

## アパレルアイテム商品紹介文中の“演出”に着目した感性語の自動推定

黒澤 義明                      竹澤 寿幸  
 広島市立大学大学院 情報科学研究科  
 [kurosawa, takezawa]@ls.info.hiroshima-cu.ac.jp

### 1. はじめに

アパレルアイテムは、個人の主観を反映した嗜好によって売れ行きが左右される。このため、個人の主観的な印象を対象とするセンチメント解析の重要な研究対象と言える。

その一方で、その嗜好は様々な要素が入り組んで構成されているため、基本的な課題～positive か negative かを判定する～だけでは不十分な対象とも言える。例えば、素材の肌触りのよいからなのか、あるいは、アイテムから受ける印象 (ex. 女性らしい, マニッシュな) からなのかわからないからである。もしわかれば、アイテム推薦やコーディネート診断などの用途に有効に使用することができる。

そこで本研究は、アパレル販売サイトの商品紹介文を題材とし、その紹介文がどのような印象を喚起するか推定することを目的とする。なお、印象を表す語を感性語と呼ぶ。

今回は特に、印象喚起の手がかりとしてキーワード“演出”に注目し、共起する感性語による機械学習実験を行う。“演出”は紹介文の書き手が読み手に期待する印象を明示的に記述していると考えられ、印象喚起の推定に重要と考えるからである。

### 2. 商品紹介文と関連研究

#### 2.1. 商品紹介文とは？

まず、商品紹介文の例を示す<sup>1</sup>。

シフォン素材のワンピースです。流行のパステルカラーで用意しました。バックに添えたりボンも **ガーリー**さを演出します。リボンは取り外しできますので、甘すぎが好みでない方にも使っていただけます。

図 1 商品紹介文の例

z この例はワンピースの例であり、実際のサイトでも、ブランドにより記述量の多少はあるものの、

<sup>1</sup> 本研究が言う「商品紹介文」は、複数の文から構成される一連のテキストを指す。以降、実験等で「商品紹介文」中の一文を指すことが必要な場合は、そう明示することとする。

同様の記述が見られる。

この例からも明らかなように、紹介文には様々な角度からの記述がなされる。ここでは、素材 (シフォン), 色 (パステル), デイテール (リボン) に加えて、印象 (ガーリー, 甘すぎ), 効果 (●●を演出) が挙げられよう。

本研究ではこのような記述すべき特徴を、アパレルアイテムの構成要素と呼ぶ。

#### 2.2. アパレルアイテム構成要素

アパレルアイテムを構成する要素については、様々な文献 (鈴木ら 2005, 富田 2004, 小林 2003, 工藤ら 2001, 飯塚ら 1985) で検討されている。黒澤ら (2015) はこれを整理し、22 種類の構成要素を提案している。

##### 2.2.1 客観的要素

22 種類の多くは、客観的に記述可能な特徴である。上の例では『シフォン素材のワンピースです』『パステルカラーで用意しました』の 2 文内に要素「素材」や要素「色」が登場する。客観的な事実～シフォンで作られたワンピース～～パステルカラーのワンピース～があるだけである。

##### 2.2.2 主観的要素

実際には事実の記述だけにとどまらず、『シフォンなので●●』のように、要素「印象」や要素「効果」を含む主観的な文となることも多い<sup>2</sup>。視点も主観的になるため、同一のアイテムであっても、使われる感性語そのものが大きく変わりうる。例えばシフォンの場合、“透け感がある”, “涼し気”, “軽やか”, “女性らしい” が視点として選ばれることが多い。こうした視点は担当者・季節・流行などによっても変わるため、商品紹介文の記述が多様になると予想される。

このことは出現回数数の分析からも明確である。黒澤ら (2015) によれば、要素「色」要素「柄」は少なく (1,116, 1,593), 一方、要素「印象」は非常に

<sup>2</sup> 主観的要素を排除するサイトもある。『サイドスリット, バックジッパー。』のように非常にシンプルな商品紹介文となる。

多い (5,017)。また同様に、機械学習 (ナイーブベイズ) の精度分析からも多様性がうかがえる (色:0.966, 柄:0.948, 印象:0.836)。通常、データ数が増えれば精度も上がることが予想される一方で、要素「印象」はさほど上がっていないからである。

例に挙げた“シフォン”に複数の視点があったとしても、それを全ての記述・公開はできない。“軽やか”と書けば、他の視点に基づく検索は困難になる。このため、非明示の感性語でも検索や提案に使えるよう機械学習が必要となる。次に、学習のための提案～要素「効果」を使用する方法～について述べる。

