

소비자 참여행동에 영향을 미치는 Short-form 영상의 요소 : 컴퓨터 비전 알고리즘을 활용한 이미지 분석

방유빈

한양대학교 일반대학원 비즈니스 인포매틱스학과
(bin0261@hanyang.ac.kr)

이상용

한양대학교 경영대학 경영학부 교수
(tomlee@hanyang.ac.kr)

최근 소셜 네트워크 서비스(SNS)와 모바일 중심의 미디어 소비가 확대되면서 사용자의 자체적인 생성 콘텐츠 데이터가 급증하고 있다. 특히 숏폼(Short-form)은 짧은 시간내에 시각적 체험을 제공하여 콘텐츠를 전달하는 특성을 가지며, 이는 오늘날 정보 소비 형태에 가장 부합한 콘텐츠로 각광받고 있다. 다양한 플랫폼에서 생성되고 있는 숏폼 콘텐츠는 이제 개인을 넘어서, 기업의 가치 향상과 마케팅 전략으로까지 그 범위가 확대되고 있다. 일반적으로 숏폼은 추천 알고리즘을 통해 소비자에게 전달되는 특성을 가진다. 따라서 소비자는 짧은 영상을 통해 콘텐츠의 소비 여부를 결정해야 하기 때문에, 시각적 요소가 다른 매체보다 중요하게 작용된다. 따라서 본 연구는 숏폼을 프레임 단위로 분할하여 이미지의 시각적 요소가 소비자 참여행동에 미치는 영향을 분석하고자 한다. 유튜브 쇼츠를 중심으로 컴퓨터 비전을 활용한 이미지 분석을 진행하였으며 숏폼을 프레임별로 이미지화한 후 심미도, 품질, 색상, 추출된 객체 등의 시각적 요소가 소비자 참여행동에 미치는 영향과 이러한 요소가 중요한 영향을 미치는 시점을 파악하였다. 본 연구는 소셜 미디어 콘텐츠, 특히 숏폼을 프레임 단위로 분석하여 각 프레임의 시각적 요소가 소비자 참여행동에 미치는 중요성과 그 영향이 발생하는 시점을 파악한 최초의 연구이다. 또한 향후 이미지의 미적 분석에 기반한 콘텐츠 자동화, 소비자 행동과 반응을 예측하는 최적화 전략에 이론적 토대와 방향을 제공한다.

주제어 : 숏폼, 심미도, 소셜 미디어, 이미지 분석, 컴퓨터 비전

논문접수일 : 2025년 10월 13일 논문수정일 : 2025년 12월 17일 게재확정일 : 2025년 12월 22일
원고유형 : Regular Track 교신저자 : 이상용

1. 개요

소셜 네트워크 서비스(SNS)와 모바일 중심 미디어 소비의 보편화로 사용자 참여와 콘텐츠 데이터 생성이 급증하고 있다. 특히 숏폼(Short-form) 영상 콘텐츠 시장은 다양한 플랫폼을 중심으로 빠르게 성장하고 있으며, 숏폼이 제공하는 직관적인 시각적 경험은 개인 및 기업의 마케팅에서도

핵심 요소로 부상하고 있다. 특히 짧은 시간 내에 핵심 정보를 효과적으로 전달하는 특성은 현대 소비자들의 정보 소비 방식과 밀접하게 부합하며, 일상적인 콘텐츠 소비 구조에 깊숙이 침투하고 있다. 2023년 로이터에 따르면, 인스타그램 릴스의 일일 재생 수는 2천억 뷰에 달할 정도로 사용량이 지속적으로 증가하는 등 숏폼 콘텐츠 수요의 지속적인 증가 추세를 확인할 수 있다. 1)

* 이 논문은 2020년 대한민국 교육부와 한국연구재단의 지원을 받아 수행된 연구임 (NRF-2020S1A3A2A02093277)
1) Reuters (2023, July 27). Meta's Reels revenue narrows in on TikTok, boosted by AI. Reuters. <https://www.reuters.com/technology/metas-reels-revenue-narrows-tiktok-boosted-by-ai-2023-07-27>

숏폼 콘텐츠는 영상 중심 커뮤니케이션이 주류인 미디어 환경 속에서 소비와 재생산이 활발히 이루어지는 대표적인 콘텐츠 형태로 자리매김하고 있다. 이진(2020)은 스마트폰 중심의 모바일 미디어는 참여자와의 거리를 더욱 좁히고, 능동성과 접근성을 강화하는 메타미디어로 기능한다고 분석하였다. 이는 물리적 장소 기반 커뮤니케이션에서 존재 기반으로의 변화를 이끌었으며, 고정된 장소 중심 시스템이 사람 기반 시스템으로 확장되었음을 의미한다. 이러한 추세 속에서, 숏폼 콘텐츠는 빠르게 도입되는 동시에 세계적으로 널리 퍼지며 사용자들의 기대와 습관을 재편하였다 (Rajendran et al., 2024).

이러한 추세는 콘텐츠 소비를 수동적 행위에서 능동적 참여로 전환시키고 있으며, 간편한 제작 및 공유 환경은 사용자 주도의 콘텐츠 생산을 활성화시키는 기반이 되고 있다. 특히 ‘챌린지(Challenge)’ 콘텐츠와 같이 참여 기반 구조를 띤 숏폼은 사용자 간의 상호작용을 유도하며 좋아요, 댓글, 공유 등의 반응을 통해 자발적인 확산 효과를 일으킨다. 이는 단순히 콘텐츠 소비를 넘어, 디지털 마케팅 및 브랜드 커뮤니케이션의 전략적인 패러다임 변화에 기여하는 요인으로 작용한다. 이처럼 많은 참여 행동과 공감을 받는 숏폼 콘텐츠는 브랜드 인지도 제고 및 수익성 증대에도 긍정적인 영향을 미친다.

현재 숏폼 콘텐츠를 가장 활발히 소비하는 플랫폼은 틱톡, 인스타그램의 릴스, 구글의 유튜브 쇼츠로 대표된다. 국내 소셜 미디어 트렌드 조사에 따르면, 전체 응답자의 68.9%가 숏폼 플랫폼 이용

경험이 있다고 응답했으며, 이는 전년 대비 12.4% 증가한 수치이다.²⁾ 이 중 틱톡은 가장 먼저 서비스를 개시한 숏폼 플랫폼으로, 중국을 기반으로 단기간에 산업 전반에서 급부상하였다. 반면, 유튜브 쇼츠는 기존 유튜브 플랫폼에서 파생된 숏폼 서비스로, 15초에서 60초 이내의 짧은 영상을 중심으로 유튜브의 특성을 계승하고 있다. 박병언 & 임규건(2015)에 따르면 유튜브는 기존 폐쇄형 콘텐츠 서비스와는 다르게 개방형 콘텐츠 서비스로 이용자들의 참여와 공유를 통하여 많은 인기를 유지하고 있다. 이러한 기존 유튜브 구조와 사용차이를 자연스럽게 흡수한 유튜브 쇼츠는 틱톡의 대안으로 빠르게 부상하며 뚜렷한 성장세를 보이고 있다. 특히 MZ세대 중심으로 높은 인기를 얻고 있으며, 현재 전 세계 월간 사용자 수는 15억 명을 넘어섰다.³⁾ 국내 시청 경험 비율도 74.9%로 가장 높게 나타났으며, 틱톡은 43.4%로 상대적으로 낮은 수치를 보인다. 2023년 기준으로 유튜브 쇼츠 이용률이 87.1%까지 증가한 반면, 틱톡은 34.5%로 감소 추세를 보이고 있다.

현재 유튜브 쇼츠는 전체 유튜브 조회수의 약 80%를 차지하며⁴⁾, 2023년 수익화 정책 도입 이후 이용률이 급격히 증가하고 있다. 기존 TV 매체와 대기업들 또한 쇼츠를 활용한 디지털 마케팅 전략을 강화하며 소비자와의 상호작용을 확대하는데 주력하고 있다. 대표적인 성공 사례로는 삼양 식품의 ‘불닭 시리즈’가 있다. ‘파이어 누들 챌린지’와 ‘플레이 불닭’ 같은 챌린지형 숏폼 콘텐츠는 글로벌 확산 전략과 결합하여 2024년 1분기 매출의

2) 오픈서베이 (2023). 소셜 미디어 검색포털 트렌드 리포트. (p.12).

3) 센트럴월드뉴스 (2022., 유튜브 쇼츠, 숏폼 영상 플랫폼계 틱톡 최대 라이벌로 급부상...그 전략은? CWN. Retrieved from <https://cwn.kr/article/179564942422863>

4) 경향신문 (2023년 2월 1일). 유튜브 시청자 빨아들이는 블랙홀, 쇼츠'...조회수 88% 차지. Retrieved from <https://www.khan.co.kr/article/202302011446001>

57.1% 증가, 영업이익 235% 상승으로 이어지며 숏폼 콘텐츠의 영향력을 보여주었다.⁵⁾ 이처럼 숏폼 콘텐츠는 디지털 미디어 환경에서 중심적인 역할을 지속적으로 수행할 것으로 전망되며, 콘텐츠 제작 및 유통, 소비자와의 커뮤니케이션 방식 등에서 새로운 패러다임을 제시하고 있다.

이처럼 유튜브 쇼츠는 국내 콘텐츠 소비 환경 전반에 점차 중요한 영향력을 미치고 있음에도 불구하고, 이에 대한 학술적 분석은 아직 충분이 이루어지지 않은 실정이다. 기존 숏폼 콘텐츠 연구는 주로 틱톡을 중심으로 전개되어 왔으며, 중국 이용자를 대상으로 한 사용자 경험이나 중독성에 관한 정성적 접근이 대부분을 차지한다. 특히 숏폼 콘텐츠의 핵심이라 할 수 있는 시각적 요소의 효과에 대한 실증적 분석은 현저히 부족하다. 시각 정보의 효과는 광고, 이커머스, 마케팅 분야에서 활발히 다루어져 왔지만, 숏폼 콘텐츠 맥락에서의 분석은 아직 초기 단계에 머물러 있다.

특히 유튜브 쇼츠와 같이 추천 알고리즘을 기반의 소비자 노출 구조에서는 짧은 시간 내 강한 시각적 인상을 각인시키는 방식이 중요한 역할을 하며, 이는 디지털 마케팅, 소비자 행동 분석, 소셜 미디어 전략 등 다양한 분야에서 학문적, 실무적 분석 가치가 크다. 따라서 유튜브 쇼츠의 국내 소비 형태와 시각적 콘텐츠 특성을 분석하는 연구는 실증적, 이론적 측면 모두에서 학문적으로 기여할 수 있을 것으로 보인다.

이에 본 연구는 숏폼 영상을 프레임 단위로 분해하여, 각 프레임의 시각적 요소가 소비자 참여 행동에 미치는 영향을 정량적으로 분석할 것을 제안한다. 숏폼 콘텐츠는 추천 알고리즘에 의해 소비자에게 자동적으로 제시되기 때문에, 시각적

요소가 주는 자극은 소비자 의사결정 과정에서 중요한 변수로 작용할 가능성이 크다.

본 연구는 컴퓨터 비전 알고리즘을 활용하여 이러한 시각적 요소를 정밀하게 분석하고, 해당 요소들이 소비자 참여행동에 미치는 영향력을 실증적으로 파악하고자 한다. 최근 컴퓨터 비전 기술의 발전은 이미지 및 비디오 분석의 정확도가 크게 향상시키고 있으며, 본 연구는 이러한 기술을 기반으로 숏폼 콘텐츠 분석에 새로운 접근을 시도하고자 한다. 이를 통해 숏폼 콘텐츠의 효과적인 활용 방안, 소비자 참여 유도 전략, 콘텐츠 마케팅 전략 수립 등에 실질적인 인사이트를 제공할 수 있을 것이다.

2. 이론적 배경

2.1. 숏폼 영상의 정의 및 특징

숏폼 영상 콘텐츠의 가장 큰 특징은 영상의 길이이다. Xie et al. (2023)은 일반적으로 영상의 길이가 15분 미만인 콘텐츠를 포함하며, 대부분 1분에서 5분 사이의 간결한 내용과 명확한 주제를 가지고 있는 영상들을 의미한다고 밝혔다. 또한 강정빈 & 이상원 (2021)은 기존 미디어의 형태로 생산되던 롱폼(Long-form) 콘텐츠와의 상대적인 개념으로서, 온라인 플랫폼 기반의 짧은 영상을 의미하기도 한다고 보았다.

