

영화 리뷰의 감정어휘 공간 및 영화 관람의 상황분석 연구

A Study on Analysis of Affective Words in Movie Reviews and the
Situation of Watching Movies

주저자

하효지 (Ha, Hyo Ji)

아주대학교 미디어학과 통합디자인연구실 연구원

공동저자

김기남 (Kim, Gi Nam)

아주대학교 미디어학과 통합디자인연구실 연구원

교신저자

이경원 (Lee, Kyung Won)

아주대학교 미디어학과 교수

kwlee@ajou.ac.kr

목차

1. 서론

- 1-1. 연구의 배경 및 필요성
- 1-2. 연구목적 및 연구과정

2. 관련연구

- 2-1. 영화추천
- 2-2. 감정어휘
- 2-3. 상황분류

3. 연구 방법 및 분석 결과

- 3-1. 영화 관련 감정어휘 분포맵 제작
- 3-2. 영화를 보는 상황에 따른
감정어휘 빈도 분석

4. 결론

참고문헌

(요약)

본 연구는 다수의 주관적 정보이며 집단지성이라고 할 수 있는 영화 리뷰들을 데이터베이스로 하여, 사용자의 프로파일 정보가 거의 없는 상태에서도 사용자의 이용 동기에 따라 선호 성향을 분석해 영화를 추천하는 새로운 방법을 제안하고자 한다. 이를 위해 본 논문에서는 영화 리뷰의 감정어휘 분석과 영화를 보는 상황 분석이 수행되었다. 첫 번째로 감정어휘를 확보하기 위하여 834개의 한국어 정서용어 중에서 전문가의 자문과 서베이를 통해 최종 68개의 감정어휘를 선정하고 유사성/비유사성의 정도를 알아보기 위하여 다차원척도분석법을 이용하여 감정어휘간의 상관관계를 분석하였다. 두 번째로 해당 영화를 보는 가장 적합한 상황에 대한 서베이를 시행하여 추천 상황을 수집하고 해당 영화의 리뷰에서 감정어휘를 추출하여 최종적으로 상황에 따라 어떤 감정어휘 빈도를 가진 영화를 추천해야 사용자들이 만족하는지에 대한 근거를 마련하였다. 본 연구는 상황에 따라 추천할 수 있는 영화의 감정어휘 분포를 제시하는 과정을 보여줌으로써 향후 제작 될 영화 추천 시스템에 많은 도움이 될 것으로 예상된다.

(Abstract)

This research suggests another ways to recommend movie by analyzing preference inclination depending on usage motivation of users with little profile information, using movie reviews. For the research, the analysis of affective words in movie reviews and the analysis of the situation that shows watching movies are conducted in this thesis. First of all, for procuring affective words, totally 68 affective words are chosen by professional consultations and surveys from 834 Korean affective words. Also, for investigating the degree of similarity or dissimilarity among several affective words, the correlation among several affective words is analyzed using multi-dimensional scaling. Second, by conducting surveys about the most accurate situation for watching applied movie, this research collects recommendation situation classified by applied movie. Also, by extracting each affective word from the review of applied movie, this research consequentially finds grounds that show what kind of affective word in movie, which makes users feel satisfied, is worth to be recommended depending on users' situation. Therefore, this research is expected to be a great help for movie recommendation system that is going to be produced by showing the procedure that suggests the dispersion of affective words from the movies which is worth to be recommended depending on users' situation.

(Keyword)

Affective Words, Movie Reviews, MDS, Situation Analysis

1. 서론

1-1. 연구의 배경 및 필요성

오늘날 영화 정보를 제공하는 여러 수단 중 온라인 커뮤니티는 감독, 배우, 영화의 장르 등의 객관적 정보뿐만 아니라 영화를 관람한 사람들이 느끼는 감성 및 의견 공유가 가능하다. 따라서 온라인 커뮤니티에서는 영화에 대한 느낌과 의견 공유를 위해 작성하는 사용자 코멘트 및 영화 리뷰, 댓글 등은 영화의 질을 암시적으로 재현하고 있으며, 여기에서 나타나는 사용자들의 의견은 관람할 영화를 선택할 때 가장 많이 참고 되는 메타데이터라 할 수 있다.¹⁾

또한 추천 시스템이라는 측면에서 볼 때 효과적인 추천을 위해서는 사용자가 원하는 바가 무엇인지 파악할 필요가 있는데, 이 때 사용자가 처하게 된 상황이나, 특정 상황을 해결하는 행위 및, 행위를 하기 위한 사용자들의 접근 의도를 파악한다면 그 의도에 맞는 제품이나 서비스를 적절하게 제공할 수 있게 된다. 이러한 사실을 뒷받침하는 이용과 충족 이론은 이용자가 왜 특정 미디어를 선택하고 이용하는가에 대한 설명을 제공해 준다.²⁾ 특히 사용자의 이용 동기는 수용자의 미디어 선택과 만족도에 대한 설명을 해주는 핵심 개념으로 사용되고 있다.³⁾ 이러한 배경을 통틀어서 본다면 영화를 추천하기 위해서는 주관적인 정보 이외에도 영화를 보는 상황의 분석 또한 중요하다는 것을 볼 수 있다.

따라서 영화 추천 시스템을 제작하기 위해서는 해당 영화가 가지고 있는 객관적인 정보와 더불어 사용자들이 영화를 보게 되었을 때 느끼는 감성의 종류가 어떠한 것이 있는지를 살펴본 다음에 영화를 보는 상황과 함께 연결시키는 것이 필요하다는 것을 파악할 수 있다.

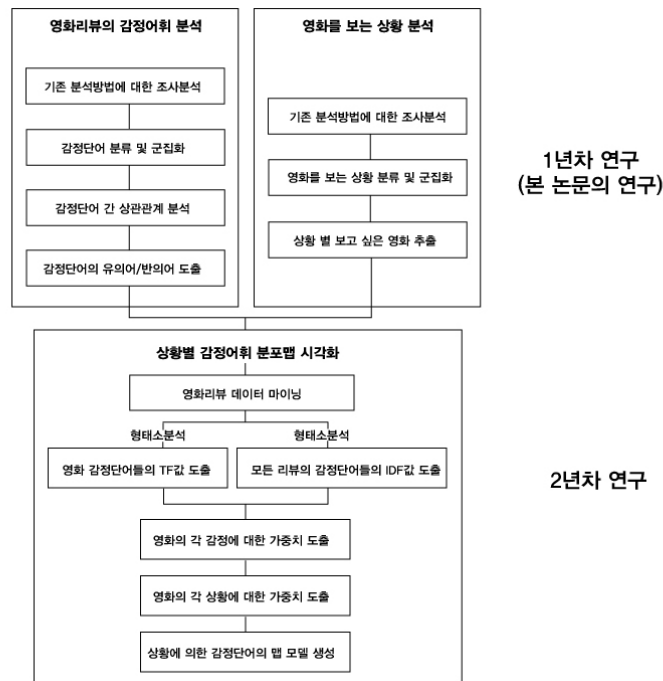
1-2. 연구목적 및 연구과정

본 연구는 영화를 관람할 때 느낄 수 있는 감정어휘를 데이터베이스로 하여 어휘들 간의 거리를 분포맵으로 표현하고, 사용자의 프로파일 정보가 거의 없는 상태에서도 사용자의 상황별 선호 성향을 분석하여 영화를 추천하는 방법을 제안하는 것을 연구의 목적으로 한다. 이처럼 감정어휘에 관한 분포맵을 이용하여 영화 추천 시스템을 제작한다면, 여러 감정이 복잡하게 얽혀있는 영화에 대해서 직관적이면서도 정확하게 감정어휘를 시각화할 수 있기 때문에 사용자들이 해당 영화에 대한 판단을 보다 쉽게 할 수 있을 것이라 예상된다.

