

iIgMV - interactive Interface for generative Music and Visualizations

Damian T. Dziwis, Felix Hofschulte, Michael Kutz, Martin Kutz

Fachbereich Medien und Design
FH Düsseldorf

Abstract

Das iIgMV ist ein modulares System zur Generierung von Musik und Visualisierungen durch Bewegungen und Gesten in Echtzeit. Es ermöglicht gleichermaßen das generieren von visuellen Darstellungen, Tönen in musikalischem Kontext und interagieren mit Musikern durch Bewegungen. Das System setzt sich aus 3 Komponenten zusammen: generative Visualisierungs- und Klangerzeugung sowie algorithmische Musikkomposition. Bindeglied zwischen den Komponenten bildet hierbei die Bewegungs- und Gestenerkennung von Microsofts Kinect. Aufgrund seiner Skalierbarkeit erstrecken sich die möglichen Einsatzgebiete von kleineren Installationen für Laien bis hin zu professionellen, konzipierten und komplexen Liveperformances und Aufführungen.

1 Motivation

Mit seinem Aetherophone (später dann Theremin) weckte Lew Termen im Jahre 1920 auch das Interesse an nicht-taktilen Interfaces für Instrumente.[B1] Diese erwecken oft Neugier beim Anwender und Zuhörer und erlauben einen spielerisch-experimentellen Zugang zum Instrument durch Bewegung. Die damit verbundene Faszination beim Spiel führte zu einer immer wiederkehrenden Motivation, diese Art der Interfaces zu entwickeln wenn neue Trackingtechnologien wie z.B. Kamera-Tracking, Infrarotsensoren oder Spielecontroller (Wii, Kinect, Sony Move) zur Verfügung standen. Bereits 1997 stellten J. Paradiso und F. Sparacino 3 fortschrittliche Interfaces auf Basis verschiedener Tracking Technologien vor die sich bereits gut in musikalischen Anwendungen und Aufführungen einsetzen ließen.[P1] In den meisten Fällen erlauben derartige Interfaces dem Anwender zwar einfach, schnell und spielerisch damit Töne zu produzieren, doch ist ein genaues tonales Reproduzieren von Tönen und Spielweisen anhand notierter Kompositionen umso schwieriger. Dabei mangelt es solchen Instrumenten auch oft an der notwendigen Komplexität, z. B. ausreichenden Parametern für Intonationen um genügend Spielraum für interessante, unikat Spielweisen zu ermöglichen. Das Erlernen einer virtuellen Beherrschung eines berührungslosen Instruments

gestaltet sich oftmals als sehr schwierig bis schlichtweg unmöglich, was zu einer bisher geringen Akzeptanz und seltenem Einsatz in Solo- oder Ensemblewerken geführt hat. Voran gegangene Arbeiten wie das Paper „Interactive Music: Human Motion Initiated Music Generation“ [P2] befassten sich mit der komplexe Bewegungs-, aber auch Gestenerkennung der Kinect zur Tonerzeugung und mit seinem Kinectar [I1] zeigte Chris Vik, dass sich solches Interface auch für komplexe Kompositionen mit Performances einsetzen lässt.[Y1] iIgMV erweitert die oft realisierte 1-zu-1-Übersetzung von Bewegungswerten zu Soundparametern um einen algorithmischen MIDI-Prozessor, der ergänzende Töne und Sequenzen im musikalischen Kontext berechnet. Zusätzlich bietet eine visuelle Komponente ein bildliches Feedback der resultierenden Komposition. Durch eine modulare Entwicklung bildet iIgMV ein skalierbares und komplexes Instrument, das Visualisierungen als auch Bewegung/Tanz instrumentalisiert und kompatibel zu mit Ensembles gespielter Musik macht und damit zum Teil der Partitur und der gesamten Komposition werden lässt.

