

# 歩行中のマイクロタスク処理におけるデータ品質向上手法とその評価

太田 千尋<sup>†</sup> 森嶋 厚行<sup>††</sup> 寺田 努<sup>†††</sup> 中村 聡史<sup>††††</sup> 北川 博之<sup>†††††</sup>

<sup>†</sup> 筑波大学院システム情報工学研究科コンピュータサイエンス専攻 〒305-8573 茨城県つくば市天王台 1-1-1

<sup>††</sup> 筑波大学知的コミュニティ基盤研究センター 〒305-8550 茨城県つくば市春日 1-2

<sup>†††</sup> 神戸大学工学研究科電気電子工学専攻 〒667-8501 兵庫県神戸市灘区六甲台町 1-1

<sup>††††</sup> 明治大学総合数理学部先端メディアサイエンス学科 〒164-8525 東京都中野区中野 4-21-1

<sup>†††††</sup> 筑波大学システム情報系 〒305-8573 茨城県つくば市天王台 1-1-1

E-mail: <sup>†</sup>chihiro.ota.2014b@mmlab.info, <sup>††</sup>mori@slis.tsukuba.ac.jp, <sup>†††</sup>tsutomu@eedept.kobe-u.ac.jp,

<sup>††††</sup>satoshi@snakamura.org, <sup>†††††</sup>kitagawa@cs.tsukuba.ac.jp

あらまし クラウドソーシングの抱える重大な課題の一つとして、タスクを処理するワーカの確保が挙げられる。通常、クラウドソーシングシステムはPCやスマートフォン、タブレット等を通してワーカを確保する。つまり、ワーカにとっては、タスクにアクセスするためには上記のようなデバイスが必要となる。本論文は、日常空間にいる歩行者をワーカとして確保する新しいアプローチを提案する。具体的には、非常に小規模なタスクとして「ナノタスク」を定義し、これを床に投影することで日常空間でのタスク処理を可能にする。我々は実際にこのアプローチのシステムを開発し、その投影画面付近を通る歩行者の挙動を観察した。観察の結果から、本アプローチのデータ品質の問題が浮上した。投影画面上を通過する歩行者のなかにはタスクを処理する意図がない者がおり、意図せずタスクに回答する場面がある。そこで、本論文は、意図のあった回答のみを選び出すために、機械学習を用いたアプローチを提案する。実験により、回答意図があると判定されたタスク結果のみを抽出したほうがデータ品質が向上するケースを明らかにした。以上により、設置個所での状況に応じて動的にデータ品質の向上手法を切り替える事が有効である事をしめす。キーワード クラウドソーシング, データ品質, HCI

## 1. はじめに

クラウドソーシングの抱える重大な課題の一つとして、タスクを処理するワーカの確保が挙げられる。ワーカを確保する手法はタスクの粒度や規模によって異なる。例えば、それぞれのタスクが難しくまた処理に時間がかかる場合には、金銭の報酬をワーカに与えることでワーカの確保が見込める。一方、タスクの規模が小さくまた十分に簡単である場合には、タスクを他の作業の中に埋め込むなど、上記とは異なるワーカ確保戦略を採用することができる [1][7]。既存の戦略の共通点として、パソコンやスマートフォン、タブレット等の端末を通してワーカの確保を行うという点がある。したがって、ワーカを確保する時に、ワーカが上記のような端末を使用している必要があった。

本稿では、ワーカがタスクにアクセスする機会を増やすために、日常空間上に存在する歩行者を、クラウドソーシングによる問題解決に重要なリソースとみなし、その歩行者をワーカとして確保する新しいクラウドソーシングのアプローチを提案する。我々の仕組みでは、マイクロタスクの一つである「ナノタスク」を床に投影し、その歩行者から得られるタスクへの回答を集める(図1)。ナノタスクとは、マイクロタスクの中でも1,2秒程度のごく短時間で処理できる程度の内容のマイクロタスクである。

図3は、我々が開発した「Task-on-the-floorシステム」[2](以下 ToF システム)の構成を示した図である。本システムはクラ



図1 Task-on-the-floor システム

ウドソーシングプラットフォーム Crowd4U [8] からナノタスクを受け取り、それを床に一つずつ投影する。床に投影されたナノタスク画面には、タスクの質問文に加えて、タスクの質問に回答するための回答選択肢も投影する<sup>(注1)</sup>。本システムは歩行者の動作を認識し、歩行者が回答選択肢の上を通過すると、クラウドソーシングプラットフォームへ回答を送信する。

(注1): タスクの内容や画面デザインは限定しない。例えば、ワーカに提示した写真に含まれる特定のオブジェクトを踏ませる、という内容のナノタスクを投影することも可能である

