

미디어 초개인화 추천을 위한 YCrCb 컬러 모델 분석을 통한 영상의 메타데이터 추출에 대한 연구

박효경* · 용성중 · 유연휘 · 문일영
한국기술교육대학교

A Research on Image Metadata Extraction through YCrCb Color Model Analysis for Media Hyper-personalization Recommendation

Hyo-Gyeong Park* · Sung-Jung Yong · Yeon-Hwi You · Il-Young Moon
Korea University of Technology and Education

E-mail : sjshk@koreatech.ac.kr / zeros@koreatech.ac.kr / wlwhs3258@koreatech.ac.kr /
iymoon@koreatech.ac.kr

요 약

최근 높은 접근성을 기반으로 다양한 콘텐츠가 양산됨에 따라 미디어 콘텐츠 시장이 더욱 활성화되고 있다. 사용자들은 취향에 맞는 콘텐츠를 찾고자 하며, 각 플랫폼에서 콘텐츠의 개인화 추천을 위해 경쟁하고 있다. 효율적인 추천시스템을 위해서는 양질의 메타데이터가 필요하다. 기존의 플랫폼들은 영상의 메타데이터를 사용자가 직접 입력하는 방식을 취하고 있다. 이는 많은 양의 데이터를 처리하는데에 시간과 비용을 낭비하게 할 것이다. 본 논문에서는 미디어 초개인화 추천을 위해서 영화예고편을 바탕으로, 영상의 YCrCb 컬러 모델을 기반으로 키프레임을 추출하고, 인공지능의 지도학습을 통해 영화 장르를 구별하고 추후 메타데이터 생성을 위한 활용방안을 제안하고자 한다.

ABSTRACT

Recently as various contents are mass produced based on high accessibility, the media contents market is more active. Users want to find content that suits their taste, and each platform is competing for personalized recommendations for content. For an efficient recommendation system, high-quality metadata is required. Existing platforms take a method in which the user directly inputs the metadata of an image. This will waste time and money processing large amounts of data. In this paper, for media hyperpersonalization recommendation, keyframes are extracted based on the YCrCb color model of the video based on movie trailers, movie genres are distinguished through supervised learning of artificial intelligence and In the future, we would like to propose a utilization plan for generating metadata.

키워드

YCrCb, 영상데이터 추출, 메타데이터, AI

1. 서 론

최근 높은 접근성을 기반으로 한 다양한 콘텐츠가 양산됨에 따라 미디어 콘텐츠 시장이 더욱 활성화되고 있다. 콘텐츠는 곧 상품이며, 더 많은 사

용자를 모으기 위해 TV 방송 제작사나 기획사, OTT(Over-The-Top) 서비스별 오리지널 콘텐츠, SNS(Social Networking Service) 등을 중심으로 콘텐츠의 다양성은 끊임없이 확장되고 있다.

사용자들은 취향에 맞는 콘텐츠를 찾고자 하며, 각 플랫폼에서 콘텐츠의 개인화 추천을 위해 경쟁이 더욱 치열해지고 있다. 넷플릭스(Netflix)의 경

* corresponding author

우, 사용자의 관심사, 시청 기록들을 바탕으로 콘텐츠 기반 필터링과 협업 필터링 기술을 융합한 독자적인 알고리즘을 적용하고 있다[1]. 유튜브(Youtube) 또한 심층 신경망을 기반으로 하여 자체 추천 알고리즘을 이용하고 있다[2].

효율적인 추천시스템을 위해서는 양질의 메타데이터가 필요하다. 양질의 메타데이터를 기반으로 개인의 데이터와 매칭 시켜 초개인화 맞춤형 서비스를 진행할 수 있기 때문이다. 기존의 플랫폼들은 영상의 메타데이터를 사용자가 직접 입력하는 방식을 취하고 있다. 이는 많은 양의 데이터를 처리하는 데에 시간과 비용을 낭비하게 할 것이다.

따라서 본 논문에서는 영상 미디어의 메타데이터를 추출하기 위해서 영화의 장르를 구분하는 데에 영상 키프레임 요소를 활용 가능 여부를 확인한다. 이를 위해 영상을 YCrCb 컬러모델을 기반으로 키프레임을 추출하고, 영상 미디어의 키프레임을 통해 영화의 장르를 구분할 수 있는 과정에 관한 연구를 진행하였다.

II. 본론

본 논문에서는 영화의 영상 키프레임을 통해 메타데이터 생성에 있어 활용 가능 여부를 연구한다. 우선 영화 예고편에서 키프레임을 추출한 후, 영화 예고편의 키프레임을 통해 장르를 구분할 수 있는 요소로 활용할 수 있는지와 인공지능에 적용하기 위한 방안을 제시하고 결과를 확인하고자 한다.

