

# vdt MAGAZIN

Die Fachzeitschrift für Tonmeister

## MUSIKPRODUKTION



**10**

Has music streaming killed the instrumental intro?

**12**

Der Schlüssel für einen emotionalen Song

**20**

Ein Hit braucht die Überraschung

**32**

KI in der Markenkommunikation



Foto: ©semisatch - stock.adobe.com

KI kann alles und nichts

# KI IN DER MARKENKOMMUNIKATION

Interview und Text **Elke Wisse**

Die Sound Agentur WESOUND in Berlin beschäftigt sich u.a. mit der klassischen Brand Sound Entwicklung. Um die Musikauswahl für den Einsatz in der Markenkommunikation künftig noch besser gestalten zu können, nutzt das Unternehmen das Potenzial künstlicher Intelligenz. In einem durch das „Zentrale Innovationsprogramm Mittelstand (ZIM)“ des BMWi (Bundesministerium für Wirtschaft und Energie) geförderten FuE-Projektes mit dem Arbeitstitel „iMozart“ wird die algorithmische Musik-Komposition im Kontext von Audio Branding nutzbar und zukunftsfähig gemacht. Lars Ohlendorf spricht über den Einsatz von KI, die Vor- und Nachteile sowie über seine Erfahrungen.

## WIE FUNKTIONIERT KI? WAS KANN KI LEISTEN IM BEREICH DER MUSIKKOMPOSITION?

Der Begriff KI ist nicht genau abgegrenzt. Im Prinzip geht es darum, dass sich ein Algorithmus in verschiedenen Situationen selbstständig zurechtfindet und sich bestenfalls auch weiterentwickeln kann. Die grundsätzliche Funktion einer KI könnte man an einer einfachen Kompressor Schaltung verdeutlichen: Das Eingangssignal wird gemessen und

entsprechend die Regelschaltung angepasst – eine simple Schaltung mit einem immer gleichen, reagierenden Verhalten. Wenn nun zusätzlich ein kleines Mathematikmodul ergänzt wird, das auf Basis der Messwerte laufend eine einfache Regressionsanalyse rechnet, dann kann der Kompressor bereits in Maßen vorausschauend agieren. Es entstehen die ersten Muster und damit auch die ersten Anzeichen einer intelligenten Schaltung.

In einem nächsten Schritt könnte der Kompressor nicht nur die Statistik kennen, sondern sich auch mit einer externen Datenbank abgleichen. Zu den Mustern kommt ein Gedächtnis. Er kann so das anliegende Signal und das eigene Regelverhalten mit den in der Datenbank gespeicherten Regelmustern vergleichen und daraus vorausschauend für das anliegende Signal die beste Einstellung ableiten – eine schon intelligentere Vorgehensweise.

Wenn die Schaltung dann noch selbstständig in diese Datenbank schreibt, sich also nicht nur bedient, sondern auch Daten ablegt, dann berühren wir schon die Bereiche des die Bereiche des maschinellen Lernens. Das Programm kann sich damit prinzipiell weiterentwickeln.

### WAS HEISST DAS ÜBERTRAGEN AUF MUSIK?

Musikprogramme, die in Echtzeit Musik generieren, gibt es schon sehr lange. Diese generativen Systeme gehen unter anderem auf die von Clarence Barlow entwickelte Software Autobus zurück. Diese Software tut im Prinzip nichts anderes, als unter vorgegebenen Grenzen zufällige MIDI Werte für die Ansteuerung von Klängen auszugeben. Ich kann zum Beispiel definieren, wie groß der melodische Ambitus ist, welcher Skala die Melodie entstammen soll, wie stark die Bindung an diese sein soll und einiges mehr. Musik nach bestimmten assoziativen oder aktivierenden Parametern allerdings, zum Beispiel „heiter“ oder „aufregend“, kann dieses Programm noch nicht generieren, weil ihm dazu die entsprechenden Operationalisierungen fehlen. Wenn das Programm also wüsste, was „heiter“, „aufregend“ oder „beruhigend“ in der Musik konkret bedeuten, könnte es die musikalischen Parameter durchaus entsprechend einstellen. Nichts anderes täte ja ein Komponist.

