

インターネット通販におけるカタログ再構成手法 洋服購入における評価プロセスモデル

白田 由香利
学習院大学 経済学部 経営学科
〒171-8588 東京都豊島区目白 1-5-1
yukari.shirota@gakushuin.ac.jp

インターネットの普及によりインターネット通販を行うサイト数が急増している。消費者の立場からすると、商品数が多いことは嬉しい反面、多数のサイトから購入するサイトを決定し、多数の商品群の中から購入商品を選択することは非常に時間と手間のかかる煩雑な作業となってしまう。特にデザインに依存して購入を決定する商品の場合、手間がかかる。本稿では対象商品を婦人服とし、インターネットでの婦人服購入を支援する高機能ナビゲーションシステム e-Stylist について論じる。本 e-Stylist システムは、多数の HTML カタログを入力すると、消費者の嗜好とセンス、及び購入条件に合った商品を内部評価機能によって選択し、それらの HTML カタログを再構成し、見易い形式で消費者に呈示してくれる機能をもつ。また、その実現のため、e-Stylist は HTML カタログファイルから消費者が評価を行うために必要な意味構造を抽出し、比較が容易な統一形式表現でデータベースに格納する。本稿では、まず婦人服選択・決定プロセスのモデル化を行う。そして、そのモデル化に基づき設計した、e-Stylist のデザイン評価部システム構成を説明する。

Catalog Restructuring in Internet e-Shopping Sites

-- Evaluation Process Model to Select and Purchase Dresses --

Yukari Shirota
Faculty of Economics, Gakushuin University
1-5-1 Mejiro, Toshima-ku, Tokyo, Japan
yukari.shirota@gakushuin.ac.jp

The number of e-shopping sites has been drastically increasing by the growth of the Internet. It is, however, distressing for consumers to select and purchase goods from a lot of sites and from a pile of catalogs because the comparison tasks are time-consuming and tire them. Especially for ladies to select their dresses in the Internet shopping, the tasks are time-consuming. To solve the problem, the paper proposes an intelligent navigation system named e-Stylist to select dresses. Input HTML catalogs, the e-Stylist system evaluates them and selects good ones based on consumer's requirements. Then the e-Stylist system restructures the catalogs to display them in a unified format. Internally, the system extracts semantic structures from HTML catalog files and expresses the semantic information on the same scheme. In the paper, we first make an evaluation process model to select and purchase dresses. Then we describe a system architecture of a design evaluation part in our e-Stylist system which was designed based on the evaluation process model.

1 はじめに

インターネットの普及によりインターネット通販を行うサイト数が急増している。消費者の立場からすると、商品数が多いことは嬉しい反面、多数のサイトから購入するサイトを決定し、多数の商品群の中から購入商品を

選択することは非常に時間と手間のかかる煩雑な作業となってしまう。特にデザインがキーとなる商品の場合、現在のキーワードによる検索機能だけのフィルタリングでは不十分であり、一々ブラウザして内容を確認する必要があるからである。

一般にインターネット通販で売れる商品のひとつとし

て「実物を見なくとも、商品内容が容易にイメージできる商品」があげられる。例えば、有名ブランドの電化製品や自動車、タイトルや識別子がついている書籍、音楽CD、ビデオゲームソフト、及び航空券やコンサートチケット等である。これらの商品を選択・購入する場合、消費者が初めに買うものさえ決めていれば、カタログに商品写真や詳細説明が無くてもすむ。消費者は主に、価格とその店舗の信頼性とサービス機能だけで、商品を選択することが可能となる。こうした商品の選択に対しては、既存のキーワード検索及び各種ソート機能の組合せで、殆どの場合解決できる。

本稿では、上記のように検索が簡単な商品は対象としない。以下のような条件をもつ、簡単に検索できない商品のためのナビゲーション機能をどうするかについて論じていく。

- (A) 商品選択は、センスや嗜好性に強く依存し、キーワード検索だけでは不十分であること。
- (B) 商品内容が個々の商品に依存する度合いが高く、個々のカタログをブラウズしないと商品イメージが掴めないこと。

上記2つの条件を満たす商品として、我々は婦人服を事例として取り上げ、そのナビゲーション機能について論じることとする。

現状のナビゲーション機能を見てみよう。例えば、現在利用者の多いショッピングモールである、楽天市場[楽天01]においては、(1)キーワードと価格帯域指定による商品検索機能、(2)検索結果を(商品名、価格、店舗名)でリストアップする機能、(3)検索結果を(写真、商品名、価格)でリストアップする機能、等が用意されている。(3)の機能においては、サムネイルがリストアップされるため、イメージの把握は容易になってはいるが、やはり、候補のカタログを一々チェックし評価する必要がある、候補数が非常に膨大であるため、時間のかかる、疲れる作業となっている。

