

差分ステレオを用いた未知立体環境における人流計測

寺 林 賢 司^{†1} 橋 本 優 希^{†2}
星 川 佑 磨^{†2} 梅 田 和 昇^{†1}

「差分ステレオ」を用いた未知の立体環境における人流計測手法を提案する。差分ステレオは、マッチング処理を運動領域に限定することでロバスト化および高速化を図った、距離画像計測法である。本手法は、計測環境やカメラの位置姿勢などの情報を予め与えることなく、人物検出やシーン理解に有効である3次元位置や大きさの情報を取得することが可能であり、設置簡易な人流計測に適する。運動領域に着目することで、ラベリングと大きさ情報のみを使い、人物検出を簡単化することができる。各人物の時系列データを計測するために、拡張カルマンフィルタを用いて、各時刻に観測される位置情報の対応付けを行う。提案手法の有効性を検証するために、未知の立体環境を対象とした計測実験を行い、計測環境と整合する時系列データが取得できることを確認した。

Measurement of Pedestrian in Unknown 3D Environment Using Subtraction Stereo

KENJI TERABAYASHI,^{†1} YUKI HASHIMOTO,^{†2}
YUMA HOSHIKAWA^{†2} and KAZUNORI UMEDA^{†1}

This paper shows a method to measure pedestrians in unknown three-dimensional environment for applications such as surveillance cameras. For this measurement, we have proposed a stereo camera system that focuses on moving regions. In the system, the correspondence problem of the stereo cameras is reduced by restricting the stereo matching process to moving regions. By focusing on moving regions, pedestrians can be detected easily based on standard labeling and information of size that is directly obtained from subtraction stereo without any given information about measurement environments. Tracking each pedestrian, the extended Kalman filter technique is employed to address occlusion problem usually seen in many people environment. The proposed method is verified by an experiment in unknown three-dimensional environment.

1. はじめに

都市環境におけるカメラを用いた人の流れの計測は、監視システムや都市設計など多くの応用が考えられ、人の移動した軌跡、大きさ情報および人数を得ることが重要であり、単眼カメラを用いたアプローチが様々ななされている^{1)–4)}。一方、ステレオカメラによるアプローチ⁵⁾では、計測環境やカメラの位置姿勢などの事前情報なしに、距離や大きさ情報を直接的に得ることができ、計測システムを色々な場所へ移動させる場合など、設置簡易な人流計測に適すると考えられる。

ステレオカメラでは、テクスチャが弱い場合や繰り返しパターンがある場合に画像間の対応付けが難しくなる、対応点問題が不可避であることが知られている。この問題に対し、複数台のカメラを用いたマルチベースラインステレオ⁶⁾やランダムドットなどのパターンを投影する手法^{7),8)}など、ステレオマッチングのロバスト化が行われてきた。しかし、これらの手法は計測対象が近い場合に限定され、人の流れの計測という、広範囲に分布する遠い対象には適さない。

そこで我々は、不審者検出や人の流れの計測を行う監視カメラなどの用途を想定し、運動情報へ陽に注目した「差分ステレオ」を提案している⁹⁾。本手法は、マッチングにおける探索範囲を運動領域へ制限することによりロバスト化および高速化を図るものであり、これまでに計測精度の検討などを行ってきた¹⁰⁾。

本研究では、設置簡易な計測を想定し、計測環境やカメラ設置に関する事前情報のない、未知の立体環境における人流計測を目的とする。そのために、差分ステレオを用いた人物の検出および3次元情報の計測を行う(第2章)。人物が多数存在する場合、複数人がまとまって移動するグループやすれ違い時において、カメラから観測することのできないオクルージョン領域が生じるなど、計測値が得られないことが考えられる。このような問題に対し、拡張カルマンフィルタを用いたトラッキングにより、各人物の時系列データを取得することを目指す(第3章)。また、未知の立体環境を対象とした計測実験を行うことで、提案手法の有効性を検証する(第4章)。

