

画像情報を用いた人の追跡とその応用例 第二報 ～ みんなで考える人流解析グランドチャレンジ ～

大西 正輝[†] 依田 育士[‡]

[†] 産業技術総合研究所 〒305-8560 つくば市梅園 1-1-1 中央第 1

E-mail: [†] onishi@ni.aist.go.jp

あらまし 筆者らはこれまでカメラによって人流計測を行うことによって、新しいサービスを生み出すための研究を行ってきた。横軸が人数、縦軸が計測している期間のダイアグラムを作成し、これまでの研究をマッピングすることで人流解析研究のグランドチャレンジを考える一助としたい。

キーワード 人流解析, グランドチャレンジ, 人数-期間ダイアグラム

Human Tracking and its Application using Image Recognition Part 2 — Think about Grand Challenge Together for Pedestrian Flow Analysis —

Masaki ONISHI[†] and Ikushi YODA[‡]

[†] National Institute of Advanced Industrial Science and Technology (AIST)

1-1-1 Umezono, Tsukuba, 305-8560 Japan

E-mail: [†] onishi@ni.aist.go.jp

Abstract Researches aiming at creating new services have been carried out by measuring pedestrian flow by (RGB-D) cameras. A diagram of the number of people on the horizontal axis and the measurement period on the vertical axis was created. We mapped our previous work on this diagram. This results will help to think about the ground challenges of pedestrian flow analysis research.

Keywords Pedestrian Flow Analysis, Ground Challenge, People - Period Diagram

1. はじめに

人は情報の 9 割程度を視覚から得ていることから、画像認識の研究は重要とされ、盛んに研究されている。一般に画像認識の研究は人が行うような、あるいは人が行うことができないような視覚情報処理や分析をコンピュータで自動化し、アプリケーションを作ることによって新たなサービスや価値を生み出すことを目的としている。中でも人を対象としたサービスを構築するためには人を抽出・追跡する研究が重要である。これまでに行われている人の追跡に関する研究が想定しているアプリケーション例としては、講義や会議映像の解析、スポーツ映像の解析、コミュニケーション分析、マーケティング分析、異常・徘徊検出などが挙げられる。

筆者はこれまでに通常のカメラや RGB-D カメラを用いて人を抽出・追跡し、その結果を利用して新しいサービスを生み出すことを目的とした研究を大学生時代に 5 年間、および産業技術総合研究所で 10 年以上にわたって継続している。

本稿では、これまでの研究を振り返ると同時に他の関連した研究をサーベイしながら、人を抽出してその結果を利用する研究を 2 次元のダイアグラムにマッピ

ングすることによって、キラーアプリケーションを考え、今後 10 年間の人の追跡研究のグランドチャレンジを考える手助けとしたい。

2. 人に関する画像認識研究

人に関する画像認識の研究は人の位置追跡だけではなく顔や視線、行動などを認識の対象としている。ここでは筆者が人の位置のみにこだわって研究を続けている理由を簡単に説明する。

2.1. 以前のできごと

10 年ほど前のこと、筆者はあるフィールドにて人の位置情報を抽出したいという依頼を受けた。そこで、実験を行い、その結果を提示したところ、その結果を見た依頼者から人の位置だけでは不十分で体や顔の向きを抽出したいという追加の依頼を受けた。そこで、その結果を提示したところ、今度は結果を見た依頼者から顔の向きだけでは不十分で視線の方向を抽出したいというさらに追加の依頼を受けた。そこで、その結果を提示したところ、結果を見た依頼者から視線の方向だけでは不十分でその人が本当に視線の先のものを

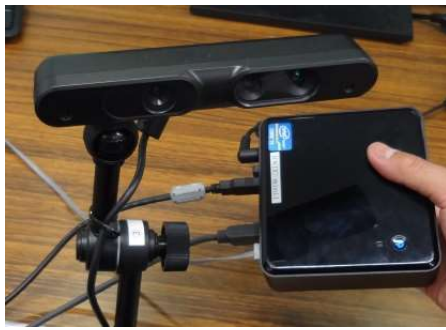


図 1 : RGB-D カメラと小型 PC

見ているかを知りたいという依頼を受けた…

このように人に関する認識は、人の位置だけではなく、顔や体の方向、視線、さらには行動やその意図など多岐にわたる。例えば行動の認識や意図の認識はそもそも行動や意図を正解としてどのように記述するのかさえも分かっていないため（本稿からは逸脱するため多くは説明しないが **walking, running, fight** などのラベルのつけられた行動認識データセットを見るととても残念に思う）グランドチャレンジの最たるものと言えるが、筆者はむしろ人の位置情報だけに徹底的にこだわることでアプリケーションとして何ができるのか、また何ができないのか、さらにできることを増やすためには位置以外にどんな情報が必要になるのかを明らかにする必要があると考え、今の研究に至っている。