### WIE BRINGT MAN DAS DER DATENBANK BEI?

Gerade sehr aktuell ist der Deep Learning Ansatz. Hier arbeitet man mit meist größeren Datenmengen und einem lernenden Algorithmus, einem neuronalen

Netzwerk. Es werden Musikstücke eingespielt, und Versuchspersonen geben dazu einfache, subjektive Reaktionen ab. Gleichzeitig werden die Eigenschaften des musikalischen Materials untersucht. Der Algorithmus sucht dann nach noch so winzigen Korrelationen zwischen diesen sehr verschiedenen Messdaten aus Befragung und digitalem Signal. Je größer dabei der Datensatz, desto besser die Ergebnisse. Für uns hätten die so zustande gekommenen Daten allerdings noch keinen unmittelbaren Wert. Denn der Algorithmus kann nun zwar vielleicht erkennen, welche Wirkung ein Musikstück hat, er kann aber noch lange kein solches erzeugen. Auch wüsste man so nicht, welche konkreten Bestandteile der Musik zu welchen Attributen führen. Es bleibt eine Black Box.

Ganz im Gegensatz dazu stehen die klassischen, empirischen Ansätze: quantitative und/ oder qualitative Befragungen. Für uns besonders vielversprechend ist hier die Expertenbefragung. Wie schätzen Komponisten, sowie Design- und Kommunikations-Experten bestimmte Musiken ein, wie würden sie bestimmte Stimmungen instrumentieren, wie komponieren?

Für iMozart nutzen wir zum einen die bereits verfügbaren empirischen Studien aus diesem Bereich und zum anderen eigens durchgeführte qualitative Interviews. Hiermit versuchen wir, Zusammenhänge zwischen emotionalen Attributen mit bestimmten musikalischen Parametern zu finden.

Drittens trainieren wir unseren Algorithmus mit Kompositionen, die von erfahrenen Komponisten zu einem bestimmten Markenbegriff komponiert wurden. Der Algorithmus analysiert dann die musikalischen Parameter, um so Bezüge zu den Markenbegriffen herzustellen. Auf diese Weise kann sich das System auch nach seiner Fertigstellung noch selbst verbessern.

### WAS KÖNNEN PROGRAMME IN DIESER HINSICHT BEREITS LEISTEN?

Auf der ganzen Welt wird zum Thema KI im Zusammenhang mit Musik geforscht. Abgrenzen kann man da insbesondere die Bereiche Music Information Retrieval, Musical Performance und Music



**Lars Ohlendorf** ist Tonmeister, Sound Designer und Komponist mit zwanzig Jahren Berufserfahrung. Er ist Spezialist für 3D Audio, interaktives Design und generative Systeme. Bei WESOUND ist Lars Ohlendorf als Creative Director | Head of Design verantwortlich für die Entwicklung von Sound-Konzepten und Musik-Strategien. Lars Ohlendorf hat zunächst die SAE besucht, danach an der HAW in Hamburg den Master Studiengang Sound/Vision belegt.

Foto: Christoph Mannhardt

Generation. Und die Ergebnisse sind mitunter bereits sehr vielversprechend. Problematisch aber ist seit Jahrzehnten eine Frage ganz außerhalb von KI: Welche Parameter in der Musik lösen welche Emotionen beim Menschen aus? Damit habe ich mich bereits im Studium auseinandergesetzt und musste schnell lernen, dass dies ein Feld geringer Valenz ist; es ist schlicht nicht möglich, pauschale Vorhersagen zu treffen. Da gibt es viel zu viele Störvariablen.

