

딥러닝 기술을 활용한 촬영 영상 시간동기화 서비스

홍순기
SBS 미디어기술연구소
연구원
sghong@sbs.co.kr

개요

SBS 미디어기술연구소는 지난 몇 년 동안 딥러닝 기술을 이용한 제작 편의 향상 방안을 제작진과 협의해 왔다. 그리고 드디어 협업의 첫 번째 결과물로 ‘딥러닝 기술을 활용한 촬영 영상 시간 동기화 서비스’가 탄생했다. ‘딥러닝 기술을 활용한 촬영 영상 시간 동기화 서비스’는 SBS 예능 프로그램의 VJ캠, ENG캠 및 거치캠 레이저 쌓기 작업에 적용되어 제작 시간 단축에 기여하고 있다. 감사하게도 이러한 성과를 인정받아, 지난해 말 과학기술정보통신부가 주최하고 한국전파진흥협회가 주관한 <제23회 전파방송 기술대상>에서 과학기술정보통신부 장관상을 받게 되었다. 수상을 계기로 SBS 미디어기술연구소가 개발한 ‘딥러닝 기술을 활용한 촬영 영상 시간 동기화 서비스’의 주요 기술에 대해 본 기고를 통해 소개하고 공유하려고 한다.



그림 1. 제23회 전파방송 기술대상 수상 사진
/ 출처 : 전자신문 김민수 기자,
www.etnews.com/20221121000219



그림 2. <제23회 전파방송 기술대상>에서 전시된 ‘딥러닝 기술을 활용한 촬영 영상 시간동기화 서비스’

레이어 쌓는 작업은 편집을 위한 필수 작업이다. 같은 시간에 촬영된 장면이 편집기 타임라인의 같은 시점에 위치하도록 쌓아야 하기에 어렵고 시간이 오래 걸린다. 특히 카메라를 다수 사용하는 제작 환경에서는 더 많은 시간과 노력이 필요하다.

‘딥러닝 기술을 활용한 촬영 영상 시간동기화 서비스’는 딥러닝 기반 자동화로 제작진의 노력과 시간을 절약해준다. 촬영 현장에서 제작진이 스마트폰에 설치한 앱을 실행하고 카메라에 스마트폰 화면을 보여주는 것만으로, 타임코드 삽입이 가능하다. 그러면 타임코드 추출 서버는 딥러닝 기술로 촬영 영상 안에 찍혀 있는 타임코드를 스스로 찾아낸다. 또한 찾아낸 타임코드로 영상이 촬영된 시간을 정확하게 예측한다. 이제 인제스트(트랜스코딩)를 시작하기만 하면, 제작진이 ‘딥러닝 기술을 활용한 촬영 영상 시간동기화 서비스’를 사용할 준비는 모두 끝난다. 인제스트 완료 후, 제작진은 편집기에서 플러그인을 실행하고 ‘조화’, ‘등록’ 두 번의 버튼을 누르는 것만으로 레이어를 자동으로 쌓을 수 있다.



그림 3. ‘딥러닝 기술을 활용한 촬영 영상 시간동기화 서비스’ 개요

딥러닝 기반 촬영 영상 시간동기화 기술 개발

SBS 기술연구소는 스마트 단말과 딥러닝 기술을 이용하여 편집기에서 촬영 영상 시간동기화를 수행하는 기술을 개발하였다. 촬영 환경에서 쉽게 사용할 수 있게 스마트 단말(폰,패드)을 사용하는 타임코드 표출 앱을 개발하여, 제작 환경에서 휴대 및 조작이 용이하도록 하였으며, 스마트 단말 간 시간동기를 위해 별도의 비용이나 장비가 필요 없게, 네트워크 타임 프로토콜(NTP)을 통해 공개된 NTP 서버 풀을 사용하여 단말 간 시간동기화를 수행할 수 있게 하였다. 또한 촬영된 타임코드의 추후 활용 가능성을 높이기 위해 타임코드는 연-월-일-시-분-초-밀리초 양식을 사용하였다.



