

# ファッション雑誌を用いたコーディネート推薦システム Clothes Coordinate Recommender System using Fashion Magazines

岩田 具治<sup>†</sup>  
Tomoharu Iwata

渡部 晋治<sup>†</sup>  
Shinji Wanatabe

澤田 宏<sup>†</sup>  
Hiroshi Sawada

## 1. まえがき

多くの人がファッションに関心を持つ。彼ら彼女らは、定期的にファッション雑誌でトレンドをチェックし、服を購入する際やコーディネートを決める際に参考にする。ファッション雑誌には参考にすべきコーディネートをしたモデルの写真が多数掲載されている。本稿では、これら雑誌の写真画像から、コーディネートに関する情報を自動抽出し、コーディネートを紹介するシステムを提案する。具体的なタスクは、上半身の服装(上衣・トップス)の写真が与えられたとき、その服装に適した下半身の服装(下衣・ボトムス)の写真を選択すること、もしくは逆に、下衣の写真が与えられたとき、適した上衣の写真を選択することである。提案法は二つのステップから成る。まず、雑誌から全身が写っている写真を抽出し、その上衣および下衣の領域を抽出する。次に、それらの領域の画像特徴量を用いて上下衣の関連性を学習する。本稿では、二つの推薦手法を提案する。一つ目は類似画像検索に基づく手法、二つ目はトピックモデルに基づく手法である。直感的には、与えられた上衣(下衣)と類似した雑誌写真を抽出し、その抽出写真の下衣(上衣)と類似した服を推薦する。これまでに、多くのコーディネート推薦システムが提案されているが [5, 6, 8], 全身写真から自動的に推薦する技術に関する研究はない。提案法では、既存の雑誌やオンラインストアにある写真を用いればよいから、メタデータを人手で作成する必要はないという利点がある。提案法は、毎朝持っている服のなかから着る服を選ぶ場合だけでなく、オンラインストアで服を購入する場合でも適用可能である。

## 2. 提案法

### 2.1 上下服装領域抽出

雑誌に含まれる全写真集合から、全身が写っている写真を特定し、それぞれの全身写真において上衣および下衣の領域を抽出する手法について説明する。本研究では、顔検出システムを用いて簡単に領域抽出を行う。まず、全写真で顔検出を行い、顔の領域を特定する。顔領域が含まれるか、および、顔領域の大きさ・位置が適切であるか(上下衣が写るスペースがあるか)の条件によって、全身写真を特定する。上衣領域は顔領域の2倍の幅、2.5倍の高さ、下衣領域は顔領域の2倍の幅、3.5倍の高さとし、顔領域の下を上衣領域として、その下を下衣領域として抽出する。なお、条件を満たす顔領域が複数ある場合、中心線に最も近い領域をその写真の顔領域とした。図 2.(a) に上下衣領域抽出例を示す。

### 2.2 データ形式

雑誌の全身写真集合を  $D_{\text{mag}}$ 、所有している服の写真集合を  $D_{\text{own}}$  とする。なお  $D_{\text{own}}$  は、全身に限らず、上衣もしくは下衣だけの写真でもよい。ここでは、所有している服でのコーディネートの推薦を想定して  $D_{\text{own}}$  は所有している服の写真集合としたが、オンラインストアでの推薦の場合、 $D_{\text{own}}$  を販売している服の写真集合とすることで、適用できる。

それぞれの写真集合において、前節の手法により、上下衣領域が抽出されているとし、写真  $m$  の上衣および下衣の特徴量をそれぞれ  $u_m, l_m$  とする。特徴量として、色、テクスチャ、形状など任意の特徴量を用いることができる。ここでは、 $u_m, l_m$  とともに、bag-of-words 表現されているとする。つまり、 $u_m = \{u_{mv}\}_{v \in V}$ 、 $l_m = \{l_{mv}\}_{v \in V}$  であり、 $u_{mv}$  は写真  $m$  の上衣領域で特徴  $v$  が出現する回数、 $l_{mv}$  は写真  $m$  の下衣領域で特徴  $v$  が出現する回数、 $V$  は特徴集合を表す。

