

コーディネート投稿サイトのユーザタグを用いた コーディネート間の類似度に基づく検索ナビゲーションシステム

吉越 優美[†] 北山 大輔[†]

[†] 工学院大学情報学部 〒163-8677 東京都新宿区西新宿 1-24-2

E-mail: †j112120@ns.kogakuin.ac.jp, kitayama@cc.kogakuin.ac.jp

あらまし 近年, wear のようなファッションコーディネート投稿サイトが登場してきている. このようなサイトでコーディネート検索する際には, ブランドや特定のアイテム, ユーザ, タグなどから検索するのが一般的である. しかしユーザの中には, あるコーディネートに対し「もう少しカジュアルにしたらどうなるだろう」といった検索意図が存在するが, 現状の検索方法ではそういった検索意図には対応していない. そこで, ユーザが任意のコーディネートを選択後, 「シンプル」や「カジュアル」などのスタイルを選ぶことで, 選択コーディネートより該当スタイルに近づく類似コーディネートを提供するシステムを提案する. そのことにより, ユーザが徐々に自分のイメージに合うコーディネートを見つけられるようにナビゲートする. また, 本システムの有用性を検証した評価実験の結果を報告する.

キーワード キーワード検索システム, ファッションコーディネート, ユーザタグ, タグ類似度

1. はじめに

近年 wear^(注1) や CodeNote^(注2) などのファッションコーディネート投稿サイトが登場してきている. これらはユーザが自身のコーディネート写真を撮影しサイトにアップロードしたり, 他人のコーディネート閲覧できるシステムになっている. 図1に示すように他人のコーディネートを検索する際は, 使用ブランドや年齢や性別などのユーザの属性といった項目を指定する. しかし, この検索方法は絶対的な意図であり, 今のコーディネートに対して「ややカジュアル寄りにしたい」など相対的な検索意図の場合には不向きである. そこで, 本研究ではそのような検索意図に対して, 任意のコーディネートからユーザの遷移させたいスタイルを選択してもらい, そのスタイルかつ元のコーディネートに似ているものをユーザに提示する検索ナビゲーションシステムを提案する.

2節で提案する手法の概要と, ファッションやユーザタグに関連する研究の紹介を行う. 3節ではコーディネート間の類似度の定義と算出方法について述べる. 4節でスタイルの代表コーディネートの抽出方法とシステムのアルゴリズムについて述べる. 5節ではプロトタイプシステムのユーザインタフェースを示し, 6節で評価について述べる. 最後に7節でまとめと今後の課題について述べる.

2. 本研究のアプローチ

2.1 対象とするコーディネートの構造

本研究ではファッションコーディネート投稿サイト「wear」を利用する. wear ではコーディネート写真と共に投稿ユーザ自身が付与している検索キーワードであるユーザタグや, 使用ブ

ランドおよびアイテムの他, 性別を含むユーザ情報が付与されている. 本研究では, コーディネート写真, ユーザタグの2点を使用し, 類似度の算出ではユーザタグのみを使用する. 例えば図2ではユーザタグはワイドパンツ, スニーカー, UNIQLO などの9個のタグを使用する.

2.2 本研究の概要

本研究では, ユーザが選択したスタイルに基づき, そのスタイルかつ閲覧中のコーディネートに類似するコーディネートにユーザに提示する検索ナビゲーションシステムを提案する. 閲覧中のコーディネートを選択したスタイルに近づくためには, コーディネート間の類似度を定義する必要がある. 我々はユーザタグのコーディネート表現能力を重みとした類似度を定義する.

この類似度に基づき, 最もスタイルを表しているコーディネートに決定することで, 閲覧中のコーディネートとスタイルの代表コーディネートの位置関係が決定され, 近づくことが可能になる. ユーザタグによる距離空間は疎であるため, 次元削減を行うことで, 位置関係をより明確にする.

本研究の貢献は以下の通りである.

