

差分ステレオを用いた人流計測

Measurement of pedestrian flow using subtraction stereo

星川佑磨† 生形徹† 有江誠† モロアレサンドロ†
寺林賢司† 梅田和昇†

Yuma Hoshikawa†, Toru Ubukata†, Makoto Arie†, Alessandro Moro†

Kenji Terabayashi†, Kazunori Umeda†

†中央大学 / JST CREST

†Chuo University / CREST, JST

E-mail: hoshika@sensor.mech.chuo-u.ac.jp

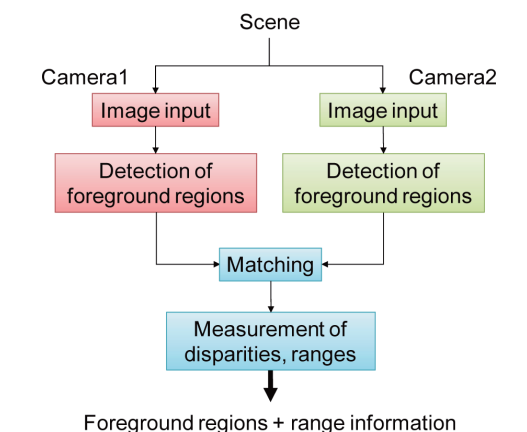
Abstract

我々は、背景差分により検出した前景領域に対しステレオマッチングを行う差分ステレオ法を提案している。本研究では、差分ステレオ法により得られる三次元情報を用いて検出領域内の人数を推定し、この検出領域をカルマンフィルタによりトラッキングすることで人の流れを計測する手法を提案する。また、検出領域毎における人数の推定では、差分ステレオと Harris のコーナー検出から得られる三次元情報を持つ特徴点の分布に着目し、画像上での人物同士の重なりを考慮した人数推定手法を提案する。

1 序論

近年、交通網の混雑緩和や災害時の人々の誘導などを目的に、カメラやレーザレンジファインダなどのセンサを用いて街やオフィスなどの人の流れを計測する研究が盛んに行われている[1][2]。これらの研究の多くでは、人の形や大きさなどの特徴から人物の識別器を作成し、これを用いて個々の人物を検出し、その位置と速度から方向毎の人数を算出して人の流れを計測している。しかし、その精度は個々の人物の検出精度に依存するため、人物同士が重なり検出が困難な混雑した環境下ではパフォーマンスが低下する。

また、近年 CPU や GPU の発達によりリアルタイムでの距離計測が可能となったステレオカメラを用いた人流計測手法が提案されている[3][4]。これらの手法の多くでは、ステレオカメラにより得られる三次元情報を用いて人物を検出し、フレーム間での三次元座標の推移から人の流れを計測する。しかし、これらの手法では人物同士の重なりが生じないよう人物を真上から俯瞰するように天井にステレオカメラを設置したり、複数のカメラを要



(a) Basic algorithm of the subtraction stereo



(b) Input Scene

(c) Output disparity image

Fig.1 Subtraction Stereo

するなどカメラの設置環境を限定されるという課題がある。

我々は、先行研究において移動物体を三次元計測する差分ステレオと呼ばれる手法を提案している[5]。また、本手法を用いた歩行者の検出・追跡手法を提案している[6]。しかし、検出領域内の人の数を推定する際に、画像上での人物同士の重なりを考慮していなかった。本研究では、差分ステレオにより検出される各領域をトラッキングし、領域毎の人数を画像上での人物同士の重なりを考慮しながら推定することで人の流れを計測する手法を提案する。

2 差分ステレオ

差分ステレオの流れを Fig.1(a)に示す. まず左右の画像それぞれにおいて, 背景差分により前景領域を抽出する. ここで得られた前景領域のみに対してステレオマッチングを行うことで, 前景領域のみに限定した視差画像を得る. Fig.1(b)のシーンに対して得られる視差画像を Fig.1(c)に示す. ここで Fig.1(c)において, 視差画像の色はカメラから対象までの距離を示し, 緑が近く, 黄色, 赤になるにつれて遠くを表している. Fig.1(c)より, 人物の領域のみにに対して視差情報が得られていることが分かる.

