

Erstellung von Pop-Video Remixen unter Einsatz von Typ-3 Grammatiken

Studienarbeit

Im Rahmen der Prüfung:
Bachelor of Science (B. Sc.)

des Studienganges Informatik
an der Dualen Hochschule Baden-Würtemberg Karlsruhe

von
Leonhard Zeller, Lukas Bailey

Abgabedatum 15. September 2025

Bearbeitungszeitraum 30.06.2025 - 15.09.2025

Matrikelnummern, Kurs 2833211, 8232296, TINF23B2

Ausbildungsfirma SAP SE

Dietmar-Hopp-Allee 16
69190 Walldorf, Deutschland

Betreuer der Dualen Hochschule Prof. Dr. Sebastian Ritterbusch

Eidesstattliche Erklärung

Ich versichere hiermit, dass ich meine Studienarbeit mit dem Thema:

Erstellung von Pop-Video Remixen unter Einsatz von Typ-3 Grammatiken

gemäß § 5 der „Studien- und Prüfungsordnung DHBW Technik“ vom 29. September 2017 selbstständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe. Die Arbeit wurde bisher keiner anderen Prüfungsbehörde vorgelegt und auch nicht veröffentlicht.

Ich versichere zudem, dass die eingereichte elektronische Fassung mit der gedruckten Fassung übereinstimmt.

Karlsruhe, 28. Februar 2026

gez.: Leonhard Zeller, Lukas Bailey

ABSTRACT

- Deutsch -

Inhaltsverzeichnis

1 Einleitung	1
1.1 Motivation	1
1.2 Zielsetzung	1
1.3 Aufbau der Arbeit	1
2 Theoretische Grundlagen	1
2.1 Struktur populärer Musik	2
2.2 Formale Sprachen und Typ-3 Grammatiken	2
2.3 Verfahren zur Gesangstrennung	2
2.4 Methoden der digitalen Signalverarbeitung zur Mustererkennung	2
3 Praktische Umsetzungsversuche	2
3.1 Datenbasis	3
3.2 Segmentierung und Mustererkennung	3
3.3 Grammatik-Erstellung	3
3.4 Generierung der Remixe	3
3.4.1 Musik-Remixe	3
3.4.2 Video-Remixe	3
4 Auswertungsstrategie	3
4.1 Evaluation der Mustererkennung	4
4.2 Bewertung der Remix-Qualität	4
5 Betrachtung der Ergebnisse	4
5.1 Ergebnisse der Segmentierung	5
5.2 Präsentation der erzeugten Beispielremixe und Beispielvideos	5
5.3 Diskussion der Herausforderungen	5
6 Fazit und Ausblick	5
Literaturverzeichnis	a
Anhangsverzeichnis	A

1 Einleitung

1.1 Motivation

Mit der Quantifizierung von Musik durch moderne Sequenzer und Digital Audio Workstations (DAWs) ist die meiste populäre Musik exakt in mehrere Takte und Strukturen wie Intro, Refrain, Strophe, Bridge und Outro unterteilt. Durch das erkennen dieser Strukturen können unter Einsatz einer einfachen Typ-3 Grammatik Remixe bestehender Songs generiert werden. Dabei können dann neue beliebig lange, situative reagierende Abmischungen von Songs und Videos erzeugt werden.

1.2 Zielsetzung

Ziel dieser Arbeit ist es ein System zu erstellen, dass im ersten Schritt die Struktur sinnvoll gewählter Pop-Musik Videos erkennen soll und aus diesen durch Zuordnung der Strukturelemente zu einer Typ-3 Grammatik für die menschliche Wahrnehmung stimmige Remixe zu generieren. Dabei soll eine beliebige Länge angegeben werden können, die ungefähr eingehalten wird. Zudem soll es möglich sein weitere nicht aus dem ursprünglichen Song stammende Elemente in den Remix einzubringen. Für die Videos sollen Abschnitte mit einem Crossfade oder anderen Übergängen ineinander greifen.

1.3 Aufbau der Arbeit

- Kapitel 2 Theoretische Grundlagen
- Kapitel 3 Praktische Umsetzungsversuche
- Kapitel 4 Auswertungsstrategie
- Kapitel 5 Betrachtung der Ergebnisse
- Kapitel 6 Fazit und Ausblick

2 Theoretische Grundlagen

2.1 Struktur populärer Musik

2.2 Formale Sprachen und Typ-3 Grammatiken

2.3 Verfahren zur Gesangstrennung

2.4 Methoden der digitalen Signalverarbeitung zur Mustererkennung

3 Praktische Umsetzungsversuche

3.1 Datenbasis

3.2 Segmentierung und Mustererkennung

3.3 Grammatik-Erstellung

3.4 Generierung der Remixe

3.4.1 Musik-Remixe

3.4.2 Video-Remixe

4 Auswertungsstrategie

4.1 Evaluation der Mustererkennung

4.2 Bewertung der Remix-Qualität

5 Betrachtung der Ergebnisse

5.1 Ergebnisse der Segmentierung

5.2 Präsentation der erzeugten Beispielremixe und Beispielvideos

5.3 Diskussion der Herausforderungen

6 Fazit und Ausblick

Literaturverzeichnis

Anhangsverzeichnis