

감정 온톨로지를 활용한 노래 가사의 감정 분류*

김민호*, 권혁철*

*부산대학교 컴퓨터공학과

e-mail:karma@pusan.ac.kr

Emotion Classification in Song Lyrics using the Emotion Ontology

Minho Kim*, Hyuk-Chul Kwon*

*Dept of Computer Science, Pusan National University

요 약

음악 감정 분류에 관한 기존의 연구들은 템포, 박자, 음정, 음표, 리듬 등과 같은 음악의 멜로디와 관련된 자질을 이용하여 음악 감정을 분류하였다. 그러나 노래(Song)와 같이 가사를 포함한 음악은 같은 스타일의 멜로디라도 가사의 내용에 따라 음악에 대하여 청자가 느끼는 감정이 크게 다르다. 본 논문에서는 감정 온톨로지를 활용하여 노래 가사를 감정에 따라 분류하는 방법에 대하여 제안한다. 기구축된 감정 온톨로지를 바탕으로 네 가지 통사적 규칙을 적용하여 노래 가사로부터 감정 자질을 추출한다. 추출된 감정 자질을 이용하여 Naive Bayes, HMM, SVM과 같은 기계학습 기법을 이용하여 8개 감정 그룹에 대해 58.8%의 정확도를 보였다.

1. 서론

인터넷과 컴퓨터의 발달로 디지털 음원의 생산성이 향상되었고, 오늘날과 같이 음악(Music)의 보편화가 이루어졌다. 스마트폰의 등장은 이들의 접근성을 더욱 높여, 음악이 우리의 생활 속으로 더욱 깊게 자리매김하도록 하였다. 지금도 수많은 음악이 만들어지고 배포되고 있다. 거대한 양의 음악으로부터 우리는 원하는 음악을 선택하여 들을 수 있고, 그 방법의 하나가 특정 분위기, 감정의 음악을 선택하는 것이다. 음악 추천 시스템은 사용자의 이러한 선택을 도울 수 있다. 음악 추천 시스템에서는 사용자의 상황, 기분에 따라 적합한 음악을 추천하는데, 이를 위해 음악에 대한 분류가 필요하다. 클래식, 재즈 등 장르에 따라 혹은 기쁨, 슬픔, 사랑 등의 감정에 따라 음악의 분류가 되어 있으면 사용자의 요구와 사용자의 상황에 적합한 음악을 추천할 수 있다[1].

음악 장르의 분류는 이미 오래전부터 연구됐으나, 음악 감정의 분류는 최근 들어서야 활발히 연구되고 있다 [2][3][4]. 감정 분류에 관한 기존의 연구들은 템포, 박자, 음정, 음표, 리듬 등과 같은 음악의 멜로디와 관련된 자질을 이용하여 음악 감정을 분류하였다. 그러나 노래(Song)와 같이 가사를 포함한 음악은 같은 스타일의 멜로디라도 가사의 내용에 따라 음악에 대하여 청자가 느끼는 감정이 크게 다르다[5]. 노래 감정에서 가사는 중요한 감정 전달의 역할을 하므로 노래 가사를 기반으로 한 감정 분류에

대한 연구가 필요하다.

가사는 텍스트 기반의 감정 분류의 한 도메인이다. 텍스트로부터 감정을 추론하는 방법은 신문, 블로그, 영화, 대화 등 특정 도메인에서 기계학습, 사전기반 등의 방법으로 연구되고 있다. 그 중엔 감정 분류를 정의하고, 감정 어휘와 그에 대한 감정 명세, 속성값으로 구성된 감정 온톨로지가 있다.

본 논문은 노래 가사 기반의 감정 분류에 관한 연구를 담고 있으며, 그 구성은 다음과 같다. 먼저, 2장에서 지금까지 연구된 국내·외 음악 감정 모델을 살펴보고, 본 논문에서 선택한 감정 온톨로지에 대하여 살펴본다. 3장에서는 본 논문에서 제안하는 감정 온톨로지의 사용하여 감정 자질을 추출하는 방법과 추출된 감정에 대한 성능을 평가한다. 4장에서 제한한 모델에 대한 성능 평가를 하고, 5장에서 결론을 내린다.

