FUDMA Journal of Sciences (FJS)
ISSN online: 2616-1370
ISSN print: 2645 - 2944
Vol. 3 No. 4, December, 2019, pp 296 –304

EQUALIZATION OF BROADCAST LOUDNESS USING FEED-FORWARD LOUDNESS CONTROL AND ADAPTIVE PARAMETER CONFIGURATION

Tekanyi, A. M. S., Kaatpo, C. and Usman, A. D.

Department of Communications Engineering, Ahmadu Bello University, Zaria

Author's email: amtekanyi@abu.edu.ng kaatpo@gmail.com aliyuusman1@gmail.com

ABSTRACT

Loudness is a subjective measure of how piercing an audio signal is perceived. Due to commercial pressures loudness has been exploited in broadcasts to attract and reach viewers and listeners. Sudden changes in loudness between television channels and programs are significant causes of nuisance for the consumers. With the transition from analogue to digital TV, loudness related issues are on the rise hence, the need for objective and accurate techniques for digital broadcast. The measurement of perceived loudness is difficult yet an important task. Many research efforts have been introduced for objective measurement and equalization of the loudness of audio channel or program in order to facilitate program delivery. In this paper, broadcast contents were examined to discover where loudness occurs in a program and how to achieve equal average loudness between broadcast programs. A modified scheme was developed based on a standard set by ITU- Radio and European Broadcasting Union by incorporating a feed forward loudness control (FLC) mechanism and an adaptive parameter configuration (APC) scheme for the purpose of addressing loudness in long and short form contents.

Keywords: Loudness, Audio Signal, Long Form Content, Short Form Content, Audio/Loudness Normalization, Program Transition.

INTRODUCTION

Audio loudness is subjective and difficult to measure because the characteristic of sound is primarily a psychological correlation of physical strength (Lee *et al.*, 2014). Legislation for audio normalization standard for broadcast and radio became a concern for broadcast organizations such as European Broadcasting Union (EBU), International Telecommunications Union (ITU), and Advanced Television Standards Committee (ATSC). Its recommendation specifies design methods for audio level meters to ensure loudness standards are met and TV viewers experience a more even loudness of content across different programs and commercials (ITU-R, 2012). With the availability of these new standards, it is now feasible to control loudness objectively as consumers do not expect large changes in audio loudness from program to interstitials and from channel to channel (EBU R, 2014). This research work developed a modified adaptive loudness equalization scheme, which employed integrated descriptors (short and long form contents) to achieve equal average loudness between programs without affecting the dynamics of the contents. The adaptive parameter configuration is used to adjustably equalize content loudness while retaining the dynamic range during smooth transitions between programs with different loudness levels.

Due to commercial pressures, loudness has been exploited in broadcasts to attract and reach viewers and listeners. These commercials are termed short-form content or momentary as the case maybe depending on its duration. Many research works were carried out to control loudness of these short descriptors neglecting the long-form content descriptors such as programs where loudness is equally prone to exist. This necessitated the need to explore more efficient ways to address loudness in broadcasting by using FLC and APC scheme to address loudness in both short-form and long-form contents of broadcasting so as to achieve average equal loudness between programs.

SIMILAR WORKS

Robin & Chris, (2014) performed an analysis that was aimed at developers who validated digital signal processing (DSP) during development, as well as sound-engineers who mix and master according to commercial constraints. Results indicated that volume stability problem during broadcast was still inherent, which affected the quality of service of the received signal at the receiver. Paulus, (2015) proposed a method for estimating the change in the overall loudness using loudness information of the partial mixes that rendered its description. Results showed that the signal still suffered from volume instability during broadcast. This was because, different sound loudness were usually generated from different channels which usually resulted into noise or unpleasant sounds to the listener. Ward *et al*., (2015) presented a real time excitation based binaural loudness meters using the combination of a dynamic auditory loudness model and the IEEE ICA SSP, 2012 model. In the work, there was no consideration for either short or long term loudness as this could have enhanced the results obtained. Fenton *et al.*, (2016) evaluated a perceptual based model of 'punch' in musical material. This was done by combining signal separation, onset detection and low level parameter measurements which produced a perceptually weighted 'punch' score. By using scores derived through a forced pair wise comparison listening test, poor results could be obtained since one of the drawbacks of paired comparison tests with no ties allowed was circular error rates when sounds are perceived to be similar by listeners. Lee *et al*., (2016) proposed an Automatic Loudness Control (ALC) algorithm in order to decrease the audio degradation. The ALC was based on the program loudness budget. The proposed ALC can degrade the audio quality when the target length is within a short time. Ward & Reiss, (2016) highlighted state-of-the-art loudness algorithms and provided insights into their differences when applied to multi-track audio.

