

大域的なグラフマッチングに基づく 教師無し Cosegmentation

玉那覇 貴紀† 中山 英樹†

†東京大学 大学院情報理工学系研究科

E-mail: {tamanaha, nakayama}@nlab.ci.i.u-tokyo.ac.jp

Abstract

Cosegmentation は入力として与えられた複数の画像から、共通する物体の領域を抽出するタスクである。近年、インターネット上の画像から一般物体認識のための知識ベースを構築する研究が盛んに行われているが、インターネットから収集された画像の多くは雑多な背景を含んでおり、知識ベースを構築する上でのノイズとなっている。Cosegmentation はそれらのノイズを除去するための手法として注目されている。

本研究では、大域的なグラフマッチングに基づく新たな教師無し Cosegmentation の手法を提案する。本手法では、従来手法で用いられている二つのグラフ間の共通部分のみをマッチングする局所的なマッチングではなく、全ての頂点について対応付けを行う大域的なマッチングアルゴリズムを用いることで、背景である可能性が高い領域のみを除去することを可能にしている。性能評価では、従来手法と比較してより高精度な結果が得られることを示した。

1 はじめに

コンピュータビジョンの分野では Never Ending Image Learner[1] 等のようなインターネット上の画像から一般物体認識のための知識ベースを構築する研究が盛んに行われている。インターネット上の画像の収集は主に Google 画像検索などを用いて行われるが、収集された画像の雑多な背景は知識ベースを構築する上でのノイズとなっている。教師無し Cosegmentation は複数の画像から、共通する前景（物体領域）を抽出するタスクであり、この問題を解く手法を用いることで、より高精度な知識ベースの構築が可能となる。

Cosegmentation は Rother ら [2] によって初めて研究が行われ、その後多くの手法が提案されてきた。教師ありの Cosegmentation として Vicente ら [3] によって提案された手法では、問題を全ての入力画像における物体の候補領域間の類似度を最大化するエネルギー最大化問題として定式化しており、Random Forest によ

て物体の候補領域間の類似度を回帰している。Meng ら [4] によって提案された教師無し Cosegmentation の手法では、最短経路問題として Cosegmentation を解くことで共通する物体領域の抽出に成功している。Yu ら [11] が提案した手法はグラフマッチングに基づいており、画像をスーパーピクセルへ変換し、共通部分のみのマッチングを行う局所的なマッチングアルゴリズムによって前景領域を得ている。しかし、局所的なマッチングによって得られた前景の候補領域は、僅かな背景を含んでしまう傾向にある。そこで、本研究では教師無し Cosegmentation のための大域的なグラフマッチングを基にした新たな手法を提案する。共通部分のみではなくグラフ全体に対してマッチングを行う大域的なマッチングアルゴリズムを用いることで、背景である可能性の高い領域を除去することが可能となり、より質の高い前景領域が得られる。マッチングを行っただけでは前景領域の部分グラフが抽出できなくなってしまうが、提案手法では連結成分分解に基づいた方法で前景として妥当な部分グラフを選択する。

2 提案手法

提案手法は大きく分けて次の四つのステップで構成される。

1. 画像のグラフ化
2. グラフマッチング
3. 前景を表現する部分グラフの抽出
4. 前景領域の抽出

それぞれのステップの詳細について以下に述べる。

2.1 画像のグラフ化

画像をグラフで表現する単純な手段として、画像のそれぞれの画素をグラフの頂点とし、隣接する頂点間を辺で接続するという方法が考えられるが、その方法だと各グラフの頂点数が膨大になってしまい、グラフマッチングに時間が掛かり過ぎるという問題がある。そこで、グラフの頂点数を抑えるために画像をスーパーピクセルによって表現する。

画像をスーパーピクセルに変換する手法として SLIC[5]を用いる。Achantaらによって提案された SLIC は K-means をベースにした手法で、単純な K-means によるスーパーピクセル化との相違点として、初期化時にクラスタの中心を色の勾配が緩やかな場所に移動するという点と、各クラスタごとに走査するピクセルの数の制限があることが挙げられる。

スーパーピクセル化された画像をグラフで表現すると、形の情報は失われてしまうため、各スーパーピクセルの形状が均一である方が元の画像を良く表現しているグラフになると考えられる。SLIC でスーパーピクセル化された画像は、他のスーパーピクセル化の手法に比べ、各スーパーピクセルが同じような形状になり易い傾向があるためグラフマッチングに基づいた手法においては有効性が高いと考えられる。