2. 감정 분류 모델과 텍스트 감정 분류 연구

음악의 감정을 분류하기 위해 어떠한 감정 모델을 사용할 것인지 결정하여야 한다. 감정 모델을 선택할 때 고려할 사항은 감정 모델이 얼마나 감정을 잘 표현하며, 각 감정이 음악의 감정을 잘 표현할 수 있는가이다. 기존의 음악 감정 분류연구에서는 Thayers 모델[6]과 The Tellegen-Watson-Clark 모델[7]이 주로 사용되었다.

Thayers 모델은 간단하면서도 효율적인 감정 모델로서, Stress와 Energy에 대한 두 개의 축을 가진다. 이 값을 이용하여 음악이 가진 소리의 강함이나, 음조, 템포에

* 이 논문은 2010년도 정부(교육과학기술부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임(No. 2010-0028784)

관한 느낌을 2차원으로 표현한다.

텍스트 기반 감정 처리에 관한 연구는 다양한 말뭉치를 대상으로 감성 극성(긍정, 부정, 중립)에 대하여 분류하는 것과 더욱더 세분된 감정으로 분류하는 것으로 나누어 진행되고 있다. 감정 극성분류에 관한 연구로는 블로그, 신문, 영화평, 상품평을 대상으로 Naive Bayes, SVM[8][9] 등의 기계학습 기법을 사용하여 분류하는 연구가 있으며, 정확도를 높이기 위해 WordNet[10][11]과 같은 사전을 활용하여 감정 분류를 하거나 자연언어처리 기법 등을 이용하여 감정 극성을 분류하는 연구가 진행 중이다. 더욱 세분된 감정 분류에 대한 연구로는 감정의 표현이 풍부한 대화형 말뭉치를 대상으로 기계학습 기법을 이용하여 9-12개 정도의 감정을 분류한 연구가 수행되고 있고[12][13]. 이에 따라 분류의 성능을 높이기 위하여 부정어 처리 및 단어 가중치 적용 기법[14], 효과적인 자질 선택에 관한 연구가 활발히 진행 중이다[15]. 기계학습 기법을 사용하며, 사용되는 자질 집합으로 어휘, 선행적 감정 어휘, 문맥 자질 등을 사용한다. 이러한 자질의 사용은 어느 정도의 만족할 만한 성능을 위해서는 많은 양의 학습 데이터가 필요하다는 단점이 있으며, 사용하는 자연어 처리 기법에서도 부정어 처리에 한정적이다. 언어 자원을 활용한 연구는 영어권에서 WordNet 2.0 어의에 3가지 값(긍정, 부정, 중립)을 태깅하여 구축한 SentiWordNet이 있어 이를 활용하는 연구가 있다[16]. 하지만, 이는 긍정, 부정, 중립의 분류로 세분된 감정분류에 대한 연구는 아니다. 한국어에서는 이러한 언어 자원 구축이 없다가, 최근에 YOON&KWON에 의해 감정 프레임워크를 정의하고 감정 라이브러리를 구축한 텍스트 기반 감정 온톨로지가 있다[17].

본 연구에서는 YOON&KWON의 감정 온톨로지를 활용하여 통사적 규칙을 적용하여 추출한 감정을 자질로 사용하여 기계학습 기법을 이용하여 가사를 기반으로 한 음악 감정 분류를 한다.

3. 감정 온톨로지를 활용한 감정 자질 추출

YOON&KWON의 감정 온톨로지는 크게 감정 분류 프레임워크와 감정 표현 라이브러리로 나눈다. 감정 분류 프레임워크는 인간 감정의 공학적 처리를 가능하도록 정의한 분류체계를 뜻하며 감정 표현 라이브러리는 감정 분류 프레임워크를 기반으로 서술형 텍스트와 대화형 텍스트를 분석 및 태깅하여 구축한 데이터베이스를 뜻한다. YOON&KWON은 기존의 감정 연구를 바탕으로 텍스트 기반 감정 온톨로지를 구축하였다.

