

## 二分探索型 Belief Propagation の 多次元拡張による領域分割

関 真規人<sup>†1</sup> 和田 俊和<sup>†1</sup>

近年では、特徴空間でのクラスタリングと画像空間での離散最適化を反復的に行い、各画素に最適なクラスタ中心値を割り当てる領域分割手法が提案されている。この方法は、特徴空間と画像空間の統合処理というだけでなく、新しい領域分割の枠組みへの見通しを示唆している。しかしながら、この方法は安定性と高速性に課題がある。この問題を解決するために、本研究では、特徴空間の分割と画像空間における離散最適化を行う方法を提案する。提案手法では、特徴空間の分割と Belief Propagation による最適化をいずれも階層的に行う。さまざまな画像を用いた実験では、精度と速度の観点で従来手法に対する有効性を確認することができた。

### Image Segmentation using Hierarchical Belief Propagation

MAKITO SEKI<sup>†1</sup> and TOSHIKAZU WADA<sup>†1</sup>

Recently, a new image segmentation method has been proposed, that iteratively performs clustering in the feature space and discrete optimization in the image space, which assigns optimal cluster center values to pixels. This method provides us a vistas on novel image segmentation framework as well as a unified processing both in feature and image spaces. Unfortunately, this method has problems on stability and processing speed. For solving this problem, we propose a method that performs feature space partitioning and discrete optimization in the image space. The optimization by Belief Propagation and the feature space partitioning are done both in hierarchical manner. Through extensive experiments, we confirmed that our method performs better than existing method both in terms of accuracy and speed.

### 1. はじめに

一般物体認識<sup>1)</sup>におけるアプローチのひとつとして、画像から得られる局所特徴量の集合(例えば Bags of keypoints<sup>2)</sup>と呼ばれる特徴)により画像を分類する方法がある。これは画像検索などの用途では有効なアプローチであるといえるが、空間的な情報を使わないため一般的な物体認識への適用には限界があると考えられる。

別のアプローチに、画像の領域分割を行い、領域の形状や色のほか、各領域における SIFT 特徴量などの高次元特徴を手がかりに物体認識あるいは領域の分類を行うアプローチがある<sup>3)-5)</sup>。かなり古くから行われてきた方法であるが、上位の枠組みとして領域間の隣接関係、包含関係などの依存関係をも扱った認識方法を考えれば発展性が十分にあるのではないかと思われる。実際に、昨今の研究事例においても、領域の隣接関係を Markov Random Fields (MRF) によりグラフ構造で表し、依存関係を考慮しつつ領域(ノード)間でのメッセージパッシングを行って物体を認識する試みがなされている<sup>6),7)</sup>。

以上のような考えのもと、我々は領域分割に基づいた物体認識の研究に取り組んでいる。本稿では、こういった物体認識への適用を想定した領域分割手法について述べる。なお、ここでは物体ごとに分割された領域を求める必要はなく、やや過剰気味に分割されていればよいと考える。ただし、過剰に分割されすぎると領域が小さくなって有効な特徴量が得られず、また領域ごとのクラス分類のために処理時間がかかってしまうため、適度な分割であることが求められる。

以下では、2章において、領域分割の関連研究について概観する。次に3章において提案手法を述べ、4章で実験結果を考察する。最後に、5章にてまとめる。

### 2. 関連研究

昨今の典型的な領域分割手法は、大域的な最適解を得るために、画素値に対する画像空間での連結性評価と特徴空間でのクラスタリングによって行われている。そのひとつに、画像における画素の隣接関係を MRF で表現し、画素値分布に対する統計モデルのもとで各画素の事後周辺確率を最大にするクラスラベルを付加するといった方法がある<sup>8)-10)</sup>。まず次節にてこの従来手法を説明する。

