

지능형 패션 코디네이션 시스템에서 유사의류 추천방법

김 정 인[†]

요 약

인터넷을 이용하는 패션/코디 쇼핑몰은 그 시장이 매년 큰 폭으로 증가하고 있다. 하지만 대부분의 쇼핑몰들은 아직도 카테고리에 의한 상품 분류를 제외하면 특정한 의류를 찾을 수 있는 기능이 포함되어 있지 않아 사용자들이 불편을 겪고 있다. 본 연구에서는 여성복의 패션 코디네이션 시스템 구축 시 내용 기반의 방식 중 휴리스틱 기반 방식으로 시스템을 구축하기 위해 여성복의 특성을 가장 잘 나타낼 수 있는 속성들을 상속관계로 정의하고, 상품 관리자가 속성 값을 입력할 수 있도록 하였으며 가장 유사한 의류를 추천하기 위한 방법을 소개하고 분석하였다.

A Recommendation Method of Similar Clothes on Intelligent Fashion Coordination System

Jung-In, Kim[†]

ABSTRACT

The market for Internet fashion/coordination shopping malls has been enormously increased year by year. However, online shoppers feel inconvenient because most of Internet shopping malls still rely on item classifications by category and do not provide the functionality in terms of which shoppers can find clothes they want. In an effort to build a fashion/coordination system for women's dress adopting the Heuristic-based method, one of the Context-based methods, we present a method for defining characteristics of a woman's dress as attributes and their inheritance relations, which can be input by a product manager. We also compare and analyze various methods for recommending the most similar clothes.

Key words: Internet Shopping Mall(인터넷쇼핑몰), Attribute(속성), Clothes(의류), Coordination(코디네이션), Recommendation(추천)

1. 서 론

인터넷이 대중화되고 인터넷 쇼핑몰의 편리함 때문에 인터넷 쇼핑몰을 이용하여 자신에게 필요한 물건을 구매하는 사람이 빠르게 증가하고 있다. 하지만, 기존의 인터넷 쇼핑몰은 보유하고 있는 상품을 단순히 나열하여 정보를 제공하는 형태여서 고객이 원하는 상품을 적절하게 제공하지 못했다. 이런 단점을 해결하고자 최근에는 고객이 원하는 상품을 추천

해주는 추천시스템에 대한 연구가 활발히 진행되어, 추천시스템은 인터넷 쇼핑몰을 재구성하는 필수적인 비즈니스 도구가 되었다[1,2]. 또한, 현재의 인터넷 쇼핑몰은 고객의 만족도 증가, 효율적인 고객관리, 마케팅 비용의 절감, 서비스 및 제품의 개선 등 경쟁력 강화를 위해 개인별 맞춤정보(Personalization)를 제공하는 형태로 발전해 가고 있다[3].

인터넷 쇼핑몰에서 상품이나 서비스를 추천하기 위한 연구는 크게 사용자들 간의 유사도를 활용한

※ 교신저자(Corresponding Author): 김정인, 주소: 부산광역시 남구 용당동 535(608-711), 전화: 051)629-1174, FAX: 051)629-1169, E-mail: jikim@tu.ac.kr
접수일: 2008년 9월 8일, 완료일: 2009년 3월 23일

[†] 종신회원, 동명대학교 컴퓨터공학과 부교수

※ 본 과제(결과물)는 교육인적자원부와 산업자원부의 출연금 및 보조금으로 수행한 산학협력중심대학육성사업의 연구결과입니다.

협력적 필터링(Collaborative filtering) 방식과 상품 및 사용자의 특성 값들을 활용하는 내용기반 필터링(Content-based filtering) 방식으로 구분할 수 있다. 협력적 필터링 방식은 사용자들의 평가 정보를 얻기 어려운 경우가 많고, 평가 정보의 수가 적을 경우 추천 성능이 매우 감소한다. 그리고 내용 기반 필터링 방식에서도 로그인을 하지 않은 상황에는 사용자에게 대한 정보를 얻을 수가 없어 협력적 필터링 방식에서의 평가 정보와 마찬가지로 구하기 어려운 경우가 많다. 이것을 해결하고자 구매했거나 사용자가 관심을 표명한 상품의 특성을 바탕으로 사용자 프로파일을 구성하는 방식이 많이 사용된다[4]. 그리고 내용기반 필터링 방식에서는 상품의 특성을 가장 잘 나타낼 수 있는 속성을 파악해 내는 것이 매우 중요하다. 본 연구에서는 여성복의 패션 코디네이션 시스템 구축을 위해 여성복의 특성을 가장 잘 나타낼 수 있는 속성을 정의하고, 내용 기반 필터링 방식을 이용하여 가장 유사한 의류를 찾아내는 방법과 그에 따른 알고리즘을 작성하여, 어떤 의류를 선택한 후 가장 유사한 의류를 추천할 수 있도록 시스템을 구축하고 그 결과를 분석하였다. 각 속성의 분류와 속성에 사용된 속성값의 지정은 유사한 의류를 찾아내는데 결정적인 역할을 담당하지만, 어떤 분류방법과 속성 값이 가장 적절한지는 경험을 통해서 꾸준히 조절해 나가야하며, 차후 연구주제로 다루기로 한다.

2. 관련연구

2.1 추천시스템의 분류(Taxonomy) [5]

추천시스템은 소비자가 원하는 상품을 추천하고자 상품에 대한 요구와 공동체(Community)의 정보를 결합하여 추천엔진에 입력으로 입력하면, 추천엔진에서 소비자에게 개인화 수준과 정보제공 방법에 따라 그림 1과 같이 적절한 형태로 상품을 추천한다. 그리고 이 상품을 사용한 소비자가 피드백(사용 만족도 및 후기)을 제공함으로써 다음번 추천을 위한 추가적인 데이터로 입력하여 사용할 수도 있다.

일반적으로 추천시스템의 추천 방법(Recommendation Method)은 다음과 같다.