Plutchik[18]의 연구를 수용하되, 문제점 분석을 통하여 명칭 변경 및 강도 구분을 통하여 기본 감정 8개와 조합 감정 17개로 재정의하였으며 기본감정에 대하여 다섯 단계의 강도 변화, 조합감정에 대하여 세 단계의 강도 변화

를 두었다. 감정의 속성으로 극성(긍정, 부정, 중립), 감정의 경험자(화자, 청자), 기술대상(언어적, 비언어적), 기술방식(설명적, 표현적, 도상적)을 가진다. 기술대상, 기술방식에 따라 6개의 범주로 분류할 수 있으며, 25개의 감정 명세, 강도, 극성의 속성으로 감정의 차별성을 두어, 경험자, 기술대상과 방식, 언어적 자질의 속성으로 텍스트에 나타난 감정 표현을 명시적으로 구분할 수 있도록 하였고, 이는 멀티모달 환경에 적합하도록 설계되어 텍스트 기반 감정 분류 프레임워크로 선택하기에 알맞다. 노래 가사의 분류는 텍스트에서 감정을 추출하며, 텍스트에서 나타난 감정 표현을 나타내기에 알맞은 YOON&KWON의 감정 프레임워크를 사용하였다.

감정 라이브러리의 활용성을 높이기 위해 감정 표현 어휘를 성격에 따라 기본형(canonical form)과 활용형(surface form)으로 구성하였다. 기본형은 어휘 표현의 어미, 조사를 제외하여 구성하였고, 표면형은 활용형 형태를 그대로 사용한다. 기본형과 활용형의 구분은 언어의 생산적 특징에 따라 구분하였다. 예를 들어 기본형 “사랑하다”의 활용형은 문장을 끝맺는 종결형 “사랑해”, “사랑한다.” 등이 있고, 문장을 연결해주는 연결형에는 “사랑하고, 사랑하는데” 등이 있고, 문장에서의 기능을 바꿔주는 전성형에는 “사랑하는, 사람함” 등 이러한 활용형태는 무수히 많으며, 이들 모두를 활용형 형태의 라이브러리 구축은 상당히 비효율적이며, 자료 부족 현상(data sparseness)으로 감정 라이브러리에 모든 활용형이 들어 있음을 보장할 수 없다. 이들은 기본형 형태의 사전으로 구성하고, 시스템에서는 형태소분석을 통한 기본형 비교를 통해서 이루어진다. 활용형으로 구성하는 경우는 특정 활용형은 하나의 감정을 표현하는 경우에 구성하였으며 활용형의 예는 “내 꿈꾸다.”의 경우 감정을 느끼기 어렵지만, 특정 어미 “어”가 붙음으로써 “내 꿈꿔”가 되어 특정 감정 <LOVE>를 표현하는 경우가 그 예이다.

감정 라이브러리를 이용하여 감정을 추출할 수 있으나, 이때 중요한 것은 평가 측도에서 recall보다 precision에 있는데. 이는 기본형의 감정 추출 시 문제가 될 수 있다. “내 사랑 하나 못 찾고” 또는 “믿지 않겠어.”라는 문장에서 “사랑, 믿다”는 감정 표현 어휘로 기본형 감정 추출 시 감정이 추출되지만, 실제 문장이 나타내는 감정은 각각 <LOVE>, <TRUST>가 아니다. 음악 추천 시스템에서의 사용자의 만족도는 감정 추출의 정확도가 좌우할 수 있으며, <LOVE> 감정의 곡을 원할 때 잘못된 추천은 사용자의 만족도를 떨어뜨릴 수 있다.

본 논문에서는 한국어의 언어학적 특징을 이용한 규칙을 생성하여 적용함으로써 감정 추출의 정확도를 향상하였다. 생성한 규칙으로 기본형의 감정 추출을 제어하는 것은 주변 공기 어휘의 관계에 따라 감정이 없어지거나 달라지는 표현에 대해 감정 추출 오류를 줄일 수 있다.

