

# 청각장애인을 위한 지상파 UHD 기반 감정표현 자막 송출 시스템 개발

이준 안충현

(주)에어코드

junelee@aircode.com

## Development of Emotion Subtitles Broadcast System based on Terrestrial UHD TV for the Hearing-Impaired

June Lee, Chunghyun Ahn

AIRCODE

### 요약

최근 지상파 UHD 방송에서는 자막을 비디오 패킷에 삽입하여 전달하는 기존 방식이 아닌 별도의 전송로로 전달하는 폐쇄형 자막(closed caption) 방식을 적용하고 있다. 또한 자막에는 문자 뿐만 아니라 이미지까지 포함하여 청각장애인의 프로그램 이해도를 높이는데 활용할 수 있다. 따라서 본 논문에서는 청각장애 시청자가 방송 콘텐츠 내용의 이해도를 향상시키기 위하여 지상파 UHD 방송에서 기존의 문자 자막과 움직이는 이미지 자막(=감정표현 자막)을 송출하여 동시에 표현할 수 있는 송출시스템을 제안한다.

### 1. 서론

자막방송은 청각 장애인들을 위한 문자 방송 서비스로서, 의미있는 오디오 콘텐츠(대화 내용, 화자의 말하는 태도, 음악 효과 등)를 시각적으로 문자와 기호를 이용하여 부가 방송 형태로 제공한다. 최근 지상파 UHD 방송에서는 자막을 비디오 패킷에 삽입하여 전달하는 기존 방식이 아닌 별도의 전송로로 전달하는 폐쇄형 자막(closed caption) 방식을 적용하고 있다. [1, 2] 또한 자막에는 문자 뿐만 아니라 이미지까지 포함하여 청각장애인의 프로그램 이해도를 높이는 데 활용할 수 있다.

따라서 본 논문에서는 청각장애 시청자가 방송 콘텐츠 내용의 이해도를 향상시키기 위하여 지상파 UHD 방송에서 기존의 문자 자막과 움직이는 이미지 자막(=감정표현 자막)을 송출하여 동시에 표현할 수 있는 송출시스템을 제안한다.

### 2. 감정표현 자막 송출시스템 구조도 및 요구사항

ATSC3.0 기반 감정표현 자막 서버 시스템 구현은 다음과 같은 구성 요소 개발을 요구한다.

- AV 신호 및 감정표현 자막 데이터 송신을 위한 ATSC3.0 규격 기반 LLS, SLS 시그널링 신호 생성
- AV 신호와 감정표현 자막 데이터(문자 및 이미지(PNG & APNG))의 동기화 기능
- IMSC1 표준 기반 감정표현 자막 데이터의 ISO/IEC 14496-30 규격 기반 ISOBMFF 패키징 기능
- 감정표현 자막 데이터의 DASH/ROUTE 패키징 기능
- 감정표현 자막 데이터의 MPU/MMT 패키징 기능

### ● 실시간 동작 검증용 모니터링 기능

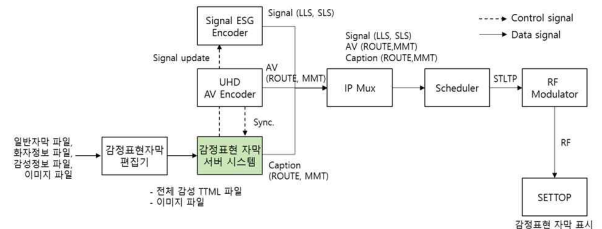


그림 1. ATSC3.0 기반 지상파 UHD 방송 및 감정표현 자막 서버 시스템 개념도

ATSC3.0 기반 감정표현 자막 서버 시스템은 기존의 지상파 UHD 방송시스템과 호환성을 유지하여 개발되어야 한다. ATSC3.0 기반 감정표현 자막 서버 시스템은 감정표현자막 편집기에서 생성된 자막 TTML과 이미지 파일을 입력 받아 AV와 동기화된 MMT/ROUTE 패킷을 생성하여 출력해야 하며, 감정표현 자막 서비스를 위한 시그널링 신호를 생성 또는 업데이트하여야 한다. 이후의 동작은 기존 방송시스템과 동일하며 스케줄러, 변조, 익사이터를 거쳐 RF로 송출된다.

