

# Optimierung des Sendebetriebs im Bereich Audio mittels teilautomatisierter Signalverarbeitung unter zu Hilfenahme von Künstlicher Intelligenz

## Einleitung

Der Bereich Sendeton und dessen Audiosignalverarbeitung benötigt aufgrund verschiedener Faktoren Optimierung und Automatisierung. Finanzielle Kürzungen, personelle Engpässe, Wirtschaftlichkeit oder auch die aktuelle Pandemielage des COVID-19 Virus zwingen Sendehäuser und Studios immer mehr zum Umdenken und Restrukturieren von Sendeabläufen, Workflows und zur Anpassung der Signalkette, um finanziellem Druck und aktuellen Hygienevorschriften gerecht zu werden. Die Zeit zur Einrichtung von Sendungen, für die Vorproduktion und die Arbeit im Sendeton selbst im Livebetrieb ist somit zeitlich und personell oft eingeschränkt. Für Toningenieur\*innen ist es immer wieder aufs Neue eine Herausforderung mit wenig Zeit und wenig Personal eine qualitative Sendung vorzubereiten, einzurichten und zu fahren.

Werden die Signalkette und die Arbeitsabläufe bei einer Livesendung zwischen verschiedenen Sendungen verglichen stellt man fest, dass sich ein Großteil der Vorgänge, Arbeitsschritte und der damit verbundene Workload wiederholt. Genau diese Abläufe machen den größten Teil der Arbeit aus. Beispielhaft erfolgt nach dem Routing der Relevanten Kanäle auf ein digitales Sendepult für Sprache, MAZ, Musik: Instrumente und Singstimme, die Grundlegenden Bearbeitungen: Dynamik, Frequenz und zum Schluss das Balancen zwischen den Kanälen um die gewünschte Mischung in Klang und Lautheit zu erreichen. Die entsprechende Frequenzbearbeitung erfolgt hier nicht nur nach kreativen Maßstäben. Ein Signal muss von Störsignalen befreit werden und so aufgearbeitet werden, um eine Sendefähigkeit in gewünschter Qualität zu gewährleisten, und das verlässlich.

## Idee und Konzept

In den letzten Jahren sind immer mehr Tools in der digitalen Audiotbearbeitung auf den Markt gekommen, welche die Beschriebene Aufgabe mittels Maschine Learning Algorithmen automatisiert oder teil-automatisiert bewältigen. Bekannteste Beispiele sind hierfür die Plugins und Nativen Softwarelösungen von iZotope wie RX9, zur intelligenten Erkennung von Störgeräuschen und Restauration von Audiosignalen in der Postproduktion<sup>1</sup>. Diese Tools erfreuen sich reger Nutzung im Bereich Filmtone, Setton, um schwierige Situationen und deren O-Ton bei Störungen zu korrigieren und aufwendige Restaurationen zu automatisieren, welche sonst einen sehr hohen Zeitaufwand für die händische Manipulation und Restauration der Signale benötigen würden. Nun entstand im Rahmen der aktuellen TV-Produktionen der Hochschule der Medien Stuttgart<sup>2</sup> die Frage, ob es möglich wäre, eine Livesendung KI gestützt fahren zu können, um eine personelle Entlastung zu ermöglichen und dem Toningenieur in der Livemischung mehr Freiheiten und kreativen Raum zu bieten.

Ebenso ob es möglich ist, durch KI gestützte Tools die Vorproduktion und Soundcheck zu beschleunigen, da Künstler\*innen und oder Gäste, oftmals wenig Zeit zur Verfügung stellen. Es ging bei dem zu beschreibenden Testaufbau aber nicht um die Frage, ob KI oder Machine Learning Algorithmen die kreative Arbeit des Personals übernehmen soll. Diese stark diskutierte Frage wird außenvor gestellt und ist nicht Bestandteil dieser Arbeit<sup>3</sup>.