通行人にタスク処理を行わせるという本システムの設計上、データ品質に関するある問題が生じる。投影画面上を通過する通行人のなかにはタスクを処理する意図がない者がおり、意図せずタスクに回答してしまう場合がある。我々が設置したケースでは、床に投影されたタスクに回答する意図のないワーカーは、通行人全体のうちおよそ2~3割であったが、これは大学キャンパスの限定された場所に設置したケースであり、一般にはより多くの通行人に回答意図が無いことが予想できる。しかし、既存のデータ品質向上手法は、タスク結果にはタスクに回答する意図のある者による回答しか含まれないことを前提としている。回答する意図のない者による回答は、ToFシステムで得られるタスク結果のデータ品質を低下させてしまう。

本稿では、通行人によるクラウドソーシングにおけるデータ品質向上の課題を取り扱う。具体的には、タスクに回答する意図のない通行人による回答を取り除くために、機械学習の適用を試みる。さらに、このようなフィルタリングがデータ品質に与える影響を分析する。

本稿の貢献は以下の通りである。

**通行人を対象としたクラウドソーシング.** 我々の知る限り、本稿は、ワーカーを確保する際にデバイスを必要としない、通行人を対象としたクラウドソーシングを扱う初めての論文である。通行人を対象とするクラウドソーシングにおいては、歩行者をマイクロタスク型クラウドソーシングのリソースとみなす。従来のようにデバイスを通してウェブブラウザ上でタスク処理をさせるのではなく、タスクを日常空間内に設置し、付近を通る歩行者にタスク処理を行わせる。

**データ品質の向上に関する検証.** 通行人を対象とするクラウドソーシングにおいては、日常空間にて歩行者にタスクを処理させるという特徴から、タスクに回答する意図のない歩行者による回答が含まれる。本稿では、回答意図のない歩行者による回答を取り除くことで、マイクロタスク結果のデータ品質を向上させることを試みる。また、実験により、提案手法を用いた場合のマイクロタスク結果のデータ品質と、回答意図の有無を考慮しない手法を用いた場合のデータ品質と比較を行った。

**長期間の設置経験からの議論.** 試験的に、ToFシステムを2013年夏から順次、筑波大学、明治大学、神戸大学、同志社大学構内の不特定多数の人が通行する場所に設置し、自然災害により被害を受けた場所を探し出したり、より良い手書きの文字を選んだりするなど、クラウドソーシングプラットフォーム Crowd4U [8] で実際に行われているタスクを配布した。我々の経験から、本稿で取り扱うアプローチは有望であり、また本稿での議論や実験結果は説得力があると考えている。

本論文の構成は以下の通りである。2.節では、関連研究について述べる。3.節では、ToFシステムについて説明する。4.節では、ToFシステムにより得られるタスクへの回答への処理について説明する。5.節では、データ品質向上のために、機械学習を用いて回答意図の有無を判別する提案手法について説明する。6.節では、5.節で説明した手法の評価実験について述べる。最後に、7.節にて、本稿のまとめと今後の課題について述べる。

## 2. 関連研究

本稿で提案する ToF システムと同様に、日常空間にクラウドソーシングを行う環境を構築したシステムや研究が存在する。Community-Sourcing Vending-Machine [5] は、自動販売機に搭載されたタッチスクリーンを用いてクラウドソーシングを行うシステムである。タッチスクリーン上にタスクが表示され、ワーカーはその場で回答する。また [6] は、公共の場所に設置したディスプレイによってクラウドソーシングを行うことを試みる研究である。これらのシステムや研究では、タスクへの回答を送信するワーカーは必ず回答する意図を持つ。一方、本稿で提案する ToF システムでは、回答を送信するワーカーの中にタスクに回答する意図のない者が含まれる点でこれらと異なる。

クラウドソーシングの結果のデータ品質を向上させるための研究には、ワーカーの行動に基づくスパムの検出を行ったものがある [4]。この研究では、ウェブブラウザのクラウドソーシングにおいて、ワーカーがタスク処理に要した時間、クリック回数や、キーボードを押した回数、およびマウスの動き等のワーカーの行動に基づいて、悪意あるワーカーの特定や、得られるタスク結果のデータ品質の予測、および誤った回答の予測を行うモデルの作成が行われた。一方、本稿で提案する ToF システムでは、ウェブブラウザ上ではなく日常空間に組み込まれたクラウドソーシングにおいて、ワーカーの行動に基づいて回答意図のない歩行者を検出し、データ品質の向上を試みる点でこれと異なる。

## 3. Task-on-the-floor システム

### 3.1 ナノタスク

ナノタスクとは、非常に簡単かつ単純な、数秒間で誰もがタスクの指示に理解し従うことができるようなマイクロタスクである。ナノタスクの具体例としては、以下のようなタスクが考えられる。

- (1) 回答に迷わないようなタスク内容であり、ワーカーに「Yes」と「No」の二択で回答させる単純なタスク
- (2) 床に投影した画面上に表示された画像の中から、指定したオブジェクトを足で踏ませるタスク

