

소셜네트워크에서 분위기 벡터를 이용한 멀티미디어 콘텐츠 추천 방법⁺

(Multimedia Contents Recommendation Method using Mood Vector in
Social Networks)

문창배¹⁾, 이종열²⁾, 김병만^{3)*}

(Chang Bae Moon, Jong Yeol Lee, and Byeong Man Kim)

요약 웹에서 정보 구매자들의 성향은 가성비에서 가심비 형태로 변해가는 추세이다. 멀티미디어 콘텐츠 추천에도 그러한 흐름이 있는데, 바로 폭소노미 (Folksonomy) 기반의 분위기를 이용한 추천 방법이다. 하지만 이런 방법의 경우 동의어를 고려하지 못한다는 문제점이 존재한다. 이 문제를 해결하기 위해 일부 연구에서는 Thayer모델의 12 분위기를 AV(Arousal and Valence)값으로 정의하여 그 문제점을 해결하였지만, 추천 성능이 재현 수준 0.1에서 키워드 기반 검색 방법보다 떨어지는 문제점을 보였다. 본 논문에서는 재현 수준 0.1에서도 키워드 기반 검색 방법과 동일한 추천 성능을 유지하면서 동의어 문제를 해결할 수 있도록 멀티미디어 콘텐츠의 분위기 벡터를 이용하는 방법을 제안하였다. 또한, 추천 성능 분석을 위해 기존 AV값 기반 방법과 키워드 기반 방법과 비교 분석하였다. 추천 성능 분석결과, 본 논문에서 제안한 방법이 전체적으로 기존 방법들 보다 우수한 추천 성능을 보였다.

핵심주제어: 멀티미디어 콘텐츠 분위기, 가심비, 12 분위기 벡터, 멀티미디어 콘텐츠 추천, 분위기 태그, 소셜 네트워크

Abstract The tendency of buyers of web information is changing from the cost-effectiveness to the cost-satisfaction. There is such tendency in the recommendation of multimedia contents, some of which are folksonomy-based recommendation services using mood. However, there is a problem that they does not consider synonyms. In order to solve this problem, some studies have solved the problem by defining 12 moods of Thayer model as AV values (Arousal and Valence), but the recommendation performance is lower than that of a keyword-based method at the recall level 0.1. In this paper, we propose a method based on using mood vector of multimedia contents. The method can solve the synonym problem while maintaining the same performance as the keyword-based method even at the recall level 0.1. Also, for performance analysis, we compare the proposed method with an existing method based on AV value and a keyword-based method. The result shows that the proposed method outperform the existing methods.

Keywords: Multimedia Contents Mood, Cost-Satisfaction, 12 Mood Vector, Music Recommendation, Mood Tag, Social Networks

* Corresponding Author: bmkim@kumoh.ac.kr

+ 이 연구는 금오공과대학교 학술연구비로 지원되었음(2019-104-024)

Manuscript received August 20, 2019 / revised October 07, 2019 / accepted October 08, 2019

1) 금오공과대학교 ICT융합특성화연구센터, 제1저자

2) 금오공과대학교 컴퓨터소프트웨어공학과, 제2저자

3) 금오공과대학교 컴퓨터소프트웨어공학과, 교신저자

1. 서론

웹 정보 구매자들의 성향은 콘텐츠에 대한 가격 대비 성능을 중요시하는 가성비 중심에서 가심비에 마음심을 더하여 심리적인 만족감까지 중시하

는 가심비 형태로 변해가고 있다 (Moon, et al., 2018). 멀티미디어 콘텐츠 추천에도 그러한 흐름이 있는데, 바로 분위기를 이용한 추천 방법이 그런 예이다. 분위기를 이용한 멀티미디어 콘텐츠 추천 방법은 분위기를 이용하여 멀티미디어 콘텐츠를 추천함으로써 가심비와 관련된 연구라 할 수 있을 것이다. 멀티미디어 콘텐츠 추천과 관련된 연구들로 Russel (1980), Hevner (1936), Thayer (1990) 및 Moon et al. (2013, 2014, 2015)의 연구들이 있는데, Russel (1980), Hevner (1936) 및 Thayer (1990)의 연구들에서는 분위기 벡터 정의하는 방법과 관련된 연구이고, Moon et al. (2013, 2019)의 연구에서는 분위기를 이용하여 멀티미디어 콘텐츠를 추천하는 방법이며, Moon et al. (2014, 2015)의 연구들에서는 분위기를 이용하여 멀티미디어 콘텐츠의 분위기 색상을 제공하는 방법이다.

SNS 서비스에서 멀티미디어 콘텐츠 추천 형태를 보면 4가지로 요약할 수 있고, 추천형태는 기존 정보검색 질의 방법 (Query-by-text), 멀티미디어 콘텐츠 일부분에 의한 검색 방법 (Query-by-part), 예제 멀티미디어 콘텐츠 전체에 의한 검색 방법 (Query-by-example) 그리고 분류 기반 검색 방법 (Query-by-class)이다 (Moon et al., 2019). 기존 정보검색 질의 방법은 멀티미디어 콘텐츠 데이터베이스에 저장한 서지 정보(제작자, 콘텐츠 명, 장르, 분위기 등)를 이용하여 기존 정보검색 시스템에서 질의 (Query)를 처리한다. 멀티미디어 콘텐츠 일부분에 의한 질의 방법은 사용자가 자신 단말기에 특정 멀티미디어 콘텐츠의 유무를 알고 싶을 경우, 멀티미디어 콘텐츠의 일부를 입력으로 유사한 멀티미디어 콘텐츠들을 찾아주는 방법이다. 또한, 멀티미디어 콘텐츠 전체에 의한 질의에서는 실제 멀티미디어 콘텐츠 전체를 입력 한다. 분류기반 질의 방법은 사전에 멀티미디어 콘텐츠를 텍소노미(Taxonomy) 방법에 의하여 분류하는 방식이다(Moon et al., 2019).

4가지 검색 방법 중 멀티미디어 콘텐츠 일부분에 의한 질의, 예제 멀티미디어 콘텐츠 전체에 의한 질의 방법은 특수한 상황에서 사용 가능한 방법이라 할 수 있을 것이고, 기존 정보검색 질의 또는 분류기반 질의 형태가 일반적인

방법이라 할 수 있을 것이다. 하지만 이 방법들의 경우 전문가나 운영자의 개입이 필요하다. 즉, 새로운 멀티미디어 콘텐츠가 나왔을 경우, 서지정보를 입력하거나 텍소노미에 따라 클래스(소속)를 결정해야 한다. 요즘처럼 새로운 멀티미디어 콘텐츠가 지속적으로 방대하게 발생하는 상황에서는 이러한 방식은 많은 문제를 발생시킬 수 있다. 이를 해결하기 위한 방법 중 한 가지는 텍소노미에 의한 자동태깅 방법일 것이다. 즉, 멀티미디어 콘텐츠를 자동으로 분류하여 서지정보 또는 분류코드를 자동으로 입력하는 것이다. 하지만 텍소노미 기반 분류방법은 사이트를 관리하는 일부 특정계층(사서, 운영자등)이 직접 분류하는 방법으로 특정 체계의 지식이 요구되며, 새로운 아이템 등이 추가될 경우 확장성에 대한 문제가 발생할 수 있다. 또한 하나의 범주에 하나의 의미를 부여함으로써 하나 이상의 주제를 가진 아이템에 대한 해결책이 필요하다.