こうした問題を解決するためには、高機能なe-shopping用ナビゲーション・システムが必要となる。著者が目標とするナビゲーション・システムとは、婦人服のような強くデザインに関連する商品を購入する際、支援してくれるシステムで、多数のHTMLカタログを入力すると、消費者の嗜好とセンス、及び購入条件に合った商品を内部評価機能によって選択し、それらのHTMLカタログを再構成し、見易い形式で消費者に呈示してくれる機能をもつものである。このシステムをe-Stylistと呼ぶ。実世界のスタイリストとは、予算内でイメージに合う洋服を選んでくれたり、小物まで含めてトータルにファッションをチェックしてくれたりする。そのようなスタイリ

スト機能を実現するための技術的ポイントは次のようにまとめられる。

- (a) 知識の収集と整理：HTML ファイルからの意味構造抽出と表現形式の統一
- (b) 評価及び選択：用途、嗜好等の消費者のニーズに合った商品の評価及び選択
- (c) 商品の呈示：HTML ファイル群の再構成
- (d) 対話による絞込み：消費者の反応に依存した評価指標の追加・修正

本e-Stylistシステム構築のためには、まず消費者がどのように商品进行评估し、比較検討を行っているのか、その際の評価プロセスを分析する必要がある。そこで本研究では、まずその評価プロセス及び評価属性のモデル化を行ない、次に、その評価属性値をHTMLファイルからどのように抽出したらよいか、また、それを表現できるようなスキーマはどうあるべきか、を考える。また、e-Stylistシステム中で商品評価が行われた後、それらの商品カタログは消費者にとって、見易く選択し易い形で呈示されるべきである。そこでどのような呈示方式がよいか、及び消費者が使い易い対話型の絞込み方式を探る必要がある。

次節では、婦人服選択・比較プロセスを分析し、その評価モデルを提案する。第3節では、その評価モデルを基に、どのようにシステムを構築すべきかを考察する。第4節では、関連する既存技術について述べる。第5節はまとめである。

2 婦人服の選択・決定のプロセス

本節では、婦人服を選択・決定するプロセスを分析する。まず、消費者が婦人服を選択する際の評価プロセスモデルを図1に示す。図1にあるように、婦人服選択には、6種類の評価が存在する。サイズ評価とは、自分のサイズに合っているか否かを評価する。デザイン評価は自分の望むデザインであるか否かを評価する。デザイン評価は洋服単品に対する評価であるが、コーディネート評価では、複数の洋服の組合せによる全体的なデザインバランスの評価を行う。その上位のワードローブ評価とは、現在手持ちの洋服と組み合わせた場合、どの程度まわしができるかの評価を行う。その他、該当素材が洗濯等にどの程度耐性があるかを評価する耐用性評価がある。また、耐用性評価及びコーディネート評価の上位レベルとして、用途適合評価がある。これは「子供の入学式に着ていく服」、「リゾート地で着る服」、「教師が講義の着る服」、「裁判所に行く時着る服」等の用途を始めに指定して、その用途に合致しているかを評価するプロセスである。本評価は、主に、デザインがTPOに合っ

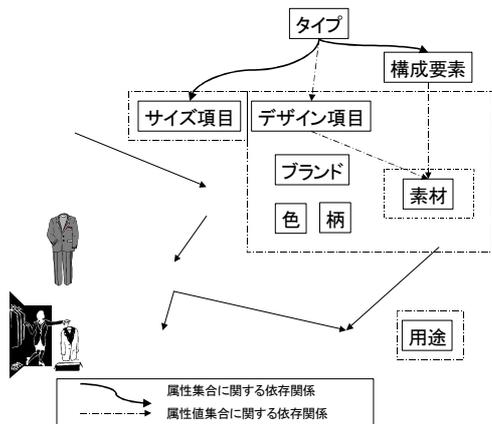


図1: 婦人服選択・決定の際の評価プロセス・モデル
 たものかを評価するが、耐用性も関係する。例えば、「油絵教室の作業服」、「子供の普段着」のような場合、何度も洗濯ができる素材である必要があるからである。

