

日本行動計量学会第36回大会（於：成蹊大学）
特別セッション 好みの計量



ECサイトにおける画像の類似性測定 ～『SUDACHI』の開発と商品推薦エンジンへの応用～

2008年9月3日（水）

アルベルト
株式会社ALBERT

〒151-0053 東京都渋谷区代々木2-22-17
TEL:03-5333-3703 FAX:03-5333-3723
www.albert2005.co.jp/

1：注目されるレコメンデーション

現在、国内外で「検索に続く次世代のコア技術」として注目を浴びています。

【日経産業新聞：2006年2月8日】

重み増す推薦エンジン

ネット通販で必須の機能に

利用者一人ひとりの好みに応じて、商品やサービスを推薦する「推薦エンジン」は、インターネット上のサービスで重要度が高まっている。家電の商品比較サイト「教えて！家電」を運営するアルベルト（東京・渋谷）の上村崇社長は「買い物に迷ったときに役立つ店員との会話が究極の推薦機能」と例え、推薦エンジン機能がネット通販では必須の機能になると強調する。

従来のネット通販では、機械的に処理できるのが購買履歴に基づく推薦。ネット通販大手のアマゾンジャパン（東京・渋谷）のサイトでは「このCDを買った人はこんなCDも買っています」というような購買履歴から商品提案をする。教えて！家電でも一部採用しているという。

関連性の高い商品提案し、まとめ買いを促す。ただ、この手法は、本やCDなど一人が多数を購買する商品には有効だが、家電など、ときどきしか購入しない商品ではうまく機能しない。

▼推薦エンジン ネット通販の利用者の属性や好み、行動履歴などに基づいて、お薦めの商品やサービスを提案する機能。ネット通販では、大量の商品を現実で見るが、キーワード検索など商品を探す手段が限られ、欲しい商品を探しづらく、その解決策の一つとして注目を集める。ネット通販大手の米アマゾン・ドットコムは「推薦エンジン」を「リコメンデーションエンジン」とも呼ぶ。

【日本経済新聞：2006年10月4日】

「検索」の次「推薦エンジン」



好みの音楽、分析し配信

「ここがゲノム解析の現場だ。一九九九年設立のバンドラ・メディアの創業者、ティム・ウェスターグレン氏に案内されたのは、米カリフォルニア州オークランド市にあるビルの一室。ヘッドホンをつけた男女が並んでパソコンに向かい、キーボードをたたく。

解析するのは生物の遺伝情報ではなく音楽。CDを聴き、メロディやリズムの強弱など四百項目を点検し、データベースにする。曲ごとに数値は異なり、それを見ればどんな曲が分る。

本業はインターネットによる音楽ラジオ運営だが、ゲノム計画と絡めて、曲を自動的に選び、次々とパソコンに流す。

選曲に使うのが「音楽マシン」は一定の音楽をもとに、ユーザーの嗜好（好み）を判断し、関連情報（バンドラのネットラジオ）を流す。オーストラリアの音楽レーベルはバンドラ設立の理由を話す。

推薦エンジンで無名アーティストや消費者を引き合わせ、一躍の著名アーティストに「賞」が贈られる音楽業界の変わり目。

解析部には「天才」が揃った。CDは、英EMIなど大手レコード会社で音楽を学んだ人たちが、料で送ってきたものを、ネットで流せばCDの宣伝になるからだ。バンドラでは一人ひとりの好み合った曲が自動的に流れ、広告効果も高い。

バンドラのサイトは、気になった曲があればクリック操作で、アマゾン・ドットコムでCDを注文したり、アップロードして飛んだりする仕組みを導く。同社の調べでは、利用者の四割が、以前より

【日本経済新聞：2007年6月1日】

戦状



検索技術、止まらぬ進歩

今にもある。近頃、インターネットの検索技術は、まるで驚異的な進歩を遂げている。同じ傾向の車は、同じような仕様の車。同じような車名を、同じような検索エンジンで検索すると、同じような結果が出てくる。これは、検索エンジンの進歩が、インターネットの普及とともに進んでいることを示している。

「検索」の次「推薦エンジン」

推薦エンジンで無名アーティストや消費者を引き合わせ、一躍の著名アーティストに「賞」が贈られる音楽業界の変わり目。

解析部には「天才」が揃った。CDは、英EMIなど大手レコード会社で音楽を学んだ人たちが、料で送ってきたものを、ネットで流せばCDの宣伝になるからだ。バンドラでは一人ひとりの好み合った曲が自動的に流れ、広告効果も高い。

