

エッジ検出と領域融合によるカラー画像のセグメンテーション

紺谷 精一 赤間 浩樹 三井 一能 串間 和彦

NTT 情報通信研究所

〒 239 神奈川県横須賀市光の丘 1-1

概要 本稿では、写真等のカラー画像からオブジェクトを抽出するため、エッジ情報による領域の差異と統計量による領域の類似性を相補的に利用する画像のセグメンテーション法を提案する。本手法は、1) Mallat の wavelet によるエッジ検出、2) リーマン計量によるカラーエッジ統合、3) エッジによる微小領域の形成、4) MDL 基準による類似領域の融合、のステップからなる。カラー写真のセグメンテーション実験を行ない有効性を例証した。また、芝や森のようなテクスチャー領域の抽出に適用できることも分かった。

Color Image Segmentation using Edge Detection and Region Merging

Seiichi KON'YA, Hiroki AKAMA, Kazuyoshi MII and Kazuhiko KUSHIMA

NTT Information and Communication Systems Laboratories.

1-1 Hikarino-oka Yokosuka-Shi, Kanagawa 239, Japan

Abstract This paper proposes a new image segmentation algorithm, which extracts objects from color photographic images. This method utilizes both edge information which describes dissimilarity between regions and statistical information which measures homogeneity of a region, complementarily. Experimental results on color photographs illustrate the effectiveness of this method and show that this method is applicable to detect texture regions such as grass and woods.

1 はじめに

筆者等は、画像内のオブジェクトを自動抽出し、オブジェクトによる類似検索を行なう画像検索システム - ExSight - を開発している [1]。日本画の素描を画像コンテンツとした場合、色情報を特徴量とするクラスタリングによりオブジェクトの自動抽出が行なえた [2]。しかし、画像コンテンツに写真を選んだ場合は、光線やテクスチャーの影響でピクセルの色情報だけの単純な特徴量ではオブジェクトを抽出できない。

領域を抽出する手法としては、画像特徴量によるクラスタリング以外に Snake [3] や Markov random field (MRF) を応用した手法 [4] がある。Snake は、画像のエッジを利用してテンプレートにマッチする領域を検出できるが、形や大きさの異なるテンプレートを予め用意しておく必要がある。MRF 法は、各点で近傍ピクセルとの共起関係から統計量を計算し、領域を推定する。テクスチャーの抽出に有効であるが、領域の境界では推定誤差が増加し、輪郭がぼやけてしまう。

本稿では、エッジによる領域間の差異と統計量による領域内の類似性を相補的に利用するカラー画像のセグメンテーション法を提案する。

2 セグメンテーションアルゴリズム

本稿で提案するセグメンテーション法は、エッジの検出、カラーエッジの統合、エッジによる微小領域の形成、類似領域の融合のステップからなる。

2.1 wavelet 変換を用いたエッジ検出

画像内の小さな構造と大きな領域の境界とを共に検出するためマルチスケールエッジ検出と呼ばれる手段がある。小さいスケールではノイズやテクスチャー等の細かい情報が検出され、大きいスケールでは全体的な色調の変化が検出される。本手法では、3 段階のスケールのエッジを利用する。エッジ位置も正確で、ノイズやテクスチャーの情報がある程度除かれている中程度のスケールのエッジをベースとする。大きなス

ケールでは、位置はずれるが主要な輪郭情報が得られるので、このスケールの情報を用いてエッジの選択を行なう。また、中程度のスケールのエッジには T 字や十字に輪郭が交わる接合部分などでエッジが欠けることがあるため、小さいスケールの情報でエッジの補間を行なう。

マルチスケールエッジの検出法として、wavelet 変換を用いた手法 [6] がある。wavelet 関数を $\psi(\cdot)$ とすると、画像 $I(x, y)$ の wavelet 変換は以下のように定義される:

$$W_a^x I(x, y) = \frac{1}{\sqrt{a}} \int I(s, y) \psi\left(\frac{x-s}{a}\right) ds \quad (1)$$

$$W_a^y I(x, y) = \frac{1}{\sqrt{a}} \int I(x, t) \psi\left(\frac{y-t}{a}\right) dt \quad (2)$$

