

A Method for Counting People at Elevator Lobby Focusing on Top of Moving Objects

Toshiaki Tsuji Member (Saitama University)

Hiroki Tashiro Non-member (Saitama University)

Shigeru Abe Member (Saitama University)

Keywords: image processing, people counting, human's head, background subtraction, optical flow, hough transform

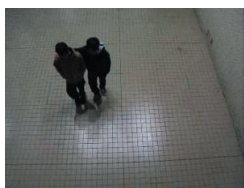
1. Introduction

In recent years, people counting systems using image processing technologies have been generally introduced because of the low-priced and high-advanced cameras and computers. Obtaining the number of people provides improvement on performance of many systems. As an example, a system to count the number of people at an elevator lobby is mentioned for efficient operation of elevators. This paper proposes a method for counting people in an elevator lobby based on two assumptions: heads of people exist around the upper edge of the images of moving objects; and shape of the head is circular. This paper shows that the method based on these assumptions acquires high recognition rate.

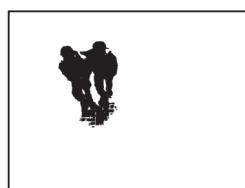
2. Procedures for Counting People

The proposed method extracts the number of people from camera images (Fig. 1(a)). The method is composed of the following procedures.

- (1) Moving objects are extracted by the background difference method (Fig. 1(b)).
- (2) Upper edge of the moving object is detected and the region W , a square region for head detection, is defined around the edge (Fig. 1(c)).



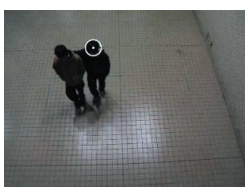
(a) An Example of input image.



(b) Extraction of moving objects.



(c) Region around top of moving objects.



(d) Head detection.



(e) Determination of human region.



(f) Segmentation of moving objects.

Fig. 1. Example of processing images

- (3) Heads are detected from the region W in frame difference image by Hough transform. Here, it is assumed that the shape of the head is circular. False detection is strongly reduced by giving a limitation on the region for head detection (Fig. 1(d)).
- (4) Heads detected by Hough transform are tracked by optical flow. Mis-detection is avoided by keep tracking the heads.
- (5) A human region is determined based on the position of the detected head. The region has a shape of an ellipse and its center position is geometrically estimated from the head position (Fig. 1(e)).
- (6) When the images of multiple people lap over, some images of moving object remain around the human region determined in (5). Then, people around the human region are counted by repeating the above procedures for the remained images. As a result, the image of the moving object is segmented to multiple human regions (Fig. 1(f)).

3. Experimental Verification

A video camera mounted on the wall above the elevator door was used for experimental verification. Two kinds of experiments in different conditions were performed. The conditions are described as follows.

- (i) The height of the camera was 500 cm and up to 10 people were waiting in the elevator hall.
- (ii) The height of the camera was 380 cm and up to 7 people were waiting in the elevator hall.

Table 1 shows the results of the experiments. Here, A denotes the real number of people manually counted from the video images. B and C denote positive and negative error. Then, counting rate D is derived by $D = 1 - (B + C)/A$. The counting rates D were 94.4% and 92.6% in conditions (i) and (ii), respectively. The main cause of the negative error was that the image in the segmented human region became too small when the images of people were strongly overlapped.

The result shows that the proposed system acquires high counting rate and high reliability in an elevator lobby with many overlapped images of people.

Table 1. Experimental result

case	(i)	(ii)
Real number of people A	179	217
Positive error B	2	3
Negative error C	8	13
Counting rate D [%]	94.4	92.6

エレベータ乗場画像の移動体上端に着目した待客数計測方式

正員 辻 俊明* 非会員 田代 浩紀*
正員 阿部 茂*

A Method for Counting People at Elevator Lobby Focusing on Top of Moving Objects

Toshiaki Tsuji*, Member, Hiroki Tashiro*, Non-member, Shigeru Abe*, Member

This paper proposes an image processing system for counting people at an elevator lobby viewed from an upper oblique angle. The method is based on two assumptions: heads of people exist around the upper edge of the images of moving objects; and shape of the head is circular. This paper shows that the method based on these assumptions acquires high recognition rate and high reliability in an elevator lobby with many overlapped images of people.

キーワード：画像処理，人数計測，人物頭部，背景差分法，オプティカルフロー，ハフ変換

Keywords: image processing, people counting, human's head, background subtraction, optical flow, hough transform

1. はじめに

カメラからの動画像を処理し、画像内の人物を検知する研究は広く行われている。その応用例には、不審者や侵入者を検知するセキュリティシステム、公共施設での入退場者数計測、小売店における顧客の動線分析などが存在する。一方、エレベータ等の自動機械にとって、運行の効率化は重要な課題である。そこで、画像処理によってエレベータ乗場の待客数を計測することができれば、待客数に応じたエレベータの配車が可能となり、運行の効率化が実現できる。エレベータ乗場の待客数計測法については依田らによる先行研究⁽¹⁾が挙げられる。依田らは人物同士の重なりを回避するためにエレベータ乗場の天井にカメラを設置し、撮影した画像内の移動体領域の面積をもとに人数の計測を行っている。また、筆者らの研究グループでは天井に設置されたカメラの画像からハフ変換で頭部を検出することにより、高性能な人数計測と車椅子の識別を実現している⁽²⁾。しかし、設置施設等の制約条件により、カメラを天井に設置できない場合には設置制限の少ない側壁上方に設置して待客数を計測する必要がある。