2.2. 人の位置の追跡に関する研究

筆者らはこれまでに図 1 のような RGB-D カメラを用いて文献[1]の手法で人を抽出するシステムを作成し、様々な実証フィールドで実験を積み重ねている。そのフィールドやデータ数は多岐にわたり、これまでの実証実験では延べ人数で数千万人規模の人の流れに関するデータを保有している。

代表的な研究としては、図 2 に示すよう救急医療の現場である初療室（ER, Emergency Room）において医師や看護師の移動軌跡を計測することでチーム医療を評価することを目指した研究、2008 年から継続しているレストラン街やオフィスフロアを持つ複合施設において人の流れを計測することでマーケティング等に利用する研究、2012 年から継続している 100 万人規模の観客数があるとされる花火大会の駅周辺の人の流れを計測・シミュレーションすることで混雑緩和施策を評価する研究、大規模な劇場において 1300 人規模の避難訓練を行い 41 台の RGB-D カメラで避難動線を計測することで避難誘導支援を行う研究などがあげられる。

3. 人数－期間ダイアグラム

これらの研究をカテゴライズし、共通性や差別化を



図 2 : 様々な環境における人流計測結果

見出すためには計測対象となる人数と計測期間の関係に分けて分類・整理する。

これまでに行われてきた人流計測を利用したアプリケーションを整理する際に、横軸に人数の対数（1 人，10 人，100 人，…，100 億人），縦軸に年月のおおよそ対数（1 分，10 分，1 時間，1 日，1 週間，1 ヶ月，1 年，10 年，100 年）のダイアグラムを作成し、そこに対象となる研究をプロットすると理解しやすい。本稿ではこの図のことを人数－期間ダイアグラムと呼ぶことにする。

例えばウルトラマンの活動は 1 人で 3 分である（怪獣を入れると 2 人かも知れない）。地球の全ての人の一生を計測対象としてサービスを創出する場合は 71 億人を平均で 71 年計測すればよい。日本の全ての人を対象とする場合は、1.3 億人の 85 年である。つまり、世界中の人を観測対象とした全ての研究はこのグラフのどこかにプロットすることが可能である。

スポーツを例にすると、100 メートル走の決勝は 10 人の 9 秒である。また、卓球は 2 人のおよそ 40 分、サッカーは 22 人の 90 分である。長い時間の競技としてはクリケットが挙げられ 22 人の 5 日間であり、多人数が参加する競技としては東京マラソンが 3 万 6 千人の 7 時間である。一方でスポーツ観戦の観点で考えると最も収容人数の多い横浜国際総合競技場は 7 万 2 千人、東京ドームは 5 万 5 千人である。

一方で国家規模の災害に目を向けると東日本大震災の被災者は震災から 5 年が経った今でも 20 万人にのぼるといわれている。また国家規模の戦争の戦没者をプロットすると日清戦争が 8 ヶ月で 4.8 万人、日露戦争が 1.5 年で 16 万人、第 1 次世界大戦が 4.3 年で 1 千万人、第二次世界大戦が 6 年で 7 千万人である。