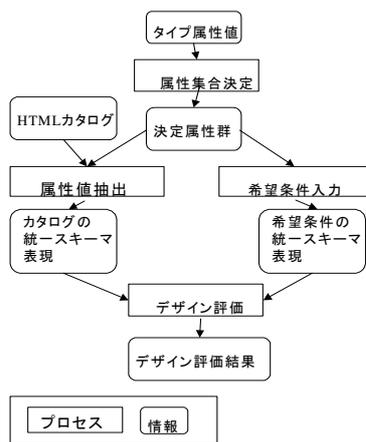


図2: デザイン評価プロセスモデル

典型的な評価手順は図1の実線直線矢印で示すように、**サイズ評価** **デザイン評価** **コーディネート評価** **用途適合評価** **ワードローブ評価**、である。耐用性評価はデザイン評価と並列して行う。

実際には、この他、以下の4種類の評価が行われるが、これは一般的評価プロセスであり、婦人服選択に固有ではないので簡単化のため以下では議論の対象からはずす。

- (1) 店舗名の信頼度評価、(2) 価格評価
- (3) 支払い手続き方法評価、(4) 配送料評価†

† 現状の電子市場で、その商品を購入した際の配送量まで提示してくれる便利なシステムはない。しかし、配送量は、(宛先、送り先、品数、商品 ID、配送会社) 及び (商品 ID、荷

次に、上記の評価を行うために必要な評価属性について説明する。評価属性は、以下の8種類に類別される。

- (1)タイプ、(2)サイズ項目、(3)構成要素、
- (4)デザイン項目、(5)ブランド、(6)色、(7)柄、
- (8)素材

図1においては、上記の8属性カテゴリは、実線の長方形で示している。点線の長方形は、ある評価に関する属性カテゴリ集合がどれであることを示している。

婦人服を選択・決定するプロセスの特徴は、婦人服のスキーマが服のインスタンス毎に変動する点である。つまり、インスタンス毎に属性が異なってくる。これを「属性集合に関する依存関係」と呼び、図1では曲線矢印で示した。このスキーマ変動の影響を受けるのは、主に、サイズ評価とデザイン評価である。そこでサイズ評価とデザイン評価では、このスキーマ変動に柔軟に対応するため、第1段階で属性集合決定、第2段階でその属性値を決定、という2段階のプロセスを取ることとした。図2に、サイズ評価及びデザイン評価のプロセスモデルを示す。

表1: ブラウスのサイズ表

サイズ	5号	7号	9号	11号	13号	15号	17号	19号
バスト	93	96	99	102	106	111	117	123
肩幅	39	40	41	42	43	44	46	47
着丈	62	63	64	65			66	
袖丈	53	54	55	56			57	

デザイン評価プロセスモデルでは、次のようにデザイン評価が行われる。

- (1) 属性集合決定プロセス: 服のタイプが決まると、属性集合に関する依存関係に基づき、サイズ項目と、構成要素の属性集合が決定される。
- (2) 属性値抽出プロセス: 対象となるHTMLカタログから属性値が抽出され、カタログの情報は、統一的なスキーマ表現に変換される。
- (3) 希望条件入力プロセス: 消費者の希望条件が入力されると、希望条件入力プロセスは、その情報を統一的なスキーマ表現に変換する。
- (4) デザイン評価プロセス: HTMLカタログ情報と、消費者の希望条件が同一スキーマで表現されると、デザイン評価プロセスが両者を比較し、評価を行い、評価結果を出力する。

サイズ、重量)の情報から、殆どの場合自動算出可能と考えられる。

以下に属性集合決定プロセス(図2参照)の例を示す。

まず消費者が、買いたい服のタイプをその値域 T= {スーツ、スカート、パンツ、ワンピース、ジャケット、ブラウス、コート} から選ぶ。例えば、ブラウスが選ばれた場合、属性集合決定プロセスはそのサイズ項目の属性として、{サイズ、バスト、肩幅、着丈、袖丈}の5個の属性を返り値とする。ブラウスの場合のサイズ表を表1に示した。表1に示すように、サイズという統括的表現は、詳細には{バスト、肩幅、着丈、袖丈}の値の組合せから決定されるからである。

また、構成要素(図1参照)に関する属性集合もタイプに依存して決まる。但し、属性集合決定プロセスはあくまでも標準的な結果を返すものであり、例外はあると考える。例えば、構成要素の例外としては、ワンピースでフードの付いたもの、尻尾の部分のあるパンツ等である。

以下に属性集合決定プロセスが返す、属性の数をまとめて示す。

- ・ タイプ：1個(属性はタイプのみ)。
- ・ サイズ：複数(属性はタイプに依存して複数決まる)。
- ・ 構成要素：複数(属性はタイプに依存して複数決まる)。
- ・ デザイン項目：1個(属性はデザイン項目のみ)。
- ・ ブランド：1個(属性はブランドのみ)。
- ・ 色：1個(属性は色のみ)。
- ・ 柄：1個(属性は柄のみ)。

