一般社団法人　電子情報通信学会
THE INSTITUTE OF ELECTRONICS,
INFORMATION AND COMMUNICATION ENGINEERS

信学技報
IEICE Technical Report
PRMU2013-127,CNR2013-35(2014-2)

空間的な人数分布推定のための記憶型回帰の検討

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　カメラを用いた群集解析は公共の安全やマーケティングなどに需要がある．群集解析の中でも，我々は小領域毎の人数を求める空間的な人数分布推定を検討する．しかし，カメラから離れた場所では人物間の遮蔽が発生するので，空間的な人数分布を正しく推定することは困難である．そこで我々は，複数視点のカメラから撮影された画像を用いて記憶型回帰により人数分布を推定する手法を提案する．本手法は，人数分布と群集の見えの対応表を用いるものであり，群集の見えは遮蔽を含んでいるため，遮蔽に頑健な人数分布推定手法が実現される．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，記憶型回帰

A study on memory-based regression for spatial people density estimation

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

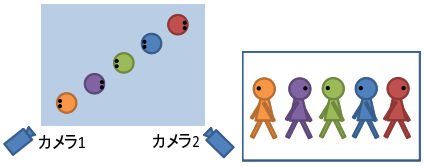
Abstract　Crowd analysis using camera is required for public safety and marketing. In this report, among the crowd analysis techniques, we focus on the estimation of spatial people density. It is difficult to accurately estimate people density far from a camera because of occlusion between people. Therefore, we propose a memory-based regression method for spatial people density estimation with images captured from cameras from multiple viewpoints. This method makes use of a table that consists of correspondences between people density maps and crowd appearances. Since the crowd appearances include situations where various occlusions occur, an estimation robust to occlusion should be achieved. In an experiment, we examined the effectiveness of the proposed method.

Key words　People density estimation, crowd analysis, multiple cameras, memory-based regression

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対する期待は高い．群集解析の中でも我々は，カメラを用いた空間的な人数分布推定に着目する．これは，図 1 のように入力画像から人数の空間的な分布を推定する技術である．これにより，群集の分布のより詳細な情報を得ることができ，マーケティングなどに利用できる．

This article is a technical report without peer review, and its polished and/or extended version may be published elsewhere.
Copyright ©2014 by IEICE



(a) カメラ 1　　(b) 上から見た図　　(c) カメラ 2

図 2　単一カメラで正しく人の存在を把握できない例

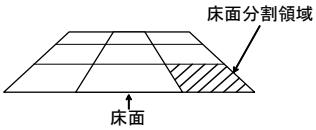


図 3　床面の分割例

図 1　画像からの空間的な人数分布推定の概念図

これまでに，群集解析を目的として人数推定や人物追跡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手法 [1]，画像特徴量と人数の関係の回帰に基づく手法 [2]，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] 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．一方，人物追跡の研究としては，人体のパーツを用いる手法 [4]，視野を共有しない複数カメラを用い，カメラ間で同一人物を推定し，対応付けることにより追跡を行なう研究 [5] [6] が挙げられる．しかし，これらの手法を直接人数の分布の推定に用いることはできない．

カメラからの空間的な人数分布推定では，カメラから離れた位置にいる人物ほど，手前にいる人物による遮蔽が発生するため，正しい推定が困難になるという問題点がある．そこで本報告では，複数視点からのカメラ画像を用いた記憶型回帰による人数分布推定手法を提案する．記憶型回帰は図 1 に示すように，群集の見えと人数分布の対応表を作成し，入力画像と対応表を比較することにより実現される．人による遮蔽を含めた群集の見えを用いることにより，遮蔽に頑健な推定を実現する．また，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で，より多くの情報をもとに推定できる．図 2 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 2(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 2(a) のようになる．このように，遮蔽の影響により，単一のカメラのみでは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 2(c) のように人の存在を正確に把握することができる．

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 [7] 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．

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.　記憶型回帰による人数分布推定

記憶型回帰による人数分布推定手法は，床面を複数の小領域に分割し，各小領域に存在する人数を推定すること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．本報告では，図 3 のように床面を 3×3 の 9 つの領域に分割する．

図 4 に提案手法の処理の流れを示す．提案手法は，(I) 学習画像生成段階，(II) 対応表作成段階，(III) 推定段階の 3 つの処理によって構成される．まず (I) の学習画像生成段階では，学習用画像を画像合成により生成する．次に (II) の対応表作成段階では，学習画像から特徴量を抽出し，抽出された特徴量と対応する人数分布を対応表に登録する．最後に (III) の人数分布推定段階では，入力画像から抽出された特徴量と対応表の特徴量を比較することにより空間的な人数分布を推定する．

本報告で提案する手法は，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 台の場合について具体的な処理手順を与える．以降，2.1 節，2.2 節，2.3 節でそれぞれの処理について詳しく述べる．