<sup>†1</sup> 和歌山大学  
Wakayama University

なお、ほかに、Mean-Shift に基づく方法<sup>11)</sup> や、Normalized Cuts を用いた方法<sup>12)</sup>、Blobworld と呼ばれる方法<sup>13)</sup> などもある。Mean-Shift に基づく方法は、特徴量のヒストグラムに対して極値を探索することで画素をクラスタリングした後、後処理として画像上での連結性を評価する 2 段階処理となっているため、前段での処理結果に大きく依存してしまうという問題がある。Normalized Cuts はスペクトラルクラスタリングと呼ばれるクラスタリング手法のひとつであり、画像空間での連結性と特徴空間でのクラスタリングを同時に扱うことができるが、基本的に固有値問題に帰着させて解くため、計算に時間がかかってしまうことが問題である。一方、Blobworld は画素の座標値と特徴量で構成される空間で Gaussian Mixture Model (GMM) を当てはめることによりクラスタリングを行うが、画像空間の座標値を含めてクラスタリングするため、大きな物体領域が分割されてしまったり、物体境界がはっきりしていても分割できないという問題がある。

### 2.1 MRF に基づく領域分割手法

従来手法として、MRF に基づく領域分割手法について説明する。

MRF は多数の変数とその変数間の相互作用からなるグラフモデルの枠組みである。文献 8)–10) などでは、個々の画素値が独立に生成されたものではなく、互いに関連しあって意味をなしたデータとなっていると考え、領域分割を MRF のもとでの最適化問題として捉えている。具体的には、画素をノードとする格子状のグラフを構築し、画素の観測値を隣接画素間で相互に作用させることによって、画素に割り当てられるべきクラスラベルについての大域的な最適解を算出する。また、このとき、特徴空間における画素値の分布として統計的な分布モデルを仮定し、そのパラメータを同時に推定することで、統計モデルのもとで各画素の事後確率を最大にするクラスラベルが求まるようになっている。なお、文献 10) では統計モデルに GMM を仮定し、そのパラメータの推定には Expectation-Maximization (EM) アルゴリズムを用いている。また、最適解の算出には Belief Propagation (BP:後述) を用いており、二重の最適化ループによって処理が行われている (図 1)。以下に処理手順を示す。

- (1) GMM の初期パラメータを設定する。
- (2) GMM のもとで、各画素が各クラスに帰属する事後確率を BP により算出する。
- (3) 事後確率をもとに GMM パラメータを更新する。
- (4) GMM パラメータが収束するまで、(2) から (3) を繰り返す。

### 2.2 従来手法の問題点

この方法は、画像空間と特徴空間の双方で最適化を図るという観点で有効なアプローチであると考えられる。しかしながら、それら最適化における主従関係がないため、条件が悪け

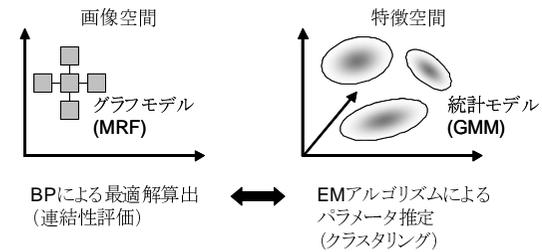


図 1 MRF に基づく従来手法の概念図  
Fig. 1 Concept of image segmentation based on MRF

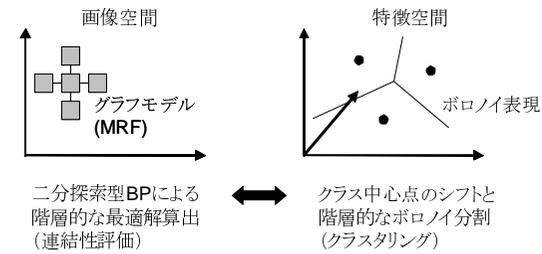


図 2 提案手法の概念図  
Fig. 2 Concept of proposed method

れば解が発散し、不安定になるという可能性がある。また、一般に、GMM には面積の大きい領域の特徴量に左右されたり、パラメータの初期値に依存して結果が変わるなどの課題もある。さらには、EM アルゴリズムや複数のクラスラベルを扱う BP の最適化計算に時間がかかるという問題もある。

