

スマートフォン端末を用いた
トラッキング手法によるリアルタイム通行量計測
福島佑樹・関本義秀

**Real-time Pedestrian Traffic Counting using a Tracking Method on
Smartphones**

Yuki FUKUSHIMA and Yoshihide SEKIMOTO

Abstract: 近年、都市計画や防災・減災の観点から、街中における人々の行動をリアルタイムに把握するニーズが高まっている。街中における人々の行動を把握する一手法として、数取器と呼ばれるカウンターを用いた通行量調査が行われている。しかし、数取器を用いた計測では、広範囲かつリアルタイムな計測が困難である。そこで本研究では、スマートフォン端末のカメラ画像を解析することで、リアルタイムに通行量を計測する手法を提案する。

Keywords: 通行量計測 (pedestrian traffic counting), 画像処理 (image processing), トラッキング (tracking), スマートフォン端末 (smartphones)

1. はじめに

街中における人々の動きを把握することは、都市計画、防災・減災やマーケティングなど様々な分野に利用可能であり、ニーズが高く必要不可欠な情報となっている。また、人々の動きは、行動範囲内の環境変化に伴って刻一刻と変化するものであるため、常に最新のデータが求められる。

街中における人々の通行量計測は、数取器と呼ばれるカウンターを用いて人手で行ってきた。人手による計測は、精度が高い反面、計測する地点ごとに計測員を配置しなければならないため、常に最新のデータを計測することは困難である。また、ビデオカメラやレーダーなどの設置型センサを用いた通行量計測に関する研究も行われてい

るが、設置個所や設置方法が限定されているため広範囲を計測できないという課題がある。このように、カウンターや設置型センサを用いた従来の手法では、広範囲かつリアルタイムなデータを得ることができない。そこで、近年では、様々な高精度センサが搭載されているスマートフォン端末を用いて人々の行動を解析する研究が行われている。その中でも歩行者の移動推定においては、デットレコニング法、Wi-Fi・Bluetoothを用いた手法とGPSを用いた手法に大別される。しかし、これらの手法では、誤差が大きいという問題や歩行者の協力が前提という課題があり、全数を把握することは困難である。

そこで、本研究では、スマートフォン端末で歩道を撮影し、逐次得られるカメラ画像から歩行者通行量を算出することで、少人数かつ広範囲において、簡易にリアルタイムな人の動きを把握可能とするシステムを提案する。

福島佑樹 〒153-8505 東京都目黒区駒場 4-6-1

東京大学空間情報科学研究センター

Phone: 03-5452-6415

E-mail: yfukus@ccsis.u-tokyo.ac.jp

2. システムの概略

本研究では、スマートフォン端末のカメラから得られる画像を順次処理し、歩行者の通行人数をリアルタイムで計測する。本研究システムの概要図を図-1に示す。

スマートフォン端末を用いた歩行者通行量の計測システムは、歩道を撮影したカメラ画像を順次取得し、一定間隔で N. Dalal (2005) らが提案している Histogram of Oriented Gradients (HOG) 特徴量の計算および Support Vector Machine (SVM) による 2 値分類を行い人物領域の抽出を行う。また、抽出されたそれぞれの人物領域に対して、次フレーム以降トラッキングを行い、個別の歩行者を判定することで歩行者通行量を計測する。

3. 歩行者通行量の計測手法

3.1 計測条件

本研究では、スマートフォン端末のカメラ画像を用いて歩道上の歩行者数をリアルタイムで計測することを目的としている。また、スマートフォン端末は、計測の利便性を考え固定せずに計測者が手で持って計測することとする。そのため、手振れが発生する可能性がある。さらに、影や明るさなどの撮影環境は、計測箇所や天候に左右される。このように、スマートフォン端末を用いて歩行者通行量を計測するには考慮すべき点が多い。一方、スマートフォン端末の計算処理速度は、一般的なパーソナルコンピュータに比べると劣っているため、スマートフォン端末でリアルタイムに処理が可能な手法を採用する必要がある。

3.2 人物領域の抽出

本研究では、撮影環境の課題とリアルタイム処理に対応するため、特徴量として HOG を用いる。HOG は、局所領域であるセルに対して輝度の勾配ヒストグラムを作成し、セルの集まりであるブロックごとにヒストグラムを正規化し、特徴量とす

