Audio Engineering Society

Convention Paper 7896

Presented at the 127th Convention
2009 October 9–12 New York, NY, USA

A Survey of Broadcast Television Perceived Relative Audio Levels

Matt Easley[1], Chris Hanna[2], Scott Skinner[3] and Matt Barnhill[4]

[1] THAT Corporation, Alpharetta, GA
mfe@thatcorp.com

[2] THAT Corporation, Milford, MA
cmh@thatcorp.com

[3] THAT Corporation, Milford, MA
sws@thatcorp.com

[4] THAT Corporation, Alpharetta, GA
msb@thatcorp.com

ABSTRACT

Perceived television volume levels can vary dramatically as audio changes both within a given broadcast channel and between broadcast channels. This paper surveys the broadcast audio levels in two large metropolitan areas (Atlanta and Boston). Both analog and digital broadcasts are monitored from multiple cable providers. Two-channel perceived loudness is measured utilizing the ITU-R Rec. BS.1770 loudness meter standard. Statistical data is presented showing the severity and nature of the perceived loudness changes. Finally, dynamic volume control technology is applied to the most severe recordings for perceived loudness comparisons.

1. OVERVIEW

During television viewing, loudness changes can be irritating and often involve manual volume adjustments by the viewer. One example would be the perceived volume change that often occurs when changing channels on a television. Another example would be the perceived volume change that can occur between a television program and a commercial. These large relative changes are typically attributed to lack of proper level control at the point of broadcast or audio compression introduced during production.

A lesser known cause of increased perceived loudness is multiple stages of spatial processing. Many television brands offer left and right channel processing that creates a virtual surround audio sensation with two speakers. The audio in some program and commercial material is pre-processed in the studio to introduce surround spatial effects (pseudo-surround) in two-channel systems. If this type of broadcast audio is then additionally processed in the television to introduce two-channel surround effects the perceived level change can be dramatic.

In all cases, automatic volume control technology can minimize listener discomfort and maintain a more consistent volume level. Over the past year we have spent a considerable amount of time recording, measuring, and documenting the perceived loudness of broadcast television audio. The purpose of this exercise was to motivate the need for automatic volume control technology in television receivers and to acquire a library of audio clips for testing purposes.

While much attention has been paid to leveling audio at the point of broadcast, it seems to have done little to alleviate the problem. In fact, with the advent of high dynamic range DTV broadcasts, larger loudness differences are now perceived by the television viewer.

We have focused our efforts on two-channel perceived loudness since viewers who utilize the two speakers mounted in their television represent a large segment of television watchers.

To gather data about perceived loudness transitions, we developed a laptop-PC-based audio/video acquisition platform. This platform was then used for two purposes. First, it allowed viewers to unobtrusively record their normal television viewing. If an unusual loudness event was perceived, the viewer simply presses a key which records the time of the event in a log file. The time stamp identifies significant broadcast audio events for later study.

Several of these acquisition platforms have been put into regular use in home viewing environments. In this manner we have acquired a large library of recordings of worst-case channel changes and program/commercial transitions.

In order to produce the survey described in this paper, we wanted to capture as many channel changes as possible. To accomplish this, we modified the platform somewhat. We modified the acquisition software to randomly change the channel of a set-top box via a Hauppauge IR-Blaster. The resulting audio/video signals were recorded for future analysis. We used a MATLAB module which implements the ITU-R Rec. BS.1770[1] two-channel perceived loudness meter to process and analyze the acquired audio

Finally, compressor-based volume control technology was developed, and implemented as a VST plug-in, to determine the improvements in perceived loudness that can be achieved with minimal processing at the television receiver.

2. ACQUISITION/ANALYSIS

The core technology utilized in acquiring the A/V recordings was the Hauppauge WinTV-HVR 1950. The HVR 1950 consists of a tuner, decoder (NTSC, ATSC or QAM), IR blaster and MPEG-2 encoder. The HVR 1950 interfaces to a laptop computer, via USB, for viewing and recording purposes. It also contains analog (L & R) Audio and (composite) Video inputs which can be used instead of its own tuner as a source of signal. To acquire significant audio events (i.e. loud commercials and/or loud channel changes) the HVR 1950's A/V input was often attached to the set-top box or television A/V output. Alternatively, some recordings utilized the Hauppauge tuner, with its RF inputs connected directly to the cable jack.

