

IMPLEMENTAÇÃO DE UM CONTROLE AUTOMÁTICO DE LOUDNESS.

Resumo: O presente trabalho visa implementar um controle de Loudness em uma emissora de televisão em áudio digital, de forma a atender a legislação brasileira, durante a inserção de medias comerciais, sem comprometer a experiência sonora seja a recepção mono, estéreo ou multicanais a partir da modelagem computacional do padrão internacional de medição de Loudness chamado de padrão BS. 1770, sem a necessidade de grandes mudanças e investimentos na infraestrutura existente, o estudo aponta uma solução teórica para o problema de forma que torne o controle viável. Seis amostras foram utilizadas, cada uma contendo um bloco comercial para validação e um bloco de programa, coletado da emissora após passar por processadores de áudio profissional. A partir de trechos de amostras gravadas, utilizando o Software para análise do Loudness "R128gain", e comparando-se as amostras, foi possível efetuar um controle de Loudness automático de forma satisfatória.

Palavras-Chaves: Controle de Loudness, BS1770, Áudio digital.

IMPLEMENTATION OF AN AUTOMATIC LOUDNESS CONTROL.

Abstract: The present work aims to implement a Loudness control in a digital audio television transmitter, in order to comply with the Brazilian legislation, during the insertion of commercial averages, without compromising the sound experience be it mono, stereo or multichannel reception from the modeling computational model of the international measurement standard of Loudness called the BS standard. 1770, without the need for major changes and investments in existing infrastructure, the study points out a theoretical solution to the problem in a way that makes control feasible. Six samples were used, each containing a commercial block for validation and a program block, collected from the transmitter after passing through professional audio processors. From recorded samples, using the "R128gain"

Loudness Analysis Software, and comparing the samples, it was possible to perform an automatic Loudness control in a satisfactory way.

Keywords: Loudness Control; BS1770; digital audio.

1. INTRODUÇÃO

O ouvido converte o som em impulsos neurais em uma forma que tem sido conhecida com alguma precisão desde o final da década de 1950. Foi observado então que, o loudness era proporcional à taxa destes impulsos neurais (J Audio Engineering Society, 1990).

A rigor, esse tipo de sonoridade deveria ser chamado de intensidade sensorial, já que há uma função do cérebro, que produz um aspecto perceptual, melhor ilustrada pelo "Efeito festa de coquetel", ou seja, todos os sons entram nas orelhas, como um ruído cacofônico, mas o cérebro processa a informação e se foca em um som, como por exemplo, voz de uma pessoa e filtra o resto (Chang & Mesgarani, 2012). Em termos técnicos o Loudness pode ser definido como: A força sonora percebida de uma amostra de áudio (música, voz, efeitos sonoros, etc) que depende de vários parâmetros, tais como conteúdos frequência, duração e nível entre outras coisas (Risskov, 2006).

Partindo do ponto que o loudness perceptual está relacionado com o interesse do ouvinte no conteúdo do sinal, nós claramente não podemos tentar qualquer forma de medição desde que o interesse irá variar de um ouvinte para outro. No entanto, a visão foi colocada de lado muitas vezes, sendo que este fator perceptual é dominante na classificação do loudness do ouvinte (J. Áudio Engineering Society, 1990).

Então, testes subjetivos foram realizados, possuindo um forte viés no sentido de encontrar a importância deste fator perceptual sob as condições restritas da televisão na casa do telespectador. Os resultados então foram mais que encorajadores, no sentido de que fatores sensoriais foram esmagadoramente dominantes nestas condições (Emmett J, 1990).

A questão das mudanças dramáticas na intensidade, percebida entre os programas de TV e comerciais, remonta pelo menos há 40 anos no Reino Unido e nos Estados Unidos, enquanto que as tentativas para resolver este problema remontam ao início da década de 1970 na Universidade de Columbia Broadcasting System (CBS) (G. Spikofski, e S. Klar, 2004). Em 1966 Bauer e Torick percebem a necessidade de uma ferramenta de medição de volume. Eles também afirmaram como as tradicionais ferramentas que têm sido usadas para medir o Loudness; Medidor VU e medidor de nível, não são as ferramentas mais eficientes para medição

