

人物モデルと縦方向フィルタリングを用いた実時間人物計数システム

中上 友宏 松尾 啓志

名古屋工業大学電気情報工学科

本稿では、複数人物が断続的にカメラ視野に侵入する環境にも対応し、かつ高速な人物検出、計数システムを提案する。様々な方向を歩く人物に対応した上半身人物モデルを構築し、画像の縦方向フィルタリングと再サンプリングを行うことにより、汎用 PC での実時間(每秒 30 回)人物検出、計数を実現した。

1. はじめに

動画像中の人物領域検出は、セキュリティ、人流調査、ヒューマンインターフェイスなどのシステムの基礎技術として重要である。特に近年では、巨大ビル、都市の雑踏[1]、競技場など、多数の人物が断続的に出入りする環境にも対応したシステムの需要も高まってきている。しかし、検出対象の増加によって、以下に挙げる 2 つの課題を解決する必要がある。

課題 1 . 検出に必要な総計算コストの抑制

課題 2 . 複数人物の重なり(交差、隠蔽)の検出

動物体検出の代表的な手法として、オプティカルフロー[2]、Snake[3]を用いた手法が従来から研究されているが、これらの手法では単一物体の抽出においても多大な計算量が必要であるため、特殊なハードウェア等を用いない限り、課題 1 の解決は非常に難しい。

それに対し、簡便かつ高速な抽出法として、抽出対象の存在しない背景画像と入力画像の差分によって物体領域を求める背景差分が挙げられる。加藤らは人物を楕円体としてモデル化し、モデルと差分画像とのマッチングによって人物検出を行う手法[4]を提案している。内田らは背景差分とテンプレートマッチングを併用した手法[1]によって、都市の雑踏での人物検出、追跡を実現している。

しかし、加藤らの手法においては、人物歩行像は侵入方向に対して様々な形状を取り得るため、一定の形状モデルによるマッチングでは誤検出の可能性が無視できない。また、背景差分では移動物体の領域情報のみが得られるため、ラベリングなどの一般的な画像処理手法では課題 2 の解決は困難である。これに対し加藤ら、内田らの両手法は検出結果のフ

レーム間の対応付け、即ち追跡処理によって課題 2 の解決を図っているが、人物数の増加に対して対応付けの候補が指数的に増加するため対応付けが困難であり、また対応付けに必要な計算量が大きくなってしまふ。逆に、対応付け処理を行わず、差分画像の領域情報だけを用いて人物交差、隠蔽を抽出できれば、追跡処理における高速性、ロバスト性の向上も期待できる。

本研究では、差分画像を用いた新たな高速人物検出手法を提案し、更にその応用として人物計数システムを構築する。提案する人物検出手法は、様々な方向を歩く人物歩行像の上半身の統計データから人物の上半身モデルを構築し、モデルと差分画像とのマッチングによって人物検出を行うモデルベースの検出手法である。このマッチング処理では、重なりを生じた複数人物の検出が可能であり、課題 2 に対応する。更に、課題 1 を解決するため、マッチング処理の簡素化を目的とした、差分画像のフィルタリング処理も同時に導入する。

以下、2 章では本研究で提案する人物検出手法の詳細を、3 章では構築した人物計数システムの詳細を述べる。そして 4 章で評価実験を行い、本システムの有効性について検討する。

2. 人物検出手法

2.1 人物モデル構築

本研究では、様々な方向を歩く人物歩行像(以下、シルエット)の上半身部分における領域変動の安定性に着目し、領域分布の統計データから検出に用いる人物モデルを構築する。

