

Künstliche Intelligenz in Audio Plugins

Andreas Kraft

Hochschule der Medien Stuttgart
andreas.kraft@mailbox.org

Abstract

In der Tonbearbeitung ist der Einsatz von Programm-erweiterungen (sogenannten Plugins) innerhalb eines Hauptprogramms ein weit verbreiteter Ansatz. Auf dem Markt werden daher von einer Vielzahl an Herstellern Plugins mit unterschiedlichsten Funktionen angeboten. Wie in vielen anderen Bereichen werden auch bei der Entwicklung von Audio Plugins zunehmend die Möglichkeiten der künstlichen Intelligenz genutzt, um bei Tonschaffenden mit neuen und verbesserten Plugin-Funktionen zu werben. Eine Auswahl KI gestützter Audio Plugins wurden untersucht und hinsichtlich ihrer Funktionen und ihres praktischen Mehrwerts in den unterschiedlichen Bereichen der Tonbearbeitung bewertet. Dabei erwiesen sich alle getesteten Plugins als sehr hilfreich und gut in ihrer Handhabung. Der Einsatz maschinellen Lernens resultierte bei allen untersuchten Programmen entweder in Zeitersparnissen oder überraschenden neuen Effekten und Möglichkeiten für die Tonbearbeitung.

Einleitung

Fand das Aufnehmen, Editieren und Mischen von Audio früher ausschließlich in professionellen Tonstudios statt, ist dies heute in überwältigendem Umfang auf einem herkömmlichen Laptop möglich. Grund dafür sind die immensen Entwicklungen der digitalen Audioteknik der letzten zwei Jahrzehnte. Wo früher schwere analoge Effektgeräte für spezielle Bearbeitungsschritte wie dem Komprimieren oder Verhallen von Audio vonnöten waren, kommen jetzt Audio Plugins, also Programm-erweiterungen zum Einsatz, die innerhalb von Minuten installiert und verwendet werden können. Inzwischen gibt es eine Fülle dieser Plugins, die sich in ihrer Funktionalität oft nur marginal unterscheiden und von den unterschiedlichsten Herstellern erworben werden können. Aktuell finden sich in deren Produktbeschreibungen immer öfter die Begriffe der künstlichen Intelligenz oder des maschinellen Lernens. Um zu untersuchen, inwieweit diese Versprechen nur marketingstrategische Alleinstellungsversuche sind, oder doch einen überzeugenden Mehrwert für die Tonbearbeitung bieten können, wurden einige Plugins mit unterschiedlichen Funktionalitäten ausgewählt und getestet. Im Folgenden soll auf die ein-

zelnen Programme eingegangen werden. Insbesondere sollen ihre Funktionalität und Handhabbarkeit, sowie potentielle Einsatzbereiche in den unterschiedlichen Bereichen der Tonbearbeitung untersucht und bewertet werden.

Untersuchung der Plugins

Untersucht wurden die folgenden Audio Plugins:

- **ATLAS** von Algonaut
- **DeRoom Pro** von Accentize
- **Chameleon** von Accentize
- **Adaptiverb** von Zynaptic
- **AI Master** von Exonic UK

Diese kommen durch ihre unterschiedliche Funktionalität in verschiedenen Bereichen der Tonbearbeitung zum Einsatz und wurden deshalb als Stichproben ausgewählt.

Samplemanagement mit ATLAS

Die Firma Algonaut bietet mit ATLAS einen KI gestützten Sample-Manager. Die Kern-Funktionalität des Werkzeugs ist die Analyse und Klassifizierung von Audiodateien. Ganze Ordnerstrukturen mit einer Vielzahl einzelner Audiodateien können mit wenigen Klicks analysiert und in einer Art Cluster dargestellt werden (siehe Abb. 1). Dieses stellt die einzelnen Sounds als kleine Kreise dar, die clusterartig um bestimmte Überbegriffe wie „Bassdrum“, „Tambourin“, etc. gruppiert sind und mit einem Klick direkt abgespielt werden können. Dies erleichtert das Durchsuchen der Audiobibliothek nach gewünschten Sounds, da sofort alle bestehenden Sounds eines bestimmten Instruments oder Effekts durchsucht und verglichen werden können und erspart das zeitintensive Durchsuchen vieler Ordner und Unterordner der Audiodatenbank. Darüber hinaus bietet ATLAS einen eingebauten Sequencer, mit dem die gefundenen Sounds direkt zu Rhythmus-Arrangements weiterverarbeitet werden und als MIDI- oder WAV-Datei ausgespielt werden können. So wird für den Bereich der Musikproduktion, bzw. des Songwritings ein äußerst effektiver Workflow ermöglicht, mit dem in sehr kurzer Zeit verschiedene Arrangement-Ideen getestet und mit verschiedenen Sounds belegt werden können. Ein weiterer denkbarer Einsatzbereich des Plugins ist die Filmvertonung. Bei dieser werden die verschiedensten Geräusche wie beispielsweise

