

群衆の映像における人数の自動計測について

芝脇 岳雄 小畑 正貴 塩野 充

岡山理科大学 工学部 情報工学科

概要

多数の人が集まったり通過するような各種施設においては、その人数や移動状況を把握することは管理や運営の上で極めて重要である。そのため、各種施設における人数計測の自動化に対するニーズは大きいと考えられる。通常、人数の計測は係員の目視で行ったり、改札口のようなゲートを設けてその中に仕組んであるバーやセンサーで計測しているが、目視は人手がかかり、ゲートを用いると混雑を増幅させる問題点がある。本研究では画像処理技術だけを用いて、多くの人の出入りする状況において、その人数を自動計測する方法について提案、実験を行なった。画像の入力にはTVカメラを用いて鉛直上方から一定時間ごとに撮像を行い、その画像集合から変化した領域を検出して得られた重心位置の変化によって各人物の移動を抽出する。そしてそれに基づいて人数の計測を行う。

Automatic counting of the number of persons in a crowd image.

Takeo SHIBAWAKI, Masaki KOHATA, Mitsuru SHIONO

*Department of Information & Computer Engineering,
Okayama University of Science*

Abstract

It is very important work to count the number of persons at a site where many people gather and walk, such as an exhibition grounds. Usually the number of persons counting is carried out by humans eyes or a physical gate with a bar or a sensor. These method, however, have some problems. In this report, a method for counting the number of persons automatically using only image processing techniques is proposed, and the result of a fundamental experiment is shown. In this method, a TV camera placed at the upper air catches crowd images every few hundreds mili-seconds. Each person can be detected by calculation of diffrentials of those images.

1. まえがき

多数の人が集まったり通過するところにおいては、その移動を把握することは管理や運営の上で極めて重要であり、そのため、その通行人数の計測が行なわれる場合が多い。例えば、公共・レジャー施設などにおいては入退場者数の把握やその移動傾向・各施設の利用状況などの情報を得る必要があり、それらの情報が管理運営上重要となる。さらに、得られた情報を混雑回避のために利用者に提供することも可能である。

従来、このような用途には入退場口に物理的なゲートを立てて赤外センサー等で計測を行うという方法や、直接に監視員をおいて肉眼で計測を行うという方法がある。しかし、ゲートを設ける方式では人の流れに制約ができて、混雑したとき支障となることや、緊急時に不都合が生じるおそれがあるなどの問題もある。また、極端に混雑した状態ではゲートを用いた方法では計数不可能になる状況も考えられる。また、監視員をおいた場合には、長時間の監視作業などであると監視員の負担も大きくなり、かつ計測を行うのが人間であるので誤りが発生することも起こり得る。このため、各種施設における人数計測の自動化に対する要求は大きいと考えられる。

そこで、本研究では画像処理技術を用いて、多くの人の出入りする状況において、その人数を自動計測する方法について提案、実験を行なう。[1] [2] [3]

2. 処理の概要

本研究では、群衆の流れを画像処理を用いてその各個人の移動方向の識別を行い、その情報を基に人数の計測をするものである。画像の入力にはTVカメラを用いて一定時間ごとにサンプリングを行い、その画像から変化した領域を検出して移動を判定する。また、それに基づいて人数を計測する。全体の処理の流れを図1に示す。以下、各処理について順に述べる。

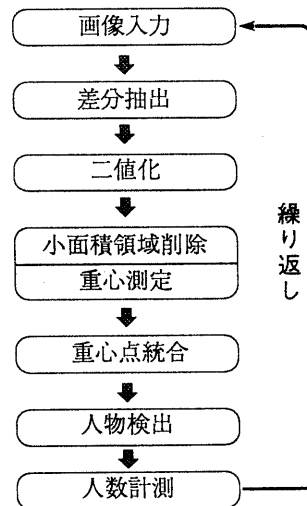


図 1: 本方式の処理の流れ

3. 画像の入力

3.1 撮影方向

測定対象となる人間の体格や服装は各個人ごとにかなり違いがある。また移動方向もさまざまなため、撮影状況によってはその画像が大幅に変化し、測定に不都合となる。このために、各個人ごとの体格や服装の変化に左右されにくい方法で画像を入力しておく、以後の処理が容易になる。そのため、本研究ではTVカメラを対象とする領域の鉛直上方に設置して、移動する人物をほぼ真上から撮影する。