2.1　学習画像生成段階

図 5 に学習画像生成の流れを示す．

（1）　各々の床面分割領域内に，0〜4 人が存在する状況をカメラで事前に撮影する．

（2）　撮影された画像から手作業で人物領域を切り出す．

（3）　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．学習画像は全ての人数分布パターンに対して生成し，次の段階である対応表作成段階で使用する．実際に生成された学習用画像の例を図 6 に示す．これからわかるように，生成した画像では前後の床面分割領域における人の重なりが考慮されている．また，各床面分割領域から切り出された画像を用いて，手前の人物ほど大きくなるように，自然な画像合成が実現されている．

（4）　この処理を両方のカメラに対して行なう．

2.2　対応表作成段階

（1）　各学習用画像から特徴量を抽出する．ここで，特徴抽出は図 7 のように床面分割領域を高さ方向に 3 分割した領域か

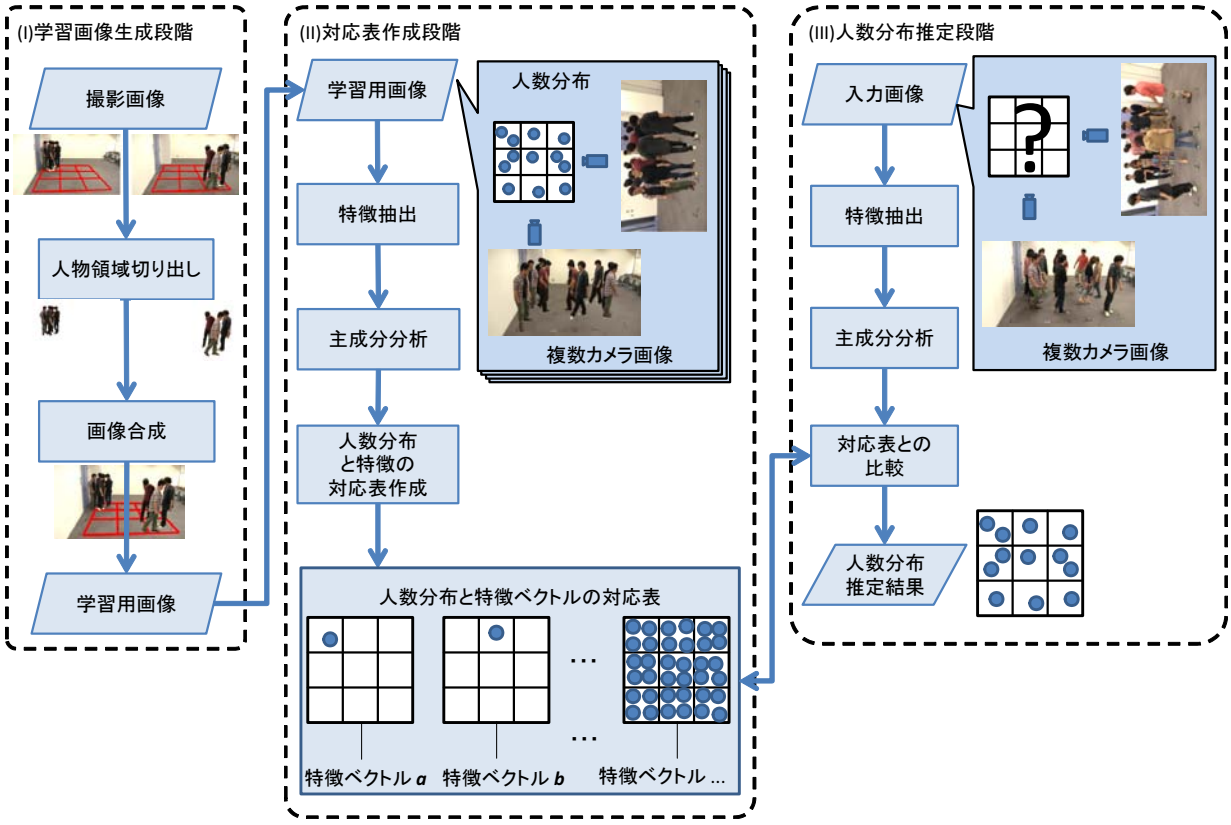


図 4　記憶型回帰による人数分布推定手法の処理の流れ

図 5　学習画像生成段階の処理の流れ

ら行なう．これを空間分割領域と呼ぶ．特徴抽出領域をこのように分割するのは，人の頭部，胴体など，部位によって特徴量が異なるためである．それぞれの空間分割領域からの前景の画素数，前景領域のエッジの画素数，前景領域と背景領域との境界の画素数の 3 つの画像特徴を抽出する．図 8 はこれらの画像特徴の例を示している．特徴ベクトルの次元数は，3（特徴量の種類）× 9（床面分割数）× 3（空間分割数）× 2（カメラ数）= 162 である．

