



빅데이터 기반의 사회연결망 분석을 이용한 관광지 이미지 인식에 관한 연구*

A Study on the Awareness of Destination Image Using Social Network Analysis based on Big Data

한 지 연** · 김 흥 범***

Han, Ji-Yeon · Kim, Hong-Bumm

요약 : 본 연구는 관광지의 경쟁력 제고에 긍정적인 영향을 미치는 관광지 이미지의 인식을 조사하기 위해 빅데이터 기반의 사회연결망 분석을 통하여 유의미한 정보를 추출하는데 그 목적이 있다. 연구목적 달성을 위해 온라인 여행 커뮤니티인 트립어드바이저로부터 서울의 117개의 관광지 및 관광상품에 해당하는 최근 5년간(2012년 1월부터 2016년 12월 31일까지) 작성된 리뷰를 분석단위로 활용하였다. 특히, 관광지 이미지를 인지적·정서적 차원으로 접근하여 텍스트 마이닝을 수행함으로써 23개의 구성개념과 200개의 이미지 요인을 도출하였다. 도출된 요인을 바탕으로 사회연결망 분석인 중심성 분석을 수행하여 관광지의 경쟁력 제고에 필요한 이미지 요인을 규명하였다. 분석결과, 관광지에 대한 긍정적 이미지 창출에 도움이 되는 구성개념은 7개, 관광지 이미지에 대하여 여행객이 심리적으로 가깝게 느끼는 구성개념은 14개, 관광지의 경쟁력 제고에 긍정적인 역할을 하는 이미지 구성개념은 8개, 여행객이 관광지 이미지를 인식하는데 있어 파급력이 높은 구성개념은 11개로 나타났다. 이러한 관광지 이미지의 세부적인 인식조사는 실현가능한 마케팅 전략 수립에 기초자료가 될 뿐만 아니라 관광분야의 빅데이터 분석에 대한 체계와 이론적 토대를 제공한 점에서 시사점을 가진다.

핵심용어 : 빅데이터, 관광지 이미지, 사회연결망 분석, 텍스트 마이닝

ABSTRACT : This study aims to extract meaningful information through a social network analysis based on Big data in order to investigate if a tourist's awareness of destination image favorably affects competitiveness of tourist destination. A web crawler was employed to gather online review data from 117 destinations and tourism products in Seoul, Korea using Tripadvisor. This web crawler is representative of the online travel community used throughout global tourism industry. Specifically, 23 image factors and 200 components composing two aspects of cognitive and affective images were derived using a system of text mining analysis. This study tried to identify and enhance destination competitiveness by investigating a tourist's awareness in terms of their destination image. This was measured using indicators of network analysis that were degree centrality, closeness centrality, between centrality and eigenvector centrality. The following results were found: First) 7 image components help to create a positive image of tourist awareness in terms of destination image. Second) 14 image components feel psychologically close to tourists in terms of destination image. Third) 8 image components favorable influence and enhance the competitiveness of destination. Fourth) 11 image components have a greater impact greater than the others when tourists have a perceived destination image. Implications and suggestions are presented along with the findings of the study, which will contribute to the

* 이 논문은 2016년 대한민국 교육부와 한국연구재단의 지원을 받아 수행된 연구임(NRF-2016S1A5A2A01022896).

** 세종대학교 대학원 호텔관광경영학과 박사과정. e-mail: bara0312@naver.com

*** 세종대학교 호텔관광대학 교수(교신저자). e-mail: kimhb@sejong.ac.kr

theoretical framework by suggesting a new perspective for measuring destination image based on Big data.

Key words : Big data, Destination image, Social network analysis, Text mining

I. 서 론

관광지 이미지에 대한 관심은 관광지의 경쟁력 제고를 위한 전략수립에 있어 긍정적인 이미지 연출이 여행객의 관광지 의사결정에 절대적인 영향력을 미치는 중요한 결정요인이라는 다양한 연구결과가 보고됨에 따라 산·학계에서 꾸준히 증가해 왔다(김민주, 2003; 서용건·서용구, 2004; 왕명명·류시영·송운강, 2011; Gallarza, Saura & García, 2002). 관광분야의 중요한 함의를 내포하는 관광지 이미지에 대하여 그동안 많은 학자들이 다각적인 차원에서 관광지 이미지를 측정하고 그 구조적 관계를 규명해오고 있으며 관광지 이미지는 여행객이 관광지를 관광/여행하면서 느끼는 총체적인 지각의 결합체의 의미로써 정리될 수 있다. 다시 말해, 여행객들은 관광지에 대하여 인식하는 이미지를 토대로 관광지를 평가하고 방문의사를 결정하게 되므로(임정우·이교은·하동현, 2013; Bigne, Sanchez & Sanchez, 2001; Tapachai & Waryszak, 2000) 시장 내에서 관광지의 경쟁적 우위를 선점하기 위해서는 긍정적인 관광지 이미지의 연출이 필수적이라 할 수 있으며 따라서, 그 중요성이 대두된다고 할 수 있다.

학계에서의 이미지 연구는 초창기 관광지 전체의 이미지 요인 파악 및 측정과 이에 근거한 포지셔닝에 관한 연구로 진행되어 오다가 2000년대 이후로는 점차 관광객의 행동요인이나 정보탐색과 같이 관광지의 선택적 관점에서 관광지 이미지를 조명하였다(김병국·박석희, 2001; 노정희, 2008; 박수완, 2004; 조명환·정선자,

2009; 채예병, 2007; 현용호·홍선영·오홍철, 2007). 최근에는 계량적 척도를 이용한 중요도-만족도 분석을 통하여 특정 관광지 또는 관광지 유형에 따른 관광지 이미지의 강·약점을 파악하여 여행객에게 긍정적으로 인식될 수 있는 관광지 이미지의 구성요인을 규명하고자 하는 연구들이 이루어지고 있으며(Jeng, Snyder & Chen, 2017; Pike, 2016), 나아가 관광지 이미지의 연구를 확장하여 관광객을 끌어들이는 이미지 요인 및 관광서비스기업에 영향을 미치는 일반적 요인(generic factors) 까지도 논의함으로써 여행객이 인식하는 관광경쟁력 요인을 도출하고자 하는 시도 또한 이루어지고 있다(Enright & Newton, 2004). 이처럼 관광지 이미지 관련 연구들이 측정상의 정교화, 개념의 확장을 위한 노력들이 행해지고 있지만, 구체적인 관광지 이미지 측정을 위해서는 현재까지 주를 이루었던 서베이 기반의 계량적 분석에서 벗어나 여행객의 관점에서 형상화되는 이미지를 내재적 차원으로 접근하여 그들의 직접적인 인식에 기반한 측정이 필요하다는 연구자들의 의견이 제기되고 있다(심영석·김홍범, 2016; Mak, 2017).

이에 발맞춰 최근 관광분야에서는 온라인상에서 실시간으로 생성되고 공유되고 있는 여행객의 경험과 느낌에 대한 의견(online customer review)이 주목받고 있다. 특히, 온라인상의 대규모 질적 정보는 여행객이 관광지를 인식하고 선택하는데 큰 영향을 미치고 있다는 다양한 연구결과들이 보고되고 있으며(이학인·정기환, 2016; Choi, Lehto & Morrison, 2007), 대표적으로 Gretzel, Yoo and Purifoy(2007)

의 온라인 고객의견의 파급력과 역할에 관한 연구에서는 여행객이 여행계획 단계에서 가장 선호하고 그 영향력이 큰 행동은 타인의 여행 후기나 여행 블로그를 읽는 것으로 나타났다. 이러한 연구결과는 온라인상의 여행객들의 경험이나 기대, 만족/불만족 등을 나타내는 솔직담백한 대규모 정보의 파급력과 활용가능성이 확대되어 여행객의 의사결정 및 관광지 이미지에 대한 인식을 내재적 차원에서 파악할 수 있다는 것을 의미한다. 따라서 관광분야의 새로운 패러다임을 해석할 수 있는 중요한 도구로써 그 기능을 수행하는 온라인상의 고객의견에 대한 실증 분석할 수 있는 체계적인 접근에 대한 필요성이 강조되고 있다(오익근·이태숙·전채남, 2015).

이러한 맥락에서 최근 관광분야에서는 실시간으로 생산되는 대규모 관광정보에 대한 새로운 지식과 정보 가치를 발견하기 위하여 데이터 마이닝, 텍스트 마이닝, 사회연결망 분석 등(김근형·오성열, 2009; 김보경·김미경, 2015)의 개념적·방법론적 접근을 통한 빅데이터 연구가 활발히 이루어지고 있다. 이와 같은 분석체계를 구체적으로 확립하여 활용가능성을 제한한다면 관광지에 대한 인식파악과 경쟁력 제고방안에 필요한 현재의 주요 이슈 및 트렌드를 파악할 수 있을 뿐만 아니라 관광지 이미지를 토대로 한 전략적 마케팅 수립에 대한 새로운 시각을 제공할 수 있을 것이다.

본 연구의 목적은 온라인 여행커뮤니티의 빅데이터를 웹 크롤러를 통해 수집하고, 텍스트 마이닝 기법과 사회연결망 분석을 활용하여 여행객이 관광지에서 직접적으로 인식하는 이미지를 구성하는 요소 간 구조적 관계(네트워크 체계)를 파악함으로써 관광지에서 여행객이 긍정적 이미지를 인식하는데 기여하는 구성개념과 경쟁력 제고에 필요한 요인을 규명하는 것이다. 이러한 연구를 통해 관광분야에 실증적으로 적용 가능한 빅데이터 분석의 체계를 확립함으로써 관광학 연

구의 학문적 시사점과 긍정적인 관광지 이미지 연출을 통한 관광지의 경쟁력 제고에 수반되는 가장 효과적인 마케팅방안을 도출하는데 필요한 실무적 시사점을 제시할 수 있을 것으로 기대된다.

II. 이론적 배경

1. 관광지의 경쟁력 제고를 위한 관광지 이미지 측정

Gallarza *et al.*(2002)가 관광지 이미지에 대한 개념적 틀을 제시한 연구에 따르면, 관광지 이미지의 복합성(complex), 다양성(multiple), 상대성(relativistic), 동적(dynamic)이라는 현상의 복합체로 설명하고 있다. 특히, 상대성과 동적이라는 특성으로 인해 전략적 변수로써 그 역할을 수행하고 있다는 주장하고 있다. 또한, Konecnik(2002)은 관광지 이미지의 긍정적 연출이 글로벌 시장에서 해당 관광지가 경쟁적 우위의 원천이 될 수 있다고 주장하였다. 따라서 관광지 이미지는 여행객의 관광의사결정에 가장 결정적인 역할을 하는 전략적 요인으로 해당 관광지의 긍정적 이미지 연출은 경쟁적 우위를 선점할 수 있는 마케팅 도구로써 그 개념적 틀을 제공한다는 것을 알 수 있다(Gartner, 1994; Rezende-Parker, Morrison & Ismail, 2003). 그러므로 관광지의 지속적인 활성화를 위해서는 잠재 관광객의 인식과 욕구를 반영한 이미지 창출이 반드시 필요하다.

초기 관광지 이미지 연구는 여행객이 관광지 자체의 자원 또는 매력성에 해당하는 자연경관, 사회간접자본, 기반시설, 사회적 환경 등의 유형적이고 직접적으로 관찰이 가능한 물리적 실체에 대한 지각에 기반 한 인지적 이미지의 구성요인을 규명하고 이에 대한 측정이 강조되었

다(Baloglu & Brinberg, 1997; Baloglu & McCleary, 1999). 그러나 관광지 이미지는 물리적 환경에 대한 인식뿐만 아니라 관광지에서의 전체적인 경험과 생각, 느낌에 대한 개인의 선호/비선호에 대한 감정으로 표현될 수 있으며, 이러한 개인의 정서는 여행객의 관광동기에 영향을 미친다는 연구결과가 보고됨에 따라(Beerli &

Martin, 2004) 관광지 이미지를 개념화하기 위해서는 관광지의 인지적 이미지와 정서적 이미지가 총체적으로 고려되어야 한다는 연구자들의 견해가 대두되었다(Chen & Uysal, 2002; Kim & Richardson, 2003; Pike & Ryan, 2004). 정리하면, 관광지 이미지는 인지적 이미지와 정서적 이미지의 두 가지 차원으로 개념구

〈표 1〉 관광지 이미지의 인식 및 경쟁력에 대한 국외연구 현황^a

연구자	측정방법	연구내용
Enright & Newton(2004)	계량적 척도를 이용한 IPA 분석	관광지 이미지 중심의 연구를 넘어서 관광지 이미지의 52개의 변수를 바탕으로 홍콩 관광지의 관광객을 끌어들이는 관광이미지 및 관광서비스기업에 영향을 미치는 일반적 요인(generic factor)을 규명함으로써, 관광경쟁력 요인을 도출한 연구임
O'Leary & Deegan(2005)	계량적 척도를 이용한 IPA 분석	아일랜드를 방문하는 프랑스 여행객으로부터 관광지 이미지를 조사하였으며 아일랜드 관광지의 긍정적 이미지를 창출하는 중요한 17가지의 속성을 도출하였으며, 특히 아름다운 경관과 친절한 지역주민이 가장 긍정적 인식의 효과가 높은 것으로 나타남
Gomezelj & Mihalic(2008)	계량적 척도를 이용한 다변량 분석	슬로베니아의 관광지 경쟁력 제고를 위해 선행연구의 모델을 통합적으로 접근하여 85개의 속성을 도출하였으며, 자연·문화 자원이 경쟁력 제고에 중요한 변수임을 규명하였음
Stepchenkova, Kirilenko & Morrison (2009)	온라인 커뮤니티 및 웹상의 관광관련 뉴스기사 내용분석	러시아의 지역별 이미지에 대한 인식을 조사하기 위해 웹상에 나타난 관광 관련 정보를 내용분석을 수행하였으며 관광분야에서 웹상의 텍스트 데이터를 분석하기 위한 활용성을 제안함
Crouch (2011)	계량적 척도를 이용한 AHP 분석	관광지 경쟁력 제고에 필요한 결정요인 규명을 위해 36개의 속성을 바탕으로 AHP분석을 수행하였으며 5개의 주요인에 대한 중요도와 결정요인을 규명하였음
Dickinger, Költringer & Körbitz(2011)	온라인 서베이와 여행 블로그 내용분석	탈린의 관광지 이미지를 조사하여 입지적 우위가 관광지로서 차별화된 경쟁력으로서의 기능을 한다는 것을 알 수 있었으며, 나아가 여행 블로그의 내용분석을 통해 온라인 정보가 현재까지의 관광지 이미지 연구를 보완할 수 있는 원천이 될 수 있다고 주장함
Dwyer, Dragičević, Armenski, Mihalić & Knežević Cvelbar(2016)	계량적 척도를 이용한 IPA 분석	세르비아의 관광지 경쟁력 제고를 위해 선행연구를 바탕으로 도출한 6개의 이미지 요인에 대한 중요도-만족도 분석을 수행하였으며, 세르비아의 경쟁력에 중요한 역할을 하는 변수는 지속가능한 개발, 만족도가 가장 높은 변수는 상품 개발 및 혁신인 것으로 나타남
Jeng <i>et al.</i> (2017)	계량적 척도를 이용한 IPA 분석	관광목적지로서 대만에 대한 여행객의 인식을 조사한 연구로, 대만에 대한 인식은 즐거움, 자유로움, 현대적인 장소, 친근한 지역, 흥미로운 문화의 정서적 이미지와 날씨, 견문을 넓힐 수 있는 기회, 안전성, 정치적 안정성, 안락함, 경관 등이 긍정적 이미지 연출에 기인하는 것으로 나타남

주: a. 연구자 정리.