Kuss & Griffiths (2017)은 사용자 생성 콘텐츠로서 플랫폼을 통해 비디오를 제작하여 공유 및 협업 등을 실현하는 웹 2.0 시대의 소셜 미디어의 한 형태라는 점이라는 것에 집중했다. 이진 (2020)은 특히 숏폼 영상 콘텐츠는 무엇보다도 참여와

5) Openads (2023). 삼양식품 불닭 브랜드가 글로벌 숏폼 챌린지로 얻은 성과. Openads. Retrieved from <https://www.openads.co.kr/content/contentDetail?contsId=13437>

상호작용에서 강한 특징을 가지며, 사용자는 적극적으로 콘텐츠와 자신의 교집합을 부연하고 의견을 반영하여 개인 사용자가 콘텐츠 제작 전문 집단이 되기도 하고, 전문 제작 집단의 개인이 콘텐츠를 독립적으로 제작하기도 한다고 밝혔다. Xiao et al. (2023)은 사용자는 흥미로운 경험이나 정보 등을 공유할 수 있으며, ‘좋아요’나 댓글을 통해 상호작용하고 자신이 좋아하는 제작자를 팔로우할 수도 있다고 언급하였다. 특히 숏폼은 추천 알고리즘의 영향을 받는다는 특징을 가지기 때문에 Park & Kaye (2021)는 적극적인 형태로서 비교적 시간이 소요되는 영상 검색에 비해 사용자의 빅데이터에 기반하여 관심여하는 풍부한 콘텐츠를 효율적으로 추천하며 ‘내가 정보를 찾는다’에서 ‘정보가 나를 찾는다’로의 전환을 실현시킨다고 보았다.

2.2. 영상에서 시각적 이미지의 중요성

이러한 영상은 사람과의 상호작용에도 중요한 영향을 미친다. 특히 숏폼의 경우 광고나 TV 등 기존의 영상 광고와는 달리 더욱 다양한 형태를 띠며, 소비자들의 능동적 참여와 구전 행동을 불러일으킨다. Addo et al. (2022)은 짧은 광고 등의 서비스 가치는 소비자들의 구매행동과 커뮤니케이션을 구축할 수 있으며, 기업의 경우 제품과 브랜드 등에 대한 소비자들의 인지도는 높이는 반면 불확실성은 낮출 수 있다는 것을 밝혔다.

따라서 숏폼 영상에서 시각적 이미지는 짧은 시간 동안 강한 인상을 남기는 것이 주된 목표인 만큼, 더욱 중요하게 작용한다. Shin et al. (2020)은 광고의 시각적 요소는 소비자의 관심을 유도하고 설득력을 높여 구매 확률을 높이는 데 중요한 역할을 한다고 밝혔으며, 박도형 외(2025)은

모바일 아이트래킹 기술을 통해 이커머스 플랫폼에서 시각적 주의와 구매 행동 간의 구체적인 연결 고리를 명확히 하기도 하였다.

Muhoho-Minni & Lubbe (2017)은 일반적으로 사용자는 광고나 매체에 대해 주의를 기울이거나, 자주 볼수록 이미지에 대한 인식이 좋아지게 된다는 것을 확인하였다. 특히 처음으로 자극을 경험했을 때 시각적 주의를 끄는 것이 중요하다. 그렇지 못할 경우 광고 자극 효과는 약화된다. Buscher et al. (2009)은 온라인 웹에서 사용자의 시선과 스크네일이 온라인에서 매우 중요한 역할을 한다고 밝혔으며, Cunningham & Nichols (2008) 역시 이러한 요소가 비디오를 볼 것인지 건너뛴 것인지를 결정하는 중요한 요소로 작용하기도 한다고 지적하였다.

따라서 본 연구에서는 숏폼 영상에서 이미지의 시각적 요소의 중요성을 파악한다. 또한 해당 요소가 영향을 미치는 시점을 파악하기 위해 각 각을 프레임으로 구분하고, 이를 분석하여 특정 시점에서 시각적 요소가 더욱 중요하게 작용하는 지점까지 파악하고자 한다.

이미지에 등장하는 객체 역시 본 연구에서 중요한 시각적 변수로 채택하였다. Guo et al. (2024)에 따르면 특정 장면에서 제품이나 사람을 포함하는 짧은 영상의 경우, 호의적인 평가를 위해 시청자의 감정을 자극하는 미학적 단서에 의존할 수 있으며, 소비자의 시청 반응에 영향을 미치는 또 다른 핵심 속성으로 작용할 수 있다고 보았다. Ling et al. (2021)에 따르면, 이미지 콘텐츠는 미적 구성과 눈길을 끄는 주제를 포함한 시각적 요소가 입소문에 영향을 미치며, 피사체를 크게 확대한 것보다 캐릭터의 얼굴, 표정을 클로즈업한 것이 더욱 긍정적인 영향을 미쳤음을 파악하였다. 이에 본 연구는 프레임에서 추출되는 객체의

수를 변수에 포함시켜, 이미지 내 객체가 소비자에게 미치는 영향도 분석하고자 한다.

2.3. 심미적 요소와 사용자 의사결정

이미지 및 영상 매체 등의 심미도는 가장 직관적이면서도 중요한 요소 중 하나이다. Jordan (1998)은 심미성은 사용자가 컴퓨터와 상호 작용할 때 경험하는 즐거움의 강력한 결정 요인이라는 것에 집중하였으며, Hoyer & Stokburher-Sauer (2012)은 일반적으로 소비자는 아름답거나 시각적으로 즐거운 대상, 사람 또는 환경을 선호하는 경향을 보인다는 것을 밝혔다. 또한 O'Brien & Toms (2008)은 특히 소셜 미디어 마케팅 같은 기술적 참여에는 시스템 사용성, 사용자 참여 및 경험 평가뿐만 아니라 심미성 같은 시각적 요소도 중요한 참여 요인이 된다고 언급하였다.

이러한 소비자 행동에 영향을 줄 수 있는 핵심적인 이미지의 미학적 요소로 심미도, 품질, 색상을 제안한다. 여러 연구에서 심미도, 품질, 색상 등은 중요한 미학적 요소로 확인되었으며, Datta et al. (2006)은 심미도는 추가적으로 다양한 시각적 요소와 함께 평가되며, 이미지의 특성(색상, 밝기 등)이 타고난 관심사와 관련 있다는 것이 많은 영역에서 알려져 있다고 언급하였다. 따라서 이와 같은 기존 연구들이 숏폼 콘텐츠를 수용하는 소비자의 인지와 참여행동에 어떠한 영향을 미치는지 분석하고자 한다.

최근 컴퓨터 비전 기술의 발달과 함께 심미도, 이미지 품질, 색상, 객체 판별 등 이미지 품질 평가(IQA)를 기술적으로 파악하기 위한 연구는 확대되고 있다. Murray et al. (2012)은 이미지의 미적 분석을 수행하기 위해 대규모 데이터 베이스인 AVA를 도입하여 주석을 통해 미적 품질 간의 상호

작용을 파악하였다. Messai et al. (2020)은 이미지 품질 평가(IQA)가 다양한 이미지 관련 프로세스 및 응용 분야에서 필수적인 역할을 한다는 점을 강조하였다. 또한, 김예은 & 이상용 (2025)은 게임 내의 다양한 이미지 특징에 대한 IQA를 통해 게임 성공에 있어 이미지가 중요한 요인으로 작용한다는 것을 파악하였으며, Hou et al. (2023)은 이미지의 채도와 밝기, 색조 등의 속성과 감정 등을 통해 이미지가 기부 및 크라우드펀딩 캠페인의 성공에 중요한 역할을 한다는 것을 입증하였다.

2.4. 소비자 참여행동에 관한 이론

이러한 숏폼 콘텐츠의 반응은 브랜드 함양 및 광고 마케팅에 대한 사용자들의 의사결정 과정에 영향을 미칠 수 있다. 숏폼 영상은 플랫폼에서 짧은 시간 내에 사용자의 지각을 이끌어 내야 하며, 사용자는 영상을 넘기지 않음으로써 영상의 소비를 결정한다.

박소진 (2022)은 좋아요, 구독, 공유, 댓글 같은 반응 행동 의도는 콘텐츠가 추천하는 제품의 구매의도에 영향을 미친다는 것을 확인하였다. Giombi et al. (2022)은 페이지 조회수, 소셜 미디어의 좋아요, 댓글, 공유 수, 클릭률 등이 일반적인 소비자 참여 측정 값이라고 설명하였다. 또한 Vinerean & Opreana (2021)은 이러한 소셜 미디어에서 소비자 참여가 고객 충성도에 영향을 미치는 요소임을 밝혔으며, Covington et al. (2016)은 시청 시간과 선택 같은 참여 지표는 유튜브의 검색 및 알고리즘에서 중요한 의미를 가진다고 지적하였다. 따라서 본 연구에서는 숏폼 영상의 사용자 참여 행동을 대표하는 지표로 영상을 시청한 조회수, 영상에 표시한 좋아요 수, 댓글 수, 그리고 소비자들이 작성한 댓글들의 감성의 긍정적인 정도를 제안한다.

이는 많은 학자들이 묘사한 소비자 구매 결정에

대한 인지-태도-행위라는 단계로 구분될 수 있다. 1974년 제시된 S-O-R 모델은 S(자극) - O(유기체) - R(반응)이라는 패턴으로 환경심리학에서 소비자 행동 영역으로 발전한 이론으로, 모든 측면에서 환경이 중요한 역할을 하여 개인의 내부 상태에 영향을 주어 행동 반응을 형성한다는 점을 지적하였다 (Mehrabian & Russell, 1974). 이후 이론은 발전을 통해 확장되었다. Richard & Chandra (2005)는 네트워크 환경, 사용자 특성, 소비자 반응 및 온라인 커뮤니케이션 결과를 조사하기도 하였다. 현재는 디지털 환경으로 확장되어, Tian et al. (2022)는 짧은 광고를 사용할 때 사용자가 자극을 인식하며, 이는 소비자의 심리적 욕구와 만족에 영향을 미쳐 소비 행동을 유발할 수 있다는 영향력으로 확장하였다. 나아가 Shi et al. (2023)은 S-O-R 모델을 기반으로 짧은 영상 마케팅 플랫폼의 특성에 따른 사용자 이용 및 만족 경험이 사용자의 반응과 마케팅 정보로 이어지는 영향을 분석하였다. Yang et al. (2024)은 숏폼 라이브 스트리밍 환경을 SOR 이론에 결합하여, 지각된 가치와 신뢰가 사용자 상호작용 행동을 매개하는 연쇄적 매커니즘을 규명하였다. 또한, Wei et al. (2024)은 숏폼의 기술적 어포던스가 사용자에게 자극으로 작용하여 내적 상태와 그에 따른 반응에 영향을 미치며, 이러한 유기체적 반응이 지속적인 참여 의도로 이어진다는 사실을 확인하였다.

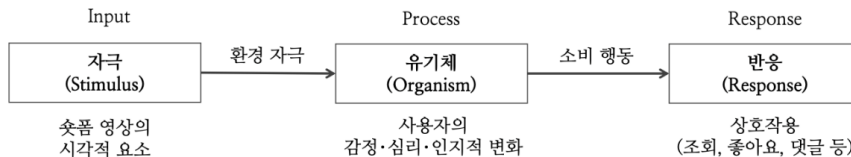
3. 연구의 구성

3.1. 연구 모형

본 연구는 숏폼 영상의 시각적 요소가 소비자의 감정·심리 반응과 행동 결정에 미치는 영향을 실증적으로 규명하고자 한다. S-O-R 모델의 이론적 틀을 확장하여 S(Stimulus)에 해당하는 숏폼 영상의 시각적 특성 (심미도, 품질, 색상, 구성 객체 등)이 O(Organism)에 해당하는 소비자의 정서적·인지적 반응을 유발하고, 이러한 반응이 R(Response)인 시청, 좋아요, 댓글 등의 상호작용 행동으로 이어지는 과정을 분석한다. 짧은 시간안에 재미·정보·핵심 메시지를 전달해야 하는 숏폼의 특성상, 시각적 인상을 형성하는 미적 요소는 소비자의 시청 여부뿐만 아니라 상호작용 지표 (좋아요, 댓글)에도 영향을 미칠 것으로 예상되며, 댓글의 긍정적 반응에도 상호 영향을 미칠 가능성을 탐구한다.