図 3 にこれらをマッピングした例を示す

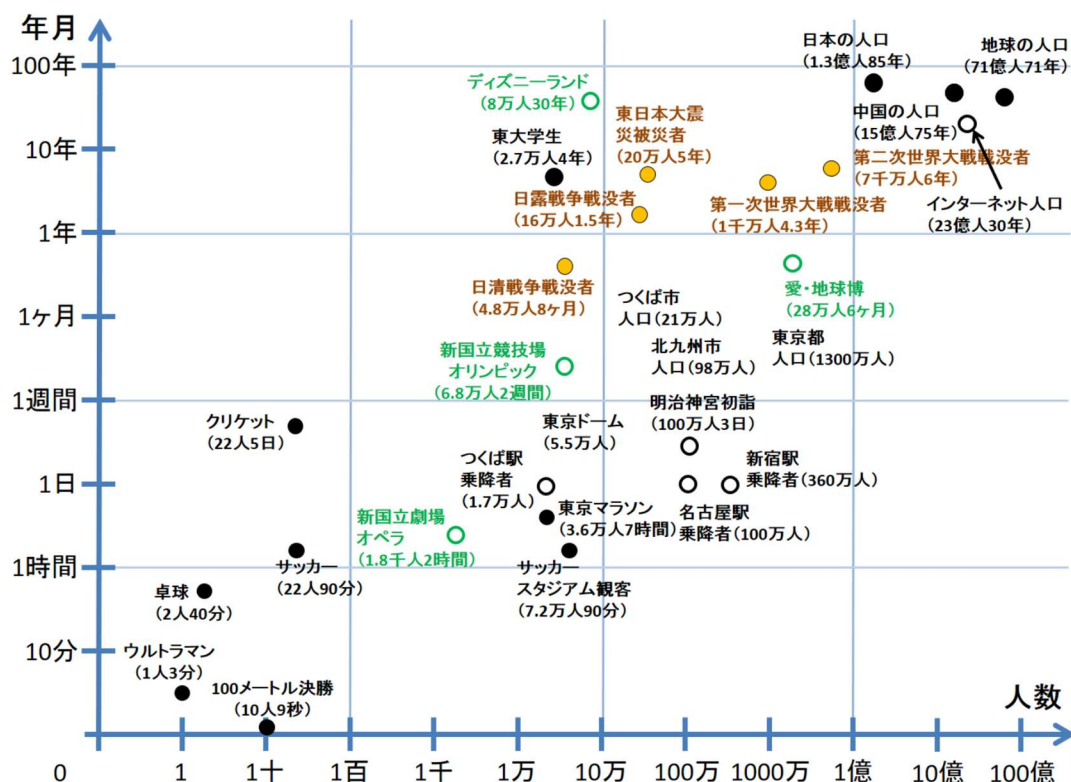


図 3：人数-期間ダイアグラム（一般的な事例）

4. ダイアグラムへの研究のマッピング

次にこれらのダイアグラムの中に図 4 のように研究をマッピングしていく。

4.1. 著者らのこれまでの研究

まず初めに、筆者らのこれまでの研究をマッピングしていく。学生時代に行っていた講義を対象とした研究[2,3]は 1 人の 60 分、会議を対象とした研究[4,5]は 4 人の 30 分である。また、独居高齢者の体調の異常検出を目的とした研究[6]は 1 人の 6 ヶ月である。本研究では電子情報通信学会の ME とバイオサイバネティクス研究会が主催した「高齢者在宅データ解析アルゴリズムコンテスト」で公開されたデータセットを利用しており、6 ヶ月の間に発生した 2 度の体調不良の日を高齢者の部屋間の移動データから推定することを目的としていた。このような異常検出や体調管理のような研究には通常は数ヶ月程度のデータ量が必要であろう。ちなみにドクター中松がイグノーベル賞を受賞したのは 35 年間にわたって自分の食べた料理写真を撮影し続け、食べたものと体調の関係を分析したことに対する功績である。

また、産総研の研究としては ER 内でのチーム医療の評価に関する研究[7]は 8 人の 60 分であり、商業施設におけるマーケティングに関する研究[8]は 3000 人の 1.5 年である。本研究は最終的におよそ 9 年間の人

流データを計測し続け、数年単位の分析を行うことで、鳥インフルエンザの影響や、近隣で起こった事故の影響、さらには東日本大震災が人の流れに与えた影響[9]などの可視化に成功している。さらには、大規模施設からの避難訓練に関する研究[10,11]は北九州芸術劇場では 570 人の 10 分[12]、東京の新国立劇場では 1300 人の 20 分[13]である。大規模施設における避難を扱う場合はマーケティングなどに比べると人数に対して計測時間は非常に短い。