## 3. 提案手法

本研究では、安定化、高速化に重点をおき、画素値分布に対する統計モデルとそのパラメータ推定方法を単純化する。具体的には、特徴空間にクラス中心点を与え、ボロノイ分割することで画素値分布のクラスタリングを行う。このとき、階層的にクラス中心点を増やすことで、ボロノイ分割を更新する。また、クラス中心点を適宜シフトすることで最適化を図る。一方、これに合わせて、MRF のもとでの最適化も階層的に行う。そのために、二分探索型 BP<sup>14)</sup> を拡張して適用する (図 2)。

### 3.1 Belief Propagation

二分探索型 BP を説明する前に、一般的な BP について述べる。

BP は、MRF における最適解を具体的かつ近似的に解く方法のひとつである。画像の領域分割を対象とするとき、BP では、画素に割り当てられる複数のラベル候補のうち、式 (1) を最小にするラベルを画素ごとに算出する。ここで、 $f_q$  は画素  $q$  におけるラベル候補、 $m_{pq}$  は画素  $p$  から画素  $q$  へ送られるエネルギー（メッセージと呼ぶ）であり、 $N(q)$  は  $q$  に隣接する画素集合を表す。また、 $D_q(f_q)$  はラベル候補  $f_q$  を画素  $q$  へ割り振るときのデータコストであり、観測された画素値とラベル  $f_q$  の差異を表すペナルティ関数として用意される。

$$b_q(f_q) = D_q(f_q) + \sum_{p \in N(q)} m_{pq}(f_q) \quad (1)$$

メッセージ  $m_{pq}$  は隣接画素間での相互作用を実現するためのものであり、BP では画像全体での大域的な最適解を得るためにメッセージパッシングを行う。具体的には、すべての画素について、隣接する 4 つの画素間で式 (2) を反復計算することになる。ここで、 $N(p) \setminus q$  は  $p$  に隣接する  $q$  以外の画素集合を表す。また、 $V(f_p, f_q)$  はラベル候補  $f_p$  と  $f_q$  を隣接した 2 つの画素に割り当てるときの不連続コストであり、隣接画素間での不連続性に対するペナルティ関数として用意される。

$$m_{pq}(f_q) = \min_{f_p} \left( D_p(f_p) + V(f_p, f_q) + \sum_{s \in N(p) \setminus q} m_{sp}(f_p) \right) \quad (2)$$

なお、厳密に言えば、画像を対象とする場合にはグラフが格子状になり、ループ構造をもつため、メッセージパッシングの結果は計算される画素の順番に依存してしまう。そのため、通常は、チェッカーボードのように 1 画素おきにメッセージを更新したり、上下左右の方向別に更新したりする。後述する本研究の実験では後者を採用している。

### 3.2 二分探索型 BP による領域分割法

提案手法では、特徴空間を階層的にポロノイ分割することで画素値分布のクラスタリングを行うとともに、ポロノイ分割のもとで各画素の事後確率を最大にするクラスラベルを二分探索型 BP を用いて階層的に算出する。提案手法は、従来手法に比べ、特徴空間でのクラスタリングを単純化し、画像空間での最適化に重きを置くことで安定化を目指すとともに、ユークリッド距離でのクラス帰属度の計算や二分探索型 BP により処理の高速化を狙ったものとなっている。また、階層的に処理することで初期値への依存性を軽減している。

二分探索型 BP は、もともと、一般的な BP の計算速度を改善するために提案された手法

である。一般的な BP の計算量は、画素数を  $n$ 、クラスラベルの数を  $k$ 、メッセージ更新回数を  $t$  とすると  $O(nk^2t)$  となるが、二分探索型 BP では  $O(nt \log_2 k)$  となり大幅な高速化が実現できることが特長である。今回の提案手法では、この高速化に加え、階層的な処理を行うためにも二分探索型 BP が有効に寄与している。

二分探索型 BP では、ラベル候補を規則的に順序づけ、まず二等分する。その 2 つの組にそれぞれ代表ラベルを設け、2 つの代表ラベルでのメッセージパッシングを行う。その後、各画素において帰属度の高い代表ラベルの組を選択し、その組のラベル候補をさらに二等分する。以後、メッセージパッシングと二等分を繰り返すことで最適解を得るものである。