### 2.3 類似領域に基づく推薦

抽出した上下衣領域の類似度に基づいてコーディネートを紹介する手法について説明する。ここでは上衣が与えられた場合の推薦について説明するが、下衣が与えられた場合の推薦も同様にしてできる。特徴量  $u_*$  を持つ上衣が与えられたとする。この服装に合う下衣を、所有写真集合  $D_{\text{own}}$  の中から選択したい。類似領域に基づく推薦では、まず、 $u_*$  に最も似ている上衣の写真  $\hat{m}$  を雑誌写真の中から検索する。

$$\hat{m} = \arg \max_{m \in D_{\text{mag}}} S(u_m, u_*), \quad (1)$$

ここで  $S(\cdot, \cdot)$  は類似度関数を表す。類似度としてコサイン類似度などを用いることができる。そして、その検索写真  $\hat{m}$  の下衣に最も似ている下衣の写真  $\hat{n}$  を所有写真の中から検索し、推薦する。

$$\hat{n} = \arg \max_{n \in D_{\text{own}}} S(l_n, l_{\hat{m}}). \quad (2)$$

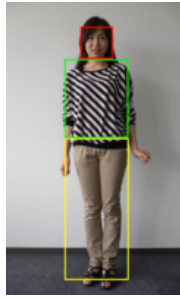
複数の下衣を推薦する場合、類似度が高いものから順に並べて提示すればよい。

### 2.4 トピックモデルに基づく推薦

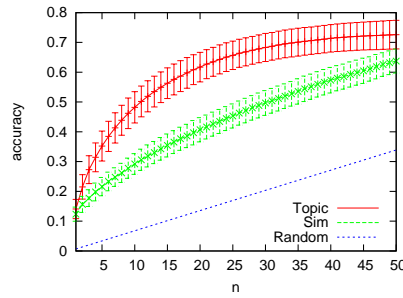
トピックモデルに基づくコーディネート推薦手法について説明する。トピックモデルとは、データに含まれるトピック(潜在意味)を抽出する手法で、文書や画像など様々な分野で応用されている [1, 2, 3, 4]。ここでは、上下衣の全身写真からトピックを抽出することで上下衣の関連を学習し、推薦に応用する。トピックモデルでは、写真  $m$  の上下衣の特徴量  $u_m, l_m$  の確率は、

$$P(u_m, l_m) = \prod_{v \in V} \left( \sum_{z \in Z} \theta_{mz} \phi_{zv}^u \right)^{u_{mv}} \left( \sum_{z \in Z} \theta_{mz} \phi_{zv}^l \right)^{l_{mv}}, \quad (3)$$

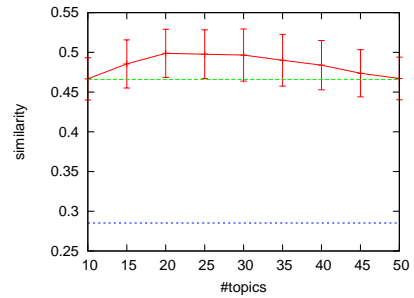
<sup>†</sup>日本電信電話株式会社 NTT コミュニケーション科学基礎研究所



(a) 上下衣領域抽出例



(b)  $n$ -best 正答率



(c) 類似度

で表される．ここで  $z$  はトピック， $Z$  はトピック集合， $\theta_{mz}$  は写真  $m$  がトピック  $z$  を持つ割合， $\phi_{zv}^u, \phi_{zv}^l$  はそれぞれ，トピック  $z$  において上衣，下衣領域で特徴  $v$  が出現する確率を表す．まず，雑誌の全身写真集合  $D_{\text{mag}}$  を用いてモデル， $\Phi = \{\phi_{zv}^u, \phi_{zv}^l\}_{z \in Z, v \in V}$  を学習する．次に，学習したモデルを用いて，与えられた上衣写真  $u_*$ ，および，所有写真集合  $D_{\text{own}}$  それぞれのトピック割合  $\theta_n = \{\theta_{nz}\}_{z \in Z}$  および  $\theta_*$  を推定する．そして，与えられた上衣写真のトピック割合  $\theta_*$  と最も類似するトピック割合を持つ下衣写真  $\hat{n}$  を所有写真の中から検索し，推薦する．

$$\hat{n} = \arg \max_{n \in D_{\text{own}}^l} S(\theta_n, \theta_*), \quad (4)$$

