

빅데이터를 활용한 영화 흥행 분석 -천만 영화의 웃음과 눈물 요소를 중심으로

황영미^{1*} · 박진태² · 문일영² · 김광선³ · 권오영²

The Box-office Success Factors of Films Utilizing Big Data-Focus on Laugh and Tear of Film Factors

Young-mee Hwang^{1*} · Jin-tae Park² · Il-young Moon² · Kwang-sun Kim³ · Oh-young Kwon²

^{1*}School of General Education Sookmyung Women's University, Seoul 04310, Korea

²School of Computer Science and Engineering, Korea University of Technology and Education, Cheonan 31253, Korea

³School of Mechanical Engineering, Korea University of Technology and Education, Cheonan 31253, Korea

요 약

이 연구는 빅데이터를 활용하여 영화흥행 요인을 분석하는 것이 목적이다. 한국의 영화산업 규모는 날로 커지고 있지만, 현재까지 진행되어온 영화 흥행 요인 분석 및 예측과 관련된 논의는 관련 데이터를 망라하지 못해 정확성을 담보할 수 없는 상황이었다. 지금까지 한국에서의 천만 영화는 총 13편이 있었고, 이 연구에서는 천만 흥행에 눈물과 웃음이 주된 텍스트 내적 요인으로 작용함을 밝혔다. 이에 빅데이터를 활용해 영화에 대한 댓글 중 웃음과 눈물과 관련된 용어를 수집한 후, 영화의 구성 5단계(발단-전개-위기-절정-결말) 중 어느 부분에 웃음과 눈물 요소가 많은지를 도표화하여 천만 영화의 장르별 구성 방식을 논증하였다. 이러한 분석 결과는 앞으로 영화 제작 전 단계에서 시나리오 상에서의 흥행 예측을 하는 종합적인 데이터베이스 구축에 기여하게 될 것이다.

ABSTRACT

The study aims to analyze factors of box office utilizing big data. The film industry has been increasing in the scale, but the discussion on analysis and prediction of box-office hit has not secured reliability because of failing in including all relevant data. 13 films have sold 10 million tickets until the present in Korea. The study demonstrated laughs and tears as an main interior factors of box-office hit films which showed more than 10 million tickets power. First, the study collected terms relevant to laugh and tear. Next, it schematizes how frequently laugh and tear factors could be found along the 5-film-stage (exposition - Rising action - crisis - climax - ending) and revealed box-office hit films by genre. The results of the analysis would contribute to the construction of comprehensive database for the box office predictions on future scenarios.

키워드 : 빅데이터, 박스오피스, 영화흥행 내적요소, 웃음 요소, 눈물 요소

Key word : big data, box office, interior factors of box-office, laugh factor, tear factor

Received 29 April 2016, Revised 03 May 2016, Accepted 17 May 2016

* Corresponding Author Young-Mee Hwang(E-mail:hym4322@sm.ac.kr, Tel:+82-2-710-9825)

School of General Education, Sookmyung Women's University, Seoul, 04310, Korea

Open Access <http://dx.doi.org/10.6109/jkiice.2016.20.6.1087>

print ISSN: 2234-4772 online ISSN: 2288-4165

©This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License(<http://creativecommons.org/licenses/by-nc/3.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.
Copyright © The Korea Institute of Information and Communication Engineering.

I. 서 론

1.1. 연구 목적

한국의 영화산업 규모는 날로 커지고 있지만, 현재까지 진행되어온 영화 흥행 요인 분석 및 예측과 관련된 논의는 관련 데이터를 망라하지 못해 정확성을 담보할 수 없는 상황이었다.

본 연구는 근원적인 차원에서 제기될 수 있는 융합 가능성을 모색함과 동시에 궁극적으로는 융합 연구 결과가 실제 산업적 측면에서 어떻게 활용될 수 있는지를 연구하는 것이 목표이다.

본 연구는 13편의 한국 천만 영화를 대상으로 흥행 요인을 분석하고 이러한 분석 결과가 실제로 어떠한 흥행 결과를 낳게 되었는지에 대한 영화 흥행 예측 모델을 구축하는 것을 목표로 한다.

1.2. 연구 필요성

오늘날 지식기반 정보화 사회에서는 기존의 분과 학문의 급격한 재구성과 새로운 통합이 이루어지고 있다. 통합의 방향은 학문간 상호 소통, 즉 학제간 연구를 통해 융합 학문 중심의 패러다임으로 변화하고 있다. 그러나 기존의 연구는 여전히 '인접 학문 간 융합', 즉 겉으로 드러난 형식적인 유사성에 주목한 융합 연구가 많았다.