텍소노미의 대안으로 폭소노미(Folksonomy)가 존재한다. 이는 텍소노미의 계층적인 방식과는 다르게 평형적인 구조로 전문가에 의한 분리체계 방식이 아닌 참여자에 의한 분리체계라 할 수 있다. 폭소노미의 대표적인 사례가 SNS서비스인 인스타그램, 페이스북, Last.fm, IPTV, 유튜브 등이라 할 수 있다. 폭소노미의 장점은 텍소노미의 확장성에 대한 문제와 배타적인 문제점 등을 보완할 수 있다. 하지만, 폭소노미는 폭소노미 태그의 유사어에 대한 문제점이 존재한다. 이를 해결하기 위해 Moon et al. (2019)의 연구에서는 AV (Arousal and Valence) 기반에서 유사어를 고려한 추천방법을 제안하였다. 하지만 재현 수준 0.1에서 키워드 기반 검색 방법보다 상대적으로 정확률(Precision)이 좋지 않았다. 다시 말해서, 관련성이 결여된 멀티미디어를 먼저 추천하는 경우가 발생 가능하였다. 또한 AV 기반 검색의 경우 분위기 벡터를 12차원에서 2차원으로 축소함으로써 분위기에 대한 정보가 일부 손실됨으로 재현 수준 0.1에서 키워드 기반 검색방법보다 성능이 저하될 수 있다.

본 논문에서는 AV 기반 검색방법을 확장하여 폭소노미의 동의어 문제를 해결하면서 재현 수준 0.1에서도 키워드 기반 추천 방법과 유사한

정확률을 얻을 수 있는 방법을 제안하였다. 이를 위해 멀티미디어 콘텐츠 분위기를 Thayer (1990) 모델의 12 분위기를 이용하여 12차원의 분위기 벡터로 표현하는 방법을 제안하고, 이를 활용하여 멀티미디어 콘텐츠의 내부 태그로 활용하는 방법을 사용하였다.

또한, 본 논문에서는 동의어를 고려하여 12 분위기 벡터와 멀티미디어 콘텐츠의 분위기 태그와의 관계를 정의하는 방법을 제안하였고, 이를 이용한 SNS서비스 기반의 멀티미디어 콘텐츠 추천 방법을 제안하였다. 제안방법의 객관적인 정확률 비교를 목적으로 기존 방법인 AV값을 이용하는 방법 및 키워드 기반 추천 방법 (last.fm의 추천 방법)과의 정확률 비교도 하였다. 본 논문에서 제안한 멀티미디어 콘텐츠 추천방법의 경우 태그가 포함된 모든 멀티미디어 콘텐츠에 적용가능 하지만, 제안한 방법과의 기존방법과의 정확률 분석을 목적으로 기존연구에서 진행한 음악 콘텐츠를 대상으로 우선 적용하였다.

2. 관련연구

분위기 정의 방법에는 Russel (1980) 모델을 이용하는 방법, Hevner (1936) 모델을 이용하는 방법 그리고 Thayer (1990) 모델을 이용하는 방법 등이 존재한다. Russell (1980)의 모델과 Hevner (1936) 모델의 경우 형용사 기반으로 구성된 모델로 각 모델에 포함된 단어들의 경우 의미가 중첩되거나 형용사적 표현이 모호하다는 단점이 있다. 이를 보완한 모델이 Thayer (1990)의 2차원 모델로, 이 모델에서는 12개의 단어(형용사)를 사용한다. Thayer (1990)의 2차원 분위기 모델에서는 멀티미디어 콘텐츠의 분위기를 Arousal과 Valence로 구성된 벡터로 표현되며, 모델에서 Arousal은 청취자가 멀티미디어 콘텐츠에서 느끼는 자극의 강도를 의미하고, Valence는 멀티미디어 콘텐츠의 안정감을 의미하며(Fig. 1 참고), 이를 이용한 멀티미디어 콘텐츠 추천 방법에는 Moon et al. (2013, 2014, 2015)의 연구들이 존재한다. 하지만 폭소노미를 활용한 소셜네트워크와 같은 환경에서는 12개

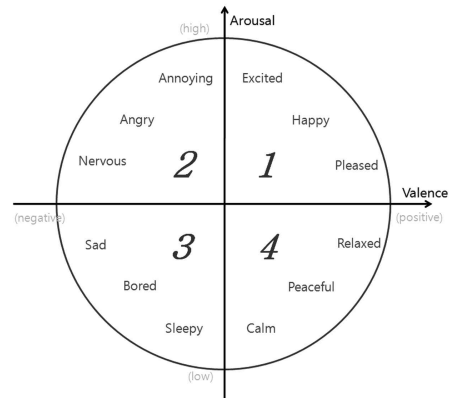


Fig. 1 Thayer's Two-Dimensional Model

분위기 이외의 동의어 및 신조어 등을 고려하는 것이 멀티미디어 콘텐츠 추천 방법에서 중요하다고 할 수 있다.

폭소노미 기반에서 동의어 및 태깅레벨, 신조어등과 관련된 문제를 해결하기 위한 연구에는 Moon et al. (2013)의 연구가 존재한다. 이 연구에서는 음악을 대상으로 분위기 벡터인 Arousal과 Valence를 음악의 내부 태그로 적용하는 방법을 제안하였다. 즉, 태그와 음악모두 AV 값으로 적용함으로써 동의어 태그를 포함하는 음악들 역시 추천이 가능하도록 하였다. 하지만 이 연구에서는 특정 지역을 대상으로 연구를 진행하였기 때문에 특정지역 기반 소량의 데이터를 이용한 추천방법이라 할 수 있고, 본 논문에서는 수년에 걸쳐 구축된 전 세계적인 소셜기반 분위기 데이터를 이용하여 멀티미디어 콘텐츠의 분위기를 정의하고, 이를 이용한 멀티미디어 콘텐츠 추천 방법을 제안한다. 즉, 기존방법의 경우 특정지역 기반 소량의 정보를 이용한 멀티미디어 콘텐츠 추천방법이라 할 수 있고, 본 논문의 경우 전 세계적 빅데이터에 근접한 정보를 이용한 멀티미디어 콘텐츠 추천 방법이라 할 수 있다. 또한 기존 연구의 경우 음악에 내포된 특징을 사용하기 때문에 이미지, 동영상 등과 같이 성향이 다른 멀티미디어 콘텐츠에 확대 적용이 힘들다는 단점이 존재하지만 본 논문에서 제안한 방법의 경우 멀티미디어 콘텐츠의 태그만을 사용하기 때문에 인스타그램, 페이스북, Last.fm, IPTV, 유튜브 등과 같이 태그가 포함된 모든 웹 서비스에 확대적용

이 가능하다는 장점이 있다.

음악 콘텐츠의 폭소노미 태그를 이용한 연구에는 Laurier et al. (2009), Ness et al. (2019) 및 Kim et al. (2011)의 연구가 존재하는데, 이 연구들의 경우 폭소노미 사이트인 last.fm에 존재하는 음악 콘텐츠의 분위기 태그를 활용하여 음악 분위기를 정의하였고, 음악의 분위기를 SVM(Support Vector Machine)을 이용하여 예측하는 연구를 진행하였다. 이 연구들의 경우 음악 콘텐츠 내포된 특징을 이용하여 하나의 분위기를 예측하는 방법이지만 본 논문에서 제안한 방법의 경우 멀티미디어 콘텐츠에 포함된 태그를 이용하여 멀티미디어 콘텐츠를 추천하는 방법이다.