The results obtained indicated that for the multiband algorithms, peak loudness was more appropriate when the sound corpus involved transient instruments. However, averaging the loudness time series tend to underestimate salient peaks, thereby affecting the loudness quality of the multi-tract audio. Ward *et al*., (2017) proposed an estimation of the loudness balance of musical mixtures using source separation technique. Results obtained from the 100 semi- professionally mixed songs showed that the relative loudness was largely dependent on the source. However the technique applied in the work has no provision for live program and this will affect the overall performance of the system. Zimmer, (2017) proposed an approach to assess loudness and dynamic range with web audio native nodes. Pires*et al*., (2017) proposed a loudness controller for broadcasting which considered momentary and short-form segments detected using a Support Vector Machine (SVM) classifier. The classifier performance was not a problem but the model did not consider classifiers above 3s which was not sufficient enough to address loudness.

From the literatures reviewed, different approaches were used in trying to normalize loudness in broadcasting. All these approaches only addressed loudness to a certain level because the techniques either addressed Short Form Contents, past or present contents without considering the Future Content which is an integral part of live program broadcast. In this paper, FLC and APC schemes were used to address both short-form and long-form contents of broadcast by using different look-ahead time durations.

LOUDNESS EQUALIZATION USING FEED-FORWARD LOUDNESS CONTROL AND ADAPTIVE PARAMETER CONFIGURATION

The developed scheme for equalization of broadcast loudness using Feed-forward Loudness Control (FLC) and Adaptive Parameter Configuration (APC) was developed based on statistical analyses of the trained input data set. The data set was obtained from recordings of commercial broadcasts from different terrestrial channels, online streaming and from cable satellites comprised of both audio and audiovisuals. The audio files were trained using Adobe Audition and Cool Edit, ensuring that it contains the same format as when broadcasted. The extraction of programs was made by reading the WAV files into vectors in MATLAB. The beginning and end of a program were located by identifying Fade-outs and Fade-ins of the signal.

To equalize the loudness level of a signal, the signal gain is adjusted proportionately to the loudness because the scheme utilizes the fact that program loudness is linearly proportional to the signal's gain. The average program loudness is measured from start to stop of the complete program to meet the aspired standard of -23 LUFS with permitted deviation of – 0.5 LU as average loudness for all audio broadcast as recommended by EBU.

Loudness Control and Parameter Adaptations

An APC was developed to essentially reset the Feed-Forward controller when a program transition is detected. This makes the Feed-forward controller adjust the output gain solely based on the loudness of the current program and not the average loudness of the previous programs.

The loudness control algorithm measures the integrated loudness of an input audio signal and adjusts its gain accordingly as measured every 0.1 seconds from start to the current time. For every iteration (0.1 seconds) the integrated loudness is calculated, resulting in a new gating block input as shown in figure 1.

