

Musikalisch intelligente Computer: Anwendungen in der klassischen und populären Musik

Gerhard Widmer

Institut für Computational Perception
Johannes Kepler Universität Linz
und
Österreichisches Forschungsinstitut für Artificial Intelligence, Wien

Grundlagenorientierte Forschung in den Bereichen Künstliche Intelligenz (KI) und Musik hat in Österreich eine beachtliche Tradition. Schon seit ca. 1990 wird am Österreichischen Forschungsinstitut für Artificial Intelligence (ÖFAI) in Wien an Methoden gearbeitet, die dem Computer so etwas wie Musikverständnis verleihen sollen und auch die Fähigkeit, seine Fähigkeiten und Kenntnisse durch Lernen selbst zu erweitern, selbst Musik zu lernen. Mit Oktober 2004 nahm auch das neue Institut für Computational Perception („Computerwahrnehmung“) an der Universität Linz seinen Betrieb auf, das in enger Kooperation mit dem ÖFAI diese Art von Forschung weiter ausbaut; verbindender Faktor ist der Autor dieses Artikels, der in beiden Instituten eine führende Rolle innehat.

In den letzten Jahren sind aus der Grundlagenforschung auch vermehrt anwendungsorientierte Projekte entstanden, und zwar sowohl im Kontext von klassischer Musik und Musikwissenschaft, als auch im neuen Anwendungsgebiet des Music Information Retrieval. Der vorliegende Beitrag gibt einen kurzen Überblick über aktuelle Forschungs- und Anwendungsprojekte an den beiden oben genannten Instituten, die zeigen oder zumindest andeuten soll, dass Computer derzeit wirklich schon intelligente musikalische Leistungen erbringen und durchaus auch nützliche Dienste leisten können.

Intelligente Computer für die Musikwissenschaft

Im Rahmen eines großangelegten und langjährigen Forschungsprojekts¹ entwickeln wir intelligente Computerprogramme, die neue Einblicke in und ein tieferes Verständnis für ein sehr komplexes musikalisch-künstlerisches Verhalten ermöglichen sollen, nämlich *ausdrucksvolle Musikinterpretation*. Ausdrucksvolle Interpretation – das willentliche und äußerst subtile Variieren von Parametern wie Tempo, Timing, Dynamik (Lautstärke), Artikulation etc. durch den ausführenden Musiker – ist ein wesentliches Element vor allem der klassischen Musik, das für die Wirkung der Musik von zentraler Bedeutung ist, gleichzeitig aber in seinen Gesetzmäßigkeiten und seiner Wirkungsweise noch kaum verstanden und deshalb für die moderne, kognitiv orientierte Musikwissenschaft von großem Interesse ist.

Zu den Fragen, die wir mit Computermethoden untersuchen, zählen etwa: Gibt es grundlegende Prinzipien, ungeschriebene Regeln der Musikinterpretation, an die sich alle Musiker mehr oder weniger (und mehr oder weniger bewusst) halten? Kann man diese Prinzipien formal quantifizieren, und in welchem Zusammenhang stehen sie mit unseren Hörgewohnheiten? Aber auch: kann man die subtilen Unterschiede im Interpretationsstil berühmter Künstler und Virtuosen quantitativ fassen und formal beschreiben?

¹ Fonds zur Förderung der Wissenschaftlichen Forschung (FWF), Projekt Y99-INF, 1998-2005 (START-Preis 1998).

