

領域の隣接関係を用いたラベル画像の生成

3 A B - 4

林 裕樹 工藤峰一 外山 淳 新保 勝

北海道大学大学院 工学研究科 システム情報工学専攻

1. はじめに

画像認識問題は画像上の領域に対して、それぞれの領域が示す対象のラベルを適切に割り当てることである。自然シーンを認識する場合、対象の形状や見え方などをあらかじめ正確にモデル化できる工業製品の認識と比較して、対象のモデル化が難しく、認識率の向上が見込みにくい。

この問題を解決するために、対象間の位置関係などの知識を導入するなど様々な試みがある。中でも、Kimら[1][2]は、節点で領域、辺で隣接関係を表すRAG(Region Adjacency Graph)上にマルコフ確率場モデルを用い、ラベリング問題をエネルギーの最小化問題として定式化した。しかし、この手法で用いる領域の特徴量や知識は比較的単純であることから、類似色対象の識別や複雑な隣接関係の評価が不得意である。また、この手法は初期ラベルの修正にシミュレーテッドアニーリングを用いるため、膨大な計算量を必要とする。従って、これらを改良することで性能の向上が期待できる。

そこで、本研究では自然シーン画像に対して、領域の特徴量を用いて初期ラベルを割り当て、シミュレーテッドアニーリングを用いずに、ラベル間の隣接関係に関する知識を適用することで、初期ラベルを高速に修正する手法を提案し、実験によりその有効性を示す。

2. ラベル画像生成

本手法では、まず、入力画像に対して対象を十分に把握されるよう領域分割し、その各々の領域に初期ラベルを割り当てる。次に、これをより適切なものに修正し、最終結果を得る。

2.1. 初期ラベルの割り当て

各領域への初期ラベルの割り当ては、あらかじめ正しいラベルを割り当てた教師画像からバックプロパゲーション法で学習させた3層のニューラルネットワークを用いる。ネットワークは対象の複雑さを表現できるようにラベル毎に作成し、ネットワークの出力 V_i が最小になるラベルを割り当てる。領域 R_i の特徴量としてKimらの F_1^i から F_9^i [1]までの他、エントロピー F_{10}^i を用いる。

$$F_1^i = R_i / (R_i + G_i + B_i)$$

$$F_2^i = G_i / (R_i + G_i + B_i)$$

$$F_3^i = B_i / (R_i + G_i + B_i)$$

$$F_4^i = (R_i + G_i + B_i) / 3$$

$$F_5^i = F_1^i - F_3^i$$

$$F_6^i = \frac{\max(F_1^i, F_2^i, F_3^i) - \min(F_1^i, F_2^i, F_3^i)}{\max(F_1^i, F_2^i, F_3^i)}$$

$$F_7^i = \frac{1}{S(R_i)} \sum_{j \in R_i} \sum_{k \in N_j} \frac{|f_j - f_k|}{4}$$

$$F_8^i = W(MB_i) / H(MB_i)$$

$$F_9^i = S(R_i) / S(MB_i)$$

$$F_{10}^i = - \sum_{s=0}^{n-1} \sum_{t=0}^{n-1} P_i(s, t) \log(P_i(s, t))$$

ここで、 R_i, G_i, B_i はそれぞれ領域 R_i における赤、緑、青の各成分の平均値、 N_j は画素 j の4近傍の画素集合、 S は領域の面積、 MB_i は領域 R_i を覆う軸平行な最小の矩形領域、 W, H はそれぞれ MB_i の幅と高さ、 f_j は画素 j の明度、 $P_i(s, t)$ は R_i の明度に関する共起行列(階調数 $n=16$)の (s, t) 成分を表す。

2.2. ラベルの修正

ラベルの修正はアニーリングを用いず、図1のアルゴリズムで行う。まず、領域 R_i と同じラベルをもつ互いに隣接している領域に R_i を含めた領域の集合 τ_i を構成する。次に、隣接領域のラベルの本手法で用いる隣接関係(表1)から、ラベルの解釈の整合性に関する評価値 E_i^I と、 τ_i の妥当性に関する評価値 E_i^S を計算して E_i を求め、この値が低いほど良いラベル割当とする[2]。このとき、 E_i が閾値以下で、かつ面積がある程度以上の領域はラベルが正しいとみなしてラベルを確定し、他の領域のラベル修正にこれらの確定領域を使う。また、 E_i は次式のように表される。

$$E_i = E_i^I + E_i^S$$

$$E_i^I = V_{\tau_i} + \sum_{\tau_j \in G_{\tau_i}} V'_{\tau_i, \tau_j}$$

$$E_i^S = \sum_{R_j \in G_i} \left\{ \frac{\max_{1 \leq k \leq 10} |F_k^i - F_k^j|}{B_{ij}^{\delta_{\lambda_i, \lambda_j}}} \right\}$$