バンドラのサイトは、気になった曲があればクリック操作で、アマゾン・ドットコムでCDを注文したり、アップロードして飛んだりする仕組みを導く。同社の調べでは、利用者の四割が、以前より

2：レコメンド方式の分類_ACKマトリクス

ALBERT独自のレコメンド方式の分類

※この分類は、「日経コンピュータ」「日経ネットマーケティング」「月間ネット販売」に掲載されました。

			レコメンデーションに用いるデータ			
レコメンド・ロジック			対象者を特定しないデータ	対象者を特定するデータ		
				アクションデータ	履歴データ	申告データ
			対象者全てに同様なレコメンデーション		パーソナライズされたレコメンデーション	
	モノ属性ベース (Attribute)	商品属性データベースが必要	A-1 スペックランキング型	A-2 商品関連性 評価型	A-3 コンテンツベース・フィルタリング	A-4 ニーズインプット型
	人ベース (Consumer)	対象者以外の購買データが必要、新発売商品に適用できない	C-1 人気ランキング型	C-2 マーケット バスケット型	C-3 協調フィルタリング	C-4 アンケートベース型
	ナレッジベース (Knowledge)	ナレッジの信頼性・妥当性の担保が必要	K-1 ナレッジベース 独断型	K-2 ナレッジベース 反応型	K-3 ナレッジベース 観察型	K-4 ナレッジベース 診断型

店頭接客における商品レコメンドの基本は、モノ属性ベースと人ベースでは「商品関連性評価型」「ニーズインプット型」。その他ナレッジベースもある。

3：レコメンド方式の魅力度_調査概要

- 調査目的 レコメンデーションへの注目が集まる中、ECサイトでのレコメンデーション機能の認知率、活用実態および、様々な種類のレコメンデーションエンジンの種類について、消費者にとっての魅力度を、インタースコープで開発した「CONEL」を用いて測定する。
- 調査手法 Web定量調査
- 対象エリア 東京 埼玉 神奈川 千葉 大阪 京都 兵庫
- 調査対象 一般生活者 20～49歳男女 300名
インターネットショッピングを2～3ヶ月に1回以上利用している人
回収数

		全体	東京都	埼玉県	神奈川県	千葉県	大阪府	京都府	兵庫県
全体		300	108	44	44	28	39	17	20
性別:	女性	150	49	21	28	16	19	10	7
	男性	150	59	23	16	12	20	7	13
年代:	20代	72	22	14	10	6	9	5	6
	30代	137	51	18	17	16	19	7	9
	40代	91	35	12	17	6	11	5	5