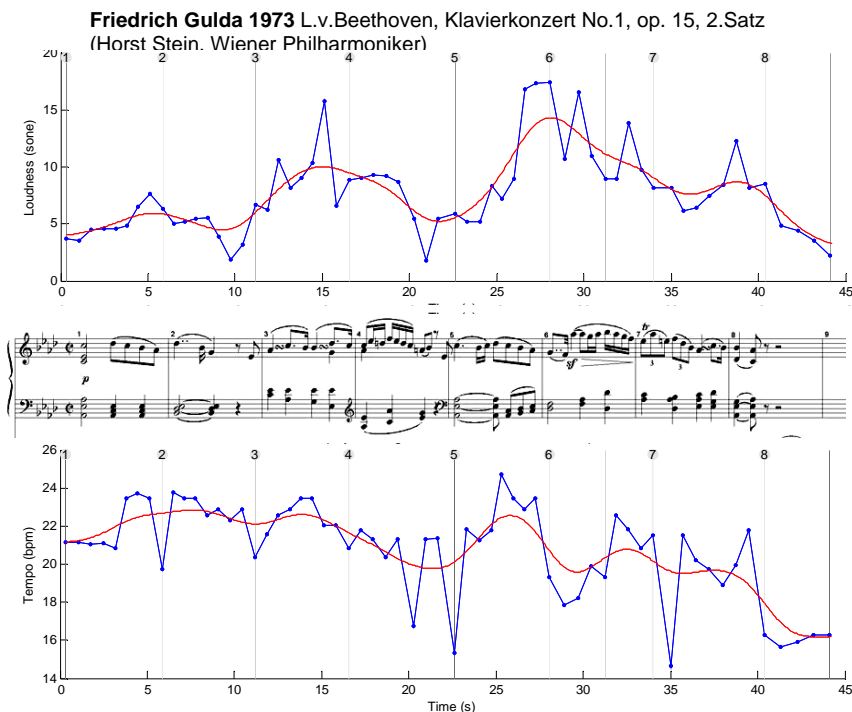


Abb.1: Friedrich Gulda spielt die ersten 8 Takte des zweiten Satzes des Klavierkonzerts No.1, op.15, von Ludwig van Beethoven. Oben: Lautstärkekurve; unten: Tempokurve. Je höher die Kurve an einem bestimmten Punkt, desto lauter (oben) bzw. schneller (unten) spielte der Pianist an dem Punkt im Stück.

Dazu müssen zunächst einmal die ausdrucksvollen Nuancen (z.B. Tempo- und Lautstärkeverläufe) in Aufnahmen von (klassischen) Musikern genauestens vermessen werden (siehe Abb. 1). Die gewonnenen Messdaten werden dann mit Methoden der KI – Maschinelles Lernen, Data Mining, Mustererkennung – vom Computer analysiert. Die Maschine soll als autonomer Entdecker agieren, der selbständig interessante Muster und Gesetzmäßigkeiten in diesen riesigen Datenmengen findet. Es handelt sich hier um aufwändige, mehrjährige Forschungen, deren Ergebnisse an dieser Stelle nur kurz skizziert werden können.²

Hier können nur einige beispielhafte Resultate kurz herausgegriffen werden. Neben Fragen der maschinellen Vermessung von Musikaufnahmen bzgl. Tempo und Dynamik (die spezielle KI-Forschungen bedingten und z.B. in einem Programms resultierten, das den Takt ("Beat") in Musikaufnahmen finden kann) befassen wir uns z.B. mit der Entwicklung neuartiger Methoden für die Visualisierung expressiver Musikinterpretation. Abb.2 zeigt einen Schnappschuss des sogenannten "Performance Worms",³ eines Visualisierungstools, das die großflächigen Tempo- und Lautstärkevariationen, die ein Musiker in einer Aufnahme eingesetzt hat, als

² Eine ausführlichere, relativ allgemeinverständliche Präsentation findet sich in: G. Widmer, S. Dixon, W. Goebel, E. Pampalk, A. Tobudic (2003). In Search of the Horowitz Factor. *AI Magazine* 24(3), 111-130 (via e-mail vom Autor erhältlich).

³ Langner, J. and Goebel, W. (2003). Visualizing Expressive Performance in Tempo-Loudness Space. *Computer Music Journal* 27(4), 69-83.

