

ファッションアイテム領域抽出技術を利用した
コーディネータ推奨ボットの改良
Improvement of Fashion Coordinator Recommendation Bot
with Fashion Item Segmentation

澤田 佑介[†]
Yusuke Sawada

田鍋 あかり[†]
Akari Tanabe

中田 洋平[†]
Yohei Nakada

1. はじめに

近年、ファッションアイテム購入の参考情報が得られるWEAR[1]等のようなコーディネートサイトが登場してきている。しかし、このようなコーディネートサイトは、膨大な情報が掲載されており、不慣れなユーザにとっては、自身の志向に適したコーディネートやコーディネータを見つけることが容易であるとは言い難い。このような背景を受け、これまで著者らのグループは、入力されたコーディネート画像から類似コーディネート検索やコーディネータ推奨を行うLINEボットを試作してきた[2]。更に、その中で用いる身体部位位置推定技術をOpenPose[3]からCDCL-HPS[4]に変更し、改良を実施してきた[2][5]。本稿では、その機能の更なる改良を目的とし、これまで使用してきた身体部位位置推定技術に加え、ファッションアイテム領域抽出技術を導入した改良方式を記す。また、本改良方式に対して実施した、類似コーディネート検索に関する初期的検証実験の結果を示す。

2. 従来方式

試作LINEボット[2][5]においては、現在は身体部位位置推定技術であるCDCL-HPS[4]を利用し、身体部位をコーディネート画像内の2次元領域として捉えることにより、類似コーディネート検索やコーディネータ推奨を実施している。以降、この従来方式について、具体的に述べる。

2.1 身体部位の抽出

本従来方式では、CDCL-HPS[4]のサイト[6]から入手した15種の身体部位の位置を推定する学習済みモデルを用いている。ただし、15種の身体部位内、コーディネートに主として関連する両上腕、両前腕、両大腿、両下腿、胴体の領域を抽出して用いている。なお、抽出された領域の境界周辺部は、他の身体部位や背景などが含まれていることも多いため、抽出した領域には収縮処理を施している。また、横向きのコーディネート画像などにも対応するため、多くのアパレルアイテムは左右対称であることから、4つの身体部位の上腕、前腕、大腿、下腿に関しては左右の部位を同一の部位として纏めている。更に、このようにして得られた5つの両上腕、両前腕、両大腿、両下腿、胴体の領域のうち、収縮処理後の面積が0より大きいものが3つ以上のものを有効な画像としている。

2.2 類似コーディネート検索法

本従来方式における類似コーディネート検索では、2つのコーディネート画像 A, B 間の距離関数を、次式のように定



(a) 検索画像 (b) 1位 (c) 2位 (d) 3位
図1 従来方式[2][5]による類似コーディネート検索結果例

めている。そして、入力画像と予め格納されているコーディネート画像間の距離を求め、類似コーディネート画像を検索している。

$$D(A, B) = \sqrt{\frac{\sum_{i \in I \cap J} \|\sqrt{f_i} - \sqrt{g_i}\|_2^2}{|I \cap J|}} + \lambda \left(1 - \frac{|I \cap J|}{|I \cup J|}\right) \quad (1)$$

ここで、 i は身体部位番号、 f_i, g_i は画像 A, B の各々における身体部位 i の領域に関する三次元色ヒストグラム（正規化済）、 I, J は画像 A, B の各々で抽出できた身体部位の集合を示している。また、 $\|\cdot\|_2$ は2ノルムを表し、 $|\cdot|$ は基数を表す。即ち、この距離関数では、第1項が、身体部位での三次元色ヒストグラムの平均ヘリンジャー距離であり、コーディネート画像間における色の不一致度を表している。また、第2項は、抽出できた身体部位の不一致度を表している。なお、正の定数 λ は、2つの項のトレードオフを調整するパラメータであり、実験的に0.2と定めている。

2.3 コーディネータ推奨法

本従来方式におけるコーディネータ推奨では、類似コーディネート画像を算出後に、次式で定められた推奨度を算出し、入力画像との距離値が小さい類似コーディネートを多く持つコーディネータを推奨できるようにしている。

$$R_k = \sum_{s \in S_k} \exp(-\eta D_s) \quad (2)$$

ここで、 k は推奨の対象となるコーディネータの番号である。 D_s は入力画像と第 s 位の類似コーディネート画像との距離(1)を示している。 S_k は、上位 N 位内の類似コーディネート画像であり、距離値の上限を表す D_{MAX} 以下の距離(1)を持つものの中で、コーディネータ k のコーディネート画像の順位を集めた自然数の集合である。即ち、 $S_k = \{s \in N | s \leq N, D_s \leq D_{MAX}\}$ である。なお、正の定数 η は距離値と推奨度の対応を調整するパラメータであり、実験的に5.0と定めている。また、順位の上限值 N は20、距離値の上限值 D_{MAX} は0.8に設定している。

2.4 従来方式の課題

文献[5]にも記したように、従来方式でのコーディネータ推奨ボットはまだ一部の改善の余地が見られる。図1は、従来方式を用いて、後述の実験と同様の条件下でWEAR[1]

[†] 明治大学 総合数理学部
School of Interdisciplinary Mathematical Sciences, Meiji University

から得たある検索画像例に対する類似コーディネートへのランキング1位~3位を示したものである。図1から分かるように、ランキング2位、3位のコーディネート画像は、色合いやテイストは似ているものの、ボトムスの種類が異なるコーディネート画像であることが確認できる。

