

소비자의 감성과 소비유형을 이용한 협업여과기반 콘텐츠 추천 기법

최 인 복[†] · 박 태 근^{**} · 이 재 동^{***}

요 약

협업여과 기법은 추천 시스템에서 널리 사용되는 기술이지만, 소비자의 참조그룹을 선정하는 방법에 따라 추천의 정확도가 달라지는 특성을 가지고 있다. 이에 본 논문에서는 콘텐츠 추천의 정확도를 높이기 위하여 소비자의 감성과 소비유형을 참조그룹으로 하여 협업여과기반으로 콘텐츠를 추천하는 기법을 제안한다. 소비자의 감성을 기쁨, 슬픔, 혐오, 행복, 이완 다섯 가지로 구분하고, 소비유형을 저실용/저쾌락, 저실용/고쾌락, 고실용/저쾌락, 고실용/고쾌락 네 가지로 구분하여 콘텐츠 추천 기법의 성능을 분석한 결과, 본 논문에서 제안하는 기법으로 콘텐츠를 추천한 경우가 소비자 감성과 소비유형을 고려하지 않은 전체 참조그룹으로 추천한 경우보다 정확도가 향상됨을 확인하였다.

키워드 : 콘텐츠 추천, 감성, 소비유형, 협업여과 추천기법

A Contents Recommendation Scheme Based on Collaborative Filtering Using Consumer's Affection and Consumption Type

In-Bok Choi[†] · Taekeun Park^{**} · Jae-Dong Lee^{***}

ABSTRACT

Collaborative filtering is a popular technique used for the recommendation system, but its performance, especially the accuracy of recommendation, depends on how to define the reference group. This paper proposes a new contents recommendation scheme based on collaborative filtering technique whose reference groups are created by consumer's affection and consumption type in order to improve the accuracy of recommendation. In this paper, joy, sadness, anger, happiness, and relax are considered as the consumer's affection. And, low-utility / low-pleasure, low-utility / high-pleasure, high-utility / low-pleasure, and high-utility / high-pleasure are considered as the consumer's shopping types. Experimental results show that the proposed scheme improves the accuracy of recommendation compared to the recommendation scheme considering neither consumer's affection nor consumption type.

Key Words : Contents Recommendation, Affection, Consumption Type, Collaborative Filtering

1. 서 론

최근 인터넷의 발달로 다양한 정보의 홍수 속에서 소비자가 원하는 콘텐츠를 찾는 것이 점차 어려워지고 있다. 따라서 소비자가 원하는 콘텐츠를 찾아내고, 찾아낸 콘텐츠를 순서대로 순위화 해주는 콘텐츠 추천 기법의 중요성이 부각되고 있다. 그러나 지금까지의 콘텐츠 추천기법들 대부분은 소비자의 선호도에 기반한 것들이고[1], 최근에 들어서 소비자의 상황

인식(Context-aware)에 기반한 추천기법들이 연구되고 있으나[10][17], 소비자의 감성을 이용한 연구는 많지 않은 상태이다. 그 이유는 소비자의 감성을 파악하는 것이 어려웠으며, 구분할 수 있는 감성의 종류도 많지 않았기 때문이다.

하지만, 유비쿼터스 시대가 도래함에 따라 BSN(Body Sensor Network)을 이용한 소비자의 다양한 감성측정이 가능하게 되었다[18]. 또한 현대의 소비자 욕구는 대중적인 수준에서 개인적인 수준으로 변화하여 자신의 감성에 맞는 콘텐츠를 요구하는 방향으로 변화해가고 있기 때문에, 소비자의 감성은 콘텐츠를 선택하는 매우 중요한 요소 중 하나로 부각되고 있다[13]. 따라서 이러한 소비자의 변화에 대응하여 소비자의 감성에 따라 적절한 콘텐츠를 제공해 주는 기술이 콘텐츠를 공급하는 측에서 요구되고 있다[9].

추천 시스템을 구성하는 방법에는 크게 내용기반 추천 기법

* 본 연구는 문화관광부 및 한국문화콘텐츠진흥원의 문화콘텐츠기술연구소(CT)육성사업의 연구결과로 수행되었음.

† 준 회 원 : 단국대학교 대학원 정보컴퓨터학과 박사과정

** 정 회 원 : 단국대학교 컴퓨터학부 멀티미디어공학전공 조교수

*** 정 회 원 : 단국대학교 정보컴퓨터학과 교수

논문접수: 2008년 1월 16일

수정일: 1차 2008년 3월 3일, 2차 2008년 3월 24일

심사완료: 2008년 4월 3일

(Content-based Filtering)과 협업여과 추천 기법(Collaborative Filtering)이 있다. 내용기반 추천 기법은 소비자의 선호 정보를 제공 받아 그와 가장 유사한 아이템을 검색해 주는 기법이며, 협업여과 추천 기법은 추천할 대상 콘텐츠에 대한 다른 소비자들의 평가를 입력받아 분석한 후, 비슷한 소비자끼리 묶어서 선호할만한 콘텐츠를 추천하는 기법이다. 개인의 선호정보를 바탕으로 한 내용기반 추천 기법은 공학적인 관점에서 볼 때 매우 상식적인 접근이지만, 사용자가 입력한 짧은 프로파일 안에서만 추천할 수 있으며, 사용자의 온전한 선호 정보 프로파일을 얻기가 쉽지 않다. 또한, 이 기법은 개인별 다량의 경험적 데이터를 필요로 하고, 경험해 보지 못한 것에 대해서는 추천할 수 없는 단점을 가지고 있다[11]. 협업여과 추천 기법은 이러한 내용기반 추천기법의 문제점을 해결하기 위해 다른 소비자의 평가를 기반으로 소비자에게 추천을 생성하는 기법이다[12]. 협업여과 추천 기법은 대상 아이템이 상세한 설명 정보를 갖고 있지 않거나 컴퓨터가 프로세싱하기 어려운 분야의 것이어도 문제없이 잘 동작하며, 소비자로부터 상세한 선호 정보가 수집되지 않은 경우에도 잘 동작한다. 또한, 사용자의 선호정보와 콘텐츠 사이의 유사성을 계산하기 위한 복잡한 과정을 수행하지 않고도 양질의 추천을 제공할 수 있다[11]. 이러한 장점으로 협업여과 추천 기법은 성공적이라 평가되어 왔지만, 추천 대상에 대한 다른 소비자의 평가가 어느 정도 축적되어 있어야 하고, 새로운 아이템이 추가되었을 때, 다른 소비자가 그에 대한 평가를 입력하기 전까지 해당 콘텐츠는 추천될 수 없는 문제점을 가지고 있다. 또한 협업여과 추천 기법은 현재의 소비자와 높은 유사도를 갖는 사람들의 그룹인 소비자 참조그룹을 선정하는 방법에 따라 추천의 정확도가 달라질 수 있다는 어려움이 있다.

