

KI UND MUSIK

„Maschinen
haben keine
Vorstellung von
der gesamten
Form eines
Musikstücks,
sie spucken
wie ein übender
Anfänger Note
für Note aus“

Wann ist Musik für uns Musik? Professor Gerhard Widmer forscht mithilfe von computergestützter Musikanalyse nach Antworten. Wir haben mit ihm über die Verbindung von Mathematik und Musik gesprochen, darüber, ob Maschinen Musik hören können, sowie über die Grenzen von KI beim Komponieren.

Von Michaela Ortis

MARIA ZIEGELBOCK



Ein Klavierstück wird auf einer Bühne in Berlin vierhändig vorgetragen, doch etwas ist anders als sonst. Auf einer Seite sitzt der Pianist, auf der anderen sitzt niemand und die Tasten bewegen sich wie von Geisterhand – denn hier spielt eine künstliche Intelligenz. Dieses eindrucksvolle harmonische Zusammenspiel von Mensch und Maschine demonstrierte erstmals Professor Gerhard Widmer im Rahmen der Auszeichnung mit dem Preis „Falling Walls Science Breakthrough of the Year 2021“. Die Falling Walls Foundation holt bahnbrechende Forschungen auf die Bühne, um Wissenschaft, Wirtschaft und Gesellschaft zusammenzubringen.

Herr Professor Widmer, abgesehen vom Musikgenuss, was lernen wir aus diesem außergewöhnlichen Konzert?

Wir wollten damit etwas Angreifbares von unseren Forschungsfragen zeigen. Denn wenn ein computergesteuerter Flügel mit einem menschlichen Pianisten gemeinsam ausdrucksvoll spielen soll, entstehen viele Fragen: Wie schaffen es Musiker, sich zu synchronisieren? Wie kann ich beim Spielen voraussagen, wann mein Musikpartner die nächste Note spielt? Wie passe ich mich an meine Partnerin an, damit das natürlich klingende Musik ist? Das alles muss der Computer können, damit es nicht mechanisch klingt. Mithilfe von KI und Machine Learning soll er stimmungsvoll spielen, genauso muss er mithilfe von Algorithmen zuhören können. Das Demo-Beispiel ist jedoch nur ein kleiner Ausschnitt der Arbeit unseres Instituts.

Professor Gerhard Widmer studierte Informatik und beschäftigt sich seit mehr als zwei Jahrzehnten mit dem Einsatz von künstlicher Intelligenz in der Musik. Er leitet an der Johannes-Kepler-Universität Linz das Institut für Computational Perception, seine Forschungen wurden bereits zweimal vom Europäischen Forschungsrat mit einem ERC Advanced Grant gefördert. Widmer erhielt 2009 den Wittgenstein-Preis, Österreichs renommiertesten Forschungspreis, und zuletzt die Auszeichnung Falling Walls Science Breakthrough of the Year 2021.

Ihr Forschungsbereich ist demnach viel umfassender?

Ja, bei unserer Grundlagenforschung geht es um akustische Wahrnehmung: Kann man einer Maschine beibringen, Musik zu „hören“, so wie wir Menschen automatisch Takt, Rhythmus oder Stilrichtungen erkennen? Ein zweiter Schwerpunkt ist das ausdrucksvolle Spielen, und das liegt mir persönlich am Herzen: Wie können Musikerinnen und Musiker durch die Art, wie sie ein Stück interpretieren, bestimmte Emotionen transportieren? Diese affektive Kommunikation liegt ja dem Wesen der Musik zugrunde. Wir vermessen mit einem speziellen Computerflügel



exakt, wie lang und wie laut ein Pianist jede Note spielt. Dann untersuchen wir mit Methoden des maschinellen Lernens und mit KI, ob es Gesetzmäßigkeiten gibt, wie man eine Passage spielen muss, damit sie für uns musikalisch und lebendig klingt.

Gibt es einen Zusammenhang zwischen Musik und Mathematik bzw. Algorithmen?

Ganz sicher, viele meiner Mitarbeiter haben musikalische Interessen. Musik und Mathematik haben beide eine abstrakte Struktur. Musik ist ein komplexes System, das aus einzelnen Komponenten wie z. B. Akkorden besteht, die nach bestimmten Regeln zu einem Ganzen zusammengefügt werden, um eine bestimmte Wirkung zu erzeugen. Auch Mathematik hat ihre Ästhetik, vielleicht nicht ganz so sinnlich wie Musik.

Warum verwenden Sie ein Klavier? Wären andere Instrumente auch geeignet?

Eine Tonaufnahme ist nicht exakt genug, um alle Daten, die wir benötigen, zu erfassen. Beim Klavier können wir Computerflügel nutzen, etwa den Bösendorfer CEUS oder von Yamaha, mit denen wir jeden Anschlag und Faktoren wie Tempo, Timing und Lautstärke genau vermessen und aufzeichnen. Für das Klavier spricht auch die

geringere Komplexität, es hat im Vergleich zu einer Violine oder der menschlichen Stimme weniger Freiheitsgrade der Interpretation. Der Datenberg, der bei einem Klavierstück entsteht, ist schon komplex genug.