側壁上方からの動画像を用いて人数計測するうえでの問題は、人物が重なり、複数の人物が一つの移動体と見なされることや人体に隠蔽部分が生じることである。人物が重なる状況で人数を計測する手法として、ある特徴量と画像内の実人数を対応させ、学習させることで人数の計測を行

う方法がある⁽³⁾⁽⁴⁾。しかし、計測結果が学習データに大きく依存するという問題がある。側壁上方からの動画像では、人物は重なるが、頭部が重なることは極めて稀である。そこで人物の肩から頭部の形状(Ω 形状)を検出し、重なりにくい頭部の検出を行う手法が提案されている⁽⁵⁾⁽⁶⁾。また、Sidlaらの研究⁽⁷⁾では駅の改札など人物が密集する状況で、 Ω 形状を検出し、人数を計測している。しかし、これらの研究では Ω 形状のテンプレートをあらかじめ用意する必要がある。 W^4 システム⁽⁸⁾では移動体領域を垂直投影し、頭部などの部位の特定に利用している。しかし、垂直投影する移動体領域は人物を横から撮影したものである必要がある。また、エレベータ乗場においては、人物はカメラの真下やカメラから離れた位置に分布し、位置によって人物の肩から頭部の形状や移動体領域の形状は大きく変化する。馬場らの研究⁽⁹⁾では頭部を円形と仮定し、ハフ変換を用いて画像内から円を検出し、検出された円を頭部としている。しかし、ハフ変換を用いて円を検出する際、頭部以外の位置に円が検出されることが問題となっている。

本研究では頭部を検出するうえで2つの仮定を設定し、エレベータ乗場の待客数を計測する手法を提案する。第一の仮定は、移動体の上端の座標が人物の頭部に位置するという仮定である。第二の仮定は、頭部は観測位置によらずほぼ円形となるという仮定である。第一の仮定により頭部が存在する領域を定め、第二の仮定によりハフ変換を用いて定めた領域内から円を検出する。2つの仮定を設けることにより、正確な頭部の検出を実現し、人数を計測する。以上の手法により構築されたシステムは、人物同士の重なりが頻繁に生じる環境において95%近くの計測率を得られることが実験的に確認された。これはテンプレートを予め用

* 埼玉大学

〒338-8570 さいたま市桜区下大久保 255

Saitama University

255, Shimo-Okubo, Sakura-ku, Saitama 338-8570

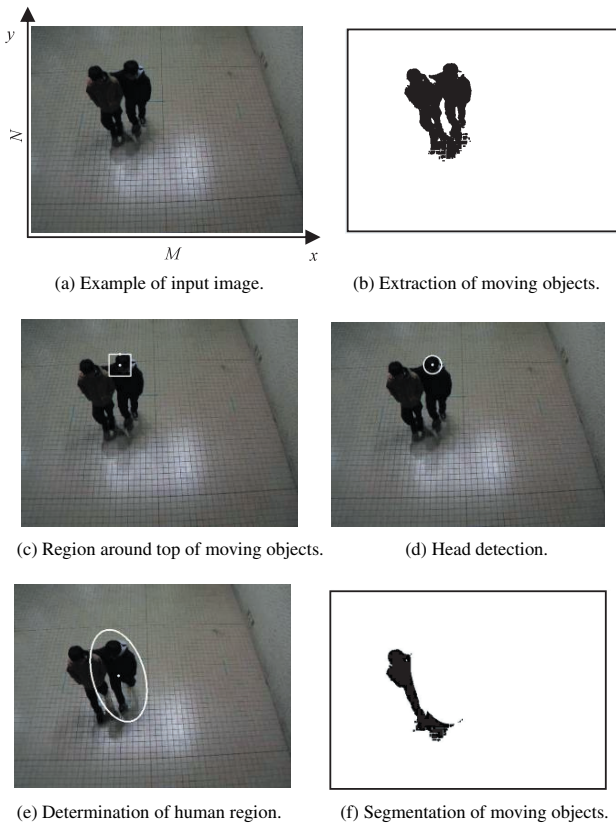


Fig. 1. Example of processing images.

意する必要のない人数計測法としては極めて高い性能である。その内容について本稿で報告する。

本論文の構成は以下のとおりである。第2章において本システムの概要について述べ、第3章で人数計測手法について記述する。さらに第4章で実験による性能評価を通じて本システムの有効性を確認する。最後に、第5章でまとめを述べる。

2. システム概要

カメラの真下の待客を撮影できるようにカメラをエレベータの扉上方の側壁に設置する。Fig. 1(a)がエレベータ乗場床面を撮影したグレースケール画像で、画素の位置は xy 平面の座標として表し、 N 、 M は撮影した画像の x 方向、 y 方向の画素数とする。また、解像度は $(N \times M)$ で表される。取得した画像に以下の処理を行い、待客数を計測する。

(1) 移動体領域の抽出 取得した画像 (Fig. 1(a)) から背景差分法により移動体領域を抽出し、ラベリング処理を行う (Fig. 1(b))。

(2) 移動体上端の周辺領域の決定 各移動体領域の上端が人物の頭部に位置すると仮定し、その周辺を頭部が存在する領域 (以下、頭部領域 W) として定める (Fig. 1(c), 2(a))。移動体領域内で y 座標が最大の画素を上端 P とする (Fig. 2(b))。また、頭部領域 W は正方形とする。

(3) 頭部の検出 頭部を円形と仮定し、ハフ変換を用いて頭部領域 W 内の画像から円を検出する。検出された

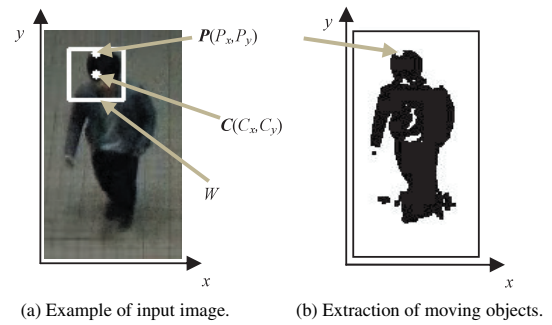


Fig. 2. Region around top of moving objects.

