

前景/背景の局在性の相違に着目した高精度前景抽出 Robust Foreground Extraction Utilizing Difference in Localization between Foreground and Background

鈴木 哲明[†] 宮野 博義[†] 石寺 永記[†]
Tetsuaki Suzuki Hiroyoshi Miyano Eiki Ishidera

1. はじめに

前景と背景の位置に関する局在性の違いを利用して、高精度に前景抽出を行なう手法を提案する。近年、注目物抽出として、グラフカットを用いた前景抽出手法の有用性が多く文献で示されている。これらの手法の多くは、前景と背景とを同一モデルで扱っている場合が多く、この影響で背景を前景として誤検出する状況が見られる。この問題を解決するため、前景の局在性を利用して、前景抽出精度の高精度化を図った。注目物を撮影した画像は、注目物が画像中に局在する場合が多く、この場合に、局在性を活用することで前景物体である注目物の抽出精度改善が期待できる。本手法では、前景は色と位置でモデル化し、背景は色のみでモデル化する。この前景/背景モデルにより、前景のみ位置の近接性を考慮することができるため、前景の局在性を反映することが可能となる。実験では、本前景/背景モデルを用いて人物を前景とする前景抽出実験を行ない、背景の誤検出が低減されることを確認した。

2. グラフカットによる前景抽出

2.1 Interactive Graph Cuts

グラフカットを用いた前景抽出処理には、Interactive Graph Cuts[1]がある。Interactive Graph Cutsは画素をノードとするグラフと、ユーザが与える前景/背景ラベルのシードとを用いて、シードが与えられていない画素に対応するノードの前景/背景ラベルを最大流アルゴリズム[2]により決定する手法である。画素に対応するノードの集合を V 、隣接ノード間のエッジの集合を E とする。ノード i のラベルを s_i (0:背景, 1:前景) とし、その集合を $S=(s_1, \dots, s_n)$ (n :ノード数) とすると、最大流アルゴリズムで最適化されるエネルギー関数 $E(S)$ は、式(1)である。

$$E(S) = \sum_{i \in V} E_1(s_i) + \lambda \sum_{(i,j) \in E} E_2(s_i, s_j) \quad (1)$$

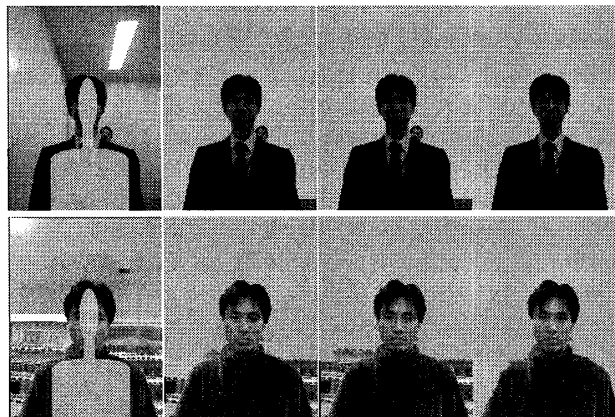
式(1)は、ノード i がラベル s_i である尤度を示す $E_1(s_i)$ と隣接するノード i とノード j がそれぞれラベル s_i, s_j であるコストを示す $E_2(s_i, s_j)$ とを実係数 λ で足し合わせた形となる。

この interactive graph cuts を高速化した手法として lazy snapping[3]がある。interactive graph cuts が画素をノードとしているのに対し、lazy snapping は watershed 法により得られた過分割領域をノードとしている点が異なるが、それ以外は本質的には同等である。そのため、本稿ではより効率的な lazy snapping の処理を前提として議論していく。

2.2 Lazy Snapping

Lazy Snapping では、式(1)の尤度 $E_1(s_i)$ は、以下のよう

[†](株)NEC 情報システムズ NEC Informatec Systems, Ltd.



