

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

理学専攻 情報科学コース 2040630 飯島 緋理 (指導教員：伊藤 貴之)

1 はじめに

本研究は印象評価結果をアノテーションとして訓練データを構築する工程の半自動化を目的とする。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してしまう。

本研究では1枚の画像あたりの印象タグ付作業をSD法 [1] を用いて、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。収集した大人数の印象回答値を参照して印象のタグ付けを半自動化する過程を可視化し、分析及びタグ付け再考が可能になるような可視化システムを提供する。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、決定木の可読性向上と、画像特徴と印象の関係を明らかにする。

2 SD法を用いた印象評価

2.1 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、印象評価や印象因子分析に関する研究を参照した結果、性差・年齢差・個人差の影響が小さいと考えられる色彩/立体感/正統性/穏健性/装飾性の5種類の因子を採用した。

2.2 パイロットテストと尺度決定

印象評価時の被験者の疲労は回答の信頼性低下を起す。これを回避するために我々は、適切な尺度抽出のためのパイロットテストを実施した。10枚の女性衣服画像を用い、5因子の中から選出した15項目の形容詞対に対して、回答者(20代女性15名)から5段階評価回答を収集した。この評価結果から、任意の形容詞対間の距離行列を算出した。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た。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。この平均値が最大となる形容詞対を因子ごとに抽出したものを表1に示す。

2.3 画像の前処理

本研究は背景が無地単色である女性の衣服画像を対象として印象評価を実施する。衣服以外の要因を印象評価から排除するために、OpenCVを用いて顔領域を検出し、顔及び背景の削除処理を実施した。

表 1: 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	ルーズ
「正統性」尺度 3	フォーマル	カジュアル
「穏健性」尺度 4	日常的	非日常的
「装飾性」尺度 5	シンプル	ゴージャス
評価段階	5 段階	

2.4 印象評価の収集と可視化データの構築

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。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14~15人である。内訳を表1に示す。

3 可視化手法

3.1 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。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への帰属を算出するため、曖昧で柔軟なクラス表現が可能である。本手法ではクラスへの帰属値をクラスタリングの確信度の値として扱う。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。
- クラスタリング結果の確信度を算出。
- 確信度をもとに決定木を構築。



図 1: 右の画像が印象項目「暗い-明るい」から「暗い」印象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。

1枚の画像の5段階評価「暗い-明るい」を例とし、本手法による印象推定の手順を図1に示す。現時点ではクラス数を例えば「暗い」「明るい」「どちらでもない」の3個に固定する。また、我々の実験では木の深さを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に指定している。

続いて、図1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する。木構造を上から下の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。各ノードは確信度の詳細を表しており、分類条件や条件判定を受けた画像数を表示する。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。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。

3.2 類似画像群の一覧表示

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画像群の一覧表示をする。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。

4 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。図2から、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が容易であることがわかる。図3、図4から、画像特徴によって策定結果の全体把握や、深さによる詳細な比較表示を可能にすることがわかる。

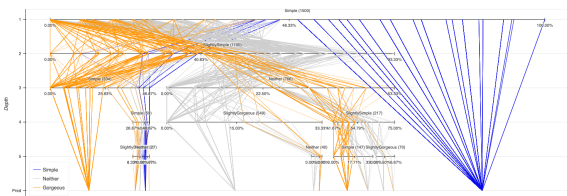


図2: 尺度「シンプル(青) - ゴージャス(オレンジ)」の決定木の可視化。「シンプル」の木の深さが浅いことから、画像群の中でも印象の確信度が高い尺度であることが推察される。



図3: 尺度「暗い-明るい」の画像一覧表示。視覚的に「明るい」は色が鮮やかな服が多く「暗い」は黒の服が多い。

5 評価実験

本手法を用いた可視化の有用性と、システムの操作性と有効性を示すための評価実験を実施した。評価者は20代女性9名で、実際に本システムを用いた操作

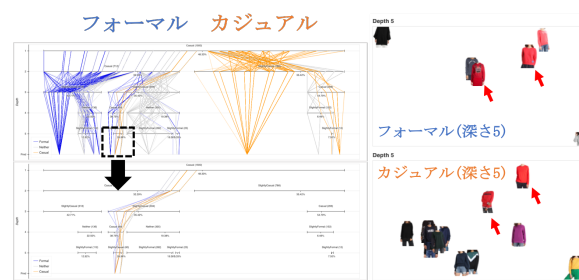


図4: 尺度「フォーマル-カジュアル」。深さ5で、異なる形容詞で策定された画像群を右に表示。フォーマルとカジュアルどちらにも赤矢印にある通り同じ赤色/形/デザインの衣服がある。着衣衣服の印象評価の際に、モデルの髪形や体格が影響する可能性が示唆される。

ログと再分類結果を記録してもらった。評価者の所属グループとしては以下の3グループに分けている。

- ML: 機械学習の専門家3人。
- VIS: 可視化の専門家3人。
- NONE: 機械学習/可視化の非専門家3人。

操作ログから、MLグループは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。VISグループは全体の画像分布を何度も眺め、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。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せず、操作は平行移動と拡大縮小のみの作業が目立った。操作や分析慣れしていないと、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いことが観察できる。

表2から、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2: 分類済みの画像から再分類された画像数の平均

	ML	VIS	NONE
【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

6 まとめ

本研究では、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。今後の展望として、印象評価者の偏りを改善し、さらなる決定木の可読性の向上を目指す。

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。

参考文献

[1] Charles E. Osgood, The nature and measurement of meaning, Psychological bulletin, pp. 197, 1952.