멀티미디어 콘텐츠를 대상으로 자동 태깅과 관련된 연구들 (Chang et al., 2003; Ji et al., 2007; Yang et al., 2007a; 2007b; Yang and Ro, 2007; Tso-Sutter et al., 2008; Vojnovic et al., 2009)이 존재하는데 Ji et al. (2007), Tso-Sutter et al. (2008) 및 Vojnovic et al. (2009)의 연구들에서는 웹 콘텐츠를 이용한 연구라 할 수 있고, Ji et al. (2007), Tso-Sutter et al. (2008) 및 Vojnovic et al. (2009)의 연구들에서는 이미지 콘텐츠를 이용한 연구라 할 수 있다. Ji et al. (2007), Tso-Sutter et al. (2008) 및 Vojnovic et al. (2009)의 연구들에서는 사용자가 어떤 웹 콘텐츠에 태그를 입력하고자 할 때, 시스템에서 해당 웹 콘텐츠의 태그를 사용자에게 추천함으로써 사용자에게 태그 입력과정을 도와주는 방법이라 할 수 있고, Chang et al. (2003), Yang et al. (2007a, 2007b), Yang and Ro (2007), Li et al. (2008), 및 Lindstaedt et al. (2009)의 연구는 Laurier et al. (2009), Ness et al. (2019) 및 Kim et al. (2011)의 연구와 유사하게 이미지 콘텐츠에 내포된 특징을 SVM과 Bayes Point Machines 등에 적용하여 이미지 콘텐츠에 자동으로 태그를 추가하는 방법이라 할 수 있다. 하지만 기존 연구들의 경우 태그를 하나의 클래스로 정의하기 때문에 동의어와 신조어에 대한 문제를 고려하기가 힘들고, 이 연구들 역시 콘텐츠에 내포된 특징을 사용하기 때문에 특정 분야의 콘텐츠만을 추천하기 위한 방법이라 할 수 있다. 즉, 기존 방법의 경우 다양한 멀티미디어 콘텐츠

에 직접적인 적용이 힘들다는 단점이 존재한다. 하지만 본 논문에서 제안하는 방법의 경우 멀티미디어 콘텐츠의 태그만을 이용하여 해당 콘텐츠의 분위기를 수치적으로 표현하고, 수치적으로 표현된 분위기를 이용하여 멀티미디어 콘텐츠를 추천해줌으로 다양한 멀티미디어 콘텐츠 추천방법에 적용이 가능하고, 분위기에 따른 동의어를 고려하기 때문에 동의어에 대한 문제를 일부 해결할 수 있는 방법이라 할 수 있다.

동의어에 대한 문제를 해결하기 위한 AV기반 추천방법 중 Moon et al. (2019)의 연구에서는 수년에 걸쳐 구축된 전 세계적 소셜기반 분위기 데이터를 이용하여 음악의 태그를 대상으로 분위기 벡터 값(Arousal과 Valence의 강도)을 멀티미디어 콘텐츠의 내부 태그로 활용하고, 태그의 AV값을 정의함으로 해당태그의 분위기를 이용하여 동의어 기반 멀티미디어 콘텐츠 추천이 가능하도록 하였다. 하지만 이 연구의 경우 멀티미디어 콘텐츠의 분위기를 AV값으로 정의하였기 때문에 다양한 분위기를 포함하는 멀티미디어 콘텐츠를 하나의 분위기로 매핑 시키는 결과를 발생할 수 있고, 재현 수준 0.1에서도 키워드 기반 추천방법 보다 좋지 않은 추천 결과를 보였다 (재현 수준 0.1과 0.2에서 1,000개 이하의 멀티미디어 콘텐츠의 정확률이 키워드기반 추천방법의 정확률 보다 떨어짐). 본 논문에서는 다양한 분위기로 매핑이 가능하고, 재현 수준 0.1에서도 키워드 기반 추천방법과 동일한 정확률을 보이는 분위기 벡터 기반 멀티미디어 추천방법을 제안하였다. 또한 제안한 방법을 Moon et al. (2019)의 AV 기반 추천방법 및 키워드 기반 추천방법과 비교분석 함으로 제안한 방법이 기존 방법보다 우수함을 보였다.

3. 분위기 태그의 동의어를 고려한 멀티미디어 콘텐츠 추천 방법

분위기 태그기반 동의어를 고려한 멀티미디어 콘텐츠 추천 방법은 네 개의 과정(Fig. 2 참고)으로 구성되며, 그 첫 번째가 멀티미디어 콘텐츠 정보 구축이고, 두 번째가 멀티미디어 콘텐

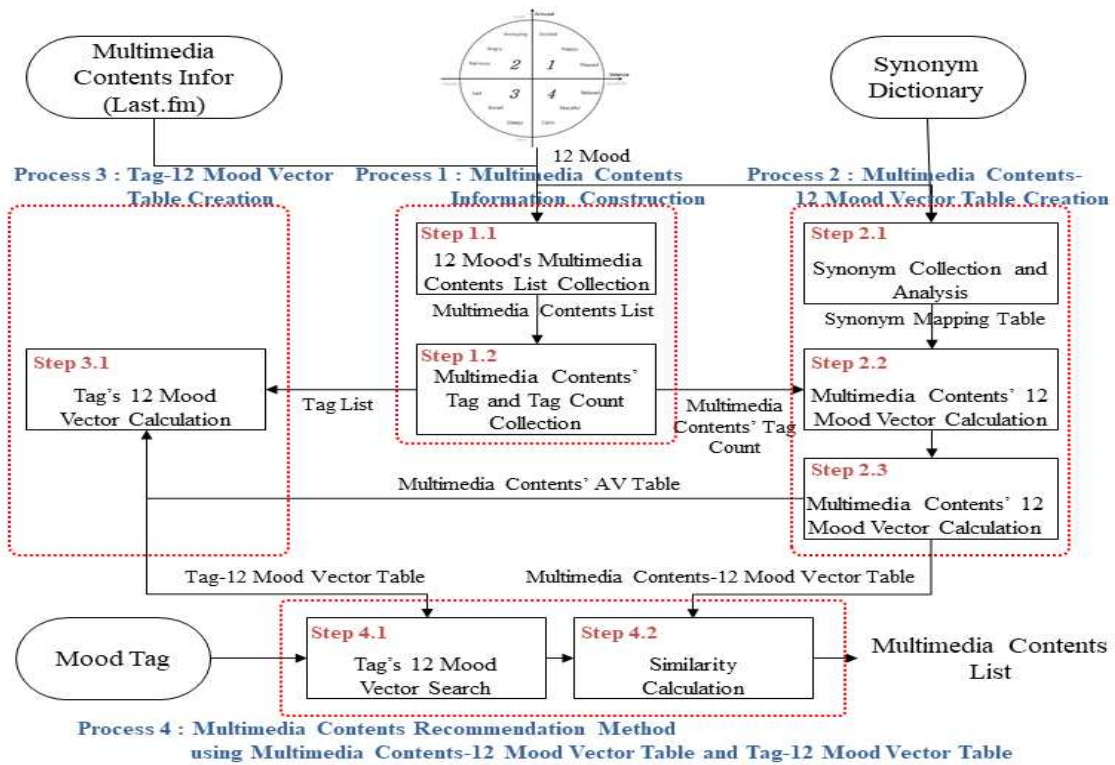


Fig. 2 Multimedia Contents Recommendation Structure

츠의 12 분위기 벡터 테이블 생성이며, 세 번째가 태그의 12 분위기 벡터 테이블 생성이고, 마지막이 멀티미디어 콘텐츠 추천이다.