この方法を用いることにより、他の方向から撮影することで発生する人物像の重なりによる影響を最小限に抑えることができる。また、各個人の体格によって発生する領域サイズの違いは比較的小さくなる。そして、鉛直上方から撮影することにより、対象となる群衆の移動がカメラの光軸に対してほぼ垂直に面する平面上になる。このため、得られた画像は2次的に処理することが可能となる。

3.2 撮影の時間間隔

本研究では、差分画像によって得られる重心移動をもとにして人数の計数を行う。そのため、時間間隔が大きくなると他の人物の重心移動領域との識別が困難になり、誤認識の増加につながる。また、この間隔があまり短すぎると、画像間の差分領域が小さくなりすぎて、領域の検出が困難になる。画像のサンプリングにはTVカメラを用いるため、その取り込み間隔はNTSC信号として得られる間隔に制限される。そこで、本研究では各画像間の取り込み間隔を約1/10秒としている。この値は実験に基づいて求めた。

3.3 撮影方法

本実験では、群衆の流れを画像として入力するのにTVカメラを用いた。設置方法は図2のとおりである。このようにして、約1/10秒ごとに512×256画素、256階調(8ビット)で6枚の連続した画像の取り込みを行った。

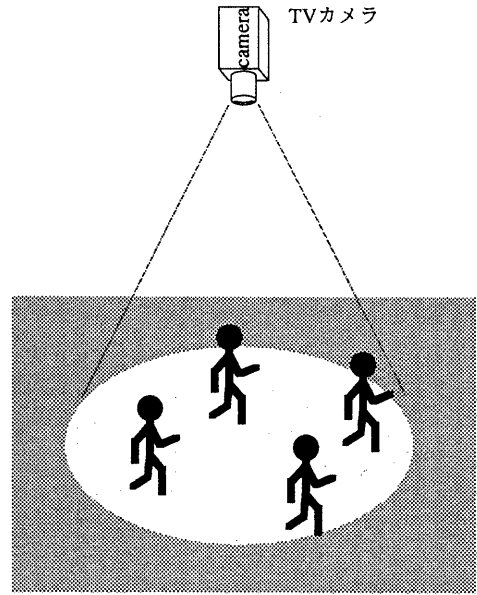


図2: カメラの設置方法

4. 差分処理

背景に対して濃度差のある物体が存在した場合、その物体が移動していると、取り込まれた各画像中において画像間に差異が生じる。本研究ではこの差をもとにして、人物の移動した領域の検出を行う。

これは、一連の画像群

$$F(f_1, f_2, \dots, f_n \dots) \quad (1)$$

に関して、任意の時刻 t_1 における画像を $f_n(x, y)$ 、それから Δt 秒後における画像を $f_{n+1}(x, y)$ とすると、差分画像 $d_n(x, y)$ は次式によって得ることができる。

$$d_n(x, y) = f_{n+1}(x, y) - f_n(x, y) \quad (2)$$

このようにして得られる差分画像群が

$$D(d_1, d_2, \dots, d_n \dots) \quad (3)$$

となる(図3)。実際のプログラムでは、8ビット(256階調)の正数で処理を行うことができるように、差分の値に128を加算して、0未満と255以上を範囲内に収めるように修正して処理を行った(式4参照)。

$$d_n(x, y) = \begin{cases} 0 & (f_{n+1}(x, y) - f_n(x, y) + 128 \leq 0 \text{ のとき}) \\ f_{n+1}(x, y) - f_n(x, y) + 128 & (0 < f_{n+1}(x, y) - f_n(x, y) + 128 \leq 255 \text{ のとき}) \\ 255 & (255 < f_{n+1}(x, y) - f_n(x, y) + 128 \text{ のとき}) \end{cases} \quad (4)$$

