

# Implementation of Audio Equalization in Video-on-Demand Broadcast Content

Myung-Kyu Kwon\*

## Abstract

In this paper, we develop the system for audio volume equalization of video on demand(VoD) content and propose the solution for it. In recent years, there has been a steady increase in the number of VoD users in addition to linear channels. However, viewers ought to sit in an uncomfortable way, adjusting the volume intermittently while they are broadcasted. Sudden changes of volume occur between the broadcasting channels, the programs from the co-channel, or the linear channels and the VoDs. Especially, upsurged dissatisfaction from the viewers has been found due to the unequalized volume when shifting between the linear channel and the VoD. In order to solve this problem, multilateral efforts were put forth, such as a system for keeping the volume at a certain level in digital broadcasting program has been legislated domestically. It leads success in equalizing linear channel volume. On contrary, too little notice has been taken for distorted volume problem of video on demand(VoD) content. In this paper, we developed and applied the volume equalization system into VoD content to achieve uniformization, a similar condition with linear channel(-24LKFS). This suggestion helped uneven current of volume which was in the stage -16 ~ -20LKFS to stable condition by lowering into the stage of -24LKFS. It also brought 20% increase in perspective of volume quality satisfaction level.

▶Keyword: Audio Loudness Control, Video on Demand, LKFS, Digital Broadcasting, Audio Level

## I. Introduction

사람들은 일상생활을 하면서 다양한 환경에 놓이며 다양한 소리에 노출된다. 사람들에 노출되는 소리는 다양한 원인에 의해서 발생하는데 사람이 청취 하였을 때 불쾌함을 발생하는 환경 소음, 사람을 즐겁게 하는 멀티미디어 사운드, 그리고 사람 사이에서 대화 및 정보를 주고 받을 때 발생하는 소리 등이 있다. 우리는 방송 시청중 오디오가 균일하지 못할때 음량을 조정하거나 사업자에게 불만의 목소리를 높인다. 방송 채널간, 실시간채널과 주문형비디오(VoD)간 또는 VoD 콘텐츠 사이에 이동할 때 음량 균일화가 되지 못해 음량 변화로 놀라는일도 있다. 동일 채널내 방송 프로그램 사이의 급격한 음량 변화로 방송 시청 도중에 수시로 음량을 조정해야 하는 시청자의 불편이 계속되고 있다[1].

국제전기통신연합(ITU)은 2010년 3월 디지털 방송의 음량 문제를 개선하기 위해 국제 디지털 방송프로그램 교환시 음량

기준값을 -24LKFS(Loudness, K-weighted, relative to Full Scale)로 할 것을 권고한 바 있고, 유럽방송연합(EBU)도 2011년 2월 오디오 프로그램의 음량 레벨을 -23LKFS(오차범위  $\pm 1\text{dB}$ )로 설정할 것을 권고(EBU Recommendation R128)한 상황이다. 국내는 방송법에 디지털방송프로그램의 표준 음량 기준을 정하여 고시하도록 명시하였고, 이에 따라 2014년 11월 28일에는 “디지털텔레비전 방송프로그램 음량 등에 관한 기준”이 제정 고시 되었으며, 이후 18개월의 유예기간을 두고 2016년 5월 29일부터 시행되었다[2]. 2016년 5월 29일부터 시행된 해당 업무지침은 예규로 제정 되었다. 해당 예규에는 방송프로그램 음량관리 및 적합여부 측정방법과 해당 방송법령 위반자 행정처분 및 과태료 부과절차 등이 명시되어 있다[3]. 라우드니스란 실제로 인간이 느끼는 음량의 크기를 의미하

\*First Author : Myung-Kyu Kwon, Corresponding Author : Myung-Kyu Kwon  
\*Myung-Kyu Kwon (mgkwon@sk.com), Media operation center, SK Broadband  
• Received: 2017. 09. 25, Revised: 2017. 10. 10, Accepted: 2017. 10. 25.

며, 인간의 귀는 낮은 음압에서 중음보다 저음과 고음에 상대적으로 둔감하게 반응하며 음압에 따라 주파수 반응이 틀려진다. 따라서 라우드니스를 측정하기 위해서는 이러한 인간 귀의 반응 특성을 보정할 수 있어야 한다. 사람마다 느끼는 오디오 음량은 매우 주관적으로, 특히 저음과 고음에 대한 부족감을 보충하기 위해 Loudness Control이 필요 하다. ITU-R BS.1770에 기반한 타겟 라우드니스 레벨과 허용 최대 트루피크 값을 설정하고, 라우드니스 레인지(Loudness Range) 컨트롤을 통하여 시청자가 프로그램 채널간에 오디오 음량 조정 없이 시청할 수 있도록 한다. 이를 달성하기 위해 솔루션으로 지정된 Target Value로 음량 자동 조정 할 수 있다[4]. 음량 조정하는 구성을 Fig. 1.에 표시 하였다.

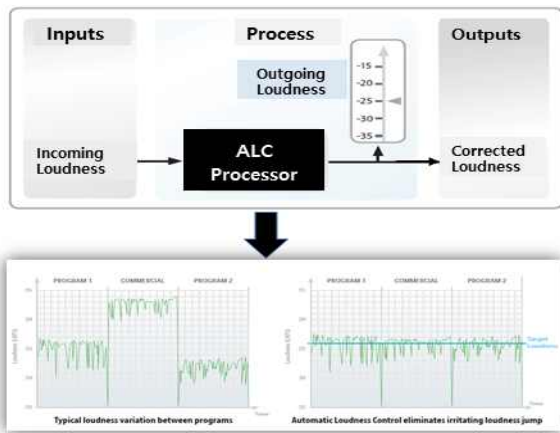


Fig. 1. Example of volume adjustment

디지털방송으로 전환되면서 음량 불균형이 심화된 기술적 배경에는 소리의 다이내믹 영역이 관여한다. 통상 다이내믹 영역은 피크레벨과 잡음레벨 사이(-65dB ~ +25dB)에 존재하는 구간이며 dB 단위로 측정 가능하다. 참고로 음악의 다이내믹 영역이 깊을수록 보다 섬세하고 풍부한 소리의 전달이 가능해진다. 다이내믹 영역을 Fig. 2.에 표시 하였다[5].

