

SD法による大規模印象評価に基づく アノテーションを支援する可視化

お茶の水女子大学大学院

人間文化創成科学研究科 理学専攻 情報科学コース

akari.iijima@is.ocha.ac.jp

飯島 緋理

(指導教員：伊藤 貴之)

要旨

画像の印象を推測する機械学習では、多数の画像へのタグ付け作業により訓練データを構築することが多い。一般的にアノテーションは少人数で実施されることが多く、1人あたりの作業項目も多い。作業項目が多いことでアノテータの疲労につながり信頼性が下がる可能性がある。また、少人数のアノテータが有する個人の印象回答が学習結果に依存するという点で運用上の問題がある。

本論文では、画像から受ける印象の個人差を解消するため、多人数の印象回答値を参照して印象のタグ付けを自動化する過程を可視化する。可視化結果や分析に基づき、人の手によるタグ付けの再策定が可能になるような可視化システムを提供する。

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生成する。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のRGB値に次元削減を適用し2次元空間に配置する。類似する画像を近くに配置した画像一覧を可視化することで、タグ付け結果を画像特徴から観察する。最後に、ユーザー操作から再策定による訓練データ作成を支援していく。本論文では、1500枚の女性着衣服画像に43人の作業者が印象評価を実施したデータを作成し、それを題材として、本手法による可視化の実行例とユーザー評価の結果を紹介する。

目次

第1章 緒言	1
1.1 研究背景	1
1.2 提案手法の要約	2
1.3 本論文の構成	3
第2章 関連研究	4
2.1 はじめに	4
2.2 印象評価結果の印象分析手法	4
2.3 機械学習の訓練データ作成のためのアノテーション	4
2.4 決定木や階層構造を用いた可視化手法	6
2.5 むすび	11
第3章 提案手法	12
3.1 はじめに	12
3.2 処理手順の概要	12
3.3 SD法を用いた大規模印象評価	12
3.3.1 尺度の因子選択	12
3.3.2 パイロットテストと尺度決定	13
3.3.3 画像の前処理	15
3.3.4 印象評価の収集と可視化データの構築	15
3.4 可視化手法	16
3.4.1 ファジィ決定木	16
3.4.2 類似画像群の一覧表示	18
3.5 ユーザインタフェース	20
3.6 むすび	22

第4章 実行例	23
4.1 はじめに	23
4.2 使用するデータ	23
4.3 可視化結果	23
4.3.1 決定木から得られる知見	23
4.3.2 類似画像一覧表示から得られる知見	25
4.3.3 決定木と類似画像一覧表示のフィルタリング機能から得られる知見	26
4.4 評価実験	27
4.4.1 評価手順	28
4.4.2 画面録画による操作ログによる評価結果	28
4.4.3 再分類結果による評価結果	29
4.4.4 操作・システムに関する評価結果	31
4.5 考察	33
4.6 むすび	33
第5章 まとめと今後の課題	35
5.1 まとめ	35
5.2 今後の課題	36
参考文献	37

第1章 緒言

1.1 研究背景

SD法 [B: Osgood 52] 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。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。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。

一般的に、アノテーション作業は少人数で実施することが多い。1人あたりの作業量が多いため、作業者の疲労が蓄積され、回答の信頼性が減少する可能性がある。そのため、近年ではアノテータのタグ付けの過程や結果を可視化し、訓練データの分布理解及びアノテーションの修正、品質保証を図る研究が活発である。Muller ら [B: Muller 21] はアノテータのタグ付け作業中を観察することによって3つのアノテーション設計におけるデザインパターンを提示しているが、実際はアノテーション自体には決まったルールが存在しない。その上、アノテータは専門家・非専門家を問わない。したがって、アノテーション設計は自由度が高い分、作業者の特性や作業者にかかる時間・負荷、タグの選定方法を吟味しなければ、訓練データが学習結果の足かせとなる問題が発生しうる。

ここで、アノテーションが個人の主観によって決められる場合を考える。Wu ら [B: Wu 21] は衣服の画像に対してアノテータから収集したキャプション・文章と、製品メタデータや属性ラベルを保持する大規模なデータセット作成をした。しかし、収集するキャプションや文章は個人によって様々な回答が得られるため、全回答をアノテーションとして採用するには学習結果に対する信頼性が低くなる。

また、画像の印象を推測する機械学習の構築を考えた時、使用する訓練データは印象のタグが付いた画像を用いる。ここで、従来のアノテーション設計に則り、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、1人あたりの作業量を減らしつつ、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。

1.2 提案手法の要約

本論文では、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、SD法による大規模な印象評価によるアノテーションのデータ構築方法と、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つの工程から構成されるものである。

1. 画像群に対して大規模な印象評価を実施。
2. 印象評価結果から自動的に画像のタグ付けをし、タグ付け結果を可視化。
3. タグ付け可視化結果をもとに人の手によるアノテーションを支援するシステムを提供。(アノテーションの半自動化)

1枚の画像あたりの印象タグ付け作業を少人数ではなく大人数で実施することで、1人あたりの作業負担を軽減しつつ、感性情報の個人差の解消につなげる。これを達成するには、大人数の印象回答値を分析できる可視化システムが要となる。

可視化では、各画像の印象値からファジィ決定木を生成し、さらに類似画像一覧表示をする。ユーザ操作によって決定木と画像群を連動させることで、決定木の可読性を向

上させ、決定木と画像特徴と印象の関係性を明らかにする。これにより、決定木がどのように画像を分類したかをユーザが理解できるようにする。

アノテーションの半自動化という課題は、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 ブームの中心的存在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。可視化システムにより印象策定の半自動化をサポートし、アノテーション作業の効率化を図る。

1.3 本論文の構成

以下、本論文の構成を示す。第2章では関連研究として、既存の印象分析手法やアノテーション手法、決定木や階層構造を用いた可視化手法を紹介する。続いて第3章では提案手法の処理手順について述べる。そして第4章で本手法の実行例とユーザ評価結果の紹介と考察について述べ、最後に第5章で結言について述べる。

第2章 関連研究

2.1 はじめに

本章では，提案手法の関連研究を紹介する．2.2節で既存の印象分析手法について，2.3節で一般的なアノテーション手法について，2.4節で決定木や階層構造を用いた可視化手法について述べる．

2.2 印象評価結果の印象分析手法

SD法による印象評価データの分析方法について紹介する．三相因子分析 [B: 代 82] はSD法の回答の個人差を考慮したデータ分析手法であり，本研究と同様に概念×回答者×尺度の三相を分析の対象とする．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．別の分析手法として探索的ポジジョニング分析 [B: 豊田 01] がある．しかし，これらの手法では，ある相の因子負荷の表と，ある相に次元削減を適用した2次元空間 (因子空間 [B: Bradley 94]) の両方の提示が必要 [B: Kanno 18] であり，単一の可視化結果としてデータを観察することができない．それに対して本研究では，ファジィクラスタリング [B: Bezdek 84] を使用し，単一の画面でデータを可視化する．

因子分解も印象評価の有用な手段である．尺度のみに注目した因子分解 [B: Stoklasa 19] を適用した分析もあれば，その尺度を基本的な因子として価値因子，活動性因子，力量性因子の3因子 [B: Osgood 64] に分解した分析も知られている．本研究では尺度に対する因子分解はまだ適用していないが，これを適用することも今後可能である．

2.3 機械学習の訓練データ作成のためのアノテーション

様々なメタデータを表現するための方法として，アノテーションがある．例えば，機

機械学習の訓練データ作成のためのアノテーションでは、少数の作業者が多くの画像に対してタグを付与する。タグ付けの作業量が多いほど、面倒で退屈な作業になる。また、作業者の疲労に繋がり、回答の信頼性が減少する可能性がある。

アノテータは専門家・非専門家を問わない。しかし一般的に、機械学習の非専門家達によるアノテーションが実施される。近年はそれを分析し、訓練データの分布理解を目指す研究が活発に行われている。Mullerら [B: Muller 21] は、アノテータを機械学習の専門家に限定して実施した。専門家によるアノテーション設計や作業経験、品質問題など、堅実な説明を提供している。その中で、アノテーション設計における3つのデザインパターンを示している。

- 原理的デザイン：タグを定義し、アノテーションの適用手順を計画した上で実施。
- 反復的デザイン：アノテーション中にタグの再定義と改良を進めながら実施。
- 即興的デザイン：多くの時間と労力、人数を必要として議論しながら実施。

複数の設計パターンがあることから、アノテーション自体には決まったルールが存在しない。本研究では原理的デザインに則り、SD法を用いた印象評価をアノテーションとして実施した。

次に、ファッション画像のアノテーション研究に着目する。

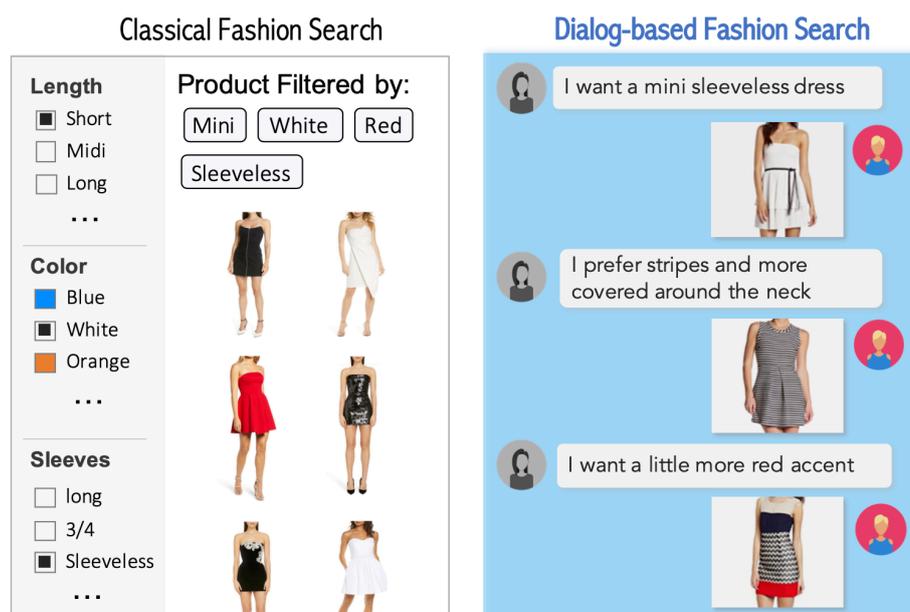


図 2.1: Wu らの可視化例 [B: Wu 21]. (左) 従来のファッション衣服に対するフィルタリング検索。(右) 自然言語表現を提供する対話的検索。

Wu ら [B: Wu 21] は画像に対してアノテータから収集したキャプションと、製品メタデータや属性ラベルを保持する大規模なデータセットである Fashion IQ を作成をした。Fashion IQ は図 2.1 にある通り、従来のフィルタリング検索ではなく、自然言語表現を提供する対話的検索の実現を目標としている。キャプションの収集として、クラウドソーシングを利用している。ユーザはオンラインショッピングのチャットウィンドウから、提供された検索候補と比較して、検索対象の視覚的特徴をショッピングアシスタントに伝えるため、対話的に自然言語表現を提供するという目標を割り当てられている。チャット形式の対話的な応答は、画像に対して詳細なメタデータを得ることが可能であるが、作業者の負担が大きくなる。本研究ではタグとして5つの形容詞対をあらかじめ定義し、作業者に提示して選択式にすることで、少しでも負担軽減を図っている。