上記のうち、デザイン項目については、本来は、タイプに依存して「属性が複数決まる」ようにするのが望ましいと考える。しかし、HTMLカタログ及び消費者の希望条件を見ると、雰囲気的な表現が多用されているため、「どの部分のデザインに対する表現であるのか明確でない」表現が多く、そのため、デザイン項目を論理的に分類分けして属性を決めることは困難となっている。また、デザイン項目を用いて行うデザイン評価においては、以下のような課題があり、簡単なパターンマッチだけでは評価できない。

- (1) デザイン用語のシソーラスの開発が不十分であること
- (2) デザイン表現は人毎に依存していて、統一性がないこと
- (3) 感性用語による表現が多いこと

ヒューリスティックなアルゴリズムを取り入れて処理を行うが、そのためには、単純な説明文の集合のほうが扱いやすい。よって、デザイン項目の論理的分類が困難であること、及び、デザイン評価部の具体的実装方法を考慮した結果、デザイン項目カテゴリの属性はデザイン項目のみとした。

婦人服を選択・決定するプロセスのもう一つの特徴は、属性に対する値域が、服のインスタンス毎に変動する点

である。つまり、「特定デザインには、それを可能とする限定された素材集合が存在する」、「特定の構成要素には、対応する素材集合が存在する」ということである。これを「属性値集合に関する依存関係」と呼び、図2では点線矢印で示した。例えば、ワンピースのデザイン項目属性値として、{バストダーツ、ノースリーブ、Aライン}等は出現可能であるが、これらはパンツのデザイン項目属性値になる可能性はない。

次に、婦人服の選択評価プロセスにおいて、どのような論理的DBが必要となるか以下に記す。

・サイズ評価：**サイズ表DB**

同じ9号サイズでも、メーカーやブランドによって若干異なるのでDBが必要。

・デザイン評価：**タイプ~デザイン項目関係DB、デザイン項目~素材関係DB、構成要素~素材関係DB**

上述した属性値集合に関する依存関係をDBとして蓄積することにより、適切でないキーワード抽出や希望条件入力をチェックすることが可能となる。

・耐用性評価：**耐用性評価ルールDB**

例：絹は耐用性が低い、デニムは耐用性が高い。

・コーディネート評価：**コーディネート評価ルールDB**

個人的なセンスに依存するが、一般的法則を格納。

・ワードローブ評価：**ワードローブDB**

既に購入済の服のDB。

・用途適合評価：**用途適合評価ルールDB**

判断基準は個人に依存する。例えば、お葬式でパールのイヤリングあるいは濃いグレーのスーツを許すか否かは人により異なる。よって、あくまでも一般

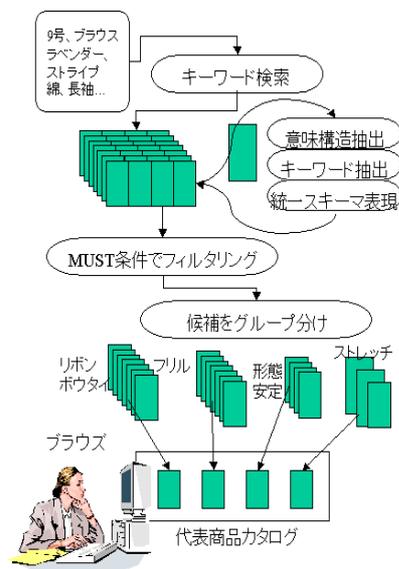


図3: e-Stylist デザイン評価部のシステム概要

的法則を格納。

さらに、欲しい服のイメージを、抽象度が高い感性用語によって検索するのであれば、以下が必要である。

- ・ブランド評価、デザイン評価、素材評価、色評価、柄評価：**感性用語辞書DB**
評価属性の特徴値と感性用語の対応関係を記録した感性用語辞書が必要。但し、人に依存するので、独自に作成する必要有り。

3 e-Stylist システム概要

前節でモデル化した婦人服選択プロセスに基づき、e-Stylistシステムではどのように評価プロセスを実装するか、そのシステム概要について述べる。本稿では、プロセス全体のうち、中心となるデザイン評価部分のシステム構成について述べる。デザイン評価部のシステムの目標は以下のようにまとめられる。システム設計にあたっては、これらを満たすように配慮する。

