Betreuende Lehrkräfte: Jakob Sudau, Prof. Dr. Tessa Taefi

Projektteilnehmer: Magdalena Lucreteanu, Kirsten Grahl

Hochschule für Angewandte Wissenschaften Hamburg

Fakultät Design, Medien und Information

Department Medientechnik

Projektkonzept Audio-Video-Programmierung

# Papier-Tastatur zur Klangerzeugung

## Inhalt

1. Einleitung

2. Projektziel

3. Anforderungsanalyse

4. Technische Rahmenbedingungen

5. Technisches Konzept

6. Bedienkonzept

7. Zeitplan

8. Teamplanung

## 1. Einleitung

Das Projekt “Papier-Tastatur zur Klangerzeugung” erfolgt sowohl im Rahmen des Wahlkurses “Audio-Video-Programmierung” als auch im Rahmen des Pflichtmoduls “Projekt B”. Die Lehrveranstaltung “Audio-Video-Programmierung” findet im Wintersemester 2021/2022 bei den Dozenten Jakob Sudau und Prof. Dr. Tessa Taefi statt. Bei dem Inhalt der Veranstaltung handelt es sich um die Programmierung von Video-Skripten mit Python und OpenCV, die Entwicklung von Algorithmen zur Video-Objekterkennung, die Programmierung von Audio-Skripten mit JavaScript und WebAudio und auch um die Entwicklung von Algorithmen zur Klangerzeugung. Das Ziel der Veranstaltung ist die Entwicklung eines Softwaresystems, das durch die Auswertung eines Live-Video-Streams einen Live-Audio-Stream erzeugt.

Dieses Projekt setzt die Inhalte und das Ziel der Lehrveranstaltung in einem “Projekt B” um. Am Ende der Projektarbeit wird das Projektergebnis präsentiert und vorgeführt. Zusätzlich wird das Ergebnis durch eine Dokumentation festgehalten. Das Projekt wird von den Teammitgliedern Magdalena Lucreteanu (Matrikelnummer: 1882817) und Kirsten Grahl (Matrikelnummer: 2435019) durchgeführt.

## 2. Projektziel

Das Ziel des Projektes “Papier-Tastatur zur Klangerzeugung” ist das Bedienen eines einfachen Klaviers oder Synthesizers, ohne dabei auf ein physisch vorhandenes Klavier oder einen physisch vorhandenen Synthesizer angewiesen zu sein. Für das Spielen wird lediglich ein Computer mit einer Videokamera und ein Blatt Papier mit der aufgedruckten Klaviertastatur benötigt.

Die Tastatur wird auf ein Blatt Papier gedruckt. Danach kann eine Videokamera an einen Computer angeschlossen und auf dieses Papier ausgerichtet werden. Der Benutzer kann nun seine Finger auf der Papier-Tastatur bewegen und durch das Berühren der Tasten, Knöpfe und Regler dementsprechende Klänge erzeugen. Die Software sorgt für die Auswertung des Videos und die Umsetzung ins dazugehörige Audio. Wie in Abbildung 1 der Tastatur zu erkennen, kann der Benutzer nicht nur eine einfache Klaviertastatur bedienen, sondern auch die Lautstärke ändern, zwischen einem Klavier und Synthesizer wechseln und zusätzlich die Parameter Attack und Release steuern.

Ein Bild, das Text, Musik, Klavier enthält.

Automatisch generierte Beschreibung

Abbildung 1: Vorlage für die Papier-Tastatur

## 3. Anforderungsanalyse

Im Rahmen des Projekts entsteht eine Software, die durch die Auswertung eines Live-Video-Streams einen Live-Audio-Stream erzeugen kann. In diesem Projekt “Papier-Tastatur zur Klangerzeugung” kann der Nutzer eine Klaviertastatur auf einem Blatt Papier mit einer, an einem Computer angeschlossenen, Videokamera aufnehmen und anschließend durch die Berührung der Tasten, Knöpfe und Regler in Echtzeit Klänge über den Computer erzeugen. Dies wird durch die in diesem Projekt entwickelte Software umgesetzt.