2.4 決定木や階層構造を用いた可視化手法

印象評価結果を分析し、アノテーションとして印象タグを付与する手法として、決定木を用いることができる。決定木とは、木構造を用いて分類と回帰を行う機械学習である。木構造を上位階層から下位階層に辿ることで、各枝に相当する特徴量が各ノードの説明変数とその閾値のルールを通過する様子を観察できる。これによって分類・学習される工程を表現する。分類結果や学習結果だけではなくその過程を観察できることから、機械学習の説明責任の点で有効な可視化手法である。しかし、決定木を用いた可視化の観察や分析の難易度は高い。分析慣れしていないことや決定木を使用する頻度が少ないことで、観察する以前に、決定木自体の理解に時間がかかってしまう。決定木の使用前には事前説明が必須となってくるため、可読性の向上を期待することが多々ある。

ファジィ決定木を用いたSD法による印象分析 [B: Matsuo 10][B: 井上 09] 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。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 [B: 太田 17] 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。図 2.2 では、Neto ら [B: Neto 20] 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。

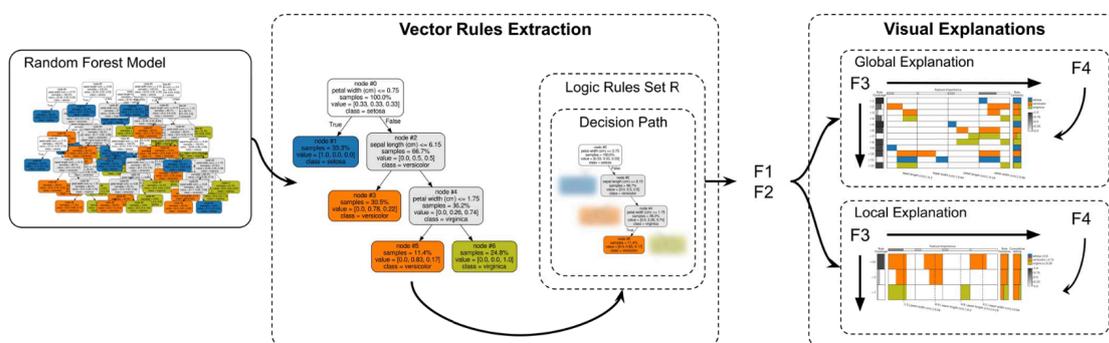


図 2.2: Neto らの可視化例 [B: Neto 20].

決定木と複数の分析画面を提示することで、決定木の可読性の向上が見込まれる。Mühlbacher ら [B: Mühlbacher 17] は、統計学の非専門家でも適切な決定木を効率的に選択できる可視化を提案している。枝切りや丸め処理を施すことで、注目して欲しい箇所を示し、効率的な分析を可能としている。しかし、実際のユーザ評価では、複数ある画面を1つずつ説明することによりかなり時間がかかっている。分析画面の提示が10エリアあるため、口頭での説明時間に2時間を要している。可視化画面の操作を実施する前の説明が長いと、操作中の作業者の集中力低下につながる。説明時間を少なくするためにも画面数を最小限に留めておく必要がある。本研究では説明資料を準備し、資料数削減のためにも、分析画面の提示は2エリアのみで実装した。

Eichner ら [B: Mühlbacher 17] は、時系列セグメンテーションのパラメータ依存性の視覚的分析を支援する手法を提示している。どのパラメータが高い影響力を持ち、どのセグメンテーション特性がパラメータの変更に対して高い感度を持つかを分析者が理解するのを助けることが目的である。左の画面には変化する階層的トピックを時間軸に

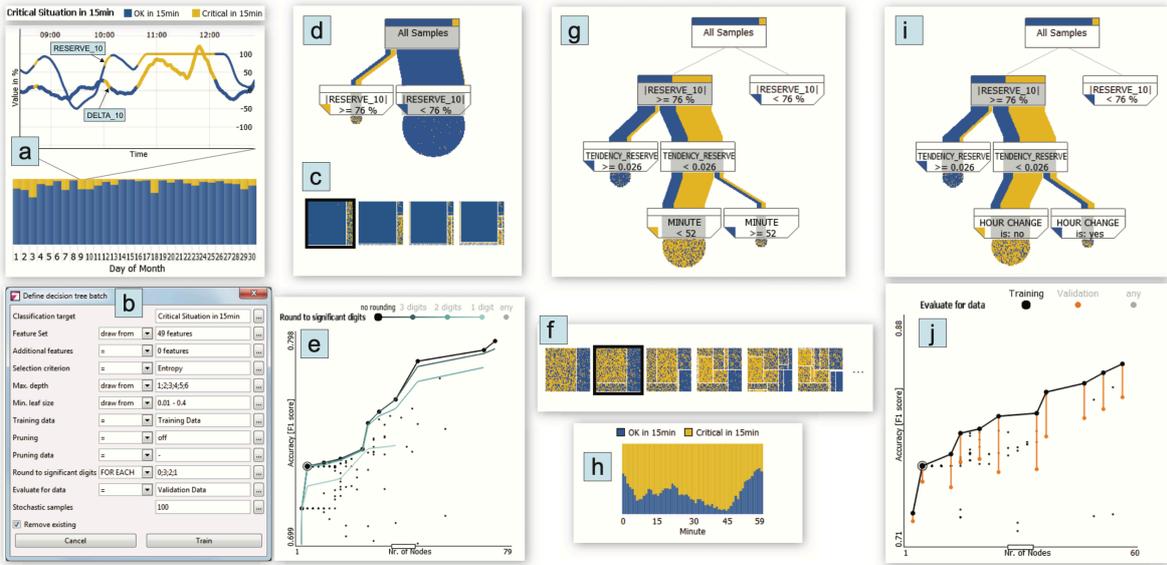


図 2.3: Mühlbacher らの可視化例 [B: Mühlbacher 17].

沿って説明するように設計・可視化し，右の画面では時系列セグメンテーションのパラメータ相関関係の可視化をしている．決定ノードの部分に色で分類されているパラメータ情報を表示しているが，細かい行列セルで難読であり，非専門家にとって理解するための説明が必須となる．我々の研究では，決定ノードを横軸目盛りとともに数値として表示させることで，非専門家にとっても親しみやすく，解読しやすい決定ノードとしている．

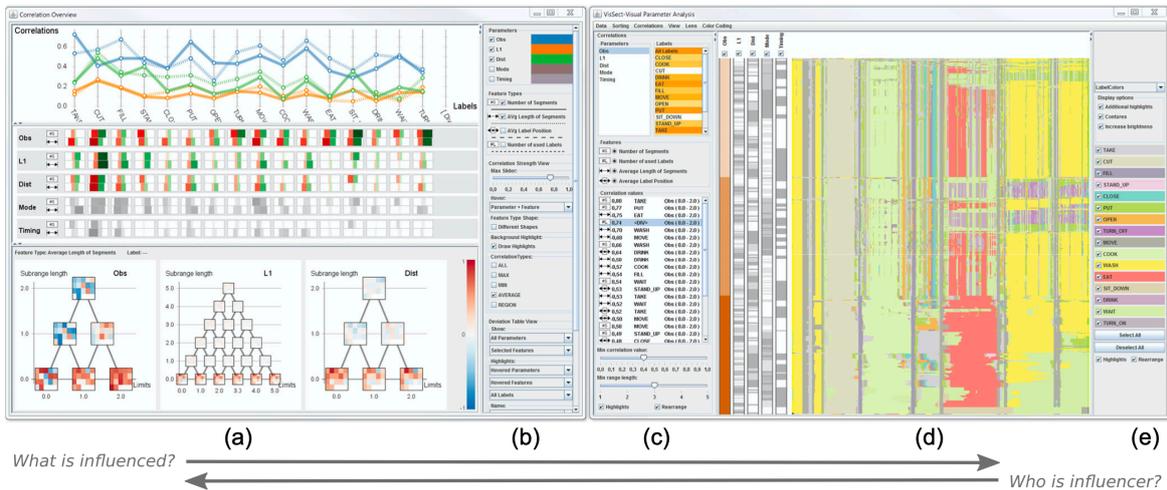


図 2.4: Eichner らの可視化例 [B: Eichner 20].

また，Cui ら [B: Mühlbacher 17] は，スキャンダルに伴う時系列テキスト分析を可能にした操作型の並行座標プロットを実装している．対話的に枝切りのパラメータを操作でき，ユーザの興味に沿ったテキスト分析ができる．本研究ではパラメータを固定して

いるが、今後は Cui らの枝切りアルゴリズムを適用することも視野に入れたい。

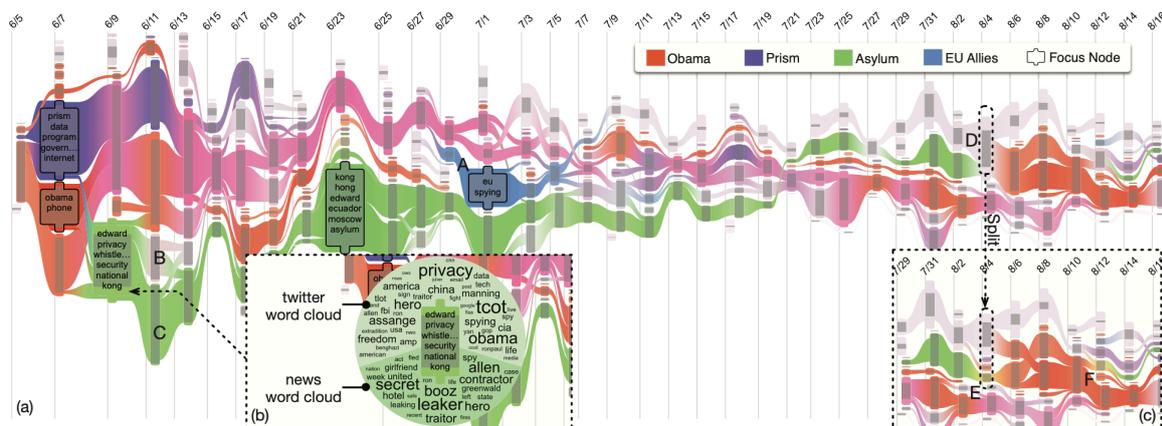


図 2.5: Cui らの可視化例 [B: Cui 16].

