

Assessment of Automatic Volume Leveling for Automotive Sound Systems

Robert Cadena and Husein Dakroub
Visteon Corporation

ABSTRACT

This paper presents an assessment of competing algorithms for normalizing volume levels between tracks and/or sources in an automotive infotainment system. Portable media players such as smartphones and iPod® devices are extremely popular for listening to music collections or streaming content from the Internet. The lack of normalization is a source of dissatisfaction if the user experiences significant changes in audio level between tracks. Several commercially available algorithms exist to solve this problem. This research includes a double-blind listening test comparing an audio sample processed with the different leveling algorithms to an unprocessed reference. The listener preference rating is recorded and results indicate which algorithm is preferred.

CITATION: Cadena, R. and Dakroub, H., "Assessment of Automatic Volume Leveling for Automotive Sound Systems," *SAE Int. J. Passeng. Cars – Electron. Electr. Syst.* 6(1):2013, doi:10.4271/2013-01-0162.

INTRODUCTION

In the automotive environment, achieving consistent playback level for music and information is increasingly difficult. Traditional audio sources such as FM and CD have achieved track to track consistency at the source by processing audio data before transmission to the airwaves or pressing on the manufacturing line, respectively. More recent sources such as portable media players and smartphones allow users to store large number of tracks and play them back in random order. This can result in large discrepancies in track to track playback level. To reduce this variation, algorithms can be employed in automotive sound systems which dynamically adapt the playback gain in response to changes in the source level.

Commercial algorithms for automatic volume leveling are available from several companies. This paper describes testing carried out to identify the best algorithm for addressing the situation described above.

AUTO-LEVELING TESTING AND RESULTS

What is automatic volume leveling? Changing sources in the car can have an effect on a listener's experience. For example, the sound from a CD has a different output level, dynamic range and frequency response than sound playing from a user's device utilizing the auxiliary input, a Bluetooth connection or traditional radio. Moreover, changing between

these sources and listening to different genres can lead listeners to adjust the volume knob multiple times during a routine drive to find a comfortable listening level. To address this issue, automatic leveling algorithms have been developed that ideally provide the following features:

- The algorithm shall level the volume from song to song, and from genre to genre.
- The algorithm shall make quieter songs louder and/or louder songs quieter, without affecting the dynamic range within every song, meaning when a song has a high dynamic peak followed by a low passage, the algorithm should not compress this range.
- The algorithm shall maintain a smooth transition between songs. When playing through a playlist, the attack and release of the volume envelope shall not be perceived by the listener. Ideally, the leveling should take place before any sound is heard.
- When leveling is in effect, there shall not be any change in quality of the sound. Any artifacts resulting in additional distortion or changes in frequency response should be avoided. The quality of the original track shall be maintained.

Test Sequence

The testing process was broken into phases which are described in detail in the following paragraphs.

- 1. Reference Test Track** A number of songs were chosen and a test file was created for the testing process.
- 2. Test Track Processing** The test file was processed by each algorithm and the results were saved in separate .wav files.
- 3. Ambient Noise** A live recording was made of the ambient cabin noise level while driving on a freeway.
- 4. Test Design** The test environment and test tool was selected.
- 5. Pilot Test** A pilot was conducted to insure the testing procedure had no major issues and to verify that the sound pressure level during testing was appropriate.
- 6. Participant Selection and Questionnaire** Participants were solicited to participate in the testing, and were requested to answer a pre-test questionnaire.
- 7. Test Administration** The test equipment was set up in a quiet listening room, and the test was conducted with participants.
- 8. Results and Analysis** The results were collected, analyzed and correlated to the answers from the pre-test questionnaire.

Table 1. Leveling Test Files

Name	Artist	Genre
"Girl"	Beck	Electronic Rock
"Life Wasted"	Pearl Jam	Hard Rock
"Lighters"	Bruno Mars	Vocal Pop
NPR News	NPR	News Station
"In and Out of Love"	Armin Van Buuren	Techno
"So What"	Miles Davis	Jazz
"Trusted"	Ben Folds	Pop Rock

Reference Test Track

A test file was created for the evaluation. This test file consisted of six songs plus one segment of talk news that represented a variety of music and talk genre. The songs covered genre which many people listen to on a daily basis; rock, jazz, talk radio, techno, and vocals. The songs were also chosen because they featured either a loud or quiet introduction. The first nine seconds of each song were used in the test file to represent a user selecting the next track from a music player in random mode. The total length of the test track' was one minute to limit listener fatigue. Test participants needed to be comfortable and able to listen to the entire track before making a judgment. After an extended

period of music listening, listeners start to fatigue and lose interest in the testing process. Because six algorithms were being tested, a test track length of one minute allowed participants to make conclusions while minimizing fatigue.

Test Track Processing

Each company provided a method of processing the reference track with their algorithm to create a test track for evaluation. Some used a PC-based graphical user interface (GUI). These applications processed the reference test track and produced an output files which was processed by the algorithm. Others used an evaluation board with digital signal processing (DSP) targeted for automotive audio processing. In this setup, analog audio of the reference test track was fed from a PC to the evaluation board. The file was processed by the DSP and the audio was sent back to the PC sound card. To record the output while processing each algorithm, Reaper audio recording software was used.