Zuerst wird die Tastatur wird auf ein Blatt Papier gedruckt, hierfür gibt es eine vorgefertigte Vorlage. Die Tastatur kann bei nicht vorhandenem Drucker auch der Vorlage entsprechend nachgezeichnet werden. Bei nicht vorhandenen Materialien gibt es zudem die Möglichkeit eine virtuelle Klaviertastatur im Webbrowser mit der Computermaus zu bedienen. Diese Möglichkeit sollte jedoch nur im Ausnahmefall in Anspruch genommen werden und stellt nicht das hauptsächliche Ziel dieses Projektes dar.

Danach wird eine Videokamera an einen Computer angeschlossen auf das Blatt Papier mit der aufgedruckten Tastatur gerichtet. Das Papier sollte gut ausgeleuchtet und in der Kamera erkennbar sein. Der Benutzer kann anschließend seine Finger auf dem Papier in die jeweiligen Bereiche der Tasten, Knöpfe und Regler bewegen. Durch das Bewegen der Finger in den jeweiligen Bereich werden die Tasten, Knöpfe oder Regler ausgelöst.

Durch die Lautstärketasten kann die Lautstärke verändert werden. Die Länge des Berührens der Taste entscheidet über die Stärke der Lautstärkeveränderung. Der Benutzer kann durch die jeweiligen Knöpfe zwischen dem aufgenommenen Klang eines Klaviers und dem Klang eines Synthesizers wechseln. Das einmalige Berühren der jeweiligen Taste reicht hierfür aus. Die Zeit der Berührung sollte mindestens eine Sekunde betragen, eine längere Berührung verändert jedoch nichts weiter. Zusätzlich kann durch zwei Regler die Attack- und Release-Zeit verändert werden. Die Position des Fingers auf dem jeweiligen Regler ist hierbei der entscheidende Faktor. So können sowohl die Attack- als auch die Release-Zeit auf einen Wert zwischen null und einer Sekunde verändert werden. Wird der Finger auf eine Taste der Tastatur bewegt, kann dadurch der jeweilige Ton erzeugt werden. Der ausgegebene Ton hängt hierbei von den vorher eingestellten eben beschriebenen Parametern ab. Dieser Ton wird in einem Webbrowser ausgegeben. Die Ausgabe kann über angeschlossene Lautsprecher oder Kopfhörer erfolgen.

Die Software soll hierbei in der Lage sein, die Videodaten in Echtzeit auszuwerten und Aktionen in unter einer Sekunde nach dem Auslösen einer Taste auszuführen. Erfolgt die Aktion zu spät nach der Bewegung des Fingers, könnte der Unterschied zwischen der Bewegung und der Wahrnehmung des Audios das Erlebnis des Nutzers beeinträchtigen. Gleichzeitig darf eine Aktion aber auch nicht zu einfach auslösbar sein. Die Tasten dürfen nicht durch eine versehentlich Bewegung sofort zu einer Aktion des Systems führen, auch wenn diese gar nicht vom Nutzer beabsichtigt war.

Des Weiteren ist der Umgang mit mehreren Fingern oder der ganzen Hand auf der Tastatur zu unterscheiden. Durch die Bewegung der Hand über die Tastatur, um die Knöpfe über der Tastatur zu erreichen, können möglicherweise mehrere Tasten bedeckt werden. Hierbei sollten jedoch nicht alle Tasten ausgelöst werden, sondern nur der gewünschte Knopf. Sind aber mehrere Finger auf der Tastatur und nicht die ganze Handfläche, sollte es möglich sein mehrere Tasten auszulösen, um dadurch einen Ton oder mehrere Töne zu erzeugen. Gegebenenfalls ist hier die Umsetzung mit farbigen Stiften, anstatt einer Hand in Betracht zu ziehen.

