

Wahrnehmungsbasierte Pegelkorrektur von Audiosignalen

Wolfgang Heß

Harman/Becker Automotive Systems, 76307 Karlsbad, E-Mail: wolfgang.hess@harman.com

Einleitung

Bei der Wiedergabe von Informationen und Unterhaltungssignalen im Fahrzeug überlagern sich Fahrgeräusche und vom Fahrzeug-Soundsystem erzeugte Signale. Damit diese, unabhängig von Fahrsituation und Signalquelle, mit einer vorgewählten Lautheit wiedergegeben werden können, sind sowohl eine Korrektur der Audiosignale auf eine mittlere Lautheit, als auch eine Anpassung an die Fahrgeschwindigkeit vorzunehmen. Die hier beschriebene Methode justiert die Audio-Eingangssignale auf eine mittlere Lautheit, sodaß die Signale innerhalb eines Lautheitskorridors ohne deutlich hörbaren Verlust von Dynamik wiedergegeben werden.

Lärmerfüllte Umgebung

Störgeräusche im Fahrzeug verändern sich mit sich ändernder Fahrgeschwindigkeit, sowohl bezüglich ihrer Lautheit als auch ihrer Klangfarbe. Bei niedrigen Geschwindigkeiten überwiegt meist das Antriebs- oder Reifen-Fahrbahn-Geräusch, bei hohen Geschwindigkeiten das Geräusch der Luftumströmung, siehe bspw. [1]. Dabei werden maximale Geräuschwerte von 70 bis über 90 dB SPL erreicht, abhängig von Fahrzeugtyp und Fahrgeschwindigkeit. Aktuelle Serien-Soundsysteme in Fahrzeugen bieten maximale Wiedergabepegel von weit über 100 dB SPL, wobei die Grenzen für Hörschädigung bei Dauereinwirkung, etwa 85 dB SPL, zu berücksichtigen sind. Effektiv stehen daher während des Fahrens für die Wiedergabe von Schall über das Soundsystem etwa 20-40 dB als Dynamikbereich zur Verfügung.

Signalquellen und Lautheit

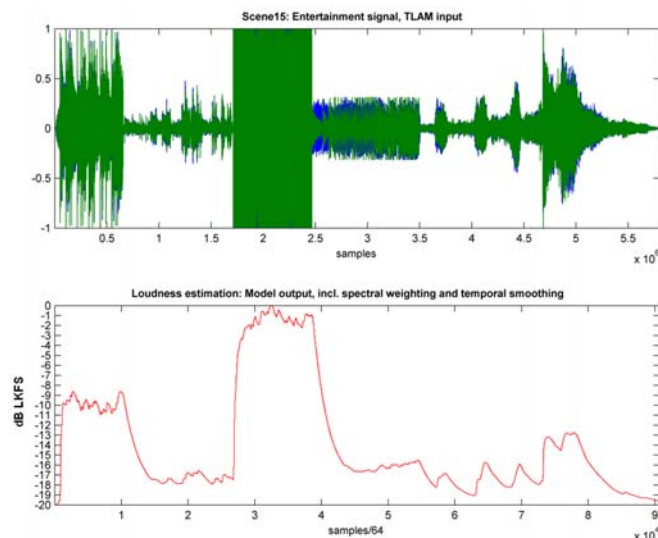


Abb. 1: Zusammenstellung von Audiodateien von CD.

Aus der Möglichkeit, Dateien von USB-Stick oder MP3-Player einzuspeisen, durch verschiedene Sendepiegel und Dynamikkompressionen von Radiostationen, durch eine Vielfalt von Modulationshüben bei Radioübertragung und andere Möglichkeiten entsteht ein weiter Bereich von Sig-

nalpegeln und -lautheiten, die vom Audiosystem verarbeitet werden müssen. Abbildung 1 zeigt eine Zusammenstellung von einigen kurzen Audiobeispielen von verschiedenen CDs. Der mittlere Teil, der in der Lautheitsschätzung als nahezu maximal laut wahrgenommen wird, siehe Abb. 1 unten, wurde einer aktuellen Pop/Rock-CD entnommen, der hintere Teil einer Klassik-CD, der vordere Teil einer Jazz-CD.

