

SD 法による画像印象のタグ付けを支援する可視化

飯島 絋理[†] 伊藤 貴之[‡]

[†]お茶の水女子大学理学部情報科学科 〒112-8610 東京都文京区大塚 2 丁目 1 番 1 号

E-mail: [†] {g1620501, itot}@is.ocha.ac.jp

あらまし 例えば画像の印象を推測する機械学習の構築において、学習に使用する訓練データの作成の際には、複数の作業者が各自の主観にもとづいて各画像の印象をタグ付けする作業が生じる。この時、タグ付け作業者が有する印象には個人差があり、学習結果がそれに依存するという点で運用上の問題がある。本研究では、多人数による印象回答値から各画像の印象を推定する決定木を構築することで印象のタグ付けを支援することを想定し、この決定木を可視化するシステムを提案する。本手法ではまず SD 法を採用した印象評価を実施し、続いて各画像の印象値を用いてファジィ決定木を作成する。これを可視化することで、各画像の印象が推定される過程を観察できる。さらに、決定木によって分類される画像群を Image Browser を用いて可視化することで、決定木の構造と画像特徴の相関を観察できる。本報告では、1500 枚の女性着衣服画像に 50 人の作業者が印象評価を実施したデータを題材として、本手法による可視化の実行例を示す。

キーワード SD 法, 可視化, 決定木, アノテーション

1. はじめに

SD 法[1]は概念の意味を測定するためにアメリカの心理学者 Osgood らが提唱した手法である。SD 法は概念×回答者×尺度の三相分析を対象としている。三相分析は具体的には任意の概念に対する印象を被験者に評価してもらう実験を指しており、形容詞対へのリッカード尺度での回答にもとづいて実施される。

SD 法を用いた印象評価は学術的にも産業的にも多くの場面で重要な印象評価手段となる。学術的には例えば心理実験や慣性情報処理などの目的で、産業的には例えば広告戦略や顧客分析などの目的で多用される。

SD 法による印象評価データの活用例として、機械学習を用いた印象推定があげられる。画像を閲覧して回答する形式の印象評価を例にすると、印象評価の対象となった画像群を訓練データとして、印象評価結果をアノテーションとして付与した形で機械学習を適用することで、テストデータを構成する画像群の印象を推定できるようになる。ここで、訓練データを構成する全ての画像の印象に対応するアノテーションを 1 人の作業者が付与するとする。この場合、作業者が各画像に対して個人的に有する印象をアノテーションとして付与するため、機械学習の挙動が作業者に依存してしまうという問題点がある。この問題を解決する一手段として、複数の作業者に各画像の印象を回答させて、その回答結果を総合的に判断して各画像にアノテーションを付与する、ということが考えられる。この手段を採用する場合にも、印象推定の適切さが訓練データの品質に大きく影響されることに変わりはない。そのため、訓練データとしての印象評価結果の分布をユーザが理解することが重要な課題となる。

これらの背景を鑑みて我々は、複数の作業者間で個

人差の出る印象評価に対して、作業者群による過去の印象回答値を参照してこれからタグ付けする画像の印象を推定することにより、印象のタグ付け作業を半自動化する研究に取り組んでいる。その一環として本報告では、ファジィ決定木にもとづく印象推定の過程を可視化する手法を提案する。本手法では以下の 3 つの処理工程から構成されるものである。

- 画像群に対して SD 法を採用した印象評価を実施する。
- 各画像の印象値からファジィ決定木を生成し、決定木の動作を可視化する。
- Image Browser によって画像と印象の関係を可視化する。

タグ付けの半自動化という課題は、印象評価結果に限らず機械学習のための訓練データ全般において重要である。第 3 次 AI ブームの中心的存である機械学習の運用現場では、その挙動に対する説明責任を問われることが多くなっている。訓練データの分布を理解することは、機械学習の説明責任という観点からも重要であると考えられる。

本研究の構成は以下の通りである。2 章では関連研究について、3 章では印象評価の内容について、4 章では提案手法について述べる。そして 5 章では本手法の実施例と考察について、6 章で本研究のまとめと今後の課題について述べる。