As noted above, we developed an application which utilized the HVR 1950 to continually record whatever material the set-top box or TV was decoding. The application also kept a companion time log with each recording. If the viewer noticed an irritating audio event they could push any key on the laptop and the time log would record when the event occurred. Later, the time log and recording could be processed to capture the audio event for further analysis and testing with automatic volume control technology. This setup allowed the insertion of A/V acquisition technology into several homes without interrupting normal television viewing habits.

The resulting long recordings were used to gather samples of program/commercial transitions. It was determined that there is no reliable way to automatically detect program/commercial transitions. Acquiring these samples also required manual video editing. Because of this limitation, the number of this type of sample available for analysis was somewhat limited.

In order to gather statistics on perceived volume change during channel changes, a modification of the setup was required. Specifically, the application was modified to utilize the IR blaster to change the channel of a set-top box every 15-25 seconds. The application utilized a text file listing all supported analog and digital channels for a given location. The list was shuffled before each channel change sequence. Short individual recordings of each channel change were made. A given channel never occurred twice in any given sequence. Sequences were run throughout the day.

In preparation for analysis, the audio was stripped from the MPEG recordings and converted into a wav file. We developed a MATLAB script that processes the wav files with a two-channel (left and right) version of the ITU-R Rec. BS.1770 loudness meter. This loudness meter is a proven algorithm for determining subjective program loudness. In our implementation of the loudness meter we chose to process data with a 5 second sliding window. A 50 second plot of a channel transition loudness meter output, as a function of time, is shown in Figure 1.

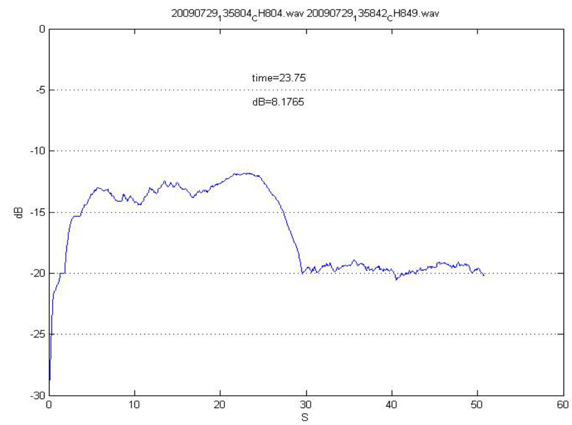

Figure 1: Channel 804 to 849 Transition Loudness Plot

In this recording the channel change takes place at 25 seconds. Each loudness meter output is analyzed within a +/- 5 second window around the channel transition point. The largest dynamic range change within that window is assumed to be the relative perceived loudness during the channel change. In this case the perceived loudness difference, upon changing channels, is measured to be 8.2 dB.

Upon analysis of the loudness data, some plots were deemed not to have a perceived loudness change. This decision is made by examining the nature of the audio material before and after the channel transition. If the loudness dynamic range of the audio material exceeds the measured range of the channel transition, the data is thought not to represent a perceived loudness change. Figure 2 is an example of such a transition.

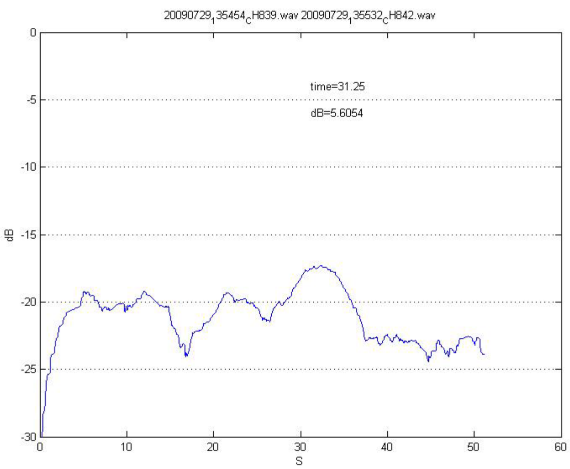

Figure 2: Channel 839 to 842 Transition Loudness Plot

In this figure, the MATLAB analysis script determined that the loudness change after the channel transition (at about 37 seconds) was larger than the any measured loudness change at the channel change point (25 seconds). This channel change was determined not to have a significant perceived level difference. We listened to numerous clips to ensure there was correlation between the undulations in the loudness data and actual perception.

