

知的画像切り出し

平田真一 白井良明 浅田稔

大阪大学工学部電子制御機械工学科

あらまし 道路シーンの単一カラー画像から、色や形状に関する知識、ならびに3次元的な拘束を用いて領域を解釈する方法について述べる。本手法の特徴は、画像から得られるエッジと領域の情報を効率的に利用することと、部分的に得られた3次元情報を拘束条件として利用し、より正確な解釈を行うことである。解釈手順は以下の通りである。まず、色による領域分割を用いて画像をよく似た色を持つ隣接領域からなる粗領域に分割、統合する。次に、対象に関する知識を用いて領域の解釈を行っていく。ここで使う知識は、対象領域抽出のための知識と対象領域検証のための知識である。最終的な解釈結果として、画像内の主要な領域に対し、正確な輪郭ならびに各領域の3次元的な相対位置関係を求める。

Intelligent Image Segmentation

Shinichi HIRATA Yoshiaki SHIRAI Minoru ASADA

Dept. of Mechanical Eng. for Computer-Controlled Machinery, Osaka University

Abstract We present a method for interpreting outdoor scenes which segments monocular color images, and identifies the segmented regions as objects in the scene such as roads, cars, trees, etc. The features of our method are efficient use of both edge and region information, and use of the 3-D geometric cues to recover the spatial relationships between objects in the scene. First, the image is segmented into connected regions of similar colors. Then, each region is interpreted by knowledge about objects. The knowledge is used in two ways: one is used to extract candidate regions, and the other is to verify and localize them in the 3-D space. The final goal is to derive the correct contours of segmented regions and spatial relationships between objects in the scene.

1 はじめに

一般の屋外シーンから所望の領域を自動的に切り出すことは、コンピュータビジョンの重要な研究対象の1つである。対象となる領域が明度やカラー情報処理によって他の領域と区別可能である場合には、シーンの解釈は必ずしも必要としないが、例えば対象が自動車や歩行者の場合、それらの領域を自動的に切り出すためには、画像内の領域を解釈することが必要である。通常、屋外シーンには人工物や自然物が含まれており、視点の変化や撮影時の天候や時刻、季節などの要因によってさまざまに変化する。このため、問題は非常に困難となるが、人間は実際に対象に関する知識を用いて画像内に含まれる領域の解釈を行い、目的の領域を切り出すことができる。

これまでに行われてきた画像理解システムの代表的なものとして、以下のものが挙げられる。

- ・ 大田の、プロダクション・システムを利用した屋外シーンの解釈 [1]
- ・ 松山、長尾の、ブラックボード・モデルを利用した航空写真の理解 [2]
- ・ Edward M. Riseman らによる The Schema System [3]

我々は最終的には対象に関する知識を用いた汎用的な画像切り出しシステムの実現を目指しているが、そのようなシステムを構築する上での問題点を明らかにするため、これまで、対象を道路シーンの単一カラー画像に絞り、実際に解釈を行うシステムを作成してきた。[4]

道路シーンの解釈は、自動走行車 (ALV) の制御や、自動車の自動操縦等の分野で数多くの研究がされている。

[5][6][7]

本システムの特徴としては、

- ・ 画像から得られるエッジと領域の情報を効率的に利用する。
- ・ 部分的に得られた3次元的情報を拘束条件として用い、さらに対象固有の知識を適用することにより正確な解釈を行う。

の2点である。

2 対象画像

対象画像としては、以下の4つの仮定をおく。

- (1) 一般道路シーンの単一カラー画像 (図1)
(道路シーン: 人工物と自然物双方を含み、一般性があると考えられる。)
- (2) 地面 (道路面) は平坦であるとする。
- (3) カメラパラメータ: 既知
(光軸まわりの回転 (roll)=0, tilt angle= 既知、焦点距離 f = 既知)



(a)



(b)

