

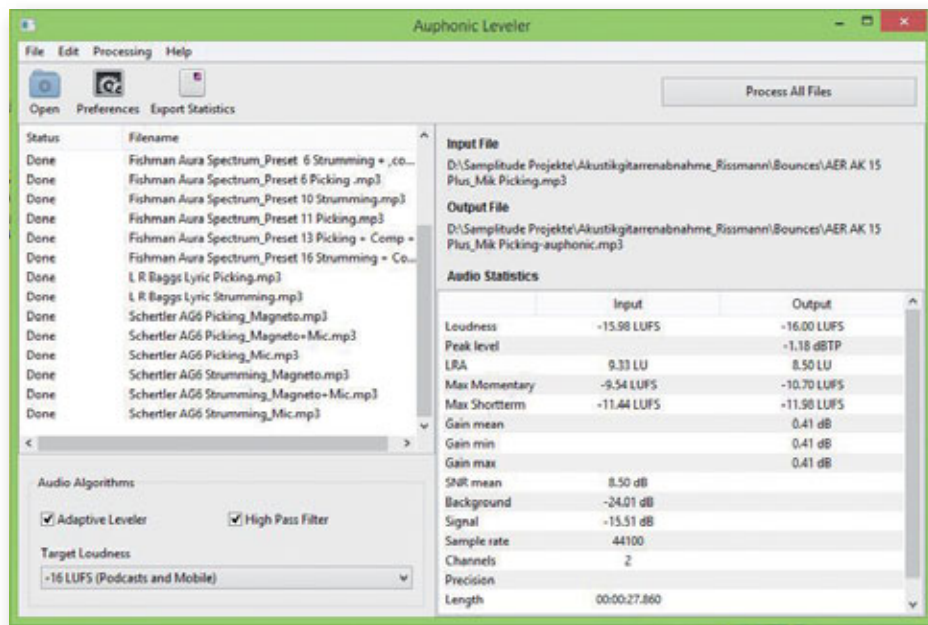
Gleichklang

Auphonic „Leveler“ Audio File Processor



Von Michael Nötges

Klar, dass die neuen Loudness-Standards für die Rundfunk- und Fernsehanstalten neben Messgeräten wie der TM-Serie von RTW (Test in Ausgabe 5/2013) auch neuen Software-Tools zur Loudness-Anpassung den Weg auf den Markt ebnet. Der „Leveler“ von Auphonic wird als intelligenter und obendrein lernfähiger Batch Audio File Processor angeboten, welcher unterschiedliches Audiomaterial analysieren, kontrollieren und vor allem dessen Loudness angleichen soll.



Die zu bearbeitenden Files lassen sich per Drag-and-Drop in das „Leveler“-Fenster ziehen; nach der Bearbeitung erscheinen im rechten Fenster zu jedem File Statistiken, die über die vorgenommenen Veränderungen informieren

Angefangen hat die junge Audio-software-Firma Auphonic aus Graz mit viel Online-Service zur automatisierten Post-Produktion von Audiomaterial. Das Prinzip ist dabei denkbar einfach: Audiofiles werden nach Registrierung auf der Aupho-

nic-Homepage hochgeladen, von den proprietären Algorithmen analysiert und anschließend mit intelligenten Bearbeitungsverfahren und in Übereinstimmung mit den Regeln aktueller Loudness-Standards (R-128, ATSC A-85) finalisiert. Dabei soll

es dem User so einfach wie möglich gemacht werden, sein Podcast, Interview, die fertige Radiosendung im Handumdrehen sendetauglich zu bekommen. Egal, ob das Material später bei YouTube, Soundcloud oder anderen Netzdiensten beziehungsweise im Rundfunk- oder bei Fernsehanstalten eingesetzt werden soll. Was bisher über ihren Online-Dienst funktionierte, wurde jetzt in einer kompakten Software-Anwendung (App) für Windows Vista, 7 und 8 sowie MAC OS 10.6 oder höher zusammengefasst. Kern des Auphonic „Levelers“, so heißt das Tool, sind dessen Algorithmen zur Analyse und Bearbeitung des Audiomaterials. Dazu gehören der Adaptive Leveler, eine Loudness-Normalisierung, ein True Peak Limiter und ein Hochpassfilter.