3. CHANNEL CHANGES

Recordings of 1,790 channel changes were analyzed in this manner. The recordings were a random mix of analog and digital channels acquired in the Boston and Atlanta metropolitan areas. Upon analysis, 23.3 percent (417) of these channel change recordings were determined not to have a significant perceived loudness change. A histogram of the perceived loudness differences of the remaining channel change recordings is shown in Figure 3.

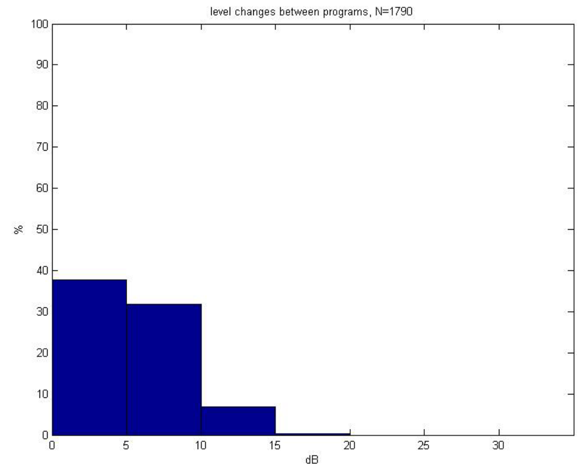

Figure 3: Histogram of Channel Change Relative Loudness

Of 1790 random channel transitions, 38% had a perceived loudness change between 0 dB and 5 dB, 32% had a loudness change in the 5-10 dB range, 6% had a loudness change in the 10-15 dB range and 1% had a loudness change between 15-20 dB.

While not part of the random data accessed for statistical purposes, perceived level changes greater than 20 dB have been observed in other data acquisition exercises. Figure 4 below shows loudness data from a recording made while changing from a Public Broadcasting channel to a Public Access channel. A perceived loudness change of 28.9 dB was measured. Listening reveals that this channel change is extremely aggravating in certain situations. Typically, the television volume would be necessarily increased to hear the dialogue in the public television broadcast. Once the television volume is set at a comfortable level, and the channel is changed, the extremely loud and distorted audio from the public access broadcast is at the level where it could potentially irritate people in rooms adjacent to the viewing area. Situations such as this are common on channels that have local programming insertion.

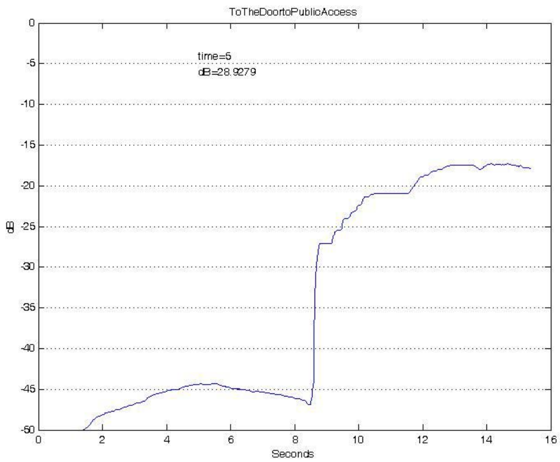

Figure 4: Extreme Channel Change Loudness Plot

4. PROGRAM/COMMERCIAL TRANSITIONS

Recordings of program-to-commercial transitions were similarly analyzed. Figure 5 show a typical example of loudness data for program to commercial transitions. In this case, a television show has a slow audio fade toward the commercial break. The commercial, a “teaser” for future programming, begins at its peak loudness level. The result is an uncomfortable loudness change.