2 Konzept/Umsetzung

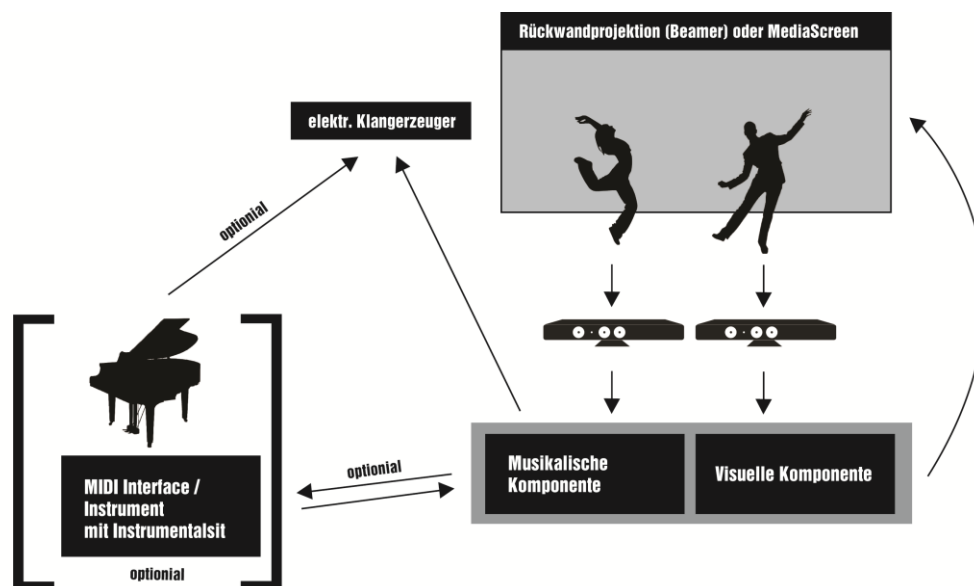


Abbildung 1: Fallbeispiel mit zwei Tänzern und Instrumentalist

Basis des Interfaces ist die Microsoft Kinect, ein Software-Setup aus SensorKinect-Modulen für die OpenNI/NITE Middleware und darauf aufsetzenden Applikationen wie OSCeolon, Synapse oder die AirKinect Extension liefern OSC-Daten über die Position von Körperknotenpunkten (z.B. linke Hand, rechtes Bein etc.) im Weltkoordinatensystem in

Relation zum Körper, der Bildschirmposition und daraus abgeleiteten Gesten wie “linkes Bein nach oben”, “rechte Hand an den Kopf” usw. Diese Daten dienen dazu in der musikalischen als auch der visuellen Komponente entsprechende musikalische und visuelle Ereignisse zu generieren und können, abgestimmt aufeinander, in einen künstlerischen Kontext gebracht werden → Musikalische und visuelle Ereignisse können über die Auslösung durch die selben Daten miteinander gekoppelt werden oder durch unterschiedliche Parametrisierung unabhängige Ereignisse und damit unterschiedliche Aussagen im künstlerischen Kontext erzeugen. Einer weiteren Synchronisierung, bis auf die Bewegungen selbst, bedarf es dabei nicht.

2.1 Umsetzung der visuellen Komponente

In der generativen als auch interaktiven visuellen Komponente werden die Daten der Kinect aus der AirKinect Extension in Adobe Flash für die entsprechende Visualisierung verarbeitet. Das visuelle Rendering erfolgt durch die Stage3D-API, bei Verwendung von 3D-Modellen mit der zusätzlichen 3D-Engine Away3D v.4. So können eine große Anzahl an 3D- und 2D-Objekten, Videos als auch komplexe Partikelsysteme mit höchster Auflösung effizient genutzt werden. Dadurch bieten sich etliche Möglichkeiten, eindrucksvolle Bildwelten zu generieren. Zudem steigt die künstlerische Freiheit ungemein und der Benutzer bekommt schnell ein eindrucksvolles Feedback seiner Bewegung auf eine Leinwand, einem Monitor und/oder einem anderem Ausgabegerät. Die Positionsdaten der Körperknotenpunkte im Weltkoordinatensystem lassen sich zum Beispiel zum Bewegen, Skalieren, Rotieren oder Morphen von Objekten nutzen, Gesten zum Auslösen visueller Ereignisse → in einem einfachen Beispiel (Abb. 2) wird in einem 3D-Raum eine direkt transformierte Übersetzung der Handbewegungen mit den jeweiligen Koordinaten im Raum (x, y, z) mit Hilfe von Würfeln visualisiert. Auch die Rotation des Körpers (Torso-Rotation) beeinflusst die Rotation der Würfel in der Bewegung. Unterschiedliche Gesten wie die Berührung der Hände mit dem Kopf oder der Kontakt beider Hände, nehmen direkten Einfluss auf die Kamerabewegung im virtuellen Raum, lassen die Visualisierung einfrieren, erzeugen wellenartige Bewegung von der Ursprungsordinate der Berührung oder ändern einfach nur die Umgebungsfarbe der Szene.