二分探索型 BP はラベル候補が順序づけられることを前提としており、ラベル候補に割り当てられる値は 1 次元のスカラーとなっている。これを領域分割問題に合うように拡張して適用する。なお、以下の説明では、特徴空間として RGB 色空間を用い、色特徴に基づいた領域分割を行うものとする。以下に、提案手法における処理手順を示す。

- (1) 特徴空間に 2 つのクラス中心点（代表ラベル）を与え、ポロノイ分割により特徴空間を二分分割する。
- (2) 画像空間において、代表ラベルでメッセージパッシングを行う。
- (3) 画素ごとに、帰属度の高い代表ラベルを選択する。
- (4) 代表ラベルごとに、それを選択した画素集合における観測値（RGB 値）の平均値を求め、その平均値の位置にクラス中心点をシフトし、ポロノイ分割を更新する。
- (5) シフト量がなくなるまで (2) から (4) を繰り返し実行する。
- (6) シフト量がなくなれば、すでにポロノイ分割されている部分空間のそれぞれにおいて、部分空間内にクラス中心点を増やし、さらに二分分割する。
- (7) 所定の分割数に達するまで、(2) から (6) を繰り返し実行する。

なお、提案手法におけるメッセージパッシングでは、式 (3)(4) に示すデータコスト  $D_q(\vec{f}_q)$  と不連続コスト  $V(\vec{f}_p, \vec{f}_q)$  を用いる。ただし、 $\vec{f}_q$  は画素  $q$  におけるラベルの RGB 値であり、 $\vec{g}_q$  は画素  $q$  で観測された RGB 値である。また  $\alpha, \beta$  は任意の重み係数である。

$$D_q(\vec{f}_q) = \alpha |\vec{f}_q - \vec{g}_q| \quad (3)$$

$$V(\vec{f}_p, \vec{f}_q) = \beta |\vec{f}_p - \vec{f}_q| \quad (4)$$

#### 4. 実験と考察

提案手法の有効性を確認するため実験を行った．実験画像には Berkeley Segmentation Dataset<sup>15)</sup>，Caltech Faces 1999 Database<sup>16)</sup>，および人工的に作成した画像を用いた．それぞれ RGB24bit カラー画像に変換して処理を行っている．

二分探索型 BP における最終的なラベル数は 8 とし，コスト関数における重み係数は標準値として  $\alpha = 1.0$ ， $\beta = 2.0$  を与えた．また，メッセージ  $m_{pq}$  の初期値は 0 とし，1 回のメッセージパッシングにおけるメッセージ更新回数は，右，左，下，上の順にそれぞれ 1 回ずつとした．

##### 4.1 処理例

まず，図 3，図 4 に領域分割結果を示す．これらの図には，比較のため，GMM のもとで標準的な BP を行う従来手法<sup>10)</sup> の結果も示している．ただし，従来手法のプログラムも自前で実装したものである．従来手法の BP においては，極力，提案手法に合わせたメッセージパッシング方法などを採用している．また，GMM におけるクラス数は 8 とし，GMM の初期値は k-means 法にてあらかじめ決定した．なお，図中の (b) および (d) は，分割された各領域を代表ラベルの値で色づけしたものである．また，(c) および (e) は，それぞれ (b) および (d) の結果を分かりやすくするため，領域ごとに任意の色をつけたものとなっている．従来手法に比べ，提案手法の方が主観に合う結果となっていると考えられる．

##### 4.2 分割性能

次に，図 4(9)-(a) を対象として，重み係数  $\alpha$  および  $\beta$  の値を 0.1 から 3.0 の範囲で変更した場合の分割性能を図 5 に示す．図中のグラフは，分割後の領域数，入力画像からの残差，処理時間からなる軸で構成されている．ここでは，過剰分割にならず，かつ，入力画像との残差が少ない方が望ましい結果である．グラフをみると，以下のことが分かる．

- 提案手法の方が過剰分割にならず，かつ，入力画像との残差が少ない結果となっている．
- 提案手法の結果は，重み係数を変更した場合のばらつきが少なく，安定しているといえる．複数の重み係数による結果から最適な係数を推定するといった上位の枠組みを導入することも可能であると思われる．
- 提案手法の計算時間は，従来手法と比較して高速である．