3 人流計測手法

本研究では, 差分ステレオにより検出した人物領域をトラッキングし, 各領域の人数を推定することで人の流れを計測する. 領域のトラッキングには, カルマンフィルタ[7]を用いる. 領域内の人数の推定には, 領域の面積とカメラからの距離を用いる. この際, Harris のコーナー検出[8]により検出される特徴点の分布から, 領域内の人物同士の重なりを考慮して人数の推定を行う.

3.1 人物トラッキング手法

人物の動きが概ね等速直線運動であるモデルを立て, カルマンフィルタに与える運動モデルとする. 状態変数 \mathbf{x} は, 世界座標系での人物の位置 (x, y, z) および速度 $(\dot{x}, \dot{y}, \dot{z})$ を用い, 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.

$$\mathbf{x}_{i+1} = \Phi \mathbf{x}_i + \boldsymbol{\omega}_i \quad (2)$$

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. $\boldsymbol{\omega}_i$ はプロセスノイズである. 観測変数には,

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

を用いる. u, v は画像上の座標, d は視差である. 観測変数 \mathbf{z} と状態変数 \mathbf{x} の関係は次式で表わされる.

$$\mathbf{z}_i = \mathbf{f}(\mathbf{x}_i) + \mathbf{v}_i \quad (5)$$

$$\mathbf{f}(\mathbf{x}_i) = \begin{bmatrix} x_i \cdot f & y_i \cdot f & b \cdot f \\ z_i & z_i & z_i \end{bmatrix}^T \quad (6)$$

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. これらの式から得られるカルマンフィルタの予測値と人物の観測値を対応づけることでトラッキングを行う.

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ユークリッド距離が, 閾値以内にありかつ最小のものを対応づける. しかし, 人物のすれ違いなど, 人物領域同士で重なるシーンでは, 対応づける観測値が複数の予測値で一致してしまうためトラッキングが出来ない. このような場合は, 観測値を用いずカルマンフィルタの更新を行い, 人物をトラッキングする.

3.2 検出領域内の人数推定手法

検出した領域内に複数の人物がいる場合には, その領域内の人数を推定する必要がある. 人物領域内の前景領域の合計画素数を面積 S , 集団までの光軸方向の距離を z とおくと, 面積と距離の関係は以下の式で表せる.

$$A = S \cdot z^2 \quad (7)$$

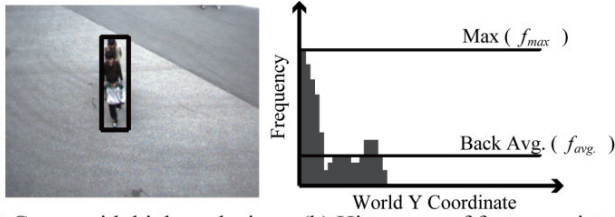
ここで A は定数である. 以後, この A を面積定数と呼ぶ. 集団の中で人物の重複がない場合, 集団の人数 n は式(7)より求まる一人のときの面積定数 A_1 を用いて以下の式で求められる.

$$n = \frac{A}{A_1} \quad (8)$$

以上はカメラの光軸を地面と平行にした場合を前提としている. しかし, 高所からカメラの光軸を下向きに設置し人物を計測すると, 画像下方に位置する人物ほど俯瞰して計測することになるため, 画像内の位置に応じて人物の見え方が異なる. 人物を俯瞰するほど画像内で人物はつぶれて見え, 取得可能な面積が減るため, 人の見え方に応じた人物の面積の補正が必要となる. 本研究では, カメラモデルを擬似透視投影と仮定し, 画像内の位置に応じた人物の面積の補正を行う. 面積の補正の式を以下に示す.

