

Sticky Annotations

Recherche im Zuge des Forschungsprojekts "Research Video"

Martin Grödl

Moritz Resl

Institute for the Performing Arts and Film (IPF)

Zürcher Hochschule der Künste

1. April 2019

Inhalt

Einleitung	2
1. Trimfunktion	3
2. Timecode	9
3. EDL	14
4. Audiospur	21
5. Untertitelspur	25
Ergebnis	28

Einleitung

Problem und Ausgangslage

Bisher wurde im Projekt Research Video davon ausgegangen, dass ein Videoschnittvorgang nur einmal stattfindet, nämlich ganz zu Beginn, noch vor der eigentlichen Analysephase mittels Videoannotation. Hier wurde von einem 'Main Video File' gesprochen, dass während des gesamten Analyse-, Annotations- und Publikationsprozesses unverändert bleibt.

Diese Annahme stellte sich als praxisuntauglich heraus, da die beteiligten ForscherInnen schon mit (noch ungeschnittenem) Rohmaterial den Annotationsprozess starten wollten. Ein realistischer Workflow muss daher einen mehrmaligen Wechsel zwischen Verdichtung bzw. Selektion des Videomaterials mittels Videoschnitt sowie Annotation des Materials beinhalten.

Das Grundproblem hierbei ist, dass die Annotationen der im Projekt zum Einsatz kommenden Annotationssysteme (Research Video Prototype, sowie Piecemaker) aus technischen Gründen zeitbasiert sind, d.h. sich auf konkrete Zeitpunkte (Timestamps) im annotierten Video beziehen. Ein Schneidevorgang stört diese Zuordnung, da sich Videomaterial auf der Zeitachse verschiebt, umordnet etc. und macht somit etwaige Annotationen unbrauchbar.

Projektziel ist daher die Suche nach bzw. das Erarbeiten von Lösungen und möglichen Workflows wie die inhaltliche Zuordnung von Annotationen über Schnittvorgänge hinweg erhalten werden kann.

Folgende Ansätze werden diskutiert:

- 1) Eingeschränkte Schnittfunktion innerhalb des Annotationstools selbst. ("Trimfunktion" ähnlich Youtube, iOS etc.)
- 2) Nachvollziehen des Schnittvorgangs mittels Timecodeformaten bzw. Spuren (z.B. SMPTE Timecode)
- 3) Nachvollziehen des Schnittvorgangs mittels EDL (Edit Decision Lists)
- 4) Einbetten von Timinginformationen als zusätzliche Audiospur
- 5) Einbetten von Timinginformationen bzw. Annotationsdaten als Untertitelspur

Für alle praktischen Tests in dem eine Videoschnittsoftware nötig ist, verwenden wir beispielhaft Adobe Premiere Pro in der aktuellen Version (CC 2019). Dieses Schnittprogramm wird auch in der Lehre an der ZHDK am Häufigsten eingesetzt.

1. Trimfunktion

Die Grundidee dieses Ansatzes ist es, eine eingeschränkte Schnittfunktion in Research Video selbst anzubieten. Somit braucht der Benutzer nicht auf eine zusätzliche, externe Schnittsoftware zurückzugreifen sondern kann das Rohmaterial direkt in Research Video laden und entsprechend kürzen. Dabei wird nicht genutztes Rohmaterial nicht gelöscht, sondern nur ausgeblendet, sodass es jederzeit wieder eingeblendet und annotiert werden kann. Da dieser Prozess in Research Video selbst stattfindet, können in jedem Bearbeitungsschritt bereits vorhandene Annotationen mitberücksichtigt werden und die feste Verbindung zwischen Annotation und Videomaterial bleibt zu jeder Zeit aufrecht.

Bei dieser Methode muss der Umfang der Schnittfunktionen stark aber sinnvoll eingeschränkt werden, da es nicht praktikabel sein wird, alle Funktionen einer gängigen Videoschnittsoftware zu implementieren. Eine solche enthält typischerweise mehrere Video- und Audiospuren und kann beliebig viele Quellvideos verarbeiten, die beliebig gestückelt und angeordnet werden können.

Ein mögliches Szenario ist die Beschränkung auf ein einzelnes (langes) Videofile mit der Möglichkeit in einen Bearbeitungsmodus zu wechseln, in dem die gewünschten Bereiche markiert werden können ("Highlighting"). Alle anderen Teile des Videos werden außerhalb des Bearbeitungsmodus versteckt, als wären sie herausgeschnitten. Das entspricht einer (nicht-destruktiven) Schnittfunktion mit einem einzigen Quellvideo ohne Umsortieren der Clips.