3.1 멀티미디어 콘텐츠 정보 구축과정

본 과정은 크게 두 단계(Fig. 3 참고)로, 그 첫 번째는 12 분위기의 멀티미디어 콘텐츠 리스트를 수집하는 단계이고, 두 번째는 멀티미디어 콘텐츠의 태그와 태그의 카운트를 수집하는 단계이다. 본 논문에서는 모든 유형의 멀티미디어 콘텐츠에 적용하기에 앞서 분위기와 밀접한 관련이 있는 음악을 대상으로 연구를 진행하였고, 한 지역에 국한된 데이터가 아니라 전 세계적으로 수집된 데이터에 멀티미디어 콘텐츠 추천방법을 적용하기 위해 Moon et al. (2019)의 방법과 동일한 방법으로 데이터를 구축하였다.

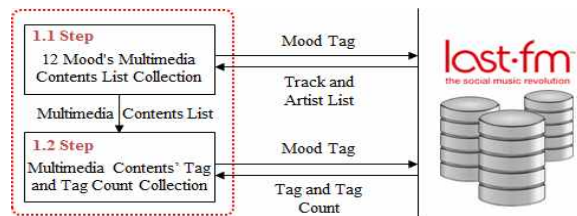


Fig. 3 Multimedia Contents Information Construction Process

3.2 멀티미디어 콘텐츠-12 분위기 벡터 테이블 생성과정

이 과정은 멀티미디어 콘텐츠의 분위기 태그를 이용하여 멀티미디어 콘텐츠 추천에 필요한 12 분위기에 해당하는 분위기 벡터를 계산하는 과정으로, 최종적으로 멀티미디어 콘텐츠의 12 분위기 벡터 테이블이 구축된다. 본 과정은 총 두 단계로 구성되며, 그 첫 번째는 동의어 수집 및 분석 단계이고, 두 번째가 멀티미디어 콘텐츠의 12 분위기 벡터 계산 단계이다.

3.2.1 동의어 수집 및 분석

태그의 동의어 분석을 위해 Moon et al. (2018)의 연구에서 진행한 동의어 분석 결과를 활용하여 멀티미디어 콘텐츠의 태그를 12 분위기로 분류하기 위한 목적으로 매핑 테이블을 구축한다(Table 1 참고).

Table 1 Synonym Mapping Table

Mood	Part	Synonym Mood List
annoying	verb	annoy, rag, get to,...
	adj	bothersome, galling, irritating,...
angry	adj	aggravated, provoked, angered, ...
nervous	adj	tense, anxious, queasy, ...
sad	adj	bittersweet, doleful, mournful, ...
bored	verb	bore, tire, drill, ...
	adj	world-weary, tired, blase, ... etc.
sleepy	adj	sleepy-eyed, sleepyheaded, asleep, ...
calm	verb	calm down, tranquilize, tranquillise, ...
	adj	unagitated, serene, tranquil, ... etc.
peaceful	adj	peaceable, irenic, nonbelligerent, ...
relaxed	verb	relax, loosen, loosen up, ...
	adj	degage, laid-back, mellow, ...
pleased	verb	please, delight, satisfy, ...
	adj	amused, diverted, entertained,
happy	adj	blessed, blissful, bright, ...
excited	verb	excite, arouse, elicit, ...
	adj	aroused, emotional, worked up, ...

3.2.2 멀티미디어 콘텐츠의 분위기 태그 카운트 값을 이용한 12 분위기 벡터 계산

멀티미디어 콘텐츠의 분위기 벡터를 계산하기 위해서는 먼저, 12 분위기 태그 카운트 값이 필요한데, 이 값은 Fig. 4와 같이 콘텐츠에 포함된 태그의 카운트와 동의어 사전을 이용하여 얻게 된다. 태그가 Thayer의 12 분위기 단어에 해당되지는 않지만 동의어 사전의 분위기 리스트에 포함되어 있다면 해당 카운트의 값을 해당 분위기 카운트에 더하는 방법이다. Fig. 4에서 콘텐츠의 태그들 중 'excited' 태그는 동의어 사전의 nervous 리스트와 excited 리스트에 포함되고, 'excited'태그의 카운트 값이 50이므로 nervous와 excited의 분위기 카운트 값에 50을 더한 다

음, 수식 (1)과 같이 계산하여 콘텐츠의 12 분위기 벡터를 구할 수 있다.

$$CV_m^k = V_m^k / s_m, \tag{1}$$

$$s_m = \sum_{i=1}^{12} V_m^i$$

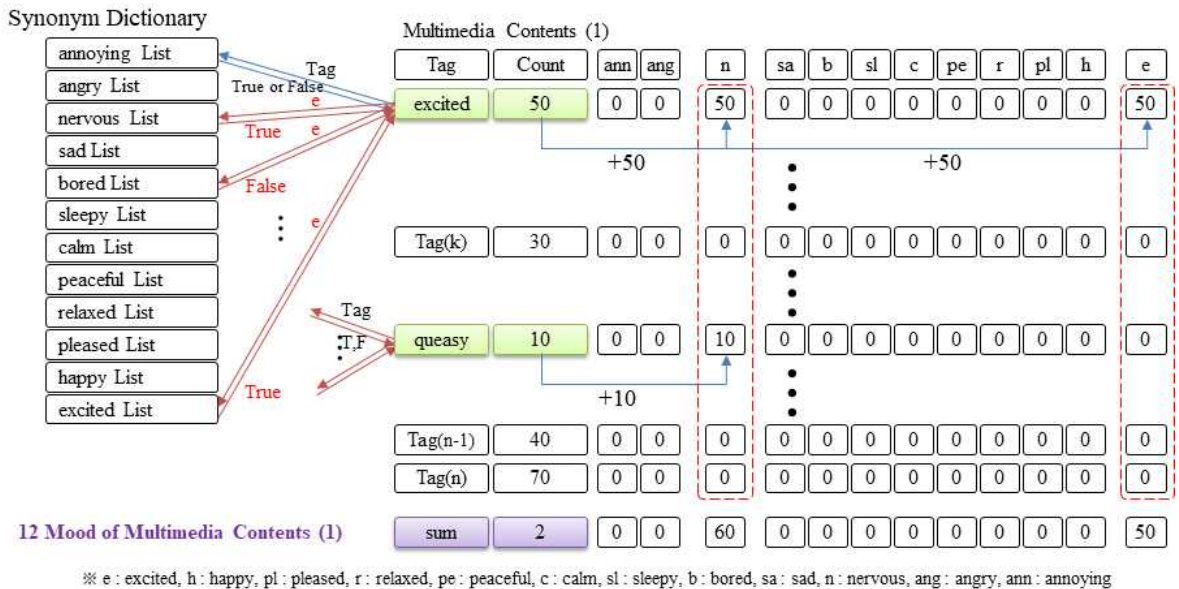


Fig. 4 12 Mood Count Value Calculation of Multimedia Content

여기서, m 은 멀티미디어 콘텐츠의 인덱스를, i 와 k 는 분위기 인덱스를, CV_m^k 는 콘텐츠 m 의 분위기 벡터의 k 번째 차원의 값을 의미하고, s_m 는 콘텐츠 m 의 12 분위기 카운트의 합을 의미한다.

