

差分ステレオを用いた複数人物計測

Measurement of Multiple Persons Using a Subtraction Stereo

中西 達也 (中央大/JST CREST) 正 寺林 賢司 (中央大/JST CREST)

橋本 優希 (中央大/JST CREST) 星川 佑磨 (中央大/JST CREST)

正 梅田 和昇 (中央大/JST CREST)

Tatsuya NAKANISHI, Chuo University/CREST, JST, tnakani@sensor.mech.chuo-u.ac.jp

Kenji TERABAYASHI, Chuo University/CREST, JST, terabayashi@mech.chuo-u.ac.jp

Yuki HASHIMOTO, Chuo University/CREST, JST, hashimo@sensor.mech.chuo-u.ac.jp

Yuma HOSHIKAWA, Chuo University/CREST, JST, hoshika@sensor.mech.chuo-u.ac.jp

Kazunori UMEDA, Chuo University/CREST, JST, umeda@mech.chuo-u.ac.jp

We have proposed a stereo camera system “ subtraction stereo ” that focuses on moving regions. This paper proposes a method to track multiple persons using a 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.

Key Words: Stereo Camera, Motion, Detection of Moving Objects, Surveillance Camera

1. 序論

近年、不審者検出や人の流れの計測を行う監視用途において、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]。単眼のカメラによる歩行者検出システム [2] や、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 [3] などが提案されているが、ステレオカメラは寸法情報を取得出来るため、これらの用途に有用であると言える。また、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。そこで、我々は運動情報に注目することでステレオカメラのロバスト化を実現する “差分ステレオ (Subtraction Stereo)” を提案している [4] [5]。

差分ステレオによって得られた運動情報から人物の行動を理解するには、人物のトラッキングを行い各人物の時系列のデータを取得する必要がある。人物トラッキングの手法としてはカルマンフィルタを用いたもの [6] やパーティクルフィルタを用いたもの [7] などがある。李ら [8] は「共連れ」を防止するシステムでカルマンフィルタを用いた人物追跡を行っているが、レーザーレンジファインダを用いて人物の足の位置を計測するため、人物の高さを計測することが出来ない。そこで、本論文では差分ステレオで得られる三次元座標や幅、高さの情報を用いて、カルマンフィルタによってオクルージョンにロバストな複数人物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。

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、その後抽出された運動領域をマッチングすることで距離情報を得る。この手法は差分で得られた領域しか計測出来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、計算コストの低減とともにマッチングを大幅にロバスト化することができる。

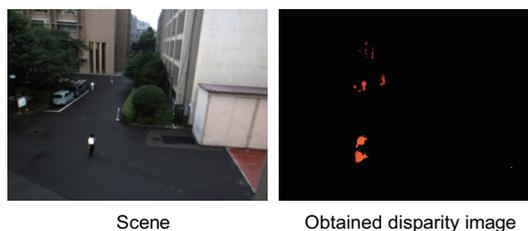
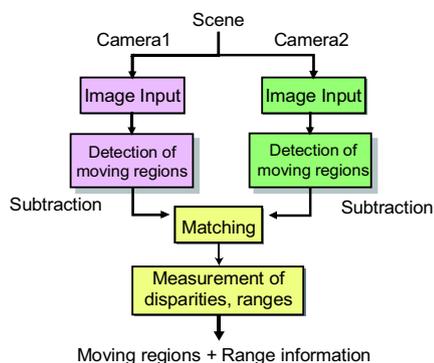


Fig.1 Flow of subtraction stereo

2.2 運動物体の三次元特徴量計測

差分ステレオで得られる視差画像から運動領域を抽出し、各領域の三次元座標と幅・高さの情報を取得することができる。

運動領域の三次元座標はラベリングによって分割された領域毎の視差の平均と重心座標から得ることができる。運動領域の座標はカメラの真下を原点とするワールド座標系に変化される。また、寸法情報を得ることが出来るため、領域の大きさを閾値として人物とそれ以外を判別し人物の領域のみの情報を得る [9]。

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、観測と予測を繰り返しながら、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 [10]。

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

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

カルマンフィルタの状態方程式は次式で与えられる。

$$X_{i+1} = \Phi X_i + \Gamma_i \omega_i \quad (2)$$

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

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

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

$$Z = f(X_i) + v_i, \quad f(X_i) = \begin{bmatrix} x_i \cdot f / z_i & y_i \cdot f / z_i & b \cdot f / z_i \end{bmatrix}^T. \quad (5)$$

f はカメラの焦点距離、 b は基線長、 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できなかった際には、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上手く行えない。そこで、複数の予測値に対応付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、人物が画像上で重なっていると判断し観測データを用いずに更新を行う。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発生するオクルージョンに頑強なシステムを構築した。