## 4. Technische Rahmenbedingungen

Für die Umsetzung der Anwendung wird unser Team mit Windows Desktop-Computer arbeiten. Die Lauffähigkeit der entstehenden Anwendung auf anderen Betriebssystemen kann auf Grund von fehlender Testbarkeit nicht garantiert werden. Die Anwendung sollte trotzdem auch auf macOS Desktop-Computern ausführbar sein.

Für die Aufnahme des Live-Videos wird eine funktionierende Videokamera benötigt, die zudem an einen Computer angeschlossen werden kann. Zur Ausgabe des Audios werden interne oder externe Lautsprecher oder Kopfhörer benötigt. Zudem wird ein Webbrowser zur Audioausgabe verwendet, der JavaScript und MIDI unterstützt. Eine Internetverbindung wird nicht benötigt. Ein Browser ohne aktiviertes JavaScript würde dazu führen, dass unser Projekt für den Benutzer nicht nutzbar ist. Unser Team wird mit Google Chrome arbeiten, da dieser Browser auch schon in der Vorlesung verwendet wurde.

Programmieren werden wir in Visual Studio Code. Darin werden wir die Live Server Erweiterung benutzen, um die Seiten in einem Webbrowser testen zu können. Die Anwendung wird mit Python, OpenCV, JavaScript, WebAudio, HTML und CSS erstellt.

Zum Austausch von Informationen zwischen der Videoauswertung und dem Audio werden wir MIDI verwenden. Sollte dieses sich nicht als ausreichend herausstellen, werden wir mit WebSockets arbeiten.

Zudem wird die auf Papier gedruckte Tastatur benötigt. Eine Vorlage dafür wurde bereits in Clip Studio Paint erstellt und wird auch späteren Nutzern zur Verfügung gestellt. Ohne eine Papier-Tastatur wird für den Benutzer nur ein Teil des Projektes bedienbar. Die Klaviertastatur und Knöpfe in Browser können aber weiterhin bedient werden. Alle Audiofunktionalitäten werden dort vollständig laufen.