지금은 학문적 성과가 분과 학문의 영역에서만 이루어지는 것으로는 부족하며 통합적인 지식 네트워크를 구축할 필요성이 있기 때문에 제반 학문 간의 융합은 그 어느 때보다 중요한 시대적 요구가 되었다. 최근 융합연구 동향은 창의적 융합인재에 관한 개념에서부터 인문학과 예술학 분야뿐만 아니라 이공계와의 융합을 통해 산업적 측면에까지 활용 가능한 연구로 발전하고 있다. 이에 본 연구는 영화 흥행을 좌우하는 텍스트 내외적 요인에 대한 객관적 지표를 산출하고, 이를 수학 및 공학계에서 활용되고 있는 최적화 이론을 토대로 객관적으로 분석한 후, 이것이 영화 관객에게 미치는 영향 관계를 빅데이터를 활용하여 검증해나가고자 한다. 이를 위해 영화 산업, 즉 영화 흥행요소를 영화 텍스트를 둘러싼 프리 프로덕션 단계에서부터 포스트 프로덕션 단계 모두를 아울러 영화 흥행 요인을 객관적으로 분석·예측하는 데에 필요한 지식 네트워크 구축을 위한 융합적 연구가 필요하다.

기존의 영화 흥행 요인 분석 논의는 영화 제작과 산업 발전에 기여하는 정도가 두드러지지 않는다는 단점이 있다. 이는 기존의 영화 흥행 요인 분석 논의가 일부 작품들에 대해 산발적으로 진행되었고, 더욱이 텍스트 외적 요인에 집중되어 있다 보니 실제 영화가 제작되는 과정에서 필요로 하는 부분을 제대로 만족시키지 못하고 있는 데서 비롯된다. 영화 흥행 요인 분석 논의가 미래의 영화 제작 및 산업에 긍정적인 영향을 끼치기 위해서는 영화 제작 전반에 걸쳐 실질적인 도움을 줄 수 있는 수준에서 제기되어야 한다. 즉, 영화 흥행 요인 분석 및 평가 체계는 천만 영화의 흥행요소를 빅데이터로 도출하여 영화 제작 사전(시나리오 개발 및 선정)에서부터 그 평가의 기준을 마련하는 것이다. 그리하여 그 결과를 통합적인 체계 속에서 데이터베이스화하고, 이를 적극적으로 활용할 수 있는 프로그램 개발이 절실히 요구된다.

천만 영화의 텍스트 내적 핵심 구성 요소를 추출하여 이를 바탕으로 영화 흥행 예측 모델 구축 및 소프트웨어 개발로 인한 흥행 예측이 가능해지면, 영화의 생산성 향상에 지대한 영향을 미치게 될 것이며, 이러한 융합 연구의 결과물을 통해 영화산업에 상당히 기여하게 될 것이므로 반드시 필요한 연구이다.

1.3. 연구 방법의 의의

'빅데이터'란 기존의 데이터베이스 관리도구로 데이터를 수집, 저장, 관리, 분석할 수 있는 역량을 넘어 대량의 정형 또는 비정형 데이터 집합으로부터 가치를 추출하고 결과를 분석하는 기술을 말한다. 빅데이터 기술은 다변화된 사회를 더욱 정확하게 예측하여 효율적으로 작동케 하고 개인화된 사회 구성원 에 맞는 맞춤형 정보를 제공해줄 수 있을 것으로 기대된다. 이같이 빅데이터는 사회, 경제, 문화, 과학기술 등의 전 영역에 걸쳐 가치 있는 정보를 제공할 수 있다는 점에서 그 중요성이 부각되고 있다.

기존의 영화 흥행 요인을 분석한 연구 결과들은 이를 통합적으로 관리되지 못하고 있다고 볼 수 있다. 또한 이들 기존 연구들은 일부 개별 영화 작품들에 대한 사후적인 평가에 그칠 뿐만 아니라 다루고 있는 영화 작품도 매우 제한적이어서 신뢰도 높은 평가 자료로 활용되기에는 한계가 있다. 본 연구는 이에 빅데이터를 활용하여 천만 영화의 흥행 요인 평가 측정 결과를 산출

한 후 흥행을 장르별로 예측할 수 있는 체계를 마련하고자 한다.

II. 본 론

2.1. 연구방법 및 내용

영화흥행 요인을 분석하기에 앞서 우선 천만 영화의 흥행요소에 대해 어떻게 접근할 것인가 대한 문제를 해결해야 한다. 그리하여 첫째로, 기존의 천만 영화 연구나 영화의 흥행 요소 연구를 바탕으로 어느 요소가 가장 핵심적인 흥행요소가 되는지를 추출한다. 둘째로는 13편의 천만 영화를 대상으로 앞서 발견된 흥행요소가 SNS 댓글이나 인터넷 포털, 기사에서 어떻게 나타나는지 키워드 버즈량으로 빅데이터 분석을 실시한다. 셋째, 영화 전체의 러닝타임을 영화의 구성 5단계(발단-전개-위기-절정-결말)로 구분하여 해당 요소가 영화별로 어느 부분에 많이 나타나는지 도표화 한다. 넷째로, 영화 구성 5단계 중 어느 부분에 흥행 요소가 많은지를 장르별로 구분하여 천만 영화의 장르별 구성 방식을 논증하고자 한다.

