

Ein Gestenbasiertes Interface zur Bewegung von Klang im Raum

Michael Schlütter¹, Cornelius Pöpel²

Multimedia und Kommunikation, Hochschule Ansbach²

Zusammenfassung

Mit dem Fortschritt im 3D-Audio Bereich werden Bedienelemente benötigt, die eine dreidimensionale Steuerung von Klangquellen ermöglichen. In dieser Arbeit wird die Nutzbarkeit bislang gängiger Instrumente zur Bewegungsumsetzung betrachtet. Anhand der sich ergebenden Diskrepanzen und eigener Überlegungen bei der Nutzung von ein- und zweidimensionalen Bedienelementen für Bewegungen im dreidimensionalen Raum erfolgt der Entwurf eines gestenbasierten Interfaces. Intuitive Handhabbarkeit und mehrparametrische Manipulation des Klanges im Raum sind zu erfüllende Anforderungen, die dem Entwurf zugrunde liegen.

1 Einführung

Aktuelle Entwicklungen in der Audiotechnik befassen sich derzeit vermehrt mit dem Thema „Klang im Raum“ auch „Spatial Audio“ genannt. Dabei beschäftigen sich die Tonschaffenden mit neuen Technologien in der horizontalen Ebene, aber vielmehr noch mit Technologien, welche die vertikale Ebene in die Wiedergabe von Klang mit einbeziehen. Die 40. Audio Engineering Society (AES) Konferenz in 2010 trug daher den Titel „Spatial Audio – Sense The Sound Of Space“.¹ Die Vorträge behandelten Themen aus dem Bereich Surround-Sound und 3D-Audio. Auf der 26. Tonmeistertagung im Jahr 2010 wurde ein Veranstaltungsblock dem Thema „Spatial Audio“ gewidmet.² Auch bei Filmproduktionen in 3D wird eine Überführung tontechnischer Inhalte in die dreidimensionale Welt immer wichtiger (Theile & Wittek 2011, S.5).

Vor dem Hintergrund der Dreidimensionalität stellt sich jedoch die Frage, wie bei der Mischung Klang im Raum bewegt werden kann. Diese Frage wird vom derzeitigen Stand der Entwicklung noch nicht ausreichend beantwortet. Welche Bedienelemente können verwendet

¹ <http://www.aes.org/events/40/program.cfm>, Stand: 23.7.2012.

² http://www.tonmeister.de/tmt/index.php?p=de2010congress_topics, Stand: 23.7.2012.

werden, um eine intuitive und effiziente Manipulation für die 3D Abmischung zu ermöglichen? Wie kann eine dreidimensionale Bedienung gestaltet werden? Bisherige Techniken mit ein- oder zweidimensionalen Bedienelementen können nur eingeschränkt Klänge auf der horizontalen und vertikalen Ebene gleichzeitig bewegen. Tonschaffende fragen nach Eingabegeräten, die für die speziellen Anforderungen zur Steuerung von Klang im Raum entwickelt werden (Peters 2010, S. 50). Die dritte Dimension erfordert neue Denkansätze bei der Entwicklung und Erweiterung von Bedienelementen, um 3D-Audiomischungen intuitiv und effizient erstellen zu können (Schlüter 2011).

Die Europäische Kommission hat hierzu das Projekt *ConGAS* (Gesture Controlled Audio Systems) der European Cooperation in Science and Technology (COST) unter der *COST Action 287* gefördert. Ziel war es, Entwicklung und Fortschritt der Datenanalyse von musikalischen Gesten voranzutreiben, und Aspekte der Steuerung von digitalem Ton und der Musikbearbeitung zu erfassen (COST 2006) (Godøy & Leman 2009).