- (1) 消費者が必要とする情報のみを見せること
- (2) 比較がし易いように同じ表示形式で見せること
- (3) 評価点によって評価が高い順に見せること
- (4) デザイン特徴により分類して見せること

表2: 評価種別ごとのキーワード検索の条件例

評価種別	厳格さ	条件例 A	条件例 B
サイズ	Must	9号	9号
タイプ	Must	ワンピース	ブラウス
色	Should	黒地に白の	ラベンダー
柄	Should	水玉	ストライプ
素材	Should	表地:ポリエステル 裏地:ポリエステル	表地:綿
ブランド	Should	シャーリーズ	
デザイン	Should	後ろボタン 胸元フリル	長袖 丸襟 フリル

図3にシステム概要を示す。以下では図3に沿って、システムフローを説明する。

- (1) キーワード検索: 消費者は荒いキーワード検索を行い、候補のカタログ群を得る(表2参照)。
- (2) 意味構造抽出: カタログごとに意味構造を抽出する。
- (3) キーワード抽出: カタログごとに、商品説明キーワ

ードを抽出する。

- (4) 統一スキーマ表現: カタログごとに、統一スキーマの表で表現する。対応する情報を他の情報から生成することも含む。
- (5) MUST条件でフィルタリング: 必須条件を満たさないカタログを捨てる。
- (6) 候補のグループ分け: 抽出キーワードのうち、重要度の高いものを選び、そのキーワードにより候補をグループ分けする。
- (7) 評価点の計算: 抽出キーワードにより、各カタログに評価点を付ける。
- (8) カタログ再構成: 写真、キーワード、各種説明文だけを抜粋し、統一表示形式の簡単なカタログ情報を構築する。
- (9) 代表商品カタログ作成: グループ内で評価点の高いカタログを代表商品カタログとし、それらを集めた代表商品カタログを作成する。
- (10) 消費者ブラウス: 消費者は関心のあるグループを選択し、選択したグループに対してさらにブラウスや絞り込み検索を行う

上記フローをさらに詳しく説明する。(1)のキーワード検索とは、現在の電子市場で行われているレベルの検索のことである。表2に条件の例を2つ示した。検索条件を与えるとき、以下の属性を指定できる。

- ・条件の厳格さ: 緩められる条件か否かを示す。

表2に示されたMustは必須条件、Shouldは緩められる条件であることを意味する。

(2)の意味構造抽出について説明する(図4参照)。現在のHTMLカタログファイルを見ると、店舗によりHTMLの書き方は大きく異なるため、商品比較検討が困難になっている。そこで、与えられたカタログから消費者がデザイン評価を行うのに必要と考えられる情報のみを抽出して、それを統一されたスキーマ形式に整形しDBに一次的に格納する。データモデルはリレーショナルとする。

タグによる文書分割(図4参照)とは、HTMLのタグを利用して与えられたHTMLファイルを、以下のような部分に分けることである。

(カタログ

(商品説明部(

デザイン説明部、サイズ説明部、

素材説明部、イメージ写真)、

店舗説明部、

操作説明部))

図5に候補カタログの統一スキーマとインスタンス例を示す。スキーマは服のタイプに依存して変化するので、カタログインスタンス毎に属性が柔軟に変更できることが必要である。しかし、それでは検索効率に影響を与えるので、参照頻度が高い属性は表「商品」に示すように固定した。説明文については、同じ文書部分にまとめて出現した複数の説明文は、まとめて一つの値として格納できるようにした。また出現場所が離れている説明文は、別の値として格納できるようにし

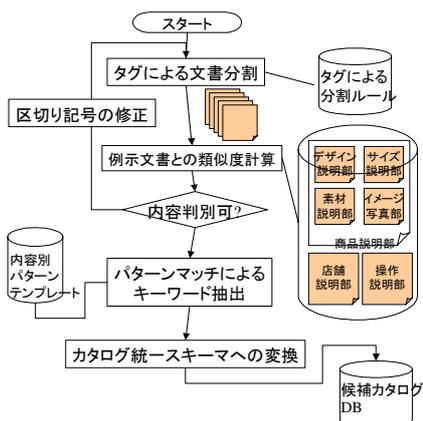


図 4:HTML カatalogファイルからの意味構造抽出プロセス

た。格納できる説明文の数は固定しない。

次に、その表を用いて、サイズ等MUST条件に対するフィルタリングを行なう。サイズ・フィルタリングは、(1)の荒いキーワード検索で行える場合もあるが、詳細なサイズ表情報と照合しないと正しい評価が行えない場合もあるので、サイズ情報を統一表形式に直して後、再度行う。