“천만 영화를 해석하는 방향은 여러 가지가 있을 수 있다. 산업적 관점, 한국 역사와 결부시키는 관점, 이와 비슷한 이데올로기 관점, 사회학적 관점, 작가론적 관점 등등이 있을 수 있다.”[1]. 그러나 본 연구는 천만 영화의 공통적인 항목요소가 있다고 보고, 이를 찾아보았다.

“천만 영화가 되려면 미학적 우수성이나 스타일적 탁월함보다는 무작위의 대중과 교감할 수 있는 ‘그 무엇’이 영화 속에 녹아있어야 한다. 결국 천만 영화 안에는 그 시대, 다양한 세대와 소통할 수 있는 집단적 무의식(Collective Unconscious)같은 것이 절묘하게 내포해있어야 한다. 이제까지 만들어진 총 17편의 천만 영화 가운데 4편만 외국 영화이고 나머지 13편이 한국 영화인 것을 보면, 이 사실을 쉽게 인지 할 수 있다. 한국 감독이 만든 영화에 한국 관객이 쉽게 동일시하는 현상. 다시 말해 자국 감독이 자국민들이 지닌 ‘그 무엇(!)’을 영화를 통해 재현하면, 다시 그 영화에 자국 관객이 공명하는 것이다.”[1].

그렇다면 과연 관객들이 공명하는 요소가 무엇인가. 이 논문에서는 많은 연구를 살펴본 결과, 그것이 눈물

과 웃음의 요소인 것을 결론내릴 수 있었다. 우선 눈물의 요소를 먼저 살펴보자.

“눈물은 영화 흥행의 필수 요소이다. 한국영화의 흥행 요건으로 액션, 웃음, 사랑 등이 없어서는 안 되는 것이지만, 언제부턴가 눈물이 영화의 중요한 요소로 자리 잡았다. 눈물 없인 흥행도 없다는 공식이 일반화된 양상이다. 이전의 한국영화에 눈물이 없었던 것은 아니다. 신파연극에 뿌리를 둔 전형적인 신파영화가 있었다. 신파라는 것이 과거 하나의 장르로 존재했고, 신파성은 한국영화에서 장르 불문하고 흥행을 위한 가장 강력한 정서적 소구 수단이다.”[2].

눈물은 한국 영화의 신파영화라는 장르에서 시작하여 한국 사람들에게 어필하는 큰 요소가 되었다. 그런데 눈물이라는 비극적 정서가 아니라 웃음이라는 희극적 정서도 많은 관객의 호응을 얻었다.

“90년대 대중문화, 좀 더 범위를 좁히자면 한국영화에서도 이러한 변화는 확연히 드러난다. 그 이전 시기까지 한국 대중영화를 지배했던 감성은 신파성이라는 뿌리에서 뻗어 나온 것이었으며, 근본적으로 비극성의 정서에 소구하는 경우가 많았다. 이른바 리얼리즘 계열의 영화로 계보화되었던 일련의 영화들에서도 비극적 엔딩이 대세였고 감정적 분출보다는 계몽이성의 진지함과 냉엄함을 중시하는 태도가 추구되었다. 그러나 1992년 <결혼 이야기>라는 기획영화의 성공은 대중영화의 감성적 헤게모니를 비극성에서 희극성으로 전화시키는 획기적 계기가 되었다. 이후 신파적 멜로드라마는 로맨틱 코미디가 되었고 갱스터 액션은 조폭 코미디가 되었으며 이른바 작가주의 감독들이 만들어내는 진지한 영화들에서도 우스꽝스러운 대목들은 필수적인 양념처럼 끼어들게 되었다[3].

90년대가 울음 대신 웃음의 정서에 편승했던 이유는 다시금 80년대가 추구했던 합목적적 계몽이성의 진지함에 맞서는 감성적 가벼움의 표현이었고 탈정치화된 주체의 율가분함을 반영하고 있으며 따라서 개인주의적 성향의 신세대에게 특히 어울리는 지배적 정서였다는 것이다[3].

희극성의 가치란 궁극적으로 대상의 지위를 격하시킴으로써 주체가 우위에 설 수 있게 해주는 매개로서의 역할이라는 것이다. 숭고와 코믹을 대극에 놓고 있는 사고 역시 이로부터 비롯된다[3].

이처럼 웃음과 눈물은 보편적 정서로서 천만 영화의

흥행요소임을 알 수 있다. 그러면 천만 영화에서 웃음과 눈물의 요소가 중요하다는 것을 어떻게 증명할 수 있을까? 이에 본 연구는 빅데이터를 활용하여 한국의 13편의 천만 영화에 대한 흥행 요소를 발견하는 데 그 첫 번째 목적이 있으므로, SNS 등의 인터넷 댓글에 웃음과 눈물의 요소가 어떻게 나오는지 빅데이터를 활용하여 도출해 보았다. 표 1은 눈물과 웃음의 요소를 추출하기 위해 정리한 것이다.

