

„Musik braucht eine bestimmte Menge Redundanz“

Der Vorarlberger Gerhard Widmer hätte eigentlich Pianist werden können, er ist aber Informatiker mit Schwerpunkt künstliche Intelligenz geworden. Der Musik ist er treu geblieben und entwickelt Maschinen, die Melodien und das variantenreiche Spiel von Interpreten erkennen können.

INTERVIEW: Peter Illetschko

STANDARD: Sie haben selbst als Kind Klavier gespielt – auch öffentlich. Sind sie noch immer aktiv?

Widmer: Schon, aber nur privat. Für eine Karriere als klassischer Pianist war ich zu faul. Ich habe maximal 30 Minuten am Tag geübt. Mit 13 hat das noch zum ersten Preis beim Landeswettbewerb gereicht. Mit 15 habe ich die klassische Musik aufgegeben und mich dem Jazz zugewandt.

STANDARD: Aber Ihre Spezialisierung auf Artificial Intelligence und Musik hat dadurch eine gewisse Logik?

Widmer: Ja, sicher, obwohl ich es so gezielt nicht angepeilt habe. Das hat sich über die Jahre ergeben. Mein wissenschaftliches Feld ist in erster Linie die künstliche Intelligenz, das maschinelle Lernen. Ich habe in den frühen Jahren, als ich die ersten Lernalgorithmen entwickelt habe, „Testprobleme“ gesucht, um zu demonstrieren, was die Algorithmen können. So habe ich begonnen mich mit Musik zu beschäftigen und wie sie eine Maschine hören kann. Der Start-Preis des Wissenschaftsfonds FWF 1998 hat mir dann quasi den offiziellen Auftrag gegeben, mich damit zu beschäftigen, die Bestätigung, dass es sich dabei um ein seriöses wissenschaftliches Thema handelt. Zuvor hat man, wenn ich von meiner Arbeit erzählt habe, immer ein wenig die Stirn gerunzelt. Heutzutage schaut niemand mehr erstaunt. Musik ist digital, das Internet ist voll davon, Suchmaschinen, die darauf spezialisiert sind, werden gebaut,

STANDARD: Was kann künstliche Intelligenz mit Musik machen?

Widmer: Sie kann Computer entwickeln, die bestimmte Aspekte von Musik wahrnehmen können. Wir bauen zum Beispiel Maschinen, die Rhythmus und Tempo von Musikstücken erkennen können, die live Stücke mitverfolgen können und dabei parallel die exakten Noten anzeigen. Davon gibt es durch die Kooperation mit dem Concertgebouw Amsterdam bereits eine Anwendung: Die geben mehrmals jährlich ein E-Magazin heraus. Abonnenten können sich Konzerte anhören. Der Clou: Der Algorithmus hört sich das Konzert ebenfalls an und zeigt dabei die Noten.

STANDARD: Inwieweit können Ihre Maschinen die Musik auch selbst auswählen?

Widmer: Wir haben ein solches Musiksystem für Bang und Olufsen entwickelt. Da ist eine Software drinnen, die entscheidet, welche Musik sie spielt.

STANDARD: Wie kann ich sicher sein, dass diese Software bei einer Cocktailparty nicht Mozarts Requiem spielt?

Widmer: Sie wählen ein Album aus Ihrer Sammlung aus, dann wählt der Algorithmus nur mehr ähnliche Musik aus. Er muss sich also jedes Stück „anhören“ und auf Ähnlichkeiten prüfen. Natürlich kann das System Fehler machen – wie jede Maschine. In der Regel sollte es aber nicht das Requiem spielen, wenn Sie nicht zuvor eine ähnliche Musik gewählt haben. Ein vergleichbares, etwas abgesehenes System haben wir für den Soundpark von FM4 gebaut. Hier gibt es ja eine große Sammlung von unbekannt Bands, die darauf hoffen, entdeckt zu werden. Der Soundpark schlägt – ebenfalls



Der Informatiker Gerhard Widmer glaubt nicht daran, dass Maschinen genauso lernen können wie Menschen. Ihnen fehlt die sinnliche Freude an der Schönheit von Musik.

nach einer Erstwahl durch den Benutzer – mehrere Songs vor. Das führte zu einem verstärkten Download von Musik, und damit hat man das Ziel auch erreicht: die Bands bekannter zu machen.

STANDARD: Viele Smartphone-Besitzer nutzen Shazam, das gerade gespielte Musik erkennt. Das gelingt natürlich nur, sofern sie in einer Referenzdatenbank liegt. Könnte man ein solches System noch smarter machen?

Widmer: Natürlich, man könnte auch unterschiedliche Interpretationen eines Liedes erkennen. Das kann Shazam nicht. Mit klassischer Musik schaffen wir das mittlerweile. Unsere Rechner können in Sekundenschnelle Musikstücke aus Liveaufnahmen identifizieren, unabhängig davon, wie oder von wem sie gespielt oder interpretiert werden. In unserem neuesten Projekt, das vom Europäischen Forschungsrat ERC gefördert wird, schauen wir uns an, welchen Ausdrucksdimension ein Musikstück über eine Interpretation bekommt: also spielerisch oder zögerlich oder schwermütig oder verhalten. Wie kann man das als Mensch erkennen? Und kann das eine Maschine auch – ohne dass sie so etwas wie ein Verständnis für Emotionen entwickeln kann?

STANDARD: Wie kann man sich das vorstellen?

Widmer: Interpretation und Ausdruck haben ganz zentral mit der Struktur eines Musikstücks zu tun und wie diese Struktur von InterpretInnen kommuniziert wird. Wenn ich Ihnen ein Musikstück vorspiele, erkennen Sie automatisch, wo eine Einheit aufhört und eine andere beginnt, Sie hören wiederkehrende Melodien, ein Motiv, das in Variationen wiederholt wird. Musik braucht eine bestimmte Menge an Redundanz, Melodien, wo man sich festhalten kann. Nicht zu viel, denn auch das wäre langweilig. Das hat mit unseren Wahrnehmungen und mit unserem Sinn für Ästhetik zu tun. Wenn in der Welt alles gleich wäre, würden wir nicht existieren können. Wenn alles egal ist, keine Struktur zu sehen ist, kann ich nichts wiedererkennen, mich nicht zurechtfinden – in der Musik und überhaupt.

STANDARD: Was ist Ihre Vermutung? Kann sie?

Widmer: Das werden wir sehen. Im Prinzip könnte man alles auf das Lernen reduzieren. Auch wir haben ja gelernt, wahrzunehmen und zu strukturieren – und erkennen deshalb eine Melodie oder einen Rhythmus. Je mehr Daten ich habe, desto mehr weiß ich über die Welt: Das ist zumindest die Google-Sicht der Welt. Wir können aber bis heute nicht genau sagen, wie unser Gehirn lernt, welche Rolle das In-der-Welt-Sein spielt, körperliche Erfahrung, Motivation, Emotion. Oder sinnliche Freude an der Schönheit von Musik. Davon sind lernende Maschinen noch weit entfernt.

GERHARD WIDMER, geboren 1961 in Bildstein in Vorarlberg, leitet das Department of Computational Perception an der Johannes-Kepler-Universität Linz und die Forschungsgruppe Intelligent Music Processing and Machine Learning Group am Austrian Research Institute for Artificial Intelligence (OFAI). 2009 erhielt er den Wittgenstein-Preis des FWF, kürzlich wurde ihm der Advanced Grant des ERC zugesprochen.