

Grabcut 適用結果の形状類似性に着目した 対話不要の画像領域分割

松尾賢治[†]

Grabcut は、画像内で前景オブジェクトが含まれる矩形領域を手動で大まかに指定し、指定された外側の領域の色分布から学習した背景と推定される色を指定された内側の領域から取り除くことによって、前景オブジェクトの正確な抽出を実現する。一方、手動による領域指定の操作を必要とするため、人手を要さず自動的に画像の領域分割を実施できるものではなかった。そこで本論文では、まず JSEG により画像を色の類似する小領域単位に過剰分割し、次に各小領域の色分布を学習して Grabcut を適用した際のそれぞれの領域分割結果の形状が類似している小領域を統合することによって、手動による領域指定の操作を必要としない領域分割方法を提案する。実験により提案法の有効性を評価した。

A non-interactive image segmentation method based on the shape similarity of the Grabcut

Kenji Matsuo[†]

Grabcut is a conventional image segmentation method that can extract correct foreground objects with human interactions. However, the Grabcut needs manual operations and cannot achieve automatic image segmentation. This paper shows a non-interactive segmentation method that can automatically extract objects from image. The first step of the proposed method is over segmentation which divides an image into many small regions based on color similarity. The second step is region integration which merges the segmented regions based on the similarity of color distribution. The comparison unit in the integration step is not pixel but the expansion region which is grown from a segmented region by the Grabcut. Therefore, the region merge operation is stable and the proposed method achieves the high performance. Simulation results show some advantages of the proposed method.

1. まえがき

一般に画像処理の実行には非常に多くの計算量を要する。近年、高性能計算機の出現により、これまで実現が難しかった高度な画像処理が可能になった。例えば、前処理で領域分割を適用することにより、画像をオブジェクト単位でハンドリングできるようになる。すなわち、符号化・伝送の対象を画像全体ではなく、オブジェクトの内容に基づいて符号量割り当ての最適化や優先度順の伝送等を実現できる。また画像認識においても、領域分割を行わない場合は画像全体を認識の対象として扱うため、正しく認識できるのは画像の中央に大きく写る単体オブジェクトに制限されるが、前処理で領域分割を適用することによって、画像内に複数オブジェクトが存在する場合でも、オブジェクト別に認識できるようになり、画像内の詳細な解析が実現可能となる。

ここで領域分割とは、本来は画像内にある家や人、また空や背景等、意味のあるまとまりを見つけ出し、オブジェクトとして抽出する処理である。しかしながら、画像中に何が表示されているか全く分からない状態でこのような分割を得るのは難易度の高い問題である。そのため、このように事前知識の無い状態では、領域分割を色や明るさ等の画像信号の性質が似ている画素同士を同じ領域へとまとめる問題に置き換え、従来から多くの取り組みがなされてきた[1]-[10]。画像信号の類似性に基づいた領域分割がある程度有効に作用するのは、画像中で同じオブジェクトを構成する画素は類似した性質を持っていることが多いという経験則に基づいている。この経験則があてはまらない場合も確かに存在し、この場合は領域分割の精度を上げるために何らかの事前知識が必要であるが、本論文では従来の取り組み同様、事前知識の無い場合を想定し、主に色と明るさの類似性に基づいた領域分割の実現について取り組むものとする。

これまでに、領域分割をエネルギー最小化問題として解く方法がいくつか提案されている。代表的なものに、動的輪郭モデル Snake[1]や Level Sets[2]、Graph Cuts[3][4]等が挙げられる。Snake や Level Sets はオブジェクトの境界線に関するエネルギー関数を作成し、エネルギー関数が小さくなるように境界線を変化させてオブジェクトを抽出する方法である。これらの方法で求められた解は局所解であることにに対し、Graph Cuts は画像内の各領域からエネルギー関数を定義して、これを最小とする大局的な解が求められるため、強力な領域分割方法として近年特に盛んに研究されている[5]-[7]。

Graph Cuts を改良した新しい領域分割方法として、Boykov らにより Interactive Graph Cuts[5][6]が提案されている。Interactive Graph Cuts では、あらかじめ前景および背景の典型的な部分をユーザが手動で指定し、それぞれの領域から前景および背景の色分布を学習する。色と明るさの画像信号だけでなく、ユーザとの対話の末学習により得られた色分布モデルを加味したエネルギー関数を作成することで、より信頼性の高い領域分割を実現する。また、Interactive Graph Cuts による領域分割結果から得られた前景と背景領域の色分布を再度学習し直し、繰り返し領域分割を行うことによって精度

を向上させた Grabcut[7]も提案されている。しかしながら、Interactive Graph Cuts および Grabcut は非常に強力な前景背景の分離を実現するものの、あらかじめ正解となる背景領域を手動で指定する操作が必要であり、ユーザとの対話無く自動的な領域分割を実現する方法ではなかった。なおこれらの方法では、前景背景だけでなく二つ以上の領域に分割する方法についても特に述べられていない。

