

複数カメラを用いた人数分布推定に関する検討

田渕 義宗[†] 高橋 友和[†] 出口 大輔[†] 井手 一郎[†] 村瀬 洋[†]

黒住 隆行^{††} 柏野 邦夫^{††}

[†] 名古屋大学 〒464-8601 愛知県名古屋市千種区不老町

^{††} 日本電信電話株式会社 NTT コミュニケーション科学基礎研究所 〒243-0198 神奈川県厚木市森の里若宮 3-1

E-mail: †tabuchiy@murase.m.is.nagoya-u.ac.jp, †ttakahashi@gifu.shotoku.ac.jp, †ddeguchi@nagoya-u.jp,
†{ide,murase}@is.nagoya-u.ac.jp, ††{kurozumi.takayuki,kashino.kunio}@lab.ntt.co.jp

あらまし 公共の安全やマーケティングを目的として、群集の解析の研究が注目を集めている。また近年、セキュリティ意識が高まり、監視カメラから得られる大量の映像を用いて群集の行動を解析する技術が求められている。しかし、人手による解析には多くの時間や労力が必要なため、このような大量の群集映像の自動解析に関する期待は高い。群集解析の中でも、場所領域毎の人数分布の推定には一定の需要がある。そこで我々は、多数の人が密集して存在する状況下での場所領域毎の人数分布の推定を目指す。本報告では、複数カメラを用いることで遮蔽の影響に対処するとともに、回帰による小領域毎の人数推定により人数分布を推定する手法を検討する。評価実験において、提案手法の人数分布推定精度を評価した。

キーワード 人数分布推定, 群集解析, 複数カメラ

A study on people density estimation using multiple cameras

Yoshimune TABUCHI[†], Tomokazu TAKAHASHI[†], Daisuke DEGUCHI[†], Ichiro IDE[†],

Hiroshi MURASE[†], Takayuki KUROZUMI^{††}, and Kunio KASHINO^{††}

[†] Nagoya University

Furo-cho, Chikusa-ku, Nagoya-shi, Aichi, 464-8601 Japan

^{††} NTT Communication Science Laboratories, Nippon Telegraph and Telephone Corporation

3-1, Morinosato-Wakamiya, Atsugi-shi, Kanagawa, 243-0198 Japan

E-mail: †tabuchiy@murase.m.is.nagoya-u.ac.jp, †ttakahashi@gifu.shotoku.ac.jp, †ddeguchi@nagoya-u.jp,
†{ide,murase}@is.nagoya-u.ac.jp, ††{kurozumi.takayuki,kashino.kunio}@lab.ntt.co.jp

Abstract Crowd analysis is focused on for public safety and marketing. Also, in recent years, it has become necessary to analyze a large volume of video data recording the activity of a crowd according to the widespread of surveillance cameras due to the increase of security awareness. Automatic analysis of such crowd videos is in demand because manual analysis takes time and effort. Among the crowd analysis, there is a demand for people density estimation. Therefore, we aim to estimate the density of people in a situation where many people exist, using multiple cameras to deal with occlusion. In this report, we study a method which estimates the people density by counting people based on regression in small areas. In an experiment, we examined the accuracy of the proposed people density estimation method.

Key words People density estimation, crowd analysis, multiple cameras

1. ま え が き

公共の安全やマーケティングを目的として、人数の計数、群集の行動調査といった群集の解析が行なわれている。また近年、

セキュリティ意識の高まりから、監視カメラから得られる大量の映像から群集を解析する技術が求められている。しかし、人手による解析には多くの時間や労力が必要なため、このような大量の群集映像の自動解析に関する期待は高い。

