

Kurzfassung

Beim Filmen von Choreographien zu vorproduziertem Audio können – besonders bei analogem Film – große Ineffizienzen bezüglich Materialverbrauch und Synchronisation entstehen. Zur Effizienzsteigerung untersucht diese Arbeit den Einsatz von Timecode in Kombination mit einem Playbacksystem. Zudem wurde aufbauend auf Vorarbeiten einer digitalen Filmklappe ein solches System unter dem Namen „ARC-Play“ konzipiert und umgesetzt. Um eine praxisnahe Entwicklung zu garantieren, wurden drei Experteninterviews durchgeführt. Diese bestätigten den Bedarf sowie den systemischen und finanziellen Nutzen einer solchen Lösung. Eine Wirtschaftlichkeitsprüfung anhand dreier hypothetischer Szenarien belegte diesen Vorteil vor allem bei aufwendigen analogen Drehs. Im Rahmen der Arbeit wurden sowohl die Hardware als auch die Software eines Prototyps entwickelt. Durch iterative Praxistests wurde dieser so optimiert, dass er ein sequenzielles Playback mit synchronem Timecode wiedergibt und Raw-Clips in der Postproduktion automatisch mit Metadaten versieht. Die Arbeit bestätigt das Potenzial von ARC-Play für choreographierte Produktionen und liefert die Basis für zukünftige Weiterentwicklungen, auch im Bereich klassischer Spielfilme.

Abstract

When filming choreography to pre-recorded audio—especially with analog film—significant inefficiencies can arise in terms of material consumption and synchronization. To improve efficiency, this thesis examines the use of timecode in combination with a playback system. Furthermore, building on preliminary work involving a digital clapperboard, such a system was designed and implemented under the name “ARC-Play.” To ensure a practical development, three expert interviews were conducted. These confirmed the need for such a solution as well as its systemic and financial benefits. A cost-benefit analysis based on three hypothetical scenarios demonstrated this advantage, particularly for complex analog shoots. As part of the thesis, both the hardware and software of a prototype were developed. Through iterative practical testing, it was optimized to provide sequential playback with synchronized timecode and automatically tag raw clips with metadata in post-production. The thesis confirms the potential of ARC-Play for choreographed productions and provides the basis for future developments, including in the realm of traditional feature films.

Inhaltsverzeichnis

ABBILDUNGSVERZEICHNIS	IV
TABELLENVERZEICHNIS.....	V
1 EINLEITUNG.....	1
2 VORARBEITEN	4
3 THEORETISCHE GRUNDLAGEN UND RECHERCHE	7
3.1 SMPTE-TIMECODE	7
3.2 DREHVERHÄLTNIS	10
3.3 ANALYSE BESTEHENDER AUDIOPLAYBACK- UND SYNCHRONISATIONSLÖSUNGEN.....	10
3.3.1 <i>Individuelle Lösungen</i>	10
3.3.2 <i>Tentacle Music Video Sync App</i>	12
3.3.3 <i>Movie Slate 8</i>	13
3.3.4 <i>Lippensynchronisation bei TikTok</i>	14
3.4 FAZIT DER RECHERCHE	15
4 EXPERTENINTERVIEWS.....	16
4.1 METHODIK DER EXPERTENINTERVIEWS.....	16
4.2 AUSWAHL DER INTERVIEWPARTNER.....	16
4.3 INTERVIEW LEITFADEN	17
4.4 INTERVIEW ERGEBNISSE	18
4.5 DISKUSSION DER ERGEBNISSE	23
4.6 TYPOLOGIE VON MUSIKVIDEOPRODUKTIONEN	26
4.7 FAZIT DER EXPERTENINTERVIEWS	29
5 KONZEPT UND ANFORDERUNG VON ARC-PLAY.....	30
5.1 GRUNDLEGENDE FEATURES VON ARC-PLAY	30
5.2 FUNKTIONSWEISE VON ARC-PLAY	31
5.3 WEITERE IDEEN UND FEATURES	33
6 BERECHNUNGEN ZUR FINANZIELLEN ERSPARNIS VON ARC-PLAY	34
6.1 METHODIK DER BERECHNUNGEN	34
6.2 AUFBAU DER BERECHNUNGEN.....	34

6.3	BERECHNUNGEN ANHAND VERSCHIEDENER SZENARIEN	35
6.3.1	<i>Low-Budget-Dreh</i>	36
6.3.2	<i>Komplexer Dreh mit digitaler Kamera</i>	37
6.3.3	<i>Analoger Dreh</i>	38
6.4	DISKUSSION DER ERGEBNISSE DER BERECHNUNGEN.....	40
7	TECHNISCHE UMSETZUNG.....	42
7.1	HARDWAREKOMPONENTEN	42
7.2	ELEKTRONIK ZUR TIMECODE-DEKODIERUNG	45
7.3	PLATINENDESIGN	47
7.4	3D-DRUCK-GEHÄUSE.....	48
8	SOFTWARE	49
8.1	FIRMWARE DES ESP32-S3.....	49
8.2	NODE.JS-SERVER	53
8.2.1	<i>ARC-Studio</i>	54
8.2.2	<i>ARC-Analyzer</i>	58
9	ANWENDUNG UND VALIDIERUNG.....	60
10	FAZIT UND AUSBLICK	63
	LITERATURVERZEICHNIS	65
	ANHANG	71
A	ANHANG: INTERVIEWTRANSKRIPT MIT EXPERTE A.....	71
B	ANHANG: INTERVIEWTRANSKRIPT MIT EXPERTE B.....	74
C	ANHANG: INTERVIEWTRANSKRIPT MIT EXPERTE C.....	78

Abbildungsverzeichnis

ABBILDUNG 1: DIE DIGITALE FILMKLAPPE ARC-SLATE IM EINSATZ	5
ABBILDUNG 2: DAS KONZEPT VON ARC-SLATE	5
ABBILDUNG 3: LTC-DATENWORT EINES FRAMES BEI 24 FPS (SMPTE, 2014, S. 22)	8
ABBILDUNG 4: ÜBERSICHT DER METADATENINTEGRATION VON ARC-PLAY	31
ABBILDUNG 5: KI-GENERIERTE VISUALISIERUNG VON ARC-PLAY MIT GOOGLE NANO BANANA 2	32
ABBILDUNG 6: SCHEMATISCHE DARSTELLUNG DER ELEKTRONIK VON ARC-PLAY	42
ABBILDUNG 7: AUFBAU DER SIMULATION DER VERSTÄRKERSCHALTUNG FÜR DAS TIMECODESIGNAL IN LTSPICE	45
ABBILDUNG 8: SIMULATION DER VERSTÄRKUNG DES TIMECODESIGNALS MIT LTSPICE OHNE SCHMITTTTRIGGER (BLAU: TIMECODESIGNAL, ROT: TIMECODESIGNAL MIT DC-BIAS, GRÜN: VERSTÄRKTES TIMECODESIGNAL)	46
ABBILDUNG 9: DIE BESTÜCKTE PLATINE VON ARC-PLAY	47
ABBILDUNG 10: DAS 3D-OBJEKT DES GEHÄUSES VON ARC-PLAY IN FREECAD	48
ABBILDUNG 11: DIE MENÜSTRUKTUR VON ARC-PLAY	50
ABBILDUNG 12: DER PROTOTYP VON ARC-SLATE IM PLAYBACK-BETRIEB	51
ABBILDUNG 13: DIE BENUTZEROBERFLÄCHE VON ARC-STUDIO	55
ABBILDUNG 14: DAS FORMULAR ZUM ERSTELLEN EINES NEUEN SHOTS IN ARC-STUDIO	56
ABBILDUNG 15: DIE SHOTLIST IN ARC-STUDIO	57
ABBILDUNG 16: DIE BENUTZEROBERFLÄCHE VON ARC-STUDIO MIT GELADENEM PREVIEW	58

Tabellenverzeichnis

TABELLE 1: INFORMATIONEN ÜBER DIE EXPERTEN.....	17
TABELLE 2: DIE ERGEBNISSE DER BERECHNUNGEN UNTER ANNAHME EINES LOW-BUDGET-DREHS.....	36
TABELLE 3: DIE ERGEBNISSE DER BERECHNUNGEN UNTER ANNAHME EINES KOMPLEXEN DREHS MIT DIGITALKAMERA.....	37
TABELLE 4: DIE ERGEBNISSE DER BERECHNUNGEN UNTER ANNAHME EINES DREHS MIT ANALOGEM FILM.....	39

1 Einleitung

In modernen Film- und Medienproduktionen hat die Digitalisierung zu einer erheblichen Effizienzsteigerung geführt. Dennoch existieren in der Praxis veraltete und nicht digitale Workflows, die die Arbeit verlangsamen und fehleranfällig machen.

Besondere Herausforderungen entstehen bei der Produktion von Inhalten, deren visuelle Ebene framegenau mit einem vorproduzierten Audiosignal synchronisiert werden muss, beispielsweise bei Musikvideos, Tanzfilmen oder sonstigen Choreographien. Hierbei ergeben sich zwei Kernprobleme:

Zum einen bieten Kameras, die keine eigene Tonspur aufzeichnen, wie es beispielsweise bei vielen analogen Kinokameras der Fall ist, in der Postproduktion keine Möglichkeit, Bild und Ton auditiv zu synchronisieren. Die Ausrichtung muss mühsam visuell anhand der Choreographie oder durch Filmklappen erfolgen. Letztere müssen allerdings framegenau mit dem Playback am Set übereinstimmen. Gerade wenn das Playback für spezifische Zeitlupeneffekte in unterschiedlicher Geschwindigkeit abgespielt und in der Postproduktion entsprechend beschleunigt oder verlangsamt werden muss, ist das Synchronisieren eine Herausforderung.

Zum anderen ist das Verfahren, den Song am Set in voller Länge durchlaufen zu lassen und kontinuierlich zu filmen, gerade bei analogem Film aufgrund der hohen Materialkosten wirtschaftlich nicht tragbar. Das selektive Filmen einzelner Segmente erfordert jedoch ein zeitaufwendiges, manuelles Suchen und Anspielen der exakten Songpositionen am Set und erhöht durch die hohe Anzahl an Einzeltakes den Aufwand für die Synchronisierung in der Postproduktion enorm.

Aus der beschriebenen Problemstellung, die auch auf eigener Erfahrung beruht, ergibt sich die Zielsetzung dieser Arbeit. Es soll untersucht werden, wie ein Timecode-basiertes Audioplayback-System konzipiert und realisiert werden muss, um die dargestellten Probleme durch Digitalisierung zu lösen und so die Effizienz von Drehs mit Choreographien in Kombination mit vorproduziertem Audio sowohl systemisch als auch finanziell zu steigern.

Daraus leitet sich die Hauptforschungsfrage der Arbeit ab:

Inwieweit lässt sich die Effizienz bei der Produktion von choreographierten, auf vorproduzierten Audioinhalten basierenden Filmen durch den Einsatz eines Timecode-basierten Audioplayback-Systems steigern?

Die Beantwortung dieser Leitfrage erfordert eine Kombination von theoretischer Konzeption und praktischer Überprüfung. Um das System unter realen Produktionsbedingungen testen und validieren zu können, ist dessen hardware- und softwareseitige Umsetzung notwendig. Daher ergibt sich folgende Zusatzfrage:

Wie muss ein Timecode-basiertes Audioplayback-System hardware- und softwareseitig realisiert werden, um die Anforderungen einer dynamischen Set-Umgebung zu erfüllen?

In dieser Arbeit soll die Beantwortung dieser Fragen methodisch umgesetzt werden.

Das Projektziel dieser Masterarbeit ist die Konzeption, die hardware- und softwareseitige Realisierung sowie die Evaluierung eines Timecode-basierten Audioplayback-Systems mit dem Projektnamen *ARC-Play*. Da bereits durch vorherige Projekte grundlegende Strukturen für die Umsetzung eines solchen Verfahrens existieren, wird das System nicht isoliert betrachtet, sondern als funktionale Erweiterung des bereits existierenden Ökosystems integriert.

Das technologische Fundament für diese Arbeit bildet das in vorangegangenen Studienprojekten entwickelte System *ARC-Slate*. Dies ermöglicht mithilfe eines QR-Codes und einer KI-basierten Analysesoftware die vollständige Automatisierung des Workflows mit einer Filmklappe. Näheres wird in Kapitel 2 erläutert.

Das im Rahmen dieser Arbeit zu entwickelnde System *ARC-Play* setzt an dieser Stelle an und erweitert das System um eine zeitgesteuerte Audiokomponente, indem es das synchrone, Timecode-gesteuerte Abspielen von vorproduzierten Audioinhalten am Set übernimmt. Es soll ermöglichen, den kreativen Prozess des Rohschnitts bereits in die Vorproduktion zu verlagern: Da das System exakt protokolliert, welcher Songabschnitt in welchem Take gefilmt wurde, soll in der Postproduktion das Bildmaterial im Anschluss vollautomatisch und lippensynchron an der richtigen Stelle der Audiospur positioniert werden.

Die Beantwortung der Forschungsfragen folgt einem iterativen und anwendungsorientierten Forschungsansatz, der theoretische Grundlagen und Recherche, empirische Erhebung und Prototypenentwicklung miteinander verknüpft.

Zu diesem Zweck werden, nachdem noch einmal ausführlich auf die bisherigen Arbeiten eingegangen wird, zunächst die theoretischen Grundlagen erarbeitet. Hierbei stehen die

Funktionsweisen von SMPTE-Timecode sowie der derzeitige Umgang in der Filmbranche mit den genannten Problemstellungen im Fokus.

Darauf aufbauend werden qualitative Experteninterviews durchgeführt. Diese dienen dazu, die in der Problemstellung identifizierten Defizite in der Realität choreographierter Filmdrehs zu validieren und deren Lösungsansätze dafür darzustellen. Um den Experten einen greifbaren Zugang zur Thematik zu bieten und gezieltes Feedback zu generieren, wird bereits vor den Interviews mit der Entwicklung der Grundstrukturen des Prototyps angefangen. Die Interviews helfen dabei, den konkreten Praxisnutzen zu erörtern und den Anforderungskatalog für ARC-Play zu schärfen.

Die Kombination aus theoretischer Recherche und empirischen Expertenerkenntnissen ergibt das finale Konzept.

Im Anschluss werden Berechnungen hinsichtlich des finanziellen Nutzens durchgeführt. Diese dienen nicht nur der Wirtschaftlichkeitsanalyse, sondern steigern auch das Verständnis des gesamten Workflows.

Im Anschluss wird die technische Umsetzung sowohl von der Hardware als auch der Software von ARC-Play dargestellt.

Im Endteil der Arbeit wird der finale Prototyp von ARC-Play im Rahmen realer Testprojekte in der Praxis evaluiert. Die daraus gewonnenen Erkenntnisse erlauben eine kritische Validierung des Systems und durch iterative Verbesserungsschleifen können weitere Verbesserungsmaßnahmen erörtert und teilweise noch direkt umgesetzt werden.

Im Rahmen der Erstellung dieser Arbeit wurden KI-gestützte Tools zur Unterstützung bei der Strukturierung und Rechtschreibung sowie der Programmierung der Software eingesetzt. Die inhaltliche Verantwortung sowie die wissenschaftliche Validierung aller Ergebnisse liegen jedoch vollumfänglich beim Verfasser.

Aus Gründen der besseren Lesbarkeit wird in dieser Arbeit auf geschlechtsspezifische Mehrfachnennungen sowie Sonderzeichen verzichtet. Alle Personenbezeichnungen und personenbezogenen Hauptwörter gelten gleichermaßen für alle Geschlechter.

2 Vorarbeiten

Wie bereits erwähnt, soll ARC-Play in das bereits existierende ARC-System integriert werden und mit der digitalen Filmklappe ARC-Slate zusammenspielen. Vieles in dieser Arbeit basiert im Ansatz auf den bereits existierenden Strukturen und soll sich darin integrieren. Daher beschreibt das folgende Kapitel ausführlich das Konzept von ARC und ARC-Slate.

In Filmproduktionen dient eine Filmklappe dazu, einzelne Takes mit Daten auszustatten, um sie im weiteren Prozess eindeutig identifizieren zu können und um das Bild und den Ton zu synchronisieren. Vor Beginn der Dreharbeiten wird hierzu eine Shotlist erstellt, die präzise definiert, welche Bildinhalte realisiert werden sollen. Längere Filme können dabei in Szenen und innerhalb dieser Szenen in Shots unterteilt werden - allerdings werden auch Variationen dieser Einteilung genutzt. Jeder einzelne Versuch, einen Shot zu drehen, wird als Take bezeichnet (Bordwell et al., 2016, S. 22–23). Zudem können auf der Klappe spezifische Parameter wie die verwendete Brennweite, die Kamera, die Lichtverhältnisse sowie der Drehort festgehalten werden (Schmidt, 2011, S. 69).

Diese Informationen werden manuell auf die Filmklappe übertragen. In der Praxis beschriftet meist eine Person aus dem Kamera-Department die Klappe oder deckt nicht benötigte Felder mit Klebeband ab. Parallel dazu dokumentiert der Script Supervisor während des Drehs, welche Takes brauchbar waren und bei welchen Fehler oder technische Probleme auftraten. In der Postproduktion werden die Takes anschließend wieder zugeordnet und vorsortiert (Bordwell et al., 2016, S. 23). Moderne, branchenübliche Schnittprogramme wie beispielsweise Blackmagicdesign DaVinci Resolve bieten hierfür vorgefertigte Informationsfelder an, in denen alle für einen Take spezifischen Metadaten eingegeben werden können.

Die manuellen Schritte – sowohl das Beschriften der Klappe mit den Daten aus der Shotlist als auch das spätere Einpflegen in die Schnittsoftware – sind fehleranfällig sowie zeitaufwendig. Dieser Workflow lässt sich durch eine digitale Automatisierung optimieren. An dieser Schnittstelle setzt das in vorangegangenen Semestern realisierte Konzept der digitalen Filmklappe ARC-Slate (siehe Abbildung 1) an: Bis auf die physische Beschriftung der Klappe verlaufen die meisten Prozesse einer modernen Filmproduktion bereits digital. Da auch die Shotlist in den meisten Fällen digital vorliegt, bietet sich daher auch eine Digitalisierung des Workflows mit einer Filmklappe an.



Abbildung 1: Die digitale Filmklappe ARC-Slate im Einsatz
Quelle: eigene Darstellung

Um dies zu erreichen, zeigt ARC-Slate neben weiteren produktionsrelevanten Informationen eines Takes auch einen QR-Code auf einem E-Ink-Display an. In dem QR-Code ist eine eindeutige ID für den spezifischen Take gespeichert. Alle Informationen über den Take, sowohl aus der Vorproduktion, also Szene, Shot, Umgebung und Tageszeit, als auch die während des Drehs erfassten Produktionsnotizen, die auf der Weboberfläche des Servers *ARC* erstellt werden können, werden im Format einer JSON-Datei auf ARC gespeichert. Nach dem Dreh wird das gesamte Material von einer KI-gestützten Software analysiert und nach QR-Codes gesucht. Sobald ein solcher gefunden wurde, wird die enthaltene ID des QR-Codes analysiert und mit den dazu passenden Informationen aus ARC verknüpft. Daraufhin wird eine spezifische Metadatendatei angefertigt, in der die Clips des Filmmaterials mit den Metadaten verknüpft sind und mit der das Schnittprogramm DaVinci Resolve in der Lage ist, alle Informationen anzuzeigen (siehe Abbildung 2).

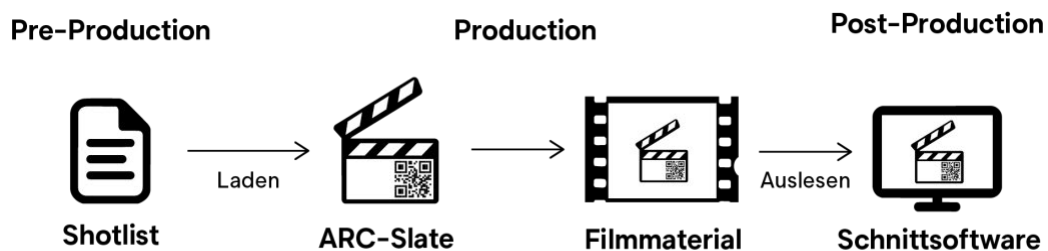


Abbildung 2: Das Konzept von ARC-Slate
Quelle: eigene Darstellung

Neben der Metadatenabwicklung ist die Klappe in der Lage, Free-Run-Timecode zu generieren (siehe Kapitel 3.1). Der Timecode wird benutzt, um Bild und Ton am Set zu synchronisieren. Um den Timecode von ARC-Slate mit weiteren Geräten wie Kameras und Audiorecorder zu synchronisieren, ist ein Audioausgang integriert, der das Timecodesignal ausgibt.

Bei Kameras ohne Timecodeeingang werden Bild und Ton mittels des Klappgeräuschs und des visuellen Ereignisses des Zusammentreffens der Balken händisch synchronisiert (Schmidt, 2011, S. 69). Bei ARC-Slate wurde das Identifizieren des Frames, an dem die Klappe betätigt wurde, ebenfalls automatisiert: Die Klappe verfügt über eine LED, die während des Schließens der Klappe aufblitzt (siehe Abbildung 1). Parallel dazu wird der aktuelle Timecode ebenfalls in ARC abgespeichert. Dieses visuelle Signal kann in der Analysesoftware von ARC detektiert werden. In Kombination mit dem vorliegenden Synchronisations-Timecode lässt sich der Anfangstimecode des Takes berechnen. Dieser wird anschließend direkt in den Video-Clip geschrieben. Zusätzlich kann die Analysesoftware bei analogen Filmproduktionen, bei denen technisch bedingt alle Takes auf digitalisierten Filmrollen hintereinander in einer Datei vorliegen, die einzelnen Takes extrahieren, zuschneiden und mit Metadaten versehen.

Die Filmklappe ARC-Slate bietet somit eine Möglichkeit, Daten im gesamten System digital zu lassen und so gerade für Filmproduktionen mit analogen Filmkameras einen Mehrwert.

Allerdings sind im bestehenden System Probleme vorhanden: Beispielsweise handelt es sich bei dem Server nur um eine lokale HTML-Datei, bei der alle Projektinformationen beim Aktualisieren des Browserfensters wieder neu in das System geladen werden müssen. Außerdem ist der Timecode-Generator der digitalen Filmklappe primitiv und ist nicht in der Lage, Drop-Frames zu erstellen oder externen Timecode zu lesen. Auch die Analysesoftware ist einfach gehalten und besitzt keine Benutzeroberfläche. Diese Nachteile sollten im Zuge der Entwicklung von ARC-Play behoben oder neu entwickelt werden und können im Umkehrschluss auch in einem neuen Prototyp von ARC-Slate Anwendung finden.

3 Theoretische Grundlagen und Recherche

Bevor das Konzept und die technischen Anforderungen für ARC-Play erstellt werden, werden technische Grundlagen dargestellt und bisherige Lösungsansätze der genannten Problemstellungen auf dem Markt recherchiert.

3.1 SMPTE-Timecode

Dem technischen Konzept von ARC-Play folgend soll die Synchronisierung des Audio-Playback-Systems mit Timecode realisiert werden. Für die Filmklappe ARC-Slate wird bereits LTC benutzt, also Linear Timecode, der von der Society of Motion Picture and Television Engineers definiert wird (SMPTE, 2014). Der Timecode besteht aus einer achtstelligen Zeichenkette mit vier jeweils zweistelligen Nummern, die jeweils mit Doppelpunkten getrennt sind. Dabei ist die erste Nummer die Stunde, die zweite die Minute, die dritte die Sekunde und die vierte der Frame beziehungsweise das aktuelle Einzelbild. Die maximale Nummer des Frames ist von der verwendeten Framerate abhängig. Dabei sind laut SMPTE für LTC 60, 59,94, 50, 48, 47,95, 30, 29,97, 25, 24 und 23,98 Frames pro Sekunde zulässig.

Da der Timecode ausschließlich mit ganzzahligen Bildern arbeitet und nicht wie das in den USA verwendete NTSC-System mit einer Bildfrequenz von 29,97, gibt es den Drop-Frame-Modus. Dabei werden zu jeder vollen Minute die ersten beiden Frames ausgelassen, außer zu jeder zehnten Minute (Schmidt, 2021, S. 708). Bei einer Framerate von 23,98 Bildern pro Sekunde werden keine Drop-Frames verwendet. Es wird rechnerisch mit konstanten 24 Bildern pro Sekunde gearbeitet, was bei längeren Laufzeiten zu einer kontinuierlich ansteigenden Asynchronität im Vergleich zur tatsächlichen Uhrzeit führt, die sich nach einer Stunde bereits auf 3,6 Sekunden beläuft (SMPTE, 2014, S. 9).

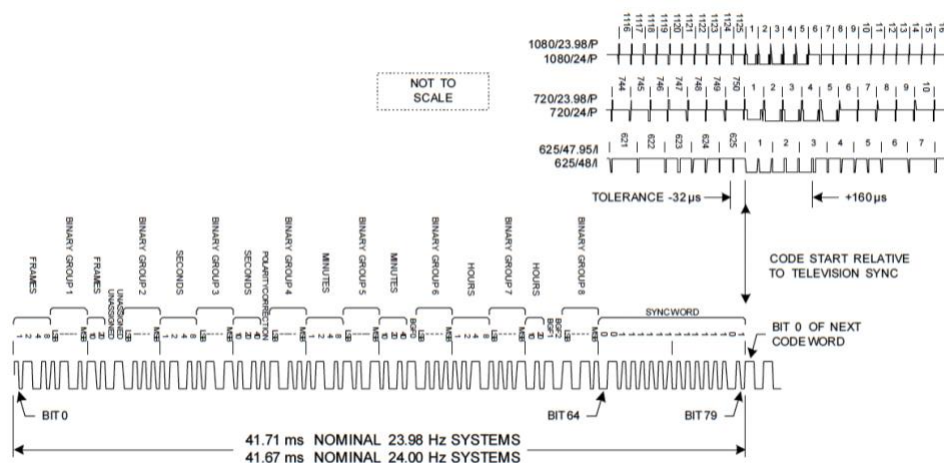


Abbildung 3: Das LTC-Datenwort eines Frames bei 24 FPS (SMPTE, 2014, S. 22)

Wie in Abbildung 3 veranschaulicht, wird Linear Timecode als Datensignal von 80 Bit übertragen. Die insgesamt acht Ziffern werden dabei mit jeweils vier Bits dargestellt. Da die Stundenwerte das Maximum von 23, die Minuten- und Sekundenwerte den Wert von 59 sowie die Frame-Ziffern den Wert von 29 definitionsgemäß nie überschreiten, werden die im 4-Bit-Binärkode nicht benötigten Datenbits als Kontrollbits beziehungsweise Flags deklariert. Diese Flags kennzeichnen beispielsweise, ob das Signal im Drop-Frame-Modus vorliegt. Um die Framenummer bei einer Framerate von über 30 Bildern pro Sekunde darstellen zu können, werden dafür zwei Frames zusammengefasst. Um die Bilder des Framepaars anschließend zu unterscheiden, wird das Field-Flag verwendet (SMPTE, 2014, S. 33–34).

Neben den Timecode-Ziffern werden zudem nochmal vier Ziffern mit je acht Bit als User-Bits für Zusatzinformationen übertragen. Diese sind frei belegbar und werden nach ASCII codiert. Neben den somit resultierenden 64 Informationsbits werden noch 16 Synchronisationsbits übertragen. Dabei werden zwölf gleiche aufeinanderfolgende logische Einheiten verwendet, die in keiner Variation aus den anderen Informationsbits auftreten können. Dadurch wird das Ende des Codeworts signalisiert. Vor und nach dem Synchronisationsblock sind noch weitere Bits vorhanden, die die Richtung des Codeworts zeigen, also ob vorwärts oder rückwärts gelesen wird (Schmidt, 2021, S. 708–710).

Die Daten werden als serieller Bitstrom nach der Bi-Phase-Mark-Codierung codiert. Dieses zeichnet sich durch drei Regeln aus: Bei jeder Bitzellengrenze gibt es eine Pegeländerung, eine logischen 1 generiert nach der Hälfte des Bits eine weitere Pegeländerung und bei einer logischen 0 gibt es innerhalb der Bitzelle keine Pegeländerung (SMPTE, 2014, S. 17). Beispielsweise ergibt sich bei 25 Bildern pro Sekunde eine Bilddauer von 40 Millisekunden. Somit ergibt sich durch das Bi-Phase-Mark-Coding bei 80 Bit pro Frame bei einer logischen 1 mit zwei Pegelwechseln pro Bit eine maximale Frequenz von 2 kHz, bei einer logischen 0 mit nur einem Pegelwechsel pro Bit ergibt sich eine minimale Frequenz von 1 kHz. Diese Frequenzen liegen exakt im Spektrum von vom Menschen hörbaren Audiosignalen, weshalb LTC oft als Audiosignal vorliegt (Schmidt, 2021, S. 710). Die Amplitude sollte zwischen 1 V und 2 V von Signalspitze zu Signalspitze liegen, es sind aber Amplituden von 0,5 V und 4,5 V zulässig (SMPTE, 2014, S. 19).