そこで本論文では、手動による領域指定の操作を必要とせず、ユーザとの対話を必要としない領域分割方法を提案する。提案法では、まず JSEG 等により画像を複数的小領域へと過剰分割して、各小領域の色分布を学習し、モデル化する。また学習した色分布モデルを用いて Grabcut を適用し、画像の二領域分割を行う。次に Grabcut による二領域分割結果間がお互いの小領域を相互に含んでいる画像は、元の小領域同士が同じ性質を持っていると判定して統合する。以上の過剰分割と領域統合によって、画像内に含まれる複数オブジェクトの分離を自動的に実現する。

2. グラフ切断エネルギー最小化による領域分割

従来法である Interactive Graph Cuts および Grabcut は、以下の手順により、前景と背景の二領域への分割を実現する。

2.1 Interactive Graph Cuts

画像 P 内の $|P|$ 個の画素を $p \in P$ とし、最終的に求めたい領域情報を $X = \{X_1, X_2, \dots, X_p, \dots, X_{|P|}\}$ とする。各 X_p は対応する画素が前景または背景のいずれに属するかを示すラベル \mathcal{F} または \mathcal{B} が与えられる。Graph Cuts は、 X を変数としてエネルギー関数 $E(X)$ を次式で定義し、エネルギーを最小とする X を求める方法である。

$$E(X) = \sum_{p \in P} D_p(X_p) + \sum_{\{p, q\} \in N} S_{pq}(X_p, X_q) \quad (1)$$

ここで第一項は、着目する画素 p 自体の有する明るさや色等の信号値とその画素に割り当てられるラベル X_p に依存して算出されるエネルギーであり、データ項と呼ぶ。データ項は与えられた画素 p の前景らしさまたは背景らしさを表す値を返す。第二項は、着目する画素 p だけでなく近傍の画素 $q \in N$ とそれぞれに割り当てられるラベル X_p および X_q から算出されるエネルギーであり、平滑化項と呼ぶ。平滑化項は着目する画素がグラフの切断点として適当か否かを表す値を、近傍の画素の信号値との差異から算出して返す。近傍の画素 q は一般に画素 p の 4 方向または 8 方向の隣接画素を用いる。

Interactive Graph Cuts では、前景と背景に属する典型的な部分を、マウスで太い線を描くと同じ要領で、ユーザが大まかに指定する。前景として指定された部分領域から、そこに属する画素の明るさ I の相対頻度を測定し、前景のヒストグラム $\theta(I, \mathcal{F})$ を作成する。また背景についても同様に背景のヒストグラム $\theta(I, \mathcal{B})$ を作成し、前景および背景それぞれの明るさの分布をモデル化する。このときデータ項は、

$$D_p(X_p) = -\log \theta[I_p, X_p] \quad (2)$$

となる。このモデルに従うと、画素 p の明るさを I_p とするとき、例えば背景と指定された部分に明るさ I_p がより多く含まれていた場合は、背景らしい明るさを持つ画素と考えられるため、 $X_p = \mathcal{B}$ と判定することでデータ項の値が小さくなる。

一方、平滑化項は次式で定義する。

$$S_{pq}(X_p, X_q) = \begin{cases} 0 & , X_p = X_q \\ \frac{\lambda \exp(-\kappa \{I_p - I_q\}^2)}{|p, q|} & , X_p \neq X_q \end{cases} \quad (3)$$

ただし、パラメータ λ および κ はバランス調整のための定数、分母は画素 p, q 間の距離である。平滑化項の値は、明るさが似ている画素間では大きく、逆に明るさの差が大きい領域間の境界では小さくなる。

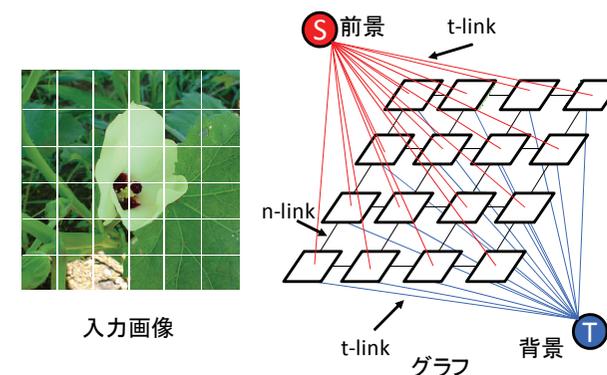


図 1. Graph Cuts のグラフの構成

式(2)および(3)で定義したエネルギーを使って、式(1)のエネルギー関数 $E(X)$ を最小化することによって、 $E(X)$ を大域的に最小とする領域情報 X を決定し、前景および背景の二領域への分割を実現する。実際に Graph Cut では、図 1 に例示するグラフを作成し、各辺を切断するときのコストをエネルギー関数と対応付けることによって、無数に存在する切断の組合せの中からグラフを最も小さいエネルギーとなる切断を決定する。グラフは、各画素に対応するノードと、特別に追加された頂点に対応するターミナル S および T から構成される。グラフでは、ノードとターミナル間は t-link と呼ばれる辺で、またノード間は n-link と呼ばれる辺で相互に接続される。t-link の切断コスト