候補カタログは、キーワードによって分類される。キーワードの重みは予め登録しておく。そして、重要なキーワードを分類用キーワードとして使い、カタログのグループ分けを行う。評価点の計算は、カタログから抽出されたキーワード集合に対して、加重平均を取る。キーワードの重みは予め登録されている。

グループ内から最高評価点のカタログを選出し、それらを集めて代表商品カタログのリストを作る。消費者がブラウズする際、代表商品カタログから、対応するカタロググループのブラウズが可能とする。カタログの情報としては、統一スキーマで格納された情報のうち、重要度が高い属性のみが表示されるようにする。消費者ブラウズにおける、絞り込み検索や希望条件の新たな入力方式については、試作実験を行った段階で考えていきたい。

今後、詳細なシステム設計に入っていきが、本節の初

めに述べた目標を満たすようにシステム設計を行っていききたい。それにより、消費者が効率的に、かつ、消費者が満足できる商品選択が行えるものと考えている。

商品	
商品ID	51110921
タイプ	ワンピース
タイトル	バイアスミドルワンピース
サイズ	9
定価	5900

説明	
商品ID	51110921
説明文	バイアスワンピースの定番カラーに春・夏カラーが加わりました。
説明文	結婚式やパーティにいかがでしょう。季節を問いません。
説明文	同じ柄のバックも用意しました。

サイズ	
商品ID	51110921
表地着丈	57
表地身幅	84
裏地着丈	76
裏地身幅	80

デザイン	
商品ID	51110921
説明文	白のトーションレース付
説明文	やさしいAライン

素材	
商品ID	51110921
素材1	ポリエステル
素材1割合	100
素材2	
素材2割合	
説明文	2枚重ねのバイアス裁断

写真	
商品ID	51110921
写真1	f2890189.jpg
キャプション1	
写真2	f78321.jpg
キャプション2	

色柄	
商品ID	51110921
色	ブラック
柄	
説明文	

図 5: カatalogの統一スキーマとインスタンス

4 関連研究

本節では、関連する3つの研究分野、(a)HTMLからの意味構造抽出技術、(b)自然言語処理、(c)感性による検索、の既存研究について述べる。

4.1 HTMLからの意味構造抽出

Webページ検索に対する研究が盛んであり、HTMLから意味構造を抽出する研究も多々行われている。その中には、HTMLを半構造データとしてとらえたアプローチ([Cha98][Miz99])や、HTMLにXMLとして意味構造を含めるアプローチ([Liu00])がある。意味階層を用いたHTML文書の変更の自動検出技術に関する研究も盛んである[Cha96,97,99, Lim99-01]。Webカタログから不完全な情報ツリーを取り出す研究もある[Abi01]。

Webページの変更検出とは、時々刻々と変更されるページ変化の様子を見つけることである。例えば、株式相場の表で、どの銘柄がどれだけ値段の動きがあったのかを知りたい場合がある。HTMLのタグに基づくシンタックス階層からの差分では、変更のあった行情報(例えば銘柄と値の羅列)を得られても、変更された値がどの列(例えば高値 or 安値)に対応するのか、分かりづらい。Limらは、前処理せずにHTML文書から直接対応する意味階層を生成することにより、具体的な変化情報を得ている

[Lim01]。その他、HTMLの文書を対象とはしていないが、半構造データからの意味構造抽出に利用できる技術として[Wan94,97]がある。

我々の e-Stylist システムの意味構造抽出プロセスにおいても、複数ある候補 HTML 文書の意味的な比較を自動的に行うため、HTML という半構造文書からの意味構造抽出機能が必要となる。

4.2 自然言語処理

HTML からの意味構造を抽出する場合、タグ、例えば表を構成するタグが、区切り記号として用いられる[Lim01]。それに対し、その区切り記号内のテキスト及び数値を最終的に抽出するためには、形態素分析等の自然言語処理が必要となる。

本 e-Stylist の意味構造抽出プロセスにおいても、例示文書との類似度計算及びパターンマッチによるキーワード抽出において形態素解析処理による前処理が必要と考えられる。例えば、「全体にスパンコール・フィルムをちりばめたキラキラ美しいチューブドレス」と言った説明文から、(スパンコール・フィルム, キラキラ, 美しい, チューブドレス)というデザインに関する重要表現用語を抽出するだけでも、かなりの内容把握が可能である。