그림 4. 타임코드 표출 형식 및 실제 촬영 예시

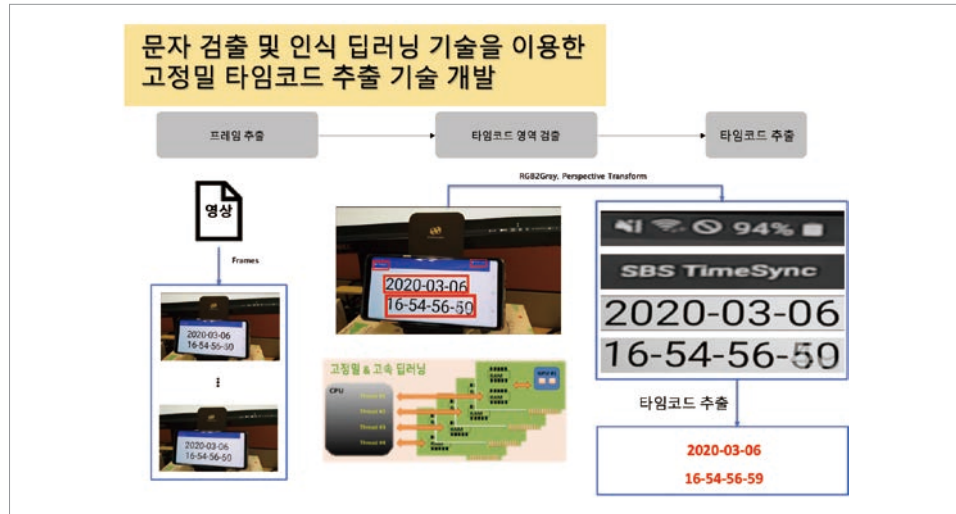


그림 5. 딥러닝 기반 타임코드 추출 시스템 구조도

영상에 촬영된 타임코드를 문자 검출 및 인식 딥러닝 기술을 이용하여 다양한 촬영 환경에서 정확하게 타임코드를 추출하는 기술을 개발하였다. 영상 내에서 문자 영역을 높은 정확도로 검출하는 Scene Text Detection 딥러닝 기술을 개발하기 위해, 공개된 오픈소스 알고리즘을 기반으로 다양한 촬영 환경에서도 높은 검출률을 보일 수 있도록 학습데이터를 자체적으로 생성하여 재학습하였다. 또한 촬영 환경에 따라 변화는 타임코드 이미지의 크기와 기울기에 강인한 이미지 처리 기술과 오픈소스 문자인식 알고리즘을 바탕으로 타임코드 패턴 매칭을 이용해 검출된 타임코드 이미지를 타임코드 문자로 변환하는 기술 역시 적용하였다.

마지막으로 추출된 타임코드를 이용하여 영상이 촬영된 시간을 추정하는 기술을 개발하였다. 통계적 분석 기법을 활용하여 문자 인식 오류에 강인한 촬영 시간 추정이 가능하게 하였다.

방송 콘텐츠 제작 패턴을 고려한 고속 트랜스코딩 및 타임코드 추출 기술 개발

일반적으로 방대한 양의 촬영 영상에 들어 있는 타임코드를 추출하기 위해서는 매우 많은 시간이 소요된다. 이를 해결하기 위해서 방송 콘텐츠 제작 과정의 일부인 인제스트(트랜스코딩)를 고속으로 수행하는 분산 처리 기반 트랜스코딩 플랫폼과 타임코드 추출 플랫폼을 연결하여 인제스트 시간 안에 타임코드 추출을 완료할 수 있게 시스템을 구성하였다. 또한 타임코드 고속 추출을 위해, 멀티 프로세스를 사용하여 촬영 영상 디코딩과 타임코드 검출을 병렬로 처리하는 기술, 타임코드가 포함되어 있는 영상 구간의 통계값에 기반하여 최적 검출 스케줄링을 수행하는 기술, 그리고 Nvidia사 TensorRT를 이용한 딥러닝 네트워크 최적화 및 멀티 GPU 기반 병렬처리 기법 적용하였다.

