

複数の GMRF モデルを用いた屋外シーンの領域分割

越後 富夫†
echigo@trl.ibm.co.jp

飯作 俊一‡
iisaku@crl.go.jp

†日本アイ・ビー・エム（株）東京基礎研究所
〒 242 神奈川県大和市下鶴間 1623-14

‡郵政省通信総合研究所
〒 184 東京都小金井市貫井北町 4-2-1

屋外の自然画像にはテクスチャ領域が多く含まれる。符号化や画像理解の目的には、テクスチャ領域を色だけでなく、テクスチャ特徴を利用して、特徴毎に領域を切り出すことが重要である。テクスチャ特徴を表現するには、画素単位ではなく、領域画像を必要とするため、領域境界近傍でテクスチャ特徴を利用するところが困難であった。本手法では、ガウシアンマルコフ確率場 (GMRF) でモデル化し、複数の定義パラメータを使い分け、領域境界で併合初期の小領域にもテクスチャ特徴を表現する GMRF パラメータを使った領域併合を行なう。さらに併合領域が拡大すると、より信頼度の高い GMRF パラメータで領域併合を行なう。以上のようにして、なめらかな境界線で画像を分割し、かつ、境界線近傍までテクスチャ特徴で領域併合を行なう手法を提案する。

Unsupervised Segmentation of Outdoor Scenes by using Multiple GMRF Models

Tomio Echigo†

Shun-ichi Iisaku‡

†IBM Research, Tokyo Research Laboratory
1623-14, Shimotsuruma, Yamato-shi,
Kanagawa 242

‡Communication Research Laboratory
Ministry of Posts and Telecommunications
4-2-1, Nukui-Kita-machi, Koganei-shi,
Tokyo 184

The goal of this research is to make region-based coding method of images effective. Texture regions implied in outdoor scene should be segmented by not only colors, but also textured features. Gaussian Markov Random Fields (GMRF) represent texture efficiently, however, the parameters of GMRF requires a large region for exactness, which means GMRF is not available in neighborhood of region boundary. Multiple models of GMRF can merge small regions in neighborhood of region boundary and at the initial merging steps, because a few parameters require small regions and regions merged reliably by more parameters can be obtained from gathering small regions.

1 はじめに

マルチメディアパソコンやインターネットのブームのおかげで、個人が簡単に画像データを所有できるようになった。データが増えてくると、管理の重要性が顕著になる。個人所有のデータだけなら、サムネールと簡単な注釈で管理するツールもあるが、一方、電子図書館や電子博物館など、膨大なデータベースを管理しているところでは、サムネールでも膨大なデータ量となるため、データを検索できることが必要になる。テキストデータでは、検索項目が内容の一部に一致する可能性が高いが、画像データの管理では、現在は管理者や所有者が付けた注釈やインデックスに基づくため、検索項目がヒットする可能性は低い。今後、データがますます増加するため、管理者の付けた注釈に依存せず、システムにより、画像の検索を支援することが重要となる。

現在、標準化している画像圧縮技術の JPEG は、画像を小ブロック毎に分割して DCT を行っている。また、動画像符号化の MPEG でも画像をブロック毎に扱っている。これらの画像符号化データに対し、インデックス、検索をシステムが支援するには、復号して画像処理することが必要になる。復号後の画像処理では、圧縮による画像の劣化も生ずるため、処理結果の信頼度は低下する。そこで、高い圧縮比だけでなく、インデックス、検索、部分画像の再利用等に有効な、次世代画像圧縮技術として、圧縮前の原画像を類似領域に分割し、領域毎に圧縮する領域分割圧縮法が有望視される。しかしながら、その前処理として重要な領域分割が、屋外自然風景を含む画像に対して、良好な結果を得る方法が確立されておらず、領域分割圧縮法の推進に障害となっている。

画像の領域分割は、画像処理の長年の研究テーマであり、特に、テクスチャ領域の分割は、テクスチャ解析を発展させたものが多い。領域分割は画像認識や解析を目的とした時、画像データからボトムアップで領域を正確に抽出するのは困難である。しかし、符号化が目的ならば、近傍で類似していれば領域として併合するほうが圧縮効率が良く、その領域の意味理解とは問題が異なる。そのとき、分割するパラメータは、色だけでなくテクスチャ形状も重要であり、テクスチャ領域で過度の分割を低減する必要がある。