Trajektorie über die Zeit sichtbar macht. Mittels Methoden der Mustererkennung versuchen wir dann, Stilmuster in diesen Trajektorien zu finden, die für bestimmte berühmte Künstler charakteristisch sind (siehe Abb.3). Wir entwickeln KI-Programme, die lernen, berühmte Interpreten an ihrem Stil zu erkennen, und die in jüngsten Experimenten schon erstaunlich hohe Erkennungsraten von bis zu 90% und darüber (auf neuen, unbekannten Aufnahmen) erreichen.⁴ Und wir arbeiten an Lernmethoden, mittels derer der Computer – durch Analyse von Aufnahmen wirklicher Musiker – selbst sogar zu *lernen* versuchen, Musik "ausdrucksvoll" zu interpretieren und zu spielen.⁵ Der Output eines dieser Programme – eine vom Computer erzeugte und gespielte "ausdrucksvolle" Interpretation einer Mozart-Klaviersonate – gewann sogar den 2. Preis in einem internationalen "Computer-Musikinterpretationswettbewerb" (RENCON 2002, Tokio).⁶

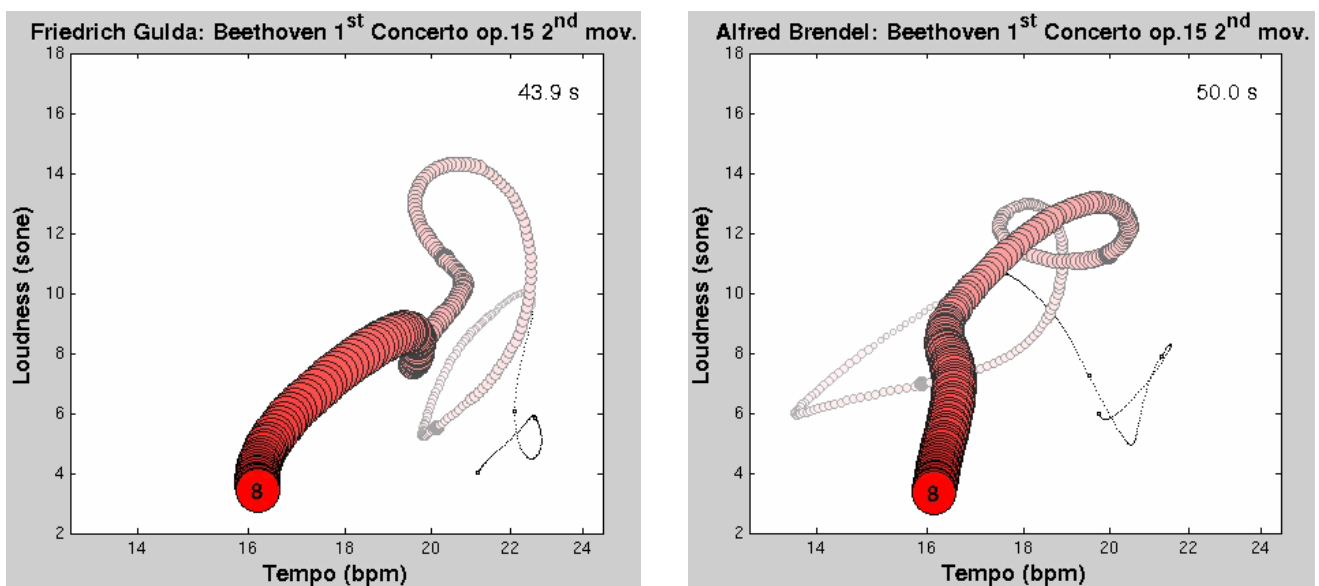


Abb. 2: Der "Performance Worm": Erste 8 Takte des 2. Satzes von Beethovens Klavierkonzert no.1, op15, gespielt von Friedrich Gulda (links) und Alfred Brendel (rechts). Waagrechte Achse: Tempo in *beats per minute* (bpm); senkrechte Achse: Lautstärke in *Sone*. Eine Bewegung des Wurms nach rechts oben beschreibt z.B. ein gleichzeitiges Beschleunigen (*accelerando*) und Lauterwerden (*crescendo*) usw. usw. Der dunkelste Punkt repräsentiert den gegenwärtigen Zeitpunkt, die Vergangenheit erscheint blasser.

⁴ C. Saunders, D. Hardoon, J. Shawe-Taylor, G. Widmer (2004). Using String Kernels to Identify Famous Performers from their Playing Style. In *Proceedings of the 15th European Conference on Machine Learning (ECML'2004)*, Pisa, Italy.

