

スマートフォンで撮影した近接群衆画像からの 混雑度推定手法に関する一検討

河中 祥吾^{1,a)} 松田 裕貴^{1,2} 藤本 まなと¹ 荒川 豊^{1,3} 安本 慶一¹

概要：観光客が快適に観光を行うためには，観光客の趣味嗜好といった比較的静的な情報と目的地及び移動経路の現在の混雑度といった動的な情報の双方を考慮した観光コンテンツを実時間で提供することが求められる．本研究では，観光中に観光客が無作為に撮影するような近接群衆画像から混雑度を推定する手法を提案する．また，観光客を主な参加者とした参加型センシングを取り入れることで，観光地全体において網羅的に情報を収集することを目指す．本稿では，近接群衆画像から混雑度を推定するための手法として，カメラのシャッター速度を調節することで生じる残像から人物の移動ベクトルを算出する手法を検討し，実際に撮影した残像画像から 23.2 %の精度で移動ベクトルの大きさを求めることができた．加えて，観光客が無作為に撮影する写真から近接群衆画像がどの程度撮影されるかを 14 名の参加者に対して観光実験を行い，無作為に撮影された画像から，45.9 [%] の割合で近接群衆画像が撮影されることが分かった．

Proposal of Crowd Density Estimation Method from Proximity Crowd Photo Image Taken with Smartphone

SHOGO KAWANAKA^{1,a)} YUKI MATSUDA^{1,2} MANATO FUJIMOTO¹ YUTAKA ARAKAWA^{1,3}
KEIICHI YASUMOTO¹

1. はじめに

2020 年の東京オリンピック開催や外国人観光客の増加を背景に，観光客がストレスなく快適に観光を満喫できるインフラやサービスが求められている [1]．観光客の限られた時間の中での満足度を向上させるためには，観光客自身の趣味や嗜好といった静的なユーザ情報や，観光スポットおよび移動経路の混雑度や限定的に行われているイベントといった動的に変化する情報を考慮した上で，実時間で観光コンテンツを提供する必要がある [2]．ここで，実時間に次の観光コンテンツを提供する上で重要となる動的な情報として，各観光スポットや移動経路などにおける混雑度があげられる．実時間に混雑度が把握できることで，動的

に観光プランを最適化でき，観光客の観光満足度向上につながると考えられる [3]．

混雑度情報を提供している既存サービスとして，Yahoo! 地図の混雑レーダー [4] や，いつも NAVI ラボの混雑度マップ [5] などがあげられる．これらのサービスでは，スマートフォンを始めとする GPS 搭載携帯電話で各社が提供するアプリケーションのユーザの中から許諾を得て送信される位置情報を収集し，可視化を行っている．それぞれ，125 m，250 m 四方といった一定範囲のメッシュレベルでの人口推定を行っているが，これらの情報から“清水寺の入り口がどの程度混雑しているか”といった具体的な観光スポットの混雑度を把握することは難しい．

既存研究として公共空間における混雑度や群衆検出手法には，固定カメラや赤外線センサなどの設置型センサを用いる手法が数多く提案されている．固定カメラを用いた手法では，カメラが設置されている範囲に関しては高精度で混雑状況が把握できると考えられるが，今回対象としている観光地において網羅的に混雑状況を把握するには，偏在

¹ 奈良先端科学技術大学院大学
8916-5, Takayama-cho, Ikoma-shi, Nara 630-0192, Japan

² 日本学術振興会特別研究員 DC1
JSPS Research Fellowships for Young Scientists, DC1

³ JST さきがけ
PRESTO

^{a)} kawanaka.shogo.kp1@is.naist.jp

的にカメラが配置されている必要があるため、普及性において問題がある [6] [7] [8] [9]. また、レーザレンジファインダなどの赤外線センサを用いた研究においても、高精度に歩行者の検出を行うことができるが、屋外への適応が難しい点や機器のコストが高い点があげられる [10].

