



Elvis aus der Maschine

Wie KI die Musikproduktion verändert

Der Name Zynaptiq ist als Hersteller von KI-Tools nur wenigen Profis ein Begriff. Doch die Werkzeuge haben es in sich und könnten die Musikproduktion, wie wir sie kennen, völlig auf den Kopf stellen. Wir sprachen mit einem der Firmengründer, Denis Goekdag, über die Möglichkeiten der KI und die künftige Rolle von Musikern.

Von Hartmut Gieselmann

Zynaptiq ist ein relativ kleiner Hersteller von Software-Tools für die Musikproduktion. Denis Goekdag gründete die Firma zusammen mit Stephan Bernsee

Ende 2011. Mit fünf Festangestellten sowie bis zu 20 freien Mitarbeitern entwickeln sie im Hauptquartier Hannover mächtige KI-Tools, die nicht nur Musikproduzenten, sondern auch Forensikern neue Möglichkeiten eröffnen.

Da wäre beispielsweise Unveil, das Hallfahnen aus einer Aufnahme entfernt, oder Unchirp, das Codec-Artefakte ausbügelt, wie sie durch eine MP3-Konvertierung entstehen. Inzwischen hat Zynaptiq auch Verfahren entwickelt, die aus einem kompletten Mix das Schlagzeug und den Gesang herausfiltern. Beeindruckend wirkt auch das Plug-in „Pitchmap“, das die Harmonien in einem fertigen Mix komplett verändern kann. Zwar hört man dabei hier und da noch einige Artefakte, aber die Ergebnisse hätte man vor ein paar Jahren noch für Voodoo gehalten.

„Die Zeit arbeitet für uns“, erklärt Denis Goekdag im Gespräch mit c’t. Denn mit steigender Leistung der Rechner könne auch die KI mit immer größeren Datenmengen in kürzeren Abständen trainiert werden.

Wie das genau funktioniert, ist ein Betriebsgeheimnis, aber laut Goekdag werden bei einigen Verfahren die KI-Tools mit Vorher-nachher-Versionen der gleichen Aufnahmen gefüttert: einmal mit Hall, einmal ohne. Einmal komplett gemixt, einmal mit getrenntem Gesang und Schlagzeug. Wenn man das oft genug macht, erkennt die KI mittels Deep Learning Gesetzmäßigkeiten, mit denen sie auch unbekanntes Musikmaterial identifizieren und reparieren kann. Das funktioniert ähnlich wie in der KI-Entwicklung zur automatischen Bilderkennung.

Für das aufwendige Training der neuronalen Netze freut sich Goekdag bereits auf die neuen Mac Pro: „Die sind vergleichsweise günstig“, denn für das Training braucht Zynaptiq sehr viel RAM – locker bis zu einem Terabyte oder mehr.

„Eine Stereo-Musikdatei belegt bei CD-Qualität pro Minute circa 10 MByte, da kommen bei Zehntausenden Mehrspur-Songs schnell solche Mengen zusammen“, rechnet Goekdag vor. „Wir lassen die neuronalen Netze deshalb oft auf den CPUs rechnen. Die sind zwar um fast eine Größenordnung langsamer als eine GPU, haben jedoch direkten Zugriff auf den Hauptspeicher. Dessen Anbindung ist bei großen Datensätzen oft das größte Nadelöhr.“

Am liebsten würde Goekdag für solche Zwecke POWER9-Systeme von IBM nutzen, die eine schnellere Speicheranbindung haben. Mit sechsstelligen Preisen sind sie jedoch deutlich teurer als die neuen Apple-Modelle. Zumieta-rechenpower sei typischerweise weniger kosteneffizient als der Kauf eigener Hardware, zumal sich die verfügbaren Konfigurationen aus RAM, CPU und GPU nicht ideal auf die jeweilige Anwendung zuschneiden lassen.

Krisenlösung

Ursprünglich engagierte sich Goekdag als Musikproduzent und Mitbegründer der TimeTools-Studios in Hannover – wo unter anderem Songs von Madonna, Tom Jones und Fury In the Slaughterhouse entstanden. Als Anfang der 2000er aufgrund der Napster-Krise die Umsätze branchenweit um bis zu 90 Prozent einbrachen,

wandte er sich der Werbemusik und dem Sounddesign für Spiele zu. Unter anderem arbeitete er viel mit Native Instruments zusammen.

Aufgrund der immer weiter schrumpfenden Zeit-Budgets für Audiotbearbeitungen sieht er einen starken Bedarf an Tools, die quasi auf Knopfdruck professionelle Ergebnisse liefern: „Wenn du [bei einer Werbeproduktion] innerhalb von kürzester Zeit eine Passage noch mal komplett umgestalten musst, weil gleich der Regisseur vor der Tür steht und eine Präsentation haben will, dann muss alles zack, zack gehen. Da hast du keine Zeit, stundenlang im Studio herumzubasteln.“

Die Bedienoberflächen der Zynaptiq-Plug-ins sind deshalb bewusst einfach gehalten. „Wenn Software die Aufgabe hat, ein Schlagzeug aus einem Gesamtmix zu entfernen, dann gibt es dafür ja eine optimale Lösung. Die sollte die KI finden, ohne dass der Anwender Dutzende Schieberegler einstellt.“ Das wissen nicht nur Einsteiger, sondern auch Profis zu schätzen.