또한, 영화의 용량 및 저작권 문제로 인한 영화 데이터 확보에 대한 한계점을 극복하고자 영화 예고편을 사용하여 예고편의 영상을 분석하고 인공지능에 적용하고자 한다.

2.1 키프레임 추출을 위한 YCrCb 컬러모델 분석

Python의 행렬이나 다차원 배열 처리를 위한 NumPy와 이미지 프로세싱을 위한 openCV, 데이터의 Peak 감지 및 검출을 위한 PeakUtils 라이브러리를 활용해 분석을 진행하였다.

YCrCb는 일종의 선형 색 공간으로, Y는 휘도 성분을 나타내고, Cr과 Cb는 적색과 청색의 농도 오프셋 성분을 나타낸다. RGB는 다음 방정식에 의해 YCrCb로 변환될 수 있다[3].

$$\begin{aligned} Y &= 0.299R + 0.587G + 0.114B & (1) \\ Cr &= 0.212R - 0.523G - 0.311B \\ Cb &= 0.596R - 0.272G - 0.321B \end{aligned}$$

영상의 키프레임 추출 과정은 [그림 1]과 같다. RGB 컬러모델을 YCrCb 컬러모델로 변환하여 Y, Cr, Cb 각 색 공간을 분리하여 변화 정도를 측정하였다. 변화 정도가 Peak 상태(높은 상태)일 경우,

장면 변환 등의 키프레임 이미지라고 판단하였고, 이때의 키프레임을 획득하였다.

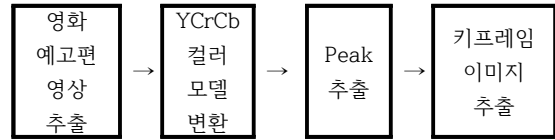


그림 1. 영상 키프레임 추출 과정

표 1. 영화 장르 구분 및 추출된 프레임 수

장르	영화예고편	추출된 키프레임 수
코믹 (Comic)	CA	118
	CB	110
	CC	67
	CD	63
액션 (Action)	AA	233
	AB	215
	AC	231
	AD	309
공포 (Horror)	HA	50
	HB	152
	HC	125
	HD	92
로맨스 (Romance)	RA	172
	RB	177
	RC	111
	RD	53

[표 1]과 같이 영화의 장르를 코믹, 액션, 공포, 로맨스로 나누었다. 장르별 4편의 대표 영화 예고편을 선정하고 영상과 오디오를 분리하였다. 그리고 분리된 영상에서 RGB 컬러 모델을 YCrCb 컬러모델로 변환하였다. 변환한 YCrCb 색 공간의 변화도를 통해 각 영화 예고편의 키프레임을 생성하여 이미지로 저장하였다. 그 후, 랜덤으로 각 영화 예고편별 50개의 키프레임을 추출하였다. 그 후, 딥러닝 학습데이터로 처리하기 위해 카테고리별 분류를 진행하였다.

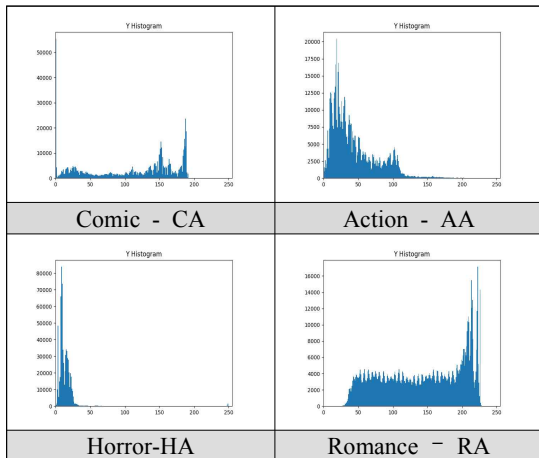


그림 2. 영화 장르별 히스토그램

[그림 2]는 YCrCb 색 공간을 활용해 추출한 Y 값의 히스토그램 이미지를 랜덤하게 뽑아 정리해 본 것이다. 영화별로 차이는 발생할 수 있으나, 장르별로 도출되는 히스토그램에 특징이 있다는 것을 확인할 수 있었다. 이를 통해 YCrCb를 통해 추출한 히스토그램 이미지를 인공지능에 적용한다면 장르별 특징을 기반으로 장르를 구별할 수 있을 것이라 예상되었다. 이를 인공지능에 적용하여 결과를 확인해 보고자 하였다.