Ambient Noise

To achieve a listening experience similar to one a subject would have in his or her car, ambient noise was added. A HEAD Acoustics binaural dummy head was placed in the front passenger seat of a 2010 Ford Fusion sedan. It contained microphones at its eardrum positions and provided line outputs for recording audio to outboard equipment. The dummy head had a sound pressure meter used for calibration. The line outputs for the binaural head were connected to the inputs of M-Audio Audiophile USB sound card to make recordings while driving.

The road noise conditions chosen for the test were driving on an interstate expressway at a speed of 70 mph. During this time, the ambient noise was measured between 84 to 86 decibel sound pressure level (dB SPL). Next, the radio was set at a volume level which allowed the music to be heard clearly above the ambient noise level. This sound level was measured to be at 93 to 94 dB SPL during music programs, and 89 dB SPL for a talk show program. This information was used to determine the appropriate ratio of music signal to noise to create the test files.

The ambient noise recording was mixed with each test file at the level of -12.7 dB relative to the test files. This was done to simulate the observed difference between the music and noise signals in during the vehicle recordings. The mixing was performed using Reaper Audio Recording Software as shown in [Figure 1](#) below.

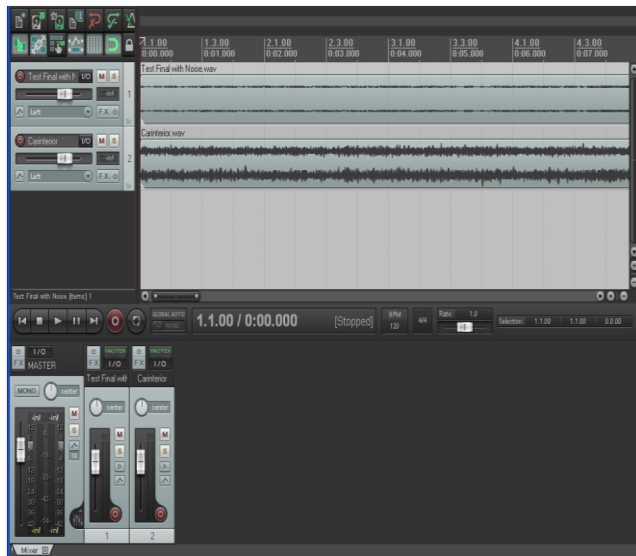


Figure 1. Reaper Audio Recording Software

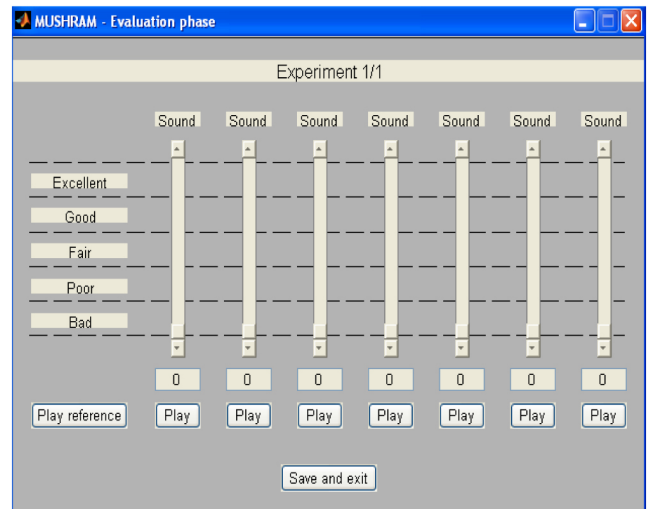


Figure 2. MUSHRAM Test Interface

Test Design

The testing was held in a listening lab located on the Visteon Corporation campus in Van Buren Township, Michigan. There was minimal ambient noise in the room, which was measured to be 56 dB SPL using the dummy head. A table and chair was located in the center of the room containing STAX Professional ear speakers and a laptop to conduct the testing.

For the evaluation, the MATLAB program MUSHRAM [1] “Multi Stimulus Test With Hidden Reference And Anchor” was used. MUSHRAM was chosen because it facilitates double blind testing in which neither the participant nor the test administrator has knowledge of which test file is being played. The MUSHRAM GUI is shown below in Figure 2. The GUI consists of several buttons and sliders. One button is used to play the reference file. The remaining play buttons are used to play the test files. The number of play buttons is one greater than the number of test files because the reference file is also hidden as one of the files that must be evaluated. While listening to each test file the listener must rate the file between 0 and 100. At least one file must be given a score of 100 which represents best in class leveling performance.

The listener was able to listen to the different tracks in any order and multiple times. Once the track was selected, the listener was required to listen to the selected track from beginning to end. The listener also had the option to listen to the reference file at any time before, between or after listening to the test files.

Pilot Test

The test was piloted by the authors and several additional volunteers to ensure that the test flow was smooth and stable. It was desired that the playback be both comfortable and representative of the playback SPL observed during the vehicle noise recording drives. At the conclusion of the pilot phase, it was agreed to set the maximum test program SPL at 91 dB SPL which resulted in an approximate average level of 86dB SPL. The pilot subjects stated the ear speakers were not too loud, and the listeners were comfortable enough to make judgements on the different tracks.

Participant Selection and Questionnaire

The test subjects for this study were Visteon Corporation employees. 33 individuals participated in the test. A pre-test questionnaire was sent to each of them asking them the following:

- Gender
- Age
- Job function
- Car make/model
- If car had premium sound

Additionally, they were asked to rate on a scale of 1-8:

- if they have/had hearing issues,
- if they were frequent music listener
- if they were experts in sound quality
- which music genre they listened to
- if they had a loud or quiet car
- if they were satisfied with the sound quality in their vehicle