

差分ステレオを用いた複数人物トラッキング

中央大学/JST CREST ○中西達也, 橋本優希, 寺林賢司, 梅田和昇

Tracking Multiple Persons Using Subtraction Stereo

○ Tatsuya NAKANISHI, Yuuki HASHIMOTO,

Kenji TERABAYASHI and Kazunori UMEDA, Chuo University / CREST, JST

Abstract: We have proposed a stereo camera system “subtraction stereo” that focuses on moving regions. This paper proposes a method to track two or more persons using the sequence of range images obtained by subtraction stereo. Tracking that is robust to occlusion, etc. is realized using Kalman filter. The effectiveness of the proposed method is verified by experiments.

1. 序論

不審者検出や人の流れの計測を行う監視用途には単眼のカメラで歩行者の検出などを行うシステムが様々な提案されている¹⁾。また、ステレオカメラとRFIDタグを用いて人流データなどを計測する手法²⁾も提案されている。このような不審者検出等の用途では、計測したい対象物は一般に動きを伴うことが多い。我々は運動情報に注目することによってステレオカメラのロバスト化を実現する差分ステレオ(以前は運動領域注視型ステレオと呼称)を提案している³⁾。本研究では差分ステレオで得られる画像に対し複数人数のトラッキングを行い人物毎の運動情報を取得する手法を提案する。

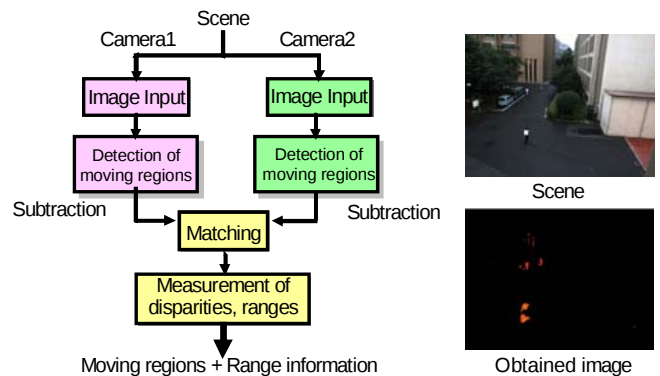


Fig.1 Flow of subtraction stereo

2. 差分ステレオの概要

2.1 基本アルゴリズム

差分ステレオの基本アルゴリズムを Fig.1 に示す。通常のステレオでは、左右カメラの画像をマッチングすることで視差画像を得る。これに対し、差分ステレオでは、左右カメラそれぞれでまず差分処理によって運動領域を抽出し、その後に抽出された運動領域を左右画像でマッチングする。この手法は運動している領域にしか適用できないという欠点がある反面、対応点探索を左右カメラの運動領域に大きく限定することができる。マッチングを大幅にロバスト化することができる。

2.2 移動物体の計測

差分ステレオで得られる視差画像から移動物体を抽出し、三次元座標と幅・高さを取得することが出来る。手法の詳細については文献⁴⁾を参照。

3. 人物トラッキング

人物の動きを理解するためには計測された複数の人物をトラッキングし、それぞれの時系列のデータを得

る必要がある。本研究では人物トラッキングにカルマンフィルタを用いる。カルマンフィルタは、誤差のある観測値から状態変数を最尤推定する手法である⁵⁾。

人物の動きが概ね等速直線運動であるモデルを立てる。ワールド座標系での人物の位置 (x, y, z) および速度 $(\dot{x}, \dot{y}, \dot{z})$ を状態変数とする。状態変数 $\mathbf{X}$ は次式で表わされる。

$$\mathbf{X} = [x \ \dot{x} \ y \ \dot{y} \ z \ \dot{z}]^T \quad (1)$$

カルマンフィルタの状態方程式は次式で与えられる。

$$\mathbf{X}_{i+1} = \Phi \mathbf{X}_i + \Gamma_i \omega_i \quad (2)$$

Φ は状態遷移行列であり

$$\Phi = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \quad (3)$$

で与えられる。 ω_i はプロセスノイズである。また、観測値として

$$\mathbf{Z} = [u \ v \ d]^T \quad (4)$$

を用いる. u, v は画像上の座標, d は視差である. 観測値と状態方程式の関係は次式のように表わされる.

$$\mathbf{Z}_i = f(\mathbf{X}_i) + \mathbf{v}_i, \quad (5)$$

$$f(\mathbf{X}_i) = \begin{bmatrix} \frac{x_i \cdot f}{z_i} & \frac{y_i \cdot f}{z_i} & \frac{b \cdot f}{z_i} \end{bmatrix}^T.$$

f はカメラの焦点距離, b は基線長, $\mathbf{v}_i$ は観測ノイズである. これらの式からカルマンフィルタを更新し予測値を得る. 得られた予測値と現フレームの人物の座標との三次元の距離を求め, 予測値から一定以内近くにいる人物を同一人物とし観測値とする. 人物のデータが取得出来なかった際には, 前フレームの状態変数から得た予測値をそのまま用いる.