### 2.3. キーワード“演出”について

“演出”は要素「効果」に属し、紹介文中に頻出する語である。文型は主に次の2種類である。型は異なっても、どちらも『対象(○)から印象(●)が喚起される』ことを意味している。

1. ○○を●●に演出  
「スタイルをガーリーに演出」等のように用いる。
2. ○○が●●を演出  
「パフスリーブがガーリー(さ)を演出」等のように用いる。「○○で」「○○することにより」等、表現形態は変わる

前節に述べたように、客観的な事実の記載でもかまわないにもかかわらず、書き手は、あえて主観的な事実を追加した。つまり、この感性語「●●」は書き手にとって大きな意味を持っていたと考えられる。したがって、感性語と周囲の単語との間の出現傾向をコーパスから学習し、未知の紹介文に対し感性語を推定することは重要な課題と言える。

そこで本研究は、この推定を目的のひとつとする。

なお、紹介文中には『女性らしく軽やかな雰囲気演出』のように複数の感性語が出現することもある。このため、学習器の判定の際に、複数の判定結果を出せるよう複数の学習器を用意する。

### 2.4. “演出”の作用範囲について

本研究の前提は『“演出”が印象を喚起させ、それにより、アイテムの印象が推定可能』ということである。しかし感性語の作用範囲に関する知見がない。

一般的には、アイテム全体の印象を示す場合と袖などのディテールの印象を示す場合との2通りが考えられる。今回は2種類のモデルを用意する。

#### 2.4.1 遠距離モデル

“女性らしい”, “ガーリー”等のように、アイテム全体の印象を記述する感性語のためのモデルである(図2)。ここで、青丸の感性語は学習時には正解ラベルとなり、学習後は推定対象となる。

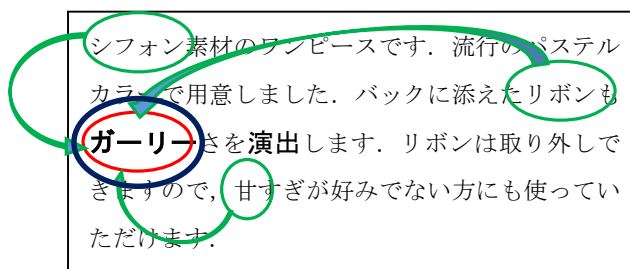


図2 長距離型推定

#### 2.4.2 短距離モデル

一方、“透け感”, “光沢感”等は素材やディテールを記述するために使われていることが多いと予想される。また、前項の感性語においても、特定の構成要素だけを修飾することがあるかもしれない。そのためのモデルが短距離モデルである(図3)。“演出”が出現する1文のみを使用するモデルである。

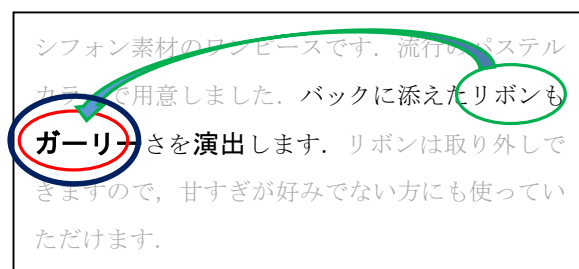


図3 短距離型推定

上記の2モデルに基づいて、本研究は2種類の実験を行う。実験1では遠距離モデル(紹介文すべて)からの、そして実験2では短距離モデル(1文のみ)からの学習を行い、両者の比較によりモデルの妥当性を検証する。

### 2.5. コーディネート推薦研究との関連

コーディネート推薦する研究は多数あり、構成要素に着目した研究がある(Liuら2012, 神間ら2011, 福田ら2011, 佐藤ら2009, 辻田ら2009)。神間ら(2011)では、色やディテール(袖の長さや形等)も利用している。しかし、あらかじめ用意した構成要素しか使用できない点、アイテムに関する知識が更新されない点が問題となる。

しかし、商品紹介文はその時々トレンドを反映しているため、動的な構成要素になる。また、そこから主観的な要素の推定を正確に行うことができれば、『タイトシルエット・サイドスリット。』というシンプルな記載をするサイトでも感性語「女性らしい」「洗練」等の推定が可能となるため、本研究の手法は重要と考えられる..