본 논문에서는 콘텐츠 추천의 정확도를 높이기 위하여 소비자의 감성과 소비유형을 이용하여 적절한 콘텐츠를 추천하는 협업여과 추천 기법을 제안한다. 제안하는 시스템에서 이용되는 소비자의 감성은 기쁨, 슬픔, 혐오, 행복, 이완 다섯 가지이고, 구분되는 소비유형은 저실용/저쾌락, 저실용/고쾌락, 고실용/저쾌락, 고실용/고쾌락 네 가지이다. 영화와 음악 콘텐츠에 대한 실험을 통하여 본 논문에서 제안하는 추천기법이 소비자 감성과 소비유형을 고려하지 않은 전체 참조그룹으로 추천한 경우보다 정확도가 향상될 뿐만 아니라, 어느 한 가지만 고려한 경우보다도 정확도가 향상됨을 확인한다.

본 논문은 다음과 같이 구성된다. 2장에서는 협업여과 추천 기법, 감성정보를 이용한 콘텐츠 추천 기법 및 소비자 유형 분류기법에 대한 기존 연구들을 살펴본다. 3장에서는 소비자의 감성과 소비유형에 기반한 추천 기법을 제안하고, 4장에서는 추천 기법의 성능을 평가한다. 마지막으로 5장에서는 결론 및 향후 연구방향에 대하여 서술한다.

2. 관련 연구

본 장에서는 본 논문에서 제안하고자하는 콘텐츠 추천 기

법의 설계와 관련된 협업여과 추천 기법, 감성에 따른 추천 기법 및 소비자 유형 분류 기법에 대하여 정리한다.

2.1 협업여과 추천 기법

협업여과 추천 기법은 비슷한 성향 및 선호도를 가지는 다른 소비자들의 평가에 근거하여 추천하는 방법으로 높은 예측능력과 추천능력을 갖는 기법이다. 협업 여과 추천 기법은 크게 모델 기반 기법과 메모리 기반 기법으로 구분된다[16]. 모델 기반 기법은 수학적 모델에 기반을 두고 각 항목에 관한 소비자의 평가를 예측하는 것으로, 크게 클러스터 모델과 베이지안 네트워크 모델로 분류할 수 있다[12]. 클러스터 모델은 소비자들을 어떤 수의 클래스들로 분류하고 클래스 멤버가 해당 아이템에 선호점수를 줄 확률을 계산하여 추천하는 방법이다. 베이지안 네트워크 모델은 노드로서 항목을 나타내고 노드의 상태로서 평가를 나타내는 방법으로 결과 모델이 오프라인에서 구성된다. 메모리 기반 기법은 추천받을 소비자에 대한 부분적 정보와 기존 소비자 데이터베이스로부터 계산된 가중치 집합에 기반하여 특정 소비자의 평가를 예측하는 방법이다.

2.2 감성정보를 이용한 콘텐츠 추천 기법

감성 전이를 위한 감성 콘텐츠 추천 시스템[4]은 우울, 불안, 분노, 지루함과 같은 노인의 부정적인 감성상태를 개선할 수 있도록 적합한 영화 콘텐츠를 추천하는 시스템이다. 이 시스템은 영화 콘텐츠를 사회적 지지, 편안함, 대리만족, 그리고 유머 콘텐츠로 분류하고, 노인의 부정적 감성에 따라 해당 분류의 영화를 제공한다. 시스템은 일상적인 감정상태일 때 사용자의 영화 선호도 정보를 이용하여 추천하고, 부정적인 감정상태일 때는 영화 선호도 정보와 영화의 감정요소를 고려한 정보를 병합해서 추천한다. 이 시스템은 영화를 분류하기 위하여 해당 클립의 해설을 기반으로 감정위드를 추출하기 때문에 분류가 부정확할 수 있다.

감성기반 온라인 쇼핑 웹 에이전트[13]는 개개인의 감성을 고려하여 소비자의 감성요인을 점수화하고 감성점수와 감성요인의 이동경로를 시각적으로 표시함으로써 고객의 구매의사결정을 도와주는 시스템이다. 시스템은 소비자가 선호할 것으로 예상되는 아이템을 감성적 특성을 바탕으로 실시간으로 분석하고 추천한다. 이 시스템은 소비자의 감성을 파악하기 위해 여러 제품을 선택하게 하고, 이를 바탕으로 감성취향을 판단한다.

2.3 소비자 유형 분류 기법

소비자 유형을 분류하는 기법에는 라이프스타일 이론, 성격유형에 따른 구분, 인구통계학적 기법, 소비유형에 따른 구분 등이 있다.

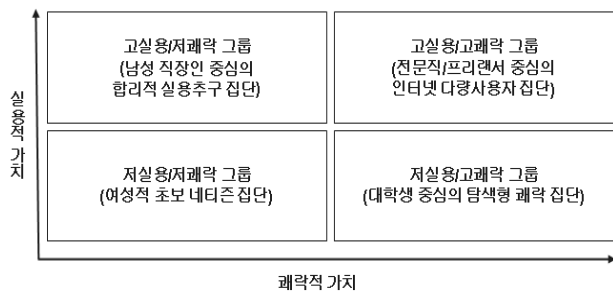
라이프스타일 이론에는 로키치 가치조사, LOV, VALS 등이 있다. 이러한 라이프스타일 이론들은 소비생활과 관련이 없는 인간의 일상생활에 관련된 항목이 너무 많고, 가치 변

수가 직접적으로 콘텐츠에 대한 구매행동을 설명하기에는 무리가 있다[5].

성격유형에 따른 구분은 성격유형검사도구(MBTI, Myers-Briggs Type Indicator)를 이용하여 사람의 경향을 외향형/내향형, 감각형/직관형, 사고형/감정형, 판단형/인식형 등으로 구분하는 것이다[2]. 이 방법은 각 개인별 성격유형검사를 추가적으로 수행해야하는 어려운 점이 있다.