연구목적을 수행하기 위해서 1년차 연구인 본 논문에서는 우선 영화 리뷰에서 볼 수 있는 감정어휘를 전문가의 분석 및 서베이를 통해 분류하고, 분류된 감정어휘를 통계적인 기법을 사용하여 어휘의 성격에 따라 군집화한 다음, 각 감정어휘들 간의 유사성/비유사성의 정도를 알아보기 위하여 다차원 척도분석법(Multi-Dimensional Scaling: MDS)을 이용하여 감정어휘 간의 상관관계를 분석하였다.

두 번째로 영화를 보는 상황에 대해서 분석하고 영화 선택 동기에 대한 설명하기 위해 감정 어휘와 연결시키는 작업을 시행하였다. 이를 위해 영화 관련 사이트에서 가져온 2년간의 흥행 영화 데이터를 기반으로 해당 영화를 보는 상황과 관련된 설문 조사를 시행하여 보편적인 상황에서 영화를 보는 상황을 수집하고 군집화를 통해 분류하였다. 그 후 상황들을 감정어휘와 연결시키고 최종적으로는 상황마다 어떤 영화를 추천해야 사용자들이 만족하는지에 대한 근거를 제공하게 된다. 본 논문의 전체 연구 체계는 <그림 1>과 같다.

- 1) 안신현, 시정곤, 문화적 메타데이터에 기반한 영화추천시스템, 한국컴퓨터종합학술대회, Vol. 35, No. 1(A), 2008, p.79.
- 2) Dimmick, J., Kline, S., Stafford, L. The Gratification Niches of Personal E-mail and the Telephone. Communication Research, Vol. 27, No. 2, 2000, pp.227-248.
- 3) 한상필, 안보섭, 박명진, 무선 인터넷 이용자의 모바일 광고 이용 동기에 관한인식 연구, 커뮤니케이션학연구, Vol. 13, No. 4, 2005, pp.68-100.



〈그림 1〉 본 논문의 연구 체계

2. 관련연구

2-1. 영화추천

영화 추천 방법에 대한 연구는 크게 ‘정보 필터링 기술을 활용한 내용기반 추천 시스템’과 ‘협력적 추천 시스템’을 중심으로 이루어지고 있는데 그 중 내용기반 추천 시스템을 연구한 Douglas W Oard, Gary Marchionini의 ‘Conceptual framework for text filtering’에 의하면 내용기반 추천시스템은 사용자의 개인정보를 기반으로 개인마다의 유형을 추출하여 이에 따른 선호도를 추정하는 것이 특징이었다.⁴⁾ 또한 사용자 개인마다의 주관적인 정보를 이용하여 선호도를 살펴본 것이 특징인데, 자신이 입력한 프로필의 양이 적을 때 좁은 정보의 틀 안에 갇히게 될 우려가 있기 때문에, 적은 양의 입력으로 효과적인 추천을 원하는 사용자들에게는 정보 획득의 용이성이 떨어진다는 것을 알 수 있다.

협업 필터링을 통한 영화 추천 시스템 방식은 Sarwar, B., G. Karypis, J. Konstan, and J.Riedl(2001), Adomavicius, G. and A. Tuzhikhin(2003), Li, P. and S. Yamada(2004) 등에 의해 연구 되어왔고, 사용자 정보와 유사한 정보를 가진 집단이 선택한 것들을 추천해 준다는 것이 특징이다.⁵⁾ 따라서 집단 지성을 활용하기 때문에 내용기반 추천 방식보다 사회학적이라 할 수 있다. 그리고 사용자 개인의 선호도 정보를 토대로 그와 유사한 의견을 보인 사람들의 정보를 반영한다는 점은 큰 장점으로 작용한다. 하지만 협업적 필터링을 활용해 추천을 하려면 일정 수준의 정보가 축적되어야만 유사도의 필터링이 가능하기 때문에 정보가 희박할 때 추천의 정확도에 문제가 발생한다는 단점이 있다.⁶⁾⁷⁾ 앞서 영화 추천에 중점적으로 고려할 요소로 주관적 정보의 전달과 사용자의

4) Oard, D.W., Marchionini, G. Conceptual Framework for Text Filtering, Technica Report CS-TR3643, 1996.

5) Sarwar, B., Karypis, G., Konstan, J., Riedl, J. Item-based Collaborative Filtering Recommendation Algorithms, Proceedings of the 10th International World Wide Web Conference, 2001, pp.285-295.

6) Li, P., Yamada, S. “A Movie Recommender System Based on Inductive Learning,” IEEE Conf. Cybernetics

동기 파악을 언급했었다. 협업 필터링 추천 방식 역시 다수의 주관적 정보와 사용자의 주관적 정보를 매치하여 추천하는 방식을 취하고 있는 것이다. 그러나 사용자의 주관적 정보가 없거나 매우 적은 경우에는 정보의 매치가 어려워 사실상 추천이 불가하다는 점이 이 방식의 문제점 중 하나로 판단된다.

2-2. 감정어휘

한글 텍스트 기반 감정언어 분류 및 활용에 관한 연구 중 이준웅, 송현주, 나은경, 김현석(2008)의 ‘정서 단어 분류를 통한 정서의 구성 차원 및 위계적 범주에 관한 연구’⁸⁾에서는 한국어 사용 말뭉치에서 도출된 정서목록을 기초로 정서 표현의 전형성에 따라 주로 사용되는 정서 단어의 집합을 규정하고, 유사성 평가를 통해 한국인이 정형적으로 사용하는 정서 용어의 분류 및 범주화를 시도함으로써 감정언어 분류를 위한 근거를 제시하였다.

그리고 온라인상의 감정언어에 대한 연구로 김명규, 김정호, 차명훈, 채수환(2009)의 ‘텍스트 문서 기반의 감성 인식 시스템’⁹⁾에서는 온라인상의 댓글에서 나타나는 감성단어 구축에 관한 시도가 이루어졌다. 이영희, 정재욱(2004)의 ‘아바타의 실시간 표정변환을 위한 감정표현 어휘 분석에 관한 연구’¹⁰⁾에서는 사용자가 입력하는 단어에 따라 아바타가 반응하는 시스템을 위해 요구되는 감정 표현 어휘들을 수량화이론III류 분석을 통해 분류, 분석이 이루어졌다.

감정어휘 공간을 나타내려는 연구로 성정연, 조광수(2013)의 ‘비주얼 햅틱 형용사의 지각적 어휘 공간 연구’¹¹⁾에서는 재질감을 표현할 수 있는 햅틱 형용사의 어휘를 도출하고 햅틱 형용사들의 관계를 MDS(Multi-Dimensional Scaling: MDS)상에 표현하였다.

하지만 본 연구에서 취급하고자 하는 데이터는 ‘영화를 보면서 느끼는 감정어휘’를 다루기 때문에 영화 추천 목적에 맞는 감정적 속성을 다양하게 반영할 수 있고, 데이터를 정제하는 과정에서 감정어휘들을 신뢰할 만한 수준으로 군집화하기 위해서 주성분 분석을 적용하기 때문에 집단지성의 질적 담론으로써의 추천 시스템으로서 그 활용성이 가장 크다고 할 수 있다. 또한 감정어휘들 간의 관계를 MDS로 나타낼 때 사용자들의 의견을 최대한 반영하고자 했다.

2-3. 상황분류

상황을 분류하기 위한 논문은 Rolf a. Zwaan, Gabriel A. Radvansky의 ‘Situation Models in Language Comprehension and Memory’¹²⁾ 및 Rolf a. Zwaan, Gabriel A. Radvansky, Amy E. Hilliard, Jacqueline M. Curiel의 ‘Constructing Multidimensional Situation Models During Reading’¹³⁾이 돋보였는데, ‘Situation Models in Language Comprehension and Memory’¹³⁾에서는 상

and Intelligent Systems (2004), pp.318-323.