## 5. Technisches Konzept

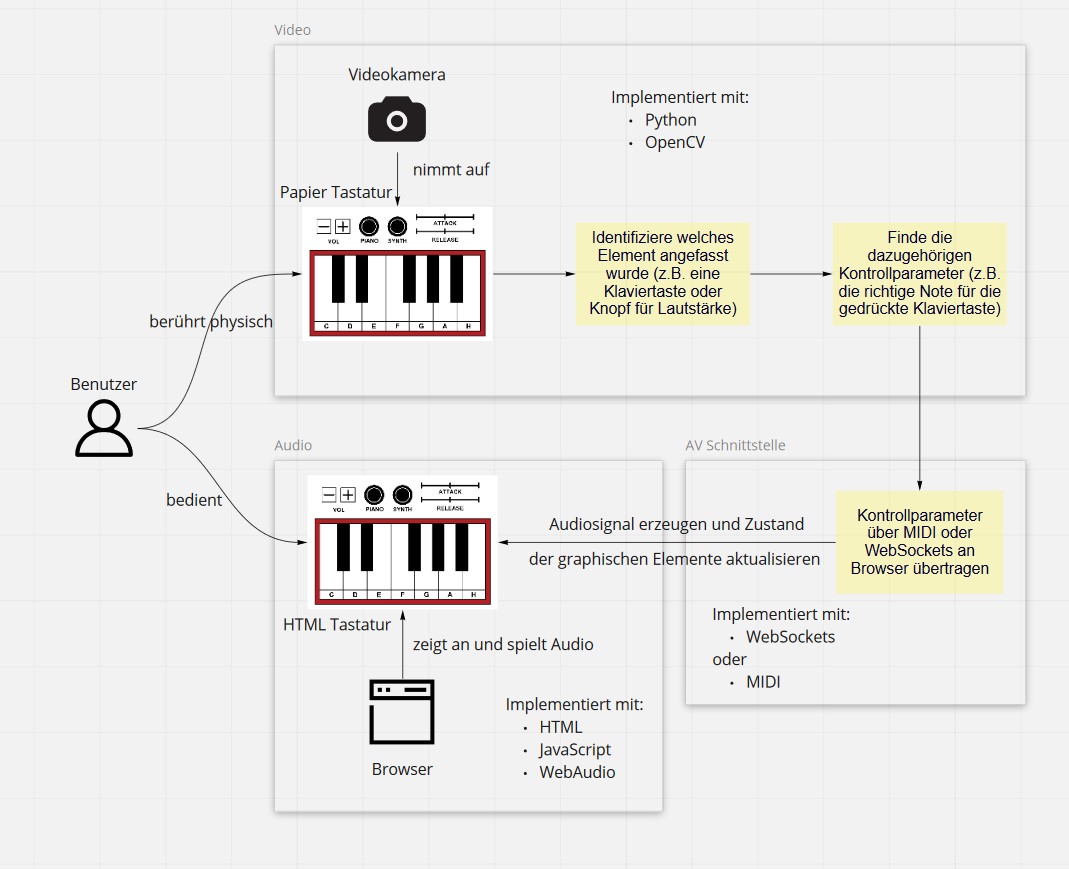


Abbildung 2: Technisches Konzept im Diagramm dargestellt

Zur Verdeutlichung des technischen Konzepts wurde ein Miro-Board erstellt, dieses ist unter folgendem Link zu finden: <https://miro.com/app/board/uXjVOahhAI0=/>. Die Abbildung 2 zeigt einen Ausschnitt des Miro-Boards, unser Architekturdiagramm. Dort sind alle Komponenten unserer Anwendung beschrieben: Video, Audio und die Audio-Video-Schnittstelle. Zudem wird die Interaktion zwischen den Komponenten gezeigt.

Wie im Diagramm zu erkennen, erfolgt die Auswertung des Live-Videos mit Python und OpenCV. Hier wird ein Algorithmus auf das Live-Video angewendet, um die Positionen der einzelnen Bereiche der Tastatur auf dem Papier in Echtzeit zu erkennen. Zudem wird die Position des Fingers ebenfalls in Echtzeit erkannt. Bei der Übereinstimmung von den Bereichen und der Position des Fingers wird die jeweilige Aktion der Taste, des Knopfes oder des Reglers ausgeführt.

Um die Aktion ausführen zu können, werden die Daten der Aktion werden mit MIDI oder bei Bedarf mit WebSockets an den Audioteil übertragen. Hier erfolgt die Programmierung mit JavaScript und WebAudio. Für den Klavierklang werden hier dann aufgenommene Klaviertastentöne abgespielt. Ist der Synthesizer ausgewählt, wird dieser hier erstellt und abgespielt. Zudem findet in diesem Teil die Kontrolle der Lautstärke und das Verändern der Attack- und Release-Zeit statt. In diesem Teil werden anschließend auch dementsprechend die Anweisungen an den Webbrowser weitergegeben, die entsprechenden Töne zu spielen.

## 6. Bedienkonzept

Die Bedienung der Software erfolgt hauptsächlich über die Papier-Tastatur. Zusätzlich wird aber auch als Alternative eine Tastatur im Webbrowser angezeigt, diese ist jedoch wie bereits erwähnt nur für Ausnahmefälle gedacht und nicht das eigentliche Projektziel. Die Papier- und die Browser-Tastatur werden aus Benutzersicht gleich bedient und beide verfügen über die fast gleichen Funktionalitäten. Ein paar Unterschiede gibt es trotzdem, denn die Papier-Tastatur wird mit den Fingern und die Browser-Tastatur mit der Maus bedient. Zudem ist das Berühren einer Taste auf dem Papier nicht grafisch darstellbar, im Browser kann dies angezeigt werden. Ein graphisches Beispiel für so eine Tastatur befindet sich in der Abbildung 1 im Abschnitt “2. Projektziel”.

