

差分ステレオを用いた歩行者計測

○橋本優希 中西達也 寺林賢司 梅田和昇 (中央大, JST CREST)
鈴木慧 高汐一紀 (慶応大, JST CREST)

Measurement of Pedestrians Using Subtraction Stereo

*Yuuki HASHIMOTO, Tatsuya NAKANISHI, Kenji TERABAYASHI, Kazunori UMEDA (Chuo Univ., JST CREST), Kei SUZUKI, Kazunori TAKASHIO (Keio Univ., JST CREST)

Abstract— In this paper, measurement of features of a pedestrian and counting of pedestrians in a group using “subtraction stereo” are discussed. Subtraction stereo is a stereo vision method that focuses on the movement of objects to make a stereo camera robust and produces range images for moving regions. Features of pedestrians such as 3D position, height and width are measured from range images obtained by subtraction stereo. Then a simple method to remove shadows and a method to count pedestrians in a group are proposed. The basic algorithm of the subtraction stereo is implemented on a commercially available stereo camera, and the effectiveness of the method to detect pedestrians with removal of shadows and the method to count pedestrians are verified by experiments using the stereo camera.

Key Words: Stereo Camera, Motion, Detection of Pedestrians, Surveillance Camera

1. 序論

ステレオ視に関しては、これまでに数多く研究開発が行われてきた [1]。最近では、CPU や GPU (Graphics Processing Unit) の高速化に伴い、比較的容易に PC を用いた実時間でのステレオ画像取得が可能となっている。さらに、車載用などに実用化されているステレオカメラもある [2]。

本研究では、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、実用的なステレオカメラの開発を行うことを目標とする。これらの用途では、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、我々は運動情報に陽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 (Subtraction Stereo)”を提案している [3]。単眼のカメラで歩行者の検出などを行うシステムは色々提案されているが [4]、ステレオカメラは、スケーラブルな利用が可能となるため、これらの用途に有意義であると考ええる。

本論文では、差分ステレオを用いた歩行者検出に関して論じる。歩行者の 3 次元特徴の計測手法を示し、さらに影の除去及び集団のおおよその人数の特定を行う。

2. 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 に示す。通常のステレオ視では、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、まず左右カメラそれぞれの入力画像から運動領域を抽出し、その後に抽出された運動領域に対して左右画像でのステレオマッチングを適用する。運動領域の抽出は、何らかの差分処理によって実現する。なお、対象は必ずしも文字通り運動している必要はなく、差分処理で検出可能であれば、静止していても差分ステレオは適用可能である。

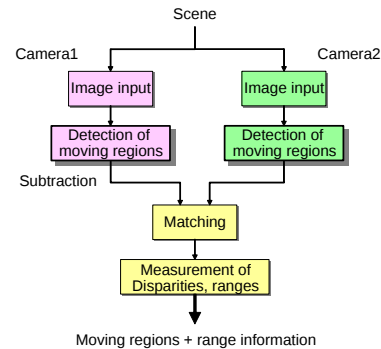


Fig.1 Flow of subtraction stereo

3. 歩行者の 3 次元特徴計測

本節では、差分ステレオによる歩行者などの物体の 3 次元特徴の計測に関して論じる。3 次元位置、高さ、幅などが計測可能である。

3.1 ラベリング

差分ステレオでは、得られる視差画像が元々運動領域に限定されているため、視差画像から移動する物体を抽出するのは容易であり、標準的なラベリング手法で行うことが可能である。その際、面積に関するしきい値処理により、ノイズの除去や抽出する対象物の大きさを限定することが可能である。

3.2 3 次元位置の計測

ある画素において、視差が与えられれば、光軸方向の距離を求めることができる。以後、距離という用語を、光軸方向の距離の意味で用いることとする。2 台のカメラが平行に配置されていると仮定し、視差を k 画素、距離を z とおくと、よく知られているように、距

離 z は以下の式で求めることができる．

$$z = \frac{\alpha}{k}, \quad \alpha = \frac{b \cdot f}{p} \quad (1)$$

ここで b は基線長（レンズ中心間の距離）， f はレンズの焦点距離， p は 1 画素あたりの横幅である．

さらに，距離 z と画像座標 (u, v) （図 2 参照）から，3 次元位置 $\mathbf{x}$ が得られる．スキューがなく，各画素のアスペクト比が 1 である場合には， $\mathbf{x}$ は

