

# 服飾オントロジを用いたECサイトにおける ユーザデザイン嗜好の推定と評価

全 泰賢<sup>1,a)</sup> 川村 隆浩<sup>1,2</sup> 中川 博之<sup>1</sup> 田原 康之<sup>1</sup> 大須賀 昭彦<sup>1</sup>

受付日 2012年1月13日, 採録日 2012年7月2日

**概要:** 近年, 商品やサービスを販売する EC サイトの市場規模がますます拡大している. しかし, 多くの EC サイトの商品検索機能はカテゴリ絞り込みやキーワード検索に限られており, 単語での特徴指定が困難な商品の検索においては難がある. たとえば, EC サイトの中でも特に売上げが伸びている服飾系の商品においては, ユーザは事前にブランド名やデザイン種別などの商品知識を頭に入れる必要があり, それによって商品群を絞り込んだ後にも何十というサムネイル画像を順番に見ていかなければならない. そこで本研究では, ユーザが閲覧した衣服 (デザイン) の履歴を, 商品画像の局所特徴量と商品説明文中のキーワードから, 我々が事前に用意した服飾オントロジ上にマッピングする. そして, オントロジ内の距離計算によってユーザのデザインに関する嗜好を推定し, ユーザ嗜好に沿った衣服を抽出・推薦するエージェントシステムを提案する. また, 推薦精度に関する評価実験を行い, 70%超の再現率を達成していることを確認した.

キーワード: Web エージェント, オントロジ, 概念検索, デザイン推定

## Estimation and Evaluation of User's Favorite Design Based on Fashion Ontology

TAIKEN ZEN<sup>1,a)</sup> TAKAHIRO KAWAMURA<sup>1,2</sup> HIROYUKI NAKAGAWA<sup>1</sup>  
YASUYUKI TAHARA<sup>1</sup> AKIHIKO OHSUGA<sup>1</sup>

Received: January 13, 2012, Accepted: July 2, 2012

**Abstract:** E-commerce market has been expanding, recently. However, search functionality for products in most of e-commerce sites are limited to keyword search and category selection, thus there is a problem to search for the product, to which it is difficult to specify its characteristics by words. In terms of a fashion product whose sales have been increasing year after year, users are required to learn the keywords like brand names and design terms in advance, and after just narrowing a list of the products they need to check small thumbnail images of enormous items one by one. Therefore, we propose an 'agent' service, which maps the user's browsing history of the product designs to the prepared fashion ontology based on local features of product images and keywords in product descriptions. Then, it estimates his/her favorite designs by calculating distances in the ontology and recommends the unseen items to the user. Also, we conducted an experiment for the recommendation accuracy, and confirmed more than 70% of recall.

**Keywords:** Web agent, ontology, search for concept, estimation of design

### 1. はじめに

2010年時点においてインターネットの世帯利用率は90%を超えており, 商品やサービスを販売する EC サイトの市場規模もますます拡大している. しかしながら, EC サイト上での商品の検索機能はあらかじめ分類されたカテ

<sup>1</sup> 電気通信大学大学院情報システム学研究科  
Graduate School of Information Systems, University of  
Electro-Communications, Chofu, Tokyo 182-8585, Japan

<sup>2</sup> 株式会社東芝研究開発センター  
Corporate Research & Development Center, Toshiba Corpo-  
ration, Kawasaki, Kanagawa 212-8582, Japan

a) z.taiken@ohsuga.is.uec.ac.jp

ゴリによる絞り込みや、キーワード検索に限られていることが多く、衣服や家具など、キーワードでの特徴の指定が困難な商品においては、ある程度絞り込んだ後は提示されたサムネイル画像群を順に見ていくしかない。また、事前にブランドやデザインに関する専門用語を知っておく必要があったり、巨大 EC サイトでは 1 カテゴリー内の商品件数が膨大であったりなど、検索にあたってのユーザ負荷が高い。その結果、ユーザの嗜好にマッチし、かつ条件を完全に満たした商品を見つけ出す前に検索を断念してしまうことも多いだろう。さらには、EC サイトごとに商品のカテゴリ分類基準が異なっているなどの問題もある。

そこで本研究では、EC サイトの中でも近年、特に売上げが伸びている服飾系の商品（衣服など）を取り扱う EC サイトを対象に、サイト上でのユーザの検索インタラクションを通じてユーザのデザインに関する嗜好を推定し、嗜好に合った適切な商品を提示するエージェントシステムを提案する。具体的には、まず（男性、女性どちらにも適用できることから）コートを対象とし、特定のサイトの分類基準に依存しないコートデザインに関する服飾オントロジを構築する。なお、本論文においてデザインとは衣服の部分的、全体的な形状、色による分類を表す。次に、EC サイトにおける各コート商品について、首周り、ボタンなど、部位ごとのデザインを商品画像の局所特徴量から分類し、商品説明文におけるキーワードも考慮しながら、服飾オントロジ上の対応するクラス（デザインを表す）にインスタンスとしてマッピングする。そして、ユーザの商品検索時には、閲覧履歴内での各コート商品（のデザイン）の出現頻度から、ユーザの嗜好をオントロジ上の特定のクラスからの距離として表現し、推定されたユーザ嗜好に合った未閲覧な商品（インスタンス）をユーザに推薦する。これにより、一定の閲覧履歴の蓄積後は、ユーザの検索負担を軽減することを目指す。