인구통계학적 기법은 같은 연령대, 같은 성별, 같은 학력 등 인구 통계학적 변수로 그룹화하는 기법이다. 이 기법은 협업여과 추천기법의 희소성 문제를 해결할 수도 있으나 [12], 공학적인 추천 기법으로는 사용하기에는 그룹화할 변수의 범위가 모호한 점이 있다[11].

소비유형에 의한 구분은 소비 목적에 따라 소비자를 그룹화하는 것이다. Babin et al., Hammond et al., Childer et al., Kevin et al. 등 많은 연구 결과에서는 소비 목적을 쾌락적 동기와 효용적 동기로 구분하였다. 이렇게 구분된 소비유형이 상품의 구매와 유의한 관계가 있는 것으로 나타났으며, 오프라인뿐만 아니라 온라인 소비자에서 유의한 결과가 나타났다. 국내에서는 농산물 온라인 쇼핑몰 분야에서 쾌락적 동기가 효용적 동기보다 구매감정과 고객충성도에 미치는 영향이 높은 것으로 나타났으며[14], 대형할인점에서도 합리적 동기뿐만 아니라 쾌락적 동기도 구매에 의미있는 영향을 미친다는 결과를 보였다[8]. 또한, (그림 1)과 같이 저실용/저쾌락 그룹은 여성적 초보 네티즌 집합, 저실용/고쾌락 그룹은 대학생 중심의 탐색형 쾌락집단, 고실용/저쾌락 가치집단은 남성 직장인 중심의 합리적 실용추구집단, 그리고 고실용/고쾌락 그룹은 전문직/프리랜서 중심의 인터넷 다량 사용자 집단의 특성이 나타났으며, <표 1>과 같은 인구통계학적 특성을 가진다[5].



(그림 1) 소비유형에 의한 소비자 그룹 분류

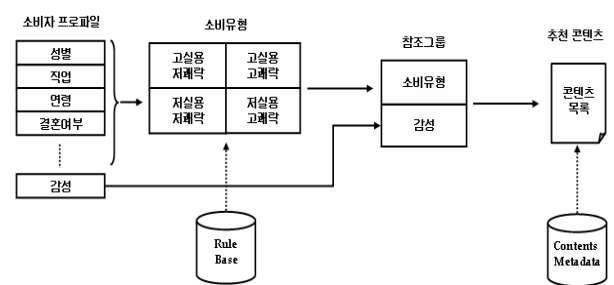
3. 소비자 감성과 소비유형에 기반한 콘텐츠 추천 기법

본 장에서는 개발하고자 하는 콘텐츠 추천 시스템에 대하여 서술한 뒤, 협업여과 추천기법에 기반하면서 소비자의 감성과 소비유형을 이용하는 콘텐츠 추천 기법에 대하여 설명한다.

3.1 콘텐츠 추천 시스템 구성도

협업여과 추천 기법에서 추천의 효과를 높이기 위해서는 참조그룹을 선정하는 것이 중요하다. 참조그룹이란 현재의 소비자와 높은 유사도를 갖는 사람들의 집합을 말한다. 본 논문에서는 참조그룹을 선정함에 있어 두 가지 요소를 고려한다. 한 가지는 소비자의 감성이고, 또 다른 하나는 소비유형이다. 본 논문에서 소비자의 감성은 기쁨, 슬픔, 혐오, 행복, 이완의 다섯 가지로 구분되며, 소비유형은 저실용/저쾌락, 저실용/고쾌락, 고실용/저쾌락, 고실용/고쾌락의 네 가지로 구분된다. 따라서 5개의 감성과 4개의 소비유형[5]을 조합하여 20개의 참조그룹을 선정한다. 콘텐츠 추천 시 한 사람의 소비자는 현재의 감성과 소비유형에 따라 20개의 참조그룹 중 하나의 그룹에 매핑되고, 해당 그룹의 소비자들이 선호하는 콘텐츠를 추천받게 된다.

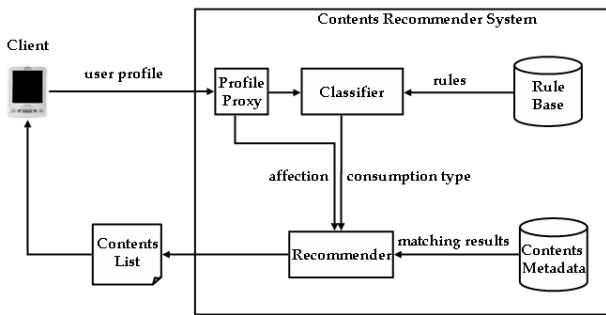
본 논문에서 제안하는 추천 시스템의 개념은 (그림 2)와 같다. 소비자는 감성정보가 포함된 프로파일 정보와 함께 콘텐츠 추천시스템에 접속한다. 추천 시스템은 소비자의 프로파일 정보를 이용하여 소비자의 소비유형을 분류하고, 소비자 감성정보와 소비유형정보를 이용하여 참조그룹을 선정한다. 그 다음 추천 시스템은 소비자의 참조그룹이 선호하는 콘텐츠를 검색하여 순위별로 소비자에게 추천 콘텐츠 목록을 전달한다.



(그림 2) 콘텐츠 추천 시스템 개념도

<표 1> 소비자 소비유형별 인구통계학적 특성

| 구분 | 저실용/저쾌락 그룹 | 저실용/고쾌락 그룹 | 고실용/저쾌락 그룹 | 고실용/고쾌락 그룹 |
|------|------------------|------------|------------|------------|
| 성별 | 여성 중심 | 관계없음 | 남성 중심 | 관계없음 |
| 직업 | 가사/자영업 | 대학생 | 직장인 | 전문직/프리랜서 |
| 연령 | 19세 이하 40세 이상 | 20대 | 30대 | 30대 |
| 결혼여부 | 관계없음 | 미혼 | 관계없음 | 기혼 |



(그림 3) 추천 시스템 구성도

이러한 콘텐츠 추천 시스템의 구성도는 (그림 3)과 같다. 소비자 유형 분류기(Classifier)는 소비자 프로파일을 이용하여 Rule Base의 규칙을 기반으로 소비자의 소비유형을 분류하고, 감정정보와 함께 해당 소비자의 참조그룹을 선정한다. 콘텐츠 추천기(Recommender)는 콘텐츠 메타데이터를 이용하여 해당 소비자의 참조그룹이 선호하는 콘텐츠를 순위화하여 추천 콘텐츠 목록을 작성하여 소비자에게 전달한다.

