Vol.2014-HCI-157 No.55
Vol.2014-GN-91 No.55
Vol.2014-EC-31 No.55
2014/3/15

ヒューマンコンピュテーションにおける非負値行列因子分解を用いたタスク割り当て手法の提案

巻口 誉宗[1,a)]　並河 大地[1]　東 正造[1]　下村　道夫[1]　金丸　直義[1]

概要：ヒューマンコンピュテーション（Human Computation, HC）において, ワーカのタスク処理結果から高品質なデータを取得する品質管理手法 (Quality Control) は重要課題の一つである. 本研究では, 少ないタスク依頼数で高い品質の処理結果を得る手法のひとつとして, タスクに対して適切なワーカの選択を行う**タスク割り当て手法**に注目する. 具体的には, ワーカの過去のタスク処理結果を元に非負値行列因子分解（NMF）によるクラスタリングを行い, タスクに適したワーカを含むクラスタを抽出する手法を提案する. 固有表現ラベリングと画像ラベリングの 2 種類のタスクに対する処理結果の実データを用いて, 提案手法によるタスク依頼数の削減と処理結果の品質向上の実現可能性を示す.

1. はじめに

ヒューマンコンピュテーション（Human Computation, 以下 HC）は, 計算機技術のみでは自動的に処理することが困難な問題を, 人の処理能力によって解決しようとする手法 [1] である. 計算機によって画像や音声, 自然言語等のメディアを認識し, 加工する技術エンジン（以下, メディア処理, 知識処理エンジン）の中には, 現時点ではサービス提供に必要な認識精度基準を満たしていないものや, 更なる認識精度向上のために学習データを大量に要するものがある.

例えば, 画像認識エンジンの場合, 光の透過や反射が起こりやすいガラス製品は認識精度が低い処理対象のひとつである. こうした処理対象がユーザから入力された際, ユーザ以外の人の処理能力を用いてエンジンが出力した認識結果を修正や補足, 認識そのものの代行を行うことができれば, エンジンを用いたサービスの品質向上が期待できる. さらに, エンジンの認識精度が低い処理対象に対する学習データも同時に収集できることから，エンジンの認識精度そのものの向上も期待できる.

現在, メディア処理, 知識処理エンジンの研究開発において, 人の処理能力を必要とする学習データ収集を行う場合は, 作業者（以下ワーカ）を専属で雇い, 作業（以下，タスク）を依頼している．このように専属ワーカをその都度確保する方式は, ワーカの選出や契約等の人集めに時間がかかり, 迅速なデータ収集が困難な点や, 大規模なデータ収集時にはワーカのタスク処理時間に応じて報酬コストが増加する点が問題となる.

これらの問題を解決する手法として, 近年は Amazon Mechanical Turk[2] 等の**クラウドソーシング（Crowdsourcing）**が注目されている. クラウドソーシングは, インターネットを通じて大量のワーカに迅速かつ安価にタスクを発注できるシステムで, HC において人の処理能力を確保する代表的な手法の一つである.

クラウドソーシングのように多数のワーカの処理能力を活用する仕組みは, 専属ワーカと比較して処理時間と報酬コストの削減の観点で優れている. 一方で, 不特定多数にタスクを依頼するため, スパムワーカ等のノイズによる品質低下や, 専属ワーカとの比較において専門知識やスキル等の専門性を必要とするタスクの品質担保が困難といった問題がある.

すなわち HC において, 専属ワーカは, 高品質にタスクを処理できるが, 報酬コストが高く（**問題 1**）, クラウドソーシングは, 報酬コストは低いが品質の担保が困難（**問題 2**）という特徴をもつ.

本研究では, 問題 1 の専属ワーカの報酬コストの高さを解消するために，クラウドソーシングのような多数のワーカへのタスク依頼を前提とし，問題 2 に関して品質の担保が実現可能な **HC プラットフォーム（以下，HC-PF）**の構築を目指す.