### 3. 実験と考察

本研究は SVM を用いた感性語の推定実験を行う。先に述べたように、2 モデルの妥当性について吟味する。

#### 3.1. コーパス

次に、本研究で使用したデータについて述べる。

##### 3.1.1 予備調査

まず、アパレルアイテム販売サイトの複数のブランドを対象とし、“演出”という語を含む商品紹介文を収集した(85,000程度)。次に、その“演出”がブランド紹介文中に出現している例を除いた(30,000程度)。残り約 55,000 件のうち 3,000 件を用いて、前述の感性語の絞り込みのため、出現傾向を調査した。出現回数が 10 以上であった 39 感性語を対象とする。なお、「綺麗」「可愛い」等の語は除いた。上位は、女性らしい:525回、こなれ感:129、カジュアル:126、上品:125、フェミニン:122、抜け感:113であった。

表 1 紹介文中の感性語出現回数

女性らしい	9,418	軽やか	662	シャープ	359
抜け感	2,733	レディ	659	スタイリッシュ	356
こなれ感	2,438	ラブ	652	ナチュラル	355
フェミニン	2,396	きちんと感	626	クラシカル	308
上品	2,354	シック	543	クール	289
カジュアル	2,331	スマート	524	女の子	279
すっきり	1,787	爽やか	508	女らしい	241
リラックス	1,579	洗練	488	キュート	234
華やか	1,535	ガーリー	465	モダン	225
大人っぽい	1,070	ふんわり	427	ヘルシー	225
エレガント	941	旬	413	セクシー	196
メリハリ	912	甘い	373	優雅	165
華奢	687	透け感	365	落ち感	152

絞り込んだ 39 語の 55,000 件中の出現回数を調査した(表 1)。3000 件による調査結果と比べ若干の順位変化はある。しかし傾向は変わらず、サンプリングとして有効であることがわかった。

なお、今回は 1,000 回以上の出現を確認した 10 感性語のみを機械学習の対象とする。

#### 3.1.2 ベクトル化

次に、MeCab<sup>3</sup>を用いて形態素解析を行った。本研究では自立語・付属語・記号にかかわらず、1 文または複数文からなる紹介文中に 1 回以上出現した形態素をすべて 1 として扱う。また、個々の推定実験においては、正解ラベルとなる感性語を 0 とし、文中には出現しなかったこととして扱う。つまり、ある意味伏字(『バックに添えたリボンも●●を演出します。』)の状態でも学習と推定を行う。

負事例には、先の 39 語のどれもが紹介文中に出現しないこととした。上同様、1 回以上出現した形態素を 1 とする。以上の作業により、5,710 件の商品紹介文が用意された。

この操作を、商品紹介文中のすべての文(実験 1)、演出”が含まれる 1 文のみ(実験 2)に適用する。正解データ数は、表 1 と同一である。

#### 3.2. 実験 1

SVM(リニアカーネル)を用いた 5 分割交差検定により実験を行う。分割の影響を減じるため 5 回繰り返した。実験結果を表 2 に示す。Accuracy は 5 回の平均、Precision は 1 回ごとの変動が大きかったため、最小値と最大値も示す。

表 2 感性語ごと実験結果(商品紹介文)

	accuracy	precision		
		min.	ave.	max.
女性らしい	0.902	0.909	0.922	0.931
抜け感	0.895	0.843	0.852	0.865
こなれ感	0.884	0.801	0.814	0.825
フェミニン	0.933	0.872	0.899	0.927
上品	0.907	0.823	0.839	0.867
カジュアル	0.928	0.877	0.888	0.892
すっきり	0.920	0.806	0.849	0.886
リラックス	0.956	0.899	0.921	0.941
華やか	0.953	0.898	0.917	0.945
大人っぽい	0.942	0.832	0.838	0.847

<sup>3</sup> <http://taku910.github.io/mecab/>

### 3.3. 実験 2

実験 1 同様、5 分割交差検定を 5 回繰り返す。結果を表 3 に示す。

表 3 実験結果 (“演出” が含まれる 1 文のみ)

	accuracy	precision		
		min.	ave.	max.
女性らしい	0.837	0.814	0.818	0.820
抜け感	0.767	0.747	0.782	0.808
こなれ感	0.788	0.712	0.747	0.800
フェミニン	0.820	0.811	0.827	0.859
上品	0.795	0.802	0.807	0.824
カジュアル	0.812	0.774	0.818	0.842
すっきり	0.845	0.774	0.831	0.880
リラックス	0.873	0.850	0.872	0.902
華やか	0.872	0.806	0.850	0.882
大人っぽい	0.895	0.725	0.785	0.836

### 3.4. 考察

#### 3.4.1 accuracy と precision について

実験 1 の結果 (表 2) は accuracy が 0.9 程度、precision で 0.8 強となり、感性語の推測が可能であることを示す。一方で、実験 2 の結果 (表 3) は、さほど高い値になってはいない。1 文だけで推測することの困難さを示唆する。

