

生地拡大画像からのドメイン適応を用いた服の生地特徴に基づく衣服コーディネート推薦 Outfit Recommendation Based on Fabric Characteristics via Domain Adaptation from High-Resolution Textile Images

森田 欣司[†]

Morita Kinji

原田 史子[†]

Harada Fumiko

島川 博光[†]

Shimakawa Hiromitsu

1. はじめに

現代社会では、WWWの普及によりSNSなどでコーディネート参考となる画像を検索する機会が増加している。しかし、多くはコーディネートの画像の視覚的要素のみに着目し、着心地の情報を得る機会は少ない。また、SNSなどでコーディネート参考となる画像を検索する際に、参考にしたと感じたコーディネートに使用されている衣服を所持していない場合が多い。そのため、自分の所持している衣服を活かすことのできるコーディネートを見つけることに時間がかかったり、新しく衣服を購入し、予想以上のコストをかけてしまったりする問題が生じると考えられる。

本研究では、ユーザが所持している衣服の生地特徴である質感を画像より予測する。これによって、ユーザに推薦するコーディネートの着心地の理解を容易にし、ユーザが所持している衣服を活用したコーディネートを推薦することを可能にする。

2. 関連研究

2.1 着心地に基づく推薦の必要性

コーディネート推薦の関連研究として、画像とキーワードの両方をクエリ入力とした、マルチモーダルな衣服画像検索システムを用いた衣服推薦が挙げられる[1]。Zhengzhongらは、衣服の画像に加えて、色、質感などの情報を文章で入力することで、衣服画像の検索精度を向上させる手法を提案した。質感を変数として扱うことで、着心地の理解を支援するだけでなく、より質の高いコーディネート推薦を実現できる。

2.2 StableNet

本研究では、衣服の質感と色を効率的にモデル学習するため、生地拡大画像を用いる。この学習済みモデルが全身画像に用いられる衣服に対して、質感と色を予測できるようドメイン適応を行う。テストデータと訓練データの分布が異なると、大幅に予測の精度が低下する可能性がある。StableNetは、訓練データの特徴間の依存関係を排除することで、偽の相関を取り除くことを可能にする[2]。よって本研究では、コーディネートの全身画像から使用されている衣服の生地特徴を抽出する全身画像からの衣服の抽出するため、生地拡大画像で学習するモデルにStableNetを用いることでドメイン適応を行う。

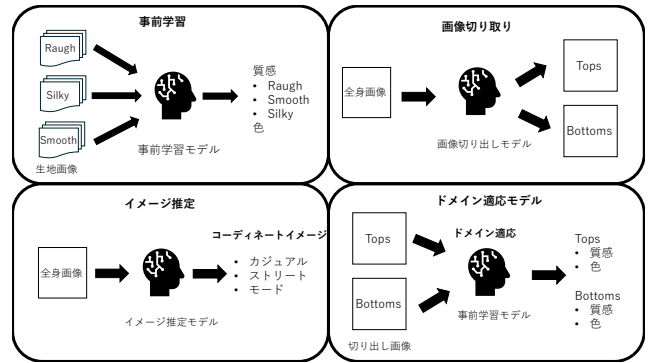


図1 全身画像データセットDの作成手法

2.3 YOLO

ファッション雑誌などで紹介されているコーディネート参考にするさい、多くの画像が衣服単体でなく全身画像である。しかし、衣服には役割や目的に応じた異なる種類がある。これらの種類を考慮せずモデル学習を行うと、ユーザが期待するコーディネートを推薦することは不可能である。よって本研究では、全身画像から確認することができるすべての服を切り出す。

物体検出によく用いられるアルゴリズムとしてYOLOがある。YOLOは画像から目的の箇所を抽出する機械学習のモデルである。RohanらはYOLOを用いて自然な全身画像から衣服を抽出する手法を提案している[3]。本研究では、コーディネートの全身画像から確認することができるすべての服を切り出す方法としてYOLOを用いる。

3. 生地拡大画像からの質感と色予測のドメイン適応による全身画像のコーディネート推薦

3.1 手法概要

本研究では、ファッション雑誌やオンライン上で公開されている全身コーディネート画像から、各衣服に含まれる生地の質感と色からなる生地特徴を抽出し、ユーザが既に所有している衣服と類似した素材や印象を持つ衣服を活用した、着心地と印象を考慮したコーディネートを提示する手法を提案する。

提案手法では、入力として衣服の拡大画像Pおよび推薦対象のジャンルを用いる。全身画像では、衣服から得られる特徴が、衣服を着用している人の特性や撮影環境によって大きく変化してしまう。また、衣服以外の情報が含まれている場合、生地の予測精度に大きく影響を与える。よって、全身画像から衣服の特徴を抽出することは困難である。そのため、本研究では、衣服の質感と色を効率的にモデル学習するため、生地拡大画像を用いる。

[†] 立命館大学 情報理工学部

College of Information Science and Engineering, Ritsumeikan University

事前にファッション雑誌などで公開されているコーディネート全体の全身画像を収集して色、生地特徴の組み合わせとジャンルをラベル推定したデータセット D を用意しておく。入力された衣服拡大画像 P と生地特徴が一致する衣服を活用したコーディネート全体の全身画像 (群 $D' \subseteq D$) を得る。 D' の中から、入力ジャンルと一致する全身画像を推薦する。データセット D の作成の手法を図 1 に示す。以降の節では当該手法について詳述する。