トはデータ項のエネルギーに、また n -link の切断コストは平滑化項のエネルギーに対応する。ターミナル S 側を前景、 T 側を背景と考えると、例えば前景と判断されラベル σ が付与された画素は、 T 側に接続された t -link の辺のコストが切断により加算される。また、 n -link では近傍の画素間で付与されるラベルが異なる場合、それらの画素間を接続する n -link の辺を切断するコストが切断により加算される。

2.2 Grabcut

Grabcut は、分割の精度をさらに向上させるため、グレイスケールではなくカラー画像を用いる点異なる。データ項における前景と背景のモデル化には、明るさのヒストグラムではなく、前景と背景の色の分布を GMM (Gaussian Mixture Model: 混合正規分布モデル) でモデル化し、モデルの適応精度を高める。また、Graph Cuts による領域分割の結果から、前景と背景の色分布を再度学習して、繰り返して Graph Cuts を適用することによって、最終的な領域分割の精度を高める。

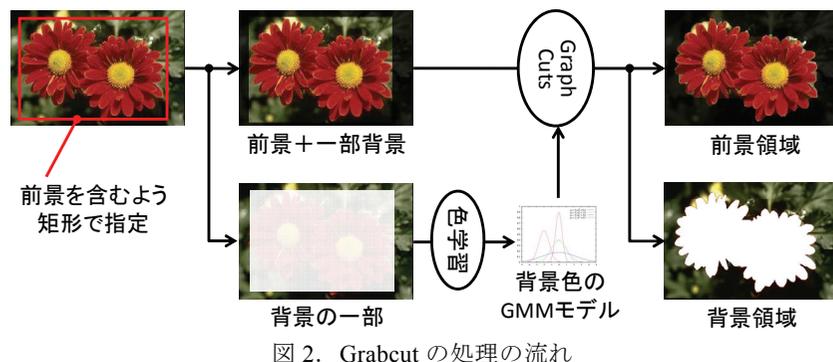


図 2. Grabcut の処理の流れ

また Grabcut では、図 2 に示す通り、最初にユーザーに求められる領域指定は前景だけで十分であり、Interactive Graph Cuts のように前景と背景の両方を指定する必要はない。すなわち、画像から抽出したいオブジェクトが含まれる大まかな領域を、ユーザーがマウス操作により矩形で指定することにより、最終的に画像を前景と背景との 2 つの領域に分離する。この際、ユーザーによって指定された矩形領域の内側には前景と背景の一部とが混在する領域と考えることができ、逆に矩形の外側の領域には背景だけが含まれると考えられる。したがって、ユーザーによって指定された矩形の外側の領域に含まれる画素の色の分布を GMM でモデル化する。生成されたモデルを背景色の分布として用いて、矩形領域の内側の混在領域に含まれる個々の画素についてその色から背景らしさを示す尤度を算出する。背景色として高い尤度を示した画素を混在領域から取り除くことによって、画像から前景を高精度に抽出する。

2.3 課題

しかしながら、従来法である Grabcut により画像の領域分割を実現するためには、ユーザがあらかじめ大まかな前景領域を手動で指定する手間が必要であった。このため、人手を要さず自動的に画像の領域分割を実施できるものではなかった。また Grabcut では、1 枚の画像を前景と背景との二領域に分割することを目的としていたため、3 つ以上の複数の領域に分割することを対象とした方法ではない。

3. 提案法

従来法の課題を解決するため、人手による領域指定を必要としない自動的な領域分割を実現し、さらに 3 つ以上の領域への分割を実現する領域分割方法を提案する。

3.1 過剰分割と領域統合による領域分割

先の 1. で述べた通り、領域分割を色や明るさ等の画像信号の性質が似ている画素同士が同じ領域となるようにまとめる問題に限定して考える。この場合、例えば図 3 に示す花の画像は、手動では多くの被験者が同図正解画像として示した通り、赤い花びら、緑の葉を含めた背景、黄色い花冠の 3 つのオブジェクトに分割することを確認している。この結果に着目すると、3 つのオブジェクトの境界は色の差異が大きい部分であることが再確認できる。このことから、画像内で色の差異が大きい部分を境界線の候補として挙げ、その中からオブジェクトの境界として相応しい境界線だけを残すことができれば、人間の判断と一致した領域分割結果が得られると考えた。



図 3. 提案法の概要

具体的には、同図に示す通り、まず画像内を似た色の画素が連続する小領域にまとめて多少過剰気味に分割する。得られた小領域間の境界線の中には、オブジェクトの境界となる真の境界線と同時に、適当でない境界線も含まれる。次に分割された小領域の中から色分布の特徴が類似した小領域同士を統合することによって、候補に挙げた境界線の中から適当でない境界線を削除する。画素単位ではなく小領域単位で性質の類似性を判定することにより安定した統合が期待できる。以上の過剰分割と領域統合により、色に基づいた領域分割を実現する。境界線に着目して考えると提案法では、過剰分割により正解となる境界線を 100% 含みリコールが 1.0 に近い状態で境界線の候補を得て、そこから領域統合によりプレジジョンを向上させることを狙っている。