以下、2 章では本研究のアプローチについて詳細を述べ、3 章ではユーザ嗜好推定とデザイン分類についての実験方法と結果を示し、考察する。また、4 章で関連研究について述べ、最後に 5 章で本論文の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

本サービスでは、事前に服飾オントロジを用意し、EC サイト上の商品を半自動的にオントロジ内の特定のデザインを表すクラスにインスタンスとして割り付けておく。そして、ユーザの EC サイト上のコート閲覧履歴からデザイン嗜好を推定し、それに基づいて未閲覧の商品をユーザに推薦する。サービスの流れを図 1 に示す。

以下の節では、それぞれ服飾オントロジの概要、コート商品のオントロジへのマッピング、ユーザデザイン嗜好の推定処理について述べる。

### 2.1 服飾オントロジの概要

本研究では、特定のサイトの分類基準に依存せず、コート商品とユーザ嗜好をマッピングするベースとして服飾オントロジを構築した。なお、筆者らの知る限り執筆時点においてコートのデザインに関するオントロジは世界的にも存在しなかったため、DBPedia オントロジなどを上位オントロジとすることを念頭に、独自のオントロジを構築した。なお、構築にはオントロジ構築ツール法造 [3] を使用した。

服飾オントロジは、Coat クラスをまず“コートの用途”によって Formal, Middle, Casual の 3 クラスに分類した。Formal クラスはサブクラスとしてスーツ、紳士服などフォーマルな衣服に合わせるようなコートのクラスを持ち、Casual クラスは逆にフォーマルな衣服には合わない、モッズコートやブルゾンなど、主に私服に合わせるようなコートのクラスを持つ。また、Middle クラス内のコートには、スーツや紳士服に合わせるように作られたものではないが、カジュアルにもフォーマルにも着こなせる P コートやダッフルコートなどをサブクラスとして持つ。上位レベルにおいてこのような分類を行ったのは、衣服によってはボタンのデザインや首のデザインが同一であっても、生地や丈、その他デザインコンセプトの違いから用途が異なる衣服が存在し、それらを同一クラス内に混在させてしまうことを避けるためである。上記の 3 クラス以下は、既存の EC サイトのカテゴリ分類や衣服の商品構成について述べているサイトを参考に、それぞれのクラス内で明確な形状の違いを基準とし、異なる観点から特殊化を行った。服飾オントロジの概要を図 2 に示す。Fashion クラス以下のサブクラスは、デザインに応じて Fashion クラスから部位に関するクラスの定義を継承している。また、ルートクラスとなる Fashion クラスには首周り、ボタンなどの衣服の部位ごとに定義されたクラス（以下、リスト参照）がプロパティとしてリンクされている。

- Neck クラス  
首周りの部位のデザインを表す。
- Button クラス  
ボタン部位のデザインを表す。
- Feature クラス  
その他の衣服の特徴的なデザイン、生地を表す。

### 2.2 コート商品のオントロジへのマッピング

コート商品の件数が少数であれば、上述した服飾オントロジクラスへの割付けは EC サイトの運営者、または第 3 者が人手で行ってもよい。しかし、件数が膨大である場合には、以下で述べるコートデザインの分類機能を用いることができる。コートデザイン分類機能は、商品説明文に含まれるキーワードと、商品画像の持つ局所特徴量を用いる。

具体的には、まず説明文（テキスト情報）内に含まれるキーワードから、コートの首、ボタンの部位に関するデザイ

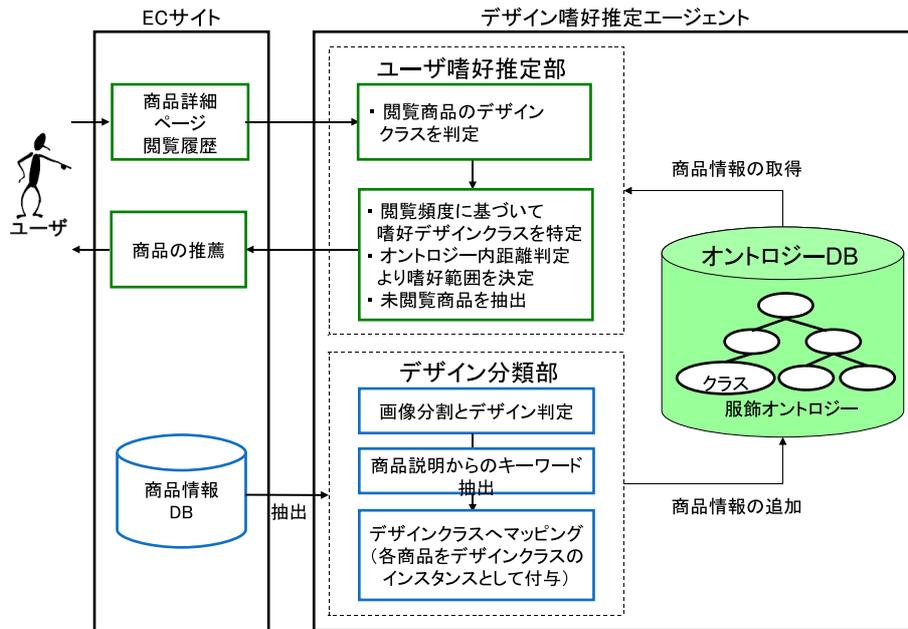


図 1 ユーザデザイン嗜好の推定の流れ

Fig. 1 Estimation process of user's favorite design.