2 „MixGlove-System“ - ein gestenbasiertes Interface

Ein Grundgedanke des gestenbasierten Interfaces liegt im Loslösen der Arbeit von festen Arbeitsplätzen und starren Bedienelementen. Ziel ist die Verschmelzung von Computing und den natürlichen menschlichen Aktionen und Reaktionen, wie z.B. Gesten. Interfaces sollen intelligenter werden und somit den User unterstützen, bzw. die Arbeit effizienter gestalten (Beilharz 2005, S. 105).

2.1 Grundlegende Idee hinter dem „MixGlove-System“

Mehrkanalton-Produktionen könnten um ein Vielfaches einfacher sein, wenn es ein Interface gäbe, das effizient die Positionierung und Bewegung von Klangquellen im Raum beeinflussen könnte. Die derzeitige Situation zeigt, dass mit bekannten Controllern und Mischpulten in dieser Hinsicht nicht effizient gearbeitet werden kann (Hamasaki et al. 2005, S. 383). Es gibt bereits verschiedene Ansätze, wie ein solches gestenbasiertes Interface sinnvoll umgesetzt werden kann und welche Funktionen damit gesteuert werden. Bekannte Beispiele sind z.B. das *Radio Baton* von Mathews (Mathews 2000) und das von Waisvisz entwickelte System *The Hands II* (Waisvisz 1985) (Waisvisz 1989) zur Steuerung von Synthesizern. Selfridge und Reiss (2011) entwickelten das *Kynan System* zur Steuerung von Equalizern und Kompressoren in einer Stereomischung.

Das Positionieren von Klangquellen erfolgt bei analogen Mischpulten und digitalen Controllern mit ein- und zweidimensionalen Bedienelementen und ist somit wenig intuitiv. Intuition spielt aber beim Mischen eine wichtige Rolle. Das *MixGlove-System* versucht, den Lernprozess für Mischungen zu verkürzen. Klänge sollen fassbar gemacht werden. Mit dem System werden intuitive, dreidimensionale Bewegungen der Hand erkannt, die Daten verarbeitet und die Parameter, wie z.B. Höhe der Klangquelle, den entsprechenden Mischaufgaben angepasst. Dies erhöht die Effizienz beim Positionieren von Klangquellen und erlaubt dadurch einen kreativeren Prozess. Dass eine Mischung intuitiv und mit mehr

Freiheit bei der Performance von statten gehen kann, zeigen z.B. *Selfridge* und *Reiss* (2011). Durch ein gestenbasiertes Interface wird die Positionierung physisch umgesetzt, was mit einem normalen Mischpult so nicht möglich ist.

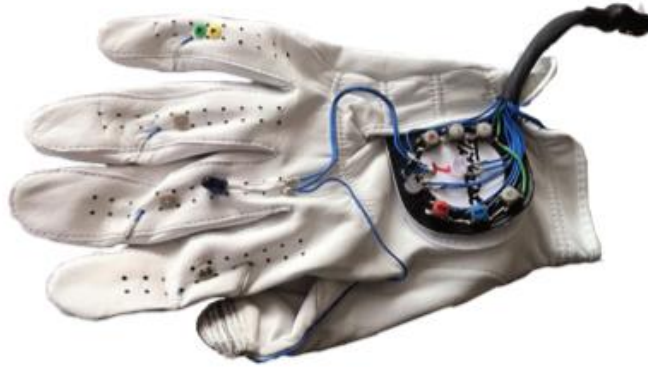


Abbildung 1: Datenhandschuh des MixGlove-Systems

2.2 Umsetzung des „MixGlove-Systems“

Das System besteht aus dem Controller (Datenhandschuh und IR-Tracking Einheiten) und der Sound Generation Unit (Software: Osculator, Max/MSP, DAW und Plug-Ins) und folgt damit der Trennung der Module Gestural Controller und Sound Production, wie bei *Miranda* und *Wanderley* (2006) beschrieben. Eine Übersicht des Mappings zeigt Abbildung 2.