- コーディネート間の類似度を定義
- スタイルの代表コーディネートの発見
- 上記に基づくコーディネートのナビゲーション

2.3 関連研究

ファッションコーディネートに関する研究は数多く行われている. 佐藤ら[1]はユーザの日頃のコーディネート写真を撮影し, 上半身と下半身に切り分けて合成することで, 様々なコーディネートを仮想的に試すことが出来るシステムを提案している. 辻田ら[2]は事前に所持している服をシステム上に保存し, 毎日のコーディネートと会う人を登録する. その選択を履歴としてその日会う人と以前会ったときと同じコーディネートになってしまうことを防いだり, 日頃着ていない服を可視化

(注1) : <http://wear.jp/>

(注2) : <http://codenote.jp/>

図 1 wear のコーディネート検索画面

するシステムを提案している。

神間ら [3] は代表アイテム, 色, 柄, 素材, デテール, 配色, シルエットの 7 つの要素によって定義されるクラシックやエレガントなどの 6 種類のイメージキーワードを用いて, コーディネートに関する気づきを促し, 所持品からのコーディネートを紹介するシステムを提案している. 牟田ら [4] は衣服の通販での購入に着目し, 現在ユーザが購入しようとしている衣服に適合するであろう衣服を推薦するシステム Intelligent Closet を提案している. このシステムは衣服の推薦の他, ユーザが選択した衣服によるコーディネートにアバターに着せることができ, ユーザは実際に着用したときのイメージを容易に把握できるようになる。

清水ら [5] はユーザのコーディネート力を向上させるために, 気づき・理解・納得の 3 段階メカニズムを取り入れたロボットによるコーディネート推薦システムを提案している. このシステムではユーザが着用した服装の履歴や, 雑誌などの最新情報を組み合わせて使うことで, より説得力のあるロボット対話を実現している. そのほかに [6] や [7] などがある. これらの研究は衣服の推薦によるコーディネート支援システムであるが, 本研究ではファッションコーディネート検索によりユーザがコーディネートを実体化できるようにすることを目的としている。

また, ユーザタグを使用した推薦に関する研究として, [8] や [9] などがあるが, これらは投稿者以外のユーザが自由に付け外しができるタグが対象であるのに対し, 本研究では投稿者のみのタグである点と, コーディネート投稿サイトの特徴を用いている点で異なる。

3. コーディネート間の類似度の算出方法

本節では, 類似するコーディネート算出するために, コーディネートに付与されたタグに重み付けを行い, その重み付けに基づいてコーディネート間の類似度を算出する手法を提案する. 提案手法は以下のステップから構成される。



図 2 wear に投稿されたコーディネートと付与情報

- (1) タグの重み付け
- (2) コサイン尺度による類似度算出

以下に対象とするコーディネートとそれぞれのステップについて説明する。

3.1 タグの重み付け

コーディネートに付与されているユーザタグには「ニット帽」や「シンプル」といったアイテムやスタイルなどの「コーディネートの特徴を表しているタグ」と, 「みんなへ感謝」や「若作りサークル」といったメッセージやユーザの属性などの「コーディネートに関係のないタグ」の 2 つが存在すると考える. タグを利用してコーディネート間の類似度を算出するためには, まず「コーディネートの特徴を表しているタグ」の重要度を大きくし, 「コーディネートに関係のないタグ」の重要度を小さくする必要がある。

我々は「コーディネートに関係のないタグ」は様々なコーディネートに付与することが出来るため, 同じコーディネートに付与されている他のタグ (共起タグ) が多種にわたるという特徴があると考えた. また, 「コーディネートの特徴を表しているタグ」は図 3 のように不特定多数のユーザが使用しているのに対し, 「コーディネートに関係のないタグ」は図 4 のように特定のユーザが複数回使用しているという特徴がある. そこで各タグの「1 ユーザあたりの共起タグの種類数」によって「コーディネートにどれくらい関係がないか」を表せると考え, その逆数を特徴の度合とする. あるタグ t_i の重みは次式を用いて決定する。

$$w_{t_i} = \frac{|T_{t_i}|}{|U_{t_i}|}^{-1} \quad (1)$$

T_{t_i} は t_i の共起タグの集合, U_{t_i} は t_i の使用ユーザの集合である. 式 (1) によって 1 ユーザあたりの共起タグの種類数が少ないタグほど重みが大きくなるようにする. 表 1 に wear 上から新着順で取得した約 34 万件のコーディネートを用いて提案手法により算出したタグの重み値の一部を示す. この結果から, 「コーディネートに関係のないタグ」と「コーディネートの特徴を表しているタグ」が差別化できていることが分かる。

