

移動ステレオカメラ画像中からの移動物体抽出と

三次元データのセグメンテーションの検討

- SIFT 特徴点を元にした前景・背景のエネルギー関数を組み込んだ手法の提案 -

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

-- Proposing a SIFT Feature Based Method Utilizing Energy Function for Foreground or Background --

立松 直倫¹
Naotomo Tatematsu大谷 淳¹
Jun Ohya

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。本稿では色確率・構造・事前確率に基づくエネルギー関数[2]に Dan らが提案する Feature-Cut[3]を併用した手法を提案する。本手法を複数の移動物体を含む時系列のステレオ画像に対し適用する実験を行い、本手法の有効性の見直しを確認した。本稿では2章で Feature-Cut の概要を説明し3章で提案手法の説明を行う。4章で実験結果を示し、5章で結論と今後の研究課題について述べる。

2. Feature-Cut の概要

Dan らの Feature-Cut の手法は、時系列画像の最初のフレームの各画素をユーザが手動でラベリングした物体領域と背景領域の結果を入力として、節 2.1 で説明するセグメンテーション処理を行い、二番目のフレームの各画素を物体領域と背景領域に分類する。この分類処理を順番に行い、時系列画像の全フレームを物体領域と背景領域に分類する。

2.1 Feature-Cut の分類処理

Feature-Cut は初めに SIFT 等の回転拡大にロバストな特徴点検出アルゴリズムを用いて特徴点の対応関係を求める。図 1 に入力画像の例と特徴点の対応関係を示す。



図1 入力画像と特徴点の関係例

特定の特徴点と画素の距離が、SIFT の Difference of Gaussian の極値の三倍の値(r)以内の画素の集合 u (図 1 赤丸枠)を、対象の特徴点に属すると定義する。直前フレームの同一特徴点の r の大きさの変化から特徴点のスケーリング行列 D を、勾配方向の変化から回転行列 R を算出し、最初のフレームの各特徴点に紐づく画素の集合 u を次フレームでの同じ特徴点に合わせて回転・拡大し重ね合わせる。

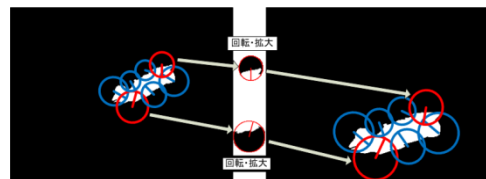


図2 直前結果の特徴点の重ね合わせ

本処理をフレーム s と t の全ての特徴点のセットに拡張し、式(1)の前景のエネルギー関数 $M'f(q+u)$ と式(2)の背景のエネルギー関数 $M'b(q+u)$ を定義する。 p, q は対応する各特徴点の組み合わせを $N(s, t)$ は対象のフレームのセットを示す。 $M'f(q+u) = \sum_{s \in S} \sum_{(p, q) \in N(s, t)} g(u, r) M_s(p + DRu) \Delta I$ (1)
 $M'b(q+u) = \sum_{s \in S} \sum_{(p, q) \in N(s, t)} g(u, r) (1 - M_s(p + DRu)) \Delta I$ (2)
 $g(u, r)$ は特徴点の位置と対象の点の距離が半径 r 以下の時は 1, 以上の時は 0 となる関数である。また ΔI は各特徴点の画素値を元に算出する特徴点の類似度を示す値である。 $\Delta I = 1 - [|I_t(q+u) - I_s(q+DRu)|]$ (3)
式(1)(2)のエネルギー関数を用いてグラフカットを行い、エネルギーが最小となる前景領域を抽出する。この処理を各フレームにおいて反復し、時系列画像の全フレームを前景と背景に分類する。

3. 提案手法

提案システムは、Kinect と WindowsPC で構成される。Feature-Cut は最初のフレームでユーザが手動で前景領域を指定する必要があるが、提案手法では TMR[1]を用いて移動物体領域の抽出を行うため、ユーザ操作は不要である。図 3 に提案手法のダイアグラムを示す。

(1) Kinect から取得した距離画像から三次元フローを生成

1 早稲田大学理工学術院国際情報通信研究科, Waseda Univ., GITS

する。

(2) 生成した三次元フローに TMR を適用してフローを複数の移動物体と背景に分類する。分類時に対象の物体が他の物体と異なる動きをし、新しい物体として分類されたフレームを生成フレームとする。

(3) 物体毎に生成フレームの解析を行い、各移動物体の領域と背景領域を分類する。

(4) 生成フレームを起点として Feature-Cut を適用して前後のフレームの各移動物体の領域と背景領域の分類を行う

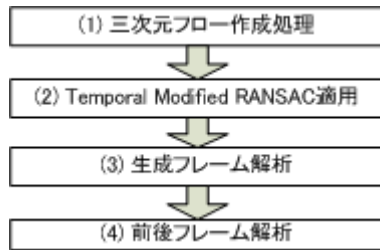


図3 提案手法のフロー

前景・背景の分類処理はグラフカットを用いて行う。提案手法では[2]の手法の色確率と構造確率に、Feature-Cut のエネルギーを組み合わせた式(4)のエネルギー式 $E(A)$ を用いて前景・背景のセグメンテーションを行う。本手法のエネルギー関数は、スケーリング定数 λ と画素 i が前景と背景のどちらに適合するかを表すデータ項 $R_i(A_i)$ と隣接画素 j との関係を表す平滑化項 $B_{i,j}$ と Feature-Cut のエネルギー式 $F_c(A_i)$ を用いて式(4)で定義される。 $F_c(A_i)$ の項は前景の場合には式(1)、背景の場合には式(2)を代入する。

