

高視点撮影映像における歩行者カウントの検討

Investigation of Detection Method for Pedestrians in the Image Sequence taken from the High Viewpoint

目黒 徹† 曲山 良佑‡ 山内 俊明† 関 靖夫†
Toru Meguro Ryosuke Magariyama Tosiaki Yamanouchi Yasuo Seki

1. まえがき

横断歩道等における歩行者の通行量の長期的データの取得は、横断歩道や歩道橋等の都市計画等の应用到有効と考えられる。歩道者を横断歩道の横方向から見ると、友人等並んで歩いている人同士を分離するのは容易ではない。しかし、交差点の電信柱上などの高い視点からカメラで撮影すれば、人間同士が並んで歩いても容易に分離することができるので、より正確なカウンティングが期待される。本研究では、このような高視点から撮影した映像に基づいて、精度良く人間（主に頭に注目）を検出することを狙いとして検討を行った。

具体的には、人間の頭を囲む四角い枠（大きさを規定）を想定し、その中の画像の分布により人間を検出するトップダウン手法を用いた。その際、影の影響や近接した人物の影響を減殺するための方法を検討した。

2. 前処理

トップダウン手法を画像に適用する前の前処理として、背景を除去するための背景差分と対象候補の部位の位置情報を把握するために必要なラベリングを行った。

まず背景差分であるが、差分を取るために必要となる背景画像は、同じカメラアングルで撮影した多数枚の画像より推定した。推定手法は、時刻 t に撮影した画像の座標 (x, y) のピクセル値を

$$Image(t, x, y) = (Image_R(t, x, y), Image_G(t, x, y), Image_B(t, x, y))$$

、推定した背景画像の座標 (x, y) のピクセル値を $BG(x, y) = (BG_R(x, y), BG_B(x, y), BG_G(x, y))$ とすると、下式のようなになる。

$$BG_R(x, y) = \text{mode}\{Image_R(t, x, y), \Lambda, Image_R(t+n, x, y)\} \quad (1)$$

但し、 $\text{mode}\{\}$ は最頻値を表す。また、 $BG_G(x, y), BG_B(x, y)$ についても同様である。

このようにして推定した背景画像と処理対象画像を用いて、以下の方式で背景差分を求めた。

(1) 背景画像と処理対象画像の同じ座標のピクセル値のユークリッドノルムを算出する。

(2) 上記で算出したノルムがあらかじめ定めておいた閾値よりも大きい場合にはノルム自身の値を、そうでない場合には0を当該座標の背景差分値とする。

なお、この背景差分は引き続き行われるラベリング及びトップダウン手法の適用対象となる。一般に、ラベリングの対象データは二値化されたものであることが多いが、ここではあえてノルムの値をそのまま残した。これにより、トップダウン手法を適用した際に、人物の影の部分や近接している人物の各部位の判別ができるものと期待されるか

らである。

次にラベリングであるが、これは前述のように閾値を越えたノルムの値を入力とし、0以外の正値を有するピクセルを通常の場合の1の値を持つピクセルと見なして処理を行った。得られた各ラベルについて、 x 軸及び y 軸の最大値と最小値を求め、その差の絶対値を当該ラベルの幅及び高さとして定義することにした。そして、この幅と高さを予め定めておいた閾値と比較し、これを下回るものについては雑音と見なして除去した。

3. トップダウン手法

直上に近い高視点からある一定の領域内を通過する人物を撮影した場合、得られた画像に含まれる人物は幅や高さがほぼ一定と見なすことができる。トップダウン手法はこれを利用し、画像中の人物をある一定の幅と高さを有する矩形としてモデル化し、その位置を特定する手法である。このトップダウン手法を適用するデータは、前処理で得られた背景差分値とラベリングの結果であり、両者を相補的に用いることにより「多値の背景差分の連結領域」を矩形モデルで探索することができるのである。この探索のために、矩形内の背景差分値の分布と人物の位置との関係を重み関数としてモデル化した。さらに、矩形内の背景差分値と重み関数の積和をモデル適合度として定義し、人物の位置推定の指標として用いることとした。

3.1 矩形モデルと重み関数

今回は、幅と高さが共に40ピクセルという矩形をモデルとした。この値は、今回の解析対象画像に起因するものであるが、近接する人物を分離して表示することも考慮してやや小さめに設定した。

一方、この矩形内の背景差分値に対する重みとして用いる重み関数に課される条件は、モデル内の中央に近いほど画素に重みを持たせ端に行くほど画素の重要性を低くすることである。これは、矩形中央に頭部が存在する場合にモデル適合度の値が大きくなるようにするためである。これにより、頭部が矩形中央にありさえすれば、矩形内の人物の向きに関係なくモデル適合度の値が大きくなるのが期待される。

このような形状の関数は多数存在するが、今回は次式のような単純な関数を正規化したものを用いた。

$$f(x, y) = \left(1 - \left|y - \frac{H-1}{2}\right| / (H/2)\right) \cdot \left(1 - \left|x - \frac{W-1}{2}\right| / (W/2)\right)$$