Schritte, Schüsse, Gegenstände etc. nachträglich vertont und hierbei auch oft aus Audiodatenbanken ausgewählt. Die übersichtliche Gruppierung nach Überbegriffen durch ATLAS könnte also auch hier zu einer enormen Zeitersparnis und Effizienz im Vertonungsprozess beitragen.



Abbildung 1: Benutzeroberfläche von ATLAS (Algonaut)¹

Hallentfernung mit DeRoom Pro

Audioaufnahmen bestehen in der Regel aus direktem und diffusem Schall, also einerseits aus dem Schallanteil, welcher auf direktem Wege von der Schallquelle zum Mikrofon gelangt und andererseits dem Schallanteil, welcher über einfache oder vielfache Reflexion an Decken, Boden, Wänden oder anderen Schallhindernissen verzögert an das Mikrofon gelangt. In vielen Fällen, wie zum Beispiel Sprachaufnahmen im Studio, ist es gewünscht nur den Direktschallanteil zu erhalten, sodass das Signal möglichst „trocken“ ist und sich gut für weitere Bearbeitungsschritte wie eine gezielte nachträgliche Verhallung eignet. Hier bietet die Firma Accentize mit einem Plugin namens DeRoom Pro (siehe Abbildung 2) die Möglichkeit, diffuse Schallanteile nachträglich aus bestehenden Aufnahmen zu entfernen. Hierfür wurde ein neuronales Netz mit vielen verschiedenen Aufnahmen darauf trainiert, zwischen Direkt- und Diffusschall zu unterscheiden. Das Plugin selbst kann schließlich als Effekt auf eine Audiospur geladen werden und entfernt in regelbarer Menge Diffusschall aus dem Audiosignal. Getestet wurde das Plugin an einer Sprachaufnahme mit einem deutlich hörbaren Raumklanganteil. Mit wenigen Klicks konnte dieser deutlich verringert und ein klareres Klangbild erreicht werden. Der Einsatz maschinellen Lernens erweist dich hier also als durchaus zielführend, da genau die in vielen Fällen störenden Charakteristika

des Klangbilds erlernt wurden und somit gezielt und unkompliziert entfernt werden können.



Abbildung 2: Benutzeroberfläche von DeRoom Pro (Accentize)

Raumklangübertragung mit Chameleon

Neben DeRoom Pro hat der Hersteller Accentize noch einige weitere KI gestützte Audio Plugins auf dem Markt. So arbeitet auch Chameleon mit neuronalen Netzen, um aus bestehenden Audioaufnahmen die natürliche Raumimpulsantwort zu extrahieren. Dies ermöglicht das Übertragen eines Raumklangs von einer Tonaufnahme auf eine andere.

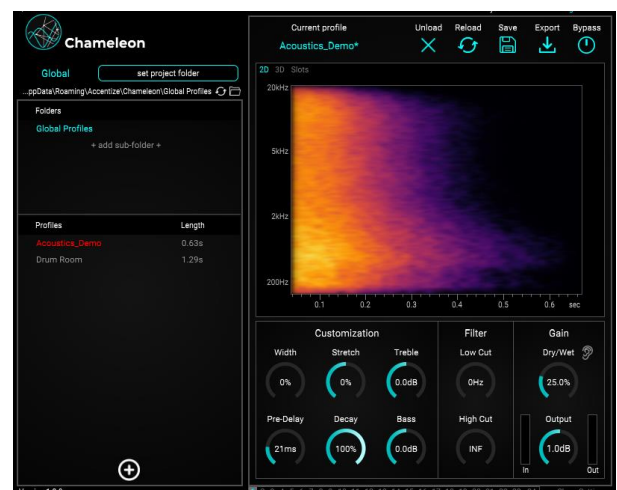


Abbildung 3: Benutzeroberfläche von Chameleon (Accentize)