3.2 소비자 프로파일 구성

소비자 프로파일은 개인에 특화된 서비스를 제공하기 위하여 소비자의 선호도, 단말기의 성능, 소비자가 처한 상황 등에 대한 정보들로 구성된다[7]. 프로파일 정보를 표준화된 형태로 구성하기 위하여 초기에는 CC/PP, UAProf, UPS 등이 제안되었으며, 클라이언트와 서버 간의 요청과 응답 시 헤더정보로 전달된다. 차세대 이동통신 서비스에 적용하기 위하여 감성을 포함한 다양한 상황정보를 프로파일로 표현하기 위한 연구들이 진행되고 있다[3][7]. 이러한 연구들의 소비자 프로파일 구성방법을 바탕으로 콘텐츠 추천에 필요한 프로파일을 구성하면 다음과 같다.

- UserProfile(성별, 직업, 연령, 결혼여부, ..., 감정, 콘텐츠URL)
 성별, 직업, 연령, 결혼여부 등과 같은 인구통계학적 특성 정보는 회원가입 등과 같이 소비자에게 입력을 요구함으로써 획득될 수 있다. 감정정보는 간단하게는 소비자의 입력으로부터 얻어질 수 있으며, BSD(Body Sensor Network)를 통하여 자동으로 획득될 수도 있다. 콘텐츠URL은 브라우저를 통하여 소비자가 특정 콘텐츠를 선택하는 것으로 얻어진다.

3.3 소비자 유형 분류

본 논문에서는 소비자의 소비유형을 판단하기 위하여 다 기준 의사결정(Multi-Criteria Decision Making) 중의 하나인 단순가중치 기법(Simple Weighted Method)을 이용한다. 단순가중치 기법은 의사결정에 필요한 요소의 평가치에 가중치를 곱해 모든 요소에 대한 합을 구하고, 각 대안에 대한 총 점수가 구해지면 가장 큰 점수를 갖는 대안을 선택하는 것이다[6]. 단순가중치 기법으로 소비유형을 판단하기 위하여 (그림 4)와 같이 Profile과 Weight 두 개의 테이블이 이용된다. Profile 테이블은 소비자의 프로파일 항목이 소비

| | G_1 (low/low) | G_2 (low/high) | G_3 (high/low) | G_4 (high/high) | G_1 (low/low) | G_2 (low/high) | G_3 (high/low) | G_4 (high/high) |
|------|--------------------|---------------------|---------------------|----------------------|--------------------|---------------------|---------------------|----------------------|
| 성별 | P_{11} | P_{21} | P_{31} | P_{41} | W_{11} | W_{21} | W_{31} | W_{41} |
| 직업 | P_{12} | P_{22} | P_{32} | P_{42} | W_{12} | W_{22} | W_{32} | W_{42} |
| 나이 | P_{13} | P_{23} | P_{33} | P_{43} | W_{13} | W_{23} | W_{33} | W_{43} |
| : | : | | | | : | | | |
| 결혼여부 | P_{1n} | P_{2n} | P_{3n} | P_{4n} | W_{1n} | W_{2n} | W_{3n} | W_{4n} |
| | Profile | | | | Weight | | | |

(그림 4) 소비자 소비유형 분류를 위한 Profile과 Weight 테이블

자 소비유형에 대하여 해당하는지에 대한 정보이며, Weight 테이블은 프로파일 항목에 대한 소비자 소비유형의 가중치이다. (그림 4)의 Profile과 Weight 테이블의 각 항목 P_{mn} 과 W_{mn} 에서 m은 소비자 소비유형이며, n은 프로파일의 항목의 개수이다.

소비자 프로파일로부터 Profile 테이블을 작성하는 방법으로는 규칙기반(Rule Base)의 형태를 이용한다. 소비자 프로파일이 입력되면, 소비자 유형 분류기(Classifier)는 Rule Base의 규칙들과 비교한다. 소비자 프로파일의 특성이 해당 규칙과 일치하면 Profile 테이블의 해당 위치에 규칙에 정해진 값을 할당한다. 규칙(Rule)들은 다음과 같은 형태로 저장된다.

$$R_{11} : \text{if sex="female" then } V_{11}=1, \text{ else } V_{11}=0$$

$$R_{12} : \text{if job="housewife" then } V_{12}=1, \text{ else } V_{12}=0$$

Weight 테이블은 오프라인에서 경험적으로 계산된다. 본 연구는 추천의 정확도 개선에 있기 때문에 Weight 테이블의 가중치는 정확도를 기반으로 계산한다. 가중치를 계산하는 방법은 각 프로파일 항목 값에 해당하는 소비자 집합을 대상으로 상위 5개의 선호콘텐츠를 선정하고, 각 소비자가 해당 콘텐츠를 선호하는지에 대하여 정확도를 계산한다. 예를 들어, 여성 전체집합을 대상으로 한 상위 5개의 선호콘텐츠 집합이 $\{C_1, C_2, C_3, C_4, C_5\}$ 이고, 여성 소비자 A의 선호콘텐츠 집합이 $\{C_1, C_2, C_3, C_{11}, C_{12}\}$ 일 때, 정확도는 3/5이고 이에 해당하는 W_{ij} 의 가중치는 0.6이 된다. Profile 테이블과 Weight 테이블을 바탕으로 네 가지의 소비자 소비유형에 대한 예측치는 단순가중치 기법에 따라 다음 (수식 1)과 같이 계산된다.