In der Praxis wird zwischen zwei unterschiedlichen Arten von Timecode unterschieden: Zum einen gibt es Free-Run-Timecode und zum anderen Record-Run-Timecode. Bei Ersterem generiert das System einen Timecode, der mit der aktuellen Uhrzeit synchron läuft

und bei Realzeitaufzeichnungen verwendet wird. Dieser wird meistens bei Dokumentationen verwendet. Bei Record-Run steigt der Timecode nur an, wenn aufgenommen wird. Der Timecode wird unterbrochen, wenn die Aufnahme unterbrochen wird. Dadurch liegt am Ende eine Sequenz mit stets ansteigendem Timecode vor (Schmidt, 2021, S. 712).

Um auf einem Filmset mehrere Geräte mit Timecode zu synchronisieren, existieren verschiedene Methoden. Eine Option sieht vor, alle Geräte auf einem Filmset, die timecodefähig sind, mit einem externen Timecode-Generator, dem Timecode-Master, zu verbinden. Das kann entweder per Kabel oder drahtlos geschehen. Bei Letzterem werden dafür kleine Boxen verwendet, wie zum Beispiel das System von Tentacle Sync (Tentacle Sync GmbH, 2025). Eine alternative Option stellt das Jam-Sync-Verfahren dar. Hierbei werden alle Timecode-fähigen Geräte nur temporär mit einem Timecode-Generator verbunden. Die Geräte lesen den externen Timecode und generieren ihn ab diesem Zeitpunkt für sich selbst (Cipher Digital Inc., 1987, S. 31).

Es ist von elementarer Bedeutung, dass Timecode keine feste Synchronisation beinhaltet. Es handelt sich lediglich um Metadaten, die einen bestimmten Frame beschreiben. Wenn ein externer Timecode-Generator angeschlossen ist, wird der Timecode des Generators nur für den ersten Frame verwendet. Sobald eine digitale Kamera oder ein Soundrecorder anfangen, aufzuzeichnen, nutzen sie ihre internen Taktgeber. Das kann dazu führen, dass die Geräte bei langen Aufzeichnungen von über 30 Minuten nicht mehr synchron sind. Um dem entgegenzuwirken, werden neben Timecode auch Genlock und Wordclock verwendet. Genlock wird verwendet, um Videosignale synchron aufzuzeichnen oder synchron abzuspielen. Das ist gerade bei Liveproduktionen oder 3D-Kameras essenziell. Bei Wordclock handelt es sich um ein vergleichbares Signal, allerdings spezialisiert für Audiorecorder. In der Praxis würde man eine Masterclock verwenden, die den Timecode und den Takt für das ganze Filmset liefert, und es werden sowohl die Kameras als auch die Recorder mit Timecode und mit Genlock bzw. Wordclock verbunden (Rapp, 2020). Für kürzere Sequenzen ist es allerdings ausreichend, nur mit Timecode zu arbeiten.

Um Timecode in digitalen Dateien zu speichern, kann man, neben LTC als Audiotrack, auch den Timecode in den Metadaten einer Videodatei speichern. Beispielsweise wird bei der MPEG-Videocodierung die Information in den Header des Picture-Layers geschrieben (Schmidt, 2021, S. 226).

Neben Videofiles kann man auch Audiofiles mit Timecode synchronisieren. Das BWF-Format ist eine WAV-Datei, die um eine Broadcast Audio Extension erweitert wurde (Schmidt, 2021, S. 767). Dabei handelt es sich um eine Erweiterung, in der mehr Metadaten gespeichert werden können. Unter anderen kann dort eine Zeitreferenz gespeichert werden. Dabei handelt es sich um einen 64-Bit-Wert, in dem die Samples seit der letzten vergangenen Mitternacht, also dem Zeitpunkt 00:00:00:00, gespeichert werden. Dieser Wert ist daher abhängig von der verwendeten Sample-Rate. Allerdings handelt es sich bei der Zeitreferenz lediglich um einen Startwert (European Broadcasting Union, 2011). Alle anderen Werte werden aus den vergangenen Samples berechnet.

3.2 Drehverhältnis

Das Drehverhältnis oder auch Shooting Ratio ist das Verhältnis zwischen der gesamten Dauer des gedrehten Materials und der finalen Laufzeit des fertigen Films oder Videos. Der Grund für ein erhöhtes Drehverhältnis ist das Aufnehmen mehrerer Takes des gleichen Shots, der Einsatz von Multikamera-Setups, bei dem dieselbe Szene aus unterschiedlichen Perspektiven gleichzeitig aufgezeichnet wird, oder das vollständige Weglassen einzelner Szenen im Schnitt. In der klassischen Ära von Hollywood lag dieses Verhältnis standardmäßig bei etwa 1:10. Dies änderte sich grundlegend, als sich die digitale Filmkamera etablierte. Inzwischen sind Drehverhältnisse von 1:50 bis 1:100 die Norm. Der Film *Deadpool* (2016) verzeichnete beispielsweise ein Drehverhältnis von 1:308 (Sharman, 2020).

3.3 Analyse bestehender Audioplayback- und Synchronisationslösungen

In der Praxis choreographierter Drehs existieren wenige kommerzielle Produkte, die die Problemstellung dieser Arbeit gezielt beheben. Etablierte Verfahren basieren meist auf individuellen Workflows und weisen keine Standardisierung auf. Dennoch lassen sich vereinzelt Ansätze und Produkte identifizieren, die versuchen, die Effizienz choreographierter Drehs zu steigern.

3.3.1 Individuelle Lösungen

Während der Recherche ist schnell aufgefallen, dass es für Playbacksysteme beim Filmen von Drehs mit vorproduziertem Audio und dem Synchronisieren keine Standardverfahren gibt. Vielmehr erstellen sich viele Produzenten eigene Tools, um einen möglichst effizienten Workflow zu erreichen.

Eine Möglichkeit ist, vor dem Dreh einen sogenannten Edit-Master herzustellen. Hierbei handelt es sich um eine mehrspurige Audiodatei, bei der das vorproduzierte Musikstück mit einem Linear Timecode kombiniert wird und sich synchronisiert auf separaten Spuren der Datei befindet. Meist werden dafür der Timecode und das Musikstück jeweils als Mono-WAV-Datei erstellt und in einer Stereo-Datei kombiniert. Um bestimmte vordefinierte Bereiche des Musikstücks, beispielsweise Strophe oder Refrain, einzeln aufzuzeichnen, können mehrere Edit-Master erstellt werden, die entsprechend zugeschnitten sind. Während des Drehs wird das Musikstück auf einem Lautsprechersystem wiedergegeben, damit die Darsteller damit interagieren und performen können. Meist gibt es noch eine gewisse Vorlaufzeit, um dem Produktionsteam und den Darstellern die nötige Vorbereitung für den kommenden Take zu ermöglichen. Das parallel dazu abgespielte Timecodesignal wird dann mit einer Timecode-fähigen Klappe oder einem Display mit Timecode-Eingang verbunden, damit diese den aktuellen Timecode anzeigen können, um so dem Bild die zeitliche Referenz zu übergeben (Ginsburg, 2024). Bei Timecode-fähigen Kameras kann das Signal auch in den Timecode-Eingang geführt werden.

Bei professionellen Filmproduktionen übernimmt die Verwaltung des Audio-Playbacks eine spezielle Fachkraft, der sogenannte Playback-Engineer. Dieser konfiguriert den Edit-Master und steuert dessen Wiedergabe auf einem Set mit beispielsweise einer DAW (Janney, 2018).

Alternativ zu den fest vorgerenderten Audiodateien wird auch spezialisierte Software wie der Timecode-Player von Haute Technique Tools verwendet. Diese ermöglicht es, ein lineares Timecodesignal dynamisch in Echtzeit an eine Audiodatei zu koppeln und beide Signale synchron über dedizierte Hardware-Schnittstellen auszugeben (Haute Technique, o. J.).

In besonderen Situationen wird das Musikstück für das Playback modifiziert. Wenn bei Drehs von Spielfilmen eine Dialogszene gedreht wird und im Hintergrund Menschen zur Musik tanzen, darf das Playback nicht über die Mikrofone, die den Dialog aufzeichnen, zu hören sein. Dafür wird ein sogenannter Thumb-Track erstellt. Dabei handelt es sich um ein Playback, das nur Frequenzen unterhalb von 50 Hz besitzt und das über einen Basslautsprecher abgespielt wird. Im Audio der Mikrofone werden alle Frequenzen unter 50 Hz entfernt, wodurch der Thumb Track nicht mehr aufgezeichnet wird. Somit können sowohl die Darsteller im Hintergrund zum Takt der Musik tanzen und der Dialog bleibt verständlich (Janney, 2018).

Eine weitere Lösung stammt vom finnischen Filmproduzenten Sami Sänpääkilä (Sänpääkilä, 2019). Für seine Lösung stand eine visuelle Darstellung sowohl des Songs in der Waveformansicht zur besseren Übersicht als auch des Timecodes im Vordergrund. Außerdem wurde der Song in Marker unterteilt, um so die einzelnen Stellen im Song zu definieren. Mit der Schnittsoftware Adobe Premiere Pro wurde dafür ein Playbackvideo erstellt, auf dem auf der oberen Hälfte der Timecode und auf der unteren Hälfte die Waveformansicht mit der aktuellen Cursorposition und den Markern zu sehen war. Somit sind an jeder Stelle der Song und der Timecode im Video synchron und die Sequenz des Songs kann für jeden Take des Videos intuitiv anhand der visuellen Darstellung abgespielt werden. Für Zeitlupeneffekte kann der Song mit dem VLC-Mediaplayer abgespielt werden, der verschiedene Abspielgeschwindigkeiten zulässt.

Der australische Filmproduzent Dan Graetz (Graetz, 2022) hat ebenfalls mehrere Verfahren vorgestellt, um das Playback in der Postproduktion zu synchronisieren. Die erste Lösung war ebenfalls, ein Timecodevideo mit hinterlegtem Audioplayback zu erstellen. Daran wurde ein drahtloser Audiosender angeschlossen, der das Signal an einen entsprechenden Audioempfänger sendet, der wiederum an einen Lautsprecher angeschlossen war. Ein weiteres Verfahren war eine Funktion von einem motorisierten Slider, dem camBLOCK Adventure. Dieser besitzt die Möglichkeit, das Timecodevideo via Netzwerk auf einem VLC-Mediaplayer parallel zu zuvor programmierter Bewegung des Sliders abzuspielen. Durch die immer gleiche Kamerafahrt kann so der Effekt erzeugt werden, dass mehrmals die gleiche Person auf einer Kamerafahrt zu sehen ist, die alle synchron zum Playback performen. Als zusätzliches Feature bietet das System eine Funktion, die Abspielgeschwindigkeit zu variieren. Dadurch wird nicht nur das Playback in unterschiedlicher Geschwindigkeit getriggert, sondern auch die Kamerafahrt unterschiedlich schnell abgespielt. Somit sind Filmaufnahmen möglich, bei denen mehrere Personen mit unterschiedlichen Zeitlupen- und Zeitraffereffekten versehen in einer Kamerafahrt synchron zur Musik performen können.

3.3.2 Tentacle Music Video Sync App

Ein Produkt, das die Synchronisierung von Playback-Audio und -Video ermöglicht, ist die Tentacle Music Video Sync App. Diese befindet sich derzeit in der Betaphase und ist nur für Android verfügbar. Tentacle Sync GmbH ist vor allem für ihre Timecode-Synchronisationsboxen Tentacle Sync E und Tentacle Track E bekannt. Diese fungieren als Timecode-Generatoren und synchronisieren ihre inneren Uhren via Bluetooth. Alle

Konfigurationen der Hardware lassen sich über eine dazugehörige App vornehmen (Tentacle Sync GmbH, 2025).

Die Tentacle Music Video Sync App erweitert den Funktionsumfang: In dieser speziellen App wird ein Audiofile mit einem festen Timecode vereint. Anders als bei den Lösungen aus dem vorherigen Kapitel, wird beim Abspielen des Stücks der Timecode direkt über die via Bluetooth synchronisierten Geräte Tentacle Sync E oder Tentacle Track E abgespielt. Alle Kameras, die mit den Empfängern verbunden sind, erhalten somit direkt den korrekten Timecode. Das Audio wird über die App abgespielt, weshalb das Smartphone mit Bluetooth oder mit Audiokabel mit einem Lautsprechersystem verbunden werden muss. Auch hier kann, bevor das Playback startet, ein Vorlauf definiert werden. Es können zudem In- und Out-Punkte gesetzt werden (Richter, 2025; Tentacle Sync GmbH, o. J.).

3.3.3 Movie Slate 8

Bei Movie Slate 8 handelt es sich um eine von PureBlend Corporation entwickelte App für iOS. Die Software fungiert als virtuelle Filmklappe und emuliert sämtliche Funktionen einer klassischen Filmklappe, wie das Anzeigen des aktuellen Shots oder Timecodes und weiterer Informationen über die Produktion.

Ein wesentliches Merkmal ist allerdings der integrierte Musikvideo-Modus. Dieser erlaubt es, Audiodateien, die sich in der iTunes-Mediathek befinden, in die Applikation zu importieren und diese parallel zu dem zuvor definierten Timecode abzuspielen. Außerdem können für bestimmte Teile des Musikstücks Marker definiert werden. Um nun eine bestimmte Sequenz des Musikstücks abzuspielen, werden Anfangs- und Endtimecodes definiert. Das Auslösen der virtuellen Klappe erfolgt durch das Berühren der digitalen Klappenbalken auf dem Touchscreen. Analog zu den zuvor dargestellten Systemen kann auch hier eine bestimmte Vorlaufzeit eingestellt werden.

Ab dem Startpunkt spielt die App die gewünschte Stelle des Songs ab und zeigt den dazugehörigen Timecode auf dem Display des Geräts an. Diese Einstellungen lassen sich auch direkt mit den Metadaten eines Shots verknüpfen. Sobald der Endtimecode erreicht wurde, gilt ein Take des Shots als abgeschlossen und die Takenummer erhöht sich automatisch.

Der Song kann über den Audioausgang des Geräts abgespielt werden. Dabei muss definiert werden, welcher Kanal des Audioausgangs den Timecode beinhalten soll. Es wird empfohlen, den zu verwendenden Song als Monomischung anzulegen und einen der beiden

Stereokanäle komplett stummzuschalten, damit der Timecode nicht überlagert wird. Alternativ kann der Song auch über AirPlay abgespielt werden. Es gibt auch die Möglichkeit, über eine Erweiterung verschiedene Geräte zu synchronisieren. Wenn auf einem weiteren Gerät ebenfalls der Song vorhanden ist, wird der Song dort ebenfalls parallel abgespielt (PureBlend Corporation, o. J.).

3.3.4 Lippensynchronisation bei TikTok

Obwohl die Social-Media-Plattform TikTok primär dem Consumer-Bereich zuzuordnen ist und auf den ersten Blick wenig mit professionellen Musikvideoproduktionen gemeinsam hat, bietet sie eine Funktion, welche die Synchronisation von Playback zu vorproduzierten Audio ermöglicht. Die als Lip-Sync bezeichnete Funktion basiert auf der Vorgänger-App Musical.ly, die 2017 von DanceByte, der Firma hinter TikTok, übernommen wurde.

Das Prinzip der Funktion ist sehr simpel gehalten: Anwender wählen einen Song aus und entscheiden sich für eine Sequenz innerhalb des Songs, die eine Dauer von 15 und 60 Sekunden hat. Daraufhin kann ein Video aufgezeichnet werden, während der Song vom Endgerät abgespielt wird. Um die Lippensynchronität zu perfektionieren, können das Audio und das Video im Nachhinein angepasst werden und es kann mit Effekten, wie beispielsweise Geschwindigkeitsänderungen, gearbeitet werden (Geyer, 2024, S. 43–50). Ergänzend besteht die Option, den Text live auf dem Display anzuzeigen (Galán, 2024).

Die Relevanz dieser Plattform für Musikvideos resultiert aus einer grundlegenden Änderung des Rezeptionsverhaltens. Jugendliche nutzen vermehrt Plattformen wie TikTok und Instagram, um mit Musik zu interagieren, indem sie dort Musik entdecken, selbst in Kurzvideos nutzen und rezipieren (Geyer, 2024, S. 61). Infolgedessen gewinnt die Erstellung von Videos speziell für diese Plattformen gegenüber dem klassischen Musikvideo an Bedeutung. Für Kunstschaffende bietet dies die Möglichkeit, den produktionellen und finanziellen Aufwand zu reduzieren und stattdessen die Verbreitungsmechanismen der sozialen Medien zur Reichweitensteigerung zu nutzen.

Die Bedienoberfläche von TikTok ist äußerst intuitiv gestaltet, um es Laien ohne Filmproduktionserfahrung zu ermöglichen, choreographierte Videos mit vorproduziertem Audio synchron zu produzieren. Allerdings ist das System auf die Hardware des verwendeten Endgeräts limitiert.

3.4 Fazit der Recherche

Keine der dargestellten Lösungen können die in Kapitel 1 aufgestellten Problemstellungen mit Zufriedenheit lösen.

Alle Ansätze bieten eine Lösung für das Synchronisieren von Playback und Timecode. Allerdings bieten wenige die Flexibilität bei der Geschwindigkeit und beim Startzeitpunkt. Beim Edit-Master-Verfahren sind spontane Änderungen mit vorgerenderten Audiodateien kaum realisierbar. Ein Workflow mit einem Audioengineer wirft, gerade für Low-Budget-Produktionen, eine finanzielle Hürde auf. Die vorbereiteten Timecode-Videos sind für viele Situationen eine gute Lösung, allerdings muss dafür in der Postproduktion immer noch der Timecode manuell ausgelesen werden und die Zuordnung der einzelnen Clips ist nicht zwingend gegeben. Die Tentacle Music Video Sync App ist nur in der Beta-Version verfügbar und es könnten sich – gerade mit der Verwendung von Bluetooth-Verbindungen zum Abspielen des Songs und den damit einhergehenden Latenzunterschieden – Synchronisationsprobleme ergeben (siehe Kapitel 4.4).

Movie Slate 8 bietet funktional die ausgereifteste Lösung für einen Musikvideoworkflow, die nicht nur die Synchronisation ermöglicht, sondern auch die Struktur einer Filmklappe in das automatische Abspielen des Playbacks integriert. Allerdings handelt es sich dabei um eine App und nicht um dedizierte Filmhardware, die ausschließlich dafür entwickelt wurde. Gerade das Aufsplitten des Audioausgangs mit dem Timecodesignal auf dem einen und das Playback auf dem anderen Kanal zeigt, dass ein Consumergerät für solche Anwendungen nicht primär gestaltet ist. Zudem bietet Movie Slate 8 keine Möglichkeit, den Song in unterschiedlicher Geschwindigkeit abzuspielen. TikTok richtet sich an eine andere Zielgruppe und bietet durch die Hardwarelimitierung keine Lösung für Einsatzmöglichkeiten bei professionellen Filmdrehen.

4 Experteninterviews

Nach umfassender Recherche in Kapitel 3 ist festzustellen, dass Drehs mit vorproduziertem Ton auf sehr unterschiedliche Weise umgesetzt werden können. Um nun auch herauszufinden, welche Lösungsansätze exemplarisch in der Praxis verwendet werden und inwiefern eine Gesamtlösung wie ARC-Play einen Mehrwert bieten kann, wurden Experteninterviews mit diversen Produzenten von Musikvideos geführt.

4.1 Methodik der Experteninterviews

Die Interviews wurden online über Zoom abgehalten, aufgezeichnet und im Anschluss mit dem Sprachmodell Google Gemini 3.1 Pro vortranskribiert. Am Ende wurden die Transkripte vollständig überarbeitet, sodass sie inhaltlich die Gespräche korrekt wiedergeben. Um die Ergebnisse der Interviews zu strukturieren, wurden die Transkripte, angelehnt an die qualitative Inhaltsanalyse nach Mayring, analysiert und mit dem dafür vorgesehenen Webtool Qcamap.org induktiv nach Kategorien codiert (Mayring, 2015, S. 120–122). Dabei erfolgte zunächst eine Erstellung von Rohkategorien, die in einem zweiten Analysedurchgang durch Abstraktion und Reduktion zu übergeordneten Hauptkategorien zusammengefasst wurden, um die Kernaussagen zu verdichten. Die Ergebnisse der Hauptkategorien werden in Kapitel 4.4 ohne weitere Interpretation dargestellt und im Anschluss in Kapitel 4.5 in Bezug auf ARC-Play diskutiert.

Da lediglich drei Experteninterviews geführt wurden, sind die Aussagen keinesfalls als allgemeingültig zu betrachten und bieten keine statistische Repräsentativität. Stattdessen liegt der Fokus dieser Untersuchung auf dem Sammeln praxisrelevanter Anforderungen.

4.2 Auswahl der Interviewpartner

Das Interviewverfahren mit der Bezeichnung Experteninterview zeichnet sich dadurch aus, dass über ein semistrukturiertes Leitfadeninterview mit einem Experten Informationen erhalten werden. Als Experten werden dabei Personen bezeichnet, die durch spezielle Tätigkeiten oder Funktionen spezielles Wissen in einem spezifischen Bereich erlangt haben und das nicht Teil des Allgemeinwissens ist (Misoch, 2019, S. 119–120). Im Rahmen dieser Arbeit soll der Workflow einer Filmproduktion von choreographierten Filmen dargestellt werden. Daher werden Produzenten interviewt, die choreographierte Filme – wobei es sich am häufigsten um Musikvideos handelt – herstellen (siehe Tabelle 1).

Allerdings unterscheiden sie sich von den vorhandenen Budgets und vom Arbeitsort. Die Experten wurden sowohl über eine gezielte Interviewrecherche als auch über Kontakte der HAW Hamburg identifiziert und via E-Mail kontaktiert.

Tabelle 1: Informationen über die Experten
Quelle: eigene Darstellung

Anhang	Primäre Rollen	Standort	Referenzen (Auszug)
A	Produzent, DoP	Wien	Bibiza, KEROSIN95
B	Geschäftsführung, Produzent	Berlin	01099, K.I.Z., SSIO, Apsilon, Kraftklub
C	Produzent, DOP	Hamburg	Disarstar, NZIRIA

4.3 Interview Leitfaden

Der Interviewleitfaden wurde nach den vier Phasen nach Misoch konzipiert (Misoch, 2019, S. 68–69). Diese starten mit der Aufwärmphase beziehungsweise Informationsphase. Dabei soll der Interviewpartner über das Projekt und die Zielsetzung informiert werden. Da zum Zeitpunkt des Interviews bereits an Vorarbeiten für den Prototypen von ARC-Play gearbeitet wurde, wurde das bisherige Konzept der praktischen Arbeit vorgestellt. Außerdem wird von eigenen Erfahrungen mit Musikvideodrehen, insbesondere mit analogem Film, berichtet. In der nächsten Phase, der sogenannten Aufwärmphase, werden lockere Fragen zum Einstieg gestellt, um den Interviewten auf die Interviewsituation einzustimmen. Es sollen offene und breite Fragen gestellt werden, die beim Experten einen gewissen Redefluss hervorrufen sollen. Konkret wurden Fragen über die Erfahrung bei bisherigen Musikvideodrehen und über das verwendete Equipment, vor allem hinsichtlich des Playbacksystems am Set, gestellt. In der darauffolgenden Hauptphase sollen die wichtigsten Fragen gestellt werden. Sie ist in verschiedene Themenbereiche gegliedert. Da es um den Workflow der Musikvideoproduktion geht, werden die Fragen thematisch chronologisch von der Vorproduktion über den Dreh mit Fokus auf das verwendete Playback und den Timecode-Workflow bis hin zur Postproduktion abgehandelt. Außerdem gibt es am Ende der Hauptphase noch Fragen über das Nutzen von Metadaten und über Finanzen. In der letzten Phase, der Abschlussphase, werden noch einfachere und weiterführende Fragen gestellt. In diesem Fall werden die Interviewpartner gefragt, ob sie selbst noch Ideen für ARC-Play

haben oder ob sie noch an anderen Stellen Defizite in der Effektivität eines Musikvideodrehs sehen.

Um einen natürlichen Gesprächsfluss zu ermöglichen, soll der Leitfaden flexibel gehandhabt werden und nicht strikt nach der Reihenfolge abgearbeitet werden. Dadurch soll es dem Interviewten ermöglicht werden, auf Themen spezifischer einzugehen und auch Informationen preiszugeben, die durch Fragen des Leitfadens nicht aufgekommen wären.

4.4 Interview Ergebnisse

Im Folgenden werden die Kernaussagen der Interviews dargestellt. Dabei werden die Kategorien, die durch das an Mayring angelehnte Verfahren der qualitativen Inhaltsanalyse beim Codieren der Interviewtranskripte induktiv erstellt wurden, systematisch durchgegangen.

Aktueller Workflow mit Playbacksystemen und Synchronisierung

Laut den Experten wird in der Praxis am häufigsten ein Bluetooth-Lautsprecher verwendet, der mit einem digitalen Endgerät wie einem Smartphone oder Tablet verbunden wird und über den der Song des Musikvideos, der teilweise noch in nicht gemasterter Form vorliegt, abgespielt wird. „Ganz am Anfang haben wir es klassisch gemacht: Wir hatten eine Bluetooth-Box dabei und haben uns auf das interne Mikrofon der Kamera verlassen“ (Anhang B).

Das Synchronisieren in der Postproduktion unterscheidet sich je nach verwendeter Technik und Drehsituation: „Das Synchronisieren in Premiere über den Autosync funktioniert bei Digitalaufnahmen eigentlich so gut, dass man sich da nicht zu viel Arbeit machen muss“ (Anhang A).

Allerdings hat ein Interviewpartner dabei die Erfahrung gemacht, dass am Set die Qualität des Audios oft sehr schlecht sein kann oder dass es durch besondere Drehsituationen mit beispielsweise größeren Distanzen, wenn sich die Kamera und die Darsteller jeweils in einem fahrenden Auto befinden, kaum zu hören war. Um das zu lösen, hat der Experte mit Bluetooth-Transmittern und -Receivern das Signal aufgeteilt und zum einen den Sound an den Bluetooth-Lautsprecher gesendet und zum anderen direkt in den Audioeingang der Kamera geleitet, um so das Audio des Songs direkt mit der Videodatei zu verknüpfen. Leider ergaben sich dabei Probleme mit der Bluetooth-Verbindung: „[...] da kam das Problem mit der Latenz. Es hat nie funktioniert, weil die Latenz je nach Entfernung nicht nur da, sondern auch noch schwankend war“ (Anhang B).

Ein weiterer Lösungsansatz, der in der Produktionsfirma desselben Experten angewendet wurde, ist das Erstellen eines Timecode-Videos: „Wir packen den Song kurz in Premiere oder Resolve und generieren einen Timecode als Video. Dann hält man am Set einfach ein iPad oder Handy mit diesem Video in die Kamera. So sieht man im Bild: ‚Minute 02:25, 15 Frames‘. So kann man es manuell synchronisieren“ (Anhang B). Bei komplexeren Drehs werden zudem mehrere Versionen dieses Videos mit verschiedenen Start- und Endpunkten oder mit verschiedenen Geschwindigkeiten für besondere Zeitlupeneffekte erstellt.

Die anderen Produzenten nutzten auch bei Drehs mit Kameras, die kein Audio aufzeichnen können, das herkömmliche System mit einem Bluetooth-Lautsprecher. Sie berichteten daher von einem erheblichen Aufwand in der Postproduktion: „Da legen wir den Ton tatsächlich manuell nach den Lippenbewegungen an. Das ist mühsame ‚Friemelarbeit‘, aber das übliche Prozedere“ (Anhang C). Ein anderer Experte meinte: „Oft ist der Ton am Set so schlecht, dass ein automatischer Sync per Timecode für viele ein riesiger Mehrwert wäre. Das Anlegen und Synchronisieren will niemand händisch machen“ (Anhang B).

Ein Experte erwähnte, dass, neben dem originalen Song, für den das Musikvideo gedreht wird, manchmal auch aus gestalterischen Gründen andere Songs abgespielt werden. „Manchmal passt der Song auch visuell nicht zum Vibe, den man erzeugen will, dann nehmen wir einen anderen Song mit der gleichen BPM-Zahl“ (Anhang B).

Aktueller Planungsaufwand für Musikvideodrehs

In der Frage, wie genau die Drehs geplant sind beziehungsweise ob bereits in der Vorproduktion bekannt ist, welche Szenen an welche Stelle des fertigen Songs zu sehen sein sollen, antworteten alle drei Experten sehr ähnlich: Der Planungsaufwand ist sehr von der Art des Projekts und der verwendeten Technik abhängig. „Bei Analogfilm wird sehr genau geplant, allein wegen der Kosten. Bei Digitaldrehs ist es eher egal, da hat man zur Not ja auch noch den Kamerasound als Referenz. Aber bei Analog wissen wir am Set genau: ‚Dieser Take ist für diesen speziellen Part im Song.‘ Besonders bei Choreographien, wo die Moves exakt auf den Beat passen müssen, ist eine genaue Planung unerlässlich“ (Anhang A).