ンの特定を行う。一方で、コート名のみではデザインが不明な部位に関しては、画像解析を用いて分類を行う。画像解析には、局所特微量抽出アルゴリズムを用いている。画像から抽出することのできる特微量には、大域特微量と局所特微量が存在し、大域特微量とはカラーヒストグラムなど画像全体から抽出することのできる特微量であり、局所特微量とは画像の局所的な領域から抽出される特微量である。いずれも特微量はベクトルとして表現され、これの照合により画像間の類似度判定を実現する。本研究では、色情報よりも形状情報を重視し局所特微量を画像間の類似度判定に用いる。なお、局所特微量はグレースケール画像からの輝度変化を抽出に用いるため色情報が考慮されていないが、画像のスケールの変化や画像の回転、隠蔽などに対し頑健性を持つ。また、局所特微量抽出アルゴリズムには SIFT [1], SURF [2] などが存在し、1つの特徴点は128次元の特徴ベクトルによって表される。本研究では、SURFを用いてコート名のみでは判別することのできない部位のデザインの分類を行った。SURF アルゴリズムは、代表的な局所特微量抽出アルゴリズムである SIFT に対し、判別精度は劣るものの処理速度に優れるという特徴があり、大量の商品画像を処理するのに適していると判断したためである。

オントロジーへのマッピング処理では、まず商品画像群を首、ボタンなどの部位ごとに分割し、それぞれの画像から局所特微量を抽出する。そして、あらかじめクラスごとに用意しておいた画像群から抽出した局所特微量とのマッチングを判定し(2画像間で特徴ベクトル間のユークリッド距離を測定し、閾値を上回った場合マッチングしたと見なす)、マッチング数の合計に基づいて対象衣服のデザイン

(クラス)を決定する。この処理をあらかじめ EC サイト上の衣服群に対して行い、服飾オントロジー上のクラスにインスタンスとして付与しておく。

### 2.3 ユーザデザイン嗜好の推定

商品検索時には、ユーザは EC サイト上で衣服の検索を開始し、気になる商品やイメージに近い商品のサムネイル画像をクリックしては詳細ページを閲覧する行為を繰り返すものとする。この間、サービスはユーザがアクセスした詳細ページ内のメイン画像を記録していき、一定以上の履歴が溜まった時点でユーザの嗜好として推定する。ユーザの嗜好としては、閲覧履歴に基づいた最も嗜好に合うデザイン(ユーザが閲覧したインスタンスが最も多いクラス)、およびそれを中心とした一定距離内のクラス群を嗜好と見なし、それらのインスタンス(コート商品)を推薦する。オントロジー内でのクラス間の距離の判定には、概念間類似度判定手法を用いる。

概念間類似度判定手法は、溝口ら [4] が定義した以下の式 (1) を用いている。溝口らの手法は、Li ら [5] が提案したオントロジー上の2クラス  $C_x, C_y$  間の類似度判定手法にオフセット値を与え、階層ごとの類似度下限値を設定することで、スーパークラスとサブクラス間のリンクを表すエッジの数により類似度が0に漸近することを防いでいる。

$$Sim(C_x, C_y) = (1 - offset)Sim_{Li}(C_x, C_y) + offset \quad (1)$$

式 (1) の  $Sim_{Li}(C_x, C_y)$  は Li らの手法である以下の式 (2) により算出される。パラメータ  $(a, b)$  は比較するクラス間の最小距離(エッジ数)  $l$ , ルートクラスからの階層の深

