

# SIFT と Graph Cuts を用いた物体認識及びセグメンテーション

## OBJECT RECOGNITION AND SEGMENTATION USING SIFT AND GRAPH CUTS

須賀晃<sup>1</sup>  
Akira Suga

福田恵太<sup>2</sup>  
Keita Fukuda

滝口哲也, 有木康雄<sup>3</sup>  
Tetsuya Takiguchi, Yasuo Ariki

<sup>1</sup>神戸大学工学部情報知能工学科  
Department of Computer Sciences and Systems Engineering,  
Kobe University

<sup>2</sup>神戸大学大学院工学研究科  
Graduate School of Engineering,  
Kobe University

<sup>3</sup>神戸大学自然科学系先端融合研究環  
Organization of Advanced Science and Technology,  
Kobe University

### 1. はじめに

物体認識は、計算機の認知能力実現のための重要な研究分野の一つであり、ロボット視覚や HMD への応用等も期待されている。また近年ではデータの大容量化によって人手によるデータの分類や検索が困難になり、計算機による物体認識の重要性がますます高まってきている。本稿では、Scale-Invariant Feature Transform (SIFT) [1] と Graph Cuts (GC) [2] を用いた物体の認識・切り出し法を提案する。本手法では、物体の回転や照明・スケール変化等に頑健な特徴量である SIFT を用いて物体を認識し、物体上に抽出されたその特徴点を物体の seed として用いることで GC を行い物体領域のセグメンテーションを行う。GC は従来手動で seed を与える必要があったが、本手法では自動的にセグメンテーションを行うことが可能となった。

### 2. 認識とセグメンテーション

最初に入力画像に対して SIFT 特徴点を抽出し、予め物体ごとに作成しておいたモデル画像の全特徴点とマッチングを行い、物体部分の対応点を得る。物体の認識には次のような投票による判定処理を行う。まず、モデル画像の各特徴点  $(x, y)$  において、あらかじめ物体の中心点からの位置ベクトル  $(\Delta x, \Delta y)$  を計算しておく。この位置ベクトルを用いて、入力画像上で対応の取れた全特徴点に対して中心候補点  $(X, Y)$  を以下の式により計算し、その候補点に投票する。

$$\begin{cases} X = x_{input} + \frac{\sigma_{input}}{\sigma_{model}} \times \sqrt{\Delta x^2 + \Delta y^2} \times \cos(\theta + \theta_{model} - \theta_{input}) \\ Y = y_{input} + \frac{\sigma_{input}}{\sigma_{model}} \times \sqrt{\Delta x^2 + \Delta y^2} \times \sin(\theta + \theta_{model} - \theta_{input}) \end{cases} \quad (1)$$

$$\text{但し、} \theta = \arctan\left(\frac{\Delta y}{\Delta x}\right) \quad (2)$$

ここで、 $\sigma$  は特徴点のスケール、 $\theta$  は輝度勾配方向である。もし入力画像中に対象物体が存在し、且つ、その特徴点が正確に物体上に取られているならば、その中心候補点は同じ位置に集まる [3]。そこで、この投票された候補点をウォード法でクラスタリングし閾値以上の投票が得られた場合、対象物体が存在すると判定する。

また、閾値以上の中心候補点に投票した特徴点は高い精度で物体上に取れるため、これを物体の seed として用いる。背景の seed は物体のモデル画像をアフィン変換したものの外側領域とする。

実際に自動で生成された seed を図 1 に示す。赤い点が物体の seed、青い領域が背景の seed である。この seed を用いて GC を行うことで物体領域を切り出す。

### 3. 実験と考察

本実験では、本やサッカーボールなど日常生活でよく見られる物体モデル 20 個を作成し、50 枚のテスト画像に対して切り出し実験を行った。モデルは 1 物体につき異なる角度から撮影した数枚の背景黒の画像を用いた。実験の結果、認識率は 82.67%、誤認識率は 0.00% となった。切り出しの成功例を図 2 に示す。オクルージョンを含む複雑な背景でも物体をそれぞれ識別し、切り出すことができた。図中の数字は正解マスクに対するセグメントのエラー率を示している。テスト画像全体におけるエラー率も 5.87% という良い結果が得られた。

しかし、SIFT で取れる特徴量もともと少ない物体が閾値以上の得票数を得ることが出来ず認識出来なかったり、seed に偏りがあるものは上手くセグメンテーションが出来ない場合もあった。今後は物体に応じて閾値を変更したり他の特徴も用いるなどして、このような物体に対しても正確なセグメンテーションが出来るように検討していく予定である。

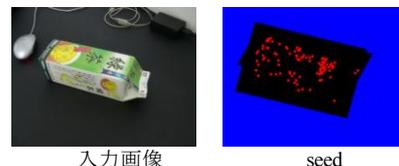


図 1 自動で作成された seed

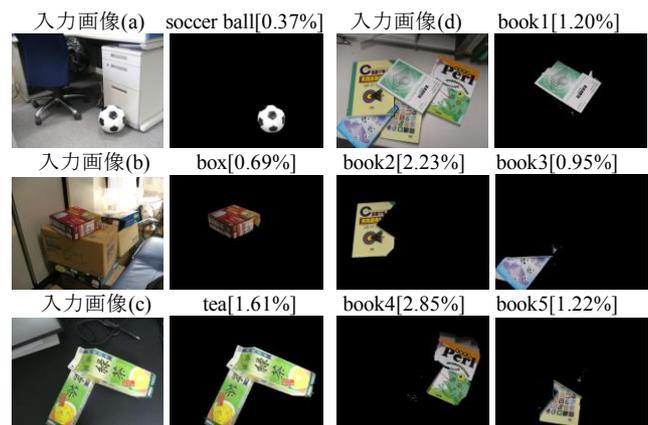


図 2 セグメンテーション結果

### 4. 参考文献

- [1] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," 2004.
- [2] Y. Boykov, M.P.Jolly, "Interactive graph cuts for optimal boundary & region segmentation of objects in N-D images,"
- [3] 高木雅成, 藤吉弘亘, "SIFT 特徴量を用いた交通道路標識認識," 画像センシングシンポジウム, LD2-06, 2007.