以上の処理を行うことによって、移動によって変化が起きた領域を検出することができる。

5. 二値化

差分処理を行って得られた画像群に対して、各画像についてそれぞれ二値化を行う。

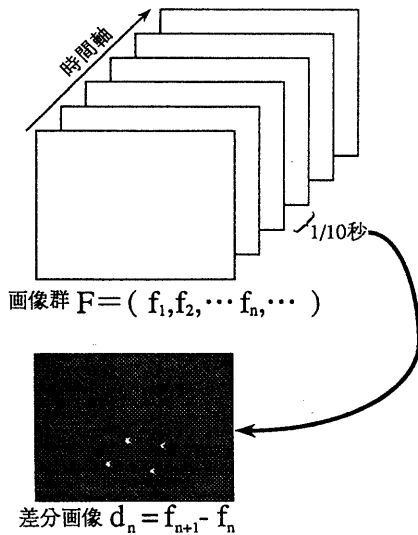


図 3: 差分処理

背景に対して移動した人物があった場合、差分処理後の画像情報には、変化のない背景領域を0としたとき、移動した人物の前後に背景に対して正と負の値をもつ2つの領域が存在する。これを適当なしきい値を用いて二値化することによって、移動している人物の前後にあらわれる2つの差分領域のうち片方の領域のみを抽出することができる。このとき、前と後ろのどちらの領域が切り出されるかは背景の色と人物の服装の色のそれぞれの濃度値によって決定される。しかし、背景と人物の服装の濃度値がほぼ同じ値でなければいずれかの領域が、背景領域の値に対して正あるいは負のしきい値を越えて検出される。このとき、背景と服装の濃度値に変化がなければ、検出される領域と人物の位置関係は同じものとなる。したがって、画像中においてしきい値を用いて得られた2値化画像の各領域と人物に対する関係は常に一定になる。

本実験で用いた値は差分値の全範囲をもとにして、その値の高い部分から一定比率をもってしきい値とした。この比率は、実際の画像をもとに実験によって求めた。

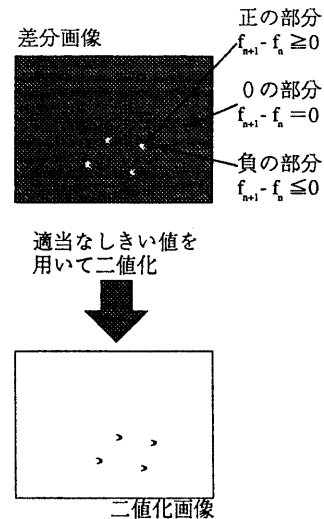


図 4: 二値化

この処理を行うことによって、得られた画像より対象となる領域を検出することができる(図4)。

6. ノイズ除去と重心検出

6.1 ノイズ除去

差分処理を二値化した画像中にはカメラの特性やその時々の日照条件・小動物の影響などによる多数の小領域がノイズとして残っているため、その除去を行う必要がある。また、それに対して一定面積以上の領域は人物の移動した差分領域として計測する必要がある。

そこで、各領域の面積を測定して、その領域をノイズとして除去の対象とするか、あるいは重心を求める領域とするかをその面積によって決定した。すなわち

- 各種の原因で発生したノイズ — 除去
- 人物の移動で発生した領域 — 6.2 の処理を実行

の2つのうちのいずれかの処理をそれぞれの領域に対して行う。

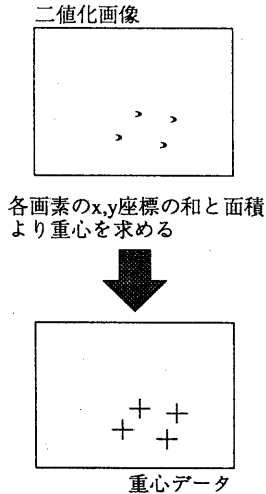


図 5: 重心計測

6.2 重心計測

各領域の重心の計測については、面積を計測すると同時に各画素の持つ (x, y) 座標の値を積算して、最後に得られた面積で割ることによって求めた。すなわち、画面中の n 番目のその領域の面積が $s(n)$ であったとき、領域の重心 $c_n(x_n, y_n)$ は次式のように求められる。

$$(x_n, y_n) = \left(\frac{\sum x(n)}{s(n)}, \frac{\sum y(n)}{s(n)} \right) \quad (5)$$