(a)入力画像と (b)RGB 特徴 (c)XYRGB シード (d)提案手法
特徴
赤：前景、青：背景

図1 従来手法と提案手法の比較

して求められる。最初に、前景/背景シードに重なる過分割領域の平均色を K 平均法によりクラスタリングし、前景色と背景色のクラスタ中心 $\{K_n^F\}, \{K_m^B\}$ を得て、これを前景/背景モデルとする。次に、ノード i の平均色 $C(i)$ と前景/背景モデルの最近傍クラスタとの距離 d_i^F, d_i^B をそれぞれ $d_i^F = \min_n (\|C(i) - K_n^F\|), d_i^B = \min_m (\|C(i) - K_m^B\|)$ として求める。ここで、 $\|A-B\|$ は、色ベクトル A, B のユークリッド距離である。前景シードと重なる過分割領域の集合を F 、背景シードと重なる過分割領域の集合を B 、それ以外の過分割領域の集合を U とすると、尤度 $E_1(s_i)$ は式(2)で得られる。

$$\begin{cases} E_1(s_i=1) = 0 & E_1(s_i=0) = \infty & \forall i \in F \\ E_1(s_i=1) = \infty & E_1(s_i=0) = 0 & \forall i \in B \\ E_1(s_i=1) = \frac{d_i^F}{d_i^F + d_i^B} & E_1(s_i=0) = \frac{d_i^B}{d_i^F + d_i^B} & \forall i \in U \end{cases} \quad (2)$$

式(1)のコスト $E_2(s_i, s_j)$ は、式(3)で求める。

$$E_2(s_i, s_j) = |s_i - s_j| \cdot 1 / (1 + \|C(i) - C(j)\|^2) \quad (3)$$

図1(a)の前景/背景シードを与えた時の Lazy Snapping の人物領域抽出処理結果を図1(b)に示す。1段目の画像は、人物の背景に他人が居る状況である。この場合、背景の人物の肌色は前景シード内にいる人物に近い肌色であるため、誤検出する。2段目の画像では、背景が様々な色で構成されており、前景モデルに近い色を持つ領域が前景として誤検出されている。

3. 前景/背景の局在性の相違に基づく尤度

3.1 従来の尤度の問題点

Lazy Snapping で人物領域を抽出した時に生じた誤検出は、色の配置に関係なく色の近さだけで尤度を設定したために生じた。これは、Lazy Snapping の尤度 $E_1(a_i)$ (式(2))

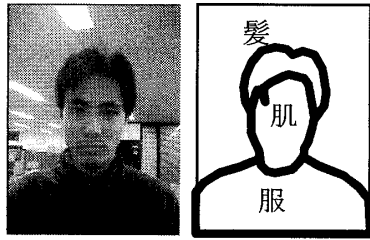


図2 前景領域の構造例

が過分割領域 i の空間的位置に寄らずその平均色が前景/背景モデルのどちらに近いのかによって決まることに起因する。図2のように、実際のデータでは、前景に局在性があるため、この局在性を利用することで、この誤検出を削減することが期待できる。前景の局在性を考慮するためには、色の近接性だけではなく、位置の近接性を考慮すれば良く、例えば特徴量を RGB から位置と色の特徴を融合した XYRGB に拡張することが、有力な対応策の1つである[4]。そこで、本稿でも特徴量として XYRGB を採用し、かつ前景と背景の局在性の違いを考慮する方法を考える。次節にて詳細を述べる。

3.2 色位置特徴空間内で定義される尤度

前景の局在性を考慮するため、色と位置を統合して扱うことのできる色位置特徴空間[4] (以降 XYRGB 空間)を用いて前景/背景尤度を定義する。この特徴空間では過分割領域の特徴は $C'(i)=[x_i, y_i, r_i, g_i, b_i]^T$ となる。ここで、 x_i, y_i は過分割領域の重心の X, Y 座標であり、 r_i, g_i, b_i は過分割領域の平均 R,G,B 階調値である。前景/背景モデル構築時には、前景は $A=[a_1, a_2, a_3, a_4, a_5]$ (a_i :実数)で、背景は $B=[b_1, b_2, b_3, b_4, b_5]$ (b_i :実数)で $C'(i)$ を変換してから K 平均法を行なう。尤度計算時にも同様に、前景尤度は A 、背景尤度は B で特徴 $C'(i)$ を変換してから計算する。ここで、 $A=B$ とすると、前景と背景とで同じ XYRGB 空間でモデル化され、 $a_1 > b_1, a_2 > b_2$ とすることで、前景は相対的に局在性が強くなる。特に、 $b_1=b_2=0$ とした場合、前景は XYRGB 空間で、背景は RGB 空間で尤度が得られ、前景は局在性が、背景は大域的性が最も強くなる。なお、 $a_1=a_2=b_1=b_2=0$ とすると従来の RGB 色空間でのモデル化となる。