Table. 1 Words related to the box office element(Tear, Laughter)

Laughters	Tear
pleasant, cool, interesting, thrilling, smiling, pleasant, brackets goodness, laughter, laugh, premature death in abdominal pain, funny, grinning, laugh with, laughter, smile, belly	tears, crying, sorrow, cried, sad, painful, pain, tears are, painful, heart, inner corner of the eye griping

아리스토텔레스는 『시학』에서 비극에서 카타르시스가 나타나는 것은 훌륭한 인물이 운명에 의해 패배하는 결말로 귀결되는 플롯에 있다고 했다. 그렇다면 영화의 구성 5단계(발단, 전개, 위기, 절정, 결말) 중 어느 구간에서 웃음이나 눈물이 가장 많이 나타나는지 그 구성이 중요하다고 할 수 있을 것이다. 이를 영화별로 빅데이터 작업을 통해 결과를 도출하고 이를 다시 장르별로 비교 분석하고자 한다.

그 결과 이러한 요소데이터를 수집 및 분석한 결과를 통합적으로 관리할 수 있는 체계를 마련할 수 있을 것으로 본다. 그렇다면 이러한 결과를 바탕으로 시나리오 상태에서도 흥행 예측이 가능하게 될 것이다[4].

영화 흥행 요인 분석 및 평가 체계는 제작 상황뿐만 아니라 프리 프로덕션(시나리오 개발 및 선정, 배우섭외, 스태프 결정, 장소헌팅, 콘터, 제작발표회 등) 과정에서부터 포스트 프로덕션(편집, 광학녹음, 홍보마케

팅, 배급 상영 등)과정에 이르기까지 영화 제작과 관련된 전방위적인 평가 및 예측이 행해질 수 있는 데로 나아가야 한다. 이를 위해서는 이 연구결과로 도출된 웃음과 눈물의 흥행 요소가 어느 구성 구간에 나타나는지를 바탕으로 하여 사회적/시대적 분위기 요인과 텍스트 내적/외적 요인 모두를 아우르는 흥행 요인 분석이 행해져야 하며, 그 결과를 영화 흥행 예측 모델을 구축하여 관리하는 프로그램 구축을 차후 연구과제로 남긴다.

본 연구의 수행은 영화에서 분석한 영화흥행 요인을 지표화한 후 빅데이터를 활용한 공학적 접근으로 구축하는 순서로 진행된다. 이 연구는 빅데이터를 공학적으로 재구성하는 기술 개발 추진 방법이 상당히 중요하다.

이 연구는 학문 간의 융합도 학문이 인간과 삶을 바라보는 방식이라는 큰 틀에서 만났다고 보는 입장에서 출발한다. 공학은 과학을 응용하여 산업과 만나게 하는 학문이다. 영화 분야에서의 융합은 영화콘텐츠를 공학과 만나게 하여 영화산업에서 이를 활용가능하게 만드는 방향이 될 것이다.

본 연구에서는 이러한 융합연구방법을 ‘영화공학(Film Engineering)’이라는 새로운 용어로 지칭하며 융합연구의 지평을 넓히고자 한다.

‘영화공학’은 ‘스토리’를 비롯한 텍스트적 요소뿐만 아니라 컨텍스트의 요소까지 모두 고려할 필요성에 의해 생긴 개념이다. 따라서 빅데이터를 활용하여 영화 흥행 요인을 분석하고 모델 구축을 통해 체계적으로 관리하여 영화 흥행을 예측할 수 있는 프로그램을 개발하고자 하는 본 연구는 ‘영화공학’의 하나의 사례가 될 것이다[5].

본 연구가 수행하고자 하는 연구는 한국 천만 영화를 대상으로 ‘① 영화 흥행 요인 분석 → ② 영화 흥행 예측 모델 및 요인 분석을 위한 빅데이터 플랫폼 구축 → 나아가 빅데이터를 활용한 흥행 예측 평가 프로그램 개발’을 염두에 두고 진행될 예정이다.

본 연구 방식과 유사한 김진욱의 연구에서는 빅데이터를 활용한 영화 마케팅 전략을 수립하기 위해, SNS에서 나타난 반응을 긍정적 단어와 부정적 단어로 구분하여 영화흥행에 미치는 영향을 살펴보고자 대량의 검색 조회 수를 데이터화 했으며, 영화 <명량>의 영화마케팅에 있어 빅데이터 분석이 영화흥행을 예측하는 데 기존의 마케팅 분석보다 가장 적은 $\pm 6\%$ 이하의 오차범위를 이끌어 냈다[6]는 것을 강조하고 있다.