ナノタスクを床に投影するためには、歩行者が床の投影画面上を通る際の進行方向を考慮してタスク画面をデザインする必要がある。以下から具体例を挙げて説明する。

図2は、ナノタスク画面の例である。ここで、仮に図2の下方から上方へタスク画面上を通過する場合を考える。まず、歩行者が近づいてくる投影画面の下方にタスクの指示文は表示されている。その次に、歩行者から見て奥にタスクに関する画像が表示されており、さらにその奥に回答選択肢が表示されている。

上記のように、ナノタスク画面は、タスク処理に必要な情報が歩行者にとってわかりやすい順序で表示されるように、歩行者にとって手前から指示文、タスク処理に必要な画像、そして回答選択肢という順序でオブジェクトを表示しなくてはならない。



図 2 床に投影するナノタスク画面の例

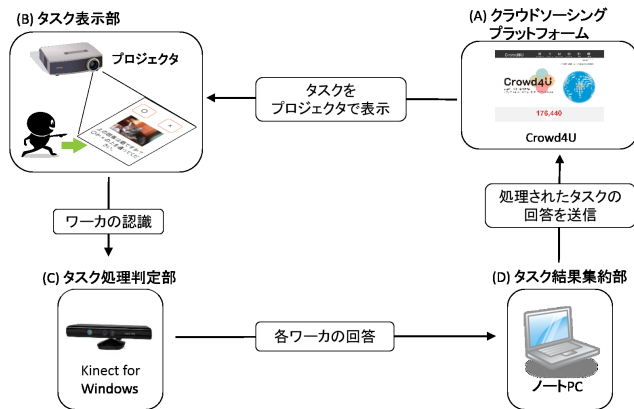


図 3 システムの構成

### 3.2 システムの構成

ToF システムの構成を図 3 に示す。ToF システムは、主に以下の 4 つの要素から構成される。

- (A) クラウドソーシングプラットフォーム (B) タスク表示部へナノタスクを発行する。
- (B) タスク表示部 (A) クラウドソーシングプラットフォームから受け取ったナノタスクを床に投影する。
- (C) タスク処理判定部 歩行者の動作を監視し、それぞれの歩行者が歩行中にどの回答選択上を通過したかを判定する。
- (D) タスク結果集約部 (C) タスク処理判定部から受け取った回答を集約し、処理を行う。

### 3.3 システムの実装

本システムでは、(A) クラウドソーシングプラットフォームとして Crowd4U [8] を用いる。(B) タスク表示部には、プロジェクタを用いてナノタスクを床に投影する。(C) タスク処理判定部には、2 台の Microsoft Kinect をセンサーとして用いる。(D) タスク結果集約部にて利用するアルゴリズムは、4. 節にて説明する。

### 3.4 システムの設置

ToF システムは、廊下のような多くの通行人があるような場所に設置されることを想定している。図 4 のようにして廊下に設置される。ToF システムは、歩行者の接近を感知すると、床に投影したタスク画面を歩行者の進行方向に合わせて回転させる。これにより、歩行者は廊下のどちらの方向から歩いてきた場合でも、ToF システムを用いてマイクロタスクを処理す

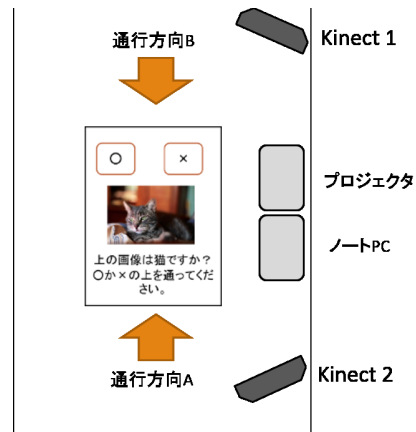


図 4 システムの設置方法

ることができる。図 4 の通行方向 A から来た歩行者は図中の Kinect1 が認識し、通行方向 B から来た歩行者は Kinect2 が認識することで、歩行者の回答を受け付ける。

しかし、常に多くの歩行者がいるような混雑した場所では、ToF システムによってタスクの指示文を読んだり回答することが困難になる。ToF システムは通行量が 1 分に 6 人程度である場所への設置が望ましい。

## 4. ToF システムによるナノタスク結果のデータ品質向上手法

ToF システムは廊下等の不特定多数の通行人がいる場所に設置されるが、床に投影されたナノタスク画面上を通過する歩行者の中には、ナノタスクを処理する意図のない者も含まれる。しかし、このシステムは、歩行者のナノタスクへの回答意図の有無に関わらず、歩行者が回答選択肢の上を通過すると、回答を受け付けてしまう。そのような回答意図のない者による回答はナノタスク結果のデータ品質を低下させるため、何らかの対策が必要となる。そこで、3.2 節で説明した ToF システムの構成要素 (D) タスク結果集約部で、ToF システムで得られたタスクへの回答を集め、データ品質を向上させることを検討している。データ品質の向上手法のうち、本稿で取り扱うものを次に示す。