Es wurde ein Testaufbau angestrebt, bei dem ein Tool Verwendung fand, das die Entzerrung der einzelnen Signale live vorgenommen hat. Die beschriebenen Störsignale sollten automatisiert bearbeitet werden, dass ein Nutzsignal entsteht, welches so flexibel wie möglich für die kreative Mischung zur Verfügung steht. Raumresonanzen, Eigenresonanzen des Mikrofons und des Mikrofonkörpers in Bezug auf die Quelle sollten intelligent erkannt und bearbeitet werden.

Ebenso wichtig war es, die Zuverlässigkeit der entstandenen Signalkette zu testen und die Laufzeit des Tools möglichst gering zu halten, um mit maximal einem Frame

---

<sup>1</sup> (iZotope, 2022)

<sup>2</sup> (HdM Stuttgart, 2022)

<sup>3</sup> (Price, 2022)

Versatz vor der Verzögerung des Bildsignals zu bleiben.

## Vorstellung: smart:EQ Live

Die sonible GmbH aus Graz, Österreich hat zuletzt ein für diesen Versuchsaufbau interessantes Tool veröffentlicht. Ein adaptiver Equalizer, welcher das zu bearbeitende Signal automatisiert erkennt und eine passende Entzerrung des Frequenzgangs live und kontinuierlich vornehmen kann<sup>4</sup>. Die Basis des Tools läuft auf VST3, ist aber auch als VST, AAX, Audio Unit und WPAPI verfügbar und damit auf allen gängigen Plattformen lauffähig.

Was kann das Tool in Bezug auf Machine Learning? Es können grundlegend 3 verschiedene Signaltypen unterschieden werden: Stimme, Instrumental und Percussion. Ebenso können im Tool selbst bestimmte Frequenzen priorisiert werden und die Geschwindigkeit, mit der sich das Tool an das Signal anpasst für 3 Frequenzbänder unabhängig eingestellt werden. Die Anpassung an das Signal läuft entweder kontinuierlich mit vorlaufendem „lernen“ oder über eine definierte Zeit.

Eine Besonderheit des Tools ist die Laufzeit, da es in nahezu Echtzeit rechnet und sich an das Signal anpasst. Dieses Feature ist grundlegend Voraussetzung für den Einsatz im Livebetrieb und ist ein Alleinstellungsmerkmal für Machine Learning basierte Plugins in der Audiosignalverarbeitung.

## Versuch 1 - Talkshow

Im Rahmen des Studieninfotages des Landes Baden-Württemberg hat die Hochschule der Medien Stuttgart eine TV-Sendung zum Thema live gesendet<sup>5</sup>. Inhalt waren neben diversen Zuspielungen von aufgezeichneten Beiträgen unter anderem auch eine Gesprächsrunde von 4 Gästen und einem Moderator. Diese Gesprächssituation sollte im Sendeton mittels smart:EQ live teilautomatisiert gefahren werden.

Alle Signale wurden mittels Audionetzwerk (Audinate Dante<sup>6</sup>) zwischen den verschiedenen Punkten des Studios verteilt. Dies beinhaltete eine Stagebox Yamaha RIO 16-8 im Studio als DA-AD Wandler, einem Yamaha QL1 Sendepult, einem Shure Microflex Wireless System und einem Server auf Win10 Basis für die

---

<sup>4</sup> (sonible GmbH, 2020)

<sup>5</sup> (HdM Stuttgart, 2021)

<sup>6</sup> (Audinate, 2022)

Aufzeichnung aller Kanäle und die Berechnung der smart:EQ Live Instanzen pro Signalweg.

Für dieses Setup wurden 8 Instanzen des smart:EQ Live Plugins verwendet, um alle Gäste inklusive Publikum, den Moderator und die Zuspieldungen zu bearbeiten. Die Latenz des Servers durch die Instanzen betrug 56 Samples bei 48kHz/24bit des Systems und war somit für die Livesendung nutzbar und unproblematisch.