Im Folgenden betrachten wir eine Reihe von gängigen Programmen und Apps die eingeschränkte Schnittfunktionen, oft "Trimming" genannt, anbieten: YouTube, Vimeo, QuickTime Player (macOS), Photos (iOS), iMovie (iOS).

YouTube

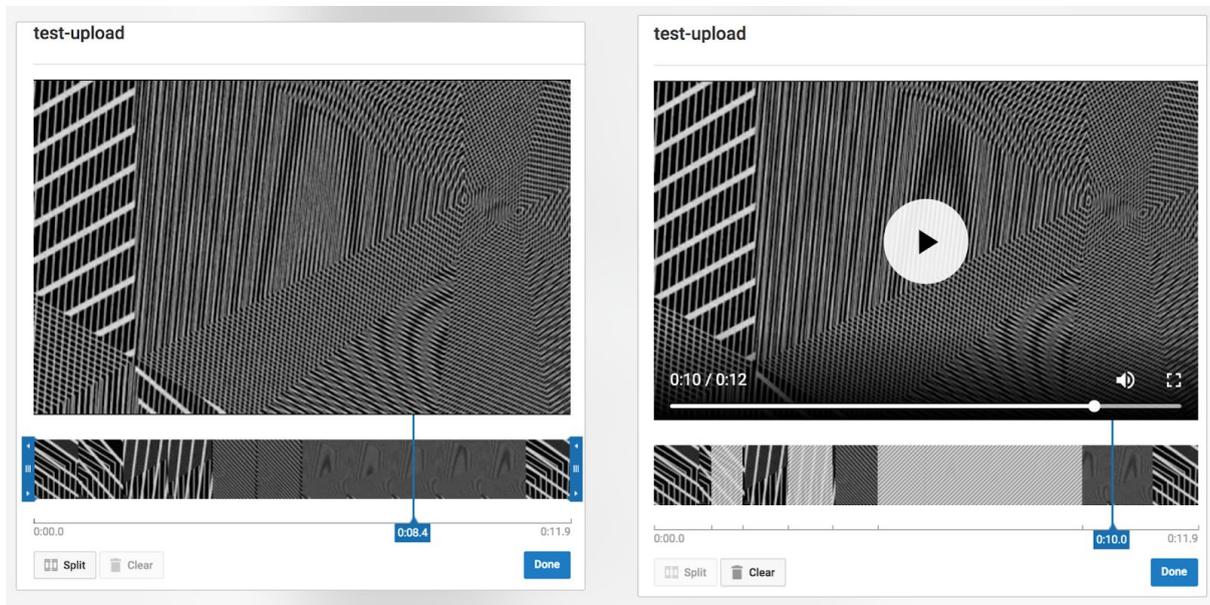
Die populäre Videoplattform YouTube bietet zwei Möglichkeiten, ein hochgeladenes Video zu bearbeiten. Diese sind in der YouTube Onlinehilfe im Abschnitt "Trim your videos" beschrieben (<https://support.google.com/youtube/answer/9057455?hl=en>).

Die beiden Optionen "Creator Studio classic" und "YouTube Studio beta" unterscheiden sich nur durch die Benutzeroberfläche, nicht durch die Funktionalität. YouTube Studio beta bietet eine überarbeitete, moderne Oberfläche und wird in Zukunft das Creator Studio classic ersetzen.

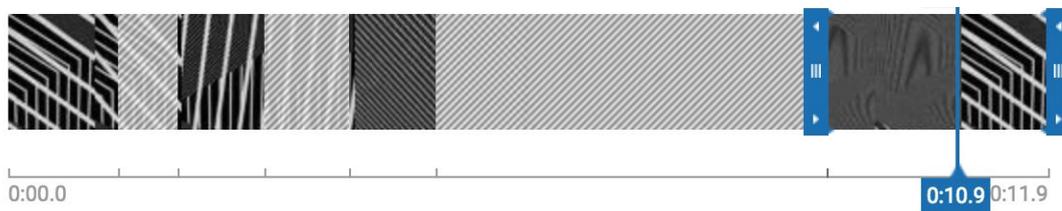
Es gibt zwei Operationen:

- **Trim:** Start und Ende kann gekürzt/beschnitten werden indem die beiden blauen Anfasser verschoben werden.
- **Split:** Aufspalten der Timeline in zwei Abschnitte, die getrennt voneinander getrimmt werden können. Das ermöglicht es, beliebige Bereiche innerhalb des Videos herauszuschneiden.