##### 4.3 処理時間

最後に処理時間について述べる．表 1 に，Intel Core2 Quad 2.4GHz CPU の汎用 PC で，それぞれの画像を処理した場合の処理時間を示す．

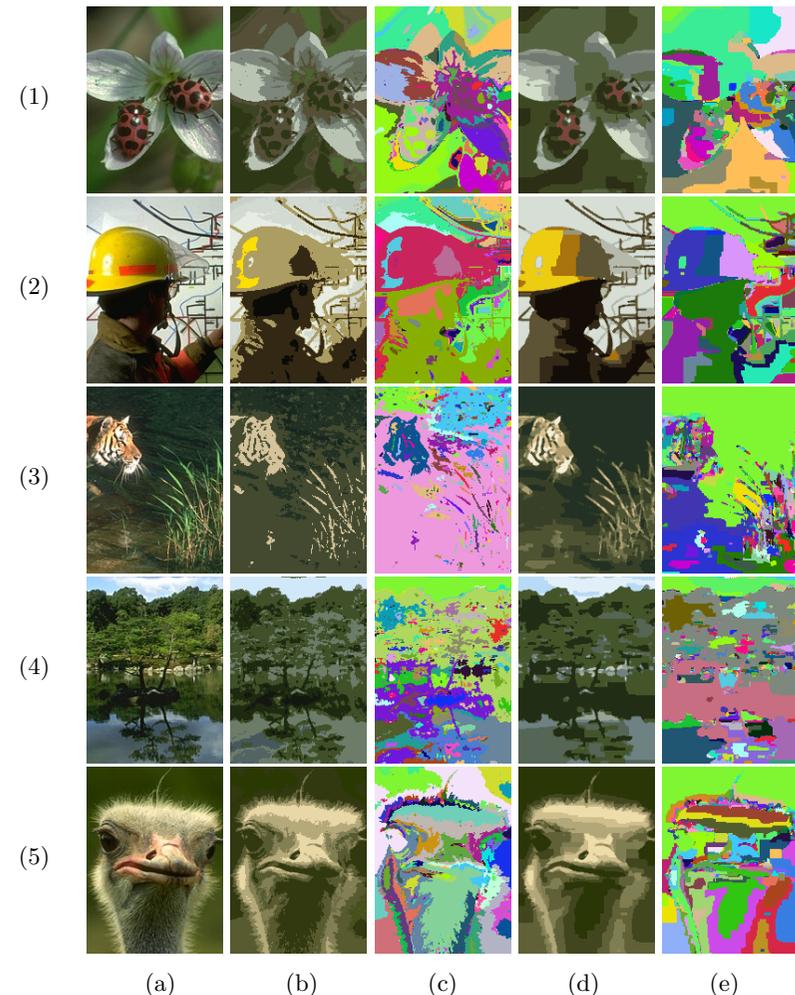


図 3 領域分割例 1 (左から (a) 入力画像，(b) 従来手法の結果，(c)(b) の結果を配色しなおしたもの，(d) 提案手法の結果，(e)(d) の結果を配色しなおしたもの)

Fig.3 Example of image segmentation results.