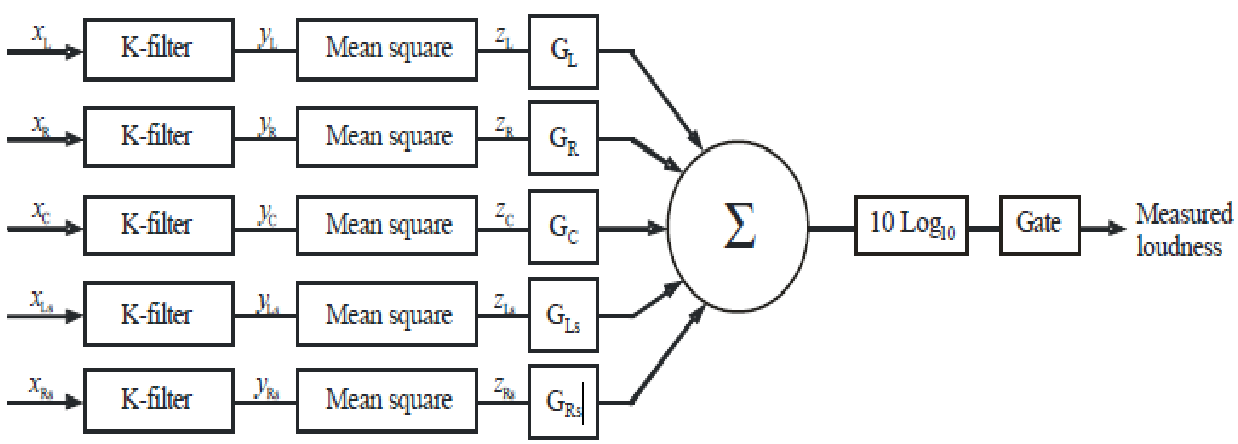

Fig. 1: Block diagram of multi-channel loudness algorithm (ITU-R BS.1770, 2014)

Fixed correction gains are applied to each channel for compensation because the same pre-filter is used in all the channels. The linear summation yields a composite loudness calculation for an equivalent loudness (Leq) expressed as (ITU-R, 2006):

$$\text{Leq} = 10Log_{10}[\frac{1}{\text{T}}\int_0^{\text{T}} \frac{X^2\text{w}}{X^2\text{ref}}\text{dt}]\text{dB} \quad (1)$$

where:

T is the integration time
X is the audio signal
W is the frequency weighting method
X^2ref is reference level of the audio

FLOW CHART OF LOUDNESS EQUALIZATION USING FLC AND APC

The flow chart of the modified broadcast loudness equalization using FLC and APC model is presented in figure 2.