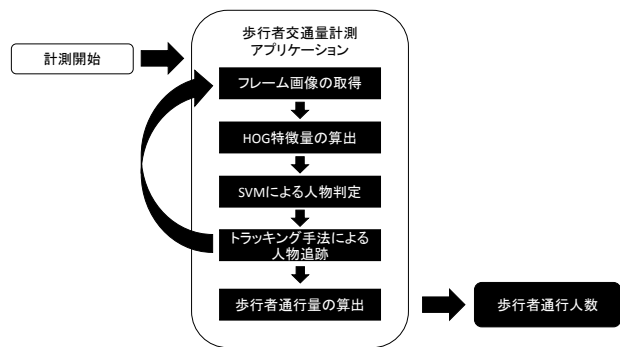


図-1 システムの概要

る手法である。画像の座標 (x, y) における輝度値を I としたとき、勾配強度 m と勾配方向 θ を次式 (1a) で算出する。

$$\begin{aligned} f_x(x, y) &= I(x + 1, y) - I(x - 1, y) \\ f_y(x, y) &= I(x, y + 1) - I(x, y - 1) \end{aligned}$$

$$\begin{aligned} m(x, y) &= \sqrt{f_x(x, y)^2 + f_y(x, y)^2} \\ \theta(x, y) &= \tan^{-1} \frac{f_y(x, y)}{f_x(x, y)} \end{aligned} \quad (1a)$$

式(1a)で算出した勾配強度 m と勾配方向 θ を用いてセルごとに勾配方向ヒストグラムを生成する。そして、セルのヒストグラムをブロック単位で L2-Norm を用いて正規化する。

以上のように HOG 特徴量は、勾配方向ヒストグラムをブロック単位で正規化するため、日照条件などの変化に頑健である。また、本研究では、人物領域を検出するための識別器として SVM を用いる。特徴量として選定した HOG は、非常に次元数が多い。HOG の次元数 \dim は、画像の幅を W 、画像の高さを H 、セルサイズを S 、ブロックサイズを B 、勾配方向数を θ とした場合次式(1b)で表される。

$$dim = \left(\frac{W}{S} - B + 1\right) \cdot \left(\frac{H}{S} - B + 1\right) \cdot B^2 \cdot O \quad (1b)$$

式(1b)より、50×50の画像において、セルサイズを5、ブロックサイズを3、勾配方向数を9とした場合の次元数は、5184となる。このようにHOG特徴量は次元数が大きい。そのため、本研究では、次元数の大きい特徴に対して識別精度の高いSVMを用いた。ただし、人物領域を検出する場合、識別回数が非常に多く計算時間がかかる。これは、画像中の人物位置及び大きさが未知であるためである。そこで、本研究では、リアルタイム計測を可能とするため、撮影方法を指定することで人物領域を既知のものとし、高速化を図った。具体的には、計測時にスマートフォン端末の画面に補助線を横方向に2本表示し、その補助線の間人が収まるように撮影を指定することで、人物の大きさを固定した。

3.3 人物領域のトラッキング

HOG特徴量とSVMを用いて人物領域は抽出可能であるが、撮影方法を指定する場合でも計算コストが高く、スマートフォン端末上でのリアルタイム処理はできない。そのため、本研究では、抽出した人物領域に対してトラッキング手法を用い、人物領域を追跡することでリアルタイム処理を実現する。

本研究では、トラッキング手法としてテンプレートマッチングを用いる。テンプレートマッチングは、探索元の画像が探索先の画像内のどの部分に出現するか判定する手法である。本来テンプレートマッチングは探索の計算コストが高いが、歩道を歩く歩行者を対象とした場合、次フレーム内での歩行者の位置は前フレームにおける人物領域の左右どちらかに限られるため、前フレームにおける人物領域付近を探索すれば良い。また、歩行者は、基本的に一方項に移動するため、次々フ