분이 가능하며, 두 가지 차원을 다각적으로 고려하여 총체적인 개념으로 접근함으로써 여행객이 관광지에 대하여 인식하는 이미지의 유의미한 결과를 도출할 수 있다.

현재까지 관광지 이미지에 대한 인식을 바탕으로 관광지의 경쟁력 제고를 위한 연구는 <표 1>과 같이 정리될 수 있다. 정리된 <표 1>을 살펴보면, 여행객이 인식하는 관광지 이미지와 이를 바탕으로 경쟁력 제고를 측정 한 연구는 특정 국가 또는 관광지로서의 도시에 대해 다양한 방법으로 접근되어 왔으며 이 가운데 특히, 관광지의 인지적 이미지와 정서적 이미지를 종합적으로 평가하기 위해 계량적 척도를 기반으로 한 중요도-만족도(IPA) 분석이 주를 이루었다. 그러나 최근 관광지에 대한 여행객들의 관심이 증가하면서 관광지 이미지에 대한 개념적, 조각적 범위가 확장되었으며 이러한 맥락에서 관광지 이미지에 대한 보다 객관·타당한 결과 도출을 위해서는 여행객이 관광지에서의 경험과 감정의 내재적 차원에서 접근해야 한다는 학자들의 의견이 대두되고 있다 (Dickinger, 2011; Mak, 2017). 따라서 최근 관광지 이미지 분야의 연구에서는 여행객의 의견과 같은 질적 데이터의 활용성이 강조되고 있으며 특히 온라인 환경에서 실시간으로 생성되고 공유되는 관광지에 대한 여행객의 솔직한 의견인 온라인 리뷰에 주목해야 한다는 의견이 제기되고 있다 (심영석·김홍범, 2016; Ye, Zhang & Law, 2009).

관광지의 경쟁력 제고를 위한 이미지 측정의 연구자들의 관점을 종합해보면, 온라인 정보에 투영된 여행객들의 의견을 토대로 관광지 이미지의 구성개념을 하나의 전략변수로써 인식하고 접근하여 속성 및 실체를 도출한다면, 관광지가 가지는 강·약점의 파악에 유용하게 활용될 수 있음을 알 수 있다. 따라서 본 연구는 국내 관광지중 여행객이 가장 많이 방문하는 서울을 관광지로 규정하여 여행객이 인식하는 이미지 조사를 온라인 리

뷰 빅데이터에 기반 하여 다각적으로 분석하고자 한다.

2. 관광분야의 빅데이터 활용

최근 정보기술의 발전으로 인한 인터넷 환경의 급격한 변화는 관광 상품의 유통과 마케팅 (Buhalis & Law, 2008; Buhalis & Spada, 2000)의 변화, 관광정보의 소비 패턴의 재설계 (Choi *et al.*, 2007; 이학인·정기한, 2016) 등 관광생태계에 여러 방향으로 변화를 이끌고 있으며, 온라인상에서 생성되는 대용량의 데이터는 여행객들의 인식에 기반 한 관광지의 이미지와 선택에 영향을 미치고 있다(Lewis & Chambers, 2000; Jeong & Jeon, 2008). 이러한 대용량의 데이터 분석을 통해 소비자들의 행동과 심리상태를 예측하여 전략수립에 활용하는 빅데이터 패러다임은 관광분야의 중요한 화두로 부상하고 있다(이영진·윤지환, 2014).

관광분야의 빅데이터는 여행객의 위치, 소비 등의 정량적 데이터와 경험에서 표현되는 인식, 감정, 경험에 대한 정성적 데이터의 결과물을 수집, 분석함으로써 관광지의 수요예측, 인식조사, 여행객 유형별 특성과 선호를 고려한 맞춤형 서비스를 제공하는데 활용되고 있다(오익근 외, 2015). 특히 비저장성이라는 관광의 특수성으로 인하여 관광지의 탐색과정에서 여행객들의 후기(review)를 통한 간접경험과 온라인망을 이용한 항공권 예매, 숙박예약 등의 일련의 활동들이 상호·유기적으로 이루지는 만큼 관광분야에서의 빅데이터는 그 활용가능성과 더불어 실무적·학문적 시사점이 상당히 크다고 할 수 있다. 실제로, 2013년부터 부산의 해운대구청은 SNS와 웹(web)상의 빅데이터를 분석을 통하여 해운대의 여행객 인식을 파악함으로써 여행객의 요구에 선제 대응하는 관광정책을 펼쳤다. 세부적으로 관광 성수기인 5월부터 10월까지 '해운대'라는 키워드 분석을 통해 여

행객들의 목소리를 수집하여 해운대의 관광매력 요인과 방해요인을 분석하였으며, 해운대의 긍정적 요인은 '해수욕장', '달맞이길', 부정적 요인은 '비싼 숙박요금', '교통체증', '복잡한 버스노선'인 것으로 나타났다. 이에 따라 구청은 숙박업소의 사전요금 표시제 및 점검, 중저가 우수 숙박업소 발굴관리 및 인증사업, 시설개선비 지원 등을 통한 부당요금 근절과 저렴한 숙박비의 게스트하우스를 확충하는 방향의 정책을 펼쳐 빅데이터를 활용한 대표적인 성공 관광지로서 발돋움할 수 있게 되었다(경남신문, 2015.07.21). 또한, 세계 최

대의 글로벌 온라인 여행 커뮤니티인 '트립어드바이저(Tripadvisor)'는 커뮤니티 내에 생성되는 리뷰, 이미지 등의 비정형 데이터를 분석하여 여행지별, 지역별, 여행객유형별 등의 선호유형을 도출하여 여행자의 특색에 맞는 여행상품을 제안하고 그에 상응하는 여행정보(관광지, 숙박업소)를 제공하는 서비스를 펼치고 있다.

빅데이터를 활용한 관광분야의 학문적 논의도 활발히 이루어지고 있다. 특히 해외에서는 <표 2>에서 정리된 바와 같이 2000년대 중반부터 빅데이터를 활용한 연구들이 이루어지기 시작하였다.

<표 2> 관광분야 빅데이터의 국외연구 현황^a

연구자	연구주제	연구내용
Stepchenkova & Morrison (2006)	러시아의 관광지 이미지에 대한 연구	미국과 러시아의 여행사와 여행안내 웹사이트 212곳에서 러시아와 관련된 키워드를 분석하여 관광목적지로서 러시아의 이미지의 구성요인 도출 및 미국과 러시아의 웹사이트 간 정보차이를 통하여 러시아의 관광수요를 충족시키기 위한 마케팅 방안을 제시함
Li, Ye & Law(2013)	호텔시장에서 고객만족의 결정요인 파악	중국 온라인 여행 커뮤니티인 다오바오닷컴(daobao.com)에서 베이징의 1063곳의 호텔에 대한 리뷰데이터를 수집하여 텍스트 분석을 수행하였으며, 고가호텔과 저가호텔에 따른 소비자 만족의 15개의 결정요인을 규명함
Költringer & Dickinger (2015)	관광지 브랜딩과 이미지 분석	온라인 여행 커뮤니티 11곳, 앵글로-아메리칸 웹사이트 162곳, DMO 웹사이트 165곳에서 5719개 문서의 빅데이터를 분석하여 UGC, DMO, 미디어에서 대표되는 오스트리아 빈의 관광지 이미지의 주요 속성을 도출하였으며, 브랜딩 방향을 제시함
Yang, Pan, Evans & Lv(2015)	중국 여행객의 행태에 대한 예측	글로벌 검색엔진인 구글(google)과 중국 최대 검색엔진인 바이두(baidu)에서 중국의 대표적인 관광지인 하이난의 시계열 분석을 통한 방문객 수를 비교분석하여 관광객 수(volume)를 측정할 수 개념적 틀을 제시함으로써 검색엔진을 활용한 수요예측의 초기 연구임
Pan & Yang (2016)	관광지의 주중 객실 점유율 예측	미국의 찰스턴, 사우스캘리포니아에 대한 검색엔진 데이터, 웹 트래픽기록, 날씨정보에 대한 빅데이터를 활용하여 시계열분석을 수행함으로써 호텔의 성과(performance)의 예측 가능한 모델을 제시함
Zhang, Lan, Qi & Wu(2017)	중국의 관광분야 연구의 평가 및 패턴분석	2003년부터 2012년의 16,024개의 중국 관광분야 연구논문에 대한 빅데이터 분석을 수행하여 10개년에 해당하는 중국의 연구추세와 현상, 평가 및 분류체계에 대한 직관적인 방법론적 체계를 제시함
Mak(2017)	관광지 이미지에 대한 관광조직(NTO)과 여행객의 지각 비교	관광지 이미지에 대하여 여행객 생성 콘텐츠(TGC)와 관광조직(NTO)의 콘텐츠의 그림과 텍스트 데이터를 분석하여 관광지 이미지에 긍정적인 영향을 미치는 요인을 도출함

주: a. 연구자 정리.

세부적으로 Choi *et al.*(2007)는 마카오의 온라인 여행 커뮤니티, 관광서 웹 페이지, 여행사, 블로그, 온라인 여행 매거진 등의 텍스트 데이터와 비주얼 데이터를 수집하여 통합적인 분석을 수행함으로써 마카오 관광지의 긍정적 이미지요인을 도출하였다. 또한 Lee, Law and Murphy(2011)는 글로벌 여행 커뮤니티인 '트립어드바이저'에서 2011년 관광지 중 가장 방문객이 많은 7개 지역의 100곳의 호텔 소비자 리뷰 데이터를 수집하여 온라인 빅데이터상에 나타난 리뷰어의 행태를 연구하였으며, 분석결과 여행경험이 많은 관광객일수록 리뷰게시에 적극적이고 상대적으로 관광지에 대하여 낮은 평가를 준다는 것을 밝혔다.

한편, 국내에서도 2010년도 이후로 빅데이터를 활용한 다양한 연구들이 진행되고 있다. 김휴찬·김미정·신현정(2014)은 빅데이터 분석을 통해 제주 관광의 트렌드를 분석할 수 있는 서비스 구축의 필요성과 지역적 차원에서 이를 관리, 분석할 수 있는 방안에 대하여 논의하였다. 또한 도해용·이애주(2015)는 한국, 중국, 일본의 외식 평가 사이트의 빅데이터를 분석하여 각 나라 레스토랑의 브랜드 개성을 측정함으로써 국내 외식기업의 현주소를 파악하고 해외시장 진출에 필요한 전략에 대하여 논의하였다. 이 외에도 조완섭·조아·권가은·류관희(2015)의 충북관광 홍보 전략에 대한 빅데이터 연구, 오익근 외(2015)의 한국관광 인식에 대한 빅데이터 분석 등 기존 관광분야의 이론적 체계를 기반으로 다양한 분석기법을 통한 빅데이터 분석이 이루어지고 있다. 최근에는 빅데이터 분석기법 가운데 정치, 경제, 조직 등의 분야에서 이슈가 되고 있는 네트워크 분석을 통하여 구조적으로 복잡한 관광지의 속성에 대한 관계규명에 관련된 연구가 활발히 진행되고 있다(김보경·김미경, 2015; 한진성·윤지환, 2016). 그러나 국내의 관광분야 빅데이터 연구들이 해외연구들에 비해 초기 수준에 머물고 있으며, 방법론적으로 동일한 체계를 답습하거나

타 분야에서 관광을 대상으로 한 연구들이 이루어지고 있기 때문에 관광분야에 실증적으로 적용 가능한 빅데이터 분석에 대한 체계적인 접근법과 개념정립의 필요성이 강조되고 있다.

