

# 差分ステレオを用いた歩行者集団の計測

## Measurement of Pedestrian Groups Using Subtraction Stereo

学 橋本 優希 (中央大/JST CREST)    ○学 有江 誠 (中央大)  
 学 星川 佑磨 (中央大/JST CREST)    学 生形 徹 (中央大)  
 正 寺林 賢司 (中央大/JST CREST)    正 梅田 和昇 (中央大/JST CREST)

Yuki HASHIMOTO, Chuo University/CREST, JST, hashimo@sensor.mech.chuo-u.ac.jp  
 Makoto ARIE, Chuo University, arie@sensor.mech.chuo-u.ac.jp  
 Yuma HOSHIKAWA, Chuo University/CREST, JST, hoshika@sensor.mech.chuo-u.ac.jp  
 Toru UBUKATA, Chuo University, ubukata@sensor.mech.chuo-u.ac.jp  
 Kenji TERABAYASHI, Chuo University/CREST, JST, terabayashi@mech.chuo-u.ac.jp  
 Kazunori UMEDA, Chuo University/CREST, JST, umeda@mech.chuo-u.ac.jp

This paper proposes a method to distinguish a state of walk of pedestrian groups by using “subtraction stereo” which is a stereo camera system that focuses on moving regions to make stereo matching fast and robust to the correspondence problem in computer vision. The state is classified into four categories: convergence, divergence, parallel shift, and stop. The state of walk is distinguished by an angle and a speed in 3D space of feature points of pedestrians obtained by the KLT Tracker. The effectiveness of the proposed method is verified by several experiments.

**Key Words:** Stereo Camera, KLT Tracker, State analysis, Pedestrian Groups

### 1. 序論

近年、凶悪犯罪の増加などをうけ、監視作業の需要が増加している。この監視作業は不審者などを検出し、未然に犯罪を防ぐために重要だが、手動で行うには多くの人手が必要となる。そのため、自動で異常を検出するシステムが期待されている。

このような状況を踏まえ、異常動作検出に関して様々な研究が行なわれており[1]~[3]、主に歩行情報が利用されている。林ら[1]は、単眼カメラを用いて比較的広範囲を対象として人物の移動軌跡や速度を利用し、監視領域での立ち止まり、うろつきなどの行動を検出している。しかし、計測しているのは2次元の動きであり、動作の理解に有効な3次元情報は用いていない。Andradeら[2]は、単眼カメラを利用し、オブティカルフローと隠れマルコフモデルを用いた群衆の異常動作検出手法を提案している。群衆の通常歩行を学習し、尤度や検出閾値を比較することによって群衆の異常動作を特定する。しかし、通常歩行のモデルは、計測環境やカメラの位置と姿勢などにより変化するため、計測システムを移動させるたびに学習をしなければならない。また、精度向上のため、通常多くの学習サンプルが必要となる。このように、これらほとんどの研究が2次元での解析にとどまっている。しかし、人の移動軌跡や寸法情報などの3次元情報を利用すると、例えば、入力シーンに階段や複数のフロアを歩く人が映り、2次元では解析できない場合など、立体的な環境における動作解析に用いる情報が増えるため有効であると考えられる。

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、運動情報に陽に注目した“差分ステレオ”を提案している[4]。これは新しいステレオ視の一手法であり、単眼カメラの場合[5]に比べて人物の3次元での動きを捉えることができるため、これらの用途に有用である。しかし、現在はまだ人物の領域抽出やトラッキングにとどまっている。

本論文では、個々の歩行者の動きに着目し、3次元で人の動きを捉え、異常動作検出の一要素技術として集団の歩行状態を識別することを目的とする。そしてそのために、差分ステレオにKanade-Lucas-Tomasi Tracker(以下、KLT)[6]を実装し、

これから得られる各特徴点の移動方向、移動速度を用いて、歩行者集団の歩行状態の判別手法を提案する。本論文では、屋外の比較的広範囲な計測環境において実験を行い、手法の有効性を検証する。

### 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図1(a)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。この手法は対応点探索を左右カメラの運動領域に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ロバスト化することができる。図1(b)の入力シーンに対し、差分ステレオを用いることで図1(c)に示す視差画像を得る。

#### 2.2 運動物体の3次元特徴量計測

差分ステレオで得る視差画像から運動領域を抽出し、代表点の3次元座標と領域の寸法情報を取得することができる。

運動領域の3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。運動領域のカメラ座標系は設置条件を用いてカメラの真下を原点とする世界座標系に変換される。また、寸法情報を得ることができるため、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[7]。

