영상 추천 서비스의 개선을 위한 영상 미디어의 메타데이터 자동생성 방법에 대한 연구

유연휘* · 박효경 · 용성중 · 문일영

한국기술교육대학교

A Research on the Method of Automatic Metadata Generation of Video Media for Improvement of Video Recommendation Service

Yeon-Hwi You* · Hyo-Gyeong Park · Sung-Jung Yong · Il-Young Moon

Korea University of Technology and Education

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

요 약

국내 OTT(Over-the-top media service) 시장의 추천 서비스에서 거론되는 대표 업체는 Youtube와 Netflix이다. Youtube는 여러 방법을 거쳐 2016년부터 이용자의 시청 시간을 기록하여 이용하는 알고리즘을 머신러닝에 도입하면서 개인화된 추천을 본격화하였고, Netflix는 사용자의 선택한 영상, 시청 시간대, 영상 시청 기기 등 정보 수집을 통해 이용자를 분류하고 비슷한 시청 패턴을 가진 사람들을 같은 그룹에 묶는 방식과 영상을 직접 시청 후 사람이 태그(메타데이터)를 직접 기록하여 사용자로부터 수집한 정보와 영상에 붙은 태그 정보를 이용한다. 본 논문에서는 수기로 작성하던 영상 미디어의 메타데이터를 자동으로 생성하여 영상미디어의 추천을 개선하기 위한 방법을 제안하고자 한다.

ABSTRACT

The representative companies mentioned in the recommendation service in the domestic OTT(Over-the-top media service) market are YouTube and Netflix. YouTube, through various methods, started personalized recommendations in earnest by introducing an algorithm to machine learning that records and uses users' viewing time from 2016. Netflix categorizes users by collecting information such as the user's selected video, viewing time zone, and video viewing device, and groups people with similar viewing patterns into the same group. It records and uses the information collected from the user and the tag information attached to the video. In this paper, we propose a method to improve video media recommendation by automatically generating metadata of video media that was written by hand.

키워드

추천서비스, 영상미디어, OTT, 메타데이터, AI

Ⅰ. 서 론

OTT 서비스는 높은 접근성과 콘텐츠의 다양성이라는 장점으로 크게 성장하고 있다. 글로벌 OTT 시장은 2012년도부터 지금까지 연평균 21.6%의 성장률을 보이며, 국내 OTT 시장 역시 2014년도부터 2020년도까지 연평균 26.3%의 가파른 성장세를 보였다. OTT 서비스를 제공하는 업체들은 서비스 이용자 개개인에게 맞춤형 영상을 추천해주는 자체적인 알고리즘을 가지고 있다. 추천알고리즘은 메타데이터 기반으로 이루어진다. 사용자가 어떠한 영상을 보고, 언제 영상을 종료했는지는 자동으로 기록하지만, 영상 정보에 관한 메타데이터는 현재 사람이 영상을 보고 직접 입력하는 방식을 사용하고 있다.

* corresponding author

본 논문에서는 수기로 작성하던 영상 미디어의 메타데이터를 인공지능을 통해 자동으로 생성하는 방안을 제안하여 미디어 추천 서비스를 개선하고자 한다.

II. OTT기반 플랫폼 및 추천알고리즘 동향

2.1 Netflix

Netflix 플랫폼의 경우 OTT 서비스 채널에서 자체 제작하여 해당 플랫폼에서만 볼 수 있는 오리지널 시리즈와 국내외의 방대한 영화 및 드라마 VOD 위주의 서비스를 제공하고 있다.

Netflix의 추천알고리즘은 사용자의 관심사, 시청기록들을 바탕으로 사용자가 어떤 상품을 구매했을 때, 그 상품과 비슷한 상품을 사용자에게 추천해주는 방식인 콘텐츠 기반 필터링과 사용자A와 비슷한 특성을 가진 사용자B에게 사용자A가 구매한 물품을 추천해주는 방식인 협업 필터링 기술을 융합한 독자적인 알고리즘을 적용하여 개인 맞춤형 추천 시스템을 제공한다. 이에 더해 사람이 영상을 보고 직접 태그를 기록해 사용자로부터 수집한 정보와 영상에 붙은 태그 정보를 활용하여 추천 시스템이 작동한다[1].