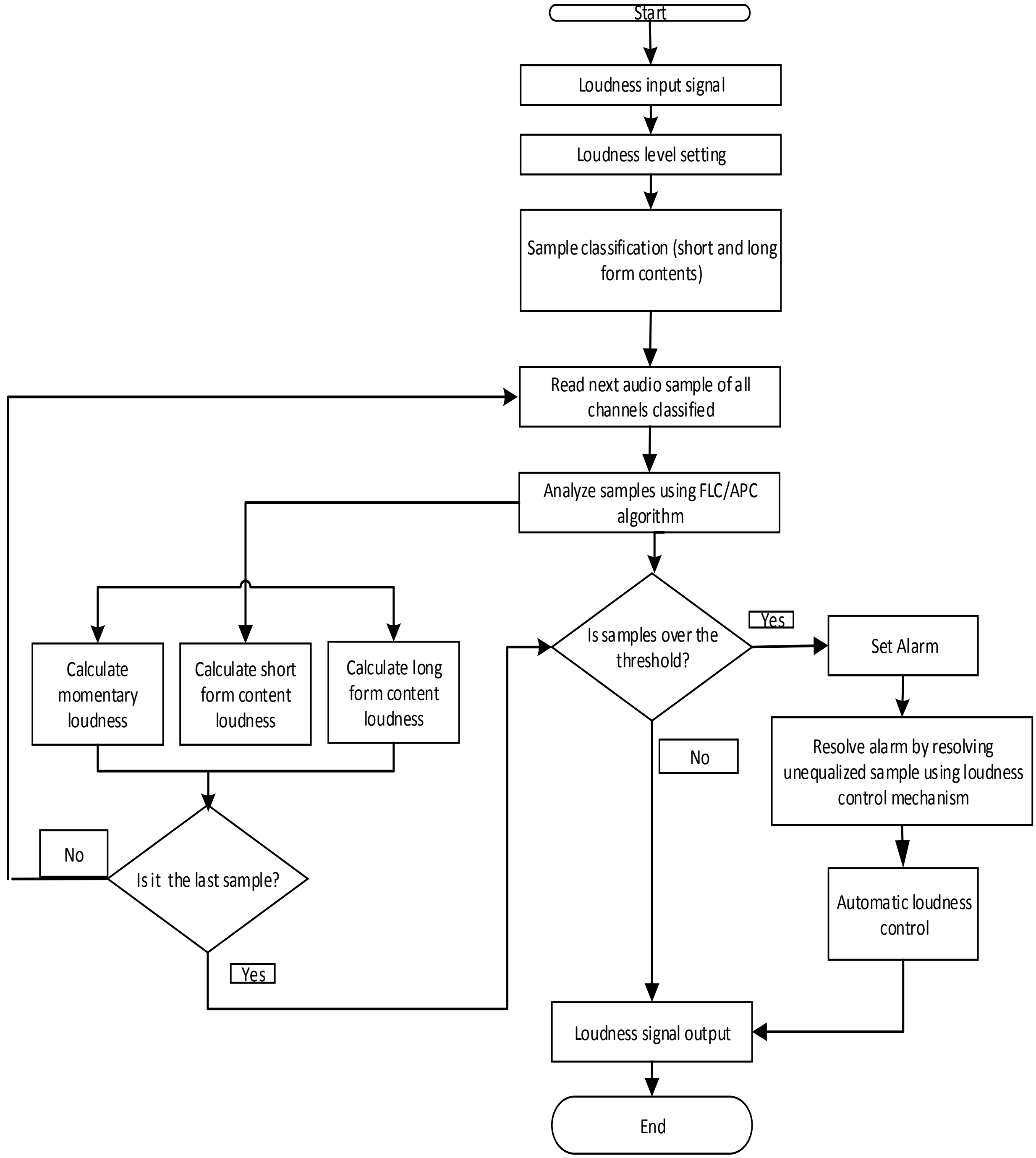

Fig. 2: Flow chart for short and long-form contents

Program Transition Identification and Simulation

When equalizing the average program loudness there is a significant advantage to know the time instant of the program transitions. By knowing the program transitions it is possible to measure the integrated loudness of each program individually so that the feed-forward controller can adjust the output gain accordingly. Signal tendencies occurring around program transitions were investigated and exploited to ease the simulation which employed parameters of table1 while the loudness difference for momentary and short contents for a given crest factor can be determined as Kalyanasundaram, (2013):

Change in loudness $(L_{\Delta i}) = L_{si} - L_{mi}$ (2)

where:

$L_{\Delta i}$ is change in loudness
L_{si} is short term loudness
L_{mi} is momentary loudness

The program transition (P_t) probability accuracy for a given set of trained data set can be determined as Olsen, (2015) :

$$P_t(\%) = \left[{P_{r(o)} - P_{r(no)}} \middle/ {P_{rT}} \right] \times 100 \quad (3)$$

where:

$P_{r(o)}$ is the probability of occurrence
$P_{r(no)}$ is the probability of non-occurrence
P_{rT} is the total probability

This can be achieved by effective identification of individual program transitions using linear correlation (r_{fg}) between two signals f_i and g_i calculated by the Pearson's correlation coefficient expressed as Pearson, (2010):

$$r_{fg} = \frac{\sum_{i=1}^{N}(f_i - \mu_f)(g_i - \mu_g)}{\sqrt{\sum_{i=1}^{N}(f_i - \mu_f)^2 \, (g_i - \mu_g)^2}} \quad (4)$$

where:

N is the total number iteration

μ_f is the mean value of f_i

μ_g is the mean value of g_i

The percentage improvement of FLC/APC scheme over MLA in input loudness accuracy (L_{la}) can be obtained from table 2 and expressed as Olsen, (2015):

$$L_{la}\ (\%) = \frac{(input(L_I) - Av.loudness)}{(InputL_I)} \times 100\% \quad (5)$$

Program Transition Identification for FLC/APC

Table 2 is the program transition identification results for both Feed-Forward Loudness Control and Parameter Configuration (FLC/APC) and short-form employing Machine Learning algorithm (MLA) as obtained during the simulation run. It is observed that the FLC/APC formed fewer errors in identifying program transitions when compared to the MLA which failed to identify more transitions. This was achieved as the data set training of the improved scheme was carried out one after the other for the over 100 individual programs. The training was done starting with fade-ins and ending with fade-outs for easier transition identification as program transition time instants were determined by signal change tendencies in a program. Correlation between two programs transiting are analyzed and indentified by the program transition identification signal as expressed by equation (2).

Data Set Realization

The input signal has duration of 34 minutes and 30 seconds documentary program with two commercial breaks running in between. The input signal had 55 trained short programs, each taken as independent program. The documentary was recorded and trained in segments as independent programs and carefully mastered to accommodate all the necessary features prone to loudness. The first commercial is an audio blocks while the second one is freebies hypes.