7) Adomavicius, G., Tuzhikhin, A. 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, No. 6, 2005, pp.734-749.

8) 이준웅, 송현주, 나은경, 김현석. 정서 단어 분류를 통한 정서의 구성 차원 및 위계적 범주에 관한 연구. 『한국언론학보』, 52(1), 2008, 85-116.

9) 김정호, 차명훈, 채수환, 텍스트 문서 기반의 감성 인식 시스템, 감성과학회지, v.12 no.4, 2009, pp.433-442

10) 이영희, 정재욱, 아바타의 실시간 표정변환을 위한 감정 표현 어휘 분석에 관한 연구. 『디자인학연구』, 2004, pp.199-208.

11) 성정연, 조광수, 비주얼 햅틱 형용사의 지각적 어휘 공간 연구. 『디자인융복합연구』, (38), 2013, pp.117-128.

12) Rolf a. Zwaan, Gabriel A. Radvansky, ‘Situation Models in Language Comprehension and Memory’, Psychological Bulletin, Vol. 123, No. 2, 1998, pp.162-185

13) Rolf a. Zwaan, Gabriel A. Radvansky, Amy E. Hilliard, Jacqueline M. Curiel, ‘Constructing Multidimensional Situation Models During Reading’, Scientific studies of reading, 2(3), 1998, pp.199-220

황에 대해서 시간(Time), 공간(Space), 특정 사건을 야기하는 요인(Causation), 목적/의도(Goal), 주체(Protagonist), 심적인 어휘(Mental lexicon) 등의 6가지 요인들로 명명하여 각각의 상황들이 무엇을 의미하는 지 서술하였고, 'Constructing Multidimensional Situation Models During Reading'에서는 앞에서 언급한 6가지의 상황 요인을 가지고 사용자들에게 여러 상황들이 얽힌 5가지 종류의 지문을 읽는 실험을 통해 각각의 상황이 미치는 영향력을 다중회귀분석을 통해 측정하는 실험을 하였다.

그러나 Rolf a. Zwaan, Gabriel A. Radvansky(1998) 및 Rolf a. Zwaan, Gabriel A. Radvansky, Amy E. Hilliard, Jacqueline M. Curriel(1998)의 연구는 여러 가지 상황 요인에 대해서 분류작업을 하고, 각각의 상황들이 미치는 영향을 실험한 논문이기 때문에, 특정한 상황 요인 하나에 대해서 여러 개의 상황을 세분화하지는 않았다는 것을 알 수 있다.

하지만 본 연구에서는 상황에 대한 요인을 정의하는 것 이외에도 각각의 상황요인 속에 어떠한 항목이 있는 지에 대해서 세분화하여 정의했기 때문에 영화를 추천하는 데에 있어서 좀 더 다양한 상황을 반영할 수 있다는 장점이 있다. 또한 영화 리뷰에서 나올 수 있는 감정어휘를 통해 영화의 감정 분포를 분석하여 맵을 생성하고 이를 상황 별 이용 동기와 결합하여 영화를 추천하는 방법을 제안하기 때문에 이 방법은 메타데이터를 이용하여 사용자의 개별적 욕구를 충족시켜 준다는 점에서 의미가 있다고 할 수 있다. 그리고 영화 데이터에 대한 리뷰들을 데이터베이스화되기 때문에 시간이 지나면 누적되는 리뷰에 따라서 보다 정확한 추천을 기대할 수도 있으며, 상황 별 이용자의 동기와 어울리는 영화들을 연계해서 다양하게 추천해 줄 수 있다.

3. 연구방법 및 분석결과

3-1. 영화 관련 감정어휘 분포맵 제작

영화 추천 시스템을 개발하기 위해 본 연구에서는 기존의 문헌 연구들을 토대로 영화와 관련된 감정 어휘들을 추출하고, 상황 별 감정 어휘들의 빈도를 시각화하기 위해 감정 어휘들을 2차원 분포맵으로 나타내었다.

3-1-1. 감정어휘의 수집

본 연구에서는 감정어휘의 수집을 위하여 두 종류의 수집과정을 거쳤다. 첫 번째는 영화 리뷰의 감정어휘 수집을 자동화하기 위해 데이터를 수집할 수 있는 웹 크롤러를 제작하였다. 크롤러는 네이버 영화 홈페이지에서 특정 영화의 댓글과 리뷰들을 정제되지 않은 데이터 형태로 수집하는 단계와 수집된 데이터를 연구에서 사용 가능한 데이터로 가공하는 단계, 마지막으로 정제된 데이터를 분석하여 감정어휘를 추출해내는 단계로 설계되었다. 크롤러를 통하여 수집되는 감정어휘는 영화를 보는 상황과 연결시켜 추후에 사용자의 이용 동기에 맞는 영화를 추천할 때도 사용된다.

두 번째는 감정어휘의 분포맵을 제작하기 위해 한덕용, 강혜자(2000)의 한국어 정서 용어들의 적절성과 경험 빈도에 관한 연구를 참고하여 834개의 정서용어 중에서 영화를 봤을 때 느낄 수 있는 감정어휘만을 분류하는 작업을 거치게 되었다.¹⁴⁾ 이 작업을 위해 본교의 국어국문학과박사 전문가 1명과 본 연구를 진행하는 연구원 2명과 함께 서로 의견취합이 가능한 감정어휘만을 골라 최종 100개의 감정어휘를 선별하게 되었다.

또한 전문가 분석을 통한 감정어휘 선별작업 이외에도 영화를 봤을 때 사용자들이 가장 많이 느끼는 감정어휘를 선별하기 위해 선정된 100개의 감정어휘를 토대로 최종감정어휘 선정을 위한 서베이

14) 한덕용, 강혜자, 한국어 정서 용어들의 적절성과 경험 빈도, 한국 심리학회지: 일반, Vol.19, 2000, pp.78-98.

를 실시하였다. 서베이는 본교의 미디어학과 학생 30명을 대상으로 영화를 봤을 때 느낄 수 있는 감정에 대한 간단한 개념 설명을 거친 뒤에, 전문가 분석을 통해 얻어진 100개의 감정어휘에 대해서 영화를 보는 상황일 때 해당 감정어휘를 느낄 수 있는 정도가 어떻게 되는 지를 조사하였다. 실제 설문에서는 ‘여러분이 지금까지 보신 여러 장르의 영화 스토리를 생각하신 뒤 해당 영화를 봤을 때 다음에 제시된 감정어휘들을 느끼는 정도가 어떠한 지 답하시기 바랍니다.’와 같이 시작하였으며, 각 감정어휘에 대해서 리커트 7점 척도(Likert-type scale)로 응답하게 하여 1점은 ‘전혀 관련 없다.’를 의미하고 7점은 ‘매우 관련 있다.’를 의미하도록 질문하였다.

본 연구에서는 사용자의 이용 동기를 이용해 영화를 추천한다는 취지에 부합하고자 영화를 봤을 때 가장 잘 느낄 수 있는 감정어휘를 수집하기 위해 전문가 분석 및 사용자 서베이를 실시하였으며, 사용자들이 설문한 리커트 7점 척도 정보를 토대로 관련성이 높은 감정어휘를 선별하기 위해 평균분석을 통해서 평균이 상대적으로 낮은 감정어휘(4.00 ‘보통이다.’를 뜻하는 수치 이하) 32개를 추가적으로 제거하여 영화 추천에 적합한 68개의 감정어휘를 최종 선정하게 되었다.