図1 原画像

3 道路シーンの解釈

解釈手順を図2に示す。

まず、画像を良く似た色を持つ隣接領域からなる粗領域に分割、統合する。

次に、対象に関する知識を用いて画像内の領域の解釈を行っていくが、この知識は、1つは候補領域の抽出のための知識であり、もう1つは候補領域の検証のための知識である。

粗領域分割の後、候補領域抽出のための知識を用いて解釈してゆく。この結果、候補領域の存在 / 非存在や画像上での位置などがわかる。

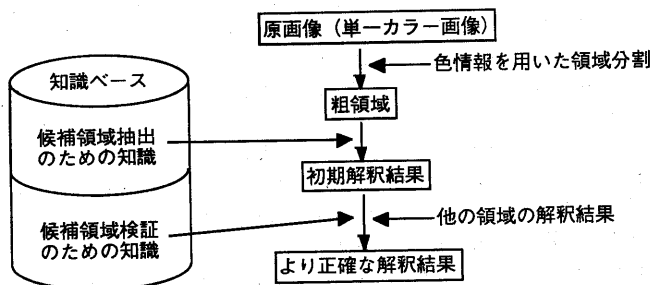


図2 領域の解釈手順

候補領域の大体の存在場所が分かると、次に個々の領域を検証のための知識を用いて詳しく調べてゆく。これらの結果は、それぞれの領域の解釈(切り出し)精度を増すばかりでなく、周囲に存在する領域の解釈を行うために利用することもできる。

最終的な解釈結果としては、画像内の主要な領域に対し、正確な輪郭ならびに各領域の3次元的な相対位置関係を求めることとする。以下、これらの処理について順に詳しく説明する。

4 色情報を用いた粗領域への分割、統合

まず式(1)-(4)により、(R,G,B)から(T,θ,S)(ここで、T: 明度、θ: 色相、S: 彩度)へ色座標変換する。

$$T = R + G + B \quad (1)$$

$$r = \frac{R}{T}, \quad g = \frac{G}{T}, \quad b = \frac{B}{T} \quad (2)$$

$$\theta_1 = \cos^{-1} \frac{2r - g - b}{\sqrt{6[(r - \frac{1}{3})^2 + (g + \frac{1}{3})^2 + (b - \frac{1}{3})^2]}} \quad (3)$$

$$\theta = \begin{cases} \theta_1 & (g \geq b) \\ 2\pi - \theta_1 & (g < b) \end{cases} \quad (4)$$

$$S = 1 - \min(r, g, b) \quad (4)$$

次に、彩度Sの1次元ヒストグラムから彩度の低い部分(無彩色部分)と彩度の高い部分(有彩色部分)に分ける。実際には、以下のように行った。

さまざまな道路シーンについて、彩度のヒストグラムをとると、普通、図3のような形状を示す。そこで、Sのヒストグラムを山ごとに分け、図3に示すような重み関数を考え、重み係数と山に含まれる頻度の積の総和がある一定値になるまで彩度の低い山から順に無彩色部分として採用する。

無彩色部分に対しては、明るさTの1次元ヒストグラムの山ごとに分割する。

有彩色部分に対しては、T-θからなる2次元ヒストグラムの山ごとに分割する。

2次元ヒストグラムの山で分割するのは、1次元ヒストグラムで分割する場合に比べて難しい。そこで、以下の方法により分割を行った。(図4)