機械学習における畳み込みニューラルネットワーク (CNN) では、学習プロセスを可視化し、デバッグや行き詰まりの阻止、モデルのパフォーマンス向上のための潜在的な方向性の発見として、近年研究が盛んである。

Liu ら [B: Liu 16] は、ノードに画像特徴量の表示をしており、対話的操作により詳細を適宜表示させることでユーザのインタラクティブ性を高くしている。クリックによってエッジの強調や拡大・縮小ができる分、画像レイアウトの複雑性が高くなる。

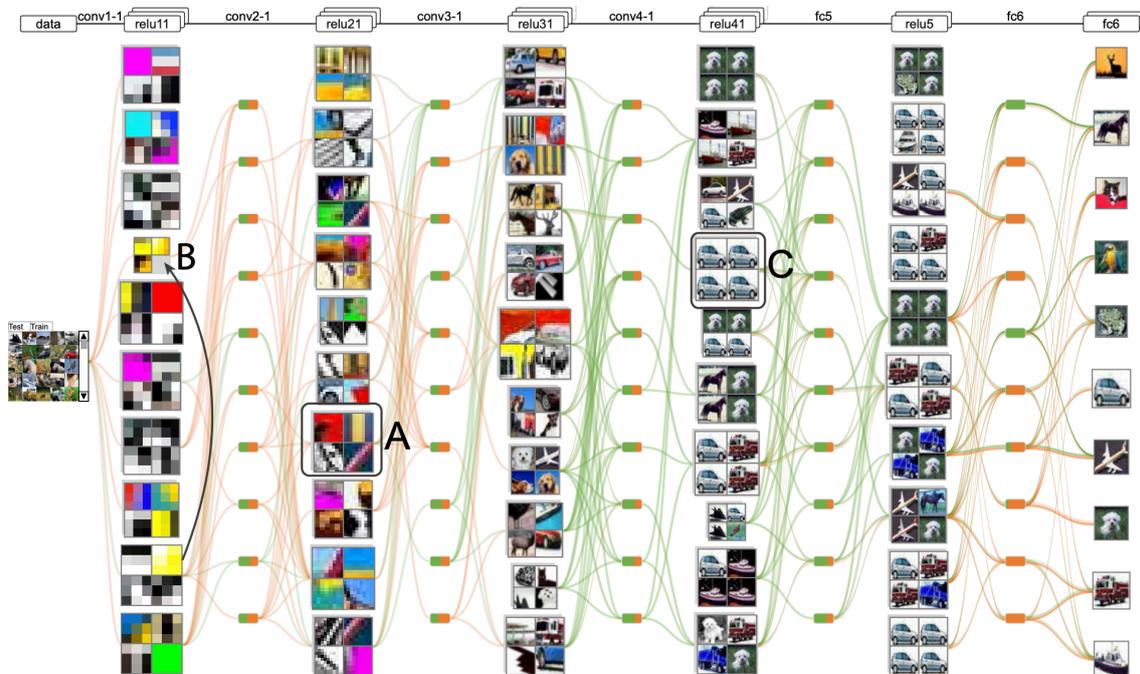


図 2.6: Liu らの可視化例 [B: Liu 16].

Jia ら [B: Jia 20] は、色分けした分類結果と、画像特徴を別の画面で表示することで

それぞれの特徴分析がし易くなっている。決定木の深さはユーザ操作によって折り畳み可能であるため、パラメータ変更も容易である。深さごとに画像表示が可能であり、誤分類される画像の推定過程を観察できる。

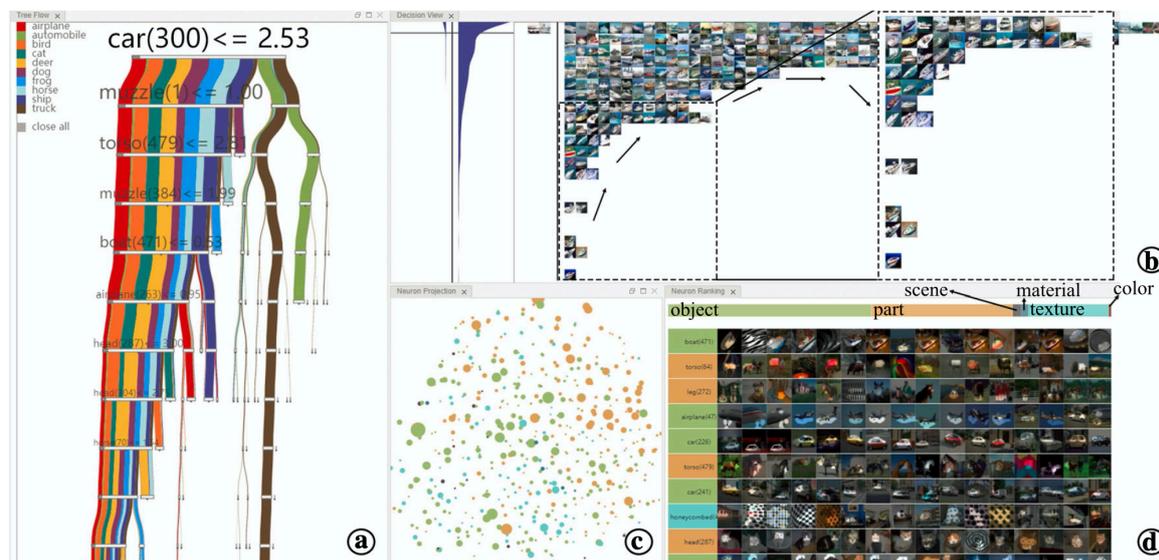


図 2.7: Jia らの可視化例 [B: Jia 20].

階層構造データの可視化に、Gou ら [B: Gou 20] は等高線を組み合わせた可視化を提案している。図 2.8 にある等高線密度マップはテストデータのカバー率を示しており、色が濃い領域ほどカバー率が低いことを表現している。等高線密度マップの上に対応して検出画像結果を表示することで、同エリア内で2つの分析結果をまとめて観察することができる。

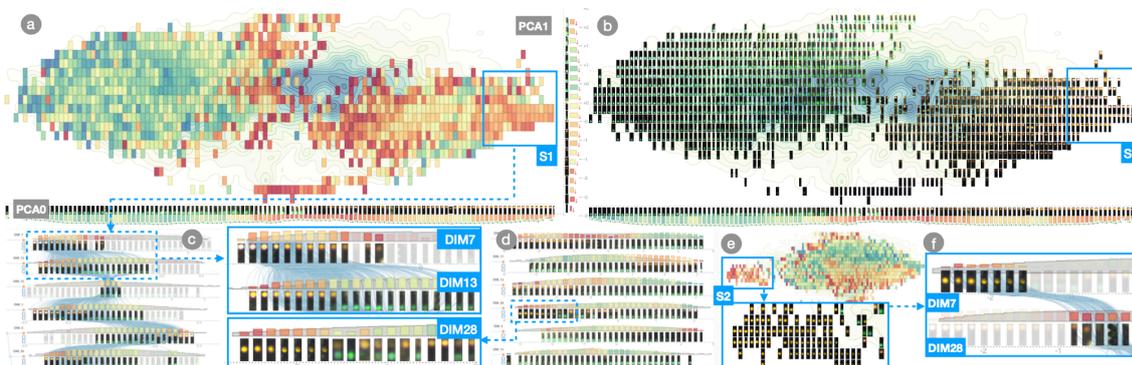


図 2.8: Gou らの可視化例 [B: Gou 20].

CNN 決定木の可視化は全体構造 (overview) が肝心である。overview を観察しながら、対話的操作により複数画面を連動させたり、詳細を適宜表示させることで決定木の理解

を容易にしている。本研究でも対話的操作を可能にし、主にダブルクリックと選択機能の2つのみを中心に使用する。本報告では、操作のしやすさと overview に対するユーザ評価の結果について説明する。

大量画像の階層型データ可視化手法に Image Browser[B: Gomi 10][B: Bederson 01]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 [B: Liu 18] 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。本研究においては画像特徴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。

2.5 むすび

本章では、提案手法の関連研究を紹介した。まず 2.2 節で印象評価結果の印象分析手法を紹介した。続いて 2.3 節で機械学習の訓練データ作成のための一般的なアノテーションについて紹介した。最後に 2.4 節で決定木や階層構造を用いた可視化手法を紹介した。以上の研究に基づき、印象の個人差を解消するために大規模な印象評価を実施する。そこから、ファジィ決定木を採用し、アノテーション作業の疲労軽減を図る。また、階層構造の可視化を応用し、決定木の可読性の向上を目指す。

第3章 提案手法

3.1 はじめに

本章では、本論文における提案手法について説明する。まず、3.2節では提案手法の処理手順についての概要を述べる。続いて3.3節でSD法を用いた大規模印象評価について、3.4節で可視化手法について説明する。3.5節では本手法におけるユーザインタフェースについて述べる。

3.2 処理手順の概要

2.2節でも述べたとおり、SD法は概念×回答者×尺度の三相を分析の対象としている。本研究ではまず、「尺度」にあたる形容詞対の選出をし、「概念」は画像であるとして、画像群にSD法を採用した大規模な印象評価を実施した。次に、印象評価結果を可視化システムにインプットさせて決定木生成と類似画像一覧表示をすることで、自動的に画像のタグ付け結果を可視化する。また、決定木から各画像の印象が策定される過程を観察でき、画像一覧からは印象策定結果を画像特徴から観察できる。

3.3 SD法を用いた大規模印象評価

3.3.1 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。ファッション・衣服におけるSD法属性について、多くの文献[B: 雙田 08][B: 松田 19]がある。堀尾ら[B: 堀尾 02]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、村山ら [B: 村山 08] 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 [B: 古川 17] 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。

3.3.2 パイロットテストと尺度決定

3.3.1 節で選択した因子から列挙した形容詞対を表 3.1 に示す。

表 3.1: 選定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
	フィットした	ルーズな	5.42
「正統性」	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、回答する項目数を増やしすぎることによって、作業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しう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、回答者(20代女性15名)から、表3.1にある15項目の形容詞対に対して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ユークリッド距離計算により算出した。ヒートマップを用いて距離行列を可視化した結果を図3.1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 3.1: 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3.1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3.2に示す。

表 3.2: 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

3.3.3 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いる。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 3.2 に示す。



図 3.2: 画像の前処理手順.

3.3.4 印象評価の収集と可視化データの構築

表 3.2 に示す形容詞対と 3.3.1 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 43 名、尺度を 5 項目として、5 段階評価で回答を収集した。43 人をランダムに 3 グループに分けることによって、1 人あたり 500 枚の画像の印象評価を回答している。これは回答者の作業量を減らすためである。したがって、画像 1 枚あたりの回答人数は 14~15 人である。内訳を表 3.3 に示す。

回答者には、図 3.3 にあるように、画像を 1 枚ずつランダムに表示させた。この時、回答者には自身の PC を使用してもらい、ブラウザは Chrome を使用するよう指定した。また、作業中は画面の明るさを変更しないように指示をした。これは画面の明るさが変化することで衣服画像に対する印象が変わってしまうことを防ぐためである。また、回答者には作業時間分の謝金を用意した。

¹<https://www.kaggle.com/dqmonn/zalando-store-crawl>

²<https://ja.clippingmagic.com/>

表 3.3: 印象評価データ

画像	女性着衣画像 1500 枚	
回答者	43 人 (全員 20 代 / 女性 37 人男性 6 人 / 日本国籍 35 人中国国籍 8 人)	
「色彩」尺度 1	暗い	明るい
「立体的」尺度 2	フィットした	ルーズな
「正統性」尺度 3	フォーマルな	カジュアルな
「穏健性」尺度 4	日常的な	非日常的な
「装飾性」尺度 5	シンプルな	ゴージャスな
評価段階	5 段階	



図 3.3: ユーザテストに用いた画面.