次に類似検索機能について考えてみる。自然言語処理のひとつである類似検索機能は、与えられた文書と類似した主題概念の文書を探してくれる。例示文書を分類項目の例示となるように選べば、類似検索を文書検索の主題分類として使うことができる[藤田 99]。本 e-Stylist システムでの利用としては、ひとつの HTML 文書を HTML タグにより細かい段落に分割し、その後、その各部分かどのような内容であるか決定するために利用できる。具体的には、予め「サイズ説明部」や「デザイン説明部」の例示文書を作成しておくことにより、その例示文書との類似度を計算して、類似度が高ければ、その部分が何を説明しているのか、判別できる。

また、自然言語処理のひとつであるパターンマッチング手法を使うことにより、テキストから特定の情報を抽出することができる[関根 99]。この情報抽出においては予めパターンを情報抽出の課題ごとに作成しなければならない点が最大の問題となるが、目指す対象領域を限定することにより実装可能と考えている。

4.3 感性情報検索

婦人服などの検索の場合、「さわやかな雰囲気」「フォーマル」等といった感性用語による検索ができると消費者によって有効性が増す。本 e-Stylist システムにおいても将来拡張として、感性用語による検索が必要と考え

る。例えば、「婚約者のご両親に挨拶に行くので、さわやかな清楚なイメージのワンピースを選べ」というような検索を行ないたい。

加藤は、Artificial Kansei による感性シミュレーションでの情報の取捨選択における課題として以下の3つをあげている[加藤 01]。

- (ア) 画像特徴量の設計
- (イ) 主観的な基準の教示
- (ウ) 統計的な分析法・学習法

e-Stylist システムをターゲットとして、これらの課題を考察してみる。まず、婦人服の Web カタログにおける写真は、洋服単体、あるいはモデルがその服を着ている写真である。大枠が決まっているため、画像処理が比較的容易に行えると考えられる。服のデザインを画像処理技術により抽出することはまだ困難な段階であるが、色や柄情報の抽出は実現可能である。特に、服の柄は繰り返しパターンが多いので複雑度が少ない場合が多い。生地柄の検出手法としては、[多田 01]の研究がある。多田らが提案した手法は、ウェーブレットを用いた手法に比較して計算量が少なく、実装が手軽に行えるという利点をもつ。

次に(イ)の主観的な基準について考察する。婦人服のデザイン感性用語と各種のデザイン特徴とのマッピング関係を作るためには、何かしらのデザイン基準が必要となる。感性用語検索の研究では、一般的なデザインに対する「デザインの分布図」がよく使われる。このデザインの分布図とは、warm—cool, soft—hard を対極とする2軸を取り、その二次元平面上に、デザインを分布させるものである。例えば、formal はこのあたり、pretty はこのあたりに分布というように対応をつける。宮川らは、建具製品のデザインの分析にこの図を使っている[宮川 01]。婦人服のデザインの基準としても、このデザインの分布図が適切と考える。つまり、デザインの分布図を仲介として、

- ・感性用語とデザイン分布図の関係
- ・服のデザイン特徴とデザイン分布図の関係

の2段階マッピングを作ることで、感性用語と服のデザイン特徴のマッピングを表せる。感性用語検索のためには、利用者のプロファイルを作る必要がある。利用者に、各種のデザイン特徴(例:ブラウスのボウ、スカートのスリット)を「デザインの分布図」上に配置してもらうことにより、その利用者の感性に依存した、服飾デザイン上の特徴と感性用語(例:プリティ)のマッピング情報を作成できる。

5 まとめ

本稿では、婦人服をインターネットで購入する際の、商品選択及び決定を支援するナビゲーションシステム e-Styleist について説明した。本稿では、まず、婦人服選択・決定プロセス全体のモデル化を行なった。モデル化においては、デザイン評価に関する評価属性及び属性値の間の依存性を、属性集合に関する依存関係、及び、属性値集合に関する依存関係の2つにより、明確化した。

そして、そのプロセスモデルのうち、最も核となる部分であるデザイン評価部分を e-Styleist でどのように実装するか、そのシステム概要について述べた。また、これらのシステム構築にあたり、どのような既存技術が利用可能であるかを、HTML からの意味構造抽出、自然言語処理、感性情報検索、の3分野に対して行った。

本 e-Styleist の研究では、消費者が e-shopping を楽しく、効率的に行えることを目標とする。現段階は、まだデザイン評価プロセスのシステム概要を決めたのみで、これから試作システムの開発に取り掛かる。順次システム設計、試作を行い、最終的にはトータルで消費者を支援するシステムを作っていくたい。