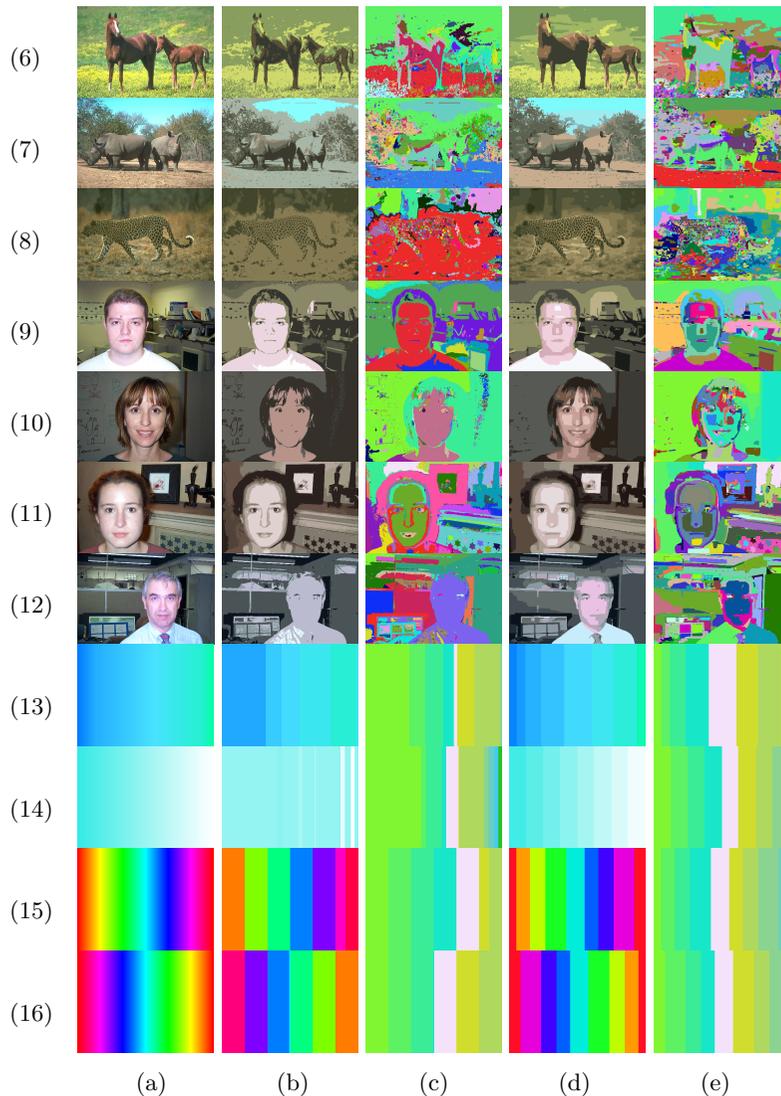


図4 領域分割例2 (左から (a) 入力画像, (b) 従来手法の結果, (c) (b) の結果を配色しなおしたもので, (d) 提案手法の結果, (e) (d) の結果を配色しなおしたもので)

Fig.4 Example of image segmentation results.

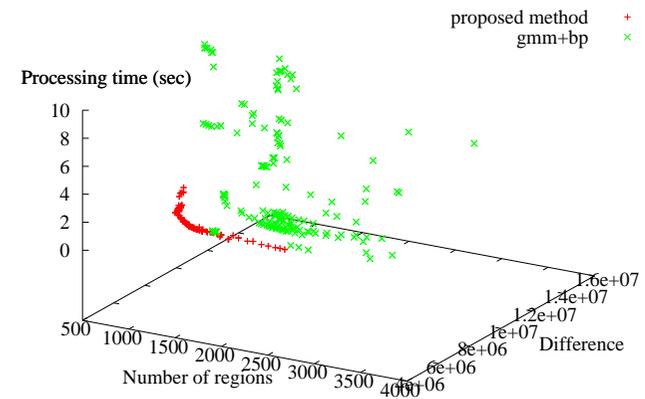


図5 分割性能

Fig.5 Segmentation performance.

表1 処理時間

Table 1 Processing time

No.	resolution (pixel)	gmm+bp (sec)	proposed (sec)
1	148 × 200	0.52	0.24
2	148 × 200	1.13	0.33
3	148 × 200	0.64	0.30
4	148 × 200	0.70	0.32
5	148 × 200	0.91	0.26
6	480 × 320	4.33	2.14
7	480 × 320	3.59	1.12
8	480 × 320	3.14	1.85
9	448 × 296	3.94	1.03
10	448 × 296	4.71	1.04
11	448 × 296	3.32	1.20
12	448 × 296	4.37	1.37
13	320 × 240	39.61	0.55
14	320 × 240	65.95	0.41
15	320 × 240	23.69	0.59
16	320 × 240	58.82	0.59