- 단순검색(Raw retrieval)은 고객이 찾고자 하는 내용의 검색 결과를 제공한다.
- 수동선택(Manual selection)은 고객이 추천을

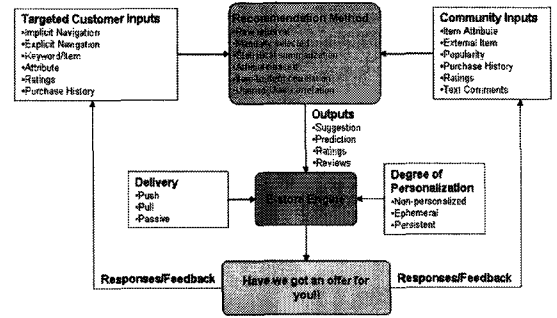


그림 1. 추천시스템의 동작과정

요구하면 해당 분야의 전문가가 고객이 좋아할 제품을 예측하여 제공한다.

- 통계적 요약(Statistical summaries)은 고객이 추천을 요구하면 고객이 속하는 공동체에서 가장 인기있는 제품을 추천한다.
- 속성기반(Attributed-based)은 아이템의 속성에 대한 고객의 선호도에 따라 추천을 수행한다.
- 아이템간 관계(Item-to-item correlation)는 상품을 작은 집합으로 구분하여 고객이 특정 집합에 포함된 상품을 구입하면 같은 집합에 포함된 상품들을 고객에게 추천하는 방식이다.
- 사용자간 관계(User-to-item correlation)는 추천을 요구하는 고객과 제품을 구매한 다른 고객과의 관계에 기초하여 고객에게 제품을 추천한다.

2.2 협력적 필터링(Collaborative filtering) 방식

협력적 필터링 방식은 많은 연구에서 성공적인 결과를 보여주는 것으로 제시되고 있으나, 협력적 필터링에 사용되는 사용자들의 평가 정보를 얻기 어려운 경우가 많고, 또한 평가 정보의 수가 적을 경우 추천 성능이 매우 감소하는 희소성(Sparsity)의 문제를 갖고 있다. 따라서 이러한 경우 내용 기반 필터링 방식이 현실적인 추천도구가 될 수 있다[4].

2.3 내용기반 필터링(Content-based filtering) 방식

내용 기반 필터링에서는 상품의 특성과 사용자의 특성이 모델링되고 각 특성을 연결해 주는 학습 결과나 추천 모델에 따라 상품의 추천이 이루어진다. 상품 및 사용자의 특성은 인구 통계학적 특성, 사용자가 명시한 선호 특성 등 다양한 정보가 존재할 수

있으나 일반적인 상황에서는 등록되지 않은 사용자에 대해서 인구 통계학적 특성을 얻기 어려운 경우가 많고, 사용자가 명시한 선호 특성들도 협력적 필터링에서의 평가 정보와 마찬가지로 구하기 어려운 경우가 많다. 따라서 그 보다는 사용자가 구매했거나 사용자가 관심을 표명한 상품의 특성을 바탕으로 사용자 프로파일을 구성하는 방식이 많이 사용된다[4,6].

그러나 비교적 제품의 속성이 간단한 경우를 제외하면 인터넷 쇼핑몰에는 많은 종류의 상품과 서비스가 존재하고, 이로 인해 다양하고 복잡한 종류의 속성 값을 갖게 된다. 예를 들어, 여성복의 속성은 색상과 치수뿐 아니라 상품에 대한 서술적인 설명도 상품의 속성이 될 수 있다. 그러나, 텍스트 형태의 상품 설명을 사용하는 경우, 서술적인 문장에 포함된 키워드의 수만큼 속성의 수가 많아질 수 있다. 이로 인해 적절하지 못한 속성이 많이 생길 수 있으며 이것 때문에 추천 성능이 떨어질 수 있다. 그리고 너무 많은 수의 자질을 유지하고 검색해야 하는데서 오는 계산 및 시간부담이 있을 수 있다. 따라서 이런 문제점을 해결하고자 상품의 특성을 가장 잘 나타낼 수 있는 속성을 파악해 내는 것은 매우 중요한 일이다.

2.4 추천시스템의 유형에 따른 적용 기술

표 1은 추천시스템의 유형에 따라 일반적으로 사용되는 기술을 정리한 것이다[7].

관련연구를 분석해본 결과, 내용 기반의 방식 중 휴리스틱 기반 방식으로 시스템을 개발하려면 상품의 특징을 잘 나타낼 수 있는 속성을 정의하는 것이 필수적이다. 본 연구에서는 프로토타입 제작을 위해 시스템 개발시 파급 효과가 가장 큰 분야인 여성복을

도메인으로 선택하였다. 그리고 효과적인 속성 분류를 위해 3명의 의류 전문가와 함께 옷의 특징을 가장 잘 나타낼 수 있는 속성과 속성 값을 정의하였다.

3. 시스템 설계

3.1 패션 코디네이션 시스템을 위한 의류의 주요 속성

패션 코디네이션(fashion coordination)은 의복 및 액세서리 등의 새로운 착장법을 나름대로 개성 있게 조화시켜 입는 방식이다[8]. 패션 코디네이션은 크게 스타일, 컬러, 이미지, 아이템, 소재, 체형, 소품에 따른 코디로 구분할 수 있다[9,10].