본 논문에서는 임의의 노래 300곡의 가사를 수집하여

감정 어휘와 주변 공기 어휘의 관계에 따라 감정의 변화를 관찰하여 네 가지 통사적 규칙을 생성하였다. 네 가지 통사적 규칙은 다음과 같다.

부정어 결합: 감정 표현 어휘에 부정을 나타내는 어휘가 나타나면 감정 추출에서 제외하고 다른 추출된 감정에 따라 노래 가사의 감정을 선택하는 것이 더 효과적이다. 이러한 부정 구문의 패턴을 정의하고 감정 표현 추출에 제한을 두었다.

감정의 시점: 한국어에서 시제는 선어말어미 ‘-는-’, ‘-았-’, ‘-겠-’을 통해 현재, 과거, 미래로 표현한다. 현재 시제를 나타내는 ‘-는-’이 함께 쓰인 감정 표현 어휘는 현재 화자의 감정을 나타내는 표현이 될 수 있지만, 과거시제를 나타내는 ‘-았-’ 또는 미래시제(추측, 의지 또한 표현)를 나타내는 ‘-겠-’은 감정 표현의 오류가 될 수 있다.

감정 상태 변화: 감정 표현 어휘와 감정 상태를 변화시키는 동사 ‘지나다, 떠나다, 사라지다.’ 등과 결합하여 현재의 감정 변화를 표현하는 구문이 있으며, 이는 현재 화자의 감정 변화를 나타내어 감정 표현의 오류가 될 수 있다. 감정 상태 변화를 주는 동사에 25개에 대해 정의하고 감정 어휘와 공기하면 감정 추출에 제한을 두었다.

의문문: 노래 가사 문장이 의문문일 때는 화자의 감정을 표현하는 것이 아니라 감정에 대한 사실을 묻는 것으로 의문문의 확인이 필요하다.

4. 감정 자질을 이용한 노래 가사 분류

가사에서 추출한 감정 자질을 활용하여 음악 감정 분류에 사용하였다. 실험 말뭉치를 구성하기 위해 한국어 가사를 포함하는 임의의 음악 525곡에 대하여 한국어를 모국어를 사용하는 태가가 가사 텍스트에서 느낄 수 있는 대표적인 감정값 하나씩 선정하여 평가 말뭉치를 구축하였다. 같은 가사에 대하여 다른 감정을 느낄 수 있기 때문에 세 명의 실험 참가자에 의하여 실시하였다. 세 명이 모두 같은 감정으로 태깅한 경우가 전체 곡의 213곡(41%)이며, 두 명 이상 일치한 경우가 425곡(82%)이었다. 실험 말뭉치의 신뢰성을 위해 두 명 이상 일치한 곡 425곡을 실험 말뭉치를 사용하였다. 그리고 감정 온톨로지에 명세한 25개의 감정을 (그림 1)과 같이 8개의 감정 그룹으로 나누어 학습과 평가를 시행하였다.

NB(Naive Bayes) 모델과 SVM(Support Vector Machine)을 사용하여 전체 말뭉치의 90%로 모델을 학습하고, 나머지 10%로 평가를 하였다. 평가 지표는 전체 음악 수에서 정확하게 감정을 분류한 음악 수의 비율이 얼마인지를 측정한 정확도(accuracy)를 이용하였다. <표 1>은 8개로 그룹화한 25개 감정에 대한 분류 정확도이다.

NB, SVM 모델은 BOW(Bag of Words) 모델로 시간, 위치 정보의 학습이 어렵다. 하지만, 가사에서 감정이 나타난 위치가 처음이나, 중간이나, 끝이나에 따라서 전체적인

<표 1> 사용된 자질 집합에 다른 NB와 SVM 비교

사용된 자질 집합	NB	SVM
① : 출현 감정의 빈도	44.9%	51.3%
② : ① + 극성(긍정, 부정) 정보	44.2%	53.6%
③ : ② + 제한된 감정 정보	45.1%	58.8%