2. 関連研究

本研究では印象評価結果に対してファジィ決定木を生成し、これを可視化することで、画像に対する印象のタグ付け作業を支援する。本章では関連研究として、ファジィ決定木およびそれ以外の手法を適用した

印象分析手法と、本研究に関係ある可視化手法について紹介する。

ファジィ決定木を用いたSD法による印象分析[2][3]では、リッカート尺度に基づいた印象値をクラスタリングすることで、確信度を持った属性値を得ることができる。さらに、得られた属性値を決定木によって可視化することで、属性値の決定する過程やルールを理解できる。ファジィ決定木は人の感性に合った解釈が可能であり、その過程やルールの説明性が高いことから、印象分析に適した手法として知られている。本研究はファジィ決定木による印象評価の構造化を支援する可視化として位置づけられる。

ファジィ決定木以外の手法を適用した例として、ニューラルネットワークを用いた画像の感性検索[4]が報告されている。この手法では、印象評価データと画像特徴データをそれぞれ別のニューラルネットワークに入力し結合することで、画像に対する印象を推定している。しかし、ニューラルネットワーク内部で印象評価結果が変換される過程がブラックボックス化されるため、学習結果に対する説明性が十分ではない。それに対して本研究ではファジィ決定木を用いることで、属性が決定されるまでの過程やクラスタリングによってルール化された軌跡を辿ることで、説明性を担保することができる。

決定木は必ずしも可読性の高い状態で生成されるとは限らない。専門家のみならず初学者にも決定木から情報を読み取らせるには、可読性の高い視覚表現が有用である。Netoら[5]は決定木を単純かつ分析が容易な行列に変換することで、その解読性を高めている。本研究では平行座標法を用いて決定木を可視化しているが、Netoらの手法を適用することも可能である。

大量画像の階層型データ可視化手法にImage Browser[6][7]が挙げられる。本手法ではファジィクラスタリングにより画像を分類し、決定木によって階層型データを構築する。このようにして構造化された画像群を閲覧するにはImage Browserが有用である。Liuら[8]は、Image Browserによって区画化された画像群に対して、画像特徴が類似する画像を隣接させるように再配置することによって、より視覚的に画像特徴の分布を表現する手法を提案している。本研究においても画像特徴にもとづいて、あるいは印象値にもとづいて画像群を配置することで、本研究の目的に合致した分析が可能となる。

3. SD法を採用した印象評価

1章でも述べたとおり、SD法は概念×回答者×尺度の三相を分析の対象としている。本研究では「概念」は画像であるとして、画像群にSD法を採用した印象

評価を実施した。本章では「尺度」にあたる形容詞対の選出方法を3.1, 3.2節で説明する。次に3.3節で画像の前処理について、3.4節で印象評価の収集と可視化するデータの構築の手順を説明する。

3.1 尺度の因子選択

本研究では使用する画像群を女性衣服の画像に限定した。堀尾ら[9]は衣服画像の印象は「活動・嗜好性」の因子、「力量性」の因子、「質感・立体感」の因子の3因子から構成されていると報告している。「質感・立体感」の因子は媒体の表現方法の影響が大きく、衣服のみが写っている衣服画像ではなく人物が着衣している衣服画像において構成可能である。したがって本研究では「質感・立体感」の因子を採用するため、衣服画像を人物が着衣している衣服画像に限定した。一方、村上ら[10]は衣服画像の印象は「自己主張」の因子、「他者からの推薦」の因子、「穏健性」の因子、「審美性」の因子の4因子から構成されていると報告している。また、古川ら[11]は衣服画像の印象は「色彩」の因子、「正統性」の因子、「年齢・性別」の因子、「装飾性」の因子の4因子から構成されていると報告している。以上の複数の因子から、性差・年齢差を考慮しない因子であり個人差が明らかに大きく出ない因子である「色彩」「立体感」「正統性」「穏健性」「装飾性」の5種類の因子を本研究では採用する。