3.2 処理の流れ

提案法の処理の流れを図4に基づいて説明する。

(1) 過剰分割処理

入力画像を色の類似する小領域単位に過剰分割する。

(2) 領域拡大処理

小領域の色分布を GMM により学習する。学習した色分布モデルを用いて入力画像に Grabcut を適用し、元の小領域と類似した性質を持つ拡大領域を得る。

(3) 領域統合処理

お互いの拡大領域により包含される小領域同士は、類似した色分布を持っていると判定して統合する。

(4) 収束判定処理

新たに領域が統合されれば、領域拡大以降の処理を繰り返す。逆に、統合された領域が一つもなければ、収束していると判定し、最終的な領域分割結果を出力する。

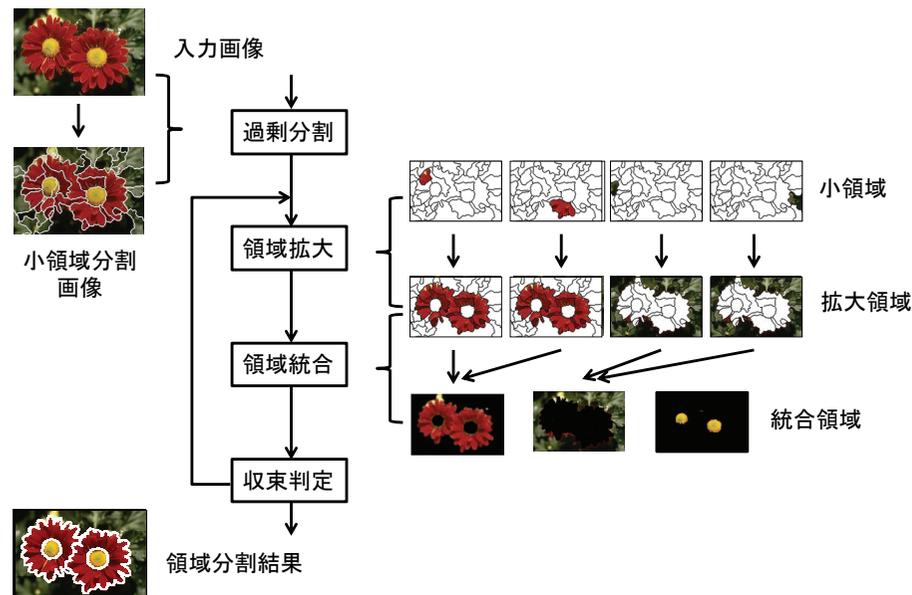


図4. 提案法の処理フロー

3.3 領域拡大による領域統合の安定化

前節2.で述べた通り、Grabcut は画像の前景と背景とを高精度に分割する方法であった。一方、前景側ではなく背景側に着目して考えると、ユーザが指定した背景領域の

色分布を GMM で学習し、学習されたモデルを使って画像内の別の場所で同じ色分布に従う画素で背景と考えられる領域を新たに見つけ出す、つまり図5に示す通り、学習領域を拡大させる方法であるとも捉える事ができる。

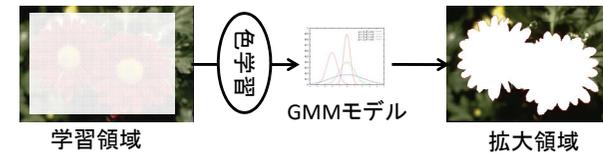


図5. Grabcut による領域の拡大

領域統合処理は、画素単位ではなく小領域単位で性質の似た小領域同士を統合することによって、より安定した統合が実現できると考えられる。提案法ではさらに安定的な小領域の統合の実現を目指し、領域統合処理の前段で Grabcut による領域拡大処理を追加する。すなわち、Grabcut により小領域 a を拡大させた拡大領域 a が、別の小領域 b を包含し、反対に小領域 b の拡大領域 b も小領域 a を包含すれば、元の小領域 a と小領域 b は相互に似た色分布の特徴を持ち、同じオブジェクトに含まれると考え、小領域 a と小領域 b を統合する。

なお、前段の過剰分割処理は色に基づく手段であれば、JSEG[8]や Normalized Cut[9]による Superpixel[10]表現等により実現できる。ある程度まとまった画素数を含む小領域が得られれば、GMM による色分布の学習と Grabcut による領域拡大が可能となる。この際、正解となる境界線を 100%含みリコールが 1.0 に近い状態で境界線の候補が得られることが望ましい。これは、後段の統合処理は適切な境界線を前段の過剰分割の際に得られた境界線候補の中から選ぶことになり、過剰分割の時点で正しい境界線が候補に含まれていなければ、良好な領域分割結果は期待できないためである。

3.4 具体例

図6に示す通り、42個の小領域に過剰分割された例を使って、領域統合の詳細について説明する。まず42個の小領域のそれぞれについて Grabcut による領域拡大を行う。例えば、小領域 10, 11 および 22 を領域拡大した結果を図7に示す。小領域 10 の拡大領域 10 には小領域 11 が包含され、逆に小領域 11 の拡大領域 11 には小領域 10 が包含される。すなわち、小領域 $10 \subseteq$ 拡大領域 11 かつ小領域 $11 \subseteq$ 拡大領域 10 である。このように小領域 10 および 11 はそれぞれの拡大領域に相互に包含されるため、色分布の性質が類似した領域と判定して統合する。一方、小領域 22 の拡大領域 22 に小領域 10 および 11 は包含されるが、逆は成り立たない。つまり、小領域 10 および 11 のそれぞれの拡大領域 10 および 11 のいずれにも小領域 22 は包含されない。この場合、小領域 22 は小領域 10 および 11 とは色分布の性質が異なる領域と判定する。