Getestet wurde das Plugin mit der im letzten Kapitel erwähnten Sprachaufnahme, sowie einer weiteren Sprachaufnahme desselben Sprechers, jedoch in

¹ Alle Bildquellen sind eigene Screenshots.

einem wesentlich trockenerem Raum. Die Raumpulsantwort der „halligeren“ Sprachaufnahme wurde mit Chameleon extrahiert und dann als Audioeffekt auf die trockene Aufnahme angewendet. Mit erstaunlich geringem Aufwand konnte so mithilfe einer intuitiven Benutzeroberfläche (siehe Abb. 3) ein glaubhaft ähnlicher Raumeindruck bei der Sprachaufnahmen erreicht werden.

Hallsynthese mit Adaptiverb

Während die bislang vorgestellten Werkzeuge eher technische Hilfsmittel sind, ist der Adaptiverb des Herstellers Zynaptic ein durch und durch kreatives Werkzeug, welches in die Kategorie der künstlichen Hallerzeugung einzuordnen ist. Das Plugin arbeitet mit künstlicher Intelligenz, um tonale und geräuschartige Klanganteile eines Signals zu identifizieren und nutzt folglich nur die tonalen Anteile, um daraus eine Hallfahne zu generieren. Durch diesen Prozess entsteht ein Hall, welcher sich sehr gut zum trockenen Signal mischt, da geräuschartige Schallanteile nicht verhallt werden und die Mischung somit weniger „verschmutzen“. Auch können mithilfe des resynthetisierten Halls sehr lange und atmosphärisch klingende Hallfahnen erzeugt werden, die deutlich von ihrem tonalen Charakter profitieren. Die Benutzeroberfläche (siehe Abb. 4) bietet wahlweise eine reduzierte Ansicht mit wenigen wichtigen Einstellungsmöglichkeiten oder eine vollständige Übersicht aller steuerbaren Parameter. Dadurch kann in der reduzierten Ansicht in kürze ein gutes Ergebnis erzielt und in der vollständigen Ansicht weiter verfeinert und ausgearbeitet werden können.



Abbildung 4: Reduzierte Ansicht der Benutzeroberfläche von Adaptiverb (Zynaptic)

Sowohl in der Musikproduktion, besonders im Bereich der Filmmusik, als auch in der Filmvertonung

eröffnet das Plugin gewaltige kreative Möglichkeiten. Wo immer immersive, atmosphärische Klänge auch aus einfachem Ausgangsmaterial wie wenigen Pianotönen kreierte werden sollen, ist der Adaptiverb ein äußerst geeignetes Werkzeug.

Mastering mit AI Master

Das Mastering ist der finale Bearbeitungsschritt in der Musikproduktion und anderen Audioproduktionsformen. Die fertige Mischung wird hierbei als Ganzes nochmals überarbeitet, um ein möglichst stimmiges Gesamtprodukt zu erhalten. Schon seit längerem wird in diesem Bereich mit künstlicher Intelligenz gearbeitet, um weite Teile des Masteringprozesses automatisieren zu können. So bieten Onlineplattformen wie LANDR vollautomatisierte Masteringdienste im Internet an. Auch im Bereich der Audioplugins finden sich hier einige Produkte, welche den Masteringprozess mithilfe künstlicher Intelligenz vereinfachen sollen. Ein Beispiel ist der hier getestete AI Master des Herstellers Exonic UK. Dieser bietet eine extrem simple Benutzeroberfläche (siehe Abb. 5), auf welchem zwischen dem Analyse- und dem Mastermodus umgeschaltet werden kann.

