

繰り返し色分布マッチングによる画像セグメンテーション

ファンヴェトクオク[†] 高橋 桂太^{††} 苗村 健[†]

[†] 東京大学大学院情報理工学系研究科 〒113-8656 東京都文京区本郷 7-3-1

^{††} 東京大学 IRT 研究機構 〒113-8656 東京都文京区本郷 7-3-1

E-mail: †{viet,keita,naemura}@nae-lab.org

あらまし 画像から特定の領域を切り出す画像セグメンテーションの問題において、画素単位に定義される尤度の代わりに、領域全体に対して定義される色分布の類似性を用いる手法が近年に提案され、その有効性が示された。しかし、抽出対象の正確な色分布が与えられることが前提とされたため、適用領域は極めて限定されていた。本研究では、この仮定を緩めるため、与えられた色分布が必ずしも正確ではなくても、それを手がかりに真の色分布と画像中の対応領域を同時に推定する手法を提案する。本研究の貢献は、色分布同士の Bhattacharyya 距離によって定義される新しいエネルギー関数を導入し、それを補助関数を用いてグラフカットで近似的に最小化するアルゴリズムを示したことにある。本稿では、外包矩形を入力とする物体領域抽出を例として議論するが、提案手法は背景差分や co-segmentation にも適用可能である。

キーワード 画像セグメンテーション, 色分布マッチング, エネルギー最適化, グラフカット, Bhattacharyya 距離

1. はじめに

画像セグメンテーションとは、画像から特定の領域を切り出す処理を指しており、物体認識や画像編集にとって不可欠な過程である。画像セグメンテーションは一般的に曖昧な問題であるため、人手によるヒント、あるいは対象領域の事前知識が必要である。例えば、人手でマークされた画素をハードな制約とする手法 [1-4] や、対象領域（またはその相補領域）の色分布を与える手法 [5-8] がある。

本研究では、色分布を入力として、画像中の対応領域をセグメンテーションする問題を扱う。入力の色分布（参照色分布）は、セグメンテーション問題の目的に応じて、同じ画像から計算されるか、別の画像から推定される。この問題を扱う手法のほとんどは、エネルギー関数を最適化する 2 値ラベリングの問題に帰着する。エネルギー関数は一般的にアピアランスの適合性と空間的な連続性に対応する 2 つの項から成る。アピアランス適合性は、画素単位の尺度とグローバルな尺度に分類される。

画素単位の尺度に基づいた手法 [1-3, 6] では、各画素が参照色分布にどう当てはまるかによって、アピアランス適合性を評価する。アピアランス適合性は、画素ごとに独立な unary term の和として表わされ、尤度項とも呼ばれる。この場合、全体のエネルギー関数はこの unary term と、空間的な連続性を表す pairwise term のみを含むため、グラフカット法 [9, 10] が直接適用できる。すなわち、エネルギーの最小化は対応するグラフの最大流の計算に相当するので、特定の条件が満たされれば大域最適性を保証する効率的なアルゴリズムが存在する。しかし、これらの手法では、アピアランスのグローバルな適

合性を考慮しないため、縮小バイアス (shrinking bias) により物体の細端部が切断されやすい (shortcutting) などの問題点がある [4]。

一方、近年の研究 [5, 7] により、セグメンテーションにおいては、グローバルなアピアランス適合性の尺度が、画素単位の尺度よりも有効であることが示された。グローバルな尺度としては、領域全体の色分布が用いられるが、直接的に unary term として定式化できないため、数学的な扱いはやや複雑である。文献 [7] では、参照色分布と適合する画像領域をセグメンテーションするという目的で、ヒストグラム間 (参照色分布 抽出領域の色分布) の差分の L1 ノルムを用いてアピアランス適合性を定式化し、Trust Region Graph Cut (TRGC) と呼ばれる一連のパラメトリック線形近似法により最適化する手法を示した。しかし、TRGC 法には、パラメータが領域の大きさに強く依存するという根本的な問題点があることが、I. Ben Ayed ら [5] によって示された。I. Ben Ayed らは、自らの提案手法 [5] で、Bhattacharyya 距離を用いてアピアランス適合性を定式化した。また、Bhattacharyya Measure Graph Cut (BMGC) と呼ばれる手法を導入し、エネルギー関数の上界として算出した補助関数をグラフカットで最適化する手法を提案した。彼らの実験結果により、参照色分布が正解データから正確に入手できる場合、BMGC 法は精度が非常に高く、TRGC 法 [7] や active contour 法 [11-13] をはるかに上回ることが示された。

しかし、真の色分布が入力として得られるという仮定は、多くの応用において現実的ではない。例として、外包矩形を入力とするセグメンテーションを考える。入力はユーザによって与えられた対象物体を囲む矩形である。

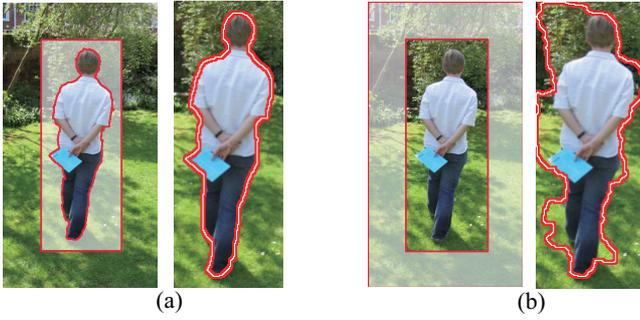


図 1 Segmentation using BMGC method [5] from (a) ground-truth distribution and (b) user-supplied background distribution. The reference distribution is computed from the white-masked region.