図 6. 過剰分割された小領域

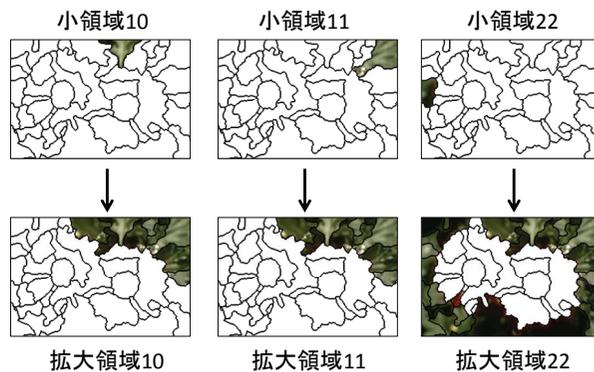


図 7. 小領域とその拡大領域

上記の判定を全 42 個の小領域に対して行った結果、表 1 に示す通り、5 つのグループに統合される。この際、小領域 4, 6, 8, 9, 23, 24, 26, 28, 29, 31, 33, 34, 35, 36 の 14 個は所属の無い状態に残る。これらの 14 個の所属の無い小領域は次の手順でいずれかのグループに仕分ける。各グループに属する小領域を拡大し、それぞれの拡大領域に所属の無い小領域が包含される回数をカウントして、最もカウント数の多いグループに所属の無い小領域を帰属させる。例えば、所属の無い小領域 8 はグループ A に属する 7 つの小領域 10, 11, 12, 18, 19, 37, 38 の拡大領域のそれぞれに含まれ、他はグループ B に属する 4 つの小領域 5, 7, 22, 25 の拡大領域のそれぞれに含まれる。この場合、小領域 8 は含まれた回数の多いグループ A に所属させる。仕分け手順でも残った小領域は孤立領域として扱い、別グループを追加する。表 1 の括弧内に結果を追記する。

以上の統合処理の結果、図 8 に示す通り、5 つの領域に統合される。領域統合により 42 から 5 つへと領域数が増えたため、統合された 5 つの領域を新たな小領域として、領域拡大処理に戻って領域統合を再度適用する。なお、この画像は次の統合処

理によって人間同様正しく 3 つのオブジェクトを示す領域に統合される。

また、Grabcut による領域拡大は画素単位で行われる。そのため、あらかじめ過剰分割された小領域の中に、拡大した際の画素が $\alpha\%$ 以上含まれれば、その小領域は全て拡大領域として塗り潰す処理を追加し、画素単位ではなく小領域領域単位で包含状況を判定する。この領域判定パラメータ α の値によって、統合処理の結果が変化する。

表 1. 統合されたグループ

グループ	小領域番号
A	10, 11, 12, 18, 19, 37, 38 (8, 9, 36)
B	5, 7, 22, 25, 27 (23, 24, 26, 28, 29, 31, 34, 41)
C	2, 13, 14, 15, 16, 20, 21 (4, 6, 17, 33, 35, 40, 42)
D	1, 30 (39)
E	3, 32

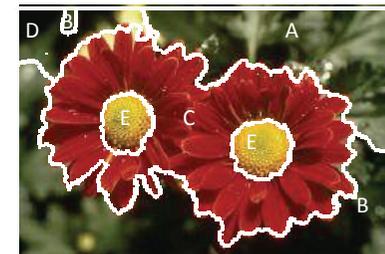


図 8. 1 回目の領域統合結果

4. 実験

提案法の有効性を確認するため、実験により従来法との比較評価を行った。

4.1 条件

実験には Berkley Segmentation Dataset and Test[11]に含まれる自然画像を使用する。領域分割方法の性能を測るための標準的なデータセットであり、様々な領域分割方法がベンチマークテストを行っている。また、その結果は Web で公開されている[12]。データセットに含まれるそれぞれの画像について 5 名以上の被験者による手動の領域分割情報が格納されており、これを領域分割の Grand Truth データとして使用できる。解像度は縦 321 画素、横 481 画素または縦 481 画素、横 321 画素の二種類である。ここで提案法は色の類似性に基づく領域分割方法であり、オブジェクト形状等の事前知識を利用しないため、同系色の中にオブジェクトが混在している場合、例えばサバン

ナの草原と同化した雌ライオンが写っている画像等の高精度な分割は期待できない。本実験は Berkley の画像の中から、画像内のオブジェクトがそれぞれ特徴的な色を持っている画像 30 枚を実験画像として選択した、特定の条件下での結果である。