このようにして求められた面積と (x, y) 座標の和によって、重心の座標を計算して内部のテーブルに記憶しておく。(図 5) 実際のプログラム中では高速化のために画面を 1 走査線毎に分解して、1 ライン中の各小領域ごとに面積と各座標値の計測を行い、それらの領域の上下の連結関係を見て接続していれば下の領域に加え、接続していない場合は得られた面積を出力することによって逐次的な処理に変換して求めた。

7. 移動の追跡

ここまでの処理で得られた複数の重心座標値より、各重心点の移動方向と距離に関して検討を行う。

ある時刻 t における重心座標群

$$C_t(c_{t,0}, c_{t,1}, \dots, c_{t,n}, \dots) \quad (6)$$

と、時刻 $t+1$ における重心座標群

$$C_{t+1}(c_{t+1,0}, c_{t+1,1}, \dots, c_{t+1,n}, \dots) \quad (7)$$

について、 t の各重心座標値 $c_{t,n}$ に対して $c_{t+1,n}$ の中で最も近い座標値を持つものを次の画面の座標群中から求める。このようにして得られた座標値の組合せは、連続した画面中において同一の領域が移動したものと推定できる。このとき、移動量の大きさが移動速度を示すため、人間の移動速度を越える大きさの移動量はカメラぶれなどによるノイズか、人間以外の移動物体による差分領域であると考えることが出来る。

したがって、求めた移動量がある値を越えていた場合は、その移動速度が規定外であると見なして削除する。

8. 結果

実験に用いたシステムは以下の通りである。

- ビデオカメラ
- ビデオデッキ
NEC PV-S98(VHS モードで使用)
- コンピュータ
NEC PC-98RL (80386+387-20MHz)
- フレームバッファ
MT98-CVFM(512 × 256 × 256 階調 × 6 画面)