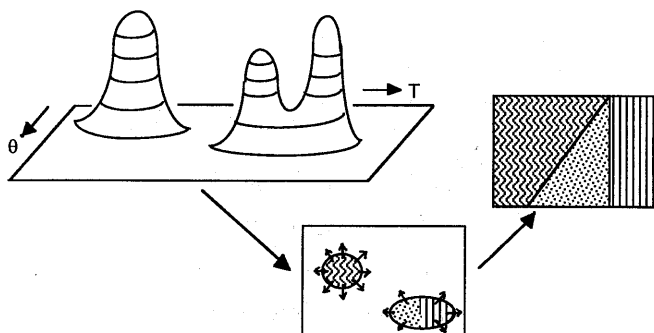


図4 2次元ヒストグラムの分割

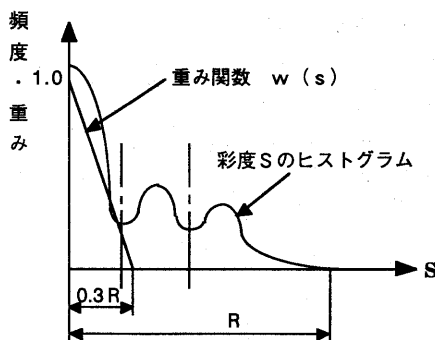
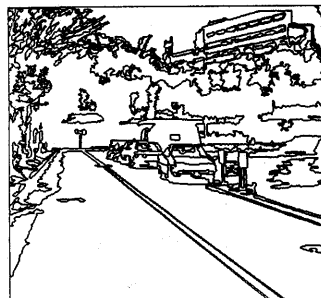


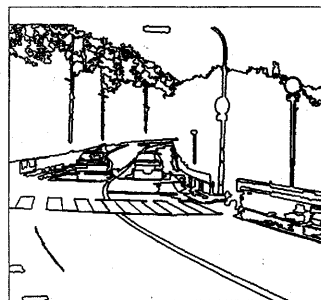
図3 彩度による分割

- (1) 2次元ヒストグラムに対し、Gaussianのスムージング($\sigma=1.5$)をかける。この時、ヒストグラムは(高さ)=255になるように正規化しておく。
- (2) 大まかなピークを求めるため、ある高さ($H=128, 64, 32, 16, 8, 4, 0$)で切断し、それぞれの切断面に存在するピークごとにヒストグラムを分割していく。
- (3) 最後に、ピーク部分の周囲だけでなく2次元ヒストグラム全体にわたって分割できるように、各ピーク部分を膨張させる。

そして、無彩色領域と有彩色領域の分割結果を合わせ、面積の小さな小領域を大領域に統合して、後の処理を行いやすいようにある程度まとまった領域(図5)にする。ここで、分割、統合の際のしきい値は、後の処理に合った最適な値を選択する必要がある。ここでは、まず、おおまかな領域分けをするように設定している。



(a)



(b)

図5 粗領域分割結果

5 道路領域の解釈

画像が粗領域に分割されると、次に道路候補領域を抽出する。一般に道路領域は画像内にある領域を抽出するための最大の手がかりとなると考えられるためである。

5.1 道路候補領域（初期候補）の抽出

道路候補領域（初期候補）を抽出するために用いた知識は以下の2つである。

- 道路領域（アスファルト）は、やや青っぽい灰色（ほぼ無彩色）。また、極端に明るすぎたり暗すぎたりする可能性は少ない。
- 道路領域は画像内で大きな面積を占めているので、ヒストグラムを用いることにより抽出できると考えられる。

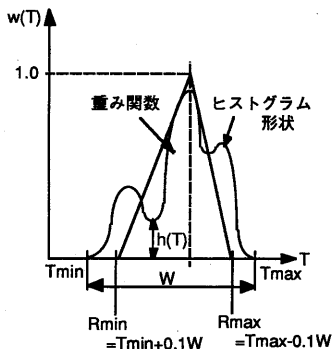
これらの知識を用いて、以下のようにして初期候補領域を抽出する。

- 図6に示すように、重み関数を考える。無彩色部分の明るさのヒストグラムの各山に対し、ヒストグラムの頻度と重み係数との積の総和、すなわち道路領域である可能性を求める。可能性が最も高い山を中心として、左右で可能性の高い山から順に可能性を足してゆき、あるしきい値になるまでの山を、道路候補領域を含む山とする。
- (i)で得られた山に含まれる点が一定割合（95%）以上含まれる粗領域を、道路候補領域（初期領域）とする。（図7）

5.2 消失線情報を用いた道路候補領域拡張

道路領域は、周囲の領域の影になっていたり、暗い領域が存在したりして先に述べた手法だけでは道路領域は完全には得られない。そこで、以下のように道路候補領域を拡張する。