가사의 감정은 달라질 수 있다. 예를 들어 가사에서 처음에 <LOVE>에 관한 감정이 나타났고, 마지막에 <SADNESS>에 관한 감정이 나타났다면? 또는 <SADNESS>에 관한 감정이 나타나다가 <LOVE>의 감정이 나타났다면? 가사에서 감정은 텍스트 내의 위치, 시간적 흐름에 의미가 있을 수 있다. [13]에서는 하나의 문장 내에서 표현되는 여러 감정을 분석하기 위해 HMM(Hidden Markov Model)을 사용하였지만, 본 연구에서는 시간적 흐름의 정보를 반영하기 위하여 HMM을 사용하였다. HMM은 상태(state)와 상태 간의 전이 확률(state transition probabilities)을 이용하여 출력을 결정하며 이는 시간적 흐름 정보를 반영할 수 있다. <표 2>는 HMM 모델과 NB 모델의 실험 결과를 비교한 것이다. NB 모델보다 약 9%의 성능 향상이 있으며, 이는 시간적

<표 2> 사용된 자질 집합에 다른 NB와 HMM 비교

사용된 자질 집합	NB	HMM
① : 출현 감정의 빈도	44.9%	48.6%
② : ① + 극성(긍정, 부정) 정보	44.2%	52.2%
③ : ② + 제한된 감정 정보	45.1%	54.2%

흐름 정보에 의미가 있음을 보여준다.

감정을 자질로 사용하였을 때 감정 분류 성능이 높다는 것을 보여주기 위해 일반 어휘를 자질로 사용하였을 때의 연구와 비교하였다. 최근 텍스트 감정 분류 연구와 비교하였다. [15]에서는 SVM 모델을 사용하여 문장 자질, 감정 어휘 자질, 문맥 자질을 사용하였다. 각 자질의 성능을 나타내었으며 이중 precision에서 가장 성능이 높게 나온 감정 어휘 자질을 사용하여, 본 논문의 제안 시스템과 비교하였다. 감정 어휘는 감정 온톨로지의 감정 표현 어휘를 사용하여 실험하였다. <표 3>은 비교 시스템과의 성능 비교이다.

비교 시스템에서는 감정 어휘를 자질로 사용하게 되어 많은 감정 표현 어휘로 말미암아 많은 노이즈가 나타났다. 실제 실험 결과를 보면, 음악 감정의 많은 비율을 차지하는 <LOVE>, <TRUST> 그룹인 그룹 3에서 recall 99.1%, precision 29.2%의 결과가 나왔다. 이는 감정 어휘를 자질로 사용하게 되어 감정 표현이 많은 가사에서 이를 이용한 방법으로 학습이 어려움을 나타낸다.

그룹 1	- <ANTICIPATION>, <PRIDE>, <OPTIMISM>
그룹 2	- <JOY>, <GRATEFULNESS>
그룹 3	- <LOVE>, <TRUST>
그룹 4	- <FEAR>, <GUILT>, <SURPRISE>, <SHAME>, <SENTIMENTALITY>
그룹 5	- <DISAPPOINTMENT>, <SADNESS>, <PESSIMISM>
그룹 6	- <REMORSE>
그룹 7	- <DISGUST>, <ANXIETY>, <BOREDOM>, <CYNICISM>
그룹 8	- <CONTEMPT>, <ANGER>, <AGGRESSION>, <ENVY>, <CURIOSITY>

(그림 1) 감정 그룹별 감정

<표 3> 기존 시스템과 제안 시스템의 성능 비교

Class	비교 시스템			제안 시스템		
	Precision	Recall	F-Measure	Precision	Recall	F-Measure
그룹 1	0	0	0	48.2	65.1	55.4
그룹 2	0	0	0	0	0	0
그룹 3	29.2	99.1	45.2	56.7	70.8	63.0
그룹 4	0	0	0	0	0	0
그룹 5	76.2	15.0	25.0	63.3	64.5	63.9
그룹 6	57.1	12.1	20.0	66.7	60.6	63.5
그룹 7	0	0	0	0	0	0
그룹 8	0	0	0	0	0	0
Micro Average	40.3	32.9	23.0	53.7	58.8	55.8

5. 결론 및 향후 연구

본 논문에서는 감정 온톨로지를 활용하여 노래 가사를 감정에 따라 분류하는 방법에 대하여 제안하였다.