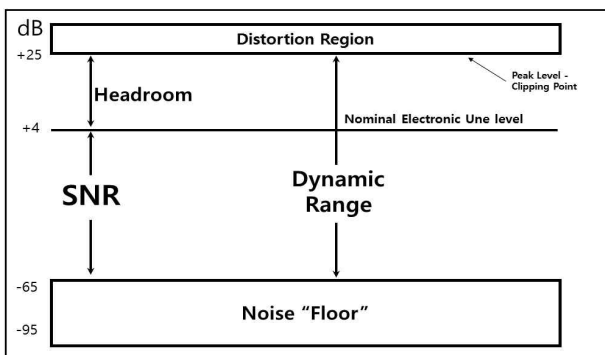


Fig. 2. Dynamic Range

한편 다이내믹 영역은 소리의 재생이나 저장 방식과 매체에 따라서도 차이가 있으며, 구형 포노그래프는 약 30~40dB, 카

세트테이프는 약 50dB 정도이다.

2011년도에 국내 지상파 방송사들을 대상으로 음량 레벨 측정이 이루어졌으며, 이때 조사된 4개 실시간채널 누적 음량 레벨 평균치는 -17.4 LKFS로 ITU-R 국제기준치 보다 평균 6.6 dB 높은 것으로 조사된 바 있다. 이때 가장 높은 채널이 -16.7 LKFS로 국제 기준치 대비 7.3dB, 가장 낮은 채널은 -18.3 LKFS로 5.7dB가 높았다. 누적 평균과 함께 음의 최고치를 나타내는 True Peak치의 경우 역시 국제 표준인 -1dBTP 보다 평균 1.5dB 높게 방송이 되고 있어 음량 표준의 적용이 요구되는 상황이었다[6]. 이러한 음량문제는 법 시행이후 방송사를 중심으로 개선되었다. 음량 기준을 반영한 방송법과 기술기준 등이 본격적으로 시행되는 2016년 5월 29일 이후에는 현저하게 개선이 된 것으로 분석되었다. 이후 측정 결과를 보면 법 기준치인 -24dB ±2dB를 만족함을 알 수 있다. 특히 기준 적용 이전에는 -13.8dB로 기준치보다 무려 10dB 높은 음량레벨로 방송을 하던 채널이 24dB로 측정되어 지상파 방송사들의 음량 기준 준수가 잘 이루어지고 있다고 판단된다[6]. 실시간채널의 음량 균일화는 개선되고 있지만 VoD, 방송광고 콘텐츠의 음량 균일화도 필수적이다. 최근 방송 시청형태가 실시간과 VoD, 그리고 광고 등 채널간 빈번한 이동이 이루어진다. 따라서 본 논문에서는 VoD 및 광고 콘텐츠에 대해 음량을 실시간채널과 같이 -24LKFS로 유지하는 시스템을 개발하고 적용하여 텔레비전 시청시 제기된 음량 문제를 개선하고자 한다.

구성은 2장에서 관련연구와 일반기술들에 대해 기술하고, 3장에서는 Audio Loudness Control 시스템개발에 대해 기술하고 4장에서는 실험과 결과분석 그리고 5장에서는 결론과 향후 계획에 대해 설명한다.

## II. Related works

### 1. Loudness measurement

소리를 평가하는 방법 중 사운드 크기(loudness)는 어떠한 소리가 사람 귀에 전달되었을 때 사람의 청각시스템이 인지하는 주관적인 사운드 크기이며, 소리의 강도(intensity)는 사람의 청각시스템에 전달된 객관적인 사운드의 강도로서 소리의 파워를 뜻하며, 일반적으로 잘 알려진 데시벨로 측정된다. 일반적으로 사람들 사이의 대화는 60 ~ 70dB이고, 교통량이 많고 소음이 심한 대로변은 약 80dB 정도이며, 일반적으로 약 70dB 범위에서 사람들이 편안함을 느낀다.

ITU(국제전기통신연합)에서 라우드니스를 측정하기 위해 ITU-R BS.1770이라는 ITU 표준규격을 설정하여 dB LKFS 또는 dB LU 단위로 라우드니스를 객관적으로 측정할 수 있게 되었다. 라우드니스 측정은 단순히 오디오 레벨을 측정하는 것과는 명확히 다르다. 라우드니스는 인간의 인지와 관련되어 있어 측정이 전압 또는 전류와 같이 객관적이기보다는 주관적이

기 때문이다. 인간이 듣는 구조는 각각의 사람마다 동일하지만 인지하는 것은 매우 다를 수 있다.

ITU(International Telecommunications Union)에서는 객관적으로 라우드니스를 측정하는 방법에 대한 해답을 위해 조사하는 연구그룹을 만들었고, 그 결과로 ITU-R BS.1770 이라는 규격을 정의 했다. 라우드니스는 콘텐츠의 길이에 걸쳐 모든 채널에서의 오디오 신호의 가중치가 적용된 파워를 누적하여 측정된다. BS.1770 방법은 모노, 스테레오 및 멀티채널 프로그램의 상대적인 주관적 라우드니스를 비교하는 리스닝 테스트를 통하여 확인되었다. 측정된 음의 크기는 LKFS로 표시 된다. LKFS의 단위는 데시벨과 같은 단위이다. -15LKFS의 프로그램을 9dB만큼 낮추는 것은(attenuating) 뉴스앵커 부분(일반적으로 대사가 콘텐츠의 전반적인 라우드니스를 대표하기 때문에)의 라우드니스가 -24LKFS인 뉴스 프로그램과 라우드니스를 같게 만들 수 있다. 뉴스 앵커 부분(대사레벨) 라우드니스의 정확한 측정은 작업자로 하여금 시청자에게 일정한 라우드니스 레벨로 콘텐츠를 제공할 수 있도록 하는 데 있어서 필수적인 것이다. 방송콘텐츠 멀티채널(5.1채널)의 주관적인 측정예를 Fig. 3.에 표시 하였다.

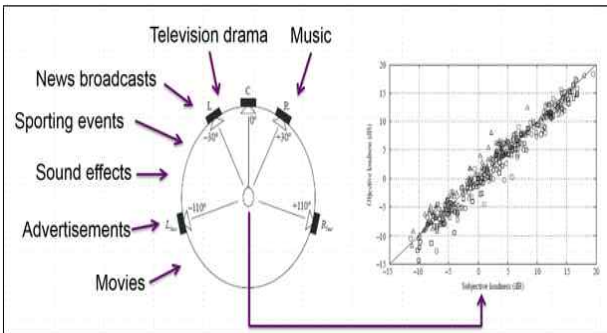


Fig. 3. A subjective measure : ITU-R BS.1770

ITU에서 라우드니스 측정을 위한 ITU-R BS.1770의 표준 규격을 설정 하였으며 dB LKFS / dB LU단위로 라우드니스를 측정한다. Music, Drama, News, Sports event, Sound effects, 광고 및 Movie를 포함한 평균 결과값을 라우드니스 측정 알고리즘의 시퀀스로 대입하여 측정한다.