〈표 1〉 감정어휘 분포맵 제작을 위해 선정된 최종 68개의 감정어휘

68개의 감정어휘					
가련하다	가슴아프다	감격스럽다	감동적이다	거북하다	겁나다
격분하다	경악하다	경이롭다	경쾌하다	공포스럽다	굉장하다
구역질나다	그립다	기겁하다	기쁘다	끔찍하다	나른하다
놀라다	늘어지다	달콤하다	담담하다	대단하다	더럽다
두렵다	등골이 서늘하다	무료하다	무섭다	무시무시하다	분노하다
분하다	불결하다	불쌍하다	불쾌하다	불편하다	서글프다
서럽다	섬뜩하다	소름끼치다	속상하다	슬프다	신경질나다
신나다	쓸쓸하다	안타깝다	암울하다	애석하다	애절하다
역겹다	오싹하다	우울하다	유쾌하다	으스스하다	잔인하다
재미있다	즐겁다	증오스럽다	지루하다	징그럽다	차분하다
통쾌하다	평온하다	행복하다	혐오스럽다	환상적이다	활기 있다
황홀하다	흥겹다				

3-1-2. 감정어휘의 분류, 군집화

감정어휘의 수집과정에서 선정된 68개의 감정어휘를 가지고 2차원 분포맵을 제작하기 위해서는 선정된 68개의 감정어휘들이 어떠한 성격으로 나뉘는 지 분류하고, 이를 군집화 하는 과정이 필요하다. 따라서 본 연구에서는 68개의 감정어휘에 대해 주성분 분석(PCA: Principal component analysis)을 실시하였다.

주성분 분석을 위한 서베이 대상자는 최종감정어휘 선정을 위해 시행된 서베이 대상자와 동일하며 특정 영화 장르가 표출하는 대표적인 감정에 대해서 본 연구에서 선정된 68개의 감정어휘가 어떠한 관련성을 갖고 있는지를 리커트 7점 척도(Likert-type scale)로 응답하게 하였다. 그런 다음 응답한 데이터를 토대로 주성분 분석을 실시하였으며, 요인구조의 회전은 직각회전법(Varimax)을 사용하였다. 그리고 요인추출과정의 평가기준은 요인적재량(factor loading)이 0.6 이상인 것을 선택하였으며, 아이겐 값(Eigen value)을 기준으로 1 이상인 요인에 대해서만 요인화를 하였다.¹⁵⁾¹⁶⁾ 또한 응답대상이 된 68개의 감정어휘 중에서 특정 대표어 범주와의 관련성에 대한 리커트 척도 평균값이 6.5점 이상이고 분산값이 0에 근접하여 주성분 분석에 오류를 범할 수 있는 감정어휘 17개를 제외하고 51개의 변수를 대상으로 주성분을 결정하였다. 이 과정에서 제외된 17개의 감정어휘는 51개의 감정어휘에 대한 주성분 분석이 끝난 뒤에 만들어진 대표어 군집을 토대로 따로 분석을 하여 군집화 범주

15) Challaglla, Goutam N. & Shervani, Tasaddug A., Dimensions and Types of Supervisory Control: Effects on Salesperson Performance and Satisfaction, Journal of Marketing 60(1), 1996, pp.89-105

16) <http://1206sun.blog.me/130100667620>

에 추가시켰다.

55개의 감정어휘에 대해서 주성분 분석을 한 결과 아이겐 값이 1 이상이 된 성분이 총 7요인으로 도출되었다. 7가지의 요인에 대한 연구모형 적합도는 표준형성 적절성의 측도(Kaiser-Meyer-Olkin: KMO 측도)가 0.652로 측정되어 비교적 적절한 모형을 제시하고 있다. 또한 요인범주 내에서 속성들의 결합이 내적일관성(Internal consistency)을 갖고 있는지 알아보기 위해 내적일관성 검증을 실시하여 Cronbach's α 계수를 제시하였다.¹⁷⁾ 그 결과 7가지의 요인에 대해서 Cronbach's α 의 계수가 최소 0.717에서 최대 0.917을 나타내고 있기 때문에 각 요인에 대해 신뢰도가 충족됨을 알 수 있다. 이는 사회과학분야에서 일반적으로 Cronbach's α 의 계수가 0.6 이상이 될 때에는 신뢰성이 검증된다는 사실에 의거했다.¹⁸⁾ 주성분 분석에 대한 결과표는 아래의 <표 2>와 같다.

<표 2> 감정어휘의 분류 및 군집화를 위한 주성분 분석

감정 어휘	Factors							Cronbach's α
	1	2	3	4	5	6	7	
우울하다	.930							.756
가련하다	.919							
애절하다	.877							
속상하다	.862							
애석하다	.847							
쓸쓸하다	.821							
불쌍하다	.808							
그립다	.773							
안타깝다	.741							
암울하다	.702							
재미있다		.878						.819
흥겹다		.866						
황홀하다		.821						
유쾌하다		.820						
활기있다		.786						
달콤하다		.732						
경쾌하다		.717						
통쾌하다		.699						
겁나다			.868					.917
무섭다			.851					
오싹하다			.841					
무시무시하다			.786					
소름끼치다			.726					
섬뜩하다			.707					
등골이서늘하다			.692					
으시시하다			.675					
구역질나다				.818				.743
혐오스럽다				.784				
불결하다				.755				
불쾌하다				.739				
끔찍하다				.717				
더럽다				.701				
징그럽다				.682				
잔인하다				.670				
광장하다					.769			.779
경이롭다					.757			
감격스럽다					.741			
경악하다					.733			

17) 성정연, 조광수, 비주얼 햅틱 형용사의 지각적 어휘 공간 연구, Design Convergence Study 38 Vol.12. no.1, 2013. pp.123-125

18) Cortina, J.M , What is coefficient alpha? An examination of theory and applications. Journal of Applied Psychology, 78, 1993, pp.98-104.

감동적이다	.712	
대단하다	.701	
환상적이다	.688	
기겁하다	.626	
나른하다	.888	.727
늘어지다	.852	
평온하다	.839	
담담하다	.802	
차분하다	.774	
분하다	.840	.717
신경질나다	.818	
증오스럽다	.793	
불편하다	.703	

그 결과 주성분 분석에 의해서 도출된 7가지 요인에서 가장 높은 값을 보인 감정어휘는 각각 ‘우울하다’, ‘재미있다’, ‘겁나다’, ‘구역질나다’, ‘굉장하다’, ‘나른하다’, ‘분하다’로 나타났는데, 이러한 감정어휘들을 대표어로 하여 51개의 감정어휘들을 분류할 수 있게 되었다.

따라서 본 연구에서는 7가지의 대표어 요인 범주로 분류된 51개의 감정어휘들과 해당 대표어 범주로서 분류가 가능한 17개의 감정어휘들을 합쳐, 최종 68개의 감정어휘들에 대해 7개의 대표어 범주로 분류 및 군집화 하는 작업을 진행하였다. 이 과정에서 7개의 대표어 카테고리의 용어를 선정하였는데 W. Parrot(2001)의 Emotion Table¹⁹⁾의 6가지 기본 감정에서 ‘잔잔한(지루한)’을 더하여 영화에 적합한 7가지의 새로운 카테고리를 통해 군집화 되었다. 대표어에 의해 최종 군집화된 68개의 감정어휘 표는 아래의 <표 3>과 같다.