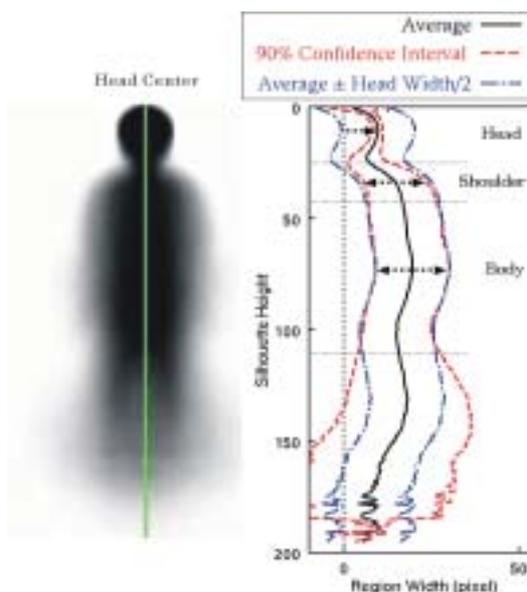
まずカメラを床面の水平方向に設置し、カメラに対して前後、左右、斜めの全 8 方向に歩く成人男性の動画を撮影する。そして各歩行方向で 2 歩分の動

画(約1秒)から時間等間隔に16枚のシルエットを抽出する。このデータを6人分、合計768のシルエットを重ね合わせ、領域の頻度を濃淡によって示した画像(重畳画像)を図1(a)に示す。このシルエットの重ね合わせでは、一定の頭部幅による各シルエットのサイズの正規化および、各シルエットの頭部中心の位置合わせを行った。

次に図1(b)は、図1(a)の右半身部分の各高さにおける領域分布を示したグラフである。図1(b)の各線は以下のデータを表す。

- ◆ 人物領域分布の平均 (Average)
- ◆ 人物領域分布の両側90%信頼区間 (90% Confidence Interval)
- ◆ 平均に対して正規化に用いた頭部幅の半分を加えた(引いた)値 (Average \pm Head Width/2)

ここで、図1(b)の上半身部分での領域分布の特徴について考える。まず第1の特徴として、肩から胴体部分にかけて、両側90%信頼区間(90% Confidence Interval)と、平均に対して頭部幅の半分を加えた(引いた)値(Average \pm Head Width/2)がほぼ一致している。次に第2の特徴として、上半身の主要な部位間(頭、肩、胴体)での90%信頼区間の変化に着目する(図1(b)の矢印)。頭部と肩にかけては、90%信頼区間の変化量が頭部幅以内に収まっている。そして、肩から胴体部分にかけては、90%信頼区間が非常に緩やかに変化している。



(a) シルエット重畳画像 (b) 人物領域の分布

図1 人物領域分布の様子

以上の特徴から、人物の上半身部分において、次の性質を導くことができる。

- 上半身の主要な部位間での領域変化量は、ほぼ人物の頭部幅以内に収まっている。

そこで本研究では、この上半身の領域分布の性質を上半身モデルと表現し、モデルと差分画像のマッチングによって複数人物検出を行う手法を提案する。このマッチングは、差分画像の1行分を1つの処理単位(1ライン)とし、差分画像の上部から下部方向に処理を行う。マッチング処理の手順を以下に示す。

- Step 1. 処理対象ラインに動領域が存在する場合、以下の処理を行う。
- (a) 真上のラインに動領域が存在しなければ、この動領域を新たな人物候補領域と認識し、次のラインの処理へ移る。
 - (b) 真上のラインで既に認識された領域があれば、Step 2へ。
- Step 2. ライン間での動領域の変化量(増分のみ)を求める。そして変化量の大きさと人物の頭部幅を比較し、処理対象ラインの動領域を次のように認識する。
- (a) 変化量が頭部幅未満の場合、各ラインの動領域は同一の人物の領域。
 - (b) 変化量が頭部幅以上である場合、その変化部分は別の人物の領域(複数人物が重なりを生じた状況)。

Step 3. 次のラインの処理に移る。

2.2 差分画像の縦方向フィルタリング

次に差分画像のフィルタリング手法について述べる。提案手法では、構築した人物モデルと差分画像とのマッチングによって人物検出を行う。しかし実際に得られる差分画像には、影や太陽、照明などの光、人物と背景色の類似による領域欠損など、多くのノイズが混入する。このため、画素単位での厳密なマッチング処理は、検出精度の点で困難であると考えられる。そこで、差分画像のノイズの影響を軽減するために、差分画像に対するフィルタリングが必要となる。しかし、一般的な膨張、収縮フィルタリングは計算量が大きく、実時間検出に用いることは難しい。一方、画素単位でのマッチング処理は計算量の点でも負荷が大きい。しかし、構築したモデルのマッチング処理に必要な情報は、人物の主要部位における領域の横幅情報であるため、差分画像の縦方向の分解能をある程度低下させ、マッチング処