Nach dem Circumplex Modell von Russell lassen sich immerhin einfache Grundstimmungen in das zweidimensionale Valence-Arousal-Modell einordnen. Daraus lassen sich beispielsweise folgende Schlüsse ziehen: Je schneller eine Musik wird, desto anregender ist sie auch. Das kann dann entweder fröhlich und ausgelassen wirken oder auch stressig und überfordernd. Je langsamer hingegen eine Musik wird, desto

beruhigender, größer oder auch missmutiger wirkt sie gemeinhin.

Der Volksmund allerdings, bestimmte Skalen ließen bestimmte Affekte assoziieren („Dur fröhlich, moll traurig“) gilt nur bedingt. Diese Tongeschlechter sind stark kulturell abhängig und damit auch geographisch geprägt. Versuche mit Kleinkindern, denen Dur/moll unbekannt sind, haben gezeigt, dass die große oder kleine Terz überhaupt keine Bedeutung für die Stimmungszuschreibung hat. Viel hilfreicher sind hier die Parameter Tempo und Rhythmus, wie sich schon im Valence-Arousal-Modell andeutet. Diese lassen am ehesten eine Vorhersage einer attribuierten Emotion zu, weil sie direkt

die Anhänger bestimmter Kulturkreise mit hoher Wahrscheinlichkeit gelernt haben und die damit auch abrufbar sind. Auf diese kann man aufbauen.

Wir sprechen also über musikalische Stereotype – man darf auch Klischees sagen. Zum Beispiel verbindet man die Klänge einer schrammeligen Akustikgitarre oft mit Natur und Lagerfeuerromantik, Big Band Swing durchaus mit Mafia, Prohibition und Bubikopf Frisuren, oder Hip Hop mit Urbanität, Battles und Underdogs. Diese Liste kann man endlos weiterführen, und man kann dabei auch



▲ Lars Ohlendorf im Studio

Foto: Jungblut

und mit körperlicher Bewegung erfahrbar sind.

Es ist aber durchaus nicht so, dass man exakt wüsste, was in der Musik bestimmte Emotionen auslöst. Deshalb ist es auch so schwierig, einen allgemeinen Algorithmus zu entwickeln, der musikalische Reize generieren kann, die uns in bestimmten Stimmungen abholen oder uns in bestimmte Stimmungen versetzen.

### GILT DAS FÜR ALLE MUSIKGENRES?

Genres sind ein gutes Stichwort! Denn hier verlässt man den körperlich geprägten Circumplex Bereich und konzentriert sich auf kulturelle Besonderheiten, also Klänge und klangliche Zusammenhänge,

munter anecken, denn es sind natürlich grobe Vereinfachungen, die auch durchaus nicht immer für alle und zu jeder Zeit gelten. Klassische Musik, dieser so indifferente Container, kann alles sein: abgehoben, elitär und arrogant, aber genauso auch exklusiv, anspruchsvoll und intelligent. Es kommt immer auf den Blickwinkel an. Genau diese Stereotype versuchen wir in eine Datenbank zu portieren, um sie dann abrufbar zu machen.

### WELCHE SOFTWARE GIBT ES BEREITS?

Es gibt eine ganze Reihe von Software, die Musik in Echtzeit generieren kann. Auf der ganzen Welt arbeiten Komponisten und Forscher an den wunderlichsten Kreationen, viele davon auf Plattformen wie max/msp oder puredata.

Aber auch für den Konsumentenmarkt gibt es die ersten Applikationen. Im Apple App Store und auf Google Play ist beispielsweise die Software Endel verfügbar. Das ist eine App, die die individuelle Situation des Users durch Gesundheitsangaben aus Fitness-Apps, dem Pulsschlag, Standortdaten und der Tageszeit auswertet, um eine dazu passende Musik zu generieren. Dabei kann der User Musik zum Entspannen, Konzentrieren, Einschlafen oder Bewegen auswählen.

