

## 음원 손실을 최소화한 음량정규화

진현진<sup>1</sup>, 이덕규<sup>2</sup><sup>1</sup>서원대학교 소프트웨어학부 소프트웨어응용전공 학부생<sup>2</sup>서원대학교 소프트웨어학부 소프트웨어응용전공 교수

sjkqtf@gmail.com, deokgyulee@seowon.ac.kr

## Loudness Normalization with Minimal Loss of Sound Source

Hyeon Jin Jin<sup>1</sup>, Deok Gyu Lee<sup>2</sup><sup>1</sup>Dept of Software Applications, Seowon University<sup>2</sup>Dept of Software Applications, Seowon University

## 요 약

음원별 음량 차이가 심하여 청각적 불편함을 초래하는 상황을 음량정규화 기술을 통해 완화하려고 한다. 그러나 현재 상용화된 관련 기술을 보면 대부분 음량을 조정하는 경우에 음원의 품질을 보장할 수 없거나, 반대로 음원의 품질을 보장하는 경우에 음량 조정이 지극히 미세한 수준이다. 또한 음원 제공 플랫폼이나 음원의 형태에 따라 적용 가능한 대상이 제한되는 문제도 존재한다. 이러한 품질 문제를 해결하기 위해 본 논문에서는 쾌적최대음량(CML)과 절대음량 체계를 도입하고 음원 손실을 최소화한 음량평준화 시스템을 제안한다. 그리고 일반적인 음원 파일뿐만 아니라 실시간 스트리밍 콘텐츠의 음원을 음량정규화할 수 있도록 설계하여 기존 상용 기술의 대상 제한 문제를 완화하고자 한다. 본 논문은 청자 혹은 서비스 이용자의 만족스러운 음원 감상에 기여한다는 점에서 의의를 가진다.

## 1. 서론

음량 정규화는 이미 다양한 분야에 폭넓게 사용되어 왔다. 사용자의 편의성을 높이기 위해 디지털 방송, 디지털 보청기 등에 적용되었으며 변화하는 시장에 발맞춰 점차 그 영역을 확장하고 있다. 2020년 12월, 국내 음향전문기업 ‘가우디오랩’에서 발표한 음량 정규화 기술이 국내 표준으로 승인받아 네이버 VOD 동영상 플랫폼과 여러 음악 스트리밍 플랫폼에 적용되었다. 이를 기점으로 사람들이 자주 접하는 각종 서비스에 음량 정규화 기술이 스며들었고 청각 편의성이 전체적으로 향상되었다. 그러나 여전히 기술적 결함이 존재한다. 음량을 정규화하는 과정에서 음원의 손실이 필연적이기 때문이다.

따라서 본 논문의 연구를 통해 앞서 설명한 문제점을 보완하게 된다면 청각적 스트레스를 겪는 사용자의 불편을 해소하는 동시에 음원의 질을 최대한 보존할 수 있는 하나의 방도가 될 것이라고 기대한다.

## 2. 관련연구

네이버 주식회사와 가우디오랩 주식회사에서 2019

년에 출원한 특허 “라우드니스 정규화 방법 및 시스템”은 서버 혹은 외부서버에서 수신한 콘텐츠로부터 정보를 추출하여 플레이어의 음량 출력 레벨이 조정되도록 하는 방법을 제시했다.[1] 기존의 트랜스코딩할 때에 음량을 조정하는 방식이 아닌, 클라이언트의 콘텐츠 재생 준비 과정에서 플레이어에 대해 음량 레벨을 최적화함으로써 다이내믹 밴드의 훼손을 막는 것에 의의가 있다.

“절대음량을 이용한 음량제어 체계의 개념”은 출력 기기마다 제각기 다른 음량 단위를 가지고 있다는 문제의 해결을 위해 60 dB(S)라는 동일한 단위를 적용하여 음량을 평준화시키는 절대음량 기술을 제안했다.[2] 해당 시스템의 특징은 청자가 듣고자 하는 음량 레벨인 목표음량과 제작자의 의도로 감소하는 음량의 정도인 차등감소수치에 따라 실제 출력되는 절대음량이 결정되는 것이다.

“딤러닝을 이용한 오디오 콘텐츠 분석 기반의 자동 음량 제어 기술 개발”은 오디오 콘텐츠 분석 모듈을 개발하기 위해 Database를 구축하였다.[3] 출력 결과는 무음, 음성, 음성/배경음악 혼합, 배경음악의 총 4개로 정의하였다. 영화, 뉴스, 스포츠 중계 장르 등의 방송 콘텐츠 음원을 직접 청취하고 데이

터라벨링하여 학습 데이터셋을 만들었다.