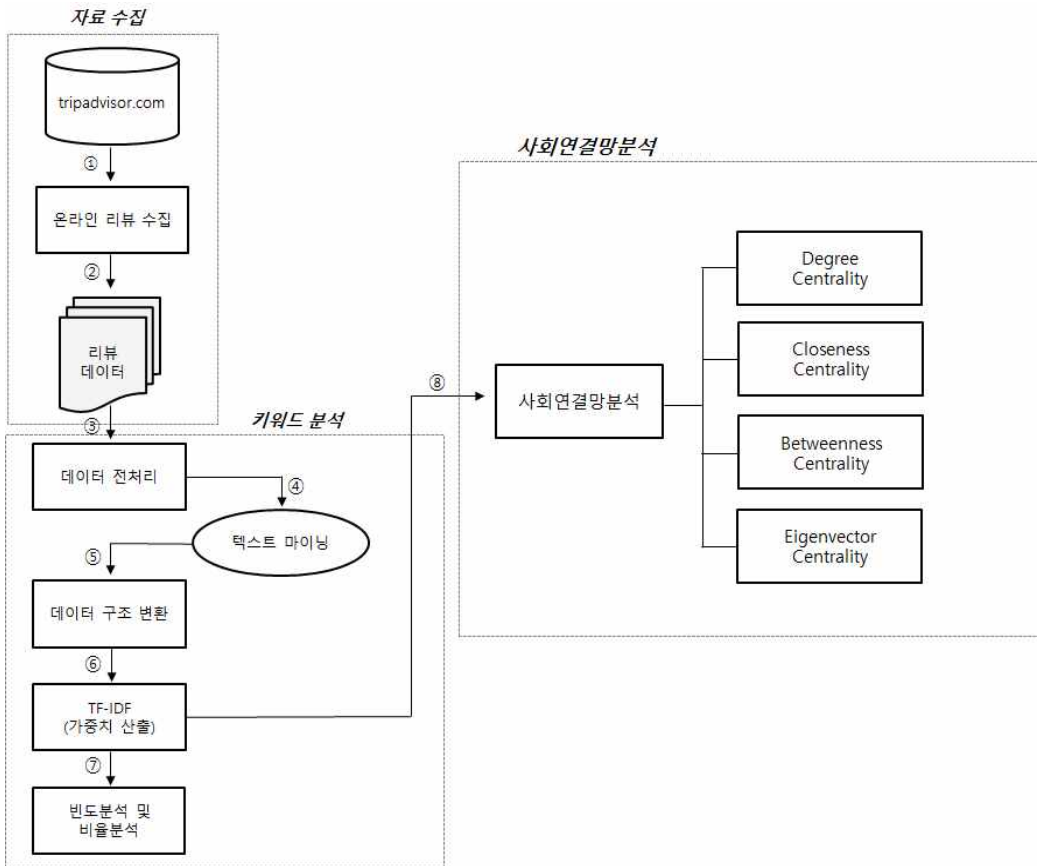
Ⅲ. 연구방법

1. 연구의 절차 및 체계

관광지 이미지에 대한 여행객들의 직접적인 인식을 조사하기 위해 온라인 여행 커뮤니티의 리뷰 데이터를 수집하여 빅데이터 분석(키워드 분석, 사회연결망분석)을 수행하였다. 자료 수집은 래퍼기반 웹 크롤러를 이용하였으며 수집된 리뷰 데이터를 바탕으로 품사 추출기와 형태소 분석기를 활용하여 명사와 형용사를 추출하였다. 추출된 명사와 형용사에 대하여 텍스트 데이터의 가중치 부여방법 중 하나인 TF-IDF(term frequency-inverse document frequency)를 적용하여 가중치를 산출한 후 가중치가 부여된 명사와 형용사를 토대로 사회연결망분석을 실시함으로써 관광지 이미지에 대한 인식을 측정하였으며 세부적인 연구의 절차는 <그림 1>과 같이 총 세 단계의 체계로 진행하였다.

첫 번째, 자료 수집 단계에서는 온라인 여행 커뮤니티 중 '트립어드바이저(tripadvisor.com)'에서 2004년 1월부터 2016년 12월31일까지 작성된 여행객 리뷰를 웹 크롤러를 통해 수집하여 전처리가 가능한 형태로 데이터베이스에 저장하였다(①,②).

두 번째, 텍스트 마이닝을 통한 키워드 분석 단계에서는 데이터베이스에 저장된 리뷰 데이터에 대하여 기능어(관사, 전치사, 조사, 접속사 등), 불용어, 구두점, 숫자 등을 제거하는 전처리 단계를 거쳤으며(③), 형태소 분석과 품사 분석을 통해 명사와 형용사를 추출하였다. 특히,



〈그림 1〉 온라인 리뷰 분석을 위한 연구 절차

이 단계에서는 어간추출분석(stemming analysis)을 수행하여 텍스트 처리의 효율성을 높이도록 하였다. 어간추출분석은 단어의 어근을 추출하여 접사를 제거함으로써 동일한 의미의 다른 형태의 단어를 하나의 단어로 처리하는 방법으로 예컨대, place, places 또는 exciting, excited 등과 같은 동일한 의미에서 파생되는 단어를 자연어 처리하여 하나의 단어로 그 결과 값을 반환하는 것이다[④]. 자연어 처리한 명사와 형용사는 비구조화 된 자료형태이기 때문에 분석에 적합하도록 구조화 자료 형태인 단어-문서 행렬(term-document matrix)로 변환하였으며, 선행연구를 바탕으로 관광지 이미지의 구성개념

에 해당하는 세부요인을 추출하였다[⑤]. 이후, 텍스트 가중치 산출법인 TF-IDF기법을 적용하여 해당 텍스트의 가중치를 산출하였다. TF-IDF는 하나의 리뷰에서 특정 단어의 중요도는 그 리뷰 내의 출현 빈도에 대하여 비례하고 전체 리뷰에서 나타난 단어의 출현 빈도와는 반비례하도록 가중치를 산출하는 방법이다[⑥]. 다음으로 TF-IDF 가중치가 부여된 텍스트의 결과 값들의 비교를 위해 빈도분석을 실시하였다[⑦].

세 번째, 가중치가 부여된 관광지 이미지 세부요인(텍스트)을 바탕으로 빅데이터 분석기법 중 하나인 사회연결망분석(SNA: social network analysis)의 중심성 분석(centrality analysis)

을 수행하여 관광지 이미지의 구조적 관계를 파악하였다. 중심성은 네트워크 각각의 개체에 의해서 지속되는 관계(relation)의 수를 계산함으로써 측정되는 것으로, 어떠한 개별요인이 네트워크 내에서 중요한 역할을 하는지를 네트워크 내에서 구성요인(액터)이 차지하는 중심적 위치의 관점에서 설명하는 것이다(Hanneman & Riddle, 2005; Knoke & Yang, 2008; Prell, 2012). 중심성을 측정하는 방법은 연결정도 중심성(degree centrality), 근접 중심성(closeness centrality), 매개 중심성(betweenness centrality)의 세 가지 유형이 있으며(Galaskiewicz & Wasserman, 1994), 본 연구에서는 3가지의 중심성 측정방법과 이에 더하여 연결정도 중심성의 확장된 개념인 아이겐벡터 중심성(eigenvector centrality)을 포함한 4가지 측정방법을 통해 중심성 분석을 수행하였다(⑧).

2. 조사 설계

1) 변수의 측정항목 구성

텍스트 마이닝을 통한 키워드 분석 과정에서는 단어를 문맥에 대치시키기 위해 단어가 가지는 실제 개념의 범주를 선정하는 것이 필수적이다(Dickinger & Költringer, 2011). 본 연구에서는 그 동안 관광지 이미지와 관련 연구 가운데 해당 구성요인 및 하위요인을 탐색적으로 규명하거나 구성개념간의 구조적 관계를 밝힌 심영석·김홍범(2016), Beerli and Martin(2004), Mak(2017), Martín and Bosque(2008), Stylidis, Shani and Belhassen(2017)의 연구를 바탕으로 여행객이 남긴 온라인 리뷰의 텍스트 마이닝 분석을 통해서 문맥의 대치를 수행하였으며, 이에 앞서 5개의 선행연구에 대한 관광지 이미지 구성개념을 <표 3>과 같이 재정립하였다. 관광지의 인지적 이미지의 구성요인은

자연 자원, 사회적(일반) 기반시설, 관광객 기반 시설, 여행객의 여가 및 레저 활동, 문화·역사 및 예술, 사회·경제적 요인, 자연적 환경, 사회적 환경, 장소의 분위기 등으로 크게 나뉘어져 측정되어 왔으며 최근에는 이에 더하여 정보적 자원, 먹거리, 교통수단(접근성) 등이 구성요인으로서 확립되어 인지적 이미지를 측정하였다. 이러한 맥락에 따라 정리된 내용을 살펴보면 관광지의 인지적 이미지는 ‘자연 자원’, ‘기반시설’, ‘여가·홍미적 요소’, ‘교통(접근성)’, ‘먹거리’, ‘문화·역사 및 예술’, ‘사회경제적’, ‘사회적 환경’, ‘정보자원’, ‘자연환경’, ‘관광지에서의 활동’, ‘적절한 일정/계획’과 같이 총 12개의 구성요인으로 정리될 수 있었다. 관광지의 정서적 이미지는 Russell, Ward and Pratt(1981)가 제안한 자극적, 흥미로운, 기쁜, 편안하게 하는, 피곤한, 우울한, 불편한, 괴로움을 주는 등의 구성요인을 바탕으로 그동안 측정되어져 왔으며 이와 더불어 최근 연구에서는 감명 깊은, 믿을만한, 환기되는, 명성이 있는 등의 요인이 정서적 이미지의 구성요인으로 규명되었다. 따라서 관광지의 정서적 이미지에 대한 구성요인은 ‘자극적인’, ‘흥미로운’, ‘기쁜’, ‘편안하게 하는’, ‘불편한’, ‘감명 깊은’, ‘최신 유행의’, ‘믿을만한(진정한)’, ‘환기되는’, ‘명성이 있는’과 같이 총 11개로 정리될 수 있었다.

앞서 정리된 관광지 이미지의 구성요인과 이를 설명하는 세부요인을 근거로 본 연구에서 수집한 여행객이 남긴 온라인 리뷰를 대치시켰으며 <표 4>와 같이 정리될 수 있다. 항목 수를 통해 인지적 이미지에 대하여 ‘사회적 환경’, ‘문화·역사 및 예술’, ‘기반시설’, ‘교통수단(접근성)’, ‘먹거리’ 등의 순으로 여행객이 그 중요성을 인식하고 있다는 것을 잠정적으로 추측해 볼 수 있다. 정서적 이미지는 ‘기쁨을 주는’에 해당하는 구성요인의 항목이 가장 높게 나타났으며 이 외의 구성요인은 고루 분포되어 있어 여행객은 관광지에

〈표 3〉 관광지 이미지의 구성요인^a

구성요인	세부요인
<i>Cognitive Image</i>	
natural resources	scenery, landscape, natural park, river, beaches, lakes, mountains, flora & fauna
infrastructure	general infrastructure(roads, commercial infrastructure, extend of building development, city & towns), tourist infrastructure(hotel, restaurant, shopping facilities, accommodation)
recreational entertainment	theme parks, zoo, aquarium, outdoor activity, festival, ceremony, spa & wellness
transportation(accessibility)	motorcycle, train, car/bus, airport, station, route
food & beverage	restaurant, food, dinner, lunch, breakfast, coffee, tea, meal, beverage, desserts, fruits, bread/cake, vegetable
cultural, history & arts	cultural attractions, cultural activities, traditional attractions, general arts, temple, museum, farm
socioeconomic	economic benefit, good value for money
social environment	local people, service, language barriers, social characteristic, safety
information resources	general information, tourists information sign, traffic sign, direction
natural environment	weather, atmosphere, seasonal characteristic
specific activities	shopping, walking/walk, eat/eating, hike/hiking, cycling, fishing, camping, land activity, water activity, relaxing
proper schedule/plan	attraction operating hours, travel schedule
<i>Affective Image</i>	
arousing	amazing, intense
exciting	interesting, exhilarating, exciting
distressing	tired, crowded, complex, waiting
pleasant	good, beautiful, great, nice, pretty, best, fun, enjoy/enjoyable, pleasant
relaxing	relaxed/relaxing, refreshing, peaceful
unpleasant	bad, dangerous
impressive	impressive, special, worth
trendy	young, cool, modern, changing
authentic	unique, attractive
evoked	kind, missing, memorable
reputational	famous, popular, clean, convenience

주: a. 심영석·김홍범(2016), Beerli & Martin(2004), Mak(2017), Martin & Bosque(2008), Styliadis *et al.*(2017)의 연구를 바탕으로 연구자가 정리함.

〈표 4〉 관광지 이미지의 구성요인 - Tripadvisor -^a (2017. 01. 기준)

구성요인	텍스트 마이닝을 통해 분류된 세부요인	항목 수
<i>Cognitive Image</i>		
natural resources	park, garden, mountain, stream, river, hill, tree, nature, blossom, island, scenery, sunset, sight	13
infrastructure	street, tower, city, building, road, alley, district, path, store, market, mall, stall, hotel, room, accommodation	15
recreational entertainment	souvenir, ticket, cosmetic, ride, entertainment, ceremony, theater, aquarium, outdoor, trail, gift, festival, spa	13
transportation (accessibility)	car, station, center, location, bus, subway, metro, accessibility, train, nearby, taxi, driver, airport, distance	14
food & beverage	food, restaurant, cooking, tea, drink, coffee, seafood, lunch, taste, snack, dinner, meal, dish, cafe	14
cultural, history & arts	palace, museum, war, photo, temple, culture, art, picture, exhibition, hanok, music, university, gallery, hanbok, budda	15
socioeconomic	price, expense, fee, cost, money, purchasing, bargain, quality, discount, buying	10
social environment	foreigner, chinese, language, nation, country, brand, Japanese, Korean, resident, local, service, village, tourist, staff, style, fashionable, life, folk, guard	19
information resources	guide, english, information, booking, knowledge, map, itinerari, sign, direction, review	10
natural environment	season, summer, spring, autumn, hot, weather, winter, cold, atmosphere	9
specific activities	eating, viewing, playing, climbing, hiking, biking, walking, shopping, strolling	9
proper schedule/plan	time, day, hour, plan, afternoon, morning, weekend, week, schedule, night, holiday	11
<i>Affective Image</i>		
arousing	amazing, fantastic	2
exciting	interesting, exciting, surprise	3
distressing	crowded, complex, waiting	3
pleasant	great, good, nice, enjoyable, lovely, beautiful, wonderful, fun, pretty, awesome, pleasant, delicious, happy	13
relaxing	resting, refreshing(fresh), peaceful, comfort	4
unpleasant	bad, disappointment, tired	3
impressive	impressive, special, worth	3
trendy	young, cool, modern, changing	4
authentic	traditional, historical, unique, attractive, ancient, real	6
evoked	kind, missing, memorable	3
reputational	famous, popular, clean, convenience	4

주: a. 트립어드바이저에 나타난 실제 여행객의 의견을 토대로 도출된 200개 관광지 이미지 세부요인을 정리함.

서 다양한 정서적 감정을 드러내고 있다는 것 또한 추측해 볼 수 있다.