また、花火大会のような大規模な誘導による混雑緩和に関する研究[14,15]は 9 万人の 6 時間である。オリンピックが行われる予定の新国立劇場の定員は 6.8 万人であり、混雑度の観点から比較的類似した対象であることが分かる。

4.2. その他の研究

Fukuhara らは居酒屋の接客係の動きを計測し、QC 活動によって現場に介入することで従業員の動きがどのように変化するかの、どのように上達していくかを定量的に評価している[16]。本研究では 15 人の従業員を介入前と介入後の 3 日ずつ(合計 6 日)計測している。

また、文献[17]では人数は不明であるが、大阪のショッピングセンターで取得した 84 日分の人の流れのデータを用いて平日や休日、身長や密度で人の歩く速度や場所がどのように変化するかを検証している。

一方、公開されているデータセットの一つ

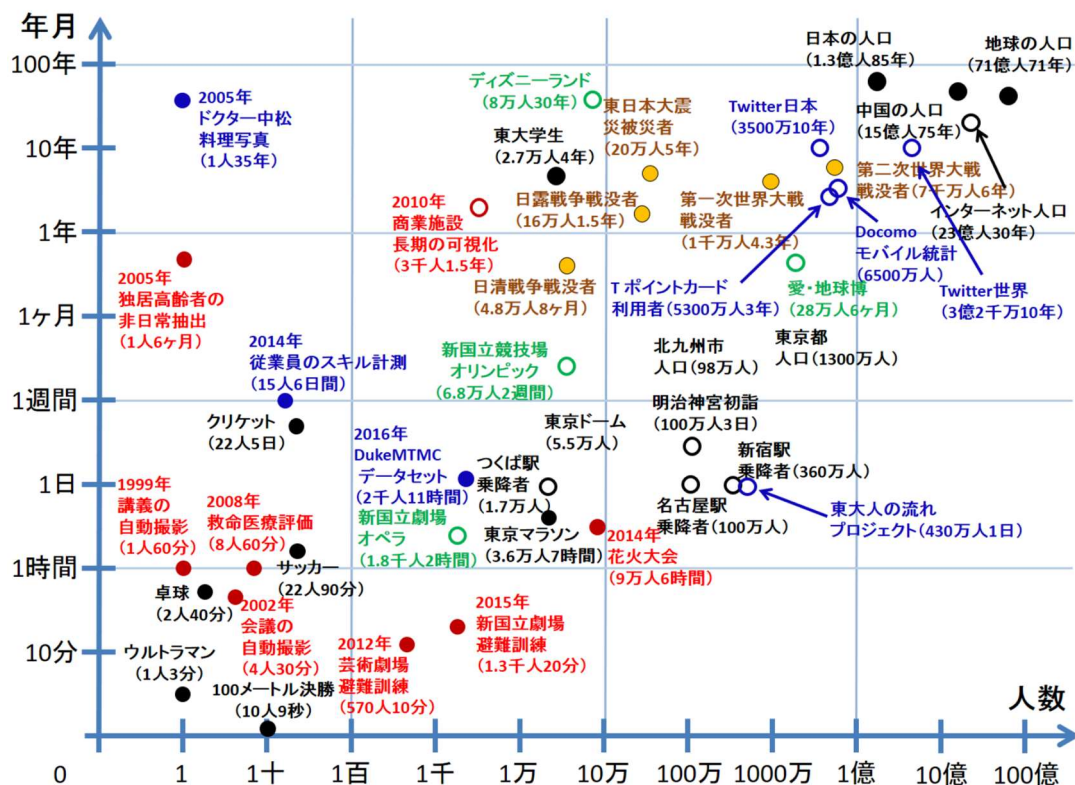


図 4：人数-期間ダイアグラム（研究の重量）

DukeMTMC Data Set では 8 つのカメラから得られる 85 分の映像（11 時間）に 2000 人の人流が収められている[18]. さらに文献[19]では期間に関する言及はないものの駅での 4200 万人の人の流れのデータベースが紹介されている.

さらには実データではないものの、アンケートベースのデータとしては東京大学の人の流れプロジェクトでは数百万人規模のトリップデータを取得している.