$$G_i = \sum_{j=1}^n P_{ij} \times W_{ij} \text{ for } i=1,2,3,4 \quad (1)$$

여기에서 P_{ij} 는 소비유형에 대한 소비자의 프로파일 항목의 평가치이고, W_{ij} 는 해당 항목에 대한 가중치이다. 또한, G_i 는 소비자가 해당 소비유형에 속할 예측치이다. 소비자의 프로파일이 해당 소비유형에 포함되는 항목이 많고, 해당 항목의 가중치가 높을수록 G_i 는 큰 값을 갖는다. G_i 의 값이 크면 해당 소비유형에 속할 가능성이 크다는 의미이며, 따라서 본 논문에서는 G_i 값이 최대인 i가 해당 소비자의 소비유형을 지시하는 색인(index)이 된다. (그림 5)는 30대의 미혼인 직

| | G ₁ | G ₂ | G ₃ | G ₄ | G ₁ | G ₂ | G ₃ | G ₄ | G ₁ | G ₂ | G ₃ | G ₄ |
|------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|-----------------------------|----------------|----------------|----------------|
| 성별 | 1 | 0.5 | 0 | 0.5 | 0.52 | 0.50 | 0.53 | 0.50 | 1 * 0.52 | 0.5 * 0.5 | 0 * 0.53 | 0.5 * 0.5 |
| 직업 | 0 | 0 | 1 | 0 | 0.57 | 0.55 | 0.54 | 0.40 | 0 * 0.57 | 0 * 0.55 | 1 * 0.54 | 0 * 0.4 |
| 나이 | 0 | 0 | 1 | 1 | 0.65 | 0.52 | 0.48 | 0.48 | 0 * 0.65 | 0 * 0.52 | 1 * 0.48 | 1 * 0.48 |
| 결혼여부 | 0.5 | 1 | 0.5 | 0 | 0.50 | 0.55 | 0.50 | 0.54 | 0.5 * 0.5 | 1 * 0.55 | 0.5 * 0.5 | 0 * 0.54 |
| | Profile | | | | Weight | | | | | | | |
| | | | | | | | | | 0.77 0.8 1.27 0.73 | | | |

(그림 5) 소비유형 판단의 예

장인 여성의 소비유형을 판단하는 예이다. G₃의 값이 1.27로 다른 소비유형의 예측치보다 크기 때문에, 이 경우에는 소비자의 소비유형을 고실용/저쾌락 판단한다.

이러한 방법으로 소비자의 소비유형을 판단하는 User Classification 알고리즘은 다음과 같다.

Algorithm UserClassification

- **Input** : UserProfile(U₁, U₂, U₃, ..., U_n, affection, content)
- **Output** : ShoppingType(1:low-utility/low-pleasure, 2:low-utility/high-pleasure, 3:high-utility/low-pleasure, 4:high-utility/high-pleasure)

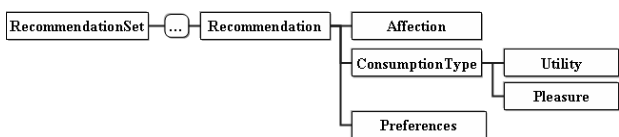
```

1 for (i=1; i<=4; i++) {
2   for (j=1; j<=n; j++) {
3     if (Uj matches Rij) Pij=Vij;
4     else Pij=0;
5     Gi = Gi + (Pij * Wij)
6   }
7 }
8 for (i=1; i<=4; i++) {
9   find (i where Max(Gi))
10 }
11 return i
    
```

3.4 추천을 위한 콘텐츠 메타데이터 설계

소비자의 감성과 소비유형을 포함하는 참조그룹에 대하여 적절한 콘텐츠를 추천하기 위해서는 개별 콘텐츠에 대하여 소비자 참조그룹에 대한 추천등급(선호도)을 메타데이터로 구성해야 한다. (그림 6)은 소비자의 참조그룹에 대하여 콘텐츠 추천을 위한 RecommendationSet Description Scheme이다. 소비자 감성(Affection)과 소비유형(ConsumptionType)에 대한 콘텐츠 선호도(Preferences)를 유지함으로써 소비자의 감성과 소비유형에 따라 추천이 가능하도록 한다. 소비유형은 효율적 동기(Utility)와 쾌락적 동기(Pleasure)에 따라 low/high 값을 갖는다.

(그림 7)은 RecommendationSet을 포함하는 Contents Recommendation Description Scheme의 예를 보인다. 콘텐츠 'Movie1'(4번 라인)은 소비자의 감성이 '행복(happiness)'이고(10번 라인), 소비유형이 '저실용/고쾌락(low-utility/



(그림 6) RecommendationSet Description Scheme

```

1 <?xml version="1.0" encoding="UTF-8"?>
2 <CRDS>
3   <ContentDescription>
4     <ContentName>Movie1</ContentName>
5     <ContentFormat>video</ContentFormat>
6     <Creator>ABC</Creator>
7   </ContentDescription>
8   <RecommendationSet>
9     <Recommendation>
10      <Affection>happiness</Affection>
11      <ConsumptionType>
12        <Utility>low</Utility>
13        <Pleasure>high</Pleasure>
14      </ConsumptionType>
15      <Preferences>10</Preferences>
16    </Recommendation>
17  </RecommendationSet>
18 </CRDS>
    
```

(그림 7) Contents Recommendation Description Scheme 예

high-pleasure)'일 때(11-14번 라인), 선호도(Preferences) '10'으로 추천될 수 있다(15번 라인).

3.5 콘텐츠 추천 알고리즘

콘텐츠 추천 알고리즘에서 먼저 고려되어야 할 사항은 언제 소비자에게 콘텐츠를 추천할 것인가와 콘텐츠에 대한 소비자의 선호정보(Preferences)를 어떻게 최신으로 유지할 것인가이다.

본 연구에서는 소비자 프로파일 항목 중에서 content 항목에 값이 없는 경우를 콘텐츠 추천 시점으로 한다. 소비자 프로파일의 content 항목에 값이 있다는 것은 소비자가 해당 콘텐츠를 요청했다는 것을 의미하며, 이 경우에는 해당 콘텐츠를 전달한다. 반대로 content 항목에 값이 없는 경우에는 소비자가 콘텐츠를 요청하지 않았다는 것을 의미하며, 이 경우에는 추천 콘텐츠 목록을 전달한다.

콘텐츠에 대한 소비자의 선호정보(Preferences)를 최신으로 유지하기 위해서는 추천된 콘텐츠를 소비자가 요청했을 경우 또는 콘텐츠를 소비한 후에 선호정보를 입력했을 경우, 선호정보를 갱신해 주어야 한다. 따라서 콘텐츠 추천 알고리즘에서는 소비자가 추천된 콘텐츠를 요청하였을 경우 또는 선호정보를 입력했을 경우에 선호정보를 갱신한다.