従来からの代表的な領域分割手法として、Ohlander [1]、富田と辻 [2] は色情報を数多くのパラメー

タで表現し、そのヒストグラムが明白なピークを持つ属性から優先的に分割する再帰的分割を行った。富田等の方法では、領域境界のウインドウの大きさの選択を工夫している。また、内山 [3] は領域内部と境界周辺でスケールスペースの変更により、フィルタを使い分けて領域の併合を行った。以上の方法は、主に色情報に基づいた領域分割および併合を行っている。

一方、テクスチャの空間的情報を利用したモデルとして、マルコフ確率場 (MRF) を利用した研究が多い。Besag [4] は MRF パラメータの最尤推定に coding 法を提案し、Geman and Geman [5] は隠れ層としてエッジの存在を明に表すライン過程を導入し、MRF とギブス分布の同値性により、エネルギー最小化法で MRF パラメータの推定を行った。山崎等 [6] は、雑音は領域に依存すると仮定して、領域分割を行なった。以上は、一般的なマルコフ場を扱っていたが、実際の領域分割では、対象画像に制約を置かないため、ガウス型の分布で十分である。金子 [7] はガウスマルコフ場 (GMRF) を用いて、テクスチャの分類が有効であることを示した。また、Panjwani と Healey [8] は、GMRF におけるノイズが色プレーン間で影響しないとして、GMRF パラメータを線形演算で求め、併合のための 2 つの候補領域の疑似尤度と、併合後の疑似尤度の比が小さい候補から領域を併合した。Panjwani 等の方法は、自然画像に対し良い分割結果を示しているが、最小分割単位を 4×4 の矩形としたため、領域境界で矩形形状が目立ち、さらに境界近傍では、GMRF を用いるのではなく、色情報だけの領域併合しかできない欠点があった。

そこで本研究では、Panjwani 等と同様に GMRF を用いて線形演算でパラメータを求めるが、画像を RGB プレーンではなく、大田 [9] が提案した変換式により、色プレーン間で相関が少なくなるようにする。さらに、最小分割単位を矩形ではなく任意の形狀になるようにし、最小分割単位から色による領域併合を行なわずにテクスチャ形状が利用できるようにする。そのためには、複数の GMRF モデルを利用し、併合初期や領域境界近傍における小領域に対しては、少数のパラメータで併合を進め、領域が大きくなるに従い多数のパラメータで信頼度の高い領域併合を行う。

2 色特徴の変換

RGB 色プレーンは、各々に輝度情報を持つことから明らかのように、プレーン間での相関が高い。領域分割では、色パラメータに相関が無い軸を用いる方が効率的である。Karhunen-Loeve (KL) 変換は、相関の強いパラメータ軸を、相関が最小となる軸に変換するには最適な変換であるが、最適な変換は画像に依存するため、領域分割のための前処理として全ての画像に対して KL 変換を行うには、処理が重すぎる。そこで、大田は複数枚の画像の選択した領域に対し固有ベクトルを計算し、その主成分から次式

$$\begin{aligned} I_1 &= (R + G + B)/3 \\ I_2 &= (R - B) \\ I_3 &= (2G - R - B)/2 \end{aligned} \quad (1)$$

が領域分割に有効であることを示した。対象とした画像は、テクスチャを含む自然画や人物画、人工的模様を含む室内などである。大田の変換式では、全ての画像において最適な変換が施されるわけではないが、少なくとも大田が用いた画像と同種の画像に対するは、前処理として KL 変換を行ったのとほぼ同等の結果が得られる。

本研究でも、画像の前処理として式(1)を採用する。符号化を目的としたとき対象となる画像は特定できないが、領域分割が最も困難なテクスチャ領域に対し、式(1)が有効で、テクスチャ領域以外でも、RGB プレーンを利用するより優位であると判断したためである。