Das ist spannend und gleichzeitig merkwürdig, denn man fragt sich, warum man diese Musik hören soll. Es gibt nichts, worauf ich als Hörer Bezug nehmen kann, nichts, was mich inhaltlich emotional ansprechen könnte. Wenn wir hier von Popmusik sprechen – und mehr als das kann eine KI-Komposition inhaltlich nicht sein – dann ist doch die Komposition selbst nur ein ganz kleiner Teil des popmusikalischen Ganzen. Es geht auch um Texte, die Künstlerpersönlichkeit, die Videokunst, das Artwork im Album, die Rezeption usw.

Allein auf Basis einer KI erstellte Musik wird sehr funktional, und so werden diese Programme auch beworben. Damit wird aber auch unser Hören, und letztendlich der ganz Mensch funktional. Musik hat keinen Selbstzweck mehr, sondern soll mich in die ideale Situation für konzentrierte Arbeit versetzen, sei es im Büro, an der Hantelbank oder was man sonst noch in den funktionalen Kosmos einbeziehen möchte. Schlafen zum Beispiel!

**GIBT ES AUCH PROGRAMM FÜR DIE PROFESSIONELLE ANWENDUNG?**

Es gab zum Beispiel ein großes Projekt am Universitätskrankenhaus Eppendorf (UKE) in Hamburg zusammen mit der Hochschule für Musik und Theater HfMT mit Prof. Georg Hajdu und Prof. Dr. Eckhard Weymann [1].

Auf Basis des Autobusk Algorithmus (s. o.) wurde hier ein Tool für die Hintergrundbeschallung von Wartezimmern im UKE entwickelt. Die Frage war, wie man auf große Dauer Musik generieren kann, die zum einen therapeutischen Zwecken genügt und zum anderen wiederholungs-frei ist.

Der Autobusk Algorithmus wurde dafür so modifiziert, dass er auf Sensor-daten wie Lautstärke im Raum, Tages-zeit u. a. reagiert und die musikalischen Komposition daraufhin anpasst. Das ist zwar noch keine echte KI Anwendung, aber schon ein intelligent reagierendes System.

**WAS STECKT HINTER IMOZART?**

iMozart ist ein Kompositions-Tool, das Musik in Echtzeit auf Basis einer gege-benen Taxonomie generieren kann. Die Software kann einfache Attribute wie sportlich, rau oder edel einer bestimm-ten Musikalität oder Klanglichkeit zuwei-sen. Daraus erstellt das Programm ein-fache musikalische Strukturen, die diese definierten Attribute unterstützen.

Wir wollen das Tool in Markenklang-Projekten als Schnittstelle zwischen uns als Agentur und dem Kunden nutzen. Als geschulte Designer haben wir über vie-le Jahre gelernt, über Ton und Klang zu sprechen, und wir wissen genau, wie man mit musikalischen Reizen eine Marke ideal repräsentieren kann. Aber im kreativen Prozess geht es natürlich darum, die Vor-stellungen des Kunden zu erfassen und umzusetzen, nicht unsere eigenen. In der Regel nähert man sich dann per Trial & Error einer einvernehmlichen Vorstellung eines Designs an.

Für uns ist iMozart definitiv ein De-sign-Tool, mit dem wir zusammen mit un-seren Kunden Ideen entwickeln wollen. Und es ist ein ganz pragmatischer An-satz. Wir verhandeln sozusagen täglich die Bedingungen, unter denen das Tool



▲ Relax-Modus      ▲ Sleep-Modus      ▲ Focus-Modus      Fotos: Endel

Wenn wir – und das ist die Idee dabei – ein Tool hätten, mit dem man spiele-risch im Gespräch zusammen mit un-se-rem Kunden grundsätzliche musikalische Ideen direkt hörbar machen und er auch selbst Parameter verändern könnte, dann könnten wir viel schneller ein gemeinsa-mes Verständnis erzeugen. Und mit dem Komponisten könnten wir daraufhin ein musikalisch konkretes Briefing kommu-nizieren. Es geht also keinesfalls darum, den Komponisten zu ersetzen, sondern darum, eine Kommunikationsbarriere zu überbrücken.

trotzdem funktionieren kann. Wir lernen sehr viel dabei, zum einen über unsere eigenen Wahrnehmungen, aber auch wie man kreativ mit Algorithmen umgehen kann. Die Ergebnisse der KI immer wieder zu hinterfragen und die Prozesse zu opti-mieren macht die Arbeit so reizvoll.