첫째, 스타일에 따른 코디네이션은 칼라, 소재, 디테일, 실루엣이 나타내는 여러 가지 이미지를 하나의 공통된 연관성에 따라 분류하고 다시 통일감있는 룩으로 표현된 옷차림을 뜻한다. 스타일 코디에는 플러스 원 코디네이션(하나를 더해서 전체 의상에 변화를 주어 새로운 감각을 느끼게 한다)과 토털 코디네이션(단순히 의복만이 아니라 헤어, 메이크업, 액세서리, 구두 등 머리에서 발끝까지 조화를 이루도록 계산된 스타일링을 한다)과 크로스 오버코디네이션(반대적인 모양이나 이미지를 조합시켜 거기에서 오는 기묘함과 의외성을 추구한다)으로 구분할 수 있다. 둘째, 컬러에 따른 코디네이션은 다양한 색상을 가지고 다양한 배색을 만드는 매칭 방법으로 디자인, 소재와 함께 가장 중요한 디자인 포인트가 되고 있다. 컬러 코디 방법에는 크게 같은 계통의 색상을 배색하여 무난한 조화를 추구하는 하모니 컬러 코디네이션과 서로 반대되는 계통의 색상을 배색하여 대조적인 이미지를 추구하는 콘트라스트 컬러 코디네이

표 1. 추천시스템의 유형에 따른 적용 기술

Recommendation Approach	Recommendation Technique	
	Heuristic-based	Model-based
Collaborative	<ul style="list-style-type: none"> • Nearest neighbor(cosine, correlation) • Clustering • Graph theory 	<ul style="list-style-type: none"> • Bayesian networks • Clustering • Artificial neural networks • Linear regression • Probabilistic models
Content-based	<ul style="list-style-type: none"> • TF-IDF(Information retrieval) • Clustering • Nearest neighbor 	<ul style="list-style-type: none"> • Bayesian classifiers • Clustering • Decision trees • Artificial neural networks

선 방법이 있다. 셋째, 이미지에 따른 코디네이션은 색상, 형태, 질감 등에서 느껴지는 분위기, 감각, 연상을 결합하여 새로운 이미지를 만들어낸다. 이미지 코디는 크게 캐주얼, 모던, 로맨틱, 내추럴, 엘레강스, 댄디 등으로 나눌 수 있으며 각각의 테마에 따른 의상 및 액세서리의 색상, 아이템, 소재 등을 다양하게 코디 하여 분위기를 연출한다. 넷째, 아이템에 따른 코디네이션은 스커트, 블라우스, 자켓, 코트, 스웨터 등의 각종 의상 아이템을 상하로 코디 하는 방법으로 단품위주의 경제적이고 저렴한 구매로 각자의 개성과 분위기를 연출할 수 있다. 다섯째, 소재에 따른 코디네이션은 각 소재가 가지고 있는 고유한 질감, 패턴, 드레이프성 등을 고려하여 의상의 용도와 유행에 맞는 코디를 하는 방법으로 최근에는 가공기술의 발달로 색다른 합성섬유 및 특수첨단소재의 개발로 다양한 연출을 가능하게 하고 있다. 여섯째, 체형에 따른 코디네이션은 착용자의 사이즈와 체형의 장단점을 정확히 판단하고 이를 커버하거나 드러내기도 하여 인체와 의상을 최대한 조화롭게 코디 하는 방법을 말한다. 일곱째, 소품에 따른 코디네이션은 구두, 모자, 핸드백, 스카프와 목걸이, 귀걸이, 반지 등의 주얼리 제품을 이용하여 의상의 분위기를 살려주는 코디방법을 말한다[11]. 실제로 이와 같은 다양한 코디 방법이 사용되고 있으나, 기존의 패션 코디네이션 시스템에 대한 연구는 3D 인체모델 기반으로 웹상에서 옷을 입혀보면서 사용자가 코디네이션 하는 형태가 주를 이룬다[11,12]. 본 논문에서는 의류에 속성을 부여하는 방식과 부여된 속성값을 계산하여 의류간의 유사도를 찾아내는 방법으로 유사의류를 추천하는 방식을 제안한다.

3.2 의류의 속성 설계

본 연구에서의 패션 코디네이션 시스템은 사용자

가 취향에 따라, 혹은 검색 속성을 지정하여 원하는 의복을 선택하면, 이와 유사한 옷들을 컴퓨터가 추천해주는 시스템이다. 이를 위하여 여성복을 아이템별로 그룹화 하였으며, 상의는 재킷, 블라우스, 셔츠로 구분하고 하의는 팬츠와 스커트로 구분하였다. 유사의류 추천 시스템을 구축하고자 정의한 속성을 정리하면 다음과 같다. 우선 모든 여성복의 공통속성으로 소재와 문양, 색상을 정의하였다. 그 후, 상의(자켓, 블라우스, 셔츠)와 하의(팬츠, 스커트)에 대한 각각의 속성은 따로 분류하였으며, 상의의 공통속성은 옷깃(칼라), 소매의 길이, 기장(전체길이)으로 구분하였고 재킷과 블라우스는 추가적인 속성을 정의하였다. 재킷은 버튼의 수와 소매모양을, 블라우스는 소매모양을 들 수 있다. 하의의 경우는 기장을 공통 속성으로 정의하였다. 팬츠의 추가 속성으로는 무릎 윗부분의 바지통과 무릎 아랫부분의 바지통, 돌출형 포켓 등을 정의하였다. 그리고 스커트의 추가 속성에는 주름모양과 전체 모양(디자인)을 들 수 있다. 이것을 정리하면 표 2와 같다.

위와 같은 의류의 다양한 속성을 표현하기 위하여, 본 시스템에서는 여성복의 특징을 핵심속성, 보통속성, 상세속성으로 나누었다. 핵심속성은 색상, 문양, 소재, 기장 속성이며 시간적으로 가장 크게 차이를 체감할 수 있는 속성이다. 블라우스에만 적용되지 않는 기장속성을 제외하면 의류의 공통속성에 해당한다. 보통 속성은 라인, 네크라인, 소매길이, 옷깃(칼라), 라인, 버튼수 속성이며 상의나 하의의 부분적인 차이를 나타내는 속성에 해당한다. 상세속성은 커프스, 포켓, 디테일 속성으로 의류의 국소적인 차이를 나타내는 속성이다. 유사한 의류를 찾기 위해서는 핵심속성, 보통속성, 상세속성의 순서로 우선되어야 함을 뜻한다.