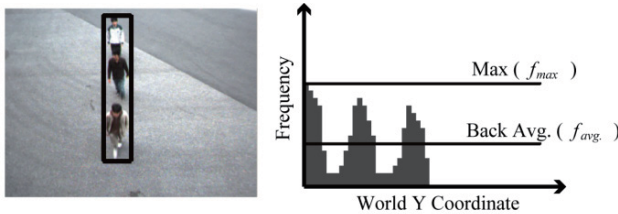
$$S' = \frac{S(1+k)}{\cos \theta - v \cdot \sin \theta + k} \quad (9)$$

$$v = \frac{y_c}{z_c} \quad (10)$$



(a) Group with high occlusion (b) Histogram of feature points

Fig.2 Example of histogram with high occlusion



(a) Group with low occlusion (b) Histogram of feature points

Fig.3 Example of histogram with low occlusion

ここで θ はカメラ光軸と水平軸のなす角度で, k は補正係数である. 補正係数 k を本研究では 0.5 とする. また, v は人物の正規化画像座標の縦方向の成分を表し, y_c にはカメラ座標系における人物の足下(領域の下端)の y , z_c にはカメラ座標系における人物の頭頂部(領域の上端)の z を用いる.

3.3 人物の重なりを考慮した集団の人数推定手法

Fig.2, Fig.3 に示す様に, 一つの検出領域内に複数の人物がいる場合, 人物同士が部分的に重なることがある. しかし, 3.2 節の人数推定手法では, 重なりで見えない部分の面積の補正は行っていないため面積の補正が必要となる. 本研究では, 3 次元情報を持つ特徴点の奥行き方向の分布から人物同士の重なり具合を定量化し, 見えない部分の面積の補正を行う手法を提案する.

差分ステレオにより検出される前景領域に対して三次元情報を持つ特徴点を抽出する. 本研究では, Harris のコーナー検出[8]により特徴点を抽出する. 各領域毎(集団毎)に特徴点の奥行き方向の分布を調べるために, まず領域毎に世界座標系(カメラの真下を原点)における $x-y$ 平面に特徴点を投影する. 領域の y 成分(奥行き方向)を等間隔に分割してセルを作成し, セルに対応する特徴点の数からヒストグラムを計算する. 人物同士の重なりが多い集団と重なりが少ない集団のヒストグラムの例を Fig.2(b), Fig.3(b)に示す. Fig.2(b)

に示す様に, 重なりが多い集団のヒストグラムでは手前の人物の位置に対応する位置にピークが発生し, 手前から二人目以降(領域の奥)に対応するセルでは全体的に頻度が少なくその平均値も低くなる. これは集団手前の人物からは多くの特徴点が算出され, 重なり隠れた奥の人物からは特徴点が算出されにくいからである. 一方, Fig.3(b)に示す様に, 重なりが少ない集団のヒストグラムでは手前から二人目以降の人物にも特徴点が算出されやすくなるため, 領域の奥に対応するセルでも頻度が多くその平均値も高くなる. このヒストグラムの特徴から領域内における人の重なり度を算出する.

ヒストグラム全体の最大頻度を f_{max} , 手前の人物に対応する部分を除いたヒストグラムの平均頻度を f_{avg} とし, 領域内における人の重なり度 o を次式により定義する.

$$o = \frac{f_{max} - f_{avg}}{f_{max}} \quad (11)$$

重なり度 o は小さい値ほど重なりが少なく, 大きいと重なりが多いことを表す. 本研究ではこのヒストグラムにおいて, 領域の手前の位置から 100cm の位置(ヒストグラムにおける手前から 10 個目のセル)に閾値を設け, これより手前に対応するセルで構成されるヒストグラムを手前の人物に対応するヒストグラムとしている.

式(11)より算出される重なり度 o を用いて見えない部分の面積を補正する. 重なりが生じている奥の人物の面積を S_{back} とし, これを重なり度 o を用いて次式で補正する.

$$S'_{back} = (1 + o) \cdot S_{back} \quad (12)$$

ここで S'_{back} は補正後の面積である. この面積 S'_{back} を用いると, 補正後の面積 S_c は手前に位置する人物の面積 S_{front} と合わせて,

$$S_c = S_{front} + S'_{back} \quad (13)$$

となる. なお, 手前の人物の面積は領域の下端(手前の人物の足下)の三次元位置と擬似透視投影モデルから算出し, 残りの面積を重なっている奥の人物の面積とする. この S_c を式(6)の S に代入し, 重なりが生じている際の集団の人数の推定を行う.