理の高速化を図ることも可能であると考える。

そこで本手法では、これらの点に対応した差分画像のフィルタリング手法を導入する。まず得られる差分画像を縦方向に分割する。この処理では、上半身の各主要部位が、差分画像の各局所領域に含まれるように分割する。次に局所領域内での動領域ヒストグラムを求め、閾値処理によって領域内の高頻度部分を抽出する。このフィルタリング手法では、差分画像の各画素に対するアクセスが1回のみであるため、マスクを用いる一般的なフィルタリングより高速に処理することができる。また、マッチング処理に必要な情報を保持した上で分解能を低下させる、再サンプリング処理であるため、検出精度の低下を避けるのと同時に、マッチング処理の高速化が可能になる。このフィルタリングにより、2.1節で示したマッチング処理手順の処理単位が分割された局所領域、求める変化量の対象が抽出された高頻度部分にそれぞれ置き換えられる。

フィルタリング時の適切な分割幅は、カメラの設置条件、人物の身長などにより変化する。また、カメラにある程度の俯角をつけることにより、同じ人物でも画像中の位置によってその像の大きさが変化する。そこで、分割幅を画像中の位置によって変化させることで、画像中の位置に関係なく、人物の主要部位の横幅を得ることができる。処理の例を図2に示す。

また、人物間の距離が離れているにも関わらず、得られる画像では互いの人物領域が画像の縦軸方向に重なってしまう状況が生じる(図3(a))。しかし、俯角を考慮したフィルタリングにより、人物一人当りの分割領域数が画像中の位置に関係なくほぼ一定数となるため(図2(b))、この領域数を用いて縦軸方向の重なりを検出することが可能となる(図3(b))

以上の検出処理を、多数の人物がカメラ視野に存在する状況に適用した一例を図4に示す。

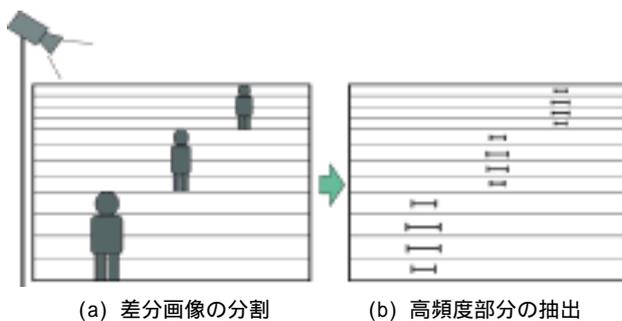


図2 差分画像のフィルタリング



(a) 縦軸方向の重なり (b) 検出結果

図3 縦軸方向の重なり検出



(a) 元画像

(b) 差分画像



(c) フィルタリング画像

(d) 検出結果

図4 検出の処理過程

3. 人物計数システム

前節で述べた人物検出手法の応用として、本研究では人物計数システムの構築を行った。

動画画像を用いた人物計数システムの研究も従来から盛んに行われているが、その代表的なものとして、画像の任意の位置に計測線(あるいは計測ゾーン)を配置し、画像処理によって実画像中の人物領域を抽出、追跡してその通過人物をカウントする手法[6][8]や、計測線上で得られる時空間画像を用いる手法[5][7]が挙げられる。これらの手法はいずれも人物の交差、隠蔽による影響を軽減するため、通行人の上空からほぼ鉛直下向きにカメラを設置することを前提としている。しかし、屋外環境ではこの前提条件を満たせない状況も十分に考えられるため、様々な環境に柔軟に対応できる手法とは言えない。これに対し、本研究で提案した人物検出手法は、一般的なカメラ視点による検出を行っているため、様々な環境での適用が可能である。以下、人物計数手法の詳細を述べる。

3.1 計測線通過人物の対応付け

実画像中の人物検出を利用して計測線通過人物をカウントする手法では、人物の計測線通過が複数の

フレームに渡るため、計測線通過人物の対応付け処理が必要となる。従来手法では人物の位置情報のみを用いて対応付けしているが、本手法ではこれに加えて、人物の色情報の類似性も対応付けに利用する。以下に、計数処理の手順を示す。