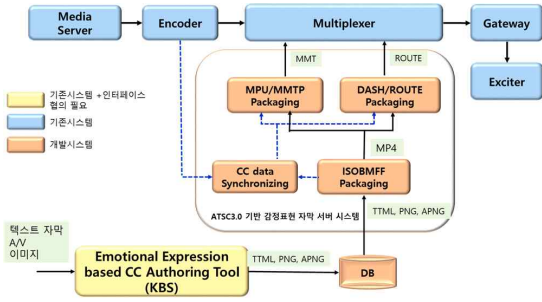


그림 2 감정표현 자막 서버 시스템 구성도

ATSC3.0 기반 감정표현 자막 서버시스템은 그림 2와 같이 자막 편집 기로부터 텍스트와 이미지의 TTML과 해당 이미지(PNG, APNG)를 입력 받아 ISO/IEC 14496-30 규격 기반 ISOBMFF 패키징을 수행한다. 이어 DASH/ROUTE 또는 MPU/MMTP으로 패키징 된 후에 IP MUX로 출력된다. 이때, 감정표현 자막 동기화 모듈에서 AV 신호와 자막 사이의 동기화 작업을 수행한다. 감정표현 자막 서버 시스템의 입출력 형태는 아래와 같다.

- 입력: 텍스트 TTML & 이미지 TTML + 이미지(PNG, APNG)
- 출력: ROUTE 또는 MMT 패킷

감정표현 자막 서버 시스템의 주요 기능

- 감정표현 자막을 위한 SLS 시그널링 신호 생성 기능
- 입력받은 감정표현 자막의 1초 또는 2초 단위 ISOBMFF 패키징 기능
- AV 신호와 감정표현 자막 데이터(문자 및 이미지(PNG & APNG))의 동기화 기능
- 감정표현 자막 데이터의 DASH/ROUTE 패키징 기능
- 감정표현 자막 데이터의 MPU/MMT 패키징 기능
- 정상 동작 모니터링 기능
- AV 신호와 동기화를 위한 수동 타이밍 조절기능(-2, +2 이내)
- 최소 2개 이상의 복수 자막 데이터 처리 지원

### 3. 감정표현 자막 송출시스템 주요 기능

1) 감정표현 자막을 위한 ROUTE SLS (USBD, S-TSID, MPD) 송출 기능

- Signal&ESG Manager에서 감정표현 자막의 ROUTE SLS 등록 기능 구현

1) Text 자막과 Image 자막을 SLS에 등록

2) MPD에 자막 정보 등록

- Profile 값이 0일 경우 Text자막이며 1일 경우 Image 자막임



그림 3 SLS에 텍스트 자막 정보 등록



그림 4 SLS에 이미지 자막 정보 등록

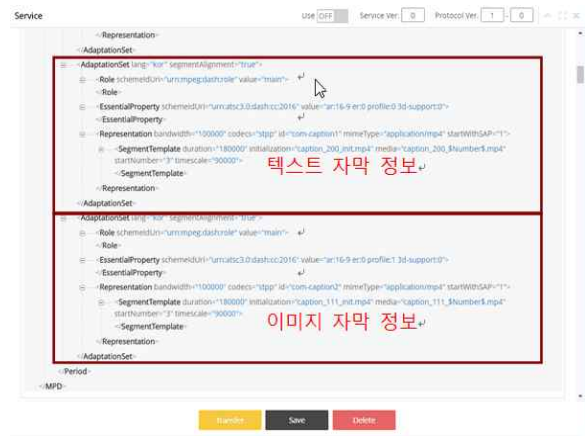


그림 5 MPD에 자막 정보 등록

2) 감정표현 자막을 위한 MMT SLS (MPT Message, Caption Descriptor) 송출 기능

- Signal&ESG Manager에서 감정표현 자막의 MMT SLS 등록 기능 구현

1) Text 자막과 Image 자막을 SLS에 등록

2) Caption Asset Descriptor에 자막 정보 등록

- Profile 값이 0일 경우 Text자막이며 1일 경우 Image 자막임



그림 6 SLS에 텍스트 자막 정보 등록

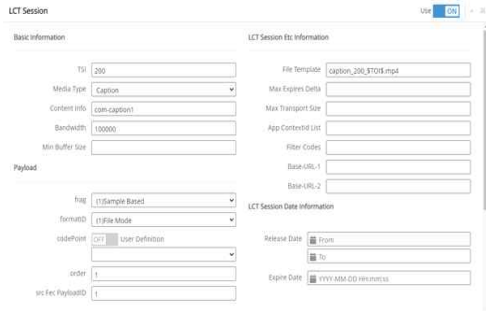


그림 7 SLS에 이미지 자막 정보 등록

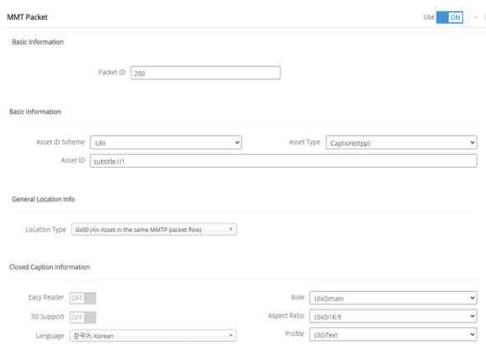


그림 8 Caption Asset Descriptor에 자막 정보 등록

- 3] Text 자막, Image 자막, Image 파일(Animated PNG)을 입수  
 1) 자막편집기로부터 UDP 수신을 받아 해당 경로의 파일을 자막인코더 폴더로 가공(Image 자막) 및 복사(Text 자막, Image 파일)하여 전달  
 2) 가공 및 복사된 자막을 자막 인코더에서 ISOBMFF로 인코딩 후 송출