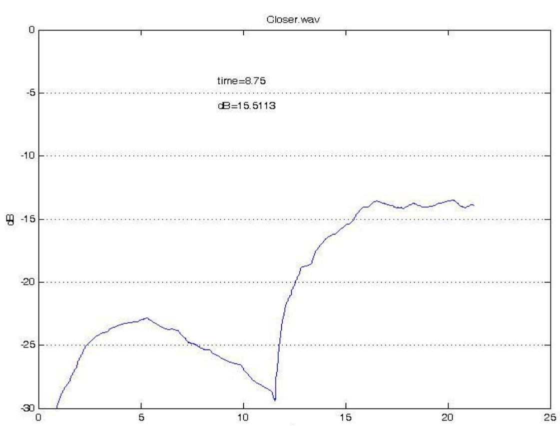

Figure 5: Television Program to Commercial Transition Loudness Plot

Figure 6 is a histogram of the magnitude of the perceived loudness changes during 82 random commercial transition recordings.

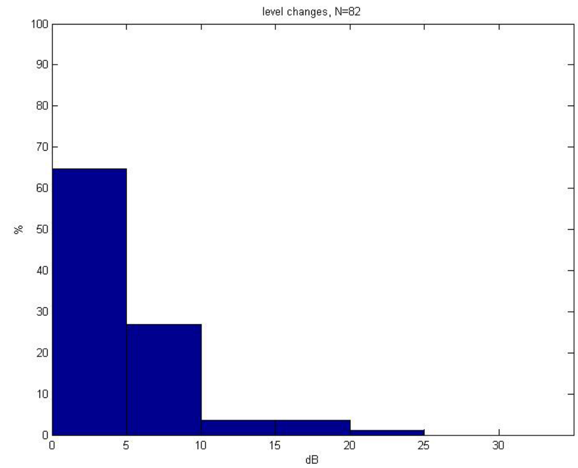

Figure 6: Histogram of Program-to-Commercial Transition Relative Loudness

Almost 65% of the commercial transitions had a loudness change of 5 dB or less. About 27% had a level change of more than 5 dB. A significant percentage, 9%, had over a 10 dB change in perceived loudness.

5. VOLUME CONTROL

Figure 7 shows an ITU perceived loudness plot of a channel “surfing” recording made from an analog cable broadcast in the Boston, MA area.

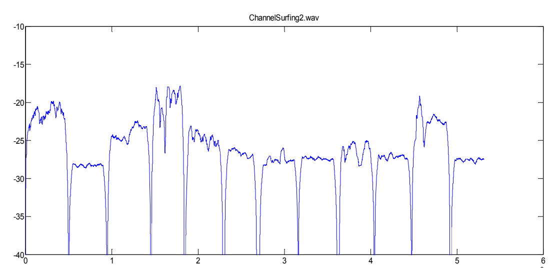

Figure 7: Channel “Surfing” Loudness Plot

The gaps in the plot show when the audio is muted during each channel change. One can see that between the first and second programs an approximately 8 dB difference in perceived loudness exists. We have acquired recordings with channel-to-channel differences in excess of 20 dB. While techniques have been proposed for normalizing audio levels in digital broadcasts, there still appears to be wide disparity in broadcast audio

levels. The situation may worsen as DTV broadcasts become more prevalent.
There is also a need, in certain situations, for volume control technologies to compress the dynamic range of television audio. Such technologies should have night settings that enable the viewer to listen to audio at a low volume and hear all the intended audio (both loud and soft sounds) while not bothering other people nearby. Figure 8 shows an ITU perceived loudness plot of the same channel change recording after being level adjusted by compressor-based volume control technology. Note that while the overall loudness level has been adjusted the nature of the audio transients within each program has been substantially maintained.

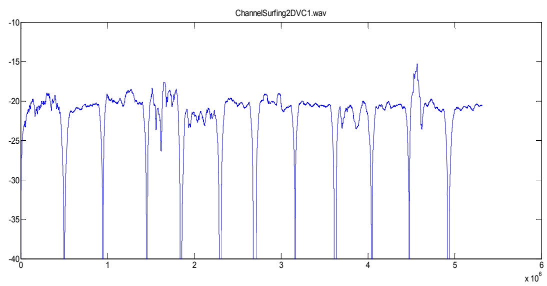

Figure 8: Channel "Surfing" After Volume Control Loudness Plot

Volume control technology of this type has minimal processor throughput and memory requirements and integrates well into current television audio processing platforms. Care must be taken to choose characteristics of the compressor such as compression ratio, time constants, etc. to achieve a leveling of the perceived audio while avoiding traditional artifacts such as pumping and breathing.