3.4 可視化手法

3.4.1 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。

- 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 3.4 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タ数を3に固定し、印象が「暗い」「明るい」「どちらでもない」の計3クラスが各クラスタ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。

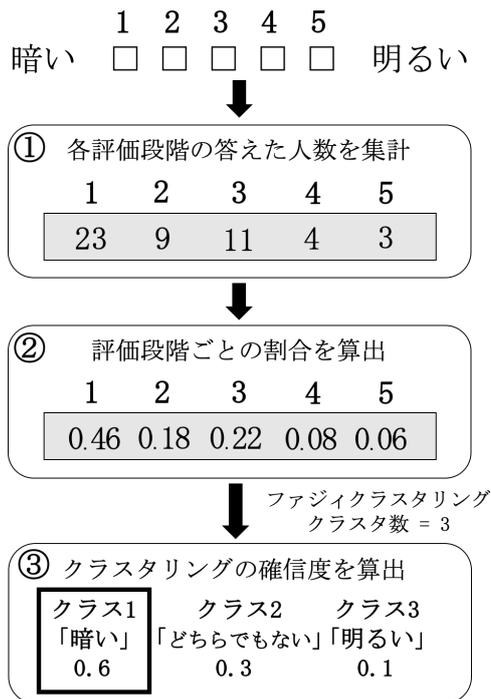
各画像に対して複数の参加者が回答することを前提にして、図 3.4①に示すように5段階評価の各値を選択した人数を集計し、続いて図 3.4②に示すように各値の人数比を算出する。この値から、Python の `fcmeans` ライブラリ³を用いて、図 3.4③に示すようにファジィクラスタリングの確信度を算出する。そして確信度が一番大きい値をクラスタリング結果として採用することで、印象推定を完了する。

続いて、図 3.4 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 `iparacoord`⁴を応用して決定木を表示する。図 3.5 に決定木の構造のイメージ図を示す。

ここでは、1つの枝は1画像を表しており、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。決定ノードは、分類条件と条件判定を受けた画像数(図 3.7①)を表し、各画像に対して範囲 $[0,1]$ で算出される確信度の詳細(図 3.7②)が表示される。分類条件には評価時のリッカート尺度5段階を使用している。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。色はクラスごとに異なり、色自体には意味を持たない。この可視化結果から、各画像が

³<https://pypi.org/project/fuzzy-c-means/>

⁴<https://github.com/intellygenta/InteractiveParallelCoordinates>



「暗い」印象
と策定された画像

図 3.4: 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.

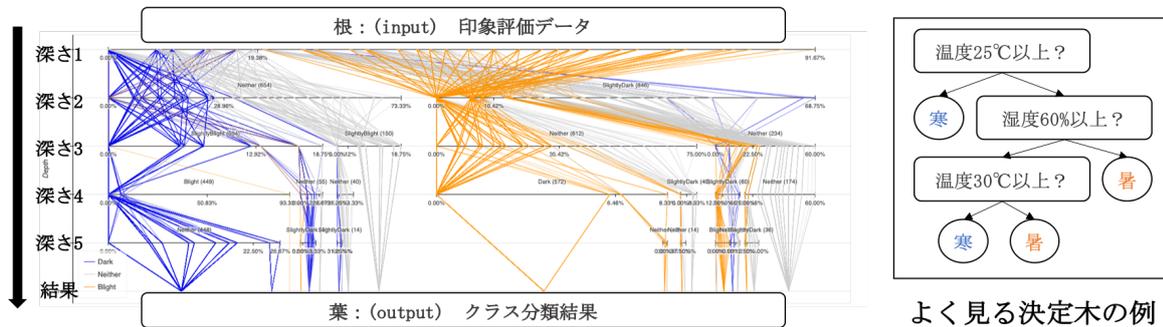


図 3.5: 決定木の構造図. (左) 本手法の決定木. (右) よく見る決定木の例.

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。

3.4.2 類似画像群の一覧表示

本節では画像と印象の関係を観察するために類似画像群の一覧表示を構築する手順を説明する。3.4.1節で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。本手法では、決定木のクラスタリング結果ごと及び深さごとに枠

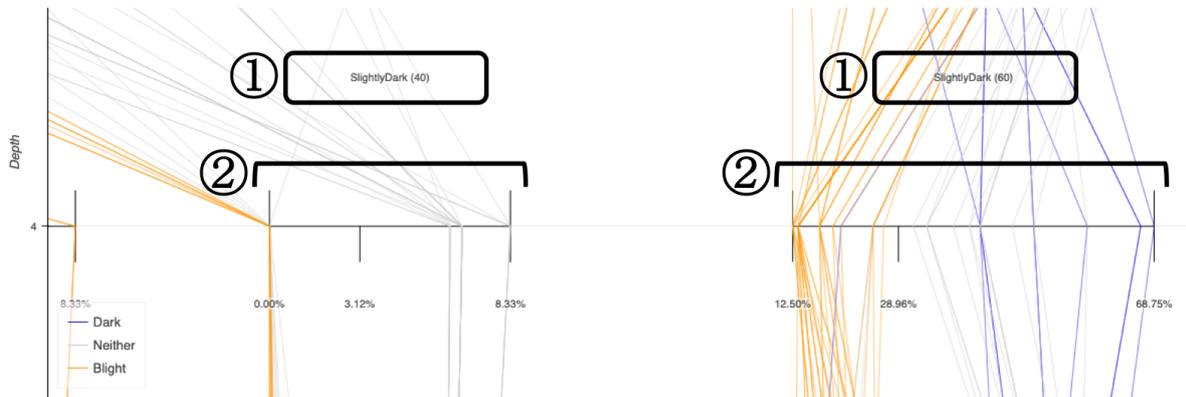


図 3.6: 決定ノードの詳細. ①分類条件 (条件判定を受けた画像数). ②確信度の詳細.

を生成し、木構造の葉ノードに対応する領域に画像群を表示する. さらに各枠に所属する画像群を一覧表示するため、画像の RGB 値を多次元ベクトルとし、これに主成分分析 (PCA) を適用して各画像を 2 次元空間に配置する. これにより、色分布が類似する画像を近くに配置する.

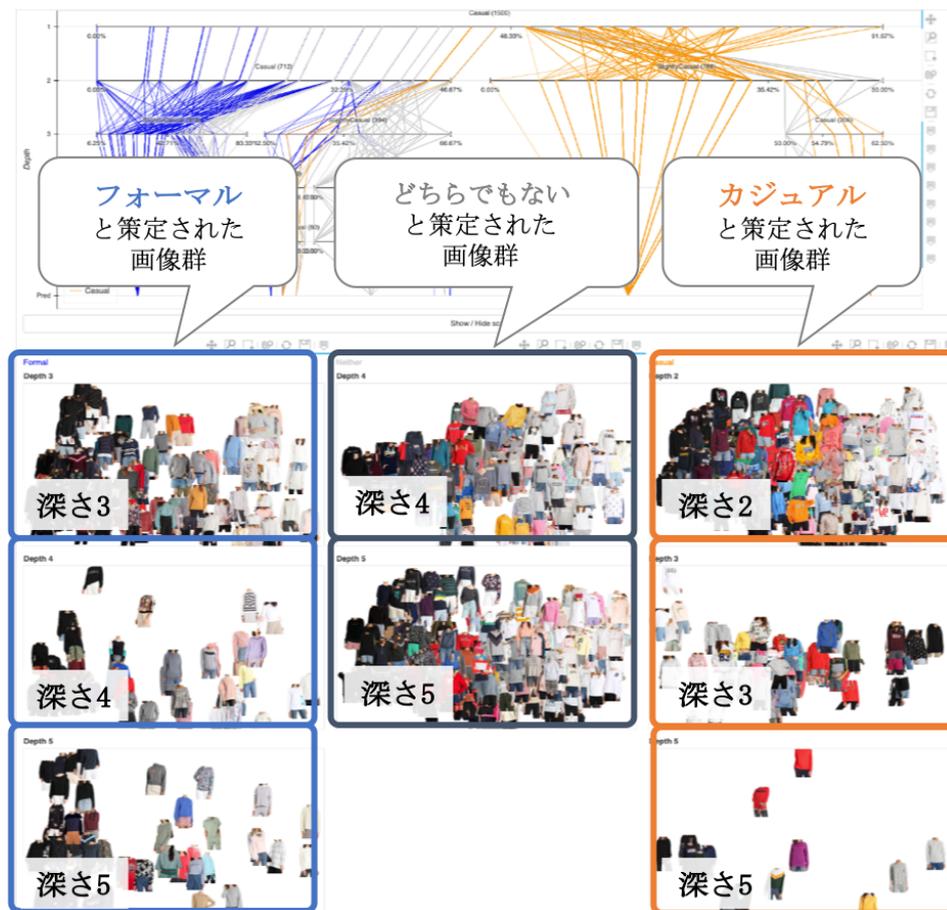


図 3.7: 類似画像群の一覧表示の詳細.