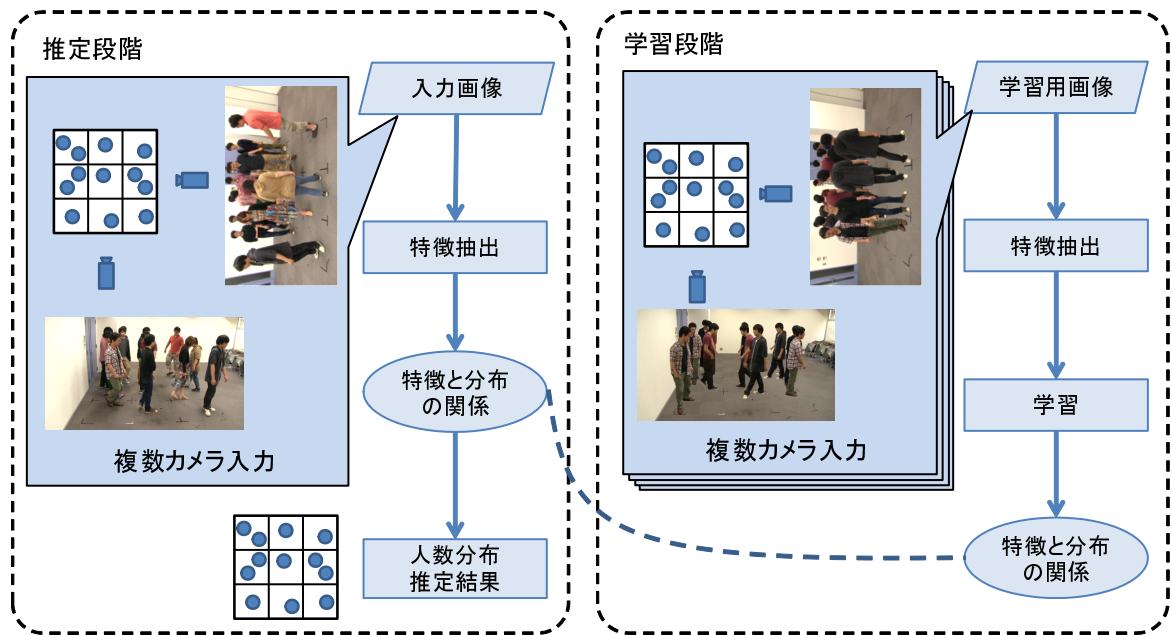


図3 人数推定を用いた人数分布推定手法の処理の流れ

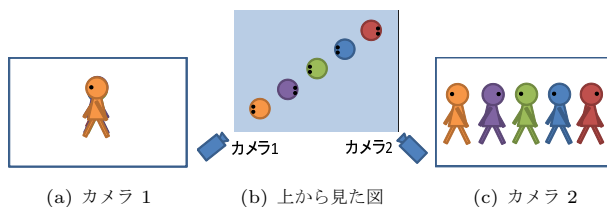


図1 単一カメラで正しく人の存在を把握できない例

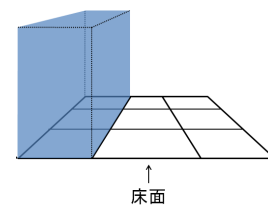


図4 撮影面分割領域

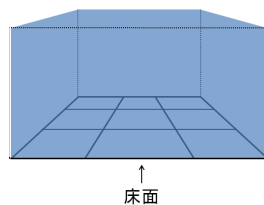


図2 床面の分割

これまでに、群集解析を目的として人数推定や群衆の行動解析といった様々な研究が行なわれている。カメラを用いた人数推定の研究として、人の顔など特定の形状を検出して計数する「検出に基づく手法」[1]、画像特徴量と人数の関係の回帰に基づく「回帰に基づく手法」[2]などが提案されている。しかし、これらはカメラの撮影範囲内に存在する人数の推定を目的としているため、撮影範囲内の人数がどのように分布しているかを考慮していない。また、仮想的なゲートを設定し、通過する人を検出して人数を計数する手法[3]もあるが、人の分布のばらつきは考慮されていない。

群衆の行動を解析する研究としては、広域環境における視野を共有しない複数カメラを用い、人々が形成するグループの対応付けにより人流の計測を行なう研究[4]、人物の移動量を推定することで人流の計測を行なう研究[5]などが挙げられる。しかし、これらの研究は人の流れの解析が目的であり、これらを

直接人数の分布の推定に用いることは出来ない。

図1は単一カメラで人の存在を正しく把握できない例を表している。ここで、図1(b)のような人の配置をカメラ1から撮影すると、得られる画像は図1(a)のようになる。このように、単一のカメラのみでは、遮蔽の影響により、人の存在を正確に把握することが困難である。一方、同じ人の配置をカメラ2から撮影した場合は、図1(c)のように人の存在を正確に把握することができる。このように、視点が異なり、視野重複がある複数のカメラを用いることにより、より多くの情報をもとに推定が可能となる。同様の問題を上方に設置された単一の監視カメラにより解決するアプローチ[6]も提案されているが、上方にカメラを設置できる状況は限られる。

