



HOCHSCHULE
DER MEDIEN

Künstliche Intelligenz in der Musikproduktion

Till Gostner
45505

Einleitung

KI hat sich in den letzten Jahren immer weiter verbreitet und in vielen Branchen eine unterstützende Funktion eingenommen. In der Medienproduktion kann zB. schon heute Midjourney genutzt werden um hochqualitative Bilder verschiedener Art zu generieren. Auch für die Audioproduktion gibt es immer mehr Tools welche den Produktionsprozess unterstützen und teilweise schon in großen Teilen ersetzen. Die Nutzung von KI in der Musikproduktion lässt sich grob in drei Bereiche aufteilen.

Produktion

Mixing

Mastering

Dieses Paper soll einen Überblick über den aktuellen Stand der Technik von KI in diesen Bereichen bieten. Es werden hierbei nur Tools und Services betrachtet welche bereits öffentlich verfügbar sind.

Produktion

Bei der Produktion entsteht die Musik grundlegend. Es werden Texte geschrieben, und eingesungen. Melodien, Harmonien und Rhythmen werden erfunden und auf verschiedenen Instrumenten gespielt und aufgenommen. Töne und Sounddesigns werden mit Synthesizern erzeugt. Die Aufgenommenen Klänge werden arrangiert und in der Zeit passend angeordnet. Dies ist der wichtigste Teil der Musikproduktion ohne den die weiteren Schritte nicht stattfinden können.

Chat GPT

Chat GPT ist ein KI-Tool welches, durch vom Nutzer eingegebene Prompts, Texte generieren kann. Dabei lassen sich auch spezifisch Songtexte erstellen. Chat GPT kann dabei jegliches Thema in den Text einfließen lassen und versteht die grundlegenden Strukturen von Pop-Songs. Die KI kann den Text auch mit Reimen generieren. Sollte einem der Text nicht gefallen, kann man Chat GPT Verbesserungsvorschläge geben. Die Genauigkeit des Prompts ist entscheidend für die Qualität des generierten Texts. Das Tool ist gut geeignet um Schreibblockaden zu umgehen. Oftmals ist ein grobes Thema für einen Songtext bereits vorhanden aber die praktische Umsetzung gestaltet sich schwierig. Hier kann Chat GPT in kürzester Zeit viele verschiedene Vorschläge machen, welche man dann selbst oder wiederum mit Chat GPT verbessern und dem eigenen Geschmack anpassen kann.

Aiva

Aiva ist ein Tool welches Melodien und Rhythmen generieren kann. Die KI wurde mit über 30000 verschiedenen Musikstücken trainiert. Die KI hat gelernt Muster in der Musik zu erkennen, um somit vorhersagen zu können welche Noten sich als nächstes am besten anhören. Zusätzlich wurden den verschiedenen Musikstücken Stimmungen zugeordnet, sodass die KI versteht welche Kombination von Noten welche Stimmung erzeugt. Es lassen sich Stil, Stimmung, Geschwindigkeit, Tonart und die Instrumentierung auswählen. Für die

einzelnen Instrumente lassen sich auch die Notendichte und Komplexität der Melodien angeben. Für jedes Instrument generiert Aiva dann Midi Noten welche man sich herunterladen kann. Generiert man mehrere Musikstücke mit den gleichen Parametern entsteht jedes mal ein leicht anderes Musikstück. Es ist auch möglich die Melodien und Rhythmen als Audiospuren auszugeben. Hierfür werden Plugin-Instrumente verwendet. Was die KI noch nicht kann, ist die Umsetzung von Textprompts. Es müssen jedesmal in einer Eingabemaske, alle Parameter manuell festgelegt werden. Das Tool ist außerdem dazu in der Lage hochgeladene Musikstücke zu Analysieren. Basierend darauf kann die KI Musikstücke erstellen welche durch Wahl der Instrumentierung und Noten eine ähnliche Atmosphäre erzeugen. Das Tool bietet verschiedene Zahlungsmodelle an. Diese bestimmen darüber ob Aiva als Urheber angegeben werden muss, wer das Copyright hält und wie das Musikstück monetarisiert werden darf. Die Website von Aiva zeigt ein beeindruckendes Beispiel, bei dem von der KI generierte Noten von einem echten Orchester gespielt werden.

Emergent Drums

Emergent Drums ist ein VST-Percussion Plugin. Die KI welche im Hintergrund des Plugins arbeitet wurde mit einer großen Anzahl Percussion Sounds trainiert. Das Tool kann unendlich viele verschiedene Kicks, Claps, Snares usw. erzeugen. Die Samplegeneration funktioniert nur Online und findet somit wahrscheinlich auf einem externen Rechensystem statt. Generierte Samples können aber lokal gespeichert

werden. Es können auch vorhandene Samples in das Programm geladen werden um dann ähnliche Sounds in demselben Stil zu erzeugen. Das Tool kann daher zB. genutzt werden um Urheberrechtlich geschützte Samples leicht abgeändert neu zu generieren um das Copyright Problem zu umgehen.