**WIRD DER OUTPUT VON IMOZART ZUM ENDPRODUKT?**

Das Ergebnis der Software ist immer nur die Basis für die weitere Bearbeitung. iMozart liefert einfache Strukturen, die nach viel klingen, weil man über Genre und Instrumentierung sehr viel beein-flussen kann, aber der musikalische In-halt bleibt stets recht überschaubar. Aus dem Algorithmus entstehen keine ge-schliffenen Sinfonien. Und das bräuchte es auch gar nicht.

„Hello World“ ist der Name des ersten Pop Musikalbums, das mithilfe einer KI Software produziert wurde. Der Franzose Benoit Carré hat das Album unter dem Künstlernamen SKYGGE produziert, zusammen mit der KI Flow Machines.

**Warner Music** hat als erste Plattenfirma die App „Endel“ unter Vertrag genommen, deren Algorithmus passend zur Stimmung des Hörers Musik komponiert. Die KI Software „Endel“ hat bereits fünf von 20 avisierten Alben produziert. Erhältlich sind die fünf „Sleep“ Alben auf Apple Music und iTunes. Es werden jeweils fünf Relax-, Focus- und On-the-Go Alben folgen.

Sobald die Struktur und die Idee eines Markenklangs auf die beschriebene Art und Weise ermittelt wurden, ist stets der Komponist gefordert. Auf diese Weise bekommt er eine sehr genaue Vorstellung davon, was der Kunde wünscht. Er kann dann darauf aufbauend seine Kreativität einbringen und auch von den Computereergebnissen aktiv abweichen. Nur so entstehen Ideen, die eben kein Algorithmus erzeugen kann. Denn der wird uns immer nur das liefern, was wir in irgendeiner Form vorher eingespeist haben, bzw. was wir schon kennen. Wie eine Echokammer. Die kennen wir ja schon aus anderen Bereichen.

### **WIE REAGIEREN KOMPONISTEN AUF DIE „KÜNSTLICHE MUSIK“? SEHEN SIE ES ALS HILFSMITTEL ODER SEHEN SIE SICH IN IHRER KREATIVITÄT EINGESCHRÄNK?**

Bis jetzt habe ich nur positives Feedback erhalten. Es ist eine grundsätzliche Frage, wie man mit dem Kunden ein Briefing erarbeiten kann, das später auch eine gewisse Verbindlichkeit hat. Gerade das Hören ist ja – im Gegensatz zum Sehen – eine sehr individuelle Angelegenheit. Ein gezeichnetes Quadrat nimmt jeder als

solches wahr, bei Musik relativiert sich die vermeintlich exakte Beschreibung des Gehörten sehr schnell. Wir müssen als Agentur schauen, dass der Kunde am Ende genau das bekommt, was er möchte und was er erwartet.

### **ES GIBT SCHON SYSTEME, DIE ORCHESTERMUSIK SCHREIBEN KÖNNEN...**

Ja, richtig, die gibt es, und es entsteht Musik, die einfach, bzw. mathematisch erfassbar ist. Mein Vater hat mir vor vielen Jahren einmal eine Software geschenkt, die Fugen und Choräle schreiben kann. Man gibt einen Cantus Firmus ein und, zack, ist die Fuge fertig. Verwendet man einen Deep Learning Algorithmus, kann dieser beispielsweise die zehn Sinfonien von Mahler analysieren und ein Stück generieren, das den Originalen sehr ähnlich ist. So geschehen bei Schuberts Unvollendeter Sinfonie, die von einer KI „vollendet“ wurde. Ich habe diese quasi Fortsetzung noch nie gehört, aber sie interessiert mich, ehrlich gesagt, auch nicht. Interessant ist doch vor allem die Frage, was das Werk in der ursprünglichen Form beim Hörer auslöst. Und auch hier geht es wieder um die Künstlerpersönlichkeit und ihre natürlichen Grenzen. Wenn wir die aufheben, fällt doch alles in sich zusammen.