ここで  $D_{\text{own}}^l$  は所有している下衣の写真集合を表す．類似度として，負の KL ダイバージェンスなどを用いることができる．

### 3. 実験

#### 3.1 領域抽出

評価のため，女性雑誌 32 巻のデータを用いて実験した．このデータには 14,813 枚の画像が含まれている．このなかには，本研究で必要となる全身写真だけでなく，小物や記号，文字の画像も含まれる．2.1 節に記述した手法を用いて，全身写真の抽出，および，上下衣領域の特定を行った結果，2,062 枚に絞られた．ここで，顔検出システムとして `opencv` を用いた．これらの抽出写真を人で判断したところ，1,502 枚 (73%) が適切に領域抽出ができていた．提案領域抽出手法は，単純であるが，高い精度で領域を抽出できると言える．以後の実験では，適切だと判断された写真集合から重複写真を除いた 1,475 枚の全身写真を用いた．

#### 3.2 推薦

全写真から 10% をテスト写真としてランダムに選択し，それ以外の写真を学習写真として 100 データセットを作成した．画像特徴として，色ヒストグラムを用いた [7]．評価尺度として  $n$ -best 正答率，および，類似度を用いた． $n$ -best 正答率とは，テスト写真の上衣 (下衣) を入力し  $n$  枚推薦したとき，正解，つまりテスト写真の下衣 (上衣)，が含まれる割合を表す．類似度とは，推薦した服がどれだけ正解に類似しているかを表す．類似度としてコサイン類似度を用いた．類似度に基づく手法 (Sim)，トピックモデルに基づく手法 (Topic)，および，ランダム

提示手法 (Random) を比較した結果を，図 2.(b)(c) に示す．(b) の Topic はトピック数 20 のモデルの結果である．Sim, Topic とともに，Random に比べ高い性能を示しており，両提案手法ともに効果があると言える．Topic は Sim に比べ高い．これは，トピックを抽出することにより，上下衣の関連を適切に学習できているためと考えられる．実際に推薦された服を見てみると，入力した服に関連する写真を適切に推薦されていた．

### 4. おわりに

本稿では，ファッション雑誌の写真を用いたコーディネートを推薦するシステムを提案し，実験により提案手法の効果を示した．今回は顔検出システムを活用し，上下衣領域を抽出したが，上下衣領域を直接抽出することにより，より高精度に領域抽出することが可能になる．上下衣だけでなく，帽子，アクセサリ，カバンなど，他のファッションアイテムの推薦や，髪型や顔のつくりに適した推薦に発展することが考えられる．トピックモデルに基づく推薦手法は，複数の領域のモデル化に容易に拡張可能である．また，ファッション雑誌は多様であるため，自分の好みの雑誌を学習データとすることで，より適した推薦が可能となる．今後，実際にユーザに利用・評価してもらい，システムを発展させていきたい．

### 参考文献

- [1] D. M. Blei, A. Y. Ng, and M. I. Jordan. Latent Dirichlet allocation. *JMLR*, 3:993–1022, 2003.
- [2] T. Hofmann. Collaborative filtering via Gaussian probabilistic latent semantic analysis. In *SIGIR '03*, pages 259–266, 2003.
- [3] T. Iwata, T. Yamada, and N. Ueda. Probabilistic latent semantic visualization: topic model for visualizing documents. In *KDD '08*, pages 363–371, 2008.
- [4] D. Mimno, H. M. Wallach, J. Naradowsky, D. A. Smith, and A. McCallum. Polylingual topic models. In *EMNLP '09*, pages 880–889, 2009.
- [5] S. Nagao, S. Takahashi, and J. Tanaka. Mirror application: Recommendation of clothes coordination in daily life. In *HFT '08*, pages 367–374, 2008.
- [6] E. Shen, H. Lieberman, and F. Lam. What am I gonna wear?: scenario-oriented recommendation. In *IUI '07*, pages 365–368, 2007.
- [7] M. J. Swain and D. H. Ballard. Color indexing. *Int. J. Comput. Vision*, 7(1):11–32, 1991.
- [8] Y. Yonezawa and Y. Nakatani. Fashion support from clothes with characteristics. In *HCI International '09*, pages 323–330, 2009.