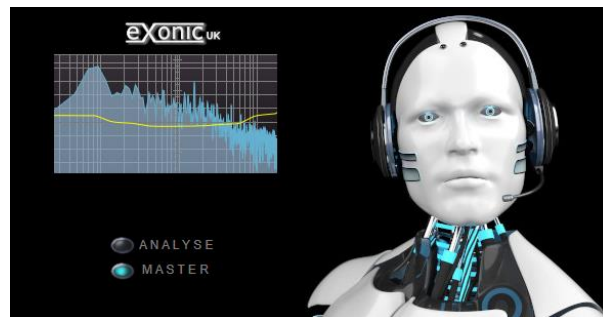


Abbildung 5: Benutzeroberfläche von AI Master (Exonic UK)

Im Analysemodus findet eine vorab mit vielen Musikstücken trainierte KI die passenden Parameter der nötigen Bearbeitungsschritte. So zum Beispiel welche Filter gesetzt und welche Kompressorstellungen vorgenommen werden müssen, um dem Musikstück eine Art Feinschliff zu geben. Im Mastermodus werden diese Bearbeitungsschritte schließlich auf das Signal angewendet und das gemasterte Audio wird hörbar und kann ausgespielt werden. Es ist nun diskutabel, ob das Endprodukt als perfekt gemastert angesehen werden kann bzw. ob ein professioneller Mastering-Ingenieur dieselben Entscheidungen getroffen hätte. Ist eine

schnelle Ein-Klick-Lösung gefragt, die eine fertige Mischung auf ein gutes Lautheitslevel bringt, ohne sie allzu sehr zu verfärben, so eignet sich das Plugin hervorragend und liefert dafür auch sehr gute Ergebnisse. Da im Mastering im herkömmlichen Sinne jedoch auch gezielt auf Kundenwünsche eingegangen werden muss und oft nicht nur technische sondern auch kreative Entscheidungen gefragt sind, ersetzen Plugins dieser Art keinesfalls den klassischen Mastering-Prozess.

Ergebnisse und Fazit

Eingangs wurde die Frage gestellt, ob der Einsatz künstlicher Intelligenz in Audioplugins tatsächlich einen signifikanten Mehrwert bringt oder nur marketingstrategischer Natur ist. Als Diskussionsgrundlage wurden Stichprobenartig fünf Audioplugins getestet und hinsichtlich ihrer Handhabbarkeit und ihres Mehrwerts für die Tonbearbeitung getestet. Alle der fünf getesteten Plugins stellten mit einem effizienten Arbeitsablauf und überraschend hilfreichen Funktionen einen klar erkennbaren Mehrwert für verschiedene Bereiche der Tonbearbeitung dar. Maschinelles Lernen wurde dort eingesetzt, wo es dem Tonschaffenden Arbeit erspart und nicht versucht diesen zu ersetzen. Einzig im Fall des AI Master kann hier diskutiert werden, ob der Name des Produkts, welcher ein Mastering verspricht, auch das halten kann, was ein professioneller Masteringprozess eigentlich erreichen soll.

Als Fazit kann grundsätzlich festgehalten werden, dass wo immer künstliche Intelligenz dem Tonschaffenden eine Zeitersparnis verschafft und so mehr Zeit für kreative Arbeit ermöglicht, von einem Mehrwert für die Tonbearbeitung gesprochen werden kann. Die eingangs formulierte Frage, ob KI in Audioplugins nur ein hohles Verkaufsargument sei, kann also auf der Grundlage der hier getesteten Plugins eindeutig mit nein beantwortet werden. Von einer Ersetzung des Tonschaffenden durch künstlicher Intelligenz kann gleichwohl nicht die Rede sein. Es bleibt abzuwarten, wo die Grenze des Machbaren für die künstliche Intelligenz liegt, wenn es um kreative Bearbeitungsschritte in der Tonbearbeitung geht.

Abbildungsverzeichnis

Abbildung 1: Benutzeroberfläche von ATLAS (Algonaut).....	2
Abbildung 2: Benutzeroberfläche von DeRoom Pro (Accentize).....	2
Abbildung 3: Benutzeroberfläche von Chameleon (Accentize).....	2
Abbildung 4: Reduzierte Ansicht der Benutzeroberfläche von Adaptiverb (Zynaptic) ...	3
Abbildung 5: Benutzeroberfläche von AI Master (Exonic UK).....	3