

色および輝度勾配を用いたグラフカットによる画像セグメンテーション

1 序論

静止画像から重要度の高い領域を抽出（セグメンテーション）する画像前景抽出は，一般物体認識などの前処理として利用され，画像処理において重要な課題の一つとなっている．近年，画像セグメンテーション手法のなかでもグラフカットを用いて前景抽出を行う手法 [1] や，画像から得られた顕著性マップを二値化することで前景抽出を行う手法 [2] が注目されている．しかし，グラフカットを用いた手法では初期情報をあらかじめ与えなければならないため，ユーザの負担となることが課題である．また，顕著性マップを用いた手法では輪郭に沿った物体の抽出が困難であり，二値化処理が抽出結果に大きく影響を与える．

そこで，本論文では輝度，色相成分，輝度勾配，画像空間中の距離から生成した顕著性マップをグラフカットの初期情報として与える事で，高精度な画像セグメンテーションの完全自動化を実現する．

2 提案手法

2.1 顕著性マップ

本手法では，前処理として原画像を SLIC (Simple Linear Iterative Clustering) 法 [3] を用いてスーパーピクセルと呼ばれる知覚的に意味のある小領域に分割する．SLIC 法は原画像を Lab 表色系に変換し，色の類似度と画素の位置関係に基づいてクラスタリングを行うため，画像内の物体を考慮した領域分割が可能である．

原画像を SLIC 法によりスーパーピクセルに分割し，各スーパーピクセルについて輝度 $h(l')$ ，色差 $h(a')$ ， $h(b')$ ，勾配方向毎に対応する勾配強度を投票した重み付き勾配方向 $h(g')$ の4つのヒストグラムを生成する．ここで， l' ， a' ， b' ， g' はヒストグラムのビン数である．次に，注目するスーパーピクセルと他のスーパーピクセルのヒストグラム間の類似度を求め，その平均値をスーパーピクセルの顕著度とする．このヒストグラムの類似度の平均値を輝度，色差，重み付き勾配方向ヒストグラムについて求めたものをそれぞれ R_l ， R_a ， R_b ， R_g として以下の式で結合し，これを注目スーパーピクセル k' の顕著度と定義する．

$$R_f(k') = \frac{R_l(k') + 0.5(R_a(k') + R_b(k')) + R_g(k')}{3} \quad (1)$$

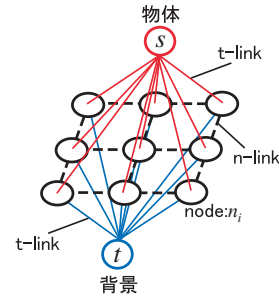


図 1: グラフ G の概念図

2.2 グラフカットによる画像セグメンテーション

本手法では，グラフカット処理に基づいてセグメンテーションを行う．図 1 に示すのは提案するグラフの概念図である．グラフ G は画像 I のスーパーピクセルに対応したノード (n_i) と，物体・背景を表現するソース (s) とシンク (t) と呼ばれるターミナルから構成される．加えて，隣接エッジ (n-link) と呼ばれる隣接スーパーピクセルとの関係を表するエッジと，ターミナルエッジ (t-link) と呼ばれるソースおよびシンクとの関係を表する 2 種類のエッジを作成する．

まず，隣接ターミナルの関係を表する n-link のコストとして，隣接ターミナル間の輝度，色，輝度勾配の類似度を与える．次に，t-link(s, n_i) に対して，前節で求めた顕著度 R_f を物体のコストとして与える．また，t-link(n_i, t) に対して，画像の中心とスーパーピクセルの重心間の距離を背景のコストとして与える．画像空間での距離はスーパーピクセル k' の重心の座標を (x_g, y_g) としたとき，以下の式で定義される．

$$w(k') = \frac{x_g}{dist_{\min}} (x_g \leq y_g) \quad (2)$$

$$w(k') = \frac{y_g}{dist_{\min}} (y_g < x_g) \quad (3)$$

ここで， $dist_{\min}$ は画像の中心から画像の端までの最短距離を表し，縦 H 画素，横 W 画素の画像において $H > W$ のとき $dist_{\min} = W/2$ ， $H \leq W$ のとき $dist_{\min} = H/2$ となる．しかし， $w(k')$ を t-link の背景コストとして直接用いると，物体コストと背景コストのスケール差が問題となるため， R_f の最大値と最小値により $w(k')$ の正規化を行い，これを背景コスト R_b とする．