Creator Studio classic Trim-Ansicht (links) und Ansicht mit ausgeblendeten Abschnitten (rechts):

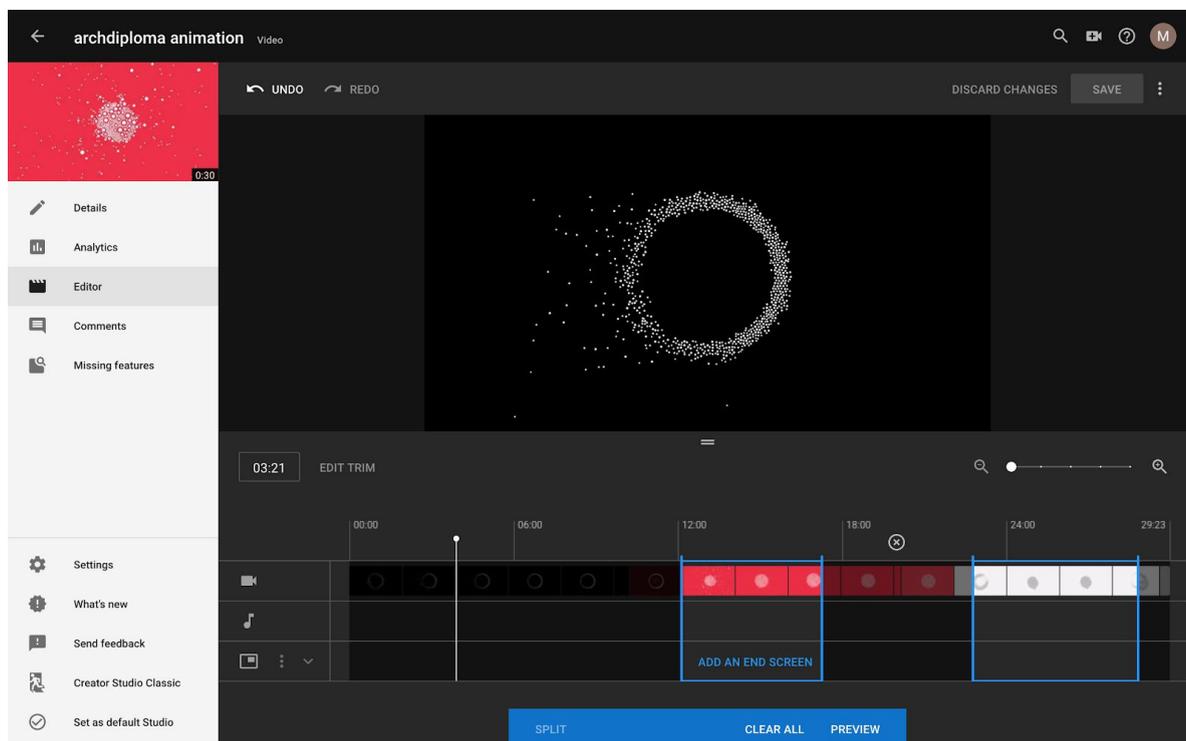


Detailansicht der Timeline mit ausgeblendeten Abschnitten (hellgrau schraffiert):



Die Bearbeitung wird mit einem Klick auf den "Done" Button bestätigt und die grau markierten Bereiche werden aus dem Video entfernt.

YouTube Studio (beta) bietet im Menüpunkt "Editor" die selben Funktionalität wie gerade beschrieben, mit neuem Interface und Komfortfunktionen wie Undo/Redo:



