

# 移動 Kinect 画像中からの移動物体の抽出と 三次元形状のセグメンテーションに関する検討

— Probabilistic Boundary Edge Map と Supervoxel を組み合わせた手法の提案 —

Study of Extracting Moving Objects from Dynamic Kinect Images and Segmenting the Moving and Still Objects' Reconstructed 3D Shape

-- Proposing a combined method Probabilistic Boundary Edge Map with supervoxel --

立松 直倫† Naotomo Tatematsu  
大谷 淳† Ohya Jun  
Larry Davis‡ Larry Davis

## 1. まえがき

筆者らは、複数の移動物体を含むシーンを移動ステレオカメラにより撮像することにより獲得されるステレオ動画画像から、シーン内の特徴点を物体毎に識別し、識別した物体毎の三次元構造の復元と、ステレオカメラの姿勢推定が可能な Temporal Modified-RANSAC (以下, TMR) を用いた手法[1]及び TMR の特徴点の移動物体毎の分類結果を入力とした三次元点群データのセグメンテーションの検討を進めている。TMR ベースの手法はカラー画像と距離画像の時系列画像から個々の移動物体と背景のフローを分類することが可能であるが、分類結果は特徴点のみの疎なデータに限られるという問題が存在する。筆者らは[1]において、Probabilistic Boundary Edge Map を用いた Fixation ベースの手法を提案したが、セグメンテーション処理がフレーム毎に行われ、時系列での情報が有効に利用できないという問題が存在した。そこで、本稿では Probabilistic Boundary Edge Map とスーパーボクセルを組み合わせた方法を提案する。本手法を複数の移動物体を含む Kinect を用いて撮像したカラー画像と距離画像の時系列画像に対し適用する実験を行い、本手法の有効性の見通しを確認した。

## 2. 提案の概要

提案手法は、1) TMR の適用、2) スーパーボクセルの生成、3) セグメンテーション処理の3つのプロセスで構成される。TMR の適用のプロセスでは、距離画像と特徴点の追跡結果を元に生成した三次元フローを各移動物体及び背景にクラスタリングし、各フレームでの物体の動きの推定を行う。スーパーボクセルの生成のプロセスでは、時系列画像にまたがり類似した画素を結合したボクセルを生成する。セグメンテーション処理のプロセスでは、生成したスーパーボクセルを物体毎にセグメンテーションし、物体領域の抽出を行う。

## 3. スーパーボクセルの生成

提案手法では、初めに SLIC[2]を用いて初期スーパーボクセルを生成する。SLIC が生成するスーパーボクセルは、図 2(a)に示すように過分割であり、この結果をそのまま

セグメンテーションに用いると精度が低下する。そのため、[3]の複数のラベルを用いたグラフカット処理を用いて初期スーパーボクセルの統合を行う。

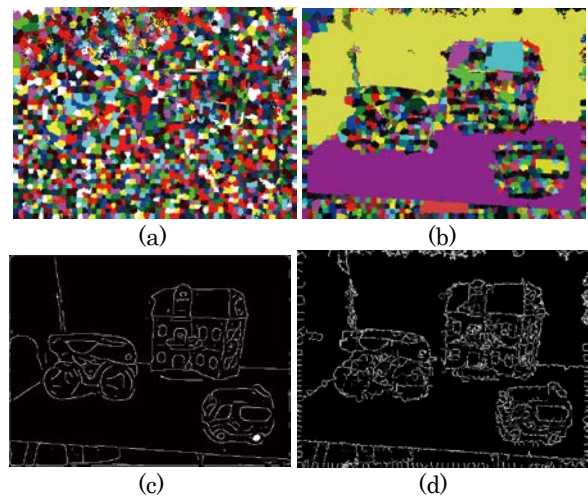


Fig.2. スーパーボクセル

グラフの平滑化項は、[1]の画素が物体の境界である尤度を示す Object probability boundary map を元に定義する。図 2(c)は Object probability boundary map の例であるが、この尤度が高い画素とスーパーボクセルのエッジは完全には重ならない。そこで、スーパーボクセルのエッジの画素と Object probability boundary map の各画素に対し、画像平面での位置(u,v)とエッジの外側・内側の色情報(r,g,b)の8次元のベクトルを定義し、各ベクトルの最近傍の点を求め投票を行い、スーパーボクセルのエッジの物体境界尤度を算出し平滑項のエネルギーとして用いる(図 2(d))。データ項は、各スーパーボクセルに対し、スーパーボクセルとラベルが同じ場合には高いエネルギーを設定し、その他のラベルに対しては低いエネルギーを設定する。構築したグラフに対して[3]のアルゴリズムを適用し、スーパーボクセルを統合する(図 2(b))。

## 3. セグメンテーション処理

各物体処理の抽出は[3]のマルチラベルのグラフカット処理を適用して行う。提案手法では TMR により、移動物体の数、及び物体の各フレームでの回転・併進移動が推定され