<표 3> 감정어휘의 군집화

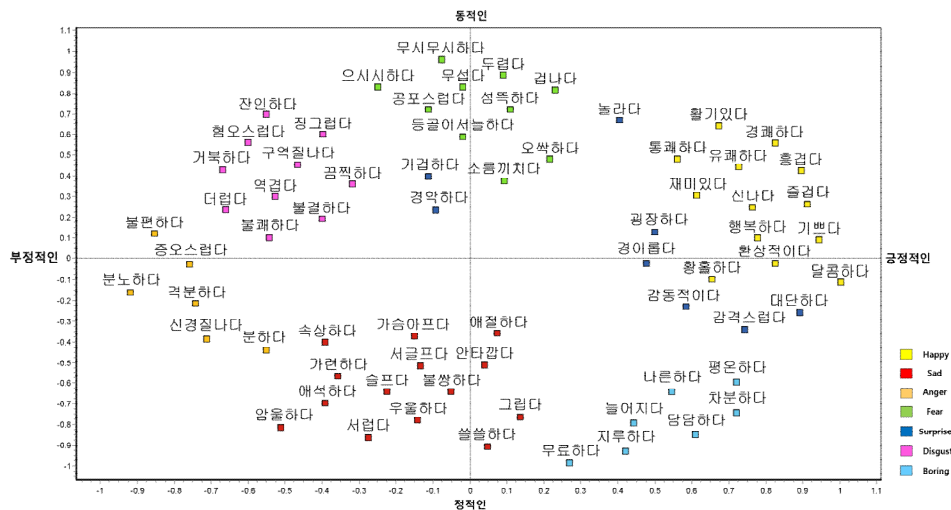
대표어	감정단어	대표어	감정단어
Happy	경쾌하다	Anger	격분하다
	기쁘다		분노하다
	달콤하다		분하다
	신나다		불편하다
	유쾌하다		신경질나다
	재미있다		증오스럽다
	즐겁다	Surprise	감격스럽다
	통쾌하다		감동적이다
	행복하다		경악하다
	활기 있다		경이롭다
	황홀하다		굉장하다
	흥겹다		기겁하다
Sad	가련하다	Disgust	놀라다
	가슴아프다		대단하다
	그립다		환상적이다
	불쌍하다		거북하다
	서글프다		구역질나다
	서럽다		끔찍하다
	속상하다		더럽다
	슬프다		불결하다
	쓸쓸하다		불쾌하다
	안타깝다		역겹다
	암울하다		잔인하다
	애석하다		징그럽다
애절하다	혐오스럽다		
우울하다	Boring	평온하다	
Fear	겁나다	Boring	나른하다

19) Parrott, W. Emotions in Social Psychology. Philadelphia: Psychology Press, 2001.

	공포스럽다		차분하다
	두렵다		담담하다
	등골이서늘하다		늘어지다
	무섭다		지루하다
	무시무시하다		무료하다
	섬뜩하다		
	소름끼치다		
	오싹하다		
	으시시하다		

3-1-3. 감정어휘 간 상관관계 분석 및 분포맵 제작

본 연구에서는 최종 군집화된 68개의 감정어휘를 토대로 유사하거나 상이한 감정어휘를 근거로 한 영화추천시스템 방법을 고안하기 위해 감정어휘 간의 거리도를 측정하여 상관관계를 분석한 다차원척도 분석(Multi-Dimensional Scaling: MDS)을 실시하였다. 다차원척도 분석이란 개체들 간의 상대적인 거리를 계산하여 사람이 인지 할 수 있는 평면상에 상대적인 거리로 나타내는 통계와 연관된 기술로 정보시각화에서는 데이터 내의 유사성/비유사성을 측정하기 위해 종종 사용된다. 다차원척도법의 장점은 상대적인 거리만을 알고 있는 개체들의 위치도를 작성할 수 있으며 물리적인 거리뿐만 아니라 심리적인 거리에 근거하여서도 위치도를 작성할 수 있다는 것이다.²⁰⁾ 다차원척도 분석을 위해 경기도 및 서울 소재 대학교 20대 남학생 11명, 여학생 9명으로 총 20명을 실험 대상으로 하여 68개의 감정어휘에 대해 의미상 거리 서베이를 실시하였는데, 서베이는 가로축 세로축 68개의 감정어휘를 배치한 설문지를 만들고(68X68) 감정어휘간의 거리가 가장 가깝다고 느껴지면 3점, 가장 멀다고 느껴지면 -3점을 주는 방식의 리커트척도를 이용하여 체크하는 형식으로 구성하였다. 20명이 기록한 데이터를 바탕으로 다양한 네트워크 분석기법이 활용 가능한 UCINET 프로그램을 사용하였고 이를 통해 영화 리뷰 감정어휘들 간의 의미상의 거리 기반으로 <그림 2>와 같은 Metric MDS를 형성할 수 있었다.



<그림 2> UCINET을 이용한 68개의 감정어휘 분포맵

그 결과 X축의 양의 방향으로서는 대표어 “Happy”, “Surprise”와 관련된 감정어휘가 분포되었으며, X축의 음의 방향으로서는 대표어 “Anger”, “Disgust”와 관련된 감정어휘가 분포되었다. 그리고 Y축의 양의 방향으로서는 대표어 “Fear”, “Surprise”와 관련된 감정어휘가 분포되었으며, Y축의 음의 방향으

20) 김민정, 다차원척도법과 MCLUST를 이용한 다변량 자료의 군집분석, 이화여자대학교, 통계학과, 2006, p.3

로는 대표어 “Sad”, “Boring”과 관련된 감정어휘가 분포되었다. 따라서 감정어휘의 성격 상 X축의 양의 방향은 긍정적인 느낌의 감정어휘들이 분포되었고, X축의 음의 방향으로는 부정적인 느낌의 감정어휘들이 분포됨을 알 수 있다. 또한 Y축의 양의 방향은 동적인(감정을 느낄 때 비교적 큰 제스처를 취할 수 있는) 느낌의 감정어휘들이 분포되었고, Y축의 음의 방향으로는 정적인(감정을 느낄 때 비교적 작은 제스처를 취할 수 있는) 느낌의 감정어휘들이 분포됨을 알 수 있다. 그리고 ‘Happy’, ‘Sad’, ‘Anger’, ‘Fear’, ‘Disgust’, ‘Boring’ 대표어와 관련된 어휘들은 각각의 단어가 뚜렷하게 군집이 되는 것을 볼 수 있는데, 대표어 ‘Surprise’에 대해서는 ‘Happy’ 대표어 군집과 ‘Fear’ 대표어 군집에 나뉘어서 분포한다는 것을 알 수 있다. 이는 사용자들이 영화를 봤을 때 ‘빅찬 기쁨’으로 인해 놀라운 감정이 생기는 경우’와 ‘갑작스럽게 등장하는 공포로 인해서 놀라운 감정이 생기는 경우’가 지배적이기 때문인 것으로 해석된다.

3-2. 영화를 보는 상황에 따른 감정어휘 빈도 분석

3-2-1. 영화를 보는 상황 분류

관람하고자 하는 영화를 선택하는 과정은 영화의 객관적인 정보뿐 아니라 사용자가 처한 상황과 목적에 영향을 받기도 한다. 따라서 사용자가 영화를 보려는 이유가 애인과의 데이트 때문인지 아니면 단지 혼자서 시간을 보내기 위함인지 등 상황변수에 따라 최종적으로 선택하는 영화가 다르게 나타날 수 있다. 먼저 영화를 관람하는 상황을 이루는 환경 변수를 구성하기 위하여 Zwaan의 연구를 참고하여 [표 4]와 같이 ‘시간’, ‘공간’, ‘주체’, ‘의도’의 총 네 가지의 분류로 구분하고 그에 따라 적절한 상황변수를 추가하였다.