The simulations are performed for look-ahead of duration 0, 1, 2 and 3 seconds where appreciable results were record based on the simulation parameters shown in table 1.

Simulations parameters are listed in Table 1.

Table 1: Simulation Parameters

Parameter	Value
Gain threshold	0.5 Db
Power threshold	0.3
Maximum Attack	2 dB/sec
Maximum Release	1 dB/ sec
Minimum Release = Minimum Attack	0.5dB/sec
Decreased Attack = Decreased Release	0.5dB/sec/sec
Update Rate	10Hz
Sampling Frequency	48Khz

RESULTS

The results obtained from the simulation are analyzed based on system accuracy in determining program transition as well experimental choice of look-ahead durations. Several simulations were carried out where an input signal specifically recorded for this work is fed to the adaptive loudness equalization scheme. The simulations were classified into A0-A3 to specify the input signal and a corresponding number specifying the duration of the look-ahead. The simulations were performed for different values of look-ahead ranging from 0-3 seconds respectively. Thus simulation A2 refers to the simulation of input signal A and look-ahead of 2 seconds.

Table 2: Simulation Result of Input Signal

Simulation	**Input**	**A0**	**A1**	**A2**	**A3**	**AV.OUTPUT**
L_I (LUFS)	-24.88	-23.40	-23.48	-23.54	-23.49	-23.48
MTPL(dB/TP)	-8.96	-6.92	-7.03	-7.06	-7.07	-7.03
LRA_1(LU)	4.93	4.89	4.87	4.90	4.88	4.89
LRA_2(LU)	4.69	5.38	5.32	5.24	5.27	5.30
σ (LU)	1.33	1.30	1.20	1.16	1.16	1.20

The input values for different descriptors and the corresponding average program transition results are shown in table 2 as obtained during the simulation run. In Figures 2 to 5, the simulation output gain signals were plotted against time with values from table 1 as obtained during the simulation run. From the plots, it is observed that all simulations of the input signal gave a maximum true peak level below -1 dB/TP which has no risk of clipping the audio sound. The average program loudness is calculated for each program which is determined by obtaining the standard deviation σ, of the average program loudness for each input and output signal. The output variation in average program loudness decreased in all simulations compared to the input. The variation in average program loudness also decreased as the look-ahead increases until the look-ahead reaches a duration of 2seconds. The main target of the adaptive loudness equalization algorithm is to decrease σ so that the variation in average program loudness will reduce.