Mixing

Beim Mixing werden die im ersten Schritt aufgenommenen oder synthetisierten Spuren bearbeitet. Es werden die Lautstärkeverhältnisse der einzelnen Spuren angepasst und verschiedene Effekte eingesetzt um den Gesamteindruck des Stücks zu verbessern. Die wichtigsten Parameter sind hierfür die Lautstärke, das Panning im Stereofeld und die Anpassung des Frequenzgangs durch Equalizer. Je nach Musikgenre ist dieser Schritt mehr oder weniger für das Endprodukt ausschlaggebend. Bei einem Singer-Songwriter Stück welches nur aus Gesang und einer Akustikgitarre besteht und authentisch klingen soll wird das Mixing sehr rudimentär ausfallen. Bei elektronischer Musik ohne Text liegt der Inhalt der Stücke im Klang der Sounds, welcher maßgeblich durch die Nachbearbeitung im Mixing entsteht. Der Übergang zwischen Produktion und Mixing ist fließend und je nach Genre unterschiedlich.

Roex Automix

Roex Automix ist eine KI welche teilautomatisiert den Mixingprozess übernimmt. Das Tool hat jedoch mehrere Schwierigkeiten welche die praktische Nutzbarkeit stark einschränken. Zuerst müssen die Einzelspuren hochgeladen werden. Für jede Spur muss dann eine von drei Prioritäten, eine grobe Panningvorgabe, die Stärke des Halls und das Instrument um welches es sich handelt angegeben werden. Es werden somit genau die Entscheidungen welche einen Großteil des Gesamteindrucks ausmachen von der KI nicht abgenommen. Zudem liefert die KI selbst mit gutem Quellmaterial keine guten Ergebnisse. Nur bei optimal vorbereiteten Spuren kann das System passable Ergebnisse erzielen. Der Vorteil des Tools erschließt sich im aktuellen Zustand noch nicht. Die Arbeit welche nötig ist um Spuren in einen Zustand zu bringen mit dem die KI gute Ergebnisse abliefert, ist deutlich größer als die Arbeit welche nötig wäre den Mix danach manuell fertigzustellen. Auch kann die KI keine künstlerisch wichtigen Entscheidungen treffen, wie einen passenden Hall für die Stimmung des Stücks auszuwählen. Die Parametrisierung ist sehr grob, so kann es dazu kommen das ein exotisches Instrument nicht in der Liste auswählbar ist. Das Tool ist in diesem Zustand nicht als brauchbar zu beurteilen, da es keine wichtigen Entscheidungen abnimmt und die Zeitersparnis aufgrund des großen Vorbereitungsaufwands minimal ist.

Mastering

Das Mastering ist die letzte Instanz vor der Veröffentlichung eines Musikstücks. Es wird der Frequenzgang und die Lautheit der finalen Stereospur angepasst um sich grob in das Klangbild anderer Stücke des gleichen Genres einzufügen. In der Regel ist dieser Schritt der unwichtigste im Hinblick auf das Endprodukt. Dies ist jedoch wie beim Mixing von vielen Faktoren Abhängig. Vor allem von der Qualität des Mixes. Ein Mix welcher auf einer Abhöre mit ungeradem Frequenzgang erstellt wurde, wird wahrscheinlich mehr Nachbearbeitung nach sich ziehen als ein Mix welcher auf einer gut eingestellten Abhöre erstellt wurde. Ein guter Mix kann unter Umständen auch gar kein Mastering benötigen.

Es gibt zahlreiche Online Mastering Services, welche lediglich den Upload der Stereospur und die Auswahl eines Mastering Presets erfordern. Beispiele hierfür sind LANDR, eMastered und Bandlab. Die Funktionalität der Services ist sehr ähnlich weshalb im Folgenden auf KI Mastering im allgemeinen eingegangen wird. Die Presets tragen bei den verschiedenen Services unterschiedliche Namen, bedienen aber ähnliche Geschmäcker. Die Anbieter haben Presets für lautes Mastering, bei dem vermutlich stärkere Kompression und Saturation verwendet wird um in Clubs oder im Auto für maximale Hörbarkeit zu sorgen. Ebenso gibt es Presets welche einen eher natürlichen Klang und einen größeren Dynamikumfang erreichen möchten. Die meisten Anbieter haben auch eine Art mittleres Preset welches einen Kompromiss der ersten beiden genannten darstellt. Das Ergebnis ist ähnlich

wie bei KI Mixing Tool stark abhängig vom Quellmaterial. Bei einem guten Mix ist das Ergebnis mit manuellem Mastering auf Augenhöhe. Auch hier sich allerdings je nach Quellmaterial Insgesamt kann man aber sagen dass KI Mastering bereits heute in einem in der Praxis nutzbaren Zustand ist.