そこで我々は、複数カメラを用いることで遮蔽に頑健な人数分布推定を目指す。本報告では、回帰により小領域毎の人数分布を推定する手法を検討する。

以下、2節で人数分布推定手法について述べ、3節で実験と結果を述べる。最後に、4節でまとめと今後の課題を述べる。

2. 回帰による小領域毎の人数分布推定手法

ここでは、図2のように床面を 3×3 の9つの領域に分割し、各領域に存在する人数の推定により全体の人数分布を求める。以降、この分割された床面の各領域を床面分割領域と呼ぶ。

回帰による小領域毎の人数分布推定手法は、学習段階と推定

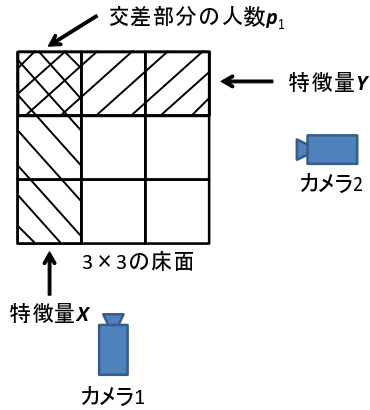


図5 学習段階に用いる特徴と人数の関係の例

段階から構成される。図3にこの処理の流れを示す。本報告で検討する手法は複数台のカメラに適用可能であるが、説明の簡単化のため、以降ではカメラ2台の場合の学習段階と推定段階について具体的な処理手順を与える。また、図4のようにカメラの撮影面を左右方向に3分割した領域を撮影面分割領域と呼ぶ。

2.1 学習段階

- (1) 2台のカメラから学習用画像を取得し、正解の人数分布を手作業で入力することにより、学習データを作成する。
- (2) 各学習用画像から、撮影面分割領域毎に特徴を抽出する。
- (3) 各画像の撮影面分割領域における特徴量とその交差領域における人数（例えば、図5における特徴量 X, Y と人数 p_1 ）との関係を回帰により学習する。

2.2 推定段階

- (1) 学習段階と同じ配置の2台のカメラを用いて人数分布推定に用いる画像2枚を撮影する。
- (2) 入力された画像から撮影面分割領域ごとに学習段階と同じ手順により特徴を抽出する。
- (3) 各画像の撮影面分割領域から算出した特徴量に基づく回帰により床面分割領域毎の人数を推定する。各床面分割領域に対する推定結果を組み合わせることにより、人数分布の推定結果を得る。

3. 人数分布推定実験

提案手法の有効性を調査するため、人数分布推定実験を行った。

3.1 実験条件

本実験では $300 \times 300 \text{ cm}^2$ の床面を 3×3 の9つの領域に分割し、人数の分布の推定を行なった。各床面分割領域に存在する人数は、最大で4人である。図6に示すように、2台のカメラを約90度の位置に設置し、両カメラの注視点が同じになるように調整した。図6にカメラと推定領域の位置関係を示す。人数分布推定における回帰の特徴量として、(1) 前景の画素数、(2) 前景領域のエッジの画素数、(3) 前景領域と背景領域との境界の画素数、の3つの特徴を用いた（図7）。本実験では前