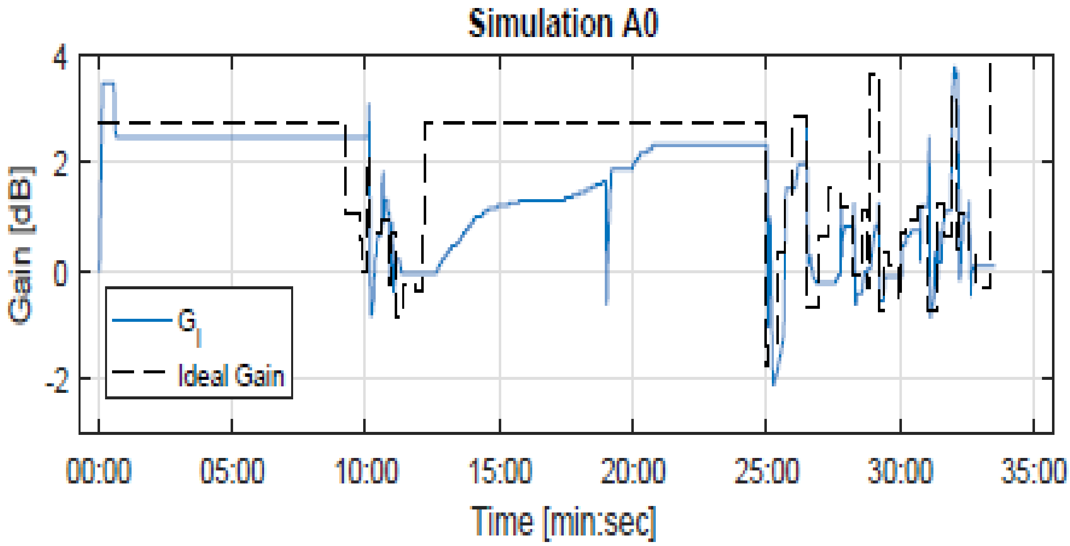

Fig. 2: Output gain signal of simulation A1

It can be observed that average program loudness determined by the standard deviation reduced from 1.33 to 1.20 and -24.88 to -23.48 for the integrated loudness equalization between input and output signifying 5.6% equal average loudness. This is one of the main advantages of measuring loudness employing long-form content over short form.