소비자의 프로파일로부터 소비자가 어떠한 소비유형에 속하는지를 분류하고, 소비자의 감성정보를 이용하여 선호도가 높은 콘텐츠를 순위화하여 추천하는 콘텐츠 추천 알고리즘은 다음과 같다.

```

Algorithm ContentsRecommendation
• Input : UserProfile(U1, U2, U3, ..., Un, affection, content)
• Output : ContentsList or a content

1 UserGroup = UserClassification Alg. (UserProfile);
2 if (UserProfile.content is null){
3   select (ContentName from Contents Metadata DB
      where ConsumptionType = UserGroup
      and Affection = UserProfile.affection
      order by Preferences Desc);
4   return (ContentsList);
5 }
6 else {
7   update ( Contents Metadata DB
      set Preferences = Preferences + 1
      where ContentName = UserProfile.content);
8   return (content);
9 }
    
```

ContentsRecommendation 알고리즘에서 소비자가 특정 콘텐츠에 대한 서비스를 요청하였는지 또는 콘텐츠를 추천할 것이지를 확인하는 방법은 소비자 프로파일의 특성 중에서 content 항목의 값을 이용한다. content 항목의 값이 없는 경우(2번 라인)에는 콘텐츠의 추천이 필요한 것이므로 추천 콘텐츠 목록을 전달하고(3-4번 라인), content 항목의 값이 있는 경우(6번 라인)에는 특정 콘텐츠를 요청한 것이므로 해당 콘텐츠의 선호정보를 갱신하고(7번 라인) 해당 콘텐츠를 전달한다(8번 라인).

4. 성능 평가

본 장에서는 3장에서 제안한 콘텐츠 추천 알고리즘의 성능을 평가하기 위하여 두 가지 실험을 수행한다. 먼저, 영화/음악 콘텐츠에 대하여 소비자의 감성과 소비유형에 따른 소비자 선호도 측정 실험을 수행한다. 그 다음 추천 알고리즘의 정확도를 측정하기 위하여 소비자 선호도 측정 실험을 기반으로 소비자의 감성과 소비유형에 따라 콘텐츠를 추천하는 실험을 수행한다.

4.1 소비자 선호도 측정 실험

본 논문에서는 소비자의 선호도 측정 및 추천 실험을 위하여 영화콘텐츠 20개와 음악콘텐츠 35개를 선정하였다. 영화콘텐츠는 드라마, 코미디, 액션, SF, 공포 장르로 구분하고 각 장르별 과거 10년 동안 가장 인기 있었던 4개를 선정하였다. 음악콘텐츠는 발라드, 댄스, 힙합, 락, 트로트, 뉴에이지, 클래식 장르로 구분하고, 각 장르별 5년 단위(최근, 2000년대 전반, 1990년대 전/후반, 1980년대 후반)로 인기순위 1위를 기록했던 음악을 5개를 선정하였다.

선정된 콘텐츠들에 대한 소비자의 선호도를 측정하기 위하여 다음 (그림 8)과 같은 절차를 수행하였다. 소비자의 감성을 유도하기 위하여 국제정서사진체계(International Affective Picture System : IAPS)[15]에서 분류된 사진을 기쁨, 슬픔,



(그림 8) 콘텐츠 선호도 조사 절차

혐오, 행복, 이완 5가지 감성별 60개를 준비하였다. 약 3분에 걸쳐 유도하고자 하는 감성 사진을 보여준 후, 해당 감성이 유도되었는지 질문을 통해 확인하였다. 확인결과 감성 유도가 부족했다면, 다시 사진을 보여줌으로써 해당 감성을 유도하였다. 원하는 감성이 유도되었다면, 위에 선정된 콘텐츠들에 대한 선호도를 측정하였다.

콘텐츠 제목에 대하여 선호여부를 좋음/보통/싫음을 선택하도록 하였다. 선호여부에 따라 ‘좋음’은 +1을 ‘보통’은 0을 ‘싫음’은 -1을 부여하여, 콘텐츠 메타데이터에 소비자의 감성 및 소비유형에 대한 콘텐츠의 선호도(Preferences) 값에 해당 점수를 합하였다.

이 실험에는 150명이 참여하였으며, 총 270회의 실험이 수행되었다. 피실험자는 각기 다른 감성에 대하여 참여하도록 했으므로 1인당 평균 1.8개의 감성에 대하여 참여하였다. 또한, 감성유도의 정확성을 높이기 위하여 선호도 측정 실험은 각 피실험자에 대하여 1일 1회의 실험만을 실시하도록 하였다. 예를 들어, 피실험자 A가 ‘기쁨’ 감성에 대하여 실험에 참여하였다면, 하루 이상의 시간이 흐른 이후에 ‘기쁨’이 아닌 다른 감성에 대한 실험에 참여하도록 하였다. 이 실험을 통해 55개의 영화/음악 콘텐츠에 대하여 감성과 소비유형에 따른 선호도를 측정하였고, Weight 테이블을 경험적으로 획득하였으며, 이를 콘텐츠 추천 실험에 활용하였다.

4.2 콘텐츠 추천 실험 및 결과

추천시스템이 얼마나 효과적인지를 평가하기 위해서는 적절한 평가 척도가 필요하다. 효과적인 성능평가를 위해 다양한 방법들이 연구되어 왔다. 그 중 Precision과 Recall은 가장 대표적으로 사용되고 있는 평가 척도이다. 본 연구는 추천의 정확도 측정 및 개선에 있기 때문에 여기에서는 Precision을 평가척도로 사용한다. Precision의 산출방식은 (수식 2)와 같다.

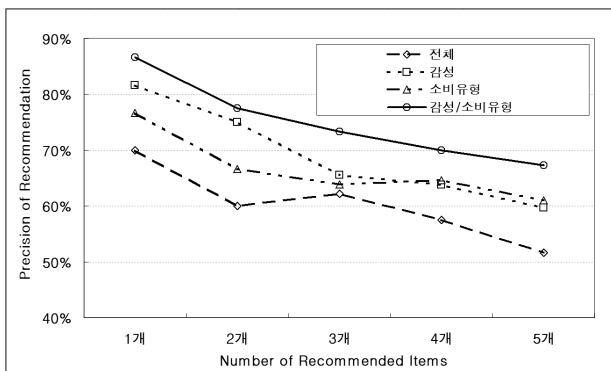
$$Precision = \frac{R^c}{R} \tag{2}$$

여기에서 R은 추천시스템에 의해 추천된 콘텐츠의 수를 나타내며, R^c는 추천된 콘텐츠 중 사용자가 실제로 선호한다고 응답한 콘텐츠의 수를 나타낸다. 따라서 Precision은 추천시스템에 의해 추천된 콘텐츠 중 피실험자가 실제로 선호하는 콘텐츠의 비율로서 추천의 정확도를 의미한다. 본 연구에서는 추천되는 콘텐츠의 수 R을 15개까지로 고정시키고 Precision을 측정하였다.