まず、TV カメラによって得られた連続画像のうちの 2 枚を図 6, 図 7 に示す。

図 6 と図 7 の差分画像が図 8 である。

図 8 を二値化した画像が、図 9 となる。

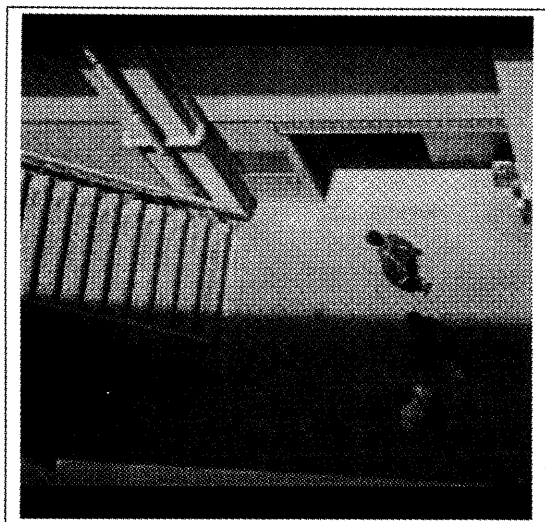


图 6: 画像 1

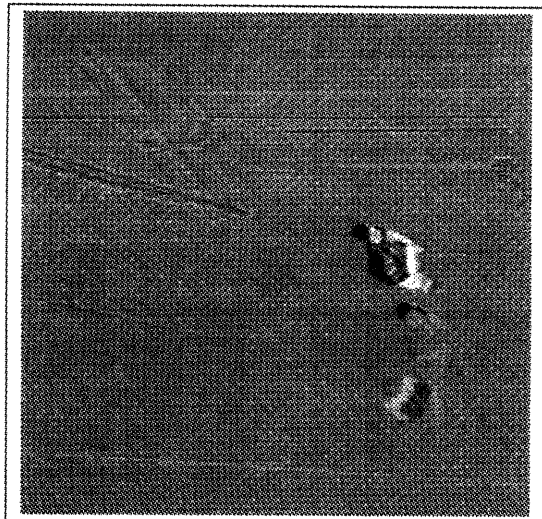


图 8: 差分画像

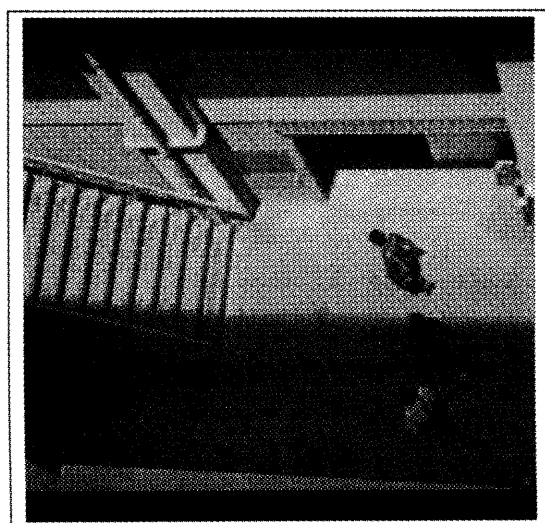


图 7: 画像 2

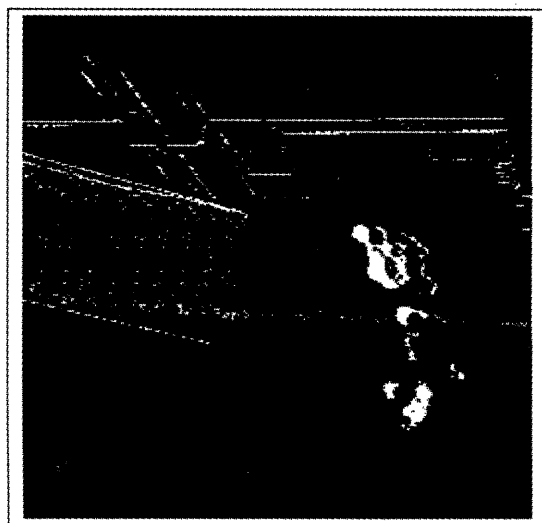


图 9: 二值画像

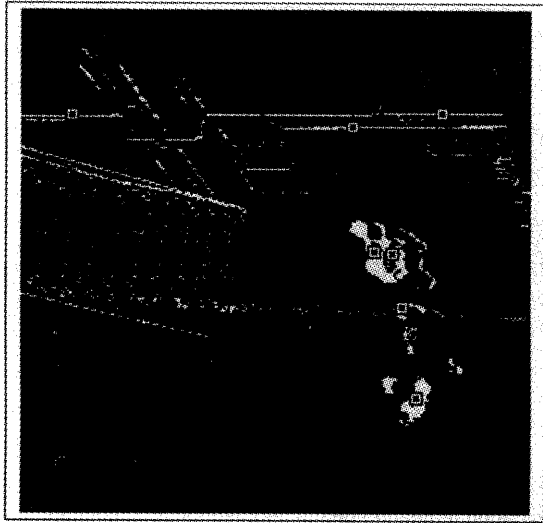


図 10: 重心画像 1



図 11: 重心画像 2

図 9 からノイズ除去を行ない、重心を求めた結果が図 10 である。また、それに続く画面間で求められた重心が図 11 である。それぞれの図において、重心は黒と白の四角で示している。