全ての入力画像をスーパーピクセルへ変換し、各スーパーピクセルをそれぞれグラフの頂点とする。頂点間の類似度として、各頂点に対応するスーパーピクセル内に含まれる SIFT 特徴量の平均ベクトル間の類似度を用いる。

図1は iCoseg[9] のデータセットの画像をスーパーピクセルに変換したもの(二段目)とグラフ化したもの(三段目)をそれぞれ表している。ここから更に左右のグラフ間でマッチングを行い、左右の画像間で共通している物体である熊の前景を抽出することが最終的な目標となる。

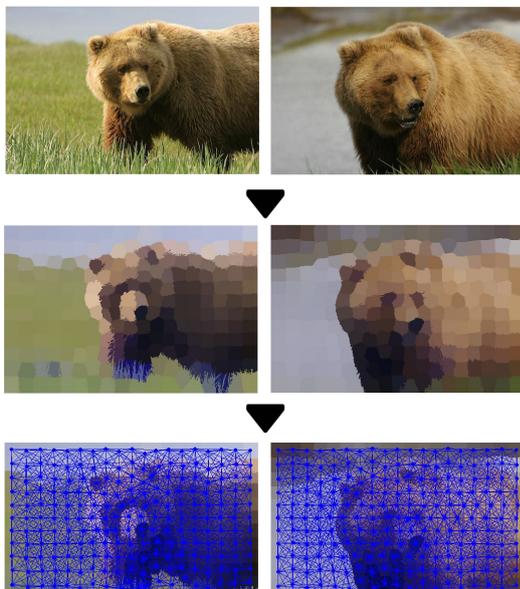


図1 画像グラフ化するまでの流れ。一段目が元画像、二段目がスーパーピクセルに変換した画像、三段目がグラフ化したものをそれぞれ表している。

2.2 グラフマッチング

グラフ化した画像のマッチングには SPINAL[6]を用いる。SPINALは Aladağらによって提案された、タンパク質間相互作用ネットワークのマッチングを目的としたアルゴリズムで、バイオインフォマティクスの分野において、IsoRank[7], MI-GRAAL[8]など他の手法より高いスケーラビリティを持つことが示されている。SPINALはグラフの全ての頂点についてマッチングを求める大域的なマッチングアルゴリズムである。

SPINALは入力グラフを $G_1 = (V_1, E_1)$, $G_2 = (V_2, E_2)$ と置いたとき、以下の目的関数 *global network alignment score* を最大化するアラインメントネットワーク $A_{12} = (V_{12}, E_{12})$ を求める。

$$GNAS(A_{12}) = \alpha|E_{12}| + (1 - \alpha) \sum_{\forall \langle u_i, v_j \rangle} seq(u_i, v_j) \quad (1)$$

ここで、 $|E_{12}|$ は E_{12} のサイズ、 $seq(u_i, v_j)$ は二つの頂点間の類似度を表し、 $u_i \in V_1, v_j \in V_2$ である。 V_{12} は頂点对 $\langle u_i, v_j \rangle$ の集合で、頂点間のマッチングを表しており、任意の二つの頂点对 $\langle u_i, v_j \rangle \in V_{12}$ と $\langle u'_i, v'_j \rangle \in V_{12}$ において $u_i \neq u'_i$ かつ $v_j \neq v'_j$ である。アラインメントネットワークの辺は $(u_i, u'_i) \in E_1$ かつ $(v_j, v'_j) \in E_2$ のとき、 $(\langle u_i, v_j \rangle, \langle u'_i, v'_j \rangle) \in E_{12}$ として存在する。

SPINALは頂点 u_i と v_j 間の近傍の類似度を計算する Coarse-grained フェーズと、アラインメントネットワークを構築する Fine-grained フェーズの二つのフェーズで構成される。

Coarse-grained フェーズでは各頂点对 $\langle u_i, v_j \rangle$ について、 u_i の近傍と v_j の近傍をそれぞれ独立集合とした重み付き完全二部グラフ \mathcal{NBG} を構築し、その最大重みマッチングを基に各頂点の近傍の類似度を計算する。図2に Coarse-grained フェーズの処理を示す。