Generell hat Zynaptiq sich von dem alten Paradigma verabschiedet, dass Software alte Instrumente und Effekte möglichst originalgetreu simulieren sollte. Viel spannender ist es für Goekdag, wenn Software in neue Bereiche der Klanggestaltung vorstößt. „Bestes Beispiel ist unser Hall Adaptiverb. Der simuliert keinen realen Raum, sondern passt sich der harmonischen Struktur des Eingangssignals an, was zu einem deutlich klareren Klangbild führt.“

Anwender seien bei KI-Software jedoch generell anspruchsvoll und ungeduldig: „Früher galt zum Beispiel das Entfernen von Nachhall aus komplexen Signalen als nicht machbar, bis wir 2012 Unveil vorstellten.“ Kurz darauf erwartete man jedoch bereits eine perfekte, artefaktfreie Performance für alle Subtypen des Problems. In zehn Jahren werde das Enthalten und Demixen von Songs laut Goekdag zum Standard gehören. Wie bei jeder anderen technologischen Entwicklung müsse man aber auch der KI Zeit für den Entwicklungsprozess geben.

Gebrauchsmusik

Da sich Platten und CDs heutzutage kaum noch verkaufen, haben viele ehemalige Produzenten umgesattelt und komponieren jetzt Musik für Werbetrailer, Filme oder Computerspiele. Solche Gebrauchsmusik soll in jeder Szene die passende

Emotion triggern. Auch hier kann KI ihnen künftig die Arbeit abnehmen: „Ein mögliches Zukunftsszenario wäre, dass der Anwender nach Hause kommt, eine Kamera per KI am Gesicht seinen Gefühlszustand erkennt und gleich on the fly die passende Musik generiert.“ Denn ein erheblicher Teil der Musik habe heute einen funktionalen Charakter, erklärt Goekdag. Sie existiere im Kontext einer anderen Hauptbeschäftigung: Sei es ein Film, ein Computerspiel oder das Einkaufen im Supermarkt. Und hier könne KI in absehbarer Zukunft durchaus auf Augenhöhe mit von Menschen hergestellter Musik spielen.

Einen weiteren Einsatzzweck sieht Goekdag bei Streaming-Diensten wie Spotify. Ähnlich wie Künstler auf ihren Alben Hits mit weniger elaborierten Werken kombinieren, müssen auch bei Streaming-Playlists nicht alle Titel Highlights sein. „Dieses Füllmaterial könnte künftig immer mehr von KI produziert werden, das mit entsprechenden Genres trainiert wurde“, orakelt Goekdag.

Das wirft neue Fragen auf: Hat eine solche Musik einen Urheber? Ist der Urheber dann vielleicht der Entwickler der KI? Oder liegt die Urheberschaft bei den Urhebern des Trainingsmaterials?

„Das ist eine komplexe, hochdifferenzierte zu führende Debatte, die in akademischen und juristischen Kreisen zunehmend an Fahrt gewinnt und dringend auch öffentlich geführt werden müsste“, sagt Goekdag. Die aktuelle öffentliche

Diskussion in Europa sei da noch nicht angekommen: „Die EU-Urheberrechtsnovelle erweckt den Eindruck, von Leuten geschrieben worden zu sein, die nicht auf dem neuesten Stand sind, was solche Entwicklungen angeht. Sie macht aus meiner Sicht viele Dinge komplizierter, ohne die Weichen für absehbare zukünftige Entwicklungen zu stellen.“

Maschinenstürmer

Doch die Musik ist damit nicht dem Tode geweiht, im Gegenteil: „Es wird eine Gegenbewegung geben: rebellierende Jugendliche, die sich vom Status quo ihrer Eltern abwenden und organische Musik komplett ohne KI zelebrieren. Aber es wird primär eine Subkultur sein. Im Mainstream werden wir künftig sehr viel KI-generierte Musik hören“, ist sich Goekdag sicher.

Der Vocoder von Kraftwerk und Autotune von Antares waren erst der Anfang. „Analog zu den Deep Fakes im Videobereich, wird es Software geben, die die eigene Stimme in die eines anderen verwandelt – Elvis oder Madonna oder auch Yoda oder Barack Obama.“ Der Vocaloid von Yamaha zeige da bereits, in welche Richtung das gehen könne.

Hier liegt die Dialektik der technischen Entwicklung: Auf der einen Seite freuen sich Künstler, wenn sie immer mächtigere Werkzeuge bekommen. Auf der anderen Seite haben sie jedoch Angst davor, durch diese Technik wegrationalisiert zu werden. (hag@ct.de) **ct**



Dank KI separiert „Unmix: Drums“ Schlagzeug-Spuren aus einem kompletten Mix. Für Magix Acid Pro hat Zynaptiq die Technik inzwischen auch auf Gesangsspuren ausgedehnt.