Figures 9 and 10 show the before and after plots of volume control processing on an extreme channel change acquired separate from the random data gathered for statistical processing.

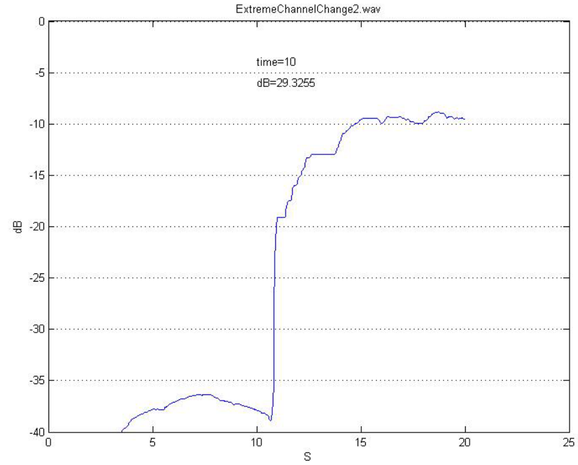

Figure 9: Loudness Plot of Extreme Channel Change Before Volume Control

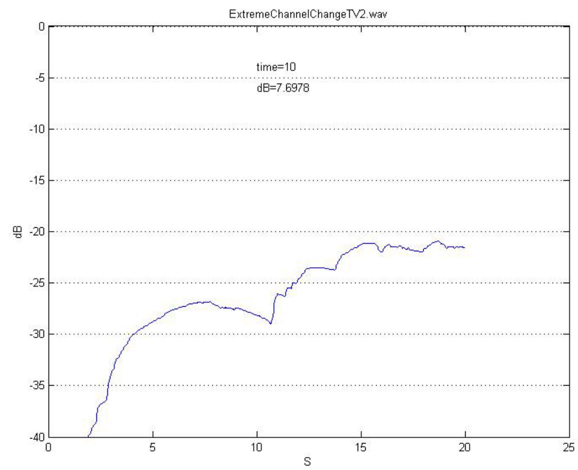

Figure 10: Loudness Plot of Extreme Channel Change After Volume Control

6. CONCLUSION

Despite efforts to normalize broadcast television audio at the source, large differences in perceived audio levels continue to plague television viewers. Irritatingly large loudness changes are often perceived during program-to-commercial transitions and while changing channels. Many times these changes are large enough to necessitate quick adjustment of the television volume control to avoid long-term discomfort. These level differences have been observed and an attempt has been made to quantify the relative loudness levels and frequency of occurrence.

This common problem has been demonstrated to be minimized with appropriate volume control technology in the television receiver.

Our effort to acquire and quantify two-channel television broadcast audio levels is an ongoing work. We have identified several areas where we may improve our acquisition and analysis. First we would like to acquire and analyze more program material before and after the channel transition. Given the necessary 5-second loudness meter measurement window, longer recordings (at least 60 seconds) seem more appropriate. Some of our recordings only had 15 seconds of pre and post transition audio. Hundreds more recordings were rejected because of their shortness. In addition, we should prescreen each channel change recording to verify that within each program there is not a commercial transition. The presence of commercial transitions can skew the channel change data.

We also plan to continue the manual work of acquiring a large number of program-to-commercial transition recordings for further analysis. To demonstrate the effects of multiple spatial processing and perceived loudness, we plan to acquire recordings after virtual surround processing is performed on the television audio output. With the slow phase-out of analog television broadcasts, follow-up efforts will also concentrate on digital broadcasts.

7. REFERENCES

[1] Algorithms to measure audio programme loudness and true-peak audio level. International Telecommunication Union RECOMMENDATION ITU-R BS.1770-1. Geneva, Switzerland: 2006-2007.