3.2 類似度の算出

決定した重みを各次元の要素の値としたベクトルを作成し, コーディネート間の類似度をベクトルのコサイン類似度によって測る. ベクトル C_x, C_y のコサイン類似度とは, 以下の式で計算される, 2 つのベクトルのなす角度のコサイン値であり, 値

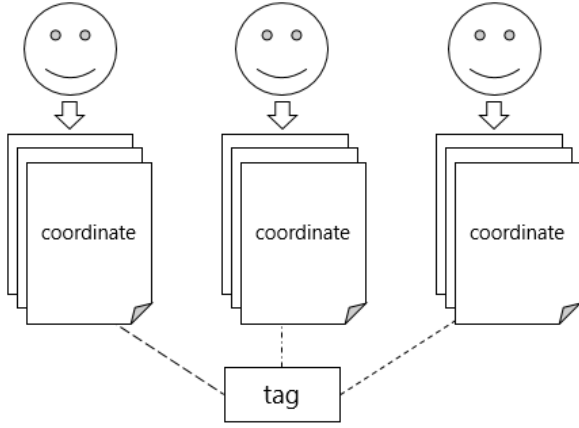


図 3 特徴を表しているタグ

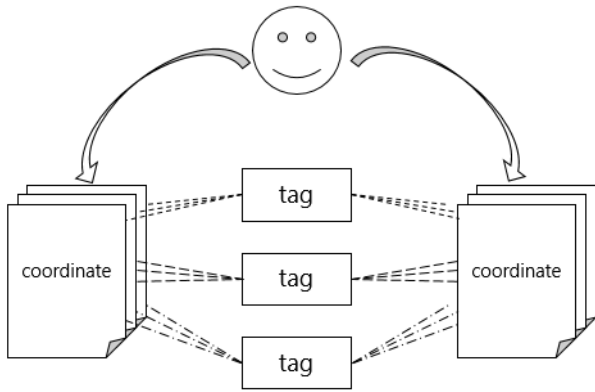


図 4 コーディネートに関係のないタグ

表 1 タグの重み付け結果

タグ名	重み値	共起タグ数	使用ユーザ数
みんなへ感謝	0.011	182	2
若作りサークル	0.094	2465	231
チビーズ*	0.113	1558	176
ニット帽	0.848	8982	7614
シンプルコーデ	0.810	10861	8797
タートルネック	0.907	6032	5470

が大きいほど 2 つのベクトルの類似度が大きいと判断できる。

$$\mathbf{C}_i = (w_{i1}, w_{i2}, \dots, w_{in}) \quad (2)$$

$$\text{Similarity}(\mathbf{C}_x, \mathbf{C}_y) = \frac{\mathbf{C}_x \cdot \mathbf{C}_y}{|\mathbf{C}_x| |\mathbf{C}_y|} \quad (3)$$

「カジュアル」タグが付与されているコーディネート集合を対象に行った、あるコーディネートについて類似度の高いコーディネート 5 件の算出結果を図 5 に示す。この結果から、類似するコーディネートが抽出できていることが分かる。

4. 検索ナビゲーションシステム

本節では、ユーザに提示するコーディネートを決めるために、類似度に基づいて全てのコーディネートを実空間にマッピングし、選択スタイルの代表コーディネート方向の、現在のコーディネートと最も近い点のコーディネートを表示するシステムを提案する。提案システムは以下のステップから構成される。



図 5 類似度の算出例

- (1) スタイルの代表コーディネートの決定
- (2) 多次元尺度法による次元削減
- (3) コーディネートの提示

以下にそれぞれのステップについて説明する。

4.1 スタイルの代表コーディネートの決定

座標上のユーザが選択したスタイルの方向を決定するために、各スタイルの代表となるコーディネートを決定する必要がある。我々は代表コーディネートを図 6 のように「スタイルタグが付与されたコーディネート（スタイルコーディネート）の多くと類似している（類似度合計の高い）コーディネート」と定義した。各スタイルコーディネートの類似度合計は次式を用いて決定する。

$$\text{Sum}(\mathbf{C}_x) = \sum_{k=1}^n \text{Similarity}(\mathbf{C}_x, \mathbf{C}_k) \quad (4)$$