（2） 抽出された特徴量のうち有効な特徴量のみを用いるために主成分分析により次元削減する．この際，固有値が 1 以上の特徴量を主成分として使用する．

図 6　生成された学習用画像の例

（3） 次元削減後の特徴ベクトルと人数分布のパターンを対応付けた表を作成する．

2.3　人数分布推定段階

（1） 対応表作成段階と同じ配置の 2 台のカメラからの画像を入力する．

（2） 対応表作成段階と同じ手順により，入力された画像から空間分割領域ごとに特徴を抽出し，次元削減を行なう．

（3） 入力特徴ベクトルと対応表内の特徴ベクトルを比較し，k 近傍を見つける．異なる人数分布のパターンでも特徴量が似た場合があるため，ここでは最近傍のかわりに k 近傍を用いる．

（4） 推定結果として，k 近傍の特徴ベクトルに対応する人数分布パターンの重み付け和を計算する．この重みは特徴ベクトル間の距離に基づいて計算する．

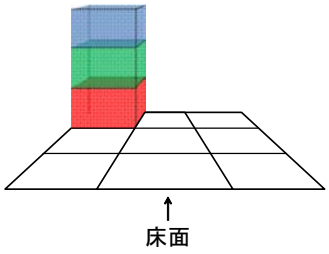
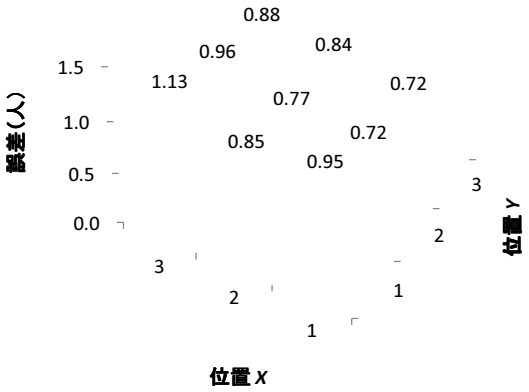


図 7　空間分割領域の例

図 8　画像から抽出される特徴量の例



(a) 比較手法

(b) 提案手法

図 11　推定誤差の比較

図 9　2 台のカメラと推定領域の位置関係

図 10　評価用データの例

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なった．

3.1 実験条件

図 9 にカメラと推定領域の位置関係を示す．本実験では 300×300 cm^2 の床面を 3×3 の 9 つの領域に分割し，人数分布の推定を行なった．本実験では前景抽出に正規化距離を用いた背景差分法 [8] を用い，エッジ検出手法として Canny エッジ検出器 [9] を用いた．評価尺度は，推定された人数分布との平均絶対値誤差を用いた．評価用データとして，2 台のカメラから同時に撮影された 1〜13 人が写った画像 184 枚を用いた．各床面分割領域に存在する人数は，最大で 4 人とした．実際に評価用データとして使用した画像において 13 名の人物が撮影された例を図 10 に示す．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．つまり，学習画像は合計 $5^9 = 1,953,125$ 枚である．

比較手法として小領域毎に人数推定を行なうことにより人数分布を推定する手法を用いる．具体的には，学習段階では提案手法と同様に抽出した特徴量と正解人数との関係を 2 次多項式として回帰し，推定段階では回帰で得られた 2 次多項式を用いて床面分割領域ごとの人数を推定した．

3.2 実験結果と考察

図 11 に提案手法と比較手法の各床面分割領域の推定誤差を示す．この図よりどの位置 (x, y) においても提案手法が比較手

数分布を推定した．評価実験では，提案手法の人数分布推定精度を評価した．実験の結果から，比較手法より提案手法の方が精度が良いことがわかった．今後の課題として，学習用データの種類の増加，有効な特徴量の検討などが挙げられる．

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. on 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. 2008 IEEE Computer Society Conf. on Computer Vision and Pattern Recognition, pp.1–7, June 2008.

[3] K. Terada, D. Yoshida, S. Oe, J. Yamaguchi, “A method of counting the passing people by using the stereo images,” Proc. 1999 International Conf. on Image Processing, vol.2, pp.338–342, Oct. 1999.

[4] B. Wu, R. Nevatia, “Detection and tracking of multiple, partially occluded humans by Bayesian combination of edgelet based part detectors,” International Journal of Computer Vision, vol.75, no.2, pp.247–266, Nov. 2007.

[5] O. Javed, Z. Rasheed, K. Shafique, M. Shah, “Tracking across multiple cameras with disjoint views,” Proc. 9th International Conf. on Computer Vision, vol.2, pp.952–957, Oct. 2003.

[6] O. Javed, K. Shafique, M. Shah, “Appearance modeling for tracking in multiple non-overlapping cameras,” Proc. 2005 IEEE Computer Society Conf. on Computer Vision and Pattern Recognition, vol.2, pp.26–33, June 2005.

[7] 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．

[8] 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．

[9] J. Canny, “A computational approach to edge detection,” IEEE Trans. on Pattern Analysis and Machine Intelligence, vol.8, no.6, pp.679–698, 1986.