Fig. 4.는 사람이 같은 크기로 느끼는 음의 음압 레벨과 주파수의 관계를 표시한 것으로, 같은 음압의 소리라도 주파수에 따라 크기를 다르게 인식하고 저주파 대역의 감도가 많이 떨어지는 것을 알 수 있다. ITU-R 에서는 Fig. 4.와 같은 사람의 청각 특성을 고려하여 사람이 실제 느끼는 소리의 크기를 정확하게 측정할 수 있는 방법을 제시 하였다. 또 LKFS라는 측정 단위로 방송 프로그램의 음량을 측정하도록 하는 권고 문서와 함께 국제간 교환 되는 방송 프로그램의 음량 레벨을 -24LKFS로 할 것을 권고하는 문서를 발표하였다. 이에 따라, 주요 국가들도 ITU-R 권고를 기초로 한 규정을 제정하여 방송 음량 규제를 시행하고 있다.

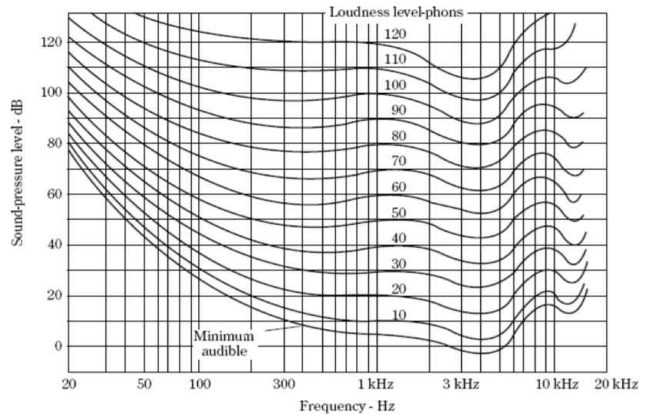


Fig. 4. Equal Loudness Contours

Fig. 4.의 등라우드니스 국제표준규격(ISO 226:2003)은 18 ~30세의 정상 청각 보유자에 대해 적용할 수 있는 그래프를 제시한다. 이 규격의 데이터는 자유 음장(Free Sound Field)에서 청음자에 대해 전방 입사되는 지속적인 순음을 양쪽 귀를 이용해서 얻은(Binaural Listening) 결과를 나타낸다[7].

Fig. 5.는 멀티채널 라우드니스 측정 알고리즘을 규정하며 알고리즘의 다양한 구성 요소에 대한 블록다이어그램을 보여준다. 라벨은 신호흐름 경로를 따라 서로 다른 지점에서 제공되며 이는 알고리즘의 설명을 돕기 위해 사용된다. 블록 다이어그램은 5개의 메인 채널(왼쪽, 중앙, 오른쪽, 왼쪽 서라운드 및 오른쪽 서라운드)이 입력되는 것을 보여준다. 이를 통해 1개에서 5개 채널을 포함하는 프로그램의 모니터링이 가능하다. 채널이 5개 미만인 프로그램의 경우 일부 입력은 사용되지 않는다. 저주파수 효과채널(Low Frequency Effect)은 측정에 포함되지 않는다.

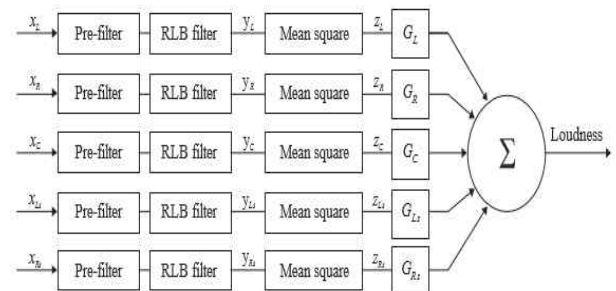


Fig. 5. Block diagram of multichannel loudness algorithm[4]

각 채널에는 Fig. 6.과 같은 구조의 'K' frequency weighting 2단(2nd Order) 필터가 포함된다. 상기의 알고리즘 및 측정방법에 의해 묵음 구간을 포함한 음량(Ungated Loudness)은 다음 식(1)에 의해서, 묵음 구간을 제외한 음량(Gated Loudness)은 식(2)에 의해 계측한다.

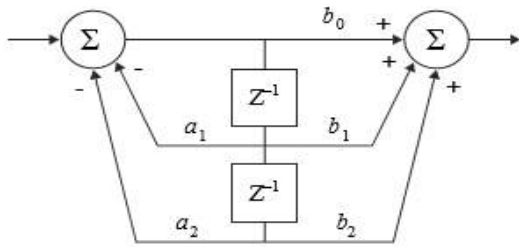


Fig. 6. Signal flow diagram as a 2nd order filter

$$\text{Loudness, } L_K = -0.691 + 10 \log_{10} \sum_i G_i \cdot z_i \quad \text{LKFS} \quad (1)$$

$$\text{Gated loudness, } L_{KG} = -0.691 + 10 \log_{10} \sum_i G_i \cdot \left( \frac{1}{J_g} \cdot \sum_{j \in J_g} z_{ij} \right) \text{LKFS} \quad (2)$$

- $G_i$  : 각 채널의 가중치 계수
- $Z_i$  : 주기 T 동안 필터에 입력되는 mean square 값
- $J_g$  : 게이팅 문턱치  $\Gamma$  에 대한 게이팅 블록
- $J_g = \{j : |z_j| > \Gamma\}$ ,  $J_g$  값은  $|J_g|$ .

방송음량을 측정하기 위한 시스템은 상기의 구성과 알고리즘을 적용하여 측정한다. 그리고 ITU-R BS. 1770 기준에 따라 묵음구간을 포함하는 음량과 묵음구간을 제외하는 구간의 음량을 측정하여 전체 음량을 평균하게 된다.

ATSC Document A/85:2009에는 디지털 TV를 위한 오디오 라우드니스를 정의하고 유지하기 위한 기술을 Fig. 7.에 기술하였다 [8]. “컴포트 존(Comfort Zone)”은 시청자가 받아들일 수 있는 정도의 프로그램 또는 프로그램내의 라우드니스 변화에 대한 가청 범위이다. “음량 미조정 존(Volume Unchanged Zone)”은 시청자가 프로그램값 또는 프로그램 내의 라우드니스 변화에 대해 음량조정 없이 들을 수 있는 라우드니스 범위이다.

이 범위 및 다른 라우드니스 허용범위를 정의하기 위해 주관적인 청취 테스트 실험이 행해졌으며, 이 결과로 Fig. 7.과 같은 기준점이 도출 되었다.