重なりが生じる集団(三人)に対し面積の補正を行った際の面積定数の変動を Fig.4 に, 実際に算出した人数の変動を Fig.5 に示す. 計測シーンは Fig.2(a)に示す集団が画像奥から手前に歩くシーンを選んだ. なお, 本検証では一人の際の面積定数を 32000 としたため,

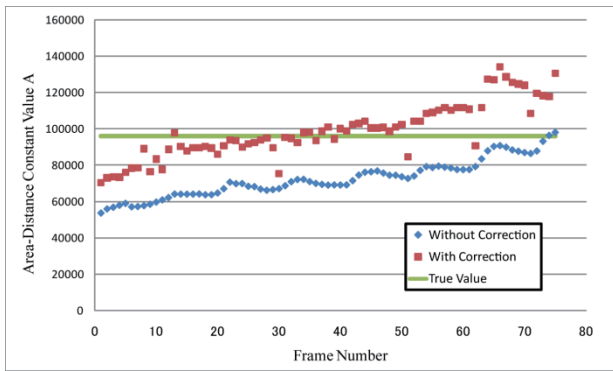


Fig.4 Example of correction of constant value A



Fig.6 Experimental Scene

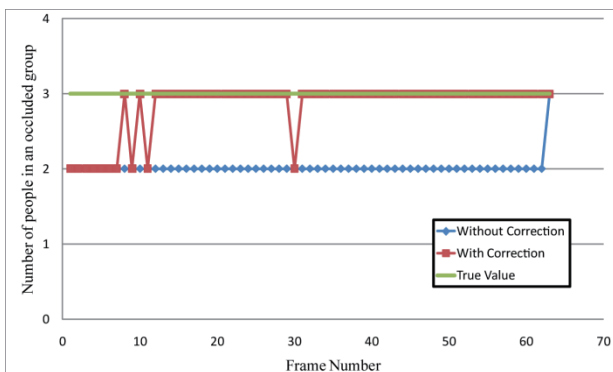
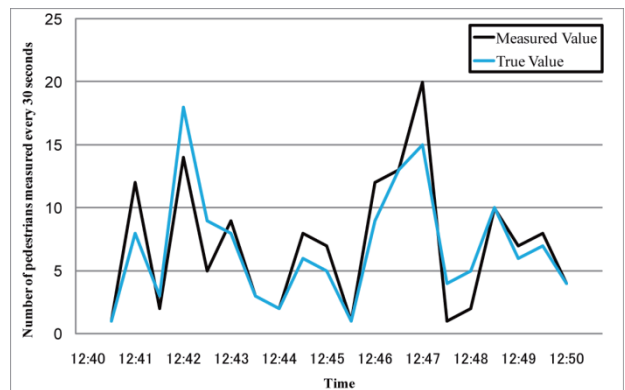


Fig.5 Output number of people in a group with occlusion

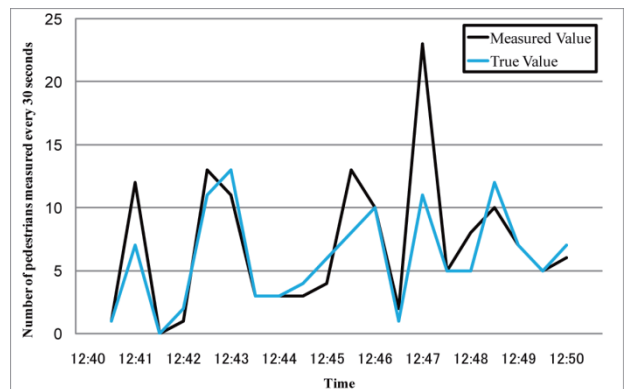
その三倍の 96000 を三人の際の面積定数の True Value として表示した. Fig.5 に示す人数は, 算出した面積定数を一人分の面積で割り小数第一位で四捨五入した結果である. Fig.4 の結果より見えなくなり減少した面積を補正し, 面積定数がより真値に近づいていることが分かる. また, Fig.5 より, 補正前は二人と計測されていた人数も補正により正しく三人と出力され, 人数推定の精度が向上していることが分かる.