一方、車載カメラの映像や比較的高い場所からスマートフォンにより撮影された俯瞰画像から人流や群衆人数を推定する手法が提案されている [11] [12]. これらの提案では、クラウドソーシングなどの利用を想定されており、網羅的に混雑度情報を収集可能と想定されるが、観光地においては、歩行者天国や寺社仏閣のような車両が侵入できない領域や俯瞰的に撮影することができない環境が多々存在するため、今回の観光地における混雑度推定にそのまま適応することは難しい。

前述した混雑度推定に関する先行研究では、設置型センサを用いた手法と、モバイル端末を用いた手法の2種類に大別することができる。今回の目的として、観光地といった比較的広いエリアにおいて網羅的に混雑度情報を得るためには、コスト面や実現性を考慮するとモバイル端末を用いた手法が適していると言える。また、モバイル端末を用いて網羅的に情報を収集するための方法として、参加型センシングがあげられる [13]. 観光地における参加型センシングでは、観光客が主な参加者となることが想定される。そして、観光客のセンシング行為への参加頻度や継続可能性を考慮すると、観光行動の一部に付随していることが理想的であるといえる。

本研究では、観光行動の一つである写真撮影に着目し、観光客が所持しているスマートフォンから観光中に撮影される写真の中から、人物が写っている近接画像を用いて混雑度を推定する手法を検討する。そこで、我々はカメラのシャッター速度を調整することで撮影可能な移動体の残像を用いて混雑度を推定する手法を提案する。

本手法では、連続的に移動している物体を、シャッター速度を遅くしたカメラで撮影することにより得られる残像画像と、その撮影時間から移動体のベクトルを算出できると考えた。その移動ベクトルの方向および大きさからその撮影エリアの混雑度を推定することを目指す。

本稿では、提案手法の実現に先立ち、撮影した残像画像からどの程度の精度で移動ベクトルの大きさを推定できるかの事前調査を行った結果を報告する。その結果、カメラと被写体間の距離が3 m の時は、平均誤差 27.5 %, 5 m の時は平均誤差 18.9 % で推定可能であることが分かった。また、観光客が無作為に撮影する写真の中に、どの程度近接群衆画像が含まれるかを明らかにするために、観光実験を行った。被験者は、男女含め 14 名とし京都東山区において2時間半程度、写真を取りながら観光してもらった。その結果、合計 595 枚の写真が集まり、そのうち 273 枚、すなわち全体の約 46 % が近接群衆画像であることが

わかった。

2. 関連研究

2.1 設置型センサを用いた人流計測手法

公共施設において人流計測を行う方法として、固定カメラから撮影される映像を用いて人物を検出する手法があげられる。文献 [14] では、ステレオカメラを用いて、実空間を3次元的に把握することで、様々な方向の通過人数及び歩行動線を計測可能かつ屋内外問わず様々な場所でロバストに人数計測が可能となっている。一方で、人物を一人ひとり検出するのではなく、群衆を個の集まりではなく“群”として扱い、従来手法では混雑により人の重なりが大きく解析が難しい状況や、追跡すべき人数が多すぎて処理が重くなる状況でも、群衆の状況を高精度に把握可能な手法も提案されている [9]. これらの手法では、人同士の重なりに強く安定して人流を把握することが可能であるが、カメラが定点で固定されているため、観光地全域を把握するには、多数のカメラを配置する必要がある。

また、レーザ光の伝搬遅延時間を利用して、距離を正確に計測することが可能なレーザレンジスキャナを用いて人流計測手法も数多く提案されている。文献 [10] では、屋内の複数箇所にレーザレンジスキャナを配置し、データを統合することでオクルージョンが頻繁に発生するような密度の高い群衆においても歩行者の検出が可能な手法が提案されている。特定の場所に限定した場合は非常に有用であるが、広いエリアを対象とする場合は多数配置する必要がある上、コストが高いといった問題がある。

2.2 モバイル端末を用いた混雑度推定

モバイル端末を用いて混雑度を推定する手法として、スマートフォンに搭載されている加速度センサやマイクから得られる人間の歩行動作や周囲の環境音の特徴を用いて混雑度判定を行う手法が提案されている。[15] では、スマートフォンをズボンのポケットに入れている状況で、加速度センサの値を用いて歩幅を検出し、その違いから街中の混雑度を空いている状態、混んでいる状態の推定が可能な手法が提案されている¹。また、加速度センサに加えてマイクから周囲の音収集し、これらの特徴量を合わせて混合判定を行う手法も提案されており、屋内環境で手にスマートフォンを把持した状態で歩行を行うことで、群衆密度(低)、群衆密度(中)、群衆密度(高/直線)、群衆密度(高/交差)の4カテゴリに平均 70 % の精度で分類可能であることが報告されている [16].