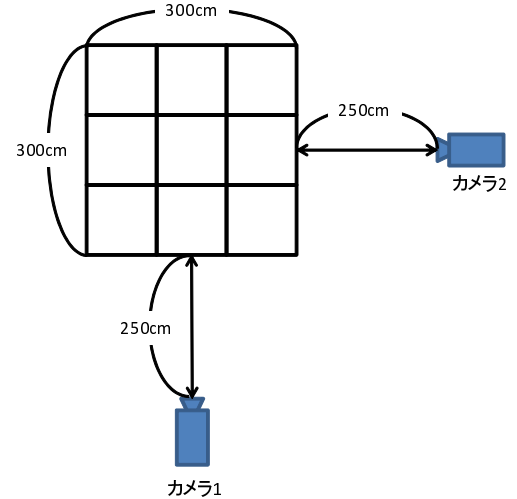


図6 カメラと推定領域の位置関係

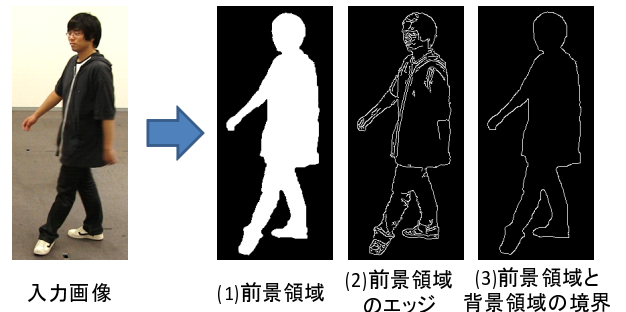


図7 特徴量の例

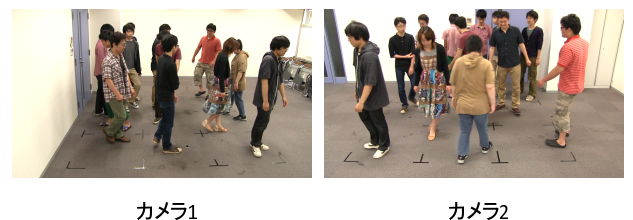


図8 評価用データの例

景抽出に正規化距離を用いた背景差分法 [7] を使い、エッジ検出手法として Canny エッジ検出器 [8] を用いた。評価尺度は、推定された人数分布と正解との平均絶対誤差を用いた。評価用データとして、カメラ1台あたり1~12人が写った画像230枚を用い、画像4,500枚を生成し学習データとして用いた。画像の生成方法を次節で詳細に述べる。また、実際に評価用データとして使用した画像の例を図8に示す。

3.2 学習用画像の生成

図9に学習画像の生成方法の流れを示す。まず、床面を 3×3 の小領域に分割する。それぞれの床面分割領域内に、0~4人が存在する状況を2台のカメラで事前に撮影する。そして、手作業で人物領域を切り出す。切り出された複数の人物画像と背景領域を合成することにより、学習用画像を生成する。この作業をカメラ毎に行ない、学習用画像を得る。実際に生成した学習用画像の例を図10に示す。図10(a)からわかるように、生成

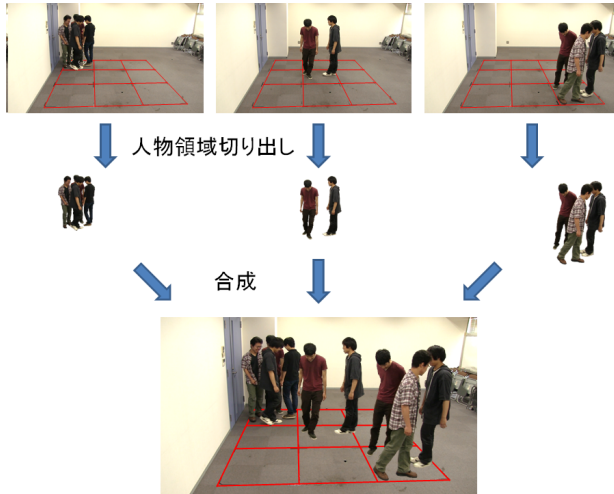


図 9 学習用画像の生成例

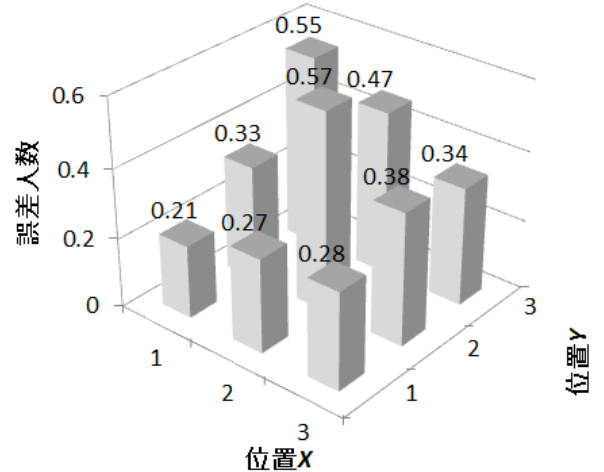


図 11 分布位置と誤差人数

$(x, y) = (1, 3)$ の誤差が 0.55 人と推定精度が低下していることがわかる。これは、カメラに対してより近い位置にいる人物による遮蔽が発生し、人数推定の精度が低下したためと考えられる。

