

Typ-3 Grammatiken für Pop-Video Remixe

Themenvorschlag für eine Studienarbeit

Prof. Dr. Sebastian Ritterbusch

7. August 2025

1 Motivation

Mit dem Einzug von Sequenzern ist einige populäre Musik inzwischen außerordentlich stark und exakt in mehrere Takte (4, 8, 12, 16, ...) umfassende Teile wie beispielsweise Intro, Refrain und Strophen strukturiert. Durch eine automatisierte Identifikation von Teilen in Symbolen einer Sprache kann eine möglichst einfache Typ-3 Grammatik aufgestellt werden, die das ursprüngliche Lied und auch sinnvolle Umsortierungen akzeptiert. Damit können dann neue Abmischungen von Liedern und Videos in fast beliebiger Länge oder situativ reagierender Ausgestaltung erstellt werden.

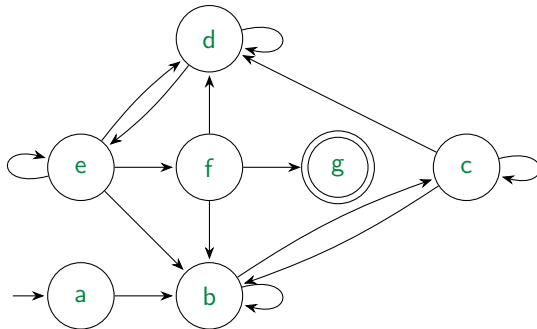


Abbildung 1: Automatisiert durch Autokorrelation bestimmte, unvollständige und nicht minimierte Struktur des Videos „Tied by Sound“ aus 214 Takten mit ursprünglicher Form **abcdefg**, wobei **a** hier beispielsweise das Intro in Takten 1-16 bezeichnet. Jeder Zustand wird in diesem Modell genau nur mit einem Symbol erreicht, daher ist der Zustand hier mit dem jeweiligen Symbol bezeichnet.

2 Aufgabenstellung

Auf Basis sinnvoll gewählter digital vorliegender populärer Musikvideos sollen automatisiert wiederkehrende Muster in Ton und Bild erkannt werden, zu denen nach durchzählender Vergabe von Symbolen [1] eine reguläre Grammatik wie in Abbildung 1 bestimmt wird, die neben dem ursprünglichen Stück auch andere sinnvolle Sequenzen akzeptiert, die wieder digital zusammengesetzt werden können. Eine große Herausforderung liegt dabei in der Segmentierung und der Mustererkennung [2], für die verschiedene Verfahren und Methoden wie Autokorrelation, gefensterte Fouriertransformation, Wavelets und das in der Spracherkennung bewährten und entsprechend inzwischen auch für Musik erfolgreich eingesetzte Mel-Frequency-Cepstrum (MFC) [4] zum Einsatz kommen können. Die Mustererkennung sollte sowohl mit manuell analysierten Strukturen als auch Robustheitsanalysen evaluiert werden und verschiedene automatisiert erzeugte Beispielvideos erzeugt werden.

Literatur

- [1] Dannenberg, Roger B., and Ning Hu. "Pattern discovery techniques for music audio." *Journal of New Music Research* 32.2, 2003.
- [2] Maddage, Namunu C. "Automatic structure detection for popular music." *IEEE Multimedia* 13.1, 2006.
- [3] M. Müller. „Fundamentals of music processing: Audio, analysis, algorithms, applications.“ Springer, 2015.
- [4] Cheng, Yu-Huei, et al. "Automatic Music Genre Classification Based on CRNN." *Eng. Letters* 29.1, 2020.