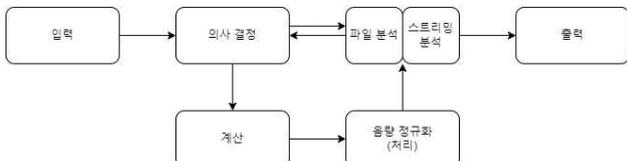
“사용자 컨텍스트 기반 음량 제어 기술”은 주변 소음 크기, 사용자 심박수, 주변 조도와 같은 환경 정보를 기반으로 하여 음량 최적화 알고리즘을 구성하였다.[4] “다자간 음성통화 품질 향상을 위한 오디오 믹서 알고리즘”은 잡음에 강인한 음성구간검출기를 통해서 연속되는 음성구간단위로 개인값을 구하여 적용함으로써 안정적인 음량을 유지할 수 있다는 특징이 있다.[5] “다중 디바이스를 지원하는 스마트 미디어 커넥터 개발에 관한 연구”는 스마트폰의 오디오/비디오 데이터를 TV나 오디오 기기들로 전송 및 재생하기 위하여 미디어커넥터 시스템을 구현하였다.[6]

이처럼 본 논문과 유사한 연구가 진행되어 왔으나, 청각의 보호를 위한 음량 레벨 조정을 제공하는 동시에 만족스러운 감상을 위해 기존 음원의 의도를 해치지 않을 정도의 품질 보장이 되는 것에 초점을 맞춘 연구는 찾아보기 어렵다. 또한 음원을 제공할 수 있는 경로가 다양해짐에 따라 스트리밍 서비스에 적용 가능한 음량정규화 기술도 필요하다.

### 3. 제안방식

#### 3.1 시스템 구조

시작 단계에서 음원 파일, 스트리밍 URL, 사용자 설정 음량 등을 입력받는다. 음량 정규화 모듈을 통한 음량 증감 조절은 global gain field에 계산된 gain 값을 곱하는 방식이다. 이에 필요한 데이터를 얻기 위해 파일 분석 및 스트리밍 분석 모듈에서 음원을 디코딩하여 음량 관련 값을 추출한다. 정규화가 완료되었다고 판단되면 정규화된 음원 파일 혹은 정규화된 스트리밍 음원 파일을 출력한다.



(그림 1) 제안방식 시스템 구조도

#### 3.2 흐름도

쾌적최대음량(CML: Comfortable Maximum Level)이란 일반적으로 청자가 어떠한 음원을 적극적으로 감상할 때 청각적으로 안전하며 불쾌감이 없는 범위의 최대음량을 의미한다. 차등감소태그란 절대음량 체제의 절대음량태그를 구성하는 요소 중 6번째 바이

트에 기록되는 차등감소수치를 의미한다.[7]

제작 과정에서 음원의 제작자에 의해 기록되는 해당 차등감소태그의 도입을 통해 음원 손실에 대한 개선을 기대할 수 있다. 청자의 입장에서는 다소 주관적인 음원 손실의 기준을 제작자의 의도에 맞추으로써 쾌적최대음량으로 음원의 음량을 조절할 때 음원의 품질이 손실되는 것을 방지하고자 한다.

실시간 스트리밍 서비스를 음량정규화하는 경우, 음량 메타데이터가 존재하지 않을 때의 딜레이를 최소화하기 위해서 임의로 선행 출력하는 과정을 도입하였다.

### 4. 결론 및 기대효과

본 논문에서는 청자 중심의 음원 손실을 최소화한 음량정규화 기술을 제안하였다. 일반적인 음원의 형태인 로컬 저장소의 파일뿐만 아니라 실시간 스트리밍 형태의 음원도 듣기 편안한 음량으로 정규화할 수 있도록 하고, 이를 실행하기 위한 각종 기준을 도입하였다. 이미 상용화된 음향 정규화 기술을 살펴보면 음원의 품질까지 고려하여 정규화시키는 기술은 대부분 그 범위가 특정 플랫폼 내부로 매우 한정적이기 때문에, 본 논문에서 제시한 음량정규화 시스템이 음원 콘텐츠 소비자에게 필요한 일률적인 음량정규화 기능을 제공할 수 있다.

#### 참고문헌

[1] 네이버 주식회사, 가우디오랩 주식회사, 임채린, 이지현, "라우드니스 정규화 방법 및 시스템", 2019.  
 [2] 견두현, 배명진, "절대음량을 이용한 음량제어 체계의 개념", 한국음향학회지 제33권 제1호, 2014.  
 [3] 이영한, 조충상, 김제우, "딥러닝을 이용한 오디오 콘텐츠 분석 기반의 자동 음량 제어 기술 개발", 한국방송·미디어공학회 2018 추계학술대회, 2018.  
 [4] 소재남, 권태익, 주상현, "사용자 컨텍스트 기반 음량 제어 기술", 대한전기학회 CICS' 17 정보 및 제어 학술대회, 2017.  
 [5] 류상현, 김형국, "다자간 음성통화 품질 향상을 위한 오디오 믹서 알고리즘", 한국음향학회지 제32권 제6호, 2013.  
 [6] 박호균, "다중 디바이스를 지원하는 스마트 미디어 커넥터 개발에 관한 연구", 융복합지식학회논문지 제6권 제1호, 2018.  
 [7] 견두현, 배명진, "절대음량을 이용한 음량제어 체계의 개념", 한국음향학회지 제33권 제1호, 2014.