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

で与えられる．ここで， (u_0, v_0) は画像中心の画像座標である．

3.3 座標変換

ワールド座標系におけるカメラの位置・姿勢が既知であれば，対象点のワールド座標系における位置 $\mathbf{x}_w$ が次式で計算される．

$$\mathbf{x}_w = R\mathbf{x} + \mathbf{t}. \quad (3)$$

ここで， $\mathbf{t}$ はカメラの位置を表す 3 次元ベクトル， R はカメラの姿勢を表す 3×3 回転行列である．

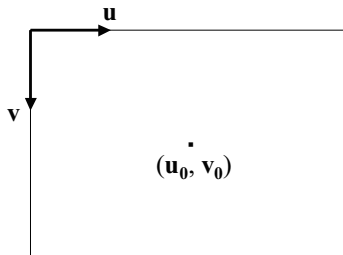


Fig.2 Parameters of an object in the image coordinate system

4. 影の除去

本研究では運動領域として歩行者を対象としている．日向の歩行者には影が生じる．図 3(a) では影が右側に伸びており，ラベリング処理で歩行者の領域と影がひとまとまりのラベルとして認識されるため，このままでは歩行者だけの高さ・幅の情報を取得することができない．そのため，影の除去が必要となる．

我々は単純な影の除去手法を提案する．本手法は影が歩行者の足元に生じるという経験則を用いる．まず画像座標系において，図 4(a) に示すように歩行者の領域の上端から下端までの長さを求め，上端から長さが $1/3$ の箇所における領域の左端，右端を求める．図 4(b) に示すように，この左端，右端で囲まれた箇所より外側の部分を除去する．図 3(a) に本手法を適用した結果が図 3(b) である．影が適切に除去されていることがわかる．

5. 集団の人数特定

本節では，差分ステレオによる集団の人数特定の手法に関して論じる．集団がステレオカメラで検出され



Fig.3 Removal of shadows ((a) Moving region in the sun (b) Moving region after the method to remove shadow is applied)

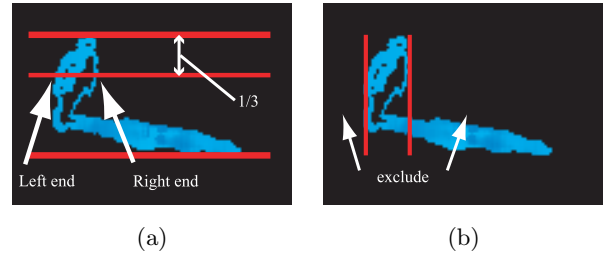


Fig.4 The method to remove shadows ((a) Detection of left and right ends (b) Removal of shadows)

ている間ずっと同一のラベルであると仮定する．このときの差分画像について，同一ラベルの画素の合計を面積 S ，集団を代表する点（例えば重心）までの距離を z とおくと，面積 S と距離 z の関係は以下の式で表せる．

$$constant = S \cdot z^2 \quad (4)$$

ここで $constant$ は定数である．式 (4) より 1 人のときの定数 $constant$ をあらかじめ求めておき，これを c とおくと，集団の中で人物内の領域の重複がない場合，集団の人数 $number$ は以下の式で求められる．

$$number = \frac{S \cdot z^2}{c} \quad (5)$$

ただし， $number$ は小数第 1 位を四捨五入した整数値とする．

本研究では距離 z を画素全ての距離の平均値とする．

6. 検証実験

本節では，歩行者検出手法を検証するための実験結果を示す．

6.1 市販ステレオカメラへの実装

提案した差分ステレオの基本アルゴリズムを，市販のステレオカメラに実装した．カメラには，Point Grey Research 社製の Bumblebee2（カラー， $f=3.8\text{mm}$ ）を用いた．画像サイズは 320×240 画素とした．

現段階では，運動領域の検出には単純な背景差分を用いている．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．距離の計測誤差については [5] で評価している．

Table 1 Success rate of removing shadows

time period	total[%]	one person[%]
11- 12	75	100
12- 13	80	100
13- 14	90	100
14- 15	90	100
total	83.8	100

図 5 に示すように、建物から歩行者を見下ろせるようにステレオカメラを設置した．具体的には、図 6 に示すように、高さが 8.3m の位置で、下向きに 50° だけティルトさせている．