2 章では，クラウドソーシングにおける品質低下を防ぐ品質管理手法について，先行研究を述べる. 3 章では品質管理手法の中でも，ワーカへのタスク割り当て手法に注目

[1] NTT サービスエボリューション研究所
1-1 Hikarinooka Yokosuka-Shi, Kanagawa 239-0847, Japan
[a)] makiguchi.motohiro@lab.ntt.co.jp

ⓒ 2014 Information Processing Society of Japan

Vol.2014-HCI-157 No.55
Vol.2014-GN-91 No.55
Vol.2014-EC-31 No.55
2014/3/15

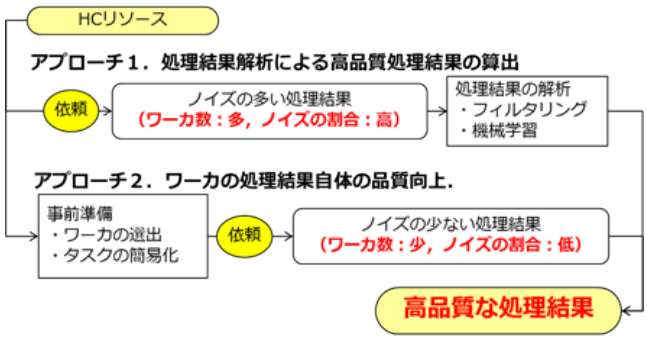


図 1 品質管理の 2 つのアプローチ

し，従来手法の整理と提案手法について述べる．4 章では提案手法の有効性確認のために行った実データ収集実験，ならびに解析について述べる．5 章では実験結果，および考察を述べる．6 章では今後の課題，最後の 7 章はまとめである．

2. 品質管理手法

ワーカのタスク処理結果には様々なノイズが含まれる可能性があり，クラウドソーシングをはじめとした HC に関する研究では，ノイズを除去し，より高品質な処理結果を得る**品質管理手法（Quality Control）**が，重要な要素として注目されている（[3], [4], [5], [6], [7]）．

処理結果の品質を低下させるノイズとして，下記のワーカが考えられる．

- 意図的に低品質な処理結果を返すスパムワーカ．
- タスク処理に対する能力や知識が不足している適性外ワーカ．
- 外的・内的要因で本来のパフォーマンスを発揮できないヒューマンエラーワーカ．

2.1 品質管理手法の先行研究

これらのノイズを除去するための品質管理手法の先行研究は，主に以下の 2 つのアプローチから研究されている（図 1）．

アプローチ 1

ワーカの処理結果解析による高品質処理結果の算出．

アプローチ 2

ワーカの処理結果自体の品質向上．

アプローチ 1 は，ノイズが多く含まれる大量の処理結果から，高品質な処理結果を算出するアプローチである．本例としては，同じタスクを複数のワーカが処理する「処理結果の冗長性確保」を前提とし，ワーカの属性を推定して正解を算出する手法 [3] や，機械学習の識別機を処理結果から直接推定する手法 [4] が提案されている．このアプローチでは，処理結果の冗長性確保のために，タスク依頼数の増加に伴う報酬コストと，必要な処理結果が揃うまでの時間コストが増加し，さらに一般正解率が低いタスクから正解を算出することが困難といった問題がある．

アプローチ 2 は，タスクをワーカに依頼する前に，タスクの簡易化や報酬設計の工夫，適切なワーカの選出といった処理を行うことで，処理結果自体の品質を向上させるアプローチである．中でも適切なワーカを選出し，タスクを割り当てる手法は，少ないワーカのタスク処理結果から高品質な処理結果が得られるため，アプローチ 1 のデメリットの解消，すなわち報酬コスト，時間コストの削減，正解率が低いタスク処理の実現が期待できる．ワーカへのタスク割り当ての先行研究として，例えば芦川らは，タスク依頼者がワーカの情報を管理可能な Private Crowdsourcing System(以下 PCSS) によって，ワーカのスキルや経験値に基づきタスクをマッチングすることで品質向上が実現される可能性について述べている [7]．

次章では，タスクの割り当て手法について既存手法を整理することで，具体的な要件を定義し，提案手法を述べる．

3. タスク割り当て手法

タスクとワーカを適切にマッチングさせるためには，**ワーカとタスクの属性推定（属性推定段階）**，**タスクに最適なワーカの選定（ワーカ選定段階）**の 2 段階の過程が必要である．アプローチ 1 のワーカ数増加（タスク依頼数の増加）のデメリットを解消するため，**タスク依頼数の削減**がまず要件に挙げられる（**要件 1**）．また，属性推定段階には，人による属性推定と，計算機による属性推定の 2 パターンがある．以下にそれぞれのパターンのメリットとデメリットをまとめ，デメリットを解消するための要件を定義する．