Abbildung 2: Visualisierungsbeispiel in Aktion

2.2 Umsetzung der musikalischen Komponente

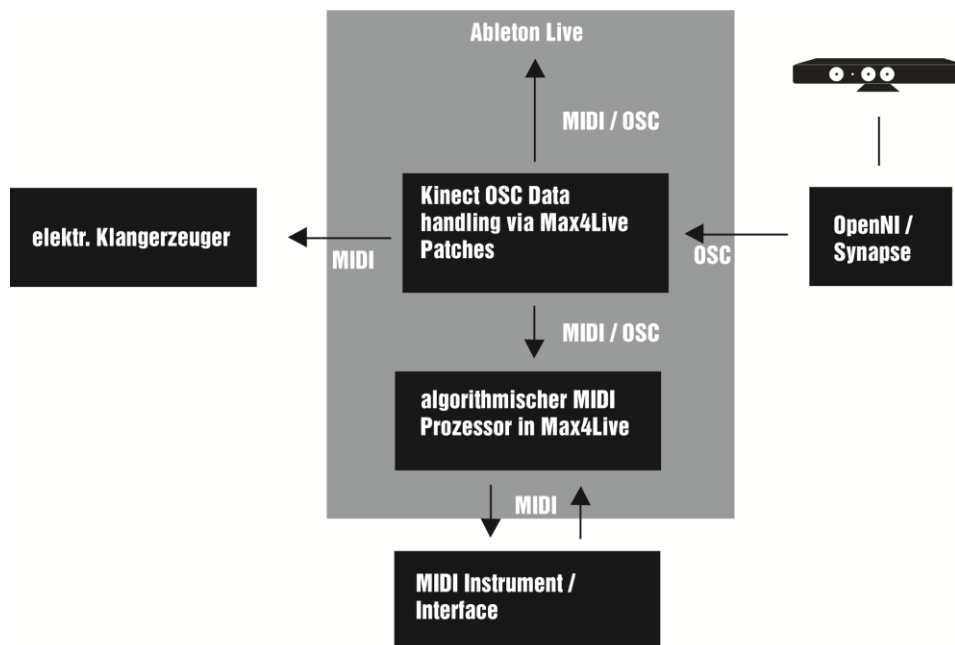


Abbildung 3: Aufbau und Routing der musikalischen Komponente

Host und Audio Workstation der musikalischen Komponente bildet die Software Live von Ableton. Hier findet sowohl die Auswertung der Kinect-Daten, das gesamte MIDI-Processing und Routing als auch die Audioerzeugung statt. Synapse (alternativ OSCeleton) liefert alle Kinect-Daten und Events als OSC über UDP an Max4Live (eine Schnittstelle von Cycling '74's Max/MSP für Live). Die damit programmierten Patches ermöglichen es, die Kinect-Daten nach Gesten oder auch Koordinaten einzelner Körperteile zu filtern, um anschließend (in MIDI übersetzt) zur Steuerung von MIDI-Klangerzeugern/Instrumenten, aber auch Effekten und sonstigen Elementen innerhalb Ableton Live genutzt zu werden. Diese MIDI-Daten können direkt als Töne, inklusive dessen Parametern, umgesetzt werden → zum Beispiel bestimmt die Höhe der einen Hand die Tonhöhe, die Entfernung zum Körper ihre Lautstärke, während man mit der anderen Hand das Vibrato kontrolliert und Ähnliches.

Diese MIDI-Daten können aber auch in komplexere musikalische Strukturen umgesetzt werden, indem man sie mit dem algorithmischen MIDI-Prozessor kombiniert, einem weiteren wesentlichen Teil der musikalischen Komponente. Auch dieser besteht aus in Max4Live programmierten Patches, die beliebig programmiert, eingestellt und miteinander

kombiniert werden können, um die verschiedensten Ergebnisse zu ermöglichen. Über einzelne MIDI-Noten können in Echtzeit Akkorde gebildet, Intervalle und Tonfolgen nach vorgegebenen Skalen ergänzt (dies sowohl nach Harmonieregeln als auch atonalen Vorgaben für z.B. zeitgenössische E-Musik), oder auch Zufallsereignisse für aleatorische Kompositionen erzeugt werden. Diese MIDI-Noten können direkt von zu MIDI übersetzten OSC-Daten der Kinect stammen, aber auch von externen MIDI-Instrumenten/Interfaces, wo der algorithmische MIDI-Prozessor seine wahre Stärke zeigt und das ilgMV auch ensemblefähig mit anderen Instrumentalisten macht.

In Verbindung mit einem MIDI-Instrument und dessen Interface ermöglicht der MIDI-Prozessor eine Mischform aus menschlich gespielten Tönen und in Echtzeit ergänzter algorithmischer Komposition. Somit lassen sich z.B. auf einem MIDI-Klavier komplexe Kompositionen realisieren, die über das menschlich spielbare hinausgehen → in Echtzeit können zu gespielten Noten passende Intervalle ergänzt werden, z.B. ein Echo das einen Ton Oktaviert oder beliebig anders Transponiert in einem Decrescendo in die tieferen Register hin wiederholt. Zwischen einem gehaltenen Akkord können Töne interpoliert werden, riesige Cluster-Akkorde simultan gespielt werden, u.v.m. Diese Patches können beliebig mit einander kombiniert werden um eine ganzheitliche ergänzte algorithmische Komposition zu ergeben. Wann und wie der Prozessor reagieren soll lässt sich dabei über Parameter einstellen → bei bestimmten Tönen oder Kombinationen/Akkorden, in bestimmten Oktaven, beim überschreiten einer bestimmten Anschlagsstärke/Velocity, oder eben durch Daten von der Kinect. Ebenso lassen sich so zur Laufzeit über Bewegungen die Parameter des Algorithmus wie z.B. Geschwindigkeit oder Metrik verändern um Variationen zu erschaffen oder unabhängig vom Pianisten mit zu spielen.