do Loudness. Durante a década de 1980 foi iniciada uma tentativa de desenvolver um sistema de controle de intensidade operando na etapa da inserção de materiais feita pela Thames Television juntamente com o ITC (Incorporated Television Company), no entanto, este esforço que foi feito durante o período que o Reino Unido adotava o som NICAM (Near Instantaneous Companded Audio Multiplex), e a distribuição de áudio digital, foi interrompida em 1992 quando a Thames Television perdeu a sua franquia (E. L. Torick, E. L., R. G. Allen, R. G. & B.B. Bauer, 1968). Em 2007, a ITU-R tomou a iniciativa de criar um padrão internacional de medição de Loudness chamado de padrão BS. 1770, que definiu o Loudness pela equação (1), onde “T” se refere ao tempo da integração e “x” o sinal de áudio. (RECOMMENDATION ITU-R BS 1770-1, 2006-2007).

$$Leq (W) = 10 \log\left[\frac{1}{T} \int_0^T \frac{x^2 W}{x^2_{Ref}} dt\right] \text{ dB} \quad (1)$$

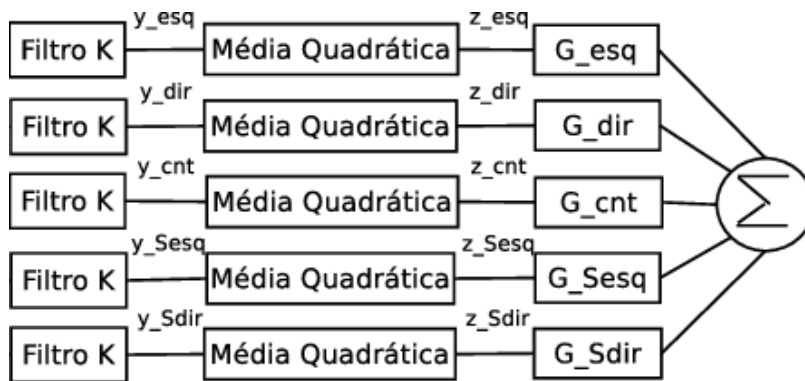
2. METODOLOGIA

2.1 A intensidade média subjetiva do áudio

Na primeira fase um algoritmo foi desenvolvido e simulado no ambiente computacional usando o Matlab para checar sua funcionalidade. O algoritmo incluiu a estimativa do nível do Loudness para um sistema multicanal (figura 1), e um controle adaptativo de Loudness.

As medidas de Loudness foram realizadas nos canais estéreo e 5.1. No caso do áudio 5.1, os 05 canais principais (esquerdo, central, direito, esquerdo traseiro e direito traseiro) foram considerados. O canal de efeitos de baixa frequência (LFE) foi excluído da medição.

Figura 1. Modelagem do Loudness em multicanais, baseado no padrão BS.1770.



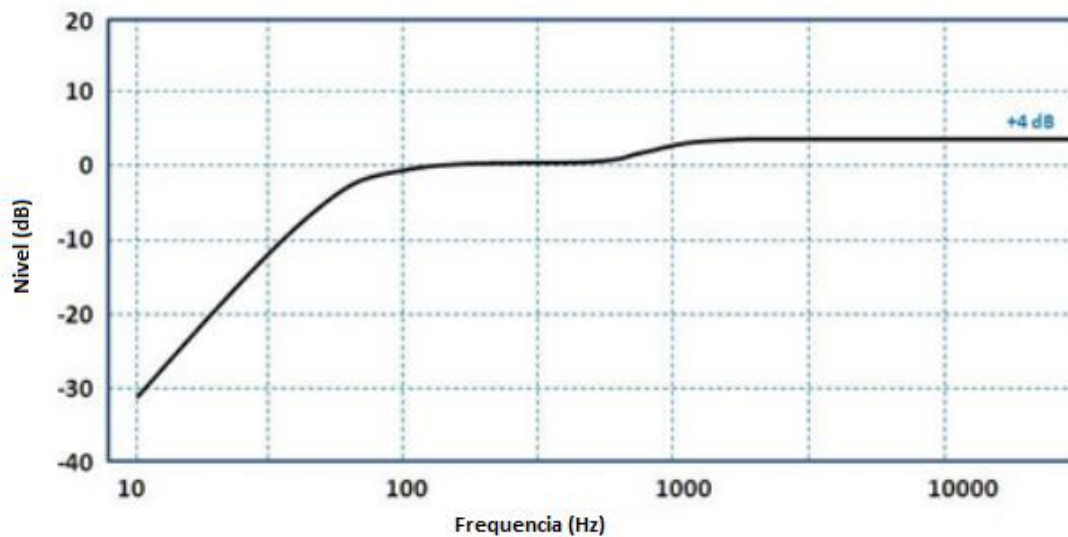
A Principal vantagem desse algoritmo está na sua simplicidade e escalabilidade. Ele é simples pelo fato de ser constituído de blocos básicos de processamento de sinal que podem ser implementados no domínio do tempo sem a necessidade de hardwares robustos. E também tem boa escalabilidade desde que o mesmo processamento é aplicado a cada canal, além da fácil implementação para o projeto proposto por poder acomodar qualquer quantidade de canais proposta.

