの領域に対する重複度を、例として木領域を比較する場合について述べる。基本的には、2 枚の画像が重なっている部分で入力画像と学習画像が共に木領域である部分 (AND) の面積を入力画像または学習画像のどちらかが木領域である部分 (OR) の面積で割ることによって、重複度を計算する (図 7 参照)。

ただし、入力画像と学習画像の木領域の認識結果のうち一方が木の可能性のある領域で他方が木領域である部分については、木の可能性のある領域を木領域と仮定して計算した重複度が一定値以上であれば木領域と考えて AND の部分に加える。したがって木領域の重複度 ( $s_T$  と表す) を

$$O_T = (\text{入力画像または学習画像が木領域である部分の面積})$$

$$P_T = (\text{入力画像, 学習画像とも木領域である部分の面積})$$

$$Q_T = (\text{入力画像, 学習画像の一方が木領域で他方が木の可能性のある領域である部分の面積})$$

として、次式で計算する。

$$s_T(dx, dy) = \frac{P_T + Q_T}{O_T}$$

同様に一樣領域 ( $s_U(dx, dy)$ ) や人工物領域の場合 ( $s_A(dx, dy)$ ) も重複度を計算する。

これらの重複度がすべてある一定値以上であれば、これらの重複度の平均を、 $s(dx, dy)$  とする。そうでなければ、 $s(dx, dy)=0$  とする。

このようにして、移動量 ( $dx, dy$ ) を変えながら  $s(dx, dy)$  を計算し、最大の  $s(dx, dy)(=s_{max})$  を 2 枚の画像の照合度とする。

### 3.2 評価例

ここで、ある入力画像 (図 3(a)) と学習画像 (図 8) をマッチングした例を示す。ここでは  $(-45 \leq dx \leq 45)$ ,  $(-35 \leq dy \leq 35)$  の範囲で平行移動を許容し、5 画素ずつ移動させてマッチングを行っている。本来、移動量の範囲は学習画像をどれくらい位置、方向を変化させるごとに撮影したかによるが、ここでは簡単のため経験的に決定した。また、認識結果の割合が入力画像、学習画像とも  $3000/(256 \times 135) (\approx 0.09)$  以上であるものだけでマッチングを行い、それぞれの評価値が 0.3 以上でなければならぬとしている。以下に学習画像を認識した結果 (図 9) を示す (入力画像の認識結果については図 4(a), 図 5(a), 図 6(a) をそれぞれ参照のこと)。

これらの認識結果を用いてマッチングを行った時、入力画像と学習画像の照合度 ( $=s_{max}$ ) とその時の移動量 ( $dx, dy$ ),  $s_U, s_A, s_T$  を表 1 に示す。またそれぞれの領域の認識結果をこの移動量で重ね合わせたものを図 10 に示す。図 10 では、AND の部分を黒で、OR の部分を灰色で表している。図 10 を見ると入力画像と学習画像の物体の見え方が良く一致していることがわかる。

この入力画像と学習画像のマッチングでは、入力画像、学習画像ともに木領域と判断された部分の面積が小さいため、木領域の認識結果でマッチングを行っていない。

表 1: 入力画像と学習画像のマッチングの度合

$s_{max}$	$dx$	$dy$	$s_U$	$s_A$	$s_T$
0.673	-20	-30	0.614	0.731	—



図 8: 学習画像



(a) 一様領域 (b) 人工物領域 (c) 木領域

図 9: 学習画像に対する認識結果



(a) 一様領域 (b) 人工物領域

図 10: 重ね合わせた図

#### 4 実験

学習画像を冬の曇の日に 8ヶ所の地点 (図 11参照) からそれぞれ 1 枚ずつを撮影し, 領域の認識を行って, データベースを作成した. 次にそれぞれの地点付近から, 夏の晴れ, 曇, 秋の晴れ, 曇, 冬の晴れの日に撮影した画像 (図 12参照) をデータベースとマッチングを行ったところ, 入力画像 24 枚すべて望ましい位置の推定結果を得ることができた.