Es wurde von weiteren Aspekten berichtet, die die Effizienz solcher Drehs steigern: „Je erfahrener die Leute sind, desto genauer teilt man das auf. Am Anfang dreht man oft noch den ganzen Song an jeder Location durch. Später weiß man genau: ‚Hier brauche ich nur diesen Satz‘. Man schneidet den Film quasi schon im Kopf vor. In der Shotlist steht dann

zum Beispiel für Shot 5B: ‚Spiel Datei 5B ab‘ und da ist dann wirklich nur dieser Part des Songs drin“ (Anhang B).

Außerdem kann sich das Genre auf den Planungsaufwand auswirken: „Bei Techno-Videos kann man im Schnitt wilder schieben als bei klaren Lyrics“ (Anhang C).

Geschwindigkeitsveränderung als kreatives Mittel

Alle drei Interviewpartner bestätigten, dass sie Erfahrung mit Musikvideodrehs mit unterschiedlicher Geschwindigkeit haben. Dabei wird am Filmset der Song schneller abgespielt und mit hoher Bildwiederholungsrate gefilmt, um im fertigen Video eine Zeitlupensequenz, die dennoch mit dem Ton des Audios synchron ist, zu erhalten. Für die Künstler ist das allerdings eine Herausforderung: „Bei 1,25- oder 1,5-facher Geschwindigkeit wird das oft schon sehr schwierig für die Künstler“ (Anhang B). Ein Experte war außerdem Teil eines Drehs, bei dem der Song rückwärts gedreht wurde (Anhang C).

Mehrkamerasysteme

Die Befragten haben zum Großteil keine Erfahrung mit Mehrkamerasystemen, allerdings wurde von B-Kameras berichtet, mit deren Material etwaige Lücken oder fehlerhafte Parts des Hauptmaterials vertuscht werden können (Anhang B und C).

Zudem ist ein Mehrkamerasystem bei Musikvideos nicht zwingend sinnvoll: „Da es meistens um eine Performance direkt in die Kamera geht, ergibt sich ein Mehrkamerasystem selten“ (Anhang B). Ein Experte erwähnte, dass Takes, die eigentlich für die Performance verwendet werden sollten, auch als B-Roll-Shots verwendet werden: „[...] Danach sichte ich das Material noch mal nach ‚B-Roll-Shots‘ - oft passiert in einem Performance-Take zufällig etwas Cooles.“ (Anhang B).

Umgang mit Notizen und Metadaten

Der Workflow mit Metadaten und Notizen, die auf dem Set erstellt werden, variierte innerhalb der Interviewpartner. Häufig wurde aber betont, dass bei Musikvideos Metadaten nicht so wichtig sind, da aufgrund der Dauer das Material trotzdem übersichtlich bleibt (Anhang B), „da die Regisseure oft selbst schneiden und das Material kennen“ (Anhang C) oder da man versucht, die Drehs möglichst simpel zu gestalten: „[...] Die Zeit am Set ist viel kostbarer als in der Postproduktion. Ich plädiere oft dafür, keine Klappe zu benutzen und es möglichst einfach zu halten, um den Flow nicht zu unterbrechen“ (Anhang B). Es wurde aber bestätigt, dass auch hier die Art der Produktion entscheidend ist: „Wenn man eine

chronologische Story unchronologisch dreht, ist Dokumentation wichtiger als bei einem – Run-and-Gun-Musikvideo, das man intuitiv zusammenschneidet“ (Anhang B). Allerdings wurde auch erwähnt, dass „mehr Infos“ immer besser sind und somit das Vorhandensein von Metadaten „definitiv ein Mehrwert“ wäre (Anhang A).

Anzahl der Takes pro Einstellung

Die Anzahl der Takes pro Einstellung ist ebenfalls projektabhängig: „Bei Analogfilm meistens nur einen bis drei Takes. Wir proben dafür vorher extrem viel“ (Anhang C). Viele Takes können außerdem Nachteile beinhalten: „Das ist teilweise eine Abwägungsfrage: Geht nach vielen Takes irgendwann die Stimmung nach unten oder geht man lieber auf Nummer sicher?“ (Anhang C).

Budget und finanzielle Einschätzung zu ARC-Play

Auch die Budgets für Musikvideodrehs sind sehr projektabhängig und reichen von Non-Budget-Produktionen bis in den „niedrigen fünfstelligen Bereich“ (Anhang C). Es wurde auch berichtet, dass gerade Newcomer und kleinere Künstler weniger komplexe Musikvideos drehen: „Oft werden heute keine kompletten Videos mehr für YouTube gemacht, sondern nur noch Social-Media-Snippets als Reels oder TikToks, weil die einen höheren Promo-Stellenwert haben“ (Anhang C).

In Bezug auf ARC-Play wurde erwähnt, dass es eher für Projekte mit höheren Budgets geeignet ist. Es soll aber dennoch nicht „mehrere Tausende Euro“ (Anhang A) kosten. Aus einem Interview ging hervor, dass der Nutzen des Systems aber trotzdem gegeben sei: „[...] das wäre eine Sache, bei der Mehrwert ziemlich schnell erkannt werden würde. Spart auf jeden Fall Stunden in der Postproduktion“ (Anhang C).

Anforderungen an die Bedienung und Features von ARC-Play

Aus den Interviews kamen verschiedene Vorschläge und Anstöße für Anforderungen von ARC-Play: „Wichtig ist eine Benutzeroberfläche, die sich wie Consumer-Technik anfühlt. Viele Leute im Musikvideobereich kommen nicht von der technischen Seite und brauchen eine intuitive Bedienung“ (Anhang C). Außerdem wurden die Flexibilität und die unabhängige Handhabung als Anforderung betont. Bei Letzterem wurde erwähnt, dass es bei professionellen Drehs inzwischen sehr viele Geräte gibt, die sich nur über eine App vollständig bedienen lassen. Das kann aber bei zu vielen Geräten umständlich werden. Daher wäre eine „Stand-alone“-Bedienung mit einem eigenen Display wünschenswert (Anhang C).

Auch ein eingebauter Lautsprecher als „All-in-One“-Lösung und ohne externen Bluetooth-Lautsprecher wurde als „riesiger Vorteil“ (Anhang A) gewertet. Zudem wurde auch ein Mehrlautsprechersystem erwähnt: „Manchmal könnten mehrere Lautsprecher, die gleichzeitig den Track abspielen, praktisch sein, wenn man über größere Distanzen filmt. Oder auch für die Regie als Playback“ (Anhang B).

Laufzeitprobleme

Die Experten haben selbst bei großen Distanzen keine Laufzeitprobleme erlebt: „Bisher nicht. Wir haben mal mit 250 mm Brennweite aus großer Entfernung gedreht. Da hört man den Ton an der Kamera ohnehin nicht mehr und muss das Material später komplett nach Sicht anlegen“ (Anhang C).

Workflow im Schnitt

Zum Thema Schnitt gibt es auch unterschiedliche Aussagen. Ein Experte fügt beispielsweise alle Clips in die Timeline ein und schneidet dann direkt dort das Video (Anhang A), ein anderer Experte arbeitet mit Multicam-Sequenzen (Anhang B).

Interaktion mit Künstlern

Aus den Interviews gingen auch Hinweise für das Interagieren mit den Darstellern hervor: „Ich habe gelernt, dass drei Sekunden Vorlauf nicht reichen. Man muss ihnen [den Künstlern] mehr Zeit geben, damit sie in den Flow kommen“ (Anhang B). Außerdem wurde ein gewisses Erwartungsmanagement mit vor allem jüngeren Künstlern empfohlen, da diese öfter aus mangelnder Dreherfahrung mit zu hohen Erwartungen an die Musikvideoproduktion gehen und daher nicht einschätzen können, welche Probleme während eines Drehs auf sie zukommen können (Anhang C).

Aktueller Trend in der Musikvideobranche

Ein Experte gab auch Auskünfte über den aktuellen Trend von Musikvideodrehs mit analogem Film: „Wir drehen auch mehr und mehr analog, da gibt es auf jeden Fall einen Aufwärtstrend [...] Ich denke, vor allem durch AI wird das immer beliebter, weil das eine Art Gegenpol ist, da du dadurch beweisen kannst, dass das echt ist“ (Anhang B). Dies steht im Gegensatz zur Aussage eines anderen Experten, der, gerade bei unbekannten Künstlern, eher eine Verschiebung der Musikvideoproduktionen hin zu Social Media sieht (Anhang C).

4.5 Diskussion der Ergebnisse

Im Folgenden werden die aus Kapitel 4.4 dargestellten Ergebnisse nach Kategorien sortiert und in den Kontext dieser Arbeit gestellt und diskutiert.

Aktueller Workflow mit Playbacksystemen und Synchronisierung

Der von den Experten dargestellte Workflow scheint häufig im „klassischen Fall“ nur aus einem Bluetooth-Lautsprecher und einem Smartphone zu bestehen. Inwiefern das Synchronisieren ein Problem ist und somit ARC-Play eine Lösung bieten könnte, ist davon abhängig, ob das Playback vom Set im Rohmaterial zu hören ist, da es in modernen Schnittprogrammen Funktionen gibt, mit denen Ton automatisch synchronisiert werden kann. Bei einer Kamera mit Mikrofon und kleinem Drehort gibt es daher kaum Probleme, weshalb es in solchen Situationen für ARC-Play kaum Mehrwert zu geben scheint. Bei Szenen mit größeren Distanzen, bei denen der Ton schlecht in der Kamera zu hören ist, hätte das System deutlich mehr Potential. Die Distanz müsste allerdings ebenfalls überwunden werden. Den größten Nutzen hätte das System aber bei Kameras, die keinen Ton aufzeichnen können, wie beispielsweise viele analoge Filmkameras oder Filmdrohnen. Da das Synchronisieren dabei ausschließlich anhand der Bewegungen der Protagonisten erfolgt, führt das zu erheblichem Aufwand in der Postproduktion. Mit ARC-Play könnte bereits alles im Vorhinein synchronisiert werden. Allerdings muss dabei, wenn kein Timecode-Eingang vorhanden ist, ARC-Slate oder ein anderes Timecode-Display verwendet werden. Den Nutzen sahen auch die Experten und würden ein solches System begrüßen. Da es laut den Interviews Probleme mit der Latenz von Bluetooth-Verbindungen gegeben hat, soll dies möglichst vermieden werden und auf analoge Audioausgänge oder interne Lautsprecher gesetzt werden.

Andere Lösungen, bei denen, ähnlich wie bei Kapitel 3.3.1, eine Klappe oder ein iPad mit Timecode in das Kamerabild gehalten werden, funktionieren auch, scheinen aber durch das manuelle Eingeben des Timecodes in der Postproduktion ebenfalls mit Aufwand verbunden zu sein, den das System, solange die Filmklappe ARC-Slate verwendet wird, ebenfalls eliminieren würde.

Alternative Songs als Playback zu verwenden, sollte als Feature in das Playback-System eingebaut werden. Dadurch ist eine größere Flexibilität gegeben. In der Praxis hat keiner der Experten kommerzielle Systeme wie in Kapitel 3.3 beschrieben verwendet, was ein weiterer Hinweis auf die hohe Individualität bei der Lösung der Problemstellung ist. Außerdem

verwendet keiner der Experten Click- oder Thumbtracks. Da diese Features aber relativ leicht umzusetzen sind, sollten sie trotzdem implementiert werden.

Aktueller Planungsaufwand für Musikvideodrehs

Musikvideodrehs, bei denen jede Szene und jeder Shot genau geplant werden, sind die, bei denen ARC-Play den größten Nutzen ausspielen könnte, da die Struktur einer digitalen Shotlistverwaltung durch ARC-Slate bereits vorliegt (siehe Kapitel 2). Dies kommt den Interviews zufolge vor allem bei analogen Drehs vor, da das sequenzielle Aufsplitten des Playbacks teures Filmmaterial spart. Es gibt aber auch digitale Drehs, bei denen die Szenen genauer geplant werden, da ein gezielteres Filmen von bestimmten Stellen im Song während des Drehtags Zeit sparen kann. Bei einer erfahrenen Crew wird für bestimmte Parts, ähnlich wie beim Master-Edit-Verfahren aus Kapitel 3.3.1, die Audiodatei zugeschnitten. Diese Vorbereitung könnte mit ARC-Play effizienter und schneller abgewickelt werden und könnte zudem auch am Set spontan geändert werden.

Geschwindigkeitsveränderung als kreatives Mittel

Die Validierung durch alle Interviewpartner belegt, dass die gezielte Geschwindigkeitsmanipulation ein essenzielles gestalterisches Stilmittel im modernen Musikvideo darstellt. Daher soll dieses Feature in die Anforderung an ARC-Play aufgenommen werden. Das System muss in der Lage sein, die Wiedergabegeschwindigkeit des Audios flexibel modifizierbar zu machen, um so auch spontan auf die Tagesform der Künstlerperformance reagieren zu können.

Das von einem Experten erwähnte Rückwärtsabspielen eines Songs ist ein sehr spezieller Fall, der nicht unbedingt als Hauptfeature in das System implementiert werden muss. Wenn dennoch ein solches Verfahren gewünscht ist, kann der Song, bevor man ihn in ARC-Play hochlädt, in einer Audiosoftware rückwärts exportiert werden.

Mehrkamerasysteme

Laut den Experten gibt es Mehrkamerasysteme nur in der Form, bei der es eine Hauptkamera gibt, in die die Performance des Künstlers gerichtet ist, und dann zusätzlich eine B-Roll-Kamera.

Ein Experte erwähnte, dass während der Performance oft ungeplante Dinge passieren, die er auch als B-Roll-Material verwenden könnte. Für diese Funktion könnte man in der Dokumentation des Takes ein zusätzliches Optionsfeld erstellen, bei dem man auswählen kann, ob bei diesem Take so etwas passiert ist. Dies könnte man in der Analysesoftware von

ARC-Slate wieder aufgreifen und dann eine zweite Version des Takes rendern, die als B-Roll gekennzeichnet wird.

Umgang mit Notizen und Metadaten

Metadaten und Notizen sind laut Experten bei Musikvideodrehs aufgrund der kurzen Lauflänge und doppelt besetzter Positionen während des Drehs und der Postproduktion weniger gebräuchlich, aber dennoch gerade bei aufwendigeren Produktionen von Nutzen. Da ARC-Play auf dem in Kapitel 2 beschriebenen ARC-System aufbaut, in dem ein vollständig automatisierter Metadatenworkflow integriert ist, wird dennoch versucht, möglichst viele Metadaten mit ARC-Play zu sammeln und in die Daten zu integrieren.

Ein Experte erwähnte zudem, dass eine Klappe in einem dynamischen Arbeitsumfeld zu zeitaufwendig ist. Das würde gegen das Benutzen der ARC-Slate sprechen. Alternativ müsste man versuchen, die Metadaten, die über die Weboberfläche des ARC-Systems erhoben werden, ohne Klappe mit den Clips zu übertragen. Dazu könnte einfach der bereits integrierte Timecode verwendet werden. Allerdings können bei choreographierten Drehs mehrere Shots und Takes pro Stelle im Song entstehen, wodurch die Zuordnung nicht eindeutig wäre. Alternativ könnten die in Kapitel 3.1 dargestellten User-Bits des SMPTE-Timecodes genutzt werden, um Informationen zu übertragen.

Budget und finanzielle Einschätzung zu ARC-Play

Das Playback-System darf laut Experten nicht zu teuer sein, da die Budgets bei Musikvideodrehs nicht sonderlich hoch sind. Wichtig für die Vermarktung eines solchen Produkts wäre, es möglichst moderat zu bepreisen, um die Zeitersparnis durch den Kunden nicht mit den Kosten des Systems obsolet zu machen.

Anforderungen an die Bedienung und Features von ARC-Play

Den Experten war es wichtig, dass die Bedienung von ARC-Play sehr einfach gestaltet werden muss. Darüber hinaus sollen alle Funktionen auch direkt auf dem Gerät einstellbar sein. Daher soll der Prototyp des Playbacksystems über ein Display und einen Drehencoder verfügen, mit dem sich alle Einstellungen vornehmen lassen können. Für die Notizen, die während des Drehs gemacht werden, könnte die Weboberfläche, die bereits bei ARC-Slate verwendet wurde, genutzt werden.

Da der Bau eines eigenen Lautsprechers recht aufwendig ist, soll sich für ARC-Play auf einen analogen Audioausgang beschränkt werden, mit dem sich sowohl kleine Lautsprecher

mit Audioeingang als auch größere Soundanlagen in Studios betreiben lassen. Bluetooth wird aufgrund der genannten Latenzprobleme nicht integriert.

Zudem gibt es laut einem Experten Fälle, in denen man das Audio an mehreren Orten gleichzeitig braucht. Dafür könnte man das System so gestalten, dass man mehreren Einheiten des ARC-Play-Systems synchronisieren kann, damit diese gleichzeitig den Ton abspielen. Bei Drehs mit Autos könnte der Audioausgang von ARC-Play direkt mit dem Lautsprechersystem des Autos verbunden werden.

Laufzeitprobleme

Obwohl die Experten keine Laufzeitprobleme hatten, scheint das Synchronisieren bei großen Distanzen aufwendig zu sein. Auch hier könnten verschiedene Lösungen wie das Nutzen von ARC-Slate oder mehreren ARC-Play-Modulen angewendet werden. Wichtig ist, dass das Netzwerk diese Distanz abdeckt.

Workflow im Schnitt

Da der Workflow im Schnitt sehr individuell ist, sollen bereits möglichst viele Informationen außerhalb des Schnittprogramms in die Rohvideodatei oder in externe Metadateiendateien integriert werden, um so unabhängig von der Schnittplattform zu sein. Die aus dem bisherigen ARC-System bestehende Unterstützung für DaVinci Resolve soll beibehalten werden und langfristig auf weitere Schnittprogramme erweitert werden.

Interaktion mit Künstlern

Den Interviews zufolge scheinen Künstler etwas mehr Vorlaufzeit zu brauchen, bevor der eigentliche Shot beginnt. Um das zu erreichen, soll die Vorlaufzeit von ARC-Play flexibel gestaltet werden. Gerade wenn echte Instrumente verwendet werden, könnte es für Musiker schwierig sein, in der Mitte des Songs anzufangen. Daher würde es Sinn machen, dass die Vorläufe für die eigentlichen Takes nicht willkürlich gewählt werden, sondern einem bestimmten Raster wie zum Beispiel Taktgrenzen folgen. Außerdem könnte der Song gezielt in bestimmte Bereiche untergliedert werden.

4.6 Typologie von Musikvideoproduktionen

Viele Befunde der Experteninterviews sind stark kontext- und produktionsabhängig. Um für den weiteren Ablauf in dieser Arbeit eine Grundlage zu schaffen, insbesondere für die Berechnungen in Kapitel 6, wird im Folgenden versucht, verschiedene Typen von

Musikvideodrehs zu unterscheiden. Diese lassen sich aufgrund der Aussagen der Experten anhand von vier Kategorien unterscheiden:

Kameratechnik

Eine Kategorie wäre die verwendete Kameratechnik, insbesondere die Verwendung von analogen: „Wenn wir digital drehen, probieren wir meistens, erst einmal komplette Durchläufe des Songs an jeder Location zu machen. [...] Bei Analogdrehs machen wir keine kompletten Durchläufe, sondern planen die Takes sehr genau abgesetzt“ (Anhang C). Eine weitere verwendete Kameratechnik sind Kameras ohne eigene Tonaufnahmemöglichkeit: „Besonders wenn man mit Drohnen oder alten Alexa-Kameras ohne internes Mikro dreht, landet man oft in der Situation, dass man nach Lippenbewegungen synchronisieren muss“ (Anhang B). Bei einer solchen Kamera gibt es folglich ebenfalls große Probleme mit der Synchronisierung.

Planungsaufwand

Man kann die Produktion auch nach der Art des Filmens unterscheiden: Es gibt den bereits erwähnten „Run-and-Gun“-Stil, bei dem spontan entschieden wird, was gefilmt wird, oder den szenischen Dreh, bei dem jede Szene klar durchgeplant ist. Dies wirkt sich häufig nicht nur auf die Vorproduktion aus, sondern auch darauf, wie ausführlich auf dem Dreh dokumentiert wird (Anhang B).

Genre des Musikstücks

Eine weitere Kategorie ist das Genre der Musik. Bei Musikrichtungen mit immer wiederholenden Elementen, keinen Taktänderungen und ohne lippensynchronen Text, beispielsweise Techno, kann das Material beliebig an jeder Stelle des Songs eingesetzt werden. Dadurch fällt der Planungsaufwand dafür deutlich geringer aus (Anhang C). Es handelt sich dann aber auch um keine feste Choreographie mehr.

Budget

Auch durch das Budget des Projekts lassen sich Musikvideos kategorisieren: Bei niedrigen Budgets wird oft mehr Content für Social Media erstellt, für den deutlich weniger Planungsaufwand notwendig ist (Anhang C).

Da sich die genannten Kategorien teilweise stark überschneiden, werden nun drei Annahmen für unterschiedliche Idealtypen abgeleitet, die versuchen, die Bandbreite an verschiedenen Produktionsarten wiederzugeben. Da es für ARC-Play die höchste Relevanz besitzt, wurde als Hauptunterscheidung die verwendete Kameratechnik verwendet.

A Digitalkamera mit Tonaufzeichnung:

Der erste Typ ist ein Musikvideodreh mit einer digitalen Kamera, die in der Lage ist, Ton aufzuzeichnen. Mit dieser Kamera macht das Synchronisieren, solange der Ton über das interne oder ein externes Mikrofon synchron aufgezeichnet wird, wenig Probleme. Teilweise kann man in moderner Schnittsoftware den Ton nur anhand der Tonspur automatisch synchronisieren.

Des Weiteren muss in einem solchen Szenario nicht zwingend in der Vorproduktion ausführlich geplant werden. Durch die im Vergleich günstigen Speicheroptionen von digitalem Filmmaterial kann theoretisch jede Szene die volle Länge des Musikstücks aufgezeichnet werden und dann im Schnitt entschieden werden, welcher Part für welche Stelle passt. Der bereits erwähnte „Run and Gun“-Style, bei dem man sehr spontan das Musikvideo dreht, wird von diesem Typ begünstigt. Dadurch kann eine gewisse zeitliche Ineffizienz entstehen, in der die Kosten für Schauspieler, Mieten oder durch die größere Materialmenge die Kosten in der Postproduktion steigen. Außerdem müssen, gerade bei professionelleren Kameras, bei denen große Datenmengen anfallen, die Speicher ebenfalls finanziert werden. Dennoch ist diese Art zu filmen oft deutlich günstiger und könnte im Extremfall mit einem Smartphone für Social Media umgesetzt werden.

B Digitalkamera ohne Tonaufzeichnung

Der zweite Typ umfasst digitale Produktionen, bei denen aus technischen und situativen Gründen kein Referenzton aufgezeichnet werden kann. Dadurch ergeben sich in der Synchronisation erhebliche Probleme. Dazu gehören Drohnen oder Situationen, bei denen am Set das Aufzeichnen von Ton aus verschiedenen Gründen, wie beispielsweise räumlicher Distanz, nicht möglich ist. Das Automatisieren über den Ton ist somit systembedingt nicht möglich und es muss sehr aufwendig auf die Choreographie und Lippenbewegung zurückgegriffen werden.

Da weiterhin digital gefilmt wird, ist durch geringeren Planungsaufwand dennoch eine hohe Flexibilität gegeben.

C Analoge Kameras

Die Verwendung von analogen Filmkameras schränkt die Musikvideoproduktion deutlich ein. Da analoge Kameras meistens keinen Ton aufnehmen, ergeben sich für die Synchronisation in der Postproduktion die gleichen Gegebenheiten wie bei Typ B. Die Flexibilität ist durch die hohen Kosten von analogem Film deutlich eingeschränkt. Um die

maximal effiziente Nutzung des Filmmaterials zu gewährleisten, wird mit einer analogen Kamera der Dreh deutlich genauer geplant und sequenziell gedreht.

4.7 Fazit der Experteninterviews

Insgesamt ist festzustellen, dass es bei Musikvideoproduktionen reale Probleme gibt, für die ARC-Play eine Lösung sein kann. Die Experten sahen alle ein gewisses Potential in dieser Idee und könnten sich vorstellen, dass das in Zukunft – je nach Preisgestaltung – Anwendung finden wird. Ein Experte hat bereits Interesse gezeigt: „[...] sag Bescheid, wenn ihr da irgendeinen Prototypen habt, dann würden wir den sehr gerne auch verwenden“ (Anhang B). Daher sind die Interviews eine gute Motivation, die Anforderungen, die die Experten an das System gesetzt haben, umzusetzen und gegebenenfalls die Entwicklung nach dieser Arbeit fortzuführen.

5 Konzept und Anforderung von ARC-Play

Die Formulierung des Konzepts und die technologische Anforderung an ARC-Play basieren auf Aspekten, die aus den Vorarbeiten (siehe Kapitel 2), aus der Theorie und Recherche (siehe Kapitel 3) und den Experteninterviews (siehe Kapitel 4) stammen. Zusätzlich fließen eigene Ideen mit in das Konzept.

5.1 Grundlegende Features von ARC-Play

Für die in Kapitel 1 dargestellten Problemstellungen bietet ARC-Play folgende Lösungen:

Synchronisiertes Abspielen von Timecode und Playback

Um Video und Ton zu synchronisieren, wird der Song auf dem Gerät abgespielt und gleichzeitig der dazu passende Timecode generiert, der als LTC ausgegeben wird. Geschwindigkeitsänderungen können spontan am Set angepasst werden. Wenn eine Kamera ohne Timecode-Eingang verwendet wird, kann eine Timecode-Anzeige oder eine Filmklappe mit Timecode-Anzeige verwendet werden. Für volle Automatisierung soll zu einem späteren Zeitpunkt die Filmklappe ARC-Slate synchron verwendet werden können.

Strukturierung durch Shotlists mit In- und Outpunkten

Ein weiteres Feature soll die intuitive Aufteilung des Songs beinhalten. Dabei wird die Shotliststruktur von ARC-Slate aufgegriffen und mit Start- und Endpunkten je Shot versehen. Dadurch lässt sich der Dreh gezielt vorbereiten und somit die Effizienz – gerade bei analogem Film – enorm steigern. Zur besseren Planbarkeit soll der Song, sobald er auf den Server hochgeladen wurde, mit einem Raster ausgestattet werden, bei dem die Start- und Endpunkte eines Shots automatisch an den Taktgrenzen des Songs einrasten.

Das Gerät soll zudem weitere Kernfeatures beinhalten, die über die genannten Problemstellungen hinausgehen:

Automatisierte Metadatenintegration

ARC-Play soll in der Lage sein, ohne den Nutzen der digitalen Filmklappe ARC-Slate die Metadaten automatisch mit den Rohvideos zu kombinieren. Dazu wird in die User-Bits des SMPTE-Timecodes (siehe Kapitel 3.1) die ID, die ARC-Slate auf dem QR-Code anzeigt, integriert. In der Postproduktion wird das Material analysiert und kann daraufhin gezielt nur auf den Start- und End-Timecode zugeschnitten werden und mit allen gesammelten Metadaten versehen werden (siehe Abbildung 4). Es kann bei sehr kleinteiliger Shotplanung

ein Preview eines choreographierten Films entstehen. Diese Funktion ist bei analogem Filmmaterial nur mit ARC-Slate gegeben.