Filtro K

O filtro K é um estágio de pré-filtro (figura 2). Conforme a figura 1 ele é aplicado a todos os canais (com exceção do canal LFE), e foi modelado dentro do padrão BS1770 para simular os efeitos acústicos da audição humana. Ele considera que, as altas frequências são percebidas pela audição humana mais altas (e mais irritantes) que o normal[11]. Por esse motivo a curva desse filtro é modelada de forma que:

- Atenua as frequências mais baixas (menores que 100 Hz).
- Preserva as frequências médias (entre 100 Hz e 1 KHz).
- Amplifica as frequências altas (maiores que 1 KHz).

Figura 2. Filtro do fator K



Media quadrática

Calcula a potência sonora das amostras do pré-filtro, considerando o intervalo de integração (T) em 400 milissegundos. Considerando uma amostragem de sinal de áudio de 48 KHz, 400 ms corresponde a 19200 amostras por canal. A amostragem de 48 KHz foi escolhida pois é uma frequência comum, e de alta fidelidade em amostragem de áudio, usadas em muitas mídias digitais em telecomunicações. Dessa forma a média quadrática de cada canal será:

$$L = \frac{L_1^2 + L_2^2 + \dots + L_{19200}^2}{400}$$

(2)

Ganhos dos canais (G)

A potência sonora da soma dos canais processados é então dada da seguinte forma:

- Os canais estéreos esquerda e direita são amplificados em 1,5 dB (1,41).
- No caso da recepção 5.1, os canais central, direita traseiro e esquerda traseiro não recebem amplificação (0 dB)

Dessa forma, retornando à equação (1), o resultado do Loudness instantâneo será:

$$LK = -0.691 + 10 \log_{10} (L_t + R_t + C + (1.41 \times L_{\text{stereo}}) + (1.41 \times R_{\text{stereo}}))$$

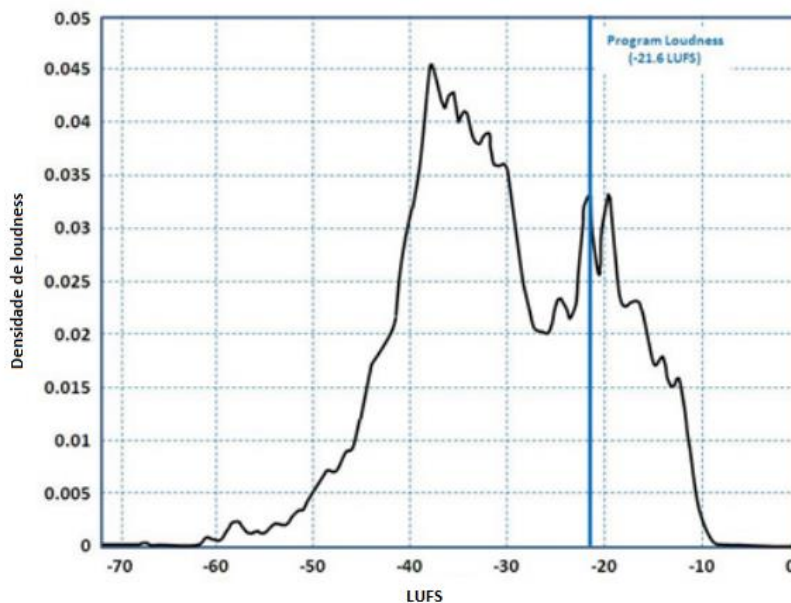
A unidade de medida do Loudness é o LKFS, que se refere ao fator K relativo a escala completa nominal (Loudness K full scale). O LKFS é similar a escala de decibéis, onde o aumento do nível de um sinal de áudio em 1 dB resulta em um aumento da leitura do Loudness em 1 LKFS.