Fine-grained フェーズでは V_{12} に含まれていない頂点对のうち、近傍間の類似度が最大のものについて \mathcal{NBG} を構築し、その最大重みマッチングを求め、マッチングに含まれていない辺の端点の集合についてさらに \mathcal{NBG} を構築して同様の処理を繰り返す Seed-and-Extend 戦略に基づいた手法でアラインメントネットワークを求める。図3に Fine-grained フェーズの処理を示す。

2.3 前景を表現する部分グラフの抽出

SPINALによるグラフマッチングで得られたアラインメントネットワークから、入力された複数の画像間で共通する前景を表現するような部分グラフを抽出する。アラインメントネットワークを深さ優先探索によって連結成分分解し、それぞれの連結成分を SPINAL の目的関数 (1) で評価する。その結果得られた最もスコアの大きい連結成分を、前景を表現する部分グラフと見なす。

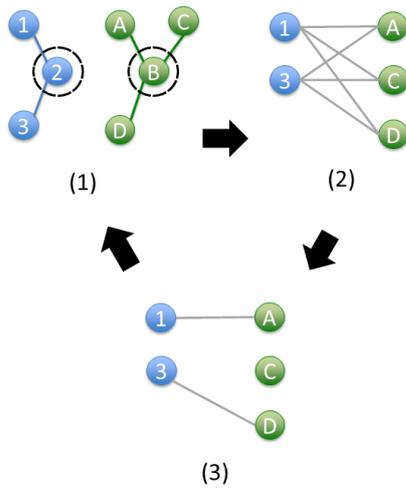


図2 Coarse-grained フェーズの処理. 各グラフから頂点の一つずつ選択し (1), それぞれの近傍から完全二部グラフを構築する (2). 最大重みマッチングを計算し (3), 得られたマッチングに基づいて近傍間の類似度を求める.

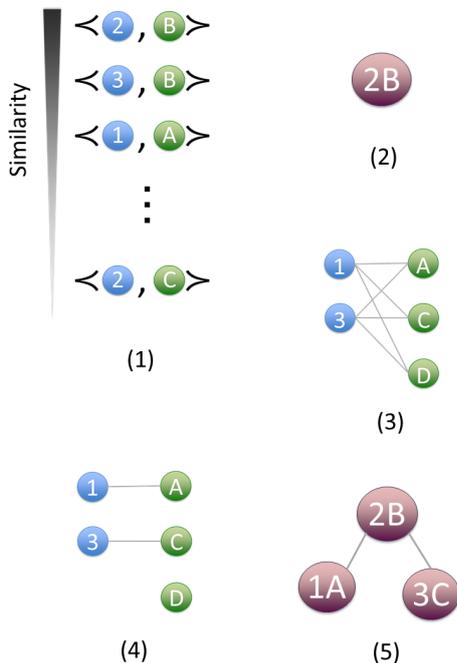


図3 Fine-grained フェーズの処理. Coarse-grained フェーズで計算した近傍間の類似度が高い順に頂点对をソートし (1), 最も類似度の高い頂点对をアラインメントネットワークに追加する (2). 追加した頂点に対応する二つの頂点の近傍から完全二部グラフを構築し (3), 最大重みマッチングを行う (4). マッチングに対応する頂点对をアラインメントネットワークに追加し, 解を拡張する (5).

2.4 前景領域の抽出

前景の抽出には Rother らの Grabcut[10] を用いる. Grabcut は前景または背景の一部を入力として与え, セグメンテーションを行うアルゴリズムである. 提案手法では 2.3 で得られた前景を表現する部分グラフの頂点を全て含むような領域を, 前景の候補領域として与えることで前景の抽出を行う. これにより, グラフマッチングの結果得られた荒い前景が精緻化される.

3 実験結果

3.1 従来手法との比較

本分野で広く用いられている iCoseg[9] のデータセットを用いて性能評価を行う. 表 1 に実験結果を示す. それぞれの性能は, 各手法によって得られたマスクと Ground Truth の一致率 (全画素に対する true positive 及び true negative の割合) によって評価している.

表 1 iCoseg データセットにおける実験結果 (%)

Grabcut	従来手法 [11]	提案手法
79.3	83.0	84.1

表 1 の 1 列目は各々の画像に Grabcut のみを適用した結果を, 2 列目はスーパーピクセルマッチングに基づく従来手法の結果を, 3 列目は提案手法の結果をそれぞれ表している. 従来手法では複数の画像間から共通部分のマッチングを行なっているが, 前景部分だけではなく背景部分まで共通部分としてマッチングが行われる場合が多々ある. 一方, 提案手法では全ての頂点間についてマッチングを行い, 前景らしくない連結成分を取り除いているため, 従来手法より高い精度が得られたと考えられる.