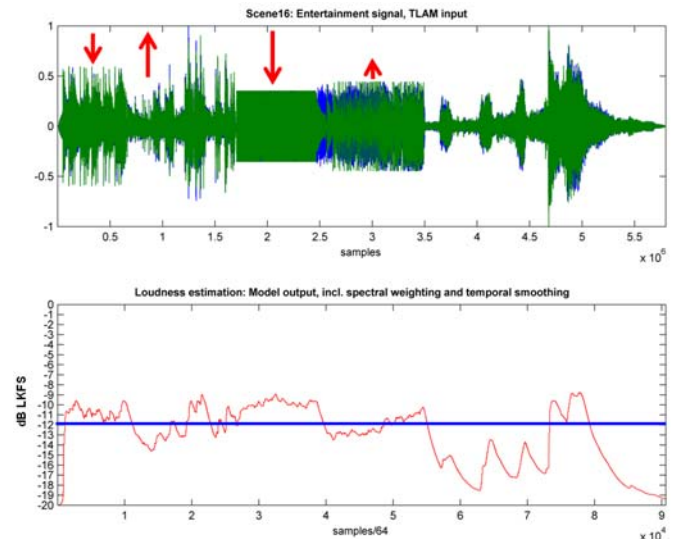


Abb. 2: Manuelle Korrektur für eine gleichbleibende mittlere Lautheit bei vollem Erhalt der Dynamik.

Ist die dynamische Struktur bekannt, so kann nach Berechnung der Lautheit durch einfache manuelle Pegelkorrektur bei vollständigem Erhalt der Dynamik eine gleichbleibende mittlere Lautheit erzielt werden, wie in Abb. 2 dargestellt.

Dynamische Lautheitskorrektur

Bei unbekannter dynamischer Struktur bzw. kurzer Vorschau, systembedingt in der Echtzeit-Signalverarbeitung, ist der Erhalt der Dynamik nicht ohne weiteres möglich, da nur der aktuelle Teil des Musikstücks analysiert werden kann. Eine weitere Herausforderung bei sehr leisen Eingangssignalen ist die Unterscheidung zwischen sehr leisem Signal und Signalpausen: Rauschen soll nicht verstärkt werden.

Lokalisations- und Lautheitsschätzung

Zur Lösung dieser Herausforderungen wurden Annahmen gemacht, die die menschliche Lokalisation als wesentlichen Eingangsparameter nutzen: Kann Schall lokalisiert werden, so ist Nutzschall vorhanden. Systembedingtes Rauschen kann nicht lokalisiert werden. Wie in [3], basierend auf [4], gezeigt, ist die Lokalisation in Räumen von Raumform etc. abhängig. Wie in real vorhandenen Räumen kann Schall in virtuellen Räumen wiedergegeben und lokalisiert werden, siehe [2], d.h., ebenso können die in den Audiosignalen kodierten räumlichen Informationen genutzt werden. Dadurch entsteht binaurale Aktivität, die durch räumliche Information eine genaue Anpassung der Lautheit erlaubt. Ab-

Abbildung 3 zeigt das Ergebnis der Berechnung eines auditorischen Bandes eines Lokalisationsmodells (in Abbildung 4 als „Psychoacoustical model“ bezeichnet). Im dargestellten Beispiel sind bei einer Laufzeit von 0 bis ca. 300 ms, und von ca. 800 ms bis 1000 ms Signale lokalisierbar, dazwischen ist eine Signelpause. Über die ganze Dauer wurde Rauschen hinzugefügt, das jedoch nicht lokalisierbar ist, und folglich in der Modellberechnung zu keiner binauralen Aktivität führt. Nach einer bestimmten Dauer binauraler Inaktivität wird eine Signelpause detektiert. Ebenfalls wird durch die Auslenkung der binauralen Aktivität, siehe Abb. 3 linke Seite, die Lautheit eines Signals geschätzt, äquivalent zu [5].

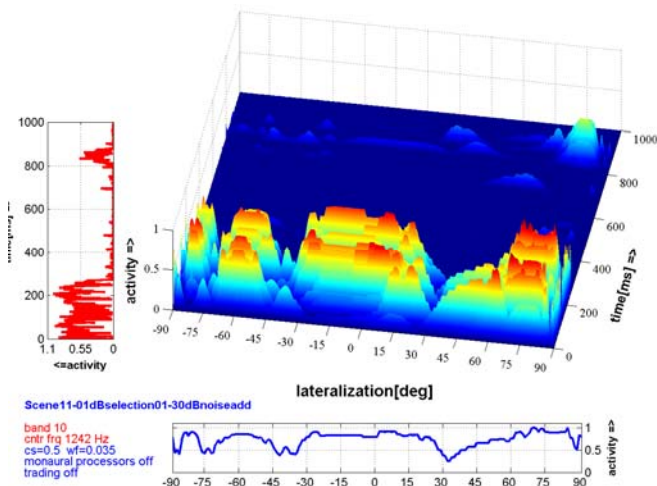


Abb. 3: Lokalisationsschätzung durch das binaurale Modell: Binaurale Aktivität und zeitliches Verhalten.

Dynamik

Detektierte Signelpausen führen zum Zurücksetzen von adaptiven Zeitkonstanten, die für eine sehr träge Änderung bei abnehmender Lautheit und eine schnelle Reaktion bei zunehmender Lautheit sorgen. Mit zunehmender Dauer des Musikstücks werden diese Zeitkonstanten, angewandt auf die Lautheitsschätzung, träger, sodaß die dynamische Struktur des Musikstücks weitestgehend erhalten bleibt.