多くの場合，矩形外部の色分布を背景の参照色分布とする^(注1)が，その分布は一般に真の色分布とは異なる．我々の実験によれば，このようなケースにおいて，BMGC法は破綻する．図 1 は，我々が BMGC 法を実装して実験した結果の一例である．人物を囲むように外包矩形を設定し，人物領域と背景領域の分離を試みたが，図 1 (a) では予め用意した正解データから背景の参照色分布を与え，図 1 (b) では外包矩形外部から参照色分布を与えた．前者のケースでは正確なセグメンテーションができたのに対して，後者では大きく失敗した．

そこで，我々は，参照色分布が必ずしも正確でなくても，その参照色分布を手がかりとして真の色分布と対応領域を同時に推定できる，新しいセグメンテーション手法を提案する．我々の手法でも，BMGC 法 [5] と同様に，色ヒストグラム間の Bhattacharyya 距離に基づき，グローバルなアピランス適合性を評価する．提案手法のエネルギー関数は，参照色分布を考慮しつつ，推定される色分布が抽出領域の色分布にできるだけ類似し，同時に相補領域の色分布とできるだけ異なるようにモデル化され，色分布と抽出領域を交互に更新することで最適化される．更新の過程では，補助関数を導入し，これをグラフカットで近似的に最小化する．この更新の過程は，BMGC 法にヒントを得たものであるが，提案手法のエネルギー関数のほうがより複雑であり，Bhattacharyya 距離の数学的扱いにおいても新たな貢献を含むものである．実験では，GrabCut データベース [6] の多数の画像を用いて，提案手法が外包矩形によるセグメンテーションにおいて有効であることを確認した．また，提案手法が背景差分や co-segmentation に適用できることも示した．

2. アルゴリズム

2.1 真の色分布の推定

入力として与えられた参照色分布を \mathcal{M} とする．外包矩形を入力とするセグメンテーション [6] の場合では， \mathcal{M}

(注 1): それに加えて，GrabCut 法 [6] のように，矩形内部の色分布を対象物の参照色分布とすることもある．

は矩形外部の色分布である．本節では， \mathcal{M} を手がかりとして背景の真の色分布 \mathcal{H} を推定するアルゴリズムを説明する．画像を 2 つの領域に分類する滑らかなセグメンテーション L を仮定し，一方の領域を $R_1^L = \{p | L_p = 1\}$ (前景)，他方を $R_0^L = \{p | L_p = 0\}$ (背景) とする．真の色分布 \mathcal{H} は以下の特徴を持つものとする．

- 1: \mathcal{H} は \mathcal{M} に近い
- 2: \mathcal{H} は R_0^L 領域の色分布に近い
- 3: \mathcal{H} は R_1^L 領域の色分布と異なる

3 の条件がない場合，真の色分布 \mathcal{H} の最適解は \mathcal{M} となる．これは，TRGC 法や BMGC 法と同じモデルであり， \mathcal{M} が誤差を含む場合には，図 1 (b) に示すように，望ましいセグメンテーションが得られない．本研究の貢献は， \mathcal{M} の誤差に対応するため，3 の条件を導入したことである．この条件は 1, 2 の条件と符号が逆になるため (式 (1) を参照)，Bhattacharyya 距離の数学的扱いに新たな工夫が必要になる．

提案手法では，セグメンテーションの問題をエネルギー最小化の問題として解く．導入するエネルギーは上記の条件のバランスをとるものである．

$$\mathcal{E}(\mathcal{H}, L) = -\mathcal{B}(\mathcal{M}, \mathcal{H}) - \mathcal{B}(\mathcal{P}_0(L), \mathcal{H}) + \mathcal{B}(\mathcal{P}_1(L), \mathcal{H}) + \lambda \mathcal{S}(L) \quad (1)$$

ここで， $\mathcal{P}_i(L)$ は R_i^L の色分布， $\mathcal{B}(f, g)$ は分布間の類似度を表す．本手法では，その類似度を表す尺度として，先行研究 [5, 13] で有効に使われた Bhattacharyya 距離を用いることにする．

$$\mathcal{B}(f, g) = \sum_{z \in \mathcal{Z}} \sqrt{f(z)g(z)} \quad (2)$$

$\mathcal{S}(L)$ はセグメンテーション L の空間的な連続性を評価するペナルティであり，pairwise term として表される．

我々は， $\mathcal{E}(\mathcal{H}, L)$ を最小化するため， \mathcal{H} と L を交互に最適化する処理を反復する．

$$\begin{cases} L^{(t)} = \arg \min_L \mathcal{E}(L, \mathcal{H}^{(t-1)}) & (3) \\ \mathcal{H}^{(t)} = \arg \min_{\mathcal{H}} \mathcal{E}(L^{(t)}, \mathcal{H}) & (4) \end{cases}$$

\mathcal{H} の初期値としては，参照色分布を用いる ($\mathcal{H}^{(0)} = \mathcal{M}$)．この手法を繰り返し色分布マッチングと呼ぶことにする．式 (3) については 2.1.1 節で，式 (4) については 2.1.2 節で説明する．全体のアルゴリズムを 2.1.3 節で述べる．

2.1.1 L の最適化

L の最適化においては \mathcal{H} は固定されるから，式 (3) の $\mathcal{E}(L, \mathcal{H}^{(t-1)})$ の最小化は，次の関数の最小化に等価である．

$$\mathcal{E}_1(L) = \underbrace{\mathcal{B}(\mathcal{P}_1(L), \mathcal{H})}_{\mathcal{F}_1(L)} - \underbrace{\mathcal{B}(\mathcal{P}_0(L), \mathcal{H})}_{\mathcal{F}_2(L)} + \lambda \mathcal{S}(L) \quad (5)$$

エネルギー関数に含まれる項 $\mathcal{F}_1(L)$ と $\mathcal{F}_2(L)$ は Bhattacharyya 距離によって定義されるため，unary term の

形式になっておらず、グラフカット法を直接適用できないことが明らかである。

我々の手法では、I. Ben Ayed ら [5] が提案した補助ラベリングのアイデアを用いて、 $\mathcal{F}_1 + \mathcal{F}_2$ の上界関数を推定する。求める上界関数は unary term のみの形式で表せるため、グラフカット法により式 (5) を近似的に最小化することが可能になる。