2.2 Faixa de intensidade subjetiva de áudio

A medida da faixa de intensidade subjetiva de áudio LRA (Loudness range), é uma medida quantitativa da variação do Loudness dentro de um programa. Genericamente é medida em LU (Loudness Unit) e ajuda a decidir quando uma compressão dinâmica é necessária[10]. Não existe um valor fixado para o LRA, pois a medida pode variar muito a depender do conteúdo, porém o valor de 15 LRA é um valor comum para emissoras de televisão aberta brasileira[12], porém a Rede Globo especifica valores mais rígidos para intervalos comerciais, que devem ficar abaixo de 8 LRA.

A medida do LRA é baseada na distribuição estatísticas do Loudness, usando uma janela de corte de integração de 3 segundos (figura 3).

Figura 3. Distribuição de Loudness.

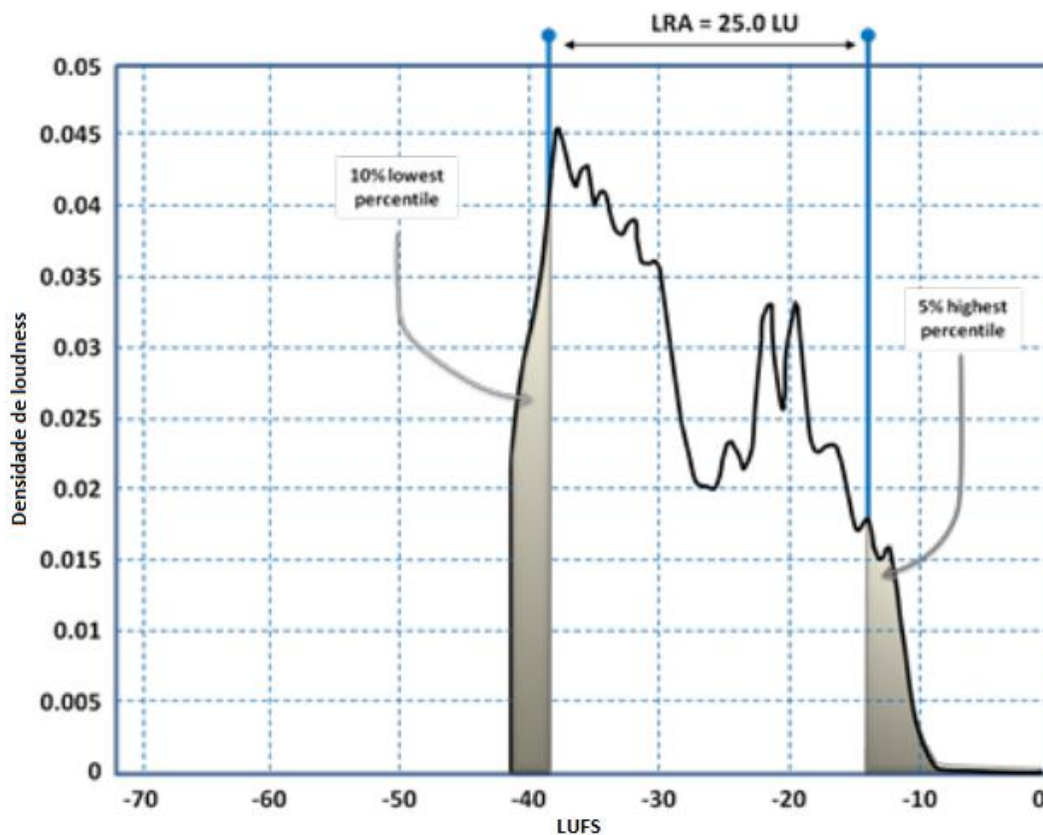


Usando a distribuição da figura 3 como exemplo, toda medida de Loudness coletada abaixo desses pontos serão excluídas: -70 LUFS, devido ao fato de nessa faixa não serem encontrados nenhum sinal de áudio relevante; -20 LU relativo ao programa medido, o objetivo é não inserir na medição períodos silenciosos ou com pequenos ruídos de fundo.