## 5. おわりに

本研究では、二分探索型 BP を拡張・適用した領域分割手法について述べた。提案手法では、特徴空間を階層的にポロノイ分割することで画素値分布のクラスタリングを行うとともに、ポロノイ分割のもとで各画素の事後確率を最大にするクラスラベルを二分探索型 BP を用いて階層的に算出する。提案手法は、従来手法に比べ、特徴空間でのクラスタリングを単純化し、画像空間での最適化に重きを置くことで安定化を目指すとともに、ユークリッド距離でのクラス帰属度の計算や二分探索型 BP により処理の高速化を狙ったものとなっている。また、階層的に処理することで初期値への依存性を軽減している。

今後は、領域分割結果をもとに物体認識あるいは領域のクラス分類を行う上位の認識手法についても、高精度・高速化を図りたいと考えている。

謝辞 プログラムの一部をご提供いただきました同研究室の浦田賢人さんに感謝致します。

## 参 考 文 献

- 1) 柳井啓司：一般物体認識の現状と今後，情報処理学会論文誌コンピュータビジョンとイメージメディア， Vol.48, No.SIG16(CVIM19), pp.1-24 (2007).
- 2) Csurka, G., C.R.Dance, L.F., Willamowski, J. and Bray, C.: Visual categorization with bags of keypoints, *Proc. ECCV Workshop on Statistical Learning in Computer Vision*, pp.1-22 (2004).
- 3) Tenenbaum, J.M. and Barrow, H.G.: Experiments in Interpretation Guided Segmentation, *Artificial Intelligence*, Vol.8, pp.241-274 (1977).
- 4) Barnard, K. and Forsyth, D.: Learning the Semantics of Words and Pictures, *IEEE Int. Conf. on Computer Vision*, pp.408-415 (2001).
- 5) 長橋知行，藤吉弘巨，金出武雄：領域分割に基づく SIFT 特徴を用いた物体認識，電気学会 システム・制御研究会， No.SC-07-8, pp.39-44 (2007).
- 6) He, X., Zemel, R.S. and Carreira-Perpinan, M.A.: Multiscale Conditional random fields for image labeling, *Proc. IEEE Computer Vision and Pattern Recognition*, pp.695-702 (2004).
- 7) 奥村健志，滝口哲也，有木康雄：大域的特徴として BoF を導入した CRF による一般物体認識，画像の認識・理解シンポジウム， No.OS4-2, pp.95-102 (2009).
- 8) Zhang, J.: The mean field theory in EM procedures for Markov random fields, *IEEE Trans. on Signal Processing*, Vol.40, pp.2570-2583 (1992).
- 9) 阿南泰三，博幸，斎藤恒雄：階層化 EM アルゴリズムを用いたテクスチャー・セグメンテーション，情報処理学会論文誌， Vol.36, No.3, pp.601-613 (1995).
- 10) Chen, F., Tanaka, K. and Horiguchi, T.: Image Segmentation Based on Bethe Ap-

- proximation for Gaussian Mixture Model, *Interdisciplinary Information Sciences*, Vol.11, No.1, pp.17-29 (2005).
- 11) Comaniciu, D. and Meer, P.: Mean Shift: A Robust Approach Toward Feature Space Analysis, *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol.24, No.5, pp.603-619 (2002).
  - 12) Shi, J. and Malik, J.: Normalized cuts and image segmentation, *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol.22, No.8, pp.888-905 (2000).
  - 13) Carson, C., Belongie, S., Greenspan, H. and Malik, J.: Blobworld: Image Segmentation Using Expectation-Maximization and Its Application to Image Querying, *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol.24, No.8, pp.1026-1038 (2002).
  - 14) 浦田賢人，Sandy, H.，和田俊和：ラベル数の削減による Belief Propagation の高速化に関する研究，画像の認識・理解シンポジウム， No.IS2-18, pp.968-975 (2009).
  - 15) Martin, D., Fowlkes, C., Tal, D. and Malik, J.: A Database of Human Segmented Natural Images and Its Application to Evaluating Segmentation Algorithms and Measuring Ecological Statistics, *IEEE Int. Conf. on Computer Vision*, Vol.2, pp.416-423 (2001).
  - 16) California Institute of Technology: Faces 1999 Database. <http://www.vision.caltech.edu/archive.html>.