Georg Holzmann, Audio Engineer und Computerspezialist bei Auphonic, verrät den Clou der neuen Anwendung: „Der speziell entwickelte Adaptive ‚Leveler‘ analysiert den Inhalt von Audiosignalen mittels Techniken des maschinellen Lernens und führt danach intelligente Lautheitsanpassung innerhalb einzelner Dateien durch.“ Was im Klar-

text heißt, dass Musik, Hintergrundgeräusche und beispielsweise ein Sprecher als Einzelteile oder Segmente eines Audio-Files erkannt werden.

Je nach Klassifizierung kommen dann die unterschiedlichen Bearbeitungsalgorithmen zum Zug. Damit wird Sprache anders gehandhabt als Musik, Hintergrundgeräusche wieder anders als Rauschen oder Netzbrummen. Das Ergebnis besteht laut dem zugrunde liegenden Ansatz darin, dass der „Leveller“ Lautstärkeunterschiede innerhalb eines Files ausgleicht und gleichzeitig die Dynamik, wenn sinnvoll, durch Kompression begrenzt.

So werden auch Loudness-Unterschiede zwischen Musik- und Sprachsegmenten intelligent angeglichen. Hintergrund und Ruheperioden dagegen sollen erkannt und eben nicht verstärkt werden, um das Anheben des Noisefloors und ablenkender Geräusche zu minimieren.

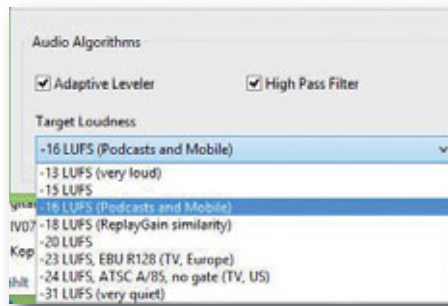
„Unser Algorithmus“, erklärt Holzmann, „wurde über zwei Jahre mit Audiodaten unseres Webservices trainiert und lernt jeden Tag neu dazu. Das ermöglicht eine komplett neuartige Kombination von maschinellem Lernen, Signalverarbeitung und Big Data in einer Audiosoftware.“ Dabei soll der „lernfähige Gleichmacher“ vor allem Lautstärkeunterschiede zwischen unterschied-

lichem Programm-Material ausgleichen, wobei die Software besondere Vorsicht bei Musik walten lassen und die natürliche Dynamik beibehalten soll.

Loudness und True Peak

Für die Loudness-Normalisierung lässt sich ein geeignetes Target Level definieren (siehe Screenshot), wobei -23 LUFS nach der R-128 genauso im Angebot sind wie -24 LUFS für den ATSC A/85-Standard oder -16 LUFS für Podcasts oder mobile Endgeräte. Zur Auswahl stehen aber auch Extremeinstellungen von -13 (sehr laut) bis -31 (sehr leise) LUFS. Der maximale True Peak Level ist zusätzlich einstellbar (Preferences), um eine wirklich übersteuerungsfreie Übertragung gewährleisten zu können. Im Auto-Betrieb wird der Parameter automatisch den einheitlichen Standards (R-128 = -1dBTP und ATSC A/85 = -2 dBTP) angepasst. Händisch sind weitere Werte zwischen -9 und 0 dBTP wählbar. Der integrierte Ausgangs-Limiter bietet zwei Messverfahren: das einfache Sample Peak und zusätzlich das präzise True Peak Measurement mit vierfachem Oversampling.

Da in der Verordnung ITU BS. 1770-2 das Verwenden eines LUFS-Gates untersagt ist, kann dieses komplett ausgeschaltet werden. Ansonsten ist ein Wert von -10 LUFS vorge-



Die „Target Loudness“ lässt sich zwischen -31 und -13 LUFS wählen; die Werte von -23 und -24 LUFS sind für den EBU R-128- und ATSC A/85-Standard voreingestellt

sehen. Im Auto-Modus wiederum wird es dem ausgewählten Standard entsprechend zugelassen oder nicht, um die Standardkonformität zu gewährleisten.