ここで、 a はスケール係数。スケール係数を 2 進に離散化 ($a = 2^j, j \in Z$) した wavelet 変換を離散 2 進 wavelet 変換、その係数 ($W_{2^j}^x I(x, y), W_{2^j}^y I(x, y)$) $_{j \in Z}$ を wavelet 係数と呼ぶ。wavelet 関数 $\psi(\cdot)$ としてある平滑化関数 (例えば、Gauss 関数) の一次微分を選ぶと原画像のエッジ情報に相当する wavelet 係数が得られる。

本手法では、小さいスケール、中程度、大きいスケールとして、各々、 $2^{L-1}, 2^L, 2^{L+1}$ のスケールを選ぶ。例えば、 2^2 スケールのエッジをベースとして利用し、 2^3 スケールの情報でエッジの選択、 2^1 スケールの情報でエッジの補間を行なう。

2.2 カラーエッジの統合

前節の手法では、RGB 等の各カラープレーン毎にエッジが検出される。エッジを領域の境界に利用するためには、カラープレーンのエッジを統合し、1 つにする必要がある。各プレーンのエッジを単純に加えるとノイズの影響を受けやすい。リーマン空間上の計量を応用して多次元のエッジ情報からエッジ強度、及び、方向を計算する手法 [7] が提案されている。この手法は、信号の色成分がノイズよりも相関が高いことを利用してノイズを低減させることができる。カラー

プレーン u, v, w の x 方向 (y 方向) へのエッジ強度を各々 $\frac{\partial u}{\partial x}, \frac{\partial v}{\partial x}, \frac{\partial w}{\partial x}$ ($\frac{\partial u}{\partial y}, \frac{\partial v}{\partial y}, \frac{\partial w}{\partial y}$) とすると、計量 g_{ij} は以下の様に定義される。

$$g_{ij} = \begin{bmatrix} p & t \\ t & q \end{bmatrix} \quad (3)$$

$$p = \left(\frac{\partial u}{\partial x}\right)^2 + \left(\frac{\partial v}{\partial x}\right)^2 + \left(\frac{\partial w}{\partial x}\right)^2 \quad (4)$$

$$t = \frac{\partial u}{\partial x} \frac{\partial u}{\partial y} + \frac{\partial v}{\partial x} \frac{\partial v}{\partial y} + \frac{\partial w}{\partial x} \frac{\partial w}{\partial y} \quad (5)$$

$$q = \left(\frac{\partial u}{\partial y}\right)^2 + \left(\frac{\partial v}{\partial y}\right)^2 + \left(\frac{\partial w}{\partial y}\right)^2 \quad (6)$$

エッジ強度 λ と方向 θ は、以下の式で与えられる:

$$\lambda = \sqrt{\frac{1}{2}\{p+q + \sqrt{(p+q)^2 - 4(pq-t^2)}\}} \quad (7)$$

$$\theta = \tan^{-1} \frac{\lambda^2 - p}{t} \quad (8)$$

2.3 エッジの細線化と初期領域の生成

多次元のカラーエッジを統合して得られたエッジは山脈状に分布しているので、尾根に相当する部分を選ぶことでエッジを細線化する。細線化を行なうことにより領域の境界位置が明確になる。ここでは、エッジ方向 θ に沿ったエッジ強度 λ が極大値となるもの(以後、極大エッジ)を選ぶことで細線化を行なう。極大エッジを連結成分に分類し、連結成分から長さ、および、平均エッジ強度が指定した閾値 (Min_{len}, Min_{mag}) 未満のものを削除する。

次に、極大エッジで囲まれた区間を初期領域とする。画像のセグメンテーションは、均質な部分を領域化することで達成される。領域の均質さは、一般に、領域内のピクセル値 (RGB 等) が同一の確率分布に従うと定義される。画像を異なる確率分布を持つ部分に分けることで領域の分割が行なわれる。本手法では、エッジによって差異のある区間が分離されているので、個々の微小領域は均質性が高く、確率分布の推定が正確に行なえる。各微小領域の確率分布を推定し、次節の類似領域の融合を適用することで画像から均質