Text to Music

Die nächste Stufe der KI Unterstützung bei Musikproduktion ist „Text to Music“. Hier werden alle bisher beschriebenen Schritte übersprungen und es wird nach Eingabe eines Prompts ein fertiges Musikstück generiert. Musicgen.com ist ein solcher Text to Music Generator. Musicgen generiert etwa 15s lange instrumentale Musikstücke. Angaben von Genre, Stimmung und Instrumentierung werden meist deutlich erkennbar umgesetzt. Bei einer Neugenerierung der Musik mit demselben Prompt erhält man jedesmal ein anderes aber stilistisch ähnliches Ergebnis. Die KI ist auch in dazu in der Lage eine angegebene BPM in vielen Versuchen genau umzusetzen. Manchmal verändert sich die BPM für kurze Zeit wodurch der Rhythmus verschoben wird, vor und nach diesen Aussetzern bleibt die BPM aber meist des Prompts entsprechend konstant. Obwohl einzelne Instrumente erkennbar sind haben die Musikstücke immer einen rohen oder diffusen Klangcharakter. Dieser hört sich an als wäre ein weißes Rauschen über der Musik welche die Melodie und Rhythmen ineinander verschwimmen lässt. So lassen sich bei den Percussion-Elementen die Snares, Toms und Hi-Hats oft nicht eindeutig auseinander halten. Musicgen ist noch weit davon entfernt

fertige Musikstücke mit Text und Gesang zu generieren, gibt aber einen Vorgeschmack auf das was noch folgen wird.

Schlagwort KI als Marketing

Sucht man im Internet mit Stichworten wie „KI Musik Generator“ oder „AI Plugins“ finden sich zahlreiche Anwendungen und Tools welche das Schlagwort KI lediglich als Marketing verwenden. In fast allen Artikeln welche über KI Plugins berichten wird zB. das Plugin Soothe 2 von Oeksound aufgelistet. Das Plugin hat jedoch nichts mit KI zu tun. Es handelt sich um einen spektralen Kompressor welcher nicht mithilfe von KI funktioniert. Interessant ist dass die Herstellerseite das Wort KI nicht verwendet. Hier nutzen die Seiten, welche über Musiktechnik berichten, das Schlagwort um mehr Traffic zu generieren.

Soundraw.io bezeichnet sich selbst „AI Music Generator“. Das Tool kann nach Auswahl von Stimmung, Geschwindigkeit und Genre in wenigen Sekunden mehrere professionell klingende Musikstücke bereitstellen. Blogger Axel Napolitano hat jedoch herausgefunden dass es sich auch hier nicht um KI handelt. Die Website basiert auf einer großen Sample und Loopbibliothek, die je nach ausgewählten Parametern Musikstücke zusammenschneidet. Eine KI ist jedoch zu keinem Zeitpunkt im Spiel.

Bedeutung für die Musikindustrie und Musik als Kunstform

KI hat schon heute einen festen Platz in der Musikproduktion gefunden. Es existieren Tools und Plugins zur Generation von Samples, Melodien und Harmonien. Text to Music ist bereits für jeden kostenlos nutzbar, die Komplexität und Soundqualität ist allerdings noch eher gering. Extrapoliert man vom aktuellen Stand der Technik einige Jahre in die Zukunft, ist es leicht anzunehmen dass sich Musikproduktion stark verändern wird. Es ist absehbar dass man in naher Zukunft Musikstücke aus jedem Genre mit beliebiger Instrumentierung und Komplexität durch einfache Prompts generieren können wird. Inwiefern dies die Musikindustrie verändern wird ist unmöglich genau vorherzusehen. Die Frage, welche sich die Konsumenten hier stellen können, ist ob der Entstehungsprozess der Musik ein Rolle für den Genuss spielt. Diese Frage wird darüber entscheiden zu welchem Grad KI in die Musikproduktion integriert werden wird. Es könnte sein dass sich in der nächsten Generation von Musikern diejenigen durchsetzen werden, welche die neuen Möglichkeiten von KI am kreativsten nutzen.

Quellen

<https://napolitano.de/2021/11/12/soundraw-ai-music-creation-tool-or-just-a-fake/>

<https://soundraw.io/>

<https://www.aiva.ai/>

<https://huggingface.co/spaces/facebook/MusicGen>

<https://twostorymelody.com/best-online-mastering-services/>

<https://chat.openai.com/>

<https://audialab.com/>

<https://musixon.org/2020/09/30/soothe2-spi/>

<https://oeksound.com/plugins/soothe2/>

<https://www.landr.com/de/>