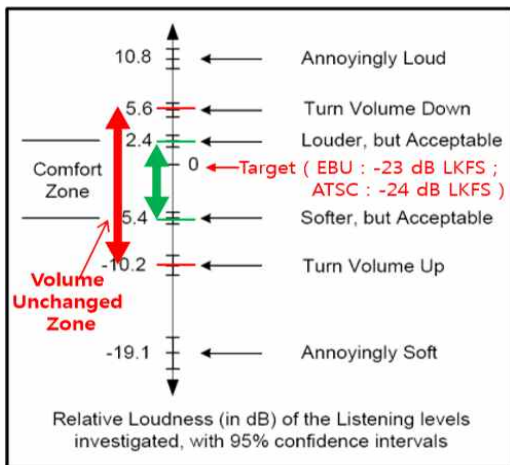


Fig. 7. Critical loudness level

현대 디지털 오디오 시스템은 오디오 신호의 처리와 분배를 매우 단순화시켰다. 이러한 신호의 피크 미터링은 일반적으로 측정기값의 최대 샘플 값을 표시하는 형태이다. 이러한 피크 샘플값에 대한 좁은 관점은 연속적인 과정에 기초해야 한다는 것을 간과하기 쉽다. 이는 예기치 못한 오디오의 오버로드, 부정확한 피크 리딩(peak reading) 및 다른 드러나지 않은 문제를 야기할 수 있다. 트루 피크(True-peak) 레벨 측정은 이러한 문제들을 미연에 방지할 수 있는 보다 정확한 오디오 신호의 변화를 확인할 수 있게 한다.

## 2. Audio Loudness Technology

### 2.1 Loudness improvement system

음량 균일화를 위한 HeadEnd 시스템을 Fig. 8.에 표시 하였다. 채널시스템의 기능 블록별로 기술하였는데 입력스트림에 대해 오디오신호를 별도 분리하여 오디오 신호를 균일화 한다. 이때 나머지 신호는 오디오 신호 균일화 하는동안 만큼 지연시켜 오디오 신호와 Remux한다. 멀티채널에 대해서도 채널별로 상기와 같은 방식으로 오디오 음량을 실시간으로 균일화 한다.

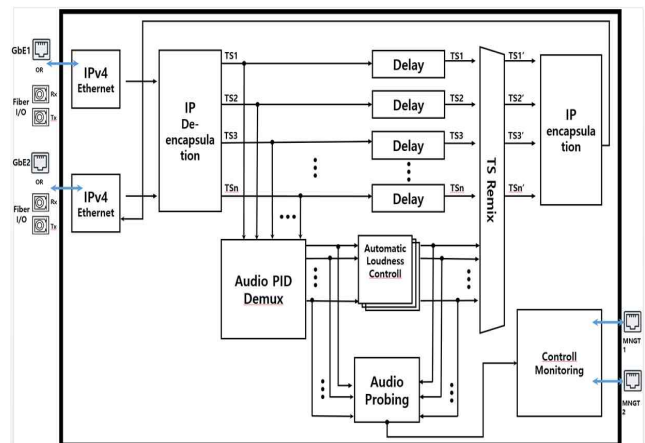


Fig. 8. Function block diagram of channel system[9]

- ▶ 시스템의 주요기능은 다음과 같다.
  - ITU-R BS.1770-1의 국제표준으로 디자인 하여 100개의 Audio가 지원 된다.
  - 자동적으로 Audio에 문제가 발생할 경우 경보를 띄운다.
  - 그래픽을 이용한 Log와 분석 내용을 확인한다.
  - Input대비 Output 지연율을 최소화(400 msec) 한다.
- ▶ 지원 Format은 아래와 같다.
  - Transport : MPEG over IP, UDP, RTP
  - Multicast : Supports IPV4 multicast
  - Audio Codec : Dolby Digital AC-3 (all mode)
  - Video Codec : Supports all video codecs

### 2.2 Flow of loudness equalization

Fig. 9.는 오디오 신호의 크기 제어에 따른 음질 열화를 보상



하기 위한 방법에 대한 예를 나타내는 그림이다. 오디오 신호 크기를 컨트롤하는 방법은 전술한 바와 같이 입력되는 데이터의 조건에 따라 다양하게 이루어 질 수 있다. 다만 오디오 신호 크기를 Target LKFS와 오차범위에 맞추게 되면 오디오 신호의 구성이 평평한 느낌이 강해질 수 있다. 이는 오디오 신호 크기의 균일화에 따른 역효과이다. 오디오 신호 크기 균일화의 목적을 달성하면서 균일화에 의한 역효과를 해결해야 오디오 균일화의 과급력과 사용자의 만족도가 향상 될 수 있다. 이에 따라 역효과를 보상하기 위한 청각적 열화 보상 모듈을 더 구비할 수 있다. 즉 Fig. 9.를 참조하면 디멀티플렉서(demultiplexer)는 기존에 녹화된 방송 프로그램 데이터 또는 라이브 방송 프로그램 데이터를 디믹스(demux)하여 오디오 데이터를 선택할 수 있다.

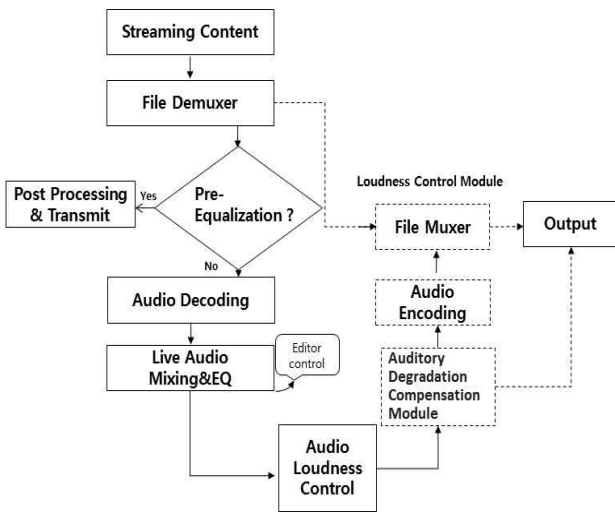


Fig. 9. compensating the sound quality deterioration[10]

그리고, 균일화 판단부는 오디오 데이터가 사전 균일화 되었는지 판단할 수 있다. 만약, 오디오 데이터에 사전 균일화가 수행되었다면 균일화가 수행된 오디오 데이터에 대한 이후 절차를 진행할 수 있다. 만약 오디오 데이터에 사전 균일화가 수행되지 않았다면 오디오 디코더는 오디오 데이터를 디코딩할 수 있다. 그리고, Live Audio Mixing & EQ와 같은 편집자 컨트롤이 수행될 수 있다. 그리고, 오디오 신호 크기 컨트롤러는 디코딩된 오디오 데이터를 이용하여 오디오 신호 크기의 균일화를 수행할 수 있다. 또한 청각적 열화 보상 모듈은 오디오 신호 크기 컨트롤러에서 수행된 균일화에 따른 역효과를 보상할 수 있다. 그리고, 멀티플렉서는 인코딩된 오디오 데이터를 디멀티플렉서에서 선택되지 않은 다른 데이터와 멀티플렉싱할 수 있다.

