

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 を用いた Fixation ベースのセグメンテーション手法の提案 —

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

-- Proposing a Fixation-based Segmentation Method Utilizing Probabilistic Boundary Edge Map --

立松 直倫† Naotomo Tatematsu 大谷 淳† Jun Ohya 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]では、特徴点の凸包領域を前景としてラベリングした GraphCut を用いた手法を提案したが、物体の形状が複雑な場合に、凸包領域では前景としてラベリングした画素に、背景の画素が含まれ、誤学習を行うという問題が存在した。そこで、本稿では Ajay らが提案するエッジが物体と背景および他の物体との境界である確率を示す Probabilistic Boundary Edge Map を用いた Fixation ベースの手法[2]をセグメンテーション処理に用いた方法を提案する。Ajay らの手法では、GraphCut のように領域ではなく、ユーザが指定した復元対象の物体の 1 画素をセグメンテーション処理の入力情報として用いるため、物体の形状に関係なくセグメンテーションを行うことが可能である。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、本手法の有効性を見直しを確認した。

2. Proposal Overview

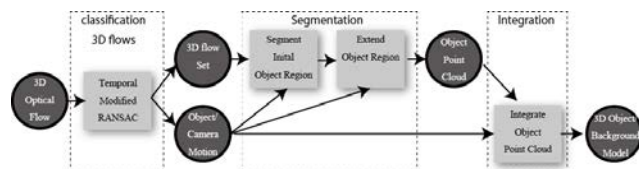


図1. 提案手法の概要

提案手法では、カラー画像の特徴点の追跡結果と距離画像を元に生成した三次元フローに対して、TMR を適用し

† 早稲田大学国際情報通信研究科 GITS
‡ The University of Maryland

て個々の移動物体及び背景の各フレームの特徴点と、直前のフレームからのカメラ・各物体の回転・併進移動を推定する。TMR の分類結果は特徴点のみの疎なデータであり、個々の物体の三次元モデルを復元する為には、密なデータが必要となる。そのため、Ajay らの手法を複数の特徴点に対して適用し、移動物体の初期領域を推定する。この初期領域には、対象の物体の一部の領域において特徴点が存在せず、かつ追跡した特徴点の間にエッジが存在する場合、物体領域に含まれないケースが存在する。提案手法では、こういった領域を物体領域にマージする為には、初期領域の拡張処理を行う。拡張処理の結果得られた物体の領域とカラー画像と距離画像から、対象のフレームの物体の三次元点群データを生成し、TMR で推定したカメラ・各物体の回転・併進移動データを用いて、初期フレームの位置に重ね合わせ、物体の三次元モデルを生成する。

3. 物体の初期領域のセグメンテーション

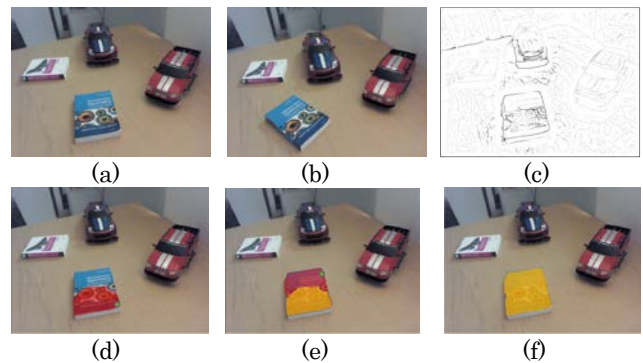


図2. 物体の初期領域のセグメンテーション

提案手法のセグメンテーション処理は、Ajay の手法を用いて行う。初めに対象フレームと隣接フレーム(図 2(a) (b))から作成したオプティカルフローと距離画像からエッジが物体と背景や他の物体の境界である確率を示す probabilistic boundary map を生成する(図 2(c))。TMR の分類結果の特徴点の一つを選択して、Ajay の手法を適用して特徴点に対応した領域(図 2(d)(赤色部分))を取得する。取得した領域に全ての特徴点が含まれているかどうかチェックを行い、含まれていない特徴点が存在した場合には、その特徴点を用いて Ajay の手法を適用し、新しい領域を追加する(図 2(e))。全ての特徴点新しい領域に含まれるまでこの処理を繰り返して、物体の初期領域(図 2(f))を得る。