2) TF-IDF를 이용한 가중치 산출

관광지 이미지의 추출된 측정변수를 계량적으로 분석하기 위해 텍스트 데이터의 가중치 산출 방법인 TF-IDF(term frequency-inverse document frequency)를 적용하였다. TF-IDF는 열(term)과 행(document)의 수준에서 분석 단위가 이루어지며, 문서 d 에서 단어 w 에 대한 가중치 값을 산출하는 방식으로 해당 수식은 다음과 같다(Salton & Buckley, 1988).

$$TFIDF(w,d) = TF(w,d) \times \log\left(\frac{N}{DF(w)}\right)$$

〈식 1〉¹⁾

$TF(w,d)$: 문서 d 에 단어 w 가 나타나는 횟수

$DF(w)$: 단어 w 가 들어가는 문서의 총 수

N : 전체 문서의 총 수

TF는 term frequency의 약어로, 단어의 출현 횟수를 의미한다. IDF는 inverse document frequency의 약어로, 단어가 출현한 문서 수의 역수이다. 예컨대, 호텔의 서비스에 해당하는 문서를 분류할 때, 이 분류 문제에 출현할 주요 단어는 'hotel'과 'service'일 것이다. 만약, 주어진 문서의 집합(set)이 모두 호텔분야의 문서라면 hotel은 대부분의 문서에 출현할 것이다. 이에 반하여, 'service'는 호텔의 서비스 관련 문서에만 출현할 것이다. 따라서 'hotel'과 'service'의 두 단어의 중요도를 봤을 때 대부분의 문서에 출현하는 'hotel'은 문서 분류에 큰 효용이 없지만 일부 문서에만 출현하는 'service'는 큰 도움이 된다. 즉, 해당 단어의 중요도는 그 단어가 문서에 나온 총 횟수(term frequency)에 비례하고, 그 단어가 있는 모든 문서의 총 수에 반비례하여

하나의 문서 중 가장 가중치 값이 높은 단어가 그 문서의 키워드로 채택되는 것이다(심영석·김홍범, 2016). 이와 같이, TF-IDF는 개별문서에서 단어의 중요도를 표현할 수 있는 기법으로(이태원·홍태호, 2015), 본 연구에서는 TF-IDF를 적용하여 텍스트 데이터의 효율성을 제고함으로써 객관·타당한 계량분석이 가능하도록 하였다.

3) 사회연결망 중심성 분석

사회연결망분석의 목적은 구조나 연결망 형태의 특징을 도출하고, 관계성을 바탕으로 체계의 특성을 설명하거나 체계를 구성하는 단위의 행위를 설명하는 것이다(김용학, 2007). 이러한 사회연결망분석에서 관계 및 구조적 특성을 분석하기 위해 널리 사용되고 있는 기법으로 중심성(centrality), 밀도(density), 구조적 공백(structural hole), 중심화(centralization) 등이 있으며, 이 가운데 중심성은 하나의 노드(node), 즉 한 개체가 전체 사회연결망의 구조에서 중심에 위치하는 정도를 표현하는 광의적인 지표와 네트워크에서 개체와 개체의 인접과 영향을 나타내는 협의적 지표로 접근할 수 있다(Stokowski, 1994). 중심성의 대표적인 측정 방법은 연결정도 중심성(degree centrality), 매개 중심성(betweenness centrality), 근접 중심성(closeness centrality), 아이겐벡터 중심성(eigenvector centrality)이 있다(Freeman, 1979; Kim, Choi, Yan & Dooley, 2011).

연결정도 중심성은 해당 노드가 네트워크 내의 다른 노드들과 연결되어 있는 링크의 합을 나타내는 개념으로 연결정도 중심성이 높은 개체는 많은 다른 개체들과 직접적인 관계를 맺고 있으므로 원하는 정보를 획득할 확률이 높고 네트워크 내에서의 영향력이 크다고 볼 수 있다(Knoke & Burt, 1983).

1) 심영석·김홍범(2016)의 연구에서 직접 인용함.

근접 중심성은 네트워크 내의 각 노드로부터 다른 모든 노드까지의 거리를 측정하여 한 노드가 다른 노드들에 얼마나 가까이 위치해 있는지를 나타내는 개념이다. 만약 두 노드의 연결성이 없는 경우는 거리를 산출할 수 없으므로 근접 중심성 산출 시 0의 값이 반환되며, 근접 중심성이 높은 개체는 네트워크 내의 다른 개체들과 가까운 거리에 위치하므로 주요 정보를 용이하게 획득하고 영향력을 확보한다고 볼 수 있다(Hanneman & Riddle, 2005). 예를 들어 관광의 맥락에서, 관광지 A가 수많은 도달 가능한 경로의 다른 목적지를 가지고 있으며 도달 가능한 목적지와와의 거리가 가까이 떨어져 있다면, 관광지 A는 다른 모든 관광목적지에 중심이 되고 가까이 위치하게 된다. 이 경우 관광지 A는 가장 짧은 경로로 모든 관광지에 도달 할 수 있게 되며 높은 근접 중심성을 가진다고 할 수 있다(Shih, 2006).

매개 중심성은 해당 노드가 네트워크 내의 다른 노드들을 연결하는 최단 거리 경로에 존재하는 빈도를 측정하는 것으로(Scott, 2000), 한 노드가 네트워크 내에서 담당하는 중재자(broker)역할을 측정하는 것이자 최단 경로의 독점 정도를 측정하는 것이다(손동원, 2002). 매개 중심성이 높은 개체는 네트워크 내의 다른

개체들이 해당 노드를 중재자로서 의존하는 정도가 크다는 관점에서 영향력이 큰 개체라 볼 수 있다. 이러한 지표를 관광지 네트워크에 적용하게 되면, 여행객이 다양한 관광지를 여행하는 동안 대부분의 여행객이 관광지 A에서 멈추게 될 경우, 다른 관광지의 쌍 사이에서 관광지 A는 매우 중요한 중재자가 되며, 이 경우 관광지 A는 높은 매개 중심성을 가진다고 할 수 있다.

아이겐벡터 중심성은 연결된 노드의 중요성에 가중치를 부여하여 노드의 중심성을 측정하는 방법으로 연결정도 중심성이 노드 간 연결된 다른 노드의 개수에 초점을 맞춘다면, 아이겐벡터 중심성은 연결된 노드의 개수뿐만 아니라 연결된 노드가 얼마나 중요한지도 함께 고려하는 것으로써 연결정도 중심성의 개념을 확장한 것이라 할 수 있다(Bonacich, 2007, 1987). 예를 들어, 영향력이 높은 사람과의 단 하나의 연결이 다른 평범한 사람들과의 관계보다 자신의 영향력(파워)을 증가시킨다는 것이 곧 아이겐벡터 중심성이 높다는 것을 의미하며, 자신의 연결정도 중심성으로부터 발생하는 영향력과 자신과 연결된 타인의 영향력을 합하여 아이겐벡터 중심성을 결정하는 개념이다.

본 연구에서는 중심성 분석을 통해 여행객이 인식하는 관광지 이미지를 조사하였으며, 중심성

〈표 5〉 사회연결망 중심성 분석의 유형 및 정의^a

구분	관광지 이미지 분석에서의 중심성 분석 결과의 해석
연결정도 중심성	값이 높을수록 해당 변수는 다른 변수와의 관계성이 높다는 것과 여행객이 인식하는 관광지 이미지에 직접적인 영향을 주는 요소로 해석
근접 중심성	값이 높을수록 해당 변수는 다른 변수와 쉽게 연결되어 나타날 수 있다는 것과 여행객의 인식에 다른 관광지 이미지의 구성개념과 어우러졌을 때 시너지효과를 낼 수 있는 요소로 해석
매개 중심성	값이 높을수록 해당 변수는 다른 변수들이 나타날 때 중개 역할이 상대적으로 많다는 것과 이는 곧 여행객들의 인식에 의존성이 큰 요소로 해석
아이겐벡터 중심성	값이 높을수록 해당 변수는 실질적으로 관광지 이미지에 대한 여행객들의 인식에 큰 파급력을 가지는 것으로 해석

주: a. 연구자 정리.

분석의 세부적인 개념인 연결정도 중심성, 근접 중심성, 매개 중심성, 아이젠벡터 중심성을 <표 5>와 같은 개념으로 적용하였다.

3. 표본 및 자료수집

본 연구에서는 글로벌 온라인 여행 커뮤니티인 '트립어드바이저(tripadvisor.com)'에 작성된 여행객의 리뷰를 웹 크롤러를 이용하여 수집함으로써 분석단위로 활용하였다. 자료수집 대상의 선정 이유는 다양한 온라인 여행 커뮤니티 중 트립어드바이저가 정보전달 측면에서 그 유용성 및 신뢰성이 높다는 연구들이 많은 학자들에 의해 보고됨에 따라(Li *et al.*, 2013; Liu, Schuckert & Law, 2015) '트립어드바이저'에서 획득할 수 있는 온라인 관광정보에 주목할 필요성이 강조되기 때문이다. 또한, Bjørkelund, Burnett and Nørvg(2012)는 '부킹닷컴(booking.com)'과 '트립어드바이저'가 여행객에게 주는 정보에 대한 장단점을 분석한 연구에서 '트립어드바이저'의 고객 의견이 특히 여행객의 의사결정에 큰 영향을 미친다고 주장하였으며 이러한 맥락에서 '트립어드바이저'의 관광정보는 분야에서 활용가능성이 상당히 높다는 것을 알 수 있다. 따라서 본 연구에서는 자료수집대상을 '트립어드바이저'로 한정하였으며, 세부적으로 관광객이 직접적으로 인식하는 관광지 이미지의 정확한 측정을 위해 서울의 관광지 및 관광 상품에 대한 온라인 리뷰를 조사·분석에 활용하였다.

자료수집은 래퍼 기반 웹 크롤러(web crawler)를 이용하여 서울의 관광지 및 관광상품, 온라인 리뷰에 대한 내용을 2016년 12월 29일부터 2017년 1월 6일까지 수집하였으며, 수집된 내용은 2004년 8월부터 2017년 1월 6일까지 작성된 관광지 및 관광 상품에 대한 12개의 카테고리(명소/랜드마크, 박물관, 자연/공원, 쇼핑, 나이트라이프, 야외활동, 투어, 콘서트/쇼, 음식/음

료, 위락시설, 스파/웰니스, 여행자리소스)에 해당하는 온라인 리뷰 전체이다. 본 연구에 활용된 분석범위는 <표 6>과 같이 각 관광지 및 관광상품에 대하여 누락된 리뷰가 발생하지 않는 시점인 2012년 1월을 시작으로 2016년 12월 31일까지의 12개의 카테고리에 해당하는 117개의 관광지 및 관광 상품에 대한 34,746개의 영문 리뷰이다. 세부적으로, 반어법으로 인해 발생하는 편향(bias)을 통제하고 형태소 분석과정에서 명사와 형용사의 구조적 배치가 명확한 영문리뷰를 분석단위로 활용함으로써 텍스트 분석결과에 대한 타당성을 확보하고자 하였다. 또한 자료수집 방법에 해당하는 웹 크롤러는 빅데이터 분석 과정의 자료수집 단계에서 활용되는 수단으로서, 방대한 양의 웹 문서를 자동으로 수집할 수 있도록 하는 중요한 기법이다(김광영·이원구·이화목·신성호·이민호, 2011). 이러한 웹 크롤러 중 관광지의 평가와 같이 특정한 주제로 작성되어 있는 웹 문서를 정확하게 수집하기 위해서는 래퍼 기반 웹 크롤러를 사용한다(강한훈·유성준·한동일, 2010). 래퍼 기반 웹 크롤러의 특징은 본 연구에서 활용된 트립어드바이저와 같이 수집 대상이 되는 웹 사이트의 전송방법, url의 구조 등을 사전에 분석하여 해당 프레임에 맞도록 웹 크롤러를 설계함으로써 해당 웹 페이지에 접근하여 사용자가 원하는 데이터를 정확하게 수집할 수 있다(Kang, Yoo & Han, 2009). 이러한 웹 크롤러를 통해 자료를 수집한 관광분야의 빅데이터 연구는 도해용·류기상(2014)이 미국 레스토랑 평가 사이트인 엘프닷컴(yelp.com)의 고객리뷰를 수집하여 외식분야의 브랜드 개성을 측정한 연구가 있으며, Schukert, Liu and Law(2015)는 '트립어드바이저'의 홍콩 185개 호텔을 대상으로 고객평가(rating)를 수집하여 호텔산업의 고객만족 결정요인을 파악하였다. 심영석·김홍범(2016) 또한 '트립어드바이저'의 서울의 관광지를 대상으로 관광지 이미지의 구성요인을 규명하였으며, 이외에

〈표 6〉 서울의 관광지 및 관광상품 목록^a

(2017. 01. 기준)

구분	관광지 및 관광상품		
명소/랜드마크	경복궁	인사동	동대문디자인플라자
	명동	북촌	조계사
박물관/전시관	창덕궁	홍익아트거리	
	남산N서울타워	봉은사	
	한국전쟁기념관	덕수궁	한가람박물관
	국립중앙박물관	국립민속박물관	국립현대미술관서울관
자연/공원	삼성미술관 리움	서울역사박물관	
	트립아이 박물관	서대문형무소역사관	
	북한산국립공원	여의도공원	여의도한강공원
	한강공원	동대문역사문화공원	서울숲
쇼핑/마켓	창계천	낙산공원	
	남산공원	서울어린이대공원	
	남대문시장	이태원	강남지하상가
	롯데백화점본점	신세계백화점본점	타임스퀘어몰
나이트라이프	노랑진시장	홍대프리마켓	
	스타필드코엑스몰	동대문쇼핑몰	
	서울곤와월드	울렛제츠	마익스캐빈
	바우드스탁	르챔버	리술베르시가샵&시가라운지
야외활동	오케이퐁퐁	골목바이널&펍	
	믹스&몰트	클럽옥타곤	
	잠실야구장	명동사격장	아띠인력거
	서울에스케이프프롬홍대점	스윗트래블코리아	이랜드크루즈
투어	메가박스 코엑스몰	원데이코리아	
	노이스케이프	올림픽공원	
	탐코리아투어	투어스바이아론	DMZ스파이투어
	익스클루시브투어코리아	얼티미트코리아투어	코스모진투어
콘서트/쇼	컬러오브코리아	노바랜드투어스코리아	
	서울시티투어	코리도어	
	명동난타극장	충정난타극장	드럼켓
	예술의전당	클라이브	오페라하우스자유소극장
음식/음료	점프!코믹 무술 퍼포먼스	세종문화회관	
	홍대난타극장	정동극장	
	디스이즈코리아	오미요리연구소	딜렉터블트래블스
	온고코리아컬리널리	HaB코리아	자넷쿠킹스튜디오
위락시설	온고푸드커뮤니케이션즈	젠김치코리아안푸드	
	히어코리아	서울푸드투어	
	롯데월드	테지움서울	
	코엑스아쿠아리움	롯데월드아쿠아리움	
스파/웰니스	디큐브시티	63시월드	
	롯데월드스타에비뉴	디보빌리지	
	실로암스파	SK-II 부띠끄스파	레비쉬스파
	캣카페오양이놀이터	스파렉스	더스파그랜드하얏트서울
여행자 리소스	드래곤힐스파&리조트	해피데이스파	
	드래곤힐	스파레이	
	한국관광공사관광안내소	인사동홍보관광안내소	서울도서관
	코엑스센터	K-Style Hub	
공항철도트래블센터	명동관광정보센터		
서울글로벌문화정보센터	한국지역정보센터		

주: a. 자료수집시점인 2017. 01월 기준으로 트립어드바이저에 등록된 서울의 관광지 및 관광상품 목록임.