한편, 전술한 방법에 따른 오디오 신호 크기 균일화는 디지털 방송 소비자에게 상당한 청취 환경의 변화를 불러 일으킨다. 또한 오디오 신호 크기가 균일화 됨으로써 디지털 방송 단말기에서 새로 요구되는 서비스, 기능들이 생성될 수 있다. 즉 디지털 방송 단말기에서 방송 오디오 음량 관련 기능들을 제공할 수 있다. 방송 음량 기준을 도입하고 방송 음량에 효과적으로 적용하기 위해

서는 방송 프로그램의 제작과 송출 모두에 음량 측정과 균일화 가능 시스템을 도입하는 것이 바람직하다. 그러나 현행 기술기준 체계로는 방송 프로그램을 제작 공급하는 프로덕션까지 규제할 수는 없다. 또한 기술기준의 효력이 발효되기 이전에 제작된 프로그램들이 방송될 경우 음량 제어를 하기 위해서는 방송 송출 단계는 적용이 되어야만 한다. 또한 생방송이 아니고 프로덕션이나 방송국에서 프로그램의 사전 제작을 하는 경우는 전체 프로그램을 대상으로 음량 균일화 전에 음량 레벨을 측정하여 해당 프로그램에 대한 균일화를 수행하는 소프트웨어 기반의 솔루션들이 적용 가능하다. Fig. 10.은 사전 제작 프로그램들을 대상으로 한 음량 균일화의 예를 보여준다[11].

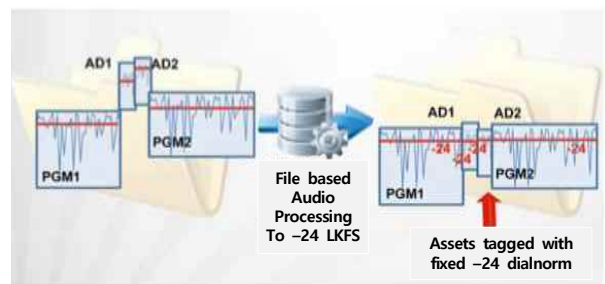


Fig. 10. Equalize volume for pre-production programs

Fig. 11.은 Infinite mode를 활용한 Integrated Loudness와 Short-term mode를 사용한 라우드니스 레인지(LRA) 영역을 보여준다. Infinite mode는 일반적으로 프로그램 전체 구간을 측정할 때 사용한다. 예를 들어 30초 광고 또는 2시간짜리 영화 등의 전체 구간에 적용하여 라우드니스를 계산한다. 이것은 프로그램 전체에 대한 평균 라우드니스를 계산할 수 있어 가장 정확한 라우드니스를 측정 할 수 있다.

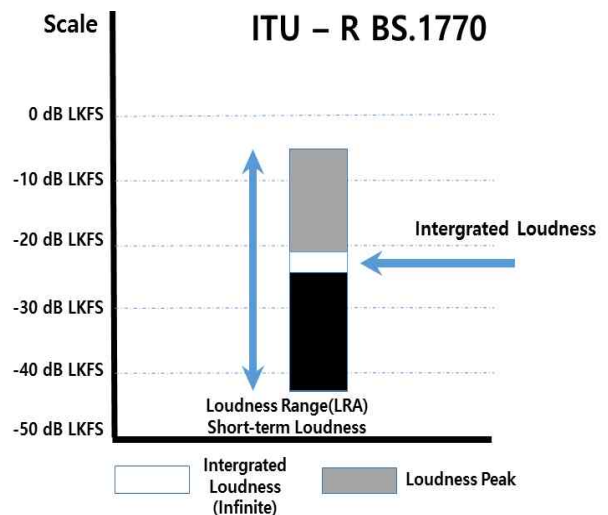


Fig. 11. Integrated Loudness & LRA

### 3. Standardization trend of loudness equalization

ITU-R에서는 2000년 9월부터 디지털 오디오 생성에 사용

될 수 있는 오디오 측정 특성에 대한 연구를 시작하였다. 2006년에 처음 발표된 권고안(ITU-R Rec. BS.1770)은 오디오 프로그램의 음량과 음성 신호의 실제 최고 레벨(true-peak audio level)을 측정할 수 있는 알고리즘을 포함하고 있다. ITU-R 음량 측정 방법은 모노, 스테레오 및 다채널 오디오 신호에 적용 가능하며, 음량 측정 단위로 LKFS(Loudness K-weighted relative to Full Scale)를 사용한다. 2007년에는 부가적인 설명이 추가되어 ITU-R Rec. BS.1770-1로 수정되었으며, 2010년 10월에는 오디오 신호가 긴 묵음 구간을 포함할 때 묵음 구간을 제외한 정확한 음량 측정을 할 수 있는 게이팅(Gating)방법이 추가된 ITU-R Rec. BS.1770-2 권고안이 발표되었다. 2012년 8월에는 실제 피크 오디오 레벨 측정 알고리즘을 수정한 ITU-R Rec. BS.1770-3이 발표되었다.

EBU(European Broadcasting Union)에서는 방송 프로그램의 제작, 배급, 송출에 있어서 오디오 신호 레벨의 필요성을 연구해왔다. 오디오 레벨링 패러다임이 라우드니스의 측정에 기반해야 할 필요가 있다는 의견에 의해서 수행되었다. EBU에서는 프로그램의 평균 라우드니스('프로그램 라우드니스')와 함께 "라우드니스 레인지"와 "최대 트루피크 레벨(True Peak Level)"을 측정하여 오디오 신호의 노말라이즈에 사용할 것을 권장하고 있다. 2010년 8월 오디오 음량의 측정 및 균일화를 위한 권고안인 "Loudness normalization and permitted maximum level of audio signals"이라는 명칭의 "EBU 권고 R128"을 발간하였다.

북미에서는 ATSC(Advanced Television System Committee) A/53 표준을 사용하여 음량을 측정하고 조정하였으나, 2006년 ITU-R에서 음량을 측정하기 위한 권고안이 발표되면서 음량 문제를 해결하기 위한 방법을 보완하기로 하였다. 이에 따라, ATSC는 2008년 부터 사용자가 듣기에 일관성이 있도록 하는 음량 레벨을 표시하는 방법에 대한 표준화 작업을 시작하여, 2009년 11월 ATSC Recommended Practice(RP) A/85 "Technique for Establishing and Maintaining Audio Loudness for Digital Television" 문서를 발간하였다.