Der kreative Gehalt dieser Ergebnisse ist in meinen Augen fraglich. Mahler war ein großartiger Komponist, warum sollte jetzt etwas Ähnliches erstellt werden? Wir sollten immer das Ziel verfolgen, etwas Neues zu kreieren, und da ist der erfahrene Künstler gefragt, der originäre Musik mit Bedeutung schreibt und nicht nur Stereotype kopiert.

Warner Music hat kürzlich ein Album, produziert von KI Software Endel, auf den Markt gebracht. Auch das macht für

mich überhaupt keinen Sinn. Denn generative Musik mit KI ist doch gerade dadurch spannend, dass sie sich immer neu schreibt und nie wiederholt. Wird die Musik in das klassische Albumformat übertragen, verliert sie diese Besonderheit. Wohlgermerkt die einzige, die sie hat.

### **WÜRDEN DER NORMALE HÖRER DEN UNTERSCHIED ZWISCHEN EINER KI UND EINER HANDGEMACHTEN MUSIK HÖREN?**

Ich weiß nicht, ob ich selber den Unterschied hören würde. Wenn zum Beispiel eine KI Komposition von einem Orchester intoniert wird, kann dabei durchaus etwas künstlerisch Wertvolles entstehen, allein schon, weil erfahrene Musiker am Werk sind.

Bei popmusikalischen Zusammenhängen ist das ähnlich. Ohne Gesang hört man wahrscheinlich keinen Unterschied, denn Popmusik ist in Teilbereichen so stark formatiert, dass sie algorithmisch relativ einfach zu erfassen ist.

### **WO GEHT DAS KOMPONIEREN MIT KI IN DEN NÄCHSTEN JAHREN HIN?**

Wenn es gut läuft, werden wir immer mehr Tools bekommen, die uns helfen, uns musikalisch weiter zu entwickeln. iMozart sehe ich als einen kleinen Baustein dazu. Wenn wir die KI Kreationen als Basis nehmen und darauf aufbauend mit „echten“ Komponisten weiterarbeiten, dann kann nur etwas Besseres dabei herauskommen. Wenn wir es als Werkzeug begreifen, dann ist KI eine sehr gute Unterstützung. Es könnte natürlich auch passieren, dass das Ergebnis aus der KI einfach als Endprodukt übernommen wird. In dem Fall würde die künstlerische Qualität langfristig massiv in den Keller gehen. Wahrscheinlich wird es am Ende eine Kombination aus beidem sein. Als Gestalter und Hörer gleichermaßen sind wir alle gefragt diese Entwicklung in die richtigen Bahnen zu lenken.

#### Referenzen

[1] <http://georghajdu.de/wp-content/uploads/From-Atmosphere-to-intervention-final.pdf>.

## PRAXISNAHE, ANWENDERORIENTIERTE UND FACHLICH FUNDIERTE BERICHTE AUS DER PROFESSIONELLEN AUDIOBRANCHE

### Abopreise:

Inland:	<b>35,00 Euro</b>
Ausland:	<b>42,00 Euro</b>
Studentenabo:	<b>29,75 Euro</b>
Einzelheft:	<b>8,00 Euro</b>

Interviews mit Profis  
und Erfahrungsberichte  
aus den Bereichen

- Tonträgerproduktion
- Hörfunk
- Film
- Fernsehen
- Multimedia
- Bühne
- Theater
- Forschung
- Entwicklung und Lehre

Jede Ausgabe mit  
wechselndem  
Themenschwerpunkt!