画面内で消失線よりも上にある領域は、地面に対し高さのある領域と考えられるので、消失線の位置の情報を用いて、粗領域を高さのある領域とそうでない領域に分ける。（図8）



(山M(j)の評価関数値)

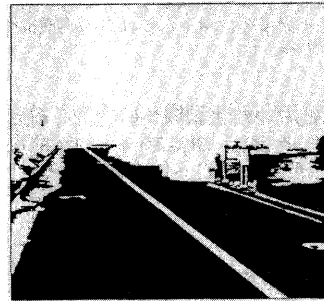
$$= \frac{1}{A_0} \sum_{T=M(j)} \{h(T) \cdot w(T)\}$$

ここで

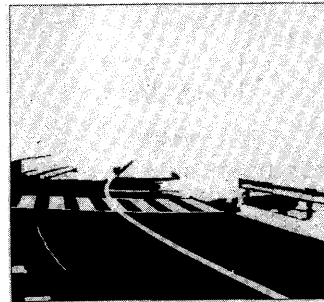
$$A_0 = \sum_{T=R_{min}}^{R_{max}} h(T)$$

図6 道路候補領域抽出アルゴリズム

次に、高さのある領域以外の領域に対して、5.1の結果を用いて、面積があるしきい値よりも大きい道路候補初期領域を核領域として、画面の下側から順に明るさの差の絶対値がしきい値以下なら ($|\Delta R| + |\Delta G| + |\Delta B| \leq \text{thrd}$)、道路領域を拡張してゆく。（図9）

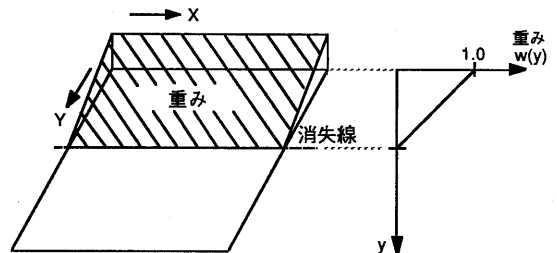


(a)



(b)

図7 道路候補初期領域



各領域に属する画素に対し、重みで加算。あるしきい値以上のものは、高さのある領域とみなす。

(地面より上にあると考えられる、領域R(j)の有効画素数)

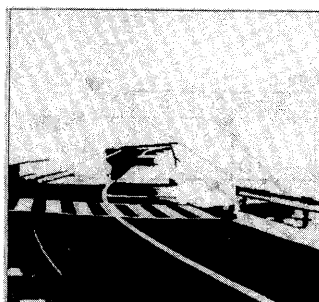
$$A_{R(j)} = \sum_{(x,y) \in R(j)} w(y)$$

$$A_{R(j)} \geq \text{thrd} \rightarrow \text{高さのある領域}$$

図8 高さのある領域の判断

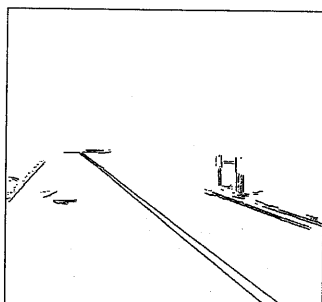


(a)

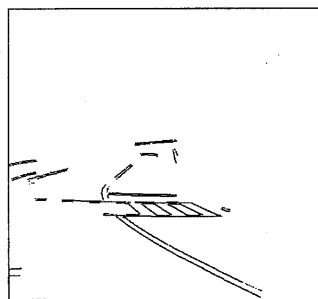


(b)

図9 拡張された道路候補領域



(a)



(b)

図10 通行区分の検出結果

5.3 道路の通行区分の抽出と車線幅の推定

先の3.2, 3.3の処理により、道路候補領域が得られると、その情報を用いることにより通行区分を抽出することができる。

ここでは、通行区分としてはセンターライン及び横断歩道を仮定している。

センターラインは、道路内にある白色あるいは黄色の平行線であるという知識を用いて、以下のように抽出する。

道路領域内または道路の境界に含まれる、非極大値抑制により細線化した強いエッジに対し、ある長さ以上のエッジセグメントに対して、明るい部分を挟み込むような平行なエッジセグメントのペアを求めてそれらを連結する。(図10)