Dessa forma o LRA será a diferença entre as estimativas das decimas e das nonagésimas quintas porcentagens da distribuição. As porcentagens menores que 10% correspondem as atenuações comuns, como a atenuação no final de uma música. As porcentagens maiores que 95% correspondem a ruídos muito altos que ocorrem de forma repentina, como um ruído de disparo de armas de fogo.

Voltando então a distribuição da figura 3 e aplicando essas condições, temos então que para o valor de Loudness de -21,6 LUFS do programa aferido nesse exemplo, o valor do LRA será de 25 LU.

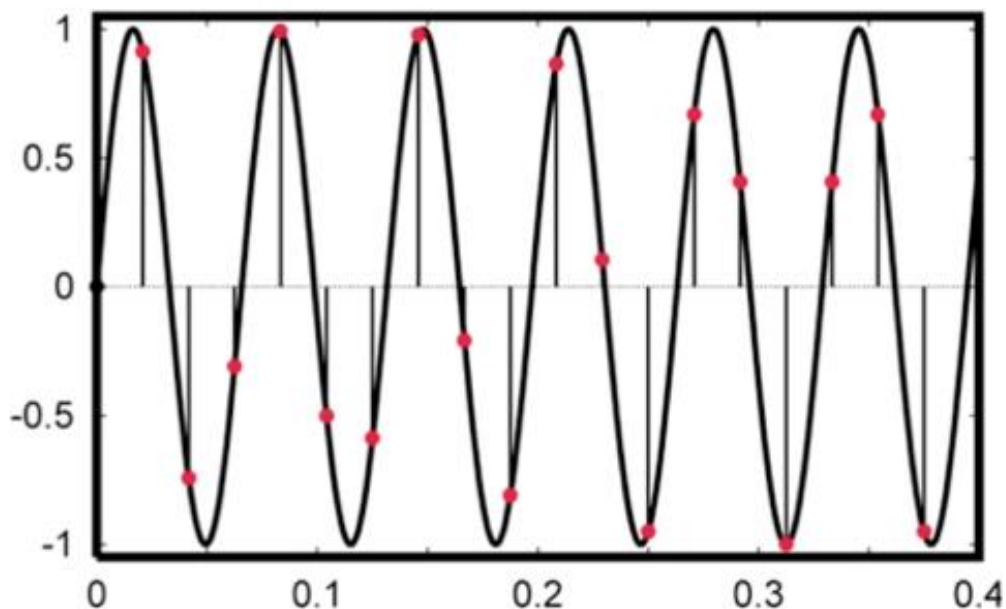
Figura 4. Faixa de intensidade subjetiva de áudio (LRA).



2.3 Nível de pico verdadeiro.

A medida do pico verdadeiro indica o valor máximo do módulo do gráfico da senoide de um sinal no domínio do tempo. Esse valor não deve exceder um nível recomendado para evitar a saturação do áudio e o *clipping* de áudio[11]. O pico máximo verdadeiro permitido em programas televisivos não deve exceder -1 dBTP. O valor do pico verdadeiro pode exceder o valor da amostragem de 48 KHz no domínio do tempo.

Figura 5. Pico verdadeiro.



Devido ao fato do valor de pico verdadeiro poder exceder o valor amostragem, podendo ficar entre as amostras como ilustrado nos pontos vermelhos da figura 5, será considerado no cálculo do pico verdadeiro uma sobre amostragem de 4 vezes o valor da amostragem original de 48 KHz, o que significa que no cálculo do valor de pico verdadeiro será considerada uma amostragem de 192 KHz. Porém essa sobre amostragem pode induzir distorções no sinal[12], então o sinal será atenuado em 12 dB, o que corresponde a 4 vezes menos que o sinal original, e depois passara por um filtro do padrão BS1770 para então termos o valor do pico verdadeiro, figura 6.

Figura 6. Calculo do Pico verdadeiro.