まず、 $\mathcal{E}_1(L)$ の補助関数 $\hat{\mathcal{E}}_1(L, L^*)^{(\text{注2})}$ を導入する。具体的な定義は後から与えるが、補助関数は以下の2つの条件を満たすものである。

$$\hat{\mathcal{E}}_1(L, L^*) \geq \mathcal{E}_1(L) \quad (6)$$

$$\hat{\mathcal{E}}_1(L, L) = \mathcal{E}_1(L) \quad (7)$$

文献 [14] の図 1 に補助関数のこれらの性質が図示されているので参照されたい。 L の更新は、以下の式にしたがうものとする。

$$L^{(\tau+1)} = \arg \min_L \hat{\mathcal{E}}_1(L, L^{(\tau)}) \quad (8)$$

上記の繰り返し更新において、 \mathcal{E}_1 は増加しないことが式 (6) ~ 式 (8) を使って証明できる。

証明 $\mathcal{E}_1(L^{(\tau+1)}) \leq \hat{\mathcal{E}}_1(L^{(\tau+1)}, L^{(\tau)}) \leq \hat{\mathcal{E}}_1(L^{(\tau)}, L^{(\tau)}) = \mathcal{E}_1(L^{(\tau)})$ ■

式 (8) を収束するまで繰り返し、 \mathcal{E}_1 の極値を得る。

次に、上界関数 Q と補助関数 $\hat{\mathcal{E}}_1$ を具体的に定義する。数学的な定式化に入る前に、用いる記号を説明する。まず、セグメンテーション範囲の画像領域を I とする。外包矩形に基づくセグメンテーションの場合では、 I は矩形内部の全領域である。 I_p は画素 p の色ベクトルを表す。最後に、 \mathcal{P}_i^L ($i = 0, 1$) は分布 $\mathcal{P}_i(L)$ のカーネル密度推定 (KDE) とする：

$$\mathcal{P}_i^L(z) = \frac{\sum_{p \in R_i^L} K_z(I_p)}{A(R_i^L)} \quad \forall z \in Z \quad (9)$$

$A(R)$ は領域 R の画素数である。 K_z として任意のカーネル関数を使えるが、本手法では以下のようにする：

$$K_z(I_p) = \begin{cases} 1 & \text{if } I_p \in z \\ 0 & \text{if } I_p \notin z \end{cases} \quad (10)$$

すなわち、この場合 \mathcal{P}_i^L はヒストグラムである。 z はヒストグラムのピンを表す (例えば、RGB 空間で $128 \times 128 \times 128$ のヒストグラムを用いる場合では、 z は $2 \times 2 \times 2$ サイズの立方体である)。 Z は z の集合である。

補題 2.1 (上界関数) 固定されたラベル L^* に対し、 $R_1^{L^*} \supset R_1^L$ を満たす L を考える。すなわち、 L によって定義される前景領域が L^* によって定義される前景領域

(注2): L と L^* の間に一方的な関係があり、この場合では $R_1^L \supset R_1^{L^*}$ 。

を含む。このとき、下記の関数は $\mathcal{F}(L) = \mathcal{F}_1(L) + \mathcal{F}_2(L)$ の上界関数である。ここで、 α は任意の正の係数である。

$$Q(L, L^*, \alpha) = \sum_{p \in R_1^L} m_p(1) + (\text{sign}(\mathcal{F}(L^*))\alpha + 1) \sum_{p \in R_0^L} m_p(0) \quad (11)$$

$m_p(1)$ と $m_p(0)$ は各画素 p に対する unary term である。

$$\begin{aligned} m_p(1) &= \frac{\delta_{L_p^*=0}}{2A(R_1^{L^*})} \sum_{z \in Z} K_z(I_p) \sqrt{\frac{\mathcal{H}(z)}{\mathcal{P}_1^{L^*}(z)}} \\ &+ \frac{\delta_{L_p^*=0}}{A(R_0^{L^*})} \left(\sum_{z \in Z} K_z(I_p) \sqrt{\frac{\mathcal{H}(z)}{\mathcal{P}_0^{L^*}(z)}} + \mathcal{F}(L^*) \right) \\ m_p(0) &= \frac{\mathcal{F}(L^*)}{A(R_0^{L^*})} \end{aligned} \quad (12)$$

ここで δ はクロネッカーのデルタである。(注3)

証明 \mathcal{F}_1 と \mathcal{F}_2 の和の上界を求めるには、Bhattacharyya 距離 B の上界と下界の両方が必要になる。 B の下界はすでに I. Ben Ayed らの研究 [5] で導出された。我々は、その結果を用いるとともに、自らの方法で B の上界を算出する。Bhattacharyya 距離の上界の推定は本研究の新たな数学的貢献である。証明の詳細は付録に記述する。

補題 2.2 (補助関数) $\alpha = 0$ の場合には、次の関数は $\mathcal{E}(L)$ の補助関数となる。

$$\hat{\mathcal{E}}_1(L, L^*, \alpha) = Q(L, L^*, \alpha) + \lambda S(L) \quad (13)$$

証明 補題 2.1 は $\mathcal{F}_1 + \mathcal{F}_2$ の上界を求めたものであるから、式 (13) が条件式 (6) を満たすのは自明である。また、 $L^* = L$ のとき、 $\forall p \in R_1^L$ に対して $\delta_{L_p^*=0} = 0$ である。したがって、式 (11) で、 $\sum_{p \in R_1^L} m_p(1) = 0$ であるから、

$$\hat{\mathcal{E}}_1(L, L, 0) = \sum_{p \in R_0^L} \frac{\mathcal{F}(L)}{A(R_0^L)} + \lambda S(L) = \mathcal{E}_1(L) \quad (14)$$