3.2 パイロットテストと尺度決定

3.1節で選択した因子から列挙した形容詞対を表1に示す。

表1 選択された因子から形容詞対の選出。

因子	形容詞対		相違度
「色彩」	暗い	明るい	5.03
	鈍い	鮮やか	4.78
「立体的」	直線的	曲線的	5.01
	平面的	立体的	4.74
「正統性」	フィットした	ルーズな	5.42
	高級な	質素な	5.84
	洗練された	無粋な	5.57
	フォーマル	カジュアル	6.02
「穏健性」	日常的な	非日常的な	5.08
	普通	異常	4.74
	無個性	個性的	4.61
	自然	不自然	4.86
「装飾性」	くすむ	目立つ	4.47
	地味な	派手な	4.37
	シンプル	ゴージャス	4.77

印象評価を実施するにあたり、被験者が回答する項目数を増やしすぎること、被験者の疲労につながり、ひいては回答の信頼性が下がる可能性がある。これを回避するために我々は、本研究の印象評価に適用する形容詞対を列挙した上で、類似する意味を持つ形容詞対のいずれかを削除する、あるいは本研究で使用する画像に適切な尺度といえない形容詞を削除する、と

いう手順で印象評価のための形容詞対を選出した。この工程のために我々はパイロットテストを実施した。画像には印象評価にも用いる 10 枚の着衣画像を用い、表 1 にある 15 項目の形容詞対に対して、回答者（20 代女性 15 名）から各評価項目に対して 5 段階評価の回答を収集した。この評価結果をもとにして我々は、任意の形容詞対間の距離行列を python の scipy ライブラリを用いて算出した。ヒートマップを用いて距離行列を可視化した結果を図 1 に示す。このヒートマップでは、距離が大きく類似性が低い形容詞対ペアに彩度の高い赤色が割り当てられている。



図 1 パイロットテストにおける 10 枚の画像の印象評価結果を距離行列として可視化した結果。距離が大きく類似性が低い尺度ほど彩度の高い赤色となる。

続いて各形容詞対について、他の形容詞対との距離の平均値を求める。この平均値が大きいほど、他の尺度との相関がないことを意味する。この平均値を「相違度」と称して表 1 に示す。本研究では、この相違度が最大となる形容詞対を因子ごとに抽出する。以上の工程によって選出された形容詞対を表 2 に示す。

表 2 各因子に対して選出された形容詞対。

因子	形容詞対	
「色彩」	暗い	明るい
「立体的」	フィットした	ルーズな
「正統性」	フォーマル	カジュアル
「穏健性」	日常的な	非日常的な
「装飾性」	シンプル	ゴージャス

3.3 画像の前処理

本研究は背景が無地単色である女性の着衣画像^{*1}を対象として印象評価を実施してする。ここで、人物の人相・表情・髪型などが統一されていないと、衣服以外の要因が印象評価に影響する。そこで本研究では、着衣画像から顔を削除する。具体的には、まず OpenCV を用いて顔領域を検出し、検出された顔領域から上を

背景色と同じ色で隠す。更に背景削除処理^{*2}を実施することで、全画像の背景色を統一させる。以上の工程を図 2 に示す。



図 2 画像の前処理手順。

3.4 印象評価の収集と可視化データの構築

表 2 に示す形容詞対と 3.3 節に示す画像群を用いて、印象評価を実施した。画像を 1500 枚、回答者を 20 代女性 50 名、尺度を 5 項目として、5 段階評価で回答を収集した。1 人あたり 500 枚の画像の印象評価を回答しているため、画像 1 枚の回答人数は 16~17 人である。印象評価の内訳を表 3 に示す。

表 3 印象評価データ。

画像	女性着衣画像 1500 枚	
回答者	20 代女性 50 名	
尺度 1	暗い	明るい
尺度 2	フィットした	ルーズな
尺度 3	フォーマル	カジュアル
尺度 4	日常的な	非日常的な
尺度 5	シンプル	ゴージャス
評価段階	5 段階	

4. 可視化手法

各画像の印象が推定される過程を観察するために、本研究では印象値を用いたファジィ決定木を構築し、それを可視化する。それと同時に、クラスタリングされた結果を画像特徴から観察するために、本研究では Image Browser を用いて画像群の分布を可視化する。本章ではこれらの可視化手法の処理手順について述べる。まず 4.1 節では、ファジィ決定木の構築と印象推定の手順を説明する。続いて 4.2 節では、Image Browser を構築する手順を説明する。