^{†1} 中央大学/JST CREST
Chuo University/CREST, JST

^{†2} 中央大学大学院/JST CREST
Chuo University/CREST, JST

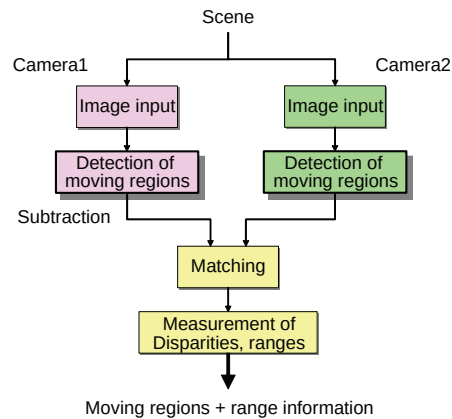


図1 差分ステレオの流れ
Fig.1 Flow of subtraction stereo

2. 差分ステレオによる人物計測

2.1 差分ステレオの概要

差分ステレオは、通常のステレオ手法と異なり、あらかじめ差分処理によって抽出した領域のみにマッチング処理を行う。図1は、差分ステレオの処理の流れを表したものである。これにより探索範囲を制限することができ、対応点探索のロバスト化および高速化を図ることができる。具体例として、図2へ示したシーンを入力とした場合、通常のステレオ手法との処理結果の違いは図3、図4のようになる。差分ステレオの出力（図3）では、人物領域に着目して距離情報が得られていることが確認できる。

2.2 人物の検出と3次元特徴量計測

差分ステレオでは、視差画像が運動領域に限定されるため、基本的なラベリングによって人物領域を抽出することができる。ノイズに対しては、大きさ情報の閾値処理により人物領域を判定する。

人物の3次元位置は、一般の平行ステレオに基づき、以下のように計算される。カメラから光軸方向の距離 z_c は、

$$z_c = \frac{b \cdot f}{d \cdot p} \quad (1)$$



図2 入力シーン
Fig.2 Input scene



図3 差分ステレオ
Fig.3 Subtraction stereo

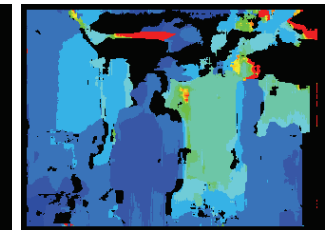


図4 通常のステレオ
Fig.4 Standard stereo

で与えられる。ここで、 d は視差、 b は基線長、 f は焦点距離、 p は1画素の横幅である。スキューがなく各画素のアスペクト比が1である場合、 z_c を用いて3次元位置 $\mathbf{x}_c$ は、

$$\mathbf{x}_c = z_c \cdot \begin{bmatrix} \frac{p}{f}(u - u_0) & \frac{p}{f}(v - v_0) & 1 \end{bmatrix}^T \quad (2)$$

のように計算できる。 (u, v) は計測点の画像座標、 (u_0, v_0) は画像中心の画像座標である。

人物領域の幅と高さは、距離が近い場合には、領域内の各点の3次元位置から求められるが、本研究で想定するような距離が遠い場合、計測誤差が問題となる。そこで、視差画像における人物領域の外接矩形の幅 u_s と高さ v_s から、次式を用いて計算する。

$$width = \frac{z_c \cdot p}{f} u_s \quad (3)$$

$$height = \frac{z_c \cdot p}{f} v_s \quad (4)$$

3. 人物トラッキング

人の流れである歩行軌跡を計測するためには、各時刻における位置情報を時系列で人物ごとに対応付ける必要がある。このとき、歩行者すれ違い時のオクルージョンなどから、位置情報を観測できないことが問題となる。本研究では、拡張カルマンフィルタを用いた人物トラッキングにより、観測できない場合に推定位置を用いることで解決を図る。