円を頭部とする (Fig. 1(d))。

(4) 頭部の追跡 頭部が検出された後、頭部の追跡を行う。重なりにくい頭部を追跡することにより、人物が重なった場合でも頭部の数を計測できる。追跡には頭部のオプティカルフローを利用する。

(5) 人物領域の決定 検出した頭部の位置をもとに人物領域を定める (Fig. 1(e))。人物領域は楕円とする。

(6) 移動体領域の分割 複数の人物が重なった移動体領域を分割し、移動体領域内の人数を計測する。(1)~(5)の処理で人物領域を定めた後に Fig. 1(f) のように人物領域外に残った移動体領域を抽出する。

3. 人数計測

〈3・1〉 移動体抽出 移動体領域の抽出には背景差分法を用いる。背景差分法は移動体がない状態の背景画像の輝度値 B_t と移動体がある状態の画像の輝度値 I_t との差分値 D_t に閾値処理を行い、移動体を抽出する方法である。差分値 D_t は (1) 式で表される。

$$D_t = |I_t - B_t| \dots \dots \dots (1)$$

本システムでは輝度値の最小値を 0、最大値 255 とし、閾値 D_{th}^b を 45 とした。閾値 D_{th}^b 以上の場合を移動体画素として黒で、閾値 D_{th}^b 以下の場合を背景画素として白で表示する。さらに 1 回の膨張・収縮処理でノイズを除去し、移動体領域を抽出する (Fig. 2(b))。

〈3・2〉 移動体上端の周辺領域の決定 移動体の上端が人物の頭部に位置すると仮定し、頭部が存在する領域 W を定める。抽出された移動体領域にラベリング処理を行い、移動体領域の画素数が閾値 S_{th} 以上の移動体領域に対し、頭部領域 W を定める。

まず、頭部領域 W の形状と大きさを定める。形状は正方形とし、その一辺の長さは画像内で頭部が最も大きく映る位置 (人物がカメラの真下に存在するとき) での頭部の直径 p_{head} とする。Fig. 3(a) はカメラの真下に存在する人物が映る画像であり、Fig. 3(b) はカメラの真下に存在する人物を横から観察した様子である。カメラに映る頭部の直径 p_{head} はカメラの設置高さ H 、人の平均身長 h 、人の頭部平均直径 r_{head} から推定する。Fig. 3(b) より p_{head} と r_{head} の関係は (2) 式で表される。

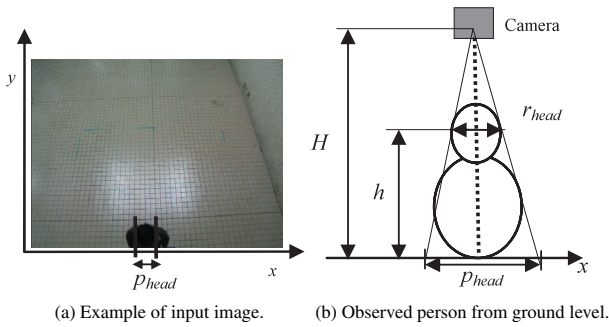


Fig. 3. Relation between p_{head} and r_{head} .

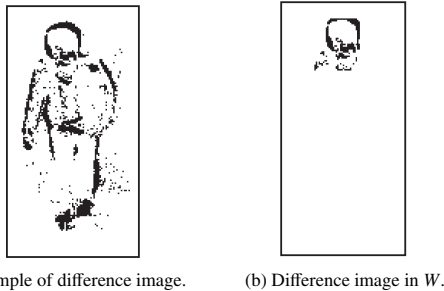


Fig. 4. Head detection.

$$p_{head} = \frac{H}{H-h} \cdot r_{head} \dots\dots\dots (2)$$

さらに、 p_{head} の画素数 q_{head} を求め、 q_{head} を頭部領域 W の一辺の画素数とする。撮影画像で、 x 軸上（画像の下端）で撮影される床面の範囲はカメラの水平画角 θ_{hori} を用いて、 $2 \cdot H \cdot \tan(\theta_{hori}/2)$ と表される。 q_{head} と p_{head} の関係は (3) 式で表される。

$$q_{head} = \frac{M}{2 \cdot H \cdot \tan(\theta_{hori}/2)} \cdot p_{head} \dots\dots\dots (3)$$

本研究では $h = 165$ [cm], $r_{head} = 24$ [cm], $\theta_{hori} = 63$ [°] とした。

次に、頭部領域 W の位置を定める。上端 P の位置から y 軸方向に $-q_{head}/2$ 移動させた位置を頭部領域 W の中心（以下、領域中心 C ）とし、頭部領域 W を定める（Fig. 2(a)）。

