

멀티뷰를 활용한 라이브 커머스 플랫폼 설계

*우예지, 윌애령, 윤정원, 이신화, 전수민, 이상운
 동아방송예술대학교
 *yejiu0828@gmail.com

Design of a Live Commerce Platform Using a Multiview

*Yeji Woo, Aeryeong Won, Jeongwon Yun, Shinhwa Lee, Sumin Jeon, Sangun Lee
 Dong-Ah Institute Of Media And Arts

요약

코로나 19로 인한 사회적 거리 두기가 계속되면서 온라인 쇼핑을 이용하는 고객이 증가했다. 그중 원활한 소통이 가능한 라이브 커머스 시장이 크게 성장했다. 모바일 기기만 있으면 시간과 장소의 제약 없이 라이브 커머스를 이용할 수 있지만 제한된 정보제공과 장애인을 위한 서비스가 없다는 것이 단점이다. 따라서 본 논문에서는 다양한 정보를 제공하기 위한 멀티뷰 화면을 송출하고 TTS, 딥러닝 기반의 STT 기술을 활용해 시청각 장애인을 위한 기능을 포함한 새로운 형태의 라이브 커머스 플랫폼 및 시스템 구조를 제안한다.

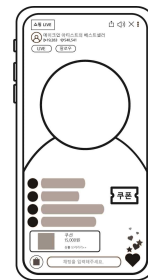
1. 서론

모바일 기기의 발전과 코로나 19로 인해 온라인 쇼핑을 이용하는 고객이 급증하고 있다. 이에 따라 일방향성 정보 전달을 하던 기존의 홈쇼핑은 쇠퇴하고 실시간으로 쌍방향 소통이 가능한 라이브 커머스 시장이 커지고 있다. 또한, 1인 미디어가 발전하면서 기업뿐만 아니라 개인이 라이브 커머스에 뛰어들기도 한다.

이처럼 많은 사람들이 라이브 커머스를 이용하고 있지만 다양한 사람들이 이용하지는 못하고 있다. 즉, 방송 소외계층이 존재하는 것이다. 시각 장애인인 쇼호스트의 설명을 통해 정보 취득이 쉬울 수 있지만, 자세한 설명이 제공되지 않으면 현재 방송되고 있는 상황을 이해하기 어렵다. 청각장애인은 실시간 자막 기능이나 수어 영상이 필요하지만 서비스를 지원하지 않아 정보 취득에 어려움을 겪는다. 그리고 현존하는 라이브 커머스 플랫폼은 실시간으로 방송되고 있는 한 화면만 시청할 수 있어 많은 시청자가 불편함을 호소하고 있다.

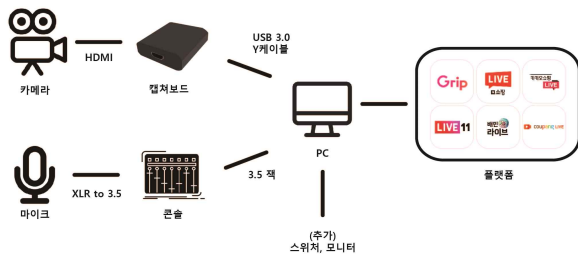
본 논문에서는 멀티뷰로 다양한 화면과 콘텐츠를 제공해 시청자가 원하는 화면과 콘텐츠를 선택해 볼 수 있고, TTS와 딥러닝 기반의 STT 기술을 활용한 음성 안내와 실시간 자막을 지원해 시·청각 장애인도 편리하게 이용할 수 있는 새로운 형태의 라이브 커머스 플랫폼을 제안한다. 그래서 현재 라이브 커머스 방송의 정보 전달에 있어서 한계를 극복하고 시청자의 불편 해소 및 방송 접근권을 보장하고자 한다. [1]