본 실험에서는 30명의 피실험자를 선발하였고, 2회에 걸쳐 다른 감성에 대하여 실험이 수행되었으며, 1인당 15개

<표 2> 피실험자에 30명에 대한 실험 결과(2회 Precision 평균)

| 피실험자 | 전체 | 감성 | 소비유형 | 감성/소비유형 |
|------|-------|-------|------|---------|
| 1 | 79.8 | 52.5 | 79.2 | 50.5 |
| 2 | 83.2 | 62.8 | 85.2 | 62.8 |
| 3 | 43.5 | 72.8 | 60.8 | 80.7 |
| 4 | 48.0 | 80.3 | 48.0 | 80.3 |
| 5 | 25.7 | 39.2 | 25.7 | 32.7 |
| 6 | 53.3 | 67.3 | 75.2 | 100.0 |
| 7 | 73.2 | 80.7 | 77.2 | 80.7 |
| 8 | 93.5 | 79.8 | 98.0 | 87.7 |
| 9 | 79.3 | 70.7 | 82.7 | 72.7 |
| 10 | 63.3 | 68.3 | 54.7 | 66.3 |
| 11 | 27.7 | 53.3 | 47.0 | 100.0 |
| 12 | 50.0 | 48.0 | 50.0 | 50.0 |
| 13 | 61.3 | 83.7 | 54.7 | 77.8 |
| 14 | 73.2 | 72.8 | 75.2 | 83.2 |
| 15 | 53.3 | 90.2 | 93.5 | 100.0 |
| 16 | 55.5 | 31.8 | 52.2 | 37.2 |
| 17 | 61.3 | 57.5 | 77.3 | 64.8 |
| 18 | 78.7 | 82.3 | 79.3 | 85.7 |
| 19 | 70.3 | 71.3 | 74.3 | 87.0 |
| 20 | 45.7 | 54.7 | 49.7 | 61.3 |
| 21 | 34.7 | 74.3 | 45.3 | 74.3 |
| 22 | 12.8 | 40.2 | 77.8 | 41.7 |
| 23 | 20.7 | 57.8 | 16.8 | 92.2 |
| 24 | 61.3 | 83.7 | 54.7 | 77.8 |
| 25 | 50.0 | 77.3 | 48.0 | 77.3 |
| 26 | 96.0 | 100.0 | 96.0 | 100.0 |
| 27 | 69.8 | 77.8 | 75.3 | 72.8 |
| 28 | 100.0 | 95.5 | 98.0 | 100.0 |
| 29 | 72.3 | 80.7 | 74.3 | 80.7 |
| 30 | 70.8 | 66.3 | 70.8 | 70.8 |
| 평균 | 60.3 | 69.1 | 66.6 | 75.0 |



(그림 9) 피실험자 30명에 대한 콘텐츠 추천 정확도(2회 평균)

<표 3> 추천기법의 Paired Sample t-Test 분석결과(95% 신뢰수준)

| 구분 | 전체 - 감성 | 전체 - 소비유형 | 전체 - 감성/소비유형 |
|---------|---------|-----------|--------------|
| t-value | 2.730 | 2.381 | 3.444 |
| p-value | 0.011 | 0.024 | 0.002 |

의 콘텐츠를 추천한 후 결과를 추출하였다. <표 2>는 피실험자별 추천의 정확도 결과를 나타내며, (그림 9)는 추천 콘

텐츠의 개수에 따른 추천 정확도의 평균을 보여준다.

추천의 정확도에서 소비자 전체를 참조그룹으로 한 추천 보다는 소비자의 감성 또는 소비유형을 참조그룹으로 한 추천 결과가 우수함을 보인다. 또한 소비자의 감성과 소비유형을 함께 참조그룹으로 한 추천 결과가 가장 우수함을 보인다. 소비자의 감성만을 참조그룹으로 하는 것과 소비자의 소비유형만을 참조그룹으로 하는 것을 비교해 보면, 두 기법간의 우수함을 판단하기가 어려워진다. 따라서 콘텐츠를 추천함에 있어서 소비자의 감성과 소비유형을 함께 고려하는 것이 바람직하다. 소비자의 감성과 소비유형을 참조그룹으로 하여 콘텐츠를 추천하였을 때, 콘텐츠 추천의 정확도는 소비자 전체를 참조그룹으로 추천한 경우보다 평균 14.7%, 소비자의 감성만을 참조그룹으로 추천한 경우보다 평균 8.4%, 소비자의 소비유형만을 참조그룹으로 추천한 경우보다 평균 5.8%의 정확도가 향상됨을 보인다.

본 연구에서 제안한 감성과 소비유형을 고려한 추천기법이 통계적인 유의성을 가지는지 확인하기 위해 Paired Sample t-Test 분석을 수행하였다. <표 3>은 소비자 전체를 참조그룹으로 한 추천기법과의 Paired Sample t-Test의 결과이다. 감성, 소비유형 및 감성/소비유형을 고려한 경우 모두 95% 신뢰수준에서 유의한 것으로 나타났으며, 이때의 t-value는 각각 2.73, 2.381, 3.444로 나타났다.

5. 결론 및 향후 연구방향

본 논문에서는 콘텐츠 추천의 정확도를 높이기 위하여 소비자의 감성과 소비유형을 참조그룹으로 하여 협업여과기반으로 콘텐츠를 추천하는 기법을 제안하였다. 실험결과 소비자의 감성과 소비유형을 참조그룹으로 추천한 경우 소비자 전체를 참조그룹으로 추천한 경우보다 14.7%, 소비자의 감성만을 참조그룹으로 추천한 경우보다 8.4%, 소비자의 소비유형만을 참조그룹으로 추천한 경우보다 5.8%의 정확도가 향상됨을 보였다. 향후에는 가변적인 상황에 대처 가능하도록 감성과 소비유형을 세분화하고, 추천의 정확도를 더욱 향상시킬 수 있는 소비자의 참조그룹에 대하여 연구할 것이다. 또한, 감성의 변화에 대하여 콘텐츠 선호도가 어떻게 변화하는지에 대하여 연구할 예정이다.