3.5 ユーザインタフェース

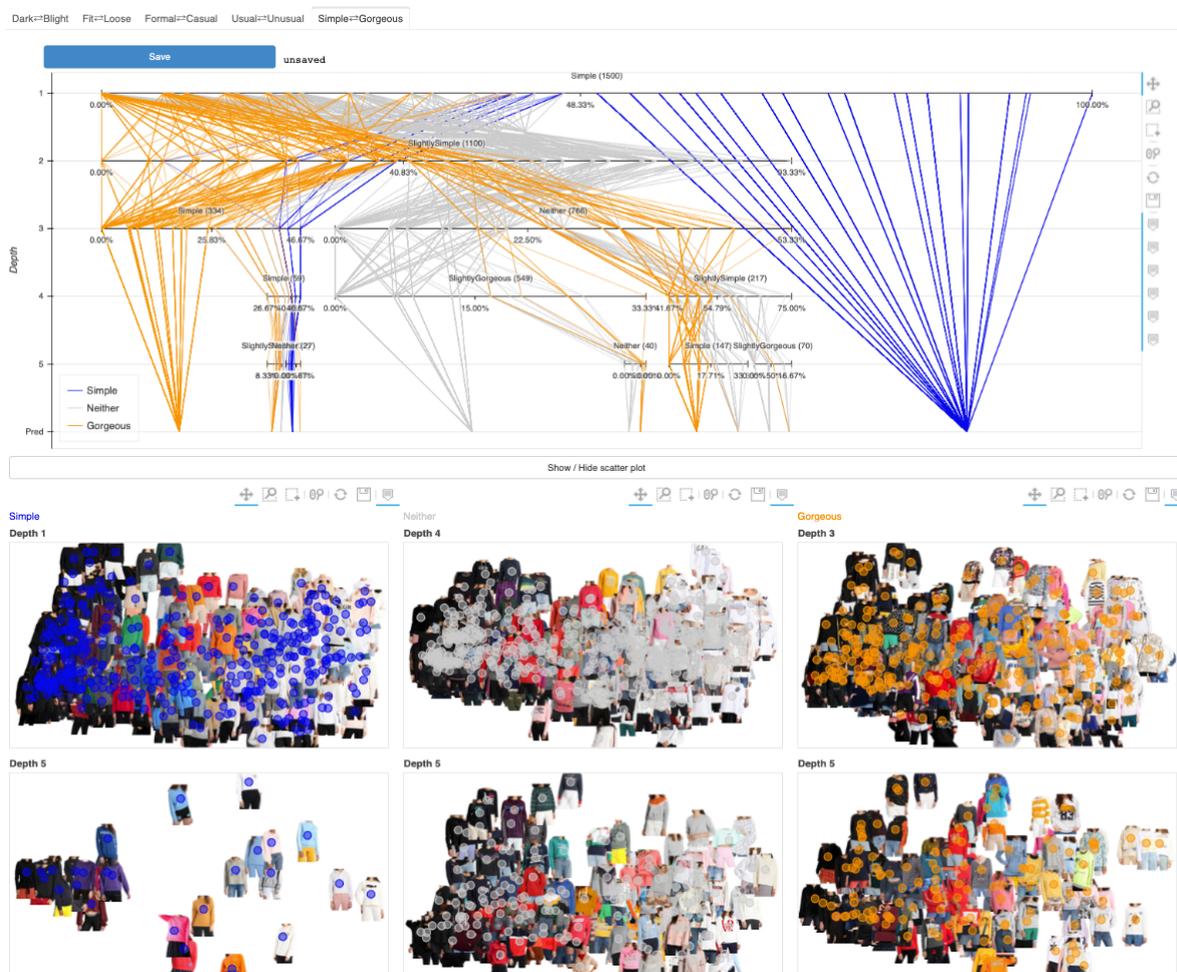


図 3.8: 本手法を操作するためのユーザインタフェース.

本手法を操作するためのユーザインタフェースを図 3.8 に示す. 可視化とユーザ操作によって達成したいことはアノテーションの半自動化である. 決定木により画像 1500 枚がすでに分類済みで, この分類可視化結果をもとに人の手で根拠のある再分類ができる. クラスの再分類は画像中央にある点描画をダブルクリックすることででき, クラス変更は図 3.9 にあるように, 点描画の色が変わることによって確認できる. この点描画は決定木と画像群一覧表示の間にあるボタンによって表示・非表示の操作ができる. 画像特徴をよく観察したい時にユーザはこのボタンを活用する.

操作後は, 「Save」の青いボタンをクリックすると, 再分類結果のダウンロードが可能であり, JSON 形式で再分類結果を取得できる.

ウィンドウ上部にあるタブを切り替えることで尺度ごとの可視化結果を観察すること

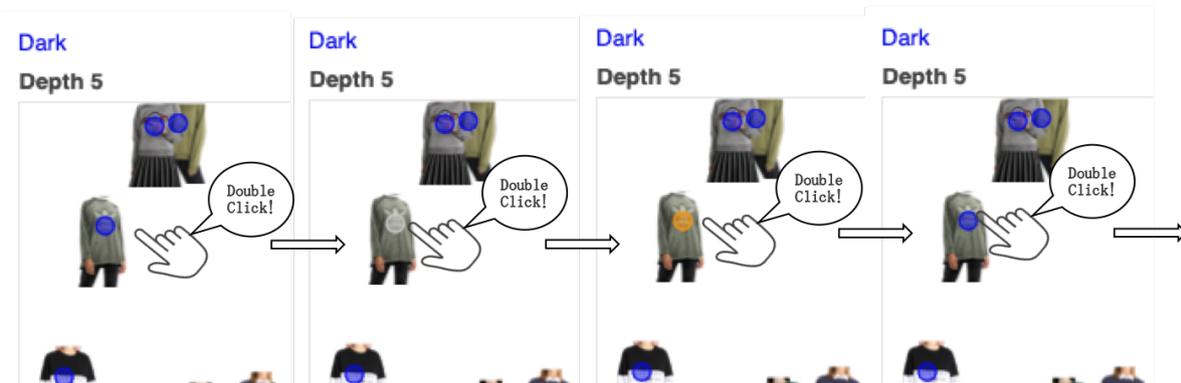


図 3.9: ダブルクリックで人の手による再分類方法.

ができる。平行移動、拡大縮小、再描画、ホバー機能やフィルタリング機能は、ウィンドウ右側のアイコンを選択することで使用でき、これら全てはどのプロットエリアでも操作可能である。

ホバー機能は、図 3.10 にあるように、決定木の枝部分と、画像の中心にある点描画に操作マウスを乗せることで、各画像が拡大されたものと初期分類結果のクラス色を点で表示している。

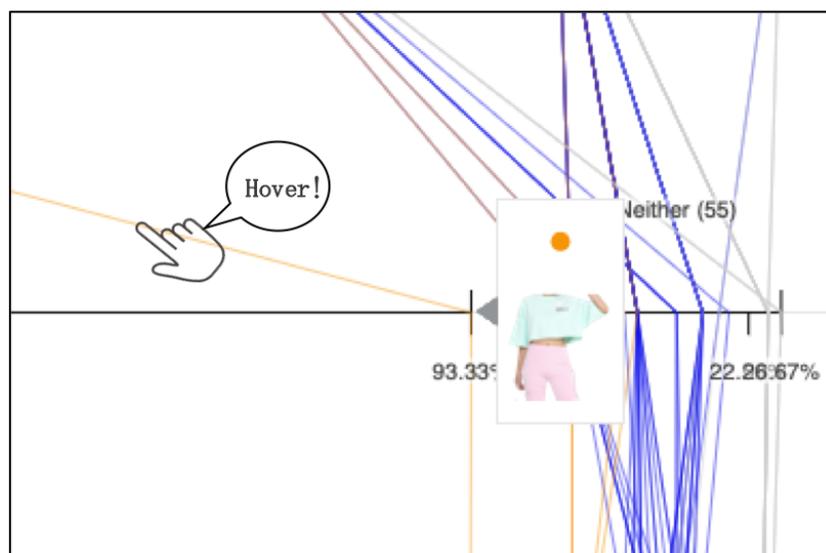


図 3.10: ホバー機能の詳細.

フィルタリング機能は、例えば、図 3.11(左)で深さ 2 の決定ノードをドラッグアンドドロップで選択すると、図 3.11(右)のように、選択された枠の中の決定ノードを通過する枝と、その枝に該当する画像が画像群一覧表示と連動してフィルタリング表示される。これを活用することによって、ユーザは気になる決定ノードや枝と画像特徴を見比べたり、分析したいクラスのみを絞ってクラス分類の傾向を観察することができる。

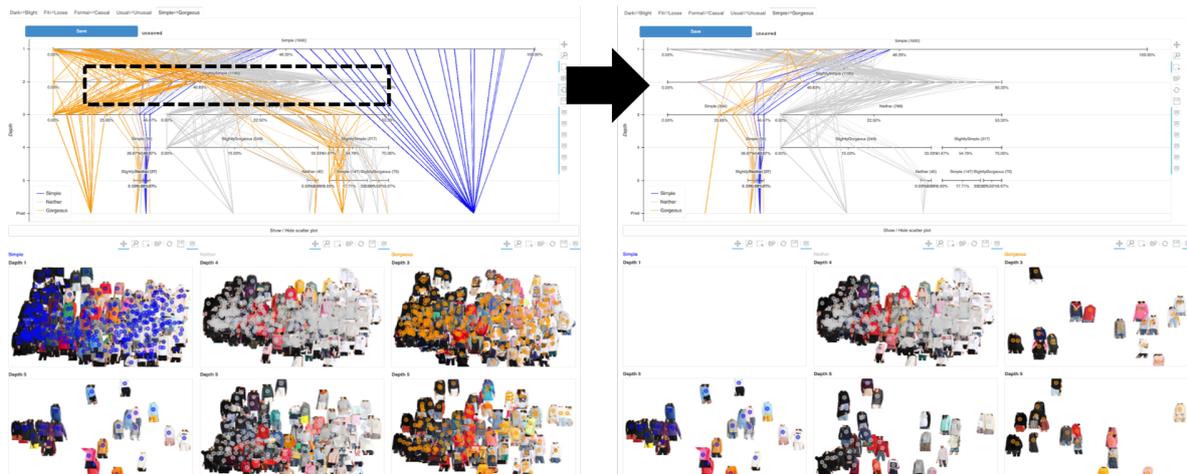


図 3.11: フィルタリング機能の詳細。(左) フィルタリング前。(右) フィルタリング後。

3.6 むすび

本章では、本論文の提案手法を説明し、各手順の詳細を述べた。3.2節では、提案手法の処理手順についての概要を述べた。3.3節では、SD法を用いた大規模印象評価と可視化データの構築方法について述べた。3.4節では、ファジィ決定木と類似画像群の一覧表示をする可視化手法について述べた。3.5節では、本手法のユーザインタフェースについて述べた。

第4章 実行例

4.1 はじめに

本章では、本手法による可視化の実行例を紹介する。まず4.2節で使用するデータについて説明する。続いて4.3節では、決定木や類似画像群の一覧表示に関しての実行結果を紹介する。そして4.4節では、4.2節で説明したデータを使用した評価実験の概要と結果を紹介する。最後に4.5節で全体の考察を述べる。

4.2 使用するデータ

本事例では表3.3に示す印象評価データを使用し、可視化している。

我々は本手法をPython 3.9.5の可視化ライブラリBokeh 2.4.0¹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。

また、3.4.1節にあるように、ファジィ決定木を構築する際の分類クラス数は3つで、深さは最大5までに固定している。

4.3 可視化結果

本節では、決定木のみ可視化結果、類似画像一覧表示のみ可視化結果、決定木と類似画像一覧表示のフィルタリング機能を用いた可視化結果を示す。

4.3.1 決定木から得られる知見

図4.1から図4.5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、

¹<https://bokeh.org/>

3クラス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 3.3 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。しかし、図 4.5 「非日常的な-日常的な」と図 4.4 「ゴージャス-シンプル」の決定木に注目すると形が似ていることがわかる。パイロットテストによって印象因子ごとに項目を決定したが、この2つの尺度に関しては衣服においてやや同義であることが観察できた。

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 4.3 における「カジュアル」、図 4.4 における「シンプル」と図 4.5 における「日常的な」に関しては、決定木の深さが浅いところでクラスが策定されていることから、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。

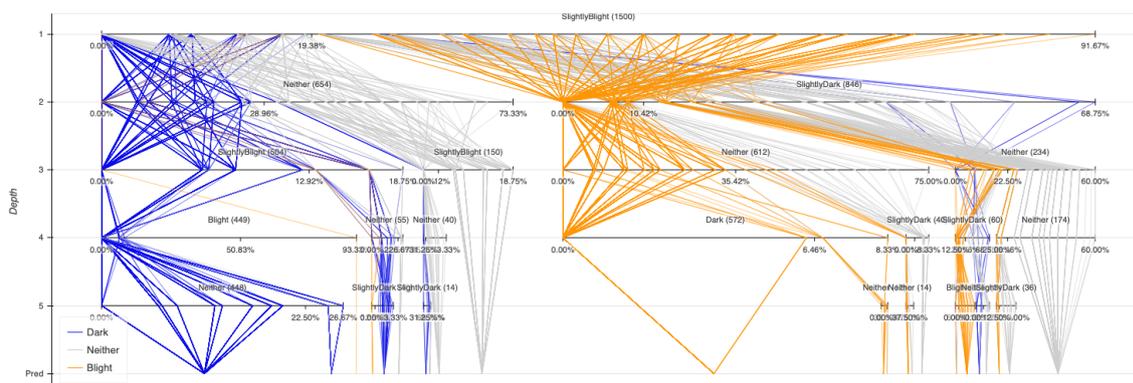


図 4.1: 尺度「暗い-明るい」のファジィ決定木.