2.2 CNN 기반의 학습 및 평가

[그림 3]과 같이 YCrCb 컬러 모델 기반 분석 결과를 바탕으로 히스토그램을 CNN(Convolutional Neural Networks)과 로지스틱 회귀모델에 적용하여 장르별 분류 결과를 확인하였다. 이때, 이미지 분류용 CNN 중 계층이 16개로 구성된 VGG-16을 적용하였고, 활성화 함수로는 ReLU를 사용하였다.

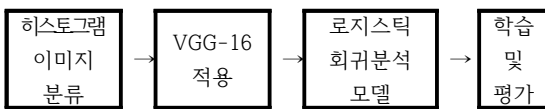


그림 3. 인공지능 적용 프로세스

III. 연구결과

[표 2]와 같이 Confusion Matrix 결과 코믹, 액션, 공포, 로맨스 이미지가 분류된 모습을 확인할 수 있었다. 공포와 로맨스의 경우 분류가 잘되지 않는 문제가 발생한 것을 확인하였다.

결과적으로 정확도는 88.9%와 정밀도는 69.5% 정도로 나타났다. 학습데이터의 양이 비교적 적기 때문에 이러한 결과가 나타난 것으로 보인다. 이에

따라, 추후 연구에서는 학습데이터에 대한 문제를 보완한다면 더 높은 정확도를 나타낼 수 있을 것이라 기대할 수 있다. 영화 예고편의 키프레임을 추출하여 영화의 장르를 구분한다면, 장르에 대한 메타데이터를 자동생성이 가능하다는 것을 기대해 볼 수 있을 것이다.

표 2. Confusion Matrix 결과

		Predicted					
		코믹	액션	공포	로맨스	Total	
Actual	A						
	c	코믹	559	11	66	44	680
	t	액션	50	437	126	57	670
	u	공포	47	92	448	100	687
	a	로맨스	43	72	128	440	683
l	Total	699	612	768	641	2,720	

IV. 결론

최근 높은 접근성을 기반으로 한 다양한 콘텐츠가 양산됨에 따라 미디어 콘텐츠 시장이 더욱 활성화되며, 소비자들의 욕구를 충족시키기 위해 각 플랫폼에서는 개인의 메타데이터와 개인 맞춤형 서비스에 대한 연구를 활발하게 진행하고 있다. 본 논문은 사람이 직접 메타데이터를 입력하던 방식에서 벗어나 인공지능을 통해 메타데이터를 자동으로 생성하는 방법에 대해 제안하였다. 첫 번째 요소로, YCrCb 컬러 모델 분석을 통해 영상의 키프레임을 추출하였다. 추출한 키프레임 이미지를 Y값을 기반으로 히스토그램 이미지를 생성하였다. 장르별로 Y값 변화의 차이가 있다는 것을 알게 되었고, 이를 CNN과 로지스틱 회귀모델에 적용함으로써 자동으로 메타데이터를 생성하는 것을 검토하였다. 향후 연구에서 학습데이터의 종류와 양을 확장시키고, Y값 뿐만 아니라 Cr값과 Cb값을 활용하여 키프레임 이미지의 특징을 추출하여 학습을 진행함으로써 인공지능의 정확도를 높이기 위해 보완하는 연구가 필요할 것으로 보인다. 추후 연구를 통해 메타데이터를 자동 생성하기 위한 시스템을 구축한다면 미디어 초개인화 추천 시스템을 더욱 발전시킬 것이다.

Acknowledgement

이 논문은 2021년도 정부(교육부)의 재원으로 한국연구재단의 지원을 받아 수행된 기초연구사업 (No. 2021R111A3057800) 과제 지원에 의하여 연구되었음.

References

- [1] C.A. Gomez-Urbe and N. Hunt. “The Netflix Recommender System: Algorithms, Business Value, and Innovation.”, *ACM Transactions on Management Information Systems*, Vol. 6, No. 13, pp.1-19, Jan. 2016.
- [2] P. Covington, J. Adams, and E. Sargin, “Deep Neural Networks for YouTube Recommendations.”, in *Proceedings of the 10th ACM Conference on Recommender Systems (RecSys '16)*, New York, NY, USA, pp. 191 - 198, Sep. 2016.
- [3] Y. Tan, J. Qin, X. Xiang, W. Ma, W. Pan, and N. Xiong, “A Robust Watermarking Scheme in YCbCr Color Space Based on Channel Coding”, *IEEE Access*, Vol. 7, pp. 1-1, Jan. 2019.