ここで n は、スタイルタグが付与されたコーディネート数である。各スタイルでこの類似度合計が最大のコーディネートを代表コーディネートとする。

4.2 多次元尺度法による次元削減

多次元尺度構成法 (MDS: multi-dimensional scaling) [10] とは、対象間の類似度をもとに、低次元空間に類似したものを近く、そうでないものを遠くに配置する方法である。ユーザタグは種類が非常に多いのに対し、1 つのコーディネートには多くても 10 個程度のタグしか付けられない。そのままでは疎な空間になってしまうため、次元削減を行いコーディネートの位置関係を明確にする。3.2 節で算出した類似度をもとに多次元尺度構成法を用いて次元削減し、各コーディネートを三次元上にマッピングを行う。

4.3 コーディネートの提示

三次元座標上の現在提示されているコーディネートから、ユーザが選択したスタイルの代表コーディネート方向かつ、現在のコーディネートに近いコーディネートを抽出する。選択スタイルの代表コーディネートと現在のコーディネートの距離を直径とする円を範囲とし、ユークリッド距離によって範囲内のコーディネートとの距離を算出する。2 つのコーディネート x, y 間の距離 $d(x, y)$ は次式によって算出される。算出結果から、 d の値が最も小さいものをユーザへ提示する (図 7)。

$$d(x, y) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2} \quad (5)$$

5. プロトタイプシステム

5.1 データセット

システムが使用するスタイルを「カジュアル」、「シンプル」、

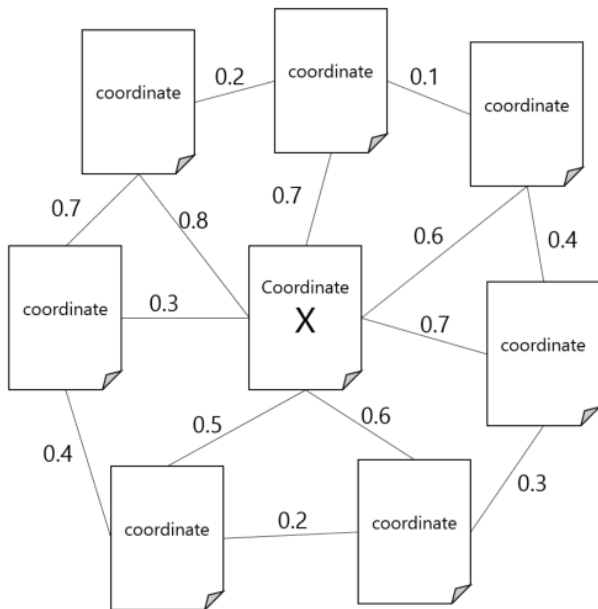


図 6 代表コーディネート

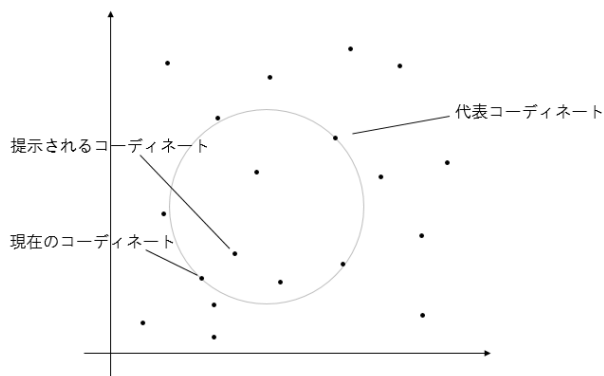


図 7 マッピング結果と提示コーディネートの決定

「ガーリー」、「フォーマル」、「ストリート」、「マニッシュ」、「春」、「夏」、「秋」、「冬」の 10 種類とし、各スタイルの wear 上の人気上位のコーディネート約 100 件、合計約 1000 件のコーディネートで構築を行った。各スタイルのコーディネートは各スタイルがユーザタグとして付与されているものを対象とした。システム内のタグの重みの計算ではプロトタイプシステムで使用する 1000 件のコーディネートの他に、wear 上から新着順で取得した約 34 万件のコーディネートも使用している。