Im Gegensatz zu *Selfridge* und *Reiss* (2011), die für das *Kynan-System* auf die Accelerometer der *Wiimotes* zurückgreifen und somit nur Bewegungen erfassen können, werden beim *MixGlove-System* die Bewegungen der Hand durch Infrarot Tracking über zwei *Wiimotes* erfasst. Der Vorteil liegt in der genauen örtlichen Bestimmung des Datenhandschuhs. Sehr langsame Bewegungen können erfasst werden. Parameter können gehalten und feinjustiert werden, was beim *Kynan-System* nicht gelingt. Allerdings muss sich der Datenhandschuh innerhalb des Tracking-Bereichs befinden, um Bewegungen aufzuzeichnen zu können. Die Auflösung der Infrarot-Kameras ist nach Auffassung der Verfasser und nach *Purkayastha* (Purkayastha, S.N. et al. 2010) ausreichend genau. Klänge können ohne Probleme im Raum platziert werden. Die Daten werden per Bluetooth an den Rechner geleitet und dort von dem Programm *Osculator* in UDP Daten und Midi-Daten übersetzt. Diese Daten werden vom Programm *Max/MSP* verwaltet und an die DAW weitergeleitet, in der die Funktionen gesteuert werden. Zur Umsetzung von 3D-Audio innerhalb der DAW wird das Plug-In *ViMiC*³ verwendet.

³ <http://www.music.mcgill.ca/~nils/ViMiC/AudioUnit/VimicAU-1.0.0-b5-2011-07-12.dmg>, Stand 23.7.2012.