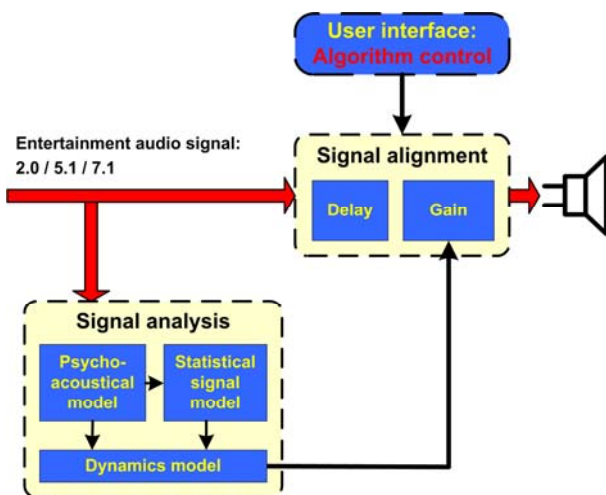


Abb. 4: Blockschaltbild der dynamischen Lokalisations- und Lautheitsschätzung mit automatischer Steuerung der Pegelkorrektur, einstellbar in der Wirkung durch Nutzer.

Diese Zeitkonstanten werden im Dynamikmodell, siehe Abbildung 4, nachgeführt. Durch die Verzögerung der Signale durch Verzögerungselemente kann nach erfolgter Signalanalyse, siehe linker Teil Abb.4, eine Nachregelung des Verstärkungsfaktors, mittels Rampen geglättet, vorgenommen werden, siehe Abb. 4 rechter Teil. Ein mittlerer Zielpegel sorgt dafür, daß ein bestimmter Dynamikbereich zur Verfügung steht. Weitere Abstimmung ist nicht notwendig.

Die Stärke des Eingriffs der automatischen Pegelkorrektur kann vom Benutzer in Bereich 0 bis 100% eingestellt werden. Bei voller Wirksamkeit des Algorithmus entfällt manuelles Nachjustieren der Wiedergabelautstärke durch Benutzer. Eine Anpassung an das Fahrgeräusch wird im Postprocessing durch die geschwindigkeitsabhängige Lautstärke (GAL) realisiert. Da keine Filter im Signalpfad wirksam sind, finden keine Klangfarbenänderungen statt; fehlende nichtlineare Dynamikkompression (Kompressor/Limiter) und eine quasi-lineare Pegelkorrektur sorgen dafür, daß keine Artefakte wahrnehmbar sind.

Zusammenfassung

In dieser Arbeit wird eine wahrnehmungsbasierte Methode zur automatischen Pegelkorrektur von Signalen präsentiert, die in ein Audiosystem eingespeist werden. Das Ziel, Signale auf eine mittlere Ziel-Lautheit zu bringen, ohne dabei die dynamische Struktur eines Musikstücks deutlich zu verändern, wurde erreicht. Ein binaurales Lokalisationsmodell nimmt Signelpausendetektion und Lautheitsschätzung vor. Gerade bei sehr leisen Signalen an der Schwelle zum Rauschen hat sich diese Methode als sehr brauchbar erwiesen: Nur wenn binaurale Aktivität in der Lokalisationsebene des binauralen Modells vorhanden ist, liegt ein Nutzsignal vor.

Der Algorithmus regelt kontinuierlich den Verstärkungsfaktor der Signalquelle innerhalb eines Toleranzbereiches nach, sodaß weder Modulation noch ein deutlicher Verlust von Dynamik entstehen. In Fahrzeugen gewährleistet dieses Verfahren eine an das Fahrgeräusch angepaßte gleichbleibend hohe Verständlichkeit von Audiosignalen.

Literatur

- [1] Helfer, M. (2007): „Reifen-Fahrbahn-Geräusch und Umströmungsgeräusch von Kraftfahrzeugen“, Fortschr. der Akustik, DAGA 2007, Stuttgart, 743-744
- [2] Hess, W., J. Braasch und J. Blauert (2003): „Acoustical evaluation of virtual rooms by means of binaural activity patterns“, Proc. of the AES 115th Convention, New York
- [3] Hess, W. (2004): „Examination of Binaural Activity Patterns“, Fortschr. der Akustik, SFA/DAGA 2004, Strasbourg
- [4] Lindemann, W. (1986): “Extension of a binaural cross-correlation model by contralateral inhibition. I. Simulation of lateralization for stationary signals”, J. Acoust. Soc. Am., Vol. 80, 1608-1622
- [5] ITU-R Recommendation BS.1770-1 (2006-2007): “Algorithms to measure audio programme loudness and true-peak audio level”, International Telecommunications Union, Geneva.