5.2 ユーザインタフェースとシステムの流れ

画面上には 10 種のスタイルがボタンとして設置され、中心に現在のコーディネートが表示される。ユーザはスタイルのボタンを押すことでコーディネートを遷移させることができる。

ユーザはまず 1 件任意のコーディネートを選択し、遷移させたいスタイルを選択する。本システムは現在のコーディネートに対し、4.3 節の処理を行い提示するコーディネートを決定し、出力する (図 8)。そのコーディネートに対し、またユーザがスタイルを選択し、ユーザが好みのコーディネートを見つけるまで繰り返される。



図 8 UI とコーディネートの遷移

表 2 類似コーディネートの実験結果

	提案手法	比較手法
1 組目	81.6%	18.4%
2 組目	86.7%	13.3%
3 組目	87.5%	12.5%

6. 評価

本研究で提案する類似度算出手法および代表コーディネート決定手法の有効性とシステムの精度を確認するための評価実験を行った。本節では、実験の手順と結果、考察について示す。

6.1 類似コーディネートの抽出手法の評価

CrowdWorks^(注3) を用いて集めた 100 人と 20 歳から 45 歳の男女 20 人を被験者として用い、本手法のうち、コーディネートの類似度について評価を行った。プロトタイプシステムで使った約 1000 件のコーディネートを用い、類似するコーディネートの抽出を行った。実験データとして 3 件のベースとなるコーディネート X を用いた。そのコーディネート X に関して、提案手法で抽出したコーディネート A と、重みを考慮せずジャカード係数によるタグの類似度だけで抽出した B を比較することで評価を行う。被験者はどちらがコーディネート X に似ているかを判断した。なお、被験者は各コーディネートが何の手法で抽出したかわからない状態で判断した。

類似コーディネートの評価結果を表 2 に示す。表 2 から、提案手法はどの組でも 8 割を超えており、比較手法より高い値となった。このことより、提案手法は比較手法より人の感性に近い類似するコーディネートの抽出ができるといえる。

6.2 代表コーディネートの抽出

6.1 節と同じ被験者に対し、本手法のうち代表コーディネー

(注3) : <http://crowdworks.jp/>

表 3 代表コーディネートの実験結果

類似度合計値	シンプル	カジュアル	秋
高	13.3%	8.3%	20.8%
中高	66.7%	36.7%	14.2%
中	17.5%	27.5%	34.2%
中低	0.8%	22.5%	5.8%
低	1.7%	5.0%	25.0%

表 4 検索ナビゲーションの精度の実験のスコア

思う	2
すこし思う	1
どちらともいえない	0
あまり思わない	-1
思わない	-2

ト抽出について評価を行った。プロトタイプシステムで使した約 1000 件のコーディネートを用い、代表コーディネート抽出を行った。実験データとして、シンプル、カジュアル、秋の 3 つのスタイルを用いた。それぞれのスタイルについて類似度合計値が最小のコーディネートから最大のコーディネートまで均等になるように 5 件抽出した。被験者は 5 件の中で、一番各スタイルを表していると思うものを判断した。なお、被験者は類似度合計値がわからない状態で判断した。

代表コーディネートの評価結果を表 3 に示す。表 3 は各スタイルに対し上から類似度合計値の高いものになっている。表 3 より、提案手法である類似合計値が最大である上段の値はどのスタイルでも下位となっている。しかし、比較的類似度合計値が高いコーディネートを多くの被験者が選んでおり、類似度合計値による提案手法により判定できる可能性がある。これは、今回代表コーディネートの抽出の際に各スタイルタグが付与されているコーディネート集合を用いている。この集合のコーディネートに必ず含まれるスタイルタグを類似度算出に用いたために、不適切なコーディネートの類似度合計値が高くなったと考えられる。

6.3 検索ナビゲーションの精度

大学生・院生 20 人を被験者として用い、プロトタイプシステムによる、コーディネートの遷移を評価した。実験データとして 10 件のベースコーディネートを用いた。各ベースコーディネートから各スタイルに 2 回連続して遷移させた。被験者はベース、1 回目、2 回目の 3 件について各スタイルに近付いたかを判断した。