な領域を得ることができる。本稿では、簡単のため、領域の確率分布として多次元正規分布を仮定する。多次元正規分布は以下の式で表される:

$$P(I|\mu_i, \Sigma_i) = \frac{1}{(2\pi)^{\frac{n}{2}} |\Sigma_i|^{\frac{1}{2}}} \exp \left\{ -\frac{1}{2} (I - \mu_i)^T \Sigma_i^{-1} (I - \mu_i) \right\} \quad (9)$$

ここで、 I はカラーベクトル (RGB 値等)、 μ_i, Σ_i は各領域の平均ベクトル、共分散行列である。 $|\cdot|$ は行列式を、 $(\cdot)^T$ は転置行列を示す。

2.4 MDL 基準による類似領域の融合

エッジによって形成される初期領域は、ノイズやテクスチャーの影響で1つの領域とされるべき部分が複数に分断されることがある。そこで、隣接した類似の領域を融合して適切な領域の抽出を図る。この場合、融合する領域の判定が問題となる。本稿では、モデルとデータの一致性の指標とされる Minimum Description Length (MDL) 基準 [5] を融合の判定に利用する。MDL の減少する場合に領域を融合して行くことで最適なセグメンテーションが達成できると考えられる。MDL は以下の式で定義される:

$$MDL = -LL + \frac{1}{2} N_p \cdot \log M \quad (10)$$

ここで、 LL は対数尤度、 N_p はモデルのパラメータ数、 M はデータ数を示す。

各領域の確率分布として、多次元正規分布を仮定しているので、

$$LL = -\sum_{i=1}^{N_r} \frac{M_i}{2} \log |\Sigma_i| - \frac{1}{2} \sum_{i=1}^{N_r} \sum_{j=1}^{M_i} (I_{ij} - \mu_i)^T \Sigma_i^{-1} (I_{ij} - \mu_i) \quad (11)$$

M_i は、領域 i のピクセル数、 I_{ij} は、領域 i の j 番目のピクセルのカラーベクトル。

$$N_p = \text{Edge_length} + N_r \left\{ n + \frac{1}{2} n(n+1) \right\} \quad (12)$$

$Edge_length$ は領域の境界にあるエッジの長さ、 N_r は領域数、 n は次元数 (RGB ならば $n = 3$)、また、 M は画像の総ピクセル数である。

領域の類似性には、2つの領域の平均ベクトル間の距離を用いている。隣接した領域対から距離の小さい順に融合判定を行ない、MDLの減少する領域対を融合する。

3 実験結果

提案法を用いて写真画像の領域分割を行なった結果を図1から6に示す。画像は、PhotoDISC[8]の24017を用いた。まず、元画像をwavelet変換し、カラーエッジの統合を行なう。スケール 2^2 のエッジ強度を図2に示す。極大値を選ぶと図3の様に細線化される。極大エッジの選択のための閾値 Min_{len} 、 Min_{mag} は共に10とした。エッジの細線化により主要な特徴が得られているが、輪郭の接合部分でエッジが欠けているため、このままではエッジが欠けた部分でつながり有効な領域が得られない(図4)。小さいスケール(2^1)の情報を用いてエッジを補間することで図5に示す有効な初期領域が形成できる。図5は、各領域を平均色によって表示している。最後に、類似領域の融合を行ない、図6に示すセグメンテーションを得る。

次に、テクスチャーのある画像(PhotoCD 19070: 図7)に対してセグメンテーションを行なった。図8に示すように、芝や森のようなテクスチャー領域の抽出が行なえることが分かる。

4 おわりに

写真等のカラー画像のセグメンテーションを行なう手法を提案した。本手法は、1) Mallatのwaveletによるエッジ検出、2) リーマン計量によるカラーエッジ統合、3) エッジによる微小領域の形成、4) MDL基準をによる類似領域の融合のステップからなり、エッジ情報による領域の差異と統計量による領域の類

似性を相補的に利用することで適切な領域の抽出を図っている。カラー写真のセグメンテーション実験を行ない、有効性を例証した。また、芝や森のようなテクスチャー領域の抽出に適用できることも分かった。