〈3・3〉 頭部の検出 頭部は観測位置によらず円形となると仮定して、ハフ変換を用いてフレーム間差分画像内から円を検出し、検出された円を頭部とする。

まず、フレーム間差分法により、フレーム間差分画像を抽出する。フレーム間差分法とは一定のフレーム間の画像同士の輝度差を求める方法である。現フレームの輝度値を I_t とすると、差分値 $D_{t,t-1}$ は (4) 式で表される。

$$D_{t,t-1} = |I_t - I_{t-1}| \dots\dots\dots (4)$$

差分値 $D_{t,t-1}$ に閾値処理を行い、Fig. 4(a) のように移動体の輪郭部分を差分画素として抽出する。本システムでは閾値 D_{th}^f を 20 とし、 D_{th}^f 以上の場合を移動体画素として黒で、 D_{th}^f 以下の場合を背景画素として白で表示する。

次にハフ変換を用いて、フレーム間差分画像内から円を検出する。抽出したフレーム間差分画像に中心位置 (a, b) ,

半径 r_n の円が存在すれば、その円は (5) 式で表される。

$$(x-a)^2 + (y-b)^2 = r_n^2 \dots\dots\dots (5)$$

つまり、 a, b, r_n , 3つのパラメータで円を表現できるので、これらのパラメータを座標軸にした空間 ($a-b-r_n$ 空間) では、画像中の円をある1点で表現できる。ハフ変換とはこの $a-b-r_n$ 空間中の一点を求める手法である。以下の手順で円を検出する。

- (1) 対象画像と投票空間（3次元の $a-b-r_n$ 空間）を用意する。
- (2) 対象画像内の画素を走査し、その画素が黒い場合には周りに半径 r_n の円を描き、描かれた円の円周上の点を投票空間に投票する。
- (3) 対象画像全ての画素に対して(2)の手続きを実施する。
- (4) 半径 r_n を変化させ、(2)~(3)を繰り返す。
- (5) 投票数が最も多い点を求める円の中心とする。

Fig. 4(a) のフレーム間差分画像にハフ変換を行った場合、頭部以外の位置に円が検出される。そこで、Fig. 4(a) の頭部領域 W 内を抽出した画像 (Fig. 4(b)) から円を検出することにより頭部以外の位置での検出を防ぐ。また、検出する円の半径を限定し、極度に大きな円や小さな円の検出をなくすことにより、頭部以外の位置での検出を防ぐ。円の半径の最大値は $q_{head}/2$ とし、最小値は $q_{head}/4$ とした。

ハフ変換により検出された円を頭部とみなし、その中心を頭部領域の中心座標 C とし、頭部領域 W を更新する。

〈3・4〉 頭部の追跡 検出した頭部を追跡する。人物が密接に重なった場合に人数の誤計測が不可避であることが多くの研究で報告されているが、追跡処理を加えることにより、一時的に重なった人物の頭部が持続的に検知される。そのため人数を正確に数えることが可能となる。頭部の追跡には頭部のオプティカルフローを利用する。オプティカルフローとは異なる時間に撮影された2枚の画像間で同じ対象を対応付けし、その移動量をベクトルデータとして表現するものである。

オプティカルフローの抽出方法として勾配法が挙げられる。勾配法は、連続する2枚の画像での対象物の移動量が微小であることを前提にオプティカルフローを求める方法である。画像上の画素 (x, y) の時刻 t における輝度値を $I(x, y, t)$ とし、画素の移動すなわちオプティカルフローを (6) 式で表す。

$$(u, v) = \left(\frac{\partial x}{\partial t}, \frac{\partial y}{\partial t} \right) \dots\dots\dots (6)$$

画像上の画素値の勾配を I_x, I_y とし、画素値の時間微分を I_t とすると、(7) 式が得られる。

$$I_x u + I_y v + I_t = 0 \dots\dots\dots (7)$$

(7) 式をオプティカルフローの拘束条件式という。KL法ではこの式にさらに「オプティカルフローは滑らかに変化する

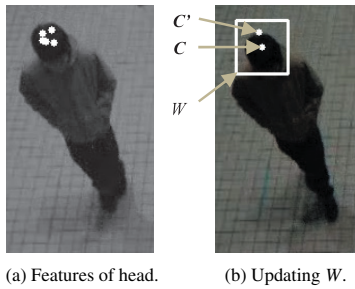


Fig. 5. Head tracking.

る分布をもつ」という過程を設け、 $(I_x u + I_y v + I_t)^2$ を最小化すると同時に、滑らかさ $((u_x^2 + u_y^2) + (v_x^2 + v_y^2))$ も最小にするようにオプティカルフローを求める。すなわち (8) 式を最小にする (u, v) を求めることになる。

$$e = \iint ((u_x^2 + u_y^2) + (v_x^2 + v_y^2)) + \lambda (I_x u + I_y v + I_t)^2 dx dy \dots\dots\dots (8)$$

KL 法を用いてより安定した追跡を行うため、画像内で抽出される特徴点を利用するものが KLT⁽¹⁰⁾⁽¹¹⁾ 法である。本システムでは KLT 法を用い、以下の手順で頭部の追跡を行う。

