

Straßenmusik

Projektkonzept für Audio-/Videoprogrammierung

von

Fabian Hensel	2364627
Jakob l'Etienne	2365009
Daniel Friesen	2340482

Einleitung

Dieses Projekt entsteht im Rahmen der Veranstaltung "Audio-Video-Programmierung". Das Ziel ist es eine Software zu erstellen, welche aus einem Video-Feed (z.B. Live-Kamera/Webcam/Videoaufnahme) Töne oder Musik erzeugt und diese am Ende der Veranstaltung zu präsentieren.

Projektziel

Das Ziel der Software ist, das Videobild einer viel befahrenen Straße zu analysieren und in musikalische Klänge zu verwandeln. Die Kamera ist dabei frontal und möglichst weit von oben platziert, zum Beispiel auf einer hohen Autobahnbrücke möglichst mittig zur Straße ausgerichtet. Die vorbeifahrenden Fahrzeuge repräsentieren einzelne Noten, die gespielt werden sobald sie eine bestimmte Linie im Bild passieren. Die Länge des Fahrzeugs repräsentiert dabei den Notenwert, also die Länge des klingenden Tons. Das Aussehen des Fahrzeugs, genauer gesagt die durchschnittliche Helligkeit, bestimmt die Tonhöhe. Um Dissonanzen zu vermeiden, werden nur Töne aus einer festgelegten Tonleiter ausgespielt. Weil kein Einfluss darauf genommen werden kann, wann ein Ton erklingt, wird der musikalische Gesamteindruck einen eher "jazzigen" oder experimentellen Charakter haben.

Anforderungsanalyse

Die Software soll in der Lage sein, via Kameraaufnahme verschiedene Fahrzeugformen und Fahrzeug Helligkeiten einer befahrenen Straße aus nahezu "Top-Down"-Sicht zu detektieren. Die Detektierung soll tagsüber geschehen. Die Software soll sowohl mit Live-Aufnahmen als auch mit aufgezeichneten Aufnahmen umgehen können. Des weiteren soll sie in der Lage sein verschiedene Töne abzuspielen sobald die detektierten Fahrzeuge eine vorab festgelegte Linie auf der Straße überqueren. Diese Töne sollen den Formen und Helligkeiten der detektierten Fahrzeuge entsprechen. D.h. bei der Detektierung eines größeren Fahrzeug soll z.b. ein tieferer Ton abgespielt werden als bei der Detektierung eines kleineren Fahrzeug. Genauer: Die Länge des Fahrzeugs soll den Notenwert repräsentieren und die Helligkeit des Fahrzeugs soll die Tonhöhe bestimmen. Die Geschwindigkeit der Fahrzeuge soll keinen Einfluss auf den Ton haben. Außerdem soll die Software die einzelnen Spuren in unterschiedliche Oktavlagen einteilen können. Die rechte Spur, also die die hauptsächlich von Lastwagen befahren wird, soll eher tiefe und lange Töne produzieren. Die Überholspur soll für kurze und hohe Töne zuständig sein.

Technische Rahmenbedingungen

Die Software wird für Desktop-Computer erstellt.

Der Video-Teil der Software wird mit der Programmiersprache Python erstellt. Die konkrete Analyse des Video-Feeds geschieht mithilfe von OpenCV.

Der Audio-Teil der Software wird mit Javascript und HTML/CSS erstellt. Die Tonausgabe bzw. Klangerzeugung geschieht mithilfe der Web Audio API.

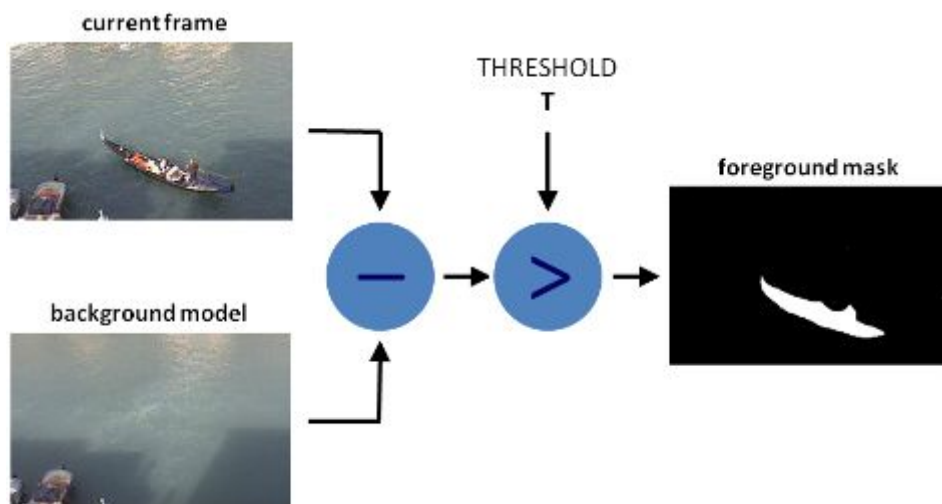
Die Verbindung der beiden Teile wird über das MIDI-Protokoll realisiert.