現在の類似領域の融合ステップでは、融合判定を隣接した2つの領域間に限定しているため局所最適解に捕らわれ易い。融合する領域の選択に筆者等が提案したクラスタリング法[2]を適用して類似領域の融合ステップの改善を行なうことが今後の課題である。

参考文献

- [1] 赤間, 三井, 紺谷, 串間, “画像内オブジェクトの自動抽出を使った画像検索システム-ExSight-,” DEWS '97.
- [2] 紺谷, 串間, “クラスタ数自動判定クラスタリングによる画像の領域分割,” 情報全 54, 3Q-02, 1997.
- [3] A. Witkin and D. Terzopoulos, “Snakes: Active Contour Models,” Int. J. Computer Vision, Vol.1, No.4, p.321, 1987.
- [4] S. Geman and D. Geman, “Stochastic Relaxation, Gibbs Distributions, and the Bayesian Restoration of Images,” IEEE trans. on PAMI, Vol.6, No.11, 1984.
- [5] 韓, 小林, 「情報と符合化の数理」, 岩波講座応用数学 13, 岩波書店 1994.
- [6] S. Mallat and S. Zhong, “Characterization of Signal from Multiscale Edges,” IEEE trans. on PAMI, vol.14, No.7, July 1992.
- [7] H.-C. Lee and D. R. Cok, “Detecting Boundaries in a Vector Field,” IEEE trans. on SP, vol.39, No.5, May 1991.
- [8] <http://www.photodisc.com/>