本稿の実験では、前景と背景に同一モデルを用いる場合として、従来の RGB 空間と、XYRGB 空間を用いて尤度をモデル化した場合と、異なるモデルを用いる場合として、提案手法の前景が XYRGB 空間、背景が RGB 空間で尤度をモデル化した場合についてみていく。各モデルの傾向を比較するため、図3に各特徴空間の前景/背景モデルの学習結果の例を示す。図3は、図1(a)下段の前景/背景シードを用いて学習した前景/背景モデルが出力する尤度の傾向を XR 空間で示している。図(a)は学習データの分布、図(b), (c), (d)はそれぞれ RGB 色空間、XYRGB 色空間、前景が XYRGB 色空間背景が RGB 色空間で前景/背景モデルを学習した時の尤度の出力傾向である。図(b), (c), (d)が、図(a)と同じ色の分布になると前景/背景の分布の学習に成功したことになる。図(b)からは RGB 色空間を用いた場合は、位置を参照せず R 軸だけで判別を行なうため、十分な判別精度は期待できないことが分かる。図(c)では、XYRGB 色空間を用いた場合は、図(a)と似た色の分布が得られているが、学習データの無い領域(図(a)の白い領域)

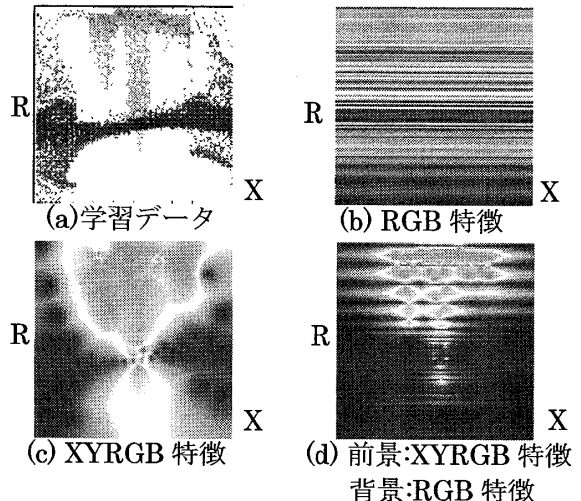


図3 特徴空間ごとの前景/背景尤度分布

(赤：前景、青：背景)

は赤(前景)と青(背景)領域に半々に分割される。この結果、学習データがない領域の特徴は色と位置に近いモデルに対する尤度が高くなる。一方、図(d)では、前景の局在性を利用したモデルは、前景の学習データがある領域のみで前景尤度が高くなり、それ以外の領域は背景の尤度が高くなる傾向にある。この傾向から、前景の局所性を高めると、より確実に前景と分かる領域のみが前景となり、背景の誤検出が低減することが期待できる。

4. 実験結果と考察

実験では、図3の3種類の特徴空間で得られる尤度を用いて実施した人物領域抽出の精度を比較した。図1に実験結果を示す。図1(d)のように、提案手法である前景と背景とで異なる色空間を用いた場合に、背景の誤検出が削減され最も抽出精度が高いことが分かる。

図3(d)のように提案手法では、前景/背景モデル構築時にデータがない領域は背景尤度が高くなるため誤検出は低減するが、逆に照明変動による階調値の変動などを吸収することは難しくなる。今後、このような領域の補完方法について、さらに改善を進めていく。

5. まとめ

本稿では、前景と背景の位置に関する局在性の違いを利用して、高精度に前景抽出を行なう手法を提案した。前景は色・位置特徴空間(xyrgb 空間)で、背景は色特徴空間(rgb 空間)でモデル化することで前景は局所的、背景は大域的なモデルとなる。本モデルを用いることで背景の誤検出を削減し、高精度な前景抽出を実現した。

参考文献

- [1] Y. Boykov, M.P. Jolly, "Interactive Graph Cuts for Optimal Boundary & Region Segmentation of Objects in N-D Images", ICCV, Vol. 1, pp. 105-112 (2001).
- [2] Y. Boykov, V. Kolmogorov, "An Experimental Comparison of Min-Cut/Max-Flow Algorithms for Energy Minimization in Vision", IN Energy Minimization Methods in Computer Vision and Pattern Recognition, (2001).
- [3] Y. Li, J. Sun, C-K. Tang, H-Y. Shum, "Lazy Snapping", SIGGRAPH (2004).
- [4] Y. Shekhi, M. Shah, "Bayesian Object Detection in Dynamic Scenes", CVPR, (2005)