- (1) 回答意図の有無による回答の選別. ToF システムにより得られる回答それぞれについて、その回答をした歩行者にタスクに回答する意図があったかどうかを判別し、回答意図があったもののみを有効な回答とする。タスクへの回答意図がない者による回答は正答しない確率が高いと考えられるため、これを取り除くことでタスク結果のデータ品質を向上させる。
- (2) 多数決による回答の決定. 多くのクラウドソーシングシステムにおいて、多数決に基づくデータ品質の向上が行われているため、ToF システムにおいても、多数決による回答の決定を検討する。ToF システムにより得られる回答には回答意図のある者となない者が含まれるが、回答意図のある回答は、より正答する確率が高いと考えられるため、少ない人数で、より品質の高い結果が得られると予想される。一方、回答意図がある回答

は、全体と比較して数が減るため、回答意図の有無でフィルタリングを行った方が良いのかどうかは、トレードオフの関係となる。6.章では、これらの関係について実験とシミュレーションを用いて評価する。

本稿では回答意図の有無の判定を機械学習を用いて行うアプローチを提案し、多数決によって回答を決定する手法との比較を行う。

次節から機械学習を用いた回答意図の判別手法について説明する。

## 5. 回答意図の判定

### 5.1 概要

本節では、ToFシステムによるナノタスク結果のデータ品質を向上させるために、機械学習による回答意図の有無を判断する手法について説明する。本稿では、Support Vector Machine (以下、SVM)を用いて、ToFシステムにより回答を送信した歩行者それぞれについて、「質問への回答意図があった歩行者」と「質問への回答意図がなかった歩行者」の二値分類を行う。

### 5.2 回答意図の判定に用いる要素

本システムによって床に投影されたタスク画面上を通過する歩行者のうち、質問に回答する意図のある者となし者との姿勢や動作には違いがあると考えられる。本稿では、これらの違いから歩行者の回答意図の有無を判断することを検討する。今回は、歩行者の回答意図の有無の判断に有効であると考えられる次の4種類の要素を用いる。

- (1) 歩行者の進行方向の変化
- (2) 歩行者の頭の角度
- (3) 歩行者の歩行速度
- (4) 歩行者の歩いた位置

それぞれの根拠は次の通りである。(1)について、回答意図のある歩行者は歩きながら正しいと考える選択肢に向かっていく。したがって、回答意図のある歩行者は、タスク画面上にて進行方向を変えながら歩くと考えられる。(2)について、本システムにおいて、ナノタスクは床に投影される。したがって、タスクの質問に回答しようとする歩行者は多くの場合、床を見るために頭を下げるため、歩行者の頭の角度は手がかりになる。(3)について、歩行者が本システムにより質問に回答する際、床に投影されたタスク問題文を読み選択肢を選ぶ必要がある。そのため、質問に回答する意図のある歩行者はタスクを確認するためにタスク画面上、もしくはタスク画面手前で立ち止まる、あるいは歩行速度が低下する等の行動をする可能性がある。(4)について、タスク画面上の回答選択肢は常に同じ位置に表示されており、回答意図のある歩行者は必ずその上を通過するはずであるため、同様の位置を通行するケースが多いと予測した。

本研究では、(1)から(4)の歩行者の姿勢や動作のデータを、ToFシステムの構成要素であるKinectを用いて取得する。具体的な取得方法については次節から説明する。

### 5.3 特徴量の取得方法

提案手法では、各歩行者について5.2節で説明した要素を表す特徴量を求め、Support Vector Machine (以下SVM)を



図5 特徴量取得のためにマイクロタスク画面上に想定する水平線

用いて分類器を作成し、「質問への回答意図があった歩行者」と「質問への回答意図がなかった歩行者」の二値分類を行う。

具体的な特徴量は、次のように計算する。まず、床に投影したマイクロタスク画面上とその手前に、図5のような10本の水平線があると想定する。また、タスク画面の左上の角を原点にとり、ローカル座標系を作成する。このローカル座標系は、タスク画面の左上の角から右上の角へ向かう方向をX軸方向、左下の角へ向かう方向をY軸方向とする。次に、Kinectを用いて各水平線上に乗った瞬間の次の(1)から(4)の4種類のデータを求める。これらのデータは、それぞれ5.2節の(1)から(4)の要素を特徴量として表したものである。下記の「歩行者の立ち位置」は、Kinectにより臀部の中心と検出された座標とする。

- (1) 前の水平線での立ち位置と比較した、ローカル座標系でのX座標の変化量の絶対値
- (2) 水平線上での歩行者の頭の角度(水平面と、頭の頂点と肩の中心を結んだ直線のなす角度)
- (3) 前の水平線を通過してからの経過時間
- (4) 水平線上での立ち位置の座標