하나의 예로서, m 번째 콘텐츠의 태그와 카운트 값이 Table 2와 같다고 하자.

Table 2의 데이터에 Table 1의 동의어 매핑 테이블을 적용 시키면 annoying의 동의어는 annoy와 bothersome이므로 annoying의 카운트 값은 40이 되고, angry의 동의어는 angered이므로 카운트 값이 15가 된다. 마찬가지로 nervous의 동의어는 tense와 anxious이므로 카운트 값이 17이 되며, excited의 동의어는 excite이므로 카운트 값이 15가 된다. 이를 Thayer 2차원 모델에 적용한다면 annoying, angry, nervous 그리고 excited 분위기는 모델에서 4, 5, 6 그리고 3 번째에 해당하므로 콘텐츠 m 의 카운트 벡터는 $(0, 0, 15, 40, 15, 17, 0, 0, 0, 0, 0, 0)$ 와 같이 구할 수 있고, 이를 식 (1)을 적용하면 최종적으로 $CV_m = (0.0, 0.0, 0.17, 0.46, 0.17, 0.20, \dots, 0.0)$ 의

콘텐츠의 분위기 벡터를 구할 수 있다. 여기서, CV_m 의 전체 합은 1.0이다.

3.3 태그의 12 분위기 벡터 테이블 생성 과정

멀티미디어 콘텐츠의 12 분위기 벡터 값을 이용하여 멀티미디어 콘텐츠를 검색하기 위해서는 태그의 12 분위기 벡터 값도 필요한데, 태그의 12 분위기 벡터 값은 Fig. 5와 같은 과정으로 계산된다. 본 과정에서의 최종적인 결과물로 태그의 12 분위기 벡터 테이블이 생성된다.

태그의 12 분위기 벡터 값 해당 분위기 태그를 포함하는 멀티미디어 콘텐츠들의 12 분위기 벡터 값을 평균함으로 계산된다. Fig. 5에서 'Mood Tag (1)'의 12 분위기 벡터는 식 $average(CV_1^1, CV_2^1, \dots, CV_i^1, \dots, CV_{n-1}^1, CV_n^1)$ 로 계산할 수 있다. 여기서, n 은 멀티미디어 개수를 의미하고, CV_i^1 은 분위기 태그 'Mood Tag (1)'을 포함하는 i 번째 멀티미디어 콘텐츠를 의미한다.

3.4 분위기 벡터기반 멀티미디어 콘텐츠 추천

분위기 태그를 이용하여 멀티미디어 콘텐츠를 추천하는 방법은 Fig. 6과 같이 해당 태그의 12 분위기 벡터와 멀티미디어 콘텐츠들의 12 분위기 벡터와의 유사도를 계산한 후, 그 유사도에

Table 2 Tag and Tag Counter

Tag	Tag Count	Tag	Tag Count
anxious	7	excite	15
tense	10	angered	15
annoy	30	bothersome	10

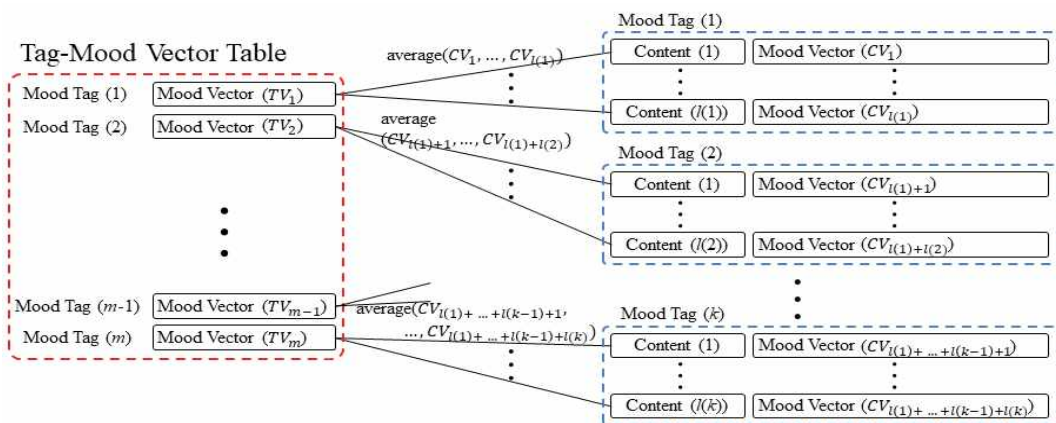


Fig. 5 Tag-12 Mood Vector Table Creation

따라 추천하는 것이다. 유사도가 높은 멀티미디어 콘텐츠들이 우선 추천된다.

유사도 계산은 Table 3의 코사인 유사도 공식을 사용한다. 제안한 방법의 정확률을 비교분석하기 위해 거리유사도를 이용한 방법을 추가로 실험하였다.

코사인 유사도를 구하는 식은 Table 3의 수식 (2)와 같고, 식에서 $TV_k^{12 \text{ Mood Vector}}$ 는 태그의 12 분위기 벡터 테이블에서 태그 k 의 12 분위기 벡터를 의미하고, $CV_l^{12 \text{ Mood Vector}}$ 는 멀티미디어 콘텐츠의 12 분위기 벡터 테이블에서 멀티미디어 콘텐츠 l 의 12 분위기 벡터를 의미한다. 또한 유사도를 0~1으로 정규화 함으로 쿼리 태그 k 에 대한 멀티미디어 콘텐츠 l 의 코사인 유사도 $s_{k,l}^c$ 를 구한다. 거리 유사도는 수식 (3)과 같고, 식에서 $TV_{k,1}$ 와 $TV_{k,2}$ 그리고 $TV_{k,12}$ 는 태그의 12 분위기 벡터 테이블에서 태그 k 의 12 분위기 벡터의 해당 차원의 값을 의미하고, $CV_{l,1}$ 와

$CV_{l,2}$ 그리고 $CV_{l,12}$ 는 멀티미디어 콘텐츠의 12 분위기 벡터 테이블에서 멀티미디어 콘텐츠 l 의 12 분위기 벡터의 해당 차원의 값을 의미하고, $s_{k,l}^d$ 는 쿼리 태그 k 에 대한 멀티미디어 콘텐츠 l 의 거리 유사도 d 를 의미한다.

4. 실험 및 결과

멀티미디어 콘텐츠 추천 방법의 정확률분석을 목적으로 12분위기에 해당하는 약 50,000개의 멀티미디어 콘텐츠 정보와 그들의 태그를 Last.fm 웹사이트의 API를 이용하여 구축하였고, 세부적으로 Table 4와 같다.