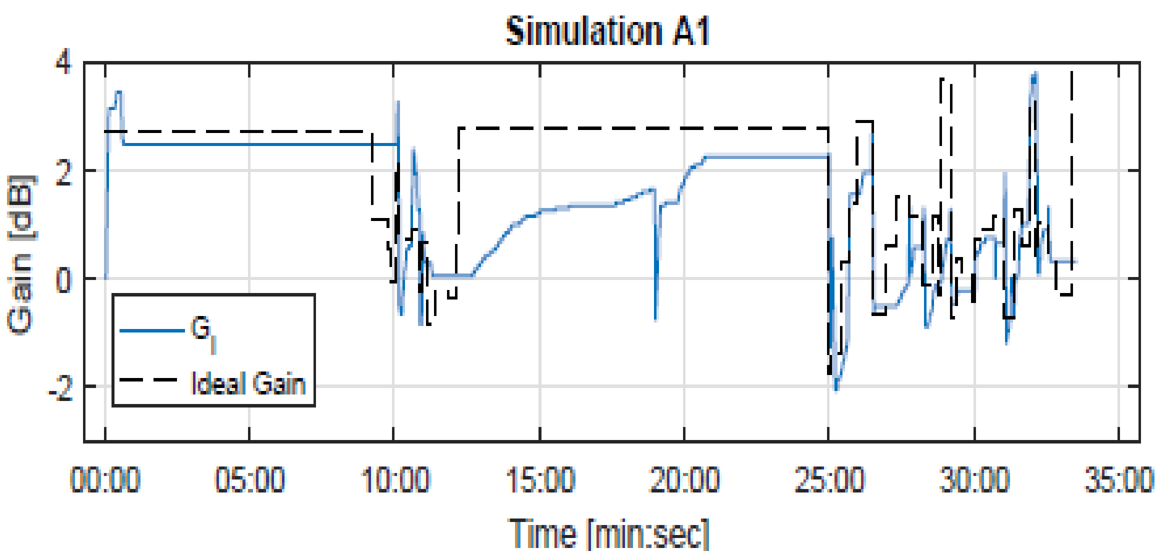

Fig. 3: Output gain signal of simulation A1

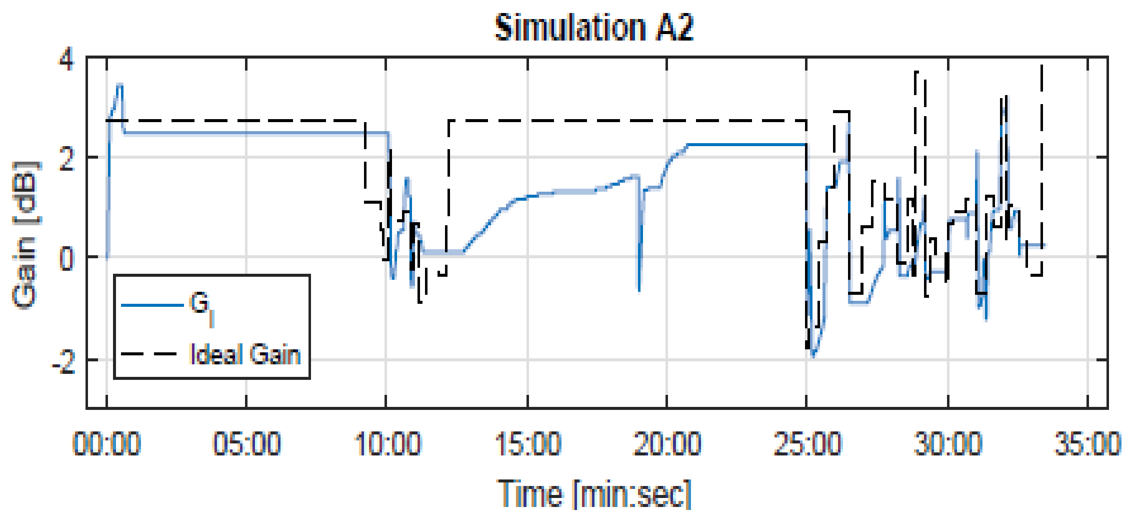

Fig. 4: Output gain signal of simulation A2

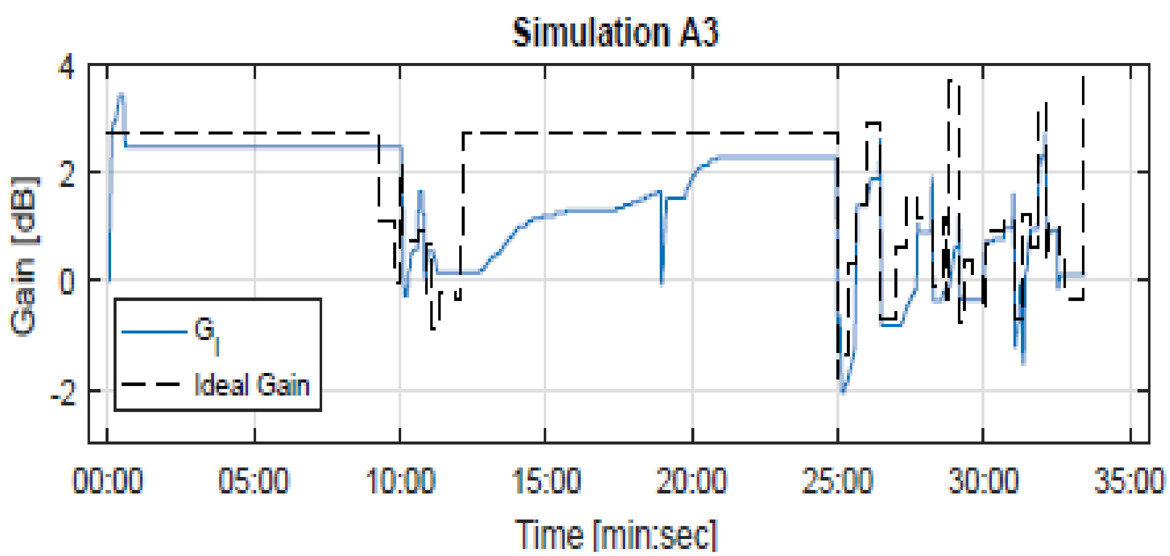

Fig. 5: Output gain signal of simulation A3

The plots shown in Figures 1 to 4 described the 9 minutes of a long play of the documentary before the first commercial break while the second commercial was played from the 26th minutes of the play time. From table 2, the loudness range LRA1 is almost unchanged in all simulations with input signal A because the output gain of the first 9 minutes is almost constant. When LRA2 is compared with LRA1, there is slight change in simulation time from A0 to A3 because of the output gain inconsistencies within the second long-term program before the second commercial.

Program Transition Identification Performance

The statistical analysis of correlation between two programs was carried out by Program Transition Identification Signal (PTIS). Square pulse signal was used to realize input signal correlation with the program transitions. The PTIS is correlated with the program transitions by calculating the cross-correlation between each PTIS. Two PTISs highly correlated have less probability of program transition than just one with significant peak, so also is the probability of program transition higher with two PTISs uncorrelated. The program probability transition occurrence is initiated once a value between 0.5 and 1 is detected as shown in table 3.

The modified FLC and APC scheme achieved 84.00% system accuracy in program transition identification obtained from results of table 3 with a program transition (P_t) probability accuracy expressed determined by equation 2. The developed FLC and APC algorithm achieved 5.6% loudness reduction when compared to MLA which can be determined using equation 5. This was achieved as a result of good choice of look-ahead time as well as prompt program transition which are key parameters to the developed algorithm. It is observed that the FLC/APC formed fewer errors in identifying program transitions when compared to the MLA which failed to identify more transitions. This was achieved as the data set training of the improved scheme was carried out one after the other for the over 100 individual programs. The training was done starting with fade-ins and ending with fade-outs for easier transition identification as program transition time instants were determined by signal change tendencies in a program. Correlation between two programs transiting are analyzed and indentified by the program transition identification signal as expressed by equation (2).