$E(A) = \lambda \sum_{i \in I} R_i(A_i) + \sum_{(i,j) \in N} B_{i,j} \cdot \delta(A_i, A_j) + FC(A_i)$ (4)

生成フレームの解析時には、物体と回転・並進行列を用いてフローの起点を終点に移動させた誤差が閾値以下の三次元フローの領域を前景領域として使用する。前景領域の色ベクトル $(c[i])$ ・構造ベクトル $(p[i])$ を K-means 法でクラスタリングし、個々のクラスタ毎に平均と共分散行列を計算して混合ガウス分布 (F_g, B_g) を作成する。式(4)の各項は[2]より式(5)(6)(7)(8)で定義する。

$$R_i(f) = -\log(\alpha(1-\lambda) \sum_1^{15} Bg(c[i]) + \lambda \sum_1^{15} PreBg(c[i])) + (1-\alpha) \sum_1^{15} Bg(p[i]))$$
 (5)

$$R_i(b) = -\log(\alpha(1-\lambda) \sum_1^{15} Fg(c[i]) + \lambda \sum_1^{15} PreFg(c[i])) + (1-\alpha) \sum_1^{15} Fg(p[i]))$$
 (6)

$$\beta_{ij} = Y \exp(-(\alpha\beta_c \|c[i] - c[j]\|)) + (1-\alpha)\beta_s \|p[i] - p[j]\|$$
 (7)

$$\beta_c = \frac{1}{2\text{sum}(\|c[i] - c[j]\|)} \quad \beta_s = \frac{1}{2\text{sum}(\|p[i] - p[j]\|)}$$
 (8)

以上のエネルギー関数を元に算出されたエネルギーをグラフに設定し、minflow/maxflow アルゴリズムを適用して切断し、物体領域の抽出を行う。

4. 実験結果

図4の時系列画像に対して、提案手法と[2]の論文の手法を適用し復元結果を比較する実験を行った。図4の時系列画像では、フレーム0から4の間、右の車は奥に移動する。フレーム4から8の間は、右の車は奥に移動し左の車は手前に移動する。全フレームを通じて Kinect は真上に上昇する。図4(d)の特徴点の分類は正常に行われている。提案手法と従来法の復元結果の比較を図5に示す。従来法では右側の車の右下の部分が欠けているが提案手法では正常に分類出来ている(赤丸枠)。また、従来法では背景部分の

復元結果に右側の車の先頭部分(青丸枠)が含まれている。

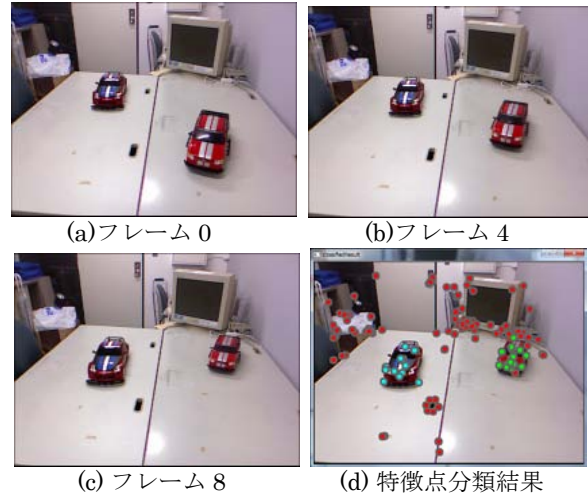


図4 入力画像と特徴点分類結果

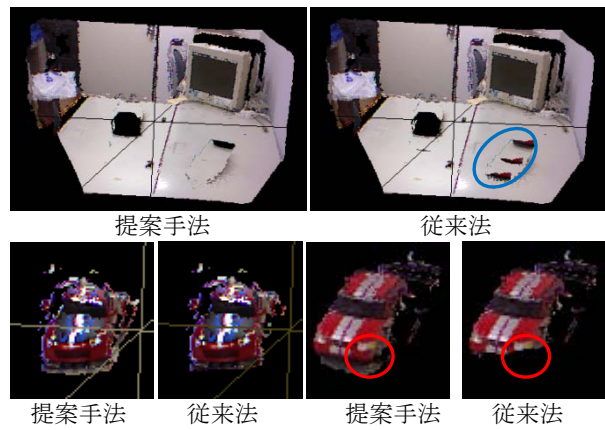


図5 復元結果

5. まとめ

本研究では筆者が今まで研究していた手法[2]のエネルギー計算に Feature-Cut のエネルギー項を追加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。今後はオクルージョンが発生する等のより複雑なケースの実験を行い、より詳細な検討を進める予定である。

参考文献

[1] N. Tatematsu and J. Ohya. Study of Temporal Modified-RANSAC Based Method for the Extraction and 3D Shape Reconstruction of Moving Objects from Dynamic Stereo Images and for Estimating the Camera Pose, IS&T-SPIE, Electronic Imaging 2011, Vol. 7878, 2011

[2] 移動ステレオカメラ画像中からの移動物体抽出と三次元データのセグメンテーションの検討--Temporal Modified-RANSAC と色・構造・事前確率を用いたグラフカットを組み合わせた手法の提案, 立松直倫, 大谷淳 電子情報通信学会技術研究報告 110(467), 253-258, 2011-03-10

[3] Dan Ring; Kokaram, A.; , "Feature-Cut: Video object segmentation through local feature correspondences," Computer Vision Workshops (ICCV Workshops), 2009 IEEE 12th International Conference on , vol., no., pp.617-624, Sept. 27 2009-Oct. 4 2009