4.3. 仮想空間を用いた研究

一方、カメラなどのデバイスを用いて実空間のデータを取得するのに比べて、仮想空間のデータはインターネットを利用してデータを取得することが可能であり比較的データの収集が容易である. 例えば twitter や facebook は日本に数千万人、世界に数億人規模のアクティブユーザがいてと推測されており、それらの人のデータが毎日収集されている.

また仮想空間のデータではないが既にインフラが整備されているものに関しては大規模な人数のデータが収集され、処理され始めている. 例えば T ポイントカードや携帯電話のモバイル統計、IC カードの使用履歴は既にインフラが整備されており、数千万人規模のデータが取得され、研究レベルでは分析されている.

図 4 にこれらの研究をマッピングした例を示す.

5. 考察

本発表の基礎となるアイディアは 2016 年に浜松で行われた画像の認識・理解シンポジウム (MIRU) においてデモ発表を行い、図 4 をポスターに印刷し、自由に書き込めるようにして議論を行った. その時に出てきたアイディアをいくつか紹介する.

5.1. どちらかの軸にとがった研究

横軸（人数）と縦軸（年月）のどちらかにとがった研究を行うことが新しい研究につながるという意見が複数聞かれた.

横軸（人数）にとがった研究例としては犯罪捜査の絞り込みのように例えば数十万～数百万人の中から犯人の 1 人を特定するという問題があがった. 羽田空港と成田空港を合わせた 1 日の平均利用人数は 30 万人程度であり空港利用者からリアルタイムでテロリストを捜索するような問題がこれに相当する.

一方で縦軸（年月）にとがった研究としては顔やファッションが経年によってどのように変化していくかを調べるといった研究があがった. 特に長期間のデータを必要とするのは健康や医療の分野である. 洗面台の鏡に取り付けられたカメラを用いて長期間の顔の変化を計測し続けることで顔色の些細な変化から例えば生活習慣病や体の異変の兆候を抽出する研究などがあげられたが、現実問題として長期間のデータが必要と

筆者の人を対象とした画像認識に関する研究は1999年頃に1人を対象とした60分の講義映像を自動で撮影する研究に始まった。10年から15年を経て人数は10万人、期間は1.5年を対象とした研究へと変化している。この対数速度で研究や技術が進歩すると仮定すると理論上は10年後には地球上の全ての人を平均寿命の100年程度計測できることとなる。そうだとすると世界中の人の移動を通して地球レベルの現象が観測できるようになるはずである。事実、商業施設で9年程度人の流れを計測した実験では鳥インフルエンザやSARSの流行によって外出を控えるようすなどが計測されている。もし、世界レベルでの人の動きが計測できれば、このような感染症の動きなど地球規模の問題の解明にも期待できるかもしれない。

6. まとめ

画像情報を用いた人の追跡とその応用例を考え、これらの関係性を可視化するための人数-期間ダイアグラムを提案し、これまでの研究のマッピングを行った。本発表ではこれらのダイアグラムを見ながら人流解析のキラアプリケーションを議論することでグランドチャレンジについて考えた。

ニュートンはりんごが落ちるのを見て万有引力を思いついたという逸話がある。ニュートンがリンゴを見て運動方程式 $F=ma$ を思いついたように、人を計測する研究は人の移動や行動を計測しながら、それに代わる人の移動の方程式を見つけることが究極のゴールであると思っている。