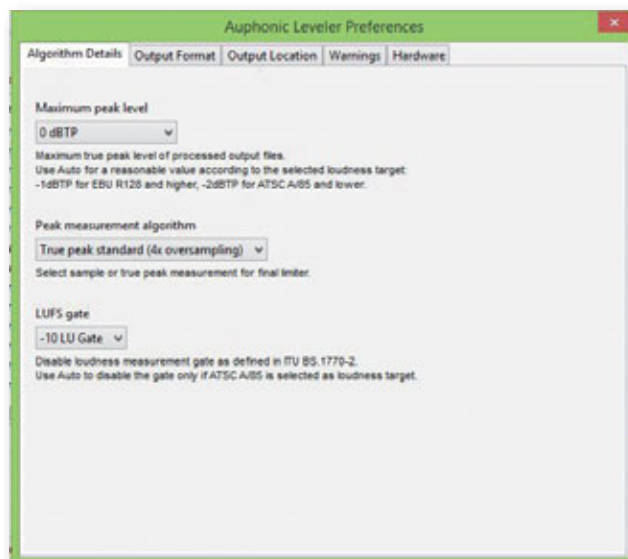
Filter und Formate

Je nach Kontext werden tiefe Frequenzen automatisch von einem HPF weggefiltert. Außerdem verfügt der „Leveller“ über eine integrierte Noise- und Hiss-Reduction. Für die Analyse wird das Audiomaterial in Segmente unterteilt, die unterschiedliche Hintergrundgeräusch-Charakteristiken haben. Für jede dieser Regionen wird dann ein sogenannter Noise-Print erstellt, um anschließend die unliebsamen Störgeräusche aus dem Signal herausrechnen zu können.

Zur Minimierung von Netzbrummen kommen neben dem Noise-Reduction-Verfahren noch Notchfilter



Für fünf Werte (Mean Leveler Gain, Input SNR, Output Max S, Output Max M und die LRA) lassen sich Thresholds definieren; werden diese bei der Bearbeitung überschritten, erscheint ein Warnhinweis



Es stehen unterschiedliche Werte für den Maximum Peak Level zur Auswahl; unter anderem -1 dBTP für den R-128 und -2 dBTP für den ATSC A/85-Standard

GLOSSAR

Fakten

Hersteller: Auphonic

Modell: „Leveler“

Typ: Desktop Batch Audio File Processor

Betriebssysteme: Win Vista, 7, 8 (32/64 Bit); MAC ab OS 10.6 (64 Bit)

Algorithmen: Adaptive Leveler, Loudness Normalization nach neuesten Loudness-Standards (EBU R-128, ATSC a/85), True Peak Limiter, High Pass Filter

Funktionen: Batch-Processing, detaillierte Bearbeitungsstatistiken, Threshold-Einstellungen für Warnhinweise

Input-Formate: WAV, AIFF, FLAC, MP3, Ogg Vorbis (Mac OS X auch: MP4/M4AM4B, AAC, ALAC, CAF, AC3, MP2, 3GP)

Output-Formate: WAV, AIFF, MP3 (via Lame), FLAC, Ogg Vorbis (Mac OS X auch: AAC/M4a)

Sampling-Frequenzen: 44,1 und 48 kHz

Bitraten: 16, 24, 32 Bit

Lizenz für Privatperson: 59 Euro

Kommerzielle Lizenz: 279 Euro

www.auphonic.com

Pro & Contra

- + Algorithmus verbessert sich kontinuierlich durch maschinelles Lernen
- + Anpassung zur optimalen CPU- und RAM-Auslastung verhindern Aussetzer
- + Bearbeitungsstatistiken und Warning-Thresholds zur Qualitätskontrolle
- + intuitive selbsterklärende Bedienung
- + konform mit R-128- und ATSC A/85-Standard
- + Loudness-Anpassung und Komprimierung, ohne Störgeräusche zu betonen
- + schnelle Batch-Bearbeitung möglich
- + vorsichtige Bearbeitung von Musik unter Beibehaltung der natürlichen Dynamik
- Probleme bei der Analyse und Anpassung mehrerer gleichwertiger Hintergrundgeräusche

ATSC A-85: Die Abkürzung steht für eine Empfehlung des Advanced Television Systems Committees (nordamerikanisches Pendant zu DVB: Digital Video Broadcasting/Digitaler Videorundfunk). Die Empfehlung A-85 bietet Richtlinien zur Audio-Messung in der Produktion und Postproduktion sowie Methoden und Anzeigetechniken, um die Loudness für gelieferten oder ausgetauschten Content effektiv zu kontrollieren und anzupassen. Außerdem beschäftigt sie sich mit Metadaten-Systemen und beschreibt moderne Möglichkeiten der Kontrolle des Dynamikumfangs.