3.2 生地拡大 Web 画像からの生地特徴予測モデル

生地の拡大画像を Web 上より収集しラベル付けを行う。本研究では、実際の着心地の理解を支援するため、生地の視覚的な質感ではなく、ユーザが着た際に感じる質感に着目する。質感のラベルの種類は、Raugh(ガサガサ)、Smooth(ツルツル)、Silky(サラサラ)の 3 種類を用いる。WWW 上の衣服の販売ページから生地の説明情報と拡大画像を収集し各画像にラベル付けする。Web 上の説明文に用いられている「さらりとした」などの単語を参考にラベル付けを行う。また、RGB の計算方法は、画像全体の RGB 値の平均を用いる。収集した生地拡大 Web 画像でモデルを訓練し、生地の拡大画像から生地の質感ラベルの所属確率と色を推定する生地特徴予測モデル M を作成する。生地が各クラスに属する確率と色の RGB 値を出力する。各クラスに属する確率が最も高いものをその生地の質感として扱う。モデルは StableNet を転移学習したものを用いる。

3.3 全身 Web 画像データの収集

ユーザにコーディネートを紹介するためコーディネート画像を大量に収集する必要がある。大量にデータを収集することで、偏りのないデータを使用することができ、ユーザのニーズに広く応えることが可能である。本研究ではファッション雑誌で紹介されているコーディネート画像を用いる。ファッション雑誌は大規模な媒体であるため、大量のデータを集めることが可能である。また、多くのユーザが参考にしたいと感じるコーディネートの画像を紹介しているため、本研究が提案する手法として適している。本研究では、多くのユーザのニーズに応えるため、「カジュアル」～「モード」のジャンル (後述) に分けることのできるコーディネートの画像を収集する。また、最新のトレンドを取り入れた、女性向けと男性向けの両方のコーディネート画像が必要である。よって、「VOGUE JAPAN」、「ELLE DIGITAL」、「GQ JAPAN」からコーディネート画像を収集する。

3.4 ドメイン適応モデルの作成

本研究では、コーディネートの全身 Web 画像から使用されている衣服の生地特徴を抽出する。しかし、全身 Web 画像では、衣服から得られる特徴が、衣服を着用している人の特性や撮影環境によって大きく変化してしまう。また、衣服以外の情報が含まれている場合、生地特徴の予測精度に大きく影響を与える。よって、全身 Web 画像から衣服の特徴を抽出することは困難である。本研究では全身 Web 画像から切り取られた各アイテム Web 画像 (後述) の生地と

色を推定できるよう、生地特徴予測モデル M にドメイン適応を行う。 M の訓練データは生地の拡大画像であるが、入力をアイテム画像に変更した際に分布シフトが起こると考えられる。事前学習した StableNet M に対して、入力画像の分布が異なっても予測を行うことができるようチューニングを行いモデル M' を得る。モデルに StableNet を用いることで訓練データの偽の相関を取り除くことができ、生地の質感が生地の色で決定されるような状況を防ぐことができると考えられる。

3.5 全身 Web 画像からのアイテム画像の切り取り

コーディネートの全身画像に使用されている衣服についての生地特徴を抽出するため、全身の画像からそれぞれの衣服を切り出す必要がある。本研究では、YOLO モデル Y を用いて各衣服の切り取りを行う。コーディネートに用いられている衣服の組み合わせを提案するため、全身の画像から確認することができるすべての服を切り出す。

3.6 全身 Web 画像のジャンル推定

本研究では、Vision Transformer を転移学習して全身 Web 画像のジャンルを分類するモデル V を作成する。コーディネートの提案を行う際に、あらかじめジャンルを設定しておくことでユーザが求めているコーディネートの提案を可能にする。分類カテゴリとしては「カジュアル」「ストリート」「きれいめ」「スポーティ」「ガーリー」「モード」を用いる。この 6 種類のジャンルを用いることで、多くのコーディネートについて説明することができ、ユーザのニーズに広く応えることが可能である。コーディネートの全身画像について各クラスに属する確率を出力する。

4. おわりに

本研究ではユーザの所持している衣服と、ユーザが求めているコーディネートのジャンルからコーディネートを紹介する手法を提案した。本研究のポイントは、コーディネートの全身画像から衣服の特徴を抽出することは困難であるため、拡大画像で訓練したモデルに対してドメイン適応を行い、全身画像から衣服の特徴を抽出することにある。そのため、実際のコーディネート画像を用いた実験を通じて、ドメイン適応により質感の分類器の精度向上を確認する。また、本研究の提案する手法が、実際にユーザが満足するコーディネート推薦手法であるか検証する必要がある。

参考文献

- [1] Zhengzhong Zhou, Yifei Xu, Jingjin Zhou, and Liqing Zhang, "Interactive Image Search for Clothing Recommendation", Proceedings of the 24th ACM international conference on Multimedia (MM '16), p.754-756 (2016)
- [2] Zhang, Xingxuan and Cui, Peng and Xu, Renzhe and Zhou, Linjun and He, Yue and Shen, Zheyang "Deep Stable Learning for Out-of-Distribution Generalization", Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), p.5372-5382, (June,2021).
- [3] Jadhav, Rohan Indrajeet and Stynes, Paul and Pathak, Pramod and Haque, Rejwanul and Hasanuzzaman, Mohammed, "An Instance Segmentation Model to Categorize Clothes from Wild Fashion Images", Proceedings of the 2022 6th International Conference on Deep Learning Technologies, Association for Computing Machinery, p.75-83, doi:10.1145/3556677.3556690 (2022).