각 속성은 옷의 분류에 따라 적용되는 속성과 그

표 2. 여성복의 대표적 특징

분 류		속 성
옷의 공통 속성		소재, 문양, 색상
상의	재킷(원피스, 외투)	라인, 네크라인, 소매길이, 옷깃(칼라), 버튼수, 기장, 포켓
	블라우스	라인, 네크라인, 소매길이, 옷깃(칼라), 버튼수, 소매모양
	셔츠	라인, 네크라인, 소매길이, 옷깃(칼라), 버튼수, 기장
하의	바지	기장, 라인(무릎 윗부분의 바지통과 무릎 아랫부분의 바지통 형태), 돌출형 포켓
	치마	기장, 라인(전체 모양), 디테일(상세모양)

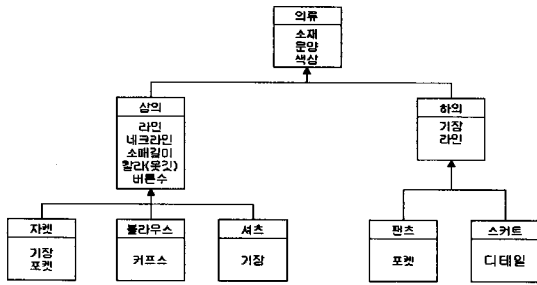


그림 2. 여성복 속성의 상속관계

량지 않은 속성으로 나눌 수 있다. 그림 2에 각 의류들의 속성들을 객체 모델링화하여 속성의 상속관계를 정의하였다.

3.3 핵심 속성과 속성 값

본 시스템에서 제안하는 유사의류 추천방법은 속성값을 세밀하게 잘 지정하여야 한다. 같은 속성 속에서도 비슷한 속성일수록 그 속성 값의 차이가 적도록 설계하여야 유사한 의류를 찾아낼 때, 속성값의 계산량으로 사용할 수 있게 된다. 그 차이가 얼마이어야 가장 적절한지는 많은 시행을 통해서 경험적으로 결정되어야 할 문제이다. 핵심속성의 속성값을 지정한 예를 나타내면 다음과 같다.

3.3.1 소재

소재는 옷을 만드는 재료를 나타내며 모든 옷의 공통 속성에 해당하고 원료섬유의 특징, 제직방법, 원단가공법 등에 따라 다양하게 분류될 수 있으나, 궁극적으로 본 시스템은 의류쇼핑몰에서도 특히 패션 코디네이션에 초점을 맞추었으므로 시각적으로 보이는 질감에 기준을 두어 분류하였다. 면, 마, 데님, 모피, 캐시미어, 알파카, 새틴, 쉬폰, 벨벳, 가죽, 실크, 레이스, 레이온, 나일론, 폴리에스테르, 폴리우레탄 등으로 분류되며 이들을 적당히 혼합한 형태의 소재도 많이 있다. 상의 하의에 공통적으로 적용되는 속성이며, 어떤 소재로 옷을 만드느냐에 따라 옷의 형태가 다르게 표현될 수 있으므로 본 시스템에서 중요하게 다루어야 할 속성이다. 표 3과 같이 식물성 재질감 소재를 1x속성으로, 동물성 재질감 속성을 2x로, 합성 재질감 속성을 3x로 표현하였다.

3.3.2 문양

의복학에서 문양은 구성적인 면과 소재적인 면으

표 3. 소재 속성의 상세

구 분	속성이름(name)	속성값(value)
식물성 재질감 소재	면(Cotton)	10
	데님(Denim)	11
	마(Hemp)	15
동물성 재질감 소재	울(Wool)	20
	모피(Fur)	23
	캐시미어(Cashmere)	24
	알파카(Alpaca)	25
	벨벳(Velvet)	26
	실크(Silk)	27
	가죽(Leather)	29
	레이스(Lace)	33
합성 재질감 소재	레이온(Rayon)	34
	나일론(Nylon)	35
	폴리에스테르(Polyester)	36
	폴리우레탄(Spandex)	37
	새틴(Satin)	38
	쉬폰(Chiffon)	39

로 구별하여 분류된다[10]. 구성적인 구분은 다시 단독구성, 복수구성, 방향성구성, 전면구성, 회화풍구성으로 나누어지며, 소재적인 구분은 자연문양으로 식물문양과 동물문양, 기하학문양으로 점문양, 추상문양, 전통문양으로 나누는데, 이 구분 방법들은 유사한 의류를 찾아내는 분류로 사용하기에는 적당하지 않다. 본 연구에 사용되는 문양의 분류는 없거나

표 4. 문양 속성의 상세

구 분	속성이름(name)	속성값(value)
문양없음	없음	10
문양있음	도트(Dot)	21
	식물(Plant)	22
	동물(Animal)	23
	캐릭터(Character)	24
	컴플라주(Camouflage)	25
	문자열(String)	26
	심벌(Symbol)	27
	가로(Horizontal Line)	35
	세로(Virtical Line)	37
	타탄(Tartan)	41
	작은 체크(narrow)	43
	큰 체크(Wide)	47

표 5. 색상 속성의 상세

구분	속성이름(name)	밝기	색상	속성값
빨강 (Red)	밝은 색(Light)	10	1	11
	선명한 색(Vivid)	20	1	21
	어두운 색(Dark)	30	1	31
핑크 (Pink)	밝은 색(Light)	10	2	12
	선명한 색(Vivid)	20	2	22
	어두운 색(Dark)	30	2	32
주황 (Orange)	밝은 색(Light)	10	3	13
	선명한 색(Vivid)	20	3	23
	어두운 색(Dark)	30	3	33
노랑 (Yellow)	밝은 색(Light)	10	4	14
	선명한 색(Vivid)	20	4	24
	어두운 색(Dark)	30	4	34
부라운 (Brown)	밝은 색(Light)	10	5	15
	선명한 색(Vivid)	20	5	25
	어두운 색(Dark)	30	5	35
초록 (Green)	밝은 색(Light)	10	6	16
	선명한 색(Vivid)	20	6	26
	어두운 색(Dark)	30	6	36
파랑 (Blue)	밝은 색(Light)	10	7	17
	선명한 색(Vivid)	20	7	27
	어두운 색(Dark)	30	7	37
남색 (Navy)	밝은 색(Light)	10	8	18
	선명한 색(Vivid)	20	8	28
	어두운 색(Dark)	30	8	39
보라 (Purple)	밝은 색(Light)	10	9	19
	선명한 색(Vivid)	20	9	29
	어두운 색(Dark)	30	9	39
흰색 (White)	흰색(White)	0	0	00
회색 (Gray)	밝은 색(Light)	10	0	10
	선명한 색(Vivid)	20	0	20
	어두운 색(Dark)	30	0	30
검정색 (Black)	검정색(Black)	40	0	40
기타색 (Other)	금은색(Gold, Silver)	50	0	50