- (1) 前フレームの画像 (グレースケール画像) 内から特徴点を 3000 個抽出する。特徴点の抽出にはハリスのコーナー検出⁽¹²⁾を用いる。
- (2) 前フレームの画像から、頭部の特徴点を選択する。前フレームの領域中心 C' を中心として追跡する頭部に検出された円 (最後に検出された円) の半径以内の距離に存在する特徴点を頭部特徴点とする (Fig. 5(a))。
- (3) 頭部特徴点を始点とするフロー (大きさが 1 pixel 以上) を抽出し、それらの平均値を頭部フロー $V(V_x, V_y)$ とする。
- (4) 前フレームの領域中心 C' と頭部フロー $V(V_x, V_y)$ を用いて、現フレームの領域中心を $C = C' + V$ とし、頭部領域 W を更新する (Fig. 5(b))。
- (5) ハフ変換を用いて、(4) で更新された頭部領域 W 内の画像から円が検出された場合、領域中心 C の座標を検出された円の中心座標として更新させ、頭部領域 W を更新する。

〈3・5〉 人物領域の決定 人物の映り方は Fig. 6(a) のように人物の位置により異なるため、定める人物領域もそれに合わせ変化させる必要がある。人物領域は頭部の位置から人物の中心位置を推定し、定める。また、人物領域は楕円とする。

人物の中心 $G(G_x, G_y)$ を領域中心 C の位置から推定する。Fig. 6(b) は人物を床面上から観測した様子であり、領域中心 C 、人物の中心 G 、人物の足元の座標 U の x 成分の位置関係について表したものである。また、 G は C と U の中点とする。カメラの真下に相当する点の x 座標を O_x とする。 $x = O_x$ の線上 (Fig. 6(a) の白の直線にあたる) に人が立つ

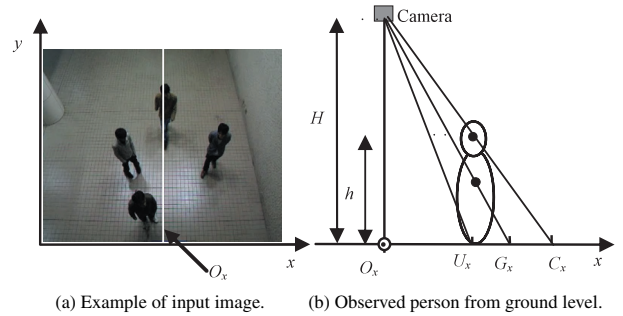


Fig. 6. Relation between C and G .

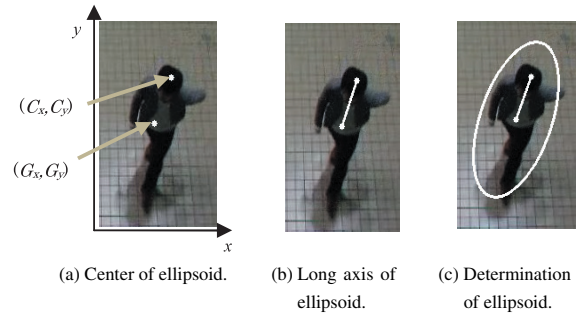


Fig. 7. Human ellipsoid model.

と G, C, U が画像上で垂直に並ぶことになる。なお、領域中心 C は頭部の中心に位置しているものとする。Fig. 6(b) から G_x は (9) 式で表される。また同様に y 軸方向にも O_y を定めると、 G_y は (10) 式で表される。

$$G_x = \left(1 - \frac{h}{2H}\right) (C_x - O_x) \dots\dots\dots (9)$$

$$G_y = \left(1 - \frac{h}{2H}\right) (C_y - O_y) \dots\dots\dots (10)$$

Fig. 7(a) は人物の中心 G と領域中心 C を表示したもので、 G と C を結ぶ直線を作成する楕円の長軸 (Fig. 7(b)) とする。本研究では $O_x = 320, O_y = 0$ とする。

楕円の長径 l_1 を定める。 l_1 は人の厚み、人物の中心 G と領域中心 C 間の距離で表される。Fig. 8(a) はカメラの真下に存在する人物が映る画像であり、Fig. 8(b) はカメラの真下に存在する人物 (厚みを持った人物) を直方体で表し、床面上から観察した様子である。Fig. 8(b) よりカメラに映る人の厚み p_{thick} は実際の人の厚み r_{thick} を用いて (11) 式で表される。

$$p_{thick} = \frac{H}{H - h} \cdot r_{thick} \dots\dots\dots (11)$$

p_{thick} の画素数 q_{thick} を求める。撮影画像内の中心線上 (Fig. 6(a) の O_x) に映る床面の範囲はカメラの垂直画角 θ_{ver} を用いて、 $H \times \tan(\theta_{ver})$ と表される。 p_{thick} の画素数 q_{thick} は (12) 式で表される。なお、 q_{thick} は人物の位置によらず一定の値とした。

$$q_{thick} = \frac{N}{H \cdot \tan \theta_{ver}} \cdot p_{thick} \dots\dots\dots (12)$$

l_1 は領域中心 C と楕円の中心 G 間の距離と q_{thick} を用い

て(13)式で表される。

$$l_l = \frac{h}{2H} \sqrt{(C_x - G_x)^2 + (C_y - G_y)^2 + q_{thick}} \dots (13)$$

本研究では人が歩行する場合や荷物を持っている場合を想定し、 $r_{thick} = 55$ [cm]とし、 $\theta_{ver} = 46$ [°]とした。