- Step 1. 検出処理から、計測線通過人物を検知する。
 Step 2. 人物の計測線通過位置を記録し、更に上半身に一定の大きさの長方形を当てはめ、この長方形領域内の色ヒストグラムを取得する。
 Step 3. 過去 T フレーム以内に検知された人物が存在するならば、処理フレームで検知した人物との相関値を算出する。
 Step 4. 相関値と設定した閾値を比較し、閾値以上ならば同一人物として対応付け、閾値未満ならば新たな通過人物として認識する。
 Step 5. 過去 T フレーム以内に対応付けされなかった人物を対応付けの対象から除外し(タイムアウト)、カウント数を 1 つ増加する。

計測線は画像の Y 軸に平行な直線とした。処理 2 で記録する人物位置は、人物の頭部(最上部)と計測線の交点の Y 座標、取得する色ヒストグラムは、RGB の各軸を Q 分割した要素数 Q^3 のヒストグラムである。処理 3 の相関値には、色ヒストグラムの類似尺度として、色ヒストグラムを Q^3 次元のベクトルとし、ベクトルのなす角の \cos を用いる。そして人物の位置情報として、人物通過位置の差の絶対値を、画像の縦幅で割った値を用いる。この 2 つの値から、人物の対応付けの相関値を算出する。以下に相関値の算出式を示す。

$$correlation = \cos\theta - \frac{abs(P_{1Y} - P_{2Y})}{H}$$

(P_{1Y}, P_{2Y} : 検出人物 1, 2 の計測線通過位置, H : 画像の縦幅)

3.2 2本の計測線を用いた人物計数

提案した検出手法は、複数人物の重なりにも対応できるが、複数人物の交差によって、一方の人物がもう一方の人物にほぼ完全に隠蔽されてしまう状況では検出が困難である。よって、設定した計測線上でこの状況が生じた場合には正しく計測できない(図 5)。

この問題を解決するため、2本の計測線を用いる手法を導入する。これにより、もし一方の計測線で人物交差によって正しく計測できない場合でも、も



図 5 計測線上での人物交差

う一方の計測線では正しく計測できるようになる。この場合、各計測線の計測結果が異なるが、各計測線での計測結果の最大値をシステム全体での計測結果とすることで、人物交差も正しく計測できる。

4. 実験

多数の人物が断続的にカメラ視野に侵入する環境での人物検出、および計数実験を行った。各実験で使用した処理システムは Pentium 866MHz、画像は 320×240 のフルカラーである。また、背景差分には SAD を用い、閾値を 60 に固定した。

4.1 人物検出実験

人物検出実験の撮影環境及び検出パラメータは次の通りである。

- 撮影条件
 - 天気：曇り
 - カメラ設置位置：地上から約 4.5 メートル
 - 俯角：約 20 度
 - 水平画角：約 28 度
- パラメータ
 - 動画フレーム数：2550 (約 8.5 秒間)
 - フィルタリングの分割幅：画像最上部で 5 pixel、最下部で 14 pixel

表 1 に計算時間、表 2 に検出率の実験結果を示す。計算時間の検討では、提案手法の比較対象として、楕円体モデル手法[4]による検出も併せて行った。検出における正解の判定は目視によって行い、検出結果が人物の上半身領域の約 7 割以上を含む場合を正解とした。検出における失報、誤報はそれぞれ、人物が検出されなかった事例、人物領域以外を検出した事例を表す。

まず計算時間の結果では、本手法は楕円体モデル手法と比較しても、非常に高速に処理できているのがわかる。また、人物数が 10 人に達する状況においても、計算時間の増加はわずかであり、動画キャプチャや画面出力などの入出力処理を含めても、ビデオレート(1/30 sec)での処理が可能であった。

検出率に関しては、総人物数の約 88% を正しく

表 1 検出実験結果(計算時間)

計算時間 (単位:msec)	フィルタリング	マッチング
楕円体モデル(人物数 1~3)	46 ~ 136	4.5 ~ 13.5
提案手法(人物数 3~10)	1.1 ~ 1.2	0.8 ~ 1.1