Fig.5 Experimental scene

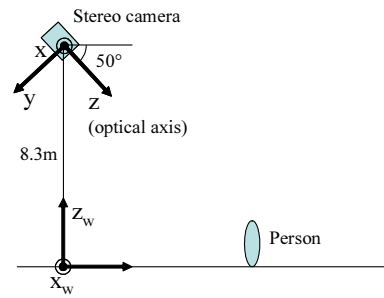


Fig.6 Experimental setup of the experiments

6.2 影の除去手法精度評価実験

日向において、影の除去手法を用いて、本手法の精度評価実験を行った．実験内容は以下のとおりである．

- キャンパスを歩く学生を撮影する．
- 歩行者をリアルタイムで検出する．
- 時間帯を分けて実験する：11～12 時，12～13 時，13～14 時，14～15 時の各 20 回，合計 80 回行う．
- 日向にいる間影を除去し続ければ成功とする．

時間帯を分けたのは、時間によって影の長さや向きが変わるためである．11 時から 12 時の間はほとんどが日向だったが、14 時から 15 時の間は日向が 1/4 に減少した．結果を表 1 に示す．これより、1 人のときは完全に影の除去が成功していることがわかる．失敗要因としては、図 7 に示すように、画像座標系において複数の人物の体や影が重なり、同一のラベルとなることが挙げられる．

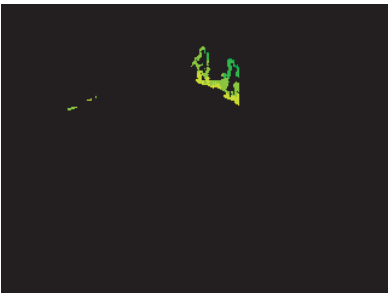
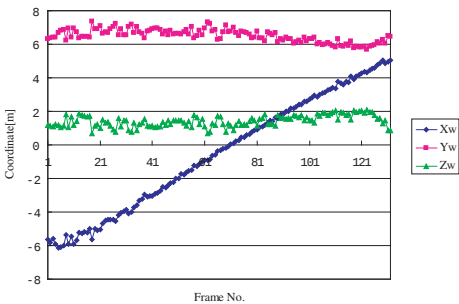


Fig.7 An example of failure of removing shadows

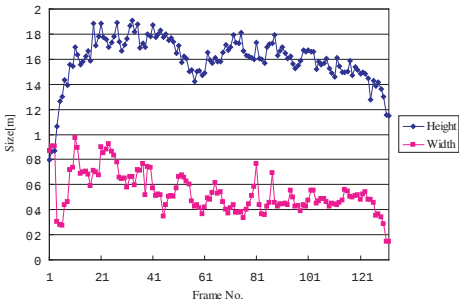
6.3 移動物体計測実験

日向における歩行者の計測実験結果を示す．左から右へ歩く歩行者の 3 次元位置，高さ，幅を時系列で計測した．結果を図 8 に示す．45 フレーム目以降が日向である．図 8(a) は歩行者の 3 次元位置を、図 8(b) は高さ・幅を示している．なお、3 次元位置は、図 6 に示すワールド座標系での値に変換してある．また、カメラが 50° ティルトしている分を補正するため、得られた高さは $\cos 50^\circ$ で除している．歩行者の運動が適切に計測されており、得られた高さ・幅も妥当な値と言える．幅が振動しているのは、歩行時の足の振れに対応したものである．

実験結果より、45 フレーム目以降に生じた影による影響を受けていないことがわかる．本手法を用いることによってほぼ安定して影を除去することができた．



(a) 3D Position



(b) Height and width

Fig.8 Experimental results of measuring a pedestrian in a general outdoor scene

6.4 人数特定の精度検証実験

差分ステレオによる人数特定の精度検証実験を行った。まず、定数 c を求めるにあたり、身長 180cm、幅 40cm の人物がカメラに対してほぼ垂直方向に手前から奥へと歩いたときの定数 $constant$ の時系列の変化を調べた結果を図 9 に示す。これより定数 $constant$ は大体 140,000 前後の値を取っているため、定数 c を 140,000 とする。ここで距離 z と 1 画素あたりの横幅 p がわかれば式 (2) より画像座標における人物領域の画素数、つまり面積 S が求まり、定数 $constant$ の理論値が求まる。 p は $7.4\mu\text{m}$ であり、定数 $constant$ の理論値を 140,000 とすると、断面積 0.52m^2 、身長 180cm なら平均幅 29cm に相当する。