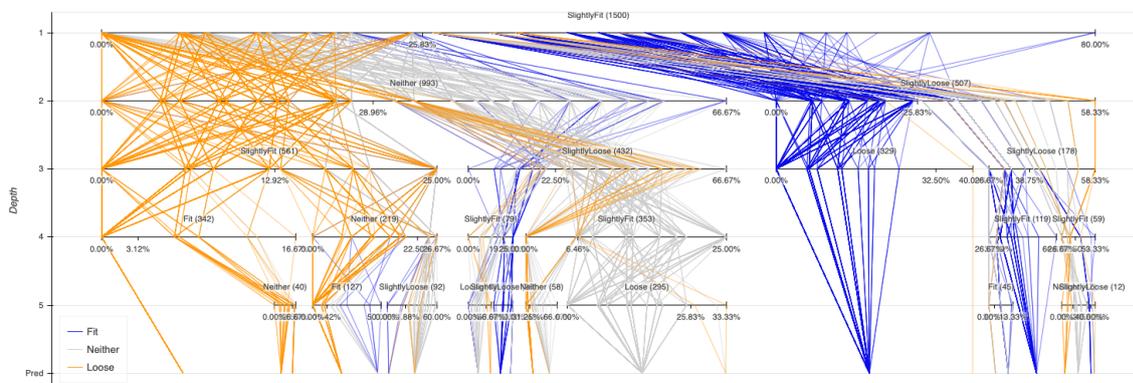


図 4.2: 尺度「フィット-ルーズ」のファジィ決定木.

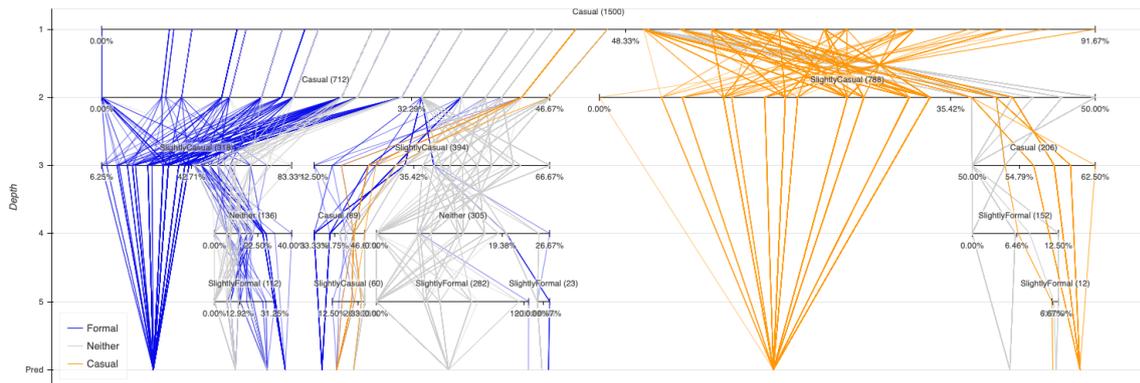


図 4.3: 尺度「フォーマル-カジュアル」のファジィ決定木.

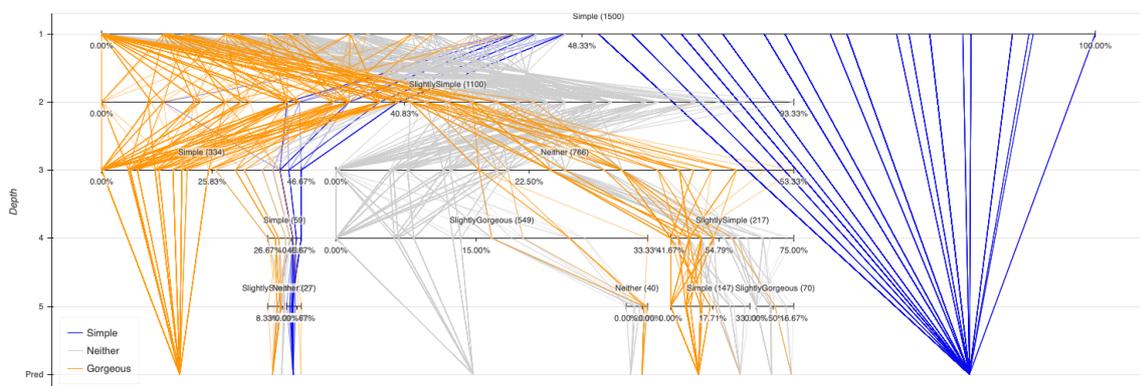


図 4.4: 尺度「シンプル-ゴージャス」のファジィ決定木.

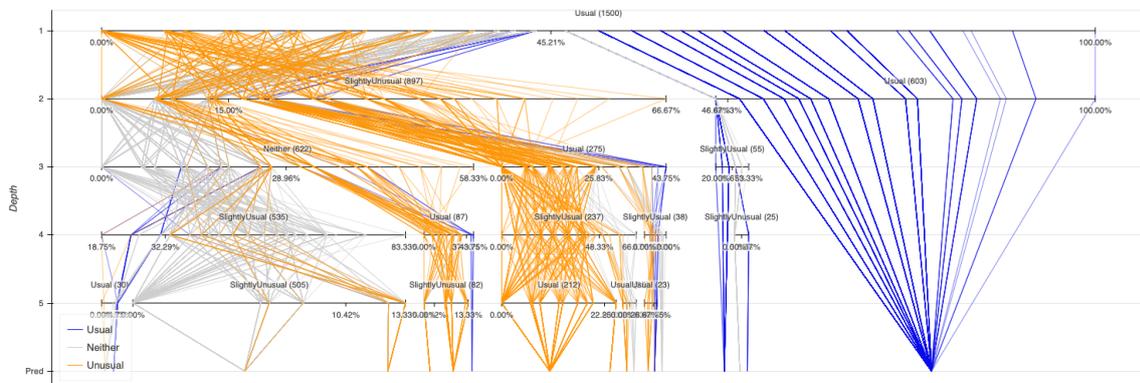


図 4.5: 尺度「日常的な-非日常的な」のファジィ決定木.

4.3.2 類似画像一覧表示から得られる知見

図 4.6 は尺度「暗い-明るい」に対する類似画像群の一覧表示例である。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 4.6: 尺度「暗い-明るい」の類似画像群一覧表示. (左)「暗い」と深さ 5 で策定された画像群. (中央)「どちらでもない」と深さ 3, 4, 5 で策定された画像群. (右)「明るい」と深さ 4, 5 で策定された画像群.

4.3.3 決定木と類似画像一覧表示のフィルタリング機能から得られる知見

図 4.7(左) は尺度「フォーマル-カジュアル」の可視化結果である. 決定木から対となるはずの「フォーマル」と「カジュアル」の混合分類が見られたので, その部分をフィルタリングすると, 図 4.7(右) にあるようなフィルタリングされた画像群を観察できる. 図 4.7(右) 中の赤矢印で示している衣服で, 同じ深さで同じ色/形/デザインが似ている衣服が存在した. 2つの画像群に所属する衣服画像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. 衣服の印象に限定するために顔削除や背景削除して印象評価を実施したが, 着衣画像であればある程度, モデルの髪形や体格が影響する可能性が示唆される. 尺度「フォーマル-カジュアル」に関わらず, 他の尺度でも決定木において混合分類が見られるが, それらはどれも深さが深いところでクラス策定されている. このことから混合分類が見られる画像群はクラス分類の確信度が低いものであり, ファジィクラスタリングの最終的なクラス分類の判断基準は正しく働いていることがわかる.

4.4.1 評価手順

評価手順としてまず、評価者に説明資料を読んでもらってから本システムを用いたタグ付けを開始してもらった。本システム²はリモートサーバに置かれているため、評価者には URL からアクセスして作業を開始してもらう。ブラウザは Chrome で固定し、画面の明るさを作業中に変更しないように評価者に指示をした。その際に、画面録画による操作ログを残してもらい、さらに分類結果をダウンロードしてもらった。最後に可視化や操作に関する評価項目に答えてもらい、評価者には作業時間分の謝金を用意した。評価項目は以下の 4 つであり、これら全てに関してコメントや感想を自由記述として回答してもらった。これらの手順には特に制限時間は設けていない。

- 決定木について。
- 類似画像一覧について。
- 画面操作について。
- タグ付け作業について。

4.4.2 画面録画による操作ログによる評価結果

画面録画による操作ログから観察できたことを述べる。どのグループも、1 尺度にかかる作業は約 10～20 分で、5 尺度あるため大体その 5 倍の約 50～100 分ほど作業に時間がかかっていた。作業時間平均の詳細を表 4.1 に示す。

表 4.1: 本システムを用いた再分類作業にかかった時間の平均

	ML グループ	VIS グループ	NONE グループ
【色彩】 暗い-明るい	15m11s	15m14s	15m00s
【立体感】 フィット-ルーズ	14m14s	17m52s	17m03s
【正統性】 フォーマル-カジュアル	14m55s	15m04s	12m37s
【穏健性】 日常的な-非日常的な	14m32s	16m18s	15m53s
【装飾性】 シンプル-ゴージャス	11m52s	19m12s	16m55s

ML グループは機械学習や予測モデルの理解があるからか、決定条件の確信度が低いものを中心にクラス分類する傾向があった。また、フィルタリング機能を多用しており、

²<https://user-test-imp.sakura.ne.jp/labo/evaluation/en/decition.html>

決定木による分類過程と画像との関係性を注意深く観察している様子が見られた。深さごとにクラス分類された画像特徴のみの観察でも再分類可能ではあるが、初期状態のクラス分類結果が同じで、かつ、画像特徴があまりにも似ている/異なると、再分類の判断に根拠を持たすことはできず、さらに作業が滞ってしまう可能性がある。決定木と画像群一覧表示の連動操作により、画像特徴に加えて、決定ノードに表示される確信度の詳細が再分類の判断の根拠になる。これはフィルタリング機能を用いることによって分析が容易であるため、MLグループは本研究の狙い通りに決定木を参照してくれた。