〈표 4〉 영화를 보는 상황의 분류

분류	항목	
시간 관련 상황	시간대	오전, 오후, 심야
	날짜	평일, 공휴일, 영화개봉일, 이성친구와의 기념일, 가족과의 기념일, 명절
공간 관련 상황	영화관, 자택, DVD방, 교통수단(자동차, 기차, 비행기 등)	
주체 관련 상황	부모님 또는 가족, 애인, 동성친구, 이성친구, 혼자	
의도 관련 상황	킬링타임, 여가활동, 데이트, 기분전환, 교육(학습), 친구의 권유	

3-2-2. 대표영화 선정 및 영화별 추천 상황조사

시간, 공간, 주체, 의도로 그룹화된 상황들은 그 조합을 달리함에 따라 영화를 보는 동기가 달라질 수 있다. 이러한 동기를 파악하기 위해 각 상황마다 어떤 감정을 느끼고 싶어 하는지에 대한 서베이를 진행하였다. 조사대상 영화를 선정하기 위해 포털사이트의 영화 순위를 참고로 개봉 연도와 상관없이 9개의 장르(가족, 드라마, 멜로/애정/로맨스, SF, 공포, 코미디, 스릴러, 판타지/어드벤처, 애니메이션)별 상위 20위까지 총 180개의 영화 가운데 서베이 참여자를 대상으로 영화를 보았다는 응답이 가장 많은 20개의 영화를 최종적으로 선정하였다. 사용자에게 선정된 20개의 영화를 제시하고 각각의 영화에 가장 추천할만한 상황변수를 선택하도록 하였다. 추천 상황의 신뢰도를 높이기 위하여 실제로 관람한 영화에만 응답하도록 하였다. 서베이는 경기도 소재 대학의 대학생 60명을 대상으로 진행되었다. 서베이 결과 영화를 보는 상황을 분류하는 요인인 ‘시간’, ‘공간’, ‘주체’, ‘의도’ 등과 관련된 상황 중에서, 대부분의 응답이 ‘주체’ 및 ‘의도’와 관련된 상황에 대해서만 각각의 영화샘플에 따라 다양한 의견을 보임을 알 수 있었다. ‘시간’에 관련된 사항은 영화의 종류에 상관없이 대부분 ‘오후’ 및 ‘평일’, ‘공휴일’을 선택했기 때문에 제외하였고, ‘공간’에 관련된 사항도 마찬가지로 영화의 종류에 상관없이 대부분 ‘자택’을 선택한 관계로 제외하였다. 여기에서는 ‘주체’ 및 ‘의도’와 관련된 상황에 대해서 뚜렷하게 상이한 결과를 보인 ‘아이언맨 3’과 ‘타이타닉’에 대한 예시를 제시하고자 한다. 그에 대한 결과는 아래의 〈표 5〉와 같다.

〈표 5〉 영화 “아이언맨 3”의 상황추천결과와 영화 “타이타닉”의 상황추천결과

영화	요인	항목	계	가장높은항목
아이언맨 3	영화를 함께 보는 대상	가족	0	동성친구
		애인	3	
		동성친구	15	
		이성친구	0	
		혼자	1	
	영화를 보는 의도	기타	2	여가활동,기분전환
		킬링타임	5	
		여가활동	9	
		데이트	1	
		기분전환	8	
타이타닉	영화를 함께 보는 대상	교육	0	애인
		기타	0	
		가족	1	
		애인	20	
		동성친구	0	
		이성친구	2	
	영화를 보는 의도	혼자	0	데이트
		기타	0	
		킬링타임	2	
		여가활동	1	
		데이트	18	
		기분전환	0	
		교육	0	
기타	0			

〈표 5〉를 보면 ‘아이언맨 3’은 주로 ‘동성친구’라는 주체적 상황 및 ‘여가활동’, ‘기분전환’이라는 의도적 상황일 때 해당영화를 추천하면 좋을 것이라는 결과가 나왔고, ‘타이타닉’은 ‘애인’이라는 주체적 상황 및 ‘데이트’라는 의도적 상황일 때 해당영화를 추천하면 좋을 것이라는 결과가 나오게 되었다. 이처럼 본 연구에서 서베이 대상이 된 영화의 장르에 따라서 실험자들은 ‘주체적 상황’ 및 ‘의도적 상황’에 대해 비교적 다양한 의견을 보인다는 것을 알 수 있었다. 하지만 〈표 5〉에서 나타나는 ‘상황에 대한 빈도 표’ 만으로는 사용자들의 ‘상황조합에 따른 만족도’를 뒷받침하는 게 부족하다는 결론이 나오게 되었다. 따라서 본 연구에서는 ‘각 상황조합에 대한 만족도를 조사하면 〈표 5〉에서 높은 값을 나타낸 항목이 만족도 점수가 높게 나타날 것이다.’라는 가설을 전제하여 ‘아이언맨 3’, ‘타이타닉’, ‘엽기적인 그녀’ 등의 영화샘플을 가지고 ‘주체적 상황’ 및 ‘의도적 상황’에 대한 사용자 만족도 서베이를 실시하게 되었다. 여기에서 ‘엽기적인 그녀’는 영화를 봤을 때 느끼는 감정어휘와 연결했을 때 로맨스 장르로써는 서로 겹치는 ‘타이타닉’과 어떤 차이점을 보이는 지 연구하기 위해 추가되었다. 서베이는 경기도 및 서울 소재 대학의 대학생 50명(남자 25명, 여자 25명)을 대상으로 진행되었으며, 설문에 대상이 된 3개의 영화를 이용하여 본 연구에서 정의한 각각의 상황 항목을 조합하여 문항을 만들고, 실험자들에게는 해당 조합 문항을 읽고 느끼는 만족도가 어떠한 지를 테스트 하였다. 실제 설문에서는 ‘귀하는 본 영화를 본 적이 있습니까?’와 같은 문항을 통해 영화 관람 여부를 물어보는 문항을 작성하게 하였으며, ‘본 적이 없다.’ 라고 응답한 실험자들에게는 해당 영화의 예고편 및 영화정보를 볼 수 있는 사이트 링크를 제공하여 해당 영화를 이해하도록 했다. 그런 다음에는 상황 요인(영화를 함께 보는 대상: 가족, 애인, 동성친구, 이성친구, 혼자 / 영화를 보는 의도: 킬링타임, 여가활동, 데이트, 기분전환, 교육목적)에 대한 조합 문항을 “상황 1. 가족과 단순히 시간을 보내기 위해 이 영화를 본다면.”, “상황 2. 애인과 데이트를 위해 이 영화를 본다면” 과 같이 1편의 영화 당 총 24문항 제작하였으며, 각 문항에 대해서 리커트 7점 척도(Likert-type scale)로 설정해 1점은 ‘매우 불만족’을 의미하고 7점은 ‘매우 만족’을 의미하도록 질문하였다.

서베이 결과를 바탕으로 3개의 영화에 대해서 사용자들이 만족할 만한 상황을 도출해 내기 위해 각

항목에 대한 통계분석을 실시하였다. 우선 각 문항에 대하여 신뢰도를 분석한 결과 신뢰도 통계량의 Cronbach' α 계수가 0.928로써 모든 문항의 모형이 신뢰도가 있음이 검증되었고, 각 항목을 제거했을 때 Cronbach' α 계수가 0.928보다 높게 나오게 되는 문항 5개는 제거를 하고 통계분석을 하였다. 본 연구에서는 사용자가 만족할 만한 상황 도출을 위해 리커트 7점 척도 중에서 5점에 해당하는 '만족'보다 평균값이 높은(즉, 평균값이 5.00점 이상이 나오는) 상황 문항을 선별하였다. 그에 대한 결과표는 아래의 <표 6>과 같다.