上記のうち、(1)(3)については、一つ手前の水平線での値と比較する必要があるため、一本目の水平線では特徴量をとらない。よって、一回の歩行者によるタスク処理で合計 $10 \times 4 - 2 = 38$ 個の値を取得する。以上のようにして取得した歩行者の姿勢や動作のデータを、(1)から(4)の4種類のデータそれぞれで正規化し、特徴量とする。

## 6. 評価実験

提案手法により作成した分類器が、ToFシステムによるナノタスク結果のデータ品質向上にどの程度効果があるかを評価するために、次の2点の評価実験を行った。

- 評価実験1: 二値分類の性能による評価
- 評価実験2: ナノタスク結果のデータ品質による評価

### 6.1 機械学習および評価に用いるデータの収集

分類器の学習、および学習した分類器の評価を行うために、次のような方法で歩行者のデータを収集した。



表 1 収集したデータの内訳

分類クラス	データ数
正例 (回答意図あり)	176
負例 (回答意図なし)	34

表 2 回答意図の有無の場合それぞれのタスクへの正答率

回答意図がある回答の正答率	0.886364
回答意図がない回答の正答率	0.676471

表 3 実際のクラスと予測クラスの関係

	実際が正例	実際が負例
予測が正例	<i>TP</i>	<i>FP</i>
予測が負例	<i>FN</i>	<i>TN</i>

実際に TOF システムを筑波大学の不特定多数の人が利用する廊下に設置し、床に投影したタスク画面を通過した各歩行者について 5.3 節にて説明した特徴量データを収集した。また、各歩行者のデータについて、歩行者の視線等に注目し、目視で回答意図の有無のラベル付けを行うことで、訓練データを作成した。ただし、ラベル付けが 1 人の主観に依存しないよう、2 人の観察者が目視でのラベル付けを行い、双方の付けたラベルが一致する歩行者のデータのみを記録した。ただし、2 人の観察者のうち一方は、本実験で歩行者の回答意図の判断に用いる 4 つの要素を知らない者が担当した。

収集したデータの内訳を表 1 に示す。本実験においては、「タスクへの回答意図あり」クラスに分類される回答を正例、「タスクへの回答意図なし」クラスに分類される回答を負例として評価を行う。

また、目視によるラベリング結果と各歩行者の回答の正誤から、回答意図のある歩行者による回答と、回答意図のない歩行者による回答について、タスクへの正答率を計算した結果を表 2 に示す。ここでの正答率とは、得られた回答のうち、実際にタスクの正解と等しかった回答が占める割合とする。

## 6.2 評価実験 1：二値分類の性能による評価

### 評価方法

学習した分類器の評価を、5 分割交差検定により行った。評価には、正解率と、正例の適合率と再現率、F 値を用いる。実際の分類クラスと分類器による予測クラスの関係を表 3 に表す。

このとき、分類器の正解率、適合率、再現率、F 値を、それぞれ数式 (1)~(4) によって求める。

$$Accuracy = \frac{TP + TN}{TP + FP + FN + TN} \quad (1)$$

$$Precision = \frac{TP}{TP + FP} \quad (2)$$

$$Recall = \frac{TP}{TP + FN} \quad (3)$$

$$F - measure = \frac{2 \times Precision \times Recall}{Precision + Recall} \quad (4)$$

正解率は、学習した分類器が正しく分類した回答の割合を表す。適合率は、分類器が「タスクへの回答意図あり」と分類した回答のうち、実際に回答意図があった回答の割合を表す。回答意図がある歩行者による回答のみを正しく取り出すことがで

表 4 評価実験 1 結果：分類器の性能評価実験結果

正解率	適合率	再現率	F 値
0.876744	0.881075	0.988889	0.931143

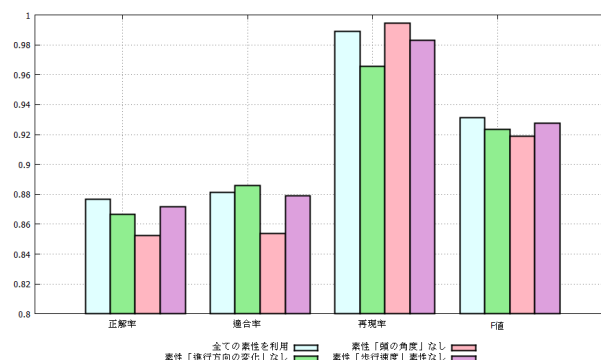


図 6 評価実験 1 結果：5 交差検定による分類器の性能

ければ、本研究の目的であるデータ品質の向上を実現することができると考えられるため、適合率は本研究において分類器の評価に重要である。再現率は、実際に回答意図がある歩行者による回答のうち、分類器が「タスクへの回答意図あり」と分類した回答の割合を表す。

また、考察のために、5.3 章で述べた (1)~(4) の素性をそれぞれ一つずつ除いて学習を行った分類器についても同様に 5 分割交差検定を行う。これにより、(1)~(4) の素性それぞれが、実際に分類器の学習に効果があったか否かを確認する。