혹은 있는 형태이다. 있는 형태는 다시 크게 3가지 정도로 나누어진다. 스트라이프, 체크, 프린트 문양 정도인데 옷을 구별하는 속성으로 중요하다. 스트라이프는 다시 초크, 트리플, 펜슬, 블록, 헤링본, 라인

보우 등으로 나누어지고 체크는 타탄, 글렌, 윈도우, 블록, 기업, 스티치트로 분류할 수 있다. 그러나, 스트라이프와 체크는 큰 특징점만 살려서 보다 단순화하여 속성값을 부여하였다. 프린트 문양은 다양하며 크게 아가일 도트, 식물, 동물, 캐릭터(만화), 킴플라주(밀리터리), 문자열, 심벌 등으로 나누어 표 4와 같이 속성값을 부여하였다.

3.3.3 색상

의복에서 색상은 우선 무지개 색을 기본으로 하며, 그기에 보다 옅은 색상과 진한 색상으로 구분한다. 또한, 흰색, 검정색, 그리고 회색이나 브라운, 핑크, 자주 등은 자주 등장하는 색상이므로 따로 관리한다. 색상 속성은 고유색상과 명암에 따라 적절한 속성을 부여하여야 한다. 표 5의 속성 부여방식을 따르면 밝은 노랑색과 흰색의 차이는 4가 되며, 밝은 보라색과 흰색과의 차이는 9가 된다. 또한 밝은 노랑색과 어두운 노랑색은 그 차이가 20이 된다.

3.3.4 기장 속성

기장은 옷의 길이를 나타내는 속성이다. 블라우스를 제외하면 다른 종류의 옷들에 공통적으로 적용되는 속성이다. 상의는 대체로 가슴선길이, 허리선길이, 엉덩이 길이, 엉덩이 아래 길이, 무릎길이, 발목길 이로 나뉘어 진다. 하의는 팬츠와 스커트에 따라 다양한 형태의 길이가 존재하며, 본 시스템에 적용하기 위한 속성을 정리하여 표 6에 나타낸다.

표 6. 기장 속성의 상세

구분	속성이름(name)	속성값(value)
상의	가슴선 길이(Bast Line)	10
	허리선 길이(Waist Line)	13
	엉덩이 길이(Hip Line)	16
	엉덩이 아래길이(Low Hip Line)	18
	무릎근처길이(Knee Line)	22
	발목근처길이(Ankle Line)	25
하의	아주짧음(Very shorts)	10
	짧음(Shorts)	13
	무릎근처(Knee Line)	15
	장단지길이(Midi Line)	20
	발목길이(Ankle Line)	25

4. 속성 값을 이용한 유사의류 검색

4.1 유사도 계산방법

두 의류의 유사도를 계산하는 가장 간단한 방법으로는 정의된 속성별로 속성 값을 비교하여 같으면 1, 같지 않으면 0으로 할당하여 그 값들을 모두 더하는 것이다. 재킷의 경우는 10가지 속성을 부여하였으므로 모든 속성이 같을 경우, 최대 10점을 획득하고 속성이 모두 다를 경우 최소 0점을 획득하여 0에서 10 사이의 유사도를 결정할 수 있다.

$$S = \sum_{i=1}^n (A_i \Delta B_i) \tag{식 (1)}$$

$$(A_i \Delta B_i) \begin{cases} = 1, & \text{if } (A_i = B_i) \\ = 0 & \text{otherwise} \end{cases}$$

S: 두 옷의 유사도

A_i: A옷의 i 번째 속성 값

B_i: B옷의 i 번째 속성 값

n: 속성의 개수

그러나, 식(1)은 속성의 중요도가 무시되므로 정확한 유사도라 보기 어렵다. 즉, 단추의 개수가 1개 차이 나는 옷과 색상이 현저하게 다른 옷을 같은 점수로 처리하게 된다. 이를 개선하는 방법으로 가중치를 도입할 수 있다. 식(2)는 가중치가 도입된 것이다. 적절한 가중치를 지정하여 단추개수의 차이보다 색상 차이를 몇 배로 중요하게 생각하도록 처리할 수 있다.

$$S = \sum_{i=1}^n (w_i \times (A_i \Delta B_i)) \tag{식 (2)}$$

$$(A_i \Delta B_i) \begin{cases} = 1, & \text{if } (A_i = B_i) \\ = 0 & \text{otherwise} \end{cases}$$

w_i: i 번째 속성의 가중치

이렇게 의류 간의 유사도를 구하는 식을 바꾸면 최대 유사도는 가중치를 더한 값이 되고 최소 유사도는 0이 된다. 속성에 가중치를 도입하여 조금 더 개선한 방법이라 할 수 있으나, 식 (2)도 같은 속성 안에서 유사한 정도를 나타내는 것은 불가능하다. 즉, 가장 속성에서 허리길이와 엉덩이 길이는 유사하지만 허리길이와 발목길이는 유사하지 않다. 위의 식으로 계산하게 되면 가장 속성이 같지 않을 경우는 그 차이의 미세함과는 상관없이 무조건 0이 된다.