이렇게 구축한 데이터를 이용하여 본 논문에서는 멀티미디어 콘텐츠의 12 분위기 벡터 테이블 및 태그의 12 분위기 벡터 테이블을 구축 하였다. 또한 구축한 테이블을 이용하여 멀티미디어 콘텐츠 정확률을 분석하였다. 정확률 평가 및 참값 구축은

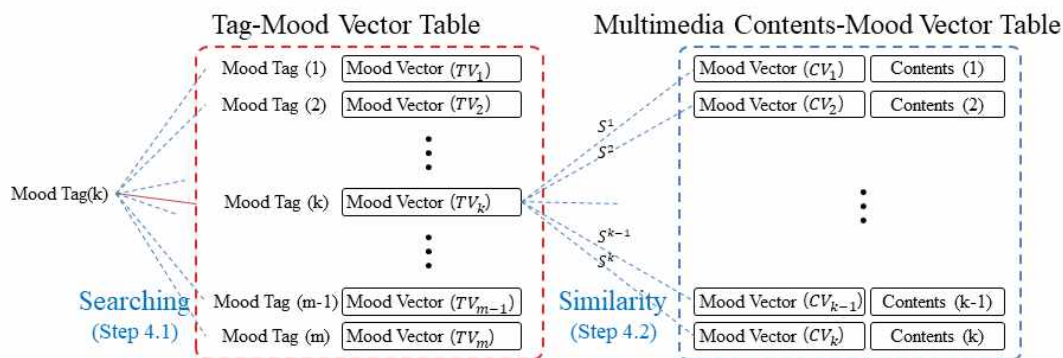


Fig. 6 Multimedia Contents Recommendation Method

Table 3 Similarities and Their Formula

Similarity	Formula
Cosine Similarity	$s_{k,l}^c = \frac{\cos\theta_{k,l} + 1}{2}$ $\cos\theta_{k,l} = \frac{TV_k^{12 \text{ Mood Vector}} \cdot CV_l^{12 \text{ Mood Vector}}}{\ TV_k^{12 \text{ Mood Vector}}\ \ CV_l^{12 \text{ Mood Vector}}\ }$ (2)
Distance Similarity	$s_{k,l}^d = \sqrt{(TV_{k,1} - CV_{l,1})^2 + (TV_{k,2} - CV_{l,2})^2 + \dots + (TV_{k,12} - CV_{l,12})^2}$ (3)

Table 4 Dataset Detail

Mood Tag	Number of Dataset	Mood Tag	Number of Dataset
annoying	1,900	calm	10,000
angry	3,200	peaceful	4,800
nervous	170	relaxed	1,900
sad	10,200	pleased	10
bored	40	happy	10,300
sleepy	7,500	excited	70

Moon et al. (2019)의 방법과 동일하게 재현율 (Recall) 및 정확률(Precision)을 계산하였고, 이를 이용하여 보간된 정확률(Interpolated precision)을 구한 후 12개의 분위기 단어에 대하여 평균 정확률 (Average precision)을 계산하였다 (Powers, 2011).

재현 수준과 정확률을 구하는 식은 수식 (4)와 같고, 보간된 정확률을 구하는 식은 수식 (5)와 같으며, 평균 정확률을 구하는 식은 수식 (6)과 같다.

$$P = \frac{|B|}{|B| + |C|}, \tag{4}$$

$$r = \frac{|B|}{|A| + |B|}.$$

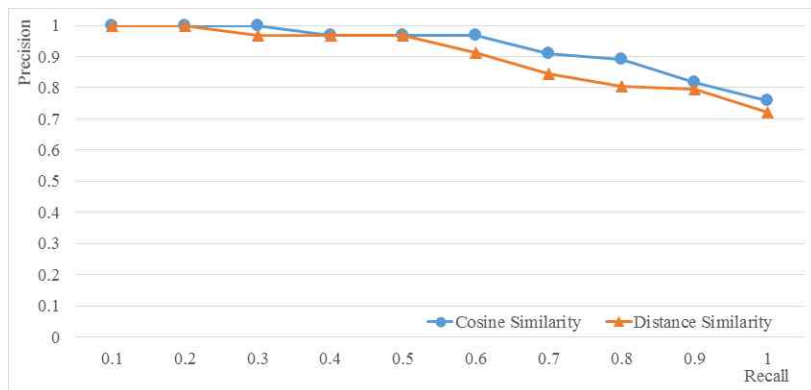
여기서, r 은 태그재현율을 의미하고, P 는 정확률을 의미하고, A 는 False Negative를 의미하고, C 는 False Positive를 의미하고, B 는 True Positive를 의미한다.

$$P(r_j) = \max_{r_j \leq r \leq r_{j+1}} P(r). \tag{5}$$

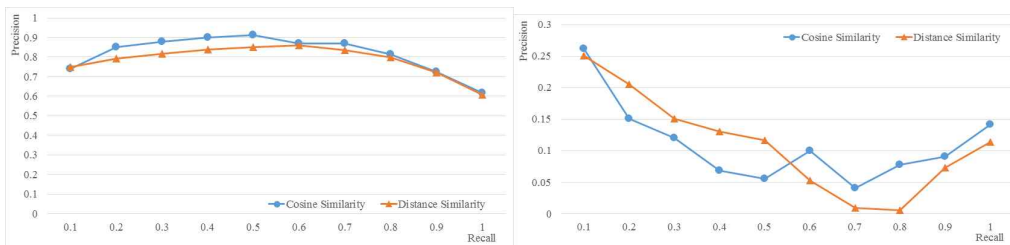
여기서, $P(r)$ 은 재현수준이 r 일 때의 정확률을 의미하고, $r_j \leq r \leq r_{j+1}$ 은 재현율의 범위이다.

본 논문에서는 r 값을 0.1 단위로 사용하였다

$$\overline{P(r)} = \sum_{i=1}^{N_q} \frac{P_i(r)}{N_q}. \tag{6}$$



(a) with Synonyms



(b) without Synonyms

(c) calculated by (a)-(b)

Fig. 7 Recommendation Performance Comparison by 12 Mood Tag

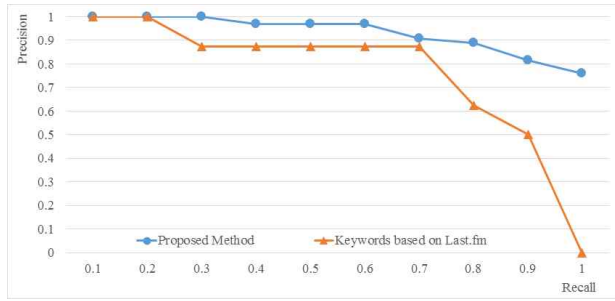


Fig. 8 Recommendation Performance Comparison between Proposed Method and Last.fm Method

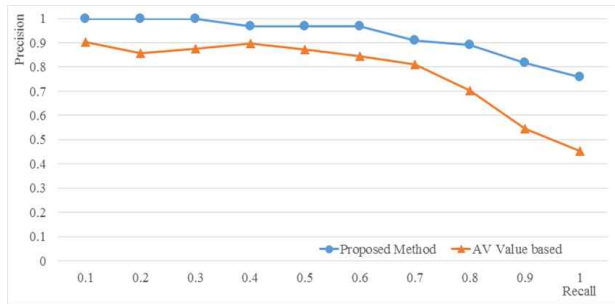


Fig. 9 Recommendation Performance Comparison between Proposed Method and AV Value based Method (Moon et al., 2019)

여기서, N_q 는 질의 개수를 의미하고, $P_i(r)$ 은 i 번째 질의에 대한 재현율 r 에서의 정확률을 의미한다.