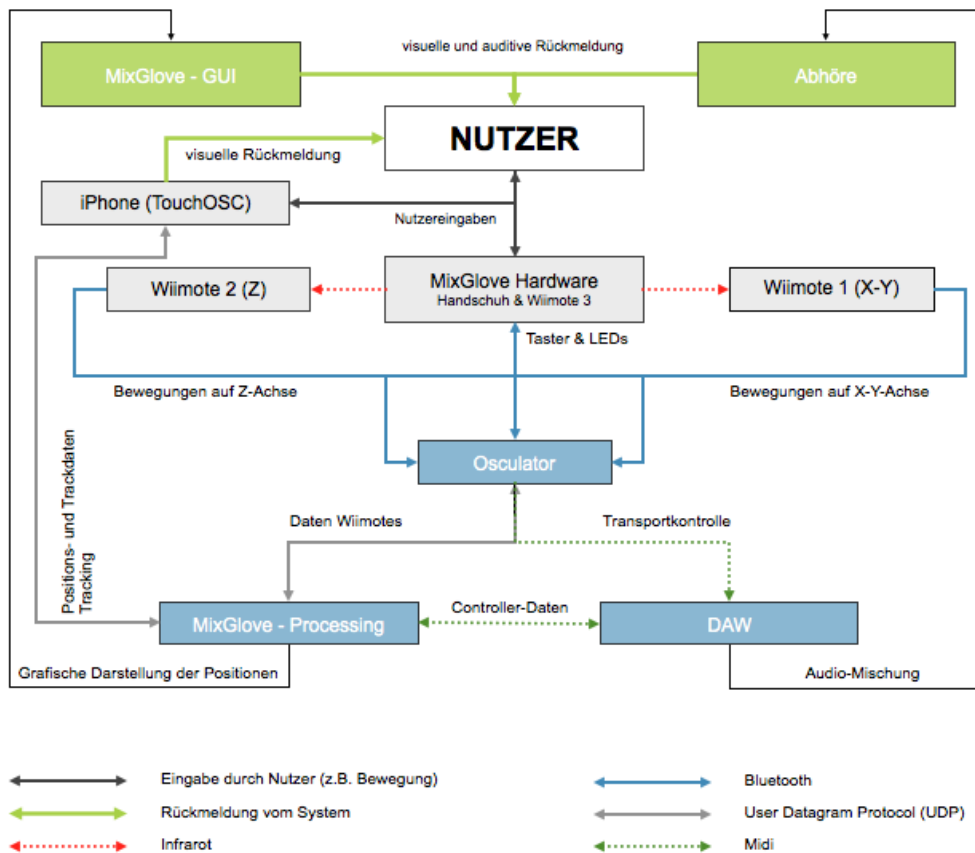


Abbildung 2: Übersicht des Mappings im MixGlove-System

Am Datenhandschuh (Abbildung 1) sind IR-LEDs angebracht, die die Erfassung durch die Infrarot-Kameras der *Wiimotes* ermöglichen. Zudem sind verschiedene Tastpunkte vorhanden, die verschiedene Funktionen, wie etwa Kanal- oder Effektwechsel steuern. Zusätzliche farbige LEDs geben dem Nutzer Rückmeldung, ob er sich im Tracking-Bereich befindet und ob der Anwender das Tracking der Handbewegungen aktiviert hat. Die Tastbefehle werden über einen Transceiver – eine modifizierte *Wiimote*, die über ein Kabel mit dem Datenhandschuh verbunden wird – an den Rechner gesendet. Gleichzeitig empfängt der Transceiver die Steuersignale für die LEDs.

3 Diskussion

Der Bereich 3D-Audio wird sich weiter entwickeln und der Markt wird nach geeigneten Bedienelementen zur Bewegung und Positionierung von Klangquellen im Raum fragen. Umso wichtiger ist es, Bedienelemente zu entwickeln, die den gestellten Anforderungen

gerecht werden. Hierzu könnte als nächster Schritt im Anschluss an diese Arbeit eine Untersuchung erfolgen, die Anwender zum bestehenden *MixGlove-System* befragt. Wo sind Schwachstellen im System? Welche Anforderungen stellt der Nutzer an ein solches System? Wie können die Rückmeldungen des Systems verbessert werden? Welche Elemente fehlen, um die Nutzung zu vereinfachen? Dies sind nur einige der Fragen, die untersucht werden sollten.

Eine technische Alternative zur gestenbasierten Steuerung via *WiiMotes* könnte im Einsatz der *Kinect*⁴ Kamera liegen. Der Open Source Code⁵ ist verfügbar und wird beispielsweise von *Pelletier* als External Object für *Max/MSP* verwendet.⁶ Hier wäre zu untersuchen inwieweit der Einsatz dieser Trackingtechnologie Vorteile bietet. Ein weiterer Ansatz wäre die Verwendung des *LEAP*. Es unterstützt das Tracking von Bewegungen im 3D-Raum und kann die einzelnen Finger unterscheiden. Laut Hersteller soll es dabei um ein vielfaches genauer sein als bisherige Geräte.⁷

4 Ausblick

Das Tracking von Handbewegungen könnte durch die Verwendung der *Wii Motion Plus* Erweiterung verfeinert und ausgeweitet werden. Dies würde eine erweiterte Gestensteuerung ermöglichen und beispielsweise Rotationen von Schallquellen oder ganzen Szenen erlauben. Eine Miniaturisierung des Senders und das direkte Anbringen am Handschuh könnten zudem zu einer größeren Bewegungsfreiheit führen und das An- und Ausziehen des Handschuhs erleichtern. Die Software könnte so erweitert werden, dass verschiedene DAWs durch das *MixGlove-System* gesteuert werden. Zudem könnten die visuellen Rückmeldungen verbessert werden. Bislang erfolgt eine Darstellung der Positionen der Klangquellen und des Datenhandschuhs nur auf zwei Sphären, welche die Draufsicht und Frontansicht repräsentieren. Eine Darstellung eines Raumes in 3D würde das Erfassen der aktuellen Szene erleichtern, und somit den Anwender entlasten.

Literaturverzeichnis

Beilharz, K. (2005): Wireless gesture controllers to affect information sonification. In: *Proceedings of ICAD 05-Eleventh Meeting of the International Conference on Auditory Display*. Limerick, Ireland, July 6-9, 2005, S. 105-112. Stand 22.09.2011.