컴퓨터 비전 기술의 발전으로 영상 콘텐츠의 시각적 요소를 분석하고 수치화할 수 있게 되었으며, 많은 연구를 통해 이미지의 심미도와 품질, 색채, 객체 분류 등을 파악할 수 있는 기술 역시 확대되고 있다. 이를 통해 소비자 참여행동에 영향을 미치는 주요 미적 요소와 강하게 작용되는 지점을 도출할 수 있다. 이는 소셜 네트워크에서 인간과 컴퓨터의 상호작용에 대한 보다 실증적인 결과를 도출하는 데 기여할 것이다.

본 연구는 국내에서 가장 영향력 있는 숏폼



〈그림 1〉 숏폼의 소비자 반응 S-O-R 모델

플랫폼인 유튜브 쇼츠의 영상과 데이터를 API와 크롤링을 통해 수집해 온다. 다음으로 수집한 영상을 프레임별 이미지로 나누어 Google의 NIMA, Colorfulness, YOLO를 활용하여 이미지의 심미도와 품질, 색상의 다채로움, 이미지 내의 객체에 관한 수치를 도출한다. 변환된 정보를 바탕으로 다중회귀 분석을 실시하여 각 정보들이 소비자 참여행동과 상호작용 (조회수, 좋아요수, 댓글수, 댓글의 감성)에 얼마나 영향을 미치고, 어떤 시점에서 영향을 미치는지 분석하였다.

3.2. 연구 문제와 가설

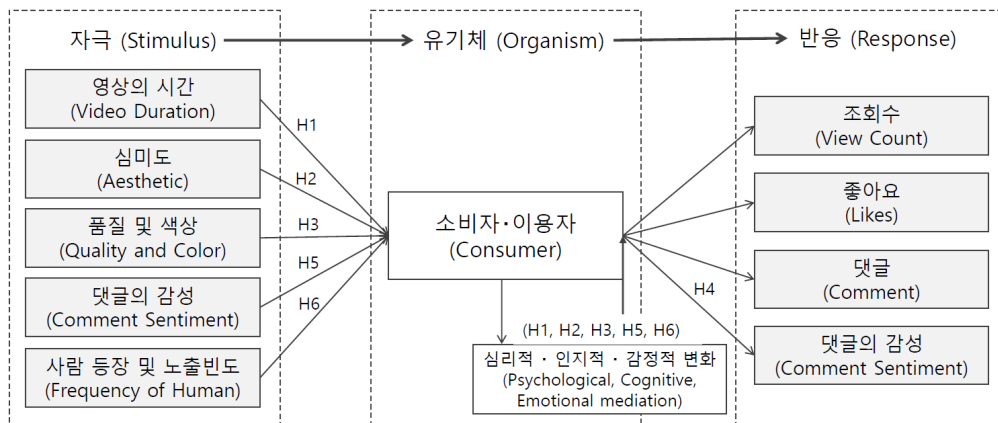
S-O-R 모델은 전통적인 시각에서부터 디지털 상호작용의 영역까지 확대되어 왔다. 숏폼 역시 짧은 시간 동안 자극을 통해 소비자의 감정적 변화를 이끌고, 영상과의 상호작용을 위해 행동을 보이는 동일한 프로세스를 가진다고 볼 수 있다. 특히 마케팅 및 광고, 이커머스 시장에서 시각적인 요소가 소비자 참여행동에 중요한 영향을 미치는 요소이기 때문에, 숏폼에서도 이러한 영향은 동일하게 중요성을 가질 것이다.

본 연구에서는 이러한 배경을 바탕으로 숏폼 영상의 시각적 요인들과 이미지의 미학적인 특징들을 최근 발전하고 있는 컴퓨터 비전 기술을 통해 실증적으로 분석하고자 한다. 또한 시각적 요소가 가장 강력하게 소비자 참여행동에 영향을 미치는 지점을 파악하고, 이를 통해 영상에서 이미지가 가지는 중요성이 주는 인사이트를 도출하고자 한다.

연구 문제 1. Short-form 영상의 이미지적 특징은 소비자의 참여행동에 어떠한 영향을 미치는가?

연구 문제 2. Short-form 영상에서의 소비자 참여행동에 중요한 영향을 미치는 지점은 어디인가?

S-O-R 모델의 관점에서 <그림 2>의 ‘Stimulus (자극)’는 숏폼 환경 특유의 감각적 요소로 개념화된다. 구체적으로, 첫 1-3초 내 고채도 색상, 굵은 자막, 얼굴 클로즈업, 빠른 컷 전환, 댓글 관련 텍스트 등은 자극 강도 (intensity)와 참신함 (novelty)을 증대시켜 주의 포착을 극대화한다.



<그림 2> 가설에 따른 연구 모형 도식화

이러한 자극은 ‘Organism(유기체)’ 단계에서 호기심과 기대, 긍정적 각성(즐거움·흥분) 및 인지적 부하의 완화/집중을 유도한다. 그 결과 ‘Response(반응)’ 단계에서 시청 지속·완주, 재시청, 좋아요·공유·팔로우, 추가적 댓글의 작성 등으로 이어지는 경로가 형성된다. 이를 위해 연구에서 설정한 가설은 다음과 같다.

H1 : Short-form 영상의 시간은 소비자의 참여행동에 영향을 미칠 것이다.

숏폼은 짧은 영상 속에서 소비자의 선택을 받는 것이 핵심이며, 주어진 시간 안에 콘텐츠의 핵심을 명확하게 전달하는 것이 필수적이다. García-Avilés & de Lara (2018)의 연구는 긴 영상은 관심을 끌기 어려우며, 스마트폰에서는 짧은 영상의 선호가 높다는 것을 언급하였다. 반면, Smith et al. (2018)은 플랫폼 추천 알고리즘은 더욱 길고 인기 있는 영상을 유도한다는 경향이 있다고 밝혔다. 따라서 숏폼 영상의 길이는 소비자가 영상의 재미와 정보를 수용하는 방식에 차이를 가져올 수 있기 때문에 본 연구는 숏폼 영상의 길이가 소비자 참여행동에 유의한 영향을 미칠 것이라는 가설을 제시한다.

H2 : Short-form 영상의 이미지 심미도는 소비자의 참여행동에 영향을 미칠 것이다.

Jordan (1998)에 따르면 이미지의 심미도는 사용자가 콘텐츠와 상호작용할 때 경험하는 시각적 즐거움과 관련이 있다. 높은 심미적 가치는 소비자의 관심을 끌고 긍정적인 행동으로 유도될 수 있다. 또한 Hoyer & Stokburher-Sauer (2012)은 미적 가치가 높은 상품과 서비스의 경우 소비자 취향에 따른 판단과 결정 과정에서 중요한 역할을 할 수 있다고 보았다. 따라서 숏폼 영상을

이루는 이미지 요소들의 심미도 역시 소비자 참여 행동과 행동에 영향을 미칠 것이라는 가설을 제시한다.

H3 : Short-form 영상의 이미지 품질 및 색상 다채로움이 높을수록 소비자 참여행동에 영향을 줄 것이다.

이러한 영향력은 추가적인 시각적 미학에 동일하게 작용할 수 있다. Maros et al. (2019)은 고품질의 이미지와 다채로운 색상은 소비자의 시각적 관심을 유도하고, 긍정적인 브랜드 인식을 형성하는 데 중요한 역할을 한다고 말하며, 전자상거래 플랫폼 등지에서 제품 이미지의 품질과 색상 같은 미학이 클릭률에 영향을 미친다는 결과를 도출하였다. 따라서 이미지의 품질과 색상 역시 소비자 참여행동에 영향을 미칠 것이라는 가설을 제시한다.

H4 : Short-form 의 댓글 감성은 심미도와 품질 및 색상이 높을수록 긍정적일 것이다.

본 연구에서는 숏폼에 대한 사용자들의 상호작용적인 요소로 댓글의 감성을 함께 고려한다. 비디오 콘텐츠에 대한 여론이나 태도 등에 대한 감정 분석의 역할을 사용자뿐만 아니라, 콘텐츠 제작자 및 마케터에게도 중요한 요소이다. Sheng et al. (2020)에 따르면, 소셜 미디어에서 상상력을 자극하고 감정적으로 몰입적인 콘텐츠는 고객에게 영감을 주고, 긍정적인 행동과 의도를 크게 높일 수 있다고 보았다. 따라서 댓글의 감성은 조회수나 좋아요 수, 댓글 수와 같은 숏폼에 대한 사용자의 행동 및 참여 요소로, 앞서 제시된 심미도, 품질, 색상 등이 높을수록 긍정적인 영향을 미칠 것이라는 가설을 제시한다.

H5 : Short-form의 댓글 감성이 긍정적일수록 소비자 참여행동에 더욱 강한 영향을 줄 것이다.

또한 댓글의 감성은 조회수, 좋아요 수, 댓글 수 같은 샷폼의 소비자 참여행동 및 상호작용에도 동일하게 영향을 주는 변수일 수 있다. Guo et al. (2020)은 기분 좋은 온라인 고객 리뷰가 불쾌한 온라인 고객 리뷰에 비해 구매 가능성이 높다는 것을 보여준다고 말하였다. 따라서 댓글의 감성은 상호작용의 영향을 평가받는 변수일 뿐만 아니라, 소비자 참여행동에 영향을 줄 수도 있는 변수가 될 것이라는 가설을 제시한다.

H6 : Short-form에서 사람의 등장 및 노출 빈도는 소비자 참여행동에 영향을 줄 것이다.

광고 및 이커머스 등에서 객체의 효과는 연구를 통해 증명된 결과로, 특히 소셜 미디어 등에서 얼굴 이미지는 사람의 특성, 성격, 의도 등 감정을 이해하는 시각적 주의를 끄는 자극으로 간주되기도 한다. Bakhshi et al. (2014)은 대규모 소셜 미디어 데이터에서 사람 얼굴이 포함된 사진은 좋아요와 댓글을 유의미하게 증가시킨다는 결과를 확인하였다. 이는 샷폼 콘텐츠에서도 동일하게 적용될 수 있다. Guo et al. (2024)은 제품이나 사람을 포함하는 짧은 영상의 경우 시청자의 긍정적인 감정을 자극하는 미학적 단서에 의존할 수 있다고 보았다. 따라서 샷폼이 사람과 컴퓨터 간의 상호 관계와 참여를 유도한다는 점에서, 사람의 등장 및 노출은 소비자 참여행동에 영향을 미칠 것이라는 가설을 제시한다.

본 연구는 다음의 가설들을 바탕으로 샷폼 영상의 시각적 요소가 소비자 참여행동에 미치는 영향을 분석하고, 이를 통해 시각적 요소의 중요성을 밝히고자 한다. 이 결과는 디지털 콘텐츠 전략 수립 및 마케팅에서 시각적 요소를 활용한

소비자 참여 극대화에 실질적인 도움과 방안에 대한 중요한 인사이트를 제공할 수 있다.

3.3. 데이터 수집 및 방법론

3.3.1 쇼츠 데이터 수집 및 변환

본 연구는 샷폼의 시각적 요소가 소비자 참여 행동에 미치는 영향을 분석하기 위해 유튜브 쇼츠 데이터를 수집하였다. 로그아웃 상태에서 자동 추천 알고리즘을 통해 다양한 주제를 포함한 3,009개의 국내 샷폼 영상을 추출하였고, 유튜브 API를 활용해 조회수, 좋아요 수, 댓글, 영상 제목, URL, 카테고리, 태그 등의 메타 데이터를 확보하였다.

댓글의 경우, API와 자체 크롤링을 통해 수집한 후, 감성 분석을 수행하였다. KNU 감성 사전을 활용한 렉시콘 기반 분석을 통해 긍부정 단어의 점수를 합산하였으며, -1과 1사이로의 스케일링하여 수치화하는 과정을 거쳤다. KNU감성 사전은 군산대학교에서 개발한 국내에 대표적인 감성 사전 중 하나로, Bi-LSTM을 바탕으로 도메인에 영향을 받지 않는 보편적인 기본 감정 표현을 긍부정으로 구성한 사전이다(Park et al., 2018).