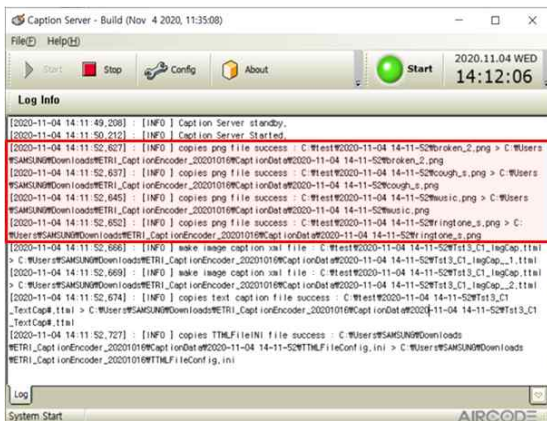


그림 9 자막 서버에서 자막 데이터 입수 화면

- 4] 감정표현 자막의 1초 또는 2초 단위 ISOBMFF 패키징 기능  
 ● 감정표현 자막을 정해진 초단위로 분리하고 ISOBMFF 형식으로 인코딩  
 1) 자막편집기로부터 UDP 수신을 받아 해당 경로의 파일을 자막 서버에서 정해진 초단위로 자막(Text, Image) 분리 및 복사(Image 파일)함  
 2) 분리된 자막을 자막 인코더에서 ISOBMFF로 인코딩 후 송출

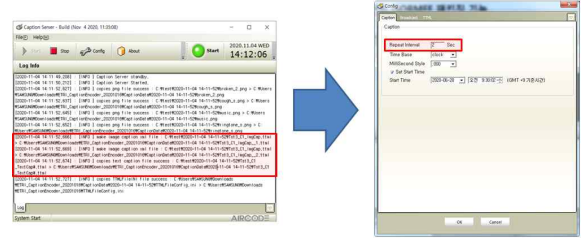


그림 10 자막 서버에서 이미지 자막을 2초 단위로 자르는 화면과 자막 인코더에서 이미지 자막을 2초 단위로 송출하도록 설정하는 화면

- 5] AV 신호와 동기화를 위한 수동 타이밍 조절 기능  
 ● AV 신호와 동기화를 위해서 사용자가 수동으로 Offset값을 조절할 수 있는 기능  
 1) 자막 인코더 설정창에 TTML Offset 항목을 추가하여 밀리초 단위로 -,+ 값 입력 가능  
 예) TTML Offset에 '?2000' 값 입력 : 원래 자막 시간보다 2초 빠른 시간으로 자막 시작시간과 종료 시간을 수정



그림 11 자막 인코더 설정 창에서 TTML Offset 설정

- 6] 감정표현 자막 데이터의 DASH/ROUTE, MPU/MMT 인코딩 기능  
 ● 자막의 ISOBMFF 인코딩 프로세스 정의  
 1) TTML자막을 ROUTE 프로토콜용 ISOBMFF 인코딩  
 ① Init Segment 생성 (자막 파일에 대한 메타데이터 정보)  
 ② Init Segment용 Box 생성  
 ③ Media Segment 생성 (자막 파일 삽입)  
 ④ Media Segment용 Box 생성  
 2) TTML자막을 MMT프로토콜용 ISOBMFF 인코딩  
 ① MPU Metadata 생성 (자막 파일에 대한 메타데이터 정보)  
 ② MPU Metadata용 Box 생성  
 ③ MFU 생성 (자막 파일 삽입)  
 ④ MFU용 Box 생성  
 ⑤ Movie Fragment Metadata 생성 (자막 파일에 대한 메타데이터 정보)  
 ⑥ Movie Fragment Metadata용 Box 생성