<http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.100.2079&rep=rep1&type=pdf>

COST (2006): ConGAS – *Gesture controlled audio systems, COST Action 287, Progress Report, Revision 1.7*, 2006. Stand: 25.11.2011.

⁴ <http://www.xbox.com/de-DE/Kinect/>, Stand 23.7.2012.

⁵ http://openkinect.org/wiki/Main_Page, Stand 23.7.2012.

⁶ <http://jmpelletier.com/freenect/>, Stand 23.7.2012.

⁷ <https://live.leapmotion.com/about.html>, Stand 23.7.2012.

- http://w3.cost.eu/fileadmin/domain_files/TIST/Action_287/progress_report/progress_report287.pdf
- Godøy, R. & Leman, M. (Editors) (2009): *Musical gestures – sound, movement, and meaning*. New York: Routledge.
- Mathews, M. (2000): *Radio Baton Instruction Manual, Preliminary Edition*. San Francisco. Stand: 25.11.2011.
<http://www.csounds.com/max/manuals/BatonManual.pdf>
- Miranda, E. R. & Wanderley, M. M. (2006): *New digital musical instruments: control and interaction beyond the keyboard*. Middleton: A-R Editions.
- Peters, N. (2010): *Sweet [re]production: Developing sound spatialization tools for musical applications with emphasis on sweet spot and off-center perception*. Montreal: Diss. McGill University.
- Purkayastha, S.N. et al. (2010): Analysis and Comparison of Low Cost Gaming Controllers for Motion Analysis. In: *2010 IEEE/ASME International Conference on Advanced Intelligent Mechatronics (AIM)*, S. 353-360. Stand: 16.08.2011.
<http://mahilab.rice.edu/sites/mahilab.rice.edu/files/publications/899-AIM%20final%20conference%20version.pdf>
- Schlütter, M. (2011): *Entwicklung eines gestenbasierten Interfaces zur dreidimensionalen Bewegung von Klang im Raum*. Ansbach: Bachelorarbeit an der Hochschule Ansbach.
- Selfridge, R. & Reiss, J. (2011): *Interactive mixing using Wii Controller*. London: Audio Engineering Society Convention Paper 8396, presented at the 130th Convention 2011 May 13–16. Stand: 22.09.2011.
<http://www.elec.qmul.ac.uk/people/josh/documents/SelfridgeReiss-2011-AES130.pdf>
- Theile, G. & Wittek, H. (2011): *Die dritte Dimension für Lautsprecher- Stereophonie*. Stand: 30.08.2011.
http://www.hauptmikrofon.de/doc/Auro3D_D_0311.pdf
- Waisvisz, M. (1985): The Hands, a set of remote MIDI-controllers. In: *Proceedings of the 1985 International Computer Music Conference (ICMC '85)*. Vancouver, S. 313-318. Stand: 26.11.2011.
<http://www.vasulka.org/archive/ExhFest11/STEIM/Steim-MIDIessay.pdf>
- Waisvisz, M. (1989): *The Hands II*. Stand 26.11.2011.
<http://www.crackle.org/The%20Hands%202.htm>
- Hamasaki, K. et al. (2005): 5.1 and 22.2 Multichannel Sound Productions Using an Integrated Surround Sound Panning System. In: *NAB BEC Proceedings 2005*, S. 382-387. Stand: 23.08.2011.
http://www.nhk.or.jp/digital/en/technical_report/pdf/nab200502.pdf

Kontaktinformationen

Michael Schlütter, Charles-Lindbergh Str. 11, 90768 Fürth, mschluetter@onlinehome.de
www.thierra-productions.de

Cornelius Pöpel, Hochschule Ansbach, Residenzstr. 8, 91522 Ansbach,
cornelius.poepel@hs-ansbach.de, Tel. 0981 4877-359