EBU: Die European Broadcast Union oder Europäische Rundfunkunion ist ein Zusammenschluss von 74 Rundfunkanstalten aus Europa, Nordafrika und Vorderasien mit Sitz in Genf. Ziel des 1950 ins Leben gerufenen Zusammenschlusses war es, ein Netzwerk zum Austausch von Nachrichtenfällen aufzubauen. Die EBU treibt technische Entwicklungen im Radio- und Fernsbereich voran und versucht diese zu standardisieren. Beispielsweise stehen R-68 oder R-128 für festgelegte Normen der EBU.

EBU R128: Die Empfehlung R-128 der EBU ist ein technisches Regelwerk, welches die Tonaussteuerung von Hörfunk- und Fernsehprogrammen neu definiert und damit den Paradigmenwechsel im Bereich des Audio Meterings von PPM zur Loudness-Anzeige einleitet.

ITU: Die International Telecommunication Union (Internationale Fernmeldeunion) beschäftigt sich und weltweit mit technischen Aspekten der Telekommunikation.

ITU-R: Der Radiocommunication Sector der ITU beschäftigt sich mit der internationalen Festlegung der Zuweisung von Frequenzbereichen an Funkdienste durch die Vollzugsordnung für den Funkdienst.

ITU BS.1770-3 und BS.1771-1: Die Abkürzung „BS“ steht jeweils für Broadcasting Service (sound) des Radio-communication Sectors der ITU. Die Empfehlung 1770-3 vom August 2012 beschreibt die Algorithmen zur Messung der Programm-Loudness und des True Peak Audio Levels. In der Empfehlung 1771-1 sind die Anforderungen für Loudness- und True Peak Meter festgelegt.

Lautheit (Loudness): Die Lautheit ist eine festgelegte Größe zur proportionalen Abbildung des menschlichen Lautstärkeempfindens mit der Maßeinheit Sone. Der Begriff Lautheit stammt aus der Psychoakustik und beschreibt, wie eine Anzahl von Testpersonen die empfundene Lautstärke von Schall überwiegend beurteilt. Im Wesentlichen hängt die Lautheit dabei vom Schalldruckpegel, dem Frequenzspektrum und dem Zeitverhalten des Schalls ab. Die Schallverarbeitung im Innenohr bestimmt dabei das Lautheitsempfinden eines Signals. Je stärker die Erregung der Nervenzellen ist, desto lauter wird das Geräusch wahrgenommen. Heute gibt es Modelle zur Lautstärkewahrnehmung des Menschen und damit technische Geräte, die die Lautstärkewahrnehmung des Menschen simulieren und dadurch auch messbar machen. Die genormten Messverfahren zur Lautheitsmessung sind in DIN 45631 und ISO 532 B zu finden.

LRA: Die Loudness Range beschreibt einen statistisch ermittelten Wert, der die Dynamik eines Programmabschnitts beschreibt und anzeigt, ob ein Signal komprimiert werden sollte oder aber zu stark komprimiert ist.

LUFS: Die Einheit LUFS (Loudness Unit Full Scale) ist eine Einheit für absolute Lautheit, die in der EBU-Empfehlung R-128 eingeführt wurde.

zum Einsatz, um das Material zu bereinigen. Das geschieht automatisch, ohne dass der User Einfluss auf die Art der Bearbeitung hat.

Der Leveler beschränkt sich auf die Bearbeitung von Standard-Formaten. Eingangsseitig lässt er WAV-, AIFF-, FLAC-, MP3- und OGG-Files zu, wobei auf dem Mac zusätzlich die Formate MP4/M4A/M4B, AC3, AAC, ALAC, CAF, MP2 und 3GP unterstützt werden. Ausgegeben werden allerdings nur WAV-, AIF-, MP3-, FLAC und OGG-Files, wobei auf dem Mac noch das AAC-Format (M4A) dazu kommt. Der Leveler unterstützt Sampling-Frequenzen von 44,1 und

48 kHz und Wortbreiten von 16, 24 und 32 Bit. MP3s werden mit 32 bis 320 kbps ausgegeben. Bei den Resampling-Algorithmen setzt der Hersteller auf die Erfahrung des Open Source Software Projekts SoX.