が成り立ち、条件式 (7) が満たされる。■

2.1.2 \mathcal{H} の最適化

\mathcal{H} に最適化において L は固定されるため、式 (4) の $\mathcal{E}(L^{(t)}, \mathcal{H})$ の最小化は下記の関数の最小化と同等である。

$$\mathcal{E}_2(\mathcal{H}) = -\mathcal{B}(\mathcal{M}, \mathcal{H}) - \mathcal{B}(\mathcal{P}_0(L), \mathcal{H}) + \mathcal{B}(\mathcal{P}_1(L), \mathcal{H}) \quad (15)$$

補題 2.3 $\mathcal{E}_2(\mathcal{H})$ の最小解は次のようになる

$$\mathcal{H}(z) = \begin{cases} 0 & (\forall z \in Z^-) \\ a_z^2 / \sum_{Z^+} a_z^2 & (\forall z \in Z^+) \end{cases} \quad (16)$$

ただし、定数 a_z は

$$a_z = \sqrt{\mathcal{M}(z)} + \sqrt{\mathcal{P}_0^L(z)} - \sqrt{\mathcal{P}_1^L(z)} \quad (17)$$

$Z^+ = \{z | a_z > 0\}$ 、 $Z^- = \{z | a_z \leq 0\}$ と定義される。

(注3): $R_1^{L^*} = \emptyset$ (空) の場合、 $A(R_1^{L^*}) = \mathcal{P}_1^{L^*}(z) = 1$ とする。

証明 式 (15) と式 (17) に基づき,

$$-\mathcal{E}_2(\mathcal{H}) = \sum_{z \in Z} a_z \sqrt{\mathcal{H}(z)} \quad (18)$$

$$\leq \sum_{z \in Z^+} a_z \sqrt{\mathcal{H}(z)} \quad (19)$$

$$\leq \sqrt{\left(\sum_{z \in Z^+} a_z^2 \right) \left(\sum_{z \in Z^+} \mathcal{H}(z) \right)} \quad (20)$$

$$\leq \sqrt{\sum_{z \in Z^+} a_z^2} \quad (21)$$

となる．式 (20) は Cauchy-Schwarz 不等式から導かれ，式 (21) は $\sum_Z \mathcal{H}(z) = 1$ による証明される．等式が成り立つ必要十分条件が式 (16) である． ■

2.1.3 アルゴリズムのまとめ

真の色分布 \mathcal{H} を推定する手続きは次のようになる．

手続き 1: 繰り返し色分布マッチング

- 1: \mathcal{H} の初期化: $\mathcal{H}^{(0)} = \mathcal{M}$
- 2: 収束するまで繰り返し
 - 2.1: 手続き 2 の実装により
 $L^{(t)} = \arg \min_L \mathcal{E}(L, \mathcal{H}^{(t-1)})$ と更新
 - 2.2: 式 (16) の実装により
 $\mathcal{H}^{(t)} = \arg \min_{\mathcal{H}} \mathcal{E}(L^{(t)}, \mathcal{H})$ と更新
- 3: $\tilde{\mathcal{H}} = \mathcal{H}^{(t)}$ を出力

行 2.2 に示す \mathcal{H} を更新する処理は補題 2.3 により直接に計算される．また，行 2.1 に示す L を更新する処理は補題 2.1 と 2.2 によって導かれる．我々は，この処理を実装するため，I. Ben Ayed らの手法 [5] と同様な手続きを適用する．

手続き 2: 補助ラベリング

- 1: L^* の初期化: $L_p^* = 0 \quad \forall p \in I$
- 2: α の初期化: $\alpha = \alpha_0$ (十分に小さな正実数)
- 3: 収束するまで繰り返し
 - 3.1: グラフカットで補助関数を最小化
 $L^{(\tau)} = \arg \min_{L: R_1^L \supset R_1^{L^*}} \hat{\mathcal{E}}_1(L, L^*, \alpha)$
 - 3.2: L^* の更新: $L^* = L^{(\tau)}$
 - 3.3: α を減少させる: $\alpha = \alpha^\rho$ ($\rho > 1$)

α が 0 に近づくと，上記の反復は必ず収束する．パラメータ α を追加する理由は，領域 R_0^L の各画素 p に追加のペナルティを加え，前景領域を拡大する方向にバイアスをかけるためである．行 3.1 の関数 $\hat{\mathcal{E}}_1$ は unary term と pairwise term のみを含んでいるため，グラフカット [9] を用いることにより正確に最適化できる． $R_1^L \supset R_1^{L^*}$ の条件を満足するために， $R_1^{L^*}$ の領域が確実に前景になるようにハードな制約を加える．ただし，ノイズが伝搬しないように，領域 $R_1^{L^*}$ の外周部の数画素（本稿では 2）には制約を加えない．

2.2 後処理 ラベルの高精度化

$\tilde{\mathcal{H}}$ は手続き 1 により推定された背景の真の色分布とする．図 2 に示すように，推定された色分布 $\tilde{\mathcal{H}}$ は，入力として与えられた参照色分布 \mathcal{M} を大きく改良し，ground-truth の色分布に近づくことがわかる．この節では，推定された色分布に基づき，セグメンテーションを高精度する手法を説明する．

直接的なアプローチは，手続き 1 から得られたラベル L ，つまり関数 $\mathcal{E}(L, \tilde{\mathcal{H}})$ (式 (1) を参照) を L に関して最小化したものを，セグメンテーション結果として採用することである．しかし， $\tilde{\mathcal{H}}$ が精度よく推定されたこの時点では，誤差を含む参照色分布 \mathcal{M} を除外して，再度セグメンテーションのラベルを推定したほうが良い．ここでは，ラベルの空間的な連続性を表す項 $S(L)$ の寄与を増加させる．また，相補領域との違いを表す項 $-\mathcal{B}(\mathcal{P}_1(L), \tilde{\mathcal{H}})$ の寄与を減少させる．したがって，後処理においてはコスト関数を次のように定義する．

$$\tilde{\mathcal{E}}(L) = \gamma \mathcal{B}(\mathcal{P}_1(L), \tilde{\mathcal{H}}) - \mathcal{B}(\mathcal{P}_0(L), \tilde{\mathcal{H}}) + \lambda' S(L) \quad (22)$$