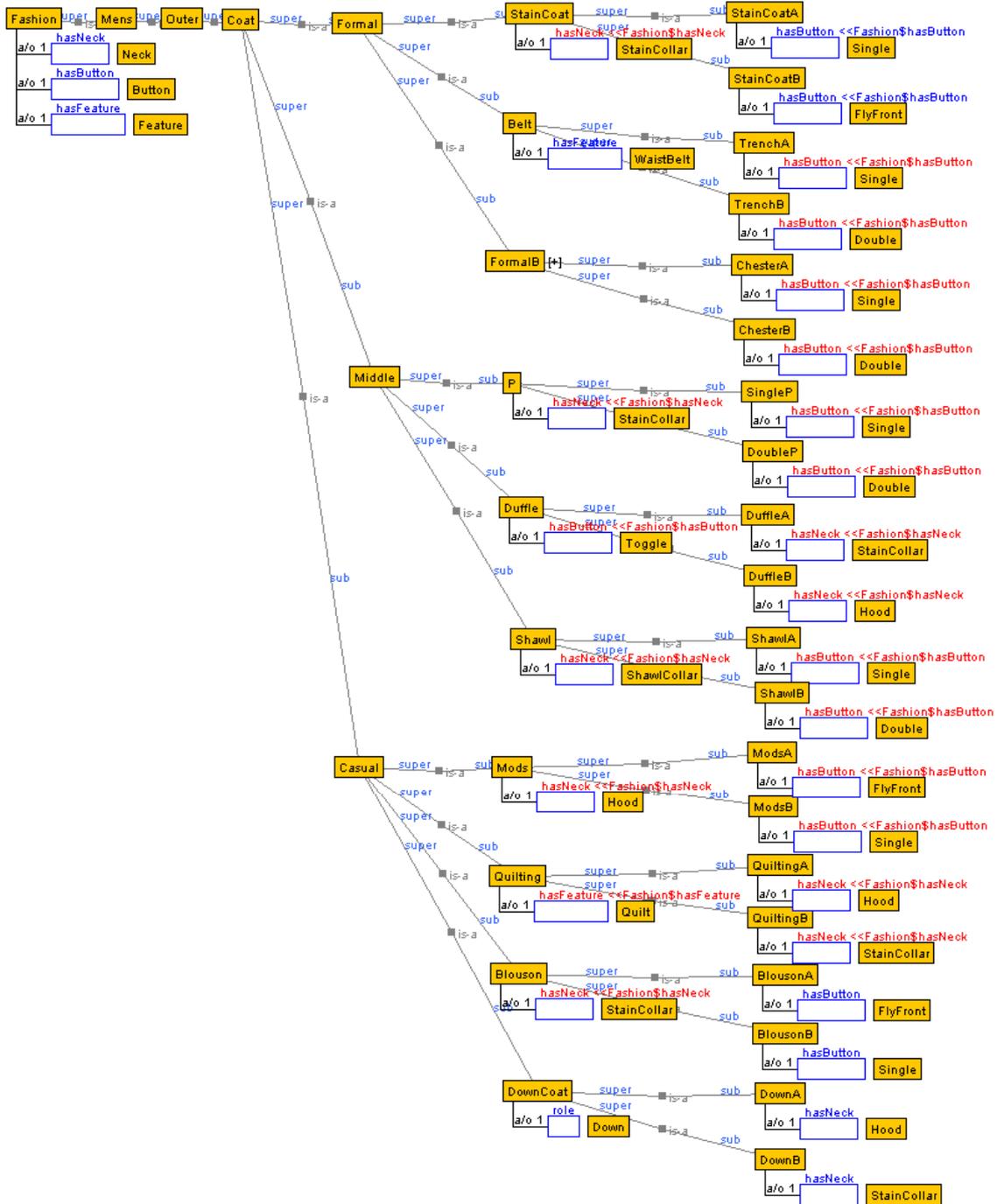


図 2 服飾オントロジ (一部)  
Fig. 2 Part of fashion ontology.

さ  $h$  に対する重みであり、本実験では Li らの実験において最も良い成績を残した組合せを用いている。

$$Sim_{Li}(C_1, C_2) = \exp^{-al} \frac{(\exp^{ah} - \exp^{-bh})}{(\exp^{bh} + \exp^{-bh})},$$

where  $a = 0.2$  and  $b = 0.6$  (2)

同様に、式 (1) の  $offset$  は Li らの手法に重み付けを行う変数  $w$  を与えた以下の式 (3) により算出される。変数  $w$  は比較するクラス間の共通の親クラスである  $LCS$  の階層ごとにオフセット値による最小類似度を定めるための重み

(0 から 1) であり、溝口らが実験に基づき設定した値 0.4 を用いる。

$$offset = w \times \exp^{-a \times 3} \frac{(\exp^{b(h-1)} - \exp^{b(h-1)})}{(\exp^{b(h-1)} + \exp^{b(h-1)})} \quad (3)$$

### 3. 評価実験

#### 3.1 ユーザ嗜好の推定実験

本章では、まずユーザー嗜好に合ったコート商品の推薦結果について述べる。

3.1.1 実験方法および実験結果

データセットとしては、ECサイト SELECTSQUARE [6], および楽天市場 [7] のコートカテゴリから収集した計 288 件の衣服画像を用いる。なお、各衣服のデザインはあらかじめ人手でオントロジ内の該当クラスにマッピングしておく（次節では、コートデザインのオントロジへのマッピング、デザイン分類について評価するが、ここでは両者を切り分けて評価する）。本実験で定義したデザインを部位ごとにそれぞれ図 3, 図 4 に示す。また、各デザインの定義を以下に述べる。

- ボタンのデザイン
    - (a) シングルフロント  
ボタンが 1 列のもの
    - (b) ダブルプレスト  
ピーコートなどのように、ボタンが 2 列のもの
    - (c) フライフロント  
ボタンやジップを布と布の間で挟み、表からはボタンが見えないようにしたもの
    - (d) トグルフロント  
ダッフルコートに用いられているトグルという留め木と、トグルと対となるループが付属しているもの
  - 首周りのデザイン
    - (a) ステンカラー  
ワイシャツなどに用いられている一般的な襟
    - (b) ショールカラー  
丸みを帯びた襟
    - (c) ノッチドラペル  
背広などに用いられている、刻みの入った襟
    - (d) フード  
頭巾風のかぶりものが付属しているもの
- まず、20 代男性 9 名の被験者に 288 件の衣服のうち、自

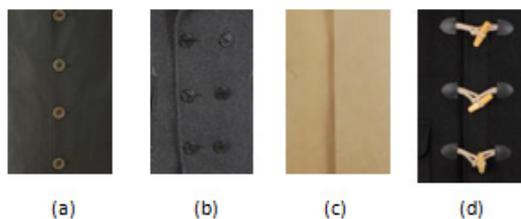


図 3 ボタンのデザイン分類  
Fig. 3 Classification of button design.