### 3. 歩行者集団の計測

#### 3.1 歩行者集団の計測の概要

個々の歩行者の動きに着目し、歩行者集団の歩行状態を計測する。まず、KLTを用いて集団に属する個々の歩行者の特徴点の変化を計測する。このとき、特徴点としてコーナー点を用いる。次に、この特徴点の変化から速度と方向を算出し、集団の歩行状態を判別する。

以降、歩行状態の概念、歩行者集団の定義、処理の流れ、

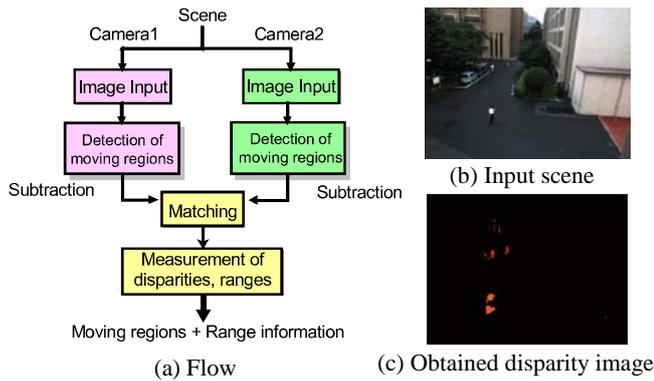


Fig.1 Subtraction stereo

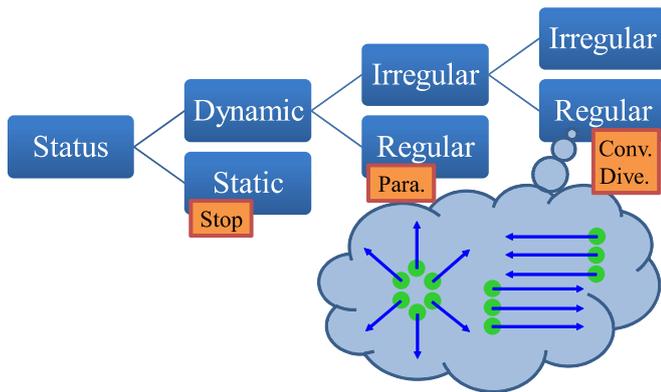


Fig.2 Classification of the state of human groups

速度と方向の算出方法，歩行状態の判別方法について説明する。

### 3.2 歩行状態の概念

図 2 に示すように，集団の歩行状態として，動いている場合と静止している場合がある。次に，動きの中でも乱雑な場合と規則性のある場合がある。この 1 つ目の規則性とは，集団全体が同じ方向に動いている状態を指す。そして，乱雑さの中でも乱雑な場合と規則性のある場合がある。この 2 つ目の規則性とは，一見乱雑なように見えるが，図 2 の右下に示すように実はある規則性を持って動いている状態を指す。また，ここで言う乱雑とは，まったく規則性を持たずに動いている状態を指す。

以上より，集団の動作に着目し，歩行状態を以下に示す 4 つに分類する。1.収束(Convergence), 2.発散(Divergence), 3.平行移動(Parallel Shift), 4.停止(Stop) (以下, Conv., Dive., Para., Stop)。ここで, Andrade ら[2]は集団の平行移動以外を異常動作としており, Boghossian ら[3]は集団の円形, 分散の動作, 動作のない領域を発見し, 異常動作としている。このように, 集団の異常動作検出には集団が平行移動しているかどうか判別することが重要であり, この 4 つの分類を押さえることでほとんどの異常動作に対応することができると思われる。

### 3.3 歩行者集団の定義

本論文では, 歩行者集団として, 以下に挙げる(1)か(2)を満たすものを定義する。(1) ラベリングによるラベルが同一である(画像上で繋がっている)歩行者, (2) 重心間の距離が一定

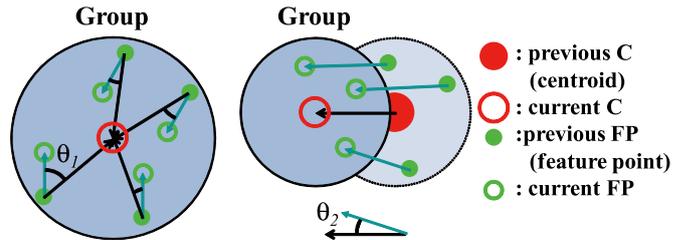


Fig.3 Relation between  $\theta_1$ ,  $\theta_2$ , and the flow of feature points