속성 값을 적절히 부여하면 같은 속성 속에서의 유사도를 차별하여 계산하는데 활용할 수 있다. 예를

들면, 밝은 빨강 색과 빨강 색, 그리고 어두운 빨강 색은 모두 빨강 색 계열이다. 속성 값은 다르지만 상당히 유사하다고 할 수 있다. 또한, 밝은 노랑색, 밝은 녹색 등도 그 차이가 크지 않다고 볼 수 있다. 이런 경우, 속성 값을 근사하게 부여하여 두 옷의 속성 값을 계산할 때, 그 값의 차이가 작게 나도록 구성한다. 식(2)를 변형하여 같은 속성 안에서 닮은 정도를 고려하는 Nearest neighbor 방식의 식(3)을 생각해볼 수 있다. 식 (3)은 속성 값의 차이만큼 그 정도를 반영하여 유사도를 계산하도록 개선한 것이다. 속성 값의 거리 만큼 유사도에 반영하기 위하여 차의 제곱 값을 구하였다. 이는 두 가지 이상에서 속성 값이 조금 차이나는 경우와 한 가지에서 속성 값이 크게 차이나는 경우 속성 값이 크게 차이나는 경우를 더욱 유사하지 않다고 판정하는데 알맞은 방식으로 사용된다.

$$S = \sum_{i=1}^n (w_i \times (A_i - B_i)^2) \tag{식 (3)}$$

식(1)과 식(2)는 속성의 개수 혹은 가중치의 합이 만점인 계산에서 높은 점수를 획득할 수록 두 옷의 유사도가 높다고 판단하지만, 식(3)은 두 옷의 차이를 속성 값에 반영하여 제곱 계산을 하였으므로, 속성 값이 차이가 없는 경우 (A_i - B_i)는 0이 되고 따라서, 유사도를 %로 보여줄 수는 없지만 S값이 작을수록 속성들의 거리가 가깝고 유사도가 높다는 것을 나타낸다.

4.2 의류 간의 속성 값 계산 실험

실제 의류의 유사한 정도를 파악하기 위하여 하나의 기준 셔츠를 두고 모양이나 색상 등이 다른 5개의 티셔츠를 대상으로 유사도를 측정하였다. 현실감을 높이기 위하여 인터넷 상에서 판매되고 있는 의류로 골랐으며[13,14], 위의 3가지 수식으로 계산한 결과를 보이도록 한다. 실험1의 데이터로 사용된 그림 3은 유사한 옷들이 2그룹으로 구별되도록 구성하였다. 육안으로 보면 A,B,C와 D,E,F가 서로 유사한 옷 처럼 보인다. 표 7에 이들 옷들의 속성을 나타낸다.

A를 기준으로 B부터 F까지의 의류에 대한 유사도를 3개의 식으로 구하였다. 식1과 식2는 속성의 일치에 1점씩 점수를 부과하였으며, 식2는 속성에 대한 가중치를 곱한 것이다. 식3은 같은 속성 내에서 속성의 차이(거리)를 제곱 계산하고 거기에 가중치를 곱