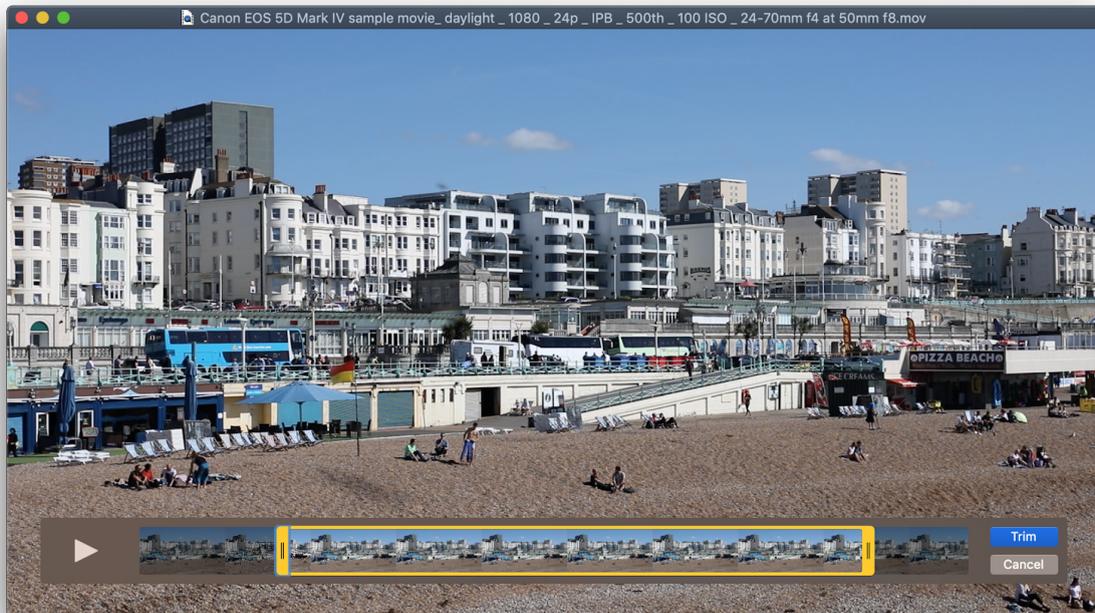
Die Bearbeitung ist nur mit je einem einzigen Video gleichzeitig möglich, es gibt keine Möglichkeit Material aus einem anderen Video einzufügen. Bereiche des Videos können auch nicht verschoben oder vertauscht werden, die Bearbeitung beschränkt sich auf Ausblenden von unerwünschten Abschnitten, wobei keine Lücken entstehen und die Chronologie der Aufnahme immer erhalten bleibt.

Vimeo

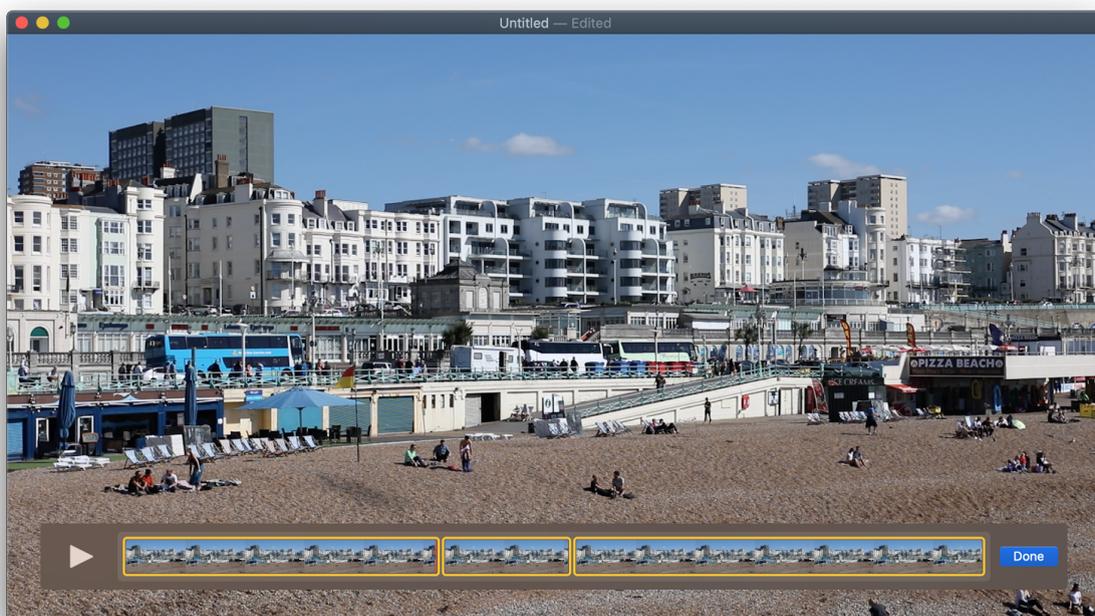
Die nach YouTube populärste Videosharing-Plattform bietet keinerlei Schnittfunktionen an. Lediglich ein Blogeintrag aus dem Jahr 2010 gibt einige Tipps zum Umgang mit Schnittprogrammen (<https://vimeo.com/blog/post/video-101-editing-basics>).