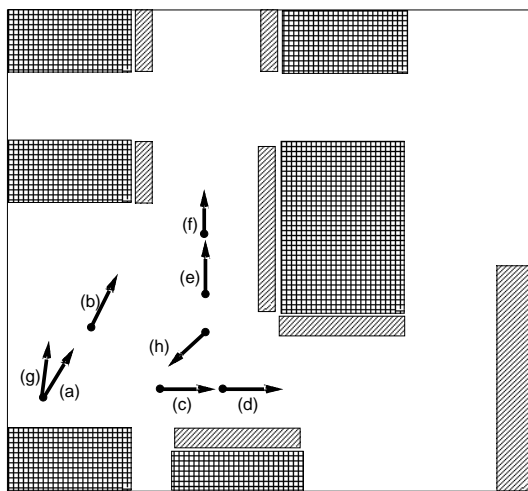


(a) (b) (c) (d)



(e) (f) (g) (h)

図 12: 学習画像



■ 建物 ■ 木

図 11: 撮影位置

ここで秋の曇の日に撮影した図 13を入力画像として図 12の各学習画像とマッチングして位置を推定した例を示す. 表 2に入力画像と各学習画像との照合の結果を表す. 一 の項はその領域で重複度を計算していないことを示す. また入力画像と図 12(a)-(g) のマッチングにおい

て  $s_{max}$  は 0 であるが, どの領域が合っていないかを示すために, 許容される移動量内で各領域の重複度の平均が最大のときの各領域の重複度 ( $s_U, s_A, s_T$ ) を示している. そのうち重複度が一定値 (0.3) 以下の場合にはイタリックで示している.



図 13: 入力画像

表 2より, 図 13の撮影位置は図 12-(h) の撮影位置であると推定できる. 学習画像の撮影した季節が冬であるため, 図 12-(h) の木の一部は落葉していて, その部分の色は大きく変化してい

表 2: 入力画像と各学習画像のマッチングの割合

	$s_{max}$	$dx$	$dy$	$s_U$	$s_A$	$s_T$
(a)	0	35	30	0.107	0.355	0.000
(b)	0	0	30	0.150	0.350	0.000
(c)	0	-45	-25	—	0.180	0.367
(d)	0	-45	-20	0.031	—	0.481
(e)	0	-45	-35	—	0.174	0.641
(f)	0	-35	-30	—	0.120	0.603
(g)	0	-10	35	0.170	0.276	0.018
(h)	0.545	15	0	0.421	0.491	0.724

るが、高い照合度を示している。

## 5 おわりに

本研究では、あらかじめ屋外の多数の地点で撮影した画像データベースと現在位置で撮影した画像をマッチングして位置を推定する手法について述べた。屋外では物体の色や見かけの色が天候、日照条件、季節等により変化するので単純な画像同士の比較ではうまくいかない。そこで画像中の建物と木を日照条件等の変化にも関わらずできるだけ簡単な方法で認識し、この認識結果を用いてマッチングを行う手法について提案した。また実験により本手法の有効性を確かめた。

今回一様領域の認識において、色の違いをまったく考慮に入れていないので色の違う建物をマッチングすると良い評価が得られる可能性がある。そこで色の違いを考慮した認識とマッチングを行いたい。これには日照条件等の変化で一様領域の色がどれだけ変化するかをモデル化することを考えている。

また本研究をもとに、屋外シーンを一度誘導して学習させ、その学習経験に基づいて動く自律移動ロボットに関する研究を行う予定である。

## 参考文献

[1] 松本, 稲葉, 井上: “視野画像列を利用した経

路表現に基づくナビゲーション,” 日本ロボット学会誌 Vol. 15, No. 2, pp. 236-242, 1997.

[2] 平田, 白井, 浅田: “単一カラー画像から得られる 3 次元情報を利用したシーンの解釈,” 電子情報通信学会論文誌, Vol. J75-D-2, No. 11 pp. 1839-1847, 1992.

[3] S. Belongie, C. Carson, H. Greenspan and M. Hebert: “Color- and Texture-Based Image Segmentation Using EM and Its Application to Content-Based Image Retrieval,” Proc. CVPR, 2000.

[4] T.M. Strat and M.A. Fischler: “Context-Based Vision: Recognizing Objects Using Information from Both 2-D and 3-D Imagery,” IEEE Trans, PAMI, Vol. 13, No. 10, pp. 1050-1065, 1991.

[5] Y. Takeuchi and M. Hebert: “Finding Images of Landmarks in Video Sequences,” Proc. CVPR, 1998.

[6] Y. Shirai: “Three-Dimensional Computer Vision,” Springer-Verlag, 1987.