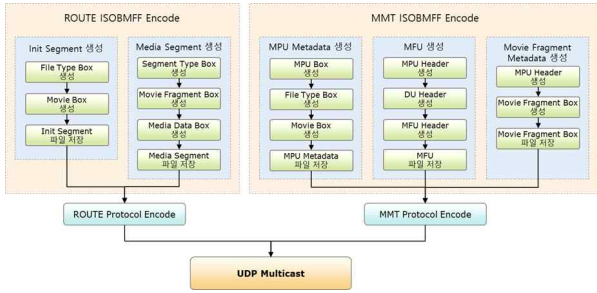


그림 12 ISOBMFF 인코딩 모듈 구성도

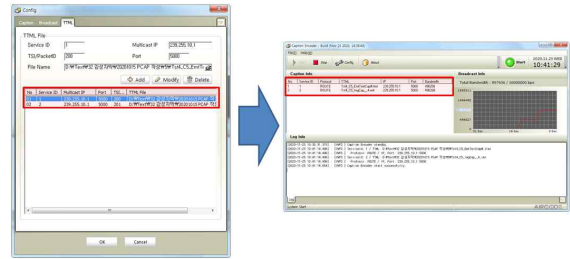


그림 15 자막 인코더의 Text 자막과 Image 자막 송출 설정 화면과 자막 인코더의 Text 자막과 Image 자막 송출 화면

기 감정표현 자막 데이터의 DASH/ROUTE, MPU/MMT 패키징 기능

● ROUTE/MMT ISOBMFF 인코딩 기능 구현

- ① Text 자막과 Image 자막(이미지 파일도 같이 첨부)을 ROUTE 혹은 MMT 프로토콜용 ISOBMFF 인코딩 수행
- ② Image 자막의 경우 Animated PNG 파일의 정보를 Movie Fragment Box의 첫번째 Track Fragment Box에 Sub-Sample Information Box ('subs')로 추가함
- ③ Image 자막의 Animated PNG 파일은 Media Data Box에서 TTML 바로 뒤에 추가함

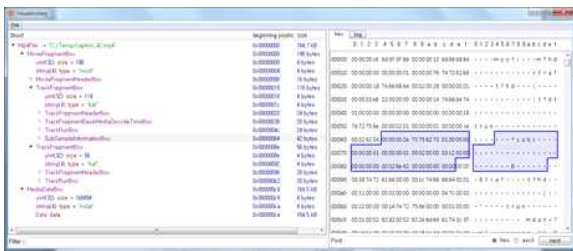


그림 13 MMT ISOBMFF 인코딩 결과

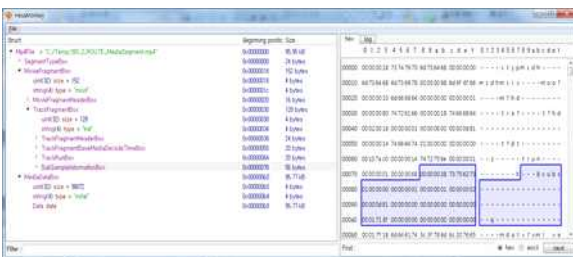


그림 14 ROUTE ISOBMFF 인코딩 결과

8) 최소 2개 이상의 복수 자막 데이터 송출 기능

● 자막 인코더에서 Text 자막과 Image 자막을 동시에 송출할 수 있는 기능 구현

- 1) 자막 인코더 설정 창에서 Text 자막과 Image 자막을 등록할 수 있는 화면 제공
- 2) 자막 인코더 송출 모듈에서 Text 자막과 Image 자막 동시 송출 지원
- 3) 자막 서버에서 자막 인코더의 Text 자막과 Image 자막 송출 설정을 변경할 수 있게 지원

4. 결론



그림 16 감정 표현 자막 송출 시스템에서 송출한 신호를 셋탑 박스에서 수신해서 텍스트 자막과 이미지 자막을 재생 화면

본 논문에서는 지상파 UHD 방송에서 기존의 문자 자막과 움직이는 이미지 자막(감정표현 자막)을 송출하여 동시에 표현할 수 있는 송출시스템을 제안하고, 실제 실험 환경을 구축하여 기능 검증을 수행하였다. 사용자가 문자 자막과 이미지 자막을 등록하고 자막서버에서 시그널링 신호와 자막을 인코딩하고 송출해서 셋탑에서 정상 수신 및 화면에 표출하는 것을 확인하였다. 이를 통하여 청각장애 시청자가 자막 방송을 통한 방송 콘텐츠의 내용을 이해하는데 도움이 되고, 일반 시청자와의 시청 선택권을 확보할 수 있다.

ACKNOWLEDGEMENT

본 연구논문은 과학기술정보통신부 및 정보통신기획평가원의 출연금으로 수행하고 있는 한국전자통신연구원 시청각장애인의 방송시청을 지원하는 감성표현 서비스 개발(2019-0-00447)의 연구결과입니다.

참 고 문 헌

- [1] 안충현, "장애인방송 기술개발 현황", 전자통신동향분석 제34권 제3호, 2019.06
- [2] "지상파 UHD TV 방송 송수신 정합", TTAK.KO-07.0127/R5, 2020.12
- [3] "ATSC Standard: Caption and Subtitles", ATSC A/343:2016, 2016.12
- [4] "Information technology - Coding of audio-visual objects - Part 12: ISO base media file format" ISO/IEC 14496-12:2005(E)