QuickTime Player

QuickTime ist der Standard Video-Player in macOS seit 2009. Er bietet unter "Edit / Trim" die Möglichkeit Anfang und Ende des Videos zu beschneiden:



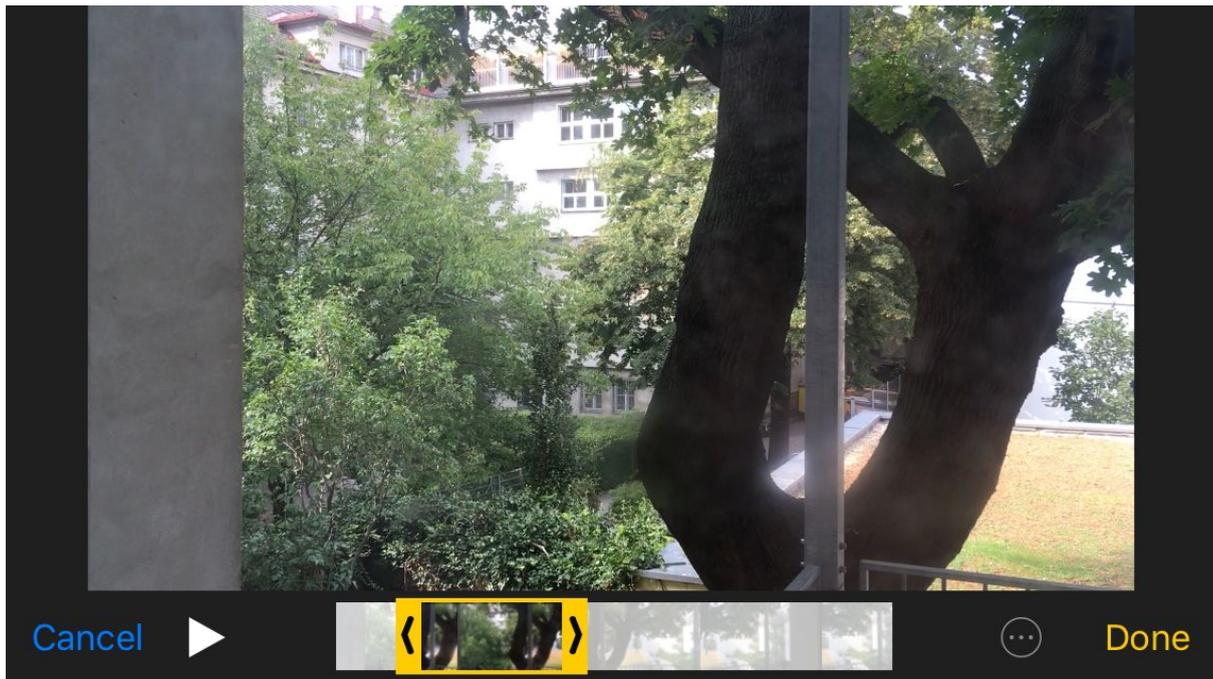
Mittels "Edit / Split Clip" kann das Video in mehrere Abschnitte aufgeteilt werden, die anschließend mittels Drag & Drop umsortiert werden können:



Photos (iOS)

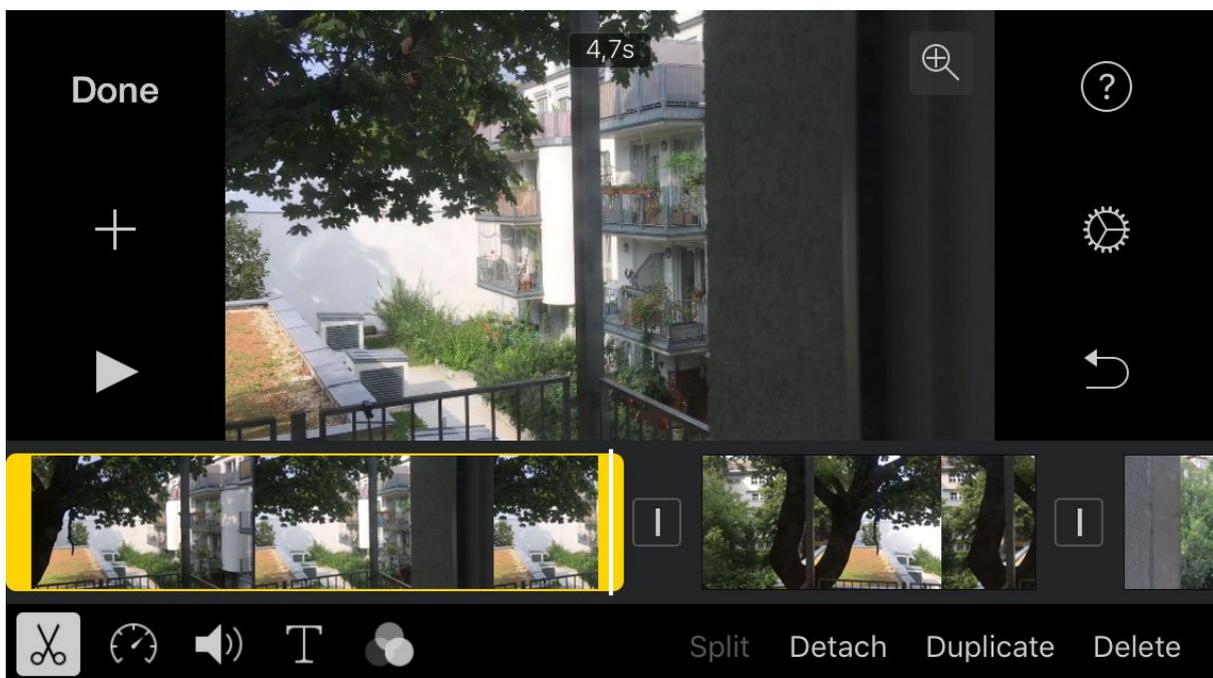
Die Photos App (hier Screenshots vom iPhone SE) in der trotz des Namens auch die mit dem Handy aufgenommene Videos verwaltet werden, bietet nur eine Möglichkeit des Schneidens: "Trim Clip".