2.2 Youtube

Youtube 플랫폼은 개인 채널을 통한 라이브 방송, 사용자 자체 제작 영상, 쇼트 비디오 등 플랫폼 이용자가 자체적으로 생산한 콘텐츠를 공유하는 형태의 서비스를 제공하고 있다.

최근 Youtube 영상추천 알고리즘 또한 Netflix와 마찬가지로 콘텐츠 기반 필터링과 협업 필터링을 기반으로 한 독자적인 추천알고리즘을 이용한다. Youtube 초기에는 조회 수 기반의 추천알고리즘을 사용했다. 하지만 자극적인 제목과 썸네일로 제작된 영상으로 인해 저품질의 영상이 추천되는 경우가 잦았고, 2012년 Youtube는 추천알고리즘을 영상을 클릭해 끝까지 시청하는 비율 기반의 추천알고리즘으로 변경했다. 지금과 같은 콘텐츠 기반 필터링과 협업 필터링을 기반으로 한 추천알고리즘은 2016년부터 구축되기 시작했다[2].

III. 요구분석

OTT 플랫폼의 메타데이터 생성은 영상 제작자 또는 배포자가 영상을 직접 본 후 태그와 같은 방식을 이용해 메타데이터를 직접 입력하는 방식을 사용하고 있다. 이러한 방법은 사람이 영상을 모두 봐야한다는 시간적 제약과 메타데이터를 직접 입력해야 하는 번거로움이 있고, 영상 제작자가 잘못된 메타데이터를 입력할 시에 사용자와 관련 없는 영상이 추천될 수 있다는 문제점을 가지고 있다.

따라서 메타데이터를 생성하는 과정을 자동화하여, 사용자가 데이터를 직접 입력하는 번거로움을 없앨 수 있다, 또한 영상 내부에서 더욱 다양한 데이터 정보를 추출할 수 있어 데이터의 다양화를 통해 더욱 정밀한 추천 서비스를 기대할 수 있다.

IV. 시스템 제안

심층 신경망 학습 모델 중 합성곱신경망(Convolution neural network)을 이용한 학습 방법은 이미지 분리에서 좋은 결과를 보여주며 대표적으로 딥러닝 분야에 사용되고 있다[3].

심층 합성곱 신경망 모델 중 VGG 모델은 옥스퍼드 대학의 연구팀 VGG(Visual Geometry Group)에 의해 개발된 모델로써, 2014년 ImageNet Challenge(ILSVRC)에서 준우승을 하였고, 이미지 분류를 위한 딥 러닝 관련 대표적 논문 중 하나로 자리매김하였다. 구성된 신경망 계층이 16개 인지 19개 인지에 따라 VGG16, VGG19로 불린다[4].

본 논문에서는 CNN 모델 중 VGG 심층 합성곱 신경망 모델을 사용하여 YCbCr의 키프레임과 오디오 스펙트럼 이미지를 추출하여 딥러닝을 구현하려고 한다. 그 후 지도학습을 통해 영상의 메타데이터를 자동생성하고 개인 정보와 매칭을 통해 초개인화 맞춤형 콘텐츠를 제공하고자 한다.