제안한 방법의 정확률을 분석하기 위해 세 가지 관점에서 비교를 하였고, 그 첫 번째가 코사인 유사도와 거리 유사도의 정확률과 비교, 두 번째가 제안한 방법과 Last.fm의 키워드 기반 정확률 비교이고, 마지막으로 Moon et al. (2019)에서 제안한 AV값 기반 추천 방법과의 정확률 비교이다. 첫 번째 실험 결과, 유사어를 고려한 경우 Fig. 7 (a)와 같이 코사인 유사도를 이용하는 방법과 거리 유사도를 이용하는 방법 모두 재현 수준 (Recall level) 0.1에서 모두 1.0의 정확률을 보였고, 재현 수준 0.2에서는 코사인 유사도를 이용하는 방법만 1.0의 정확률을 보였으며, 전체적으로 코사인 유사도를 이용하는 방법이 좋은 정확률을 보였다. 또한 유사어를 고려

하지 않은 경우 역시 코사인 유사도를 이용하는 방법이 거리 유사도를 이용하는 방법보다 좋은 정확률을 보였다 (Fig. 7 (b) 참고).

두 번째 실험은 Last.fm의 키워드 기반 정확률과 본 논문에서 제안한 방법의 정확률 비교인데, 실험결과는 Fig. 8와 같이 재현 수준 0.1과 0.2에서 정확률 1.0으로 동일한 정확률을 보였지만 재현 수준 0.3부터는 제안한 방법이 키워드 기반 정확률보다 좋은 정확률을 보였다. 또한 키워드 기반 정확률의 경우 재현 수준 1.0에 근접하는 경우 정확률 0.0에 근접하는 것을 알 수 있고, 제안한 방법의 경우 재현 수준 1.0에 근접하더라도 정확률 0.7이상을 유지하는 것을 알 수 있었다. 즉, 본 실험에서는 제안한 방법이 키워드 기반 추천 방법보다 우수한 정확률을 보임을 알 수 있었다.

마지막으로 Moon et al. (2019)에서 제안한 AV 값 기반 추천 방법과 본 논문에서 제안한 방법의 정확률 비교결과는 Fig. 9와 같고, 본 논문에서 제안한 방법이 전체적으로 우수함을 보였다.

세부적으로 본 논문에서 제안한 방법의 정확률을 분석하기 위해 12 분위기별 정확률을 분석하였다. 분석 결과, Fig. 10와 같이 전체적으로 제안한 방법의 멀티미디어 콘텐츠 추천방법에서 우수한 정확률을 보였고, 특히 AV 값 기반 방법 Moon et al. (2019)에서 문제가 되었던 1,000개 이하의 멀티미디어 콘텐츠(pleased, excited, bored 그리고 nervous)에 대해서도 제안한 방법이 전체적으로 우수한 정확률을 보였다. 즉, 태그 pleased 및 excited, annoying, nervous, bored 그리고 sleepy의 경우 재현 수준 0.1과 0.2에서 제안한 방법이 기존 AV 값 기반 방법과 키워드 기반 정확률보다 좋은 정확률을 보임을 알 수가 있었다. 즉, AV 기반 검색의 경우 분위기 12차원 벡터를 2차원으로 축소하였기 때문에 분위기에 대한 정보가 일부 손실되어 재현 수준 0.1에서 좋지 않은 성능을 보였지만 제안방법의 경우 12차원 벡터를 모두 고려하기 때문에 재현 수준 0.1에서도 좋은 성능을 유지한다고 할 수 있다. Fig. 10에서 삼각형 표기는 AV 값 기반 방법의 정확률이고, 원형 표기는 키워드 기반 정확률이며 마지막으로 사각형 표기는 제안 방법의 정확률을 의미한다.

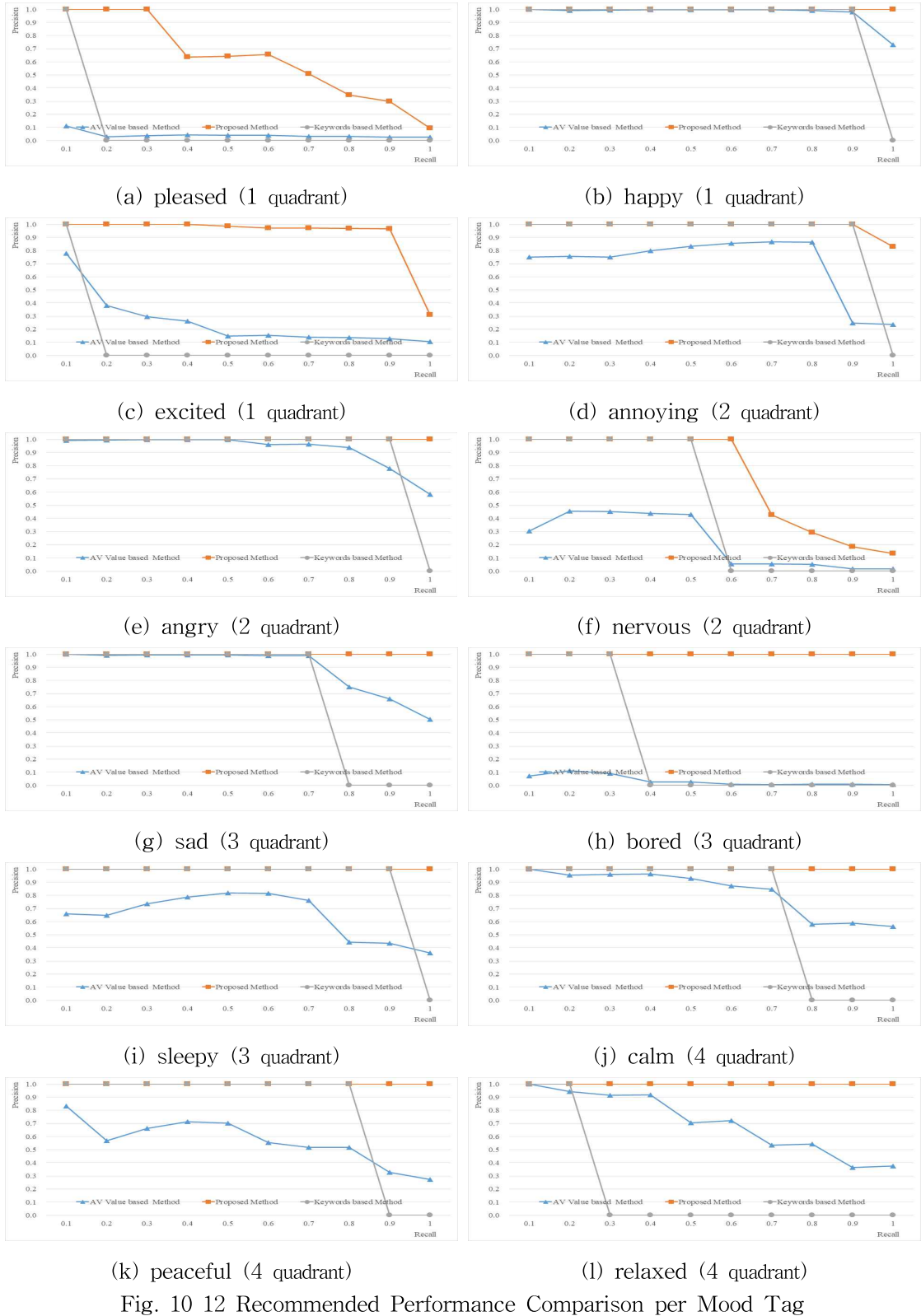


Fig. 10 12 Recommended Performance Comparison per Mood Tag

5. 결 론

본 논문에서는 SNS서비스에는 폭소노미에 존재하는 문제점인 동의어에 대한 문제점을 일부 해결하고, 멀티미디어 콘텐츠 추천에서 가심비를 극대화하기 위한 분위기 태그를 이용한 멀티미디어 콘텐츠 추천 방법을 제안 하였다. 본 제안 방법에서는 Thayer 모델에 기초한 분위기 벡터를 이용함으로써 재현 수준 0.1에서 정확률이 떨어지는 문제점을 해결하고자 하였다.