3 GMRF モデルと疑似最尤法

3.1 複数の GMRF モデル

画像 X_{ij} のポテンシャルをガウス分布とすると、GMRF は次式のようになる。

$$P(X_{ij}/\varepsilon_{mn}) = K \exp(-U(i,j)) \quad (2)$$

K は正規化定数で、

$$U(i,j) = \frac{1}{2}[u_1 u_2 u_3] \Sigma^{-1} [u_1 u_2 u_3]^t \quad (3)$$

ここで $[u_1 u_2 u_3]$ は平均が 0 のガウスノイズベクトルとし、 Σ はノイズ共分散行列で

$$\Sigma = \begin{pmatrix} \nu_{11} & \nu_{12} & \nu_{13} \\ \nu_{21} & \nu_{22} & \nu_{23} \\ \nu_{31} & \nu_{32} & \nu_{33} \end{pmatrix} \quad (4)$$

であるが、2. で色プレーン間の相関を小さくしたので、 Σ は対角要素だけを採用し、その他は 0 とする。そのとき、画素 (i,j) における空間的相互作用は次式のように表される。

$$u_1(i,j) = \{I_1(i,j) - \mu_1\} \quad (5)$$

$$- \sum_{\lambda=1}^3 \left[\sum_{(m,n) \in \epsilon_{1\lambda}} \alpha_{1\lambda}(m,n) \{I_\lambda(i+m, j+n) - \mu_\lambda\} \right]$$

$$u_2(i,j) = \{I_2(i,j) - \mu_2\} \quad (6)$$

$$- \sum_{\lambda=1}^3 \left[\sum_{(m,n) \in \epsilon_{2\lambda}} \beta_{2\lambda}(m,n) \{I_\lambda(i+m, j+n) - \mu_\lambda\} \right]$$

$$u_3(i,j) = \{I_3(i,j) - \mu_3\} \quad (7)$$

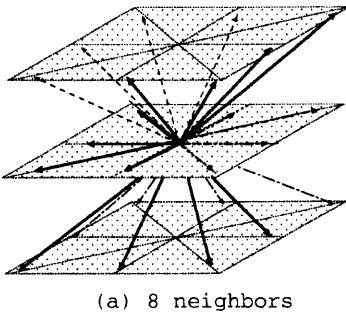
$$- \sum_{\lambda=1}^3 \left[\sum_{(m,n) \in \epsilon_{3\lambda}} \gamma_{3\lambda}(m,n) \{I_\lambda(i+m, j+n) - \mu_\lambda\} \right]$$

本研究では、近傍系 ε を 4 近傍と 8 近傍の 2 種類とした。この様子を図 1 に示す。ただし、4 近傍では色プレーン間の相互作用が無いとし、式(5),(6),(7) の α, β, γ の色プレーン間の相関パラメータを 0 にする。これは式(1)により、色プレーン間で相関が小さくなっている、同一色プレーンにおけるパラメータに比較して影響が小さいと考えたからである。したがって 4 近傍では、対称性を考慮すると、6 つの GMRF パラメータで表現できる。一方、8 近傍では、近傍距離を少し大きくしたので、色プレーン間の相関も無視できなくなり、相関パラメータを 8 近傍全てに利用する。したがって 8 近傍では 36 のパラメータで表現する。

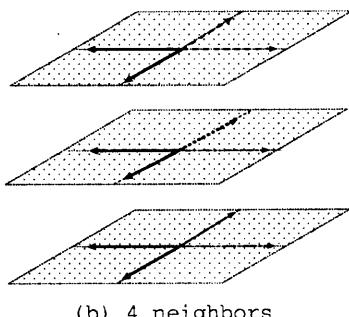
3.2 疑似最尤法によるパラメータ推定

パラメータ α, β, γ を求めるために、本研究では、疑似尤度を用いる。GMRF でモデル化した領域 S の疑似尤度は、

$$\prod_{(i,j) \in S} P(X_{ij}/\varepsilon_{mn}) \quad (8)$$



(a) 8 neighbors



(b) 4 neighbors

図 1: 8 近傍と 4 近傍のパラメータとなる。対数疑似尤度を最大にするパラメータを求めるため、全てのパラメータに関する微分を行ない、その値を 0 とする。式(4)でノイズ共分散行列を対角行列としたので、パラメータは線形結合となり、連立演算から求められる。