図 4 首周りのデザイン分類  
Fig. 4 Classification of design around the neck.

身の嗜好に合うと思われる衣服をチェックしてもらい、正解データの作成を行う（チェック数に制限は設けない）。次に、データセット 288 件を 9 つに分割し、8 つのデータセットを訓練データとし、残り 1 つのデータセットを実験対象とする。そして、訓練データ内の正解データの中で 20% 以上を占めるコートデザインをユーザの嗜好とし、該当するクラス（複数可）と、上述した概念間類似度判定に基づいて類似度が 0.7 以上のクラスに含まれるインスタンス（衣服）を抽出する。そして、その中に正解データが含まれているかどうかを適合率、再現率で評価した。上記の方法を訓練データを切り替えながら 9 回繰り返す交差検定により、被験者ごとに適合率、再現率を算出した。実験用に作成した衣服検索システムのインターフェース、結果画面をそれぞれ図 5, 図 6 に示す。実験システムでは、インターフェース画面でチェックを行った衣服からユーザの嗜好を推定し、結果画面で嗜好にあてはまる衣服を提示している。実験結果を表 1 に示す。

3.1.2 実験結果の考察

実験結果（表 1）では適合率は平均で 16% という結果であったが、再現率は平均でも 73%, 80% を超えた被験者も 9 名中 3 名おり、良い結果であったといえるだろう。EC サイトにおける商品の推薦システムは、多少嗜好に沿わない商品を含めてもユーザの嗜好に沿う商品を確実に提示することが重視されるため、本実験結果はそれに沿う形になっているといえる。現状、多くの EC サイトがキーワード検索とカテゴリ検索しか提供していないのも、同じ理由（ユー



図 5 実験システムのインターフェース  
Fig. 5 Screenshot of experimental system.



図 6 実験システムの結果画面  
Fig. 6 Screenshot of experimental system.

表 1 実験結果

Table 1 Experimental result.

被験者	再現率	適合率
A	0.94	0.23
B	0.61	0.06
C	0.63	0.11
D	0.87	0.30
E	0.75	0.23
F	0.96	0.10
G	0.44	0.07
H	0.63	0.13
I	0.72	0.21
平均	0.73	0.16

ザに多くの商品を見る負荷を強いても、機会損失を最小にする戦略)であると考えられることから、再現率を重視した今回の結果は実展開時に EC サイト側に受け入れられやすいものであるだろう。しかし、推薦結果が膨大になってしまうとそもそもの研究の意義を失ってしまうため、今後はランキング形式を取り入れることなどを検討したい。

次に、実験システムの内部ログなどから適合率と再現率の低下要因について考察する。まず、適合率を下げた要因について、被験者 1 人 1 人の正解データを見ていくと、同一の色、首、ボタンのデザインでもチェックされていない衣服が多く見られた。このことから、現状のシステムが特徴として取得できていないデザイン（シルエット、風合いなど）が主な原因として考えられる。たとえば、色はユーザ嗜好の大きな要素を占めていると考えられるが、同じグレーであっても、明めのグレー、暗めのグレーなど実際には多くの種類が存在する。次節に示すデザイン分類においてあらかじめテキスト情報から色名を取得しておくことができるが、予備実験によれば黒、白、赤といった名前による分類では精度があまり向上しないことを確認している。そこで現在、テキスト情報から大まかな色を特定したうえで、商品画像からカラーヒストグラムをとるなどといったアプローチを検討している。そのほかにも、商品画像の写り方（暗い、小さい）によってチェックされなかった場合もあると考えている。この点に関しては、EC サイト側でも改善の動きが見られると同時に、ユーザの閲覧履歴が多くなれば精度の向上が期待できるだろう。ただし、あまり長期間にわたる履歴収集は、ユーザ嗜好の変遷、季節の移り変わりなど外部要因の変化に影響されるため、一定期間で区切るなどの仕掛けが必要と考えている。

また、再現率を下げた要因としては、9 つに分割したデータセット内のデザインの偏りがあげられる。分割は各デザインが均等に入るよう意識して行ったが、同じ服ではないため完全に均一ではなく、対象データにだけ存在するデザインや、訓練データでの出現数が少ないデザインも存在した。さらに、被験者の嗜好に合ったデザインが、データセッ

ト内にそもそもわずかしかが存在しない場合も、今回は出現割合により嗜好を推定しているため、抽出することができていない。ユーザ嗜好と判定する出現割合に関しては、閲覧履歴の数、EC サイト上の商品数などによって変更していく仕組みが必要だろう（一方で、上述したように適合率を考慮せず、1 回でも出現したデザインは嗜好に合うものと見なすという方法も検討の余地があるだろう）。また、これらはいわばブートストラップ問題でもあるため、データセットの大規模化や履歴の蓄積によって改善する可能性がある。