Im Signalfluss wurden die Instanzen als Insert im Sendepult nach dem digitalen Trim im Kanaleingang angelegt. Als Fallback-Lösung wurde im Ausfall der Insert-Punkt automatisch auf Bypass gesetzt und die smart:EQ Live Instanzen somit aus der Kette gelöst. Alternativ könnte hier auch ein Redundanter Server angelegt werden, auf welchen umgeschaltet werden könnte. Die Sendung hatte eine Länge von 1:31 (HH:MM) und es gab im gesamten Testlauf und in der Generalprobe keinen Ausfall des gesamten Systems<sup>7</sup>.

Fazit Versuch 1: Auffällig war, wie schnell die verschiedenen Instanzen die Sprache als Quelle erkannt haben und passende

Filter gesetzt wurden. Ebenso wurden binnen ca. 5 Sekunden störende Resonanzen erkannt und gedämpft. Das Resultat waren deutlich bessere Sprachverständlichkeit, angenehm ausgewogener Klang der Stimmen und das Übersprechen der verschiedenen Quellen wurde ebenso gemindert. Um die beschriebenen Effekte selbst zu erzielen wären für einen einzelnen Toningenieur gut 1 Stunde Arbeit notwendig, wenn nur ein Typ von Mikrofon zum Einsatz kommt und die Einstellungen auf die weiteren Quellen übernommen werden können. Ist dies nicht der Fall wächst der Aufwand wiederum mit der Anzahl der verschiedenen Quellen. Negativ aufgefallen ist folgendes: Ist aus dem Soundcheck nicht bereits ein angelernter Algorithmus vorhanden, startet das Tool in der Sendung bei null. Bis die erste Einstellung angelernt ist vergehen bis zu 10 Sekunden. In dieser Zeit sind die Veränderungen der Filterkurven als Verfärbung des Signals wahrnehmbar und können als störend empfunden werden. Ein Soundcheck von wenigen Sekunden mit der Person oder einem Stand-In können das Problem aber mindern. Dieses Problem

---

<sup>7</sup> (sonible GmbH, 2020)

ist bekannt und lässt sich nicht lösen, da es immer ein „Lernen“ des Algorithmus gibt<sup>8</sup>. Dennoch ist positiv aufgefallen, dass sich der Toningenieur deutlich mehr auf das Geschehen und die Mischung konzentrieren konnte, ohne sich um Störende Resonanzen oder Übersprechen zwischen den Handmikrofonen kümmern zu müssen. Hier gab es vor und während der Produktion eine starke Entlastung.

## Versuch 2 – Musikshow

In einer zweiten TV-Show wurde der Systemaufbau in einer Musikproduktion getestet ob auch hier eine Entlastung des Personals möglich ist und der Workflow beschleunigt werden kann.

Luna Chiara aus Stuttgart, eine Popsängerin, hat mit Ihrem DJ den Inhalt für eine TV-Sendung der Hochschule der Medien Stuttgart gestellt. Es wurden eigene Popsongs und Coversongs Live-On-Tape aufgezeichnet und direkt auf Master, ohne Postproduktion. Besondere Anforderung an den Sendeton war zum einen der sehr kurze Soundcheck von 30 Minuten und das Monitoring für die Künstler. Das Monitoring lief auch über das gleiche System, über das Sendepult.

Insgesamt kamen erneut 8 Instanzen zum Einsatz. 2 Instanzen für die Sängerin und diverse Instanzen für die Mehrkanalzuspielung des DJ der zur Mischung mehrere Stereo-Stems aus seinem DJ-Setup übergeben hat.

Fazit – Versuch 2: Auch hier hat sich das Tool als hilfreich und zuverlässig erwiesen. Für die Stimme war der gleiche Effekt wie in der Talkrunde zu beobachten. Bessere Verständlichkeit und Transparenz im Klang. Resonanzen und Balance der Stimme, welche sich dadurch gut in die Mischung verarbeiten ließ. Für die Bearbeitung der unterschiedlichen Stems für Rhythmus, Drums und Synthesizer war auch hier ein Mehrgewinn zu beobachten. Durch das Reduzieren von Resonanzen und langen sinusförmigen Moden, war eine Mischung der verschiedenen Quellen mit einer guten Durchhörbarkeit deutlich schneller zu erreichen.