状態変数は、人物領域の重心位置およびその速度として、

$$\mathbf{x} = \begin{bmatrix} x & \dot{x} & y & \dot{y} & z & \dot{z} \end{bmatrix}^T \quad (5)$$

のように定義する。人物の運動に対して等速直線運動を仮定し、状態遷移モデルを

$$\mathbf{x}_{i+1} = \Phi \mathbf{x}_i + \omega \quad (6)$$

とする。ここで、 Φ は状態遷移行列であり、等速直線運動モデルとして具体的には、

$$\Phi = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \quad (7)$$

である。また、 ω はプロセスノイズである。観測変数は、画像上の対応する座標値 (u, v) および、視差 d として、

$$\mathbf{z} = \begin{bmatrix} u & v & d \end{bmatrix}^T \quad (8)$$

のように定義し、観測モデルを以下のとおりとした。

$$\mathbf{z}_i = f(\mathbf{x}_i) + \mathbf{v} \quad (9)$$

$$f(\mathbf{x}_i) = \begin{bmatrix} \frac{x_i \cdot f}{z_i} & \frac{y_i \cdot f}{z_i} & \frac{b \cdot f}{z_i} \end{bmatrix}^T \quad (10)$$

$\mathbf{v}$ は観測ノイズ、 f はカメラの焦点距離、 b は基線長である。ここでは、オクルージョンによって観測できないときの予測モデルという観点から、拡張カルマンフィルタを導入したが、視差 d の誤差ひいては光軸方向の距離 z_c の誤差に対して、本来のフィルタリング効果が期待できる。

各時刻における計測位置情報は、以上のように計算される推定位置との距離の閾値処理によって人物ごとに対応付ける。

4. 未知立体環境における計測実験

本節では、差分ステレオを用いた人流計測の有効性を検証するために、未知の立体環境を対象として実験を行う。単眼カメラのように、環境情報やカメラの設置位置姿勢を与えることなく、人流を計測することが出来るかを確認する。

4.1 実験条件

図5に示す立体環境において、被験者がスロープを上がってくるシーンを対象として計測実験を行った。このとき、ステレオカメラはスロープ上にある平坦部分へ、光軸が水平と

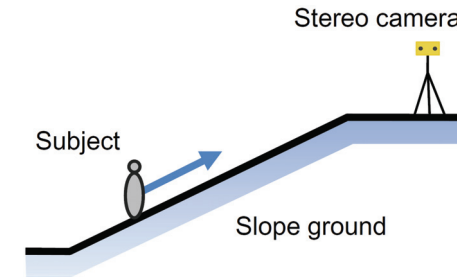


図5 未知の立体環境における実験
Fig. 5 Experiment in unknown 3D environment



図6 計測画像の例
Fig. 6 Example of captured image



図7 出力結果の例
Fig. 7 Example of processing result

なるように設置した。

ステレオカメラは、Point Grey Research 製の Bumblebee2 (color, VGA, 48[fps], $b = 120[\text{mm}]$, $f = 3.8[\text{mm}]$, $p = 14.8[\mu\text{m}]$) を使い、実験での画像サイズは $320 \times 240[\text{pixel}]$ とした。処理速度は、DELL XPS 710 Extreme (2.93[GHz] $\times$ 2) を用いて約 29[fps] である。

拡張カルマンフィルタのパラメータは、プロセスノイズ ω を速度成分に対して $0.2[\text{m/frame}]$ 、観測ノイズ $\mathbf{v}$ を視差 d に対して $0.04[\text{pixel}]$ とした。人数推定に必要な定数 c は、予め実験から求めておき $c = 140000$ とした。

4.2 実験結果

図6は、実験中に計測した画像の例であり、これに対応するシステムの出力結果を示したものが図7である。図7は、視差画像に計測状況を付加したものであり、黄色い枠線およびその上に表示された記号は、人物個別のIDを示しており、人物トラッキング適切に動作し時系列データが計測できていることを表している。これらの図から、人物が適切に検出

