

사용자 검색 효율 향상을 위한 빅데이터 처리 기반 영상 분석 시스템 설계

이현섭 · 김준호 · 차주형 · 나보아 · 김진덕
동의대학교

Design of Image Analysis System Based on Big Data Processing for Increasing Search Efficiency

Hyoun-Sup Lee · Jun-Ho Kim · Joo-Hung Cha · Bo-Ah Na · Jin-Deog Kim

Dong-Eui University
E-mail : jdk@deu.ac.kr

요 약

콘텐츠 제공 시스템의 사용자들은 영상의 제목 및 해시 태그, 영상 설명 등의 내용을 활용하여 필요한 영상을 검색한다. 영상 길이가 매우 길거나 여러 가지 이슈가 포함된 경우에는 원하는 영상을 찾기 위한 부가적인 비용이 발생한다. 본 논문에서는 이러한 문제를 해결하기 위해 영상 내에서 많이 발생한 정보를 기반으로 랭킹 시스템을 구축하여 사용자에게 정확한 영상 정보를 제공하는 시스템에 대하여 제안한다.

제안된 논문은 키워드 분석과 영상의 이미지 분석을 통해 사용자의 검색 결과의 정확도를 높여 효율적인 영상 검색을 수행할 수 있도록 지원한다.

키워드

Image Processing, Frequency analysis, Service Platform, BigData, Ranking System

1. 서 론

최근 유튜브, 웨이브, 넷플릭스 등 다양한 영상 제공 서비스가 사용자에게 제공되고 있다.

많은 사용자가 CP(Content Provider)가 제작한 영상이나 방송, 영화 영상들을 위와 같은 서비스를 통해 시청하고 있다. 이러한 영상 제공 서비스들은 일반적으로 영상의 제목 및 영상을 설명하는 여러 가지 키워드를 통해 사용자의 영상 검색을 지원한다. 영상 검색을 수행하는 과정에서 단순 키워드 검색의 결과는 별도의 프레임 인덱스를 제공하지 않고 영상의 도입부를 통해 검색 결과를 제공하게 된다. 만약 영상의 길이가 길고 찾고자 하는 내용이 전체 콘텐츠의 일부일 경우에는 이러한 검색 결과 이후 별도의 검색 비용을 들여 사용자는 영상을 시청한다.

별도의 비용이란 필요한 영상 정보를 찾기 위하여 구간 탐색, 재생속도 제어 등의 기능을 통해 사용자가 원하는 장면을 탐색하는 과정에서 발생하

는 시간 요소를 뜻하며 사용자는 원하는 영상 정보를 찾기 위한 부가적인 작업을 수행해야 하는 것을 의미한다.

이러한 문제를 해결하기 위해 본 논문에서는 두 가지 기법을 제안한다. 첫 번째 기법은 영상에 나오는 단어의 빈도수를 추출하여 구간별 클립정보와 함께 랭킹 테이블을 구성하고 이를 사용자에게 제공하여 검색을 지원하는 방법이다.

두 번째 기법은 머신러닝을 통해 첫 번째 방법의 클립정보를 학습시키고 학습 영상을 사용자에게 제공하여 유사한 영상을 검색할 수 있도록 지원하는 방법이다.

이 두 가지 기법을 구현하기 위해 빅데이터 분석과 영상 학습의 인공지능 요소가 필요하며 이어오는 관련 연구에서 영상 학습 관련 최신 연구인 구글의 Vision AI에 대하여 설명하고 3장에서 제시된 기법에 대한 설계를 제시한다.

II. 관련 연구

구글은 이미지 검색을 지원하기 위하여 AutoML Vision과 Vision API를 구현하여 제공하고 있다.

또한, 많은 연구에서 영상 분석 기법을 활용하고 있다[1].

AutoML Vision[2]은 사용하기 쉬운 그래픽 인터페이스로 간단하게 이미지를 업로드해 커스텀 이미지 모델을 학습시킬 수 있으며 정확성, 지연 시간, 크기에 맞춰 모델을 최적화한 후 클라우드의 애플리케이션 또는 에지의 다양한 기기에 내보낼 수도 있는 영상 처리 전용 API이다.

Vision API[3]는 REST 및 RPC API를 통해 선행 학습된 강력한 머신러닝 모델을 제공하고 이미지에 라벨을 할당하고 사전 정의된 수백만 개의 카테고리 빠르게 분류할 수 있게 구현되어 있다.