Überwachung

Der „Leveler“ ist auch ein Batch-Converter, der große Mengen Audiofiles bearbeiten kann. Um die CPU- und RAM-Auslastung in der sogenannten „task queue“ dem jeweiligen System anpassen zu können, lässt sich die Anzahl der parallel zu bearbeitenden Files festlegen. Bei zwei CPUs empfiehlt sich nicht mehr als

eine doppelte Parallelverarbeitung. Außerdem lässt sich die Blockgröße je nach zur Verfügung stehendem RAM bestimmen, um Fehler bei der Bearbeitung zu vermeiden.

Zur Überwachung der Bearbeitungsvorgänge gibt es detaillierte Statistiken (siehe Screenshot) der Bearbeitungsvorgänge, welche zur Information und Qualitätsüberwachung hilfreich sind. Zu diesem Zweck sind die festlegbaren Warnschwellenwerte von großem Nutzen. Soll zum Beispiel die durchschnittliche Lautstärke 40 dB nicht überschreiten, lässt sich der Threshold dementsprechend festlegen. Ein

Warnhinweis informiert dann über den eingestellten Regelverstoß. Gleiches ist für die Input-SNR (Störabstand), die Werte der Max Shortterm Loudness (Max S), Max Momentary Loudness und die Loudness Range (LRA) möglich. Ist beispielsweise die Loudness Range für ein bestimmtes Programm-Material festgelegt, lässt sich mit dem „Leveler“ die Vorgabe sehr gut einhalten und definieren.

Das ist gerade für Podcast-Plattform, Online-Radio und auch Streaming-Portale interessant, um eine gleichbleibende Qualität und ausgewogene Lautstärkeverhältnisse selbst bei unterschiedlichem Programm-Material anbieten zu können. Im Idealfall bekommt der Hörer eine ausgewogene Abstimmung unterschiedlicher Audio-Segmente angeboten: Die Musik dröhnt eben nicht nach der Anmoderation oder dem zuvor gehörten Interview aus den Lautsprechern, da sie über die gleiche Lautheit verfügt.

Praxis

Grundsätzlich ist der Auphonic „Leveler“ ein übersichtliches und leicht zu bedienendes Tool, dessen GUI auf schlichte Einfachheit und Performance setzt. Per Drag-and-Drop ziehe ich zunächst alle Soundfiles, die ich für die unterschiedlichen Pickup-Systeme (Praxisworkshop Verstärker-Gitarrenabnahme, Teil 5) aufgenommen habe, in das „Leveler“-Fenster. Das Target-Level stelle ich auf Broadcast-Norm (R-128: -23 LUFS) und lasse sowohl den Adaptive Leveler als auch HPF aktiviert. Nachdem ich unter „Preferences“ noch das Ausgangsformat für die WAV-Files auf MP3 gestellt und den Destination-Ordner gewählt habe, klicke ich auf den virtuellen Button „Process All Files“ und die Maschine läuft. Bereits beim Öffnen der Files erscheint auf der rechten Seite des GUIs ein Informationsfenster.