文献 [12] では、スマートフォンを有する協力者が比較的高所から群衆を撮影した俯瞰画像内の群衆人数を推定する手法が提案されている。0.7 秒の時間間隔で撮影された2枚の画像の差分をとることで群衆が存在する動体領域を抽出し、あらかじめ構築した歩行者面積モデルから動体領域



図 1 混雑度情報収集システムの流れ

における1人あたりの面積を算出することで高精度な群衆人数推定が実現されている。東京や大阪など高層ビルが立ち並ぶ環境においては非常に有効な手法であるが、京都や奈良といった歴史的建造物の多い観光地などにおいては、俯瞰画像が撮影できない環境が多々あると考えられる。また、人がすべて動いている前提で考えられているため、待機行列ができてしまうような状況においては、検出精度が著しく下がってしまうことが考えられる。

3. 提案手法

3.1 システム概要

提案する混雑度情報収集システムの流れを図1に示す。本システムでは、観光客を主な参加者とした参加型センシングにより観光中に撮影される写真を収集する。参加者がスマホアプリを用いて写真を撮影した時に、写真、コメントおよびスマートフォン内蔵のセンサから取得可能な撮影時の位置座標、加速度、地磁気それぞれの値をサーバへ送信する。収集された様々なジャンルの写真の中から、近接群衆画像のみを抽出し、3.2節で述べる推定手法を用いて画像に写っている人の移動ベクトルを算出する。この時、撮影時の位置情報、加速度、地磁気から撮影方向、画角、撮影範囲を算出することが可能であり、撮影された写真に対してアフィン変換を適応することで、鳥瞰図に変化することが可能となる。得られた図から混雑度を判断し、地図上に混雑度情報をオーバーレイすることで混雑度状況を提示する。

以上の手法から実時間で得られる観光スポットや移動経路などの混雑度状況に応じて、観光プランを最適化し、動的に快適に観光可能なスポットを推薦することで、観光客の満足度向上をめざす。

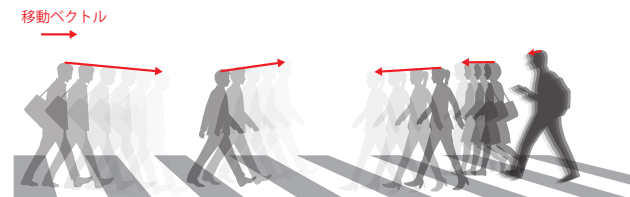


図 2 混雑度推定手法のイメージ図

3.2 混雑度推定手法

本研究で提案する近接群衆画像を用いた混雑度推定手法のイメージ図を図2に示す。本手法は、カメラのシャッター速度を調整することで撮影される残像から移動ベクトル算出し、移動ベクトルの方向、大きさ、数から混雑度を推定する。一般に、移動量が多いほど残像が長く薄く残り、移動量が小さいほど短く濃く残る。この特性を用いることで、撮影画像に写っている残像のピクセル数から移動ベクトルの大きさを算出する。

4. 事前調査

4.1 残像画像を用いた移動量推定

4.1.1 実験内容

本実験は、提案手法の実現に先立ち、撮影した残像からどの程度の精度で移動ベクトルが算出可能かを検証することを目的に行った。今回は、基礎調査として横方向に移動する1名の被写体を対象に行う。すなわち、ベクトルの方向は横方向のみで固定とした時の、大きさ（被写体の移動量）を算出する。実験環境を図3に示す。スマートフォンには、Apple社製のiPhone6Sを用い、被写体距離 l だけ離れた位置に三脚で固定し、カメラの設定はシャッター速度を0.25秒、ISO感度を4とした。本実験では、歩行区間 s を5mとし、被験者には、時速2mから時速6mの間の速度で歩行してもらい、カメラの前を通過する時に撮影した。この時、歩行時間を被験者自身で測定してもらった。被写体距離 l は3mと5mとして、それぞれ9回ずつ歩行した。