横断歩道は、道路内にあるゼブラ・パターンであるという知識を用いて、以下のように抽出する。

道路内の白色領域を抽出し、収縮、膨張処理により四辺形部分に切り離す。その四辺形部分がゼブラ・パターンの一部であれば、ほぼ同じ形状をしていると考えられる。そこで、領域の境界を抽出し、各領域ごとに平行移動して周囲の領域とマッチングがとれたときの平行移動量(dx,dy)を2次元パラメータ空間へ投票する。そして2次元パラメータ空間のピークを検出してピークに含まれる領域のうち、一定数以上連なる領域群を抽出する。(図11)

通行区分線や横断歩道領域の情報から、消失点を求めることにより3次元の道路の配置や幅を推定することができる。また、それらの情報を用いることによって、自動車候補領域や、ガードレールなどの領域を得る手がかりとすることもできる。

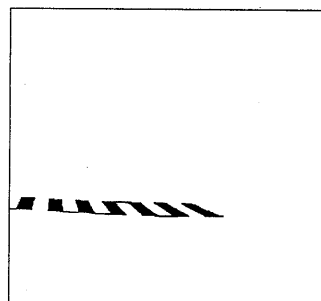


図11 横断歩道の抽出結果

5.4 道路方向の推定

道路候補領域並びに通行区分線の候補が求められたので、これらの情報を用いて道路方向を推定する。ここでは、道路は直線であると仮定する。図12のように、カメラの光軸とZ軸を一致させる。座標原点を画像中心に置いたとき、道路の方向ベクトルは

$$\vec{a} = (x', y', f) \quad (5)$$

ここで、

(x', y') : 道路の消失点の画面上での座標

f : 焦点距離

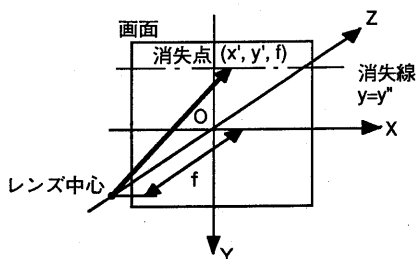


図 12 光学系の幾何学

今、画像に対する仮定より、 y', f は既知なので、通行区分線と、直線 $y = y'$ (y' : 消失線の位置) との交点を求めることにより、 x' は求められる。地面を XZ 平面に一致するように変換すると、道路の方向ベクトルは

$$\vec{a} = (x', 0, \frac{f}{\cos \alpha}) \quad (6)$$

で与えられる。

道路の消失点の x 座標 x' を求めるには、次の方法を用いた。通行区分の各点を、それぞれが持つエッジの方向へ延長し、消失線との交点を求め、それらのうち最大頻度を持つ交点を求める点とした。得られた道路の方向を図 13 に示す。

6 道路の周囲の領域の解釈

道路候補領域が求められると、次に、道路の周囲の領域を解釈する。これは、道路の周囲の領域を先に解釈しておくことにより、後の自動車候補領域の解釈時に解釈範囲を絞れるからである。道路の周囲の領域としては、空、木(林)の領域、建物領域、電柱等のボール領域等が考えられる。

6.1 色がほぼ均一な大領域の抽出

道路領域以外で、色がほぼ均一と考えられる領域は、道路候補領域と同様の方法で求めることが出来る。ここでは、空と木(林)を抽出した。空は、地面に対し高さのある領域のうち、非常に明るい画面内の上方の部分にある大領域を拡張することにより求められる。また、木(林)の部分は、彩度一定値以上で色相が木の葉の色(緑色)の部分を基準色とし、その色を持つ領域のまわりで、よく似た色相値を持った領域を統合することにより求められる。

6.2 建物の抽出と向きの解釈

建物は人工物であり、その輪郭は直線部分が多い。逆に、木などの自然物領域は輪郭が複雑になることが多い。また、建物はある程度高い領域であるので、以下の方法で抽出する。