2. 기존의 라이브 커머스 플랫폼 및 TTS, STT



[그림 1] 기존 라이브 커머스 플랫폼 형태

네이버, 카카오, 배달의 민족 등이 라이브 커머스 플랫폼을 출시했다. 네이버 쇼핑 라이브는 쉬운 접근성과 낮은 수수료로 소상공인, 공공기관에 특화되어 농수산물, 특산품 등이 많이 판매되고 있다. 카카오 쇼핑 라이브는 높은 수수료를 갖고 있지만, 전용 스튜디오와 라이브 커머스 담당팀을 구성하여 차별화된 쇼핑 콘텐츠를 제공할 수 있다. 현존하는 라이브 커머스 플랫폼은 판매자에게 부담이 될 수 있는 스튜디오나 장비 기기 등을 대여하는 방식으로 제공하고 시청자 수, 채팅 기능, 쿠폰, 결제 기능을 지원한다. 이처럼 편리하게 방송할 수 있어 많은 개인과 기업이 플랫폼을 이용해 방송을 진행하고 있다. 현재 사용되고 있는 라이브 커머스 시스템을 개인과 기업으로 분리해 볼 수 있다. [2]



[그림 2] 기존 라이브 커머스 시스템 구성도 (기업)

기업에서는 [그림 2]와 같이 시스템을 구성한다. 카메라를 통해 입력된 영상 소스는 캡처 보드로 입력되고 마이크를 통해 입력된 음성 소스는 콘솔로 입력된다. 캡처 보드와 콘솔에서 출력된 각각의 소스는 PC의 OBS와 같은 송출프로그램을 거쳐 라이브 커머스 플랫폼을 통해 실시간으로 방송된다. 더 완성도 높은 라이브를 제공하려면 조명을 설치하고, 듀얼 모니터를 추가하여 송출프로그램 관리와 실시간 댓글, 라이브 방송을 같이 확인할 수 있다. [3]



[그림 3] 기존 라이브 커머스 시스템 구성도 (개인)

개인은 [그림 3]과 같이 설계한다. 스마트폰, 무선 마이크, 라이브 커머스 플랫폼으로 구성한다. 먼저, 스마트폰에 라이브 커머스 플랫폼을 설치한다. 스마트폰에 무선 마이크를 연결한 후, 설치된 플랫폼을 열어 스마트폰의 카메라 기능을 활용하여 촬영한다. 스마트폰으로 입력된 영상 소스와 무선 마이크에 입력된 음성 소스는 플랫폼을 거쳐 시청자에게 방송된다. 휴대폰 1대를 더 추가하면 시청자의 반응을 실시간으로 확인해 더욱 원활한 소통이 가능하다. [4]

본 논문에서 적용할 TTS는 Text to Speech의 약자로, 문자를 음성으로 바꿔주는 기술이다. 사전 녹음된 목소리를 전용 엔진이 단어별로 쪼개 재조합하고 사용자에게 들려준다. IT 기업들은 시각장애인들이 휴대폰을 편리하게 사용할 수 있도록 TTS 기술에 기반한 기능을 제공하고 있다. 애플은 아이폰에 'Voice Over 및 화면 말하기'를 제공한다. 이는 모든 메뉴와 애플리케이션 제목, 텍스트를 읽어주고 원하는 메시지나 웹 페이지 등 텍스트를 하이лай트로 표시하면 해당 항목을 읽어주는 '선택 항목 말하기' 기능을 제공한다. 삼성은 갤럭시에서 휴대폰을 여는 순간부터 화면 내용을 음성으로 안내해주는 'TalkBack' 기능을 기본으로 내장하여 시각장애인 또는 고령자가 스마트폰을 편리하게 사용할 수 있도록 보조하고 있다. STT는 Speech to Text의 약자로, 음성을 컴퓨터가 해석해 문자로 바꿔주는 기술이다. 답러닝은 인간의 뇌 신경 네트워크를 흉내 낸 인공 뉴런들로 이루어진 신경망에 의한 학습 시스템으로 인간의 뇌와 같은 구조를 기계에 구현한 기술이다. 인공지능은 구분하고 이해하는 감각적인 업무를 힘들어한다. 하지만 답러닝은 특징을 파악할 수 있

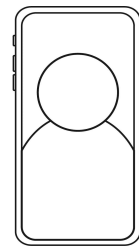
고 인공지능망을 통해 학습되기 때문에 개발자가 데이터를 따로 지정하지 않는다. 또한, 이러한 답러닝 기술을 음성인식 기술에 적용하면 음성인식의 신뢰성 확보가 가능한 것이 장점이다. 이를 기반으로 만들어진 기술이 캡션 기능이다. 청각장애인은 영상을 시청할 때 자막이 필요한데 캡션 기능을 통해 실시간 자막을 이용할 수 있다. 사용자가 음성인식 디바이스로 인터넷을 통해 접속하면, 서버에서 답러닝 기반의 STT로 전환되어 음성이 문자로 표현되면서 서비스를 제공한다. [5], [6], [7], [8], [9]