4.1 ファジィ決定木

本研究ではクラスタリング手法にファジィクラスタリングを採用する。一般的なクラスタリング（いわゆるハードクラスタリング）では、各クラスターへの帰属は 0 か 1 かの 2 値で判断されるが、ファジィクラスタリングは [0, 1] の範囲の実数で各クラスターへの帰属を算出するため、曖昧で柔軟なクラスタ表現が可能である。本手法ではクラスターへの帰属値をクラスタリングの確信度の値として扱う。印象評価にはリッカート尺度を用いているため、間隔尺度の水準を満たしてい

*1 使用画像：<https://www.kaggle.com/dqmonn/zalando-store-crawl>

*2 背景削除：<https://ja.clippingmagic.com/>

ると仮定する。

ファジィ決定木の構築手順は以下の通りである。処理工程から構成されるものである。

- クラスタリング数と決定木の深さを決定。
- ファジィクラスタリング結果からクラスタリングの確信度を算出。
- 確信度をもとに決定木を構築。

本手法による印象推定の手順を図3に示す。ここでは1枚の画像の5段階評価「暗い-明るい」を例とする。現時点での実装ではクラス数を3に固定し、印象が「暗い」「明るい」「どちらでもない」の計3クラスが各クラスに対応するものとする。また、ファジィ決定木の構築に際して、木の深さの最大値をあらかじめ指定しておく。我々の実験では、決定木の深さを6以上に指定すると決定属性の再帰や繰り返しが頻繁に出現したため、最大深さを5にした。

各画像に対して複数の参加者が回答することを前提にして、図3①に示すように5段階評価の各値を選択した人数を集計し、続いて図3②に示すように各値の人数比を算出する。この値から、Pythonのfcmeansライブラリを用いて、図3③に示すようにファジィクラスタリングの確信度を算出する。そして確信度が一番大きい値をクラスタリング結果として採用することで、印象推定を完了する。

1 2 3 4 5
暗い □ □ □ □ □ 明るい

① 各評価段階の答えた人数を集計
1 2 3 4 5
23 9 11 4 3

② 評価段階ごとの割合を算出
1 2 3 4 5
0.46 0.18 0.22 0.08 0.06

ファジィクラスタリング
クラス数 = 3

③ クラスタリングの確信度を算出
クラス1 クラス2 クラス3
「暗い」 「どちらでもない」 「明るい」
0.6 0.3 0.1



「暗い」印象
と策定された画像

図3 右の画像が印象項目「暗い-明るい」から「暗い」印象であると策定されるまでの手順。3つのクラス「暗い」「どちらでもない」「明るい」から確信度の高いクラスに策定される。

続いて、図3に示した手順で算出される確信度をもとに決定木を生成し、対話的に平行座標プロットを表示するPythonモジュールiparacord^{*3}を適用して決定木を表示する。図4に決定木の構造のイメージ図を示す。

ここでは、木構造を上位階層から下位階層にたどることで、各画像がクラスタリングされる工程を表現する。木構造の各ノードは長方形で描かれており、クラスの種別によって異なる色を割り当てられている。そして、各画像に対して範囲[0, 1]で算出される確信度のヒストグラムによって、画像数を表示する。木構造の葉ノードは、クラスタリング結果となる各クラスを表現している。この可視化結果から、各画像がクラスタリングされる過程を観察できるだけでなく、その確信度も推察可能である。概して、浅いノードでクラスが決定される画像は確信度が高い傾向にあり。逆に深いノードでクラスが決定される画像は確信度が低い傾向にある。

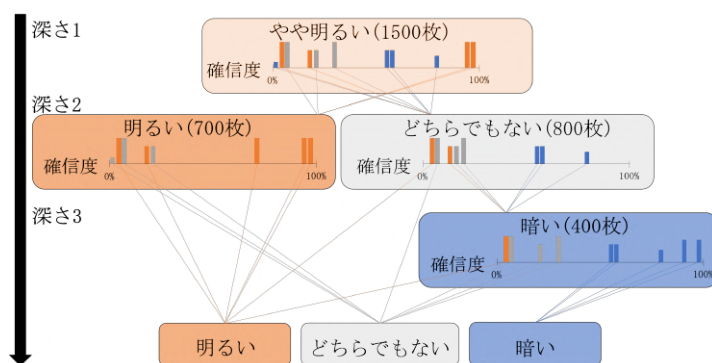


図4 印象が策定される過程を可視化する決定木

4.2 Image Browser

本節では画像と印象の関係を観察するためにImage Browserを構築する手順を説明する。4.1節ではクラスタリング結果による画像の木構造を得ることから、これを画像群に適用する。