実験で撮影された画像の残像から被験者の移動ベクトルの大きさを算出し、歩行時間から算出される正解移動量と比較を行う。

4.1.2 実験結果

実験により得られた残像画像の一例を図4に示す。まずは、得られた残像画像から被写体が移動した距離を算出する。初めに、残像画像とは別に、背景のみの写真も撮影し、各残像画像に対して背景差分を施すことで被験者の残像のみを抽出する。背景差分法を適用し得られた残像画像を図5に示す。図5中、残像画像における頭部の最大画素数 x_p を静止画像時の頭部画素数 B_p で引くことで、実際に移動した分の画素数 Δh_p が算出される。

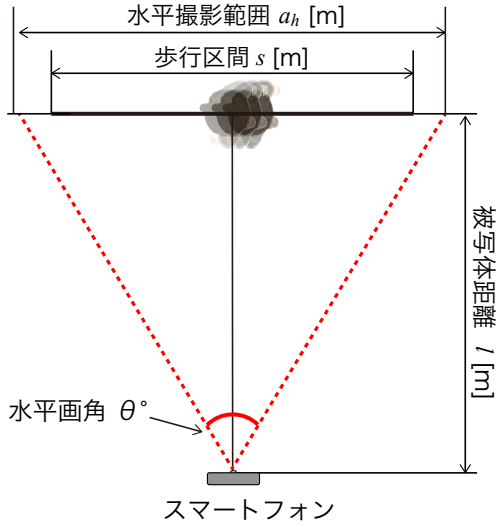


図 3 残像検出実験における実験環境

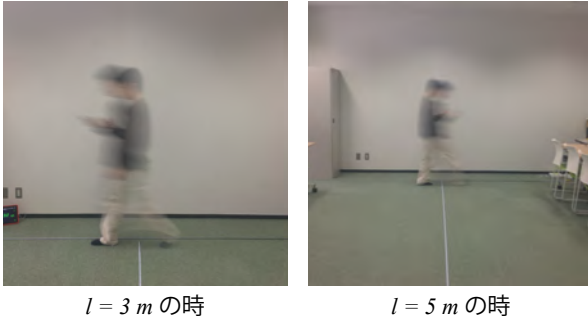


図 4 撮影された残像画像の一例

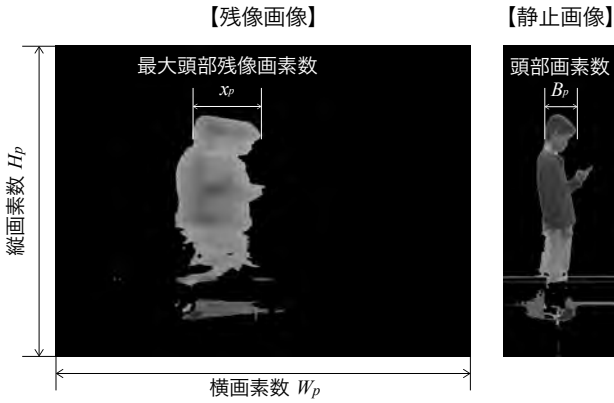


図 5 背景差分法により得られた残像画像

$$\Delta h_p = x_p - B_p \quad (1)$$

次に、画像内での移動距離と実空間での移動距離を対応付けるために、図 3 における水平撮影範囲 a_h を算出する。被写体距離 l における水平撮影範囲 α_h は、式 (2) で示すように水平画角 θ および被写体距離 l から算出できる。また、水平画角は、式 (3) で示すように、水平イメージセンササイズ α および、焦点距離 f から算出することが可能で