- サンプル抽出方法 ヤフーバリューインサイトパネル「ボイスポート」から抽出 =パネル調査
- 実査期間 2008年8月21日（木）～8月23日（土）

詳細結果は、13:15～ A会場「CSマーケティングセッション」で発表します。

3：レコメンド方式の魅力度_調査設問

* 【Q4】 インターネットショッピングサイトを利用するにあたり、以下のおすすめ機能について、あなたの考えを教えてください。

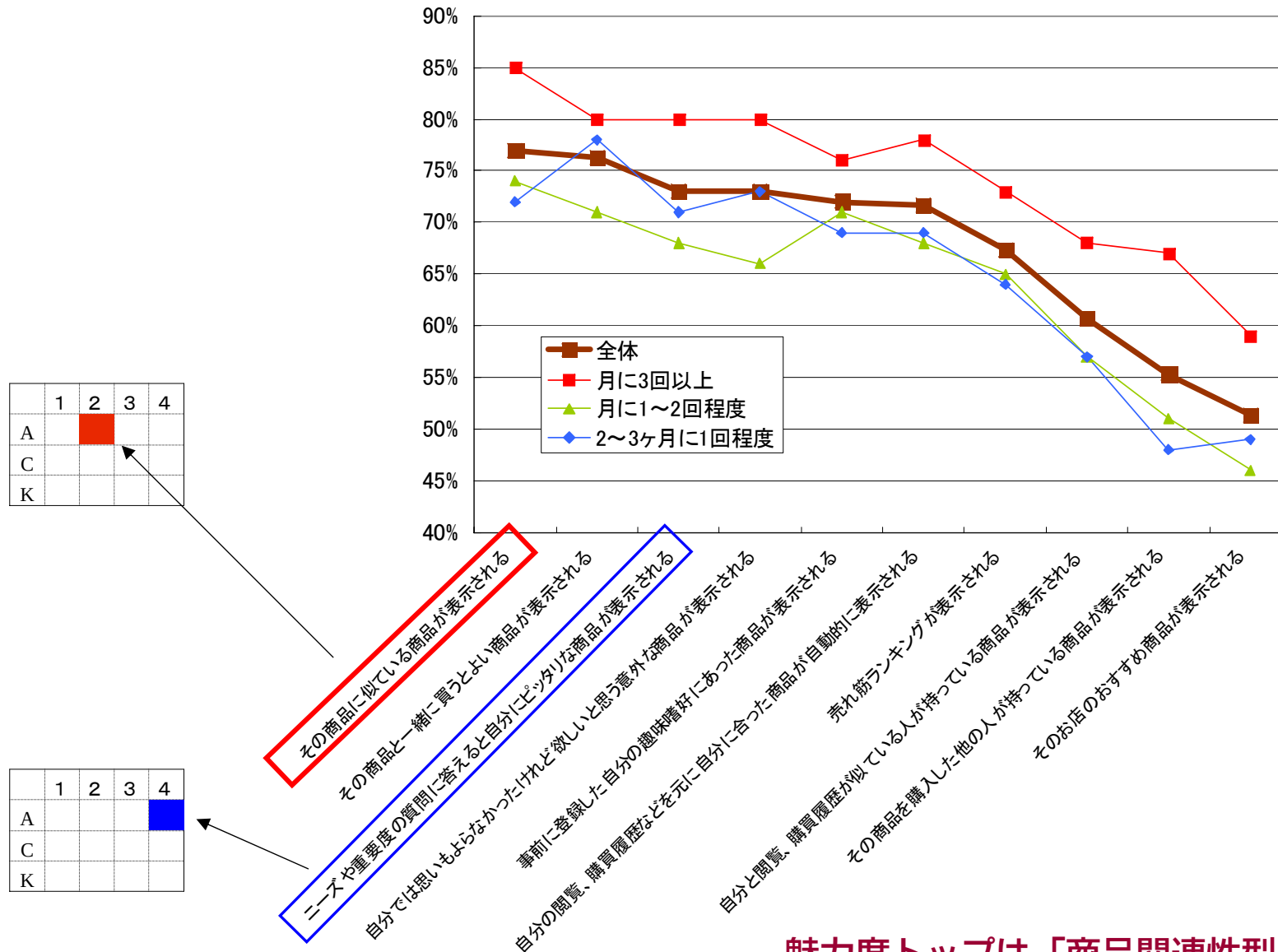
		1. あると非常に魅力的	2. あるとかなり魅力的	3. あるとやや魅力的	4. あって当たり前	5. あってもなんとも思わない(無関心)
	→	○	○	○	○	○
* (1) 売れ筋ランキングが表示される	→	○	○	○	○	○
* (2) その店のおすすめ商品が表示される	→	○	○	○	○	○
* (3) 自分の閲覧、購買履歴などを元に自分に合った商品が自動的に表示される	→	○	○	○	○	○
* (4) その商品に似ている商品が表示される	→	○	○	○	○	○
* (5) その商品と一緒に買うとよい商品が表示される	→	○	○	○	○	○
* (6) ニーズや重要度の質問に答えると自分にピッタリな商品が表示される	→	○	○	○	○	○
* (7) その商品を購入した他の人が持っている商品が表示される	→	○	○	○	○	○
* (8) 事前に登録した自分の趣味嗜好にあった商品が表示される	→	○	○	○	○	○
* (9) 自分と閲覧、購買履歴が似ている人が持っている商品が表示される	→	○	○	○	○	○
* (10) 自分では思いもよらなかったけれど欲しいと思う意外な商品が表示される	→	○	○	○	○	○

	1	2	3	4
A				
C				
K				

	1	2	3	4
A				
C				
K				

種々の方式のトップ3についての集計をした。

3：レコメンド方式の魅力度_調査結果



魅力度トップは「商品関連性型」レコメンド

4：モノ属性ベースと人ベースのレコメンド

店頭接客における商品レコメンドの基本、
「ニーズインプット型」と「商品関連性評価型」

	1	2	3	4
A				
C				
K				

	1	2	3	4
A				
C				
K				

モノ属性ベースのレコメンドには深い商品理解、
商品データベースが必要となる

代替え案としての人ベースの手法である、
マーケットバスケット、協調フィルタリングが登場

	1	2	3	4
A				
C				
K				

	1	2	3	4
A				
C				
K				

モノ属性の「自動的取得技術」として
テキストマイニング、画像認識の登場

5：画像認識技術を用いた『SUDACHI』使用例

The screenshot shows the SUDACHI website interface. At the top, there's a navigation bar with categories like トップ, パソコン, デジカメ, テレビ, 家電, アロマ, ネイル, コスメ, and ファッション. Below this, the 'ファッション' (Fashion) section is active, showing 'メンズ' (Men's) and 'レディース' (Women's) sub-sections. A central grid displays various shoes and clothing items. Three callouts highlight specific recommendations:

- 色が似ている 別カテゴリの商品を提示** (Color similarity, recommend items from a different category): Points to a white shirt in the 'メンズ' section.
- 色が似ている 同一カテゴリの商品を提示** (Color similarity, recommend items from the same category): Points to a white shoe in the 'メンズ' section.
- 色が違う 同一カテゴリの商品を提示** (Color difference, recommend items from the same category): Points to a blue shoe in the 'メンズ' section.