その結果、ノイズ共分散行列の対角要素は次式となる。

$$\nu_{11} = \frac{1}{M_S} \sum_{(i,j) \in S} \{u_1(i,j)\}^2 \quad (9)$$

ただし、 M_S は領域 S の面積。 ν_{22}, ν_{33} についても同様である。

3.3 疑似尤度比による併合

画像が領域 S_1, \dots, S_n で分割されているとして、その中から S_k と S_l の併合について考える。 S_k と S_l が併合する前の疑似尤度と、仮に S_k と S_l が併合されて領域 S_m になったときの疑似尤度の比は、

$$\begin{aligned} R_p(k, l) &= \frac{\prod_{r=1,..,n} P(S_r)}{P(S_m) \prod_{r=1,..,n; r \neq k, l} P(S_r)} \quad (10) \\ &= \frac{P(S_k) P(S_l)}{P(S_m)} \end{aligned}$$

両辺の対数をとり、整理すると、

$$h(k, l) = \frac{M_k + M_l}{2} \ln(|\Sigma_{S_m}|) - \frac{M_k}{2} \ln(|\Sigma_{S_k}|) - \frac{M_l}{2} \ln(|\Sigma_{S_l}|) \quad (11)$$

となり、画像全体から隣り合う領域の組合せを全て求め、その最小値 $h(k, l) = \min_{i, j} h(i, j)$ となる k と l を併合する。4 近傍と 8 近傍に対しても、同じように対数疑似尤度比を求め、同じ評価で最小値となる領域の組合せを併合する。

4 分割併合法

本手法では、分割併合法による領域分割を行う。その処理の流れを図 2 に示す。

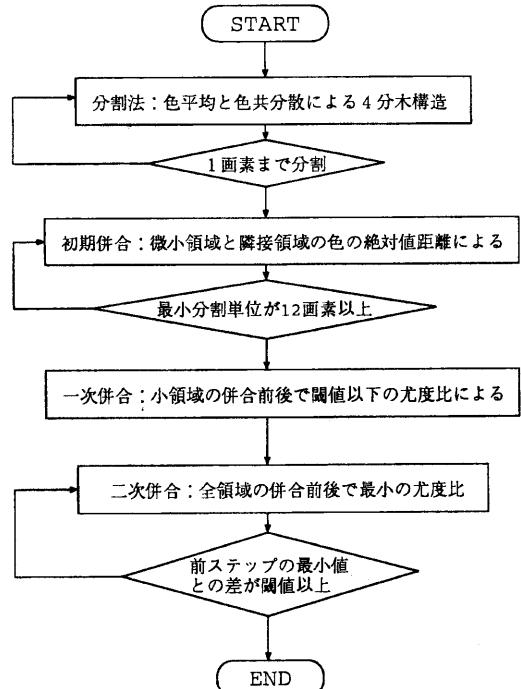


図 2: 処理の流れ

4.1 分割法

画像全体を、大きさが 32×32 のブロックに分割し、初期ブロックとした。初期ブロックの色プレーンにおける平均値と共分散行列を求め、同様にそのブロックを 4 分割したと仮定したときの平均値と共分

散行列を求める。もし4分割した全てのブロックの平均値と、初期ブロックにおける平均値の差が、ある閾値 $ThreshColor$ 以内のとき、かつ、全4ブロックの共分散行列の全要素と、初期ブロックの共分散行列の全要素の各々の差が、ある閾値 $ThreshCo-val$ 以内のとき、初期ブロックの分割を停止する。それ以外のとき、初期ブロックを4分割する。ここで、分割されないブロックは、以後の処理で分割されることがないため、閾値 $ThreshColor$, $ThreshCo-val$ はできるだけ小さくしておく。分割されたブロックに対しては、同様の手順で、再帰的に処理を繰り返し、4分木構造で最終的に1画素になるまで分割を行う。

この分割法では、4分木構造としたため、水平または垂直でない領域境界線の近傍では細かく領域が分割される。初期ブロックの大きさに近い分割されない領域は、最終的に併合される領域の中央部に多く、それがテクスチャ領域であっても、平滑フィルタで、色情報による併合が行われたのと等価である。