楕円の短径 l_s を定める。 l_s は人の横幅 $r_{shoulder}$ 、人の肩までの平均高さ $h_{shoulder}$ を用いて定める。カメラに映る頭部の直径 p_{head} を求めた場合と同様にカメラに映る人物の横幅 $p_{shoulder}$ を求めると(14)式で表される。

$$p_{shoulder} = \frac{H}{H - h_{shoulder}} \cdot r_{shoulder} \dots (14)$$

l_s の画素数は $p_{shoulder}$ の画素数 $q_{shoulder}$ とし、(15)式で表される。なお、 $q_{shoulder}$ は人物の位置によらず一定の値とした。

$$l_s = q_{shoulder} = \frac{M}{2 \cdot H \cdot \tan(\theta_{ver}/2)} \cdot p_{shoulder} \dots (15)$$

本研究では荷物を持つ人物などを想定し、 $r_{shoulder} = 70$ [cm]、 $h_{shoulder} = 140$ [cm]とした。

定めた楕円の中心(人物の中心 G)、長軸、長径、短径で楕円を作成し、人物領域とする(Fig. 7(c))。

〈3・6〉 移動体領域の分離 Fig. 9(a)のように重なった人物の移動体領域を分離し、人数を計測する。

- (1) 上端 P が頭部に位置している人物に対し、頭部領域 W を定める(Fig. 9(a))。
- (2) ハフ変換を用いて頭部領域 W 内のフレーム間差分画像から頭部を円として検出し、1人目の人物を計測する(Fig. 9(b))。
- (3) 検出した頭部から人物領域(楕円)を作成する。人物領域外に移動体領域を抽出し、移動体領域を分割

する(Fig. 9(b))。

- (4) 移動体領域の分割の前後では上端 P の位置は変わり、検出されていない人物に位置する。Fig. 9(b)の移動体領域に(1)~(2)の処理を行うことにより頭部を円として検出し2人目の人物を計測することができる(Fig. 9(c))。

上記の手順を行うことにより、3人目以降の人物も計測することができる。

4. 性能評価

〈4・1〉 実験装置 本システムを構成する機器をTable 1に示す。撮影に使用するカメラは一般的なデジタルカメラを用い、動画撮影モードで撮影を行った。カメラをエレベータの扉上方の側壁に設置した。また、俯角を27°(真下が撮影できる角度)とした。

〈4・2〉 評価方法 Fig. 1(a)のようなエレベータ乗場の映像を用いてシステムの性能評価を行う。エレベータ到着時に全身が映っている人物を待客とし、エレベータ到着時の計測人数と待客数の差異の絶対値を計測誤差とする。実験は(i)と(ii)それぞれの条件で実施した。

- (i) カメラの設置高さ500cm、待客が1~10人存在する状況で待客を計測した場合。閾値 S_{th} は3000画素とした。
- (ii) カメラの設置高さ380cm、待客が1~7人存在する状況で待客を計測した場合。閾値 S_{th} は5500画素とした。

なお、実験箇所はともに同一の場所で行った。これらの実験条件を定めるにあたり、実用を想定した環境を選んでいく。まず、乗場の人数については条件(i)で10人乗りのエレベータを想定し、条件(ii)では(i)と同程度の密集度となるよう7人までの実験結果を取っている。照明については鏡面反射や影の発生による障害が想定されるため、反射性

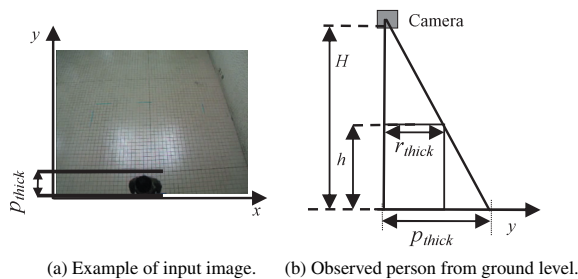


Fig. 8. Thickness of human.

Table 1. Experimental apparatus.

OS	Microsoft Windows Vista
CPU	Intel Celeron 2.8 GHz
Camera	FUJIFILM FinePix S9000
Resolution	640×480
Frame rate	10 fps

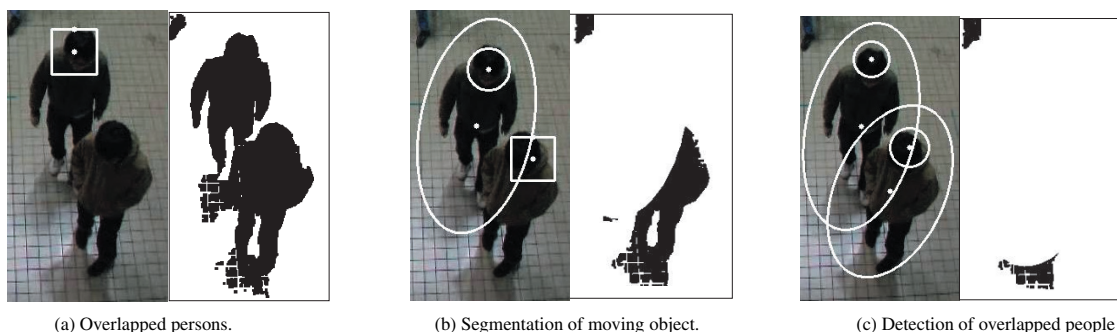


Fig. 9. Segmentation of moving object.

Table 2. Evaluation index.