Other elements include a 'Ranking' section for 'メンズコスメ(スキンケア)' (Men's Cosmetics (Skincare)) and a '家族団らん 見つける!' (Family Time Find!) section.

5： 画像認識技術を用いた『SUDACHI』使用例

The screenshot shows the 'レディース水着' (Ladies Swimwear) section of the SUDACHI website. The main banner features the text 'Sexy & Cute' and '真夏の Beach Style'. Below the banner, there are buttons for '人気商品' (Popular Items) and 'ビキニ' (Bikini). A large grid of bikini images is displayed, with a central image of a woman in a pink polka-dot bikini. Callouts explain the recommendation logic:

- 選択した商品** (Selected Item): Points to the central image of the woman in the pink polka-dot bikini.
- 色が似ている 同一タイプの商品を提示** (Recommend items of the same type with similar colors): Points to a vertical column of pink bikinis on the right side of the grid.
- 色が似ている 別タイプの商品を提示** (Recommend items of a different type with similar colors): Points to a vertical column of pink bikinis on the left side of the grid.
- 色が違う 同一タイプの商品を提示** (Recommend items of the same type with different colors): Points to a vertical column of bikinis in various colors (blue, white, black) on the right side of the grid.

At the bottom, there is a 'Checked!' section with a list of recommended items and a 'Bikini Coll' label.

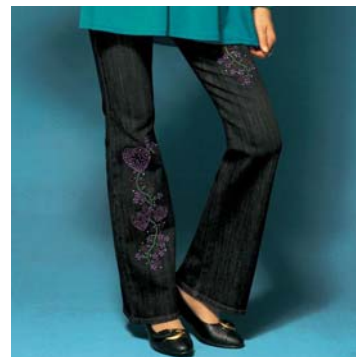
SUDACHI

6： 類似画像検索システム『SUDACHI』

●ECサイトにおける商品画像の特徴

ECサイトにおける商品画像の形態には以下の様にさまざまなものが存在し、単に商品のみが写っているものもあるが、背景部分に景色が写っていたり人物が着用していたりする画像もある。

今回開発した類似画像検索システム『SUDACHI』は、背景部分は無視し、対象となる商品の部分のみから特徴量抽出を行ない、多数の商品の中から、色、テクスチャ（模様）、形状、カテゴリの重み付けをした上で類似画像を検索することができるシステムである。（特許出願済）