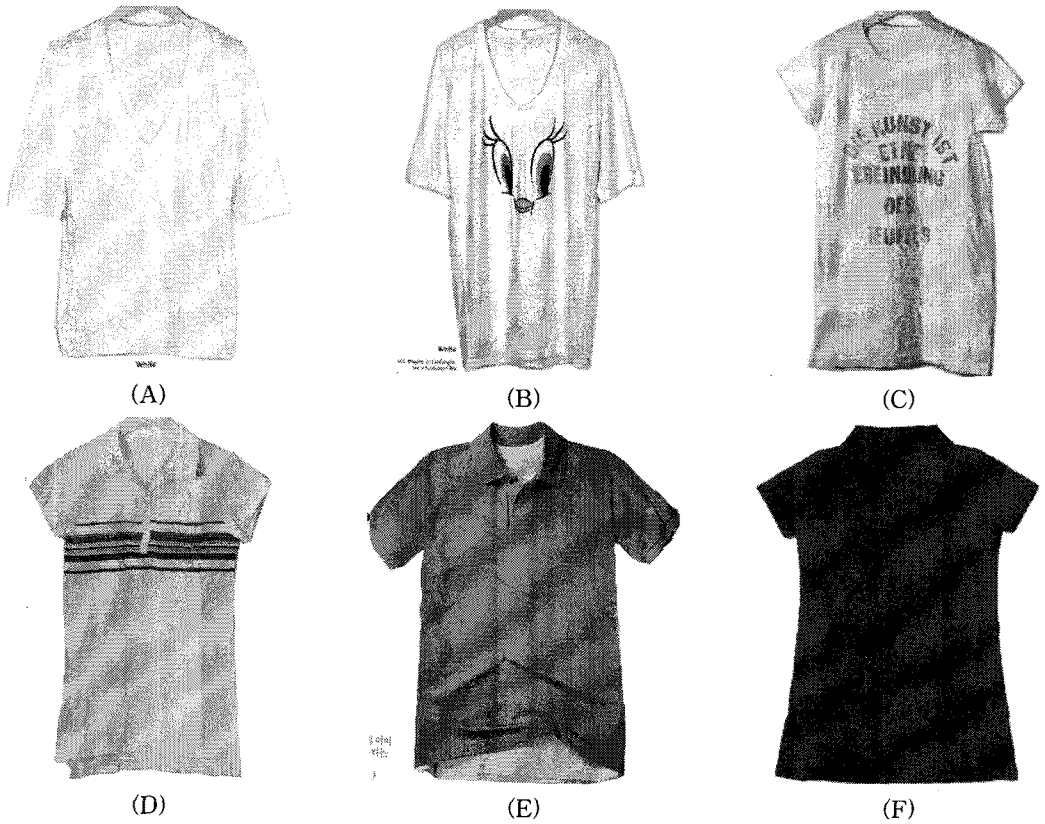


그림 3. 실험에 사용된 셔츠 6벌

표 7. 6벌의 셔츠에 대한 속성 및 속성 값

유형	A		B		C		D		E		F	
소재	Cotton	10	Cotton	10	Cotton	10	Cotton	10	Cotton	10	Cotton	10
문양	None	10	Character	24	String	26	Stripe	35	None	10	None	10
색상	White	0	White	0	White	0	White	0	Pink	22	Orange	23
라인(형태)	Loose	14	Loose	14	Loose	14	semi-fit	12	semi-fit	12	semi-fit	12
넥라인	V Line	26	Round	20	Round	20	Basic	0	Basic	0	Basic	0
소매길이	3/4Length	20	Plain	12	Cap	10	Cap	10	Plain	12	Cap	10
칼라(옷깃)	None	0	None	0	None	0	Shirts	30	Shirts	30	Shirts	30
버튼수	None	0	None	0	None	0	2Buttons	10	1Buttons	5	3Buttons	15
기장	Hip	16	Hip	16	Hip	16	Hip	16	Hip	16	Hip	16
포켓	None	10	None	10	None	10	None	10	Kangaroo	7	None	10
커푸스	None	0	None	0	None	0	None	0	Band	3	None	0

한 것이다.

실험1의 테이터에서 3가지 식으로 계산한 결과 값을 표 8로 정리하였다. 식1에 의하면 A와 가장 유사한 옷으로 똑같이 11개의 속성 중에서 8개의 속성에

서 일치율을 보여 73%(8 / 11 * 100)의 유사도를 보인 B와 C를 뽑았으며, 가장치를 고려한 식2로도 59%(10 / 17 * 100)의 유사도로 같은 결과를 보였다. 그러나, 식3은 B와의 거리를 496, C와의 거리를 664

표 8. A기준 5벌의 셔츠에 대한 속성값 및 유사도 실험 결과(라인, 네크라인, 소매길이 속성의 가중치가 3인 경우)

유형	가중치	A			B			C			D			E			F			
		1식	2식	3식	1식	2식	3식	1식	2식	3식	1식	2식	3식	1식	2식	3식	1식	2식	3식	
소재	1	10	1	1	0	1	1	0	1	1	0	1	1	0	1	1	0	1	1	0
문양	1	10	0	0	196	0	0	256	0	0	625	1	1	0	1	1	0	1	1	0
색상	1	0	1	1	0	1	1	0	1	1	0	0	0	484	0	0	529	0	0	529
라인(형태)	3	14	1	3	0	1	3	0	0	0	12	0	0	12	0	0	12	0	0	12
네크라인	3	26	0	0	108	0	0	108	0	0	2028	0	0	2028	0	0	2028	0	0	2028
소매길이	3	20	0	0	192	0	0	300	0	0	300	0	0	192	0	0	300	0	0	300
칼라(웃깃)	1	0	1	1	0	1	1	0	0	0	900	0	0	900	0	0	900	0	0	900
버튼수	1	0	1	1	0	1	1	0	0	0	100	0	0	25	0	0	225	0	0	225
기장	1	16	1	1	0	1	1	0	1	1	0	1	1	0	1	1	0	1	1	0
포켓	1	10	1	1	0	1	1	0	1	1	0	0	0	9	1	1	0	1	1	0
커프스	1	0	1	1	0	1	1	0	1	1	0	0	0	9	1	1	0	1	1	0
합계	17		8	10	496	8	10	664	5	5	3965	3	3	3659	5	5	3994			
유사도			73%	59%	-	73%	59%	-	45%	29%	-	27%	18%	-	45%	29%	-			

로 계산하여 B가 좀 더 유사하다고 판정하였다. 어느 식도 육안으로 비슷하게 보이는 옷들에 대하여 유사하다고 판단한 것을 알 수 있다. 또한, 비슷하지 않은 그룹에 대해서도 유사도를 계산하였는데, 식1을 이용하면 A에 대하여 D와 F가 각각 45%의 유사도를 보였고 식2로도 똑같이 29%의 유사도를 보였다. 가장 유사하지 않은 옷으로는 식1, 식2의 유사도가 각각 27%, 18%를 보인 E를 골랐다. 그러나, 식3의 유사한 거리 결과는 3659의 E, 3965의 D, 3994의 F순으로

훨씬 정밀한 차이를 구할 수 있었다.

식1의 결과 : A → (B, C) → (D, F) → E

식2의 결과 : A → (B, C) → (D, F) → E

식3의 결과 : A → B → C → E → D → F

같은 실험1의 데이터지만, F를 기준으로 유사도를 계산한 결과 값을 표 9로 정리하였다. 식1에 의하면 F와 가장 유사한 옷은 E(73%), 그 다음으로 D(64%), A와 C(같은 45%), B(36%)의 순이었다. 식2로는

표 9. F기준 6벌의 셔츠에 대한 속성값 및 유사도 실험 결과(라인, 네크라인, 소매길이 속성의 가중치가 3인 경우)

유형	가중치	F			A			B			C			D			E			
		1식	2식	3식	1식	2식	3식	1식	2식	3식	1식	2식	3식	1식	2식	3식	1식	2식	3식	
소재	1	10	1	1	0	1	1	0	1	1	0	1	1	0	1	1	0	1	1	0
문양	1	10	1	1	0	0	0	196	0	0	256	0	0	625	1	1	0	1	1	0
색상	1	23	0	0	529	0	0	529	0	0	529	0	0	529	0	0	529	0	0	1
라인(형태)	3	12	0	0	12	0	0	12	0	0	12	1	3	0	1	3	0	1	3	0
네크라인	3	0	0	0	2028	0	0	1200	0	0	1200	1	3	0	1	3	0	1	3	0
소매길이	3	10	0	0	300	0	0	12	1	3	0	1	3	0	0	0	0	0	0	12
칼라(웃깃)	1	30	0	0	900	0	0	900	0	0	900	0	0	0	1	1	0	1	1	0
버튼수	1	15	0	0	225	0	0	225	0	0	225	0	0	25	0	0	25	0	0	100
기장	1	16	1	1	0	1	1	0	1	1	0	1	1	0	1	1	0	1	1	0
포켓	1	10	1	1	0	1	1	0	1	1	0	1	1	0	1	1	0	1	1	9
커프스	1	0	1	1	0	1	1	0	1	1	0	1	1	0	1	1	0	1	1	9
합계	17		5	5	3994	4	4	3074	5	7	3122	7	13	1179	8	12	131			
유사도			45%	29%	-	36%	24%	-	45%	41%	-	64%	76%	-	73%	71%	-			

D(76%), 그 다음으로 E(71%), C(41%), A(29%), B(24%)순이었다. 마지막으로 식3의 결과는 E(131), D(1179), B(3074), C(3122), A(3994)임을 알 수 있다.