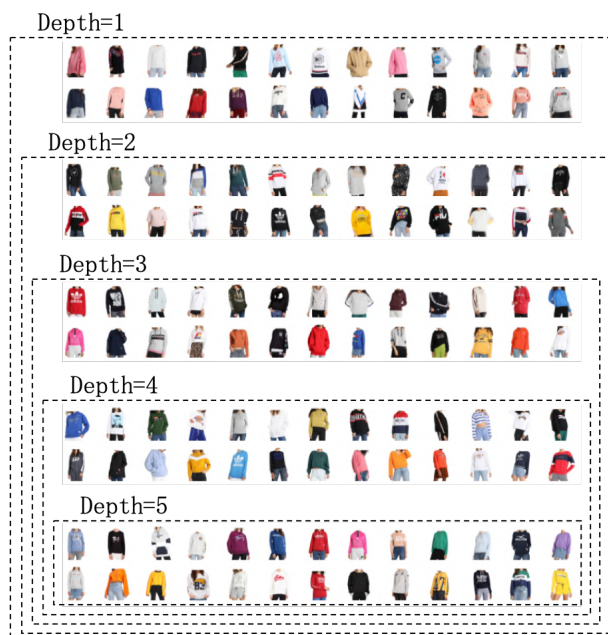


図5 各尺度のクラス別に決定木の深さにもとづいて画像群を表示するImage Browser。

^{*3} iparacord: <https://github.com/intellygenta/InteractiveParallelCoordinates>

本手法で構築する **Image Browser** のスナップショットを図 5 に示す。本手法では、決定木の深さごとに枠を生成し、点線で表示する。枠の中にさらに枠を作る入れ子構造にもとづいており、決定木の深い部位が **Image Browser** での内側に相当する。そして、木構造の葉ノードに対応する領域に画像群を表示する。現時点での実装では、画像の並び順は単にデータ中での出現準となっており、特に意味をもつものではない。

5. 実行例

我々は本手法を **Python** の可視化ライブラリ **Bokeh** を拡張することにより実装した。**Bokeh** は平行移動や拡大縮小といった視点処理、クリック操作等によるオブジェクト選択機能、画像として可視化結果保存、といった対話操作機能が充実した可視化ライブラリである。本事例では表 3 に示す印象評価データを可視化し、以下の観点から結果を観察する。

- クラスタリング分類の確認
- 尺度の適性度
- 決定木の深さによる画像群の関係性

図 6～10 の全体を通して、各尺度のクラスタリング分類結果について、3 つのクラスの間には大きな偏りの存在は確認されなかった。画像群は全ての尺度について、3 段階の印象をほとんど網羅していることがわかる。今回の実行例では、表 3 に示す印象評価を実施する前に、少人数の参加者によるパイロットテストを実施して、尺度が適切であるか、尺度間に強い相関がないか、といった点をあらかじめ見積もっており、これが効果的であったものと考えられる。

また、決定木を観察することで、印象を回答しやすい画像群を発見することも容易である。例えば図 8 における「カジュアル」と図 10 における「シンプル」に関しては、決定木の深さが浅いことから、画像群の中でも印象の確信度が高い尺度であったことが推察される。今回の画像群は衣服の中でもトップスが多く、さらに無地の衣服やトレーナーが多く含まれており、これらが印象を回答しやすい衣服の例であったと考えられる。