3.1 人によるワーカの属性推定

本例として，リサーチ会社のスクリーニングに代表される「自己申告」がある．自己申告のメリットは，詳細な属性情報を容易に取得できる点等が挙げられる．一方，デメリットとして，スパムワーカや，ヒューマンエラーワーカ，自らの状態を正しく把握していない適性外ワーカからは，信頼できる申告結果が得られないという点等が考えられる．ワーカが申告する自信度について，小山ら [5] は，ワーカが申告した「タスク処理結果への自信度」と実際の正解率の間にばらつきがあることを報告している．このデメリットを解消するために**ワーカの自己申告情報以外の情報利用**を要件として定義する（**要件 2**）．

また，人によるワーカの属性推定の別の例として，タスクを処理したワーカの評価を，リクエスタや別のワーカが行う「他者評価」[6] がある．この手法は，タスクの処理結果が大量にある場合，評価するワーカが増加し，報酬コスト，時間コストが増加するといったデメリットがある（導かれる要件は要件 1 と等しい）．

ⓒ 2014 Information Processing Society of Japan

3.2 計算機によるワーカの属性推定

本例として, テストタスクが挙げられる. 正解が既知のテストタスクをワーカに処理させ, 正解率を算出することで, そのタスクに適したワーカの選定や, スパムワーカ, 適正外ワーカを除外できると考えられる.

デメリットとして, テストタスクの処理にかかる報酬コストや時間コストの増加が挙げられる（導かれる要件は要件 1 と等しい）. 他の属性推定手法の例として, 処理時間の解析がある. 例えば NTT コム リサーチ [8] では, アンケートの回答時間が短いワーカ 5% を解答集計に含めないという品質管理が可能である. この手法では, アンケートのように一定以上の想起時間を要するタスクにおいては, 処理時間の短さからスパムワーカを判断し, 除去することができると考えられるが, 画像のカテゴリ分類のようにワーカの得意不得意によって処理時間にばらつきが出るようなタスクにおいては, 処理を早く正確に行えるワーカを排除してしまう可能性がある.

我々が検討するマッチング手法は, HC-PF へ適用するという観点で, 処理するメディアの種類や難易度, 正解の有無等, 異なる種類のタスクで共通的に利用できることが望ましい. そこで, **複数種類のタスクに適用可能**という要件を定義する（**要件 3**）.

3.3 提案手法

前述の属性推定はいずれも, ワーカ個人を対象として属性を推定するものである. しかし, 例えばテストタスクによる属性推定を個人単位で行う際には, 推定したい人数分, テストタスクを依頼するため, タスク依頼数が膨大になり, 要件 1 を満たせない. また, アンケートのような正解がないタスクにおいては, 個人の処理結果よりも, ワーカの集団としての処理結果が重要視され, 個人単位の属性推定は効果的ではない.

そこで本研究では**ワーカの集団に対する属性推定**に注目し, ワーカのクラスタリングによるタスク割り当て手法を提案する. 具体的には, ワーカの過去のタスク処理結果から類似した回答傾向を持つワーカ群を作成（属性推定）し, これから依頼しようとする新規タスクに適した**エキスパートワーカ**群を選定する.

提案手法は, ワーカの過去のタスク処理結果からクラスタを作成する点で要件 1, 2 を満たすことができると考えられる. また, ワーカのクラスタリングと同様にタスクをクラスタリングすることで, 新規タスクと類似した過去のタスクを抽出し, そのタスクに対するエキスパート群の算出が可能となる. これにより, 種類の異なるタスクへの適用という要件 3 も満たすことができると考えられる. クラスタリングによるワーカ群の属性推定の先行研究として, Gomes ら [9] は, 画像のカテゴリ分類をワーカの主観によって行わせ, その処理結果を元にワーカをクラスタリングする手法を提案している. 我々の手法は, 1 種類のタスクではなく, 複数種類のタスクを対象とする点, クラスタを分析し, エキスパート群の選定を行う点で Gomes らの先行研究と異なる. 次章以降では, 提案手法が要件 1,2,3 を満たす可能性があることを, タスク処理結果の実データ解析により示す.