Table 1 Definition of the state of groups and its condition

state	condition
Conv.	$0[^\circ] \leq \theta_1 \leq 45[^\circ]$
Dive.	$135[^\circ] \leq \theta_1 \leq 180[^\circ]$
Para.	$0[^\circ] \leq \theta_2 \leq 45[^\circ]$
Stop	$v \leq 0.5[m/s]$

以内の距離の歩行者。集団を定義するため上記のいずれにおいても, 合計人数が 2 人以上であることを前提とする。このとき, 我々が提案している手法[7]を用いて集団の人数を検出する。

### 3.4 処理の流れ

歩行者集団の計測手法の流れについて説明する(図 3 参照)。計測は, 世界座標系における  $X_w-Y_w$  平面 (以下,  $X_w-Y_w$  平面) において行う。以下に挙げる手順より, 集団ごとに処理を行う。(1) KLT による特徴点の移動量から, 集団全点の速度  $v$  と方向  $\theta$  を取得する, (2) 図 3 の左側の図に示すように, 特徴点から集団の重心へのベクトル  $\mathbf{v}_1$  と, 特徴点の移動ベクトル  $\mathbf{v}_2$  の間の角度  $\theta_1$  ( $0[^\circ] \leq \theta_1 \leq 180[^\circ]$ ) を計算する, (3) 図 3 の中央の図に示すように, 集団の重心の移動方向と各特徴点の移動方向の角度の差の絶対値  $\theta_2$  ( $0[^\circ] \leq \theta_2 \leq 180[^\circ]$ ) を計算する, (4)  $\theta_1, \theta_2, v$  を特徴量として状態を判別する (3.6 節)。

### 3.5 特徴点の速度と方向の算出

連続する 2 枚の画像において, KLT により求まる同一の特徴点の 3 次元情報を用いて, 速度と方向を算出する手法について述べる。

まず, 連続する 2 枚の画像において, 同一の特徴点の世界座標をそれぞれ求める。ここで, ステレオカメラを用いるから世界座標が求まることに留意されたい。次に, 特徴点の移動量とこの連続する 2 フレーム間の時間から速度を求める。ここで,  $X_w-Y_w$  平面において移動量を計算する。また, 特徴点の移動ベクトルと  $X_w$  方向の間の角度から方向を求める。具体的には,  $X_w$  方向を  $0[^\circ]$  とし, 反時計回りを正とする。

本論文では精度向上のため, 過去 10 フレーム分の速度の値を平均した値を速度として用いる。このとき, 10 フレーム分を途切れることなく取得できた特徴点のみ, 集団の状態判別に利用する。

### 3.6 歩行状態の判別

歩行状態の分類は, 3.2 節で説明した。そこで, KLT を用いて算出する速度と方向を使い, 各々の状態について, 条件を満たすかどうか判別する。各状態とその条件を表 1 に示す。条件については, 予備実験を行いパラメータを決定した。この表 1 に示す条件と特徴点全点の  $\theta_1, \theta_2, v$  を用いて歩行状態の判別を行う。

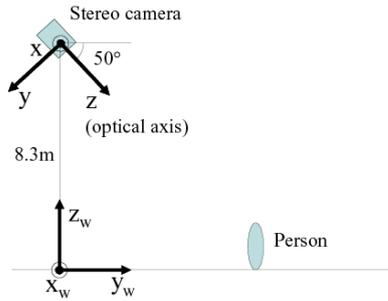


Fig.4 Experimental setup of the experiments

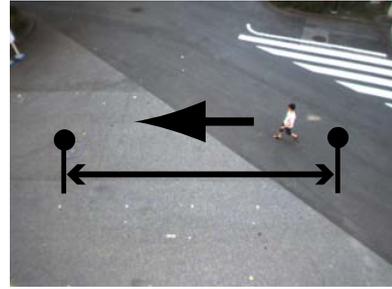


Fig.5 Experimental scene 1



Fig.6 Scene of feature points trajectory

Table 2 Result of accuracy evaluations of the velocity and the direction

	Velocity[m/s]	Direction[°]
True value	1.39	180.00
Ave.	1.74	184.61
S.D.	0.44	11.76
Max.	3.28	215.05
Min.	0.92	153.88