### 結果と考察

実験 1 の結果を表 4 に示す。表 4 から、再現率と比較して正解率が低いことがわかる。再現率が高いということは、実際に正例である回答の多くを正しく分類できたことを表す。一方で、正解率が低いということは、正例と負例双方を合わせた分類精度はそれほど良くないことを示す。このことから、正例の分類精度に対して負例の分類精度が低いことがわかる。また、再現率と比較して適合率が低いことから、実際は負例である回答を誤って正例と分類した場合が多いことがわかる。

また、5.3 章で述べた (1)~(4) の素性をそれぞれ一つずつ除いて学習を行った分類器で 5 分割交差検定を行った結果を図 6 に示す。図 6 より次のことがわかった。5.3 章の (1) 「歩行者の進行方向の変化」のみを除いて学習した分類器は、全ての素性で学習した分類器と比較して、適合率は上昇し、一方で再現率は低下している。このことから、「歩行者の進行方向の変化」は、再現率の上昇に貢献する素性であることがわかった。同様に、5.3 章の (2) 「歩行者の頭の角度」のみを除いて学習した分類器を全ての素性で学習した分類器と比較すると、適合率は低下し、再現率は上昇している。このことから、「歩行者の頭の角度」は、適合率の上昇に貢献する素性であることがわかった。5.3 章の (3) 「歩行者の歩行速度」のみを除いて学習した分類器は、全ての素性で学習した分類器と比較して、性能に大差が見られなかった。したがって、「歩行者の歩行速度」は、分類に有効でない素性であることがわかった。最後に、5.3 章の (4) 「歩

行者の歩いた位置」のみを除いて学習した分類器は、全ての素性で学習した分類器と比較して、全ての評価指標において性能が低下している。このことから、「歩行者の歩いた位置」は、正例、負例ともに分類に有効であることがわかった。

### 6.3 評価実験 2：ナノタスク結果のデータ品質による評価評価方法

分類器によるフィルタリングがナノタスク結果のデータ品質に及ぼす影響を評価する。ナノタスク結果のデータ品質は、歩行者のナノタスクへの正答率で評価する。

分類器を用いて回答意図があると分類された回答のみを得る場合と、分類器による回答の分類を行わない場合のタスクへの 1 回答あたりの正答率を比較する。また、それぞれの場合について、同一のナノタスクに対する複数人の回答から多数決によって 1 つの回答を決定する場合の正答率も比較する。ただし、ToF システムによって床に投影されたタスク画面上を通過する歩行者のうち、回答意図のある歩行者の占める割合によって、1 回答あたりの正答率が異なることが予想される。そこで、本実験では、タスク画面上を通過する歩行者のうち回答意図のある歩行者の占める割合が異なる場合でシミュレーションを行う。このとき、回答意図のある歩行者とない歩行者それぞれの回答の正答率をもとに歩行者一人あたりの正答率と、一人あたりの正答率をもとにシミュレーションした多数決を行う場合の 1 回答当たりの正答率をシミュレーションし、分類器の有無で正答率にどのような影響があるかを評価する。また、回答意図のある歩行者の占める割合については、9 割、5 割、1 割でシミュレーションを行う。

一人あたりの正答率は、回答意図のある歩行者の占める割合が  $m$  であるとき、回答意図のある回答の正答率を  $P1$ 、回答意図のない回答の正答率を  $P2$  とすると、分類器を用いる場合と用いない場合それぞれを、次の数式 (5)、(6) によって計算する。*Precision* には、表 4 より、分類器の適合率を用いる。

$$\begin{aligned} & \text{分類器を用いる場合の一人あたりの正答率} \\ & = P1 \times \textit{Precision} + P2 \times (1 - \textit{Precision}) \end{aligned} \quad (5)$$

$$\begin{aligned} & \text{分類器を用いない場合の一人あたりの正答率} \\ & = P1 \times m + P2 \times (1 - m) \end{aligned} \quad (6)$$

この値をもとに、多数決を行った場合の 1 回答あたりの正答率をシミュレーションする。

また、回答意図がない歩行者の回答のタスクへの正答率は、表 2 の値だけでなく、0.5 である場合でもシミュレーションを行う。回答意図がない回答の正答率は 0.5 を予測していたが、表 2 より、実験で収集したデータから得られた正答率はそれを大きく上回っていることがわかった。今回の実験にて回答意図がない歩行者の回答の正答率が予想よりも高い原因として、目視によるラベリングが正確でなかった可能性が考えられるが、サンプル数が少なくまだ判断ができないため、今後も継続して調査する。