4. タスク処理結果の実データ解析

提案手法の実現性, および要件を満たす可能性を検証するため, まずはタスク処理結果のクラスタリングによってエキスパート群の選定が可能であることを, 実データ解析によって確認した. 以下に実データ収集, 解析について詳細を述べる.

4.1 タスク

メディア処理, 知識処理エンジンの研究で使われるラベル付き学習データのうち, 過去に専属ワーカによってラベリングされたデータを用いて以下の 2 種類のタスクを作成した.

タスク 1:固有表現ラベリング

単語単位で提示された固有表現に対し, 該当するクラスを一覧から選択する. クラスは関根ら [10] の拡張固有表現階層を用いた.

タスク 2:画像ラベリング

提示された写真に対し, 写っているオブジェクトを一覧から選択する. 写真は PASCAL Visual Object Classes Challenge 2010[11] を用いた.

4.2 クラスタリング手法

クラスタリング手法には, **非負値行列因子分解（Non-negative Matrix Factorization：以下 NMF）**を用いた. NMF は I 行 $\times J$ 列の行列 $\mathbf{X}$ を, I 行 $\times K$ 列の特徴行列 $\mathbf{T}$, K 行 $\times J$ 列の特徴行列 $\mathbf{V}$ に分解し, $\mathbf{T}$ と $\mathbf{V}$ の積 $\mathbf{TV}$ と, 分解前の行列 $\mathbf{X}$ の距離 D が最小となるよう, $\mathbf{T}$ と $\mathbf{V}$ を最適化する手法である. NMF は文章データや購買データにおける頻出パターンを算出する用途で広く利用されている [12]. NMF の処理結果から得られる特徴行列 $\mathbf{T}$, $\mathbf{V}$ は, それぞれタスク, ワーカがクラスタ $k \in \{1, \ldots, K\}$ に属する強さを表す. 代表的なクラスタリング手法である K-means と比較した場合, ワーカとタスクを同時にクラスタリングし, 関係性を把握しやすい点を重視し, 初期検討として NMF を選択した.

$\mathbf{T}$ と $\mathbf{V}$ の積 $\mathbf{TV}$ と, $\mathbf{X}$ の距離 $\mathbf{D}$ の定義は, 一般化 Kullback-Leibler divergence（以下, KL divergence）を用いた [12].

ⓒ 2014 Information Processing Society of Japan

Vol.2014-HCI-157 No.55
Vol.2014-GN-91 No.55
Vol.2014-EC-31 No.55
2014/3/15

例）画像ラベリングタスク		選択肢	ワーカ（j in {1,...,J}）		
			ワーカ 1	ワーカ 2	ワーカ J
タスク (i in {1,...,I})	1	犬	1	0	0
		猫	0	0	0
		鳥	0	1	1
	2	犬	0	0	0
		猫	1	1	1
		鳥	0	1	1

処理結果行列X

図 2　処理結果行列の作成

$$D*(X,TV)=\sum_{i=1}^{I}\sum_{j=1}^{J}d*(x_{ij},\widehat{x_{ij}}) \tag{1}$$

$$d*(x_{ij},t_i^T,v_j)=x_{ij}\log\frac{x_{ij}}{t_i^T v_j}-x_{ij}+\widehat{x_{ij}} \tag{2}$$

ここで $\widehat{x_{ij}}=t_i^T v_j$ であり, KL divergence を用いた $\mathbf{T}$, $\mathbf{V}$ の各要素 t_{ik}, v_{kj} の最適化は, 以下の更新式で行われる.

$$t_{ik}\leftarrow t_{ik}\frac{\sum_j \frac{x_{ij}}{\widehat{x_{ij}}}v_{kj}}{\sum_j v_{kj}},\ v_{kj}\leftarrow v_{kj}\frac{\sum_i \frac{x_{ij}}{\widehat{x_{ij}}}t_{ik}}{\sum_i t_{ik}} \tag{3}$$