### 3.2 コートデザインの分類実験

次に、コート商品群を服飾オントロジの各クラスにマッチングするデザイン分類について述べる。なお、データセットは上記と同じ 288 件のコートを対象とする。

#### 3.2.1 局所特徴量によるデザイン分類実験

ここでは、まず商品画像の局所特徴量のみを用いてデザイン分類を行う。

デザイン分類の処理手順は以下のとおりである。まず EC サイトから収集した画像群を、それぞれ首周り、ボタンの 2 つの部位ごとに分割し、人の目で見ると思われるデザインを付与する（正解データ）。次に分類対象の衣服画像とは別の EC サイトから首周り、ボタン部位のデザインごとに 30 件ずつ収集した画像群と 1 : 30 の特徴点のマッチングを各画像ごとに行い、画像間のマッチング数の最大値と平均値をカウントしておく。そして最も最大値が大きかったデザインを、その衣服のデザインであると分類する。最大値が同値であるデザインが 2 つ以上存在した場合は、平均値が最も大きかったデザインとする。そして分類されたデザインがあらかじめ付与しておいたデザインと同一であるかによって正解率の算出を行う。なお、特徴点どうしがマッチングしたと見なされるユークリッド距離の閾値は、過去の実験の結果から首周りを 0.3、ボタン部位を 0.5 に設定し、上記の実験方法に従いそれぞれ特徴点のマッチングを行った。EC サイトから収集した画像、デザインの画像 30 件間のマッチングのイメージを図 7 に示す。これはユークリッド距離の閾値を上回った特徴点どうしを線で結んだものである。

正解率は首周りが正解かつボタン部位が正解であった場合、首周りのみ正解であった場合、ボタンのみ正解であった場合をそれぞれ算出した。また、ボタン部位から特徴点がいっさい抽出されなかった場合はフライフロントデザインと見なした。これは、別途行ったデザイン分類実験の結果から、フライフロントデザインはボタンがないという特徴から特徴量がいっさい抽出されないことが分かっているためである。デザインの分類結果を表 2 に示す。

#### 3.2.2 テキストおよび局所特徴量によるデザイン分類実験

次に、コートの詳細情報ページに含まれているテキスト

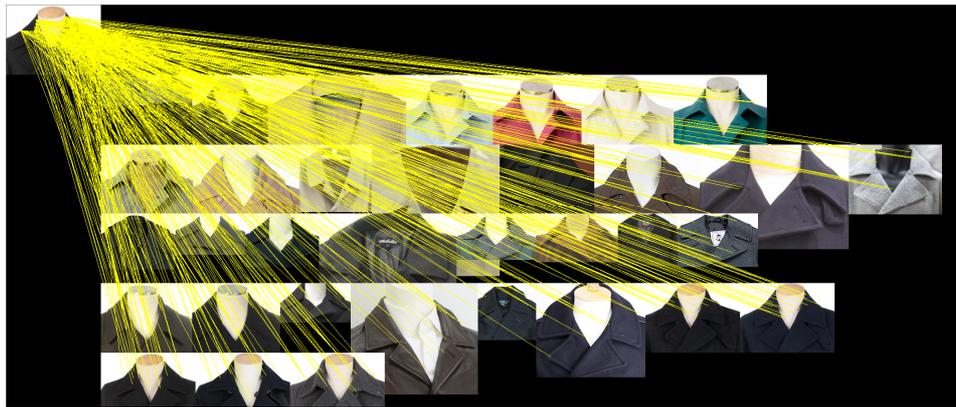


図 7 画像間の特徴点マッチングイメージ  
Fig. 7 Image matching of local features.

表 2 局所特徴量を用いたデザイン分類結果

Table 2 Design classification by local features of product images.

部位	正解率
首周り∩ボタン	0.23
首周り	0.49
ボタン	0.42

表 3 テキストおよび局所特徴量を考慮したデザイン分類結果

Table 3 Design classification by local features of product images and keyword of product descriptions.

部位	正解率
首周り∩ボタン	0.68
首周り	0.94
ボタン	0.70

情報から抽出したキーワードと商品画像の局所特徴量に基づいてデザイン分類を行う。

具体的には、ダッフルコートは首周りがステンカラー（もしくはフード）、ボタンはトグルフロントであるというように、コートの種類ごとに対応する部位デザインをあらかじめ定義しておき、“ダッフルコート”というキーワードが得られた際に対応表に基づいてデザイン进行分类する。そして、キーワードでは確定することのできないデザインについて、上述した商品画像の局所特徴量に基づいて判別する。正解率は前項と同様の方法で算出した。デザインの分類結果を表 3 に示す。

### 3.2.3 実験結果の考察

デザイン分類機能は、商品件数が膨大であった場合に各商品のオントロジへのマッピングをサポートするものがあるが、平均して約 7 割程度の精度で分類に成功しており、一定程度の有効性が確かめられたといえるだろう。なお、今回はすべて 1-best の結果であるが、2, 3-best の解を出し、最終的に EC サイト側（商品販売側）に 2 択, 3 択を行ってもらう形とすることも現実的である。