4. 実験

4.1 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各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 \times 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。カメラの位置姿勢を Fig.2 に示す。また、プロセスノイズを $\dot{x}, \dot{y}, \dot{z}$ に対

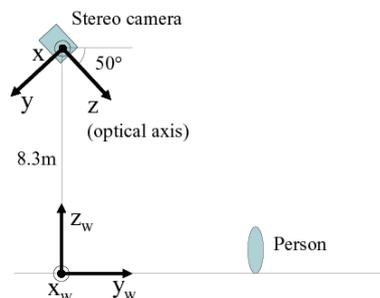


Fig.2 Experimental setup of the experiments

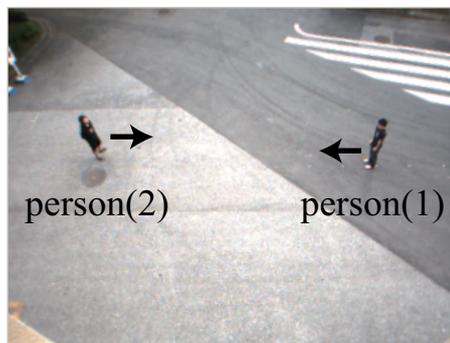


Fig.3 Experimental scene 1

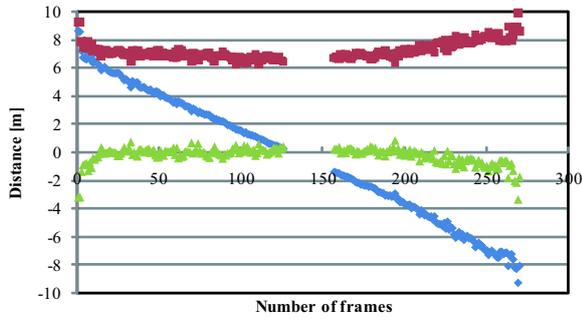
して 0.2[m/frame], 観測ノイズを d に対して 0.04pixel に設定した。

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.3 に実験の様子を示す。図中の横方向を x 軸、中央を原点として右方向を正とする。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。それぞれの人物の計測値を Fig.4, 5, x 座標の計測値と予測値の関係を Fig.6 に示す。Fig.4, 5 から人物 1 は左へ、人物 2 は右へ移動していることがわかる。また、Fig.6 から 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 ともにトラッキングが上手く行えていることがわかる。

4.2 複雑なシーンでのトラッキング実験

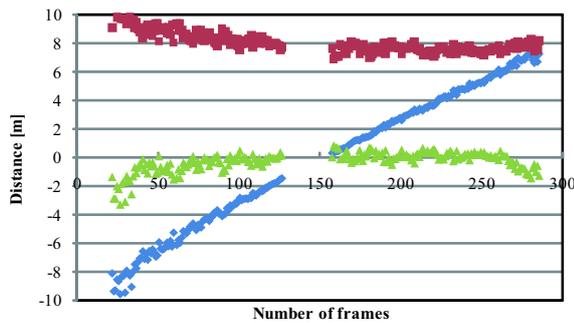
提案手法のロバスト性を検証するため 4.1 節の実験より複雑なシーンでの人物トラッキング実験を行った。実験に用いるステレオカメラや PC, カメラの位置姿勢は 4.1 節と同一条件で実験を行った。本実験では複数人物が画面上でランダムな方向に進んでいくシーンを 60[s] 間計測した。実験風景を Fig.7 に示す。

Fig.8 に示すように観測値は 3 次元情報で得られる。実験で計測された人物の中から前半約 15[s] 間で計測された ID_1 ~ 10 の人物の観測値を Fig.9 に示す (ID_5 は一瞬しか観測されなかった為除く)。一人の人物が観測されている時間の 80% 以上が同一の ID として認識されていた場合をトラッキング成功とすると、人物トラッキングの成功率は約 75% だった。また、計測中に同時に観測された人物の平均人数は 4.5 人であった。この事から複数人物が同時に計測されるような環境でもそれぞれの人物の座標情報の取得が行えていることがわかる。ID_2, 3, 7, 10 では人物の観測データが途切れている部分があるが、これは人物のすれ違いによって観測値が得られなかった部分である。したがって前節で示したすれ違いが起こった際のトラッキングが本実験においても上手く行えていることがわかる。



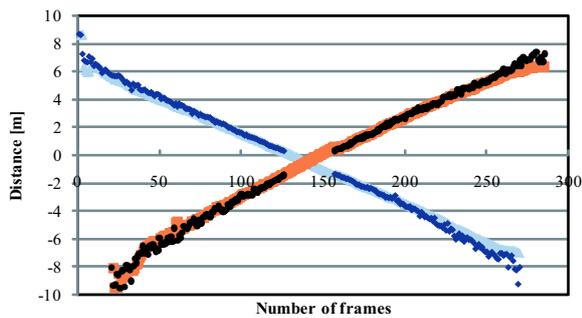
◆ measured value of x coordinates ■ measured value of y coordinates
▲ measured value of z coordinates