#### 3.4.2 2 モデルの適切性について

どの感性語についても実験 1 の結果がよかった。2 章に述べたように、1 文からのみで決まる感性語があると考えていた。しかし、実験の結果はそのような感性語の存在を明らかにすることができなかった。

さらに確認のため、素材の特徴を有すると考えられる感性語“透け感”について、10 回の繰り返し実験を行った。この結果、precision の値は「遠 (min:0.794, ave:0.843, max:0.903)」、「近 (0.714, 0.811, 0.892)」となった。この観点からも紹介文全体を使用するモデルがより有効であるとの結論となった。

ただし、表 1 に見られる感性語はそのほとんどがアイテム全体の印象を表す語になっているように思われる。キーワード“演出”は特定の感性語 (すなわち、遠距離型) を発見する機能を持つだけかもしれないため、別の構成要素「効果」に関連するキーワード (“叶える”, “魅せる” 等) を使ってさらなる検討をする必要がある。

#### 3.4.3 precision の高さについて

理由のひとつには、「同一ブランドの商品紹介文は似ている」ことが挙げられる。中にはコピー同様の紹介文も見られる。このため複数の別のアイテムを収集しても、紹介文の内容はほとんど同じで、実験デザイン上は交差検定であっても実際には限りなくクローズドテストになっている可能性もある。

サイト特有の癖の学習は重要である。しかし、ほとんどコピーのデータからの学習は重要とは言えない。この点は検討できなかったため、今回は単に見かけ高いだけなのかについての検討を加えたい。

### 4. おわりに

本研究は SVM を用い、商品紹介文中に含まれる“演出”に着目することで、アパレルアイテムが喚起する印象の自動推定を試みた。実験の結果、高精度で感性語の推定が可能であることがわかった。一方で、本研究が想定した 2 モデルの存在を明らかにすることはできなかった。

今回は要素「効果」の中でも“演出”だけを対象としたことにも原因があると考えられるため今後、様々なキーワードにより広範な検討を行いたい。

#### 謝辞

この研究は、広島市立大学特定研究費 (先端学術研究費平成 26 年度) の補助を得ている。

#### 参考文献

- 福田未央, 中谷善雄(2011). “服の着用経験に基づき服自身がコーディネートを推薦するシステム”, 情報処理学会全国大会講演論文集, pp.161-163.
- 小池恵里子, 伊藤貴之, 渡辺知恵美(2013), “提示量を適正化した女性向け商品検索支援システム”, DEIM Forum 2013.
- 飯塚弘子, 内田八重子, 香川幸子(1985). 服飾デザイン論. 文化出版局.
- 神間唯, 丸谷宜史, 梶田将司, 間瀬健二(2011). “ファッションイメージキーワードに基づいたコーディネートシステムの提案”, 情報処理学会研究報告 2011-HCI-142(26), pp. 1-7.
- 小林茂雄(2003). 装いの心理. アイ・ケイコーポレーション
- 工藤勝江, 山本由紀子, 東陽子, 古御堂誠子, 野原美香(2001). 文化ファッション大系 ファッション流通講座⑤ コーディネートテクニック アパレル編 I (商品 知識), 文化出版局.
- 黒澤義明, 小川湧真, 竹澤寿幸(2014), “ファッションアイテム紹介文の構成要素自動分類”, 言語処理学会年次大会.
- 黒澤義明, 村上大志, 竹澤寿幸(2015), “ファッションアイテム構成要素自動抽出のためのデータ収集と分類”, 言語処理学会年次大会.
- Liu, S., Feng, J., Song, Z., Zhang, T., Lu, H., Xu, C., and Yan, S. (2012), “Hi, magic closet, tell me what to wear!” in Proceedings of the 20th ACM international conference on Multimedia, pp. 619-628.
- 佐藤彩夏, 渡邊恵太, 安村通晃(2009). “姿を利用したファッションコーディネート支援システム suGATALOG の提案と評価”, 情報処理学会論文誌, 53(4), pp.1277-1284.
- 瀬古沢照治, 三橋宏行, 小澤幸夫(2008). “アパレルオンラインショッピングにおけるワン・トゥ・ワンのリコメンドシステム”, 電気学会論文誌 C, 128(8), pp.1333-1341.
- 鈴木洋子, 天野豊久 (2005). 文化ファッション大系 ファッション流通講座⑦コーディネートテクニック 演出編, 文化出版局.
- 富田明美(2004). アパレル構成学 着やすさと美しさを求めて. 朝倉出版.
- 辻田眸, 北村香織, 神原啓介, 塚田浩二, 椎尾一郎(2009). “Asa1-coordinator: 履歴情報を利用したファッションコーディネート支援.” ヒューマンインタフェースシンポジウム 2009 論文集, pp.85-88.