3. RESULTADOS E DISCUSSÃO

1ª Amostra

```
DigitemUrl successfully loaded.  
SoX successfully loaded.  
FFmpeg successfully loaded.  
analyzing ...  
[1/2] "001_programa_track1.wav": -23.1 LUFS (<0.1 LU)  
peak: -3.7 TPFS, range: 4.4 LU  
[2/2] "01_bloco comercial_track1.wav": -23.4 LUFS (<0.4 LU)  
peak: -7.6 TPFS, range: 7.0 LU
```

Trecho do Programa
Trecho do Comercial

Na primeira amostra, os arquivos de áudio digital possuem:

Bloco de Programa: Intensidade subjetiva média de áudio no valor de -23,1 LUFS (ou -23,1 LKFS), 0,1 LU abaixo da norma, pico verdadeiro no valor de -3,7 TPFS e faixa de intensidade subjetiva de áudio no valor de 4,4 LU

Bloco de Comercial: Intensidade subjetiva média de áudio no valor de -23,4 LUFS, 0,4 LU abaixo da norma, pico verdadeiro no valor de -7,6 TPFS e faixa de intensidade subjetiva de áudio no valor de 7,0 LU.

2ª Amostra

```
DlgItemUrl sucessfully loaded.  
SoX sucessfully loaded.  
FFmpeg sucessfully loaded.  
analyzing ...  
[1/2] "002_programa_track1.wav": -23.8 LUFS <0.8 LU>  
    peak: -6.9 TPFS, range: 4.6 LU  
[2/2] "02_bloco comercial_track1.wav": -22.9 LUFS <-0.1 LU>  
    peak: -6.6 TPFS, range: 6.6 LU
```

Trecho do Programa

Trecho do Comercial

Na segunda amostra, os arquivos de áudio digital possuem:

Bloco de Programa: Intensidade subjetiva média de áudio no valor de -23,8 LUFS, 0,8 LU abaixo da norma, pico verdadeiro no valor de -6,9 TPFS e faixa de intensidade subjetiva de áudio no valor de 4,6 LU.

Bloco de Comercial: Intensidade subjetiva média de áudio no valor de -22,9 LUFS, 0,1 LU acima da norma, pico verdadeiro no valor de -6,6 TPFS e faixa de intensidade subjetiva de áudio no valor de 6,6 LU.

A diferença na intensidade média subjetiva do áudio entre os blocos (Programa e Comercial) foi de 0,9 LU.

3ª Amostra

```
DlgItemUrl sucessfully loaded.  
SoX sucessfully loaded.  
FFmpeg sucessfully loaded.  
analyzing ...  
[1/2] "003_progama_track1.wav": -23.2 LUFS <0.2 LU>  
    peak: -8.8 TPFS, range: 4.4 LU  
[2/2] "03_bloco comercial_track1.wav": -23.6 LUFS <0.6 LU>  
    peak: -7.3 TPFS, range: 6.5 LU
```

Trecho do Programa

Trecho do Comercial

Na terceira amostra, os arquivos de áudio digital possuem:

Bloco de Programa: Intensidade subjetiva média de áudio no valor de -23,2 LUFS, 0,2 LU abaixo da norma, pico verdadeiro no valor de -8,8 TPFS e faixa de intensidade subjetiva de áudio no valor de 4,4 LU.

Bloco de Comercial: Intensidade subjetiva média de áudio no valor de -23,6 LUFS, 0,6 LU abaixo da norma, pico verdadeiro no valor de -7,3 TPFS e faixa de intensidade subjetiva de áudio no valor de 6,5 LU.

A diferença na intensidade média subjetiva do áudio entre os blocos (Programa e Comercial) foi de 0,4 LU.

4ª Amostra

```
DlgItemUrl sucessfully loaded.  
SoX sucessfully loaded.  
FFmpeg sucessfully loaded.  
analyzing ...  
[1/2] "004_programa_track1.wav": -22.7 LUFS <-0.3 LU>  
      peak: -5.2 TPFS, range: 4.4 LU  
[2/2] "04_bloco comercial_track1.wav": -23.6 LUFS <0.6 LU>  
      peak: -6.8 TPFS, range: 6.1 LU
```

Trecho do Programa
Trecho do Comercial

Na quarta amostra, os arquivos de áudio digital possuem:

Bloco de Programa: Intensidade subjetiva média de áudio no valor de -22,7 LUFS, 0,3 LU acima da norma, pico verdadeiro no valor de -5,2 TPFS e faixa de intensidade subjetiva de áudio no valor de 4,4 LU.