- (i) 空や木の輪郭線から、一定値以上の長さを持った直線部分を求める。
- (ii) それらの直線部分に隣接する領域で、高さのある未解釈領域を建物領域とする。(図 14)

建物には、窓が規則的に並んでいることが多く、これらは建物の壁の向きを推定するのに役立つ。

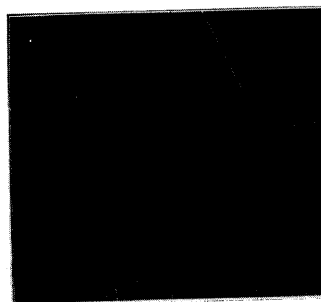


図 13 図 1(a) における道路、建物の方向
(灰色: 道路、白色: 図 1(a) 右上の建物)

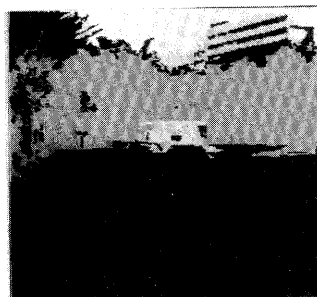


図 14 建物、緑、空領域

今、建物の壁は地面に対し垂直な平面であるとする。このとき、壁の法線ベクトル $\vec{e}$ は、

$$\vec{e} = (-\sin \theta, -\cos \theta \sin \alpha, \cos \theta \cos \alpha) \quad (7)$$

θ : 地面を XZ 平面に置いたときの、X 軸からの角度

α : tilt angle

と表せる。 $\vec{e}$ は、建物内の水平線の消失点を (x'', y') とすると、ベクトル (x'', y', f) に垂直であるから、(ベクトルの内積 = 0) より、次の関係式が成り立つ。

$$\theta = \tan^{-1} \left(\frac{f}{x'' \cos \alpha} \right) \quad (8)$$

つまり、 x'' を求めれば、 θ は定まる。

x'' を求めるには、先の道路方向の推定で消失点を利用したときと同じ方法を用いる。ただし、画面上で垂直なエッジは除外する。以上の方法から得られた壁の向きを図 13 に示す。

6.3 ボールの抽出

電柱、交通標識などのボールは、領域の前後関係を得る手がかりとして重要である。

抽出手順は、粗領域の中で、細長くかつ平行な輪郭のベアを一定割合以上持った未解釈領域を抽出することにより得られる。

ボールが得られると、一部の領域間の定性的な 3 次元的位置関係を推定することができる。すなわち、ボールに隠された背後の領域(ボールの両側の領域の特徴量が似ているような領域。例えば林や自動車など)を抽出することができる。(図 15)



図 15 ボールとその背後の領域

7 自動車の抽出

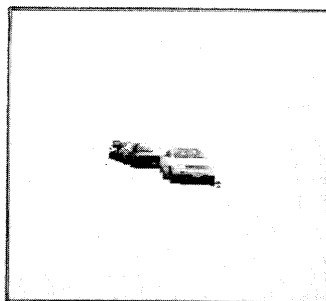
道路領域、道路の周囲の領域の解釈がある程度得られたところで、自動車の抽出を行う。

7.1 自動車候補領域の抽出

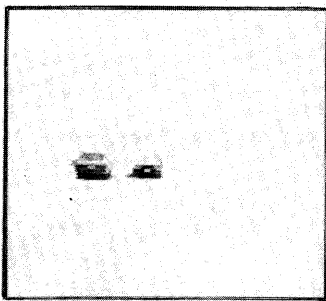
自動車は、車種により色、形は様々で、また、見る方向によっても形は変化する。そこで、初めに自動車候補領域を求める際には、以下のような簡単な知識を用いた。すなわち、

自動車候補領域は道路領域内に影を落とす。その影の上側に隣接したまとまった未解釈の領域が自動車候補領域と考えられる。

ここで、先の結果より推定される道路の幅から、自動車の大体のスケールを制約条件として与える。(図 16)



(a)



(b)