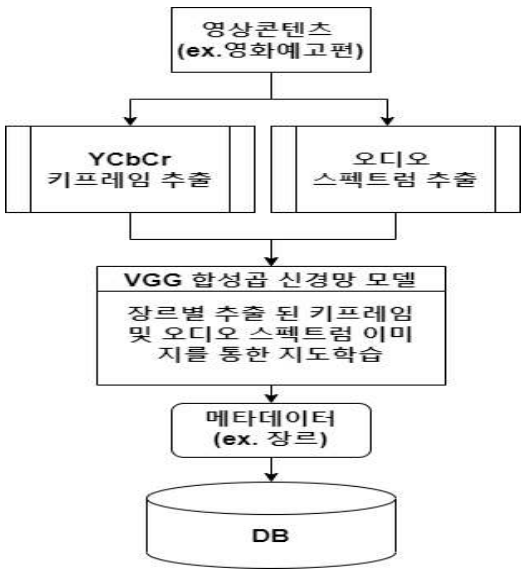

그림 1. 제안 시스템 구성도

그림 1과 같이 초개인화 맞춤형 시스템은

(1) YCbCr을 통해 영상 콘텐츠의 키프레임과 오디오 분석을 통한 스펙트럼 이미지를 추출한다.
- 키프레임 및 오디오는 영상의 화면, 음악, 효과 등 다양한 특징들을 요소로 활용된다.

(2) 영화예고편의 YCbCr 키프레임 및 오디오 스펙트럼 이미지를 장르별 카테고리로 분류하여 사전 학습데이터를 확보한다. 합성곱 신경망 VGG 모델을 통해 사전 확보한 학습 데이터로 지도학습을 진행한다.

(3) 입력받은 영상 데이터의 키프레임 및 오디오 스펙트럼 데이터를 학습된 인공신경망에 전달하여, 메타데이터를 자동 생성하여 DB에 저장한다.

(4) 개인의 콘텐츠 시청 선호도 및 데이터를 바탕으로 자동 생성된 영상의 메타데이터와 매칭하여 영상 콘텐츠를 추천한다.

본 논문에서는 AI를 통해 자동 생성된 메타데이터와 개인의 누적된 콘텐츠 성향 데이터를 통해 콘텐츠 추천에 대한 초개인화 시스템 구현을 목표로 한다.

Ⅴ. 결 론

OTT 서비스는 높은 접근성과 콘텐츠의 다양성이라는 장점으로 크게 성장하고 있다. OTT 서비스를 제공하는 업체들은 서비스 이용자 개개인에게 맞춤형 영상을 추천해주는 자체적인 알고리즘을 가지고 있다. 추천알고리즘은 메타데이터 기반으로 이루어진다. 사용자가 어떠한 영상을 보고, 언제 영상을 종료했는지는 자동으로 기록하지만, 영상 정보에 관한 메타데이터는 현재 사람이 영상을 보고 직접 입력하는 방식을 사용하고 있다. 본 논문에서는 사람이 영상을 보고 직접 입력하는 메타데이터의 수동방식을 AI와 영상의 키프레임 및 오디오를 통해 자동으로 구현하는 시스템을 제안하고, 추후 구현을 통해 초개인화 OTT 시스템을 제작하고자 한다. 앞으로 영상의 정확한 메타데이터 생성에 필요한 요소에 대한 연구와 실시간 영상처리에 관한 추가적인 연구를 진행하여 기존 추천 모델보다 더욱 개인화된 시스템을 구현하고자 한다.

Acknowledgement

이 논문은 2021년도 정부(교육부)의 재원으로 한국연구재단의 지원을 받아 수행된 기초연구사업 (No.2021R1I1A3057800) 과제 지원에 의하여 연구되었음.

References

[1] C.A. Gomez-Uribe and N. Hunt. “The Netflix Recommender System: Algorithms, Business Value, and Innovation.” *ACM Transations 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] J.D. Kim, S.H. Lee, H.S. Ko, “CNN Structure optimize for Facial Recognition”, *The Institute of Electronics and Information Engineers*, pp. 1172-1175, Jun. 2019.

[4] K.S. Yoon, J.Y. Choi, “Compressed Ensemble of Deep Convolutional Neural Networks with Global and Local Facial Features for Improved Face Recognition”, *Korea Multimedia Society*, Vol. 23, No. 8, pp. 1019-1029, Aug. 2020.