4.3　処理結果行列の作成

NMF の対象とする行列 $\mathbf{X}$ は，タスク処理結果から以下の手順で作成した．それぞれのタスク番号を $i\in\{0,\ldots,I\}$, ワーカの番号を $j\in\{0,\ldots,J\}$ とし, 各タスクの処理結果を $[0,1]$ の 2 値で表現する．すなわちワーカ j の i 番目のタスクへの処理結果ベクトルを $\mathbf{y_i^j}$ する．例えば画像ラベリングタスク i において, オブジェクトの選択肢が [犬, 猫, 鳥] であり, ワーカ j が犬と鳥を選択した場合, $\mathbf{y_i^j}=[1,0,1]$ と表現される．この処理結果ベクトルを結合し, 処理結果行列 $\mathbf{X}$ を作成した (図 2).

4.4　実データ収集と解析

実データは社内にクラウドソーシングウェブサーバを作成し, 33 名の被験者へタスクを依頼, 処理結果を収集した (図 3，図 4).

タスク処理結果を基に NMF によってワーカをクラスタリングし, 得られたクラスタと正解率を解析した.

5.　実験結果

5.1　正解率とクラスタとの関係

図 5 に, タスク 1（固有表現ラベリング）のワーカごとのタスク正解率のグラフを示す．横軸はワーカ ID であり, 縦軸は正解率（0 ～ 1）である．また, 図 6 に NMF によってワーカを 3 つのクラスタに分類（$K=3$）したうち, 各ワーカが 2 つめのクラスタ（$k=2$, クラスタ 2）に属する強さ（帰属度）をプロットしたグラフを示す．図 6 から, ワーカ $\{3,9,11,19,20,21,22\}$ はクラスタ 2 に同程度の強さで属する．また, 図 5 と図 6 の比較から, クラスタ 2 への帰属度が高いワーカは正解率も高い傾向にあると言える.

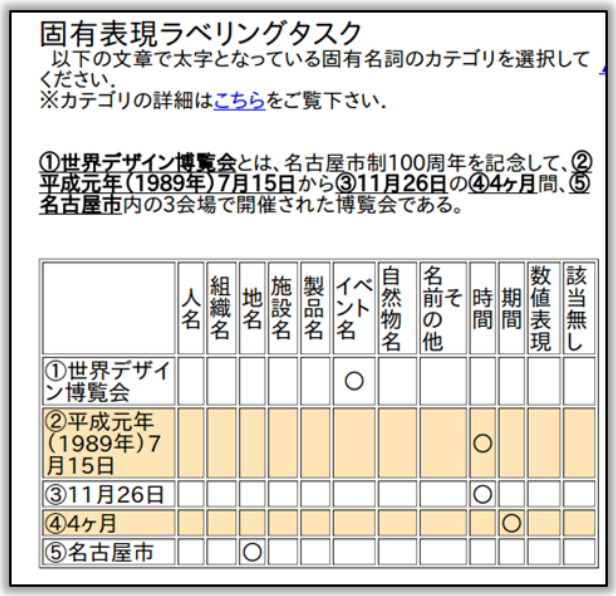


図 3　固有表現ラベリングタスクの処理画面



図 4　画像ラベリングタスクの処理画面

5.2　多数決人数と正解率の推移

NMF によるクラスタリングを用いることで，多数決による正解算出におけるタスク依頼数削減効果を検証した．それぞれのクラスタ $k\in\{1\ldots K\}$ について，帰属度が高いワーカを上位 n 人選出し，多数派となった処理結果をそのクラスタの代表値として，正解率を算出した．クラスタ数が増加するほど似た回答をするワーカ同士をクラスタリングできるため，正解率の高いワーカ群を抽出しやすいと考えられるが，クラスタ数の増加は，その後のエキスパート群選定において評価が必要な対象の増加につながり，要