Bloco de Comercial: Intensidade subjetiva média de áudio no valor de -23,6 LUFS, 0,6 LU abaixo da norma, pico verdadeiro no valor de -6,8 TPFS e faixa de intensidade subjetiva de áudio no valor de 6,1 LU.

A diferença na intensidade média subjetiva do áudio entre os blocos (Programa e Comercial) foi de 0,9 LU.

5ª Amostra

```
DlgItemUrl sucessfully loaded.  
SoX sucessfully loaded.  
FFmpeg sucessfully loaded.  
analyzing ...  
[1/2] "005_programa_track1.wav": -24.0 LUFS <1.0 LU>  
    peak: -3.3 TPFS, range: 8.8 LU  
[2/2] "05_bloco comercial_track1.wav": -22.9 LUFS <-0.1 LU>  
    peak: -8.5 TPFS, range: 5.6 LU
```

Trecho do Programa
Trecho do Comercial

Na quinta amostra, os arquivos de áudio digital possuem:

Bloco de Programa: Intensidade subjetiva média de áudio no valor de -24,0 LUFS, 1,0 LU abaixo da norma, pico verdadeiro no valor de -3,3 TPFS e faixa de intensidade subjetiva de áudio no valor de 8,8 LU.

Bloco de Comercial: Intensidade subjetiva média de áudio no valor de -22,9 LUFS, 0,1 LU acima da norma, pico verdadeiro no valor de -8,5 TPFS e faixa de intensidade subjetiva de áudio no valor de 5,6 LU.

A diferença na intensidade média subjetiva do áudio entre os blocos (Programa e Comercial) foi de 1,1 LU.

6ª Amostra

```
DlgItemUrl sucessfully loaded.  
SoX sucessfully loaded.  
FFmpeg sucessfully loaded.  
analyzing ...  
[1/2] "006_programa_track1.wav": -22.8 LUFS <-0.2 LU>  
    peak: -3.0 TPFS, range: 13.5 LU  
[2/2] "06_bloco comercial_track1.wav": -23.5 LUFS <0.5 LU>  
    peak: -7.9 TPFS, range: 5.6 LU
```

Trecho do Programa
Trecho do Comercial

Na sexta amostra, os arquivos de áudio digital possuem:

Bloco de Programa: Intensidade subjetiva média de áudio no valor de -22,8 LUFS, 0,2 LU acima da norma, pico verdadeiro no valor de -3,0 TPFS e faixa de intensidade subjetiva de áudio no valor de 13,5 LU.

Bloco de Comercial: Intensidade subjetiva média de áudio no valor de -23,5 LUFS, 0,5 LU abaixo da norma, pico verdadeiro no valor de -7,9 TPFS e faixa de intensidade subjetiva de áudio no valor de 5,6 LU.

A diferença na intensidade média subjetiva do áudio entre os blocos (Programa e Comercial) foi de 0,7 LU.

TABELA DOS RESULTADOS OBTIDOS

AMOSTRA	BLOCOS	LOUDNESS MÉDIO (LUFS)	RANGE (LU)	TRUE PEAK (TPFS)	Δ LOUDNESS (LU)
1 ^a	PROGRAMA	-23,1	4,4	-3,7	0,3
	COMERCIAL	-23,4	7,0	-7,6	
2 ^a	PROGRAMA	-23,8	4,6	-6,9	0,9
	COMERCIAL	-22,9	6,6	-6,6	
3 ^a	PROGRAMA	-23,2	4,4	-8,8	0,4
	COMERCIAL	-23,6	6,5	-7,3	
4 ^a	PROGRAMA	-22,7	4,4	-5,6	0,9
	COMERCIAL	-23,6	6,1	-6,8	
5 ^a	PROGRAMA	-24,0	8,8	-3,3	1,1
	COMERCIAL	-22,9	5,6	-8,5	
6 ^a	PROGRAMA	-22,8	13,5	-3,0	0,7
	COMERCIAL	-23,5	5,6	-7,9	

Tabela-01: Resultados obtidos na análise das amostras pós controlador.