도 빅데이터를 활용한 연구에서는 자료수집 단계에서 웹 크롤러를 통해 자료를 수집하여 분석을 수행한 체계를 따르고 있다(Wang, Chan, Ngai & Leong, 2013).

IV. 분석 결과

1. 키워드(key-word) 분석결과

수집된 34,746개의 리뷰를 분석하기 위해 텍스트 마이닝 기법을 활용하여 자연어처리(기능어 제거, 구두점과 공백, 숫자 제거)를 수행하였으며 형태소분석을 통해 명사, 형용사를 추출한 결과 37,753개의 단어가 자연어처리 되었다. 그리고 어간추출분석(stemming analysis)을 수행하여 동일한 의미의 형태가 다른 언어를 하나의 단어로 처리하였고, 그 결과 31,846개의 단어로 축소되었다. 이를 바탕으로 텍스트 데이터의 객관·타당한 계량적 분석이 가능하도록 TF-IDF 가중치 산출법을 적용하였으며, 선행 연구를 바탕으로 관광지 이미지의 구성요인에 대하여 확립한 분류기준에 따라 200개의 관광지 이미지 변수를 추출하여 키워드 분석을 수행하였다. <표 7>은 텍스트 마이닝 결과에 따른 키워드 분석과 TF-IDF가 적용된 단어들의 키워드 분석에 대한 상위 60개에 해당하는 결과이다. 분석결과에서 나타난 바와 같이, 단어빈출(빈도수)에 기반 한 키워드 분석과 텍스트에 가중치가 부여된 값과의 순위는 다소 차이가 있음을 알 수 있다. 두 분석결과와의 차이를 통하여 키워드 분석을 수행할 때는 리뷰 내에서 해당 단어의 단순 언급(빈도 기반)수준의 분석보다 전체 문서에 대비하여 해당 단어의 중요도를 고려한 값을 결과에 반영해야 키워드 분석을 보다 정확하게 측정할 수 있다는 것을 파악할 수 있으며, 따라서 본 연구에서 적용한 TF-IDF 가중치 산출법은 분

석단위로써 질적 데이터인 텍스트 데이터를 객관·타당한 수준의 값으로 변환할 수 있는 간단하지만 성능은 우수한 기법임을 알 수 있다(이태원·홍태호, 2015).

텍스트에 가중치가 부여된 TF-IDF에 기반한 키워드 분석결과를 살펴보면, 'shopping', 'walking', 'viewing' 등과 같이 관광지에서의 특정한 활동과 관련된 키워드의 언급이 높게 나타남을 알 수 있다. 또한, 'palace', 'museum', 'street', 'tower', 'city', 'park' 등과 같은 관광지의 매력물과 관광지의 일반적인 기반시설과 관련된 키워드와 'day', 'time', 'night', 'hour' 등의 여행객이 관광지에서 보낼 수 있는 일정과 관련된 키워드 또한 그 언급도가 높은 것으로 나타났다. 최근 관광지 이미지 연구에서 하나의 구성요인으로 포함된 먹거리(Mak, 2017)와 관련된 'food', 'restaurant' 등의 키워드도 여행객이 관광지의 이미지를 인식하는 중요한 키워드로 나타났다. 흥미로운 것은 'resident', 'Korean' 등과 같이 관광지에서의 인적서비스 및 사회적 환경을 대변하는 키워드의 언급수준이 높은 것으로 나타났으며 이러한 결과를 통해 관광지에서의 지역민의 역할이 여행객들이 관광지의 이미지를 인식할 때 중요하게 작용할 수 있다는 것을 추측해 볼 수 있다.

더불어, 관광지의 전반적인 분위기에 대해 느끼는 감정인 'great', 'good', 'nice', 'beautiful', 'historical', 'traditional' 등의 키워드와 관광지의 고유한 특색을 경험하고 느낄 수 있는 감정인 'worth', 'memorable' 등의 키워드도 여행객이 관광지의 이미지를 인식하는데 중요하게 작용한다는 것을 추측해 볼 수 있다. 키워드 분석을 종합해 보면 여행객들이 관광지에서 느끼는 관광지의 자원, 매력 등 물리적 속성에 대해 느끼는 인지적 이미지와 관광지에서의 경험을 바탕으로 관광지에 대한 생각과 느낌을 대변하는 심리적 상태인 정서적 이미지가 균형 있게 인식되고 있

〈표 7〉 텍스트 마이닝 분석 결과 비교

Rank	빈도기반 키워드분석(Stemming words) ^a (n=34,746)				가중치기반 키워드분석(TF-IDF words) ^b (n=34,746)			
	Variables	Freq	Variables	Freq	Variables	Amount	Variables	Amount
1	shopping	15,319	46 art	2,353	1 shopping	31,339.8	46 culture	10,566.3
2	great	11,972	47 souvenir	2,336	2 palace	29,477.7	47 memorable	10,403.6
3	time	11,839	48 location	2,199	3 korean	24,813.1	48 art	10,237.2
4	korean	11,731	49 ticket	2,183	4 resident	24,383.1	49 temple	10,104.6
5	palace	10,532	50 eating	2,170	5 great	23,711.9	50 souvenir	9,891.5
6	walking	10,238	51 resting	2,166	6 time	23,690.8	51 ticket	9,678.3
7	good	10,126	52 picture	2,131	7 museum	22,790.8	52 exhibit	9,459.0
8	day	8,780	53 information	2,090	8 walking	22,665.0	53 guard	9,325.4
9	food	8,578	54 guard	2,085	9 food	22,208.0	54 eating	9,222.7
10	nice	8,026	55 exhibit	2,073	10 day	21,946.5	55 changing	9,206.3
11	street	7,497	56 temple	1,995	11 good	21,759.2	56 picture	9,196.4
12	museum	7,489	57 memorable	1,939	12 street	21,061.4	57 center	9,176.9
13	viewing	7,210	58 wonderful	1,875	13 lovely	20,967.4	58 location	9,127.8
14	beautiful	6,633	59 missing	1,785	14 viewing	20,859.9	59 bus	9,054.6
15	lovely	6,317	60 attractive	1,767	15 tower	20,364.5	60 information	8,914.2
16	enjoy	6,016	61 pretty	1,761	16 nice	20,027.9	61 cafe	8,435.5
17	guide	6,004	62 cosmetic	1,700	17 guide	18,530.3	62 wonderful	8,238.2
18	historical	5,961	63 booking	1,684	18 beautiful	18,270.5	63 cosmet	8,190.2
19	tower	5,798	64 subway	1,680	19 enjoy	16,732.6	64 stream	7,961.2
20	city	5,426	65 bus	1,624	20 city	16,553.6	65 attractive	7,932.6
21	interest	5,309	66 cafe	1,574	21 interest	15,916.9	66 booking	7,927.2
22	worth	4,925	67 impress	1,518	22 tourist	15,446.6	67 pretty	7,892.6
23	resident	4,872	68 stream	1,490	23 night	15,312.6	68 missing	7,866.2
24	night	4,548	69 awesome	1,386	24 historical	15,276.4	69 subway	7,761.6
25	restaurant	4,480	70 plan	1,366	25 park	15,079.7	70 hiking	7,491.6
26	hour	4,243	71 hotel	1,356	26 worth	14,631.3	71 cooking	7,217.6
27	fun	4,061	72 mall	1,326	27 restaurant	14,401.3	72 impress	7,167.3
28	park	3,900	73 service	1,286	28 fun	14,046.4	73 young	7,011.7
29	building	3,841	74 modern	1,284	29 hour	13,776.7	74 mall	7,004.0
30	traditional	3,655	75 ride	1,244	30 building	13,659.0	75 hotel	6,763.5
31	tourist	3,644	76 strolling	1,240	31 traditional	13,479.5	76 service	6,703.5
32	amazing	3,303	77 kind	1,236	32 photo	13,316.7	77 awesome	6,693.7
33	market	3,158	78 cool	1,220	33 war	13,219.6	78 plan	6,654.4
34	local	3,149	79 room	1,205	34 market	12,952.7	79 room	6,609.0
35	english	3,145	80 special	1,193	35 garden	12,690.7	80 ride	6,482.6
36	price	3,005	81 stall	1,183	36 amazing	11,982.3	81 modern	6,372.5
37	crowded	2,998	82 waiting	1,181	37 local	11,949.0	82 village	6,273.3
38	war	2,998	83 village	1,180	38 car	11,684.9	83 strolling	6,191.5
39	buying	2,977	84 young	1,174	39 store	11,672.4	84 stall	6,164.7
40	garden	2,939	85 clean	1,158	40 english	11,538.4	85 waiting	6,120.2
41	car	2,833	86 brand	1,101	41 price	11,446.1	86 cool	6,109.4
42	store	2,824	87 fantastic	1,093	42 buying	11,358.5	87 kind	6,088.5
43	culture	2,654	88 nation	1,079	43 crowded	11,292.2	88 special	5,966.3
44	photo	2,552	89 afternoon	1,067	44 resting	11,116.2	89 tea	5,896.2
45	station	2,551	90 tea	1,064	45 station	10,624.8	90 brand	5,853.3

주: a. 어간추출분석(stemming analysis)을 통하여 동일한 의미의 다른 형태(예컨대, place, places 또는 exciting, excited 등)의 단어를 하나의 단어로 처리함으로써 그 결과 값을 반환함.

b. Stemming words에 대하여 TF-IDF의 방법으로 가중치를 산출한 값이며, $TFIDF(w,d) = TF(w,d) \times \log(\frac{N}{DF(w)})$ 임.

다는 것을 알 수 있다.

2. 사회연결망 중심성 분석결과

도출된 관광지 이미지 변수에 대하여 연결정도 중심성, 근접 중심성, 매개 중심성, 아이젠벡터 중심성 분석을 수행한 결과는 <표 8>과 같이 정리될 수 있으며, 결과에 대한 해석은 각각의 중심성 분석결과의 상위 10위에 해당하는 내용을 기준으로 살펴보고자 한다. 먼저, 개체 간의 직접적인 관계를 나타내는 정보인 연결정도 중심성(degree centrality)은 'time', 'Korean', 'great', 'shopping', 'walking', 'good', 'day', 'resident', 'food', 'nice'의 순으로 네트워크 내에서 영향력이 큰 것으로 나타났다. 이러한 결과는 여행객의 일정과 관련된 구성개념, 관광지의 사회적 환경을 나타내는 구성개념, 관광지에서의 특정한 활동과 관련된 구성개념, 관광지에서 느낄 수 있는 기쁨의 정서와 관련된 구성개념, 먹거리와 관련된 구성개념 등이 여행객이 인식하는 관광지 이미지에 직접적인 영향을 주는 요소이며, 키워드 분석 결과와 비슷하게 나타났음을 알 수 있다.

다음으로 네트워크의 개체 간의 관계에서 가까이 위치하여 개체를 연결하는 준거점의 역할을 측정하는 근접 중심성(closeness centrality) 분석 결과, 'war', 'aquarium', 'garden', 'cosmetic', 'historical', 'night', 'brand', 'palace', 'bargain', 'shopping'의 순으로 개체 간 연결가능성이 큰 준거점인 것으로 나타났다. 이러한 결과는 관광지의 문화·역사 및 예술과 관련된 구성개념과 여가 및 레저 활동과 관련된 구성개념, 자연자원과 관련된 구성개념, 사회경제적 구성개념, 일정과 관련된 구성개념, 여행지에서의 특정한 활동과 관련된 구성개념 등이 관광지 이미지에서의 중심에서 다른 구성개념과 어우러져 시너지 효과를 발휘함으로써 이미지의 긍정적 연출에

도움이 될 수 있는 요소가 될 수 있는 것으로 해석할 수 있다.

셋째, 네트워크 내에서 개체들 간의 중개 역할을 측정하는 매개 중심성(betweenness centrality) 분석 결과, 'park', 'guide', 'walking', 'river', 'market', 'station', 'knowledge', 'stream', 'driver', 'service'의 순으로 다른 개체들과의 관계에서 그 의존정도가 큰 것으로 나타났다. 흥미로운 것은 자연자원의 구성개념('park', 'river', 'stream')의 상위 값에서 차지하는 비중이 높은 것으로 나타났으며 여행객이 인식하는 관광지 이미지에서 연결고리가 가장 큰 요소로 해석해볼 수 있다. 또한, 접근성(교통수단)의 구성개념, 정보자원의 구성개념 등이 여행객들이 인식하는 관광지 이미지에 대하여 그 의존성이 큰 요소로 나타났다.