식1의 결과 : $F \rightarrow E \rightarrow D \rightarrow (A, C) \rightarrow B$

식2의 결과 : $F \rightarrow D \rightarrow E \rightarrow C \rightarrow A \rightarrow B$

식3의 결과 : $F \rightarrow E \rightarrow D \rightarrow B \rightarrow C \rightarrow A$

즉, 3가지 식 모두 A,B,C를 비슷한 그룹으로, D,E,F는 다른 그룹으로 분리된 것을 알 수 있지만, 식1,2보다 식3의 결과는 유사한 정도를 훨씬 정밀한 수치로 표현할 수 있었다.

5. 결 론

인터넷을 이용하는 패션/코디 쇼핑물은 그 시장이 매년 큰 폭으로 증가하고 있다. 하지만 대부분의 쇼핑 물들은 아직도 카테고리에 의한 상품 분류를 제외하면 특정한 의류를 찾을 수 있는 기능이 포함되어 있지 않아 사용자들이 불편을 겪고 있다. 본 연구에서는 여성복의 패션 코디네이션 시스템 구축 시 내용 기반의 방식 중 휴리스틱 기반 방식으로 시스템을 구축하기 위해 여성복의 특성을 가장 잘 나타낼 수 있는 속성을 상속관계로 정의하고, 상품 관리자가 속성 값을 입력할 수 있도록 하였다. 또한 어떤 의류를 선택한 후에 가장 유사한 의류를 찾기 위한 3가지 방법을 소개하고 비교 분석하였다. 특히, 단순히 속성의 동일성만을 계산하는 두 가지 방법과는 달리 속성값의 거리를 계산하여 유사도를 구하는 방법을 제안하였고 그 결과를 같이 분석해 보았다. 컴퓨터가 속성 값 계산에 의해서 유사한 의류를 추천할 수 있다는 것은 패션 코디 쇼핑물을 지능형으로 발전시킬 수 있는 계기가 되며, 이를 응용하면 컴퓨터가 어떤 상의에 대하여 어울리는 하의를 추천할 수 있는 지능형 쇼핑물로 발전할 수 있다. 향후 본 연구를 더욱 발전시켜 지능형 패션/코디 시스템을 완성해 나갈 예정이다.

참 고 문 헌

[1] 정경용, 김진현, 정현만, 이정현, “개인화 추천 시스템에서 연관 관계 군집에 의한 아이템 기반의 협력적 필터링 기술,” 정보과학회 논문지, 소프트웨어 및 응용 제31권, 제4호 pp. 467-477,

2004. 4.
 [2] 김영설, 김병천, 윤병주, “개선된 추천시스템을 이용한 전자상거래시스템 설계 및 구현,” 정보처리학회논문지, 제9-D권, 제2호, pp. 329-336, 2002. 4.
 [3] 이건호, 이동훈, “사례기반추론과 규칙기반추론을 이용한 e-쇼핑몰의 상품추천 시스템,” 정보처리학회논문지, D. 제11-D권, 제5호, 통권 제94호, pp. 1189-1196, 2004. 10.
 [4] 안형준, 김종우, “인터넷 상점에서의 내용기반 추천을 위한 상품 및 고객의 자질 추출 성능 비교,” 정보처리학회논문지, D. 제13-D권, 제2호, 통권 제105호, pp. 279-286, 2006. 4.
 [5] J. Ben Schafer, Joseph A. Konstan, John Riedl, “E-Commerce Recommendation Applications,” *Journal of Data Mining and Knowledge Discovery*, Vol. 5, Issue 1-2 (January-April 2001) pp. 115-153, 2001. 1.
 [6] 임성신, 최창민, 이영숙, 김정인, “지능형 패션코디네이션시스템의 분석,” 한국멀티미디어학회 춘계학술발표대회 논문집, pp. 267-270, 2007.
 [7] Gediminas Adomavicius, Alexander Tuzhilin, “Toward the Next Generation of Recommender Systems: A Survey of the State-of-the-Art and Possible Extensions,” *IEEE TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING*, Vol.17, Issue 6, June 2005, pp. 734-749, 2005.
 [8] 이경순, 김희섭 “의생활과 패션 코디네이션,” 교문사, 1998.
 [9] 오희선, 박화순 “아름다운 여성을 위한 패션코디,” 경춘사, 2000.
 [10] 김영자, “복식미학의 이해,” 경춘사, 2004.
 [11] 강인애, 김효숙, 최창석, “가상 3D 패션 코디네이션 연구,” 대한가정학회지, 제40권, 6호, 통권 172호, pp. 159-171, 2002. 6.
 [12] 박윤희, 신봉기, “전자상거래에서 활용 가능한 의상 시뮬레이션 방법 설계,” 한국정보과학회 춘계학술발표대회논문집, 제30권, 제1호(B), pp. 639-641, 2003.
 [13] <http://www.style-story.com>
 [14] <http://www.funfungirl.biz>



김 정 인

1991년 4월~1993년 3월, 게이오
대학 계산기과학전공 공
학석사

1993년 4월~1996년 3월, 게이오
대학 계산기과학전공 공
학박사

1996년 5월~2월 포항공과대학교 정보통신연구소 연구
원, 기계번역시스템 개발

1998년 3월~현재, 동명대학교 컴퓨터공학과 부교수
관심분야 : 기계번역, 기계학습, 시멘틱웹