## Auswertung

Nach der gemeinsamen Auswertung ist dem Team als erstes die Zuverlässige Funktion des Systems aufgefallen. Obwohl der Algorithmus frei arbeiten konnte und der Arbeitsbereich absichtlich so offen wie

---

<sup>8</sup> (Casper, 2018)

möglich gehalten wurde, um das Tool möglichst frei agieren zu lassen, waren die Ergebnisse nach kurzer Eingewöhnung für den Toningenieur verlässlich und meist vorhersehbar. Der Mehrgewinn war bereits in der ersten Stunde im Studio vorhanden und es gab gutes Feedback von Regie und Toningenieur am Pult. Auch die Integration in das System, in Server bzw. DAW (Sequoia 15) war einfach und schnell möglich. Auch die DAW Reaper wurde zur Vergleichbarkeit eingerichtet. Auch hier war das System binnen zwei Stunden eingerichtet und spielbereit.

Die Einbindung in ein digitales Audiosystem ist Voraussetzung, um große Kanalzahlen einzeln bearbeiten zu können. Große Rundfunkanstalten und TV-Studios müssen in ihren komplexen Systemen entsprechend Platz und eine Vorbereitung zur Anbindung haben. Das kann in komplexen und starren Strukturen unter Umständen zum Nachteil werden und eine Integration behindern.

Sind die Systemanforderungen aber gegeben, ergab sich in diesen

Produktionen ein eindeutiger Mehrwert für Studiopersonal, Künstler und Konsument. Durch die Entlastung in den technischen Abläufen einer Mischung hat der Pultoperator / Toningenieur deutlich mehr Zeit und Möglichkeiten für eine kreative Mischung. Somit hat er auch mehr Zeit sich um Künstler\*innen und deren Betreuung und künstlerische Umsetzung zu kümmern. Die Qualität des gesamten Produkts kann sich so steigern lassen, wobei gleichzeitig Zeit gespart werden kann.

Machine Learning und der in diesem Beispielversuch verwendete Algorithmus von sonible kann einen Mehrwert in der Produktion darstellen und trotz angespannter Situation durch zeitlichen Druck und finanzielle Grenzen für Studiopersonal einen Ausgleich schaffen. Der Fokus für Toningenieur\*innen kann somit stärker auf kreativer Arbeit liegen und somit einen positiven Effekt für eine erfolgreiche Produktion schaffen.

## Literaturverzeichnis

Audinate. (22. März 2022). *audinate.com*. Von Dante: The Networking Standard for the Professional AV Industry: <https://www.audinate.com/meet-dante?link=hm-r2-l-dc> abgerufen

Casper, J. (07. Januar 2018). *Review: Sonible smart:EQ Live*. Von Ask Audio: <https://ask.audio/articles/review-sonible-smarteq-live> abgerufen

HdM Stuttgart. (17. November 2021). *Live TV-Show vom Studieninfotag am 17.11.2021 an der HdM*. Von YouTube: <https://www.youtube.com/watch?v=2iNQh80SViU> abgerufen

HdM Stuttgart. (2022). *Hochschule der Medien Stuttgart*. Von Die Live-Sendung zur MediaNight: <https://www.hdm-stuttgart.de/medianight/live> abgerufen

iZotope. (22. März 2022). *iZotope Produkte*. Von iZotope: <https://www.izotope.com/en/products.html> abgerufen

Price, A. (28. Januar 2022). *Artificial Intelligence in Music Production – Friend or Foe?* Von Audio Media International: <https://audiomediainternational.com/artificial-intelligence-in-music/> abgerufen

sonible GmbH. (März 2020). *Manual - Smart:EQ live*. Von sonible.com: [https://www.sonible.com/wp-content/uploads/2020/03/manual-smartEQ-live\\_download.pdf](https://www.sonible.com/wp-content/uploads/2020/03/manual-smartEQ-live_download.pdf) abgerufen