Fig.4 Measured values of person (1)



◆ Measured value of x coordinates ■ Measured value of y coordinates
▲ Measured value of z coordinates

Fig.5 Measured values of person (2)



▲ Predicted value of person(1) ◆ Measured value of person(1)
■ Predicted value of person(2) ● Measured value of person(2)

Fig.6 Measured and predicted values of x coordinates of person (1) and person (2)

4.3 一般環境での実証実験

提案手法の一般環境での有効性を検証するため秋葉原での実証実験を行った。実験では 9:00 ~ 18:00 の間に秋葉原 UDX 内の通路を歩行する人物を計測した。この実験では背景登録の際に 100 フレームの連続した画像を取得し、各画素での最頻値を背景として用いた。実験に用いるステレオカメラや PC は 4.1 節と同じものを用い、カメラの地面からの高さを 5.9[m], Tilt 角を 40° とした。実験風景を Fig.10 に示す。

計測データの中から 60[s] 単位で朝、昼の各時間帯から 5



Fig.7 Experimental scene 2

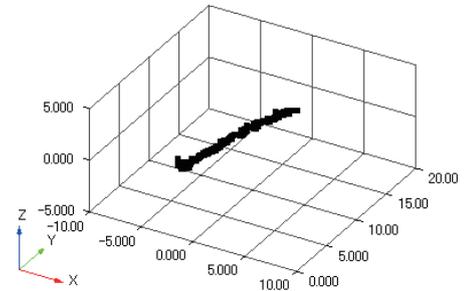


Fig.8 3D view of measured value (ID_1)

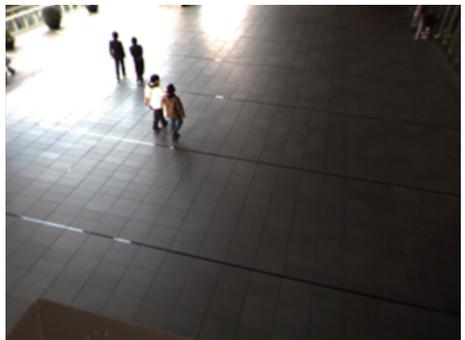


Fig.10 Experimental scene 3

Table 1 Results of tracking

Morning	62%
Daytime	58%
Evening	52%

個、夕方の時間帯から 4 個のデータを取得し、4.2 節と同様にトラッキングの成功率を調べた。結果を Table 1 に示す。ここから、一日を通してのトラッキング成功率は 60% 程度であるといえる。4.2 節と比べて成功率が下がった原因としては、背景差分の照明変化に対するロバスト性が不足していた点があげられる。照明変化によって背景部分の一部を人物と誤認識することによりトラッキングを失敗したと考えられる。そのため、照明変化にロバストな背景差分手法の導入によってトラッキング成功率の向上が可能であるといえる。また、夕方での成功率が低くなった原因としては輝度値の低下によって人物の抽出精度が低下した為と考えられる。