3.3.2 영상 데이터 변환

샷폼 영상의 시각적 요소를 추출하기 위해 영상을 프레임별로 구분하여 이미지화하였다. 약 10초에서 60초 사이의 영상에서 특정 프레임을 선택해 시퀀스 이미지로 변환하고, 이를 분석하여 시각적 요소를 파악하였다. Yan et al. (2020)은 영상에서 키 프레임을 감지하는 것은 비디오 분류, 인식, 요약과 같은 프로그램에 적용될 수 있으며, 이러한 작업은 소수의 키 프레임만을 사용

하여 효율적으로 수행할 수 있다고 보았다. 따라서 본 연구에서는 영상을 300프레임으로 나누어 이미지화 한 후 각각의 이미지를 분석하였다. 프레임 수는 샷폼 영상 길이를 대표하는 변수로 사용하였고, 이를 통해 프레임별 시각적 요소가 소비자 참여행동에 미치는 영향을 조사하였다. 각 프레임의 심미도, 품질, 색상, 객체 등을 분석하여 소비자 참여행동에 중요한 영향을 미치는 요소를 규명하였다.

3.3.3 컴퓨터 비전 알고리즘

본 연구의 핵심은 시각적 요소가 샷폼 콘텐츠 소비자에게 미치는 영향을 평가하는 것이다. 이를 객관적으로 이해하고 컴퓨터와 사람 간의 상호작용에 대한 제언을 위해, 컴퓨터 비전 기술을 활용한 이미지 분석을 진행하고자한다. 최근 컴퓨터 비전 기술은 크게 발전하였으며, 이는 샷폼 영상의 시각적 요소를 정밀하게 평가하고 객관적이며 실증적인 결과를 도출하는 데 기여할 수 있다.

1) NIMA

이미지의 미학적 요소와 품질을 평가하기 위한 방법론으로 Google의 NIMA(Neural Image Assessment)를 사용한다. NIMA 신경 이미지 평가는 deep CNN을 기반으로 특정 이미지가 기술적 혹은 미적으로 매력적인지 예측한다. 이는 심층 객체 인지 네트워크로서 인간의 인지와 높은 상관관계를 가지며, 사람이 파악하는 안정적인 이미지 점수를 도출할 수 있다(Talebi & Milanfar, 2018). NIMA는 미적 평가(Aesthetic Score)와 품질 평가(Quality Score)로 나뉘며, 1에서 10점의 척도로 평점 분포를 이룬다. 이는 다른 이미지 평가가 접근하는 방식 보다 인간의 선호도를 파악하는

것에 긍정적인 영향을 미치며, 샷폼의 소비자가 인지하는 이미지의 심미도와 품질에도 동일한 인지 결과를 미칠 수 있다.

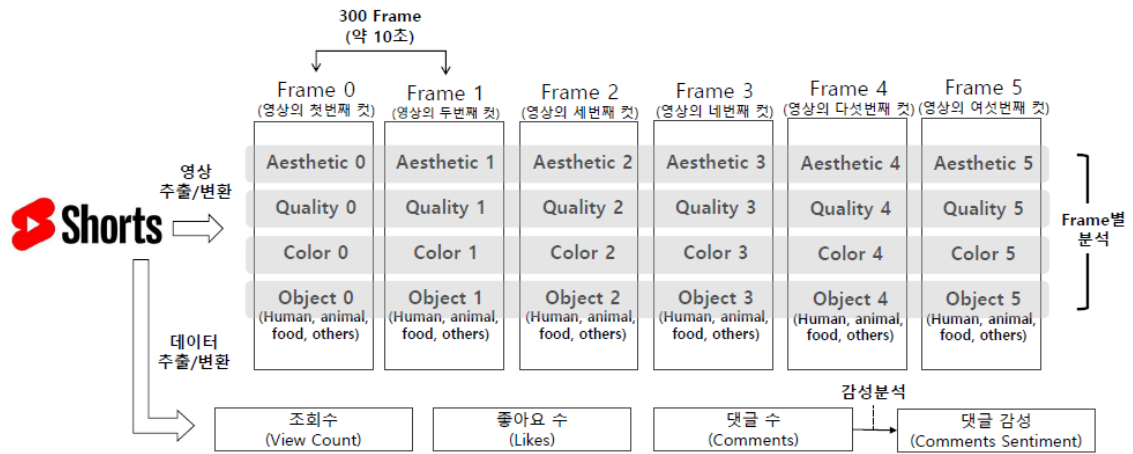
2) Colorfulness metric

Colorfulness는 자연 이미지의 색상도를 정량화하고 이미지를 평가하여 이미지 품질 평가 프레임워크에 통합하는 것을 목표로 하며, Colorfulness metric을 기반으로(Hasler & Suesstrunk, 2003) Open CV와 Python을 통해 색상의 다채로움을 정량화하여 점수가 높을수록 색상이 더욱 다채롭다는 것을 의미한다. Singh (2006)은 신중한 색상 사용은 경쟁사와의 차별성에 도움이 될 뿐만 아니라, 기분과 느낌에 긍정적 또는 부정적인 영향을 미치며 제품에 대한 태도에도 영향을 미칠 수 있다고 밝혔다. 따라서 샷폼 이미지에서 색상의 다채로움이 소비자 참여행동에 미치는 결과를 분석하는데 활용하고자 한다.

3) YOLO

YOLO(You Only Look Once)는 단일 신경망을 사용해 객체를 빠르고 정확하게 예측하고 분류하는 컴퓨터 비전 알고리즘으로, CNN 레이어가 ImageNet 데이터 세트에서 훈련되었으며 다른 객체 탐지 시스템에 비해 뛰어난 속도와 정확성을 자랑한다(Redmon et al., 2016). 본 연구에서는 개량형인 YOLOv5를 사용해 샷폼에서 추출된 객체를 분석하고, 사람, 동물, 음식 및 기타 객체의 존재 여부와 소비자 참여행동 간의 관계를 파악할 것이다.

본 연구에서 수집된 데이터를 기반으로 한 전체적인 연구 모형은 <그림 3> 과 같다.



〈그림 3〉 데이터 수집 및 변환을 통한 연구 절차 모형

3.3.4 변수 설명

본 연구에서 다중 회귀 분석에 사용한 변수는 아래 <표 1>과 같다.

종속변수는 숏폼 콘텐츠 참여행동과 상호작용을 대표할 수 있는 지표로, 조회수와 좋아요 수, 댓글 수와 댓글의 감성 점수를 사용했다. 통제변수로는

〈표 1〉 변수 설명

구분	변수명	변수 설명
종속변수	조회수	숏폼 영상의 시청 수
	좋아요 수	숏폼 영상의 좋아요 수
	댓글 수	숏폼 영상의 댓글 수
종속/독립변수	댓글 감성	숏폼 영상의 댓글 감성 점수 (-1부터 1)
독립변수	프레임 수	영상을 300프레임으로 나누었을 때 프레임 수
	이미지 심미도 (0~5)	0부터 5번째 프레임별 심미도 (NIMA)
	이미지 품질 (0~5)	0부터 5번째 프레임별 품질 (NIMA)
	이미지 색상 (0~5)	0부터 5번째 프레임별 색상 (Colorfulness)
	사람의 수 (0~5)	0부터 5번째 프레임별 사람의 수 (YOLO)
	음식의 수 (0~5)	0부터 5번째 프레임별 음식의 수 (YOLO)
	동물의 수 (0~5)	0부터 5번째 프레임별 동물의 수 (YOLO)
	객체의 수 (0~5)	0부터 5번째 프레임별 기타 객체의 수 (YOLO)
통제변수	카테고리 (11개)	숏폼 콘텐츠별 구분되는 카테고리
	제목의 글자수	숏폼 콘텐츠 타이틀의 글자 수
	태그 수	숏폼 콘텐츠에 포함되어 있는 태그의 수
	사운드 템포	숏폼의 음악 및 사운드의 템포
	배경음악 여부	숏폼 영상의 배경음악 여부

숏폼 콘텐츠의 제목 글자의 수, 태그 수, 사운드 템포, 배경음악의 여부를 사용했다. 독립변수의 경우 본 연구에서 시각적 요소가 줄 수 있는 영향을 분석하기 위한 심미도, 품질, 색상, 등장 객체, 댓글의 감성을 사용했는데, 해당 시각적 요소가 영향을 주는 시점을 파악하기 위해 영상을 프레임별로 구분하여 특정 시점에서의 영향력을 파악한다.

영상의 총 표본 수는 3,009개이며, 이를 기반으로 숏폼의 메타 데이터, 추출된 데이터, 그리고 컴퓨터 비전 알고리즘을 통한 이미지의 미학적 요소를 도출하였다. 프레임별로 구분된 이미지 데이터는 영상 길이에 따라 그 수가 다르며, 0번째 프레임은 3,009개로 동일하나, 이후 프레임으로 갈수록 표본 수가 감소하여 5번째 프레임은 940개까지 줄어든다. 추가로 숏폼 특성을 분석하기 위해 카테고리별 더미 변수로 추가했으며, 통제 변수로 제목 글자 수, 태그 수, 영상의 사운드 템포와 배경음악 여부를 포함하였다.

4. 연구 결과

4.1. 기초 통계량 분석 및 상관관계

본 연구에서는 컴퓨터 비전 알고리즘을 활용하여 소비자 참여행동에 영향을 미치는 숏폼 영상의 이미지적 요소를 분석하였다. 국내에서 가장 활발히 사용되는 유튜브 쇼츠를 기준으로, 영상을 프레임별로 구분하고 각 이미지의 미학적 요소를 추출한 후, 다중회귀분석을 진행하였다. 이를 통해 소비자 참여행동을 대표하는 종속 변수

와 이미지의 미학적 요소, 숏폼의 메타 데이터 간의 상관관계를 분석하여 영향력을 파악하였다.

변수 간의 상관관계의 경우 다른 독립 변수들 간에는 큰 영향력이 나타나지 않지만, 컴퓨터 비전 알고리즘을 통해 파악한 심미도, 품질, 색상의 경우 각 프레임 간에 일정 수준 이상의 상관관계를 가진다. 이는 한 영상에 도출된 각각의 이미지 사이에 획기적인 변화가 비교적 적으며, 그로 인한 시각적 요소들의 점수 분포가 유사하게 분포되기 때문인 것으로 보인다. 객체 추출에 관련된 변수의 경우 비교적 프레임 간 변화가 크지 않는 사람의 빈도는 약한 관계성을 보인다. 따라서 각각의 프레임을 따로 파악하여 미적 요소가 종속 변수에 미치는 영향력을 파악할 필요성이 있다. 연구에 사용한 변수의 수가 많아서, 상관관계 분석표는 생략되어 있다.

독립변수들 간의 다중공선성을 확인하기 위하여 프레임별 심미도 모형에 대해 VIF를 산출한 결과는 <표 2>와 같다. 이때의 종속변수는 조회수로 설정하여 VIF를 계산하였다. 0번째 컷(NIMA0)을 포함한 회귀모형 VIF 값은 모두 1.02~1.06 수준으로 나타나 다중공선성 문제는 존재하지 않는 것으로 판단된다. 다른 컷들에 대해서도 동일한 방식으로 검토한 결과, 모든 VIF 값이 1~2사이의 낮은 수준으로 일관되게 나타나 독립변수 간 상관성은 낮음을 확인하였다.

또한 모델 적합도 진단한 결과 소셜미디어 데이터의 특성상 발생할 수 있는 부분을 제외하면 정규성과 등분산성에 크게 위배하는 부분은 나타나지 않았으며, 잔차독립성과 극단성에서도 영향점이 발견되지 않았다. 따라서 본 연구의 회귀모형은

<표 2> 조회수에 대한 프레임별 심미도 0번째 분석에 대한 다중공선성 분석

독립변수명	프레임수	댓글감성	사람의수	동물의수	음식의수	기타	NIMA0
VIF 값	1.021	1.013	1.056	1.008	1.002	1.036	1.033

전반적으로 안정적이며, 통계적 가정 위배로 인한 왜곡 가능성은 크지 않은 것으로 판단된다.

4.2. 심미도 중심 회귀분석 결과

4.2.1 조회수에 대한 프레임별 심미도

다중 회귀 분석을 통해 심미도가 종속변수에

미치는 영향력에 대해 파악하였다. 분석 횟수마다 각각의 프레임에 대한 심미도의 영향력을 파악하여 분석을 진행하였다. 각각의 프레임별 영향력과 조회수와의 관계를 분석한 결과는 다음과 같다.