4.2 初期併合

分割法で、最小分割領域は1画素であった。テクスチャ領域を表現するには、画素の近傍領域の色の変化を保持する必要があり、そのパラメータを推定するには、少なくとも数画素からなる領域を必要とする。GMRFによる併合では、画素ではなく、最小分割単位となる領域が対象になるため、その画素数が少ないと、領域の隅々まで、テクスチャ情報を利用した領域分割が行えたと言える。しかし、GMRFパラメータを求めるには、少なくとも数画素を要するため、本研究では、GMRFパラメータが決定できる最低許容画素数として、GMRFパラメータの2倍とした。即ち、4近傍のGMRFでは6つのパラメータがあるので、最低12画素を必要とする。また8近傍36のGMRFパラメータを求めるには、72画素以上を必要とする。

初期併合においては、色情報による併合を行うため、最小分割単位を小さく設定する。本研究では、4近傍のGMRFパラメータを求めるための最小許容画素数の半分の6画素と定めた。

初期併合では、色情報を用い、最小分割単位の6画素になるまで、併合を行う。そのため、5画素以下の領域 I に対し、その隣接する周囲領域 $candidateI_k$

において、色の差の絶対値総和が最小となる

$$\min_k \left\{ \sum_i^3 |I_i - candidateI_{ki}| \right\} \quad (12)$$

のとき、その領域を併合する。隣接は4近傍だけとし、斜め方向は隣接とみなさない。図3のように、1領域に対し隣接する1つの候補を選択し、併合させるが、その候補がチェーンで継って併合することがあり、一度の初期併合で数画素が同時に併合することがある。初期併合は、全ての領域が最小分割単位の6画素以上になるまで繰り返すので、初期併合を終了後は最小分割領域の画素数は6以上になる。

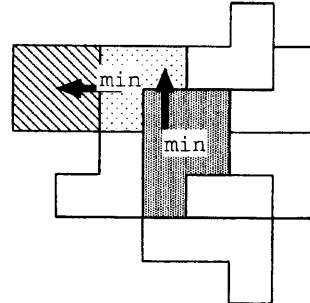


図3: 初期併合のチェーン

4.3 複数のGMRFによる併合

初期併合によって、領域内画素数は6画素以上が保証されている。しかし、6画素では、4近傍のGMRFパラメータを安定に求めることができない。4.2に記述したように、4近傍でも最小12画素を要したので、12画素未満の領域では、単独でGMRFパラメータを求めることができない。そこで、12画素未満の小領域において、隣接する領域を候補とし、そのうち色の差の絶対値総和が最小となる領域を、仮の候補領域とする。最小領域でも6画素以上が保証されているので、少なくとも2つの領域を併合した仮の候補領域から、GMRFパラメータを求めることができる。この様子を図4に示す。図のように、仮の候補領域は2つの領域の併合とは限らず、複数の領域からなることもあり、仮の候補領域から8近傍のGMRFパラメータが求められることもある。

仮の候補領域を含めると、全ての領域に対してGMRFパラメータを計算することができる。しかし、12画素程度の小領域は画像に数多く存在し、し

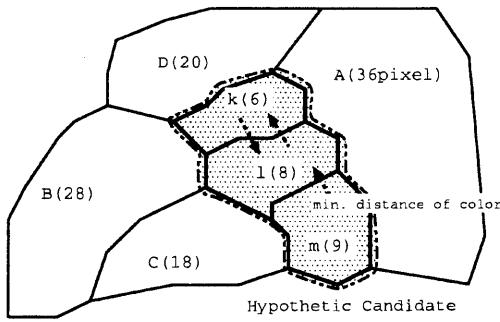


図 4: 小領域の仮の併合による候補領域

かも次のステップ毎の併合は候補が多い程、処理時間に費すので、処理時間の削減のため、ステップ毎の併合でいざれ明らかに併合されるであろう領域を、前もって一度に併合する。