⁵ Widmer, G. and Tobudic, A. (2003). Playing Mozart by Analogy: Learning Multi-level Timing and Dynamics Strategies. *Journal of New Music Research* 32(3), 259-268

⁶ Automatische Musikinterpretation mittels Computer ist absolut kein Ziel unserer Forschung. Aber dieses Ergebnis zeigt zumindest, dass die Maschine tatsächlich in der Lage ist, sinnvolle Muster und Regelmäßigkeiten in Musikaufnahmen zu finden.

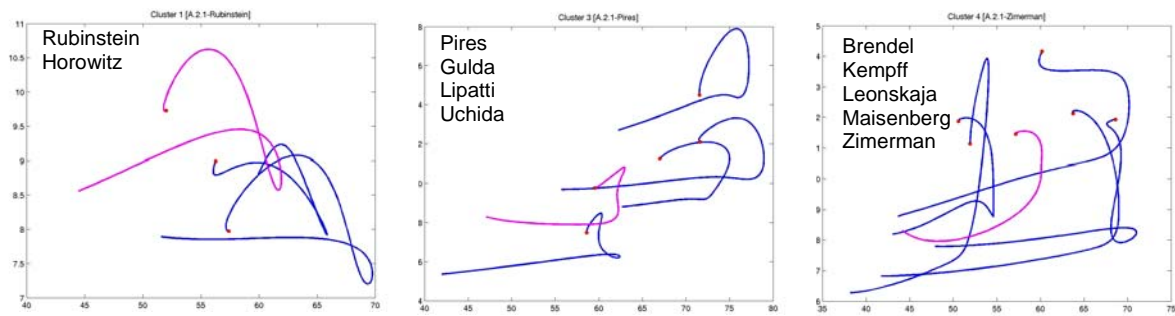


Abb. 3: Drei klar unterscheidbare Strategien, eine bestimmte Klasse von Phrasen in einem Stück von Franz Schubert (Impromptu D.899, Ges-Dur) durch Tempo-Lautstärke-Veränderung zu formen (vom Computer in Aufnahmen verschiedener berühmter Pianisten entdeckt).

„Interfaces to Music“

Aus den oben beschriebenen Forschungen entwickelten sich auch Ideen zu möglichen praktischen Anwendungen. Ein vor kurzem begonnenes Projekt mit dem Titel „Interfaces to Music“⁷ hat sich zum Ziel gesetzt, neue Arten von Computerschnittstellen zu entwickeln, die es möglich machen, Musik auf ganz neue Weise (z.B. durch Visualisierung von Musikinterpretation mittels Computeranimation) zu vermitteln, zu lehren, zu erleben und auch selbst kreativ zu formen - kurz, neue Formen der Interaktion mit Musik zu ermöglichen. Die angepeilten Anwendungsszenarien reichen von didaktischen Anwendungen (etwa in Konservatorien, wo der Computer mittels detailliertem akustisch-grafischem Feedback eine präzisere Vermittlung von Spiel- und Interpretationstechniken ermöglichen könnte) über Anwendungen im sogenannten "Edutainment" (z.B. multimediale Animationen zum ergänzenden Transparentmachen komplexer Musikaufführungen in Konzertsälen oder intelligente interaktive Musikinstallationen in Musik "museen" wie etwa dem Haus der Musik in Wien) bis hin zum Home Entertainment, wo der Computer dem Musikkonsumenten neue Möglichkeiten zum kreativen Umgang mit Musikaufnahmen - über das bloß passive Zuhören hinaus - bieten könnte.