천만영화에서 ‘웃음’과 ‘눈물’이라는 텍스트 내적 요소를 도출하고 이를 빅데이터를 활용하여 검증한 본 논문의 영화흥행 분석 역시 오차범위가 적은 방식이라는 것을 강조하여도 될 것이다.

III. 장르별 웃음, 눈물의 빅데이터 분석 결과

3.1. 장르별 빅데이터 분류 및 수집

본 논문에서 수행하고자 하는 것은 천만이상의 관객을 동원한 흥행작들의 흥행 요소를 규명하고자 하는 것이다. 이를 위해 영화를 사극, 드라마, 액션, 재난, 역사물(시대극)로 분류하고 각 장르에서 대표적인 영화, 총 13개의 흥행작을 추출하였다. 사극 장르에서는 <왕의 남자>, <명량>, <광해>를 드라마 장르에서는 <실미도>, <국제시장>, <7번방의 선물>, <변호인>, <태극기 휘날리며>, 재난 장르에서는 <괴물>, <해운대>, 액션 장르에서는 <베테랑>, <도둑들>, 역사물(시대극)로는 <암살>을 선택하였다. 각 영화에 대해 제작단계에 따른 기본 데이터(제작사, 감독, 배우, 배급사, 상영관 수, 개봉시기, 전개방식, 장르, 등급, 수상내역 등)를 포털사이트와 영화진흥위원회의 연도별 통계자료를 통해 수집하였고, 또한 각 영화에 대한 감상평을 전문가와 일반인으로 나누어 수집하였다.

연구를 위해 python, java를 이용한 웹 크롤러를 설계, 개발하였다.

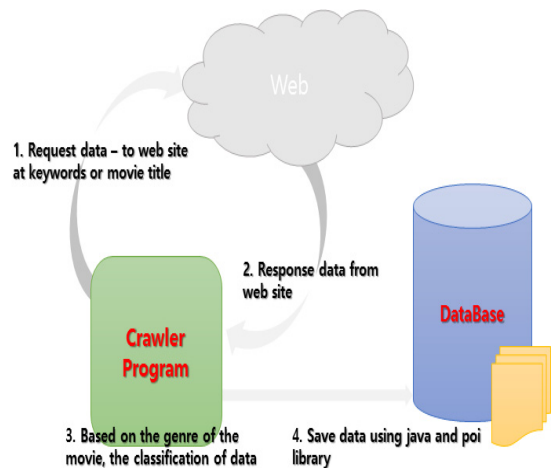


Fig. 1 Operation process of the web crawler

해당 프로그램에서는 특정 사이트의 내용을 텍스트 파일로 읽어와 영화별로 저장해 주는 역할과 별도의 키워드를 기준으로 관련된 내용을 수집해 주는 역할을 한다. 웹 데이터를 읽어 오는 것은 python과 crawler4js를 서버에 저장하는 것은 java와 poi라이브러리를 이용하였다. 그림 1은 개발한 웹 크롤러를 이용해 웹 사이트에서 정보를 수집하는 과정을 도식화 한 그림이며, 그림 2는 데이터를 수집하여 저장한 데이터베이스 구조이다.

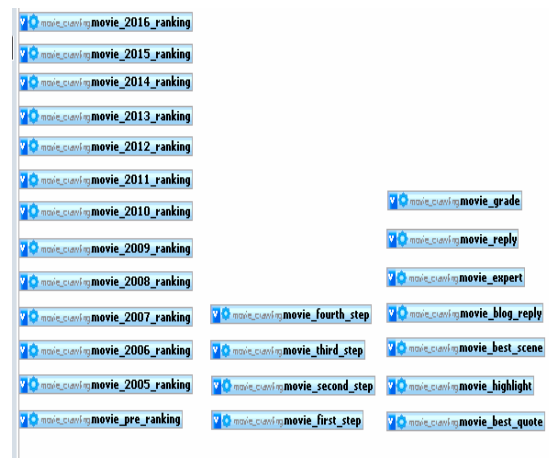


Fig. 2 Construction of Database

각 영화 마다 관객들이 추천한 수를 바탕으로 명장면을 13개씩 추출하였다. 추천 수의 기준은 네이버, 다음, 영화진흥위원회에서 관객들이 평한 장면과 대사를 조합해 추천 수가 많은 순으로 추출한 것이다. 다음 절에서 제시하는 장면에 scene1 ~ scene13으로 표시되어 있으며, 1부터 추천 수가 높은 장면이다. 각 장면을 이미지화 하여 구글 이미지 검색을 통해 검색한 결과, 장면별 5,000 ~ 6,000개 정도의 검색결과가 추출되었으며, 각 결과 페이지에 있는 눈물과 웃음의 요소를 중복을 제거하여 추출하였다. 이렇게 수집한 결과를 바탕으로 다음절에서는 흥행작에서 영화의 러닝 타임 안에서 명 장면과 눈물과 웃음의 요소의 수와 러닝 타임 구간 별 각 요소의 가 등장하는 추이를 살펴해보도록 하겠다.