<표 3>을 살펴보면 모든 프레임의 심미도가 샷폼 조회수와 유의한 관계를 보이며, 각각의 프레임에서 심미도는 조회수와 정(+)의 관계를

<표 3> 조회수에 대한 프레임별 심미도 회귀분석 결과표

종속변수		조회수 (Views)						
구분	변수	분석 0	분석 1	분석 2	분석 3	분석 4	분석 5	
표본		3,009	2,903	2,507	2,035	1,516	940	
상수항		9.649*** (0.671)	10.535*** (0.708)	11.175*** (0.755)	12.004*** (0.814)	13.167*** (1.070)	26.750*** (2.790)	
독립변수	심미도	NIMA A0	0.436*** (0.093)					
		NIMA A1		0.340** (0.102)				
		NIMA A2			0.3111** (0.105)			
		NIMA A3				0.392*** (0.109)		
		NIMA A4					0.391** (0.126)	
		NIMA A5						0.454** (0.162)
	프레임 수		0.292*** (0.031)	0.198*** (0.034)	0.105* (0.041)	0.031 (0.059)	-0.269** (0.107)	-2.678*** (0.417)
	댓글 감성		1.953* (0.935)	1.990* (0.919)	1.336 (0.900)	0.775 (0.913)	0.577 (1.045)	0.6230 (1.189)
	객체 추출	사람의 수	0.057* (0.025)	0.051* (0.025)	0.046* (0.028)	0.043 (0.029)	0.039 (0.030)	0.030 (0.041)
		동물의 수	0.201 (0.189)	-0.213 (0.173)	0.239 (0.174)	-0.165 (0.210)	0.069 (0.241)	0.141 (0.245)
		음식의 수	-0.180 (0.138)	0.032 (0.125)	0.133 (0.087)	-0.009 (0.926)	0.073 (0.108)	0.010 (0.150)
		기타 객체수	0.014 (0.024)	-0.008 (0.026)	-0.006 (0.023)	-0.057* (0.030)	-0.029 (0.034)	0.032 (0.045)
	R -squared		0.099	0.079	0.066	0.078	0.095	0.155
F		20.124***	10.455***	3.855***	2.705**	4.388***	8.707***	

$p^* < 0.1, p^{**} < 0.05, p^{***} < 0.001$

나타낸다. 특히 초기 프레임의 심미도가 큰 영향을 미치고, 이후 프레임으로 갈수록 감소하는 경향을 보이는 것을 확인할 수 있다. 이는 심미도가 소비자의 시선을 끌고 긍정적인 영향을 줄 수 있음을 의미하며, 초기 이미지가 선포 시청 결정에 특히 중요한 요소인 것으로 해석된다. 그러나 4번째 프레임부터 다시 계수가 증가하는 현상을 볼 수 있는데, 이는 프레임 수에 따라 표본이 줄어드는 시점부터 발생하는 효과로 고려된다.

선포의 프레임 수는 조회수에 유의한 영향을 미치며, 초기 프레임은 정(+)의 관계를 보이지만, 4번째 프레임 이후 감소하는 경향을 보인다. 이는 선포 영상이 적정 길이가 일정 수준까지는 조회수에 긍정적인 영향을 미치지만, 지나치게 길어질 경우 부정적인 영향을 줄 수 있음을 나타낸다.

댓글의 감성 점수 또한 조회수에 정(+)의 유의미한 영향을 미친다. 긍정적인 감정은 소비자 참여행동에 긍정적 영향을 미치며, 초기 이미지에서 높은 영향을 미치지만, 시간이 지남에 따라 인지적 작용이 점차적으로 줄어들어 이후에는 유의미한 관계를 미치지 못하는 것으로 해석할 수 있다.

한편, 객체 추출 결과에서는 초기 이미지에서 사람의 존재와 등장 빈도가 유의미한 긍정적 결과를 보였다. 초기 사람 수의 영향력이 가장 높게 나타났으며, 다른 변수들과 유사하게 이후 감소하는 경향을 보였다. 또한, 3번째 프레임에서 기타 객체 수는 약한 음의 관계를 나타내는데, 이는 프레임 수가 줄어드는 것이 효과적인 지점과 동일하며 지나친 객체의 노출이 소비자 피로도를 높이고 이탈을 초래할 수 있음을 시사한다.

4.2.2 좋아요 수에 대한 프레임별 심미도

각 프레임별 심미도와 좋아요 수의 관계를 분석한

결과는 다음과 같다. <표 4>에서 각 프레임의 심미도는 좋아요 수와 유의한 정(+)의 관계를 보였으며, 조회수와 유사하게 초기 심미도 계수는 높게 나타나고 프레임이 지날수록 감소하는 경향이 있었다. 이는 미학적으로 우수한 이미지가 소비자에게 긍정적인 인상을 주어 참여를 유도하고 좋아요 수에 긍정적인 영향을 미친다고 해석할 수 있다. 또한, 표본이 줄어드는 프레임에서는 심미도가 다시 상승하는 경향 역시 확인되었다.

프레임 수 역시 좋아요 수에 유의한 영향을 미치는 것으로 분석되었으며, 초기까지는 긍정적인 관계를 보이다가 4번째 프레임부터 감소하는 경향이 나타났다. 이는 적절한 선포 길이가 좋아요 수에 중요한 변수로 작용함을 시사한다.

댓글의 감성도 좋아요 수와 정(+)의 관계를 보였다. 긍정적인 반응이 높을수록 소비자 만족도와 상호작용이 증가하며, 초기 이미지에서 강한 인지 효과를 미친 후 점차 줄어드는 것으로 해석된다. 좋아요 수는 조회수보다 인지 효과가 더 오래 유지되며, 3번째 프레임부터는 댓글 감성과 유의한 관계가 없었다.

객체 추출 결과, 사람의 수는 좋아요에 유의미한 영향을 미치지 않았다. 사람의 존재가 시청 유지에는 기여하지만, 좋아요 수 증가와는 관계가 다르다는 해석이 가능하다. 반면, 2번째 프레임에서 동물의 등장은 소비자에게 긍정적인 인상을 주어 정(+)의 관계를 보였고, 기타 객체의 수는 조회수와 동일하게 3번째 프레임에서 음의 영향을 미쳤다.

4.2.3 댓글 수에 대한 프레임별 심미도

각 프레임별 심미도와 댓글 수의 관계를 분석한 결과는 다음과 같다.

<표 4> 좋아요 수에 대한 프레임별 심미도 회귀분석 결과표

종속변수		좋아요 수 (Likes)						
구분	변수	분석 0	분석 1	분석 2	분석 3	분석 4	분석 5	
표본		3,009	2,903	2,507	2,035	1,516	940	
상수항		5.411*** (0.643)	6.277*** (0.676)	7.326*** (0.717)	8.108*** (0.776)	9.619*** (1.028)	19.986*** (2.726)	
독립변수	심미도	NIMA A0	0.440*** (0.090)					
		NIMA A1		0.357*** (0.097)				
		NIMA A2			0.263*** (0.099)			
		NIMA A3				0.351** (0.104)		
		NIMA A4					0.312** (0.121)	
		NIMA A5						0.458** (0.156)
	프레임 수		0.301*** (0.030)	0.204*** (0.032)	0.107*** (0.039)	0.038 (0.056)	-0.253** (0.103)	-2.207*** (0.402)
	댓글 감성		2.472** (0.895)	2.336** (0.877)	1.634* (0.854)	1.203 (0.871)	1.098 (1.003)	1.189 (1.146)
	객체 추출	사람의 수	0.026 (0.024)	0.027 (0.024)	0.028 (0.027)	0.028 (0.027)	0.029 (0.029)	0.009 (0.040)
		동물의 수	0.240 (0.181)	-0.167 (0.165)	0.309* (0.165)	-0.121 (0.200)	0.160 (0.232)	0.142 (0.237)
		음식의 수	-0.253 (0.132)*	-0.019 (0.120)	0.044 (0.083)	-0.060 (0.088)	0.078 (0.104)	-0.027 (0.145)
		기타 객체수	-0.001 (0.023)	-0.014 (0.025)	-0.018 (0.022)	-0.048* (0.029)	-0.031 (0.033)	0.002 (0.043)
	R -squared		0.107	0.087	0.076	0.092	0.107	0.163
F		19.159***	9.411***	3.444***	2.706**	4.023***	7.544***	

$p^* < 0.1, p^{**} < 0.05, p^{***} < 0.001$

<표 5>에서 댓글 수와 심미도의 관계에서 초기 프레임은 유의미한 정(+) 효과를 보였으나, 그 이후의 프레임에서는 유의미한 영향을 나타내지 않았다. 특히 4번째 프레임부터는 댓글 수가 증가하는 경향이 확인되었다. 이는 첫 번째 프레임의 심미도가 시청자의 흥미를 유발해 댓글 작성에 영향을 미친다는 것을 시사한다. 초기

프레임은 시청 결정을 유도하지만, 이후 프레임들은 이미 시청을 결정한 후의 내용이기 때문에 댓글에 미치는 영향은 적을 수 있다. 프레임 수 역시 댓글 수와 유사한 경향을 보였다.

댓글의 감성과 댓글 수는 유의미한 관계를 보이지 않았다. 조회수와 좋아요 수는 긍정적인 댓글이 영상에 대한 만족도와 흥미를 나타내지만,

〈표 5〉 댓글 수에 대한 프레임별 심미도 회귀분석 결과표

종속변수		댓글 수 (Comments)						
구분	변수	분석 0	분석 1	분석 2	분석 3	분석 4	분석 5	
표본		3,009	2,903	2,507	2,035	1,516	940	
상수항		3.285*** (0.516)	3.999*** (0.548)	4.583*** (0.596)	4.465*** (0.661)	6.372*** (0.878)	14.895*** (2.316)	
독립변수	심미도	NIMA A0	0.172** (0.072)					
		NIMA A1		0.073 (0.079)				
		NIMA A2			0.054 (0.083)			
		NIMA A3				0.219** (0.088)		
		NIMA A4					0.162 (0.104)	
		NIMA A5						0.220* (0.132)
	프레임 수		0.207*** (0.024)	0.148*** (0.026)	0.100*** (0.033)	0.06 (0.048)	-0.279*** (0.088)	-1.840*** (0.341)
	댓글 감성		0.070 (0.719)	-0.033 (0.711)	-0.445 (0.710)	-0.7130 (0.741)	-1.110 (0.857)	-0.714 (0.973)
	객체 추출	사람의 수	0.003 (0.019)	0.020 (0.020)	-0.006 (0.022)	-0.007 (0.023)	0.008 (0.025)	-0.002 (0.034)
		동물의 수	0.137 (0.145)	-0.131 (0.134)	0.151 (0.137)	-0.173 (0.171)	0.112 (0.198)	-0.083 (0.201)
		음식의 수	-0.225** (0.106)	-0.033 (0.097)	0.042 (0.069)	-0.044 (0.075)	0.037 (0.089)	0.051 (0.123)
		기타 객체수	0.014 (0.019)	-0.011 (0.021)	-0.008 (0.018)	-0.041* (0.025)	-0.018 (0.028)	0.005 (0.037)
	R -squared		0.067	0.054	0.046	0.059	0.065	0.096
F		11.117***	5.021***	2.491*	1.899	4.121***	6.579***	

$p^* < 0.1, p^{**} < 0.05, p^{***} < 0.001$

댓글 수는 댓글 작성 동기에 더 큰 영향을 받는 것으로 해석된다. 댓글 수는 영상의 주제나 상호 작용 유도 특성에 더 영향을 받는 것으로 보인다.

추출된 객체가 댓글에 미치는 정(+) 영향은 크지 않았다. 3번째 프레임부터는 기타 객체의 증가가 댓글 감소의 원인이 되었고, 초기 프레임에서는 음식 수가 댓글 수를 줄이는 요인으로 나타났다.

이는 준비된 재료 이미지보다는 완성된 음식 이미지가 댓글에 더 효과적이라는 추측이 가능하다.

4.2.4 댓글 감성에 대한 프레임별 심미도

각 프레임별 심미도와 댓글 감성 간의 관계를 분석한 결과는 다음과 같다.