Der Algorithmus wirkt dabei zu keiner Zeit „statisch“ oder „computerisiert“ sondern adaptiert Eigenschaften des vom Menschen gespielten, wie z.B. die Dynamik oder das Timing zur Laufzeit um das „menschliche“ zu bewahren.

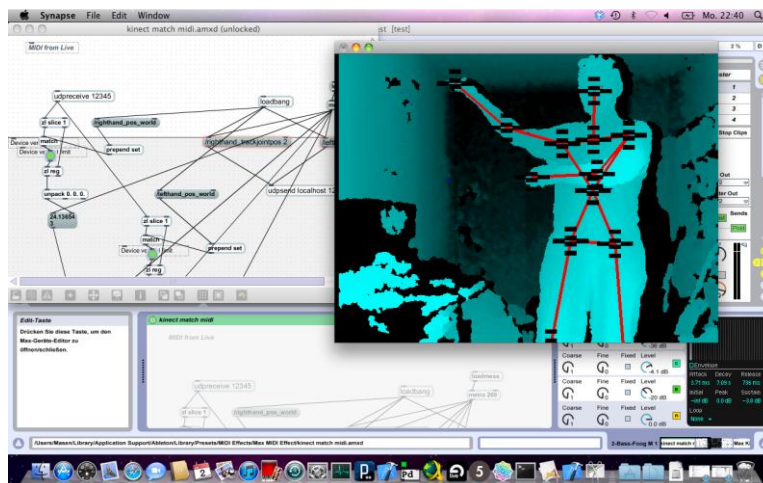


Abbildung 4: Musikalische Komponente in Aktion

3 Aktueller Stand/Aussicht

Zum gegenwärtigen Zeitpunkt besteht eine grundlegende technische Umsetzung mit der wir bereits an der Umsetzung einer Performance, sowohl mit einem international bekannten Choreographen als auch einem Pianisten, arbeiten. Das zusammen erarbeitete Konzept sieht es vor, eine interaktive Aufführung umzusetzen, die Vorzüge des iIgMV mit seinen Möglichkeiten über Tanz, Sounds als auch Visualisierungen zu erzeugen und gleichzeitig mit dem Pianisten an einem MIDI/Player Piano zu interagieren, fanden bei den Beteiligten großen Zuspruch, da es neue Umsetzungsformen ermöglicht.

In der weiteren Entwicklung werden die Komponenten des iIgMV um zusätzliche GUIs ergänzt, zudem wird die Bibliothek an parametrisierbaren Visualisierungsformen (vorgefertigte 2D- und 3D-Visualisierungen und eine Schnittstelle zum Einbinden eigener Modelle) und entsprechender Echtzeitanimationen dieser (Bewegen, Skalieren, Rotieren, Morphen usw.) erweitert, als auch eine erweiterte Bibliothek an algorithmischen Patches (verschiedene Intervall-Strukturen, Tonfolgen Generatoren, stehende Triller u.v.m.) bereitgestellt, die es auch anderen Anwendern/Künstlern erlauben, diese Module zu kombinieren und parametrisieren, um so selbst nach eigenständigen Ideen die gewünschten Ergebnisse zu realisieren.

Literaturverzeichnis

- [P2] T. L. Berg, K. Yun, J. Honorio, D. Chattopadhyay & D. Samaras (2012) „Two-person Interaction Detection Using Body-Pose Features and Multiple Instance Learning“ The 2nd International Workshop on Human Activity Understanding from 3D Data at Conference on Computer Vision and Pattern Recognition, CVPR 2012.
- [P1] J. Paradiso & F. Psaracino (1997) „Optical Tracking for Music and Dance Performance“ Fourth Conference on Optical 3D Measurement Techniques, ETH, Zurich
- [B1] A. Ruschkowski (1998), „Elektronische Klänge und musikalische Entdeckungen“, Stuttgart: Reclam
- [I1] C. Vik, „Kinectar“: <http://kinectar.org/>
- [Y1] C.Vik „Controlling a 4-story pipe organ with the Kinect“: <http://www.youtube.com/watch?v=xEMbjnTJCHM>

Kontaktinformationen

Damian T. Dziwis <d.dziwis@masen-inhabitants.com>
Felix Hofschulte <felixx1810@aol.com>
Michael Kutz <michaelkutz@pixelrebelln.de>
Martin Kutz <martinkutz@pixelrebelln.de>