3. 改良方式

上記のような課題を解決すべく、本稿では、アイテム領域抽出技術を用いた改良方式を提案する。以降、この改良方式について、具体的に述べる。

3.1 アイテムの抽出

本改良方式では、アイテム領域抽出技術として、インスタンスセグメンテーションを行う深層神経経路網モデルであるMask R-CNN[7]を用いる。より具体的には、ファッションデータベースであるModanet[8]を整備し、それを用いて学習した学習済みパラメータ[9]を持つMask R-CNNを使用する。なお、この学習済みパラメータ[9]では、13種類のアイテムカテゴリの信頼度、および、領域を抽出することが可能である。ただし、本改良方式では、13種類のアイテムの全てを対象とはせず、コーディネート印象に大きく影響するトップス、アウター、ショーツ、パンツ、スカート、ドレス(ワンピース)の6種類のみを抽出の対象としている。また、既に身体部位位置推定技術で領域の情報は考慮されているため、抽出された領域についての情報は用いず、信頼度の情報のみを用いている。

3.2 距離関数の改良

本改良方式では、Mask R-CNN[7]の学習済みパラメータ[9]を用いて抽出された信頼度を用いて、以下のような距離関数を考える。

$$d(A, B) = \|s - r\|_2 \quad (3)$$

ここで、 r , s は画像 A, B の各々における対象アイテムの信頼度を並べたベクトルである。また、 $\|\cdot\|_2$ は2ノルムを表す。即ち、この距離関数(3)は、抽出できたアイテムの信頼度の不一致度を表すものとなっている。なお、この信頼度ベクトルの算出においては、抽出されなかったアイテムの信頼度の値は0とし、同じ種類のアイテムが複数抽出された場合は、その中で最大の信頼度の値としている。

そして、本改良方式では、この距離関数(3)に正の係数を掛けたものを加え、前述の距離関数(1)を以下のように改良している。

$$D'(A, B) = D(A, B) + \alpha d(A, B) \quad (3)$$

なお、正の定数 α は、2つの距離関数(1)と(3)のトレードオフを調整するパラメータであり、後述の実験では0.5と設定している。

4. 初期的検証実験

前述のような改良方式の検証のため類似コーディネート検索に関する初期的検証実験を実施した。具体的には、コーディネートサイトWEAR[1]に掲載された女性ユーザー月間ランキングより、4年間(2016年1月から2019年12月)各月の上位100位程の画像データ計4571枚を用いて、従来方式と改良方式により、いくつかの検索画像例に対して類似コーディネートランキング上位を算出し、比較した。なお、その際に検索画像自身は検索対象から除外している。

図2には、このような検証実験の下で得られた図1に対



(a) 検索画像 (b) 1位 (c) 2位 (d) 3位
図2 改良方式による類似コーディネート検索結果例

応する改良方式の検索結果例を示す。図1と図2を比較すると、従来方式では、ボトムスがスカートであるコーディネート画像が検索されているのに対し、改良方式では、ボトムスがパンツであるコーディネート画像のみ検索されていることがわかる。また、本稿では、紙面の関係上、掲載できないが、アウターを着用している検索画像例においては、上位5位がアウターを着用しているコーディネートが検索されるなど、改良方式の有効性が確認されている。

5. まとめと今後の課題

本稿では、これまで著者らのグループが試作してきたコーディネート推奨ボットの更なる改良を目的として、ファッションアイテム領域抽出技術を導入した改良方式について記した。また、本改良方式に対しての初期的検証実験の結果を示した。今後、改良方式を用いたLINEボットを実装し、その使用感を検証するとともに、改良方式の定量評価を実施していく。また、類似コーディネート検索法やコーディネート推奨法の更なる改良を目指す。

謝辞

本研究を進めるにあたり、ZOZO Research 中村琢磨氏には貴重なご意見ご助言を頂きました。深謝いたします。

参考文献

- [1] 株式会社 ZOZO テクノロジーズ, WEAR, <http://wear.jp/women-ranking/>
- [2] 田鍋 あかり, 三宅 季実佳, 中田 洋平, “身体部位位置推定技術を利用したコーディネート推奨ボットの試作と改良”, 2020年電子情報通信学会総合大会 学生ポスターセッション, ISS-P-042, (2020).
- [3] Z. Cao, G. Hidalgo, T. Simon, S. E. Wei, and Y. Sheikh, “Realtime Multi-person 2D Pose Estimation using Part Affinity Fields”, IEEE Conference on Computer Vision and Pattern Recognition 2017, pp. 7291-7299 (2017).
- [4] K. Lin, L. Wang, K. Luo, Y. Chen, Z. Liu, and M.-T. Sun, “Cross-Domain Complementary Learning Using Pose for Multi-Person Part Segmentation”, IEEE Transactions on Circuits and Systems for Video Technology, Vol. 31, No. 3 (2020).
- [5] 澤田 佑介, 田鍋 あかり, 中田 洋平, “試作コーディネート推奨ボットに対するファッションアイテム領域抽出技術の適用”, 2021年電子情報通信学会総合大会 学生ポスターセッション, ISS-P-049 (2021).
- [6] K. Lin, “CDCL-human-part-segmentation”, <https://github.com/kevinlin311tw/CDCL-human-part-segmentation>
- [7] K. He, G. Gkioxari, P. Dollár, and R. Girshick, “Mask R-CNN”, Proc. 2017 IEEE International Conference on Computer Vision (ICCV), pp. 2961-2969 (2017).
- [8] S. Zheng, F. Yang, M. H. Kiapour, and R. Piramuthu, “Modanet: A Large-scale Street Fashion Dataset with Polygon Annotation”, In Proc. 26th ACM International Conference on Multimedia, pp. 1670-1678 (2018).
- [9] P. C. Cadoppi, “Maskrcnn-Modanet”, <https://github.com/cad0p/maskrcnn-modanet>