また、提案法の過剰分割処理には JSEG[8]を使用した。過剰分割結果からの提案法の領域統合の改善性能を確認するため、JSEG の精度との比較を行う。JSEG は色の類似する領域を画素単位でまとめる領域分割方法であり、入力画像を減色し、局所領域における色のばらつきに基づいてオブジェクトの境界らしさを示す領域境界尤度を算出して、同じ色の連続する小領域とその境界を求める領域分割方法である。JSEG の実装には提案者によって公開されている JSEG ソフトウェア[13]を使用し、デフォルトパラメータを使用して実験した。さらに、Berkley Segmentation Dataset and Test と同時に同機関にて公開されている Berkley Segmentation Engine: BSE[14]-[16]による領域分割方法との比較を合わせて行う。BSE は、明るさ、色だけでなく局所領域におけるテクスチャの類似性を考慮して、自然画像の境界を検出する方法である。

4.2 精度

領域分割の精度を二通りで評価する。1 つ目はオブジェクト間の境界線の再現性に関する評価であり、上述の Berkley Segmentation Dataset and Test で行われているベンチマークテストである。2 つ目は主にクラスタリング方法の性能評価に用いられる指標による領域分割結果間の一致性に関する評価である。

(1) 境界線に基づく精度評価

被験者の手動により描かれた Grand Truth となる境界線を領域分割方法による結果がどの程度忠実に再現できたか測定し、Recall と Precision の曲線を求める。テスト画像 1 枚につき 5 名以上被験者による Grand Truth データが存在するため平均値で評価する。提案法、JSEG、BSE の Recall-Precision 特性曲線を図 9 に示す。提案法では JSEG より R 値=1.0 となるような過剰分割を期待したが、実際には R 値=0.758 付近の低い値を取った。この値から提案法の統合領域によって P 値=0.651 の改善を狙う。提案法の領域判定パラメータ α の値により結果が変化する。グラフは、 $\alpha=25, 50, 75, 100\%$ の結果を示した。 $\alpha=25\%$ および 50% のとき BSE よりも良好な P 値=0.862 および 0.825 を示した。しかしながら、R 値=0.501 および 0.588 であったため、表 2 に示す通り F 値で比較すると、本実験で用いたパラメータでは BSE を超える精度を提案法では得られなかった。図 10 に示した通り、Grand Truth は被験者の主観により異なる。そのため、境界線で比較した場合、オブジェクトをより細かく分割する被験者 2 および 3 の作成した Grand Truth データは、色の類似性に基づいた領域分割法では分割できない単位まで分割されているため、提案法では Recall が大きく下がる傾向にあることを確認した。

(2) 領域分割結果間の一致性による評価

必ずしも全ての被験者が全く同一の Grand Truth に向かって領域を分割する訳ではなく、領域分割の詳細度合いは被験者間で異なる。したがって Martin らは、被験者に

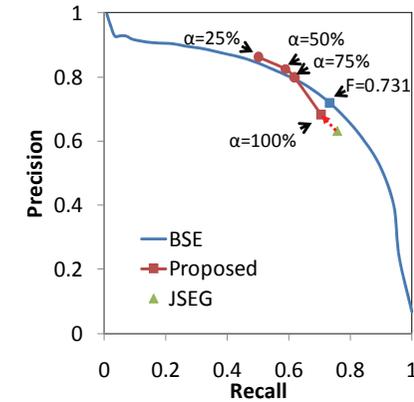


図 9. Recall-Precision 特性

表 2. 方式別の F 値比較

	R 値	P 値	F 値	
JSEG	0.758	0.631	0.688	
提案法	$\alpha=25\%$	0.501	0.862	0.634
	$\alpha=50\%$	0.588	0.825	0.687
	$\alpha=75\%$	0.618	0.799	0.697
	$\alpha=100\%$	0.705	0.682	0.693
BSE (F 値最大)	0.733	0.720	0.726	



図 10. 被験者による Grand Truth 詳細化の違い

よる分割の細かさの違いに依存しない、一貫性のある評価尺度を提案した[17]. 大局的な一貫性誤差 global consistency error: GCE と局所的な一貫性誤差 local consistency error: LCE である。これらの尺度は比較する 2 つの分割結果が近似的に等しい分割数

であるとき、初めて意味を持つ。一方、Rand はクラスタリングによる2つの分類結果の一致性を測る尺度である Rand Index を提案した[18]。Rand Index は分割数の異なる分割結果間の一致性の評価を可能にしたが、分割の詳細化には対応せず影響を受ける。これらを受けて、Unnikrishnan らは分割の詳細化に対応し、かつ任意の分割数を許容する尺度として、Probabilistic Rand Index を提案した[19]。Probabilistic Rand Index は0から1の値を取るが出力される期待値は比較的大きいことから、その問題を解決した Adjusted Rand Index[20]による評価をここでは行う。

提案法の領域判定パラメータを $\alpha=50\%$ とし、Adjusted Probabilistic Rand Index を測定した。代表的な4枚と画像30枚の平均について、JSEG からの改善度を括弧内に合わせて結果を表3に示した。提案法はJSEGより高い値を示すことが確認できる。また、対応する画像1から4の領域分割結果を図11に例示した。上からある被験者による Grand Truth, JSEG による結果, 提案法による結果, 提案法による結果の領域表示である。視覚的にもJSEGと比較して良好な分割結果が得られていることを確認できる。