마지막으로, 네트워크의 개체 간 연결 개수뿐만 아니라 연결된 노드의 중요도를 측정하는 아이젠벡터 중심성(eigenvector centrality) 분석 결과, 'time', 'great', 'Korean', 'shopping', 'good', 'walking', 'day', 'food', 'resident', 'nice'의 순으로 나타났으며 상위 값들에 대해서는 연결정도 중심성의 결과와 비슷하게 나타난 것을 확인할 수 있다. 아이젠벡터 중심성은 연결정도 중심성의 개념을 확장한 형태의 개념이기 때문에 이와 같은 결과가 나타났음을 짐작해볼 수 있으며, 해석 상 연결정도 중심성이 네트워크 내의 개체들 간의 직접적인 링크의 합을 나타내는 개념으로써 여행객이 인식하는 관광지 이미지에 직접적인 영향을 주는 요소로 해석할 수 있다면, 아이젠벡터 중심성은 연결정도 중심성의 개념과 더불어 그 중요도를 파악할 수 있는 측정방법이기 때문에 관광지 이미지에 대해 여행객이 인식에 파급력을 지니는 요소로 해석할 수 있는데 그 차이가 있다. 따라서 여행객의 일정과 관련된 구성개념, 관광지의 사회적 환경을 나타내는 구성개념, 관광지에서의 특정한 활동과 관

〈표 8〉 관광지 이미지 중심성 분석 결과 - 사회연결망 분석 -^a

순위	연결정도 중심성 (degree centrality)	근접 중심성 (closeness centrality)	매개 중심성 (betweenness centrality)	아이겐벡터 중심성 (eigenvector centrality)				
1	time	14.64	war	3.21	park	9.83	time	37.84
2	korean	13.69	aquarium	3.18	guide	9.28	great	36.01
3	great	13.45	garden	3.18	walking	5.13	Korean	35.43
4	shopping	12.51	cosmetic	3.17	river	4.81	shopping	33.17
5	walking	12.40	historical	3.17	market	4.66	good	32.96
6	good	12.37	night	3.16	station	4.27	walking	32.08
7	day	11.16	brand	3.15	knowledge	3.66	day	29.40
8	resident	9.33	palace	3.12	stream	3.60	food	26.28
9	food	9.32	bargain	3.12	driver	3.51	resident	24.86
10	nice	8.95	shopping	3.11	service	3.17	nice	24.37
11	enjoy	8.59	art	3.10	food	3.16	enjoy	23.19
12	lovely	8.28	price	3.10	bad	3.10	lovely	22.57
13	beautiful	7.48	ceremony	3.09	hiking	2.86	street	20.76
14	street	6.90	crowded	3.08	hot	2.60	beautiful	20.20
15	guide	6.83	atmosphere	3.08	resting	2.56	guide	19.47
16	worth	6.55	cafe	3.08	itinerary	2.41	interest	18.32
17	interest	6.49	fashionable	3.08	restaurant	2.17	restaurant	18.17
18	palace	6.43	cool	3.07	tower	2.15	worth	18.00
19	restaurant	6.41	guard	3.07	nice	2.15	palace	17.83
20	city	6.27	exhibit	3.07	shopping	2.13	viewing	17.36
21	viewing	6.24	hotel	3.07	seafood	2.07	city	16.95
22	hour	6.02	entertainment	3.06	booking	2.06	hour	16.43
23	museum	5.50	fun	3.06	museum	1.94	museum	16.18
24	tourist	5.29	mall	3.05	street	1.89	tourist	15.28
25	historical	4.55	viewing	3.05	traditional	1.88	historical	14.55
26	night	4.41	music	3.04	train	1.87	night	13.76
27	english	3.87	changing	3.04	beautiful	1.72	traditional	12.14
28	tower	3.84	exciting	3.04	souvenir	1.67	fun	11.96
29	fun	3.70	building	3.03	cooking	1.60	english	11.94
30	building	3.68	street	3.03	store	1.60	local	11.61
31	local	3.63	district	3.02	path	1.52	tower	11.55
32	traditional	3.61	gallery	3.02	tourist	1.38	building	11.26
33	photo	3.28	store	3.02	bus	1.35	photo	10.64
34	buying	3.04	hour	3.01	palace	1.23	amazing	10.44
35	price	3.00	memorable	3.01	island	1.22	culture	10.41
36	amazing	3.00	museum	3.01	city	1.17	buying	10.12
37	culture	2.84	alley	3.01	car	1.14	crowded	9.91
38	crowded	2.80	young	3.00	viewing	1.11	price	9.87
39	park	2.77	ancient	2.99	tea	1.06	eating	8.77
40	car	2.59	good	2.99	ticket	1.05	car	8.55
41	station	2.54	time	2.99	tree	1.04	park	8.48
42	eating	2.37	expense	2.99	festival	0.99	market	8.29
43	market	2.36	university	2.99	accommodation	0.95	store	8.23
44	store	2.35	weekend	2.99	korean	0.95	location	8.11
45	location	2.24	amazing	2.98	cosmetic	0.89	station	8.05
46	information	2.21	culture	2.98	ride	0.88	information	7.94
47	changing	2.11	gift	2.98	war	0.81	souvenir	7.72
48	souvenir	1.99	great	2.98	building	0.79	changing	7.32
49	resting	1.94	impress	2.98	price	0.78	resting	7.14
50	ticket	1.91	korean	2.98	blossom	0.78	wonderful	6.99

주: a. 관광지 이미지에 대한 200개의 세부요인 중 각각의 중심성 분석에 대한 상위 50개의 값을 정리한 결과임.

련된 구성개념, 관광지에서 느낄 수 있는 기쁨의 정서와 관련된 구성개념, 먹거리와 관련된 구성개념 등이 여행객이 관광지 이미지를 인식하는데 있어 그 파급력이 큰 요소라 할 수 있다.

3. 분석결과의 정리

네트워크 중심성 분석을 통해 도출된 관광지 이미지의 세부요인 결과 값과 평균을 바탕으로 관광지 이미지의 구성개념에 대한 중요도를 정리한 결과가 <표 9>, <표 10>에 제시되어 있다. 첫째, 연결정도 중심성의 평균값은 2.046이며, 이를 기준으로 관광지 이미지의 구성요인 중 전체 평균값보다 높은 '기쁨', '관광지에서의 활동', '적절한 일정/계획', '감명 깊은', '흥미로운', '사회적 환경', '기반시설'의 7개의 관광지 이미지 구성요인이 관광지 이미지에 대하여 여행객의 인식에 직접적으로 영향을 미치는 것을 알 수 있다. 이러한 결과는 관광지에서 여행객 스스로가 주도하여 일정을 계획하고 그 가운데 즐길 수 있는 다양한 활동과 이를 수반할 수 있는 기반시설, 사회적 환경이 여행객의 인식에 기쁨과 깊은 감명으로 위치할 때 관광지에 대한 긍정적 이미지 창출에 도움이 된다는 것을 시사하는 내용이다.

둘째, 근접 중심성의 평균값은 2.810이며, 이를 기준으로 '최신 유행의', '문화·역사 & 예술', '자극적인', '흥미로운', '환기되는', '감명 깊은', '믿을만한(진정한)', '여가·흥미적 요소', '사회적 환경', '괴로움을 주는', '사회경제적', '기쁨', '기반시설', '관광지에서의 활동'의 14개의 관광지 이미지 구성요인이 여행객이 인식하는 관광지 이미지에 있어 그 준거점 역할을 하는 것으로 나타났다. 결과를 살펴보면, 관광지의 정서적 이미지에 해당하는 다수의 구성요인이 전체 평균값보다 높게 나타나 역할의 강도가 인지적 이미지보다 크다는 것을 확인할 수 있다. 결과를 바탕으로 정리하면, 14개의 이미지 요인이 여행객과 심리적

으로 가까운 거리에서 상호·유기적으로 어우러져 관광지 이미지의 긍정적 연출에 도움이 되고 있다는 것을 알 수 있다. 더불어 관광지 이미지의 부정적 인식에 영향 미치는 '괴로움을 주는'의 값이 전체평균보다 높은 것으로 나타났으며, 세부요인 중 관광지의 복잡함을 의미하는 crowded(3.084), complex(2.949)의 값으로 인하여 이와 같은 결과가 나타난 것을 추측해볼 수 있다. 따라서 관광지 운영에 있어 복잡함을 해소할 수 있는 다양한 고민이 필요할 것으로 판단된다.

셋째, 매개 중심성의 평균값은 0.708이며, 이를 기준으로 '정보자원', '자연 자원', '관광지에서의 활동', '기반시설', '불편한', '교통(접근성)', '먹거리', '편안하게 하는'의 8개의 관광지 이미지 구성요인이 여행객들이 인식하는 관광지 이미지에 대하여 의존성이 큰 것으로 나타났다. 매개 중심성의 결과를 통해 정서적 이미지에 비해 인지적 이미지의 구성요인이 전체 평균보다 높은 비중을 차지하고 있는 것으로 확인되며, 나아가 여행객에게 의존성이 큰 '정보자원', '자연자원', '관광지에서의 활동', '기반시설', '교통(접근성)', '먹거리' 등의 구성요인을 바탕으로 관광지의 경쟁력을 제고 한다면 긍정적인 이미지의 인식에 도움이 될 것이라 판단된다.

넷째, 아이겐벡터 중심성의 평균값은 6.550이며, 이를 기준으로 연결정도 중심성의 7개의 상위요인('기쁨', '관광지에서의 활동', '적절한 일정/계획', '감명 깊은', '흥미로운', '기반시설', '사회적 환경')에 더하여 '자극적인', '믿을만한(진정한)', '정보자원'의 10개의 관광지 이미지 구성요인이 여행객이 관광지 이미지를 인식하는데 있어 그 파급력이 높은 것으로 나타났다.

마지막으로, 중심성 분석결과의 모든 부분에 공통적으로 전체평균보다 높은 것으로 나타난 '관광지에서의 활동', '기반 시설'에 주목할 필요가 있으며 결과적으로 관광지 이미지의 긍정적 인식에 가장 중요하고 큰 영향을 주는 요소는 여행객

<표 9> 관광지 이미지 차원별 중심성 분석결과 요약 - Cognitive Dimension -^a

DIMENSION	연결정도 중심성 $\mu=2.046$	근접 중심성 $\mu=2.810$	매개 중심성 $\mu=0.708$	아이겐벡터 중심성 $\mu=6.550$
<i>Cognitive Dimensions</i>				
natural resource	park(2.773)	garden(3.181)	park(9.830)	park(8.480)
	garden(1.638)	sunset(2.891)	river(4.813)	garden(6.077)
	$\mu=0.875(19)$	$\mu=2.695(18)$	$\mu=1.695(2)$	$\mu=3.179(19)$
infrastructure	street(6.902)	hotel(3.065)	market(4.655)	street(20.755)
	city(6.265)	mall(3.045)	tower(2.152)	city(16.946)
	$\mu=2.261(7)$	$\mu=2.835(13)$	$\mu=1.115(4)$	$\mu=7.260(6)$
recreational entertainment	souvenir(1.985)	aquarium(3.181)	souvenir(1.671)	souvenir(7.722)
	ticket(1.909)	cosmetic(3.174)	TICKET(1.054)	ticket(6.434)
	$\mu=0.769(20)$	$\mu=2.915(8)$	$\mu=0.483(9)$	$\mu=2.940(20)$
food & beverage	food(9.317)	cafe(3.077)	food(3.158)	food(26.276)
	restaurant(6.407)	food(2.903)	restaurant(2.170)	restaurant(18.165)
	$\mu=1.856(12)$	$\mu=2.711(17)$	$\mu=0.839(7)$	$\mu=6.058(12)$
socioeconomic	buying(3.043)	bargain(3.116)	price(0.778)	buying(10.117)
	price(3.000)	price(3.097)	buying(0.604)	price(9.873)
	$\mu=1.162(17)$	$\mu=2.886(11)$	$\mu=0.145(18)$	$\mu=4.297(17)$
transportation (accessibility)	car(2.594)	car(2.781)	station(4.272)	car(8.545)
	station(2.540)	subway(2.768)	driver(3.509)	location(8.111)
	$\mu=1.208(16)$	$\mu=2.595(22)$	$\mu=0.959(6)$	$\mu=4.325(16)$
cultural, history & art	palace(6.429)	war(3.213)	museum(1.935)	palace(17.831)
	museum(5.503)	palace(3.123)	palace(1.229)	museum(16.175)
	$\mu=1.873(11)$	$\mu=2.981(2)$	$\mu=0.391(13)$	$\mu=6.288(11)$
information resource	guide(6.830)	english(2.935)	guide(9.275)	guide(19.469)
	english(3.865)	information(2.839)	knowledge(3.659)	english(11.941)
	$\mu=2.024(8)$	$\mu=2.641(21)$	$\mu=1.986(1)$	$\mu=6.628(10)$
proper schedule/plan	time(14.639)	night(3.161)	day(0.654)	time(37.835)
	day(11.164)	hour(3.013)	morning(0.548)	day(29.402)
	$\mu=3.809(3)$	$\mu=2.783(15)$	$\mu=0.208(17)$	$\mu=10.965(3)$
Social environment	korean(13.694)	brand(3.148)	service(3.168)	korean(35.425)
	resident(9.328)	fashionable(3.077)	tourist(1.376)	resident(24.862)
	$\mu=2.345(6)$	$\mu=2.914(9)$	$\mu=0.454(11)$	$\mu=7.222(7)$
Specific activities	shopping(12.506)	shopping(3.11)	walking(5.127)	shopping(33.173)
	walking(12.396)	viewing(3.045)	hiking(2.862)	walking(32.082)
	$\mu=4.079(2)$	$\mu=2.812(14)$	$\mu=1.408(3)$	$\mu=11.545(2)$
natural environment	atmosphere(0.592)	atmosphere(3.077)	hot(2.600)	cold(3.249)
	cold(0.795)	autumn(2.619)	summer(0.416)	weather(3.025)
	$\mu=0.596(23)$	$\mu=2.570(23)$	$\mu=0.414(12)$	$\mu=2.412(23)$

주: a. 관광지 이미지의 각 구성요인별 2개의 상위세부요인과 구성요인의 평균값을 정리한 결과임.
 μ =평균(순위)에 해당하며, 굵은 글씨체는 각 중심성 분석의 평균보다 큰 값에 해당함.