4 人流計測実験

大学キャンパス内の校舎入り口における人の流れを計測した. 実験シーンを Fig.6 に示す. Fig.6 に示すように, 実験シーン内に仮想的なゲートを設置し, これを通過する人数を方向毎に 10 分間計測した. 本実験環境において, 人は集団や個人の切り分けが可能な程度の密度で歩行した. 実験に使用し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)である. カメラは高さ 5.5[m], 仰角 45[deg]に設定した.



(a) Number of pedestrians walking to the top



(b) Number of pedestrians walking to the bottom

Fig.7 Experimental result for the university campus scene

Table1 Experimental results

	Going Up	Going Down	Total
G.T.	137	121	258
Output	141	140	281
Accr. [%]	97.1	84.3	91.1

画像中で画像上方(奥)へ歩いた人数と, 下方(手前)へ歩いた人数を 12 時 40 分から 10 分間それぞれ 30 秒間ごとに集計した結果を Fig.7(a), (b)に示す. 両結果より, ゲートを通過した人数の推移が大まかに計測出来ていることがわかる. 真値との差が大きい時間帯では, ゲート付近でのすれ違いなどで人が密集した際に人物とトラッカーが誤対応を起こし, 同じ人物が複数回ゲートを通過したと判定してしまうことが人数に誤差が生じた原因の一つである. 大人数がゲート付近に集結すると, 一つの領域内で違う方向に歩いているにも関わらず大きな集団が一つの方向に移動していると認識してしまうことも, 方向毎の人数に誤差が生じる原因の一つである.

また, 10 分間合計での方向毎の人数計測の結果を真値(G.T.)と精度(Accr.)と併せて Table1 に示す. 結果より, 約 10%程度の誤差で人流計測が行えていることが分かる.

5 結論

差分ステレオにより検出される各領域をカルマンフィルタによりトラッキングし, 各領域の人数推定を行うことで人の流れを計測する手法を提案した. 領域内の人数の推定では, 前景領域から算出される特徴点の三次元分布を用いて画像上での人物同士の重なりを考慮する手法を提案した. 今後は, 個々の人物や集団の切り分けが困難な混雑環境下における高精度な人流計測を目指す.

参考文献

- [1] 中村克行, 趙卉菁, 柴崎亮介, 坂本圭司, 大鋸朋生, 鈴川尚毅: “複数のレーザレンジスキャナを用いた歩行者トラッキングとその信頼性評価”, 電子情報通信学会論文誌 D-II, Vol. J88-D-II, No.7, pp.1143-1152, 2005.
- [2] O. Sidla, Y. Lypetsky, N. Brandle, and S. Seer: "Pedestrian detection and tracking for counting applications in crowded situations," Proc. of the IEEE AVSS2006, p.70, 2006.
- [3] D.Beymer: "Person counting using stereo," Proc. of the Workshop Human Motion, pp.127-133, 2000.
- [4] K. Terada, D. Yoshida, S.Oe and J.

Yamaguchi: "A method of counting the passing people by using the stereo images," Proc. of the ICIP1999, Vol.2, pp.338-342, 1999.

- [5] 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.
- [6] 寺林賢司, 中西達也, 橋本優希, 星川佑磨, 川田昂次, 梅田和昇: “差分ステレオによる多人数歩行者の運動計測”, 2009 年度精密工学会春季大会学術講演会講演論文集, pp.185-186, 2009.
- [7] S. Thrun, W. Burgard, and D. Fox: "Probabilistic Robotics," MIT press, 2006.
- [8] C. Harris and M. Stephens: "A combined corner and edge detector," Fourth Alvey Vision Conference, pp.147-151, 1988.