그리고 객체와 얼굴을 감지하고 인쇄 및 필기 텍스트를 읽으며 유용한 메타데이터를 이미지 카탈로그에 구축해 주기 때문에 기존의 영상 처리 및 분석을 위한 학습 모델을 구축하기 어려운 환경을 해결해 줄 수 있는 좋은 대안으로 판단된다.

그러나 구글의 API는 높은 비용을 요구하기 때문에 본 논문에서 제시하는 영상 처리 및 키워드 분석 시스템에는 효율적이지 않다.

III. 시스템 설계

그림 1과 2는 본 논문에서 제시하는 키워드 분석과 영상 분석 시스템의 설계 구조이다.

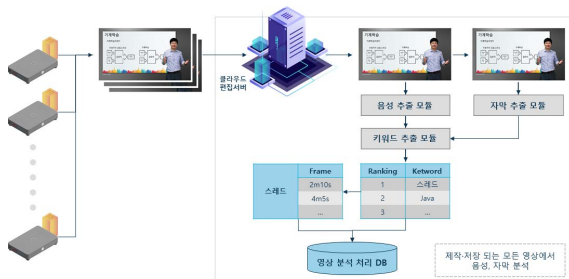


그림 1. 키워드 분석 기법

키워드 분석 기법은 초기 빅데이터 영상 분석 기법으로 영상의 음성 정보와 자막 파일을 분석하여 가장 빈번하게 발생한 키워드를 레벨링 하여 영상 분석에 활용한다.

영상 정보 뿐만 아니라 영상의 핵심 되는 정보는 일반적으로 강의자의 음성 또는 자막 정보를 통해 나타나므로 분석 키워드를(음성, 자막)으로 구성 한다.

영상 분석 기법은 일정 수준 이상의 영상 데이터가 구축된 단계에서 적용할 수 있으며 초기 단계 구축된 영상 분석 처리 DB의 키워드 정보를 토대로 학습 Video 추출하여 학습을 수행한다.

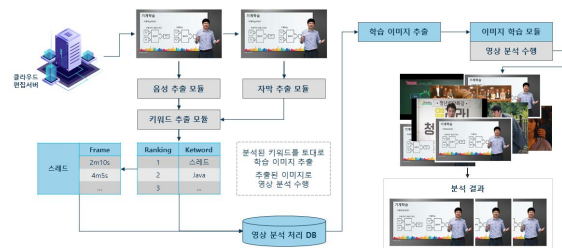


그림 2. 영상 분석 기법

이후 딥러닝을 활용하여 Video 영상 학습 수행 유사 이미지 검색 기법을 통한 이미지를 검색한다.

이미지 검색 대상은 키워드 분류가 어려운 영상 즉, 음성이 없고 자막이 프레임에 예측된 영상이나 제목 및 태그 분류만 이루어진 영상을 대상으로 한다.

그 이유는 모든 영상 패턴을 추출할 경우 분석 시스템에 상당한 자원 부하가 발생하므로 사용자의 검색 패턴에 맞춘 영상 분석을 지원하여 시스템의 자원 비용을 최적화한다.

IV. 결론

본 논문에서 효율적인 영상 검색을 위한 두 가지 기법에 대하여 설계 및 제안하였다. 사용자가 원하는 구간의 영상을 효과적으로 검색할 수 있으며 상용화 API를 통한 비용 증가를 최대한 낮출 수 있는 구조로 제시되었다.

향후 시스템 구현을 통해 제시된 기법들을 검증하고 상용화 시스템에 API 타입으로 제공할 예정이다.

Acknowledgement

이 논문은 2019년도 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원을 받아 수행된 연구임 (No.2019-0-01644, 빅데이터 영상 분석기능을 탑재한 능동형 영상 촬영 및 서비스 시스템)

References

- [1] J. H. Jun, J. O. Song, S. M. Lee, "Facial Expression Analysis System based on Image Feature Extraction", Proceedings of the Korean Society of Computer Information Conference. July 12, 2016, pp. 293 -294
- [2] Google AutoML Vision [Internet]. Available : <https://cloud.google.com/vision/?hl=ko#>
- [3] Google Vision API [Internet]. Available : <https://cloud.google.com/vision/?hl=ko#>