ここで， $\gamma < 1$ および $\lambda' > \lambda$ である． $\gamma = 0$ の場合は，正解色分布を出発点とした I. Ben Ayed らの手法 [5] と同じコスト関数が得られる．式 (22) は式 (5) と同じ形をしており，手続き 2 によって最小化できる．

3. 実験

提案手法の有効性を評価するため，外包矩形に基づくセグメンテーション，背景差分，co-segmentation の実験を行った．

まず，実装の詳細を説明する．式 (1) の平滑化項を

$$S(L) = \sum_{(p,q) \in N} \delta_{L_p + L_q} \left(\frac{1}{1 + \|I_p - I_q\|^2} + \frac{\epsilon}{\|p - q\|} \right) \quad (23)$$

と定義する． $\epsilon = 10^{-3}$ ， N は 8 近傍の隣接画素組の集合である．括弧の中で，第一項は境界線に沿ったコントラストを評価し，第二項は境界線の長さを短くする方向にバイアスをかける．式 (1) と式 (22) における $S(L)$ の重みパラメータは，それぞれ $\lambda = 2 \times 10^{-3}$ と $\lambda' = 5 \times 10^{-3}$ である．手続き 2 に用いたパラメータは $\alpha_0 = 0.85$ ， $\rho = 1.1$ とした．本実験では，RGB の色モデルを用い， $128 \times 128 \times 128$ の 3 次元ヒストグラムを色分布モデルとして使った．

本稿の全ての実験に対して，手続き 1 は 3 回の反復で，手続き 2 は 5 回の反復で収束した．式 (22) では， $\mathcal{B}(\mathcal{P}_1(L), \tilde{\mathcal{H}})$ の重みパラメータ γ を 0 にした．この設定により， $\tilde{\mathcal{E}}(L)$ の最適化は 3 回で収束した．

真の色分布の推定: 我々は，提案した繰り返し色分布マッチングの処理により，参照色分布がどう改良されるかを調べた．図 2 に，手続き 1 の繰り返し回数 t における，推定された色分布 $\mathcal{H}^{(t)}$ およびラベル $L^{(t)}$ を示す．

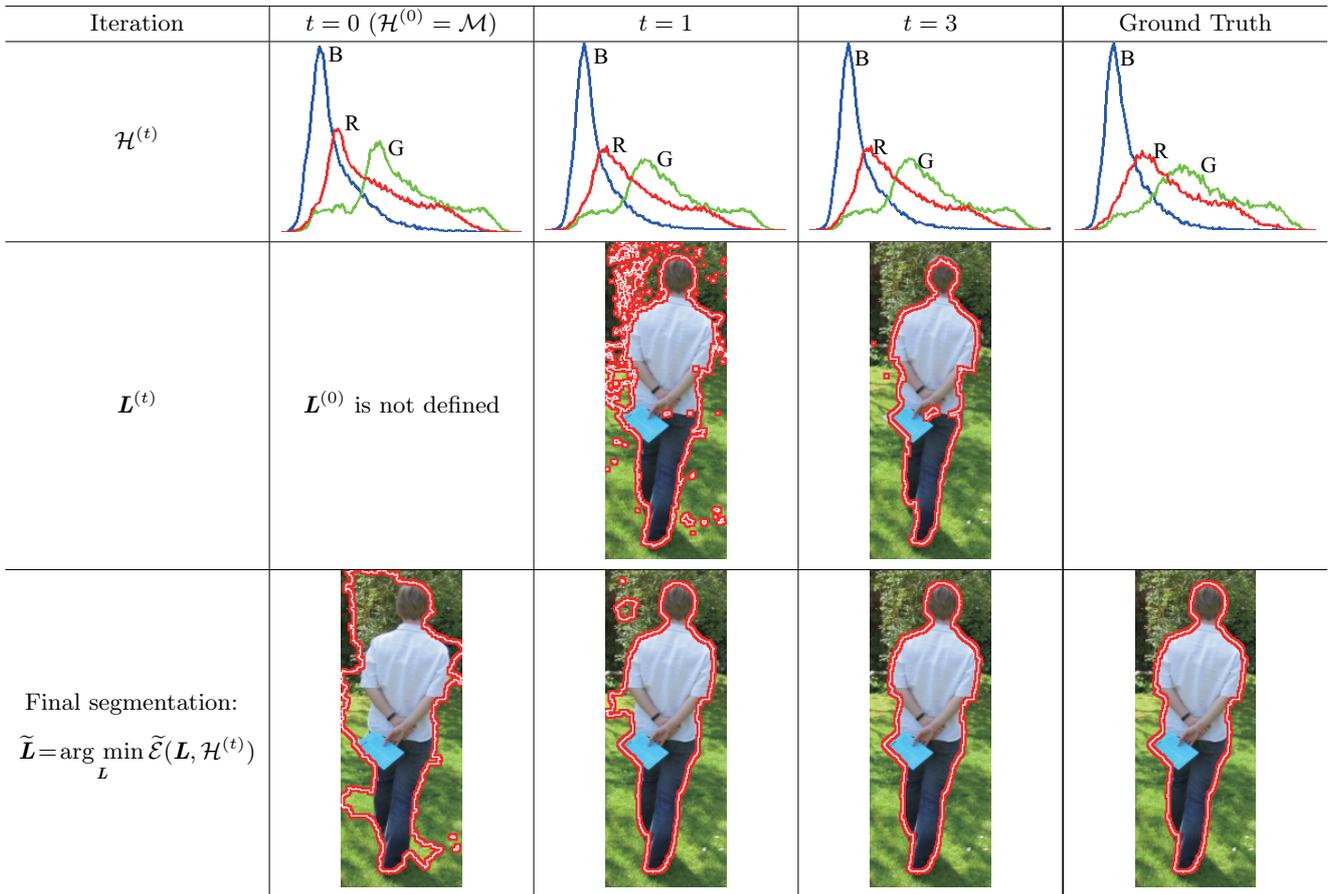


図2 Latent distribution $\mathcal{H}^{(t)}$ and label $L^{(t)}$, corresponding to the number of iterations t of Procedure 1. $\mathcal{H}^{(t)}$ closely approaches ground-truth distribution.