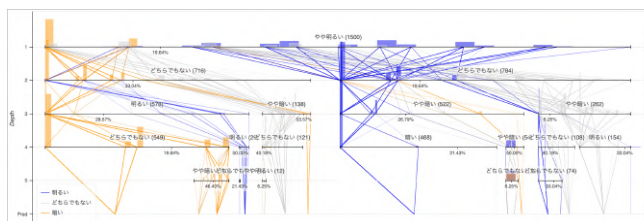


図 6 尺度「暗い-明るい」のファジィ決定木

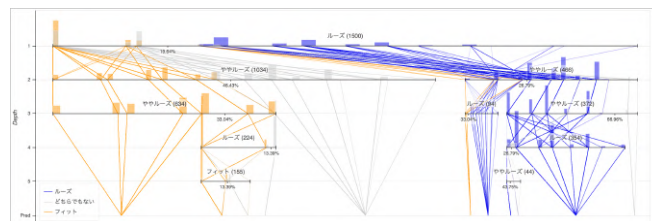


図 7 尺度「フィット-ルーズ」のファジィ決定木

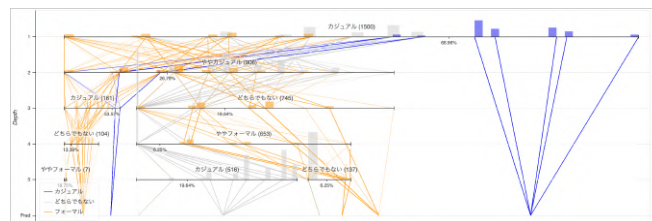


図 8 尺度「フォーマル-カジュアル」のファジィ決定木

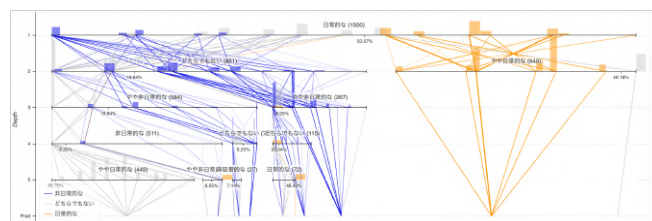


図 9 尺度「日常的-非日常的」のファジィ決定木

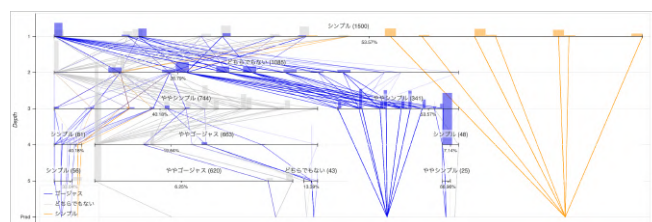


図 10 尺度「シンプル-ゴージャス」のファジィ決定木

図 11 は尺度「暗い-明るい」に対する **Image Browser** での画像表示例である。全体を観察してわかるように、「明るい」印象は色が鮮やかな衣服が多く、「暗い」印象は黒に近い衣服が多い。「どちらでもない」印象には青や緑の衣服、あるいは複数の色を組み合わせた衣服が多いことが観察できる。

図 12(左)は尺度「非日常的-日常的」の「非日常的」に対する **Image Browser** での画像表示例である。原色の衣服が目立つだけでなく、大きなデザインが衣服の中央に存在するものや、柄が多く入った衣服が観察できる。また、深さ 4 のノードに赤色の衣服が多く観察されることが特徴的である。そこで対照となる「日常的な」という印象であると推定された画像群から、同じく赤色の衣服を並べた結果を図 12(右)に示す。両者を比べると、色や形、衣服のデザインが似ているにも関わらず、印象は対照的である。2 つの違いはモデルのポーズと髪の毛が少し写っているかないかである。このことから、衣服の印象評価の際に、モデルの髪形や体格が影響する可能性が示唆される。



図 11 尺度「暗い-明るい」の Image Browser. 視覚的に「明るい」印象は色が鮮やかな服が多く、「暗い」印象は黒の衣服が多い。



図 12 左: 尺度「非日常的な-日常的な」の Image Browser. 右: 「非日常的な」と「日常的な」印象であると策定された画像群の中からそれぞれ似ている衣服を並べた。

6. まとめと今後の課題

本研究では、画像群に対して SD 法を採用した印象評価を実施し、続いて各画像の印象値からファジィクラスタリングを適用してファジィ決定木を構築し、クラスタリングされる過程を可視化した。決定木により画像群の階層型データが得られることから、クラスタリング結果と画像特徴の相関を観察するために、Image Browser を開発してその相関を可視化した。