3. 제안한 라이브 커머스 플랫폼

3.1 제안한 플랫폼 인터페이스 구조



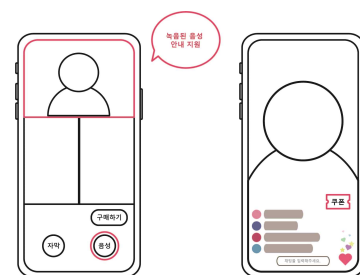
[그림 4] 제안한 멀티뷰 구조



[그림 5] 더블 클릭시 구조

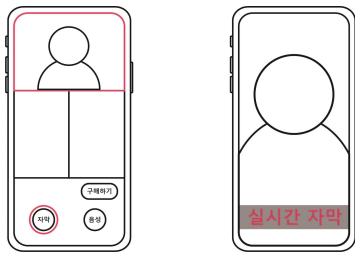
본 논문에서 제안하는 새로운 라이브 커머스 플랫폼 <Lime : Live, Multiview for Everyone>에 접속 시 첫 화면에 [그림 4]와 같이 멀티뷰가 나타나고 하나의 멀티뷰 화면을 선택해 더블 클릭하면 [그림 5]처럼 풀스크린이 된다. 기존 라이브 커머스 플랫폼처럼 첫 화면에 라이브 방송 화면을 배치하고 후면에 멀티뷰를 배치하면 시각장애인이 라이브 방송 화면에 있는 여러 아이콘으로 인해 음성 버튼을 누르기까지의 과정이 복잡해진다. 그래서 첫 화면에 멀티뷰를 배치했다.

[그림 4]에 1, 2, 3번은 멀티뷰로 다양한 화면을 구성한다. 각 멀티뷰 화면을 통해 실시간 방송 송출이 가능하고, 관리자가 미리 편집된 동영상 업로드도 할 수 있어 관리자의 선택에 따라 각 멀티뷰를 실시간 방송 또는 동영상 업로드 2가지로 활용할 수 있다. 관리자는 판매하는 제품 관련 영상(ex. 제품 구성, 활용 예시 등)을 업로드해 라이브 방송을 통한 정보 전달에 있어서 부족함을 채울 수 있다. 한 화면만 시청 가능한 것에 불편함을 느꼈던 시청자는 멀티뷰를 통해 현장의 카메라가 잡고 있는 화면 중 본인이 원하는 화면을 선택해서 볼 수 있으며, 관리자 측에서 제공하는 다양한 콘텐츠도 이용할 수 있다.



[그림 6] 제안한 시각장애인을 위한 기능

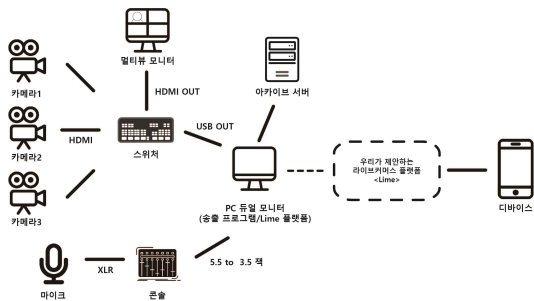
[그림 6]과 같이 멀티뷰 화면 아래에는 TTS가 적용되는 자막, 음성 버튼, 구매하기 버튼을 두어 시청각 장애인이 편리하게 이용할 수 있도록 한다. 시각장애인이 음성 버튼을 누른 후, 멀티뷰 화면을 한 번 클릭 시 해당 화면에 대한 녹음된 음성 안내를 지원하고, 두 번 클릭 시 폴스크린이 된다. 현재 라이브 커머스 플랫폼은 라이브 화면에 TTS가 적용되지 않는 이미지(ex. 쿠폰 등)를 사용하는데 이러한 이미지를 TTS가 적용되는 텍스트 버튼으로 사용해 빠짐없이 음성이 지원될 수 있도록 한다.