ⓒ 2014 Information Processing Society of Japan

Vol.2014-HCI-157 No.55
Vol.2014-GN-91 No.55
Vol.2014-EC-31 No.55
2014/3/15

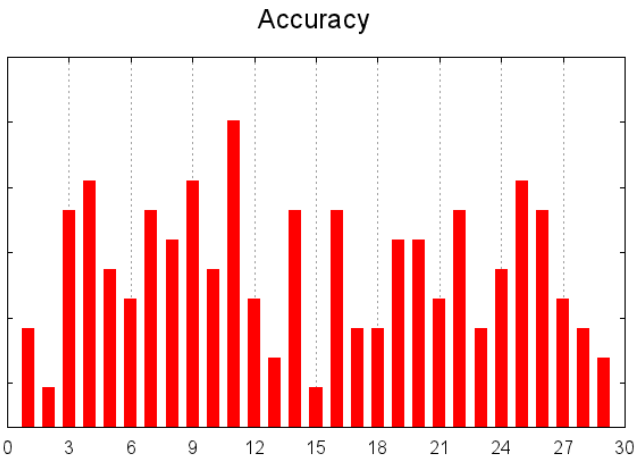


図 5 各ワーカの正解率．横軸；ワーカ ID，縦軸：正解率

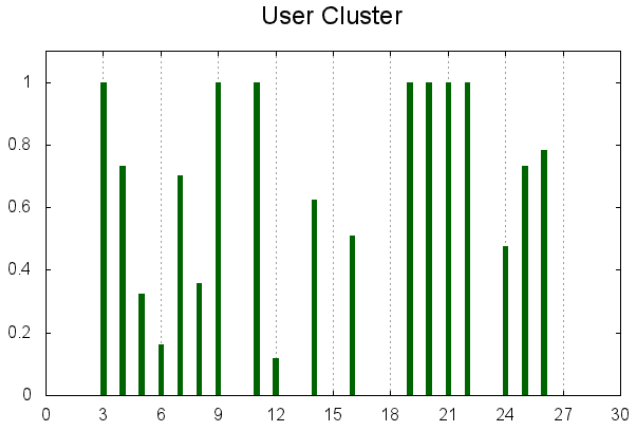


図 6 各ワーカのクラスタ 2 への帰属度．横軸；ワーカ ID，縦軸：帰属度

件 1 の観点から望ましくない．本解析ではクラスタ数を手動で $K = 3$ と設定し，クラスタリングを行った．

初期値によるクラスタリング結果の変動を排除して評価するため，33 名から 29 名のワーカをランダムに抽出する試行を 10 回繰り返し，それぞれの試行において最も正解率の高いクラスタの正解率，正解率の低いクラスタの正解率の平均と標準偏差を算出した．また，比較手法として，29 名のワーカからランダムに n 人抽出し多数決を取った正解率を算出した．

図 7 に，タスク 1 の処理結果を基にワーカをクラスタリングし，n を 1 ～ 29 まで変化させた際の多数決人数と正解率の推移を示す．クラスタ数 K は 3 とした．図 7 の青線（四角）は 3 つのクラスタの中で最も正解率が高いクラスタの正解率平均値，赤線（四角）は最も正解率が低いクラスタの正解率平均値である．ランダムでワーカを抽出し，多数決をとった正解率は黒線（丸）で示した．

このグラフから，ランダムにワーカを抽出した場合（黒線）は，多数決人数を増やしても正解率が改善されない一方で，正解率の高いクラスタは少ない多数決人数で正解率

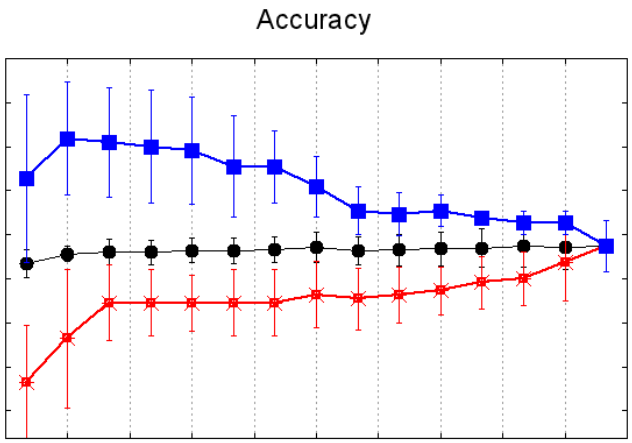


図 7 タスク 1 における多数決人数と正解率の推移（クラスタ数 3）．横軸：多数決を取る人数，縦軸：多数決結果の正解率

が最大となり，人数を増やしていく毎に正解率が低下する．

青線のクラスタにおいて少ない多数決人数で高い正解率が得られる要因としては，帰属度が高いワーカは正解率も高いことが考えられる．また，多数決人数の増加，すなわち帰属度が低いワーカの割合の増加に伴って正解率が低下することから，青線のクラスタへの帰属度が低いほどワーカの正解率も低下していると考えられる．このことから，NMF によって正解率の高いエキスパート群の抽出が可能で，ランダムにワーカを抽出する場合に比べて少ないワーカ数でより質の高い処理結果が得られる可能性が示唆される．