4. CONCLUSÃO

Considerando os parâmetros a serem fiscalizados, em nenhuma das amostras analisadas foi constatada uma ou mais irregularidades, ou seja, todas as amostras atenderam à norma do sistema brasileiro de televisão digital após passarem pelo controlador.

- I. A diferença na intensidade média subjetiva do áudio entre blocos adjacentes inferior a 2 LU;
- II. A intensidade média subjetiva do áudio do bloco de intervalo comercial dentro da margem de -23 LKFS em +/- 2 LU;
- III. A faixa de intensidade subjetiva de áudio (Loudness) medida dentro da margem de 15 LU.

Além dos parâmetros fiscalizado, também foi considerado as especificações da Rede Globo para os comerciais que são:

- I. Loudness médio em -23 LKFS com tolerância de 0 LU para mais, e 1 LU para menos.
- II. Faixa de Loudness máximo em 8 LU;
- III. Pico de Loudness máximo em -3 dBTP;

Sendo assim, conclui-se que esse trabalho atingiu com louvor a proposta inicial de efetuar um controle automático de Loudness em intervalos comerciais em uma emissora de televisão operando em áudio digital multicanal e está de acordo com a Portaria nº 354, de 11 de julho de 2012, que regulamenta a padronização do volume de áudio nos intervalos comerciais da programação dos serviços de radiodifusão sonora e de sons e imagens nos termos da Lei nº 10.222, de 9 de maio de 2001[5].

5. REFERÊNCIAS

[1] "Digital Peak Coding Level Indicator" at present under consideration by IEC.TC84.WG10.

[2] Fletcher H and Munqon W A "Loudness; Its Definition, Measurement and Calculation". J. Acoustic SOC of America 5 pp 82-108 (1933).

- [3] J Audio Engineering Society, Vol 38, NO 06, 1990.
- [4] Emmett J. "A Programme Loudness Meter". Institute of Acoustics Conference, 1990.
- [5] <http://www.anatel.gov.br/legislacao/procedimentos-de-fiscalizacao/786-portaria-559>, acessado em 15 de Agosto de 2016.
- [6] Dr. Edward Chang and Eng. Nima Mesgarani "[The cocktail part effect](http://www.hear-it.org/The-mystery-of-the-cocktail-party-effect-solved)" <http://www.hear-it.org/The-mystery-of-the-cocktail-party-effect-solved>.
- [7] Grimm E, Skovenborg E, and Spikofski G, (2010) Determining an Optimal Gated Loudness Measurement for TV Sound Normalization, in Proc. of the 128th Convention of Audio Engineering Society.
- [8] Skovenborg E and Lund T, (2008) Loudness Descriptors to Characterize Programs and Music Tracks, in Proc. of the AES 125th Convention, San Francisco.
- [9] Lee SW, Cho YS and Kim JK, (2013) A Study on analysis of digital TV loudness, The Journal of Korea Society of Communications and Space Technology, vol. 8, no. 4, pp. 105-110
- [10] Punchihewa A, Kilani S, and Haizart E, (2012) Loudness Management for Home Television Viewing, 2012 IEEE International Instrumentation and Measurement Technology Conference (I2MTC), pp. 844-848
- [11] RECOMMENDATION ITU-R BS.1770-1. (2006-2007). Algorithms to measure audio programmed level and true-peak signal level.
- [12] Monique Cruvinel <http://www.abert.org.br/web/index.php/notmenu/item/20989-loudness-comeca-a-ser-cobrado-a-partir-de-12-de-julho>.
- [13] G. Spikofski, and S. Klar, "Levelling and Loudness – In Radio and Television Broadcasting", EBU Technical Review, 2004.
- [14] E. L. Torick, E. L., R. G. Allen, R. G. & B.B. Bauer, "Automatic Control of Loudness Level", IEEE Transactions on Broadcasting. vol.BC-14, no. 4, 1968.
- [15] T Lund, "Control of Loudness in Digital TV", NAB BEC Proceedings. Risskov, Denmark, 2006.