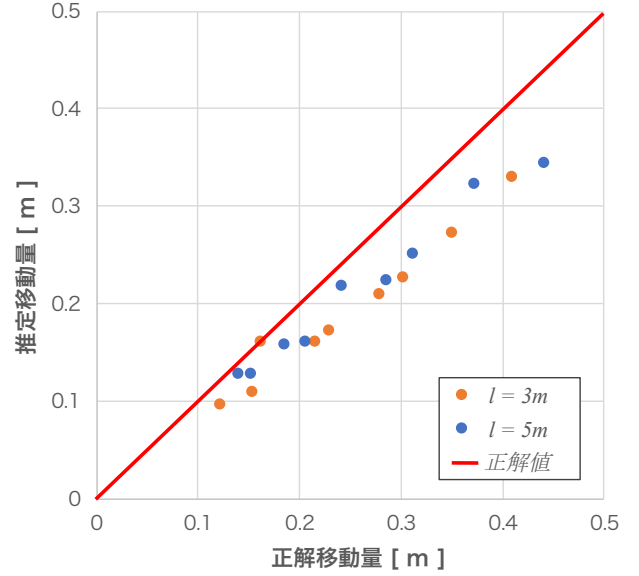


図 6 推定移動量と正解移動量の比較

ある。イメージセンサは、端末により異なる場合があるが、iPhone6S では、 4.8×3.6 mm の物が用いられており、ここでは、 $\alpha = 4.8$ とする。また、焦点距離は AndroidOS、iOS においても、API 経由で取得することが可能となっており、ここでは、 $f = 4.15$ であった。

$$a_h = 2 \cdot \tan\left(\frac{\theta}{2}\right) \cdot l \quad (2)$$

$$\theta = 2 \cdot \tan^{-1} \frac{\alpha}{2 \cdot f} \quad (3)$$

最終的に、式 1 から得られる被写体の移動画素数 Δh_p 、式 2 から得られる水平撮影範囲 α_h および画像の横画素数 W_p を用いて、式 4 を適用することで残像画像から算出される被写体の移動距離 Δh_m が得られる。

$$\Delta h_m = \frac{\Delta h_p}{W_p} \cdot \alpha_h \quad (4)$$

次に、正解値の求め方について述べる。正解移動量の算出には、実験中に計測してもらった歩行時間を元に求める。被験者は、等速で歩行していると仮定し、歩行時に測定した歩行時間 [s] を歩行区間 [m] で割ることで歩行速度 [m/s] が算出される。本実験ではカメラの、シャッター速度を 0.25 秒としたため、0.25 秒間の移動量を算出し、正解移動量とする。

これらの手順を適用し、被写体距離 l が 3 m の時、5 m の時それぞれの推定移動量および正解移動量を算出した。その結果から、縦軸に推定移動量、横軸に正解移動量としてプロットしたグラフを図 6 に示す。結果として、被写体距離が 3 m の時に平均 27.5 %、5 m の時に平均 18.9 % の誤差率で残像から被写体の実空間での移動量を推定できることが分かった。

4.2 参加型センシングによる近接群衆画像収集

4.2.1 実験内容

本実験は、観光中無作為に撮影される写真に、近接群衆群衆がどの程度含まれるかを検証するために行った。参加者は、男女14名とし、2016年11月22日と29日の2日間行った。各実験参加者には、それぞれの日に1人で2時間半を目安に写真を撮りながら観光してもらった。実験場所は京都府東山区とし、開始地点を祇園四条駅、目的地点を清水寺とした。本実験では、3.1節で述べたスマートフォンアプリケーション（以下、アプリ）を用いた。実験で用いたアプリの画面を図8に示す。実験1では、アプリ上で観光スポットをいくつか推薦した上で、各観光スポットへのチェックインとして写真を撮影するタスクを依頼した。ただし、チェックインの際、何を撮影するかは自由とした。実験2では、自由に観光を行ってもらい、任意のタイミングで写真を撮影してもらった。以下、実験1、実験2の詳細手順を記述する。

実験1：観光スポット推薦あり

- (1) 開始地点で被験者にアプリを入れた端末を渡す。
- (2) アプリから最寄りの観光スポットを3つ推薦される。
(図7中①)
- 推薦された観光スポット以外へ行くことや、任意のタイミングで写真を撮影することは自由とした。
- (3) 推薦されたスポットの中から1つ選択して行き、チェックインとして写真を撮影し、一言コメントをつけて投稿する。(図7中②)
- (4) チェックインすると、次の観光スポットが新たに3つ推薦される。
- (5) 3, 4を繰り返し、目的地点を目指す。(図7中③)