このように構築したグラフに対して，min-cut/max-flow algorithm[1] を用いてグラフカット処理を行う．

3 評価実験

提案手法により求めた画像セグメンテーション結果に対して評価を行う。評価には人手によって作成された正解画像である GT (Ground Truth) 画像を用意し、これと比較することでどの程度正しく物体を抽出できているかを調べる。評価指標として、Precision (適合率), Recall (再現率), F-measure (F 値) の 3 つを用いる。Precision は物体としてラベル付けされた画素の中で、どれだけ正解画素が含まれているかを表す指標である。Recall は全ての画素の中で、どれだけ正解画素が含まれているかを表す指標である。また、F-measure は Precision と Recall の調和平均であり、総合的な評価値を表す。これらは以下の式により定義される。

$$\text{Precision} = \frac{\text{物体とラベル付けされた中での正解画素数}}{\text{物体とラベル付けされた画素数}} \quad (4)$$

$$\text{Recall} = \frac{\text{正解画素数}}{\text{全画素数}} \quad (5)$$

$$\text{F-measure} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (6)$$

本実験では、実験画像 1000 枚と、それらに対して作成された GT 画像 [2] を用いる。また、提案手法の前処理として用いた SLIC 法のスーパーピクセル分割数は $k = 400$ 、顕著性マップを生成する際に使用するヒストグラムのビン数については $l' = 20$, $a' = 50$, $b' = 50$, $\theta' = 9$ とした。

t-link の物体コストおよび n-link の類似度の算出に用いる特徴量を変更し、各特徴量の有効性を検証する。実験に用いたのは全ての特徴量を結合した式 (1) および輝度、色差、輝度勾配である。画像 1000 枚に対して実験を行い、求めた Precision, Recall, F-measure を表 1 に示す。また、全ての特徴量を用いた前景抽出結果を図 2 に示す。ここで、n-link のバランス変数 λ は t-link に対する n-link の重みを決定する変数であり、各特徴量で最も結果が良いものを実験的に求めた。

実験結果より、F-measure は特徴量を全て用いたものが最も高くなったが、Precision と Recall は輝度を特徴量として用いたものが一番優れていることが分かる。これは、全ての特徴量を用いた顕著性マップでは各特徴量を単純に平均化しており、特定の特徴で顕著物体を高精度に表現できていたとしても、平均化によって顕著度が低くなったことが原因だと考えられる。したがって、画像によって有効な特徴量の選択や重み付けを行うことが今後の課題である。

4 結論

本論文では、色および輝度勾配を用いたグラフカットによる画像セグメンテーション法を提案した。輪郭

表 1: 実験画像 1000 枚に対する平均評価値

	Precision	Recall	F-measure
全て	0.763	0.885	0.794
輝度	0.781	0.889	0.789
色差	0.738	0.875	0.758
輝度勾配	0.598	0.809	0.646

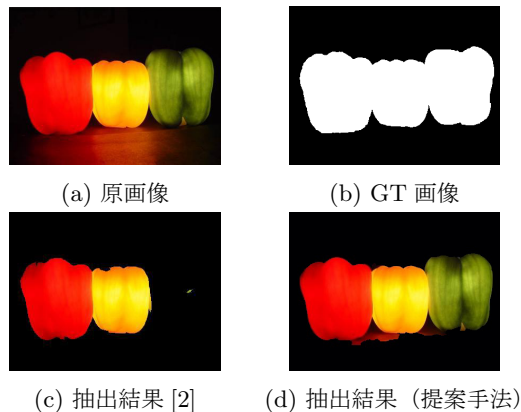


図 2: 前景抽出結果

に沿った画像セグメンテーションの完全自動化を実現するため、顕著性マップとグラフカットを組み合わせた手法を提案した。前処理として原画像をスーパーピクセルと呼ばれる知覚的に意味のある小領域に分割し、物体領域に沿った顕著性マップの生成を実現した。また、顕著物体は画像の中央に位置するものと仮定し、各スーパーピクセルに画像中央からの距離に基づいた重みを与えた顕著性マップを生成した。さらに、グラフカットの初期情報として生成した顕著性マップと各スーパーピクセル間の類似度を与え、完全自動かつ高精度なセグメンテーション結果が得られることを確認した。

参考文献

- [1] Y. Boykov and V. Kolmogorov: "An Experimental Comparison of Min-Cut/Max-Flow Algorithms for Energy Minimization in Vision," *Proc. of PAMI*, vol. 26, no. 9, pp. 1124-1137, 2004
- [2] R. Achanta, S. Hemami, F. Estrada and S. Süsstrunk: "Frequency-tuned Salient Region Detection," *IEEE International Conf. on Computer Vision and Pattern Recognition (CVPR 2009)*, pp. 1597-1604, 2009
- [3] R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua and S. Süsstrunk: "SLIC Superpixels Compared to State-of-the-art Superpixel Methods," *IEEE Trans. on Pattern Anal. and Machine Intelligence*, Vol. 34, No. 11, pp. 2274-2282, 2012