図 10 は視差画像である。図 10(a) は 2 人組が右から左へ歩いているところで、(b) は 10 人のまとまった集団が右から左へ歩いているところである。集団が赤色の矩形で囲まれているが、これは集団が同一のラベルであることを示している。これらに人数特定の手法を用いた結果を図 11(a), (b) に示す。ただし、 $number$ の変化がわかりやすいように四捨五入せず、実数で示している。

実験結果より、本手法を用いることによってある程度の人数を特定できていることがわかる。また、カメラに対して人物の重なる部分が増えるとその分面積が小さくなり、実際の人数よりも特定する人数が少なくなってしまう。

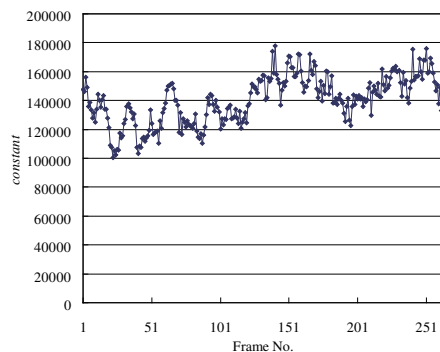


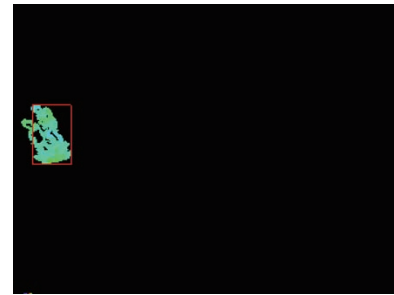
Fig.9 Time-series change of “constant”

7. 結論

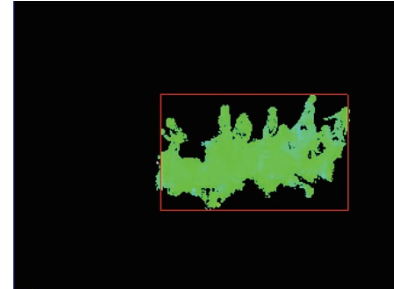
本論文では、差分ステレオを用いた単純な影の除去手法及び集団のおおよその人数を特定する手法を提案した。そして、市販のステレオカメラを利用して検証実験を行い、提案手法が歩行者を検出し 3 次元計測を行うこと及び個々人の分離が難しい集団のおおよその人数を特定するのに有効であることを示した。

今後の課題は、基本アルゴリズムなどの改良及びスタンドアロンで稼動する小型ステレオカメラシステムの構築である。

- [1] R. Hartley and A. Zisserman: “Multiple View Geometry in Computer Vision,” Cambridge Univ. Press, 2000.
- [2] 十川能之, 塙圭二: “ステレオ画像認識による車両前方監視システム”, 自動車技術, Vol.56, No.4, pp.34-39, 2002.

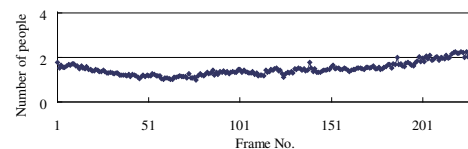


(a) Experimental scene 1 (2 persons)

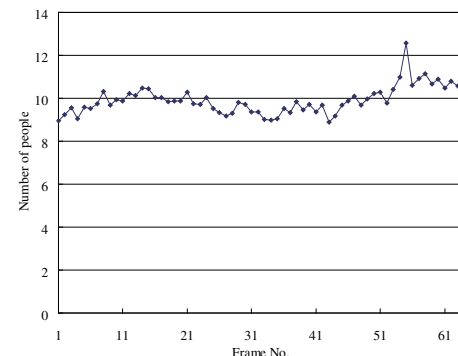


(b) Experimental scene 2 (10 persons)

Fig.10 Disparity image of the experimental scene



(a) Results for experimental scene 1



(b) Results for experimental scene 2

Fig.11 Experimental result of counting pedestrians

- [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 1 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予稿集, 1N31, 2007.
- [4] 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-D2, No.5, pp.1104-1111, 2004.
- [5] 和田正樹, 梅田和昇, 中西達也, 橋本優希: “運動領域注視型ステレオカメラシステムの開発”, 日本機械学会ロボットクス・メカトロニクス講演会 2008 講演論文集, 2P1-G10, 2008.