Nur Anfang und Ende eines Videos können verschoben werden, es gibt kein "Split" und keine Umsortierung.



iMovie (iOS)

iMovie (im Beispiel wieder auf einem iPhone SE) bietet zusätzlich die aus QuickTime bekannte Split-Funktion mit Umsortierung:



Fazit

Möglicher Workflow in Research Video

- Aktivieren eines "Edit"-Modus innerhalb Research Video
- Gesamtes Video-Rohmaterial wird angezeigt
- Markieren/Ändern von unerwünschten Abschnitten des Videos mittels "Trim + Split" Interface
- Deaktivieren des "Edit"-Modus
- Zuvor markierte Abschnitte werden versteckt. Es bleiben keine Lücken, es gehen weder Videomaterial noch Annotationen verloren.

Vorteile

- Kein externes Schnittprogramm nötig
- Intuitives Konzept: "Trim + Split" entspricht einem "Filter" der Timeline. Nicht benötigte Teile des Videos werden versteckt, die Chronologie bleibt erhalten.
- Nicht destruktiv: Material wird lediglich ausgeblendet.
- Entspricht dem Annotationskonzept: Versteckte Bereiche sind nichts anderes als Annotationen auf dem Gesamtvideo und könnten intern in einer speziellen Annotationsspur abgelegt werden.

Nachteile

- Beruht auf der Annahme des "Main Video File", einer einzelnen Videodatei, die zwar sehr lang sein kann, aber von Anfang an vorhanden sein muss. Es kann kein Material nachträglich hinzugefügt werden.
- Nicht Kernkompetenz/Ziel von Research Video: Erweiterungen in Richtung Umsortieren von Abschnitten oder Hinzufügen von neuem Videomaterial würde zu sehr in Richtung Erstellung eines eigenen Schnittprogramms gehen
- Hoher Umsetzungsaufwand
 - Braucht eigene Ansicht/eigenen Editor zum Verwalten
 - Realtime Mapping der gekürzten Timeline auf die korrekte Zeit beim Abspielen notwendig

Die Idee einer Trimfunktion in Research Video selbst lässt sich zwar elegant mit dem Konzept der Annotation verbinden und bietet Vorteile für den Endanwender, die Erfahrung zeigt aber, dass die Praxis der Annotation ein sehr dynamischer Prozess ist, bei dem mehrere Videodateien verwendet werden. Diese können sich im Laufe der Arbeit auch mehrmals ändern. Somit wird dieser Ansatz mit hoher Wahrscheinlichkeit praktisch ungeeignet sein.

2. Timecode

Der Begriff "Timecode" bezeichnet ein Verfahren, das es ermöglicht jedes Einzelbild eines Videos zu identifizieren, indem jedem Frame (Einzelbild) ein eindeutiger Code zugewiesen wird. Dieser Code besitzt typischerweise das Format HH:MM:SS:FF wobei HH für Stunde, MM für Minute, SS für Sekunde und FF für Frame steht. Dabei ist zu beachten, dass die Frameangabe (FF) von der Bildrate des Videos abhängig ist.

Timecode wird im professionellen Film- und Videobereich hauptsächlich verwendet, um in der Bild- und Tonmaterial aus verschiedenen Quellen zu synchronisieren, da oft mehrere Kameras und Tonaufnahmegeräte gleichzeitig zum Einsatz kommen. Viele semi-professionelle Video- und digitale Spiegelreflexkameras zeichnen ebenfalls Timecode auf.