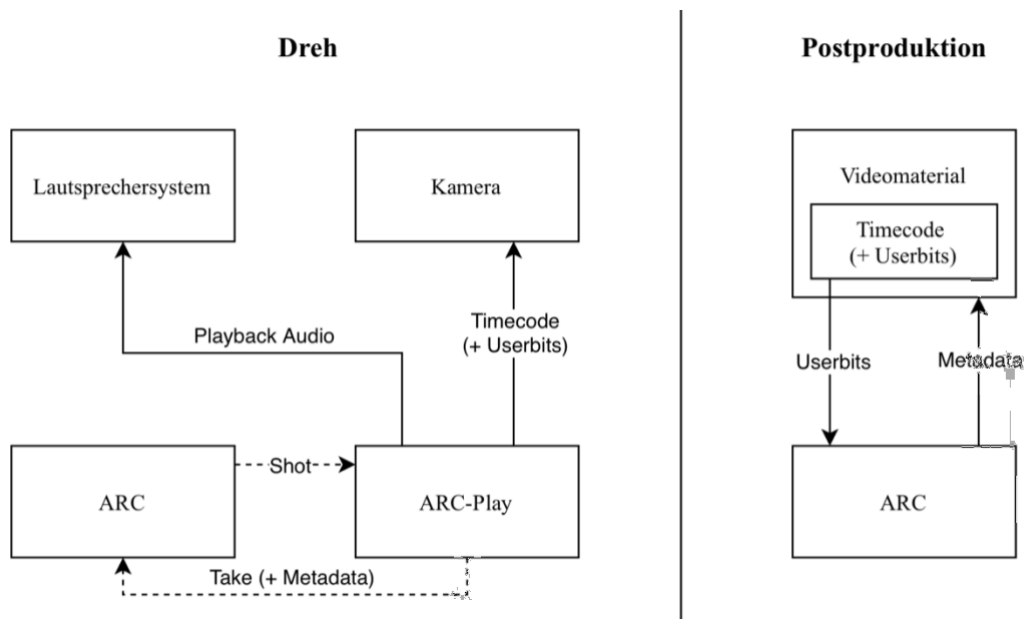


Abbildung 4: Übersicht der Metadatenintegration von ARC-Play
Quelle: eigene Darstellung

Verbesserung des Drehverhältnisses

Ein weiteres Kernfeature ist die potenzielle Verbesserung des Drehverhältnisses. Wie in Kapitel 3.2 aufgezeigt, entsteht ein hohes Drehverhältnis bei Spielfilmen unter anderem durch eine hohe Anzahl an Takes für jeden Shot, da man vermeiden möchte, fehlerhaftes Bildmaterial nutzen zu müssen. Bei Musikvideodrehs werden allerdings häufig mehrere Shots zum Beispiel an unterschiedlichen Orten vom gleichen Part des Musikstücks erstellt und von den einzelnen Shots wiederum mehrere Takes. Das somit sehr hohe Drehverhältnis könnte dadurch verringert werden, indem man einzelne, redundante Takes reduziert und auf alternative Shots als Backup zurückgreift. Um das Feature zu implementieren, soll eine Anzeige entwickelt werden, bei der jeder Part des Songs mit dessen Shot- und Take-Anzahl dargestellt wird. Dadurch wird auch vermieden, dass bestimmte Bereiche des Audios gar kein Videomaterial beinhalten.

5.2 Funktionsweise von ARC-Play

ARC-Play soll neben seiner Funktion als Playback-System auch als ganz normaler Timecode-Generator nutzbar sein. Im Gegensatz zum bisherigen Timecode-Generator von ARC-Slate soll er in der Lage sein, viele gängige Timecodeformate abzuspielen. Neben Free-Run-Timecode sollen auch Record-Run-Timecode und Jam-Sync möglich sein.

Das Gerät soll ein großes Display und einen Drehknopf zur Bedienung bekommen. Das Design könnte sich bei späteren Prototypen an das des Apple iPods orientieren. Diese sind für ihr simples Bedienkonzept bekannt und das ähnliche Design könnte beim Benutzen gewisse Assoziationen hinsichtlich Musikabspielen hervorrufen (siehe Visualisierung in Abbildung 5). Mit diesem Design könnten so sämtliche Einstellungen des Geräts direkt vorgenommen werden. Allerdings muss darauf geachtet werden, keine Patentrechte zu verletzen. Außerdem verfügt es über zwei 3,5-mm-Klinkenbuchsen, bei denen die eine eine Monoversion des Audioplaybacks ausgibt und die andere sowohl als Ein- als auch als Ausgang für das Timecodesignal fungiert.



Abbildung 5: KI-generierte Visualisierung von ARC-Play mit Google Nano Banana 2

Sobald das Gerät vom Server einen Shot erhalten hat, kann die Bedienung auf Wunsch komplett autark auf dem Gerät erfolgen. Bevor der Song abgespielt wird, wird in entsprechender Taktgeschwindigkeit mit kurzen Signaltönen eingezählt. Wenn ein Take erstellt wurde, wird die Takeziffer automatisch um eine Nummer erhöht und der Timecode wieder auf den Starttimecode des Songs zurückgesetzt. Aber auch ohne geladenen Shot kann das Gerät den Song mit synchronisiertem Timecode von vorne bis hinten abspielen.

Der Server basiert auf der vorherigen Version von ARC, soll aber strukturell deutlich aufgewertet und erweitert werden. Die Hauptbedienoberfläche von ARC-Play erfolgt nun über *ARC-Studio*, das auf die bisherige Softwareoberfläche aufbaut. Dort soll zudem die

Shotlisterstellung inklusive der Unterteilung des Songs in Sequenzen möglich sein. Sobald ein Take erstellt wird, können wie zuvor Notizen erstellt werden, und es soll markiert werden, ob es sich um einen guten Take handelt. Neu hinzukommen soll die Funktion, dass man einen Haken setzen kann, wenn es sich um einen potenziellen B-Roll-Take handelt.

Die Analysesoftware soll nun nicht mehr autark, sondern zusammen mit der ARC-Studio Teil des Front-Ends des Servers sein und den Namen *ARC-Analyzer* bekommen. Die Software soll in der Lage sein, aus einem Ordner, in dem sich Filmmaterial von einem Dreh befindet, bei dem ARC-Play verwendet wurde, automatisch zu identifizieren, den Timecode zu extrahieren und mithilfe der User-Bits die Metadaten wieder zuzuordnen. Auch die Previewerstellung soll damit möglich sein.

5.3 Weitere Ideen und Features

Es folgen nun Features, die aus den bisherigen Kapiteln der Arbeit hervorgingen, aber nicht zwingend zur Kernfunktionalität des Geräts beitragen und durch den begrenzten Zeitrahmen dieser Arbeit nicht priorisiert werden. Sie könnten aber zu einem späteren Zeitpunkt integriert werden.

Bei der Verwendung von mehreren ARC-Play-Instanzen könnten größere Distanzen besser überwunden werden. Daher soll ARC-Studio auf die Verwendung von mehreren Geräten parallel ausgestattet werden. Diese müssten vor dem Dreh synchronisiert werden. Da im Rahmen dieser Arbeit nur die Umsetzung eines Prototyps geplant ist, kann diese Funktion vorerst nicht implementiert werden.

Um auch das Audio des Songs auf die Postproduktion vorzubereiten, könnte das in Kapitel 3.1 vorgestellte BWF-File verwendet werden, mit dem bei abweichendem Start-Timecode des Songs der Timecode wieder synchronisiert werden könnte.

Außerdem wäre es, wie in den Interviews erwähnt, für manche Situationen von Vorteil, alternative Songs abzuspielen. Dafür müssten bei der Shot-Erstellung zwei Zeitstempel hinzugefügt werden: zum einen die des Songs, der abgespielt wird, und zum anderen die der finalen Position dieses Shots im Video.

Um bei der visuellen Integration von Timecode nicht ausschließlich auf eine externe Timecodeanzeige oder ARC-Slate angewiesen zu sein, könnte der Server Timecodevideos, wie in Kapitel 3.3.1 dargestellt, generieren. Diese könnten, ähnlich wie die Lösung von Dan Graetz aus Kapitel 3.3.1, über den VLC Media Player auf ein Tablet übertragen und extern getriggert werden.

6 Berechnungen zur finanziellen Ersparnis von ARC-Play

6.1 Methodik der Berechnungen

Um den Nutzen des Prototyps des Audio-Playback-Systems schon in der Theorie zu erforschen, wurden Berechnungen getätigt, um unter verschiedenen Annahmen in Bezug auf den Workflow von Musikvideodrehn und auf den Wirkungsgrad hinsichtlich der Effizienz von ARC-Play eine grobe Kostenschätzung mit und ohne die Verwendung des Systems darzustellen. Dabei wurde aus der in Kapitel 4.6 erschlossenen Typologie von Musikvideodrehn Szenarien entwickelt und dazu passend verschiedene Parameter angepasst. Im Anschluss sollen die Ergebnisse diskutiert werden. Die dargestellten Werte beinhalten viele heuristische Annahmen. Sie stellen keine gesicherten realistischen Szenarien dar und es kann in der Realität zu Abweichungen kommen. Die Berechnungen dienen vielmehr dazu, den Workflow mit dem System beispielhaft zu simulieren und dessen Effizienz klar vor Augen zu führen.

6.2 Aufbau der Berechnungen

Um ARC-Play hinsichtlich des finanziellen Mehrwerts theoretisch zu untersuchen, sollen die Tabellen mit den Ergebnissen den Workflow einer Musikvideoproduktion darstellen – allerdings nur auf die für ARC-Play relevanten Bereiche.

Dazu zählt zum Ersten die Vorproduktion: Um das System zu benutzen, muss zwingend eine Shotlist erstellt werden, die entweder bereits im Webinterface von ARC-Play oder extern erstellt und dann in das System überführt wird. Das bedeutet, das System erfordert gegebenenfalls mehr Zeit in der Vorproduktion, wodurch Kosten für die Gehälter der Personen, die den Dreh vorbereiten, in die Höhe gehen. Daher könnten in der Vorproduktion sogar Mehrkosten anfallen.

Während des Drehs wäre ARC-Play potenziell in der Lage, durch die bereits in dem System vorbereiteten Shots und Playbacks Zeit zu sparen. Es muss nicht mehr manuell die richtige Stelle gesucht werden und durch die Vorbereitung ist das Team gezwungen, sich an die Shotlist zu halten. Die reine Zeitersparnis wird nicht sonderlich groß sein, allerdings werden – je nach der Größe der Crew – nur bei wenigen Minuten durch die Gehälter der Crewmitglieder hohe Beträge gespart.

Das größte Potential hat das System in der Postproduktion, da die Synchronisation komplett automatisiert vonstattengehen soll. Dadurch kann sich der Editor im Schnitt, je nach

verwendeter Technik, bis zu mehrere Stunden Arbeit sparen, um das Material zu synchronisieren. Außerdem bietet die automatisierte Integration von Metadaten eine effiziente Strukturierung des Materials.

Bei analogen Produktionen könnte zudem das Drehverhältnis leicht verbessert werden. Bei analogem Film kann je nach Drehdauer selbst eine sehr geringe Verbesserung einen großen finanziellen Mehrwert bieten.

Der Wirkungsgrad für die einzelnen Bereiche unterscheidet sich je nach Szenario.

6.3 Berechnungen anhand verschiedener Szenarien

In den folgenden Berechnungen handelt es sich grundsätzlich immer um hypothetische Annahmen, die von der Realität abweichen können. Um in den folgenden Szenarien dennoch mit Zahlenwerten arbeiten zu können, werden die meisten Gagen aus der aktuellen Gagentabelle für Filmschaffende nach der Gewerkschaft Ver.di (ver.di, 2024, S. 19) entnommen. Diese werden dort als Wochengagen dargestellt. Allerdings werden im selben Vertrag Arbeitstage mit zehn Stunden definiert. Daher wurden die Wochengagen durch 50 Stunden geteilt, um einen angenäherten Wert für die Berechnungen zu erlangen. Die Gage eines Regisseurs ist dort nicht mit aufgeführt, daher wird eine Annahme getroffen, bei der das Gehalt eines Regisseurs mit dem eines DoP identisch ist. Wichtig ist anzumerken, dass die Gehälter in der Realität abweichen können, da gerade im Low-Budget-Filmbereich viele Positionen niedrig- oder unbezahlt besetzt werden. Auch für die Vorbereitung wurde die Annahme getroffen, dass das von einer Regieassistentin oder dem Regisseur übernommen wird.

Um die verschiedenen Szenarien möglichst vergleichbar zu halten, wird versucht, nur die relevanten Parameter zu verändern. Daher wird davon ausgegangen, dass bei allen Szenarien ein Musikvideo für einen Song produziert wird, der eine Laufzeit von drei Minuten besitzt. Außerdem wird jede Position in jedem Szenario mit den gleichen Gehältern bezahlt. Für die Annahme ohne ARC-Play wird die aus den Experteninterviews aus Kapitel 4.4 erwähnte klassische Methode verwendet, bei der man mit einem Smartphone und einem Bluetooth-Lautsprecher den Song abspielt.

Die Szenarien spiegeln nur die oben erwähnten Bereiche wider, also die Shotlist- und Playbackerstellung, den Dreh, die Synchronisierung, den Rohschnitt in der Postproduktion und, falls vorhanden, die Materialkosten für analogen Film. Die Kosten für die gesamte Produktion könnten noch deutlich höher ausfallen.

6.3.1 Low-Budget-Dreh

Dieses Szenario lehnt sich an Typ A aus Kapitel 4.6 an:

In der Vorproduktion übernimmt eine Person die Shotlist- und Playbackerstellung und wird in diesem hypothetischen Fall mit der Gage einer Regieassistentz vergütet. Das Playback ist ein noch nicht releaster Song des Künstlers und muss einfach nur auf ein Smartphone kopiert werden. In diesem Szenario wird der Song an verschiedenen Orten in voller Länge gedreht. Es sind keine Zeitlupeneffekte geplant. Dementsprechend ist die Shotlist einfach und flexibel gehalten. Die Person benötigt für diese Aufgaben eine halbe Stunde.

Die Crew am Set besteht aus zwei Personen, einem DoP und einer Assistenz, und ist auf einen Drehtag mit zehn Stunden angesetzt. Als Equipment kommt eine moderne spiegellose Kamera mit internem Mikrofon zum Einsatz und die Gegebenheiten belaufen sich immer so, dass auf der Kamera das Playback gut zu hören ist. Der Dreh läuft sehr flexibel ab und es werden viele Shots spontan im „Run-and-Gun“-Style hinzugefügt.

Der Rohschnitt wird von einem Editor übernommen und er benötigt dafür acht Stunden. Da immer der gesamte Song am Stück gefilmt wurde und auf der Kamera das Playback des Songs am Dreh gut zu hören ist, kann das gesamte Material automatisch synchronisiert werden. Da jeder Take in voller Länge des Stücks gedreht wurde und von daher das Drehverhältnis relativ hoch ist, gibt es viel Material, wodurch der Editor etwas länger für die Auswahl der finalen Takes braucht.

Tabelle 2: Die Ergebnisse der Berechnungen unter Annahme eines Low-Budget-Drehs
Quelle: eigene Darstellung; (ver.di, 2024, S. 19)

		Ohne ARC-Play		Mit ARC-Play		
	<u>Position</u>	<u>Dauer</u>	<u>Kosten</u>	<u>Dauer</u>	<u>Kosten</u>	<u>Ersparnis</u>
Vorbereitung:	Regieassistentz	0,50 h	16,32 €	1,00 h	32,64 €	- 16,32 €
Dreh:	DoP	10,00 h	662,20 €	11,00 h	728,42 €	- 66,22 €
	1. AC	10,00 h	324,20 €	11,00 h	356,62 €	- 32,42 €
Rohschnitt	Editor	8,00 h	284,96 €	6,40 h	227,97 €	56,99 €
Gesamt			1.287,68 €		1.345,65 €	- 57,97 €

Würde in diesem Szenario ARC-Play verwendet werden, würde man für die Vorproduktion länger brauchen, da der Song erst in das System geladen und die Shotlist im System implementiert werden muss. In dieser Annahme würde es eine Stunde dauern. Während des Drehs gäbe es dadurch, dass die Takes in voller Länge des Songs aufgezeichnet werden, kaum einen Mehrwert und das System würde als einfaches Abspielsystem dienen. Durch die

Dokumentation der Takes benötigt der Dreh sogar eine Stunde länger. Durch die verwendete Kameratechnik bietet Timecode gegenüber der automatischen Synchronisierung allein durch den Ton keinen Mehrwert. Einzig durch das Metadatenhandling können Takes, die bereits während des Drehs als gelungen markiert wurden, schneller identifiziert werden, wodurch etwas Zeit gespart wurde (siehe Tabelle 2).

6.3.2 Komplexer Dreh mit digitaler Kamera

Diese Annahme geht von dem aus Kapitel 4.6 hervorgegangenen Typ b hervor:

Die Shotlist ist in diesem Fall etwas aufwendiger, da an mehreren Orten und dort auch teilweise mit unterschiedlichen Kameras gedreht wird. In dem Video werden Drohnenaufnahmen, Szenen mit mehreren Autos und daher mit größeren Distanzen aufgezeichnet und es werden Zeitlupeneffekte verwendet. Daher müssen verschiedene Varianten des Playbacks in unterschiedlichen Geschwindigkeiten vorbereitet werden. Die zuständige Person, die in diesem Fall der Regisseur ist, benötigt für die Vorbereitung zwei Stunden.

Mit der Digitalkamera werden alle Takes in voller Länge aufgenommen, außer bei den Zeitlupenszenen, da diese gezielt eingesetzt werden und die Künstler den Song nicht vollständig in dieser Geschwindigkeit performen können. Die Crew ist während des Drehs deutlich größer: Neben dem DoP gibt es eine Kameraassistentz (1st AC), einen Oberbeleuchter, eine Lichtassistentz und einen Regisseur. Es sind zwei Drehtage angesetzt mit einer gesamten Drehdauer von 20 Stunden. Während des Drehs werden Notizen für die Postproduktion erstellt. Da für einzelne Szenen zwischen verschiedenen Playbackvarianten gewechselt werden muss, kommt es manchmal zu Verzögerungen, in denen das gesamte Team warten muss.

Der Editor braucht für die Synchronisierung und den Rohschnitt zwölf Stunden. Dadurch, dass nicht in jeder Szene der Ton des Playbacks zu hören ist, wie zum Beispiel bei denen mit größeren Distanzen oder mit Drohne, benötigt das Synchronisieren relativ lange und es kann nicht immer automatisch synchronisiert werden. Außerdem muss die Geschwindigkeit der Zeitlupen angepasst und ebenfalls synchronisiert werden. Da es auch hier ein hohes Drehverhältnis gibt, dauern der Rohschnitt und der Vergleich der Takes ebenfalls etwas länger.

Tabelle 3: Die Ergebnisse der Berechnungen unter Annahme eines komplexen Drehs mit einer Digitalkamera
Quelle: eigene Darstellung; (ver.di, 2024, S. 19)

		Ohne ARC-Play	Mit ARC-Play	
--	--	---------------	--------------	--

		<u>Dauer</u>	<u>Kosten</u>	<u>Dauer</u>	<u>Kosten</u>	<u>Ersparnis</u>
Vorbereitung:	Regisseur	2 h	132,44 €	2 h	132,44 €	- €
Dreh:	DoP	20 h	1324,40 €	19 h	1258,18 €	66,22 €
	1. AC	20 h	648,40 €	19 h	615,98 €	32,42 €
	Oberbeleuchter	20 h	732,40 €	19 h	695,78 €	36,62 €
	Lichtassistentz	20 h	396,40 €	19 h	376,58 €	19,82 €
	Regisseur	20 h	1324,40 €	19 h	1258,18 €	66,22 €
Rohschnitt	Editor	12,00 h	427,44 €	4,00 h	142,48 €	284,96 €
Gesamt			4985,88 €		4.479,62 €	506,26 €

Mit ARC-Play würde die Vorbereitung, also die Shotlist-Erstellung und das Vorbereiten des Playbacks, in etwa gleich schnell funktionieren, da die Shots zwar in die Software des Systems eingefügt werden müssen, aber die verschiedenen Playbackvarianten spontan am Set erstellt werden können. Während des Drehs wird bei Aufnahmen mit größeren Distanzen und mit der Drohne die digitale Filmklappe ARC-Slate verwendet, um so den Timecode zu übermitteln. Es kann minimal Zeit gespart werden, da die Crew beim Wechseln des Playbacks nicht auf das Playbacksystem warten muss. Daher dauert der Dreh nun 19 Stunden. In der Postproduktion würde der gesamte Prozess des Synchronisierens komplett automatisiert geschehen, inklusive der Zeitlupen. Auch die Takeauswahl beschleunigt sich durch die bereits integrierten Metadaten. Dadurch verkürzt sich der Rohschnitt auf vier Stunden (siehe Tabelle 3).

6.3.3 Analoger Dreh

Der dritte Fall orientiert sich an Typ C aus Kapitel 4.6:

Da analoger Film sehr teuer ist, soll bei diesem Dreh nicht jeder Take in voller Songlänge gedreht werden, sondern es soll bereits vor dem Dreh feststehen, welcher Shot an welcher Stelle des Songs zu sehen sein soll, um so Material zu sparen. Daher dauert die Shotlisterstellung etwas länger. Zudem sollen auch hier verschiedene Geschwindigkeiten verwendet werden. Die verschiedenen Playbackvarianten werden ebenfalls in der Vorproduktion erstellt. Insgesamt benötigt der Regisseur dafür drei Stunden.

Die Crew des Drehs besteht aus acht Personen und ist auf drei Drehtage, also 30 Stunden, angesetzt. Dazu gehören ein DoP mit zwei Assistenten (1st AC und 2nd AC), eine Materialassistentz, ein Oberbeleuchter mit Lichtassistentz und ein Regisseur mit Regieassistentz. Während des Drehs werden die einzelnen Stellen des Playbacks mit dem Smartphone gesucht und mit einem Bluetooth-Lautsprecher abgespielt. Um diese Stellen

herauszusuchen, wird Zeit benötigt, in der das gesamte Team warten muss. Auch die weiteren Versionen des Playbacks auszuwählen, benötigt Zeit.

Der Rohschnitt ist sehr aufwendig und der Editor benötigt lange, um das Material mit dem Song zu synchronisieren. Da aufgrund der Beschaffenheit der Kamera kein Audio mit aufgezeichnet wurde und auch kein Timecode verwendet werden kann, kann man nur anhand der Choreographie oder Lippenbewegung das Bild mit dem Ton synchronisieren. Da die Takes nicht in voller Länge des Songs gefilmt werden, dauert dieser Prozess durch das Suchen der richtigen Stelle im Song länger. Allerdings ist durch die geringere Materiallänge das Anlegen des Materials kürzer. Insgesamt benötigt der Editor acht Stunden.

Für das Musikvideo wurde auf 35 mm mit vier Perforationen gedreht. Als Filmmaterial kam Kodak Vision3 in der Version 50D und 200T, ein Farbnegativfilm, zum Einsatz. Außerdem wurde ein Drehverhältnis von 1/5 angestrebt. Es muss also für 15 Minuten Filmmaterial eingekauft werden, was sich bei 35 mm und im Vier-Perforationen-Format auf 410,40 Meter beläuft. Für die Berechnung wurde der Online-Filmrechner von Ulrich Schmidt verwendet (Schmidt, 1999). Hierbei wurde aus Gründen der einfacheren Berechnung vernachlässigt, dass beim Einlegen der Filmrolle und beim An- und Auslaufen etwas Material verloren geht. Eine Rolle des verwendeten Films mit 122 Metern kostet bei Silbersalz 434,35 € (SILBERSALZ Film GmbH, o. J.). Für die genannte Länge beläuft sich der Preis auf 1.461,12 €. Die Laborpreise von Andec Cinefilm für das Entwickeln belaufen sich auf 1,30 €/Meter. Bei dieser Länge würde das 533,52 € kosten. Der Scan kostet beim gleichen Anbieter 350 € für 10 Minuten Filmmaterial in 4K und 16-Bit-DPX-Files. Für jede weitere Minute werden 27,50 € in Rechnung gestellt. Für die benötigten 15 Minuten sind daher 487,5 € fällig (Andec Filmtechnik GmbH, o. J.). Insgesamt belaufen sich die Kosten für das Verwenden von 15 Minuten analogem Film auf 2482,14 €.

Tabelle 4: Die Ergebnisse der Berechnungen unter Annahme eines Drehs mit analogem Film.
Quelle: eigene Darstellung; (ver.di, 2024, S. 19; SILBERSALZ Film GmbH, o. J.; Andec Filmtechnik GmbH, o. J.)

		Ohne ARC-Play		Mit ARC-Play		
		<u>Dauer</u>	<u>Gagen</u>	<u>Dauer</u>	<u>Gagen</u>	
Vorbereitung	Regisseur	3 h	198,66 €	2,7 h	178,79 €	19,87 €
Dreh:	DoP	30 h	1986,60 €	27 h	1787,94 €	198,66 €
	1. AC	30 h	972,60 €	27 h	875,34 €	97,26 €
	2. AC	30 h	731,40 €	27 h	658,26 €	73,14 €
	Material-Assistenz	30 h	731,40 €	27 h	658,26 €	73,14 €
	Oberbeleuchter	30 h	1098,60 €	27 h	988,74 €	109,86 €

	Lichtassistenz	30 h	594,60 €	27 h	535,14 €	59,46 €
	Regisseur	30 h	1986,60 €	27 h	1787,94 €	198,66 €
	Regieassistenz	30 h	979,20 €	27 h	881,28 €	97,92 €
Rohschnitt	Schnitt (Editor)	8 h	284,96 €	0,8 h	28,50 €	256,46 €
	<u>Film</u>	<u>Ratio</u>	<u>Kosten</u>	<u>Ratio</u>	<u>Kosten</u>	<u>Ersparnis</u>
Filmmaterial	Kodak Vision3 35 mm (4 Perf.)	1/5	2482,14€	1/4	2.000,72 €	481,42 €
Gesamt			12046,76 €		10380,91€	1665,85€

Mit ARC-Play würde die Vorbereitung für den Dreh schneller ablaufen, da das System für aufwendig geplanten Drehs ausgelegt ist und die Shotlist direkt im System angefertigt werden kann. Die einzelnen Geschwindigkeitsvariationen sind auch hier kein Problem und lassen sich immer flexibel einstellen. Von daher benötigt die Regieassistenz nur 90 % der Zeit, die sie ohne ARC-Play benutzt hätte. Der Dreh würde durch die perfekt vorbereiteten Shots und das damit einhergehende Playback schneller ablaufen. Er dauert nun 90 % der ursprünglichen Zeit, also 27 Stunden. In der Postproduktion ist die Effizienz am höchsten, weil sich mit ARC-Play sowohl das Synchronisieren als auch der Rohschnitt komplett automatisieren lassen. Der Editor müsste sich nur noch zwischen den einzelnen Takes für jede Stelle entscheiden und bräuchte für den Rohschnitt weniger als eine Stunde. Da ARC-Play, wie in Kapitel 5.1 beschrieben, ein Feature hat, mit dem man das Schnittverhältnis genau über den Song betrachten kann und dadurch eventuell dieses minimal verbessern kann, liegt dieses unter idealen Bedingungen bei 1:4 anstatt 1:5 (siehe Tabelle 4).

Wenn die Annahme abgewandelt wird, indem davon ausgegangen wird, dass bei einem Dreh ohne das System für jeden Shot der Song in maximaler Länge aufgenommen wird, wäre der Nutzen von ARC-Play natürlich erheblich höher. Allerdings ist, wie in den Experteninterviews aus Kapitel 4.4 herauskam, in der Praxis nicht davon auszugehen.

6.4 Diskussion der Ergebnisse der Berechnungen

Durch die Berechnungen kann man erahnen, dass sich mehr Geld sparen lässt, je aufwendiger der Dreh und je unflexibler das verwendete Equipment ist. Gerade bei analogem Film unterstützt das Tool bei der Planung und kann unter bestimmten Voraussetzungen auch Material sparen. Allerdings gibt es auch Hinweise darauf, dass bei Produktionen, die ohne starres Muster auskommen und bei denen viel aus dem Moment heraus und spontan passiert, der Nutzen von ARC-Play schnell schwindet. In diesen Fällen wäre das System nicht rentabel und würde sogar eher einschränken.

Allerdings gibt es auch Vorteile von diesem System, die nicht unbedingt berechenbar sind: Da ARC-Play vermehrt auf Automatisierung setzt und das Benutzen von Strukturen wie Shotlists und Metadaten erzwingt, kann das Risiko für menschliche Fehler gesenkt werden. Beispielsweise behält man mit diesem System den Überblick über die Parts des Songs, die bereits gedreht wurden, und man läuft weniger in der Gefahr, einzelne Teile mit zu wenigen Takes auszustatten oder gar komplett zu vergessen.

Außerdem können mit diesem System die Künstler entlastet werden, weil man nicht mehr in der Versuchung ist, immer den kompletten Song zu drehen, sondern ihn in Parts aufzuteilen. Dadurch kann beispielsweise bei Songs mit Gesang die Stimme geschont werden. Ein Gegenargument wäre, dass Künstler, wenn sie den gesamten Song performen, eventuell authentischer performen.

Wichtig zu erwähnen ist auch, dass sich Zeiteinsparungen nicht unbedingt nur mit finanziellem Nutzen rechtfertigen lassen: Wenn beispielsweise der Editor weniger Zeit benötigt, um das gesamte Material zu synchronisieren, bleibt ihm mehr Zeit für kreative Entscheidungen im Schnitt. Genauso kann es die Crew am Set entlasten, wenn man sich über das Playbacksystem keinerlei Gedanken mehr machen muss, weil das bereits vorbereitet wurde, und es bleibt mehr Zeit, sich auf gestalterische Maßnahmen zu konzentrieren. Dadurch könnte die Qualität von Produktionen indirekt durch ein solches System gesteigert werden.