Die dort angezeigten Bedienelemente sind wie folgt:

* Weiße Klaviertasten: spielen volle Töne.
* Schwarze Klaviertasten: spielen Halbtöne.
* Lautstärkeknöpfe: verändern die Lautstärke. Mit dem “-” (Minus) Knopf wird die Lautstärke reduziert und mit dem “+” (Plus) Knopf wird sie erhöht.
* Piano Knopf: es wird das Klavier als Spielinstrument ausgewählt.
* Synth Knopf: es wird ein Synthesizer als Spielinstrument ausgewählt.
* Regler: die Attack oder Release Werte setzen.

## 7. Zeitplan

|  |  |  |
| --- | --- | --- |
| **Termin** | **Aufgabe** | **Aufwand** |
| 14.10 - 09.12 | Durcharbeitung der Vorlesungsinhalte, Bearbeitung der Video und Audio Aufgaben, Vorstellung der bearbeiteten Aufgaben | 12 PT |
| 10.12 | Projektideen sammeln und bewerten/vergleichen | 2 PT |
| 15.12 | Projektkonzept schreiben | 3 PT |
| 16.12 | Projektkonzept vorstellen | 0.25 PT |
| 17.12 | GitHub Projekt erstellen | 0.25 PT |
| 19.12 | Kommunikation mit WebSockets erproben (inklusive Server starten). Eventuell die Kommunikation über eine MIDI Schnittstelle ermöglichen. | 2 PT |
| 26.12 | Identifizieren welche Taste wurde gedrückt auf die Papier Tastatur mittels OpenCV | 5 PT |
| 26.12 | Klaviertastatur graphisch mit HTML erstellen | 3 PT |
| 02.01 | Informationen über gedrückte Tasten mit MIDI an HTML übergeben | 2 PT |
| 02.01 | Papier- und HTML-Tastaturen sind beide funktional | 2 PT |
| **Termin** | **Aufgabe** | **Aufwand** |
| 09.01 | Weitere Knöpfe (für Lautstärke, Effekte usw.) in HTML erstellen | 5 PT |
| 13.01 | Präsentation Inkrement | 0.25 PT |
| 16.01 | Identifizieren welcher Knopf wurde gedrückt auf die Papier Tastatur mittels OpenCV | 5 PT |
| 23.01 | Informationen über die gedrückten Knöpfe mit WebSockets/MIDI an HTML übergeben | 2 PT |
| 23.01 | Papier- und HTML-Knöpfe sind beide funktional | 2 PT |
| 26.01 | Weitere Tests eventuell Bugfixes | 2 PT |
| 27.01 | Präsentation Projektergebnis | 0.25 PT |
| 03.02 | Abgabe Projektdokumentation und Konzept | 0.25 PT |

## 8. Teamplanung

Unser Team besteht aus den beiden Mitglieder Kirsten Grahl (Matrikelnummer: 2435019) und Magdalena Lucreteanu (Matrikelnummer: 1882817). Um das Projekt innerhalb des vorgegebenen Zeitplans erstellen zu können, haben wir uns dazu entschieden agil zu arbeiten. Dies bietet sich vor allem durch unsere kleine Teamgröße und den kurzen Zeitraum für die tatsächliche Umsetzung des Projektes von ungefähr einem Monat an.

Auf Grund unsere kleinen Teamgröße werden wir beide alle Rollen in der Produktentwicklung übernehmen. Wir werden beide an diesem Projekt als Product Ownerinnen und Entwicklerinnen arbeiten. Der oder die Product Owner\*in definiert die Anforderungen und somit was noch zu erledigen ist. Dadurch wird der zeitliche Ablauf des Projektes klar definiert und es werden Deadlines festgelegt. Der oder die Entwickler\*in ist für die Umsetzung dieser Anforderungen zuständig und implementiert diese im geplanten Zeitraum. Wie bereits erwähnt, werden in diesem Projekt alle Rollen von uns beiden übernommen. Dies bietet den Vorteil, innerhalb von kürzester Zeit Anforderungen definieren oder anpassen zu können und diese dann direkt im Anschluss direkt umsetzen zu können.