表 2 検出実験結果(検出率)

検出結果 (総人物数 13438)	正答数(率)	失報数(率)	誤報数
提案手法 (人物数 3~10)	11850 (88.18%)	1588 (11.82%)	590

検出することができ、十分な精度が得られることを確認した。検出における失報は、図 5 のように、人物が完全に他の人物に遮蔽されてしまう場合や、人物と背景色の類似によって動領域が十分に検出できなかった場合(図 6 (a))が主な要因である。この問題を解決するためには、固定閾値ではなく、差分結果から動的に閾値を決定する手法[9]などが考えられる。誤報の要因としては、縦方向に重なった人物の検出ミスや、カメラに対して正面(背面)を向いた人物の肩を誤検出してしまう場合(図 6 (b))があった。前者の問題は、色情報を用いて人物の頭部位置検出を行うことで対応可能であり、後者の問題も検出領域の面積、あるいは色情報を用いることで、真に複数人物の領域かを判定する方法で対応できると考えられる。本手法は高速に処理が行えるため、検出における他の処理の併用も十分に可能であると考えられる。

次に、画像中の存在人物数(込み具合)に対する検出率を求めたグラフを図 7 に示す。人物数の増加に対して、検出率が徐々に低下していることがわかる。検出システムの実用的な目安を 80% の検出率とすると、このグラフから本手法は人物数が約 10 人存在する状況まで対応可能であると言える。

なお、今回の実験では動画撮影を曇りの日に行ったため、背景差分に対する人物の影の影響は比較的小さいものであった。しかし、晴天時は影と背景のコントラストが高くなるため、本手法を屋外で適用することを考えた場合、背景差分に大きな影響を与えてしまう。影領域の検出には、HIS 表色系を用いる手法[10]などが提案されており、これらの手法を用いることで、影の影響を軽減できると考えられる。

4.2 人物計数実験

続いて人物計数実験の結果を示す。実験の撮影環境及びパラメータは次の通りである。



(a) 失報例(差分ミス) (b) 誤報例(肩の誤検出)

図 6 検出ミスの例(失報、誤報)



図 7 存在人物数に対する検出率

- 撮影条件
 - 天気：曇り
 - カメラ設置位置：地上から約 4メートル
 - 俯角：約 15度
 - 水平画角：約 34度
- パラメータ
 - 動画フレーム数：5700(約 190秒間)
 - フィルタリングの分割幅：画像最上部で 4 pixel、最下部で 16 pixel
 - 計測線(2本)の位置：x=150, 170
 - 相関値の閾値：0.75
 - 対応付けのタイムアウト：5フレーム
 - 色ヒストグラムの要素数：64

表 3 に計算時間、表 4 に計数正解率の結果を示す。表 4 では、2本の計測線を用いる手法の効果を確認するため、各計測線単独での計数結果も併せて示す。失報、誤報はそれぞれ、人物を計測できなかった事例、人物以外を計測した事例を表す。

表 3 から、計数処理に必要な計算時間は非常に小さな値であり、入出力、検出、計数といった計数システムの全処理を含めた場合でも、本手法は実時間処理が可能であった。

表 4 の計数正解率に関しては、各計測線単独での正解率が、87.3%、90.2%であり、単独でもかなりの精度が得られた。そして 2本の計測線を用いる手法

表3 計数実験結果(計算時間)

計算時間 (単位:msec)	ヒストグラム取得	対応付け
提案手法	0.20	0.01

表4 計数実験結果(計数正解率)

	通過人数	正答数 (率)	失報数 (率)	誤報数
計測線 x=150	134	117 (87.3%)	17 (12.7%)	4
計測線 x=170	134	123 (91.8%)	11 (8.2%)	5
2本の計測線	135	126 (93.3%)	9 (6.7%)	9