Ein erstes, im wahrsten Sinn des Wortes „greifbares“ Resultat dieser Forschungen ist ein neuartiges Interface, das dem Benutzer ermöglicht, interaktiv mit Interpretationen klassischer Musikstücke zu „spielen“, indem er/sie einfach durch Handbewegungen in der Luft die Geschwindigkeit und die Dynamik (Lautstärke) einer Musikaufnahme (z.B. eines Chopin-Preludes) steuert und damit in Echtzeit eine Musikinterpretation nach seinem/ihrem Geschmack entwirft. Dazu sind keinerlei störende Sensoren oder Verkabelungen nötig: ein einfaches, bereits im Jahr 1917 (!) von einem russischen Ingenieur erfundenes und im wesentlichen nur aus 2 Antennen bestehendes elektrisches Gerät – ein sogenanntes *Theremin* –, das eigentlich als Musikinstrument zur Klangerzeugung gedacht ist, dient uns als Messgerät, das die Handbewegungen registriert und an den Computer weiterleitet (siehe Abb. 4). Das System ist mit dem Performance Worm (siehe obige Abb. 2) gekoppelt, sodass der Benutzer nicht nur akustisches, sondern auch visuelles Feedback erhält. Verschiedene Szenarien der möglichen Nutzung dieser Technologie werden derzeit angedacht.

⁷ Wiener Wissenschafts-, Forschungs- und Technologiefonds (WWTF), Projekt CI010, 2004-2007.



Abb. 4: Der Autor bei der Arbeit: Interaktives Steuern von Tempo und Dynamik einer Musikaufnahme mittels eines MIDI-Theremins.

Automatische Strukturierung und Visualisierung digitaler Musiksammlungen

Eine alternative Forschungsrichtung mit sehr direkter Nutzungsperspektive ergibt sich aus dem rapiden Wandel in Richtung digitaler Musik und web-basierter Musikservices, der sich derzeit auf dem Musikmarkt vollzieht, und dem damit einhergehenden Forschungsboom im Bereich *Music Information Retrieval (MIR)*.⁸ Auch in diesem Bereich sind unsere Institute prominent vertreten. In einem großen, ca. 2 Mio. Euro "schweren" multinationalen Forschungsprojekts im Rahmen des 6th Framework der EU⁹ arbeiten wir z.B. an Methoden zum automatisierten Strukturieren und Visualisieren großer digitaler Musiksammlungen, Tools zum Navigieren in solchen Musiksammlungen, Algorithmen für das automatische Erkennen von Musikstilen, oder auch "Personal DJs", die dem individuellen Geschmack eines Benutzers angepasste Playlists generieren.

⁸ Siehe etwa Computer Music Journal 28(2), Special Issue on Music Information Retrieval, MIT Press, 2004

⁹ "Semantic Interaction with Music Audio Contents (SIMAC)", EU-Projekt FP6-507142, Key Action IST-2.3.1.7 Semantic-based Knowledge Systems; Projekt-Homepage: www.semanticaudio.org