3.2. 장르별 웃음과 눈물 요소 추이 분석

3.1절에서 영화의 장르별 명장면이 등장하는 시간에 대해 수집한 내용은 아래의 표 2와 같다. (전체 수집내

용 중 사극 장르-<왕의남자>, <명량>에 해당하는 내용만 본 논문에 첨부)

각 영화의 명장면 선정 기준은 3.1 절에서 설명했듯이, 네이버, 다음, 영화진흥위원회의 전문가 및 일반인들이 선정한 명장면, 명대사를 조합하여, 그 중 추천 수가 가장 많은 순으로 scene1 ~ scene13까지를 선정한 것이다.

Table. 2 Scenes appeared time of costume drama movie

Costume drama			
King And The Clown	Appearance Time	ROARING CURRENTS	Appearance Time
scene1	5.44	scene1	86.35
scene2	5.42	scene2	11.12
scene3	27.39	scene3	26.5
scene4	31.41	scene4	39.47
scene5	31.54	scene5	80.49
scene6	36.41	scene6	76.04
scene7	35.25	scene7	65.52
scene8	40.26	scene8	13.21
scene9	53.01	scene9	10.19
scene10	63.53	scene10	109.35
scene11	67.29	scene11	70.05
scene12	69.02	scene12	56.37
scene13	76.44	scene13	53.15

13개 영화에서 추출한 명장면 안에서 관객들이 평한 내용 중 웃음과 눈물과 관련된 내용을 카운팅 하기 위하여 별도의 테스트 셋을 다음과 같이 지정하였다.

TestSet 1 눈물 : 눈물, 통곡, 슬픔, 울었다, 슬프다, 울음, 아프다, 아픔, 눈물이, 아프다, 마음이 아픔, 눈시울, 흑흑

TestSet 2 웃음 : 즐겁다, 멋지다, 웃기다, 통쾌, 웃긴, 방긋, 찡다, 유쾌, 멋짐, 웃음이, 웃다, 요절복통, 우습다, 씨익, 폭소, 씩, 실소, 빙그레, 웃겨, 웃음, 미소, 배꼽

위의 테스트 셋 1,2를 이용하여 각 장면에 대해 관객들이 언급한 내용 중 웃음과 눈물에 관련된 것이 몇 개나 되는지 중복되는 내용을 제거한 후 분류한 결과는 다음의 표 3과 같다.

Table. 3 Counting of Tears and Laughter at box office scene

King And The Clown		Roaring Currents	
Count of Laughter	Count of Tear	Count of Laughter	Count of Tear
284	685	24	201
354	512	34	153
124	245	102	203
234	652	152	204
256	152	38	320
123	354	112	265
153	258	103	284
205	345	204	306
102	235	105	201
135	684	106	423
102	524	157	621
135	675	42	203
112	125	50	109

아래의 그림 3에서 볼 때, <명량>이 다소 무거운 정통 사극인 탓으로 대체로 웃음 요소에 대한 버즈량(SNS+댓글+기사)은 적다고 볼 수 있었지만, 사극 장르에 있어서 웃음은 대체로 영화 러닝타임을 5구간(발단-전개-위기-절정-결말)으로 구분한 구성에 있어 2구간(전개) 부분에서 가장 높은 버즈량을 기록하고 있으며, 후반부로 갈수록 웃음의 요소가 적었다. 사극 장르에 있어서 웃음은 흥행 요소의 발단 부분에서 준 긴장감을 전개 부분에서 다소 해소시키면서 관객을 끌어들이는 요소로 작용했다고 볼 수 있다.

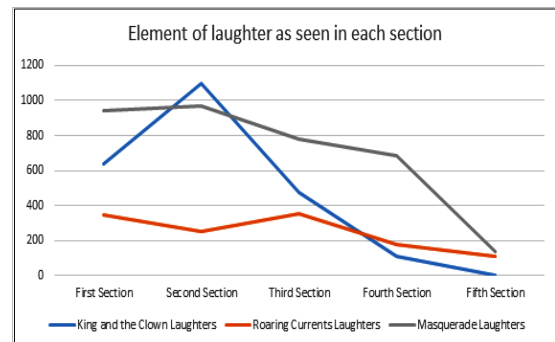


Fig. 3 Elements of laughter as seen in each section

이는 사극 장르에서 흥행하려면 전개 부분에서 웃음의 요소를 넣어야 한다고 결론내릴 수 있다. 그러면 눈물의 요소는 어떠한가를 살펴보자.