†神奈川工科大学, Kanagawa Institute of Technology

‡(株)日本デジタル研究所, JDL

3.2 モデル適合度による人物の位置推定

基本的には、前処理で求めた各ラベルの外接矩形の範囲内で矩形モデルを縦方向と横方向にずらし、ずらした位置のモデル適合度の値の大小で人物の位置を推定する。具体的には、モデル適合度の値が予め定めておいた閾値を上回るとともに、極大値となっている場合に当該矩形モデルの位置を人物の位置候補とするのである。さらに、既に検出済みの方の全ての人物位置との距離を算出し、そのいずれもが予め定めておいた閾値よりも大きければ人物位置とする。また、矩形モデルのずらし幅は最小単位の1ピクセルではなく、5ピクセルとした。これらは、いずれも近接した人物を分離し、同一人物を2人以上に誤認しないようにするための方策である。

4. 実験

建物の高所から真下に人通りのある地点を撮影して得た画像を用いて本手法の有効性を検証する。検証のための画像は建物の出入り口を高所から無作為に撮影した30枚とし、またその際に天候や時間によって明度に変化があるため、背景差分及びモデル適合度の閾値をある程度変化させ1枚の画像についても複数回の検出を行った。

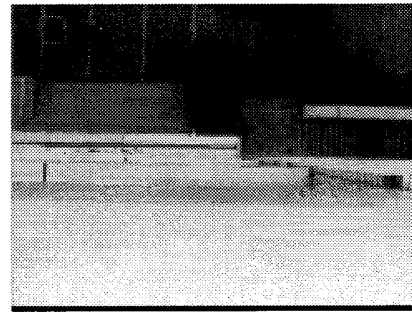
まず、対象となる人物が離れている場合については正確な検出結果が得られている。しかし、(a)の右側の人物の様に服装や、髪の色が背景と似ていた場合十分な背景差分結果が得られず人物として検出されていないことがわかる。しかしこの場合は背景差分の閾値、モデル適合度の閾値あるいはその両方を変えることで(b)の様に正確な検出結果が得られている。

次に対象となる人物が近接していた場合だが、(c)のように正確な検出ができなかった結果が多く見られた。その原因として先ほど述べた理由の他に、場所によっては影が濃く写り減殺できていない等の複数の原因が重なった為と考えられる。

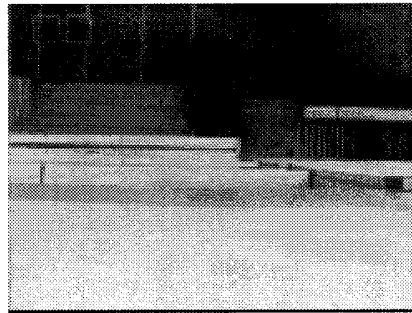
さらに(d)の画像については1人物から2箇所誤って検出されている。その原因として、背景差分を得る際に人物のある部分が背景色と似ていることで人物が分断され、それがラベリング及び人物推定に影響(e)し、結果として1人の人物から2箇所を誤って検出してしまったことが考えられる。これについても閾値を変化させることで正確な検出結果が得られている。

5. むすび

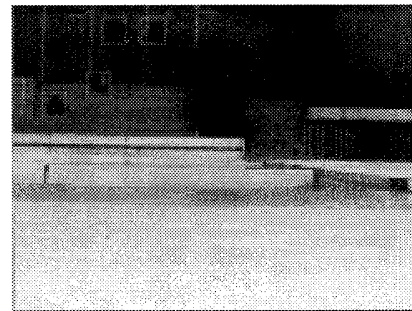
本稿では高視点からの撮影画像にトップダウン手法を用いて正確に歩行者のカウントを行う手法を提案した。実験の結果、全画像のうち約6割が正確に検出されており、本手法が歩行者のカウントについて有望であることがわかった。しかし、課題であった人物が近接した場合の検出については満足な結果が得られなかった。今後の課題として、より正確な検出の為に各パラメータの最適値の選定、新たな手法を取り入れるなどで影の減殺の精度を高めること、また今回は曇天での実験であった為、他の天候の場合への適用などが挙げられる。



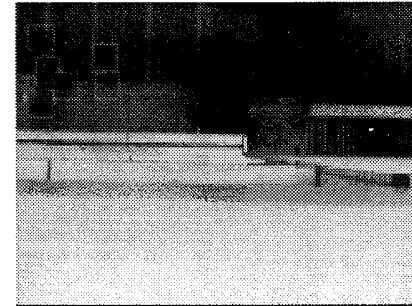
a. 閾値変化前



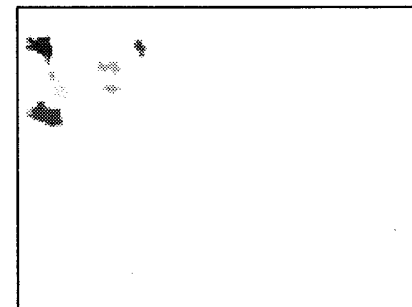
b. 閾値変化後



c. 検出されなかった画像



d. 誤って検出された画像



e. ラベリング結果