5.3 タスク種類への対応とクラスタ数について

図 8 に，タスク 2 の画像ラベリングタスクから生成した 3 つのクラスタについて，図 7 と同様にワーカを抽出し，多数決人数と正解率の推移を算出した結果を示す．このグラフから，タスク 2 においては正解率の高いクラスタであっても，ランダムにワーカを抽出した場合と大差なく，クラスタリングによるエキスパート群の抽出効果は低いと言える．

図 9 に，クラスタ数を 5 としたグラフを示す．このグラフでは，ランダム抽出よりも高い正解率のクラスタが生成されている．このことから，タスクの種類によって適切なクラスタ数やタスク依頼数削減の効果は変化すると考えられる．

5.4 クラスタによるワーカの属性推定について

あるタスクが苦手なワーカを対象としてアンケートを取りたい場合など，正解率の低いクラスタの抽出が重要となるケースもある．図 7，図 8，図 9 では，ランダム抽出に比べて正解率が低いクラスタも抽出されており，こうしたケースにもクラスタリングによる属性推定が応用可能と考えられる．特に正解がないタスクにおいては，多様な処理

ⓒ 2014 Information Processing Society of Japan

Vol.2014-HCI-157 No.55
Vol.2014-GN-91 No.55
Vol.2014-EC-31 No.55
2014/3/15

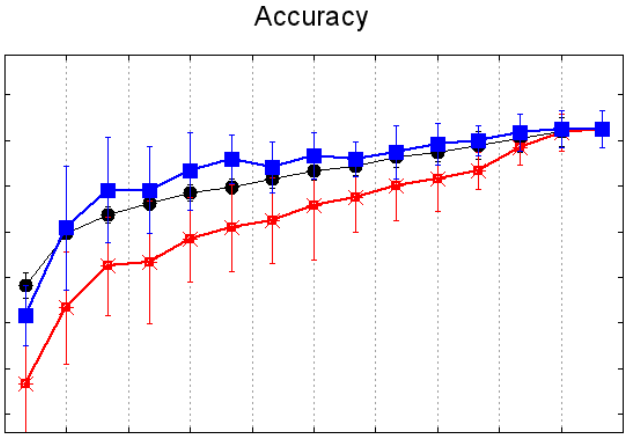


図 8　タスク 2 における多数決人数と正解率の推移 (クラスタ数 3).
横軸：多数決を取る人数，縦軸：多数決結果の正解率

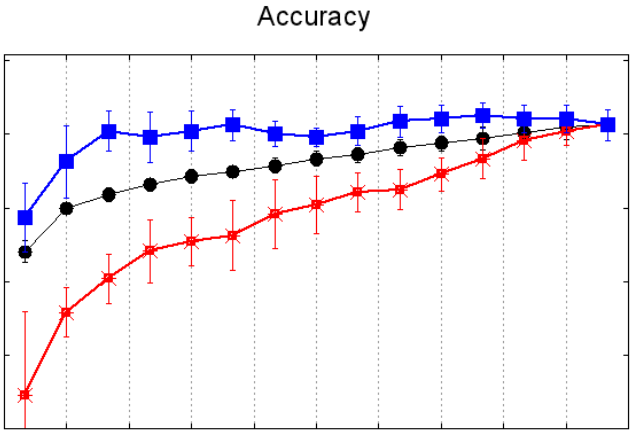


図 9　タスク 2 における多数決人数と正解率の推移 (クラスタ数 5).
横軸：多数決を取る人数，縦軸：多数決結果の正解率

結果を収集することが求められるケースもあることから，提案手法によって属性が異なるワーカのクラスタに均等に処理を依頼することで, 処理結果の多様性を少ない人数で確保できるといった利点も考えられる.

