

人物画像の衣服の形状や色を用いた印象語によるコーディネート解析

新坂歩^{†‡} 清水郁子[†] 柴原一友[§] 藤本浩司[§]

[†]東京農工大学大学院工学府情報工学専攻 [§]テンソル・コンサルティング株式会社

1 はじめに

毎日、訪問先、同席する人の年齢や嗜好、その日の気分などを考慮して適切な衣服のコーディネートを考えることはセンスが問われる難しい作業である。そのため、さまざまなコーディネート支援のシステムが提案されている [1] [2]。本研究ではコーディネート支援を印象語により行うことを目指し、コーディネート画像と印象語との関連を解析する。ここで印象語とは”派手”や”地味”のように、画像中の人物がどんなコーディネートをしているのかを示す言葉である。まず、画像から自動的に衣服の色や形状の情報を印象語から想起されるコーディネート画像を提示するために、衣服の色情報や形状から得られる特徴量をもとにサポートベクターマシンで学習を行う。実験では、特徴量による認識率の違いの評価や個人差に関する解析を行なった結果を示す。

2 コーディネート解析の概要

本章では、コーディネート解析の概要について述べる。

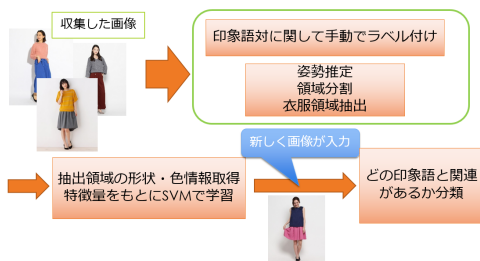


図 1: コーディネート解析の流れ

2.1 印象語について

コーディネート支援の従来研究 [1] で提案されている配色・形状印象因子を参考に、4種類の印象語を用いることにする。印象語は、画像中のコーディネートがどのようなコーディネートなのかを示す言葉である。この印象語について、それぞれの印象語対に”どちらでもない”を含めた3種類のラベル付けを手動で行う。例えば、【派手-地味】という印象語では、”派手”または”地味”または”どちらでもない”の3種類のラベルを付与する。

A method for coordinate analysis of fashion snaps based on impression words using shape and color of clothes

Ayumi NISSAKA^{†‡}, Ikuko SHIMIZU[†], Kazutomo SHIBAHARA[§], Kouji FUJIMOTO[§]

[†]Graduate School of Engineering, Tokyo University of Agriculture and Technology, Tokyo, Japan

[§]Tensor Consulting Co.Ltd., Tokyo, Japan

[‡]s160757q@st.go.tuat.ac.jp

表 1: 印象語対

印象語対
派手-地味
軽い-重い
大人っぽい-子供っぽい
涼しげな-暖かな

2.2 前処理

2.2.1 姿勢推定

まず、姿勢推定手法 [3] を適用し、頭先・首・肩の座標データを取得する。また、領域抽出の精度を向上させるために、コーディネート撮影した人物画像として明らかに適切ではな姿勢が出力されている場合には、姿勢推定を誤ったものとして排除する。具体的には、以下の条件をもとに排除する。

- 右肩と左肩の x 座標が同じもの
 - 頭先の y 座標が画像の上部 25 % に入っていない
 - 右肩・左肩の y 座標の差が 1 % 以上のもの
- また、姿勢推定で得られた頭先・首・肩の座標データから、顔の大きさおよび肩幅を計算する。

2.2.2 衣服領域抽出

領域分割手法 [4] を適用し、人物画像の領域分割を行う。そして、姿勢推定で得られた座標データをもとにトップス・ボトムスの領域を抽出する。この領域分割手法 [4] では、画像内の物体の輪郭検出を行い、輪郭に対して重み付けをし、重みの閾値を設定することで領域分割を行う。本手法では、輪郭の重みを段階的に変化させ、人物の首の位置から顔の大きさ 1 つ分～3 つ分下のピクセルが存在する領域について、”顔の大きさ×肩幅”のサイズの 1 倍～5 倍の大きさになった領域を衣服領域として抽出する。

2.3 特徴量の抽出

抽出された衣服領域について、特徴量の抽出を行う。本手法で用いる特徴量は、色情報に関するものと形状に関するものの2種類に大別できる。色情報に関する特徴量の計算には HSV 色空間を用い、トップス・ボトムス・背景における HSV の中央値および平均値、トップス・ボトムス間の HSV それぞれの中央値および平均値の差、トップス・背景間の HSV それぞれの中央値および平均値の差、ボトムス・背景間の HSV それぞれの中央値および平均値の差を扱う。全 36 種類の特徴量である。形状に関する特徴量の計算では、トップス・ボトムスの面積比および輪郭の長さの比の2種類を用いる。衣服領域の輪郭は一番外側の輪郭を用い、面積は輪郭内の領域を用いる。