In der Veranstaltung wurde der Umgang mit diesen Frameworks/Programmiersprachen näher gebracht, daher soll auch dieses Projekt mit diesen Werkzeugen erstellt werden.

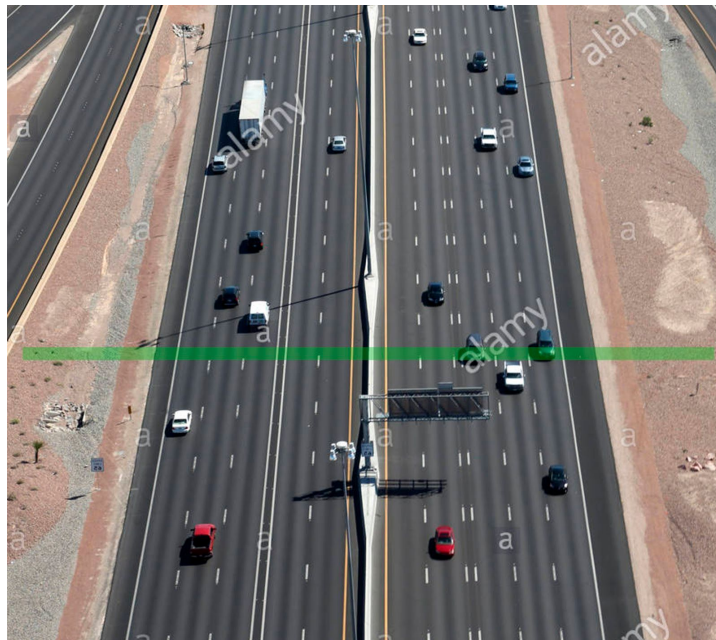
Technisches Konzept

Die Videoanalyse soll mithilfe von Funktionen und Algorithmen von OpenCV realisiert werden. Zuerst sollen die Frames des Input-Videos in Graustufenbilder umgewandelt werden. Die Erkennung der Fahrzeuge soll dann mithilfe einer Background-Subtraction geschehen, daher ist eine statische Kamera beim Live-Videomaterial/Videoaufnahmen sehr wichtig. Eine Background-Subtraction arbeitet wie folgt:

Es wird ein Frame des Videomaterials als Hintergrund festgelegt und jeder folgende Frame wird vom Hintergrund abgezogen. Pixel die bei folgenden Frames anders sind als beim Hintergrund werden dadurch markiert und aus diesen markierten Pixeln wird eine Binär-Maske erstellt.



Mithilfe dieser Maske ist es möglich die Fahrzeuge zu erkennen und ihre Position im Bild zu bestimmen. Wenn die Fahrzeuge eine vorab festgelegte Linie im Bild passieren, soll mithilfe des MIDI-Protokolls eine Note "gespielt" werden, wobei die Tonhöhe vom durchschnittlichen Helligkeitswertes des Fahrzeugs abhängt und die Tonlänge von der Länge des Fahrzeugs. Diese Note wird mithilfe der Web Audio API ausgelesen und als hörbarer Ton über Lautsprecher/Kopfhörer ausgegeben.

Bedienkonzept

Key: C D E F G A B

Minor Major Pentatonic

Sobald das Videobild in der Software initialisiert wurde, wird das Bedienkonzept minimalistisch gehalten sein.

Der Benutzer soll über drei wesentliche Bedienelemente verfügen:

1. Die Position der Linie, die das ausspielen der Noten durch vorbeifahrende Autos auslöst (Triggerline). Dadurch lässt sich auch bei unterschiedlichen Videostreams die ideale Position der Triggerline anpassen um beispielsweise Verdeckung durch Straßenschilder oder Bäume zu vermeiden.
2. Den Grundton, auf welchem die Tonleiter basiert, aus der letztendlich die Töne ausgespielt werden.
3. Die Art der Tonleiter, die wiederum auf besagtem Grundton aufbaut.

Durch die beschränkten Möglichkeiten soll einerseits die Bedienbarkeit erleichtert werden, andererseits lassen sich durch die Festlegung von verschiedenen Tonarten eine Vielfalt an unterschiedlichen Klangfarben erstellen.

Eine Erweiterung der Bedienelemente zu synthesesizertypischen Parametern ist denkbar, jedoch kein explizites Ziel im Rahmen des Projekts.

Zeitplan

Meilensteine:

22.12.2020	Objekterkennung erfolgreich
05.01.2021	Vorstellung Prototyp
19.01.2021	Projektpräsentation
26.01.2021	Projektabgabe

Teamplanung

Daniel Friesen:

- Umwandlung der ankommenden Python-Daten in Midi-Daten, die den Vorgaben der Benutzeroberfläche entsprechen
- Prototyp bis 22.12, Arbeitsaufwand ca. 10 h

Jakob l'Etienne:

- Videoanalyse mit Python
- Objekterkennungs-Prototyp bis 22.12, Arbeitsaufwand ca. 10 h

Fabian Hensel:

- Tonerzeugung mit Javascript
- Erstellung der Benutzeroberfläche