VISグループは基準を決めてから再分類する傾向があり、これはMLグループとNONEグループと大きく異なる動きであった。例えば、図4.8(上)の赤矢印の流れは、MLグループとNONEグループの再分類作業の流れの傾向を表している。これら2つのグループは左上から右下へと流れるように再分類作業していたが、VSグループは図4.8(下)の赤矢印の流れのように、「どちらでもない」から再分類を開始し、全体の画像分布を何度も眺めて、自分の基準をブレないように心がけている様子が観察できた。全体像から分析に入るといって、普段から分析慣れしている可視化の専門家らしい操作ログ結果が見れた。このことから、本システムの全体像の見やすさによって、個人の主観が首尾一貫するような再分類作業を可能としている。

NONEグループは決定木をあまり使用しておらず、操作は平行移動と拡大縮小のみの作業が目立った。このことから、操作や分析慣れしていないユーザにとって、決定木の理解や、インタラクティブな操作による分析が難しかったことが観察できる。

4.4.3 再分類結果による評価結果

評価者には作業後、再分類結果をダウンロードしてもらった。表4.2は、1500枚の自動分類された画像が作業者によって再分類された画像数の平均である。再分類された画像は最大で平均93枚であり、約100枚以内の再分類操作で作業が完結していることがわかる。決定木による自動分類の結果が再分類の支援に優位に働いていることが分析できる。

表4.3から表4.7は、各尺度の再分類された画像数の移動平均を示す。

表4.3より、尺度「暗い-明るい」は、例えば「暗い→明るい」や「明るい→暗い」といった2クラスを跨ぐクラス変更が発生した画像数が、どのグループもほぼ0である。このことから、尺度「暗い-明るい」では、自動策定結果がそのまま使用されているこ

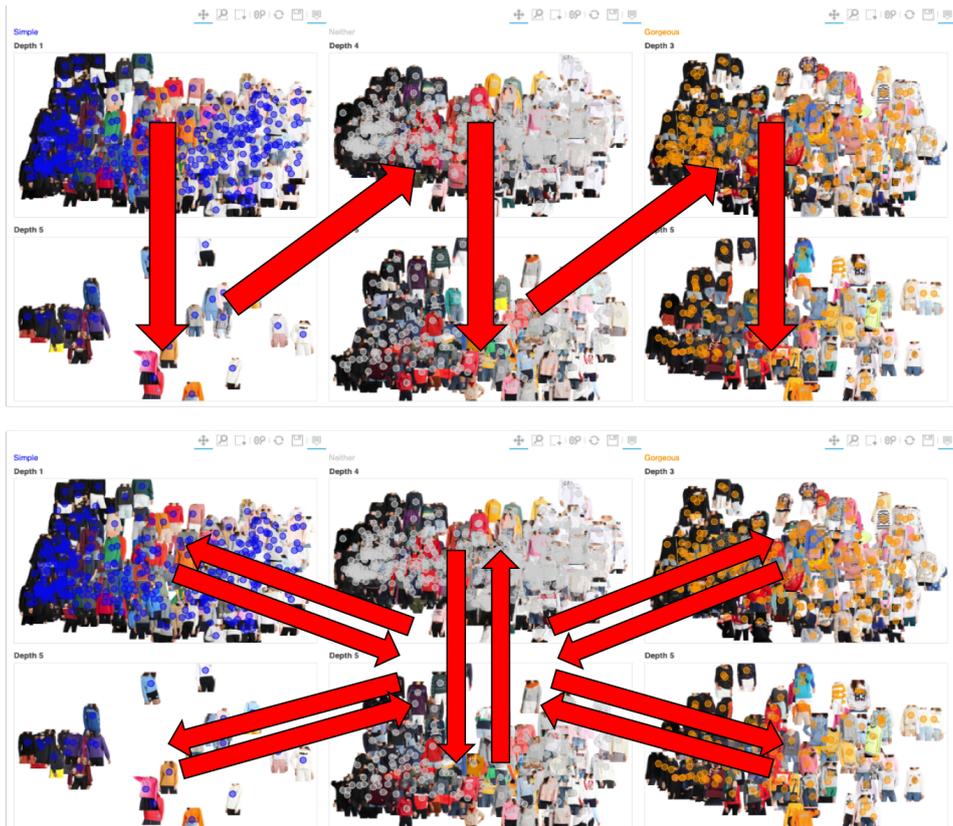


図 4.8: 画面録画の操作ログから観察できる再分類作業の流れの傾向. (上)MLグループとNONEグループ. (下)VISグループ.

表 4.2: 決定木によって分類済みの計 1500 枚の画像から再分類された画像数の平均

	ML グループ	VIS グループ	NONE グループ
【色彩】 暗い-明るい	55.98	26.65	47.99
【立体感】 フィット-ルーズ	27.65	21.99	67.98
【正統性】 フォーマル-カジュアル	56.65	26.66	93.31
【穏健性】 日常的な-非日常的な	80.97	37.65	93.99
【装飾性】 シンプル-ゴージャス	68.29	38.31	83.32

とがわかる. また, 表 4.4 における尺度「フィット-ルーズ」に関しては, MLグループとVISグループのクラス分類した画像数が少ない. 分類作業前のクラス策定結果が十分であり, クラス分類する必要性がなかったことが示唆される. これも自動策定結果がそのまま使用されている一例である.

表 4.5 から表 4.7 では, 一定方向のクラス分類の流れが目立つことが観察できる. 尺度「日常的な-非日常的な」では「どちらでもない→日常的な」の流れ, 尺度「フォーマル-カジュアル」では「フォーマル→どちらでもない→カジュアル」の流れ, 尺度「シンプル-ゴージャス」では, 「ゴージャス→どちらでもない→シンプル」の流れが目立つ.

表 4.3: 尺度「暗い-明るい」再分類された画像数の移動平均。

暗い			
	ML グループ	VIS グループ	NONE グループ
分類作業前	499		
分類作業後	492.33	512	489.66
暗い→暗い	478.33	497.33	484.33
暗い→どちらでもない	19	1.66	14.66
暗い→明るい	1.66	0	0
どちらでもない			
	ML グループ	VIS グループ	NONE グループ
分類作業前	376		
分類作業後	369	363.33	387.33
どちらでもない→暗い	14	14.66	5.33
どちらでもない→どちらでもない	345.33	356.33	357.66
どちらでもない→明るい	16.66	5	13
明るい			
	ML グループ	VIS グループ	NONE グループ
分類作業前	625		
分類作業後	638.66	624.66	623
明るい→暗い	0	0	0
明るい→どちらでもない	4.66	5.33	15
明るい→明るい	620.33	619.66	610

4.3.1 節では、決定木の可視化から得られる知見で、比較的「日常的な」と「カジュアル」と「シンプル」に関して確信度が高い衣服画像が多く含まれていることがわかる。したがって、再分類される流れと確信度の高いクラスには相関がある画像群であることが観察できる。

4.4.4 操作・システムに関する評価結果

作業後、評価者に自由記述のコメントやアンケートに回答してもらった。

まず、決定木に関しては、理解に時間がかかったため、可読性の向上を期待する声は NONE グループには多かった。その一方で ML グループと VIS グループは、フィルタリング機能を多用し、類似画像一覧と並行して観察することで理解がしやすかったとの声があった。決定木のみを使用であれば、専門家・非専門家にかかわらず可読性は低いことは予想していたため、類似画像群一覧表示とフィルタリング機能による補助が可読性の向上に有効に働いていることが示された。

次に、類似画像一覧に関しては、尺度「暗い-明るい」以外はうまくいかせなかったという意見があったため、RGB 値以外でも次元削減する必要性を感じた。また、尺度

表 4.4: 尺度「フィット-ルーズ」再分類された画像数の移動平均.

フィット			
	ML グループ	VIS グループ	NONE グループ
分類作業前	477		
分類作業後	474.33	480	458.33
フィット→フィット	467	472.66	448
フィット→どちらでもない	6.66	3	28.33
フィット→ルーズ	3.33	1.33	0.66
どちらでもない			
	ML グループ	VIS グループ	NONE グループ
分類作業前	543		
分類作業後	553	541	579.66
どちらでもない→フィット	2.33	6.33	8.33
どちらでもない→どちらでもない	538.33	531.66	528.66
どちらでもない→ルーズ	2.33	5	6
ルーズ			
	ML グループ	VIS グループ	NONE グループ
分類作業前	480		
分類作業後	472.66	479	462
ルーズ→フィット	5	1	2
ルーズ→どちらでもない	8	6.33	22.66
ルーズ→ルーズ	467	472.66	455.33

「ゴージャス-シンプル」では、衣服の生地感やデザイン・ロゴといった画像の細部まで注目する必要があったため、何度も拡大画像を必要とした。尺度ごとによって類似画像一覧の役目に一貫性がないことは、評価者には操作に慣れるまでの時間を必要とさせてしまう。ホバー機能による詳細情報表示の再設計や、画像群の特徴の見方を変更できるような新機能追加や改良の余地があると感じた。

画面操作に関しては、NONE グループでは、操作が少し難しいと感じたという意見が挙がった。普段から分析慣れしていないユーザにとっては、システムで作業する以前に、操作の理解に時間がかかったことがわかる。しかし、総じて全グループからは、インタラクティブな操作が豊富で、ダブルクリック操作でのクラス変更が便利だったとの声が多くあった。

最後に、タグ付け作業に関しては、画像特徴のみの再分類作業ではなく、決定ノードによる確信度の詳細表示を再分類作業の判断基準に加わえることができたため、タグ付け精度が上がると感じた、と回答してくれた。また、可視化により分類基準のブレはなかった、分類する判別要素が人によって違うことを実感した、などの評価をもらい、印象のアノテーションを支援するという意味合いでは、この可視化システムは全体的に効果があったと思われる。

表 4.5: 尺度「日常的な-非日常的な」再分類された画像数の移動平均.

日常的な			
	ML グループ	VIS グループ	NONE グループ
分類作業前	612		
分類作業後	663.33	629	670.33
日常的な→日常的な	605.66	607.33	606
日常的な→どちらでもない	4.66	4.33	5
日常的な→非日常的な	1.66	0.33	1
どちらでもない			
	ML グループ	VIS グループ	NONE グループ
分類作業前	556		
分類作業後	526.66	541.66	496.33
どちらでもない→日常的な	30.33	21.33	60.33
どちらでもない→どちらでもない	515.33	530.33	481.66
どちらでもない→非日常的な	10.33	4.33	14
非日常的な			
	ML グループ	VIS グループ	NONE グループ
分類作業前	332		
分類作業後	310	329.33	333.33
非日常的な→日常的な	27.33	0.33	4
非日常的な→どちらでもない	6.66	7	9.66
非日常的な→非日常的な	298	324.66	318.33

4.5 考察

以上の結果から、再分類された画像数の平均より、アノテーション作業の効率化を図れた。また、決定木からはクラス分類の偏りを発見できたり、尺度の適切さを図ることができ、類似画像一覧表示からは画像群全体の特徴を観察できた。決定木と類似画像一覧表示のフィルタリング機能により、連動して詳細を表示させることで、決定木の可読性が向上し、決定ノードにおける確信度と画像特徴の傾向が見れた。さらに、本システムの可視化によって、個人の主観によらない印象のタグ付けを可能とし、決定木の可視化の必要性を示した。一方で、分析慣れしていないユーザにとっては本システムの操作・理解に時間を要する。アノテータは専門家・非専門家を問わないため、誰もが分析・操作できるような、さらに可読性の高いシステムへと改良していく必要がある。

4.6 むすび

本章では、本手法による可視化の実行例を紹介した。4.2節で使用するデータについて説明し、4.3節で実行結果を紹介した。4.4節で評価実験の概要と結果を紹介し、4.5節で全体の考察を述べた。

表 4.6: 尺度「フォーマル-カジュアル」再分類された画像数の移動平均.