そこで、72画素未満の領域（仮の候補を含む）に對し、隣接する領域対の全ての組み合わせにおいて、併合前後の疑似尤度比が閾値 $ThreshMerge$ 以下のとき予め併合する。ただし、色の差の絶対値総和が

$$\sum_i^3 |I_i - candidateI_{ki}| < ThreshDiff \quad (13)$$

となるときとする。以上を一次併合と呼ぶ。たとえば、図4では、Aと仮の候補(HC), BとHC, CとHC, DとHC, AとD, BとD, BとCの組み合わせがあり、そのうち、CとHC, DとHCが条件を満たすと、CとDとHCが併合される。ここで閾値 $ThreshDiff$ は、領域併合において、最終的に併合したくない色の差の総和で、比較的大きな値とし、閾値 $ThreshMerge$ は、処理時間の短縮のために用いるので、小さな値とする。

一次併合が施されると、領域の数および形状が変化するので、領域の隣接関係、パラメータも更新される。仮の候補領域が一次併合によって併合されるとき、仮の候補として集められた領域全てを、眞の領域と同じように併合する。しかし、一次併合で併合されなかった仮の候補領域は、もと通り分割される。一回の一次併合で、領域分布が変わるので、12画素未満の領域に対しては、仮の候補を再度生成し、一次併合を再度施す。こうして、一次併合によって併合される領域がなくなるまで、処理を繰り返す。

画素数の大きな領域に対しては、多くのパラメータで表現するほうが、併合において信頼度が増すため、対称な8近傍で、色ブレーン間の相互関係を表

すパラメータを含む36のパラメータの計算には、2倍の画素数となる72画素以上の領域に對して適用する。したがって、72画素以上の画素数の領域同士の併合には、8近傍パラメータのGMRFを利用し、片方の併合候補が71画素以下ならば、4近傍パラメータのGMRFを利用する。領域同士の組み合わせは、72画素以上の大領域、12以上72画素未満の小領域、12未満の領域が複数集まって72画素以上となる仮の大領域、仮の小領域から組合わさったペアとして8通りある。そのうち、8近傍のパラメータが求められるのは、大領域と大領域、大領域と仮の大領域の2通りだけで、残り6通りの場合は4近傍のパラメータを求める。

全ての領域に對し、隣接する全ての候補領域を見い出し、式(11)で領域併合を仮定した疑似尤度と、併合前の2つの候補の疑似尤度の比が最も小さい候補から併合を行う。このとき、8近傍で求めた疑似尤度と4近傍で求めた疑似尤度があるが、それらは區別しない。ただし、併合する条件として候補対の領域内色パラメータの差の絶対値総和が、閾値 $ThreshDiff$ 以下のときとする。以上の処理を二次併合と呼ぶ。

二次併合においては、候補対を併合する毎にデータを更新し、ステップ的に行う。そして、そのときの最小疑似尤度比が、前回の併合ステップにおける最小値に比べ、閾値 $ThreshStop$ 以上になったとき、併合を停止し、領域分割処理を終了する。

5 実験結果

実験には、SIDBA の home3.ppm を、家のレンガ、木、芝生のテクスチャと、テクスチャの無い空を含む 128×128 画素の部分画像に切り出して使用した。前もって定義した閾値は次のようにある。

$ThreshColor:$	5	4分木ブロックの単一色平均の差
$ThreshCo-vak:$	18	4分木ブロックの色共分散の差
$ThreshDiff:$	72	併合を禁止する色の絶対値距離
$ThreshMerge:$	10	一次併合小領域の疑似尤度比
$ThreshStop:$	5000	二次併合を停止する疑似尤度比 のステップ間差分

図5は4分木による分割結果、図6は色に基づく初期併合の結果を表す。図7は、小領域を4近傍のGMRFモデルで併合した一次併合の結果、図8はステップ的に併合を繰り返す二次併合を行なった最終