次に、首周り、ボタンそれぞれの正解率の低下要因について考察する。局所特徴量のみを用いた場合の実験結果（表 2）において、首周りの部位画像は主にエッジ部分から多くの特徴点が抽出されているが、これに対して誤った部分の特徴点が数多くマッチングしていた。たとえば、ノッチドラベル、ショールカラーから抽出される特徴点は主にエッジであり、特徴的なデザインである襟からはほとんど特徴点が抽出されなかった。これは襟やフードが立体的に写されていることにより、輝度の変化から得られる特徴点が他のデザインと比較して少なかったことが原因と考えられる。結果的にノッチドラベル、ショールカラーは主にフードと分類されたが、これはフードは構造的に複雑かつ不規則であり、フード部分の特徴点と誤った部分（肩など）の特徴点とのマッチングが数多く発生してしまったためである。なお、この問題に対してはマッチングを判定するユークリッド距離の閾値を下げることで、一定程度ノイズを除去できることを別の実験で確認した。一方で、フードとステンカラーについては正しく分類可能であることが分かった。さらに、画像情報を補う意味でテキスト情報も利用した場合の実験結果（表 3）では、首周りの分類精度は高い正解率を示した。これはコート名のみで首周りのデザインを一意に特定することのできるデザインが多数存在したためである。

また、局所特徴量のみを用いた場合の実験結果（表 2）において、ボタン周りの部位画像は主にシングルフロントとダブルプレスト間のデザイン判定が難しく、全体の精度を低下させた。局所特徴量は主にボタン形状から抽出されているが、画像内でのボタンの位置は考慮していないためである。局所特徴量を用いた画像の類似度判定は、画像の拡大・回転などに頑健性を持っているが、特徴点の位置関係は考慮しない。そこで、重要な特徴点のマッチングには重み付けを加えるなどの改良が必要であると考えている。一方、トグルフロントに関しては、紐部分、トグルの複雑な形状から多くの特徴点が抽出され、正しく分類すること

ができた。また、フライフロントの場合は、特徴点がほとんど抽出されないことが特徴となり、正しく分類することができている。なお、テキスト情報も利用した場合の実験結果(表3)において、ボタン周りの分類精度は7割にとどまった。これは首周りとは逆に、コート名のみではデザインを特定することのできないコートが多数存在したためである。このようにテキスト情報は、コートの種類や部位によってはデザインを特定することに役立つが、すべての部位のデザインを特定することにはできない。画像特徴量とテキスト情報には、色やデザイン部位ごとに扱いに得手不得手があり、両者を組み合わせることで一定の精度を実現している。

#### 4. 関連研究

ユーザ意図を反映した検索手法としては以下の2つがあげられる。浅野ら[8]は、テキストベースの検索結果から得られた画像群に対し、ユーザが意図を反映できるとされる画像を選択し、それらをブロックに分割、置き換えを行い、新たなクエリ画像を生成して検索、というフローを繰り返すことでユーザの求める画像を見つけ出す手法を提案している。しかし、この手法はカラーヒストグラムのみを類似度判定の特徴量として扱っているという点に加えて、候補群をブロックに分割しクエリを構成するという点で、衣服の部位ごとのデザインの分類には適さないものと思われる。また、安川ら[9]は従来の画像検索においてユーザによる画像領域に対する注目点に着目し、選択した画像群の特徴量の分布の分散に注目し、ある画像特徴量の分散が小さければユーザはその特徴量に着目していると判断する。しかし、この手法では浅野らの手法と同様に、ユーザに自身の嗜好を入力させる操作が必要となる。そのため、今回対象としている衣服のように自身の嗜好や求める商品のイメージを明示的に指示することが困難な場合には適さないものと思われる。

次に、既存の類似画像検索エンジンのうち、形状情報を特徴量として取り扱っていると思われるエンジンの特徴と衣服検索に適用するうえでの問題点について述べる。TinEye[10]は画像を入力とし、類似した画像を提示する検索エンジンである。しかし、類似度判定に用いられる閾値は形状情報が同一である画像しか提示されないよう厳しく設定されており、類似商品を閲覧したいといった検索には適さない。また、Gazopa[11]はキーワード、画像、スケッチ画像など扱えるクエリが充実しており、画像を入力とした検索の際にはキーワードによる絞り込みも可能である。さらに、FashionNavi[12]は本研究と同様に衣服を検索対象とし、形状情報に加えて色情報も検索条件として扱うことが可能な検索エンジンである。しかしながら、いずれの検索エンジンにおいても局所特徴量の主な用途は画像内に存在する物体の分類であり、部位ごとのデザインは考慮さ

れていない。したがって、今回のようにコートという分類がすでに定まっている場合、単純な適用では十分な精度を実現することはできないだろう。

最後に、衣服のための概念検索手法として、Gefuら[13]の研究をあげる。Gefuらは、本研究と同様にECサイトにおける検索機能の問題を指摘し、衣服の商品情報に関するオントロジを構築し、それを用いた概念検索による衣服検索手法を提案している。Gefuらの提案するシステムのインタフェースはクエリとしてキーワードを入力し、システムは拡張クエリ候補を提示するなどしてユーザの検索を支援している。しかし、この手法では前もって商品のブランドやデザインに関する知識と、求める商品の明確なイメージを持つことが必要となる。一方、本研究では衣服や家具のようにデザインに左右される商品の場合、自身の嗜好や求める商品のイメージはユーザ自身明確でない場合も多いと考え、商品閲覧履歴を通してユーザの持つ潜在的なデザイン嗜好を推定し、推定した嗜好に基づいて商品を提示する手法を提案した。

#### 5. まとめと今後の課題

本論文では、衣服などを取り扱うECサイトにおける検索支援のため、ユーザのデザイン嗜好を推定し、嗜好に合った商品を提示するエージェントシステムを提案した。そして、実験を通じて一定程度の推定精度を示し、提案手法の可能性を示した。今後、取得するデザイン部位を増やすとともに、データセット・閲覧履歴を大規模化することでさらなる精度向上が図れるものと考えている。