提案手法によって抽出された前景領域の画像と元画像を並べたものを図 4 に示す. この結果から, 雑多な背景を含む画像でも良い前景領域が得られていることが確認できる.

3.2 前景を表現する部分グラフの抽出方法の比較

提案手法では, グラフマッチングの結果得られたアラインメントネットワークから前景を表現する部分グラフを抽出する方法として, アラインメントネットワークを連結成分分解し, それぞれの連結成分を式 (1) によって評価し, 最も値の大きな連結成分を部分グラフとしている. しかし, 最終的に得られた部分グラフが物体よりも非常に大きくなるケースがしばしば見られるため, 閾値処理によりグラフ間で類似度の低い頂点をあらかじめ取り除き, 残った頂点の中から最もスコアの高い連結成分を抽出する方法を検討した. 閾値処理を行わない場合との比較実験の結果を以下に示す. 3.1 の実験と同様に iCoseg のデータセットを用いている.

表2 部分グラフを抽出する方法の比較実験 (%)

閾値処理無し	閾値処理 1	閾値処理 2
84.1	75.8	76.0

表2の閾値処理1では、頂点を類似度の高い順に並べたときに、隣り合う値の差が最も大きくなるところで閾値処理をしている。閾値処理2では、大津の二値化法により、二つに分割された頂点集合のそれぞれの集合内の類似度の分散が小さく、かつ集合間の分散が大きくなるような閾値で頂点を除去している。ナイーブな手法である閾値処理1に比べて閾値処理2のほうが良い結果が得られているが、閾値処理を行わない場合が最も良い結果となった。閾値処理を行うと前景部分の頂点もわずかに除去されてしまうことがあり、その場合前景部分の連結性が失われてしまう。そのため、前景として抽出された部分グラフが非常に小さくなってしまい、精度が落ちてしまったと考えられる。

4 結論

本研究では複数の画像から共通する物体領域を抽出する Cosegmentation というタスクにおける、大域的なグラフマッチングアルゴリズムに基づいた新たな手法を提案した。従来のグラフ間の共通部分のみのマッチングを行う手法に対し、グラフ全体に対してマッチングを行う大域的なマッチングアルゴリズムを適用することで、背景領域の除去に焦点をあてた。iCoseg のデータセットにおける実験では、従来手法よりも良い精度の結果が得られることを示した。また、マッチングの結果から前景を表現する部分グラフを抽出する手法について、閾値処理により類似度の低い頂点をあらかじめ取り除いた場合の実験を行い、閾値処理を行わないほうが良い結果が得られることについて述べた。

画像認識のための知識ベースを構築するシステムに対して、Cosegmentation のアルゴリズムを適用するためには、共通の物体が存在しないようなノイズ画像を取り除く必要がある。そのため、今後の課題としてノイズ画像を除去する手法の検討などが挙げられる。

参考文献

- [1] X. Chen et al. NEIL: Extracting Visual Knowledge from Web Data. In Proc. of IEEE ICCV, 2013.
- [2] C. Rother et al. Cosegmentation of Image Pairs by Histogram Matching - Incorporating a Global Constraint into MRFs. In Proc. of IEEE CVPR, 2006.
- [3] S. Vicente et al. Object Cosegmentation. In Proc. of IEEE CVPR, 2011.
- [4] F. Meng et al. Object Co-Segmentation Based on Shortest Path Algorithm and Saliency Model. ACM Trans. Multimedia, vol. 14, no. 5, pp. 1429-1441, 2012.
- [5] R. Achanta, A. Shaji, K. Smith, A. Luchi, P. Fua, and S. Susstrunk. SLIC: Superpixels Compared to State-of-the-art Superpixel Methods. IEEE Trans. PAMI, vol. 34, no. 11, pp. 2274-2281, 2012.
- [6] A. E. Aladağ et al. SPINAL: scalable protein interaction network alignment. Bioinformatics, vol. 29, no. 7, pp. 917-924, 2013.
- [7] R. Singh et al. Global alignment of multiple protein interaction networks with application to functional orthology detection. Proc. Natl. Acad. Sci. USA, vol. 105, no. 35, pp. 12763-12768, 2008.
- [8] O. Kuchaiev et al. Integrative network alignment reveals large regions of global network similarity in yeast and human. Bioinformatics, vol. 27, no. 10, pp. 1390-1396, 2011.
- [9] D. Batra et al. iCoseg: Interactive co-segmentation with intelligent scribble guidance. In Proc. of IEEE CVPR, 2010.
- [10] C. Rother et al. "GrabCut" - Interactive foreground extraction using iterated graph cuts. In ACM Trans. Graphics, vol. 23, no. 3, pp. 309-314, 2004.
- [11] H. Yu et al. Unsupervised cosegmentation based on superpixel matching and Fastgrabcut. In Proc. of IEEE ICME, 2014.