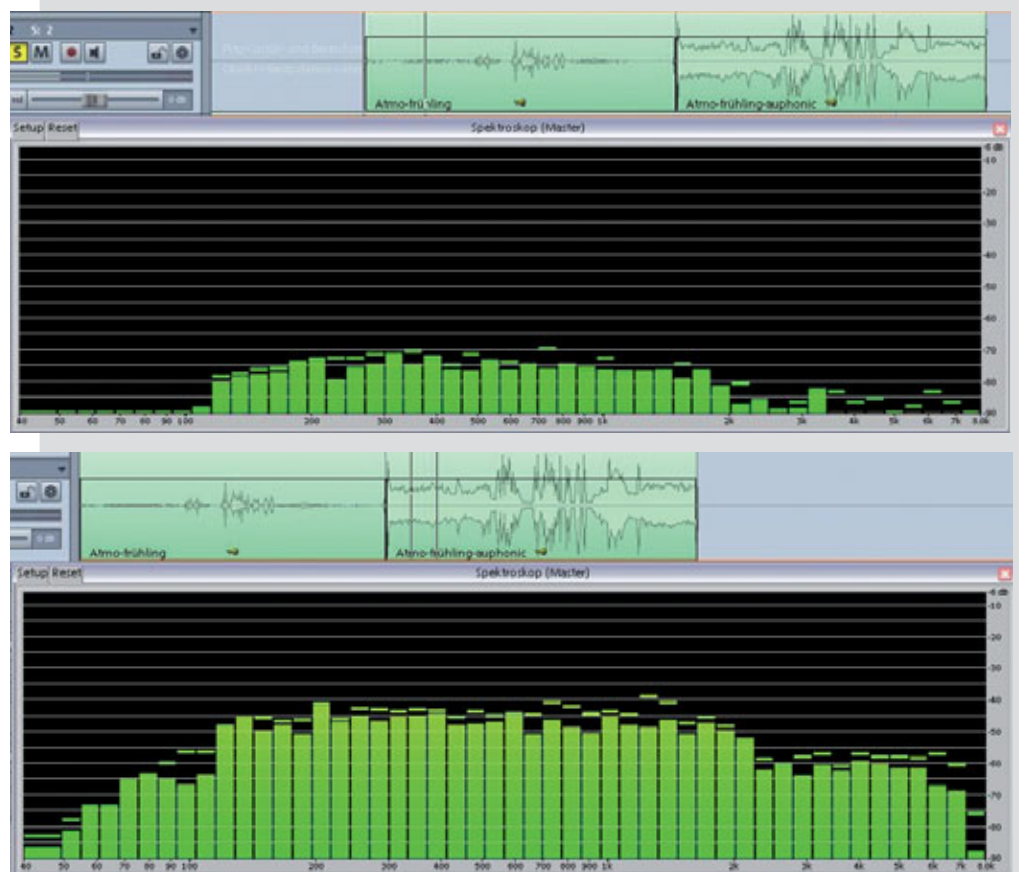
Während des Bearbeitungsprozesses wird im linken Fenster (Datei-Browser) der Status aller zu bearbeitenden Files angezeigt. Nach Abschluss der Anpassung lässt sich für jedes File die Audio-Statistik im rechten Fenster anzeigen. Auf diesem Weg ist es möglich, sehr übersichtlich die vorgenommenen Veränderungen zu kontrollieren. Neben

Standards wie der File-Größe sind Anpassungen der Loudness-, LRA- oder Max-Momentary-Werte sofort abzulesen. Um den „Leveler“ als komfortablen Batch-Converter zu betreiben und eine Audiodatenbank beispielsweise von Wave- in MP3-Files zu konvertieren, braucht man bloß die Algorithmus-Häkchen (Adaptiv Leveler und High Pass Filter) deaktivieren, und schon sind die Algorithmen deaktiviert. Bei den konvertierten Soundfiles ändert sich übrigens klanglich gar nicht so viel, außer dass die Loudness aller Files nach der Behandlung einheitlich auf -23 LUFS liegt. Allerdings handelt es sich bei den hier als Beispiel eingesetzten Gitarrenaufnahmen um sehr ähnliches und bereits normalisiertes Klangmaterial.