〈표 10〉 관광지 이미지 차원별 중심성 분석결과 요약 - Affective Dimension -^a

DIMENSION	연결정도 중심성 $\mu=2.046$	근접 중심성 $\mu=2.810$	매개 중심성 $\mu=0.708$	아이젠벡터 중심성 $\mu=6.550$
<i>Affective Dimensions</i>				
arousing	amazing(2.999)	amazing(2.981)	amazing(0.104)	amazing(10.439)
	fantastic(0.845)	fantastic(2.942)	fantastic(0.001)	fantastic(3.803)
	$\mu=1.992(9)$	$\mu=2.962(3)$	$\mu=0.052(22)$	$\mu=7.121(8)$
exciting	interesting(6.485)	exciting(3.038)	interesting(0.776)	interesting(18.324)
	surprise(0.658)	interesting(2.974)	surprise(0.001)	surprise(2.885)
	$\mu=2.510(5)$	$\mu=2.962(4)$	$\mu=0.259(15)$	$\mu=7.685(5)$
distressing	crowded(2.803)	crowded(3.084)	crowded(0.193)	crowded(9.909)
	waiting(1.143)	complex(2.949)	waiting(0.049)	waiting(4.406)
	$\mu=1.483(13)$	$\mu=2.893(10)$	$\mu=0.081(20)$	$\mu=5.448(14)$
relaxing	resting(1.937)	peaceful(2.755)	resting(2.558)	resting(7.142)
	comfort(0.952)	refreshing(2.710)	peaceful(0.473)	comfort(3.586)
	$\mu=1.012(18)$	$\mu=2.646(20)$	$\mu=0.799(8)$	$\mu=3.807(18)$
unpleasant	bad(0.792)	disappoint(2.839)	bad(3.104)	bad(3.117)
	disappoint(0.630)	tired(2.692)	disappoint(0.002)	disappoint(2.763)
	$\mu=0.638(22)$	$\mu=2.672(19)$	$\mu=1.035(5)$	$\mu=2.682(22)$
impressive	worth(6.551)	impressive(2.981)	worth(0.339)	worth(17.997)
	impressive(1.052)	worth(2.974)	impressive(0.006)	impressive(4.500)
	$\mu=2.866(4)$	$\mu=2.947(6)$	$\mu=0.115(19)$	$\mu=8.938(4)$
pleasant	great(13.445)	fun(3.058)	nice(2.151)	great(36.006)
	good(12.365)	good(2.994)	beautiful(1.718)	good(32.955)
	$\mu=5.308(1)$	$\mu=2.876(12)$	$\mu=0.470(10)$	$\mu=15.161(1)$
evoked	memorable(1.643)	memorable(3.013)	memorable(0.208)	missing(6.183)
	missing(1.434)	missing(2.942)	kind(0.011)	memorable(6.075)
	$\mu=1.392(14)$	$\mu=2.953(5)$	$\mu=0.077(21)$	$\mu=5.699(13)$
trendy	changing(2.114)	cool(3.071)	cool(0.506)	changing(7.315)
	modern(1.116)	changing(3.039)	changing(0.344)	young(4.615)
	$\mu=1.265(15)$	$\mu=3.018(1)$	$\mu=0.221(16)$	$\mu=5.011(15)$
authentic	historical(4.553)	historical(3.168)	traditional(1.882)	historical(14.551)
	traditional(3.610)	ancient(2.994)	historical(0.382)	traditional(12.137)
	$\mu=1.916(10)$	$\mu=2.938(7)$	$\mu=0.388(14)$	$\mu=6.838(9)$
reputational	famous(0.757)	clean(2.852)	clean(0.088)	clean(3.255)
	popular(0.661)	famous(2.776)	convenience(0.003)	popular(2.744)
	$\mu=0.655(21)$	$\mu=2.748(16)$	$\mu=0.023(23)$	$\mu=2.726(21)$

주: a. 관광지 이미지의 각 구성요인별 2개의 상위세부요인과 구성요인의 평균값을 정리한 결과임.
 μ =평균(순위)에 해당하며, 굵은 글씨체는 각 중심성 분석의 평균보다 큰 값에 해당함.

이 관광지에서 할 수 있는 다양한 활동과 이를 수반하는 기반시설이라는 것이다.

V. 논의 및 결론

본 연구는 온라인 여행 커뮤니티의 빅데이터에 기반한 사회연결망 분석을 통하여 긍정적인 관광지 이미지 창출을 위해 여행객에게 인식되고 있는 관광지 이미지의 핵심요소를 파악하고 구성개념 간의 관계를 기술적이고 탐색적으로 고찰하고자 하였다. 특히, 관광지의 이미지는 여행객들이 관광지를 경험하면서 느끼는 인식을 통해 표현되기 때문에(Lee, Back & Kim, 2009) 그동안의 관광지 이미지에 대한 연구들이 계량적 척도에 주안점을 두고 관광지의 경쟁력 제고를 위한 중요도와 만족도를 연구하였다면 본 연구는 여행객들의 실제 의견을 내재적 차원에서 분석함으로써 그들이 직관적으로 인식하고 있는 이미지를 파악하고 시사점을 제시하는데 그 의미가 있다.

먼저, 현재까지 관광지 이미지 관련 연구 가운데 구성요인 및 하위요인을 규명하거나 변수간의 구조적 관계를 밝힌 연구를 기준으로 실제 여행객이 남긴 의견과 비교함으로써 그 중요도뿐만 아니라 차이가 있다는 것을 알 수 있었다. 텍스트 마이닝을 통해 분류된 세부요인 분석 결과, 관광지의 사회적 환경을 나타내는 세부요인이 19개로 가장 많은 것으로 나타났으며, 기반시설(15개), 문화·역사 및 예술(15개), 교통(접근성)(14개), 먹거리(14개), 기쁜(13개), 여가·흥미적 요소(13개) 등의 순으로 여행객의 의견이 많은 항목인 것으로 나타났다. 항목 수가 가장 많이 나타난 사회적 환경에 대한 세부요인을 선행연구와 비교해보면, 기존 연구에서는 사회적 환경을 구성하는 세부요인이 지역 주민, 서비스, 언어의 장벽, 사회적 특징, 안정인 것으로 나타

났지만(Beerli & Martin, 2004; Martin & Bosque, 2008) 실제 여행객은 5개의 세부요인에 더하여 국가를 지칭하는 중국인(Chinese), 일본인(Japanese), 국가(nation), 나라(country), 외국인(foreigner) 등이 언급됨으로써 국내를 찾는 주 여행객에 대한 문화적 수용력 또한 중요하게 인식되고 있다는 것을 짐작해 볼 수 있다. 관광지 이미지는 Gallarza *et al.*(2002)가 주장한 4가지의 특성 중 상대성과 동적이라는 현상으로 인하여 시대적인 상황을 적절히 반영한 접근법이 중요하다는 것을 알 수 있다. 특히, 선행연구와의 비교·분석을 통해 도출된 구성개념에 대하여 향후 관광지 이미지를 측정하는 연구에서 보다 확장된 항목으로써 활용가능하리라 판단된다.

다음으로, 그동안 관광지 이미지에 대한 인식과 경쟁력에 관한 연구는 중요도-만족도 기법(IPA)을 통해 구성개념에 대하여 단일적 차원에서 강·약점을 파악하였다면(Enright & Newton, 2004; Jeng *et al.*, 2017; Pike, 2016), 본 연구에서는 관광지 이미지의 구성개념 간 구조적 관계(네트워크)에 숨어있는 정보를 포착하기 위해 네트워크 분석의 중심성 분석을 바탕으로 여행객이 인식하는 관광지 이미지에 대한 실증 분석을 진행하였다. 특히, 본 연구에서는 중심성 분석의 개념정립과 관광분야에 실증적으로 적용하기 위한 방안을 제시하고자 하였다. 세부적으로, 연결정도 중심성은 관광지 이미지에 대하여 여행객의 인식에 직접적으로 영향을 측정하기 위한 개념, 근접 중심성은 관광지 이미지에 대하여 여행객이 인식하는데 준거점 역할을 함으로써 심리적으로 가까운 거리를 측정하는 개념, 매개 중심성은 여행객이 관광지 이미지를 인식하는데 의존성이 큰 개념으로써 경쟁력 제고를 측정하는 개념, 아이겐벡터 중심성은 관광지 이미지를 인식하는데 그 파급력이 큰 개념으로 정리될 수 있었다.

이를 바탕으로 한 분석결과를 살펴보면, 연결 정도 중심성 분석에서는 관광지의 인지적·정서적 이미지가 고루 분포한다는 것을 알 수 있었으며, 관광지에서 여행객이 주도적으로 일정을 계획하고 그 가운데 즐길 수 있는 다양한 활동과 이를 수반할 수 있는 기반시설과 사회적 환경이 기쁨과 깊은 감명으로 인식됨으로써 관광지의 이미지에 대하여 긍정적 인식을 가진다는 것을 알 수 있었다. 근접 중심성 분석에서는 인지적 이미지보다 정서적 이미지의 점유가 높은 것으로 나타났으며 관광지에서 느끼는 최신 유행의 느낌(trendy)과 자극적인(arousing), 흥미로운(exciting), 환기되는(evoked), 감명 깊은(impressive), 믿을만한(authentic), 기쁨(pleasant)의 정서적 이미지와 문화·역사 및 예술, 여가·흥미적 요소, 사회적 환경, 사회경제적인 요소, 관광지에서의 활동이 여행객의 심리적 거리와 가깝게 위치함으로써 상호·유기적으로 어우러져 관광지의 긍정적 이미지 연출에 도움이 되는 것으로 나타났다. 또한, 매개 중심성은 정서적 이미지보다 인지적 이미지의 점유가 높은 것으로 나타났으며, 흥미로운 것은 그동안의 관광지 이미지의 경쟁력 제고를 위한 중요도-만족도 연구에서는 관광지의 교통(접근성), 기반시설, 사회적 환경 등이 강점으로 나타났으나(Enright & Newton, 2004; Jeng *et al.*, 2017) 매개 중심성 분석 결과, 선행연구에서 나타난 구성개념과 더불어 '정보자원', '관광지에서의 활동', '먹거리', '편안하게 하는'의 구성개념을 바탕으로 경쟁력을 제고 한다면 관광지의 긍정적 이미지 인식에 도움이 되는 것으로 나타났다. 특히, '정보자원'의 평균이 가장 높게 나타나 최근 여행객이 관광지에서 얻을 수 있는 경험과 지식에 대한 정보의 중요성이 고려될 필요성이 있음을 시사하였다. 아이젠벡터 중심성은 연결정도 중심성의 결과에서 나타난 7개의 상위요인에 더하여 자극적인, 믿을만한, 정보자원의 3가지 구성개념이 여행객이 관광지 이

미지를 인식하는데 그 파급력이 높은 것으로 나타났다. 네 가지의 중심성 분석에서 관광지에서의 활동, 기반시설이 공통적으로 전체평균보다 높게 나타났으며 결과적으로 여행객이 관광지 이미지를 인식하는데 가장 중요하고 큰 영향력을 주는 것은 여행객이 관광지에서 할 수 있는 다양한 활동과 이를 수반하는 기반시설이라는 것을 알 수 있었다.

본 연구는 현재까지 관광분야에서 관광지 이미지 측정을 위해 활용되었던 분석체계에서 보다 확장된 빅데이터 분석을 활용하여 내재적 차원에서 여행객이 인식하는 관광지 이미지를 측정하고자 하였다. 세부적으로, 선행연구에서 활용하였던 관광지 이미지의 구성개념과 실제 여행객들이 인식하는 이미지의 구성개념을 비교해봄으로써 시대적인 관광지 이미지의 흐름을 파악하고자 노력하였다. 또한, 서울의 관광지 및 관광상품을 분석단위로 활용하여 관광지 이미지의 구조적 관계를 네트워크 분석을 통해 실증함으로써 관광지 이미지 연출에 필요한 구성요인뿐만 아니라 경쟁력을 갖추기 위해 고려되어야 하는 요소까지 도출함으로써 향후 서울의 관광지 이미지 제고에 필요한 기초자료가 될 수 있을 것이라 판단된다. 무엇보다 관광분야에 빅데이터를 분석할 수 있는 체계와 비정형 텍스트 데이터를 보다 객관·타당한 계량데이터로 활용할 수 있는 방안을 제시하였으며 관광분야에서도 최근 활용되고 있는 네트워크 분석에 대하여 개념정립과 접근방안을 논의함으로써 학문적 성과를 제고하였다고 사료된다. 그럼에도 불구하고 관광지 이미지 또는 경쟁력에 고려되는 요인들을 규명하는데 초점을 맞춰, 이러한 요인들이 실제 여행객의 방문이나 만족에 어떠한 영향을 주는지를 증명하지는 못하였다. 또한, 향후 연구에서는 국내 관광지뿐만 아니라 다양한 국가와의 비교분석을 통하여 국내 관광지가 경쟁력을 제고 할 수 있는 보다 심층적인 논의와 내·외국인 또는 동양권과 서양권의 여행객

특성에 따라 관광지 이미지가 어떻게 인식되고 있는지를 연구한다면 학문적, 실무적으로 다양한 시사점을 제공할 수 있을 것이라 판단된다.