4. 初期領域拡張処理

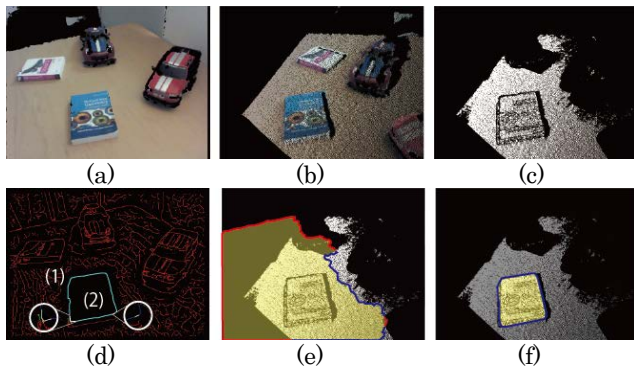


Fig.3. 物体の初期領域の拡張処理

(図 2(f)) の例では、物体の初期領域には、本の手前の部分のように特徴点の存在しない領域が含まれていない。この本の手前の面との境界のエッジを削除して再度セグメンテーション処理を行えば、手前の面も物体領域として拡張することが可能である。この拡張処理には、拡張処理を行うために削除が必要なエッジを選択する仕組みと拡張した領域が対象の物体領域であるかを評価する仕組みが必要である。削除が必要なエッジは、初期領域のエッジに他のエッジが接続する又は、近接する部分間のエッジとなる。初期エッジの周囲のエッジを検索し、削除対象候補のリストを作成する。また、拡張した領域の評価は、類似画像という特定の画素が対象の物体であるかどうかを示す画像を用いて行う。初めにフレーム 1 の三次元点群データを画像平面に射影した画像(a)と、フレーム 2 の三次元点群データを TMR で推定した回転行列・併進行列を用いてフレーム 1 の位置に移動させ、画像平面に射影した画像(b)を生成する。そして、生成した画像の各画素において、距離値と画素値の類似度を示す類似画像(Sim)(図 3(c))を生成する。類似画像において画素値が低い部分は他の物体または、背景の領域と考えられるため、エッジの内側・外側の両方の画素値が低い場合、エッジは対象物体のエッジではないと考えられる。ただし、この類似画像において、オクルージョンが発生した領域については、比較対象の画素が存在しないため、常に他の物体として評価されるという問題が存在する。提案手法ではこの問題を解決する為に、オクルージョンが発生した領域の画素値を 255 としたグレースケール画像(Oc)を作成し、エッジの周辺の画素値が高く、オクルージョンが発生した領域と考えられる場合には、仮に類似画像の画素値が低くとも、ペナルティを与えない。また、セグメンテーションした結果の境界の画素の近傍に物体エッジが存在しない場合も、ペナルティを与えない。以上の内容を数式で表現した式が(1)となる。式(1)の F_{in}/F_{out} の関数は、引数で渡された画像における、エッジから距離が r 以内の内側/外側の画素値の合計を示す。

$$EdgePenalty = \begin{cases} 1 & \text{if } ((F_{in}(Sim) < \alpha_1 \cap F_{out}(Sim) < \alpha_1) \\ & \cap (F_{in}(Oc) < \alpha_2 \cap F_{out}(Oc) < \alpha_2)) \text{ (1)} \\ & \cup HasNearEdge = false \\ 0 & \text{otherwise} \end{cases}$$

作成した削除対象候補のリストからエッジを一つ選択して、そのエッジを削除した状態でセグメンテーション処理

を行い、その結果を評価する。エッジのペナルティ値の合計が事前に定めた閾値以下の場合には、初期領域にマージする。全ての削除対象のエッジの削除・拡張・評価の処理が終了するまで、拡張処理を続ける。

図 3 の例では、(1)のエッジを削除して拡張処理を行った結果(図 3(e))では、ペナルティが発生したエッジ(赤色)が閾値以上存在するため、この領域の統合は行わない。(2)のエッジを削除して拡張処理を行った結果(図 3(f))はペナルティが発生したエッジが存在しないため、本の手前側の面は、物体領域として追加される。

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。中列の各図では、初期領域を黄色/水色で、拡張処理で追加された領域を赤色/青色で示す。特徴点が存在しない左の車の背の部分や車輪の部分が拡張処理で物体領域に統合されている。右列の図は各物体を三次元復元した結果である。各物体が背景からセグメンテーションされている。

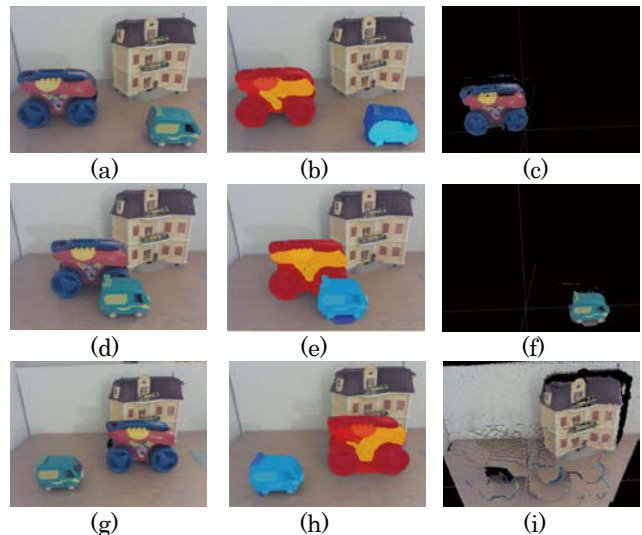


Fig.4. 実験結果

6. Conclusion

本研究では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 Fixation ベースのセグメンテーション処理を行う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] N. Tatsumatsu, J. Ohya : Accurate dense 3d reconstruction of moving and still objects from dynamic stereo sequences based on temporal modified-ransac and feature-cut. In: IS&T-SPIE Electronic Imaging 2012. Volume 8301. (2012) 05-15 10
- [2] Mishra, A.K., Aloimonos, Y.: Visual segmentation of "simple" objects for robots. In: Robotics Science and Systems conference. (2011)