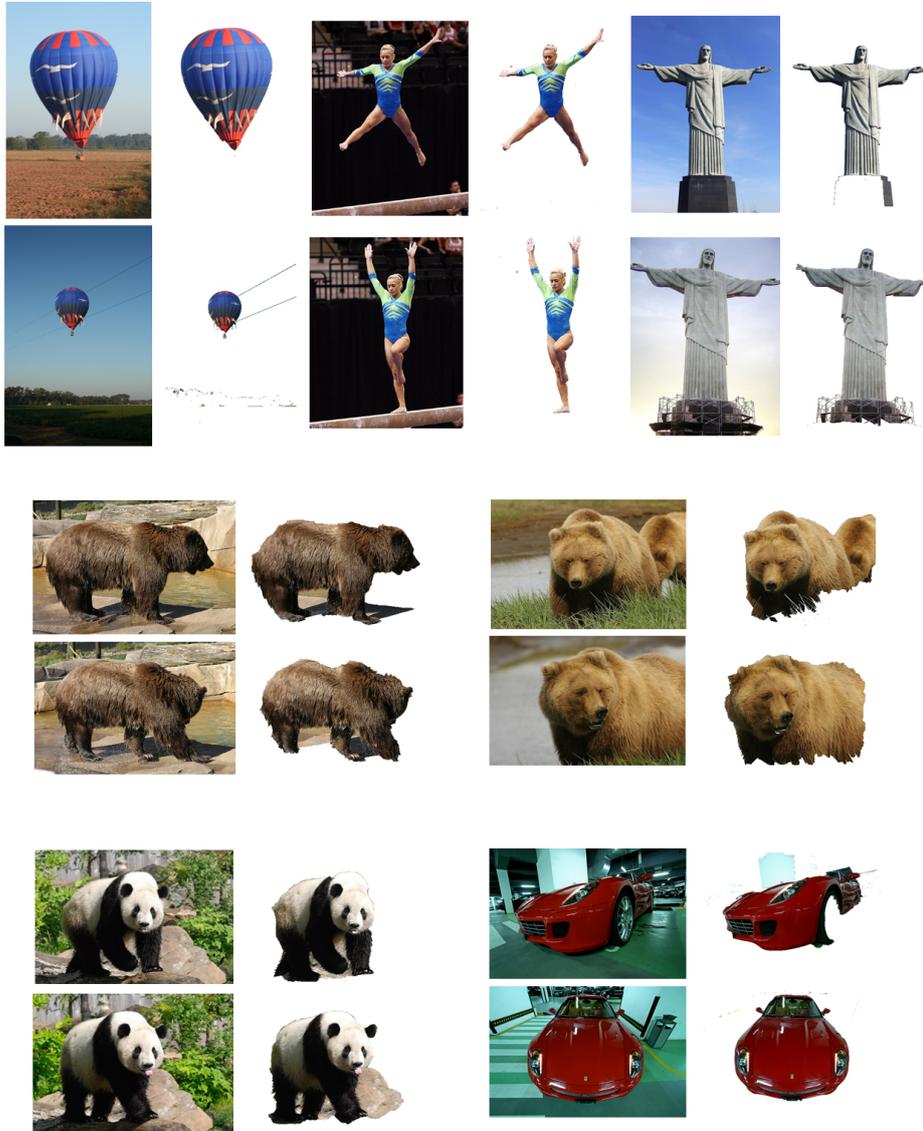


図4 提案手法によって抽出された物体領域. 各クラスの画像が四枚組みになっており, 左の二枚が元画像, 右の二枚が抽出された共通する物体の領域をそれぞれ表している.