Table 3: Simulation Result of the Average Program Transition Probabilities of input A

	$L_{\Delta i}$	L_{mi}	L_{si}	C_{Lmi}	C_{Lsi}
$L_{\Delta i}$	1	0.75	0.17	0.98	0.23
L_{mi}	0.75	1	0.64	0.91	0.85
L_{si}	0.17	0.64	1	0.64	0.77
C_{Lmi}	0.98	0.91	0.64	1	0.80
C_{Lsi}	0.23	0.85	0.77	0.80	1

CONCLUSION

In this research work, a modified loudness control algorithm that carters for equalization of broadcast loudness was developed using feed-forward loudness control and adaptive parameter configuration model. The modified loudness equalization algorithm achieved 84% accuracy in program transition identification and 5.6% compliance with recommended standard of loudness unit full scale in sound equalization. If the input signal does not contain much variation in average program loudness, the algorithm is liable to increase the variation in average program loudness. For programs with significant loudness and dynamic range, the algorithm might also change the loudness range and dynamics of such programs. The duration of the look-ahead affects both the loudness equalization and dynamic retention capability, adjusting the look-ahead proportionately improves the output results.

REFERENCES

Binaural Rendering of Multi-Channel Audio Contents”AES Convention, Paris. Broadcasting Program. Paper presented at the IT Convergence and Security (ICITCS), 2014 International Conference on convergence and security.

Dominic W., Hagen W.,Russell D. M andMark D. P (2017). Estimating the loudness balance of musical mixtures using audio source Separation. Workshop on intelligent music production, Sanford, uk.

EBU (2014). R128-s1-2016, “Loudness parameters for short-form content (advertisements, promos, etc.)”(European Broadcast Union, Geneva).

Fenton E and Steven .M, (2017). Audio Dynamics Towardsa Perceptual Model of Punch. University of Huddersfield Publication

ITU, “ITU-R BS.1770-3: Algorithms to measure audio programme loudness and true-peak audio level,” Tech. Rep., International Telecommunication Union, 2014.

Lee Y.H., Chong-Sang .C, and Kim.J.W (2015). Automatic Loudness Control Based on Program Loudness Budget. Intelligent Image Processing Research Center Korea Electronics Technology Institute Gyeonggi, Korea

Lee Y.H., Chong-Sang .C, and Kim.J.W (2016). Low Delay Automatic Loudness Control for Broadcasting Services”, Intelligent Image Processing Research Center Korea Electronics Technology Institute Gyeonggi, Korea

Lee, S., Baek, B., & Kim, C. (2014). A Study on Audio Levels and Loudness Standard for Digital

Pires, L. d. S., Vieira, M. N., &Yehia, H. C. (2017). Automatic loudness control in short-form content for broadcasting. The Journal of the Acoustical Society of America, 141(3), EL287-EL292.

Ponsot E, Déjardin.H and Roncière. E (2016) "Controlling Program Loudness in Individualized

Recommendation ITU-R BS.1770-4 (10/2015),"Algorithms to measure audio program loudness and true-peak audiolevel". International Telecommunication Union Convention, Geneva, 2017.

Vyas.A. Kannao.R and Bhargava.V.(2014). Commercial Block Detection in Broadcast News Videos. A Publication of Prithwijit GuhaDept. of EEE IIT Guwahati Assam 781039, India.

Ward J. D. Reiss, and Athwal C (2012). Multi-track mixing using a model of loudness and partial loudness. The Proceedings of the 133rdAudio Engineering Society Convention

Ward. Dand Reiss.J.D,"Loudness Algorithms For Automatic Mixing", Proceedings Of The 2nd Aes Workshop On Intelligent Music Production, London, Uk, 2016

Zimmer Sebastian (2017). An approach to assess loudness and dynamics with Web Audio native nodes. Cologne Center for eHumanities Albertus-Magnus-Platz, 50923 Köln, Germany, a publication of Queen Mary University, London.