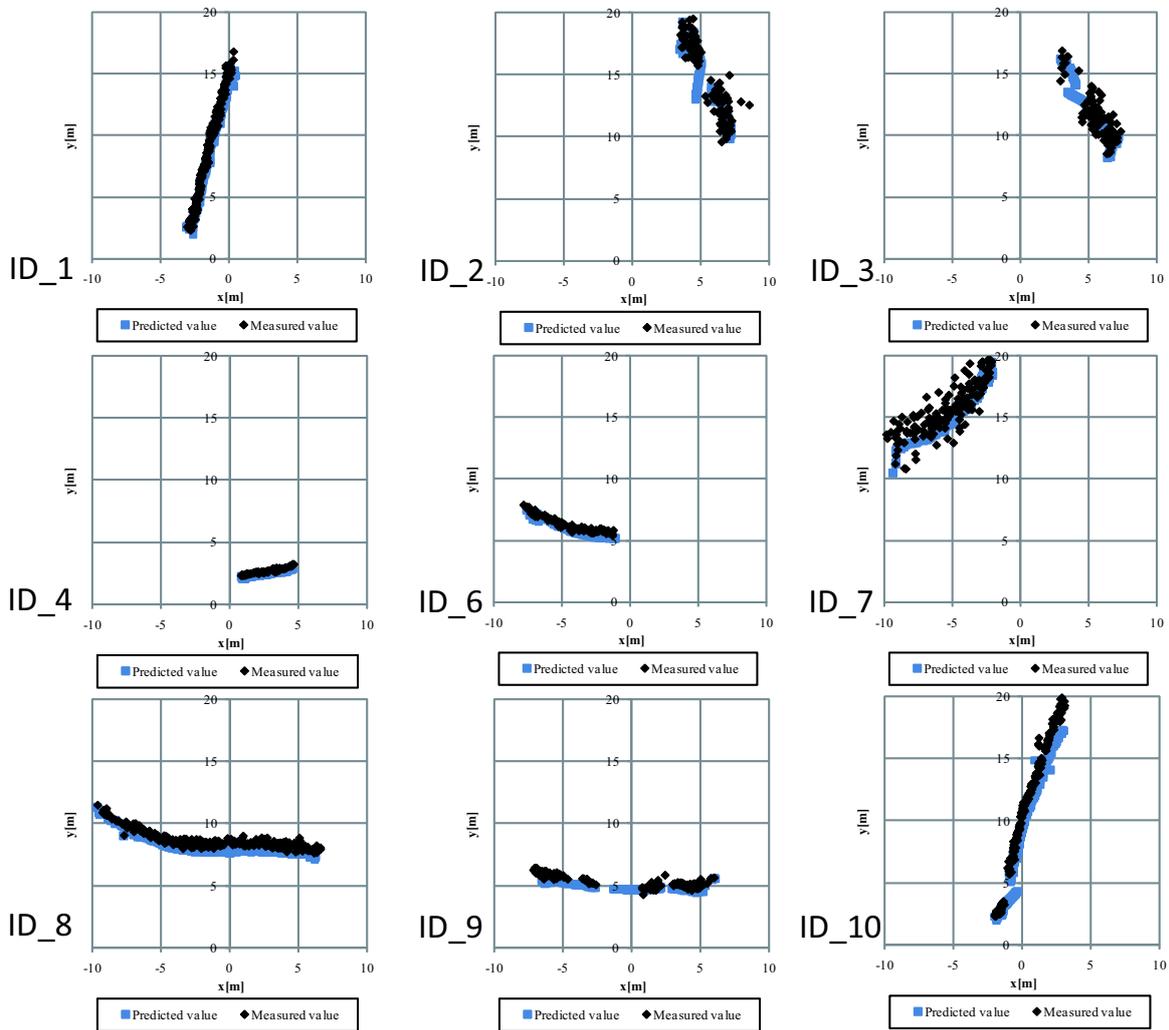


Fig.9 Measured values of person (ID_ 1 ~ 10)

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た。また、提案手法の有効性を人物のすれ違うシーンでの実験や、更に複雑なシーンでのトラッキング実験により検証し、オクルージョンに対するロバスト性を示した。更に、秋葉原での実証実験を行った。

今後は複雑なシーンでの人物トラッキングの精度向上のために、人物領域の大きさ等の新たな閾値を対応付けに用いた手法を検討する。Mean-shift法 [11] の様な色情報を用いた人物トラッキングを併用することでトラッキング成功率を高めることも考えられる。また、差分ステレオの基本アルゴリズムの改良として背景差分の照明変化に対するロバスト化や、高精度な影除去手法の実装をおこなう。

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] 羽下哲司，鷺見和彦，八木康史：“変化領域内の動きの時空間特徴に着目した屋外情景における歩行者の検出”，電子情報通信学会論文誌 D，Vol.J87-DII，No.5，pp.1104-1111，2004.
- [3] 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.
- [4] 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.
- [5] Kazunori Umeda, Tatsuya Nakanishi, Yuuki Hashimoto, Kota Irie, and Kenji Terabayashi: “Subtraction Stereo -A Stereo Camera System That Focuses On Moving Regions -”, Proc. of SPIE Vol.7239 3D Imaging Metrology, 2009.
- [6] J. H. Lee, T. Tsubouchi, K. Yamamoto and S. Egawa : “People Tracking Using a Robot in Motion with Laser Range Finder ”, Accepted to IEEE/RSJ Int. Conf.on Intelligent Robots and Systems, 2006.
- [7] 石黒勝彦，山田武士，上田修功：“複数対象トラッキングと複数ダイナミクスの同時推定法”，画像の認識・理解シンポジウム (MIRU2008)，pp.294-301，2008.
- [8] 李在勲，金容植，川田浩彦，大矢晃久，油田信一：“測域センサを用いたセキュリティーシステムの開発 - 人物追跡・計数アルゴリズム - ”，第 24 回日本ロボット学会学術講演会，2B22，2006.
- [9] Y. Hashimoto, Y. Matsuki, T. Nakanishi, K. Umeda, K. Suzuki and K. Takashio : “Detection of Pedestrians Using Subtraction Stereo ”, Proc. International Workshop on Sensor Webs, Databases and Mining in Networked Sensing Systems, 2008.
- [10] S. Thrun , W. Burgard and D. 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.
- [11] D. Comaniciu, V. Ramesh and P. Meer: “Real-Time Tracking of Non-Rigid Objects Using Mean Shift ”, Proc. IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Vol.2, pp.142-149, 2000.