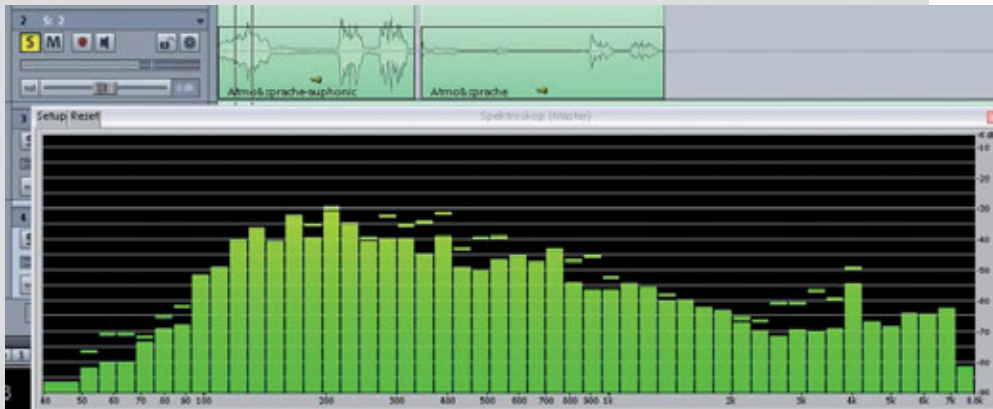
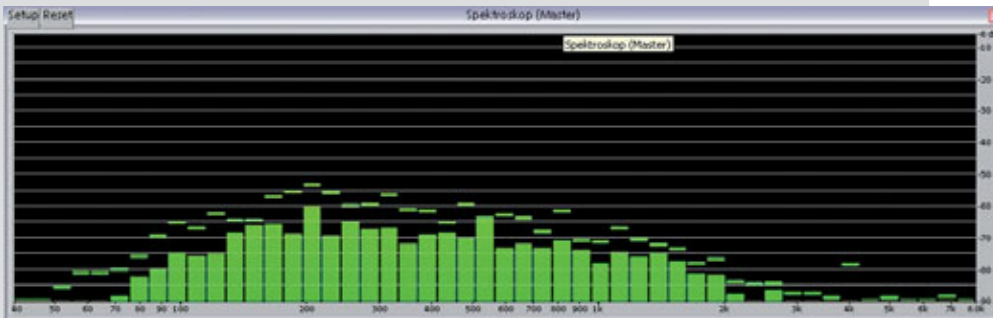
Um es dem Auphonic „Leveler“ aber nicht so leicht zu machen, nehme ich Akustikgitarre, Sprache

und eine Probe-Sequenz (Akustikgitarre und Gesang) mit dem Smartphone (Motorola MB-525) auf (alle Soundfiles finden Sie im Mehrwertbereich auf www.tools4music.de).

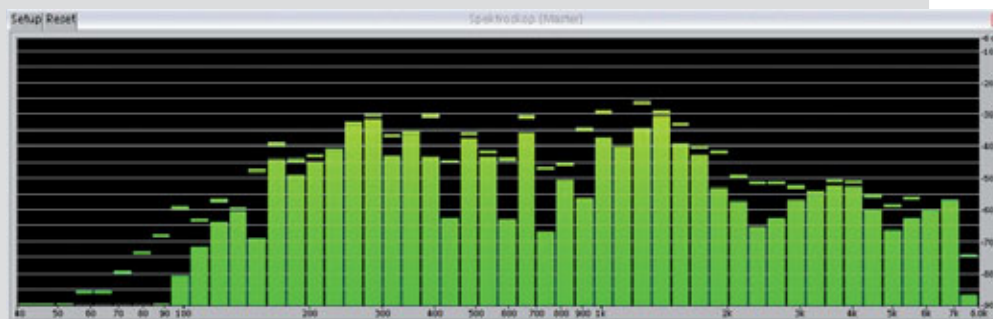
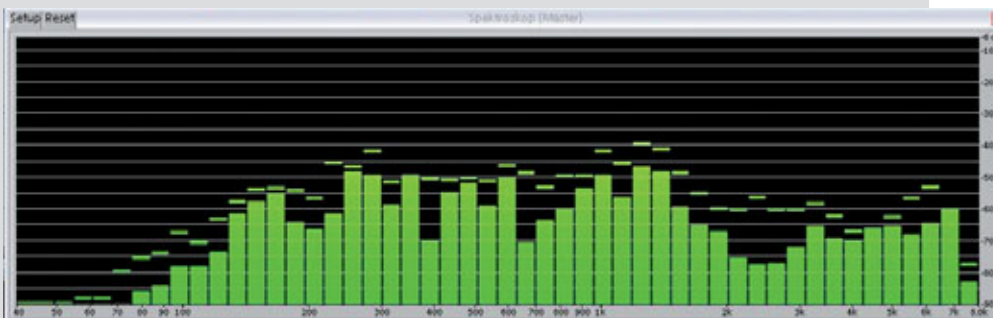
Dann lade ich die Files in den „Leveler“ und schaue mir an, was das Tool aus den rudimentären Aufnahmen „macht“: Bei der Aufnahme von Sprache im Freien hebt der „Leveler“ die Stimme deutlich nach vorne, was sie sehr präsent und gut verständlich werden lässt. Die Hintergrundgeräusche sind zwar insgesamt auch lauter, aber bleiben als Hintergrundgeräusche erkennbar. Während einer anderen Audio-Sequenz fliegt ein Flugzeug vorbei, dessen Antriebsbrummen nur wenig lauter ist als das Hintergrundgeräusch. Hier stößt der Analyse-Algorithmus an seinen Grenzen, da er das etwas lautere Flugzeug nicht mehr zum Hintergrund zählt. Die Verstärkung bringt das Turbinen-