実験2：自由観光

- (1) 開始地点で被験者にアプリを入れた端末を渡す。
- (2) 普段観光を行う際に計画を行う人は、事前に観光計画を作成してもらい計画に沿って行動する。
計画を行わない人はそれぞれ自由に観光する。
- (3) 被験者は興味を持った地点など任意のタイミングで写真を撮影しながら、目的地点へ行く。

4.2.2 実験結果

それぞれの実験により収集された写真の枚数および近接群衆群衆の撮影枚数を表1に示す。

推薦された観光スポットでのチェックインで写真撮影を依頼した実験1では、総撮影枚数が386枚となり、そのうち近接群衆画像の枚数は180枚含まれていた。また、任意のタイミングでのみ写真撮影を行ってもらった実験2では、総撮影枚数が209枚となり、近接群衆群衆がそのうち93枚含まれていた。やはり、チェックイン時に写真撮影が必要な実験1の方が、写真の総数は多いが、近接群衆画像が含まれる割合を見てみると、45%前後とほぼ同様の結果となった。



図7 推薦ありにおける被験者の行動遷移イメージ

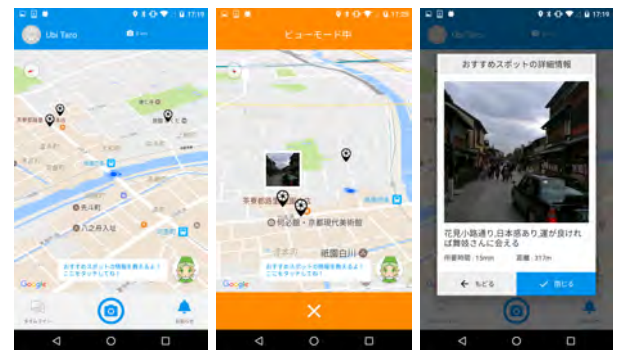


図8 実験に使用したアプリの画面

表1 収集された写真枚数

実験	撮影枚数	近接群衆枚数	割合
実験1	386	180	46.6 %
実験2	209	93	44.5 %
合計	595	273	45.9 %

具体的に、観光実験で得られた近接群衆写真の例を図9に示す。(a)や(d)は観光の開始についての投稿や、目的地に到着したことに関する投稿に、付随して、人が写っているパターンである。一方で、図9中、(b)や(c)に関しては、混雑状況に関するコメントも添えたうえで、画像が投稿されており、同様に混雑度に関するコメント付きの画像が、全体で18枚あったことがわかった。

この結果から、観光中無作為に撮影される写真に近接群衆画像を収集可能であることが分かった。しかし、無作為に撮影された写真だけでは、リアルタイムで網羅的に情報を収集するには、不足している情報の部分を参加者に依頼して情報を収集する必要があると考えられる。

5. おわりに

本研究では、スマートフォンで撮影される近接群衆画像から混雑度を推定する手法を検討した。提案手法では、カメラのシャッター速度を調整することで得られる移動体の残像からそれぞれの移動ベクトルを算出し、ベクトルの方



図 9 観光実験で得られた写真の一例

向, 大きさ, 数から混雑度を推定する. 本稿では, 提案手法を実現する前段階として, 1名の被写体が横方向にのみ移動すると仮定した状態で撮影された近接画像から, 23.2%の誤差で移動距離を推定できることがわかった.

また, 本手法を観光地において観光客を主な参加者とした参加型センシングに適応することを考えている. そこで, 観光者が無作為に撮影する写真の中から, 近接群衆画像がどの程度得られるかを明らかにするために観光実験を行った. 被験者は男女14名とし, 独自に開発したスマートフォンアプリで写真を撮りながら1人で2時間半程度京都府東山地区で観光してもらった. その結果, 合計595枚集まり, そのうち45.9%の割合で近接群衆画像が含まれていることがわかった.

今後の課題として, 移動ベクトルの大きさだけでなく方向を検出や奥行方向への移動ベクトルの検出に向けて取り組む予定である.

謝辞 本研究は, 国立研究開発法人情報通信研究機構(NICT)の委託研究「京都インバウンド観光に向けたIoT-to-Human システムの研究開発」及びJSPS 科研費 16H01721の助成によって行った.