$t = 0$ のとき、つまり参照色分布 \mathcal{M} が対応領域の抽出に直接に使われた場合、参照色分布は ground-truth 分布と大きく異なるため、セグメンテーションが失敗した。しかし、繰り返しマッチングの処理により、推定された色分布 $\mathcal{H}^{(t)}$ は \mathcal{M} から出発して、ground-truth 分布に近づいたので、セグメンテーションの結果が大きく改善された。

外包矩形に基づくセグメンテーションに対する定量的および定性的評価：GrabCut データベース [6] の 49 枚^(注4)の画像およびデータベースに定義された外包矩形に対して、提案手法を適用した。本実験では、参照色分布は矩形外部の領域の色分布とした。提案手法では、まず、繰り返し色分布マッチング(手続き 1)を実行し、背景の真の色分布 $\tilde{\mathcal{H}}$ を推定する。次に、推定された色分布に基づき、節 2.2 で述べた手法を適用して、セグメンテーションの精度を改善し、結果として出力した。セグメンテーションの精度を以下の尺度で評価した。

$$\text{誤差率} = \frac{\text{誤って分類された画素数}}{\text{矩形内部の全画素数}} \quad (24)$$

上記の誤差率の全てのサンプルに対する平均値を計算し、GrabCut 法 [6] および Dual Decomposition (DD) 法 [8]

(注4): 全 50 枚のうち、“cross.jpg”では画像全体が外包矩形とされているため、提案手法が適用できず、実験から除外した。

表 1 Comparison of average error rate and running time.

Method	Ours	GrabCut [6]	DD [8]
Avg. error rate	7.1%	8.1%	10.5%
Running time	1.6s	1.5s	> 100s

と比較した。表 1 に示すように、提案手法の精度が一番高いことがわかる。

いくつかのセグメンテーション結果を図 3 に示す。各組において、GrabCut 法 [6] の結果を左に、提案手法の結果を右に配置した。その他の画像に対する結果は添付資料を参照されたい。ほとんどの場合、提案手法は GrabCut 法より視覚的にも良い結果を得た。GrabCut 法は画素単位の尺度を用いてアピランス適合性を評価するため、図 3 に示すように、縮小バイアス (shrinking bias) により物体の細端部が切断されやすい (shortcutting)。一方、提案手法ではグローバルな尺度を用いて色分布間の類似度を定式化するため、shortcutting の問題を克服できた。外包矩形に基づくセグメンテーションに関しては、Lempitsky らの研究 [15] で、外包矩形の各辺が物体の領域に十分に近いという仮定 (tightness prior) が導入され、GrabCut 法を改善できることが示された。本研究の提案はアピランス適合性に関するものである一方、tightness prior は領域のトポロジーに関する仮定であり、性質が全く異なるものである。このトポロジー的な事前



図 3 Segmentation results: GrabCut [6] (left) and our method (right). Our method provides better results because the distribution is inferred based on global measures. (Due to limited space, we only show the region inside the bounding box.)

知識はエネルギー関数をさらに複雑にするため、本稿では採用しなかったが、将来的には本研究の提案と組み合わせられる可能性がある。

計算時間：表 1 は提案手法と従来手法の計算時間を示す。実験では、Intel Core 2 Duo, 3.0GHz の CPU と 4GB の RAM を搭載したラップトップ PC を使用した。提案手法で用いる最適化処理は最小切断（最大流）[9] の計算を少数回繰り返すのみなので、GrabCut 法と計算コストが同程度である。

他のセグメンテーション問題：提案手法は参照色分布がある程度正確であれば動作するため、様々な応用が可能である。

図 4 は背景差分の問題に提案手法を適用した例である。背景画像（図 4 左）全体から参照色分布を取得し、対象画像（図 4 右）全体に提案手法をそのまま適用した。参照画像と対象画像では視点位置が移動しており、文献 [16]

のような典型的な背景差分法が適用できない設定になっている。提案手法では、背景画像に含まれていた領域と対象画像で新しく現れた領域を正しく分離できている。

2 枚の画像から共通する領域を抽出する問題は co-segmentation と呼ばれる。Co-segmentation のアイデアは Rother ら [7] によって提案され、現在ではセグメンテーション問題の中で注目されているテーマの 1 つである [7, 17, 18]。提案手法はこの問題にも適用できる。まず、一方の画像全体から参照色分布を取得し、それをを用いてもう一方の画像全体に提案手法を適用してセグメンテーションを行う。この処理を画像の役割を入れ替えて再度行う。図 5 に co-segmentation の一例を示す。上段の入力画像では背景領域が共通している。下段に示す処理結果では、それぞれ背景からハサミとバナナが分離できている。前景色と背景色が近いにもかかわらず、正確なセグメンテーションができていることから、繰り返し



図 4 Segmentation results (right) from reference background images (left). The backgrounds are *not identical* between the input and reference images.



図 5 Background co-segmentation: The common background parts of an image pair (top row) are simultaneously segmented (bottom row).