具体的には、まず歩行者集団において、過去 10 フレーム分のデータを保持した有効な特徴点全点において、 $\theta_1, \theta_2, v$  が各条件を満たすかどうか調べる。そして、各条件を満たす特徴点を数え、特徴点全点の半数以上が当てはまる状態を、その歩行者集団における歩行状態とする。状態の優先度を高い順に平行移動、停止、収束、発散とし、複数の条件が当てはまったときは優先度の高いものを適用する。また、どれにも条件が合わないときは、状態無しとする。

#### 4. 歩行者集団の計測実験

本章では、3つの実験から、差分ステレオを用いた歩行者集団の計測手法の有効性を検証する。大学構内の計測環境において、歩行者集団を対象として集団の状態判別を行った結果を示す。

ステレオカメラは、Point Grey Research 製の Bumblebee2 (color, VGA, 48[fps],  $b = 120[\text{mm}]$ ,  $f = 3.8[\text{mm}]$ ,  $p = 7.4[\mu\text{m}]$ ) を使い、実験での画像サイズは  $320 \times 240[\text{pixel}]$  とした。また、ステレオマッチングには Bumblebee2 に付属のライブラリを利用している。背景差分で得られる左右の差分画像に対してステレオマッチングを行うことで、視差画像が得られる。PC は DELL XPS 710 Intel(R) Core(TM)2 Extreme (2.93[GHz]×2) を使用した。

##### 4.1 歩行者の速度と方向の精度評価実験

速度と方向の精度を検証するため、歩行計測実験を行った。図 4 に示すように、ステレオカメラを高さ 8.3[m]、仰角  $50^\circ$  として設置した。1 人の歩行者が画面中央付近を右側から左側へ歩くというシーンを対象とし、速度と方向それぞれの計測精度を求めた。歩行者が計測範囲を通過する間に 180 フレーム取得し、このデータを用いて平均値と標準偏差を計算した。全フレームにおいて過去 10 フレーム分のデータを保持した有効な特徴点を最低 1 点は得た。このときのシーンと計測範囲を図 5 に示す。ここで、代表点で評価を行なうため、本実験では複数の特徴点が現れた場合、各特徴点の速度と方向を平均して 1 つにまとめている。



Fig.7 Experimental scene 2

平均値と標準偏差の計算結果を表 2 に示す。また、画像座標上における有効な特徴点の 10 フレーム分の軌跡の例を図 6 に示す。特徴点は頭や肩や腹のあたりに出現した。これらより、平均値を見ると概ね真値に近い値となったが、標準偏差は大きいことが分かる。これは、歩行時の足の振れにより進行方向及び垂直方向に周期的な振動が生じるので完全な等速直線運動にはならず、このため標準偏差が大きくなったものと考えられる。また速度については、別のデータでも精度評価を行った結果、平均値が真値よりもやや大きくなることが分かった。これは、ステレオカメラで特徴点の 3 次元座標を計測するときの計測誤差により、実際の移動量よりも計測値が大きくなるためである。

##### 4.2 集団が 1ヶ所に集まり移動するシーンでの状態判別実験

状態判別の有効性を検証するため、集団が 1ヶ所に集まり移動するシーンに対し、状態判別実験を行った。本節では、ステレオカメラを高さ 5.1[m]、仰角  $40^\circ$  として設置した。図 7 に示すように、11 人の集団が画面周辺部から中央付近に集まり、全員が集まってから左側へ歩くというシーンを対象とし、歩行状態の判別条件より、集団の状態を求めた。本実験では状態の真値は最初は収束で、途中から平行移動である。

各シーンにおいて、収束、発散、平行移動、停止のいずれかの状態を計測した回数及び、システムがもれなく計測した場合の計測回数を表 3 に示す。このデータを用いて状態の適合率(Precision)と再現率(Recall)を計算した。ここで、適合率とは実際に状態を計測した結果のうち、正確な状態を計測できた割合、再現率とは計測シーンのうち、正確な状態を計測で

Table 3 Number of measurements of state in Conv. and Para. scenes

Number	Conv.	Para.
Measured	28	162
Total	62	275

Table 4 Precision and recall in Conv. and Para. scenes

	Conv.	Para.
Precision[%]	61	85
Recall[%]	27	50

Table 5 Number of measurements of each state obtained from Fig.7

Number	Conv.	Dive.	Para.	Stop
Measured	20	22	147	1
True	17	0	137	0

きた割合である。各シーンにおける適合率と再現率の計算結果を表4に示す。表4より、平行移動は85[%]と高い適合率を得ていることが分かる。収束の適合率は平行移動と比べて低い、これは同時に収束と平行移動の両者が条件を満たし、状態の判別条件から優先度の高い平行移動を集団の状態と誤って判別してしまったことが原因である。また、この4つの状態において、状態が計測された回数と正しい結果を得た回数を表5に示す。これより、誤って状態を判別していることが確認できる。この問題については、条件をより厳しくすることで誤った状態判別を減らすことができるが、同時に再現率が低下するため、どちらに重点を置くか検討する必要がある。