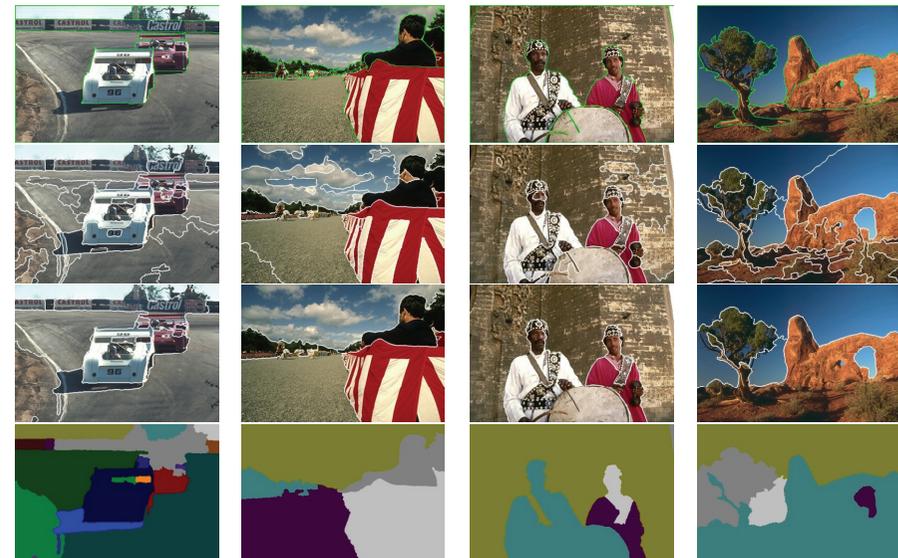
表3. Adjusted Probabilistic Rand Index

	JSEG	提案法 $\alpha=50\%$
画像1	0.216	0.436 (+0.220)
画像2	0.562	0.858 (+0.296)
画像3	0.687	0.814 (+0.127)
画像4	0.364	0.446 (+0.082)
画像30枚の平均	0.546	0.613 (+0.067)

4.3 処理量

提案法では過剰分割された各小領域に対して GMM による色分布の学習と Graph Cuts の処理を繰り返し適用する。また領域統合前後で変化が続く限り領域統合処理が繰り返し行われる。提案法の処理時間と、収束までの領域統合の繰り返し回数について計測した。表4に結果を示す。JSEGと比較すると提案法は約22倍の処理時間を必要とする。その大部分はGMMによる色分布の学習と Graph Cut に要する重い処理である。すなわち、図4に示した提案法全体の処理の中で、領域拡大に約95%と大部分の処理量が集中し、過剰分割は約5%、領域統合および収束判定にいたっては無視できるほど小さい。JSEGにより過剰に分割された34.7個の小領域が提案法によって最終的に11.6領域に統合され、これはGrand Truthの18.8領域に近い数となる。

なお、提案法の領域判定パラメータを $\alpha=25, 50, 75, 100\%$ と変化させた場合、それぞれ8.2, 11.6, 13.8, 25.7領域と、 α の値が小さいほど統合が進むことを確認した。また、繰り返し回数には2.33, 2.27, 2.27, 2.30回と α の値に関して特徴的な差異は確認できなかった。



画像1 画像2 画像3 画像4

図11. 領域分割結果

表4. 処理量 (画像30枚の平均値)

JSEG		提案法 $\alpha=50\%$			Grand Truth
処理時間	領域数	処理時間	繰り返し	領域数	領域数
5.07 秒	34.7	111.54 秒	2.27 回	11.6	18.1

5. まとめ

本論文では、従来の領域分割方法である Grabcut の課題を解決し、人手による事前入力が必要とせず、ユーザとの対話不要で実現できる領域分割法を提案した。提案法により色に基づいた領域分割の精度を改善し、Grabcutの強力な領域分割性能を活かした領域分割が可能となった。提案法は従来法であるBSEよりもRecallおよびPrecisionが良好な特性を部分的に示した。しかしながら、本実験で使用したパラメータではBSEよりも良好なF値は確認できていないことに加え、BSEは現在最良の領域分割法ではなく、BSEよりも強力な領域分割アルゴリズムが公開されているため、提案法の絶対的な優位性は確認できなかった。本論文では、画像を一度過剰分割し、分割された領

域を統合するアプローチと、画素単位ではなく小領域単位で性質の類似性を比較する方法についてその可能性を示した。今後は色以外の特徴を取り入れることで、同様の過剰分割と領域統合のアプローチによる精度改善が期待できると考えられる。

また、処理量と収束までの領域統合の繰り返し回数について計測した。表4に結果を示す。JSEGと比較すると提案法は約22倍の処理時間を必要とする。その大部分はGMMによる学習処理とGraphcut処理と負荷の重い処理であり、無駄な重複を含んでいる。無駄な処理の省略を行って高速化することが今後の課題となる。