本手法の用途の一つとして、印象を推測する機械学習の構築のための訓練データ作成があげられる。本研究では、印象推測の題材である画像群のタグ付け作業を複数の作業者に分担してもらうことを想定して、個人差のある印象評価の回答結果を総合的に判断して各画像にタグ付けすることを想定する。本研究の可視化手法により、印象評価結果の分布の観察が容易になり、機械学習の説明責任という観点で訓練データの分布を理解することも容易になると期待される。

今後の課題は以下の通りである。

- 印象評価の回答者の偏りの改善
- 決定木の可読性の向上
- Image Browser の画像表示スタイルの模索
- 本研究の評価測定

現時点で本研究に用いている印象評価結果は 20 代女性 50 名による回答にもとづいており、男性の回答も 20 代以外の世代の回答も含まれていない。データの信頼性や多様性を高めるためにも、偏りのない回答者群による印象評価を実施したい。また、可読性の高い視覚表現を決定木の可視化に採用することで、専門家のみならず、初学者にも決定木から情報を読み取れるようにしたい。加えて、Image Browser の画像群を印象値や特徴量にもとづいて並び替えることで、Image Browser から知見を読み取れる可能性をさらに模索したい。以上を踏まえた可視化手法の拡張を今後の課題としたい。これらの課題を解決した上で、本研究による可視化手法を評価し、タグ付けの半自動化という点でどのような有用性があるかを考察したい。さらに、本手法を用いて生成した訓練データを適用して機械学習を実施し、印象評価の推定結果に改善効果が出るかを評価したい。

謝 辞

本研究の一部は、日本学術振興会科学研究費補助金の助成に関するものです。

ユーザーテストにご協力していただいた皆様に感謝いたします。

参 考 文 献

- [1] Osgood, C.E., "The nature and measurement of meaning" Psychological bulletin, 49(3), p.197, 1952.
- [2] OMatsuo, Takeshi, et al., "Sensitivity Information

Analysis of Running Shoes Using Fuzzy Decision Tree and Visualization of Analytical Results” SCIS & ISIS SCIS & ISIS 2010. Japan Society for Fuzzy Theory and Intelligent Informatics, 2010.

- [3] 井上博行, 王冠, “ファジィ決定木を用いた花束の感性ルール抽出”, 日本知能情報ファジィ学会 ファジィ システム シンポジウム 講演論文集 第25 回ファジィ システム シンポジウム, 日本知能情報ファジィ学会, 2009.
- [4] 太田茂, 竹之内宏, 徳丸正孝, “深層ニューラルネットワークによる特徴抽出を用いた衣服の感性検索”, 日本感性工学会論文誌, 16(3), pp.277-283, 2017.
- [5] Neto, Mário Popolin, Fernando V. Paulovich, “Explainable Matrix - Visualization for Global and Local Interpretability of Random Forest Classification Ensembles”, arXiv preprint arXiv:2005.04289, 2020.
- [6] 五味愛, 伊藤貴之, “「何時, 何処で, 誰と」 3 つのメタ情報を用いた大量個人画像の一覧可視化手法”, 研究報告グラフィクスと CAD (CG), 2010(2), pp.1-6, 2010.
- [7] Bederson, Benjamin B, “PhotoMesa: a zoomable image browser using quantum treemaps and bubblemaps”, In Proceedings of the 14th annual ACM symposium on User interface software and technology, pp. 71-80, 2001.
- [8] Liu, Xiaotong, Yifan Hu, Stephen North, Han-Wei Shen, “CorrelatedMultiples: Spatially Coherent Small Multiples With Constrained Multi-Dimensional Scaling”, In Computer Graphics Forum, vol. 37, no. 1, pp. 7-18, 2018.
- [9] 堀尾華子, 加藤雪枝, “衣服の印象評価における媒体間の比較”, 日本家政学会誌, 53(7), pp.693-701, 2002.
- [10] 村山和弘, 油屋直子, “学生の衣服の色彩嗜好と色彩感情に関する分析”, 尚絅学院大学紀要, 55, pp.157-164, 2008.
- [11] 古川貴雄, 三浦爾子, “ラグジュアリーファッションに注目したトレンドの定量分析”, 共立女子大学・共立女子短期大学総合文化研究所紀要, 23, pp.35-48, 2017.