2.4 サポートベクターマシンによる学習

前述した4種類の印象語について、色情報および形状の特徴量を用いてサポートベクターマシンで学習を行う。本研究では、得られた特徴量で学習を行い作成した識別器において、各印象語対における分類で色情報および形状のどの特徴量が影響を及ぼしているのかを検証する。

各印象語対に対して、表2に示すような2クラスに分類する。例えば、「派手-地味」の印象語対において「派手-そうではない」と「地味-そうではない」の組み合わせがあり、「そうではない」というクラスにはそれぞれ主軸となる印象語のラベルが付いていない画像が分類される。

表 2: 学習で扱う組み合わせ

[派手/地味] - そうではない
[軽い/重い] - そうではない
[大人っぽい/子供っぽい] - そうではない
[涼しげな/暖かな] - そうではない

3 実験

3.1 コーディネート画像の収集

本実験では、ZOZOTOWNに出店しているショップごとに掲載されているスタッフコーディネートや衣服アイテムを紹介する際の人物が全身写った画像を用いた。トップス1種類、ボトムス1種類となっているコーディネート画像のみを収集した。また、帽子やバッグなど衣服以外のものを身に付けている場合は、衣服とバッグなどがトップスやボトムスの領域に被っていない画像のみを用いた。

収集したデータに対し、2名の被験者（いずれも20代女性、被験者Aと被験者Bと呼ぶ）がラベルづけを行った。各ラベルについて、学習用データに70枚、評価用データに30枚をランダムに選択し、実験を行った。

3.2 特徴量の違いによる実験

被験者Aがラベルづけしたデータのみを用い、各印象語対における分類で色情報および形状の特徴量が影響を及ぼしているかを検証するために、以下に示すような3パターンで検証した。

- all: すべての特徴量を用いた場合
- allHSV: 色情報のみを用いた場合
- allShape: 形状のみを用いた場合

表3に結果を示す。

形状のみを特徴量として学習した場合の分類正答率はすべてのパターンにおいて低くなるという結果となった。一方、色情報のみを特徴量として学習した場合は、分類正答率が全体的に70%以上となる結果が得られた。特に、【派手-そうではない】と【地味-そうではない】のパターンでは、色情報のみの場合の方がすべての特徴量を用いる場合よりも正答率が高くなった。これは、【派手-地味】という印象語対については、衣服の色が与える影響が大きいためであると考えられる。一方で、【涼しげな-そうではない】と【暖かな-そうではない】のパターンでは、すべての特徴量を用いた場合が一番正答率が高くなった。このことから、【涼しげな-暖かな】という印象語対では、色だけではなく形状も影響を与えていることがわかる。

表 3: 色情報による分類正答率の違い

組み合わせ	all	allHSV	allShape
派手-そうではない	81.1%	87.8%	55.6%
地味-そうではない	71.1%	77.8%	50%
軽い-そうではない	81.7%	81.7%	43.3%
重い-そうではない	90%	81.7%	50%
大人っぽい-そうではない	63.3%	56.7%	55.6%
子供っぽい-そうではない	64.4%	72.2%	62.2%
涼しげな-そうではない	85%	71.7%	43.3%
暖かな-そうではない	90%	78.3%	48.3%

3.2.1 個人差の検証

本実験では、2名の被験者がラベル付けしたデータを用いる。2名それぞれのラベルづけしたデータから学習データをランダムに選択し、テストもそれぞれの被験者のデータを用いて行った。

2名の結果を比較すると、被験者Bのデータを用いた実験では、全ての特徴量を用いても、色情報のみ、形状のみをそれぞれ用いても分類正答率は約55%となった。一方、被験者Aのデータではどの特徴量を用いるかで大きく結果が異なった。これにより、同じ画像についても印象語は個人により大きく異なり、識別器も個人に対応したものを生成する必要性が示唆される。

4 まとめ

本研究ではコーディネート支援を印象語により行うことを目指し、コーディネート画像と印象語との関連の解析を行った。衣服の色情報や形状から得られる特徴量をもとにサポートベクターマシンで学習を行い、特徴量の違いによる認識率の差に関する解析を行った。衣服の色または色・形状の両方に注目すると印象語からコーディネート想起できる可能性があると考えられるが、個人差が大きいことがわかった。

参考文献

- [1] 山本萌絵, 鬼沢武久, "ユーザの感性を考慮した対話型服飾デザイン・コーディネートシステム", 日本感性工学会, pp.135-143, 2016
- [2] 村上大志, 黒澤義明, 目良和也, 竹澤寿幸, "オンラインショッピングサイトにおけるコーディネートの類似度判定手法とその印象評価", 言語処理学会第20回年次大会, pp.282-285, 2014
- [3] Y.Yang and D.Ramanan "Articulated pose estimation with flexible mixtures-of-parts" IEEE Conference on Computer Vision and Pattern Recognition, pages 1385-1392, 2011
- [4] Pablo Arbelaez, Michael Maire, Charless Fowlkes, Jitendra Malik, "Contour Detection and Hierarchical Image Segmentation", IEEE Transactions on Pattern Analysis and Machine Intelligence, pp.898-916, 2010