[그림 7] 제안한 청각장애인을 위한 기능

청각장애인은 [그림 7]과 같이 자막 버튼을 누른 후, 멀티뷰 화면을 선택해 시청하면 딥러닝에 기반한 STT가 적용되어 실시간 자막 기능이 지원된다.

3.2 제안한 플랫폼의 시스템 구조



[그림 8] 제안한 시스템 구조

[그림 8]은 < Lime > 플랫폼을 적용한 시스템 구조이다. 예시로 3대의 카메라를 이용해 시스템을 설계했다. 카메라와 스위처를 HDMI로 연결하고 스위처와 PC는 USB로 연결한다. 스위처의 HDMI OUT은 멀티뷰 모니터에 입력한다. 무선마이크의 수신기를 XLR을 이용해 콘솔과 연결한다. 콘솔은 5.5 to 3.5 잭으로 PC와 연결한다. PC로 입력된 영상과 음성 소스는 송출프로그램과 < Lime > 플랫폼을 거쳐 시청자의 디바이스로 방송된다. [그림 4]처럼 총 멀티뷰 화면이 3개로 구성되어있을 경우, CAM 1과 CAM 2의 소스를 스위처를 통해 컷팅하고 CAM 3 소스를 플랫폼 멀티뷰에 바로 인가하면 [그림 4]의 1번은 CAM 1, 2가 전환되는 메인화면, 2번은 CAM 3 소스, 3번은 관리자가 업로드한 동영상이 플레이된다. 또한, 아카이브 서버를 연결하면 플랫폼을 통해 녹화된 소스를 손쉽게 관리하고 편집할 수 있다.

3.3 제안한 시스템 구성 장비

1) 카메라



[그림 9] SONY PXW-Z90

Sony의 PXW-Z90은 3G-SDI, HDMI, 멀티/마이크로 USB 단자가 탑재되어있다. QFHD로 방송 급 화질의 4K 영상까지 제공할 수 있어 생중계 촬영이 가능하고 작고 가벼운 사이즈여서 라이브 커머스에 적합하다. [10]

2) 스위처



[그림 10] Blackmagic ATEM Mini Pro

Blackmagic의 ATEM Mini Pro는 4개의 HDMI 입력이 가능하고, USB 플래시 디스크에 H.264 포맷 영상 녹화가 가능하다. 화면 효과가 위한 DVE, 트랜지션 효과, 그린 스크린 크로마키 등을 지원해 다양한 효과를 간편하게 적용할 수 있다. 카메라 4대 및 미디어, 프리뷰, 프로그램 영상뿐 아니라 녹화 및 스트리밍, 오디오 상태를 모니터링할 수 있는 멀티뷰 기능도 제공한다. [11]

3) 마이크



[그림 11] SHURE WL185

SHURE의 WL185는 SLXD1 바디팩 송신기와 SLXD4 단일 채널 수신기와 함께 사용한다. WL185는 무선마이크이며 단일지향성이다. 쇼호스트는 제품의 정보 전달을 위해 양손을 쓰는 경우가 많고, 쇼호스트만의 목소리를 집중적으로 흡음할 수 있어 적절한 장비이다. [12]

4) 콘솔



[그림 12] YAMAHA MG124CX

대부분의 라이브 커머스는 쇼호스트가 적은 개수의 마이크를 사용하기 때문에 소형 콘솔을 이용한다. YAMAHA의 MG124CX는 소형 아날로그 콘솔로 4개의 모노 마이크/라인 입력 채널과 4개의 스테레오 라인 입력 채널로 구성되어 총 12개의 입력 채널을 가지며, 4개의 버스를 가지고 있어 라이브 커머스 방송에 적합한 장비이다. [13]