参考文献

- 1) M. Kass, A. Witkin, and D. Terzopoulos, "Snakes: Active Contour Models," *International Journal of Computer Vision*, vol. 1, no. 3, pp. 321-331, 1988.
- 2) M. Sussman, P. Smereka, and S. Osher, "A Level Set Approach for Computing Solutions to Incompressible Two Phase Flow," *Journal of Computational Physics*, vol. 114, pp. 146-159, Sept. 1994.
- 3) Y. Boykov and V. Kolmogorov, "An Experimental Comparison of Min-Cut/Max-Flow Algorithms for Energy Minimization in Vision," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 26, no. 9, pp. 1124-1137, Sept. 2004.
- 4) 石川博, "グラフカット", 情報処理学会研究報告, 2007-CVIM-158-26, pp. 193-204, 2007.
- 5) Y. Boykov and M-P. Jolly, "Interactive Graph Cuts for Optimal Boundary & Region Segmentation of Objects in N-D Images," *Proceedings of IEEE International Conference on Computer Vision*, vol. 1, pp. 105-112, Jul. 2001.
- 6) Y. Boykov and G. Funka-Lea, "Graph Cuts and Efficient N-D Image Segmentation," *International Journal of Computer Vision*, vol. 70, no. 2, pp. 109-131, Nov. 2006.
- 7) C. Rother, V. Kolmogorov and A. Blake, "GrabCut: Interactive Foreground Extraction Using Iterated Graph Cuts," *ACM Transactions on Graphics*, vol. 23, no. 3, pp. 309-314, Aug. 2004.
- 8) Y. Deng, and B.S. Manjunath, "Unsupervised segmentation of color-texture regions in images and video," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 23, no. 8, pp. 800-810 Aug. 2001.
- 9) Jianbo Shi and Jitendra Malik, "Normalized Cuts and Image Segmentation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 8, Aug. 2000.
- 10) X. Ren and J. Malik, "Learning a classification model for segmentation," *Proceedings of IEEE International Conference on Computer Vision*, vol. 1, pp. 10-17, Oct. 2003.
- 11) D. Martin, C. Fowlkes, D. Tal, and J. Malik, "A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics," *Proceedings of IEEE International Conference on Computer Vision*, vol. 2, pp. 416-423, July 2001.
- 12) The Berkeley Segmentation Dataset and Benchmark, <http://www.eecs.berkeley.edu/Research/Projects/CS/vision/bsds/>
- 13) JSEG software, <http://vision.ece.ucsb.edu/segmentation/jseg/software/>
- 14) C. Fowlkes, D. Martin, and J. Malik "Learning Affinity Functions for Image Segmentation: Combining Patch-based and Gradient-based Approaches," *Proceedings of IEEE International Conference on Computer Vision and Pattern Recognition*, vol. 2, pp. 54-61, Jun. 2003.

- 15) D. Martin, C. Fowlkes, and J. Malik, "Learning to Detect Natural Image Boundaries Using Local Brightness, Color and Texture Cues," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 26, pp. 530-549, May 2004.
- 16) C. Fowlkes, J. Malik, "How Much Does Globalization Help Segmentation?," Technical Report CSD-04-1340, Division of Computer Science, University of California, Berkeley, Jul. 2004.
- 17) D. Martin, C. Fowlkes, D. Tal and J. Malik, "A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics," *Proceedings of IEEE International Conference on Computer Vision*, vol. 2, pp. 416-423, Jul. 2001.
- 18) W. M. Rand, "Objective criteria for the evaluation of clustering methods," *Journal of the American Statistical Association*, vol. 66, pp. 846-850, 1971.
- 19) R. Unnikrishnan and M. Hebert, "Measures of Similarity," *Proceedings of IEEE Workshop on Applications of Computer Vision*, pp. 394-400, Jan. 2005.
- 20) L. Hubert, and P. Arabie, "Comparing partitions," *Journal of Classification*, pp. 193-218, 1985

付録

A.1 Probabilistic Rand Index

$|P|$ 画素の画像 P に対し, K 人の被験者が作成した Grand Truth となる分割を G_1, G_2, \dots, G_K とし, またあるアルゴリズムによる領域分割の結果を分割 G とする. i 番目の画像に対し, 分割 G において割り当てられたラベルを X_i^G , 被験者 k による分割 G_k において割り当てられたラベルを $X_i^{G_k}$ と表す.

このとき, 分割 G の被験者 k の分割 G_k に対する Probabilistic Rand Index $PR(G, G_k)$ は,

$$PR(G, G_k) = \frac{1}{K} \sum_{k=1}^K \left\{ \frac{1}{|P| C_2} \sum_{\substack{i,j \\ i \neq j}} [\delta(X_i^G, X_j^G) \cdot \delta(X_i^{G_k}, X_j^{G_k}) + (1 - \delta(X_i^G, X_j^G)) \cdot (1 - \delta(X_i^{G_k}, X_j^{G_k}))] \right\} \quad (A1)$$

で定義される. ただし,

$$\delta(i, j) = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases} \quad (A2)$$

である. 尺度 $PR(G, G_k)$ は $[0, 1]$ の範囲で値を取り, 0 は分割 G と被験者 k が作成した Grand Truth となる分割 G_k の間に全く類似性が無いことを意味し, 反対に全く同じ分割が行えている場合は 1 を示す.