본 논문에서 제안한 방법을 검증하기 위하여 제안한 방법을 Last.fm의 키워드기반 추천방법과 AV값 기반 추천 방법 (Moon et al., 2019)과의 정확률을 분석을 하였다. 분석결과, AV값 기반 추천 방법에서의 문제 (재현 수준 0.1과 0.2에서 멀티미디어 콘텐츠의 정확률이 키워드기반 추천방법의 정확률 보다 떨어짐)를 해결할 수 있었고, 동의어를 고려하여 멀티미디어 콘텐츠 추천이 가능함을 알 수 있었다. 즉, 본 논문에서 제안한 방법의 경우 12차원 벡터를 모두 고려함으로써 Last.fm의 키워드기반 추천방법 및 AV값 기반 추천 방법보다 우수한 정확률을 보였다.

References

- Chang, E., Kingshy, G., Sychay, G., and Wu, G. (2003). CBSA: Content-Based Soft Annotation for Multimodal Image Retrieval using Bayes Point Machines, *IEEE Transactions on Circuits and Systems for Video Technology*, 13(1), 26 - 38.
- Hevner, K. (1936). Experimental Studies of the Elements of Expression in Music, *The American Journal of Psychology*, 48(2), 246-268.
- Ji, A., Yeon, C., Kim, H., and Jo, G. (2007). Collaborative Tagging in Recommender Systems, *Australasian Joint Conference on Artificial Intelligence*, pp. 377-386
- Kim, J., Lee, S., Kim, S., and Yoo, W. Y. (2011). Music Mood Classification Model based on Arousal-valence Values, *Proceedings of 13th International Conference on Advanced Communication Technology (ICACT)*, pp. 292-295.
- Laurier, C., Sordo, M., Serra, J., and Herrera, P. (2009). Music Mood Representations from Social Tags, *Proceedings of the 10th International Society for Music Information Conference*, Kobe, Japan, pp. 381-386.
- Li, J., and Wang, J. Z. (2008). Real-Time Computerized Annotation of Pictures, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 30(6), 985-1002.
- Lindstaedt, S., Mörzinger, R., Sorschag, R., Pammer, V., and Thallinger, G. (2009). Automatic Image Annotation using Visual Content and Folksonomies, *Multimedia Tools and Applications*, 42(1), 97-113.
- Moon, C. B., Kim, H. S., and Kim, B. M. (2013). Music Retrieval Method using Mood Tag and Music AV Tag based on Folksonomy, *Journal of The Korean Institute of Information Scientists and Engineers*, 40(9), 526-543.
- Moon, C. B., Kim, H.S., Lee, H. A., and Kim, B. M., (2014). Analysis of Relationships Between Mood and Color for Different Musical Preferences, *Color Research & Application*, 39(4), 413-423.
- Moon, C. B., Kim, H.S., Lee, D. W., and Kim, B. M. (2015). Mood Lighting System Reflecting Music Mood, *COLOR Research and Application*, 40(2), 201-212.
- Moon, C. B., Yi, J. Y., Kim, D.-S., and Kim, B. M. (2018). Analysis of Overlapping Mood Tags Based on Synonyms, *Korea Computer Congress 2018 (KCC 2018)*, June 20-22, Jeju, Korea, pp. 667-669.
- Moon, C. B., Lee, J. Y., Kim, D.-S., and Kim, B. M. (2019). Multimedia Content Recommendation in Social Networks using

- Mood Tags and Synonyms, *Multimedia Systems*, to be published, 2019.
- Moon, C. B., Lee, J. Y., Kim, D.-S., and Kim, B. M. (2019). Analysis of Mood Tags For Music Recommendation, *Journal of the Korea Industrial Information Systems Research*, 24(1), 13-21.
- Ness, S. R., Theocharis, A., Tzanetakis, G., and Martins, L. G. (2019). Improving Automatic Music Tag Annotation using Stacked Generalization of Probabilistic Svm Outputs, *Proceedings of the 17th ACM International Conference on Multimedia*, pp. 705-708.
- Powers, D. (2011). Evaluation: From Precision Recall and f-measure to Roc Informedness Markedness and Correlation, *Journal of Machine Learning Technology*, 2(1), 37-63.
- Russel, J. A. (1980). A Circumplex Model of Affect, *Journal of Personality and Social Psychology*, 39(6), 1161-1178.
- Thayer, R. E. (1990). *The Biopsychology of Mood and Arousal*, Oxford University Press.
- Tso-Sutter, K., Marinho, L., and Schmidt-Thieme, L. (2008). Tag-aware Recommender Systems by Fusion of Collaborative Filtering Algorithms, *ACM Symposium on Applied Computing*, pp. 1995-1999.
- Vojnovic, M., Cruise, J., Gunawardena, D., and Marbach, P. (2009). Ranking and Suggesting Popular Items, *IEEE Transactions on Knowledge and Data Engineering*, 21, 1133-1146.
- Yang, S., Kim, S., and Ro, Y. M. (2007). Semantic Home Photo Categorization, *IEEE Transactions on Circuits and Systems for Video Technology*, 17(3), 324-335.
- Yang, S., Kim, S. K., Seo, K. S., Ro, Y. M., Kim, J., and Seo, Y. S. (2007). Semantic Categorization of Digital Home Photo using Photographic Region Templates, *International Journal of Information Processing and Management*, 43(2), 503-514.
- Yang, S., and Ro, R. M. (2007). Photo Indexing using Person-Based Multi-feature Fusion with Temporal Context, *International Conference on Mobile Ubiquitous Computing, Systems, Services and Technologies (UBICOMM'07)*, pp. 257-262.



문 창 배 (Chang Bae Moon)

- 정회원
- 금오공과대학교 소프트웨어공학과 공학사
- 금오공과대학교 소프트웨어공학과 공학석사
- 금오공과대학교 소프트웨어공학과

공학박사

- 영풍전자(주) 선임연구원
- (현재)금오공과대학교 ICT융합특성화센터 연구교수
- 관심분야 : 인공지능, 영상처리, 지식검색, 감성 공학



이 종 열 (Jong Yeol Lee)

- 비회원
- 금오공과대학교 컴퓨터공학과 학사
- 금오공과대학교 컴퓨터공학과 석사
- 금오공과대학교 소프트웨어공학과 박사수료
- (현재)국립금오공과대학교 컴퓨터

소프트웨어공학과 시간강사

- 관심분야 : 신경망컴퓨팅, 영상처리, 정보검색



김 병 만 (Byeong Man Kim)

- 정회원
- 서울대학교 컴퓨터공학과 학사
- 한국과학기술원 전산학과 석사
- 한국과학기술원 전산학과 박사
- 미국 UC, Irvine 대학 방문교수
- 미국 콜로라도 주립대학 대학

방문교수

- (현재)국립금오공과대학교 컴퓨터소프트웨어공학과 교수
- 관심분야 : 인공지능, 정보검색, 정보보안