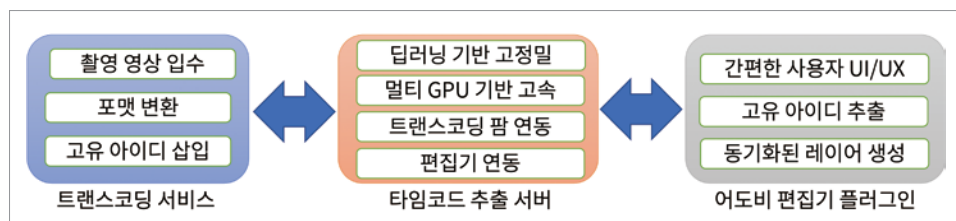


그림 6. 트랜스코딩, 타임코드 추출 서버 그리고 편집기 플러그인 간 연동

**제작진에게
익숙한 편집기
플러그인으로
자동 시간동기화
클라이언트 개발**

아무리 좋은 기술을 개발하였어도 사용자에게 익숙하지 않고 불편하다면 좋은 기술을 개발한 의미가 많이 퇴색될 것이다. 따라서 ‘딥러닝 기술을 활용한 촬영 영상 시간동기화 서비스’는 제작진이 사용하는 최종 클라이언트를 Adobe Premiere Pro 편집기의 서드파티 플러그인으로 개발하였다. 제작진이 편집에 사용해오던 편집기에 자동 동기화 클라이언트를 개발하여 제작진의 사용 편의성을 극대화하였다. 개발한 편집기 플러그인은 분산 트랜스코딩 및 타임코드 추출 플랫폼과 연동되어 제작진의 원클릭만으로 시간동기화를 가능하게 한다. 또한 제작진의 편의성을 더욱 높이기 위해, 하나의 촬영 기기에서 생성된 파일들을 묶어 편집기에 하나의 트랙으로 맵핑하는 기능 및 챗터로 영상이 나뉘어 저장되는 카메라(GoPro, Sony Eng 등) 지원을 위해 챗터로 저장된 영상 파일들을 자동으로 트랙에서 붙여 주는 기능 등도 같이 개발되어 플러그인에 포함되었다.

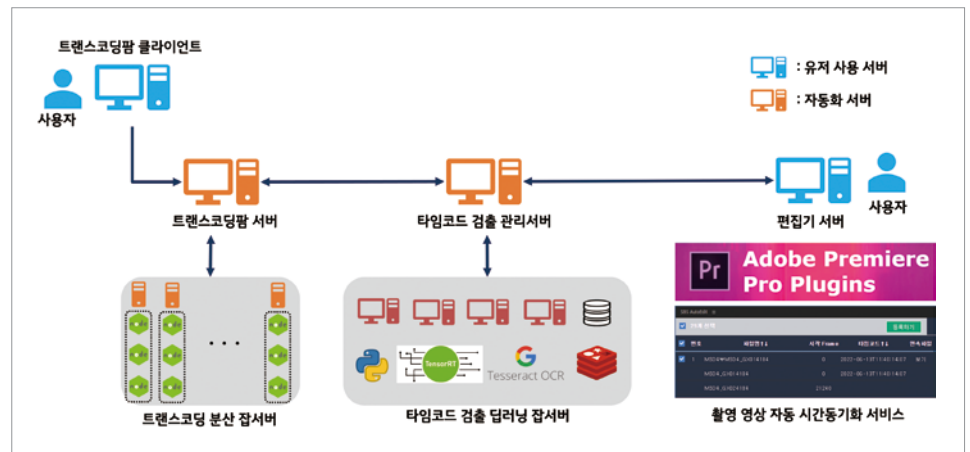


그림 7. 사용자 편의성을 극대화하기 위한 워크플로우 자동화 구조

마치며

‘딥러닝 기술을 활용한 촬영 영상 시간동기화 서비스’ 경험을 기반으로, SBS 미디어기술연구소는 ‘AI 기반 촬영 영상 메타데이터 검색 서비스’를 준비하고 있다. 레이어 자동 동기화를 위한 원포인트 솔루션 개발이 이번 서비스의 목표였다면, ‘AI 기반 촬영 영상 메타데이터 검색 서비스’는 제작 편의성 향상을 위한 토탈 솔루션 개발이 목표이다. 얼굴인식, 대사생성, 문자인식 등의 딥러닝 기술을 전방위적으로 활용하여, 편집기에서 제작진이 특정 장면을 빠르고 정확하게 찾는 솔루션을 제공할 계획이다. 📺