7 Technische Umsetzung

Für die Hardwarekomponenten wurden ähnliche Komponenten verwendet, die bereits bei der digitalen Filmklappe ARC-Slate verwendet wurden und sich dabei bewährt haben. Daher beruht ein Teil des Entscheidungsprozesses auf bereits vorhandener Erfahrung mit der Hardware. Alle Komponenten dienen dazu, den in Kapitel 5 dargestellten Anforderungen gerecht zu werden (siehe Abbildung 6).

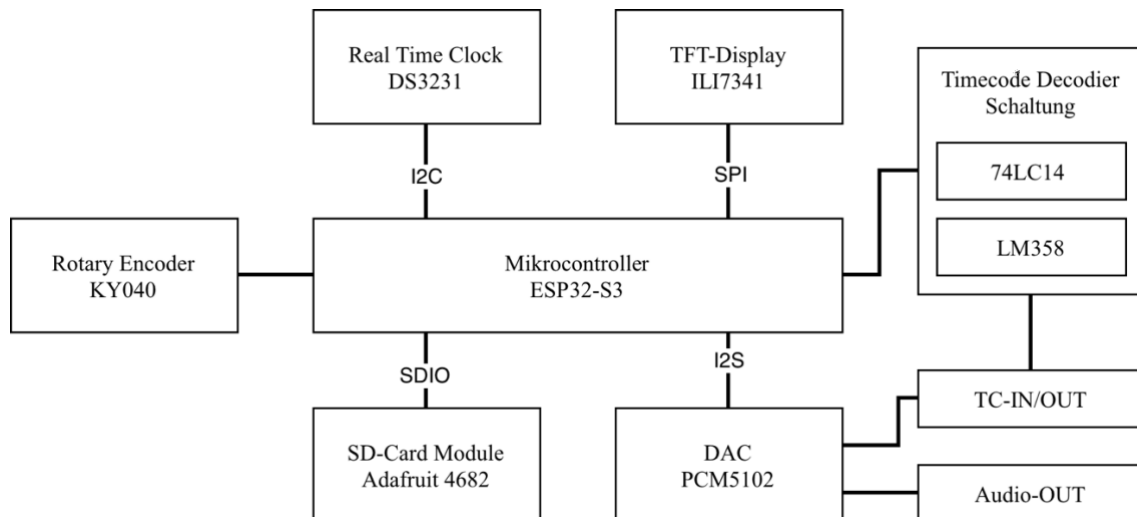


Abbildung 6: Schematische Darstellung der Elektronik von ARC-Play
Quelle: eigene Darstellung

7.1 Hardwarekomponenten

ESP32-S3

Als Herzstück des Geräts wurde ein ESP32-S3 gewählt. Im Speziellen wurde ein ESP32-S3 Dev Board ausgewählt, also ein Entwicklerboard, bei dem die GPIOs auf Pins verfügbar gemacht wurden, um mit einem Steckbrett die Elektronik für das System zu entwickeln. Der Mikrocontroller zeichnet sich durch einen leistungsstarken Prozessor mit zwei Prozessorkernen aus und verfügt über eine hohe Konnektivität via Bluetooth und WLAN (Espressif Systems, 2026). Da der Mikrocontroller für ARC-Play mehrere Aufgaben mit hoher Performance parallel ausführen muss, also das Enkodieren und Dekodieren von Timecode und gleichzeitig die WLAN-Verbindung und die Displayanzeige, aber auch ein gutes Schnittstellenangebot haben muss, eignet er sich gut für diese Aufgabe. Außerdem ist das Preis-Leistungs-Verhältnis des Mikrokontrollers sehr gut, weshalb er sich schnell gegen andere Modelle durchgesetzt hat.

DS3231

Der ESP32-S3 hat verschiedene Schnittstellen, mit denen weitere Komponenten verbunden werden können. Unter anderem verfügt er über eine I2C-Schnittstelle. Diese besteht aus einem SDA-(serial data) und einem SCL-(serial clock) Signal (Plate, 2017, S. 39–40). Bei ARC-Play wurde damit eine Real-Time-Clock (RTC) mit der Bezeichnung DS3231 verbunden. Diese enthält einen temperaturnausgleichenden Kristalloszillator, der eine sehr geringe Abweichung der aktuellen Zeit über einen langen Zeitraum verspricht. Er verfügt außerdem über eine Knopfatterie, um die Zeit noch zu speichern, falls die Stromversorgung unterbrochen wird. Neben der aktuellen Zeit im Stunden-, Sekunden- und Minutentakt wird auch das aktuelle Datum gespeichert, das auch automatisch die unregelmäßige Anzahl der Monate ausgleicht. (Maxim Integrated Products, Inc., 2015). Zusätzlich zur aktuellen Zeitinformation, die über die I2C-Schnittstelle ausgegeben wird, werden auch ein 32-kHz-Signal und ein programmierbares Rechtecksignal, das standardmäßig auf 1 Hz ausgelegt ist, ausgegeben. Die Uhrzeit kann über den ESP32-S3 eingestellt und programmiert werden. Für ARC-Play dient die Real-Time-Clock vor allem zur Berechnung des Timecodes. Bei Free-Run wird die aktuelle Uhrzeit der Real-Time-Clock für den Timecode direkt verwendet, bei Record-Run wird das 1-Hz-Rechtecksignal verwendet, um den Timecode möglichst akkurat zu halten.

PCM5102

Des Weiteren verfügt ARC-Play über einen DAC, also einen Digital-Analogwandler, mit der Bezeichnung PCM5102. Dieser ist dafür zuständig, Ton abzuspielen. Er besitzt zwei Kanäle, bei denen der eine für den LTC und der andere für das Playback verwendet wird. Es werden 16-, 24- und 32-Bit-Audiodaten unterstützt. Verbunden wird der DAC via I2S (Texas Instruments Inc., 2012). Dabei handelt es sich um eine Verbindung, die auf Audio ausgelegt ist. Für die Übertragung des Signals sind drei Verbindungen notwendig. Das PCM5102 verfügt zudem über mehrere Pins, die durch das Anlegen von Masse oder 3,3 V verschiedene Einstellungen hervorrufen können. Zu erwähnen ist hierbei der XSMT-Pin, der steuert, ob das Gerät stummgeschaltet wird, und der auf der Platine so verlötet werden muss, dass das Gerät dauerhaft aktiv ist. Das Modul, das in ARC-Play eingebaut wird, besitzt außerdem verschiedene Pins und einen 3,5-mm-Stereo-Audioausgang. Dieser wird allerdings bei diesen Prototypen nicht verwendet. Stattdessen werden zwei extra 3,5-mm-Audiobuchsen benutzt, bei denen jeweils nur die Mono-Kanäle verwendet werden.

Displaytreiber ILI9341

Das Display des Prototyps ist ein TFT-Display mit der Bezeichnung ILI9341 (ILI Technology Corp., o. J.). Dabei handelt es sich um einen Displaytreiber, der über eine SPI-Schnittstelle gesteuert werden kann. Diese wurde von Motorola entwickelt und kann Daten synchron seriell ausgeben und gleichzeitig einlesen. Sie verfügt über vier Signale: Serial Clock (SCL), Master Output/Slave Input (MOSI), Master Input/Slave Output (MISO) und Chip Select (CS). Dabei können mehrere Geräte parallel angeschlossen werden, allerdings benötigt jedes weitere Gerät einen eigenen CS-Pin (Plate, 2017, S. 35). Das TFT-Display hat eine Diagonale von 2,8 Zoll und eine Auflösung von 240 x 320 Pixeln. Es ist prinzipiell touch-fähig, darauf wird allerdings bei ARC-Play verzichtet und stattdessen auf einen Rotary-Encoder gesetzt. Außerdem verfügt es zudem über ein SD-Karten-Modul, das ebenfalls über SPI gesteuert wird, womit es allerdings bei der Entwicklung des Prototyps zu Problemen kam.

SD-Card-Reader

Da der ESP32-S3 nur einen begrenzten Speicher hat, wird für das Speichern der WAV-Datei des Audio-Playbacks das SD-Card-Modul Adafruit 4682 verwendet. Dieses soll zuerst ebenfalls mit der SPI-Schnittstelle verwendet werden. Allerdings hat das nach wiederholten Versuchen nicht funktioniert, weshalb auf das SDIO-Protokoll zurückgegriffen wurde. Das gibt es in unterschiedlichen Ausführungen. Da die Übertragungsrate nicht sonderlich hoch sein muss und auch an Pins gespart werden soll, wurde der 1-Bit-Übertragungsmodus gewählt. Dafür wird neben den CMD- und CLK-Pins nur ein Data-Pin (DAT 0) verwendet (SD Card Association, 2018). Neben dem Song werden auch aktuelle Daten über den Song wie der Songtitel, die Taktart, die BPM-Zahl und die Songdauer als JSON-Datei auf der SD-Karte gespeichert.

Rotary Encoder KY040

Zur Bedienung des Geräts wurde sich für einen Rotary-Encoder des Modells KY-040 entschieden. Dieser schaut auf den ersten Blick so aus wie ein normales Potentiometer. Allerdings hat er keinen Start- und Endpunkt, er rastet ein, wenn man ihn dreht, und er lässt sich drücken. Daher eignet sich ein Rotary-Encoder hervorragend zur Bedienung eines Softwaremenüs. Neben der Stromversorgung benötigt er nur drei weitere Kabel zur Verbindung mit dem ESP32-S3 (AZ-Delivery Vertriebs GmbH, o. J.).

7.2 Elektronik zur Timecode-Dekodierung

Da ARC-Play neben seinen Eigenschaften als Playbacksystem auch als normaler Timecode-Generator fungieren soll und somit auch via Jam-Sync synchronisiert werden soll, soll das Gerät nicht nur Timecode generieren, sondern auch lesen, also dekodieren, können. Für den Prototyp werden zunächst 24, 25, 29,97 und 30 Bilder pro Sekunde unterstützt. Um das zu erreichen, muss mit analogen Elektronikbauteilen das Eingangssignal modifiziert werden, damit es für den ESP32-S3 lesbar wird. Dieser wird mit 3,3 V betrieben und kann daher nur Signale zwischen 0 V und 3,3 V lesen (Espressif Systems, 2026). Wie in Kapitel 3.1 dargestellt, ist das Timecodesignal ein Audiosignal, womit es um 0 V schwingt und dessen Spitzenspannung zwischen 0,5 V und 4,5 V liegt. Daher muss das Timecodesignal an die Gegebenheit des ESP32-S3 angepasst werden. Die Schaltung wurde zuvor mit dem kostenlosen Programm LTspice simuliert, um so auf die geeigneten Werte zu kommen.

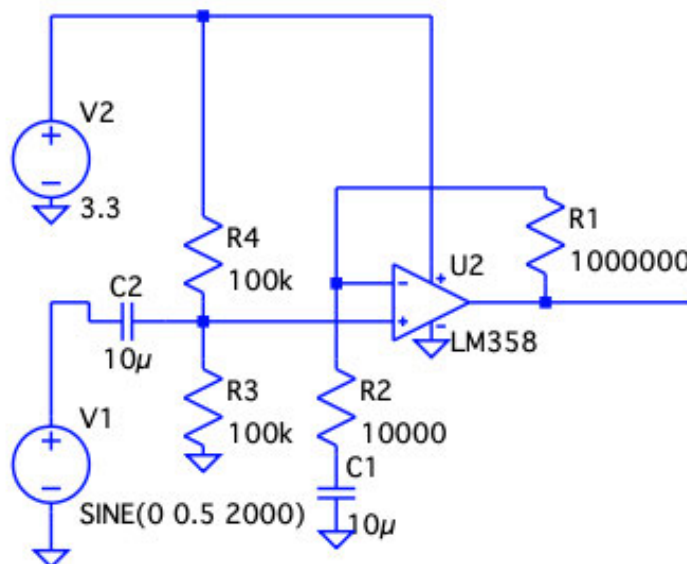


Abbildung 7: Aufbau der Simulation der Verstärkerschaltung für das Timecodesignal in LTspice
Quelle: eigene Darstellung

Um die Verstärkung auch für niedrige Signale auszurichten, wurde von einem Sinussignal mit 2 kHz und 0,5 V Amplitude für die Simulation ausgegangen. Das Signal muss zuerst angehoben werden, damit keine Spannung unter 0 V ankommt. Dies wird mit einem Spannungsteiler erreicht (Mancini, 2003, S. 7). Bei dieser Schaltung wird die Leitung, auf der das Signal liegt, über je einen 100-kOhm-Widerstand (R3) mit der Masse und über einen Widerstand (R4) mit 3,3 V verbunden (siehe Abbildung 7). Dadurch bekommt das Timecodesignal einen DC-Offset von 1,65 V (siehe Abbildung 8). Dieses Signal wird nun mit einem Operationsverstärker verstärkt, in diesem Fall mit dem Modell LM358 (Texas

Instruments Inc., 2024). In diesem Modul sind zwei Operationsverstärker eingebaut, für die Timecode-Dekodierung wird allerdings nur einer verwendet. Die Schaltung ist in Form eines nicht invertierenden Verstärkers konzipiert (Mancini, 2003, S. 19). Die beiden Widerstände R1 und R2 stellen dabei die Gesamtverstärkung ein. In dieser Konfiguration ergibt sich eine hundertfache Verstärkung des Signals. Dabei kann es passieren, dass das Signal übersteuert. Dies ist aber zu akzeptieren, da so möglichst schwache Signale noch dekodiert werden können und bei der Timecode-Dekodierung nur die Flankenerkennung ausschlaggebend ist (siehe Abbildung 8). Um sowohl das Eingangssignal als auch das Verstärkungssignal von Gleichspannung zu entkoppeln, wurde jeweils ein 10-nF-Kondensator verwendet. Die Impedanz eines Kondensators wird bei einer Frequenz von null unendlich (Mancini, 2003, S. 25–26). Das bedeutet, Gleichspannungsanteile werden blockiert.

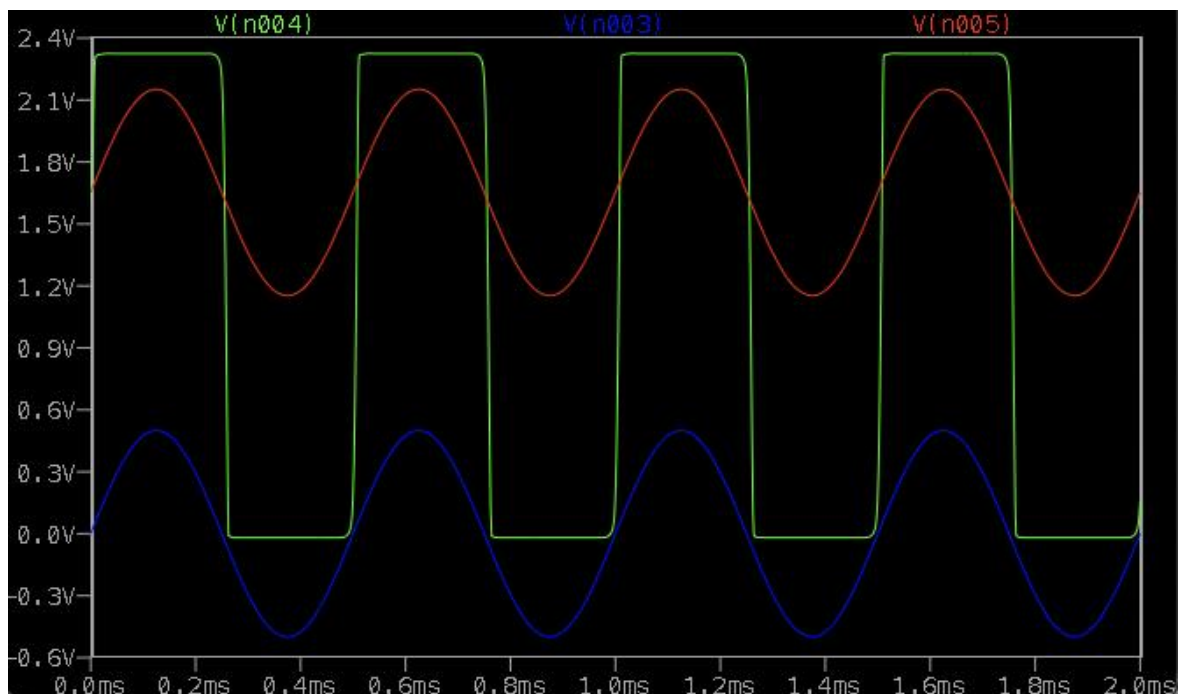


Abbildung 8: Simulation der Verstärkung des Timecodesignals mit LTspice ohne Schmitttrigger (Blau: Timecodesignal, Rot: Timecodesignal mit DC-Bias, Grün: verstärktes Timecodesignal)
Quelle: eigene Darstellung

Das verstärkte Signal wird im Anschluss noch mit einem Schmitt-Trigger für den ESP32-S3 aufbereitet. Bei einem Schmitt-Trigger handelt es sich um einen Schwellenwertschalter mit Hysterese (Reinhold, 2010, S. 241). Aus dem Datenblatt des verwendeten Schmitt-Triggers, einem 74LC14, ist zu entnehmen, dass bei einer Betriebsspannung von 4,5 V unter Normalbedingungen der positive Schaltschwellenwert bei 2,5 V liegt (Texas Instruments Inc., 2025, S. 5). Da durch die Verstärkerschaltung bei einer Betriebsspannung von 3,3 V, wie auf Abbildung 8 zu sehen, das Signal auf über 2,3 V verstärkt wird, ist davon auszugehen, dass der Schmitt-Trigger ausgelöst wird. Das Signal wird dadurch in ein

sauberes Rechtecksignal zwischen 0 V und 3,3 V umgewandelt und kann vom ESP32-S3 von einem digitalen Pin gelesen und entsprechend interpretiert werden.

Der ESP32-S3 ist mit dieser Schaltung nun in der Lage, LTC zu dekodieren. Mit den Timecodesignalen des aktuellen Prototyps der ARC-Slate und dem Audioausgang verschiedener Geräte hat das Dekodieren des Timecodes problemlos funktioniert. Einzig wenn manche Konsumergeräte Timecode abgespielt haben, wurde dieser nicht immer zuverlässig erkannt.

7.3 Platinendesign

Um die Komponenten für ARC-Play robust zu platzieren, wurde mit der kostenlosen Software KiCAD eine Platine erstellt und anschließend auf der Website aisler bestellt. Auf der Platine können in zwei Schichten Leiterbahnen gezeichnet werden. Die Rückseite wurde dabei vollständig als Masse definiert. Auf der Vorderseite wurden alle restlichen Signale miteinander verbunden. Falls es Situationen gab, an denen sich Leiterbahnen gekreuzt hätten, wurde mit Vias auf die Rückseite gewechselt, um dort von der Masse isoliert den Weg fortzusetzen. Die Stromversorgung des Prototyps erfolgt derzeit über den USB-Anschluss des ESP32-S3, allerdings ist auf der Platine ein Stromanschluss vorhanden, um die Stromversorgung für einen späteren Prototyp beispielsweise mit einem Akku zu gewährleisten. Die Platine wurde zudem so gestaltet, dass sie sowohl für ARC-Play als auch für die Filmklappe ARC-Slate benutzt werden kann. Dazu kann der Anschluss des Displays und der der SD-Karte mit anderen Steckern verwendet werden. Die Platine wurde im Anschluss mit den Komponenten bestückt, indem diese aufgelötet wurden (siehe Abbildung 9). Um dabei Fehler zu vermeiden, wurde die Platine entsprechend beschriftet.

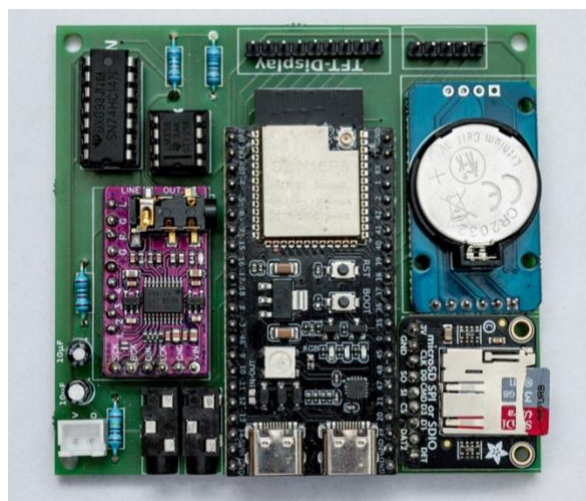


Abbildung 9: Die bestückte Platine von ARC-Play
Quelle: eigene Darstellung

7.4 3D-Druck-Gehäuse

Das Design von ARC-Play soll sich, wie bereits in Kapitel 5 beschrieben, grob an das Design des Apple iPods orientieren. Für den Prototyp soll das bereits leicht angedeutet umgesetzt werden. Der Rotary-Encoder soll dabei ähnlich wie das Click-Wheel von Apple unterhalb des Displays nach vorne zeigen. Es wurden eine Gehäuseschale, in der die Platine und der Rotary-Encoder montiert sind, und eine Gehäusedecke, in der das Display integriert ist, mit der kostenlosen Software FreeCAD erstellt und mit einem 3D-Drucker ausgedruckt. Zur einfacheren Handhabung wurde die Platine als 3D-Objekt aus KiCAD exportiert und in das 3D-Modell eingefügt (siehe Abbildung 10). Um beide Teile zu vereinen, sind Löcher für Gewindebohrungen vorgesehen, mit denen mit M3-Schrauben das Gehäuse zusammengebaut werden kann. An den Seiten wurden entsprechende Aussparungen für den USB-C-Anschluss des ESP32-S3, den Audio-Ausgang und für den Timecode-Ein- und Ausgang hinzugefügt.

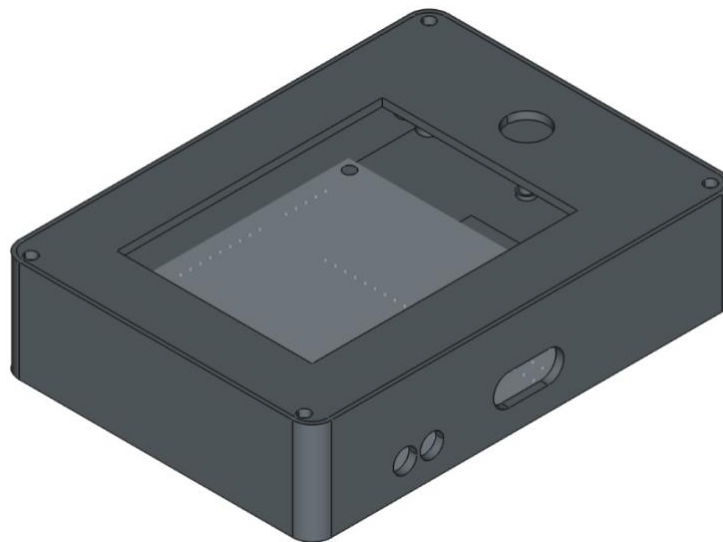


Abbildung 10: Das 3D-Objekt des Gehäuses von ARC-Play in FreeCAD
Quelle: eigene Darstellung

8 Software

Die Software von ARC-Playback ist in zwei Bereiche gegliedert: Zum einen gibt es die Firmware, die auf den ESP32-S3 geladen wird, zum anderen gibt es den Server, der auf einem Gerät, das am Dreh lokal vorhanden ist, wie beispielsweise einem Laptop, läuft. Dieser Server besitzt ein Frontend, das im Browser geöffnet werden kann und in das sich auch mobile Geräte, die sich im gleichen Netzwerk befinden, verbinden können. Für das Frontend gibt es die beiden Unterseiten ARC-Studio und ARC-Analyzer.

Für die Programmierung der Firmware und der Software wurde unterstützend das KI-Tool GitHub Copilot verwendet. Alle implementierten Codeabschnitte wurden überprüft und auf ihre Funktionalität im Gesamtkontext für das ARC-System getestet. Die Verwaltung des Codes erfolgte mittels Git auf der Plattform GitHub.

8.1 Firmware des ESP32-S3

Sobald ARC-Play mit der Stromversorgung verbunden wurde, fährt es hoch. Im Hintergrund wird automatisch, wenn verfügbar, die aktuelle Zeit des NTP-Servers abgerufen und mit dem RTC synchronisiert. Daraufhin startet das Display, auf dessen Oberfläche zu jedem Zeitpunkt alle wichtigen Informationen dargestellt werden. Nach dem Start wird direkt der laufende Timecode im Free-Run-Modus angezeigt und auch schon über den TC-Ausgang ausgegeben. Oben rechts wird dauerhaft die aktuelle Uhrzeit angezeigt und links ein Indikator, ob das Gerät gerade Timecode beziehungsweise das Playback abspielt oder pausiert. Außerdem werden immer der aktuelle Modus, die Framerate und die User-Bits angezeigt (siehe Abbildung 12). Um das Gerät verbinden zu können, werden nach dem Bootvorgang unten das verbundene WLAN-Netzwerk und die IP-Adresse des Geräts angezeigt. Wenn der Rotary-Encoder gedrückt wird, wird der Timecode pausiert oder gestartet. Wird er gedreht, wird die Lautstärke angepasst. Dazu erscheinen auf dem Display eine Balken- und Prozentanzeige, die nach kurzer Zeit wieder verschwinden. Um die Einstellungen aufzurufen, muss lange auf den Encoder gedrückt werden. Es erscheint ein Menü-Overlay, mit dem verschiedene Optionen ausgewählt werden, die in sich wiederum weitere Einstellungsmenüs beinhalten (siehe Abbildung 11). Im Einstellungsmenü können der Modus und die Framerate eingestellt werden. Außerdem kann per Knopfdruck erneut der RTC mit dem NTP-Server synchronisiert werden und die SD-Karte formatiert werden.

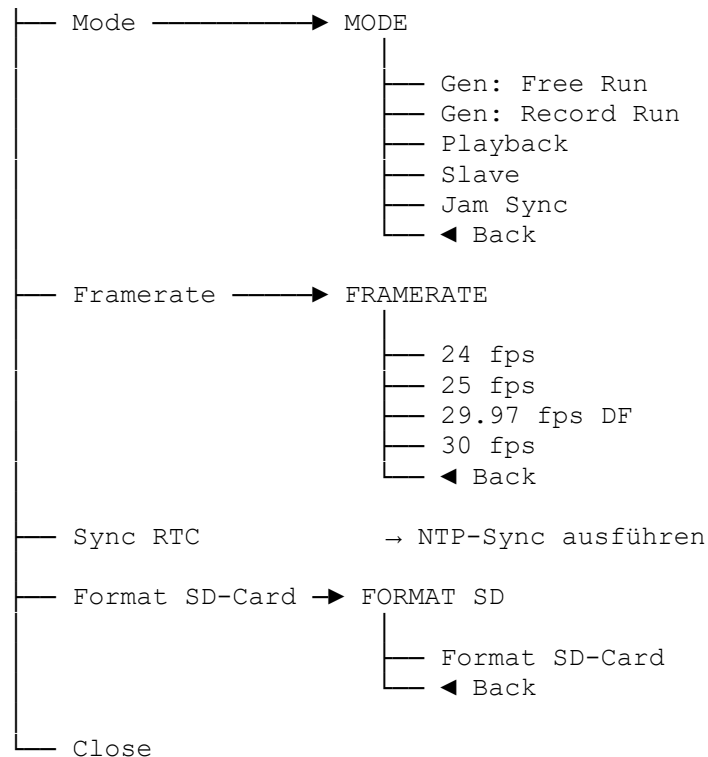


Abbildung 11: Die Menüstruktur von ARC-Play
Quelle: eigene Darstellung

Sobald ein Shot auf ARC-Play geladen wurde und sich das Gerät im Playback-Modus befindet, werden alle Informationen, die auch auf einer Klappe angezeigt werden, auf dem Gerät angezeigt. Zudem werden der Start- und Endtimecode, die Taktart und Geschwindigkeit des Songs, der Songtitel, die Gesamtlänge und der aktuelle Zeitpunkt im Song anhand eines Fortschrittbalkens angezeigt. In diesem ist auch farbig markiert, welchen Bereich des Songs der aktuelle Shot beinhaltet (siehe Abbildung 12). Wird ein Song abgespielt, wird mit vier Schlägen eingezählt. Die Geschwindigkeit orientiert sich an der des Songs und der eingestellten Geschwindigkeit.

Mit doppeltem Klicken und anschließendem Drehen des Rotary-Encoders kann zwischen den Shots gewechselt werden. Dadurch kann das Gerät fast ausschließlich ohne Zweitgerät benutzt werden. Wird ein Shot abgespielt und das Ende des Abschnitts erreicht, wird der Timecode automatisch auf den Start-Timecode des Shots zurückgesetzt und die Takenummer automatisch erhöht. Außerdem wird eine JSON-Nachricht mit allen relevanten Informationen wie der Takenummer, dem Datum und den verwendeten User-Bits an den Server geschickt.

Um eine dauerhafte Synchronität der Einstellungen zwischen ARC-Play und dem Server zu gewährleisten, wird bei jeder Statusänderung eine JSON-Nachricht an den Server geschickt. Dadurch werden alle Einstellungen jederzeit synchronisiert.