Counting rate(A)	$(1-(b+c)/a) \times 100$
Rate of P-error (B)	$b/a \times 100$
Rate of N-error (C)	$c/a \times 100$
Rate of overlapped persons (D)	$d/a \times 100$
Counting rate of overlapped persons (E)	$(1-e/d) \times 100$

Table 3. Experimental result.

case	(i)	(ii)
Real number of people (a)	179	217
P-error (b)	2	3
N-error (c)	8	13
Real number of overlapped persons (d)	77	123
O-error (e)	10	13
Counting rate A [%]	94.4	92.6
Rate of P-error B [%]	1.1	1.4
Rate of N-error C [%]	4.5	6.0
Rate of overlapped persons D[%]	43.0	50.7
Counting rate of overlapped persons E[%]	87.0	89.4

の高い床面と強い照明の場所を選んでいる。本稿の図からも床面の照明の反射によるハイライトや人の影が視認されるが、このような環境においても本手法の人数計測に支障がないことを確認する。また、人の多様性に対処できるか否かを調べるため、荷物や帽子を被った人、および走行中の人をサンプルに加えて検証を行った。

実験の評価指標として Table 2 を用いる。Table 2 におけるそれぞれの値は次のようになる。

- (a) **Real number of people** : 1 回の実験での待客数の合計
- (b) **P-error** : 1 回の実験において、エレベータ到着時に計測人数が待客数を上回る場合の計測誤差の合計
- (c) **N-error** : 1 回の実験において、エレベータ到着時に計測人数が待客数を下回る場合の計測誤差の合計
- (d) **Real number of overlapped people** : 1 回の実験において、エレベータ到着時に重なった待客の数の合計
- (e) **O-error** : 1 回の実験において、エレベータ到着時に重なった待客を対象にした場合の計測誤差の合計

〈4・3〉 評価実験 条件 (i), 条件 (ii) における実験結果を Table 3 に示す。

計測率 A の値は、設置高さ 500 cm の条件 (i) で 94.4%、設置高さ 380 cm の条件 (ii) で 92.6% となり、良好な結果となった。高さ 380 cm の条件 (ii) の計測率 A の値より高さ 500 cm の条件 (i) の計測率 A の値が下回った理由はカメラの設置高さにある。重なる人物の割合 D は高さ 500 cm の条件 (i) の値よりも高さ 380 cm の条件 (ii) の値が上回り、設置高さを低くすると重なることが多くなる結果となった。重なりが多くなることにより計測誤差が増加したことが、



(a) Miss detection. (b) Miss tracking.

Fig. 10. Example of error cause.

(ii) の計測率 A の値が (i) の計測率 A の値を下回った理由である。

両条件の実験での全計測誤差数のうち、重なりが計測誤差の原因となったものは 88.5%、重なりが原因でないものは 11.5% となった。

重なりが原因となり、計測人数が待客数を下回る場合の主な原因は、移動体領域の分割の際に、人物の検出漏れが発生することが挙げられる (Fig. 10(a))。そして、計測人数が待客数を上回る場合の主な原因は、Fig. 10(b) のように他の人物の腕など早い動きの影響を受け、頭部を追跡できなくなり同じ人物を再度検出してしまうことが挙げられる。また、実験中に画像内で頭部が重なることはなかった。

上記の実験結果のうち、重なった人物を対象とした計測率 E の値は条件 (i) で 87.0%、条件 (ii) で 89.4% となった。つまり、重なりによって計測精度が多少低下しているものの、重なりが生じたパターンの中の大多数では正しい計測が実施されたとと言える。人同士が密接に重なっている状況における誤検知の回避が困難なことは文献(7)など、斜め上方より撮影を行う他の手法でも報告されているが、本手法では先述のように頭部の追跡により誤検知の多くを回避している。提案法により重なりが多い状況において高い計測精度を得られることが確認された。

一方、重なりの影響を受けず、計測誤差が生じた原因としては、人物の影や、人物がボタンを押す動作などが人物領域内に収まらず、人物の一部を再度検出することが挙げられる。

また、計測を行った中には以下の人物が含まれていた。

- Case1 荷物 1** : 人物が荷物を持っている場合
- Case2 荷物 2** : 人物が大きな荷物を持っている場合
- Case3 走行** : 人物が走行している場合
- Case4 帽子** : 人物が帽子を被っている場合

Case1 では荷物が人物領域内に含まれ、計測誤差の原因にならなかった (Fig. 11(a))。また、Case2 としては台車を押した人物がエレベータを利用した。人物領域外の移動体領域の画素数が閾値 Th_a 以上となるが、頭部領域 W 内に円が検出されず、計測誤差の原因にならなかった (Fig. 11(b))。Case3, Case4 でも、計測誤差の原因にならなかった (Fig. 11(c), Fig. 11(d))。

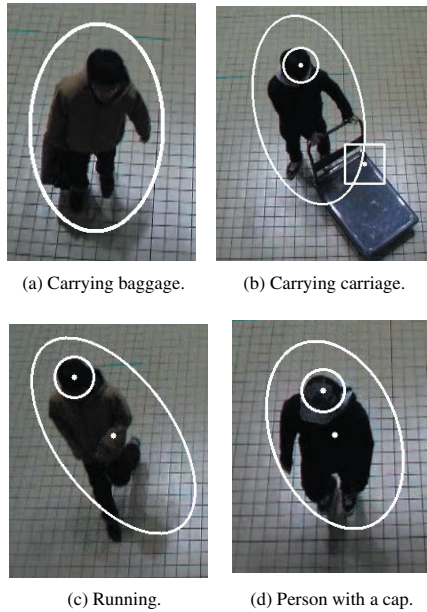


Fig. 11. Example of various persons.