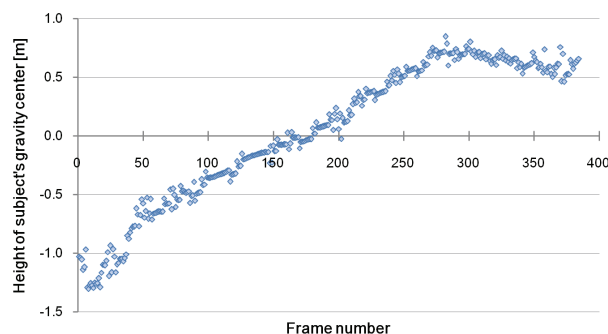


図 8 計測された高さ情報

Fig. 8 Measured height of subject's gravity center

され 3 次元情報が計測されている様子を確認することができる。

本実験で計測された人物の重心座標の高さを図 8 へ示す。図中において、横軸がフレーム番号、縦軸が高さを表しており、時系列データが計測されていることが確認できる。また、300 フレーム近辺を境界として、スロープを上っている様子とスロープ上部の平坦部分を移動していることが分かる。

一方で、平坦部分での高さが徐々に小さくなっていることも確認できるが、これは影の影響により人物の領域が大きく検出され、それによって重心座標が低くなったためと考えられる。したがって、人物の位置情報の代表値としての重心座標を精度よく計測するためには、影の除去の必要である。

5. おわりに

本稿では、設置簡易な計測を想定し、差分ステレオを用いた未知の立体環境における人流計測手法を提案した。スロープや平坦部分を含む立体環境での計測実験において、計測環境および被験者の運動と整合する時系列データが計測できることを確認した。

今後は、さらに複雑な立体環境での検証や、人物領域の正確な検出のために影の除去に取り組む予定である。

参考文献

- 1) Kilambi, P., Ribnick, E., Joshi, A.J., Masoud, O. and Papanikolopoulos, N.: Estimating pedestrian counts in groups, Computer Vision and Image Understanding, Vol.110, No.1, pp.43–59 (2008).
- 2) Xu, F., Liu, X. and Fujimura, K.: Pedestrian detection and tracking with night vision, IEEE Trans. on Intelligent Transportation Systems, Vol.6, No.1, pp.63–71 (2005).
- 3) 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).
- 4) 羽田哲司, 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検出, 電子情報通信学会論文誌 D, J87-D2(5), pp.1104–1111 (2004).
- 5) Kelly, P., O'Connor, N.E. and Smeaton, A.F.: A framework for evaluating stereo-based pedestrian detection techniques, IEEE Trans. on Circuits and Systems for Video Technology, Vol.18, No.8, pp.1163–1167 (2008).
- 6) Okutomi, M. and Kanade, T.: A multiple-baseline stereo, IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol.15, No.4 (1993).
- 7) 佐藤雄隆, 松山方大, 坂上勝彦: ステレオビジョンを搭載した電動車いすにおける近赤外パタン光投影の検討, 動的画像処理実用化ワークショップ (DIA2007) 論文集, pp.189–192 (2007).
- 8) 松木優介, 梅田和昇: マルチスポット光とステレオカメラを用いたアクティブ・パッシブ統合距離画像センサの高性能化, 日本機械学会ロボティクス・メカトロニクス講演会 (Robomec2008) 講演論文集, 2P1-G12 (2008).
- 9) Umeda, K., Hashimoto, Y., Nakanishi, T., Irie, K. and Terabayashi, K.: Subtraction stereo - a stereo camera system that focuses on moving regions -, Proc. of SPIE, Vol.7239, 7239–33 (2009).
- 10) Hashimoto, Y., Matsuki, Y., Nakanishi, T., Umeda, K., Suzuki, K. and Takashio, K.: Detection of pedestrians using subtraction stereo, Proc. of 2nd International Workshop on Sensor Webs, Databases and Mining in Networked Sensing Systems (SWDMNSS 2008), pp.165–168 (2008).