レームではさらに探索領域が絞られる。さらに、歩行者が突然大きく、または小さくなることもないため、拡大縮小しての探索も必要ない。よって、本研究におけるトラッキングにはテンプレートマッチングを採用した。具体的には、抽出された人物領域の上半分の領域をテンプレート画像とし、次フレーム以降で随時マッチングを行う。そして、5フレーム間トラッキングできた場合に歩行者として判定する。ただし、判定後もフレーム内に出現する可能性があるため、次の人物領域の抽出時にトラッキングに用いたテンプレート画像と抽出された人物領域とマッチングを行い、同一である場合は、同一人物として無視する。

4. 実証実験

4.1 実験条件

本実証実験における実施箇所を図-2に示し、実験環境を表-1に示す。実験場所を新橋虎ノ門地区とした理由は、現在2020年東京五輪開催を見据えた再開発が行われており、今後人の流れが大きく変化する可能性が高いためである。

4.1 実験結果

歩行者通行量の計測は、計測地点で5分間計測を行い、計測終了後に別の計測地点を順次移動して計測することで行った。この地点Aから地点Dの計測地点をそれぞれ2度計測したデータを本実験の計測結果とした。また、別のカメラで計測中の状況を撮影した動画から人手で歩行者人数を数えた値を真値とした。本研究のシステムを用いた計測結果と真値との比較結果を表-2に示す。表-2より、地点Aや地点Bの比較的歩行者が少ない計測地点は高精度な結果となっているが、地点Cや地点Dの歩行者の多い計測地点では精度が低くなった。歩行者の多い計測地点において精度が低い原因として、HOGとSVMによる人物領域の抽出に失敗している例とトラッキングに失敗している例があった。前者は、フレーム内で終始複数人が重なって歩行していた場合に、複数人を一人

の人物として領域を抽出していることが原因と考えられる。また、後者は、左右から来た歩行者が重なった場合など、一時的にトラッキング対象の歩行者が隠れた際に正常に追跡できていないのが原因である。人物領域の抽出および追跡の失敗については、人物領域に複数人含まれているか判定する処理などを追加することで精度向上が可能であると考えられる。

5. おわりに

本研究では、簡易かつリアルタイムに歩行者の通行量を計測することを目的として、固定せず handheld 状態のスマートフォン端末から通行量を計測する手法を示した。これにより、本研究のシステムを用いれば、少人数であっても広範囲の人々の動きを推定することが可能であることを明らかにした。しかし、歩行者の多い計測地点では誤差が大きくなるという課題がある。処理を追加することで精度向上は可能であると考えられるが、目的としているリアルタイム計測が困難となる可能性もある。今後は、リアルタイム計測と高精度が両立する解決手法を考案より良い歩行者人数計測システムを目指す。

参考文献

- J. Zhang, L. L. Presti, and S. Sclaroff, 2012. Online multi-person tracking by tracker hierarchy, *Proc. of the IEEE International Conf. on AVSS*, 379-385.
- M. D. Breitenstein, F. Reichlin, B. Leibe, E. Koller-Meier, and L. V. Gool, 2011. Online multiperson tracking-by-detection from a single, uncalibrated camera, *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 33, 9, 1820-1833.
- N. Dalal and B. Triggs, 2005. Histograms of Oriented Gradients for Human Detection, *Proc. of the IEEE CVPR*, 1, 886-893.



図-2 計測地点

表-1 実験環境

項目	詳細	
日時	2016年8月23日(火) 10時~11時	
場所	東京都港区新橋 JR新橋駅付近	
使用端末	名称	Motorola Moto G 4G
	OS	Android 4.4
	CPU	Qualcomm Snapdragon 400 processor 1.2 GHz quad-core
	Memory	1GB
	Camera	1280 x 720
処理速度(平均)	89ms/Frame	

表-2 真値とシステムによる計測結果の比較

地点	開始時刻	真値(人)		システム(人)		誤差率(%)	
		左	右	左	右	左	右
A	10:02	82	16	71	14	-13.41	-12.50
A	10:44	76	14	69	17	-9.21	21.43
B	10:10	9	65	10	52	11.11	-20.00
B	10:53	11	59	15	51	36.36	-13.56
C	10:21	21	105	17	84	-19.05	-20.00
C	11:01	29	91	12	72	-58.62	-20.88
D	10:30	46	127	29	94	-36.96	-25.98
D	11:09	33	122	35	101	6.06	-17.21