† 早稲田大学国際情報通信研究科 GITS

‡ The University of Maryland

ている。グラフのデータ項の生成は、この推定した移動情報を元に生成した類似度画像を用いて行う。

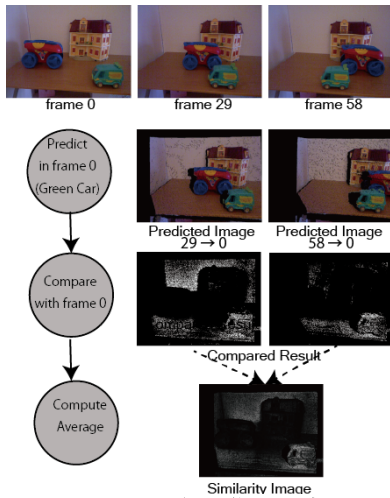


Fig.4. 類似画像の生成

図4は、0フレーム目の右側の緑の車に対する類似画像の生成例である。最初に0フレーム以外のフレームの三次元点群データに対して、緑の車の回転・併進移動の情報を用いてフレーム0での位置に射影する。そして、各フレームの射影した結果と、フレーム0の実際のデータとの各画素の類似度を計算する。最後に画素毎に全フレームの平均を計算し類似画像を生成する。

同じ処理を全フレーム及び全物体に対して行い、スーパーボクセル単位で物体毎の類似度の平均を求める。この平均値をデータ項の値として用いる。ただし、スーパーボクセルが特徴点を含む場合には、対象の類似度が低い場合でも、物体に属するスーパーボクセルと考えられるため、特徴点の属するラベルに対して定数値の小さいエネルギーを設定する(図5右の画素値が255の画素)。また、類似度の差が小さいスーパーボクセルの場合には、全ラベルに対して同じエネルギーを設定する(図5右の黒い画素)。



Fig.5. データ項

平滑化項は、オプティカルフロー画像と距離画像に対して、円形のフィルタを適用して内側と外側の画素の平均値の差を求め、式(1)に代入した結果を比較し、より大きい値を用いる。 $\beta_1$ ,  $\beta_2$ はスケールパラメータである。

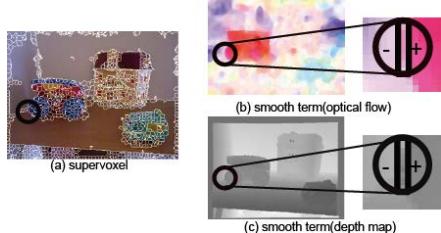


Fig.6. 平滑化項

$$g(x) = \frac{1}{1 + e^{-\beta_1(\text{diff}(\text{flowordepth}) - \beta_2)}} \quad (1)$$

生成したグラフに対して[3]のアルゴリズムを適用して、エネルギーが最小になるように切断し、各ラベルに対応した物体領域を抽出する。

## 5. 実験結果

Kinect(c)で撮像した時系列画像(図4(a)(d)(g))に対して、提案手法を適用した結果を示す。左列の各図はTMRの分類結果を示し、中列の図はデータ項の状態を示す。左列がセグメンテーション結果である。各物体が背景からセグメンテーションされている。

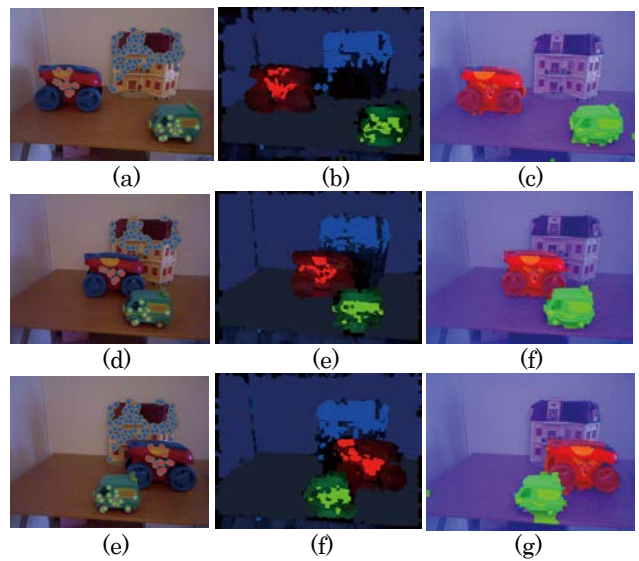


Fig.7. 実験結果

## 6. Conclusion

筆者が今まで研究していた手法のセグメンテーション処理にスーパーボクセルベースのマルチラベルのグラフカットを適用することで移動物体のセグメンテーション精度の向上を実現した。今後はより複雑なケースの実験を行い、より検討を進める予定である。

## Reference

- [1] Tatemasu, N., Ohya, J., Larry D: Detection and Segmentation of Moving Objects from Dynamic RGB and Depth Images. In: International Conference on Virtual Reality and Visualization 2012
- [2] Achanta, R.; Shaji, A.; Smith, K.; Lucchi, A.; Fua, P.; Susstrunk, S., "SLIC Superpixels Compared to State-of-the-Art Superpixel Methods," Pattern Analysis and Machine Intelligence, IEEE Transactions on, vol.34, no.11, pp.2274,2282, Nov. 2012
- [3] Delong, A.; Osokin, A.; Isack, H.N.; Boykov, Y., "Fast approximate energy minimization with label costs," Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on, vol., no., pp.2173,2180, 13-18 June 2010