実際に、 $(x, y) = (1, 3)$ における推定が失敗した例とその際の人数分布の正解とその推定結果、および絶対値誤差を図 12 に示す。結果から、概ね良好に推定できているが、 $(x, y) = (1, 3)$ でのみ推定精度が低下していることがわかる。これは、どちらのカメラにおいても $(x, y) = (1, 3)$ の位置の 2 人が、手前の人物により隠れているためだと考えられる。このように、どちらのカメラからも離れた位置の推定は遮蔽による影響を強く受けるため、正しい人数推定が困難である。

また、1 つの床面分割領域に 3 人以上が存在する例とその際の人数分布の正解とその推定結果、および絶対値誤差を図 13 に示す。結果から、床面分割領域の人密度が高い (3 人以上存在する) 領域は推定が正しく行なわれないことがわかる。これは、学習用画像において、床面分割領域の人密度に偏りがあるのが原因だと考えられる。そのため、学習用画像におけるデータの偏りを低減するしくみの検討が必要である。

また、図 14 の $(x, y) = (2, 2)$ や $(x, y) = (3, 1)$ のように、人が存在しない床面分割領域において、推定結果が 1 人以上となる例が見られた。これは、推定対象の床面分割領域の周囲の人物から得られる特徴量が影響したためである。そのため、特徴量の抽出領域を細分化する等の工夫が必要である。

4. まとめ

本報告では、複数カメラを用いた回帰による小領域毎の人数分布推定手法を検討した。具体的には、複数カメラにより得られる画像を撮影面分割領域に分け、各領域から特徴量を抽出した。そして、撮影面分割領域が交わる領域の人数の回帰により人数分布を推定した。

評価実験では、提案手法による人数分布推定精度を評価した。実験の結果から、回帰による小領域ごとの人数分布推定手法で



(a)



(b)

図 10 生成学習用画像の例

した画像では前後の床面分割領域における人の重なりを考慮した。また、各床面分割領域から切り出された画像を用いるため、手前の人物ほど大きくなるように、自然な画像合成を実現した。しかし、1 人 1 人を分けて切り出していないため、左右の床面分割領域においては人の重なり順が不自然になってしまうという問題点がある (図 10(b))。また、服装・人の種類が事前に撮影した画像に強く依存するという問題点もある。なお、本実験では各床面分割領域に対して 5 パターン (0~4 人それぞれ 1 パターン) を撮影して学習用画像の生成に用いた。

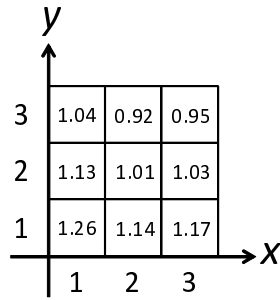
3.3 実験結果と考察

図 11 に本報告で検討した人数分布推定手法の各分布位置に対する推定誤差を示す。誤差が 0.21 人と最も小さい位置 $(x, y) = (1, 1)$ に比べて、両カメラから遠い位置である

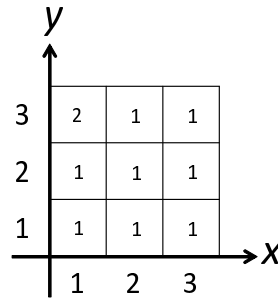


(a) カメラ 1

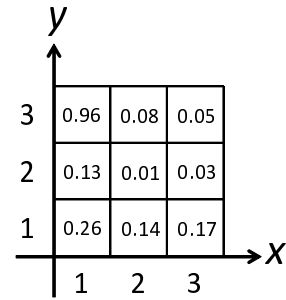
(b) カメラ 2



(c) 推定結果



(d) 正解



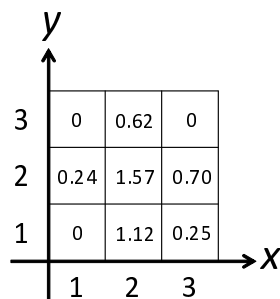
(e) 絶対値誤差

図 12 $(x, y) = (1, 3)$ の推定がうまくいかない例

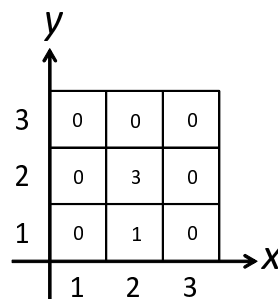


(a) カメラ 1

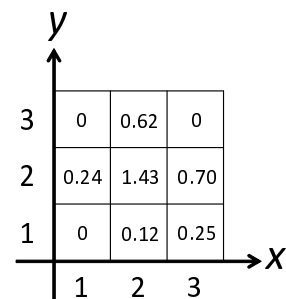
(b) カメラ 2



(c) 推定結果



(d) 正解



(e) 絶対値誤差

図 13 1つの床面分割領域に3人以上いる場合の例

は、カメラから遠くなるにつれ推定精度が低下する傾向があることを確認した。また、床面分割領域の人密度が高い場合や、推定対象となる床面分割領域の周囲の人物が特徴量計算に悪影響を及ぼす場合に、人数分布の推定精度が低下することを確認した。