フォーマル			
	ML グループ	VIS グループ	NONE グループ
分類作業前	218		
分類作業後	252	279.66	270.66
フォーマル→フォーマル	249.33	278.66	257.33
フォーマル→どちらでもない	10.66	6	18.66
フォーマル→カジュアル	25	0.33	9
どちらでもない			
	ML グループ	VIS グループ	NONE グループ
分類作業前	563		
分類作業後	560.66	550.66	541.33
どちらでもない→フォーマル	1.33	1	12.66
どちらでもない→どちらでもない	546.66	543.66	510.33
どちらでもない→カジュアル	15	18.33	40
カジュアル			
	ML グループ	VIS グループ	NONE グループ
分類作業前	652		
分類作業後	687.33	669.66	688
カジュアル→フォーマル	1.33	0	0.66
カジュアル→どちらでもない	3.33	1	12.33
カジュアル→カジュアル	647.33	651	639

表 4.7: 尺度「シンプル-ゴージャス」再分類された画像数の移動平均.

シンプル			
	ML グループ	VIS グループ	NONE グループ
分類作業前	426		
分類作業後	462.33	437	432.66
シンプル→シンプル	419.66	420	405.66
シンプル→どちらでもない	5	5.66	19
シンプル→ゴージャス	1.33	0.33	1.33
どちらでもない			
	ML グループ	VIS グループ	NONE グループ
分類作業前	679		
分類作業後	660	669.66	683.33
どちらでもない→シンプル	15.33	15.66	24
どちらでもない→どちらでもない	649.66	656	641.66
どちらでもない→ゴージャス	14	7.33	13.33
ゴージャス			
	ML グループ	VIS グループ	NONE グループ
分類作業前	395		
分類作業後	377.66	393.33	384
ゴージャス→シンプル	27.33	1.33	3
ゴージャス→どちらでもない	5.33	8	22.66
ゴージャス→ゴージャス	362.33	385.66	369.33

第5章 まとめと今後の課題

5.1 まとめ

本節では、本論文で述べた内容を列挙し、内容を総括する。

第1章では、研究背景として印象評価結果のアノテーションに関する可視化の近年の動向を述べ、その問題点や残された課題を列挙し、本論文における提案手法の概要を述べた。

第2章では、関連研究について述べた。既存の印象分析手法やアノテーション手法、決定木や階層構造を用いた可視化手法を紹介し、本手法との差異を示した。

第3章では、提案手法について述べた。SD法を用いた大規模印象評価の実施方法と、そこから収集した印象評価データを用いて、ファジィ決定木の構築方法と類似画像一覧表示の描画の可視化手法について説明した。開発した可視化システムのユーザインタフェースについても説明した。

第4章では、印象評価データを提案手法に適用した可視化の実行例を示した。また、その有用性を示すための評価実験について述べた。実行例と評価実験からわかった考察についても説明した。

上述の通り、本論文では、複数の作業員間で個人差の出る印象評価に対して、作業員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手法を提案した。これにより、印象のタグ付けを目的とした本研究の新しいアノテーション作成手順を適用することで、画像から受ける印象の個人差を解消しつつ、印象タグのアノテーション作業の効率化を実現できた。また、本システムを用いて実際にアノテーション作業を評価者にしてもらい、操作ログと再分類結果やアンケートによって、本手法の有用性を示した。評価結果から、決定木と類似画像一覧を連動して活用することで、決定木の可読性の向上を確認でき、決定木の可視化の必要性を示した。

5.2 今後の課題

今後の課題としては、まず印象評価の回答者の偏りの改善をする。表 3.3 にあるとおり、回答者はほぼ 20 代女性が占めているため、男性の回答者を増やしたり、幅広い年代の回答者を集めてもう一度大規模実験をする。また、現在の実装では決定木モデルにファジィクラスタリングを採用しているが、異なるモデルを採用して比較対象実験を実施していきたい。これは女性衣服画像も同じく、異なる画像群で印象評価も実施していきたい。

次に、現段階では、決定木のクラス数と深さを固定して可視化しているため、可変パラメータの使用可能を目指す。これにより、ユーザ操作によって自由な表現を可能とする。さらに、初学者でもわかるように、類似画像一覧表示と決定木の連動関係をより操作しやすく、理解しやすいデザイン設計を考えていく。

最後に、本研究の評価測定を追加で実施していきたい。本研究では、4.4 節にもあるように、評価者は実際に本システムを使用してアノテーション作業をした。今後はその評価者たちが再分類した結果を他者に見せ、一般的にタグ付けが適切かどうかをさらに評価してもらう。これは、本手法において印象の個人差の解消に繋がっているかを測っていく。

そして、これらの機能を実装し、評価実験を完了させた後に、実際にアノテーション支援として訓練データ作成に本システムが貢献していき、印象推定における学習結果の向上に役立てていきたい。

B：参考文献

- [B: Osgood 52] Charles E. Osgood, The nature and measurement of meaning, *Psychological bulletin*, pp. 197, 1952.
- [B: 代 82] 代喜一, 3相因子分析法による椅子の形態イメージの分析, *心理学評論*, pp. 91–103, 1982.
- [B: 豊田 01] 豊田秀樹, 探索的ポジショニング分析 セマンティック・デファレンシャルデータのため3相多変量解析法, *心理学研究*, pp. 213–218, 2001.
- [B: Bradley 94] Margaret M. Bradley, Peter J. Lang, Measuring emotion: the self-assessment manikin and the semantic differential, *Journal of behavior therapy and experimental psychiatry*, pp. 49–59, 1994.
- [B: Kanno 18] Kanno Yoichiro, Minetoki Toshitaka, Bogaki Takayuki, Toko Kiyoshi, Visualization of Flavor of Sake by Sensory Evaluation and Statistical Method, *Sensors and Materials*, pp. 1197–1206, 2018.
- [B: Bezdek 84] James C. Bezdek, Robert Ehrlich, William Full, FCM: The fuzzy c-means clustering algorithm, *Computers & geosciences*, pp. 191–203, 1984.
- [B: Stoklasa 19] Jan Stoklasa, Tomáš Talášek, Jana Stoklasová, Semantic differential for the twenty-first century: scale relevance and uncertainty entering the semantic space, *Quality & Quantity*, pp. 435–448, 2019.
- [B: Osgood 64] Charles E. Osgood, Semantic differential technique in the comparative study of cultures, *American Anthropologist*, pp. 171–200, 1964.
- [B: Muller 21] Michael Muller, Christine T. Wolf, Josh Andres, Michael Desmond, Narendra Nath Joshi, Zahra Ashktorab, Aabhas Sharma, Kristina Brimijoin, Qian

- Pan, Evelyn Duesterwald and others, Designing Ground Truth and the Social Life of Labels, Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems, pp. 1–16, 2021.
- [B: Wu 21] Hui Wu, Yupeng Gao, Xiaoxiao Guo, Ziad Al-Halah, Steven Rennie, Kristen Grauman, Rogerio Feris, Fashion IQ: A new dataset towards retrieving images by natural language feedback, Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 11307–11317, 2021.
- [B: Matsuo 10] Takeshi Matsuo, Hiroki Katayama, Tokumaru Masataka, Noriaki Muranaka, Sensitivity Information Analysis of Running Shoes Using Fuzzy Decision Tree and Visualization of Analytical Results, SCIS & ISIS SCIS & ISIS 2010, pp. 156–160, 2010.
- [B: 井上 09] 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, pp. 90, 2009.
- [B: 太田 17] 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, pp. 277–283, 2017.
- [B: Neto 20] Mário Popolin Neto, Fernando V. Paulovich, Explainable Matrix-Visualization for Global and Local Interpretability of Random Forest Classification Ensembles, IEEE Transactions on Visualization and Computer Graphics, pp. 1427–1437, 2020.
- [B: Mühlbacher 17] Thomas Mühlbacher, Lorenz Linhardt, Torsten Möller, Harald Piringer, Treepod: Sensitivity-aware selection of pareto-optimal decision trees, IEEE transactions on visualization and computer graphics, pp. 174–183, 2017.
- [B: Liu 16] Mengchen Liu, Jiaxin Shi, Zhen Li, Chongxuan Li, Jun Zhu, Shixia Liu, Towards better analysis of deep convolutional neural networks, IEEE transactions on visualization and computer graphics, pp. 91–100, 2016.

- [B: Cui 16] Weiwei Cui, Shixia Liu, Zhuofeng Wu, Hao Wei, How hierarchical topics evolve in large text corpora, *IEEE transactions on visualization and computer graphics*, pp. 2281–2290, 2014.
- [B: Eichner 20] Christian Eichner, Heidrun Schumann, Christian Tominski, Making Parameter Dependencies of Time-Series Segmentation Visually Understandable, *Computer Graphics Forum*, pp. 607–622, 2020.
- [B: Jia 20] Shichao Jia, Peiwen Lin, Zeyu Li, Jiawan Zhang, Shixia Liu, Visualizing surrogate decision trees of convolutional neural networks, *Journal of Visualization*, pp. 141–156, 2020.
- [B: Gou 20] Liang Gou, Lincan Zou, Nanxiang Li, Michael Hofmann, Arvind Kumar Shekar, Axel Wendt, Liu Ren, VATLD: a visual analytics system to assess, understand and improve traffic light detection, *IEEE transactions on visualization and computer graphics*, pp. 261–271, 2020.
- [B: Gomi 10] 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)*, pp. 1–6, 2010.
- [B: Bederson 01] Benjamin B. Bederson, PhotoMesa: a zoomable image browser using quantum treemaps and bubblemaps, *Proceedings of the 14th annual ACM symposium on User interface software and technology*, pp. 71–80, 2001.
- [B: Liu 18] Xiaotong Liu, Yifan Hu, Stephen North, Han-Wei Shen, CorrelatedMultiples: Spatially Coherent Small Multiples With Constrained Multi-Dimensional Scaling, *Computer Graphics Forum*, pp. 7–18, 2018.
- [B: 雙田 08] 雙田珠己, 村上精一, 大学生における衣服の色彩嗜好と選択理由の関連性, *繊維製品消費科学*, pp. 881–888, 2008.
- [B: 松田 19] 松田博子, 名取和幸, 破田野智美, 色の好みとパーソナリティとの関係 -色の感情的意味からの考察-, *日本色彩学会誌*, pp. 69–80, 2019.
- [B: 堀尾 02] 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*, pp. 693–701, 2002.

[B: 村山 08] 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, pp. 157-164, 2008.

[B: 古川 17] 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, pp. 35-48, 2017.