참고 문헌

- [1] 고수정, 김진수, 김태용, 최준혁, 이정현, “협력적 여과와 내용 기반 여과의 병합을 통한 추천 시스템에서의 사용자 선호도 발견,” 정보과학회논문지 제7권 제6호, 2001. 12.
- [2] 고정형, 이종호, “성격유형을 고려한 전자상거래 소비자 구매의도 영향 연구,” 한국경영교육학회 경영교육논총 43집, pp.1-19, 2006. 08.
- [3] 김경식, 이재동, “상황에 적응화된 콘텐츠 제공을 위한 클라이언트 프로파일 프레임워크,” 정보처리학회논문지C, 제 14-C권 제3호, pp.293-304, 2007. 06.

[4] 박면용, 안승민, 하성도, 정도인, 류인균, “감정 및 정서상태 전이를 위한 감성 콘텐츠 추천 시스템 개발,” 한국감성과학회 제10권 제1호, pp.1-11, 2007. 03.

[5] 박철, “인터넷탐색 가치에 의해 분류한 온라인 소비자 집단별 특성에 관한 연구,” 소비자학 연구 제12권 제1호, pp.31-52, 2001. 03.

[6] 안춘수, 강태진, 정상윤, 홍성일 “수정된 다기준 의사결정을 이용한 흐름방식에서의 작업순서 결정,” 공업경영학회지 제20권 제44호, pp.135-151, 1997. 11.

[7] 유명식, 오돈성, “차세대 이동 통신 서비스 지원을 위한 프로파일 관리 기술 동향,” 한국통신학회지 제22권 제9호, pp.77-89, 2005. 09.

[8] 윤미현, 김영신, “대형할인점 이용 소비자의 쇼핑동기에 따른 쇼핑결과,” 소비문화연구 제10권 제2호, pp.171-195, 2007. 06.

[9] 이구형, “감성공학의 개념과 연구 및 응용 방법,” 대한인간공학회지 17권, 1998.

[10] 이병욱, 송희석, 강신철, 남수현, 최준연, “유비쿼터스 환경에서의 상황인식 기반 VOD 콘텐츠 추천시스템,” Telecommunications Review 제17권 제6호, 2007. 12.

[11] 이수정, 이형동, 김형주, “사용자 경향에 기반한 동적 추천 기법: 영화 추천 시스템을 중심으로,” 정보과학회논문지 제31권 제2호, pp.153-163, 2004. 02.

[12] 이용준, 이세훈, 왕창중, “인구 통계 정보를 이용한 협업여과 추천의 유사도 개선 기법,” 정보과학회논문지: 컴퓨터의 실제 제9권 제5호, pp.521-529, 2003. 10.

[13] 임치환, “온라인 쇼핑몰에서 고객의 감성을 활용한 추천 효과,” 산업경영시스템학회지 제28권 제3호, pp.58-64, 2005. 09.

[14] 황덕란, 최영찬, 김미영, 문정훈, “온라인 쇼핑몰에서 소비자의 쇼핑 동기와 감정이 쇼핑행동에 미치는 영향,” 한국산업경영학회 경영연구 제22권 제2호, pp.347-375, 2007. 05.

[15] Bradley, M. M. and Lang, P. J., “Handbook of Psychophysiology (2rd Edition),” New York: Cambridge University Press, pp.581-607, 2006. 08.

[16] Breese, J., Heckerman, D. and Kadie, C., “Empirical Analysis of Prediction Algorithms for Collaborative Filtering” Proc. of the 14th Conference on Uncertainty in Artificial Intelligence, pp.43-52, 1998.

[17] Cuddy, S., Katchabaw, M. and Lutfiyya, H., “Context-aware Service Selection based on Dynamic and Static Service Attributes,” Wireless and Mobile Computing, Networking and Communications, IEEE International Conference, 2005.

[18] Guang-Zhong Yang, “Body Sensor Network”Springer, 2006.



최 인 복

e-mail : pluto612@dku.edu
 1999년 단국대학교 전자계산학과(학사)
 2002년 단국대학교 대학원 전자계산학과(석사)
 2002년~현재 단국대학교 대학원 정보컴퓨터학과 박사과정

관심분야: 유비쿼터스 컴퓨팅, 콘텐츠 적응화, (모바일)인터넷 기술, 분산/병렬 처리



박 태 근

e-mail : tkpark@dku.edu
 1991년 포항공과대학교 컴퓨터공학과(학사)
 1993년 포항공과대학교 컴퓨터공학과(석사)
 2004년 포항공과대학교 컴퓨터공학과

(박사)

1996년~2000년 SK Telecom 중앙연구원 선임연구원
 2000년~2001년 3Com Korea 과장
 2001년~2002년 Ericsson Korea 차장
 2004년~현재 단국대학교 컴퓨터학부 멀티미디어공학전공 조교수

관심분야: 이동 통신, QoS, 센서 네트워크, 콘텐츠 유통/서비스, 멀티미디어 통신망



이 재 동

e-mail : letsdoit@dankook.ac.kr
 1985년 인하대학교 전자계산학(학사)
 1991년 Cleveland State University(석사)
 1996년 Kent State University(박사)
 1997년 3월~현재 단국대학교 자연과학대학 정보컴퓨터학부 컴퓨터과학전공 부교수

2006년 4월~현재 단국대학교 문화콘텐츠기술연구소 소장
 2004년 7월~2006년 6월 단국대학교 정보통신원 원장(C.I.O)
 2002년 11월~현재 농협중앙회 진산고문
 2006년 7월~현재 민관확대 콘텐츠 정책 협의회 위원
 2007년 2월~현재 Dream economy leader 포럼 위원
 2005년 1월~2006년 12월 전국대학정보화 협의회 이사
 2005년 8월~2006년 8월 문화관광부 KOCCA CT포럼/전략기획 운영위원/분과위원장
 2004년 1월~현재 (사)이러닝 산업협회 이사
 2005년 1월~2006년 12월 교통안전공단 자문위원

관심분야: Ubiquitous Computing, Contents Technologies, (Mobile) Internet Technologies/Applications, GIS Technologies and Applications, Many aspects of parallel/distributed processing