Beim Vergleich der bearbeiteten und unbearbeiteten Aufnahme ist deutlich zu erkennen, wie der „Leveler“ die Loudness anpasst: Deutlich wird auch, dass sich das Frequenzspektrum ausweitet (ohne: 100 bis 3.000 Hz; mit: 40 bis 7.000 Hz)



Im Gegensatz zur unbehandelten Aufnahme werden die unteren Mitten (100 bis 400 Hz) angehoben, was das Flugzeuggeräusch in den Vordergrund rücken lässt



Auch wenn der Auphonic „Leveler“ leicht in die Dynamik der Musik (Gitarre, Gesang, Backings) eingreift, bleibt sie doch insgesamt recht natürlich, wobei die Loudness insgesamt deutlich angehoben wird

geräusch in den Vordergrund und verzerrt damit insgesamt das Klangbild. Allerdings ist das ein Grenzfall, da die einzelnen Klangelemente sich kaum voneinander absetzen. Bei den Sprach, Gesangs- und Instrumentenaufnahmen gefällt mir das Ergebnis des Algorithmus sehr gut, da auch bei

unterschiedlichem Material immer ein einheitlicher Gesamtklang berechnet wird. Selbst wenn auf Sprache Musik folgt oder zwei Stimmen in unterschiedlichem Abstand zum Mikrofon positioniert sind, gelingt es dem „Leveler“, Loudness und Dynamik so anzupassen, dass keine Lautstärke-sprünge zu hören sind und insgesamt ein gleicher Lautheitseindruck entsteht. Klangästhetisch ist das vielleicht nicht in jedem Fall die für jeden Geschmack passende optimale Lösung, doch um ein Programm oder die Lautstärke respektive Loudness vieler unterschiedlicher Audiofiles unkompliziert zu vereinheitlichen, ist der „Leveler“ empfehlenswert.

Finale

Solange die Bestandteile des Audiomaterials sich gut voneinander abheben, wie es beispielsweise bei Sprache vor einem ruhigen Hintergrund, reiner Musik oder einem klaren Geräusch der Fall ist, funktioniert der „Leveler“ ausgezeichnet und zuverlässig. Außerdem bringt er unterschiedliches Programm-Material Musik, Sprache, Gesang oder Geräusche intelligent und unkompliziert auf gleiches Loudness-Niveau. Bei komplexen Klangbildern, deren Bestandteile sich nicht klar voneinander unterscheiden, gerät der „Leveler“ an seine Grenzen, die sich durch das ständige Lernen des Algorithmus stetig verschieben.

Als Desktop Batch Audio File Processor eignet sich der Auphonic „Leveler“ besonders für all jene, die zahlreiche Audiofiles in ihrem Klang-Archiv in kurzer Zeit auf ein einheitliches Loudness-Niveau anpassen möchten. ■

NACHGEFRAGT

Georg Holzmann von Auphonic zu diesem Test:

„Im Auphonic ‚Leveler‘ ist die Noise Reduction nicht inkludiert, deswegen wird in dem Test das Rauschen auch nicht vermindert. Die Noise- und Hum-Reduction wird in den nächsten Tagen veröffentlicht. Unser Tool wurde natürlich hauptsächlich für Sprache bzw. Radioinhalte gemacht, aus diesem Grund versuchen wir, die Musik so wenig wie notwendig zu verändern. Dort passieren hauptsächlich nur Lautstärken-Anpassungen.“