<표 6> 대표영화에 대한 상황 만족도 평균분석 표 (평균 5.00점 이상, 표본 N=50)

대표영화	선별된 상황 문항	평균	표준편차
아이언맨 3	상황 11: 동성친구와 기본전환을 위해 이 영화를 본다면	5.82	1.224
	상황 9: 동성친구와 단순하게 시간을 보내기(킬링타임)위해 이 영화를 본다면	5.78	1.148
	상황 10: 동성친구와 여가활동을 위해 이 영화를 본다면	5.66	1.136
	상황 1: 가족과 단순하게 시간을 보내기(킬링타임)위해 이 영화를 본다면	5.20	1.262
	상황 2: 가족과 여가활동을 위해 이 영화를 본다면	5.18	1.082
타이타닉	상황 7: 애인과 데이트를 위해 이 영화를 본다면	6.28	.784
	상황 6: 애인과 여가활동 위해 이 영화를 본다면	6.04	.925
	상황 8: 애인과 기본전환을 위해 이 영화를 본다면	5.70	1.282
	상황 5: 애인과 단순하게 시간을 보내기(킬링타임)위해 이 영화를 본다면	5.68	.913
	상황 13: 이성친구와 여가활동을 위해 이 영화를 본다면	5.22	1.183
	상황 12: 이성친구와 단순히 시간을 보내기(킬링타임)위해 이 영화를 본다면	5.20	1.229
엽기적인 그녀	상황 9: 동성친구와 단순하게 시간을 보내기(킬링타임)위해 이 영화를 본다면	5.16	1.695
	상황 8: 애인과 기본전환을 위해 이 영화를 본다면	5.62	1.276
	상황 7: 애인과 데이트를 위해 이 영화를 본다면	5.44	1.459
	상황 5: 애인과 단순하게 시간을 보내기(킬링타임)위해 이 영화를 본다면	5.26	1.440
	상황 6: 애인과 여가활동 위해 이 영화를 본다면	5.20	1.125
	상황 11: 동성친구와 기본전환을 위해 이 영화를 본다면	5.12	1.913
	상황 13: 이성친구와 여가활동을 위해 이 영화를 본다면	5.10	1.344

사용자의 만족도 조사 결과, 영화 '아이언맨 3'는 동성친구와 같이 기본전환, 킬링타임, 여가활동을 위한 상황 순으로 가장 높은 만족도를 나타내었고, 그 뒤로는 가족과 킬링타임 또는 여가활동을 위한 상황이 높은 만족도를 나타내었다. 그리고 영화 '타이타닉'은 애인과의 데이트, 여가활동, 기본전환 등을 위한 상황 순으로 가장 높은 만족도를 나타내었고, 이성친구와 여가활동, 킬링타임을 위한 상황 및 동성친구와 킬링타임을 위한 상황 순으로 높은 만족도를 나타내었다. 마지막으로 영화 '엽기적인 그녀'는 애인과의 기본전환, 데이트, 킬링타임, 여가활동을 위한 상황 순으로 가장 높은 만족도를 나타내었고, 그 뒤로는 동성친구와 기본전환을 위한 상황 및 이성친구와 여가활동을 위한 상황 순으로 높은 만족도를 나타내었다. 따라서 영화 '아이언맨 3' 및 '타이타닉'에 대해서는 대체적으로 <표 5>의 결과와 비슷한 만족도 점수가 나오게 됨을 알 수 있었다. 하지만 <표 5>의 결과를 통해서 예상한 것과는 달리 영화 '타이타닉'에서는 '동성친구와 킬링타임'을 위한 상황 만족도가 비교적 높게 나왔고, 이와 비슷한 장르인 '엽기적인 그녀'에서도 '동성친구와 기본전환'을 위한 상황 만족도가 비교적 높게 나왔음을 알 수 있었다. 따라서 본 연구에서는 이와 같은 결과가 나오게 된 원인을 분석하기 위해 설문 대상이 된 실험자들의 성별에 따라서 해당 문항에 대한 평균비교를 위한 독립표본 t-검증을 시행하였다. 그 결과 <표 7>과 같이 영화 '타이타닉'의 '동성친구와 단순하게 시간을 보내기(킬링타임)위해 이 영화를 본다면'이라는 문항에 대해서는 남자의 평균이 3.76, 여자의 평균이 6.56으로 나타났고, 해당 문항은($t=-10.493, p<.001$) 성별에 따라 평균값이 통계적으로 유의미한 차이가 있는 것으로 나타났다. 또한 영화 '엽기적인 그녀'의 '동성친구와 기본전환을 위해 이 영화를 본다면'이라는 문항에 대해서는 남자의 평균이 3.96, 여자의 평균이 6.28으로 나타났고, 이 문항에 대해서도($t=-5.370, p<.001$) 성별에 따라 평균값이 통계적으로 유의미한 차이가 있는 것으로 나타났다. 이는 남성의 경우 여성과 달리 동성친구와 '로맨스 및 연애'영화를 관람하는 것을 기피하는 현상

이 가장 큰 원인으로 해석된다.

〈표 7〉 성별에 따른 상황 만족도 t-검증 (타이타닉 상황 9: 동성친구-킬링타임, 엽기적인 그녀 상황 11: 동성친구-기분전환)

대표영화	선별된 상황 문항	평균		표준편차		t값	유의확률
		남자(N=25)	여자(N=25)	남자(N=25)	여자(N=25)		
타이타닉	상황 9: 동성친구와 단순하게 시간을 보내기(킬링타임)위해 이 영화를 본다면	3.76	6.56	1.128	.712	-10.493	.000***
엽기적인 그녀	상황 11: 동성친구와 기분전환을 위해 이 영화를 본다면	3.96	6.28	1.837	1.137	-5.370	.000***

*p<.05, **p<.01, ***p<.001

따라서 본 연구에서는 서베이 결과로 도출된 만족도 결과를 기반으로 3개의 영화에 대한 대표적인 추천 상황을 통계분석을 통해 선정할 수 있게 되었다. 또한 같은 상황이라도 성별에 따라서 서로 다른 만족도를 보인다는 것을 확인할 수 있었다.

3-2-3. 상황별 추천 감정어휘 빈도분포

상황별 감정어휘 빈도분포를 구하기 위하여 앞서 선정한 각 대표영화에 대한 인터넷 리뷰를 수집하였다. 영화 당 평균 1만개의 리뷰가 수집되었으며, 이 리뷰 데이터를 토대로 각각의 감정어휘가 얼마나 나오는지 확인하였다. 상황에 따라 감정어휘 빈도분포가 다르게 나오는 것을 확인하기 위해 3가지의 서로 다른 가상 상황을 설정하였는데, 각각의 상황은 본 연구에서 시행한 ‘상황 만족도 서베이’ 결과를 이용하여 3개의 영화에 대해 가장 높은 평균값을 가진 항목을 선택하였다. 그 결과 〈표 8〉을 보면 동성친구와 기분전환을 위한 상황 (가)에 대해서는 감정어휘 종류가 ‘Happy’, ‘Surprise’, ‘Boring’ 순으로 높게 분포되었고, 애인과 데이트를 위한 상황 (나)에 대해서는 ‘Sad’, ‘Happy’, ‘Surprise’, ‘Boring’ 순으로 높게 분포 되었다. 그리고 애인과 기분전환을 위한 상황 (다)에 대해서는 ‘Happy’, ‘Sad’, ‘Surprise’ 순으로 감정어휘가 높게 분포되었다.

〈표 8〉 서로 다른 상황에서의 감정 어휘 빈도 분포

상황	대표영화	감정	절대 빈도	상대 빈도(%)
(가) ‘동성친구’ ‘기분전환’	아이언맨 3	Happy	3456	71.99
		Sad	60	1.25
		Fear	69	1.44
		Anger	52	1.08
		Surprise	729	15.18
		Disgust	14	0.29
		Boring	421	8.77
(나) ‘애인’ ‘데이트’	타이타닉	Happy	494	23.73
		Sad	1070	51.39
		Fear	13	0.62
		Anger	10	0.48
		Surprise	344	16.52
		Disgust	9	0.43
		Boring	142	6.82
(다) ‘애인’ ‘기분전환’	엽기적인 그녀	Happy	562	76.15
		Sad	109	14.77
		Fear	1	0.14
		Anger	3	0.41
		Surprise	42	5.69
		Disgust	1	0.14
		Boring	20	2.71