Wie kann Timecode in Research Video genutzt werden? Falls es eine Möglichkeit gibt die Timecodes der im Schnittvorgang beteiligten Quellvideos ins geschnittene Video zu übernehmen, könnte der Schnittvorgang einfach anhand des fertig geschnittenen Videofiles nachvollzogen und die Annotationen automatisch angepasst werden. Im Folgenden wird diese Möglichkeit anhand von Adobe Premiere Pro evaluiert.

Anmerkung: Es gibt mehrere Arten von Timecodes. In diesem Abschnitt sprechen wir von Timecode, der als Metadatum oder Datenspur im Videofile integriert wird. Die Möglichkeit eines als Audiosignal eingebetteten Timecodes wird in Abschnitt 4 (Audiospur) besprochen.

Zunächst zwei Definition des Begriffs "Timecode" aus den Manuals zweier führender Videoschnittprogramme:

Aus: Final Cut Pro X Help (Glossar: 'timecode')

A signal recorded with your video that uniquely identifies each frame. By default, timecode appears in Final Cut Pro in the format hours: minutes: seconds: frames. Timecode supports a variety of functions in Final Cut Pro, including timeline playback, syncing video and audio clip items, navigating through projects in the timeline, and moving and trimming clips.

Aus: Premiere Pro User Guide

(<https://helpx.adobe.com/premiere-pro/using/timecode.html>)

Many camcorders and high-end video decks record timecode, which marks specific frames with unique addresses. Timecode is important whenever you want to capture the same frames that were identified or captured previously, as in the following tasks:

- [...]
- You plan to export sequences to another system by using EDL.
- [...]
- You plan to synchronize captured video with audio recorded separately.

Unlike the numbers on time counters found in home VCRs, timecode is recorded onto videotape as part of the video signal. If footage lacks timecode, you can add it by copying it with a camera or deck that writes timecode. You can then log or capture the video from that device.

Viele Digitalkameras können Timecode aufzeichnen. Einige Beispiele:

Timecode Einstellungen einer Sony DSLR-Kamera (aus <https://youtu.be/A8Ca00P90oo?t=47>):



Timecode Einstellung der Canon EOS 5D Mark III (Auszug aus der Betriebsanleitung):

MENU Setting the Time Code

Time code	
Count up	Rec run
Start time setting	
Movie rec count	Rec time
Movie play count	Rec time
HDMI	
Drop frame	Enable

The time code is a time reference recorded automatically to synchronize the video and audio during movie shooting. It is recorded at all times in the following units: hours, minutes, seconds, and frames. It is mainly used during movie editing.

Set the time code with [**5**: Time code]. In the <[**A**+]> mode, this function is displayed in [**3**].

Timecode Einstellung der Sony Alpha 7 Mark III (Auszug aus dem Online Manual):

SONY
Hilfe

Digitalkamera mit Wechselobjektiv
ILCE-7M3 α7III

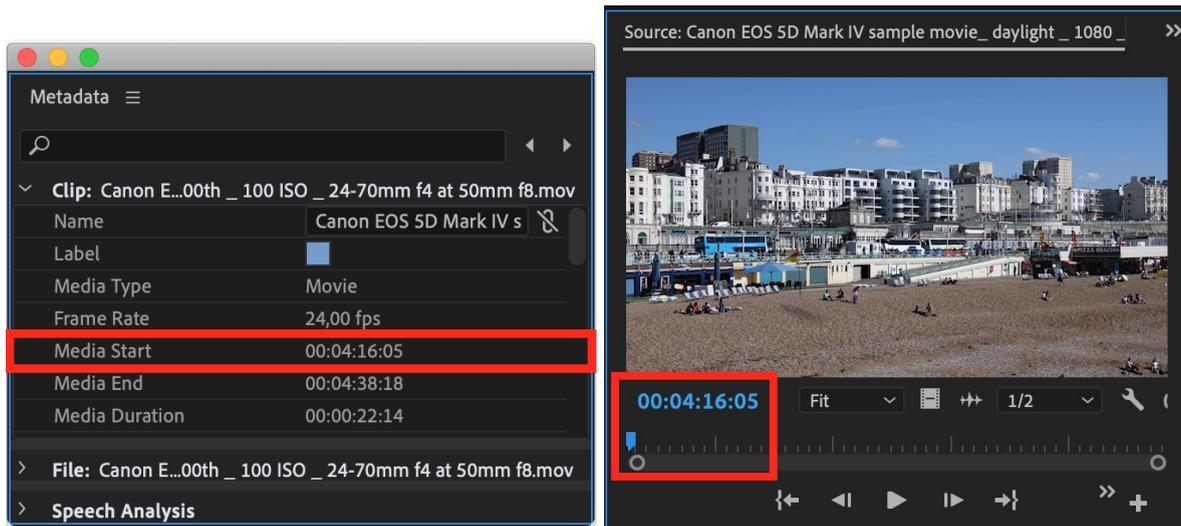
TC/UB-Einstlg.