ここで、 λ_i は R_i のラベル、 G_i は R_i に隣接する領域の集合、 $V'_{i,j}$ は本手法で新たに導入した R_i と R_j の隣接関係のベナルティー(表1)、 B_{ij} は R_i と R_j の共有境界で隣接する画素の明度差の絶対値平均、 $\delta_{\lambda_i, \lambda_j}$ はクロネッカーのデルタである。

このアルゴリズムを用いて、確定領域から順にラベルを決定することにより処理を行う。

3. 実験と考察

本手法の有効性を確かめるために、実際の屋外シーンの画像に対して実験を行った。原画像(図2)は256×256画素、各色256階調のカラー自然シーン画像を用いた。

この原画像を領域分割し(図3)、初期ラベルを割り当て、これを修正し結果を得た(図4)。図4より領域分割の段階で過分割になっている領域に対し、ほぼ適切なラベルが割り当てられていることがわかる。このときの処理時間はCPUがPentium-90MHzの計算機で70秒前後であり、同条件で数時間以上かかるシミュレーテッドアニメーリングに比べ高速である。また、本手法、Kimらの手法とも画素毎の認識率が80%前後であることから、本手法が高速化の点で有効であるといえる。

しかし、ラベルが実際の対象と一致しない部分のほとんどが影などの暗い部分や道路などの無彩色に近い領域であることから、領域を示す特徴量が色に偏り過ぎている上に、特徴量に加えたエントロピーがあまり有効ではなかったと考えられる。したがって、特徴量に領域形状など色以外の特徴や、対象の性質をより反映した知識に基づいたラベルの修正法が有効であると思われる。

4. おわりに

本研究では、自然シーン画像に対して、シミュレーテッドアニメーリングを用いずに、領域の隣接関係を用いて適切なラベル画像を高速で生成する手法を提案し、実験によりその有効性を確かめた。しかし、領域の特徴量が色に関するものに偏っていることと、用いる知識が隣接関係と簡単な存在位置だけであるため、適切なラベルを割当てられない部分が残った。

今後の課題として、より有効な領域の特徴量と隣接関係の知識を用いたラベルの修正法の構成が挙げられる。

文献

- [1] I. Y. Kim and H. S. Yang, Efficient Image Labeling Based on Markov Random Field and Error Backpropagation Network. *Pattern Recognition*, 26(1993), 1695-1707.
- [2] I. Y. Kim and H. S. Yang, An Integrated Approach for Scene Understanding Based on Markov Random Field Model. *Pattern Recognition*, 28(1995), 1887-1897.

表 1: 領域の隣接関係に関する知識

ラベル	拘束条件	ペナルティ
Sky	Road, Grass, Window とは接しない	e_b
	下半面には存在しない	1
Foliage	Grass と接する	e_b
	二つ以上の領域と接する	隣接領域数
Road	Sky と Window とは接しない	e_b
	Wall か Grass と接する	境界長
Wall	上半面には存在しない	1
	二つ以上の領域と接する	隣接領域数
Grass	Sky と Window とは接しない	e_b
	上半面には存在しない	1
Window	Sky, Road, Wall, Grass とは接しない	e_b

e_b = 共有境界長/境界長

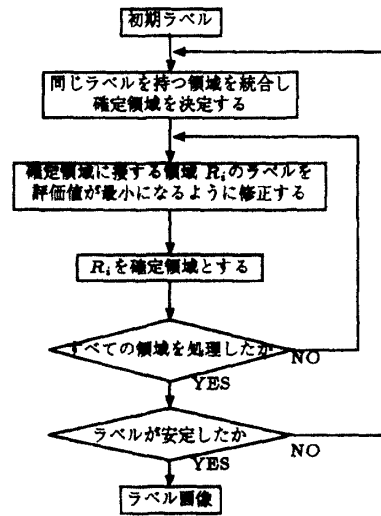


図 1: ラベルの修正法



図 2: 原画像 (256×256, 256 階調)

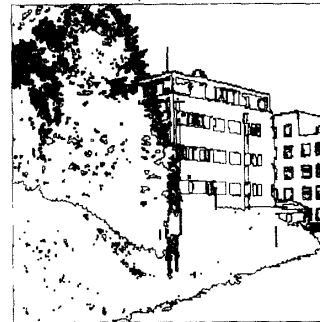


図 3: 領域分割結果 (領域数 2826)

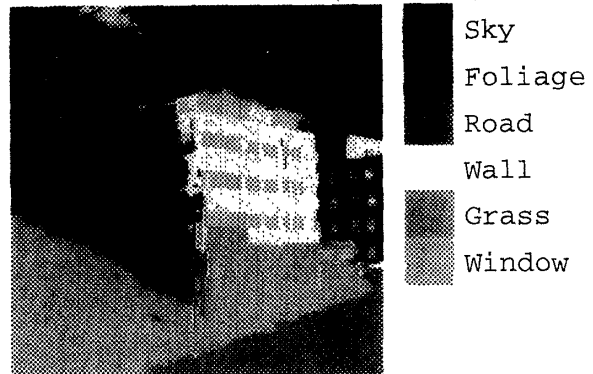


図 4: ラベルの割当結果