Abbildung 12: Der Prototyp von ARC-Slate im Playback-Betrieb
Quelle: eigene Darstellung

Auf dem ESP32-S3 wird mit der gleichen Programmiersprache programmiert wie für Arduinos, also einer Abwandlung von C++. Für die Firmware wurden daher verschiedene Bibliotheken verwendet, die über die Arduino-IDE oder mit PlatformIO eingebunden werden können. Letzteres wurde in diesem Projekt verwendet. Dazu gehören Bibliotheken, die benutzt werden, um bestimmte Funktionen des ESP32-S3 oder die verwendete Hardware zu aktivieren. Außerdem werden bestimmte Bibliotheken benötigt, um Prozesse auf dem Mikrokontroller auszuführen. Im Folgenden werden ausgewählte Bibliotheken beschrieben:

Zentral ist die umfangreiche Bibliothek „Arduino Audio Tools“ von Phil Schatzmann (Schatzmann, 2026). Mit dieser Bibliothek werden alle Prozesse der Audioverarbeitung abgehandelt. Mit AudioSourceSDMMC.h kann ein Audiofile über eine SD-Karte abgespielt werden. Dafür muss zusätzlich SD_MMC.h von Arduino eingebunden werden. Über CodecWAV.h kann die gegebene WAV-Datei decodiert werden. Außerdem wird über die Audiobibliothek die Ausgabe über die I2S-Schnittstelle mit dem PCM5102 abgewickelt.

Eine wichtige Bibliothek, die für die Timecode-Codierung und -Encodierung verwendet wird, ist die libltc.h-Bibliothek von Robin Gareus (Gareus, 2022). Diese ist ursprünglich als reine C-Anwendung erstellt worden. Mit Hilfe von Github Copilot wurde die komplette Bibliothek umgeschrieben und an die Hardware des ESP32-S3 angepasst. Somit funktioniert

sie mit der vorhandenen Technik. Mit der Bibliothek kann Timecode erstellt und über den PCM5102 ausgegeben werden. Der Timecode wird dabei ausschließlich über den linken Channel des Digital-Analog-Wandlers ausgegeben, der rechte wird für die Audioausgabe verwendet. Die Bibliothek ermöglicht zudem die Berechnung der Dropframes des Timecodes und das Modifizieren der User-Bits. Diese beinhalten standardmäßig das Datum und die Zeitzone in hexadezimaler Schreibweise, die von anderen Timecode-Monitoren gelesen werden können. Die Real-Time-Clock DS3231 liefert für diese Funktion das Datum. Im Playback-Modus werden, wie in den Anforderungen geschrieben, die User-Bits zur Take-Identifikation verwendet. Dazu werden die aktuelle Kamera, Roll und Szene und der aktuelle Shot und Take in die User-Bits geschrieben. Wie in Kapitel 3.1 beschrieben, sind dafür 32 Bits im hexadezimalen Format nutzbar. Für die Kamerabezeichnung werden 4 Bit verwendet und dabei die aus der hexadezimalen Schreibweise verwendeten Buchstaben A-F zur Verfügung gestellt. Alle anderen Werte nutzen 7 Bits und haben somit Werte zwischen 0 und 127 zur Verfügung. Dadurch lassen sich alle Informationen für spezifische Takes abdecken. Die Bi-Phase-Codierung wird ebenfalls von dieser Bibliothek übernommen.

Um eine Verbindung mit dem ARC-Studio aufzubauen, wird eine Websocket-Verbindung aufgebaut, also eine Verbindung, die dauerhaft geöffnet ist und in der beide Clients ohne Anfrage Daten verschicken können (Steyer, 2025, S. 337–338). Für die Firmware wurde dazu die ESPAsyncWebServer.h-Bibliothek eingebunden (ESP32 Asynchronous Networking, 2026).

Die Real-Time-Clock DS3231 von Adafruit wird mit der Bibliothek RTC-lib.h gesteuert (Adafruit Industries, 2024). Um diese mit der aktuellen Uhrzeit zu synchronisieren, wird sie entweder mit einer anderen Timecode-Quelle synchronisiert oder über WLAN über das Network Time Protocol (Mills et al., 2010) synchronisiert. Über den SQW-Pin gibt die RTC ein 1-Hz-Squarewave-Signal aus, das einen Hardware-Interrupt auslöst, womit der Timecode direkt auf die über einen längeren Zeitraum stabile Zeit des RTC synchronisiert wird.

Für das Display wird die für den Displaytreiber ILI9341 benötigte Adafruit-Bibliothek Adafruit_ili9341 verwendet (Adafruit Industries, 2026). Damit ist es möglich, auf dem Display verschiedene Elemente oder Text anzuzeigen und bestimmte Farben darzustellen.

Die Firmware des ESP32-S3 ist darauf ausgelegt, via Multithreading mehrere Tasks parallel auszuführen. Dafür können Funktionen als Tasks definiert werden, die dann einem bestimmten Prozessorkern zugeordnet werden und dann parallel ausgeführt werden können.

Der ESP32-S3 nutzt dafür FreeRTOS, eine Abwandlung des Betriebssystems Vanilla FreeRTOS (Espressif Systems, o. J.). Die Tasks, die explizit auf verschiedene Prozessorkerne ausgelagert werden, sind die der Display-Darstellung, die der Timecode-Kodierung und die der Timecode-Dekodierung. Letztere werden allerdings nie parallel laufen, da Timecode entweder geschrieben oder gelesen wird.

Um zeitkritische Prozesse im System nicht zu unterbrechen, wird mit Hardware-Interrupts gearbeitet. Der ESP32-S3 verfügt über 32 Hardware-Interrupts (Espressif Systems, 2026, S. 36). Diese können bestimmten Ereignissen zugeordnet werden. Außerdem kann der Modus angepasst werden, welches Verhalten des Signals den Interrupt auslösen soll. Im Fall von ARC-Play wurden drei Interrupts definiert: Dem Encoder-CLK-Pin wurde ein Interrupt hinzugefügt, um das System zu unterbrechen, wenn der Knopf gedrückt wurde. Der RTC-SQW-Pin wurde mit einem Interrupt ausgestattet, um den Timecode an den Sekundentakt des RTCs anzupassen. Der LTC-IN-Pin wurde mit einem Interrupt ausgestattet, um die Veränderung des Timecodesignals und somit dessen Flanken zu erkennen.

8.2 Node.js-Server

Node.js ermöglicht die Implementierung eines Webservers auf Basis der Programmiersprache JavaScript. Im Rahmen dieser Arbeit fungiert der Node.js-Server als Backend-Komponente. Seine Hauptaufgabe besteht darin, die für die Frontend-Applikationen ARC-Studio.html und ARC-Analyzer.html erforderlichen Dateien bereitzustellen. Dazu implementiert das Backend eine REST-API, welche als zentrale Kommunikationsschnittstelle zwischen Client und Server dient. Alle projektbezogenen Daten werden auf einer JSON-Datei mit dem Namen state.json gespeichert. Dort befinden sich alle Projektinformationen, die komplette Shotlist und die Informationen des Songs, inklusive aller berechneten Zeitstempel der Beats. Sobald sich etwas am Status ändert, wird das direkt in diese Datei geschrieben. Dadurch wird ermöglicht, dass selbst bei einem Absturz des Servers die Dateien erhalten bleiben und bei jedem Neustart die bereits vorhandenen Dateien sofort angezeigt werden, wodurch sich eine starke Verbesserung der Nutzererfahrung im Vergleich zur Vorgängerversion aus Kapitel 2 ergibt.

Eine REST-API erlaubt den standardisierten Zugriff auf die Ressourcen eines Servers. Die Interaktion folgt einem zustandslosen Client-Server-Modell: Ein Client initiiert eine Anfrage an den Server, welcher diese verarbeitet und eine entsprechende Antwort zurücksendet. Jede Ressource des Servers ist über eine eindeutige URI adressierbar. Die Kommunikation erfolgt

über das HTTP-Protokoll, wobei die Operationen auf den Ressourcen durch die standardisierten HTTP-Methoden definiert werden:

- GET: Fordert die Daten einer spezifischen Ressource an. Gemäß der HTTP-Spezifikation gilt diese Methode als sicher; sie liest den Zustand der Ressource auf dem Server lediglich aus, ohne ihn zu modifizieren.
- POST: Übermittelt Daten an den Server, um eine neue Ressource zu erstellen.
- PUT: Aktualisiert eine bestehende Ressource vollständig oder erstellt diese, falls sie unter der angegebenen URI noch nicht existiert.
- DELETE: Entfernt die spezifizierte Ressource dauerhaft vom Server.

Ein wesentliches Architekturmerkmal von REST ist die Zustandslosigkeit. Dies impliziert, dass jede Anfrage eines Clients an den Server isoliert zu betrachten ist und sämtliche für die Verarbeitung notwendigen Informationen, wie z. B. Authentifizierungsdaten, explizit enthalten muss. Auf dem Server werden zwischen den Anfragen keine clientbezogenen Sitzungskontexte gespeichert, wodurch eine serverseitige Sitzungsverwaltung entfällt (Balzert, 2011, S. 286–289).

Für die Interaktion zwischen dem Backend-Server, den webbasierten Frontends ARC-Studio und ARC-Analyzer sowie dem ESP32-S3 in ARC-Play wurde eine REST-Schnittstelle implementiert. Die Schnittstelle ist so konzipiert, dass sie sowohl die Bereitstellung der Benutzeroberfläche als auch die Echtzeit-Abfrage und -Modifikation des Systemstatus sowie die Dateiverarbeitung abhandelt. Beispielsweise kann über die API der Status im `state.json` gelesen oder geschrieben werden, es kann der Song auf den Server hochgeladen werden oder es wird die Startseite des ARC-Servers ausgeliefert. Datenströme und Zustandsänderungen werden vorwiegend im JSON-Format übermittelt.

8.2.1 ARC-Studio

Wenn sich das Gerät, auf dem der Node.js-Server läuft, mit dem gleichen Netzwerk wie ARC-Play verbunden ist, ist es mit ARC-Studio möglich, das Playbacksystem ARC-Play zu steuern. Um den ESP mit dem Server zu verbinden, wird die IP-Adresse, die beim Start von ARC-Play auf dessen Display angezeigt wird, manuell eingegeben. Über einen farbigen Punkt wird dargestellt, ob das Gerät verbunden ist. Rot bedeutet keine Verbindung, Orange bedeutet, dass eine Verbindung aufgebaut wird, und Grün bedeutet, dass eine Verbindung besteht.

Es lassen sich alle Parameter von ARC-Play über ARC-Studio steuern, also Play, Pause, TC-Reset, der TC-Modus, die Framerate, die Lautstärke und ob er mit der NTP-Zeit synchronisiert werden soll (siehe Abbildung 13). Einzig die Formatierung der SD-Karte erfolgt am Gerät, um ein versehentliches Löschen zu verhindern. Es wird außerdem bei jedem Status-Update der aktuelle Timecode mitgeschickt. Wenn der Timecode auf dem ESP32-S3 läuft, wird auf ARC-Studio ebenfalls ein Timecode angezeigt. Dieser wird allerdings vom Server berechnet und läuft nicht mit dem Timecode des ESPs synchron. Er dient lediglich zur Veranschaulichung für den Nutzer. Sobald er pausiert wird, wird er auf der Benutzeroberfläche vom Timecode, der mit der Statusnachricht von ARC-Play mitgeschickt wurde, überschrieben und beide Geräte zeigen wieder die gleichen statischen und deswegen ausgegrauten Ziffern an.

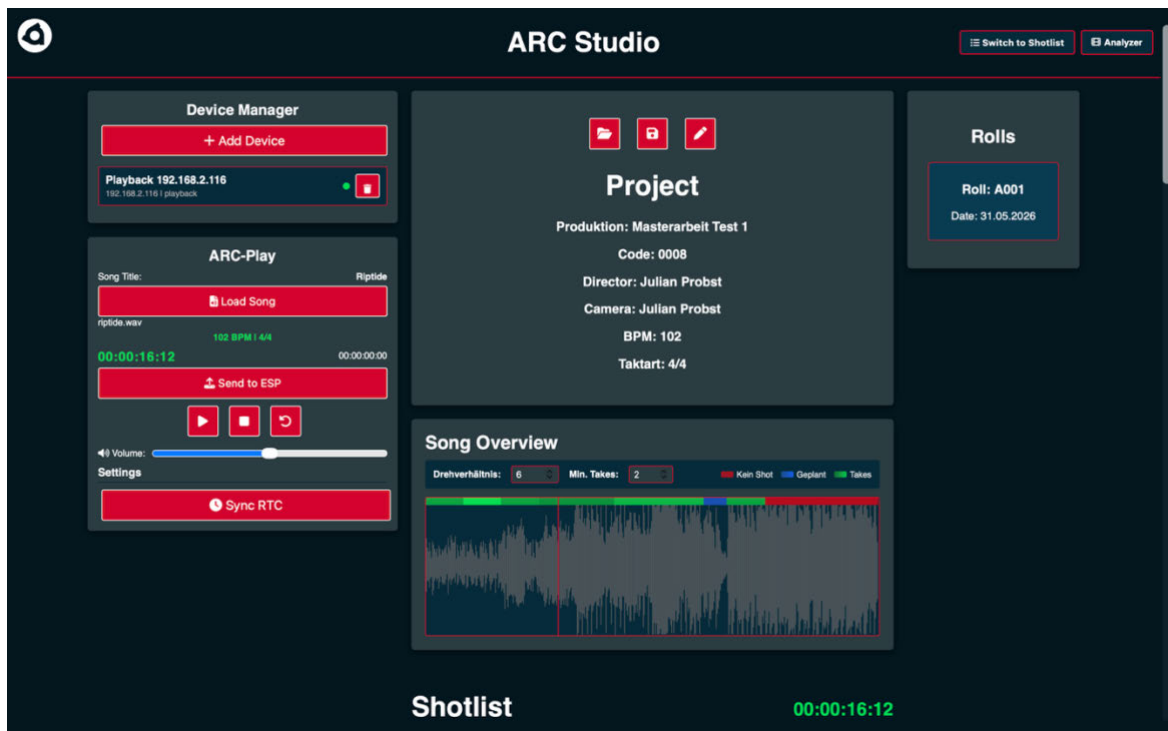


Abbildung 13: Die Benutzeroberfläche von ARC-Studio
Quelle: eigene Darstellung

Um ARC-Studio zu benutzen, muss zuerst ein Projekt erstellt werden. Hierbei werden der Name des Projekts, der Regisseur, der DOP und weitere Informationen eingestellt. Alle Informationen werden daraufhin auf dem state.js abgespeichert. Das Projekt kann aber auch ausgekoppelt und in einer separaten JSON-Datei exportiert werden. Es können auch andere Projekte importiert werden.

ARC-Studio besitzt zwei Modi: zum einen den Shotlist-Modus und zum anderen den Playback-Modus. ARC-Play fungiert im Shotlistmodus als ganz normaler Timecode-

Generator. Im Playback-Modus wird ARC-Play ebenfalls auf diesen Modus umgeschaltet. Nun kann ein Song in ARC-Studio vom Dateisystem des PCs ausgewählt und auf den Server hochgeladen werden. Im Anschluss wird der Song in ein Mono-Audiofile umgewandelt und auf ARC-Play übertragen werden. Dabei wird die Datei in kleine 4-kB-Pakete aufgesplittet und mit 10 ms Verzögerung übertragen. Dadurch bleibt die Bandbreite der Übertragung gering und es wird verhindert, dass es zu Abstürzen kommt. Zusätzlich wird ein Auszug aus state.json mit zusätzlichen Songinformationen wie Dauer, Name, BPM und Dauer des Songs ausgekoppelt und übertragen. Beide Dateien werden auf der SD-Karte des ESPs gespeichert und überschreiben dabei die zuvor vorhandenen Dateien.

Add New Shot

Scene: 01

Shot: 08

Framerate: 25

Description: Outro

Environment: ☐ INT ☒ EXT

Day/Night: ☒ DAY ☐ NIGHT

MOS: ☒

Set In/Out Markers

▶ Abspielen ☒ Snap to Grid

Start (TC): 00:00:37:16

End (TC): 00:00:47:01

Playback Speed: 1.0x

Save Shot Cancel

Abbildung 14: Das Formular zum Erstellen eines neuen Shots in ARC-Studio
Quelle: eigene Darstellung

Um nun neue Shots zu definieren, können über das Shotformular typische Informationen, die auch auf einer Filmklappe angezeigt würden, wie die Framerate, der Ort oder die Tageszeit, definiert werden. Im Playback-Modus kann allerdings zusätzlich definiert werden, von welchem Start- bis Endpunkt in der Audiodatei das Playback für diesen Shot abgespielt werden soll. Um das auszuwählen, kann dazu grafisch ein bestimmter Part im Song definiert werden. Der Start- und Endpunkt kann dabei auf Wunsch genau auf den Takt des Songs einrasten (siehe Abbildung 14). In der Shotlist-Übersicht ist dann der jeweils definierte Bereich auf einer Waveform farbig markiert. Die Shots können nun zu ARC-Play

gesendet und dort abgespielt werden. Sobald ein Durchlauf abgeschlossen und die Nachricht von ARC-Play angekommen ist, wird der Take mit allen Informationen angezeigt. Zudem können weitere, dokumentarische Informationen wie Notizen oder Markierungen, ob es sich um einen Good-Take oder einen B-Roll-Take handelt, hinzugefügt werden (siehe Abbildung 15).

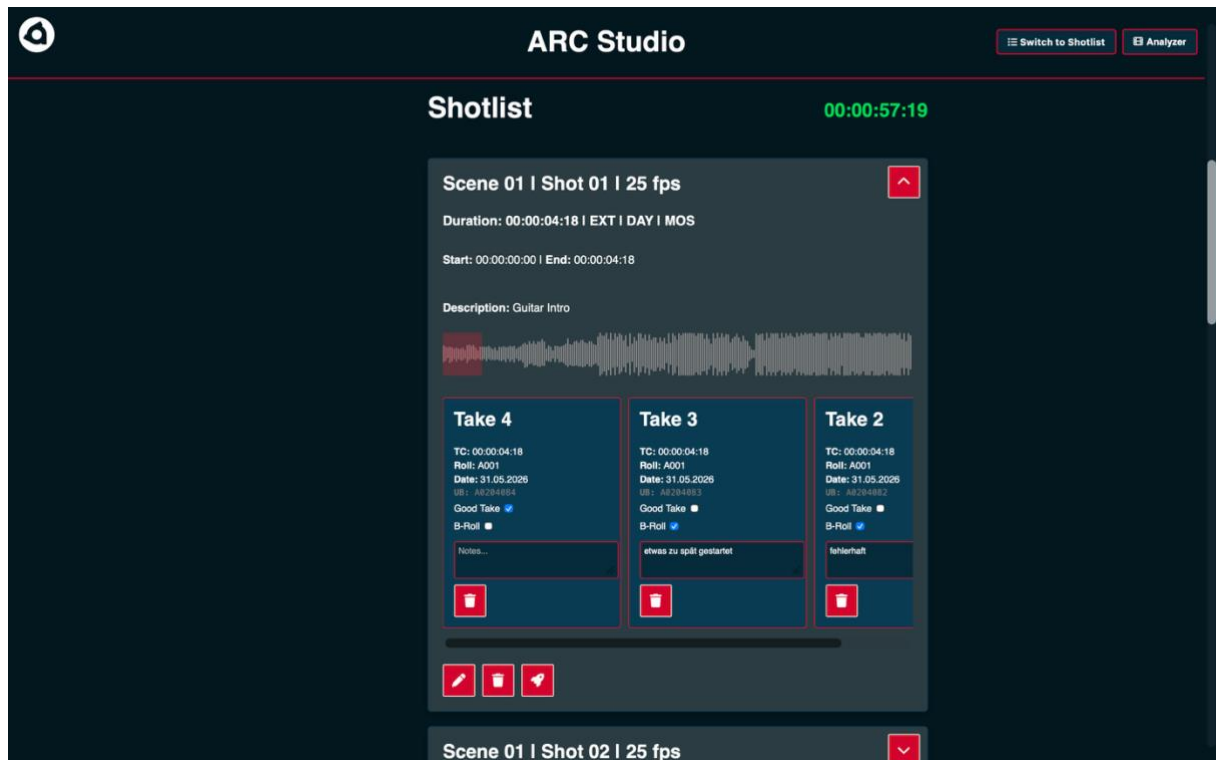


Abbildung 15: Die Shotlist in ARC-Studio
Quelle: eigene Darstellung

Zur Erfüllung des in Kapitel 5.1 definierten Features des besseren Drehverhältnisses wurde eine große Waveform des Songs implementiert, die anzeigt, was bisher geplant und was bisher gedreht wurde. Dazu sind die Bereiche neben der Waveform farblich markiert. Rot bedeutet, dass gar nichts für diesen Part definiert ist. Blau bedeutet, dass für diesen Bereich Shots geplant sind. Je mehr Sättigung das Blau hat, desto mehr Shots sind für diesen Bereich geplant. Grün bedeutet, dass es für diesen Bereich abgedrehte Takes gibt. Auch hier weist die Sättigung der Farbe auf die Anzahl der Takes hin. Dies kann modifiziert werden, indem man das Drehverhältnis anpasst: Je niedriger das Drehverhältnis, desto schneller wird eine hohe Sättigung erreicht (siehe Abbildung 13).

Ähnlich wie die Bibliotheken für Arduino können in Node.js Module via npm installiert werden (Steyer, 2025, S. 7). Dadurch ergibt sich eine große Auswahl an Software, die implementiert werden kann. Die wichtigsten für den Server werden hier in Ausschnitten dargestellt:

Das Modul `multer` (expressjs, 2026) wird verwendet, um den Song, der über das Playback zu hören sein soll, auf den Server hochzuladen. Dabei wird der Dateiname beibehalten und auf dem Server im Ordner „uploads“ gespeichert. Die Datei, die sich zuvor in diesem Ordner befunden hat, wird gelöscht.

Ein weiteres Modul ist `Essentia.js` (Correya et al., 2021), das verwendet wird, um die BPM-Zahl und Taktart des Songs zu analysieren. Es wird dazu benutzt, die taktgenaue Rasterung für die Shot-Erstellung zu gewährleisten.

Außerdem wurde `waveserver.js` eingebunden (katspaugh, 2026). Dabei handelt es sich um eine Bibliothek, die in der Lage ist, die Waveform einer Audiodatei darzustellen. Das dient vor allem der besseren Übersichtlichkeit des Songs.

8.2.2 ARC-Analyzer

Die zweite Seite des Node.js-Servers beinhaltet die Unterseite ARC-Analyzer (siehe Abbildung 16). Diese Software dient dazu, die rohen Videodaten, die während des Drehs entstanden sind, zu analysieren, um so die Metadaten, die mit ARC-Studio erhoben wurden, den einzelnen Clips zuzuordnen.

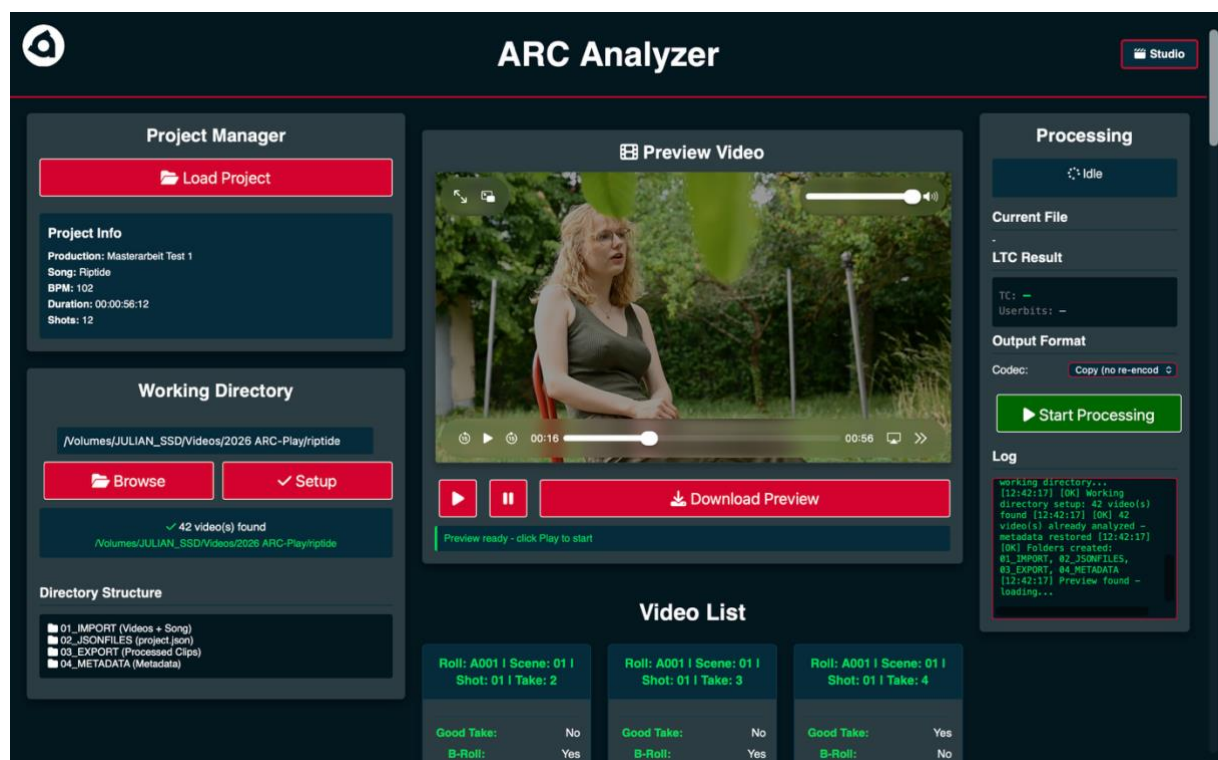


Abbildung 16: Die Benutzeroberfläche von ARC-Studio mit geladenem Preview
Quelle: eigene Darstellung

In ARC-Analyzer wird – genau wie in ARC-Studio – ein Projekt geöffnet oder erstellt. Wenn man von ARC-Studio zu ARC-Analyzer wechselt, bleibt das gleiche Projekt geöffnet.

Zunächst muss ein Working Directory definiert werden, in dem sich die Dateien befinden. Sobald diese gesetzt wurde, wird eine Ordnerstruktur erstellt, um alle Dateien vorzusortieren. Alle RAW-Videoclips werden in einen 01_IMPORT-Ordner gelegt. Aus dem geöffneten Projekt wird ein project.json exportiert, das in 02_JSON abgelegt wird. In 03_EXPORT befinden sich Unterordner, in die die exportierten Videoclips, das Audio des Songs und das Preview exportiert werden. In 04_METADATA wird eine CSV-Datei für DaVinci Resolve exportiert.

Wenn diese Ordnerstruktur erstellt wurde, können nun die Dateien untersucht werden. Dazu wird im Hintergrund die Software ffmpeg (FFmpeg, 2026) verwendet, die den Audiokanal der Videos extrahiert, damit sie von ltcDump, einem Teil der Software ltcLib (Gareus, 2022), die bereits in abgewandelter Form mit der Firmware des ESPs verwendet wird, auf den Timecode untersucht werden kann. Sobald Timecode gefunden wurde, werden die User-Bits des Timecodes analysiert, mit den Informationen aus dem project.json abgeglichen und den entsprechenden Roh-Clips zugeordnet. Nun wird jeder Clip anhand des ebenfalls im project.json abgespeicherten Start- und Endtimecodes des zuvor definierten Shots zugeschnitten und neu gerendert. Selbst wenn ein Clip mit abweichender Playback-Geschwindigkeit gedreht wurde, ist die Software in der Lage, den schnelleren Timecode zu lesen und das Video entsprechend zu verlangsamen. Beim Neuberechnen der Clips kann zwischen verschiedenen Formaten oder der Beibehaltung des ursprünglichen Formats gewählt werden. Der richtige Timecode wird in den Header der neuen Datei geschrieben. Der Dateiname beinhaltet entsprechend Roll, Szene, Shot und Take. Am Ende wird noch eine CSV-Datei geschrieben, die alle Metadaten beinhaltet und in DaVinci Resolve importiert werden kann. In DaVinci Resolve kann man durch diese Metadaten das Material sortieren. Durch den hinterlegten Timecode werden die Clips auch per Knopfdruck an die richtige Stelle verschoben.

Auf Wunsch kann in ARC-Analyzer bereits ein Preview des Musikvideos erstellt werden. Dabei werden die zugeschnittenen Clips automatisch anhand des Timecodes an die richtige Stelle des Songs platziert und als gesamtes Video exportiert. Liegen für eine Stelle im Song mehrere Takes vor, werden die Takes bevorzugt, die als „Good Take“ markiert wurden. Diese Funktion dient zur Vorschau und ist nicht als Rohschnitt des Musikvideos zu verstehen.

9 Anwendung und Validierung

Um ARC-Play zu testen, wurden iterativ mehrere Testmusikvideos gedreht. Dadurch konnte das komplette System hinsichtlich Benutzerfreundlichkeit getestet und Probleme identifiziert werden.