참고문헌

- 강한훈 · 유성준 · 한동일(2010). 다양한 계층 트리 구조를 갖는 쇼핑몰 상에서의 상품평 수집을 위한 웹 크롤러 래퍼의 설계 및 구현. 『한국지능시스템학회 논문지』, 20(3), 318-325.
- 경남신문(2015.07.21.). 빅데이터를 활용한 지방행정 (2) 부산 해운대구, SNS로 흐름을 읽다. <http://www.knnews.co.kr/news/article-View.php?idxno=1153962>
- 김광영 · 이원구 · 윤화목 · 신성호 · 이민호(2011). 웹 자원 아카이빙을 위한 웹 크롤러 연구 개발. 『한국콘텐츠학회논문지』, 11(9), 9-16
- 김근형 · 오성열(2009). 온라인 고객리뷰 분석을 통한 시장세분화에 텍스트 마이닝 기술을 적용하기 위한 방법론. 『한국콘텐츠학회논문지』, 9(8), 272-284.
- 김민주(2003). 관광지 이미지의 측정척도 개발에 관한 연구. 『관광경영연구』, 19, 93-114.
- 김병국 · 박석희(2001). 관광지 이미지의 형성에 관한 연구: 인지적, 정서적 이미지를 중심으로. 『관광학연구』, 25(1), 271-290.
- 김보경 · 김미경(2015). 빅데이터를 활용한 템플스테이 키워드 네트워크 분석-블로그와 트위터 중심. 『관광학연구』, 39(5), 27-40.
- 김용학(2007). 『사회연결망이론』. 서울: 박영사.
- 김휴찬 · 김미정 · 신현정(2014). 소셜데이터 분석방법을 이용한 제주관광 트렌드 서비스 설계 및 구현. 『e-비즈니스연구』, 15(3), 173-193.
- 노정희(2008). 관광 목적지로서의 지각된 위험과 관광지 이미지가 한국 방문 의사에 미치는 영향에 관한 연구. 『관광 · 레저연구』, 20(2), 169-188.
- 도해용 · 이재주(2015). 빅데이터를 이용한 한국 · 중국 · 일본의 레스토랑 브랜드개성에 관한 연구. 『외식경영연구』, 18, 227-248.
- 박수완(2004). 관광지이미지에 대한 정보원천과 관여수준의 역할에 관한 연구: 자율적 정보원천과 방문경험을 중심으로. 『호텔관광연구』, 16, 135-152.
- 서용건 · 서용구(2004). 한류가 한국의 관광지 이미지와 관광객 의사결정에 미치는 영향. 『관광학연구』, 28, 47-64.
- 손동원(2002). 『사회 네트워크 분석』. 서울: 경문사.
- 심영석 · 김홍범(2016). 텍스트 마이닝을 이용한 관광지 이미지 구성요인 및 측정에 관한 연구. 『관광학연구』, 40(7), 221-245.
- 오익근 · 이태숙 · 전채남(2015). 빅데이터 분석을 통한 한국관광 인식에 관한 연구. 『관광학연구』, 39(10), 107-126.
- 왕명명 · 류시영 · 송운강(2011). 한국 영상매체 접측과 관광지 이미지, 방문의도 간의 관계. 『관광연구저널』, 25(1), 129-142.
- 이영진 · 윤지환(2014). 관광분야에서 SNS 빅데이터의 활용 방법 모색. 『관광연구저널』, 28, 5-14.
- 이태원 · 홍태호(2015). Support Vector Machine을 이용한 온라인 리뷰의 용어기반 감성분류모형. 『Information Systems Review』, 17(1), 49-64.
- 이학인 · 정기한(2016). 관광지 정보원천이 관광지 이미지에 미치는 영향. 『인터넷전자상거래연구』, 16(1), 179-194.
- 임정우 · 이교은 · 하동현(2013). 한국의 국가이미지, 관광지 인지적 및 정서적 이미지, 의료관광 이미지 간의 관계. 『관광연구』, 28(3), 231-250.
- 조명환 · 정선자(2009). 관광 목적지 이미지 형성요인에 관한 연구. 『관광 · 레저연구』, 21(4), 109-127.
- 조완섭 · 조아 · 권가은 · 류관희(2015). SNS 데이터 분석을 통한 스마트 충북관광 구축. 『한국데이터정보과학회지』, 26(2), 409-418.
- 채예병(2007). 강화도 이미지에 따른 관광지 선택에 관한 실증적 연구. 『관광학연구』, 31(6), 353-369.
- 한진성 · 윤지환(2016). 소셜 빅데이터 텍스트 마이

- 닝을 활용한 부산국제영화제(BIFF) 활성화 방안 연구: 영화제 개최 전·후 비교 분석을 중심으로. 『관광학연구』, 40(1), 133-145.
- 현용호 · 홍선영 · 오홍철(2007). 인지적 이미지 구성 확인모델에 따른 관광지 이미지 형성모델과 충성도간의 구조적 관계고찰. 『관광·레저연구』, 19(4), 227-247.
- Baloglu, S., & Brinberg, D. (1997). Affective images of tourism destinations. *Journal of Travel Research*, 35(4), 11-15.
- Baloglu, S., & McCleary, K. (1999). A model of destination image formation. *Annals of Tourism Research*, 26(4), 868-897.
- Beerli, A., & Martin, J. (2004). Factors influencing destination image. *Annals of Tourism Research*, 31(3), 657-681.
- Bigne, J., Sanchez, M., & Sanchez, J. (2001). Tourism image, evaluation variables and after purchase behaviour: Interrelationship. *Tourism Management*, 22(6), 607-616.
- Bjørkelund, E., Burnett, T., & Nørvåg, K. (2012). A study of opinion mining and visualization of hotel reviews. *In Proceedings of the 14th International Conference on Information Integration and Web-based Applications & Services* (pp. 229-238). ACM.
- Bonacich, P. (1987). Power and centrality: A family of measures. *American Journal of Sociology*, 92(5), 1170-1182.
- Bonacich, P. (2007). Some unique properties of eigenvector centrality. *Social Networks*, 29(4), 555-564.
- Buhalis, D., & Law, R. (2008). Progress in information technology and tourism management: 20 years on and 10 years after the Internet—The state of eTourism research. *Tourism Management*, 29(4), 609-623.
- Buhalis, D., & Spada, A. (2000). Destination management systems: Criteria for success—an exploratory research. *Information Technology & Tourism*, 3(1), 41-58.
- Chen, J., & Uysal, M. (2002). Market positioning analysis: A hybrid approach. *Annals of Tourism Research*, 29(4), 987-1003.
- Choi, S., Lehto, X., & Morrison, A. (2007). Destination image representation on the web: Content analysis of Macao travel related websites. *Tourism Management*, 28(1), 118-129.
- Crouch, G. (2011). Destination competitiveness: An analysis of determinant attributes. *Journal of Travel Research*, 50(1), 27-45.
- Dickinger, A. (2011). The trustworthiness of online channels for experience- and goal-directed search tasks. *Journal of Travel Research*, 50(4), 378-391.
- Dickinger, A., & Költringer, C. (2011). Extracting destination representation and competitiveness from online content. *Information Technology & Tourism*, 13(4), 327-339.
- Dickinger, A., Költringer, C., & Körbitz, W. (2011). Comparing online destination image with conventional image measurement—The case of Tallinn. *Information and Communication Technologies in Tourism 2011*, 165-177. doi:10.1007/978-3-7091-0503-0_14
- Dwyer, L., Dragičević, V., Armenski, T., Mihalič, T., & Knežević Cvelbar, L. (2016). Achieving destination competitiveness: An importance-performance analysis of Serbia. *Current Issues in Tourism*, 19(13), 1309-1336.
- Enright, M., & Newton, J. (2004). Tourism destination competitiveness: A quantitative approach. *Tourism Management*, 25(6), 777-788.
- Freeman, L. (1979). Centrality in social networks: Conceptual classification. *Social*

- Networks*, 1, 215-239.
- Galaskiewicz, J., & Wasserman, S. (1994). *Advances in social network analysis: Research in the social and behavioral sciences*. Sage Focus Editions, 171. Thousand Oaks, CA: Sage Publications.
- Gallarza, M., Saura, I., & García, H. (2002). Destination image: Towards a conceptual framework. *Annals of Tourism Research*, 29(1), 56-78.
- Gartner, W. (1994). Image formation process. *Journal of Travel & Tourism Marketing*, 2(2-3), 191-216.
- Gomezelj, D., & Mihalič, T. (2008). Destination competitiveness: Applying different models, the case of Slovenia. *Tourism Management*, 29(2), 294-307.
- Gretzel, U., Yoo, K., & Purifoy, M. (2007). *Online travel review report: Role & impact of online travel reviews*. Laboratory for Intelligent Systems in Tourism. Texas A&M University.
- Hanneman, R., & Riddle, M. (2005). *Introduction to social network methods*. Riverside, CA: University of California.
- Jeng, C., Snyder, A., & Chen, C. (2017). Importance-performance analysis as a strategic tool for tourism marketers: The case of Taiwan's destination image. *Tourism and Hospitality Research*, 17(2), 1-14.
- Jeong, M., & Jeon, M. (2008). Customer reviews of hotel experiences through consumer generated media(CGM). *Journal of Hospitality & Leisure Marketing*, 17(1-2), 121-138.
- Kang, H., Yoo, S., & Han, D. (2009). Modeling web crawler wrappers to collect user reviews on shopping mall with various hierarchical tree structure. *Proceeding of the WISM 2009 International Conference* (pp. 69-73), Web Information Systems and Mining.
- Kim, H., & Richardson, S. (2003). Motion picture impacts on destination images. *Annals of Tourism Research*, 30(1), 216-237.
- Kim, Y., Choi, T., Yan, T., & Dooley, K. (2011). Structural investigation of supply networks: A social network approach. *Journal of Operations Management*, 29(3), 194-211.
- Knoke, D., & Burt, R. (1983). *Applied network analysis*. Newbury Park, CA: Sage.
- Knoke, D., & Yang, S. (2008). *Social network analysis*. Thousand Oaks, CA: Sage.
- Költringer, C., & Dickinger, A. (2015). Analyzing destination branding and image from on-line sources: A web content mining approach. *Journal of Business Research*, 68(9), 1836-1843.
- Konecnik, M. (2002). The image as a possible source of competitive advantage of the destination: The case of Slovenia. *Tourism Review*, 57(1/2), 6-12.
- Lee, H., Law, R., & Murphy, J. (2011). Helpful reviewers in TripAdvisor: An online travel community. *Journal of Travel & Tourism Marketing*, 28(7), 675-688.
- Lee, M., Singh, N., & Chan, E. (2011). Service failures and recovery actions in the hotel industry: A text-mining approach. *Journal of Vacation Marketing*, 17(3), 197-207.
- Lee, Y., Back, K., & Kim, J. (2009). Family restaurant brand personality and its impact on customer's emotion, satisfaction, and brand loyalty. *Journal of Hospitality & Tourism Research*, 33(3), 305-328.
- Lewis, R., & Chambers, R. (2000). *Marketing leadership in hospitality, foundations and practices*. New York: Wiley.
- Li, H., Ye, Q., & Law, R. (2013). Determinants

- of customer satisfaction in the hotel industry: An application of online review analysis. *Asia Pacific Journal of Tourism Research*, 18(7), 784-802.
- Liu, X., Schuckert, M., & Law, R. (2015). Can response management benefit hotels? Evidence from Hong Kong hotels. *Journal of Travel & Tourism Marketing*, 32(8), 1069-1080.
- Mak, A. (2017). Online destination image: Comparing national tourism organisation's and tourists' perspectives. *Tourism Management*, 60, 280-297.
- Martín, H., & Bosque, I. (2008). Exploring the cognitive-affective nature of destination image and the role of psychological factors in its formation. *Tourism Management*, 29(2), 263-277.
- O'Leary, S., & Deegan, J. (2005). Ireland's image as a tourism destination in France: Attribute importance and performance. *Journal of Travel Research*, 43(3), 247-256.
- Pan, B., & Yang, Y. (2017). Forecasting destination weekly hotel occupancy with big data. *Journal of Travel Research*, 56(7), 957-970.
- Pike, S. (2016). Destination image: Identifying baseline perceptions of Brazil, Argentina and Chile in the nascent Australian long haul travel market. *Journal of Destination Marketing & Management*, 5(2), 164-170.
- Pike, S., & Ryan, C. (2004). Destination positioning analysis through a comparison of cognitive, affective, and conative perceptions. *Journal of Travel Research*, 42(4), 333-342.
- Prell, C. (2012). *Social network analysis: History, theory and methodology*. India: Sage.
- Rezende-Parker, A., Morrison, A., & Ismail, J. (2003). Dazed and confused? An exploratory study of the image of Brazil as a travel destination. *Journal of Vacation Marketing*, 9(3), 243-259.
- Russell, J., Ward, L., & Pratt, G. (1981). Affective quality attributed to environments: A factor analytic study. *Environment and Behavior*, 13(3), 259-288.
- Salton, G., & Buckley, C. (1988). Term-weighting approaches in automatic text retrieval. *Information Processing & Management*, 24(5), 513-523.
- Schuckert, M., Liu, X., & Law, R. (2015). Hospitality and tourism online reviews: Recent trends and future directions. *Journal of Travel & Tourism Marketing*, 32(5), 608-621.
- Scott, J. (2000). *Social network analysis*. London: Sage.
- Shih, H. (2006). Network characteristics of drive tourism destinations: An application of network analysis in tourism. *Tourism Management*, 27(5), 1029-1039.
- Stepchenkova, S., & Morrison, A. (2006). The destination image of Russia: From the online induced perspective. *Tourism Management*, 27(5), 943-956.
- Stepchenkova, S., Kirilenko, A., & Morrison, A. (2009). Facilitating content analysis in tourism research. *Journal of Travel Research*, 47(4), 454-469.
- Stokowski, P. (1994). *Leisure in society: a network structural perspective*. London, Cassell plc: Mansell Publishing.
- Stylydis, D., Shani, A., & Belhassen, Y. (2017). Testing an integrated destination image model across residents and tourists. *Tourism Management*, 58, 184-195.
- Tapachai, N., & Waryszak, R. (2000). An examination of the role of beneficial image in tourist destination selection. *Journal*

- of Travel Research*, 39(1), 37-44.
- Wang, Y., Chan, S., Ngai, G., & Leong, H. (2013). Quantifying reviewer credibility in online tourism. In *Proceeding of International Conference on Database and Expert Systems Applications* (pp. 381-395). Springer Berlin Heidelberg.
- Yang, X., Pan, B., Evans, J., & Lv, B. (2015). Forecasting Chinese tourist volume with search engine data. *Tourism Management*, 46, 386-397.
- Ye, Q., Zhang, Z., & Law, R. (2009). Sentiment classification of online reviews to travel destinations by supervised machine learning approaches. *Expert Systems with Applications*, 36(3), 6527-6535.
- Zhang, L., Lan, C., Qi, F., & Wu, P. (2017). Development pattern, classification and evaluation of the tourism academic community in China in the last ten years: From the perspective of big data of articles of tourism academic journals. *Tourism Management*, 28(10), 114-125.

2017년 6월 12일 최초투고논문 접수
2017년 7월 18일 최종심사완료 및 게재확정 통보
2017년 7월 28일 최종논문 도착
3인 익명심사 畢