<표 6> 댓글 감성에 대한 프레임별 심미도 회귀분석 결과표

종속변수		댓글 감성 (Text Sentiment)						
구분	변수	분석 0	분석 1	분석 2	분석 3	분석 4	분석 5	
표본		3,009	2,903	2,507	2,035	1,516	940	
상수항		-0.155*** (0.013)	-0.157*** (0.014)	-0.138*** (0.017)	-0.153*** (0.020)	-0.181*** (0.026)	-0.201** (0.078)	
독립변수	심미도	NIMA A0	0.004* (0.002)					
		NIMA A1		0.005** (0.002)				
		NIMA A2			-0.001 (0.002)			
		NIMA A3				0.001 (0.003)		
		NIMA A4					0.002 (0.003)	
		NIMA A5						0.001 (0.004)
	프레임 수		0.001 (0.001)	0.000 (0.001)	0.001 (0.001)	0.002 (0.001)	0.003 (0.003)	0.004 (0.012)
	객체 추출	사람의 수	0.001** (0.000)	0.000 (0.001)	0.001 (0.001)	-0.000 (0.001)	-0.000 (0.001)	0.002 (0.001)
		동물의 수	-0.004 (0.004)	0.005 (0.003)	0.002 (0.004)	-0.002 (0.005)	0.001 (0.006)	0.006 (0.007)
		음식의 수	0.001 (0.003)	-0.000 (0.003)	-0.001 (0.002)	-0.003 (0.002)	-0.001 (0.003)	0.003 (0.004)
		기타 객체수	0.000 (0.000)	-0.000 (0.001)	-0.000 (0.001)	-0.0003 (0.001)	0.000 (0.001)	-0.000 (0.001)
	R -squared		0.072	0.071	0.066	0.063	0.088	0.097
	F		6.693***	7.700***	4.324***	2.962**	3.915***	1.4159

$p^* < 0.1, p^{**} < 0.05, p^{***} < 0.001$

<표 6>에서 솟폼의 초기 심미도가 댓글 감성에 유의미한 정(+) 영향을 미치는 것으로 확인되었다. 이는 심미도가 소비자에게 긍정적인 인상을 주고, 댓글 감성에도 긍정적인 영향을 미친다는 선행 연구와 일치한다. 하지만 시간이 지날수록 이러한 인지 작용은 감소하는 경향을 보였다. 댓글 감성 점수라는 특성으로 인해 시간적 변화

와 관련된 프레임 수와는 유의미한 영향을 보이지 않았다.

객체 추출에서는 초기 사람의 존재와 등장 빈도가 댓글 감성에 긍정적인 영향을 미치는 것으로 나타났다. 사람의 존재는 소비자들에게 인간적인 연결과 공감을 유발해 긍정적인 감정을 강화했으나, 시간에 따라 이후 프레임에서는 유의미한 결과를 보이지 않았다.

4.3. 품질 중심 회귀분석 결과

다중 회귀 분석을 통해 이미지의 품질이 종속 변수에 미치는 영향력에 대해 파악하였다. 영상의 카테고리과 함께 각 프레임 별로 미치는 영향력을

분석한 결과는 <표 7>과 같다. 분석 횟수 마다 각각의 프레임에서의 품질을 파악하고 색상의 다채로운 정도를 함께 평가하여 전반적인 품질 요소와 소비자 참여행동 간의 관계를 파악하였다.

<표 7> 조회수에 대한 프레임별 품질 및 색상 회귀분석 결과표

종속변수		조회수 (Views)						
구분	변수	분석 0	분석 1	분석 2	분석 3	분석 4	분석 5	
표본		3,009	2,903	2,507	2,035	1,516	940	
상수항		10.972*** (0.656)	12.349*** (0.666)	12.743*** (0.692)	14.491*** (0.751)	12.743*** (0.692)	29.465*** (2.740)	
독립변수	품질 및 색상	NIMA Q0	0.100 (0.078)					
		Color 0	0.005* (0.003)					
		NIMA Q1		-0.0750 (0.075)				
		Color 1		0.004 (0.003)				
		NIMA Q2			0.0023 (0.076)			
		Color 2			0.0003 (0.003)			
		NIMA Q3				-0.132 (0.082)		
		Color 3				0.002 (0.003)		
		NIMA Q4					0.002 (0.003)	
		Color 4					0.000 (0.003)	
		NIMA Q5						-0.040 (0.121)
		Color 5						0.006 (0.004)
프레임 수		0.305*** (0.032)	0.204*** (0.034)	0.115*** (0.041)	0.041 (0.059)	0.115*** (0.041)	-2.770 *** (0.417)	
넷글 감성		1.983** (0.939)	2.161** (0.923)	1.312 (0.905)	0.933 (0.919)	1.312 (0.905)	0.726 (1.197)	
객체 추출	사람의 수	0.060** (0.025)	0.053** (0.025)	0.051* (0.028)	0.051* (0.029)	0.051* (0.028)	0.028 (0.042)	
	동물의 수	0.232 (0.190)	-0.199 (0.173)	0.258 (0.174)	-0.158 (0.211)	0.258 (0.174)	0.236 (0.245)	
	음식의 수	-0.169 (0.139)	0.016 (0.126)	0.131 (0.087)	0.005 (0.093)	0.131 (0.087)	0.025 (0.151)	
	기타 객체수	0.021 (0.024)	-0.004 (0.027)	-0.003 (0.023)	-0.051 (0.030)	-0.003 (0.023)	0.035 (0.045)	
R -squared		0.087	0.077	0.063	0.074	0.090	0.149	
F		16.164***	8.355***	2.369*	1.071	2.696***	6.878***	

p* < 0.1, p** < 0.05, p*** < 0.001

숏폼의 각 프레임에서 심미도와 색상을 분석한 결과, 초기 색상의 다채로움이 유의미하게 나타났으나 품질은 조회수와 유의한 관계를 보이지 않았다. 색상은 소비자의 시각을 사로잡는 중요한 요소로, 초기 몇 초 동안 흥미를 유발하여 조회수를 증가시킬 가능성이 크다. 짧은 영상의 특성상, 색상의 자극은 즉각적이고 단발적일 수 있으며, 이후 프레임에서의 인지 효과는 빠르게 줄어들어 것으로 해석된다. 반면, 품질은 해상도와 선명도와 같은 기술적 기준에 기반하므로 전체적인 인상과 만족도에는 영향을 미칠 수 있으나, 시청자의 즉각적인 반응보다는 전체 숏폼에 미치는 영향이 크다. 따라서 품질은 조회수에 강한 영향을 미치지 못하는 것으로 분석된다. 한편, 좋아요 수, 댓글 수 및 댓글 감성에 대한 프레임별 품질 및 색상에 대한 연구도 진행하였으며, 이에 대한 결과는 심미도 분석과 질적으로는 차이가 나타나지 않아, 생략하였다. 저자에게 요청 시 제공받을 수 있다.

4.4. 통계적 결과에 대한 논의

본 연구는 급부상하는 숏폼 콘텐츠의 이미지적 요소가 소비자의 참여행동과 상호작용에 미치는 영향을 조사하고, 프레임별 이미지로 구분하여 주요 영향을 분석하였다. 심미도, 품질, 색상, 객체 등 소비자의 인지에 영향을 미칠 수 있는 미적 요인을 컴퓨터 비전 알고리즘으로 도출하였으며, 이러한 요인들이 소비자의 참여행동에 미치는 영향을 분석하였다. 숏폼 콘텐츠와 시각적 요소의 관계 및 영향에 대한 가설의 결과는 다음과 같다.

H1 : 분석 결과 프레임 수가 증가하면 소비자 참여행동에 긍정적인 영향을 미칠 수 있지만, 지나치게 많은 프레임은 부정적인 영향을

미치는 것으로 나타났다. 즉, 영상의 길이를 향한 두 가지 다른 선호도가 숏폼에서 적절한 수준으로 유지되어야 한다는 것으로 해석할 수 있다. 이는 기존 연구에서 나타난 스마트폰에서는 짧은 영상을 선호한다는 분석과 플랫폼 알고리즘은 길고 인기있는 영상을 유도한다는 연구 모두에 부분적으로 적용되는 것으로 해석된다. 따라서 해당 가설은 전부 채택한다.

H2 : 숏폼 영상의 심미도는 조회수와 좋아요 수에 대해 모든 프레임에서 긍정적인 영향을 미쳤으며, 특히 첫 번째 프레임부터 시간이 지남에 따라 영향력이 점차 감소하는 경향을 보였다. 이를 통해 숏폼의 각 프레임에서 심미도가 미치는 효과의 변화를 확인하였다. 댓글 수와 댓글 감성은 초기 프레임에서 긍정적인 영향을 미쳐 초기 시각적 인지가 참여행동에 중요한 요소임을 확인하였다. 이는 시각적 미학이 개인의 소셜 커머스 의도에 직접적인 영향을 미치며 소셜 커머스 참여 의도를 촉진할 수 있다는 결과(Yang et al., 2021)와 일치하며, 마이크로 비디오의 첫인상인 썸네일은 사용자 클릭 및 시청을 유도하는데 중요한 역할을 하며 사용자 관심과 결합될 때 강력한 영향을 미친다는 연구(Bo, 2022)와도 결과를 같이한다. 심미도가 모든 종속 변수에서 긍정적인 영향을 미친다는 것은 시각적 인지는 물론, 콘텐츠 자체의 매력과 긍정적인 반응, 상호작용까지 이끌 수 있는 중요한 요소임을 보여준다. 따라서 해당 가설은 전부 채택한다.

H3 : 품질과 색상의 다채로움 역시 숏폼에서의 소비자 참여행동에서 부분적으로 영향을

미친다는 결과를 확인할 수 있었다. Maros et al. (2019)의 연구에서 고품질의 이미지와 다채로운 색상이 소비자의 시각적 관심을 유도한다는 결과가 숏폼에서도 동일하게 적용되었다. 품질은 조회수와 댓글 수에 유의한 관계를 보이지 않은 반면, 초기 이미지에서 좋아요 수와 댓글 감성에 긍정적인 영향을 미쳤다. 고품질 이미지는 시청자에게 만족감을 주고 긍정 반응을 유도하는 것으로 해석된다. 색상은 조회수와 댓글 감성에 유의미한 영향을 미쳤으며, 초기 이미지의 색상이 가장 큰 영향력을 발휘했다. 이 제원(2023)의 연구와 같이, 짧은 영상에서 몰입을 확대하기 위해 시각적 요소, 특히 창의적인 색상과 이미지 구성이 중요하다는 결론을 확인할 수 있다. 따라서 해당 가설은 부분 채택한다.

H4 : 댓글의 감성은 심미도, 품질, 색상과 같은 시각적 요소에 긍정적인 영향을 보이는 것으로 확인되었다. 이는 숏폼의 시각적 요소가 소비자에게 긍정적인 영향을 미친다고 해석할 수 있으며, Nieto et al. (2022) 사진에 달린 댓글의 감성으로부터 유추한 감성 점수가 심미도 평가 점수와 정(+)의 상관성이 있다는 연구와도 동일한 결과를 보인다. 미학적 요소는 시각적 즐거움과 만족감을 주며, 이러한 긍정적 감정은 상호작용으로 이어져 긍정적인 댓글로 나타나는 것을 확인했다. 따라서 본 가설을 전부 채택한다.

H5: 숏폼에 작성된 댓글의 긍정적인 반응 역시 소비자 참여행동에 긍정적인 영향을 미치는 것을 확인하였다. 이는 Santini et al. (2020)의 연구와 유사하게, 소셜 미디어 플랫폼에

서 만족과 긍정적 감정을 통해 고객 참여를 높일 수 있다는 분석과 일치한다. 다만 긍정적인 댓글 감성이 실제 댓글 수에 영향을 미치는 것에는 유의미한 관계를 보이지 않았다. 따라서 본 가설을 부분 채택한다.

H6 : 숏폼 콘텐츠에서 사람의 존재와 등장 여부는 숏폼 소비자 참여행동에 부분적으로 긍정적인 영향을 미치는 것으로 나타났다. 초기 프레임에서 사람의 존재와 노출이 증가하는 경우, 조회수와 댓글 감성에 긍정적인 영향을 미치는 것을 확인하였다. 김성덕 외 (2012)의 연구에 따르면 광고 모델의 특성이 구매 의도 및 브랜드 태도에 영향을 미친다는 연구와 유사하며, 사람의 존재와 등장 빈도는 스토리텔링이나 소비자 참여를 유도하는 역할을 하고, 참여를 촉진하는 특성이 있다고 해석된다. 따라서 본 가설을 부분 채택한다.

5. 결론

5.1. 연구 결과 요약

본 연구는 숏폼 영상의 프레임별 시각적 요소가 소비자 참여행동에 미치는 영향을 분석하였다. 분석 결과 초기 프레임에서의 높은 심미도와 품질, 색상 다양성이 소비자 참여행동과 반응에 중요한 영향을 미쳤음을 확인할 수 있었다. 특히 조회수와 좋아요 수에는 시각적 요소가 더욱 강하게 작용했으며, 이는 시청자가 시청 초기에 시각적 매력을 판단하고 콘텐츠 소비 여부를 결정함을 시사한다. 반면, 댓글 수의 경우 콘텐츠의 주제나 참여 유도 요소와 같은 맥락에 더욱 영향을 받는 것으로 판단된다.