Für die Tests wurden unterschiedliche Songs ausgewählt, die sowohl musikalische Passagen als auch Textpassagen beinhalten. Um auch die Software ARC-Analyzer und die Funktion der automatischen Preview-Generierung auszuprobieren, wurden die Musikvideos sehr detailreich geplant. Das bedeutet, dass einzelne Shots sehr kurzgehalten wurden, um so bereits in der Preview ein möglichst abwechslungsreiches Video zu erlangen. Es soll somit ein Musikvideo nach dem in Kapitel 4.6 definierten Typ C imitiert werden, also ein Film mit einer analogen Kamera, bei dem die Shots sehr spezifisch auf den Song zugeschnitten sind. Es wurde aber aus Kostengründen eine Nikon Z6II verwendet, in deren Audioeingang das Timecodesignal von ARC-Play geführt wurde.

Im Folgenden werden Fehler und Auffälligkeiten in der Planung, Umsetzung und Postproduktion mit dem ARC-Play-System beschrieben und mögliche Verbesserungsvorschläge dargestellt. Teilweise wurden die Verbesserungen bereits implementiert.

In der Vorbereitung für die Drehs wurden Songs gewählt, bei denen auch die BPM, also die Taktschläge pro Minute, bekannt sind. Diese Songs wurden zugeschnitten und in ARC-Studio hochgeladen. Leider wurde dabei teilweise die BPM falsch erkannt, wodurch sich das Planen der Shots als ziemlich schwierig herausgestellt hat. Daher wurde direkt eine Funktion eingebaut, bei der man händisch die Taktart und die BPM des Songs hinzufügen kann. Normalerweise sind diese Informationen bei Künstlern, die gerade einen neuen Song produzieren, vorhanden. Im nächsten Schritt wurden die einzelnen Shots geplant. Daher wurde es schnell unübersichtlich, was auf den einzelnen Shots zu sehen sein soll. Daher wurde ein neues Feld mit dem Namen „Description“ eingeführt, auf dem der Shot beschrieben werden kann. Das Feld ist zudem bei den verarbeiteten Clips in DaVinci Resolve einsehbar. Eine zusätzliche Erweiterung dafür wäre, dass man bestimmte Parts eines Songs fest definieren kann und die automatisch in der Beschreibung des Shots eingefügt werden könnten, beispielsweise ob es sich um eine Strophe oder den Refrain handelt. Zudem muss zukünftig verhindert werden, dass zweimal der gleiche Shot erstellt wird.

Während der Drehs kam es vor allem durch das Einzählen zu Problemen. Wenn der Song anfängt, funktioniert das Einzählen leider gar nicht. Das Einzählen ist zudem etwas zu kurz. Außerdem ist es für Protagonisten schwierig, nur aufgrund des Einzählens den Einsatz parat zu haben. Daher wurde diese Funktion so modifiziert, dass nun acht Schläge zu hören sind und bereits währenddessen der Song im Hintergrund zu hören ist, sodass die Künstler schon vorher anfangen können zu performen.

Bei Änderungen der Framerate wurde durch die erhöhte Geschwindigkeit eine falsche Einstellung in der Kamera gewählt. Das könnte verhindert werden, indem das System darauf hinweist, wenn Einstellungen in der Kamera geändert werden müssen, und der Nutzer bestätigen muss, dass sie entsprechend vorgenommen wurden.

Zudem gab es kleinere Bugs: Beispielsweise hat manchmal das Playback nicht bei der richtigen Stelle angefangen, sondern ganz von vorne. Es hat sich auch als unpraktisch herausgestellt, dass der erste Shot nur ausschließlich über ARC-Studio an ARC-Play gesendet werden kann. Daher könnte durch doppeltes Klicken auf den Rotary-Encoder ein Shot von ARC angefordert werden und dieser, falls vorhanden, direkt auf das Gerät geladen werden.

In der Postproduktion gab es zudem deutliche Probleme mit der Timecode-Erkennung. Diese war leider nicht durchgehend gegeben. Daher musste die Anwendung noch deutlich robuster gemacht werden. Es wird nun beispielsweise in zwei Durchläufen das Timecodeaudio analysiert, einmal ohne Filter und falls nichts gefunden wurde, nochmal mit Tiefpassfilter bei 4 kHz, um so hochfrequentes Rauschen zu eliminieren. Inzwischen wird jeder Take korrekt zugeordnet. Außerdem waren das Bild und der Ton auf Anhieb nicht hundertprozentig synchron. Daher wurde ein Offset von fünf Frames hinzugefügt, wodurch nun alle Bilder perfekt mit dem Audio synchron sind. Das System soll aber zukünftig so optimiert werden, dass bereits am Set der Timecode und das Audio perfekt synchron laufen, ohne später einen Offset verwenden zu müssen.

Während des Tests kamen noch einige Ideen auf, die zusätzlich in das Gerät implementiert werden könnten, um die Funktionen zu erweitern: Ein Feature, das sich in manchen Situationen als nützlich erweisen könnte, wäre, dass man auf dem Display den aktuellen Text anzeigen könnte, ähnlich wie bei der Lip-Sync-Funktion von TikTok (siehe Kapitel 3.3.4). Dadurch könnte es den Künstlern gegebenenfalls leichter fallen, den richtigen Einsatz für den Shot zu finden. Ein weiteres Feature wäre, dass während des Einzählens Regieanweisungen von ARC-Play abgespielt werden können, beispielsweise „Camera

Roll“, um so einen reibungslosen Drehablauf zu gewährleisten. Zudem verfügen einige Kameras über APIs, mit denen Funktionen der Kameras gesteuert werden können, beispielsweise Kameras von Blackmagic Design (Blackmagic Design Pty Ltd, 2025). Dadurch könnte man direkt über das ARC-System die Kamera auslösen, alle Einstellungen für den Shot vornehmen oder auch die zuvor gesammelten Metadaten an die Kamera schicken, die diese dann schon direkt in die Aufnahme­datei schreibt.

Insgesamt lässt sich festhalten, dass sich ARC-Play in der Praxis größtenteils als nützlich erwiesen hat. Gerade ab dem Zeitpunkt, an dem man einen Shot auf das Gerät geladen hat, ist das Erstellen von wiederholten Takes sehr intuitiv. Das automatische Handling der Metadaten erweist sich gerade durch die Notizen am Set als großer Vorteil. Um es aber für einen dynamischen Drehalltag nutzbar zu machen, müssen sowohl die Hardware als auch die Software robuster gestaltet und die genannten Probleme eliminiert werden.

10 Fazit und Ausblick

Das Ziel der vorliegenden Arbeit war, das Ausmaß von Ineffizienzen bei choreographierten Drehs mit Audioplayback darzustellen und darauf aufbauend zu untersuchen, inwieweit ein Timecode-basiertes Audioplayback-System diese verringern kann und wie ein solches System realisiert werden sollte.

Die Ergebnisse der Arbeit bestätigen, dass in der Produktionspraxis erhebliche Probleme vorhanden sind. Dies bestätigt sich aus den sehr individuellen Lösungsversuchen sowohl aus der Marktanalyse als auch aus den Experteninterviews. Allerdings gab es bislang kein dediziertes Gesamtkonzept, das in der Lage ist, das Synchronisieren des Videomaterials mit vorproduziertem Audio und eine effiziente Drehplanung mit Metadatenhandhabung zu kombinieren.

Durch die in Kapitel 6 dargestellten Berechnungen geht hervor, dass insbesondere für professionelle Drehs oder Drehs mit analogem Film ein Produkt wie ARC-Play durchaus einen finanziellen Mehrwert bieten kann, der dessen Anschaffung rechtfertigt. Die Interviewpartner der Experteninterviews aus Kapitel 4 sahen einem solchen System ebenfalls positiv entgegen.

Durch die technische Konzeption, die Realisierung und die anschließenden Tests konnten die Funktionen des Systems erfolgreich dargestellt werden. ARC-Play befindet sich derzeit in einem Proof-of-Concept-Status und ist kein vollständig ausgereiftes Produkt. Allerdings sind die Kernfunktionen bereits vorhanden und erbrachten im praktischen Anwendungstest spürbare Effizienzgewinne für den Workflow.

Für eine valide Einordnung der Erkenntnisse ist eine kritische Reflexion dieser Arbeit notwendig. Die Belastbarkeit der qualitativen Experteninterviews ist durch die geringe Anzahl von drei Interviewpartnern limitiert. Obwohl die stichprobenartige Befragung viel Aufschluss über den Produktionsalltag von Musikvideodrehs gegeben hat und somit auch zur Grundlage für die Realisierung des Prototyps wurde, bieten sie kein vollumfängliches Bild der Branche.

Zudem weisen die Berechnungen aus Kapitel 6 große heuristische Anteile auf, wodurch unter anderen Annahmen die Ergebnisse variieren könnten. Dennoch waren die Berechnungen, gerade in Kombination mit den Experteninterviews, hilfreich, um die Relevanz des Produkts einordnen zu können und dessen Nutzen im Drehalltag zu simulieren.

Durch die Konzeptionsphase und die Anwendungstests sind viele Ideen entstanden, die im Prototypen von ARC-Play noch nicht implementiert sind und die in Zukunft das Produkt noch deutlich aufwerten können (siehe Kapitel 5.3 und 9). Zudem kann die entwickelte Hardware auch für die digitale Filmklappe ARC-Slate verwendet werden, wodurch diese ebenfalls auf ein neues Level angehoben wird.

Ein Aspekt, der den Anwendungsfall von ARC-Play deutlich erweitern dürfte, ist, die User-Bits des SMPTE-Timecodes für eine eindeutige Identifizierung zu nutzen: Die Zuordnung jedes Clips über das Timecodesignal ist zum jetzigen Entwicklungsstand nur auf den Playback-Modus des Geräts beschränkt. Allerdings könnte dies auch im Shotlist-Modus integriert werden. Dadurch würde der Anwendungsfall enorm erweitert werden, da nun nicht mehr nur ein Mehrwert bei Drehs mit Playback vorhanden ist, sondern für alle szenischen Drehs, für die eine Filmklappe verwendet wird. Da alle Informationen über den Clip über das ARC-System vorhanden sind, ließen sie sich mit den User-Bits zuverlässig zuordnen. Bei Drehs mit einer Timecode-fähigen Kamera wäre eine physische Filmklappe somit obsolet. Dadurch wäre ARC-Play kein reines Nischenprodukt für spezielle Anwendungsfälle mehr, sondern in einem breiten Anwendungsfall nutzbar. Zudem würde es das Verwalten der Metadaten und die Synchronisierung von Bild und Ton vollständig automatisieren. Dadurch ließe sich sowohl während des Drehs als auch in der Postproduktion die Effizienz und somit die Relevanz des Produkts erheblich steigern.

Mit diesen vielversprechenden Perspektiven ist eine Weiterentwicklung von ARC-Play fest eingeplant. Im nächsten Schritt sollen die bis zu diesem Zeitpunkt vorhandenen Probleme des Prototyps beseitigt werden. Daraufhin ist geplant, mit den Interviewpartnern, die Interesse an dem System gezeigt haben, in Kontakt zu treten, um es unter realen Bedingungen zu testen und sowohl die Hardware als auch die Software iterativ für den tatsächlichen Marktbedarf auszurichten.

Literaturverzeichnis

Adafruit Industries. (2024). *Adafruit/RTClib* (Version 2.1.4) [Software].

<https://github.com/adafruit/RTClib>

Adafruit Industries. (2026). *Adafruit/Adafruit_ILI9341* (Version 1.6.3) [Software].

https://github.com/adafruit/Adafruit_ILI9341

Andec Filmtechnik GmbH. (o. J.). Preislisten. *Andec Filmtechnik*. Abgerufen 4. Mai 2026,

von <https://andecfilm.de/services/>

AZ-Delivery Vertriebs GmbH. (o. J.). *Drehimpulsgeber Modul Datenblatt*. Abgerufen 5.

Mai 2026, von

https://cdn.shopify.com/s/files/1/1509/1638/files/Drehimpulsgeber_Modul_Datenblatt.pdf?349756184529908641

Balzert, H. (2011). *Lehrbuch der Softwaretechnik: Entwurf, Implementierung, Installation und Betrieb*. Spektrum Akademischer Verlag. <https://doi.org/10.1007/978-3-8274-2246-0>

Blackmagic Design Pty Ltd. (2025, August). *REST API for Blackmagic Cameras*.

<https://documents.blackmagicdesign.com/DeveloperManuals/RESTAPIforBlackmagicCameras.pdf?v=1696143610000>

Bordwell, D., Thompson, K., & Smith, J. (2016). *Film art: An introduction* (11th ed., international student edition). McGraw-Hill Education.

Cipher Digital Inc. (1987). *Time Code Handbook – A Guide for the User from Fundamentals to Technical Specifications*.

https://bitsavers.org/test_equipment/cipher_digital/Cipher_Digital_-_Time_Code_Handbook_1987.pdf

Correya, A., Marcos-Fernández, J., Joglar-Ongay, L., Alonso-Jiménez, P., Serra, X., & Bogdanov, D. (2021). Audio and Music Analysis on the Web using Essentia.js.

- Transactions of the International Society for Music Information Retrieval*, 4(1), 167–181. <https://doi.org/10.5334/tismir.111>
- ESP32 Asynchronous Networking. (2026). *ESP32Async/ESPAsyncWebServer* (Version 3.11.0) [Software]. <https://github.com/ESP32Async/ESPAsyncWebServer>
- Espressif Systems. (o. J.). *FreeRTOS (Supplemental Features)*. Abgerufen 12. Mai 2026, von https://docs.espressif.com/projects/esp-idf/en/v6.0.1/esp32s3/api-reference/system/freertos_additions.html#id15
- Espressif Systems. (2026, März 5). *ESP32-S3 Datasheet—V2.2*. https://documentation.espressif.com/esp32-s3_datasheet_en.pdf
- European Broadcasting Union. (2011, Mai). *Specification of the Broadcast Wave Format (BWF)—V2.0*. <https://tech.ebu.ch/docs/tech/tech3285.pdf>
- expressjs. (2026). *Expressjs/multer* (Version 2.1.1) [Software]. <https://github.com/expressjs/multer>
- FFmpeg. (2026). *FFmpeg/FFmpeg* [Software]. <https://github.com/FFmpeg/FFmpeg>
- Galán, A. (2024, Februar 2). *Wie man auf TikTok die Lippen synchronisiert*. Malavida. <https://de.malavida.com/faq/wie-man-auf-tiktok-die-lippen-synchronisiert>
- Gareus, R. (2022). *X42/libltc* (Version 1.3.2) [Software]. <https://github.com/x42/libltc>
- Geyer, C. (2024). *Populäre Musik im Kontext von Social Media-Kurzvideos: Musikaneignung im Rahmen von TikTok und Instagram Reels*, vol. 15. Nomos Verlagsgesellschaft mbH and Co. KG. <https://doi.org/10.5771/9783748940357>
- Ginsburg, F. (2024, Dezember 3). *Sync Playback: Revisited*. Filmtvsound. <https://www.filmtvsound.com/index.php/audio-techniques/316-sync-playback-revisited>
- Graetz, D. (2022, Januar 27). *Perfect Music Video Sync – any framerate...every time* [Videoaufnahme]. <https://www.youtube.com/watch?v=oV4AJnQ9QiA>

- Haute Technique. (o. J.). *Timecode and Audio*. Abgerufen 10. März 2026, von <https://timecodesync.com/player/>
- ILI Technology Corp. (o. J.). *ILI9341 Specification*. Abgerufen 3. Juni 2026, von <https://datasheet4u.com/pdf/769801/ILI9341.pdf>
- Janney, R. (2018, Oktober 12). *Learn Music Video Playback From An Audio Professional*. Shutterstock. <https://www.shutterstock.com/blog/learn-filmmaking-music-video-playback>
- katspaugh. (2026). *Katspaugh/wavesurfer.js* (Version 7.12.7) [Software]. <https://github.com/katspaugh/wavesurfer.js>
- Mancini, R. (2003). *Op amps for everyone: Design reference* (2nd ed.). Newnes.
- Maxim Integrated Products, Inc. (2015). *DS3231*. <https://www.analog.com/media/en/technical-documentation/data-sheets/DS3231.pdf>
- Mayring, P. (2015). *Qualitative Inhaltsanalyse: Grundlagen und Techniken* (12., vollständig überarbeitete und aktualisierte Aufl). Beltz.
- Mills, D., Burbank, J., & Kasch, W. (2010). *Network Time Protocol Version 4: Protocol and Algorithms Specification* (J. Martin, Hrsg.; Nr. RFC5905; S. RFC5905). RFC Editor. <https://doi.org/10.17487/rfc5905>
- Misoch, S. (2019). *Qualitative Interviews* (2., erweiterte und aktualisierte Auflage). De Gruyter Oldenbourg. <https://doi.org/10.1515/9783110545982>
- Plate. (2017, November 12). *Schnittstellen—Serielle Schnittstelle, USB, Feldbusse, SPI, I2C, 1-Wire etc.* [Skript]. <https://www.emmzett.at/lib/exe/fetch.php?media=info%3Aschnittstellen.pdf>
- PureBlend Corporation. (o. J.). *MovieSlate—User Guide*. Abgerufen 3. Juni 2026, von https://www.movie-slate.com/User_Guide

- Rapp, H. (2020, September 20). *The Difference Between Timecode and Genlock and Why You May Need Both On Set*. Nofilmschool. <https://nofilmschool.com/timecode-and-genlock-and-why-you-may-need-both-set>
- Reinhold, W. (2010). *Elektronische Schaltungstechnik: Grundlagen der Analogelektronik*. Carl Hanser Verlag GmbH & Co. KG. <https://doi.org/10.3139/9783446424821>
- Richter, T. (2025, Juni 25). *Timecode für Musikvideos: Kostenlose Tentacle Music Video Sync App vereinfacht die Musikvideoproduktion*. slashCAM. <https://www.slashcam.de/news/single/Kostenlose-Tentacle-Music-Video-Sync-App-vereinfac-19368.html>
- Sänpäkkilä, S. (2019, Juli 22). *Visual Music Video Timecode Slate System—How to achieve perfect lip sync*. <https://samisanpakkila.com/2019/07/visual-music-video-timecode-slate/>
- Schatzmann, P. (2026). *Pschatzmann/arduino-audio-tools* (Version 1.2.3) [Software]. <https://github.com/pschatzmann/arduino-audio-tools>
- Schmidt, U. (1999). *Filmlängenrechner*. http://www.cinematography.de/filmruntime_calc.php
- Schmidt, U. (2011). *Digitale Film- und Videotechnik* (3. Aufl.). Carl Hanser Verlag.
- Schmidt, U. (2021). *Professionelle Videotechnik: Grundlagen, Filmtechnik, Fernsehtechnik, Geräte- und Studioteknik in SD, HD, UHD, HDR, IP* (7. Aufl.). Springer Berlin Heidelberg. <https://doi.org/10.1007/978-3-662-63944-3>
- SD Card Association. (2018, Juli 5). *SD Specifications Part E1 SDIO Simplified Specification*. https://www.sdcard.org/downloads/pls/pdf/?p=PartE1_SDIO_Simplified_Specification_Ver3.00.jpg&f=PartE1_SDIO_Simplified_Specification_Ver3.00.pdf&e=EN_SSE1

Sharman, R. (2020). *Moving Pictures—An Introduction to Cinema*.

<https://uark.pressbooks.pub/movingpictures/chapter/editing/>

SILBERSALZ Film GmbH. (o. J.). *Kodak VISION3 Film – SILBERSALZ35—True*

Cinefilm. Abgerufen 27. April 2026, von <https://silbersalz35.com/product/kodak-vision3/>

SMPTE. (2014). *SMPTE ST 12-1:2014—Time and Control Code*.

<https://doi.org/10.5594/SMPTE.ST12-1.2014>

Steyer, R. (2025). *Programmieren mit JavaScript: Die vielseitige Sprache für*

Webentwicklung & mehr – Grundlagen und fortgeschrittene Techniken. Springer

Fachmedien Wiesbaden. <https://doi.org/10.1007/978-3-658-45577-4>

Tentacle Sync GmbH. (o. J.). *Tentacle Music Video Sync App (Android Beta)—First Steps*.

Abgerufen 8. Juni 2026, von [https://support.tentaclesync.com/hc/en-](https://support.tentaclesync.com/hc/en-us/articles/28360401789853-Tentacle-Music-Video-Sync-App-Android-Beta-First-Steps)

[us/articles/28360401789853-Tentacle-Music-Video-Sync-App-Android-Beta-First-Steps](https://support.tentaclesync.com/hc/en-us/articles/28360401789853-Tentacle-Music-Video-Sync-App-Android-Beta-First-Steps)

Tentacle Sync GmbH. (2025, Februar 3). *SYNC E Bedienungsanleitung*.

<https://manuals.tentaclesync.com/de/sync-e-manual.html>

Texas Instruments Inc. (2012, September). *PCM5100, PCM5101, PCM5102*.

[https://www.ti.com/lit/ds/symlink/pcm5102.pdf?ts=1776015659465&ref_url=https](https://www.ti.com/lit/ds/symlink/pcm5102.pdf?ts=1776015659465&ref_url=https%253A%252F%252Fwww.ti.com%252Fproduct%252FPCM5102)

[%253A%252F%252Fwww.ti.com%252Fproduct%252FPCM5102](https://www.ti.com/lit/ds/symlink/pcm5102.pdf?ts=1776015659465&ref_url=https%253A%252F%252Fwww.ti.com%252Fproduct%252FPCM5102)

Texas Instruments Inc. (2024, Oktober). *Industry-Standard Dual Operational Amplifiers*.

[https://www.ti.com/lit/ds/symlink/lm358.pdf?ts=1777915049369&ref_url=https%](https://www.ti.com/lit/ds/symlink/lm358.pdf?ts=1777915049369&ref_url=https%253A%252F%252Fwww.ti.com%252Fproduct%252FLM358)

[53A%252F%252Fwww.ti.com%252Fproduct%252FLM358](https://www.ti.com/lit/ds/symlink/lm358.pdf?ts=1777915049369&ref_url=https%253A%252F%252Fwww.ti.com%252Fproduct%252FLM358)

Texas Instruments Inc. (2025, Februar). *SNx4HC14 Hex Inverters with Schmitt-Trigger*

Inputs.

[https://www.ti.com/lit/ds/symlink/sn74hc14.pdf?ts=1777945063289&ref_url=https](https://www.ti.com/lit/ds/symlink/sn74hc14.pdf?ts=1777945063289&ref_url=https%253A%252F%252Fwww.google.com%252F)

[%253A%252F%252Fwww.google.com%252F](https://www.ti.com/lit/ds/symlink/sn74hc14.pdf?ts=1777945063289&ref_url=https%253A%252F%252Fwww.google.com%252F)

ver.di. (2024, Oktober 12). *Tarifvertrag für auf Produktionsdauer beschäftigte Film- und Fernsehschaffende* [Vertrag]. <https://cms-s3.verdi.de/drupal/public/documents/TV-FFS-2025-Text.pdf?VersionId=0056a225-edf5-4eba-9b37-11f6b08e0f06>

Anhang

A Anhang: Interviewtranskript mit Experte A

Datum: 31.03.2026, 14:00

Interviewer (I): Ich habe ja im Vorfeld schon ein bisschen erklärt, worum es geht, aber ich würde es noch einmal kurz zusammenfassen: Das Thema meiner Masterarbeit ist die Entwicklung eines Playback-Systems, das in der Lage ist, gleichzeitig einen Song und einen synchronen Timecode auszugeben. Ich habe hier auch einen Prototypen auf einem Steckbrett – das ist aktuell noch ein reiner Timecode-Generator, an dem ich gerade arbeite.

Die Idee dazu entstand bei einem Musikvideo-Dreh in Hamburg. Wir haben auf Analogfilm gedreht und hatten eine feste Choreographie für den gesamten Song. Da Analogmaterial sehr teuer ist, konnten wir nicht einfach den kompletten Song aus verschiedenen Perspektiven durchlaufen lassen. Wir mussten die Szenen „anstückeln“. Das war kompliziert, weil wir nur eine Bluetooth-Box und ein Handy hatten. Das hat überhaupt nicht gut funktioniert. In der Postproduktion war es dann extrem schwierig, die Aufnahmen nur anhand der visuellen Bewegungen zu synchronisieren.

Mein Gerät soll die Wave-Datei fix mit dem Timecode verknüpfen. Über eine Software kann man vorab Szenen und Shots definieren und festlegen, von welchem bis zu welchem Timecode gedreht werden soll. Diese Daten lädt man auf das Gerät, das dann Audio und Timecode parallel abspielt. Den Timecode kann man dann entweder direkt in eine Timecode-fähige Kamera einspeisen oder – für Kameras ohne Eingang wie unsere Analogkamera – über meine digitale Filmklappe synchronisieren. Die Klappe befindet sich im selben Netzwerk und bietet, neben dem QR-Code für Metadaten, eine visuelle Referenz für den Timecode.

Experte (E): Voll gute Idee! Wenn das funktioniert, hilft das sehr. In der Postproduktion ist das manuelle Synchronisieren zwar machbar, aber anstrengend.

I: Genau. Das Ziel meiner Masterarbeit ist es, das Gerät zu bauen und eine Modellberechnung zu erstellen, die simuliert, wie sehr ein solches System die Effizienz am Set steigert. Dafür brauche ich Informationen über den Status quo in der Praxis. Mein Eindruck ist bisher, dass jeder seine eigene Taktik hat und es kaum einheitliche Workflows gibt. Wie läuft das bei euch in der Produktion ab? Ist der Drehplan so strikt wie beim szenischen Film oder entscheidet man vieles erst im Schnitt?

E: Bei Analogfilm wird sehr genau geplant, allein wegen der Kosten. Bei Digitaldreh ist es eher egal, da hat man zur Not ja auch noch den Kamerasound als Referenz. Aber bei Analog wissen wir am Set genau: „Dieser Take ist für diesen speziellen Part im Song.“ Besonders bei Choreographien, wo die Moves exakt auf den Beat passen müssen, ist eine genaue Planung unerlässlich.

I: Wenn ihr solche Choreographien dreht, was nutzt ihr dann vor Ort, um den Song abzuspielen?

E: Meine Erfahrung bei fast allen Musikvideos in Wien: eine Bluetooth-Box. Man sucht die Stelle auf dem Handy raus und drückt Play. Das ist eigentlich der Standard-Workflow, weil bei Musikvideos oft das Geld fehlt.

I: Das heißt, ihr habt bei diesen Drehs gar keinen Timecode am Set?

E: Überhaupt nicht.

I: Hast du Erfahrung mit dem Aufzeichnen von Bands, bei denen Instrumente synchron sein müssen?

E: Nicht auf Analogfilm, nur digital. Da lassen wir die Musik laufen und die Band spielt dazu. Das Synchronisieren in Premiere über den Autosync funktioniert bei Digitalaufnahmen eigentlich so gut, dass man sich da nicht zu viel Arbeit machen muss.

I: Ich habe in manchen Videos gesehen, dass Zeitlupen verwendet werden, der Sänger aber trotzdem lippensynchron singt. Hast du sowas schon gemacht?

E: Ja, das habe ich gemacht. Der Song wird dabei oft schneller abgespielt. Man muss dann genau wissen, mit welchem Faktor man das Footage später in der Postproduktion verlangsamt – zum Beispiel um den Faktor 1,33 oder doppelt so schnell. Wenn du so eine Funktion in dein System einbauen könntest, wäre das super, weil es mehr kreative Möglichkeiten bietet.

I: Mein Ziel ist es, durch die Verknüpfung von Metadaten und Timecode den Rohschnitt fast zu automatisieren. Meine Filmklappe hat ein E-Ink-Display mit einem QR-Code. Die Software erkennt diesen im Clip und ordnet automatisch alle Notizen und die Bewertung des Takes zu. Man könnte dann eine Timeline exportieren, die direkt nur die „Best-Takes“ enthält. Wie dokumentierst du Shots am Set?

E: Die Regisseure, mit denen ich arbeite, machen das meistens für sich selbst auf dem iPad. Mehr als „Der war gut“ oder „Der war schlecht“ wird da selten notiert.

I: Wie viele Takes macht ihr pro Shot, bis er als erledigt gilt?

E: Das ist individuell. Bei Analogfilm sind wir natürlich eingeschränkt; da sind es meistens maximal drei Takes. Wenn der erste schon gut war, macht man zur Sicherheit noch einen zweiten „Safety“, aber viel mehr nicht.

I: Bei unserem Dreh in Hamburg hatten wir für unseren 3,5-Minuten-Song auch nur eine Rolle 35-mm-Film. Da haben wir dann auch nur zwei Takes für jede Stelle im Song verwendet.

E: Bei Bibiza hatten wir zwei Rollen 16-mm-Film. Bibiza wollte unbedingt diesen analogen Look für alle seine Videos. Der Look ist auch wirklich cool, das muss man sagen.