감정 자질을 추출하기 위해 기구축된 감정 온톨로지를 활용하였다. 감정 온톨로지를 사용하기 위해 감정 온톨로지 범주에 따라 선택하고, 사용방법을 구분하였다. 네 가지 통사적 규칙을 적용함으로써 추출을 정확성을 높이며, 정확한 자질을 구성할 수 있게 하였다.

기존 연구와 달리 학습 자질로 감정을 사용함으로써 자질의 수를 크게 줄여 적은 양의 학습 데이터를 사용하여 좋은 성능을 낼 수 있었다. 사용된 자질을 감정 분류에 사용하여 8개 그룹으로 실험하였을 때 58.8%의 정확도를 보였다.

시간적 정보의 의미를 확인하기 위해 HMM과 NB와 비교하여 시간적 정보의 의미를 확인하였다. 이는 노래 가사에서 시간적 정보가 의미가 있음을 보여준다.

향후 연구로는 더 많은 데이터를 활용한 감정 정확도 향상, SVM과 시간적 정보를 활용할 수 있는 HMM의 혼합 모델, 또는 멜로디 기반의 음악 감정 분류와 결합하는 연구로 음악 추천 시스템에서 사용자의 만족도를 높여줄 수 있을 거로 생각한다.

참고문헌

- [1] Rainer Typke, Frans Wiering, Remco C. Veltkamp "A survey of music information retrieval systems" Multimedia Information Retrieval, 2005
- [2] T. Li and M. Ogihara "Detectin gEmotion in Music" ISMIR, 2003,
- [3] Y Feng, Y Zhuang, Y Pan "Music information retrieval by detecting mood via computational media aesthetics", IEEE/WIC international

Conference on Web Intelligence, 2003

[4] 유민준, 김현주, 이인권 "감성모델을 이용한 음악 탐색 인터페이스, HCI2009, 2009

[5] 진소영, 최병철, "노래의 가사가 개인의 음악선호에 미치는 영향", 한국 음악치료학회지, 2006

[6] R. Thayers. "The biopsychology of mod and arousal", Oxford University Press. 1989.

[7] A. Tellegen, D. Watson and L. Clark. "On the dimensional and hierarchical structure of affect." Psychological Science, 1999

[8] B. Pang, L. Lee and S. Vaithyanathan, "Thumbs up? Sentiment Classification using machine Learning Techniques,", EMNLP, 2002

[9] K. Dave, S. Lawrence, D. M. Pennock, "Mining the Peanut Gallery: Opinion Extraction and Semantic Classification of Product Reviews,"Proceedings of the 12th international conference on World Wide Web, 2003

[10] WordNet (<http://wordnet.princeton.edu>)

[11] J. Kamps, M. Marx, R.J. Mokken, and M.D. Rijke, "Using WordNet to measure semantic orientation of adjectives" Proceedings of the 4th International Conference on language Resources and Evaluation, 2004

[12] 임성수, 조성배 "Computing with Words기반 SMS로부터의 사용자 감정 추론, 한국컴퓨터종합학술대회 논문집 제38권 제1호, 2007

[13] 문현구, 장병탁 "HMM을 이용한 채팅 텍스트로부터의 화자 감정상태 분석," 한국정보과학회 추계학술발표 논문집, 제28권 제2호, 2001

[14] 정유철, 최윤정, 맹성현, "감정 기반 블로그 문서 분류를 위한 부정어 처리 및 단어 가중치 적용 기법의 효과에 대한 연구", 인지과학제 19권 제 4호, 2008

[15] 박홍민 "한국어 발화 문장 감정 분류를 위한 효과적 자질 집합과범주 체계 결정", 서강대학교 석사 학위 논문, 2009

[16] Bruno Ohana, Brendan Tierney. "Sentiment classification of reviews using SentiWordNet, Dublin Institute of Technology, 2009

[17] 윤애선, 권혁철 "감정 온톨로지의 구축을 위한 구성요소 분석", 인지 과학 제21권 제1호, 2010

[18] Plutchik R, Kellerman H. EMOTION Theory, Research, and Experience, Academic press, 1980