프레임 진행에 따른 효과를 비교했을 때, 초기 프레임의 영향력이 가장 강력하며 이후 점차 약화되는 패턴이 확인된다. 이는 짧은 시간 내 주의를 끌어야 하는 숏폼의 특성이 두드러지는 지점으로, 시간이 지날수록 메시지 전달 효과가 제한됨을 시사한다. 또한 사람의 등장 여부 역시 초기 시청자의 관심을 끌고 긍정적인 감성과 상호작용을 촉진하는 요인으로 나타나, 숏폼의 대인적 속성이 소비자 참여를 강화하는데 기여함을 확인하였다.

5.2. 연구 기여도

숏폼 콘텐츠의 생산 및 소비가 급증하며, 이는 소셜 네트워크 서비스에서 콘텐츠 가치를 지속적으로 높이고 있다. 본 연구에서 다룬 유튜브 쇼츠는 향후에도 지속적인 성장이 예상되며, 개인과 기업 모두 중요한 전략적 요소로 적극 활용하고 있다. 본 연구는 소셜 미디어 콘텐츠, 특히 숏폼 영상을 프레임 단위로 분석하여, 시각적 요소가 소비자 참여행동에 미치는 영향을 심도 있게 조사한 최초의 연구로, 디지털 콘텐츠 분석의 새로운 방법론을 제시한다.

보다 구체적으로, 본 연구의 학문적 기여는 다음과 같이 정리할 수 있다. 첫째, 본 연구는 S-O-R (자극 - 유기체 - 반응) 이론을 숏폼 맥락으로 확장하여, ‘자극’에 해당하는 시각적·구성적 요인을 프레임 단위에서 조작화하였다. 기존 연구가 숏폼을 영상 수준의 평균값 혹은 메타데이터 중심으로 다룬 데 비해, 본 연구는 초기 구간의 시각적 단서가 인지적 주의 포착과 정서적 각성에 미치는 영향을 시간적으로 분해하여 보여줌으로써, 단기주의(attentional myopia) 및 선택적 주의 이론과 S-O-R의 접점을 경험적으로 강화한다.

둘째, ‘자극’의 다차원성을 심미성, 색채 풍부성,

대상(인물/사물 등)의 존재 여부와 같은 이질적 구성요인으로 구분하여 동시 투입 상황을 검증하였다. 이는 동일한 노출 시간 안에서 서로 다른 단서가 경쟁·보완하는 과정을 이론 모형에 내재화함으로써, 시각적 설계 요인이 정서·인지 경로를 통해 ‘반응’(조화·호감·발화)을 유발한다는 매커니즘을 보다 세밀히 설명한다.

셋째, 본 연구는 ‘초기 프레임 효과’(primacy window)를 개념화하였다. 숏폼 소비는 첫 수 초 내 정보처리가 집중되는 경향이 있으며, 본 연구는 자극 요인의 영향력을 시간 구간별로 분해함으로써, 노출 즉시의 심미·색채·사회적 존재감 단서가 반응 형성에 비대칭적으로 기여한다는 점을 보여준다.

본 연구의 결과는 콘텐츠 제작자와 실무자에게 소비자 관심을 끌기 위한 최적화 전략의 이론적 방향성을 제공한다. 보다 구체적으로 본 연구의 실무적 기여 또는 함의를 정리하면 다음과 같다. 첫째, 숏폼 제작·편집 단계에서 ‘초기 1-3초’의 설계가 핵심이라는 운영지침을 제공한다. 심미적 완성도와 색채 풍부성, 인물의 즉시적 제시 등 고주목 단서를 첫 구간에 집중 배치함으로써 스크롤 저항과 시청유지의 임계값을 넘길 가능성이 커진다. 이는 에디팅 순서와 썸네일/인트로 시퀀스의 재설계를 요구한다.

둘째, 브랜드·제품 특성에 따른 단서 최적화를 시사한다. 사람, 사물/텍스트 오버레이, 배경/색채 등 단서를 콘텐츠 목적에 맞게 조합할 때 성과가 극대화됨을 의미한다. 셋째, 댓글 정서 관리가 성과의 선순환을 만든다. 시각적 자극이 유도한 긍정 정서가 댓글·좋아요로 표출되고, 이는 재추천·재노출을 견인할 수 있으므로, 제작 단계에서 긍정 정서를 촉발하는 장면 구성 및 커뮤니티 매니지먼트(고정 댓글, 질문형 콜투액션 등)의 동시

설계가 요구된다. 댓글 감성은 운영 KPI로 주기적 모니터링이 가능하다는 점에서 현업 적용성이 높다.

넷째, 소비자 참여행동에 영향을 미치는 이미지에 대한 컴퓨터 비전 알고리즘을 활용한 분석은 기술의 실용성을 증명한다. 이는 심미도, 품질, 색상, 객체 등을 수치화하여 이미지 및 썸네일 자동화 과정에 기여할 수 있으며, AI의 발전으로 영상 프레임의 자동 최적화하거나 강한 시각적 인상을 주는 이미지를 선별할 수 있게 된다. 이는 효율적인 콘텐츠 제작 및 배포, 자동화 전략에 방향성을 제시하여 미래 디지털 마케팅에 기여할 수 있다.

본 연구는 기존의 주관적 평가 방법론과 차별화되며, 다양한 콘텐츠의 시각적 요소 분석으로 확대될 수 있다는 점에서 의의가 크다. 광고, 마케팅, 소셜 미디어, 이커머스 등 다양한 분야에서 효과적인 콘텐츠 전략과 시청자 반응 예측에 시각적 요소의 중요성을 강조할 수 있는 근거를 제공한다. 이는 기업과 단체가 시각적 요소를 효율적으로 배치하고 미학적 요소를 활용하여 강력한 콘텐츠 제작 및 마케팅 전략을 수립하는 데 기여할 것이다.

5.3. 연구 한계 및 향후 방향성

본 연구의 한계점은 다음과 같다. 첫째, 연구 표본이 특정 플랫폼과 국내 샷폼에 집중되어 있어, 다른 플랫폼이나 다국적 소비자를 대상으로 한 샷폼과의 차이를 파악하는 데 제한이 있다. 이로 인해 연구의 범용성과 확장성이 저하된다.

둘째, 데이터 표본 및 프레임 수에 제한이 있다. 본 연구는 최대 1분의 쇼츠를 300프레임으로 나누어 분석했지만, 더 많은 또는 더 적은 프레임 수에 대한 추가 연구가 필요하다. 전체 표본이 3,009개로, 프레임이 길어질수록 표본 수가 줄어드는 한계도 있다. 더 많은 표본을 확보하면

연구의 보편성이 향상될 수 있을 것이다. 셋째, 소비자 참여행동에 영향을 미치는 시각적 요소를 컴퓨터 비전 기술로 분석했으나, 샷폼에는 이미지적 요소 외에도 동적 요소, 콘텐츠적 특징, 유명인 등장 등 다양한 다른 요인의 영향이 존재한다는 점이다. 또한, 로그아웃 기반 자동 추천 샘플링이 완전히 무작위성을 보장하지 못한다는 이러한 요인들의 복합적인 영향을 분석하거나 통제하기 위한 추가적인 연구가 필요하다.

넷째, 본 연구는 S-O-R(Stimulus - Organism - Response) 모델을 연구 틀로 사용하지만, 종속변수는 소비자의 행동 반응만 사용하였고, Organism에 해당하는 심리적 반응(정서, 인지 부하, 몰입 등)에 관한 매개 효과에 대한 분석은 행하지 않았다. 또한, 감정 분석에 있어서는 자연어 처리 기술을 사용했지만, 감성 사전에 기반한 긍정과 부정 점수화로 인해 다양한 감정 상태를 충분히 반영하지 못했다. 또한, 소셜 미디어 사용자들의 반응이나 문맥, 뉘앙스를 완전히 반영하지 못할 수도 있다.

마지막으로, 컴퓨터 비전 기술에 의존했기 때문에 알고리즘의 정확성과 성능에 따라 결과에 차이가 생길 수 있으며, 이는 분석 결과의 신뢰성에 영향을 미칠 수 있다. 그러나 최근의 기술 발전을 고려할 때, 향후에는 시각적 요소에 대한 보다 높은 정확도와 다양한 변수를 포함한 연구가 가능할 것으로 기대된다.

참고문헌(References)

[국내 문헌]

강정빈, & 이상원. (2021). 샷폼 콘텐츠의 스토리텔링에 관한 연구-코로나 19 관련 샷폼 애니메이션 '라바(Larva)'시리즈를 중심으로. *기초*

- 조형학연구, 22(5), 13-28. <https://doi.org/10.47294/ksbda.22.5.2>
- 김성덕, 윤명길, & 김기수 (2012). 광고모델 특성이 구매의도와 브랜드태도에 미치는 영향. *유통과학연구*, 10(4), 47-53. <https://doi.org/10.15722/jds.10.4.201204.47>
- 김예은, 이상용. (2025). 게임의 첫인상이 인기도에 미치는 영향: 인디 게임의 이미지적 특성을 중심으로. *지능정보연구*, 31(1), 1 - 22. <https://doi.org/10.13088/jiis.2025.31.1.001>
- 박도형, 김한길, 손채형, 유하영, 이다연. (2025). E-Commerce 추천 시스템이 소비자의 상품 선택 행동에 미치는 차별적 영향 분석: 아이 트래커를 통한 시각적 반응 분석. *지능정보연구*, 31(1), 93 - 114. <https://doi.org/10.13088/jiis.2025.31.1.093>
- 박병언, 임규건. (2015). 일반영향요인과 댓글기반 콘텐츠 네트워크 분석을 통합한 유튜브(YouTube)상의 콘텐츠 확산 영향요인 연구. *지능정보연구*, 21(3), 41 - 58. <https://doi.org/10.13088/jiis.2015.21.3.19>
- 박소진 (2022). 유튜브 이용자의 반응행동의도 (좋아요, 구독, 공유, 댓글)와 전문성이 유튜브 추천제품 구매의도에 미치는 영향. *지역산업연구*, 45(3), 31-53. <https://doi.org/10.33932/rir.45.3.2>
- 이제원 (2023). 짧은 동영상의 사회적 시청동기가 지속시청의도 향상에 미치는 영향 연구. *서비스마케팅저널*, 16(2), 21-35. <https://doi.org/10.56352/smj.2023.16.2.02>
- 이진 (2020). 슷폼 동영상 콘텐츠의 유형 연구. *인문콘텐츠*,(58), 121-139. <https://doi.org/10.18658/humancon.2020.09.121>
- A. A., & Kulbo, N. B. (2022). Digital analytics approach to understanding short video advertising in digital marketing. *Journal of Marketing Theory and Practice*, 30(3), 405-420. <https://doi.org/10.1080/10696679.2022.2056487>
- Bakhshi, S., Shamma, D. A., & Gilbert, E. (2014, April). Faces engage us: Photos with faces attract more likes and comments on instagram. In *Proceedings of the SIGCHI conference on human factors in computing systems* (pp. 965-974). <https://doi.org/10.1145/2556288.2557403>
- Buscher, G., Cutrell, E., & Morris, M. R. (2009, April). What do you see when you're surfing? Using eye tracking to predict salient regions of web pages. In *Proceedings of the SIGCHI conference on human factors in computing systems* (pp. 21-30). <https://doi.org/10.1145/1518701.1518705>
- Bo, L. (2022). Towards Micro-video Thumbnail Selection via a Multi-label Visual-semantic Embedding Model. <https://doi.org/10.48550/arXiv.2202.02930>
- Covington, P., Adams, J., & Sargin, E. (2016). Deep neural networks for YouTube recommendations. In *10th ACM Conference on Recommender Systems* (Boston, MA). <https://doi.org/10.1145/2959100.2959190>
- Cunningham, S. J., & Nichols, D. M. (2008, June). How people find videos. In *Proceedings of the 8th ACM/IEEE-CS joint conference on Digital libraries* (pp. 201-210). <https://doi.org/10.1145/1378889.1378924>
- Datta, R., Joshi, D., Li, J., & Wang, J. Z. (2006). Studying aesthetics in photographic images using a computational approach. In *Computer Vision - ECCV 2006: 9th European Conference on Computer Vision, Graz, Austria, May 7-13, 2006, Proceedings, Part III 9* (pp. 288-301).