今後の検討課題として、ユーザ嗜好推定、デザイン分類については3.1.2項、3.2.3項においてすでに述べたが、別途、システムのインタフェースについても検討する必要があると思われる。現在は、ユーザがサムネイル画像を見て詳細ページへアクセスした商品を閲覧履歴に加えているが、詳細ページの拡大画像を見て自身の求める商品ではないと判断する場合もある。そのため、(ユーザ負荷とはなるが)「いいね!」ボタンなどを設置してユーザから嗜好に合っている旨を伝えてもらう処理や、あるいは詳細ページでの閲覧時間やスクロール数からユーザ意図を判別する処理の導入を検討していきたい。

#### 参考文献

- [1] Lowe, D.G.: Distinctive Image Features from Scale-Invariant Keypoints, *International Journal of Computer Vision*, Vol.60, No.2, pp.91-110 (2004).
- [2] Bay, H., Ess, A., Tuytelaars, T. and Gool, L.V.: SURF: Speeded-Up Robust Features, *Computer Vision and Image Understanding*, Vol.110, No.3, pp.346-359 (2008).
- [3] 法造, 入手先 (<http://www.hozo.jp/hozo>).
- [4] 溝口祐美子, 中本利明, 浅川一満, 長野伸一, 稲葉真純, 川村隆浩: オントロジーを利用した文書間のセマンティックな類似度計算手法, 電子情報通信学会技術研究報告,

- AI2008-15, pp.1-6 (2009).
- [5] Li, Y., Bandar, Z.A. and McLean, D.: An approach for measuring semantic similarity between words using multiple information sources, *IEEE Trans. Knowledge and Data Engineering*, Vol.15, No.4, pp.871-882 (2003).
  - [6] SELECTSQUARE, available from (<http://www.selectsquare.com>).
  - [7] 楽天市場, 入手先 (<http://www.rakuten.co.jp/category/mensfashion/>).
  - [8] 浅野美香, 堀田政二: 検索意図を反映したクエリの作成と類似検索, 映像情報メディア学会技術報告, Vol.34, No.34, pp.21-22 (2010).
  - [9] 安川和希, 加藤俊一: 複数の例示画像を用いる画像検索システムの構築, 映像情報メディア学会技術報告, Vol.34, No.18, pp.15-18 (2010).
  - [10] TinEye, available from (<http://www.tineye.com/>).
  - [11] Gazopa, available from (<http://www.gazopa.com/>).
  - [12] FashionNavi, available from (<http://visseeker.yahoo-labs.jp/fn/help.html>).
  - [13] Gefu, Z. and Zhao-hui, H.: Design of a Semantic Search Engine System for Apparel, *IEEE International Conference on E-Business and E-Government*, pp.1414-1417 (2010).



全 泰賢

2010年東京工科大学コンピュータサイエンス学部卒業。2012年電気通信大学大学院情報システム学研究科修士課程修了。同年 NEC ソフト (株) に入社, 現在に至る。主にソフトウェアの開発に従事。



川村 隆浩 (正会員)

1992年早稲田大学理工学部電気工学科卒業。1994年同大学大学院理工学研究科電気工学専攻修士課程修了。同年 (株) 東芝入社。現在, 同社研究開発センター主任研究員, 工学博士。2001~2002年米国カーネギーメロン大学ロボット工学研究所客員研究員。2003年より電気通信大学大学院情報システム学研究科客員准教授。2007年より大阪大学大学院工学研究科非常勤講師。主としてマルチエージェントシステム, セマンティック Web の研究・開発に従事。人工知能学会会員。



中川 博之 (正会員)

1974年生。1997年大阪大学基礎工学部情報工学科卒業。同年鹿島建設 (株) に入社。2007年東京大学大学院情報理工学系研究科修士課程修了, 2008年同大学院博士課程中退。同年より電気通信大学助教, 現在に至る。エージェントおよび自己適応システム開発手法の研究に従事。電子情報通信学会, IEEE CS 各会員。



田原 康之 (正会員)

1966年生。1991年東京大学大学院理学系研究科数学専攻修士課程修了。同年 (株) 東芝入社。1993~1996年情報処理振興事業協会に外向。1996~1997年英国 City 大学客員研究員。1997~1998年英国 Imperial College 客員研究員。2003年国立情報学研究所入所。2008年より電気通信大学准教授。博士 (情報科学) (早稲田大学)。エージェント技術, およびソフトウェア工学等の研究に従事。日本ソフトウェア科学会会員。



大須賀 昭彦 (正会員)

1981年上智大学理工学部数学科卒業。同年 (株) 東芝入社。同社研究開発センター, ソフトウェア技術センター等に所属。1985~1989年 (財) 新世代コンピュータ技術開発機構 (ICOT) 外向。2007年より電気通信大学大学院情報システム学研究科教授。工学博士 (早稲田大学)。主としてソフトウェアのためのフォーマルメソッド, エージェント技術の研究に従事。1986年度情報処理学会論文賞受賞。現在, IEEE Computer Society Japan Chapter Chair, 人工知能学会理事。電子情報通信学会, 人工知能学会, 日本ソフトウェア科学会, IEEE CS 各会員。