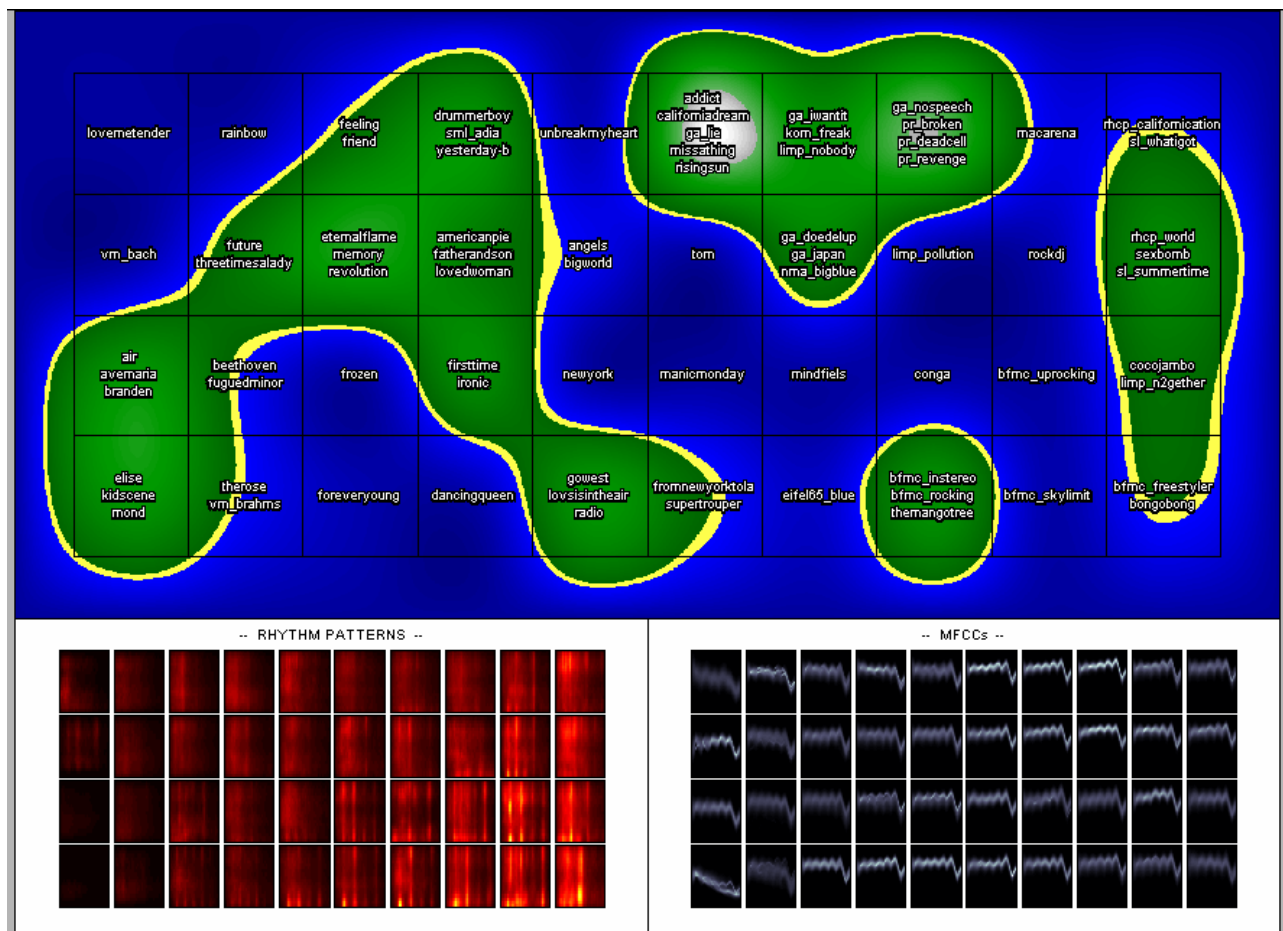


Abb. 5: Islands of Music v.2

Abb.5 zeigt als Beispiel einen Screenshot der *Islands of Music v.2*, eines Programms zum interaktiven Explorieren der Struktur von Musiksammlungen nach verschiedenen musikalischen Kriterien.¹⁰ Der obere Teil des Bildschirms zeigt die Projektion einer Sammlung von MP3-Files nach musikalischer / klanglicher Ähnlichkeit; Inseln repräsentieren zusammenhängende Bereiche mit ähnlicher Musik. In der unteren Bildschirmhälfte werden verschiedene musikalisch relevante Eigenschaften der Musik in verschiedenen Regionen angezeigt. Mit einem Slider kann interaktiv und stufenlos die Gewichtung verschiedener musikalischer Faktoren (z.B. Rhythmus, Klangfarbe, etc.) modifiziert werden. Die daraus resultierenden Änderungen in der Struktur der Musiksammlung werden dann unmittelbar sichtbar.

Jenseits von Audio und Klang: Music Web Mining

Die Islands of Music zeigen, dass ein Computer durchaus in der Lage ist, den Klang von Musikaufnahmen zu „verstehen“ und klanglich-musikalische Ähnlichkeiten zwischen Musikstücken für intelligente

¹⁰ Siehe z.B. E. Pampalk, S. Dixon, G. Widmer (2004). Exploring Music Collections by Browsing Different Views. *Computer Music Journal* 28(2), 49-62.