I: Zu Postproduktion: wie strukturierst du dir das Material im Beginn der Postproduktion?

E: Ich lege mir eine Timeline an und füge alle Clips hinein und sortiere das mir dort. Da gibt es aber sehr unterschiedliche Ansätze.

I: Würde es dir im Schnitt helfen, wenn die Metadaten und Notizen vom Set direkt in der Datei hinterlegt wären?

E: Es wäre definitiv ein Mehrwert. Je mehr Infos man hat, desto besser. Ich würde aber wahrscheinlich trotzdem noch mal alles durchsehen, um sicherzugehen, dass man nichts Wertvolles übersieht.

I: Ein wichtiges Thema für meine Effizienzanalyse ist die Preisgestaltung. Wie setzen sich die Budgets bei Musikvideos zusammen? Kann man durch Zeitersparnis Geld sparen?

E: In Österreich sind die Budgets oft winzig. Newcomer kommen mit 1.000 Euro an. Selbst bei bekannteren Sachen wie Bibiza lagen die Budgets früher bei etwa 5.000 Euro pro Video. Davon gingen allein 2.500 Euro für das Filmmaterial, die Entwicklung und das Scannen drauf. Da verdient eigentlich niemand mehr etwas; das sind oft Spaßprojekte, bei denen die Crew für 200 Euro am Tag oder komplett umsonst arbeitet. Bei Bilderbuch ist es dann vielleicht ein bisschen mehr. Ein Freund von mir hat bei Bausa in Berlin Regie geführt, selbst da waren es nur 20.000 Euro.

I: Das ist interessant. Selbst bei bekannteren Künstlern ist das Budget also oft ein Problem. Gibt es im Workflow Punkte, an denen oft Fehler passieren oder Zeit verloren geht?

E: Eigentlich nur das Abspielen der Musik. Die Box verliert die Verbindung, man erwischt die falsche Stelle im Song oder drückt zu spät Play. Das sind Kleinigkeiten, die sich über den Tag aber summieren. Wenn dein Gerät ein „All-in-One“-Ding wäre, das auch laut Musik abspielen kann, ohne Bluetooth-Probleme, wäre das ein riesiger Vorteil.

I: Genau das ist der Plan. Es soll ein fixes Gerät sein, das stabil läuft. Es gibt zwar eine Beta-App von Tentacle Sync, aber die hat eben wieder die Bluetooth-Latenzprobleme. Mein System soll das physisch lösen. Das wärs von meiner Seite! Vielen Dank für deine Zeit!

E: Sehr gerne, sehr cooles Projekt.

I: Was ich mir nur dachte: Hättest du Interesse das mal auszuprobieren?

E: Ja sicher, melde dich gerne! Ich bin viel im Norden unterwegs!

I: Ich bin inzwischen in Regensburg und bin dort auch mit einem Start-up für Filmequipment aktiv. Wir versuchen gerade, eine eigene Filmentwicklungsmaschine zu bauen, um die Kosten für Analogdrehs zu senken. Ich schicke dir mal die Website.

E: Wenn ihr es schafft, dass das nicht mehr Tausende Euro kostet, wäre das genial.

B Anhang: Interviewtranskript mit Experte B

Datum: 03.04.2026, 13:00

Interviewer (I): Ich studiere zeitabhängige Medien – Vision und Sound. Das Studium ist sehr projektbasiert; wir setzen eigenständig verschiedene Projekte um. Ich habe auch davor, schon im Bachelor, immer wieder Projekte gemacht, bei denen ich selbst etwas baue, was eher in Richtung Filmtechnik geht. Als Bachelorarbeit habe ich zum Beispiel einen Filmscanner für Analogfilm gebaut – also für 35-mm-Film.

Diese ganzen Technikrichtungen begeistern mich sehr. Ich bin auch Teil eines kleinen Start-ups; wir haben bisher zwar noch nicht viel verdient, entwickeln aber eigene Produkte. Dazu gehört auch das Playback-System, um das es heute gehen soll. Ich habe dir auch ein Bild im Chat geschickt, da sieht man das ein bisschen. Ich dachte, man sieht es direkt, aber man muss es wohl erst herunterladen. Es geht dabei unter anderem um eine Filmklappe, die ich dir, glaube ich, auch mal per E-Mail geschickt hatte.

Experte (E): Im Chat sehe ich gerade leider kein Bild. Moment, vielleicht muss ich es noch mal neu laden. Ah, okay, ja – das ist diese Filmklappe.

I: Genau. Die Idee dahinter ist, dass sie einen QR-Code hat. Wenn man vorab Szenen und Shots definiert, hängt das Ganze in einem Netzwerk. Wenn man danach das Filmmaterial mit der Software analysiert, wird der QR-Code erkannt. Alle Metadaten und Notizen werden dann entweder direkt in die Metadaten der Filmdatei geschrieben oder als extra Metadaten-Dateien erstellt. Das war ein Projekt, das ich an der Uni umgesetzt habe. Mein aktuelles Projekt soll nun mit diesem System zusammenspielen können. Das nur am Rande – ich wollte es erwähnt haben, weil es in eine ähnliche Richtung geht.

Vielleicht fange ich einfach mal an: Ich habe dir das Konzept ja schon per E-Mail kurz beschrieben. Wir hatten einen Musikvideo-Dreh in Hamburg im Rahmen der Uni, ebenfalls auf Analogfilm. Das war mein erster Versuch, das professionell umzusetzen. Da wir auf Film gedreht haben und verschiedene Perspektiven der gleichen Choreographie wollten, konnten wir nicht einfach alles von vorne bis hinten durchdrehen, sondern wollten stückeln, um Filmmaterial zu sparen, da das extrem teuer ist. Dabei war es jedoch sehr kompliziert, die bestimmten Bereiche mit der Bluetooth-Box und dem Handy perfekt zu timen. Da wir mit einer Arriflex 435 gedreht haben, konnten wir zudem kein Audio aufnehmen. In der Postproduktion war das sehr mühsam, da wir nur anhand der Bewegungen synchronisieren konnten – ohne Audio und ohne Timecode.

Deswegen dachte ich mir, es wäre cool, wenn man vorab definieren könnte: „An dieser Szene wollen wir diesen Shot haben“. Das Lied wird dann automatisch an der richtigen Stelle abgespielt, während gleichzeitig der Timecode ausgegeben wird. Das Ganze soll am Ende mit der Klappe synchronisiert sein, die das über das Netzwerk anzeigt. Die Klappe ist so designt, dass man sie für Analogfilm wiederverwenden kann, da man dort ja sonst gar keine Metadaten hat. Sie ist auch in der Lage, einzelne Takes aus der fertigen Filmrolle „rauszustanzen“. Du hast ja schon geschrieben, dass es von Tentacle Sync etwas Ähnliches gibt.

E: Ich habe das System von Tentacle selbst noch nie benutzt, aber ich erinnere mich, es mal gesehen zu haben.

I: Genau. Das ist tatsächlich sehr ähnlich. Was bei mir noch dazukommt, ist die Synchronisation mit der Klappe. Ob das mit der Tentacle-Klappe funktioniert, weiß ich nicht, das ist ja noch ein Beta-Programm. Ein Problem bei Tentacle ist – so habe ich es in Foren gelesen –, dass der Sound über eine App abgespielt wird. Da viele Geräte heute keinen Audio-Ausgang mehr haben, braucht man oft einen Adapter. Wenn das dann wieder über Bluetooth läuft, bekommt man Synchronisationsprobleme. Mein Ansatz ist ein fertiges Gerät mit zwei physischen Audio-Ausgängen: einer für den Timecode und einer für das Lied. Die Software funktioniert schon einigermaßen, aktuell ist die Technik aber noch auf dem Steckbrett aufgebaut.

Aber jetzt die Frage an dich: Wie macht ihr das bei eurer Firma? Wenn ihr Musikvideos dreht, wie spielt ihr die Musik am Set ab?

E: Ganz am Anfang haben wir es klassisch gemacht: Wir hatten eine Bluetooth-Box dabei und haben uns auf das interne Mikrofon der Kamera verlassen. Da gab es aber oft Probleme: Die Box war zu weit weg, es gab Windgeräusche oder man filmt von Auto zu Auto usw. Dann kommt der Sound bei der Kamera nicht ordentlich an. Im Schnitt sitzt man dann da und versucht, einzelne Wörter oder einen markanten Beat herauszuhören.

Ich habe dann mal Bluetooth-Dongles und Splitter geholt, um vom Abspielgerät gleichzeitig die Box und einen Empfänger am Kamera-Mikrofon-Eingang zu bedienen. Das ist dann zwar kein LTC oder so, aber ich dachte, dann hätte man zumindest immer den sauberen Song auf der Tonspur. Aber da kam das Problem mit der Latenz. Es hat nie funktioniert, weil die Latenz je nach Entfernung nicht nur da, sondern auch noch schwankend war. Man konnte nicht einfach sagen: „Es sind immer drei Frames Versatz“, sondern es war jedes Mal anders.

I: Und wie handhabst du das aktuell?

E: Aktuell mache ich selbst weniger Drehs und mehr Geschäftsführung. Aber was wir oft machen: Wir packen den Song kurz in Premiere oder Resolve und generieren einen Timecode als Video. Dann hält man am Set einfach ein iPad oder Handy mit diesem Video in die Kamera. So sieht man im Bild: „Minute 02:25, 15 Frames“. So kann man es manuell synchronisieren. Aber man muss eben immer etwas in die Kamera halten. Bei schnellen „Run-and-Gun“-Musikvideos nimmt eine Klappe oft unnötig Zeit weg. Auch bei Drohnenaufnahmen machen wir das so, weil man den Sound der Drohne ja nicht nutzen kann.

I: Spannend. Bei meinem Konzept geht es ja auch darum, Shots vorab zu definieren. Wie oft kommt das bei euch vor? Weiß man bei Musikvideos vorher schon genau: „Im Refrain will ich diese Szene oder diesen Shot“?

E: Je erfahrener die Leute sind, desto genauer teilt man das auf. Am Anfang dreht man oft noch den ganzen Song an jeder Location durch. Später weiß man genau: „Hier brauche ich nur diesen Satz“. Man schneidet den Film quasi schon im Kopf vor. In der Shotlist steht dann zum Beispiel für Shot 5B: „Spiel Datei 5B ab“ – und da ist dann wirklich nur dieser Part des Songs drin.

Manchmal drehen wir auch auf „Speed“, also lassen den Song 1,5-mal schneller laufen, um später in der Postproduktion eine leichte Zeitlupe zu erhalten. Dafür bereiten wir dann die entsprechenden Versionen des Songs inklusive angepasstem Timecode vor. Manchmal loopen wir auch einzelne Wörter in einer Datei. Dann sagt der Artist den Satz 100-mal hintereinander, wir filmen durch und können so schneller mehrere Takes abklappen.

I: Das ist sehr interessant, weil mein System genau diese Vorbereitung übernehmen würde. Man könnte sofort sagen: „Jetzt 1,5-fache Geschwindigkeit“, und der Timecode am Ende passt trotzdem. Ich habe das bei einem Video von Kraftklub gesehen – dadurch bin ich auch auf euch gekommen. Manche Artists schaffen das schnellere Performen aber wahrscheinlich gar nicht, oder?

E: Genau, deshalb haben wir immer die normale Version als Backup dabei. Bei 1,25- oder 1,5-facher Geschwindigkeit ist das oft schon sehr schwierig für die Künstler.

I: Dreht ihr auch auf Analogfilm?

E: Ja, sehr viel. Zum Beispiel sind die ganzen Apsilon-Geschichten eigentlich alle analog. Ich war da selbst zwar noch nicht am Set dabei, aber ich schätze, die machen das ähnlich mit dem iPad im Bild.

I: Es gibt zwar Analogkameras, die Ton aufzeichnen können, aber das ist meistens sehr aufwendig.

Habt ihr Erfahrung mit Livebands? Spielt man da den Song laut ab oder nutzen die In-Ear-Monitoring mit Click? Und nutzt ihr „Thump Tracks“ für Szenen, in denen im Vordergrund Dialog gesprochen wird?

E: Was sind „Thump Tracks“?

I: Das habe ich bei der Recherche gefunden. Das nutzt man zum Beispiel in Club-Szenen. Damit die Leute im Hintergrund im richtigen Rhythmus tanzen, spielt man nur die Kick-Drum bei etwa 40 bis 60 Hertz ab. Diese Frequenz kann man später aus den Dialog-Mikrofonen einfach herausfiltern.

E: Ach so. Nein, wir nutzen eigentlich nie In-Ears. Bei Bands wird der Song einfach laut abgespielt, das klappt meistens gut. Manchmal bereiten wir es vor, dass am Anfang eingezählt wird. Das ist bei unkonventionellen Einstiegen wichtig, damit die Artists den Rhythmus finden. Solche Club-Szenen, wie du sie beschreibst, hatten wir schon oft. Irgendwann kann man den Song am Set einfach nicht mehr hören. Manchmal passt der Song auch visuell nicht zum Vibe, den man erzeugen will; dann nehmen wir einen anderen Song mit der gleichen BPM-Zahl.

I: Wenn ihr den Song für die Musiker stückelt, brauchen die dann nicht auch einen gewissen Vorlauf, um reinzukommen?

E: Absolut. Ich habe gelernt, dass drei Sekunden Vorlauf nicht reichen. Man muss ihnen mehr Zeit geben, damit sie in den Flow kommen. Manchmal ist es trotzdem etwas durcheinander, wenn sich die Parts in den Dateien zu sehr überlappen, aber das ist am Ende nicht so wichtig – außer es gibt ganz spezifische Regieanweisungen auf bestimmte Takte.

I: Arbeitet ihr mit Mehrkamerasystemen oder nur mit einer Kamera?

E: Meistens ist es nur eine Kamera. Höchstens filmt jemand noch zusätzlich mit einem Camcorder für einen „B-Cam-Look“. Da es meistens um eine Performance direkt in die Kamera geht, ergibt sich ein Mehrkamerasystem selten.

I: Und wie dokumentiert ihr die Takes am Set? Habt ihr da einen speziellen Workflow?

E: Ich persönlich versuche, am Drehtag so wenig Zeit wie möglich mit Dokumentation zu verbringen. Die Zeit am Set ist viel kostbarer als in der Postproduktion. Ich plädiere oft dafür, keine Klappe zu benutzen und es möglichst einfach zu halten, um den Flow nicht zu unterbrechen. Da sitzt man lieber im Schnitt eine Minute länger und ordnet die Clips zu.

Das hängt aber stark vom Projekt ab. Wenn man eine chronologische Story unchronologisch dreht, ist Dokumentation wichtiger als bei einem „Run-and-Gun“-Musikvideo, das man intuitiv zusammenschneidet. Da die Videos kurz sind, hat man das Material meistens gut im Kopf, zumal die Regie oft direkt nach dem Dreh mit im Schnitt sitzt. Bei Kurzfilmen ist das natürlich eine ganz andere Geschichte.

I: Wie strukturiert ihr das Material? Schaut ihr alles durch oder gibt es ein System?

E: Wenn ich den ganzen Song an jeder Location in voller Länge habe, erstelle ich eine Multicam-Sequenz, synchronisiere alles und schneide nach Gefühl. Danach sichte ich das Material noch mal nach „B-Roll-Shots“ – oft passiert in einem Performance-Take zufällig etwas Cooles. Wenn der Plan aber schon vorher feststeht, habe ich für jeden Part eine kleine Multicam-Sequenz oder lege die zwei, drei Takes direkt in der Timeline übereinander.

I: Diese Unterscheidung ist spannend für mich, um zu sehen, wo mein Projekt am meisten Sinn ergibt. Bei einem Dreh, der von vorne bis hinten durchläuft, ist der Mehrwert sicher geringer als bei gestückelten Szenen.

E: Ich muss aber ehrlich sagen: Oft ist der Ton am Set so schlecht, dass ein automatischer Sync per Timecode für viele ein riesiger Mehrwert wäre. Das Anlegen und Synchronisieren will niemand händisch machen. Besonders wenn man mit Drohnen oder alten Alexa-Kameras ohne internes Mikro dreht, landet man oft in der Situation, dass man nach Lippenbewegungen synchronisieren muss. Das ist echt eine Qual. Ein System, das die Multicam-Anlegung automatisiert, wäre also definitiv hilfreich.

I: Genau das ist der Gedanke. Wenn die Metadaten durch die Klappe schon in den Dateien sind, könnte man den Rohschnitt oder die Multicam-Sequenzen automatisieren. Fallen dir sonst noch Ineffizienzen oder Fehlerquellen am Set ein?

E: Das Hauptthema in der Musikvideowelt ist wirklich: „Wie kriegt man den Timecode sauber ins System?“ Alles andere ist danach meistens nur ein Klick.

Wir bauen uns intern auch viele Tools. Mein Mitgeschäftsführer ist Informatiker und entwickelt Tools für Kalkulationen, Abrechnungen oder Verträge, damit das alles schneller geht. Was am Set auch sehr cool ist, ist die App „Shot Lister“. Das ist eine Live-Shotlist, die dir am Drehtag anzeigt, wie sehr du im Verzug bist. Man kann das auf iPads teilen, Shots verschieben und sieht sofort, wie sich das auf den Zeitplan auswirkt. Das spart das ständige manuelle Umrechnen von Uhrzeiten.

I: Genau daran arbeiten wir in meinem Start-up auch. Mein Kollege ist ebenfalls Informatiker. Die Vision ist ein Server-System, auf dem Projekte gehostet werden. Das System beinhaltet die Shotlist und die Filmklappe, sammelt alle Daten und schreibt sie automatisch in die Dateien. So hätte jeder am Set – egal ob Licht oder Regie – Zugriff auf die für ihn relevanten Informationen in Echtzeit.

E: Ja check auf jeden Fall mal Shotlister aus – das funktioniert auch mit verschiedenen Departments. Es gibt auch noch was von Arri. Außerdem gibt es Studio Binder – das fand ich aber extrem schlecht. Ich habe aber auch schon oft gehört, dass Netflix so interne Tools bauen, die in die Richtung gehen.

I: Es gibt auch noch Movie Slate 8, das ist auch ähnlich wie mein Konzept. Noch als letzte Frage: Könnte euch das Playbacksystem einen finanziellen Mehrwert bieten?

E: Also 100 %. Das wäre auch etwas, von dem man die Leute auch gar nicht so krass überzeugen müsste. Die Hemmschwelle, etwas Neues zu verwenden, ist immer hoch. Aber das wäre eine Sache, bei der Mehrwert ziemlich schnell erkannt werden würde. Spart auf jeden Fall Stunden in der Postproduktion.

I: Genau, die Stunden in der Postproduktion und das bessere Handling von analogem Film sind Verbesserungen, auf die ich abziele.

E: Ja, also sag Bescheid, wenn ihr da irgendeinen Prototypen habt, dann würden wir den sehr gerne auch verwenden. Wir drehen auch mehr und mehr analog, da gibt es auf jeden Fall einen Aufwärtstrend. Wir haben früher nie analog gedreht, aber inzwischen liegen bei uns tonnenweise Rollen in Kühlschränken rum – vor allem 16 mm und 35 mm. Es ist natürlich teuer, aber es gibt in Berlin auch eine ziemlich gute Infrastruktur.

I: Ja spannend, wir arbeiten gerade auch an einer Entwicklungsmaschine, die auch geringe Mengen entwickeln kann. Das freut mich zu hören, dass es da einen Aufwärtstrend gibt.

E: Ja, voll. Manche Artists legen da sehr großen Wert drauf. Ich denke, vor allem durch AI wird das immer beliebter, weil das eine Art Gegenpol ist, da du dadurch beweisen kannst, dass das echt ist.

I: Hättest du noch Verbesserungsvorschläge für das Audioplaybacksystem?

E: Manchmal könnten mehrere Lautsprecher, die gleichzeitig den Track abspielen, praktisch sein, wenn man über größere Distanzen filmt. Oder auch für die Regie als Playback.

I: Ein guter Vorschlag, vielen Dank! Damit ist unsere Zeit gleich vorbei, ich bedanke mich für das Interview!

E: Sehr gerne!

C Anhang: Interviewtranskript mit Experte C

Datum: 13.04.2026, 16:00

I: Ich würde dir gerne kurz mein Konzept erklären: Ich entwickle ein Audio-Playback-System für Musikvideodrehs auf Basis eines ESP32. Es soll gleichzeitig einen Song und einen absolut synchronen Timecode ausgeben. Die Idee entstand bei einem analogen Kurzfilmprojekt an der HAW. Wir hatten eine Tanzchoreographie und haben den Ton einfach über eine Bluetooth-Box abgespielt. Da Analogfilm sehr teuer ist, konnten wir nicht den ganzen Song mehrfach durchlaufen lassen, sondern mussten die Szenen anstückeln. Ohne Timecode und Ton war es extrem schwierig, das in der Postproduktion wieder synchron zu bekommen. Mein Gerät soll es ermöglichen, Szenen vorab zu definieren. Das Gerät spielt dann exakt diesen Part mit Timecode ab. Passend dazu habe ich eine digitale Filmklappe gebaut, die über einen Lichtblitz auch Analogkameras automatisch synchronisieren kann. Der Grund für das Interview ist, dass ich eine Modellberechnung machen möchte, und dazu bräuchte ich eine Annahme für einen Standard-Workflow für Musikvideodrehs, um herauszufinden, an welchen Momenten man sich Zeit sparen könnte.

E: Das klingt total sinnvoll. Es gibt Musikvideos, bei denen das extrem viel hilft, während andere Sachen eher spontan entstehen.

I: Wie spielst du aktuell bei Musikvideodrehs das Playback ab?

E: Meistens über Bluetooth-Boxen. Oft ist das noch keine gemasterte Version des Liedes, sondern der aktuelle Stand über das Handy.

I: Und wie synchronisiert ihr das im Nachhinein?

E: Wenn wir digital drehen, probieren wir meistens, erst einmal komplette Durchläufe des Songs an jeder Location zu machen. Dann hat man eine Basis für den Rohschnitt und kann weitere Takes darüberlegen. Bei Analogdrehs machen wir keine kompletten Durchläufe, sondern planen die Takes sehr genau abgesetzt. Da legen wir den Ton tatsächlich manuell nach den Lippenbewegungen an. Das ist mühsame „Friemelarbeit“, aber das übliche Prozedere. Timecode hatte ich bei Musikvideos bisher noch nie im Einsatz.

I: Wie detailliert plant ihr im Vorfeld? Steht vorher fest, welcher Songpart zu welcher Sekunde gedreht wird?

E: Das ist unterschiedlich. Oft sind nur Bilder und Settings geplant, aber die genaue Schnittfolge noch nicht. Bei Techno-Videos kann man im Schnitt wilder schieben als bei klaren Lyrics. Aber gerade im Analogbereich gibt es Produktionen, die wirklich Shot-for-Shot planen. Ein Regisseur, mit dem ich mal zusammengearbeitet hab, mit Architektur-Hintergrund, hat die Bilder sogar 3D-mäßig als Storyboard vorgeplant.

I: Nutzt ihr bei elektronischer Musik auch Taktgeber wie Click-Tracks?

E: Meistens lassen wir das echte Lied laufen, um in die richtige Stimmung zu kommen. Dass etwas anderes als der Song läuft, ist bei mir eher selten. Bei komplett szenischen Musikvideos gibt es auch Drehs ohne Musik im Hintergrund.

I: Hast du Erfahrung mit Zeitlupen, bei denen der Sänger trotzdem synchron bleibt?

E: Ja, das haben wir gemacht. Man dreht zum Beispiel mit 48 Bildern pro Sekunde und spielt den Song in doppelter Geschwindigkeit ab. Das müssen die Künstler aber vorher üben, da es bei längeren Passagen schwierig wird. Ich war mal bei einem Projekt dabei, das komplett in einem One-Take rückwärts gedreht wurde. Da wurde zwei Tage lang geprobt und am dritten Tag gedreht, weil man die Kulisse nur einmal verwenden konnte.

I: Arbeitet ihr mit Mehrkamerasystemen?

E: Selten im Sinne eines Live-Schnitts. Meistens gibt es eine B-Kamera für B-Roll, zum Beispiel auf Super 8, um eine zweite Ebene zu haben. An der HAW gab es aber mal ein Projekt von einem Kommilitonen von mir, der ein Musikvideo mit drei oder vier Kameras live geschnitten hat.

I: Gibt es bei großen Distanzen Probleme mit der Schallverzögerung, etwa bei Drohnen oder langen Brennweiten?

E: Bisher nicht. Wir haben mal mit 250-mm-Brennweite aus großer Entfernung gedreht; da hört man den Ton an der Kamera ohnehin nicht mehr und muss das Material später komplett nach Sicht anlegen.

I: Wie dokumentierst du Takes am Set für die Postproduktion?

E: Bei Musikvideos eher wenig, da die Regisseure oft selbst schneiden und das Material kennen. Bei Dokumentationen schreiben die Tonleute oft Berichte und nutzen Tentacle Sync für die Dateibenennung. Bei szenischen Produktionen gibt es eine Regieassistentin für den Schnittbericht und eine Kameraassistentin für die Klappe.

I: Wie viele Takes macht ihr pro Shot?

E: Das ist teilweise eine Abwägungsfrage: Geht nach vielen Takes irgendwann die Stimmung nach unten oder geht man lieber auf Nummer sicher? Es kommt auch ein bisschen auf das Konzept des Videos drauf an, ob es eher Run-and-Gun oder eher gestaged ist. Bei Analogfilm meistens nur einen bis drei Takes. Wir proben dafür vorher extrem viel. Teilweise filmen wir die Proben mit dem Handy durch den Kamerasucher, damit die Künstler genau sehen, wie sie im Bild wirken, bevor wir das echte Filmmaterial verbrauchen.

I: Wie hoch sind die Budgets in deiner Erfahrung?

E: Das reicht von Pro-Bono bis hin zu Indie-Artists mit etwa 1.500 bis 2.000 Euro. Oft werden heute keine kompletten Videos mehr für YouTube gemacht, sondern nur noch Social-Media-Snippets als Reels oder TikToks, weil die einen höheren Promo-Stellenwert haben. Selbst bei Major-Artists liegen die Budgets oft nur im niedrigen fünfstelligen Bereich, vielleicht 10.000 bis 15.000 Euro.

I: Wo siehst du Ineffizienzen oder typische Fehlerquellen?

E: Bei jüngeren Musikern ein Erwartungsmanagement, da die häufig nicht einschätzen können, was alles bei Drehs möglich ist und was nicht.

I: Würde mein System einen Mehrwert bieten? Und hättest du noch Verbesserungsvorschläge?

E: Gerade für analoge und klar durchgeplante Drehs auf jeden Fall. Wichtig ist eine Benutzeroberfläche, die sich wie „Consumer-Technik“ anfühlt. Viele Leute im Musikvideobereich kommen nicht von der technischen Seite und brauchen eine intuitive Bedienung. Es muss flexibel bleiben, falls man am Set spontan etwas ändert. Wie soll der Ton abgespielt sein?

I: Es soll zwei physische Ausgänge haben, um Bluetooth-Latenzen zu umgehen und stabil zu laufen. Das Gerät könnte man auch mit einem Tentacle synchronisieren. Außerdem soll es auch über ein eigenes Netzwerk mit einem Computer oder dem Smartphone bedient werden können.

E: Ich glaube, es ist wichtig, dass es auch an dem Gerät selbst bedient werden kann! Es gibt so viele Apps für verschiedene Geräte wie Lampen usw. auf dem Set. Ich bin dankbar für jedes Gerät, das man ohne Smartphone – also stand-alone - bedienen kann.

I: Ah, das ist guter Input, macht auf jeden Fall Sinn!

E: Ich glaube auch, dass dieses Gerät auch eher im High-Class-Bereich seine Anwendung findet, weil viel Low-Budget-Drehs auch mehr auf Social Media gehen. Es wird wahrscheinlich eine Abwägung zwischen „Wie viel spare ich mir im Schnitt?“ und „Wie viel kostet mich dieses Gerät?“.

I: Vielen Dank für deine Zeit!

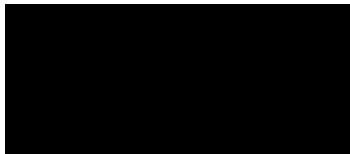
E: Sehr gerne! Ich wünsch dir viel Erfolg!

Eigenständigkeitserklärung

Hiermit versichere ich, dass ich die vorstehende Masterthesis mit dem Titel

Timecode-basiertes Audioplayback-System für choreographisches Filmen

– bzw. im Falle einer Gruppenarbeit die entsprechend gekennzeichneten Teile der Arbeit – selbstständig ohne fremde Hilfe gefertigt und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe. Wörtlich oder dem Sinn nach aus anderen Werken entnommene Stellen sind unter Angabe der Quelle kenntlich gemacht. Diese Arbeit hat in gleicher oder ähnlicher Form noch keiner Prüfungsbehörde vorgelegen.



Hamburg, 12.06.2026

Ort, Datum, Unterschrift