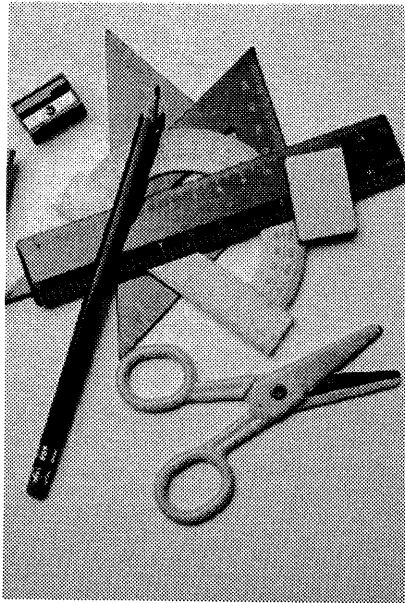


図 1: 元画像

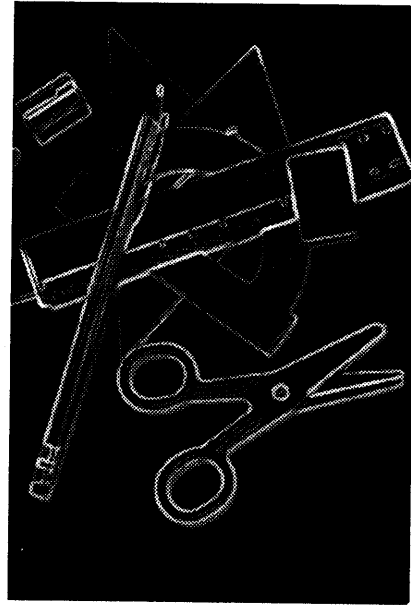


図 2: エッジ強度

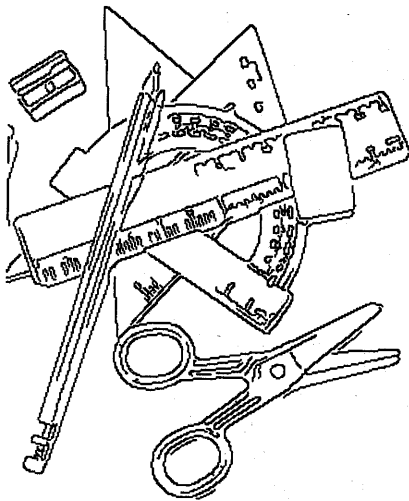


図 3: 細線化したエッジ

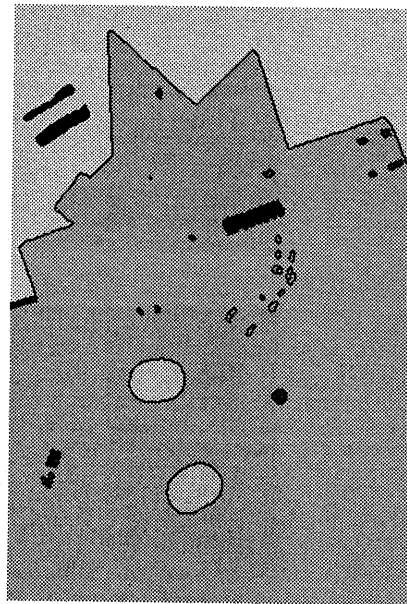


図 4: エッジ補間を行わない場合

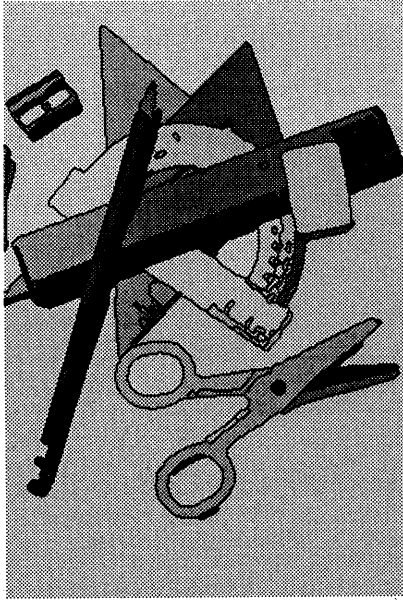


図 5: 初期領域

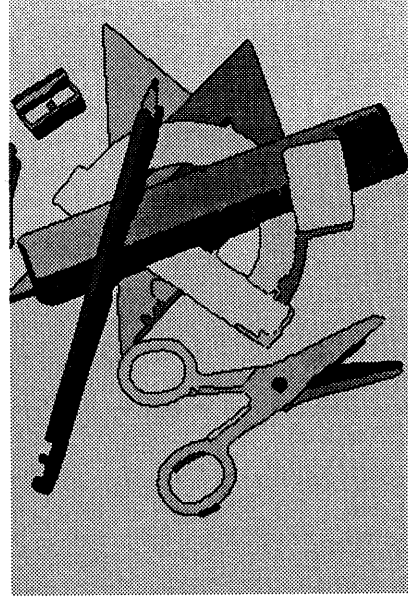


図 6: セグメンテーション結果

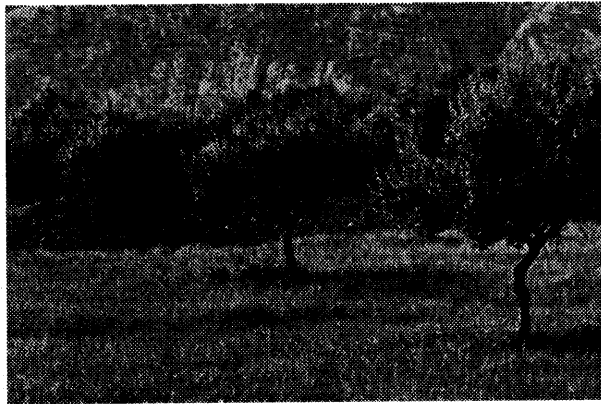


図 7: テクスチャーのある画像

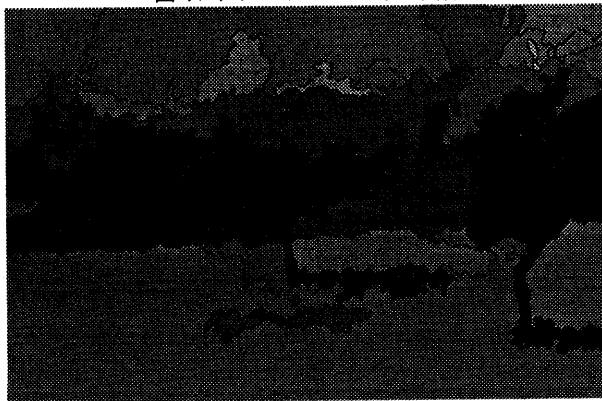


図 8: セグメンテーション結果