Anwendungen zu nutzen. Allerdings stellt sich derzeit in der Forschung immer klarer heraus, dass es hier auch strikte Grenzen gibt: nicht alles, was Musik ausmacht und was Menschen dazu veranlasst, Musikstücke oder Künstler als ähnlich oder verwandt zu empfinden, lässt sich aus der Tonaufnahme selbst erklären. Im Gegenteil, Musik ist eingebettet in ein komplexes Netz aus kulturellen, historischen und sozialen Kontexten; persönliche Ansichten und soziale Bindungen spielen eine ebenso große Rolle beim Interpretieren von Musik wie der tatsächliche Klang der Musik selbst. Eine ganz neue Forschungsrichtung und ein entsprechendes Forschungsprojekt ¹¹ versucht daher folgerichtig, solche sozialen, außermusikalischen Faktoren ebenfalls für den Computer greifbar zu machen, um noch wesentlich bessere, der menschlichen Einschätzung nähere Computermodelle von Musikähnlichkeit und Musikklassifikation entwickeln zu können. Und die beste Datenquelle dafür ist das *Internet*, in dem Musikliebhaber, -kritiker, -produzenten und -anbieter auf Abermillionen von Webseiten über Musik schreiben und diskutieren. Music Web Mining ist der Versuch, mit intelligenten Computermethoden automatisch relevante Metainformation über Musikstücke und Künstler aus dem Internet zu gewinnen.

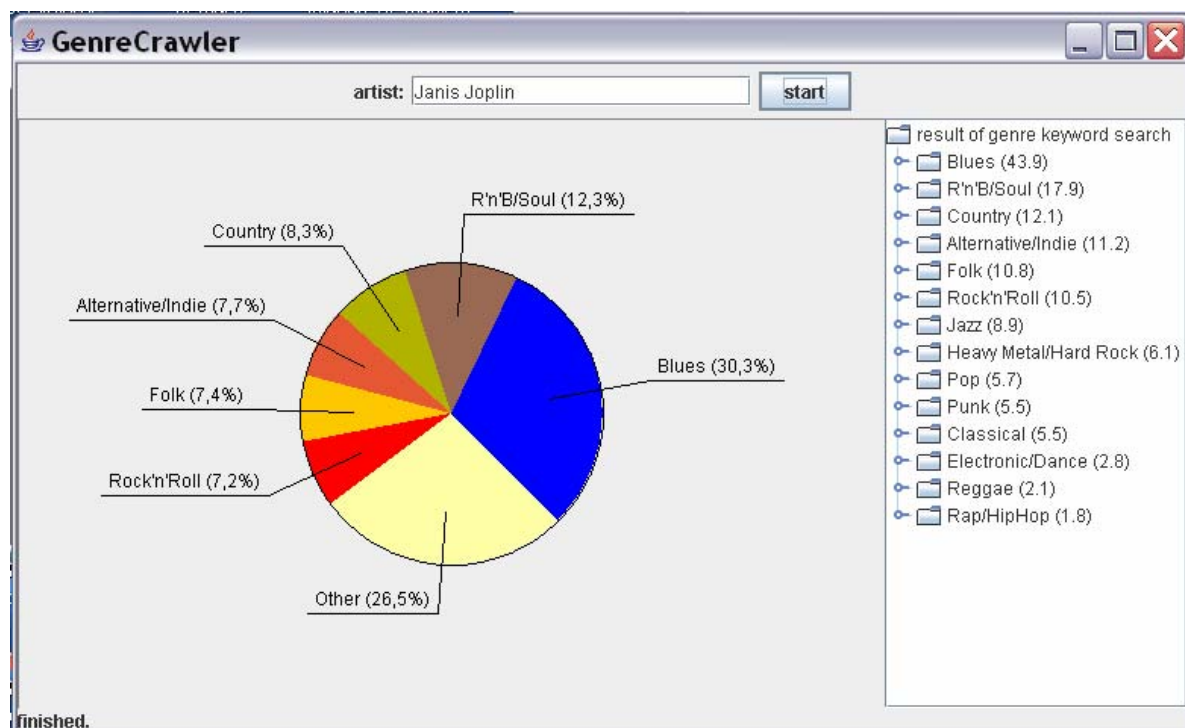


Abb. 6: Der "Genre Crawler" versucht, „Janis Joplin“ einzuordnen.