4. 결론

코로나 19로 인해 온라인 쇼핑을 이용하는 고객이 급증하였다. 라이브 커머스는 모바일을 통한 뛰어난 접근성, 판매자와 고객의 실시간 커뮤니케이션이 가능하다는 점에서 온·오프라인의 장점이 모두 내재되어 많은 인기를 끌고 있다. 기존의 라이브 커머스에 멀티뷰 기능과 시청각 장애인을 위한 기능을 추가한 <Lime> 플랫폼을 제안함으로써, 기존의 라이브 커머스의 단점을 보완하고, 앞으로 더 성장할 수 있는 라이브 커머스 시장을 기대한다. 또한, 현재 가장 유행하고 있는 라이브 커머스에서 시청각 장애인을 위한 서비스를 제공한다면 앞으로 미디어 시장에서 만들 콘텐츠나 서비스에 장애인을 위한 기능을 도입하게 하는 효과를 일으킬 수 있을 것으로 전망한다. 마지막으로, 현재 장애인을 위한 방송 서비스가 극심히 부족한 상황이다. 최근 한국정보통신기술협회는 딥러닝 기술에 기반하여 ‘음성-자막 자동변환 기술’을 개발하였으며, 한국전자통신연구원은 ‘장애인 방송 시청 지원 감성표현 서비스’ 기술을 개발했다. 한국정보통신기술협회와 한국전자통신연구원은 미디어 기술 협력을 통해 장애인 방송 서비스를 개선해 정보격차를 해소할 수 있도록 노력하고 있다. 방송통신위원회는 과학기술정보통신부와 2023년까지 음성-자막-수어 자동변환 시스템을 개발해 감성표현이 가능한 아바타 수어 방송활용을 목표로 하고 있다. 이처럼 끊임없는 기술 개발과 연구를 통해 모두가 차별 없이 편리하게 미디어 서비스를 이용할 수 있도록 미디어 산업이 발전해 나가야 한다.

5. 참고문헌

- [1] 뜨는 업계, 라이브 커머스! 시청각장애인에게는..
<https://post.naver.com/viewer/postView.nhn?volumeNo=30490826&memberNo=21473067&vType=VERTICAL>
- [2] 같은 듯 다른 플랫폼별 라이브 커머스 시장
<https://www.avocadogiant.com/portfolio/%EA%B0%99%EC%9D%80-%EB%93%AF-%EB%8B%A4%EB%A5%B8-%ED%94%8C%EB>

[%9E%AB%ED%8F%BC%EB%B3%84-%EB%9D%BC%EC%9D%B4%EB%B8%8C-%EC%BB%A4%EB%A8%B8%EC%8A%A4-%EC%8B%9C%EC%9E%A5/](https://www.youtube.com/watch?v=zjSAIQaCSIU)

[3] [알려주라방] Ep.2 모든 장비는 들어오는 곳과 나오는 곳이 있다!
<https://www.youtube.com/watch?v=zjSAIQaCSIU>

[4] 라이브 커머스, 방송 송출에 필요한 장비들
<https://m.post.naver.com/viewer/postView.nhn?volumeNo=31596246&memberNo=1738298>

[5] TTS 이것만은 알고 합시다!
<https://m.blog.naver.com/PostView.naver?isHttpsRedirect=true&blogId=skaibril&logNo=221326200258>

[6] SK텔레콤-SM엔터 협력 결실... '누구'에 '아이린 알람' 넣는다
<https://www.ajunews.com/view/20191029130051382>

[7] 시각장애인 위한 스마트폰 어디까지 왔나
<https://zdnet.co.kr/view/?no=20131015125125>

[8] STT(Speech-to-Text)
<https://woongsin94.tistory.com/333>

[9] 진짜 너무 쉬운 인공지능 원리 (ft. 머신러닝 답러닝)
https://www.youtube.com/watch?v=Udkc6rkH_CQ

[10] PXW-Z90
https://pro.sony/ko_KR/products/handheld-camcorders/pwx-z90

[11] ATEM Mini Pro
<https://www.blackmagicdesign.com/kr/products/atemmini>

[12] SLXD14/85
https://www.shure.com/en-ASIA/products/wireless-systems/slx_d_digital_wireless/slxd14-85

[13] [YAMAHA] MG124CX
<http://prod.danawa.com/info/?pcode=1316342>

[14] 온·오프라인의 장점만 모았다... 성장하는 라이브 커머스 플랫폼
<http://www.iconsumer.or.kr/news/articleView.html?idxno=12542>

[15] 방통위·과기정통부, AI 기반 음성-자막-수어 자동변환기술 시연
<http://www.epnc.co.kr/news/articleView.html?idxno=206231>