6. 今後の課題

6.1 クラスタからのエキスパート群選択について

得られたクラスタをタスク割り当てに利用するためには, 新規のタスクに対するエキスパート群を選択する必要がある. そのため, 今後はクラスタリングによる新規タスクと過去のタスクの類似性解析や, クラスタ数の決定規則手法, エキスパート群の具体的な選定手法について検討を進める.

エキスパート群の選定にはクラスタ代表者へのテストタスクの依頼が現実的と考えられるが，クラスタ数が多いとその分代表者が増え，タスク依頼数が増加する．そのためクラスタ数を増やすほど，正解率の高いエキスパートワーカ群を選出しやすくなると考えられるが，選定にかかるタスク依頼数とのトレードオフを検証する必要がある.

6.2 ワーカの属性の偏り，モチベーションについて

今回の実データ収集は, 社内でワーカを募ったため，収集したデータそのものにワーカの属性の偏りが想定される．また，無報酬での依頼のため, 金銭的な報酬が得られるクラウドソーシングやリサーチ会社によるアンケートとはモチベーションが異なり, 処理結果に影響を与えていることも考えられる．今後は, 実際のワーカとして想定されるリサーチ会社等を用いて, 多様なワーカを対象とした実データ収集実験を行い, タスク割り当て手法の有効性を実証する.

7. まとめ

本稿では HC における品質管理手法としてワーカのタスク処理結果によるクラスタリングを用いたタスク割り当て手法を提案し，固有表現ラベリングと画像ラベリングの 2 種類のタスクに対する処理結果の実データを用いて, 提案手法がタスク依頼数の削減と処理結果の品質向上の実現可能性を示した.

参考文献

[1] von Ahn, L.: *Human Computation*, Doctoral Thesis. UMI Order Number: AAI3205378, CMU, (2005).
[2] Amazon Mechanical Turk, `http://aws.amazon.com/jp/mturk` (2014.02.13).
[3] Dawid, A.P. and Skene, A. M., *Maximum Likelihood Estimation of Observer Wrror-Rates Using the EM Algorithm*, Journal of the Royal Statistical Society. Series C, Vol. 28, No.1, pp. 20-28,(1979).
[4] 梶野 洸, 坪井 祐太, 佐藤 一誠, 鹿島 久嗣, エキスパートによる訓練データとクラウドソーシングで作成した訓練データからの教師付き学習, 人工知能学会論文誌, Vol.28, No.3, pp.243-248, (2013).
[5] 小山 聡, 馬場 雪乃, 櫻井 祐子, 鹿島 久嗣, クラウドソーシングにおけるワーカーの確信度を用いた高精度なラベル統合, 人工知能学会全国大会, 2M5-OS-07b-2, (2013).
[6] 馬場 雪乃, 鹿島 久嗣, 非定型出力をもつクラウドソーシングタスクにおける成果物の統計的品質推定, 人工知能学会全国大会, 2M5-OS-07b-1, (2013).
[7] 芦川 将之, 有賀 康顕, 宮村 祐一, *PrivateCrowdSourcing* を用いた言語，音声資源の収集, 人工知能学会全国大会, 3M3-OS-07d-2, (2013).
[8] NTT コム リサーチ, `http://research.nttcoms.com/service/qpolicy4.html/` (2014.02.13).
[9] R. Gomes, P. Welinder, A. Krause, and P. Perona, *Crowdclustering*, Advances in Neural Information Processing Systems, (2011).
[10] 関根の拡張固有表現階層 *-7.1.0-*, `https://sites.google.com/site/extendednamedentityhierarchy/` (2014.02.13).
[11] PASCAL2 Visual Object Classes Challenge 2010, `http://pascallin.ecs.soton.ac.uk/challenges/VOC/voc2010/` (2014.02.13).
[12] 澤田 宏, 非負値行列因子分解 *NMF* の基礎とデータ／信号解析への応用, `http://www.kecl.ntt.co.jp/icl/signal/sawada/mypaper/829-833_9_02.pdf/` (2014.02.13).

ⓒ 2014 Information Processing Society of Japan