では 93.3%の正解率が得られ、2本の計測線を用いた場合の有効性が確認できる。2本の計測線を用いても計測できなかった状況は、複数人物がほぼ完全に隠蔽され、かつ同方向に歩く場合がほとんどであった。これに関しては、鉛直下向きのカメラ設置ではない以上、解決はほぼ不可能である。また、2本の計測線での誤報数が多くなっているが、これは各計測線の結果の最大値をシステム全体としての結果とするため、各計測線での誤報を必ずカウントしてしまうためである。各計測線での誤報は、検出処理の誤報に起因するため、検出の誤報を上述の方法で減らすことで、この問題に対応できると考えられる。

次に、画像中の存在人物数に対する計数正解率を示したグラフを図8に示す。人物数が5人までの状況では100%の正解率を示しているが、6人以上の正解率が減少している。これは人物数の増加によって、上述の計測不可能な状況が生じたためである。また、撮影した動画像では7人の場合にこの状況の頻度が高かったため、7人での正解率が落ちているが、基本的には図7の検出率と同様に、人物数の増加にほぼ比例して正解率が減少すると考えられる。

5. おわりに

本研究では、多数の人物が存在する状況に対応した新たな人物検出手法を提案し、更に提案手法の応用として、人物計数システムを構築した。

人物検出では、シルエットの統計データから人物の上半身モデルを構築し、人物の重なりに対応した。更に、差分画像のフィルタリングを導入し、マッチング処理の簡素化を行うことで、実時間検出を実現した。

人物計数では、画像の任意の位置に計測線を設定

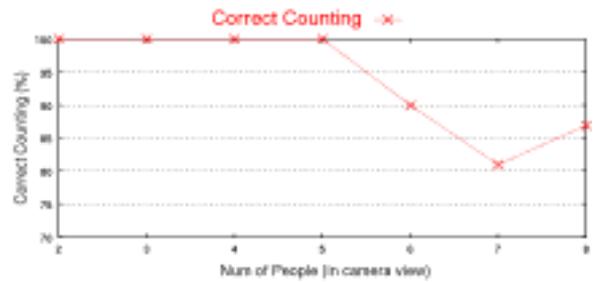


図8 存在人物数に対する計数正解率

し、人物の色情報を用いることで人物の計測線通過をカウントするアルゴリズムを導入し、実時間計数を実現した。

謝辞 本研究の一部は、財団法人日東学術振興財団の助成による。

6. 参考文献

- [1] 内田 恭輔、三浦 純、白井 良明, " 雑踏における歩行者の追跡 ", 信学技報 PRMU99-209 pp.39-45 (2000)
- [2] 長井 敦、久野 義徳、白井 良明, " 複雑変動背景における移動物体の検出 ", 信学論 D-II Vol. J80 No.5 pp.1086-1095 (1997)
- [3] 荒木 昭一、横矢 直和、竹村 治雄, " 交差判定により分裂・統合する動的輪郭モデルによる複数移動物体の実時間追跡 ", 信学論 D-II Vol. J80 No.11 pp.2940-2948 (1997)
- [4] 加藤 博一 中澤 篤志 井口 征士, " 楕円体モデルを用いたリアルタイム人物追跡 ", 情処学論, Vol140 No.11, pp.4087-pp.4096 (1999)
- [5] 間瀬 健二, "x-t 時空間画像を用いた歩行者の計数 ", 信学技報 IE90-43 pp29-36 (1990)
- [6] G Sexton, X Zhang, G Redpath, D Greaves, " Advances in automated pedestrian counting ", IEE Conf Publ, No.408 pp106-110 (1995)
- [7] 寺田 賢治、田所 英昭、山口 順一, " 2本のななめ計測線を用いた通過人数計測 ", 電気学会論文誌 C 研究開発レター Vol.120-C, No.12, pp.2105-2106 (2000)
- [8] 小川 浩太郎、杉村 宏治、菅 泰雄、辻谷 潤一: " 画像処理を用いた自動通過人数計測システムの構築 ", 第2回動画処理実利用化ワークショップ pp.18-22 (2001)
- [9] P.L. Rosin and T. Ellis, " Image difference threshold strategies and shadow detection ", 6th British Machine Vision Conf, Birmingham, pp. 347-356 (1995)
- [10] 園田 頼信、緒方 敏博、松野 洋輔, " 監視画像における対象物と影の分離 カラー情報の利用 - ", 信学技報 PRMU99-19 pp.1-6 (1999)