グランドチャレンジではそのために少しでも大規模に少しでも長期間の計測実験を行い競争力のあるデータを撮り続けていく必要があるだろう。

文 献

- [1] 大西正輝, 依田育士, “ファジィクラスタリングを用いたステレオ画像からの動線抽出,” 電気学会論文誌, vol.128, no.9, pp.1438-1446, Sep. 2008.
- [2] 大西正輝, 泉 正夫, 福永邦雄, “情報発生量の分布に基づく遠隔講義撮影の自動化,” 電子情報通信学会論文誌 (D-II), vol.J82-D-II, no.10, pp.1590-1597, Oct. 1999.
- [3] Masaki Onishi, Masao Izumi, Kunio Fukunaga, “Production of Video Images by Computer Controlled Camera Operation Based on Distribution of Spatiotemporal Mutual Information,” Proc. 15th International Conference on Pattern Recognition (ICPR), vol.4, pp.102-105, Sep. 2000.
- [4] 大西正輝, 影林岳彦, 福永邦雄, “視聴覚情報の統合による会議映像の自動撮影,” 電子情報通信学会論文誌 (D-II), vol.J85-D-II, no.3, pp.537-542, March 2002.
- [5] Masaki Onishi, Takehiko Kagebayashi, Kunio Fukunaga, “Production of Video Images by Computer Controlled Cameras and Its Application to TV Conference System,” Proc. of IEEE Conference on Computer Vision and Pattern Recognition (CVPR), vol.2, II-131-II-137, Dec. 2001.
- [6] 青木茂樹, 大西正輝, 小島篤博, 福永邦雄, “独居高齢者の行動パターンに注目した非日常状態の検出,” 電気学会論文誌, vol.125-E, no.6, pp.259-265, June 2005.
- [7] 依田育士, 大西正輝, 川島理恵, 黒嶋智美, 織田順, 三島史朗, 太田祥一, 行岡哲男, “救命救急医療におけるチーム医療行為解析のための動線と会話の同時解析システム,” 映像情報メディア学会誌, vol.66, no.5, pp.J158-J166, May 2012.
- [8] 大西正輝, 依田育士, “大型複合施設における長期間にわたる人流比較と可視化手法,” 電子情報通信学会論文誌 (D), vol.J93-D, no.4, pp.486-493, April 2010.
- [9] Masaki Onishi, “[Invited Paper] Analysis and Visualization of Large-Scale Pedestrian Flow in Normal and Disaster Situations,” ITE Transactions on Media Technology and Applications, vol. 3, no. 3, pp.170-183, July 2015.
- [10] 山下倫央, 副田俊介, 大西正輝, 依田育士, 野田五十樹, “一次元歩行者モデルを用いた高速避難シミュレータの開発とその応用,” 情報処理学会論文誌, Vol.53 No.7, pp.1732-1744, July 2012.
- [11] Tomohisa Yamashita, Shunsuke Soeda, Masaki Onishi, Itsuki Noda, “Exhaustive Testing of Evacuation Plan with High-Speed Evacuation Simulator,” International Scientific and Technical Conference Emergency Evacuation of People from Buildings, pp.357-364, March 2011.
- [12] 野中陽介, 大西正輝, 山下倫央, 岡田 崇, 島田敬士, 谷口倫一郎, “大規模な避難シミュレーションのための歩行速度モデルの精緻化,” 電気学会論文誌C, vol.133, no.9, pp.1779-1786, Sep 2013.
- [13] 大西正輝, 山下倫央, 星川哲也, 佐藤和人, “人の流れの計測とシミュレーションによる避難誘導の伝承支援 -新国立劇場における避難体験オペラコンサートを例に-, ” 人工知能学会合同研究会, 知識・技術・技能の伝承支援研究会, SIG-KST-026-06, Nov. 2015.
- [14] 山下倫央, 大西正輝, “オリンピックにおける人の流れの解析,” 情報処理, vol.55, no.11, pp.1189-1195, Nov.2014.
- [15] 熊谷章平, 大西正輝, 堀田一弘, “学習画像の適応選択による動画像中の群衆領域の推定,” 電子情報通信学会論文誌, vol.J99-D, no.8, Aug. 2016.
- [16] Tomohiro Fukuhara, Ryuhei Tenmoku, Ryoko Ueoka, Takashi Okuma and Takeshi Kurata, “Estimating skills of waiting staff of a restaurant based on behavior sensing and POS data analysis: A case study in a Japanese cuisine restaurant,” Proceedings of the 5th International Conference on Applied Human Factors and Ergonomics (AHFE), pp.4287-4299, 2014.
- [17] Dražen Bršćić, Francesco Zanlungo, Takayuki Kanda, “Density and Velocity Patterns during One Year of Pedestrian Tracking,” Transportation Research Procedia, vol. 2, pp.77-86, 2014.
- [18] E. Ristani, F. Solera, R. S. Zou, R. Cucchiara and C. Tomasi, “Performance Measures and a Data Set for Multi-Target, Multi-Camera Tracking,” ECCV 2016 Workshop on Benchmarking Multi-Target Tracking.
- [19] Alexandre Alahi, Vignesh Ramanathan, Li Fei-Fei, “Socially-aware Large-scale Crowd Forecasting,” CVPR2014.