Was möchten Sie aus diesem Datenberg erkennen?

Man spielt ja ein Stück nicht mechanisch von Note zu Note, sondern in einem Spannungsbogen mit Dramatik und Höhepunkten. Wenn ich ein Musikstück interpretiere, hat das nicht nur mit der Betonung der Noten zu tun, sondern ich habe dabei eine Vorstellung, was der Komponist wollte, was das Publikum von einer Stilrichtung erwartet und wie ich in manchen Passagen überraschen kann, damit es nicht zu langweilig ist. Das ist eine Art von Kommunikation. Im Datenberg suchen wir nach diesen größeren Mustern der Kommunikation von Ausdruck in der Musik. Bei dieser quantitativen Analyse unterstützen uns KI-Methoden, weil sie schneller sind.

Sie machen Grundlagenforschung, sind daraus dennoch Anwendungen für den Alltag entstanden?

Wir haben Kooperationen mit der Medienindustrie im Bereich Musikwahrnehmung. So waren

wir die Ersten, die mit der digitalen Stereoanlage von Bang & Olufsen ein Entscheidungssystem entwickelt haben, welches Stück als Nächstes gespielt werden soll. Ähnlich ist der automatische Music Recommender im FM4-Soundpark, der Hörerinnen neue, musikalisch ähnliche Songs vorschlägt. In der Schweiz unterstützen wir die Tantiemen-Verrechnung, indem unsere Systeme erkennen, wann im 24-Stunden-Radiobetrieb Musik gespielt wird; da geht es um Machine Listening, also Computerprogramme, die darauf trainiert wurden, zu hören. Um nochmals auf das vierhändige Klavierstück zurückzukommen: Das war ein Demo-Beispiel, aber wir haben nicht die Absicht, ein System für den Markt zu bauen, damit Menschen Klavier üben können – denn unser Forschungsinteresse sind Grundsatzfragen.

Wohin geht in Ihrem Forschungsbereich die Entwicklung von künstlicher Intelligenz?

Mit dem aktuellen Hype um maschinelles Lernen forschen natürlich auch die großen Player wie Amazon oder Google zu KI und Musik. Sie demonstrieren die Mächtigkeit ihrer Modelle und lassen Musik von Chopin oder Beethoven komponieren. Eine halbe Minute klingt das auch wie Chopin, aber das geht ewig gleich so weiter. Maschinen haben keine Vorstellung von der gesamten Form eines Kunstwerks. Sie spucken wie ein Anfänger, der ein Instrument übt, Note für Note aus. Diese Sequenzmodelle von Google halte ich für musikalisch nicht geeignet, denn Komponieren hat viel mehr Ebenen. Das ist auch der Fokus meines Forschungsprojekts für die nächsten fünf Jahre: Ich möchte einen neuen Zugang zum Modellieren von Musik zeigen und damit wieder mehr darüber lernen, was Musik wirklich ausmacht.

Ist bei der Forschung zu Musik und künstlicher Intelligenz entscheidend, wer die größten Computer hat?

Musik darf keine Frage der Rechenpower sein, sonst machen wir etwas falsch, behaupte ich einmal. Keine Uni kann mit den Rechenzentren von Google konkurrieren, aber das zeigt mir auch, dass hier etwas falsch läuft. Denn unser menschliches Gehirn ist so langsam im Vergleich zu einem Supercomputer und wir schaffen es dennoch,



Bösendorfer CEUS. Der Computerflügel verfügt über ein Mess- und Wiedergabesystem, mit dem wichtige Daten für Widmers Forschungen erfasst werden können.

Musik zu komponieren und zu spielen – es muss also ohne diese Rechenpower gehen.

Werden wir einmal in einem Klavierkonzert sitzen, wo Computer für uns spielen und Stücke neu interpretieren?

Wir haben nicht das Ziel, Computer-Interpreten zu bauen. Computer können immer etwas Neues erzeugen und spielen auch oft etwas Unerwartetes. Aber sie machen dabei Fehler, die ein Mensch nicht machen würde. Das ist das Problem beim Machine Learning: Der Computer muss allgemeine Prinzipien lernen, die er dann bei anderen Musikstücken anwenden kann. Bei Texten ist es dasselbe: Etwas ist neu, stilistisch anders, aber wer bewertet, ob das kreativ wertvoll ist? Menschen können in ihrem Schaffensprozess entscheiden, was sie für kreativ halten und was sie wegwerfen, weil es absurd oder langweilig ist. Computer haben dieses Verständnis nicht, hier kommen wir an die Grenzen der Modellierbarkeit – und das ist auch gut so. Das Interessante entsteht im Konflikt zu etwas Bestehendem, man kann daher nicht etwas Neues schaffen, ohne die Geschichte zu kennen. Das bedeutet, ich kann nicht aus dem Nichts gute Kunst machen. <<



Video-Clip zum vierhändigen Konzert anlässlich der Preisverleihung „Falling Walls Science Breakthrough of the Year 2021“.