Abb. 6 zeigt ein kleines Beispiel dazu aus unserer Forschung. Der „Web Crawler“ ist ein Demo-Applet, das zeigt, wie ein Computer lernen kann, Musikkünstler vollautomatisch möglichen Genres oder Stilrichtungen zuzuordnen, und zwar ausschließlich auf der Basis von Informationen, die der Computer sich selbst aus dem Internet – von beliebigen Webseiten – holt. Dahinter stehen Methoden des maschinellen Lernens, die anhand einiger vorgegebener Beispielmusikanten relativ einfache linguistische Profile aufbauen (d.h. Listen von Worten, die auf Webseiten mit den Künstlernamen auftreten und auf eine bestimmte Stilrichtung hinzudeuten scheinen) und anhand dieser gelernten Profile dann beliebige neue Namen einem oder mehreren Genres zuordnen können. Der Zweck des ganzen liegt in der Entwicklung immer intelligenterer Ähnlichkeitsmetriken, mit denen dann eine Vielzahl von MIR-Anwendungen möglich werden.

¹¹ Fonds zur Förderung der Wissenschaftlichen Forschung (FWF), Projekt L112, 2005-2008.

Ausblick

Musik und Künstliche Intelligenz – was ursprünglich als reine Grundlagenforschung begann, motiviert durch pure Neugierde, entwickelt sich jetzt rapide zu einem Forschungsbereich mit vielfältigsten praktischen Anwendungsperspektiven. Das gilt einerseits natürlich ganz offensichtlich für den Bereich Music Information Retrieval. Aber auch die oben genannten Forschungen zur Musikinterpretation eröffnen ganz neue Perspektiven, vor allem im Kultur-, Bildungs- und "Edutainment"-Bereich, wo Visualisierungen, Animationen und andere computerisierte Formen der Wissensvermittlung dazu eingesetzt werden könnten, komplexe musikalische Sachverhalte und Strukturen transparent zu machen.

Und schließlich wird es, auf Basis intelligenter Computermethoden, in naher Zukunft eine ganz Generation neuartiger Musikgeräte geben, die völlig neue Formen der Interaktion mit Musik für den Konsumenten möglich machen werden – weg vom passiven Hören hin zu direkter Interaktion mit und Steuerung von Musik und Audio. Hier entsteht ein großer Markt, der in der unmittelbaren Zukunft intensive innovative Forschung – auch Grundlagenforschung – im Bereich KI & Musik erfordern wird.

Danksagung

Unsere Forschungen im Bereich KI und Musik werden derzeit von folgenden Organisationen finanziell unterstützt: der *Europäischen Union* (Projekte FP6 507142 SIMAC – „Semantic Interaction with Music Audio Contents“; FP6 IST-2004-03773 S2S² – „Sound to Sense, Sense to Sound“; und COST 282 KnowEST – „Knowledge Exploration in Science and Technology“); dem Österreichischen *Fonds zur Förderung der Wissenschaftlichen Forschung* (FWF; Projekte Y99-START („AI Models of Expressive Music Performance“) und L112 („Operational Models of Music Similarity for MIR“); und dem *Wiener Wissenschafts-, Forschungs- und Technologiefonds* (WWTF; Projekt CI010 „Interfaces to Music“). Das Österreichische Forschungsinstitut für Artificial Intelligence wird außerdem von den Bundesministerien für Bildung, Wissenschaft und Kultur (BMBWK) und für Verkehr, Innovation und Technologie (BMVIT) unterstützt.

Kontakt:

Univ.-Prof. Dr. Gerhard Widmer
Institut für Computational Perception
Johannes Kepler Universität Linz
Altenberger Straße 69
A-4040 Linz
Österreich
und
Österreichisches Forschungsinstitut für Artificial Intelligence
Freyung 6/6
A-1010 Wien
Österreich

Mail: gerhard.widmer@jku.at
WWW: <http://www.cp.jku.at/people/widmer>