以上の方法で人物のトラッキングが基本的には可能となる. しかし, このままでは人物のすれ違いのようなオクルージョンの起こるシーンでは, 観測値と予測値の対応付けが困難なためトラッキングが出来ない. そこで, 複数の予測値に対応づける観測値がオクルージョンにより一致してしまった場合には, 人物が重なっていると判定し観測データを用いず更新を行う. その後予測値に対応する観測値を別々に発見すると, トラッキングを再開する. 以上の手順により人物のすれ違いでのオクルージョンに頑強なシステムを構築した.

4. すれ違う人物のトラッキング実験

3章で説明した人物トラッキングの有効性を検証するため人物がすれ違う際の運動領域の特徴量の計測実験を行った. ステレオカメラは Bumblebee2(Point Grey Research. カラー, $f=3.8\text{mm}$, $b=120\text{mm}$, VGA, 48fps), PC は DELL XPS710 Core2 Extreme 2.93GHz×2 を使用した. カメラの位置姿勢は地面から高さ 8.3[m], Tilt 角 50° とした. また, プロセスノイズを $\dot{x}, \dot{y}, \dot{z}$ に対して $0.2[\text{m/frame}]$, 観測ノイズを d に対して 0.04pixel に設定した.

画面中央付近で人物がすれ違いそのまま反対側へ歩くというシーンを計測した. Fig.2 に実験の様子を示す. 図中での横方向を x 軸とし, 中央から右を正とする. 人物 1 は画面の右側から現れて左側へほぼ x 軸の負の方向に, 人物 2 は左側から現れ右側へほぼ x 軸方向に動いた.

それぞれの人物の x 座標の計測結果と予測値を Fig.3 に示す. 140frame 前後で交差が起こり計測値を取得できていないが, カルマンフィルタによって予測値を得ることで人物 1, 2 とともにトラッキングが上手く行えていることがわかる.

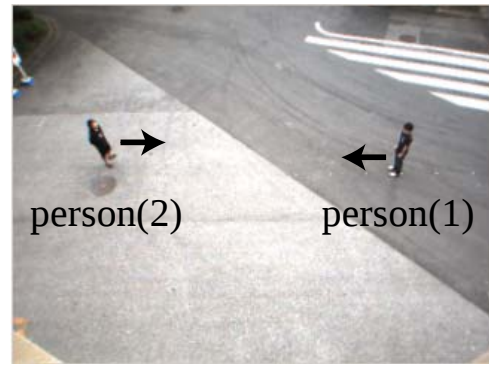


Fig.2 Experimental scene

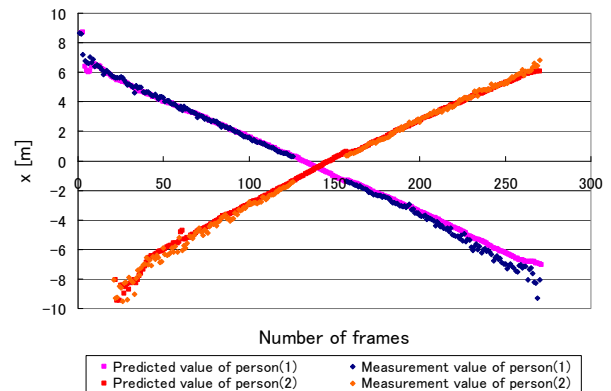


Fig.3 Measurement and predicted values of x coordinates of person(1) and person(2)

5. 結論と今後の展望

差分ステレオにおけるオクルージョンに頑強な複数人物トラッキング手法を提案し, 有効性を検証した. 今後は更に複雑なシーンでの人物トラッキングを行う.

参考文献

- 1) 寺田賢治: “車と人の監視”, 日本ロボット学会誌, Vol.22, No.7, pp.864-867, 2004.
- 2) 依田育士, 細谷大輔, 坂上勝彦: “ユビキタスステレオビジョンによる滞留人数計測システム”, 映像情報メディア学会誌, Vol.60, No.11, pp.1829-1834, 2006.
- 3) 梅田和昇, 和田正樹, 松木優介: “運動領域注視型ステレオカメラシステムの提案 第一報: 基本コンセプトの提案”, 第25回日本ロボット学会学術講演会(RSJ2007), 1N31, 2007.
- 4) Yuuki Hashimoto, Yusuke Matsuki, Tatsuya Nakanishi, Kazunori Umeda, Kei Suzuki and Kazunori Takashio: “Detection of Pedestrians Using Subtraction Stereo”, Proc. SWDMNSS2008.
- 5) Sebastian Thrun, Wolfram Burgard, Dieter Fox, 上田隆一 訳: “確率ロボティクス”, 毎日コミュニケーションズ, 2007.