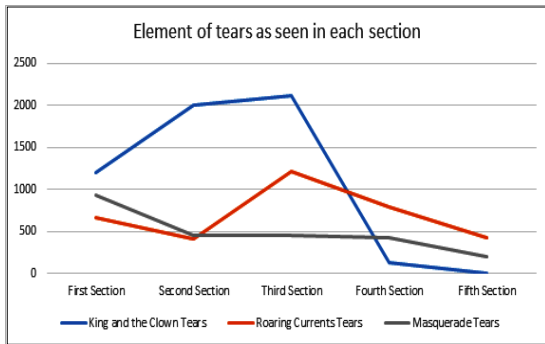


Fig. 4 Elements of tear as seen in each section

위의 그림 4를 보면 <광해: 왕이 된 남자>가 웃음 요소가 많은 탓에 대체로 눈물 요소는 적은 버즈량을 보이지만, 대체로 사극에서의 눈물 요소는 구성 5구간 중 3구간에서 가장 높은 버즈량을 기록하고 있으며 후반부로 갈수록 눈물 요소가 적은 버즈량을 보이고 있다. 사극 장르에서는 사건의 전개가 지나간 다음 눈물 요소로 관객의 감정이입을 유도하는 것이 효과적이라는 것을 알 수 있다. 이는 사극 장르에서 흥행하려면 3구간인 위기 부분에서 울음의 요소를 넣어야 한다고 결론내릴 수 있다. 웃음과 눈물의 요소는 천만 영화 13편을 전체로 놓고 보면 모두 웃음과 눈물의 요소를 지니고 있으며, 영화별로 각각 다른 특성을 보이기도 한다는 것을 알 수 있다.

그림 5에서 볼 때, <왕의 남자>인 경우 웃음 요소와 울음요소가 급격한 차이를 보이며 비극적 정서를 주제로 되어 있는 반면, <베테랑>의 경우는 웃음 요소가 울음 요소보다 훨씬 많은 버즈량을 보이고 있는데, 이는 <베테랑>의 천만 영화의 요소는 희극적 정서를 주제로 되어 있다는 것이 특징적으로 나타난다. 또한 웃음과 울음의 버즈량이 가장 비슷한 영화는 재난영화인 <괴물>과 <해운대>로 재난을 통한 울음 요소에도 그만큼의 웃음의 요소가 들어가 있어야 한다는 것이 발견된다. 이외의 대부분의 천만 영화에서는 웃음과 울음의 요소가 대체로 유사하게 나타나고 있어 영화 흥행에는 웃음과 울음을 골고루 섞어주는 것이 바람직하다는 결론을 내릴 수 있다.

IV. 결 론

본 연구는 국내에서 천만이상의 관객을 동원한 한국영화 13편의 흥행작들의 흥행 요소를 규명하고자 하였다. 이를 위해 천만 영화를 대상으로 웃음과 눈물이라는 텍스트 내적 흥행 요소를 도출하고 이러한 분석 결과가 실제로 어떠한 흥행 결과를 낳게 되었는지에 대한 종합적인 데이터베이스를 구축하는 것을 목표로 하였다. 그 결과 13편 각각의 영화가 영화별로 웃음과 울음의 요소가 다르게 나타났지만, 대체로 웃음과 울음이 골고루 섞여 있는 것으로 나타났다. 특히 장르별로 살펴볼 때, 사극 장르의 경우 구성의 구간별로 웃음은 전개 부분에서, 울음은 위기 부분에서 가장 많은 버즈량을 기록하고 있다는 것을 알 수 있어서 영화

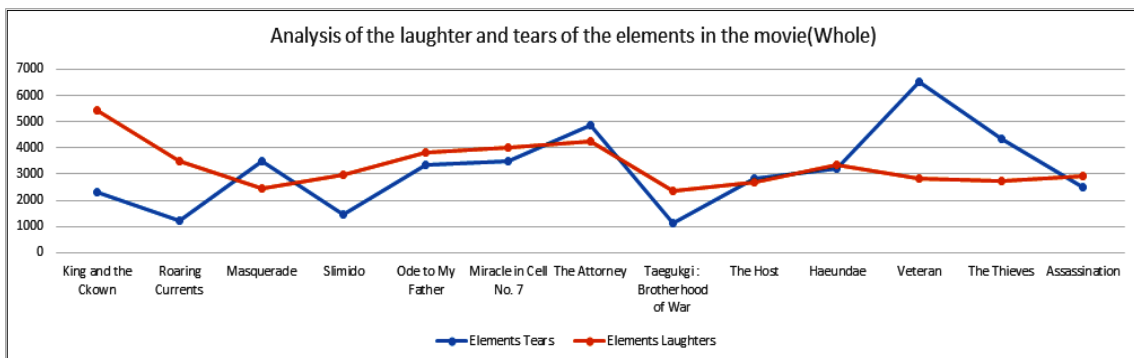


Fig. 5 Analysis laughter and tear of the elements in the movie

가 흥행하려면 구성을 어떻게 해야 할 것인가에 대한 중요한 자료를 제시한 것으로 볼 수 있다.