6： 類似画像検索システム『SUDACHI』

●背景除去

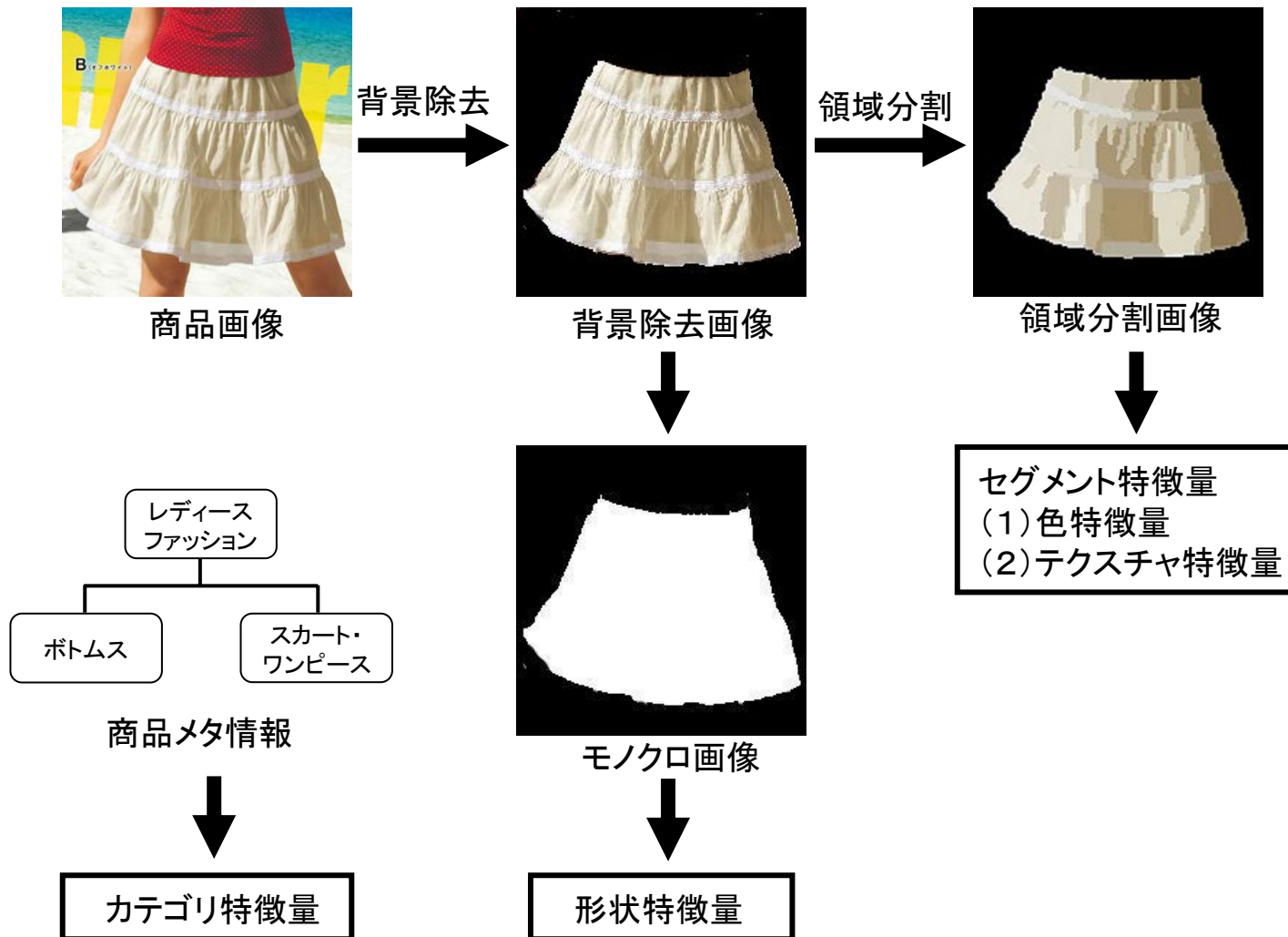
商品画像から、色やテクスチャ、あるいは形状等の特徴量を抽出する際には、背景部分は無視し、対象となる商品部分のみから特徴量抽出を行なうことが望ましい。背景部分も含め、画像全体から特徴量抽出を行なったのでは、高精度な類似商品推薦システムを構築することは困難である。『SUDACHI』では、画像からの特徴量抽出に先だち、背景部分の除去を行ない、対象となる商品部分のみの画像を抽出する。



●特徴量の抽出

背景除去した後の商品画像の類似性の解析には、商品画像から得られる色やテクスチャの情報を表す「セグメント特徴量」や「形状特徴量」を用いる。セグメント特徴量とは、画像の領域分割(segmentation)によって得られた領域ごとに抽出される特徴量で、色やテクスチャといった特徴の似ている領域をグループ化することで、元画像を類似した色やテクスチャを持った複数の領域に分割する処理を行なう。セグメント特徴量に基づく商品画像間の類似度計算では、Earth Mover's Distance (EMD)と呼ばれる距離尺度を用いる。形状の解析には独自の形状特徴量を用いている。

6： 類似画像検索システム『SUDACHI』



6：まとめ

- ・レコメンド方式の魅力度は、類似商品を提示する「商品関連性評価型」が最も高く、ニーズや重要度を質問してぴったりの商品をおすすめする「ニーズインプット型」も3位であり、店頭接客のおすすめの基本に合致している。
- ・類似商品の提示には、深い商品理解や商品データベースが必要だが、書籍やCDなど、膨大な商品数があるカテゴリでは、事実上すべてに深い知識を持つ事は不可能であり、人ベースでの関連商品提示という手軽な方式が台頭して来たと考えられる。
- ・テキストマイニングの技術が進化し、商品説明文等を自動解析することで、商品属性を定義することが可能になって来ている。またCGMデータを活用することで、関連商品をおすすめすることもできるようになっている。
- ・画像の類似性測定技術は、主にデザインで選択されるカテゴリ商品に特に有効であり、モノ属性を自動的に解析し、商品の関連性を定義することで類似商品をおすすめする手法として注目される。