検索ナビゲーションの評価結果を表 5 に示す。このとき表 5 は各スタイルについて表 4 によって点数化し、平均値を算出した結果である。表 5 より、スタイルによってバラつきがあるが、どちらともいえない (0) からすこし思う (1) の間になった。このことより代表コーディネートに近付ける本手法により、そのスタイルに近付くといえる。

7. ま と め

本研究では、あるコーディネートについて「ややカジュアルにしたい」という検索意図に対するナビゲーションシステムに

表 5 検索ナビゲーションの精度の実験結果

シンプル	0.80
カジュアル	1.00
ガーリー	0.55
フォーマル	0.30
ストリート	0.45
マニッシュ	0.30
春	0.50
夏	0.25
秋	0.85
冬	0.30

取り組んだ。提案手法ではユーザタグに重みを付け類似度を算出し類似コーディネートと代表コーディネートの抽出を行った。評価実験を行い、比較手法よりも人の感覚に近い類似コーディネートを抽出できていること、検索ナビゲーションの精度を確認した。

今後の課題として以下を考えている。本研究では代表コーディネートの抽出の際にスタイルタグを用いたために、不適切なコーディネートの類似度合計値が高くなったと考えられるため、スタイルタグを除いた類似度合計値の算出を行う予定である。また、検索ナビゲーションの精度を高めるため、タグ以外のメタデータの利用やタグのクラスタリングを行う予定である。

謝 辞

本研究の一部は、平成 27 年度科研費若手研究 (B)(課題番号: 15K16091) によるものです。ここに記して謝意を表すものとします。

文 献

- [1] 佐藤彩夏, 渡邊恵太, 安村通晃. 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, Vol. 53, No. 4, pp. 1277-1284, Apr. 2012.
- [2] 辻田暉, 北村香織, 神原啓介, 塚田浩二, 椎尾一郎. 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論文集, Vol. 2009, pp. 85-88, Sep. 2009.
- [3] 神間唯, 丸谷宜史, 梶田将司, 間瀬健二. 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. 研究報告ヒューマンコンピュータインタラクション (HCI), Vol. 2011, No. 26, pp. 1-7, Mar. 2011.
- [4] 牟田将史, 益子宗, 星野准一. 衣服の通信販売サイトにおける商品探索支援システム. 研究報告エンタテインメントコンピューティング (EC), Vol. 2012, No. 2, pp. 1-5, May. 2012.
- [5] 清水葵, 鈴木優, 上田博唯. ロボット対話型服装推薦システムにおける, ユーザの服装コーディネート力向上検証. 研究報告エンタテインメントコンピューティング (EC), Vol. 2013, No. 2, pp. 1-6, May. 2013.
- [6] Si Liu, Jiashi Feng, Zheng Song, Tianzhu Zhang, Hanqing Lu, Changsheng Xu, and Shuicheng Yan. Hi, magic closet, tell me what to wear! MM '12 Proceedings of the 20th ACM international conference on Multimedia, Vol. 2012, pp. 619-628, Dec. 2012.
- [7] 趙瑩, 荒木健治. ユーザ情報に基づくファッションコーディネート推薦システム (ライフログ, ライフログ活用技術, オフィス情報システム, 情報通信マネジメント, 一般). 電子情報通信学会技術研究報告. LOIS, ライフインテリジェンスとオフィス情報システム, Vol. 111, No. 383, pp. 13-18, Jan. 2012.

- [8] 齋藤敬, 富樫敦, 梶功夫. ソーシャルブックマークデータを用いた推薦アルゴリズムの提案. 全国大会講演論文集, Vol. 72, pp. 825–826, Mar. 2010.
- [9] 住元宗一郎, 中川博之, 田原康之, 大須賀昭彦. コンテンツ投稿型 sns における未知性と意外性を考慮した推薦エージェントの提案. 電子情報通信学会論文誌. D, 情報・システム = The IEICE transactions on information and systems (Japanese edition), Vol. 94, No. 11, pp. 1800–1811, Nov. 2011.
- [10] J.B. Kruskal. Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis. *Psychometrika*, Vol. 29, No. 1, pp. 1–27, May. 1964.