参考文献

- [Abi01] Serge Abiteboul, Luc Segoufin, Victor Vianu: Representing and Querying XML with Incomplete Information, *Proc. of PODS 2001*, 2001, pp. 150-161.
- [Cha96] Sudarshan S. Chawathe, Anand Rajaraman, Hector Garcia-Molina, Jennifer Widom: Change Detection in Hierarchically Structured Information, *Proc. of SIGMOD Conf. 1996*, 1996, pp. 493-504.
- [Cha97] Sudarshan S. Chawathe, Hector Garcia-Molina: Meaningful Change Detection in Structured Data, *Proc. of SIGMOD Conf. 1997*, 1997, pp. 26-37.
- [Cha98] Sudarshan S. Chawathe, Serge Abiteboul, Jennifer Widom: Representing and Querying Changes in Semistructured Data, *Proc. of ICDE 1998*, 1998, pp. 4-13.
- [Cha99] Sudarshan S. Chawathe: Comparing Hierarchical Data in External Memory, *Proc. of the 25th Very Large Data Bases Conf.s*, 1999, pp. 90-101.
- [Ham97] Joachim Hammer, Jason McHugh, Hector Garcia-Molina: Semistructured Data: The Tsimmis Experience, *Proc. of the First East-European Symposium on Advances in Databases and Information Systems (ADBIS'97)*, St.-Petersburg, September 2-5, 1997, pp. 1-8.
- [Lim99] Seung Jin Lim, Yiu-Kai Ng: An Automated Approach for Retrieving Hierarchical Data from HTML Tables, *Proc. of the 1999 ACM CIKM International Conf. on Information and Knowledge Management*, Kansas City, Missouri, USA, November 2-6, 1999, pp.466-474.
- [Liu00] Ling Liu, Calton Pu, Wei Han: XWRAP: An XML-Enabled Wrapper Construction System for Web Information Sources, *Proc. of ICDE 2000*, 2000, pp. 611-621.
- [Lim01] Seung Jin Lim, Yiu-Kai Ng: An Automated Change Detection Algorithm for HTML Documents Based on Semantic Hierarchies, *Proc. of the 17th International Conf. on Data Engineering*, April 2-6, 2001, Heidelberg, Germany, IEEE Computer Society 2001, pp. 303-312.
- [Miz99] Yoshiaki Mizuuchi, Keishi Tajima: Finding Context Paths for Web Pages, *HYPERTEXT '99, Proc. of the 10th ACM Conf. on Hypertext and Hypermedia: Returning to Our Diverse Roots*, February 21-25, 1999, Darmstadt, Germany, ACM, 1999, pp.13-22.
- [Wan94] Jason Tsong-Li Wang, Kaizhong Zhang, Karpjoo Jeong, Dennis Shasha: A System for Approximate Tree Matching, *IEEE Transactions on Knowledge and Data Engineering*, Vol. 6, No. 4, August, 1994, pp. 559-571.
- [Wan94] Jason Tsong-Li Wang, Gung-Wei Chirn, Thomas G. Marr, Bruce A. Shapiro, Dennis Shasha, Kaizhong Zhang: Combinatorial Pattern Discovery for Scientific Data: Some Preliminary Results. *Proc. of SIGMOD Conf. 1994*, 1994, pp. 115-125.
- [Wan97] Jason Tsong-Li Wang, Dennis Shasha, George Jyh-Shian Chang, Liam Relihan, Kaizhong Zhang, Girish Patel: Structural Matching and Discovery in Document Databases, *Proc. of SIGMOD Conf. 1997*, 1997, pp. 560-563.
- [加藤01] 加藤俊一:「感性の工学的モデル化とヒューマンメディアデータベース」、情報処理学会研究報告、2001-DBS-124-5、2001年5月21日、東京。
- [関根99] 関根 聡:「テキストからの情報抽出—文書から特定の情報を抜き出す—」、情報処理、Vol. 40, No. 4, April, 1999, pp. 352-357.
- [多田01] 多田昌裕、池田大地、杉林弘基、他:「テキスト画像検索システム—大分類と例示検索—」、情報処理学会研究報告、2001-DBS-124-6、2001年5月21日、東京。
- [藤田99] 藤田澄男:「自然言語処理を利用した情報の検索・分類へのアプローチ」、情報処理、Vol. 40, No. 4, April, 1999, pp. 352-357.
- [宮川01] 感性情報処理法によるデジタル伝統工芸プレゼンテーション、情報処理学会研究報告、2001-DBS-124-2、2001年5月21日、東京。
- [楽天01] 楽天市場、<http://www.rakuten.co.jp/>