Fig. 12.는 미국, 유럽 및 일본의 라우드니스 규제 현황을 보여준다.

ITU-R BS-1770			
	ATSC A/85	EBU R128	ARIB TR-B32
Country	Americas	Europe	Japan
Target	-24 LKFS	-23 LUFS	-24 LKFS
Short term	10 sec	3 sec	3 sec
Relative Gating	NA or -10 dB	-10 dB	-10 dB
Integrated Tol.	+/-2 dB	+/-1 LU	+/-3 dB
TP Max	-2 dB TP	-1 dB TP	-1 dB TP

Fig. 12. Specification for Loudness

현재 우리나라의 라우드니스 규격은 ITU-R BS.1770-3 기준으로 제정되어 활용중에 있다. ITU-R BS.1770-2가 기본적인

측정을 정의했다면 EBU R 128은 그것에 기반하고 확장시켰다. EBU R 128은 BS.1770-2에 라우드니스 레인지(Loudness Range)와 타겟 레벨: -23LUFS(풀스케일에 대한 라우드니스 단위)을 추가 했다.

### III. The Proposed Scheme

#### 1. Configuration of suggested system

소프트웨어기반으로 Audio Loudness Control시스템을 개발하였다. 콘텐츠 입고부터 인제스트, 인코딩 및 배포 등 전체 업무흐름을 Fig. 13.에 간략하게 구성하였다[12].

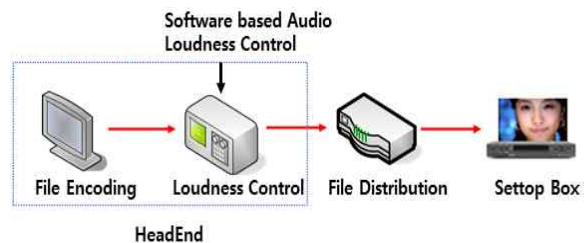


Fig. 13. Workflow scheme

Fig. 13. 을 보면 현장에서 취득된 오디오 데이터는 Ingest 서버에 저장되고, 저장된 파일은 종류에 따라 인제스트 및 인코딩 작업을 수행한다. 작업 흐름을 보면 콘텐츠 편성일 확정, 콘텐츠 계약, 수급, 콘텐츠 등록, 소재전달, 장르맵핑, 소재확인 및 수급완료 처리순으로 진행된다. 콘텐츠가 입고되면 채널번호맵핑, 인코딩 스케줄을 확정한다. 미디어형태, 해상도, 음향구분에 따라 자동으로 인코딩 및 음량 균일화, DRM 작업을 진행한다. 영상 시놉시스 등 영상을 검수하고 영상파일은 백업을 수행한다. 이 과정에서 소프트웨어기반의 라우드니스 제어 장치를 통과하게 되고 이때 모든 콘텐츠는 국제 규격인 -24LKFS 수준으로 작업이 수행된다.

개발된 라우드니스 제어 시스템의 오디오 음량 균일화 순서는 아래와 같다.

- 음량 측정모듈에서 원음을 받아서 해당 원음의 음량 측정
- 음량 제어모듈에서는 해당 원음의 음량과 목표음량을 비교
- 두 음량의 차이를 매핑 커브와 비교
- 매핑 커브를 기준으로 원음의 Gain값 결정
- Gain에 따라 원음 조절
- 상태 버퍼에 현재 음의 Gain값 저장 및 상태모델 업데이트
- Gain이 적용된 음 출력

특정기간 동안 낮은 주파수, 저주파 효과 채널을 제외하고 모든 채널의 가중 평균치를 측정한다. 측정 초당 10회의 소리크기를 측정(100ms)한다. 이후 400ms의 평균 소리 측정(Momentary)하고 K필터링을 통해 최종 음량 측정 계산한다.

최종 측정값은 LKFS단위를 사용한다.

소프트웨어기반으로 라우드니스 제어 시스템을 개발함으로써 하드웨어 장비는 신규로 도입 되어야 한다. 일반적으로 사용되는 하드웨어 시스템은 Table 1.과 같다. OS는 Linux나 Window 모두 가능하나 여기서는 Linux를 사용하였다.

Table 1. System Environments

Items	Audio Loudness Server
Operating System	Linux or Window 2008
CPU Clock speed	3.0GHz, Dual CPU
Memory Size	24GB

음량균일화 개발시스템의 기능흐름 및 모듈별 설명을 Table 2.에서 보여준다. Software based Audio Loudness Control 시스템의 기능 블록도이다.

Table 2. Functional Flow and Explanation of each module

Functional Flow	Explanation of each module
<div style="border: 1px solid black; padding: 5px; margin-bottom: 10px;">                     1. Measure level by segmentation                      - Segmentation                      - Measure level by segmentation                 </div>	<ul style="list-style-type: none"> <li>- Level measurement divided into two steps.</li> <li>- Segmentation is split based on silence between voice and voice.</li> <li>- Perform level measurement by intervals Momentary (400ms), Short-term (3second), Integrated LKFS.</li> <li>- Exclude LFE(Low Frequency Effect)channels. Apply weighted average.</li> <li>- Apply K-filter.</li> </ul>
▼	
<div style="border: 1px solid black; padding: 5px; margin-bottom: 10px;">                     2. Adjust the volume for each segmentation                 </div>	<ul style="list-style-type: none"> <li>- Calculate the gain between input volume and target volume. Based on Mapping Curve.</li> <li>- Adjust the volume for each segment to the target value of -24LKFS.</li> <li>- Adjustable in 100ms increments.</li> </ul>
▼	
<div style="border: 1px solid black; padding: 5px; margin-bottom: 10px;">                     3. Level measurement for the entire section                 </div>	<ul style="list-style-type: none"> <li>- Momentary, short-term, Integrated LKFS measurements for the entire content file.</li> <li>- Apply Gating.</li> </ul>
▼	
<div style="border: 1px solid black; padding: 5px;">                     4. Level adjustment for the entire section                 </div>	<ul style="list-style-type: none"> <li>- Adjust the Integrated LKFS value to the target value for the entire content file.</li> </ul>