그동안 영화 텍스트 외적인 요소를 통해 빅데이터를 활용하여 영화산업에 영향을 끼치는 요소에 대한 연구는 있었지만, ‘웃음’과 ‘눈물’ 같은 중요한 텍스트 내적인 요소를 추출하여 접근한 논문을 없었다. 이 연구는 천만 영화에서 가장 중요한 내적인 요소를 발견하고 이 요소가 천만 영화의 장르별 구성에 어떤 추이를 보이는가를 명장면을 통해 추출해 내었으므로, 기존의 영화 텍스트 외적 요소를 빅데이터 자료로 한 연구나 공정, 부정적 반응을 단순히 분석 요소로 검증한 연구와는 크게 차별화된다.

텍스트 내적인 요소들에 대한 정밀한 분석과 정보 구축을 통해 영화의 발상이나 구성에 있어 기초 자료들을 풍부하게 제공할 수 있으며, 영화 제작이나 기타 관련 영상물의 생산 작업 현장에도 실질적인 자료로 활용될 수 있을 것이다. 이와 같은 빅데이터를 활용한 데이터베이스 구축은 영화 제작 환경에서 사전 제작 단계에서부터 사후 단계에 이르기까지 전반적인 영향 평가가 가능한 기초 자료 마련을 뜻한다. 이처럼 영화 흥행 요소의 데이터베이스 개발을 통해 흥행 예측이 가능해지면 영화의 생산성 향상에 지대한 영향을 미치게 될 것이다.

본 연구의 분석 결과는 실제 영화 산업 현장에서 사전-제작-사후 단계 전반에 걸쳐 유의미한 영향을 줄 수 있을 것이며, 이를 토대로 제작된 영화는 향후 흥행과 부가가치를 창출해낼 수 있을 것이다. 이러한 영화 흥행 자료의 활용을 통해 영화 제작에 드는 비용의 리스크를 최소화하는 데 많은 도움이 될 것으로 보인다. 그러므로 흥행 실패에 따른 위험성을 최소화하는 데 기여할 수 있을 것이다.

이는 공학과 영화 콘텐츠, 영화 산업과의 ‘융합’ 연구가 대학 내 연구로만 그치지 않고, 더 나아가 실제 산업에 긍정적인 영향을 끼칠 수 있다는 점에서 좋은 선택이 될 수 있다는 점에서 의의를 부여할 수 있다.

ACKNOWLEDGMENTS

This work was supported by the Ministry of Education of the Republic of Korea and the National Research Foundation of Korea (NRF-2015S1A5B6037028)

REFERENCES

- [1] S. R. Kang, "A Study on the 10-million Mark Films- Focusing on Patriarchal familism and Tragedy as 'Shinpa'," *International Society for the Korean people culture*, vol. 52, pp.343-372, Dec. 2015.
- [2] J.K. Kim, "Film Editing Style and Tears in the Audience," *Society for journalism and communication studies*, vol. 19, no. 1, pp. 87-113, Feb. 2015.
- [3] S.Y. Kim, "A Study on 'the Comic Mode' in Korean Cinema Since the 1990s - Focusing on the Transitional Emergence of Romantic Comedy Genre Films-," *Film Studies Association of Korea*, vol. 60, pp. 29-55, Jun. 2014.
- [4] Y.D.Kim, I.H. Joo, Y.K.Park, I.Y.Moon, O.Y.Kwon, "A Study on Personal Experience Knowledge Evaluation Model for Knowledge Service," *Journal of the Korea Institute of Information and Communication Engineering*, vol. 17, no. 8, pp. 1865-1872, Aug. 2013.
- [5] Y.M. Hwang, "The Scientific Possibility and Direction of Film Engineering as Convergence Science," *Film Studies Association of Korea*, vol. 61, pp. 453-473, Sep. 2014.
- [6] J.W. Kim, "Big Data's Effect on its Practical Use for Box-office Success of Films : Focus on the Process of Management of Opinion Mining on *Roaring Currents*," *Journal of the Korea Entertainment Industry*, vol.9. no. 2, pp. 81-90, Jun. 2015.



황영미(Young-Mee Hwang)

국문학 박사, 숙명여대 기초교양학부 교수, 영화평론가, 한국사고와표현학회 회장, 국제영화비평가연맹 한국본부 회장
※관심분야 : 융합연구, 빅데이터, 영화 흥행, 영화공학,



박진태(Jin-Tae Park)

컴퓨터공학과 박사과정

※관심분야 : 빅데이터, 웹 표준, 네트워크 최적화



문일영(IL-Young Moon)

통신공학 박사, 한국기술교육대학교 교수

※관심분야 : 빅데이터, 무선 네트워크



김광선(Kwang-Sun Kim)

기계공학 박사, 한국기술교육대학교 교수, 한국공학교육학회 회장

※관심분야 : 데이터 최적화, 빅데이터



권오영(Oh-Young Kwon)

통신공학 박사, 한국기술교육대학교 교수

※관심분야 : 고성능컴퓨팅, 미들웨어