

コセグメンテーションによるセグメンテーションの向上 Improving Segmentation by Graph Cut Based Cosegmentation

信田 萌伽[†] 小西 功記[†]
Moeka Nobuta Kohki Konishi

1. はじめに

セグメンテーションは画像の各画素に対して物体や背景などのラベルを付与する技術である。この技術は、医療画像処理、ロボット制御、工業用検査や衛星画像処理など、詳細な領域分割が必要なアプリケーションで利用される。

画像のコントラストや着目する物体の画像上での大きさにより、セグメンテーションが困難な状況がしばしば起こりうる。この課題に対して、セグメンテーションの精度を向上させる後処理が提案されている。例えば、条件付き確率場を用いた手法[3]などがある。このような後処理の大半は、1 枚の画像情報のみから、特に物体境界での精度を向上させるものである。

コセグメンテーションは、複数の画像に対し画像間で共通する物体をセグメンテーションする技術であり、Rotherら[5]によって提案された。彼らが提案した手法はマルコフ確率場による平滑化項とマッチング項を組み合わせたエネルギー関数を最小化する方法であった。その後、マッチングの改良や機械学習を利用したコセグメンテーションが提案されている。

複数の画像で類似する領域を対応付け、対応した領域のラベルが一致するよう制約を与えることで、1 枚の画像では難しいセグメンテーションが可能になると考えられる。本稿では、セグメンテーション対象の画像に加え、類似した対象が写った画像を入力とし、それらの画像をコセグメンテーションすることでセグメンテーションを向上する方法を提案する。本研究の核となるアイデアは、コセグメンテーションをセグメンテーションの後処理として利用することである。

2. 提案手法

提案手法は、複数の画像とそのサリエンスマップを入力とし、画像間の類似情報を用いてサリエンスマップをアップデートして出力する。入力画像 (Target 画像) 対

し、その画像の類似画像 (Reference 画像) とコセグメンテーションを行うことで Target 画像のセグメンテーション精度を向上させる。コセグメンテーションは Target 画像と Reference 画像のスーパーピクセルとサリエンスマップを利用したグラフカットにより行う。

2.1 前処理

コセグメンテーションを行うための前処理について述べる。まず、Target 画像に対して類似画像検索を行い、Reference 画像を決定する。類似画像検索には画像全体における色ヒストグラムの相関を利用し、相関の高い画像から順に K 枚を抽出する。これにより Target 画像と Reference 画像の組が K 組できる。

それらの画像に対し、セグメンテーションで同一ラベルとなる領域を抽出するため、画像をスーパーピクセルに分割する。ラベルは物体と背景の 2 種類とする。スーパーピクセルへの分割には SEAL[6]を用いる。

2.2 画像の組に対するコセグメンテーション

画像の組に対してコセグメンテーションを行うため、Target 画像と Reference 画像、さらにそれらのスーパーピクセルとサリエンスマップの合計 6 つを用いてグラフを作成する。スーパーピクセルをグラフのノードとし、スーパーピクセル間に次のようにグラフのエッジを設定する (図 1)。Reference 画像で物体と判定される領域と類似する Target 画像内の領域は物体に対応する可能性が高いと考えられる。この考えに基づき、Target 画像と Reference 画像のノード、Reference 画像のノードと Source ノード、Sink ノードの間にエッジを設定する。

Target 画像と Reference 画像間のエッジはスーパーピクセル間距離が指定値以下の場合のみ設定する。エッジを設定する場合、Target 画像と Reference 画像のスーパーピクセル s_T, s_R の間のエッジの重み $w(s_T, s_R)$ は、それぞれの画像の Lab 画像としての平均輝度値 $\overline{I_{Lab}}$ とスーパーピクセル中心座標における HOG 特徴量 (Histograms of Oriented Gradients) [2]を利用して次のように求める：

$$w(s_T, s_R) = \begin{cases} \max(0, 255 - c_{Lab} |\overline{I_{Lab}}(s_T) - \overline{I_{Lab}}(s_R)| \\ -c_{HOG} |HOG(s_T) - HOG(s_R)|) \\ \text{if } M(s_R) > T_R \\ 0 \text{ otherwise} \end{cases}$$

ただし M は Reference 画像のサリエンスマップ、 c_{Lab}, c_{HOG}, T_R は領域類似度を制御するパラメータである。

Source ノード、Sink ノードは Reference 画像のスーパーピクセルに対応するノードのみとエッジで接続する。ノード s_R との間のエッジの重みはサリエンスマップにより

$$w(\text{Source}, s_R) = c_R \overline{M}(s_R),$$

$$w(\text{Sink}, s_R) = \max(0, 255 - c_R \overline{M}(s_R))$$

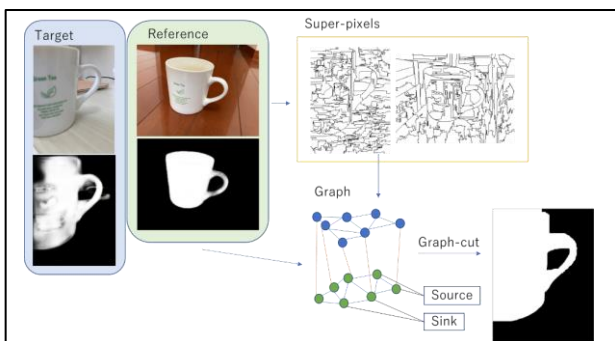


図 1 提案するセグメンテーションの流れ

[†] 株式会社ニコン (Nikon Corporation)