参考文献

- [1] 国土交通省観光庁, 明日の日本を支える観光ビジョン概要. <http://www.mlit.go.jp/common/001126601.pdf>.
- [2] Kwan Hui Lim, Jeffrey Chan, Christopher Leckie, and

- Shanika Karunasekera. Personalized tour recommendation based on user interests and points of interest visit durations. In *Proceedings of the 24th International Conference on Artificial Intelligence, IJCAI'15*, pp. 1778–1784. AAAI Press, 2015.
- [3] 日高真人, 松田裕貴, 河中祥吾, 中村優吾, 藤本まなと, 荒川豊, 安本慶一. 実時間観光コンテンツ提供に向けた観光情報収集・キュレーションシステムの提案. 研究報告高度交通システムとスマートコミュニティ (ITS), Vol. 2017-ITS-68, No. 15, pp. 1–8, 2017.
- [4] YahooJapan, Yahoo!地図 混雑レーダー. <https://map.yahoo.co.jp/maps?layer=crowd&v=3>.
- [5] ZenrinDataCom, いつも NAVI ラボ 混雑度マップ. <http://lab.its-mo.com/densitymap/>.
- [6] Jie Xu, Getian Ye, Gunawan Herman, and Bang Zhang. An efficient approach to detecting pedestrians in video. In *Proceedings of the 16th ACM International Conference on Multimedia, MM '08*, pp. 789–792, New York, NY, USA, 2008. ACM.
- [7] 服部博憲, 満上育久, 棕木雅之, 美濃導彦. 固定カメラ映像を対象とした hog 人物検出器のシーン適応手法 (テーマ関連, 顔・人物・ジェスチャ・行動). 電子情報通信学会技術研究報告. PRMU, パターン認識・メディア理解, Vol. 109, No. 470, pp. 163–168, 2010.
- [8] Huiyuan Fu and Huadong Ma. Real-time crowd detection based on gradient magnitude entropy model. In *Proceedings of the 22nd ACM International Conference on Multimedia, MM '14*, pp. 885–888, New York, NY, USA, 2014. ACM.
- [9] NEC 技報, Vol.67 No.1 群衆行動解析技術を用いた混雑推定システム. <http://lab.its-mo.com/densitymap/>.
- [10] T. Takafuji, K. Fujita, T. Higuchi, A. Hiromori, H. Yamaguchi, and T. Higashino. Indoor localization utilizing tracking scanners and motion sensors. In *2014 IEEE 11th Intl Conf on Ubiquitous Intelligence and Computing and 2014 IEEE 11th Intl Conf on Autonomic and Trusted Computing and 2014 IEEE 14th Intl Conf on Scalable Computing and Communications and Its Associated Workshops*, pp. 112–119, Dec 2014.
- [11] 原佐輔, 小島颯平, MoustafaMahmoud Elhamshary, 内山彰, 梅津高朗, 東野輝夫. 車載カメラを用いた cnn による方向別歩行者頭部検出法の提案. 研究報告モバイルコンピューティングとパーベイスブシステム (MBL), Vol. 2016-MBL-81, No. 24, pp. 1–8, 2016.
- [12] 小島颯平, 内山彰, 廣森聡仁, 山口弘純, 東野輝夫. 俯瞰画像における動体領域面積に基づく群衆人数推定法の提案. 情報処理学会論文誌, Vol. 58, No. 1, pp. 33–42, jan 2017.
- [13] Yutaka Arakawa and Yuki Matsuda. Gamification mechanism for enhancing a participatory urban sensing: Survey and practical results. *Journal of Information Processing*, Vol. 24, No. 1, pp. 31–38, 2016.
- [14] 馬場賢二, 榎原孝明, 湯浅裕一郎. 画像処理による人流計測システム. 東芝レビュー, Vol. 61, No. 12, pp. 35–38, dec 2006.
- [15] 米村淳, 大岸智彦, 井戸上彰, 小花貞夫. スマートフォンを用いた人の混雑度推定手法の提案と評価. 研究報告モバイルコンピューティングとユビキタス通信 (MBL), Vol. 2013, No. 5, pp. 1–8, 2013.
- [16] 西村友洋, 樋口雄大, 山口弘純, 東野輝夫. スマートフォンを活用した屋内環境における混雑センシング. 情報処理学会論文誌, Vol. 55, No. 12, pp. 2511–2523, 2014.