#### 結果と考察

評価実験 2 の結果を図 7, 8, 9 に示す。図 7, 8, 9 は、それ

ぞれ、ToF システムにより回答を受け付けた歩行者のうち回答意図があった歩行者が 9 割、5 割、1 割を占める場合のシミュレーション結果を表したグラフである。これらのグラフは、横軸に回答意図のある歩行者とない歩行者を含むタスク画面上を通過した人数を、縦軸に 1 回答あたりのタスクへの正答率をとる。分類器を用いない場合は、タスク画面上を通過した歩行者すべての回答を用いるので、横軸の値の人数で多数決をとった場合のタスクへの正答率をシミュレーションし、プロットした。ただし、多数決を行わない場合の横軸の値は 1 である。分類器を用いる場合は、回答意図がある歩行者の回答のみを取り出し、その回答を使って多数決を行う場合のシミュレーションを行う。横軸の値の人数がタスク画面上を通過した際、その中で回答意図があると分類された歩行者のみで多数決を行った際の正答率をシミュレーションし、プロットした。多数決を行わない場合の横軸の値は、回答意図の有る歩行者の割合から逆算し、一人の回答意図のある歩行者の回答を得られるまでに必要な人数とした。

これらの結果より、次のことがわかった。

(1) 回答意図のある人の正答率が、回答意図のある人の割合よりも高いときには、

図 7 より、回答意図がある歩行者が 9 割を占めるとき、回答意図のない回答の正答率が 0.676471 である場合と 0.5 である場合双方とも、分類器を用いない場合の方が 1 回答あたりの正答率が高いことがわかる。

次に、図 8 より、回答意図がある歩行者が 5 割を占めるとき、回答意図のない回答の正答率が 0.676471 である場合は、多数決を行わない場合は分類器を用いる方が高い正答率を得られるが、さらに高い正答率を求めて多数決を行うと、分類器の有無によって正答率に大きな差は得られないことがわかる。しかし回答意図がない回答の正答率が 0.5 である場合には、分類器を用いることで大きくタスクへの正答率を向上させることができた。このことから、回答意図のある歩行者が 5 割を占めるとき、回答意図のある回答とない回答それぞれの一人当たりのタスクへの正答率に大きな差がある場合には、分類器を用いることでより高い正答率を得ることができる。

最後に、図 9 より、回答意図がある歩行者が 1 割を占めるとき、回答意図のない回答の正答率が 0.676471 である場合は、多数決を行わない場合は分類器を用いた方が高い正答率が得られるが、多数決を行う場合は分類器を用いない方が少ないタスク画面上の歩行者数で高い正答率を得られることわかる。一方、回答意図のない回答の正答率が 0.5 である場合は、大人数で多数決を行ったとしても、分類器を用いないときの正答率は分類器を用いたときの正答率よりも低いことがわかる。

これらの実験結果より、回答意図のある回答とない回答の正答率の差が大きく、また ToF システムのタスク画面上を通過した歩行者のうち回答意図のある歩行者が占める割合が少ない場合に、分類器による回答意図の有無でのフィルタリングがデータ品質向上に効果的であることがわかった。