図 16 自動車候補領域

7.2 自動車候補領域の検証

3.8 で求めた道路候補領域は、簡単な知識のみを用いているので、得られた領域が何台の自動車を含んでいるのか分からない。そこで、自動車の影の部分のうち、道路との接点は道路面上にあると考え、自動車の影の領域のうち道路候補領域に接している点を求め、道路候補領域と共に、シーンを真上から見た図にマッピングした。(図 17)

自動車は、最大寸法が決まっているので、(小型乗用車の場合、幅 1700mm、長さ 4500mm) その寸法を持つ長方形が、真上から見た自動車のサイズを表す(図 17 右下)。図を見ると、1 台が前方にあり、約 3 台がその後方に連なっていることが分かる。自動車の影の、道路候補領域との接点と、自動車の長方形モデルとのマッチングが出来れば、個々の自動車の正確な切り出しが出来ると期待される。

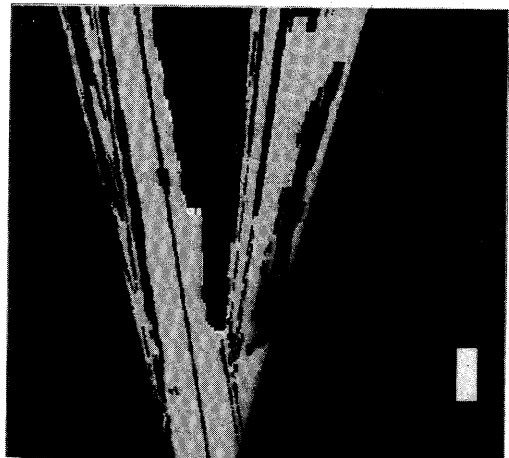


図 17 図 1(a) における道路を真上から見た図
(灰色: 道路候補領域、白色: 自動車の影と道路領域の接点)

8 おわりに

以上道路の単一カラー画像から色、形状に関する知識、ならびに 3 次元的な情報を用いて道路候補領域や建物、自動車候補領域などを解釈する方法について述べた。

また、候補領域からその形状、空間での位置に関する知識を用いて、道路、建物、自動車について 3 次元的な情報の一部を引き出すことができた。

縦列駐車している自動車の場合のように、3 次元的な情報が引き出せれば、画像を単に 2 次元的に解釈する場合に比べて、より多くの情報や正確な解釈の手がかりとなることが分かった。

3 次元的な情報は、単独の領域から、その領域の形状に関する知識を用いて抽出することもできるが、領域間の隠蔽関係をうまく抽出することができれば、画像内の領

域間の相対的な位置関係を抽出することができる。このように、今回述べた手法は、単一カラーの道路シーンに含まれる情報のうちの一部しか抽出できていないと思われるので、現在、より多くの情報を引き出す方法を検討中である。また、様々な道路シーン画像に対する解釈の信頼性の検証を行っている。

参考文献

- [1] Ohta, Y.: "Knowledge-based interpretation of outdoor natural color scenes", Research Notes in Artificial Intelligence 4, Pitman Advanced Publishing Program (1985)
- [2] 松山、長尾: "航空写真の構造解析" 情報処理, 21, 5, pp.468-480 (1980)
- [3] Edward M. Riseman et al.: "The Schema System" International Journal of Computer Vision Vol.2, Number 3, pp.209-250 (1989)
- [4] 平田、川原、白井、浅田: "知的画像切り出し - 道路シーンの解釈 -" 情報処理学会第40回(平成2年前期)全国大会講演論文集(I) pp.397-398
- [5] T. Kanade, C. Thorpe et al.: "Carnegie Mellon Navlab Vision System" International Workshop on Industrial Application of Machine Vision and Machine Intelligence (1987)
- [6] 安居院、中島: "道路情報処理とコンピュータビジョン" コンピュータビジョン 47-9 (1987)
- [7] 中井、森: "自律移動ロボットの沿目標移動と障害物回避行動" 第7回日本ロボット学会予稿集 pp.617-620 (1989)