여기에서 상황 (가)와 상황 (다)의 경우 영화를 보고자 하는 의도는 ‘기분전환’으로써 서로 같으나 주체가 동성친구에서 애인으로 바뀔 때 따라 ‘Happy’, ‘Sad’ 감정어휘 분포의 상대적 빈도는 증가하고 ‘Fear’, ‘Anger’, ‘Surprise’, ‘Disgust’, ‘Boring’ 감정어휘 분포의 상대적 빈도는 감소함을 알 수 있다. 또한 상황 (나)와 상황 (다)를 비교했을 때 주체는 서로 같으나 영화를 보고자 하는 의도가 데이트에서 기분전환으로 바뀔 때 따라 ‘Happy’ 감정어휘 분포의 상대적 빈도는 확연히 증가하고 ‘Sad’, ‘Fear’, ‘Surprise’, ‘Disgust’, ‘Boring’ 감정어휘 분포의 상대적 빈도는 감소하며 ‘Anger’ 감정어휘 분포의 상대적 빈도는 거의 변화가 없음을 알 수 있다. 이러한 결과는 영화를 보는 상황 요인이 변화함에 따라 감정어휘 분포의 빈도가 확연히 변화할 수 있다는 점을 검증한 결과이며 이는 영화의 정보와 상황이 ‘감정어휘’를 기준으로 연결되어 상황마다 어떠한 영화를 추천해야 하는지 근거를 제공해 주었다고 할 수 있다. 따라서 본 연구에서 제안하는 영화추천시스템은 사용자가 특정 상황을 입력하게 되면 해당 상황에서 일반적으로 느끼길 원하는 감정 어휘 분포와 가장 유사한 감정어휘 분포를 가지고 있는 영화를 추천할 수 있게 된다.

4. 결론

본 연구는 영화를 보면서 느낄 수 있는 감정어휘를 선정하고 전문가의 분석 및 서베이를 통해 감정어휘간의 상관관계를 분석하여 다차원적도분포맵을 제작하였다. 그런 다음 영화를 추천할 수 있는 각각의 상황을 구하고 영화 리뷰에서 감정어휘 빈도를 계산해냄으로써 사용자가 영화를 보려는 상황에 가장 적합한 감정어휘 빈도를 보이는 과정을 살펴보았다. 이러한 일련의 과정은 영화 추천 시스템에 대해 영화를 보는 상황과 그에 따른 감정어휘를 새로운 기준으로 제안함으로써 기존의 장르 중심의 영화추천 시스템 연구들과는 차별성을 가진다고 할 수 있다. 또한 기본적인 영화의 프로파일 정보를 제외한 채 영화를 볼 때 느끼는 감정어휘 및 상황 등의 이용 동기만으로도 사용자의 취향에 적합한 영화를 추천할 수 있다는 가능성을 제시했다고 판단된다.

하지만 본 연구에서 조사대상으로 선정한 영화의 표본 수가 적었고, 다양한 연령대 계층을 조사하지 못했기 때문에 일반화하기에 다소 조심스러운 점이 있다. 그러나 향후에 본 연구에서 만들게 되는 영화 추천시스템이 스마트앱(Smart App)을 기반으로 만들어질 계획이기 때문에 예상되는 주 사용층이 20대임을 고려한다면 조사한 표본만으로도 상황의 변화에 따라서 감정어휘의 빈도가 달라짐을 보였다는 것은 본 연구에 있어서 가치 있는 결과라 할 수 있겠다. 향후에는 영화 추천 방식의 체계적 검증을 위해서 다양한 영화 표본을 기반으로 한 감성 적합도를 측정하고, 성별 이외에도 연령, 기간과 관련한 서베이를 실시하여 보다 정교한 추천시스템을 만들고자 한다. 또한 상황과 감정의 유비관계를 알아내기 위하여, 통계적 방법을 통하여 특정 상황에 사람들이 어떤 감정들을 가질 수 있는가를 체계적으로 알아볼 수 있는 실험을 설계할 계획이다. 그리고 분류된 각각의 영화를 보는 특정 상황에서 선호되는 영화의 리뷰 내에 선별된 감정단어들이 얼마나 자주 중요하게 사용되는지 알아보기 위해 TF-IDF(Term Frequency - Inverse Document Frequency) 값을 통해 분석해보고 7개의 대표어에 대한 각각의 가중치를 추출하여 상황 별 감정어휘 맵을 시각화하고, 맵의 분포 패턴을 비교하여 유사한 패턴을 가진 리뷰의 영화를 추천할 계획이다.

참고문헌

도서

- Parrott, W. Emotions in Social Psychology. Philadelphia: Psychology Press, 2001.

논문

- 김민정, 다차원척도법과 MCLUST를 이용한 다변량 자료의 군집분석, 이화여자대학교, 통계학과, 2006, p.3
- 김정호, 차명훈, 채수환, 텍스트 문서 기반의 감성 인식 시스템, 감성과학회지, v.12 no.4, 2009, pp.433-442
- 성정연, 조광수, 비주얼 햅틱 형용사의 지각적 어휘 공간 연구. 『디자인융복합연구』, (38), 2013, pp.117-128.
- 이영희, 정재욱, 아바타의 실시간 표정변환을 위한 감정 표현 어휘 분석에 관한 연구. 『디자인학 연구』, 2004, pp.199-208.
- 이재신, 노동환, 시청동기가 미디어 선택과 이용에 미치는 영향 : 대학생들의 미국드라마 소비를 중심으로, 한국방송학보 22(5), 한국방송학회 (2008), 246-286
- 이준웅, 송현주, 나은경, 김현식. 정서 단어 분류를 통한 정서의 구성 차원 및 위계적 범주에 관한 연구. 『한국언론학보』, 52(1), 2008, 85-116.
- 한덕용, 강혜자, 한국어 정서 용어들의 적절성과 경험 빈도. 한국 심리학회지: 일반, Vol.19, 2000, pp.78-98.
- 한상필, 안보섭, 박명진, 무선 인터넷 이용자의 모바일 광고 이용 동기에 관한인식 연구, 커뮤니케이션학연구, Vol. 13, No. 4, 2005, pp.68-100.
- Adomavicius, G., Tuzhihin, A. 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, No. 6, 2005, pp.734-749.
- Challaglla, Goutam N. & Shervani, Tasaddug A., Dimensions and Types of Supervisory Control: Effects on Salesperson Performance and Satisfaction, Journal of Marketing 60(1), 1996, pp.89-105
- Cortina, J.M , What is coefficient alpha? An examination of theory and applications. Journal of Applied Psychology, 78, 1993, pp.98-104.
- Dimmick, J., Kline, S., Stafford, L. The Gratification Niches of Personal E-mail and the Telephone. Communication Research, Vol. 27, No. 2, 2000, pp.227-248.
- Li, P., Yamada, S. "A Movie Recommender System Based on Inductive Learning," IEEE Conf. Cybernetics and Intelligent Systems (2004), pp.318-323.
- Oard, D.W., Marchionini, G. Conceptual Framework for Text Filtering, Technica Report CS-TR3643, 1996.
- Rolf a. Zwaan, Gabriel A. Radvansky, 'Situation Models in Language Comprehension and Memory', Psychological Bulletin, Vol. 123, No. 2, 1998, pp.162-185
- Rolf a. Zwaan, Gabriel A. Radvansky, Amy E. Hilliard, Jacqueline M. Curiel, 'Constructing Multidimensional Situation Models During Reading', Scientific studies of reading, 2(3), 1998, pp.199-220
- Sarwar, B., Karypis, G., Konstan, J., Riedl, J. Item-based Collaborative Filtering Recommendation Algorithms, Proceedings of the 10th International World Wide Web Conference, 2001, pp.285-295.

학술대회 및 세미나

- 안신현, 시정곤, 문화적 메타데이터에 기반한 영화추천시스템, 한국컴퓨터종합학술대회, Vol. 35, No. 1(A), 2008, p.79.
- 이경원, 하효지 (2013.5). 영화추천 시스템 개발을 위한 감정어휘 및 영화를 보는 상황의 분석. 『한국디자인학회 학술발표대회 논문집』, 78-79.

인터넷 사이트

- 요인분석 네이버 블로그 포스팅 <http://1206sun.blog.me/130100667620>