分割結果を示す。最小分割単位を任意形状にしたので、領域境界にブロック状の歪みは無かった。

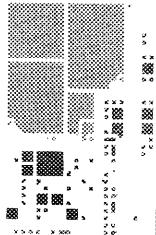


図 5: 4 分木分割

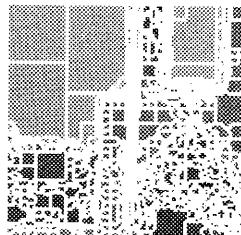


図 6: 初期併合

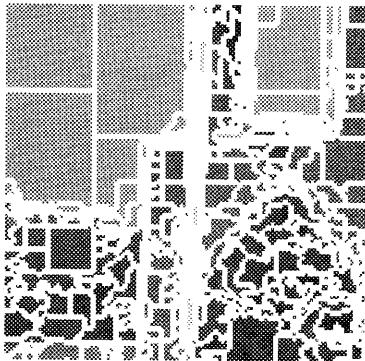


図 7: 一次併合

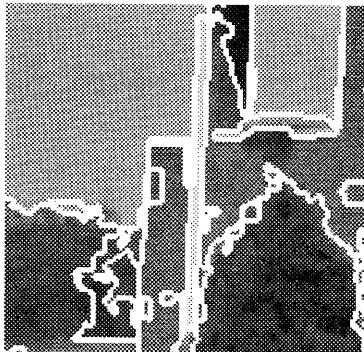


図 8: 領域分割結果

6 おわりに

GMRF パラメータの推定に信頼度を得るには、多くの画素数を必要とする。領域の中央部では、大きな領域をとることができるので、領域境界近傍では画素数が不足するため、従来では、GMRF による領域併合が行なえなかった。

そこで、本研究では複数の GMRF を用いて、領域境界近傍や併合初期段階においても、テクスチャ形状を表現した GMRF モデルが利用できることを提案した。そのためには、少ないパラメータでテクスチャ形状を表現する必要があるため、画像に前処理を施して、色プレーン間の相互作用を小さくした。さらに、4 近傍で 6 つのパラメータにより GMRF を表し、少なくとも 12 画素以上の領域であれば、GMRF による領域併合を行なうために、最小分割単位を 6 画素の領域とし、12 画素未満の領域については、複数の領域を集めた仮の候補領域とすることで GMRF パラメータを求めることが可能にした。また、領域併合の信頼度を上げるために、領域が成長し 72 画素以上になると、8 近傍で色プレーンの相関も考慮した、36 のパラメータで表現した GMRF モデルを用いた。このように本研究では、領域の隅々まで GMRF を用いたテクスチャ形状による領域分割が行えた。本研究で用いた閾値は、制約が緩く設定しており、分割数を明示する必要がないので、対象画像を限定しない汎用的な手法であると言える。

本研究で得られた領域は、テクスチャ形状、色の類似度でまとめられているため、領域分割符号化には都合が良く、領域内データの圧縮には効率が良いと考えられる。

謝辞

本研究に対し、役立つ情報を供与して下さった筑波大学大田友一教授、UC Irvine Prof. Healey に深く感謝します。また、有益な助言をいただいた郵政省通総研山崎達也氏に感謝します。

参考文献

- [1] R. Ohlander, K. Price, and D. R. Reddy, Picture Segmentation using a Recursive Region Splitting Method, Computer Graphics and Image Processing, pp. 313-333, 1978.
- [2] F. Tomita and S. Tsuji, Computer Analysis of Visual Textures, Kluwer Academic Publishers.
- [3] 内山、武川、多重スケールフィルタを用いたカラー テクスチャ画像の領域分割、信学技報、PRU95-223, pp. 55-62, 1996-03.
- [4] J. E. Besag, Spacial Interaction and the Statistical Analysis of Lattice System, J. Roy. Statist.

Soc., B36, pp. 192-236, 1974.

- [5] S. Geman and D. Geman, Stochastic Relaxation, Gibbs Distribution, and the Bayesian Restoration of Images, IEEE Trans., PAMI-6, no. 6, pp. 721-741, 1984.
- [6] 山崎, シラジ, 野田, MRF モデルを用いた劣化画像のデータ駆動型セグメンテーション, 信学論, D-II, vol. J75-D-II, No. 6, pp. 1075-1084, 1992.
- [7] H. Kaneko and E. Yodogawa, A Markov Random Field Application to Texture Classification, Proc. PRIP-82, pp. 221-225, 1982.
- [8] D. K. Panjwani and G. Healey, Markov Random Field Models for Unsupervised Segmentation of Textured Color Images, IEEE Trans. PAMI, vol. 17, no. 10, pp. 939-954, 1995.
- [9] Y. Ohta, Knowledge-based Interpretation of Outdoor Natural Color Scenes, Research Notes in Artificial Intelligence 4, Pitman Advanced Publishing Program.