소프트웨어개발 기반으로 오디오 라우드니스 시스템을 구성함에 따라 -24LKFS 국제 규격으로 오디오 균일화 하는 과정에 대한 테스트 프로그램 소스를 Fig. 14.에서 보여준다[13].

```
# pySoundFile install
# pyloudness install
# ffmpeg install

import soundfile
import pyloudness

x, fs = soundfile.read ("Guitar-16-44pl-stereo-72secs.wav")

loudness_stats = pyloudness.get_loudness ("Guitar-16-44pl-stereo-72secs.wav")

loudness = loudness_stats['Integrated Loudness'] ['T']
LRA = loudness_stats['Loudness Range']['LRA']
print("Loudness before normalization: %.1f LKFS"%loudness)

target = -24
gaindB = target - loudness;
gain = 10^(gaindB/20)
xn = x * gain

soundfile.write("Guitar_normalized.wav", xn, fs)
loudness_stats = pyloudness.get_loudness ("Guitar_normalized.wav")

loudness = loudness_stats['Integrated Loudness']['T']
print("Loudness before normalization: %.1f LKFS"%loudness)
```

Fig. 14. Procedure for Audio Loudness Control

실시간채널의 오디오 라우드니스 제어 시스템 경우는 실시간으로 범위를 벗어나면 균일화 시키는 구조이나, VoD의 경우 파일 전체 평균값이 목표값으로 나오도록 하는 구조이다. 따라서 어떤 부분에서는 조금 높이고, 어떤 부분에서는 조금 낮춰짐으로 해서 평균값을 산출한다. 배경과 대사에 음량 Gap에 의해 상대적으로 낮은 대사를 듣기 위해 오디오를 높이는 경우가 발생할 수 있다.

## IV. Experiment and analysis

### 1. System configuration and experiment

VoD 콘텐츠를 장르별 5개 그룹으로 구분하였다. 스포츠, 엔터테인먼트, 드라마, 영화 그리고 뉴스 그룹으로 분류하였다. 실험은 2단계로 나누어 진행 되는데 1단계는 측정 단계이고, 2단계는 목표값으로 변환하는 단계이다. 1단계는 국제 표준인 ITU-R BS.1770-3 표준으로 측정하며 다음과 같이 세분화 할 수 있다. 첫째, 원본 소스 콘텐츠 입력, 둘째, 원본소스를 두개의 필터를 통과하여 통과된 필터로부터 인간이 들을 수 있는

가청주파수 범위에서 에너지 분포에 가중치를 넣은 음원을 획득, 셋째, 획득된 음원을 LKFS 방식으로 계산, 넷째, 원음의 각 채널(스테레오면 2개, 5.1채널이면 6개)에 대한 LKFS값을 획득한다. 획득되는 LKFS는 400ms 음원으로부터 계산된 값이고, 이것은 100ms window단위로 업데이트 된다. 2단계는 변환단계로 1단계에서 측정된 음원값과 목표값 차이를 구한다. 이 값이 조정해야할 Gain에 해당한다. 이 값만큼을 조정한다. 그리고 Gain값이 반영된 음을 출력한다. 5개 그룹의 원음 측정결과와 음량은 누적평균 -16 ~ -20LKFS로 측정 되었다. 국제표준보다 4~8dB정도 높은 편이다. 측정결과 과형을 Fig. 15.에 표시 하였다. 일부 대사 소리가 작게 들리는 현상이 있으며 이것은 가청시 배경소리와 대사의 음량 Gap이 있기 때문이다.

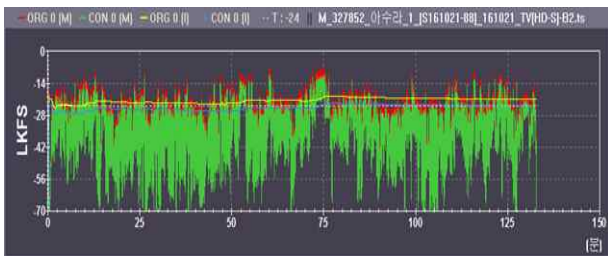


Fig. 15. Loudness analysis

Fig. 16.은 콘텐츠의 음량을 측정하고 측정값과 목표값의 차이를 보여준다. 측정값은 -21.46LKFS이며 목표값은 -24LKFS이다. 따라서 Gain은 3.5dB이며 이 차이값 만큼 조정하여야 한다[14].

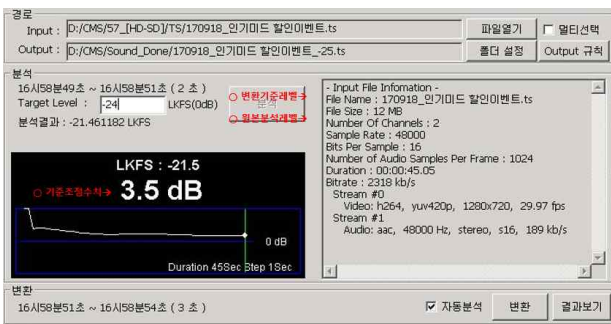


Fig. 16. Loudness Control UI

방송 VoD 콘텐츠에 대해 음량 균일화 작업(시험)을 Fig. 16.과 같이 수행하고 그 작업결과 로그를 Fig. 17.에 표시 하였다.

번호	종류	경로	ST	Duration	TargetLevel	Integrated	Channels	Container	Codec	Track count	시작(시)	종료(시)
ORIGINAL	본인 음원	J:\2014.12.24\...	6.058	02:17:13	-24.0	-18.4827	2	MPEGTS	H264/AAC	1	-	-
CONVERSION	변환 음원	C:\NewMed\N...	6.058	02:17:13	-24.0	-20.0000	2	MPEGTS	H264/AAC	1	2017-09-27 10:29:59	2017-09-27 10:29:59

Fig. 17. Conversion Test Log

## 2. Experimental results and analysis

### 2.1 Comparison of loudness equalization

VoD 콘텐츠의 라우드니스 측정값을 보면 누적평균 음량이 -16LKFS에서 -20LKFS범위에 분포한다[15]. 이것은 기준치인 -24LKFS 대비 무려 4~8dB 높아 기술기준에 부적합하다. VoD 콘텐츠의 경우 최근에 입고된 콘텐츠는 국제 규격에 맞게 조정 하지만 10년전에 도입된 콘텐츠는 분량도 많고 Hits수도 적어 음량 변경 작업에 어려움이 있다. 오래된 콘텐츠에 대해 자동으로 음량 변경할 수 있는 프로그램 개발을 통해 음량 균일화를 수행 하여야한다. VoD 콘텐츠에 대해 장르별로 음량 균일화 작업 전,후 비교를 Fig. 18.에서 보여준다. LKFS값은 누적 평균을 의미한다.