色分布マッチングの有効性が明らかである。

4. ま と め

本研究では、色分布を入力として、画像中の対応領域をセグメンテーションする問題において、与えられた色分布が必ずしも正確でなくても、それを手がかりに真の色分布と画像中の対応領域を同時に推定する手法を提案した。この手法は、従来研究で示されていた色分布間の Bhattacharyya 距離の下界に加えて、その上界を新たに導出したことにより可能になったものである。提案手法で導出された補助関数は unary term と pairwise term のみで構成されるため、グラフカットを用いて最小化できる。提案手法は外包矩形によるセグメンテーション、背景差分、co-segmentation の問題に対して効果的に適用できることが示された。

- [1] Y. Boykov and M. Jolly: “Interactive graph cuts for optimal boundary and region segmentation of objects in n-d images”, Proc. ICCV, pp. 105–112 (2001).
- [2] J. Liu, J. Sun and H.-Y. Shum: “Paint selection”, Proc. SIGGRAPH, pp. 1–7 (2009).
- [3] Y. Li, J. Sun, C.-K. Tang and H.-Y. Shum: “Lazy snapping”, Proc. SIGGRAPH, pp. 303–308 (2004).
- [4] B. Price, B. Morse and S. Cohen: “Geodesic graph cut for interactive image segmentation”, Proc. CVPR, pp. 3161–3168 (2010).
- [5] I. Ben Ayed, H. Chen, K. Punithakumar, I. Ross and S. Li: “Graph cut segmentation with a global constraint: Recovering region distribution via a bound of the bhattacharyya measure”, Proc. CVPR, pp. 3288–3295 (2010).
- [6] C. Rother, V. Kolmogorov and A. Blake: “Grabcut: interactive foreground extraction using iterated graph cuts”, ACM Trans. Graph., **23**, 3, pp. 309–314 (2004).
- [7] C. Rother, V. Kolmogorov, T. Minka and A. Blake: “Cosegmentation of image pairs by histogram matching - incorporating a global constraint into mrfs”, Proc. CVPR, pp. 993–1000 (2006).
- [8] S. Vicente, V. Kolmogorov and C. Rother: “Joint optimization of segmentation and appearance models”, Proc. ICCV, pp. 755–762 (2009).
- [9] Y. Boykov and V. Kolmogorov: “An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision”, IEEE Trans. PAMI, **26**, 9, pp. 1124–1137 (2004).
- [10] V. Kolmogorov and R. Zabini: “What energy functions can be minimized via graph cuts?”, IEEE Trans. PAMI, **26**, 2, pp. 147–159 (2004).
- [11] D. Freedman and T. Zhang: “Active contours for tracking distributions”, IEEE Trans. Image Processing, **13**, 4, pp. 518–526 (2004).
- [12] T. Zhang and D. Freedman: “Improving performance of distribution tracking through background mismatch”, IEEE Trans. PAMI, **27**, 2, pp. 282–287 (2005).
- [13] O. Michailovich, Y. Rathi and A. Tannenbaum: “Image segmentation using active contours driven by the bhattacharyya gradient flow”, IEEE Trans. Image Processing, **16**, 11, pp. 2787–2801 (2007).
- [14] D. D. Lee and H. S. Seung: “Algorithms for nonnegative matrix factorization”, Proc. NIPS, pp. 556–562 (2002).
- [15] V. Lempitsky, P. Kohli, C. Rother and T. Sharp: “Image segmentation with a bounding box prior”, Proc. ICCV, pp. 277–284 (2009).
- [16] J. Sun, W. Zhang, X. Tang and H.-Y. Shum: “Background cut”, Proc. ECCV, pp. 628–641 (2006).
- [17] L. Mukherjee, V. Singh and C. Dyer: “Half-integrality based algorithms for cosegmentation of images”, Proc. CVPR, pp. 2028–2035 (2009).
- [18] D. Hochbaum and V. Singh: “An efficient algorithm for co-segmentation”, Proc. ICCV, pp. 269–276 (2009).

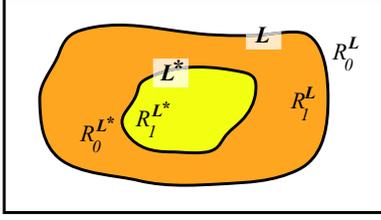


図 A.1 Given a fixed label L^* , the variable label L , which separates the image into two regions R_0^L and R_1^L , satisfies $R_1^L \subset R_1^{L^*}$.

付 録

証明（補題 2.1）まず， $\mathcal{F}_1(L)$ の上界を推定する．これは Bhattacharyya 距離の上界の推定に等価であり，本研究の独自の提案である．図 A.1 に示すように， $R_1^L \subset R_1^{L^*}$ の条件から

$$R_1^L = R_1^{L^*} \cup (R_1^L \cap R_0^{L^*}) \quad (\text{A.1})$$

が成り立つ．式 (9) の KDE 関数は次のように書き直せる．

$$\begin{aligned} \mathcal{P}_1^L(z) &= \frac{\sum_{p \in R_1^{L^*}} K_z(I_p) + \sum_{p \in R_1^L \cap R_0^{L^*}} K_z(I_p)}{A(R_1^{L^*}) + A(R_1^L \cap R_0^{L^*})} \\ &\leq \frac{\sum_{p \in R_1^{L^*}} K_z(I_p) + \sum_{p \in R_1^L \cap R_0^{L^*}} K_z(I_p)}{A(R_1^{L^*})} \end{aligned} \quad (\text{A.2})$$

この不等式を用いて $\mathcal{F}_1(L)$ を評価する（式を簡約するため， z を省略する）．

$$\begin{aligned} \mathcal{F}_1(L) &= \sum_{z \in Z} \sqrt{\mathcal{P}_1^L \mathcal{H}} \\ &\leq \sum_{z \in Z} \sqrt{\left(\mathcal{P}_1^{L^*} + \frac{\sum_{p \in R_1^L \cap R_0^{L^*}} K_z(I_p)}{A(R_1^{L^*})} \right) \mathcal{H}} \\ &= \sum_{z \in Z} \sqrt{1 + \frac{\sum_{p \in R_1^L \cap R_0^{L^*}} K_z(I_p)}{\sum_{p \in R_1^{L^*}} K_z(I_p)}} \sqrt{\mathcal{P}_1^{L^*} \mathcal{H}} \end{aligned} \quad (\text{A.3})$$