今後の課題として、学習用データの種類の増加、特徴量を抽

出する領域の検討、より遮蔽に頑健な人数分布推定手法の検討が挙げられる。また、人数分布のパターンと画像特徴量の対応表を事前に作成し、入力の特徴ベクトルとの照合により人数分布推定を行なう手法の検討も行なう予定である。



(a) カメラ 1

(b) カメラ 2

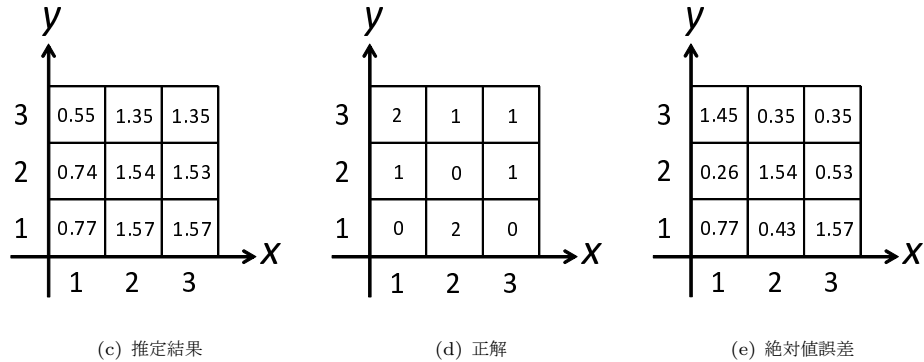


図 14 0 人の領域が 1 人以上と推定されてしまう場合の例

謝 辞

日頃より熱心に御討論頂く名古屋大学村瀬研究室諸氏に深く感謝する。本研究の一部は、科学技術研究費補助金による。また、本研究では画像処理に MIST ライブラリ (<http://mist.murase.m.is.nagoya-u.ac.jp/>) を使用した。

文 献

- [1] M. Li, Z. Zhang, K. Huang, T. Tan, “Estimating the number of people in crowded scenes by mid based foreground segmentation and head-shoulder detection,” Proc. 19th International Conf. Pattern Recognition, WeAT2.1, pp.1-4, Dec. 2008.
- [2] A. Chan, Z. Liang, N. Vasconcelos, “Privacy preserving crowd monitoring: Counting people without people models or tracking,” Proc. IEEE Computer Society Conf. Computer Vision and Pattern Recognition, pp.1-7, June 2008.
- [3] 寺田賢治, 吉田大輔, 大恵俊一郎, 山口順一, “ステレオカメラを用いた通行人数の自動カウント,” 電気学会論文誌. (C), vol.120, no.6, pp.850-856, 2000.
- [4] 芥ゆかり, 新田直子, 馬場口登, “疎分散カメラ間の人物グループの対応付けによる人流解析,” 電子情報通信学会技術研究報告. PRMU2010-272, pp.205-210, March 2011.
- [5] 中崎孝之, 新田直子, 中村和晃, 馬場口登, “疎分散カメラにおける前景領域間の移動量推定による広域人流計測,” 電子情報通信学会技術研究報告. PRMU2012-195, pp.87-92, March 2013.
- [6] 庭川 誠, 恩田寿和, “監視カメラ映像中の局所的な動き検出とイベント累積による時間的かつ空間的な混雑度調査,” 電気学会論文誌 (D), vol.124, no.10, pp.1060-1066, 2004.
- [7] 長屋茂喜, 宮武孝文, 藤田武洋, 上田博唯, 伊藤敬一, “時間相関型背景判定法による移動物体検出,” 電子情報通信学会論文誌 (D-II), vol.J79-D-II, no.4, pp.568-576, 1996.
- [8] J. Canny, “A computational approach to edge detection,” IEEE Trans. Pattern Analysis and Machine Intelligence, vol.8, no.6, pp.679-698, 1986.