### 4.3 集団2組が交差するシーンでの状態判別実験

本節では、さらに別のシーンで状態判別の有効性を検証するため、集団2組が交差するシーンに対し、実験を行った。4.2節と同一条件でステレオカメラを設置した。図8に示すように、3人の集団が画面右側から左側へ、8人の集団が左側から右側へ歩き、中央付近で交差してそのまま直進するというシーンを対象とし、3.6節で説明した条件より、集団の状態を求めた。本実験では状態の真値は最初は平行移動で、途中収束と発散を経て平行移動に戻る。

各シーンにおいて、収束、発散、平行移動、停止のいずれかの状態を計測した回数及び、システムがもれなく計測した場合の計測回数を表6に示す。このデータを用いて状態の適合率と再現率を計算した。各シーンにおける適合率と再現率の計算結果を表7に示す。表7より、適合率は概ね高いと言える。収束と発散の適合率が平行移動と比べて低い原因については、前節で述べた理由による。また、この4つの状態において、状態が計測された回数と正しい結果を得た回数を表8に示す。これより、前節と同様に誤判別の傾向が見られた。以上より、速度と方向は概ね真値に近い値を得ていることから、状態判別の条件をさらに調整することが精度向上に繋がると考えられる。

## 5. 結論と今後の展望

本論文では差分ステレオを用いた歩行者集団の計測手法を提案した。そして、集団の状態判別を目的として屋外の比較的広範囲な計測環境において3つの実験を行い、提案手法の有効性を検証した。実験は不審者や異常状態の検出に必要な歩行に関するもので、実験結果より本手法が歩行者集団の状



Fig.8 Experimental scene 3

Table 6 Number of measurements of state in intersect scene

Number	Conv.	Dive.	Para.
Measured	46	36	204
Total	101	66	304

Table 7 Precision and recall in intersect scene

	Conv.	Dive.	Para.
Precision[%]	72	72	81
Recall[%]	33	39	54

Table 8 Number of measurements of each state obtained from Fig.8

Number	Conv.	Dive.	Para.	Stop
Measured	44	55	187	0
True	33	26	165	0

態を判別するのに有効であることを示した。また、3次元で計測を行なうため、計測する特徴点の方向と速度がカメラの設置条件に依存せず、計測環境の変化に対してロバストであると考えられる。

今後の展望としては、精度の更なる向上と提案手法の更なる応用が挙げられる。後者に関しては、集団の形状や個々人の動きに着目することにより、状態の数をさらに増やすことが考えられる。また実験結果より、入力シーンに階段や複数のフロアを歩く人が映る、3次元環境においても提案手法が利用可能であると考えられるため、今後はこのような環境において実験を行うことを検討する。

## 文献

- [1] 林健太郎, 羽下哲司, 関真規人, 笹川耕一, “映像監視における人物位置行動の検出技術”, 情報処理学会論文誌. コンピュータビジョンとイメージメディア, Vol.47, No.9, pp.12-20, 2006.
- [2] Andrade, E., Blunsden, S. and Fisher, R., “Hidden Markov Models for Optical Flow Analysis in Crowds,” Proc. of ICPR'06, Vol.01, pp.460-463, 2006.
- [3] Boghossian, B. and Velastin, S., “Motion-Based Machine Vision Techniques for the Management of Large Crowds,” in The 6th IEEE International Conference on Electronics, Circuits and Systems, Vol.2, 1999.
- [4] 梅田和昇, 寺林賢司, 橋本優希, 中西達也, 入江耕太, “差分ステレオ運動領域に注目したステレオ視-の提案”, 精密工学会誌, Vol.76, No.1, pp.123-128, 2010.
- [5] 羽下哲司, 八木康史, “時空間動き特徴に着目した屋外侵入者監視技術に関する研究”, 情報処理学会研究報告 2006-CVIM-154, No.51, pp.259-274, 2006.
- [6] Shi, J. and Kanade, T., “Good Features to Track,” Proc. of CVPR, 1994.
- [7] Terabayashi, K., Hashimoto, Y. and Umeda, K., “Measurement of pedestrian groups using subtraction stereo,” Proc. of ISVC2009, Part II, pp.538-549, 2009.