そして、図 10 と図 11 から求めた移動方向の様子が図 12 である。この図中において、連続していると判断された重心間は白線で接続して示す。

同様にして、6 画面間で移動の方向を求めた結果が、図 13 である。

9. まとめ

本実験においては撮影された各人物がある程度散らばっている場合には、移動検出能力の有効性を確認できた。しかし、今回用いた画像データは基礎実験のために、画面中の人数が数人のものであり、実際にはもっと多人数のものを対象としなければならない。そのような場合にも個々の人物を正確に検出できるためにはどのような撮影条件が必要となるのかを検討しなければならない。また、人数が増えていった場合に人数計測の誤差がどのようにあられるのかの検討も行なっていきたいと考えている。また、



図 12: 3 画面での計測結果

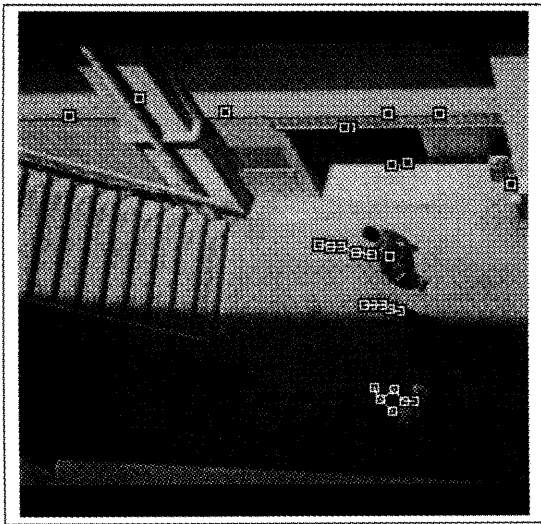


図 13: 6 画面での計測結果

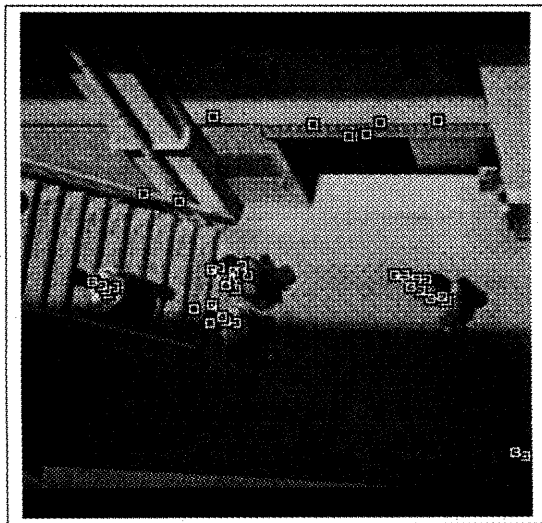


図 14: 計測結果例 2

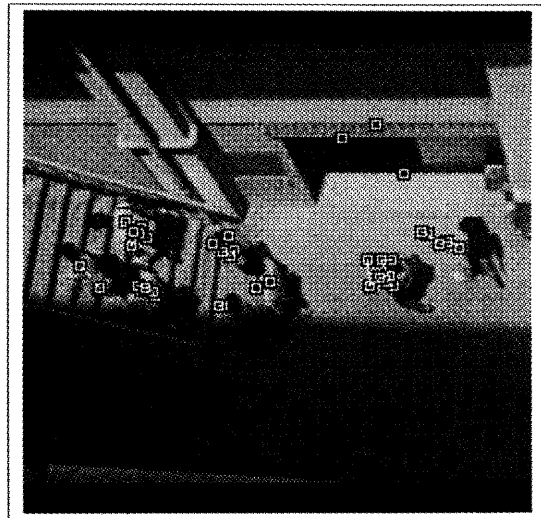


図 15: 計測結果例 3

撮影時における影の影響についても、これを極力減少させるような方法を検討したいと考える。

参考文献

- [1] 芝脇, 小畑, 塩野: “群衆の映像における人数の自動計測” 平成 5 年電気・情報関連学会中国支部第 44 回連合大会 142023 (1993-10).
- [2] 芝脇, 小畑, 塩野: “群衆の映像における人数の自動計測について” 平成 5 年電気関連学会関西支部連合大会 G8-7 (1993-11).
- [3] 竹内, 長田: “画像を利用した群衆流動解析法に関する研究” 信学技報 IE90-60, pp23-30, 1990.