このことから、回答意図のある歩行者の占める割合や回答意図のある回答とない回答それぞれの正答率を計算し、それに合

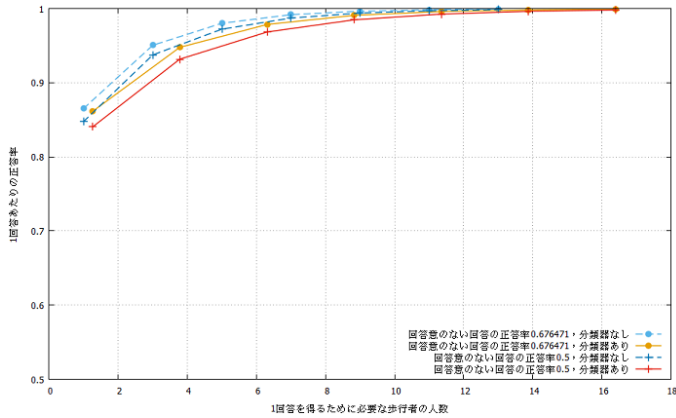


図7 評価実験2結果：回答意図の有る歩行者が9割である場合のタスクへの正答率シミュレーション結果

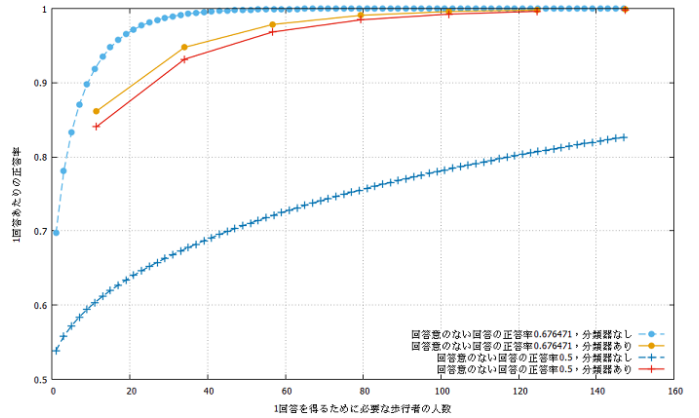


図9 評価実験2結果：回答意図の有る歩行者が1割である場合のタスクへの正答率シミュレーション結果

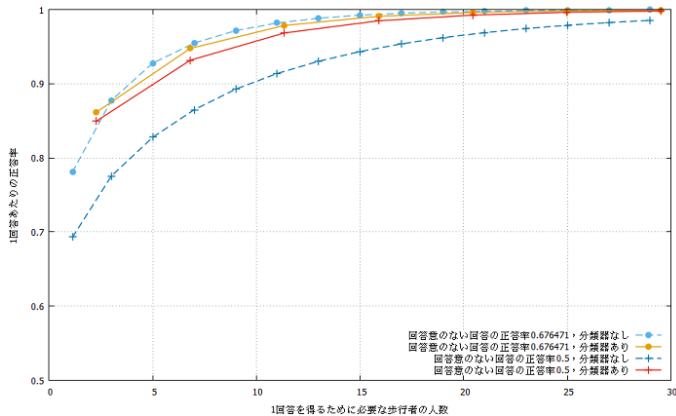


図8 評価実験2結果：回答意図の有る歩行者が5割である場合のタスクへの正答率シミュレーション結果

わけて分類器の使用の有無を切り替えることが重要であるとわかった。分類器で各歩行者の回答意図の有無を分類し、ToFシステムにより回答を受け付けた歩行者のうち回答意図のあった歩行者が占める割合を求める。その割合によって、分類器によるフィルタリングや多数決を行うか否かを切り替えることで、より高いタスクへの正答率を得ることを試みる。また、回答の正誤を確認できるタスクである場合は、回答意図のある歩行者とない歩行者それぞれの正答率を計算し、1回答あたりのタスクへの望ましい正答率に合わせて、より細かく分類器の有無や多数決の有無を切り替えることができる。

## 7. まとめと今後の課題

本研究では、歩行中のマイクロタスク処理を対象に、回答意図の有無の判定を利用したデータ品質向上について検討をおこなった。また、機械学習によって回答意図の有無の判定する分類器を構築した。実験により、回答意図の有無の判断に用いる素性としては、歩行者の進行方向の変化、頭の角度、歩いた位置が有効であることがわかった。また、回答意図の有無の判定結果を利用することがデータ品質の向上につながるケースを明らかにした。

今後の課題としては、ToFシステムにより回答を受け付けた

歩行者のうち、回答意図があった歩行者がどの程度の割合を占めるかによって、分類器の適用の有無を自動的に切り替える仕組みの開発がある。

## 8. 謝辞

筑波大学，明治大学，神戸大学，同志社大学には，平素よりToFシステムを設置させて頂き，協力を頂きました。本研究の一部は科研費(#25240012)による。

## 文献

- [1] Luis von Ahn, Laura Dabbish. :Labeling images with a computer game, CHI '04 Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp319-326
- [2] 品川有輝, 森嶋厚行, 中村聡史, 寺田努「日常空間に組み込んだ Human Computation 環境によるクラウドソーシングタスク処理」. 情報処理学会 インタラクション 2014 論文集 pp706-707
- [3] Crowd4U, 「Crowd4U とはなんですか?」, <https://crowd4u.org/ja/about>, 最終アクセス日 2015 年 12 月 22 日
- [4] Jeffrey M. Rzeszotarski, Aniket Kittur. :Instrumenting the crowd: using implicit behavioral measures to predict task performance, UIST '11 Proceedings of the 24th annual ACM symposium on User interface software and technology, pp 13-22
- [5] Kurtis Heimerl, Brian Gawalt, Kuang Chen, Tapan S. Parikh, Bjorn Hartmann. : Communitysourcing: Engaging Local Crowds to Perform Expert Work Via Physical Kiosks, CHI'12 Proceeding of the SIGCHI Conference on Human Factors in Computing Systems, 2012. pp1539-1548.
- [6] Jorge Goncalves, Denzil Ferreira, Simo Hosio, Yong Liu, Jakob Rogstadius, Hannu Kukka, Vassilis Kostakos. :Crowdsourcing on the spot: altruistic use of public displays, feasibility, performance, and behaviours, UbiComp'13, Proceeding of the 2013 ACM international joint conference on Pervasive and ubiquitous computing, pp753-762
- [7] Rajan Vaish, Keith Wyngarden, Jingshu Chen, Brandon Cheung, Michael S. Bernstein. :Twitch Crowdsourcing : Crowd Contributions in Short Bursts of Time. ACM, Proceedings of the 32nd annual ACM conference on Human factors in computing systems, pp3645-3654
- [8] Atsuyuki Morishima, Norihide Shinagawa, Tomomi Mitsuishi, Hideto Aoki, Shun Fukusumi : CyLog/Crowd4U: A Declarative Platform for Complex Data-centric Crowdsourcing. PVLDB 5(12): 1918-1921 (2012)