5. まとめ

動画画像から人物の頭部を検出する方法を提案し、本手法を用いてエレベータ乗場の待客数を計測した。カメラを高さ 500 cm, 380 cm に設置した場合の計測率はそれぞれ 94.4%, 92.6% となり、良好な結果となった。また、重なった人物のみを対象にした計測率も求めた。重なりにくい頭部を追跡し、移動体領域を分割することにより、カメラを高さ 500 cm, 380 cm に設置した場合の計測率はそれぞれ 87.0%, 89.4% となり、良好な結果となった。このことにより重なりに強い待客計測法であることが確認された。

(平成 21 年 5 月 14 日受付, 平成 21 年 8 月 3 日再受付)

文 献

- (1) H. Yoda, J. Motoike, M. Ejiri, and T. Yuminaka: "A Measurement Method of the Number of Passengers Using Real-time TV Image Processing Techniques", *Trans. IEICE*, 69-D, 11, pp.1679-1686 (1986) (in Japanese) 依田晴夫・本池 順・江尻正員・弓仲武雄:「実時間画像処理を用いた待客数検出方式」, *信学論*, 69-D, 11, pp.1679-1686 (1986)
- (2) H. Uchidate, R. Inoda, T. Tsuji, and S. Abe: "Counting people and recognition of wheelchairs at elevator lobby by real-time image processing", *Trans. IEIJ*, Vol.129-D, No.6, pp.578-584 (2009-6) (in Japanese) 内館 光・猪田良介・辻 俊明・阿部 茂:「実時間画像処理によるエレベータ乗場の人数計測と車椅子の認識」, *電学論*, 129-D, 6, pp.578-584 (2009-6)
- (3) D.B. Yang, H.H. Gonzalez-Banos, and L.J. Guibas: "Counting People in Crowds with a Real-Time Network of Simple Image Sensors", 2003 ICCV, Vol.1, pp.122-129 (2003)
- (4) S.-F. Lin, J.-Y. Chen, and H.-X. Chao: "Estimation of Number of People in Crowded Scenes Using Perspective Transformation", *IEEE Trans. Systems*,

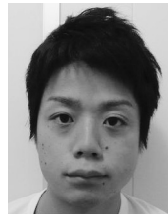
Man and Cybernetics, Part A, Vol.35, No.6, pp.645-654 (2001)

- (5) T. Zhao and R. Nevatia: "Bayesian human segmentation in crowded situations", 2003 CVPR, Vol.2, pp.459-466 (2003)
- (6) K. Nakagawa, K. Terada, and Y. Ishizuka: "The extraction of people region by two times of Hough Transform", *The Papers of Technical Meeting on Industrial Instrumentation and Control, IEIJ*, pp.11-14 (2008) (in Japanese) 中川康紀・寺田賢治・石塚善士:「2 回の一般化ハフ変換を用いた人領域の抽出」, *電学産業計測制御研究会*, pp.11-14 (2008)
- (7) O. Sidla, Y. Lypetsky, N. Brandle, and S. Seer: "Pedestrian Detection and Tracking for Counting Applications in Crowded Situations", *AVSS*, p.70
- (8) S. Haritaoglu, D. Harwood, and L.S. Davis: "W4: Real-Time Surveillance of People and Their Activities", *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol.22, No.8 (2000)
- (9) 馬場功淳・大橋 健・乃万司・松尾英明・江島俊朗:「Head Finder: フレーム間差分をベースにした人物追跡」, *画像センシングシンポジウム 2000*, pp.329-334 (2000)
- (10) J.-Y. Bouguet: "Pyramidal Implementation of the Lucas Kanade Feature Tracker Description of the algorithm", *OpenCV, Documentation, Microprocessor Research Labs, Intel Corporation* (1999)
- (11) J. Shi and C. Tomasi: "Good Features to Track", 1994 IEEE Conference on Computer Vision and Pattern, Recognition (CVPR'94), pp.593-600 (1994)
- (12) C. Harris and M. Stephens: A Combined Corner and Edge Detector, *Proc. 14th Alvey Vision Conference*, pp.147-151 (1988)

辻 俊 明 (正員) 1978 年 7 月 9 日生。2006 年 3 月慶應義塾大学理工学研究科総合デザイン工学専攻後期博士課程修了。同年 4 月東京理科大学工学部第一部機械工学科助手。現在、埼玉大学工学部電気電子システム工学科助手。博士(工学)。主としてロボティクス、モーションコントロール、画像処理に関する研究に従事。



田 代 浩 紀 (非会員) 1985 年 9 月 5 日生。2009 年 3 月埼玉大学工学部電気電子システム工学科卒業。同年、JFE エンジニアリング(株)入社。在学中は主として画像処理の研究に従事。



阿 部 茂 (正員) 1949 年 3 月 29 日生。1971 年 6 月東京大学工学部電子工学科卒業。1976 年同大学大学院博士課程修了。工学博士。同年三菱電機(株)入社。中央研究所, 産業システム研究所で電力系統, オブジェクト指向, 計算機幾何応用のシステム研究開発に従事。1997 年同社稲沢製作所エレベータ開発部長。2001 年ビルシステム事業本部技師長。2004 年 4 月埼玉大学工学部教授, 1985 年電気学会論文賞受賞。IEEE, 電子情報通信学会, 情報処理学会各会員。