[국외 문헌]

Addo, P. C., Akpatsa, S. K., Nukpe, P., Ohemeng,

- Springer Berlin Heidelberg. https://doi.org/10.1007/11744078_23
- García-Avilés, A., & de Lara, A. (2018). An investigation of science online video: Designing a classification of formats. In B. León & M. Bourk (Eds.), *Communicating Science and Technology Through Online Video*. Routledge Focus. <https://doi.org/10.4324/9781351054584-2>
- Giombi, K., Viator, C., Hoover, J., Tzeng, J., Sullivan, H. W., O'Donoghue, A. C., ... & Kahwati, L. C. (2022). The impact of interactive advertising on consumer engagement, recall, and understanding: A scoping systematic review for informing regulatory science. *Plos One*, 17(2), e0263339. <https://doi.org/10.1371/journal.pone.0263339>
- Guo, J., Wang, X., & Wu, Y. (2020). Positive emotion bias: Role of emotional content from online customer reviews in purchase decisions. *Journal of Retailing and Consumer Services*, 52, 101891. <https://doi.org/10.1016/j.jretconser.2019.101891>
- Guo, Y., Ban, C., Yang, J., Goh, K. Y., Liu, X., Peng, X., & Li, X. (2024). Analyzing and predicting consumer response to short videos in e-commerce. *ACM Transactions on Management Information Systems*, 15(4), <https://doi.org/10.1145/3690393>
- Hasler, D., & Suesstrunk, S. E. (2003, June). Measuring colorfulness in natural images. In *Human Vision and Electronic Imaging VIII* (Vol. 5007, pp. 87-95). SPIE. <https://doi.org/10.1117/12.477378>
- Hoyer, W. D., & Stokburger-Sauer, N. E. (2012). The role of aesthetic taste in consumer behavior. *Journal of the Academy of Marketing Science*, 40, 167-180. <https://doi.org/10.1007/s11747-011-0269-y>
- Hou, J. R., Zhang, J., & Zhang, K. (2023). Pictures that are worth a thousand donations: How emotions in project images drive the success of online charity fundraising campaigns? An image design perspective. *MIS Quarterly*, 47(2), 535-584. <https://doi.org/10.25300/misq/2022/17164>
- Jordan, P. W. (1998). Human factors for pleasure in product use. *Applied Ergonomics*, 29(1), 25-33. [https://doi.org/10.1016/s0003-6870\(97\)00022-7](https://doi.org/10.1016/s0003-6870(97)00022-7)
- Kuss, J. D., & Griffiths, D. M. (2017). Social networking sites and addiction: Ten lessons learned. *International Journal of Environmental Research and Public Health*, 14(3), 311. <https://doi.org/10.3390/ijerph14030311>
- Ling, C., AbuHilal, I., Blackburn, J., De Cristofaro, E., Zannettou, S., & Stringhini, G. (2021). Dissecting the meme magic: Understanding indicators of virality in image memes. *Proceedings of the ACM on human-computer interaction*, 5(CSCW1), 1-24. <https://doi.org/10.1145/3449155>
- Maros, A., Belém, F., Silva, R., Canuto, S., Almeida, J. M., & Gonçalves, M. A. (2019). Image aesthetics and its effects on product clicks in E-commerce search. In Proceedings of the SIGIR 2019 eCom workshop. CEUR-WS. <http://ceur-ws.org> (Vol. 2410).
- Mehrabian, A., & Russell, J. A. (1974). The basic emotional impact of environments. *Perceptual and Motor Skills*, 38(1), 283-301. <https://doi.org/10.2466/pms.1974.38.1.283>
- Messai, O., Hachouf, F., & Seghir, Z. A. (2020). Adaboost neural network and cyclopean view for no-reference stereoscopic image quality assessment. *Signal Processing: Image Communication*, 82, 115772. <https://doi.org/10.1016/j.image.2019.115772>

- Muhoho-Minni, P., & Lubbe, B. A. (2017). The role of the media in constructing a destination image: The Kenya experience. *Communication*, 43, 58 - 79. <https://doi.org/10.1080/02500167.2016.1226915>
- Murray, N., Marchesotti, L., & Perronnin, F. (2012, June). AVA: A large-scale database for aesthetic visual analysis. In *2012 IEEE Conference on Computer Vision and Pattern Recognition* (pp. 2408-2415). IEEE. <https://doi.org/10.1109/cvpr.2012.6247954>
- Nieto, V.D., Celona, L., & Fernandez Labrador, C. (2022). Understanding aesthetics with language: A photo critique dataset for aesthetic assessment. *Advances in Neural Information Processing Systems*, 35, 34148-34161.
- O'Brien, H. L., & Toms, E. G. (2008). What is user engagement? A conceptual framework for defining user engagement with technology. *Journal of the American Society for Information Science and Technology*, 59(6), 938-955. <https://doi.org/10.1002/asi.20801>
- Park, C. S., & Kaye, B. K. (2021). What's this? Incidental exposure to news on social media, news-finds-me perception, news efficacy, and news consumption. In *Social Media News and Its Impact* (pp. 98-121). Routledge. <https://doi.org/10.4324/9781003179580-6>
- Park, S. M., Na, C. W., Choi, M. S., Lee, D. H., & On, B. W. (2018). KNU Korean sentiment lexicon: Bi-LSTM-based method for building a Korean sentiment lexicon. *Journal of Intelligence and Information Systems*, 24(4), 219-240.
- Rajendran, P. T., Creusy, K., & Garnes, V. (2024). Shorts on the rise: Assessing the effects of YouTube Shorts on long-form video content. <https://doi.org/10.48550/arXiv.2402.18208>
- Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You only look once: Unified, real-time object detection. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 779-788). <https://doi.org/10.1109/cvpr.2016.91>
- Richard, M. O., & Chandra, R. (2005). A model of consumer web navigational behavior: Conceptual development and application. *Journal of Business Research*, 58, 1019 - 1029. <https://doi.org/10.1016/j.jbusres.2004.04.001>
- Santini, F.d.O., Ladeira, W. J., Pinto, D. C., Herter, M. M., Sampaio, C. H., & Babin, B. J. (2020). Customer engagement in social media: a framework and meta-analysis. *Journal of the Academy of Marketing Science*, 48, 1211-1228. <https://doi.org/10.1007/s11747-020-00731-5>
- Sheng, H., Yang, P., & Feng, Y. (2020). How to inspire customers via social media. *Industrial Management & Data Systems*, 120(6), 1041 - 1057. <https://doi.org/10.1108/imds-10-2019-0548>
- Shi, R., Wang, M., Liu, C., & Gull, N. (2023). The influence of short video platform characteristics on users' willingness to share marketing information: Based on the SOR model. *Sustainability*, 15(3), 2448. <https://doi.org/10.3390/su15032448>
- Shin, D., He, S., Lee, G. M., Whinston, A. B., Cetintas, S., & Lee, K. C. (2020). Enhancing social media analysis with visual data analytics: A deep learning approach. *Mis Quarterly*, 44(4), 1459-1492. <https://doi.org/10.25300/misq/2020/14870>
- Singh, S. (2006). Impact of color on marketing. *Management Decision*, 44(6), 783-789. <https://doi.org/10.1108/00251740610673332>
- Smith, A., Toor, S., & Van Kessel, P. (2018).

- Many turn to YouTube for children's content, news, how-to lessons. *Pew Research Center*, 7, 1. Retrieved September 25, 2025, from <http://www.pewinternet.org/2018/11/07/many-turn-to-youtube-for-childrens-content-news-how-to-lessons/>
- Talebi, H., & Milanfar, P. (2018). NIMA: Neural image assessment. *IEEE Transactions on Image Processing*, 27(8), 3998-4011. <https://doi.org/10.1109/tip.2018.2831899>
- Tian, K., Xuan, W., Hao, L., Wei, W., Li, D., & Zhu, L. (2022). Exploring youth consumer behavior in the context of mobile short video advertising using an extended stimulus - organization - response model. *Frontiers in Psychology*, 13, 933542. <https://doi.org/10.3389/fpsyg.2022.933542>
- Vinerean, S., Cetina, I., Dumitrescu, L., & Tichindelean, M. (2013). The effects of social media marketing on online consumer behavior. *International journal of business and management*, 8(14), 66. <http://dx.doi.org/10.5539/ijbm.v8n14p66>
- Wei, W., Li, N., & Chen, Y. (2024, June). The impact of short-video application affordances on cross-cultural user engagement behavior intention: Based on SOR model. In *International Conference on Human-Computer Interaction* (pp. 129-146). Cham: Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-60898-8_9
- Xiao, L., Li, X., & Zhang, Y. (2023). Exploring the factors influencing consumer engagement behavior regarding short-form video advertising: A big data perspective. *Journal of Retailing and Consumer Services*, 70, 103170. <https://doi.org/10.1016/j.jretconser.2022.103170>
- Xie, J., Xu, X., Zhang, Y., Tan, Y., Wu, D., Shi, M., & Huang, H. (2023). The effect of short-form video addiction on undergraduates' academic procrastination: a moderated mediation model. *Frontiers in Psychology*, 14, 1298361. <https://doi.org/10.3389/fpsyg.2023.1298361>
- Yan, X., Gilani, S. Z., Feng, M., Zhang, L., Qin, H., & Mian, A. (2020). Self-supervised learning to detect key frames in videos. *Sensors*, 20(23), 6941. <https://doi.org/10.3390/s20236941>
- Yang, L., Yuan, X., & Yang, X. (2024). Study of the influencing mechanism of user interaction behavior of short video e-commerce live-streaming from the perspective of SOR theory and interactive ritual chains. *Current Psychology*, 43(35), 28403-28415. <https://doi.org/10.1007/s12144-024-06494-4>
- Yang, Y., Tang, Y., Zhang, Y., & Yang, R. (2021). Exploring the relationship between visual aesthetics and social commerce through visual information adoption unimodel. *Frontiers in Psychology*, 12, 700180. <https://doi.org/10.3389/fpsyg.2021.700180>

Abstract

What Drives Consumer Engagement Behaviors in Short-Form Videos? An Image-Analytic Approach Using Computer Vision Algorithms

Yubin Bang* · Sang-Yong Tom Lee**

Recent growth in social networking services (SNS) and mobile-centric media consumption has led to an explosive increase in user-generated content data. Short-form video, in particular, delivers vivid visual experiences within a brief time span and has emerged as a media format well aligned with contemporary patterns of information consumption. As short-form content proliferates across platforms, its role has expanded beyond individual expression to encompass firm value creation and marketing strategy. Because short-form videos are typically distributed via recommendation algorithms, viewers must decide whether to engage with content under severe time constraints, making visual cues more consequential than in many other media. Accordingly, this study decomposes short-form videos into frame-level images and examines how visual attributes shape consumer engagement behaviors. Focusing on YouTube Shorts, we apply computer-vision - based image analytics to convert videos into frame sequences and extract visual features – including aesthetics, quality, color characteristics, and detected objects – to test their effects on engagement outcomes and to identify when these features exert the strongest influence over the course of a video. This research is among the first to analyze short-form social media content at the frame level, demonstrating both the importance of frame-specific visual elements and the temporal loci at which their effects emerge. The findings provide a theoretical foundation and future directions for content automation grounded in computational aesthetics, as well as optimization strategies to predict and enhance consumer responses in algorithmically curated media environments.

Key Words : Short-form video, aesthetic appeal, social media, image analysis, computer vision

Received : October 13, 2025 Revised : December 17, 2025 Accepted : December 22, 2025

Corresponding Author : Sang-Yong Tom Lee

* School of Business, Hanyang University

** Corresponding Author: Sang-Yong Tom Lee

School of Business, Hanyang University

222 Wangshimri-ro, Seongdong-gu, Seoul 04763, Korea

Tel: +82-2-2220-1814, Fax: +82-2-2220-1669, E-mail: tomlee@hanyang.ac.kr

저 자 소개



방유빈

한양대학교 일반대학원 비즈니스인포매틱스학에서 석사 학위를 취득하였다. 주요 관심 분야는 Big Data Analytics, Business Intelligence, Vision AI, Machine Learning/Deep Learning, AI-Driven Business Transformation 등이다.



이상용

현재 한양대학교 경영대학 교수로 재직 중이다. 서울대학교 경제학과를 졸업하고, Texas A&M University에서 박사학위를 취득하였다. 주요 관심 분야는 정보경제, 개인정보보호 및 보안, 소셜미디어, 정보통신정책, 기술경영 등이다. 관련 연구들을 MIS Quarterly, Management Science, Journal of Management Information Systems를 비롯한 다수의 저널에 관련 논문을 게재하고 있다.