和の中の第一項を評価するために，不等式

$$\sqrt{1+x} \leq 1 + \frac{x}{2} \quad \forall x \geq 0 \quad (\text{A.4})$$

を適用して， $\mathcal{F}_1(L)$ の上界関数を得る．

$$\begin{aligned} \mathcal{F}_1(L) &\leq \sum_{z \in Z} \left(1 + \frac{1}{2} \frac{\sum_{p \in R_1^L \cap R_0^{L^*}} K_z(I_p)}{\sum_{p \in R_1^{L^*}} K_z(I_p)} \right) \sqrt{\mathcal{P}_1^{L^*} \mathcal{H}} \\ &= \mathcal{F}_1(L^*) + \sum_{z \in Z} \frac{1}{2} \frac{\sum_{p \in R_1^L \cap R_0^{L^*}} K_z(I_p)}{\mathcal{P}_1^{L^*} A(R_1^{L^*})} \sqrt{\mathcal{P}_1^{L^*} \mathcal{H}} \\ &= \mathcal{F}_1(L^*) + \sum_{p \in R_1^L} \frac{\delta_{L_p^*=0}}{2A(R_1^{L^*})} \sum_{z \in Z} K_z(I_p) \sqrt{\frac{\mathcal{H}(z)}{\mathcal{P}_1^{L^*}(z)}} \end{aligned} \quad (\text{A.5})$$

この関数は unary term と定数のみを含んでいる．

次に，文献 [5] と同様に，Bhattacharyya 距離の下界によって $\mathcal{F}_2(L)$ の上界を推定する． $R_1^L \supset R_1^{L^*}$ から，

$$R_0^L = R_0^{L^*} \setminus (R_1^L \cap R_0^{L^*}) \quad (\text{A.6})$$

が成り立つ．式 (9) の KDE 関数は次のように書き直せる．

$$\begin{aligned} \mathcal{P}_0^L(z) &= \frac{\sum_{p \in R_0^{L^*}} K_z(I_p) - \sum_{p \in R_1^L \cap R_0^{L^*}} K_z(I_p)}{A(R_0^{L^*}) - A(R_1^L \cap R_0^{L^*})} \\ &\geq \frac{\sum_{p \in R_0^{L^*}} K_z(I_p) - \sum_{p \in R_1^L \cap R_0^{L^*}} K_z(I_p)}{A(R_0^{L^*})} \end{aligned} \quad (\text{A.7})$$

この不等式に基づいて， $\mathcal{F}_2(L)$ を評価する．

$$\begin{aligned} \mathcal{F}_2(L) &= - \sum_{z \in Z} \sqrt{\mathcal{P}_0^L \mathcal{H}} \\ &\leq - \sum_{z \in Z} \sqrt{\left(\mathcal{P}_0^{L^*} - \frac{\sum_{p \in R_1^L \cap R_0^{L^*}} K_z(I_p)}{A(R_0^{L^*})} \right) \mathcal{H}} \\ &= - \sum_{z \in Z} \sqrt{1 - \frac{\sum_{p \in R_1^L \cap R_0^{L^*}} K_z(I_p)}{\sum_{p \in R_0^{L^*}} K_z(I_p)}} \sqrt{\mathcal{P}_0^{L^*} \mathcal{H}} \end{aligned} \quad (\text{A.8})$$

次の不等式

$$\sqrt{1-x} \geq 1-x \quad \forall x: 0 \leq x \leq 1 \quad (\text{A.9})$$

を適用して， $\mathcal{F}_2(L)$ の上界関数を得る．

$$\begin{aligned} \mathcal{F}_2(L) &\leq - \sum_{z \in Z} \left(1 - \frac{\sum_{p \in R_1^L \cap R_0^{L^*}} K_z(I_p)}{\sum_{p \in R_0^{L^*}} K_z(I_p)} \right) \sqrt{\mathcal{P}_0^{L^*} \mathcal{H}} \\ &= \mathcal{F}_2(L^*) + \sum_{p \in R_1^L} \frac{\delta_{L_p^*=0}}{A(R_0^{L^*})} \sum_{z \in Z} K_z(I_p) \sqrt{\frac{\mathcal{H}(z)}{\mathcal{P}_0^{L^*}(z)}} \end{aligned} \quad (\text{A.10})$$

最後に，式 (A.5) と式 (A.10) を組み合わせ， $\mathcal{F}(L) = \mathcal{F}_1(L) + \mathcal{F}_2(L)$ の上界関数を計算する．

$$\begin{aligned} \mathcal{F}(L) &\leq \mathcal{F}(L^*) + \sum_{p \in R_1^L} \frac{\delta_{L_p^*=0}}{2A(R_1^{L^*})} \sum_{z \in Z} K_z(I_p) \sqrt{\frac{\mathcal{H}(z)}{\mathcal{P}_1^{L^*}(z)}} \\ &\quad + \sum_{p \in R_1^L} \frac{\delta_{L_p^*=0}}{A(R_0^{L^*})} \sum_{z \in Z} K_z(I_p) \sqrt{\frac{\mathcal{H}(z)}{\mathcal{P}_0^{L^*}(z)}} \end{aligned} \quad (\text{A.11})$$

定数項 $\mathcal{F}(L^*)$ は unary term の和として表せる．

$$\begin{aligned} \mathcal{F}(L^*) &= \frac{\mathcal{F}(L^*)A(R_1^L \cap R_0^{L^*})}{A(R_0^{L^*})} + \frac{\mathcal{F}(L^*)A(R_0^L)}{A(R_0^{L^*})} \\ &\leq \sum_{p \in R_1^L} \frac{\delta_{L_p^*=0} \mathcal{F}(L^*)}{A(R_0^{L^*})} + (\text{sign}(\mathcal{F}(L^*))\alpha + 1) \sum_{p \in R_0^L} \frac{\mathcal{F}(L^*)}{A(R_0^{L^*})} \end{aligned} \quad (\text{A.12})$$

（注： $\text{sign}(\mathcal{F}(L^*))$ が $\mathcal{F}(L^*)$ の記号， $\alpha \geq 0$ ）

式 (A.11) (A.12) により，補題 2.1 が証明される． ■