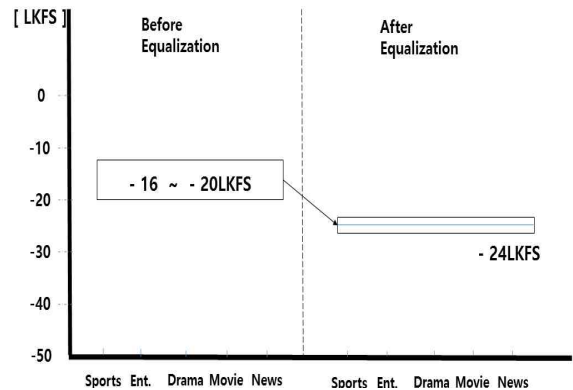


Fig. 18. Comparison before and after equalization

### 2.2 Satisfaction of loudness equalization

VoD 콘텐츠에 대해 음량 균일화 이전에는 고객들로부터 불만의 목소리가 들려왔다. 그러나 균일화 이후는 현저히 줄어들었으며 만족도 측면에서 20% 향상 되었음을 볼 수 있다. 이러한 만족도는 S社 방송품질 체험단 소속 구성원 및 Test요원 50명을 대상으로 8월 ~ 9월사이 3주간 설문조사를 수행 하였으며 척도는 5점 만점 기준이다. 특히 대사 음량이 크고, 대사가 많은 영화나 스포츠 장르에서 큰 효과를 볼 수 있었다. 만족도 결과를 Fig. 19.에 표시 하였다.

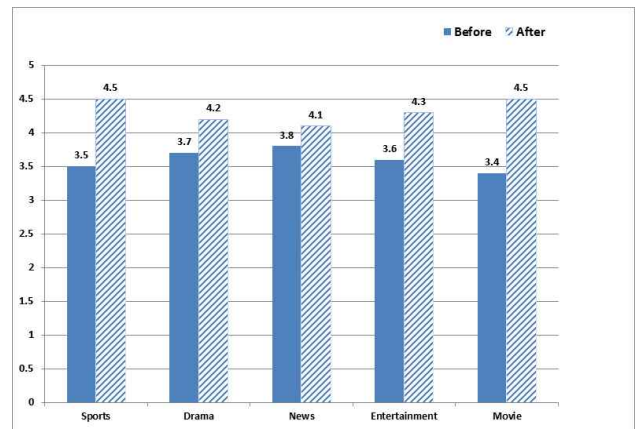


Fig. 19. Satisfaction comparison



## V. Conclusions

본 논문에서는 주문형비디오(VoD) 방송 콘텐츠의 음량 불균일화 문제를 분석하고 시스템 개발을 통해 문제를 해결하였다. 디지털 텔레비전 방송 음량 기술기준이 시행 되었다. 실시간 방송채널들은 법 규정에 따라 음량을 균일화하여 품질의 향상을 가져왔다. 그러나 VoD 콘텐츠나 광고 같은 경우 시청빈도가 높아지는 상황에서도 음량 균일화 작업이 미흡했다. 실시간 채널과 VoD간 이동이 동일 셋업에서 이루어지고 있으므로 상호 음량 균일화가 이루어져야 방송 콘텐츠를 즐길 수 있다. 실시간 채널의 경우 법적 규제가 있어 법 시행 전,후로 뚜렷한 차이를 보였다. 실시간 채널의 음량 분석 결과 법 시행 전에는 측정 대상인 모든 채널들이 기술기준보다 높은 음량 레벨로 방송을 하였으나, 시행 후는 측정 대상인 대부분의 방송 채널들이 기술기준에 적합한 음량 레벨을 유지하였다. 주문형 비디오 콘텐츠(VoD)에 대해 음량제어 시스템을 개발하여 적용하였다. 적용 콘텐츠 모두 국제기준인 -24LKFS에 부합 하였으며 콘텐츠 음량 균일화에 대한 만족도는 20% 향상을 가져 왔다. 특히 스포츠나 영화 콘텐츠의 경우 만족도가 약 30%이상 도출되어 타 장르에 비해 효과가 큼을 알 수 있다. 향후에는 오래되고 Hits수 적은 콘텐츠에 대해서도 자동으로 음량을 조정해 줄 수 있는 프로그램이 개발되어야 한다. 이를 통해 구작 VoD 콘텐츠도 음량 균일화가 이루어지고 실시간채널과 더불어 방송 콘텐츠를 즐길 수 있어야 한다.

## REFERENCES

- [1] O.S. Cho, J.Y. Jung, D.J. Choi, N.H. Hur, K.S. Oh, "Standardization and regulation of international broadcasting volume", Journal of Telecommunications Technology Association, Vol. 142, pp. 83-87, Aug. 2012.
- [2] "Standard volume of a digital broadcast program", Ministry of Science, ICT and Future Planning, Nov. 2014.
- [3] "Guideline for Digital Television Broadcasting Program Loudness Standards", Central Radio Management Station, May. 2016.
- [4] Rec. ITU-R BS.1770-1, "Algorithms to measure audio program loudness and true-peak audio level" ITU, 2006-2007.
- [5] S.O. Lee, "Stabled Television volume level", Broadcasting & Media, Vol. 21, No. 3, Jun. 2016.
- [6] S.O. Lee, O.S. Cho, J. k. Kim, "Study on Digital TV Broadcast Volume" The Journal of Korea Society of Communication and Sapce Technology, Vol. 80, No. 4, Dec. 2013.
- [7] D. H. Gejn. "Absolute Sound Level System", Ph. D. Dissertation, SoongSil University, Dec. 2012.
- [8] ATSC Document A/85:2009 "Comfort Zone"
- [9] Miranda, "Function Block Diagram of AXINO", Technical Report, May. 2015.
- [10] B. H. Choi, et al, "Audio signal size control method and device", US Patent WO 2014148848 A2, Intellectual Discovery Co., Ltd., Mar. 2013.
- [11] S. U. Lee, G. H. Nam, J. H. Na, J.Y. Han, "A Study On Digital Broadcasting Program Audio Level Regulation and Management System", The Report of Korea Society of Communication and Sapce Technology, NoV. 2013.
- [12] XEIN M&C, "Audio Loudness Control Technical Report", Jan. 2016.
- [13] EBUR128 Source Code, [https://github.com/jiixyj/libebur128/tree/master/ebur128\(ebur128.c\)](https://github.com/jiixyj/libebur128/tree/master/ebur128(ebur128.c))
- [14] NeoConvergence, "Neo Audio Loudness Control materials", Technical Report, Aug. 2016.
- [15] S. G. Park, "A Study on Real-Time Loudness Metering Algorithm for Digital Broadcasting", The Korean Institute of Electromagnetic Engineering and Science, Vol. 16, No. 4, Apr. 2005.

### Authors



Myung-Kyu Kwon received the B.S. degree in electric engineering from Kyungpook National University, Korea, in 1985 and Ph. D. degree in Convergence Engineering from Hoseo University, Korea, in 2017. He is currently working for SK Broadband.

He is interested in Distributed system, Digital content, ICT, Digital convergence, AI, Cloud computing, and Big data.