Die Timecode (TC)- und User Bit-(UB)-Informationen können als Zusatzdaten zu Filmen aufgezeichnet werden.

1 MENU → (Einstellung) → [TC/UB-Einstlg.] → zu ändernder Einstellwert.

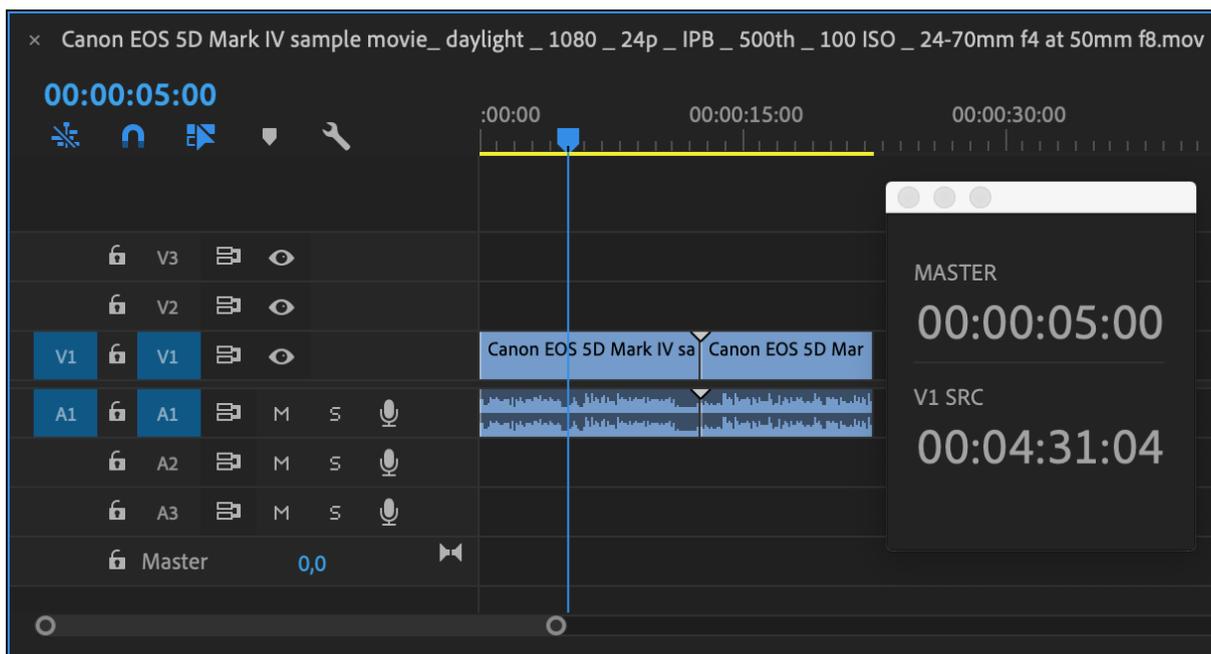
Um Format und Eigenschaften eines solchen Timecodes zu untersuchen haben wir ein unbearbeitetes Videofile einer Canon EOS 5D Mark IV getestet. Das Video kann unter <https://vimeo.com/182286275> (Download -> Original) heruntergeladen werden.

In Premiere Pro zeigt sich die Einbettung des Timecode folgendermaßen:



In den Metadaten scheint ein "Media Start" von "00:04:16:05" auf, und auch im Source Monitor des Clips hat das erste Bild des Videos einen Timecode von "00:04:16:05". Bei einem Video ohne Timecode würde man einen Beginn bei 00:00:00:00 erwarten.

Im Folgenden Screenshot testen wir das Verhalten des Timecodes beim Schnitt:



Wir haben ungefähr in der Mitte des 22 Sekunden langen Videos geschnitten und die beiden entstehenden Teile vertauscht. Der Abspielknopf steht nun bei 5 Sekunden (MASTER). V1 SRC zeigt den korrekten Timecode des Quellmaterials an dieser Stelle. Falls es möglich ist, den Timecode aus V1 SRC zusammen mit dem geschnittenen Video zu exportieren, könnten die Schnitte in Research Video nachvollzogen werden.

Zur genaueren Inspektion der Ausgangs-Video-datei verwenden wir das Programm 'ffprobe' (Version 4.1.2) aus der FFmpeg Software Suite (<https://ffmpeg.org>), einer Sammlung an Werkzeugen zur Verarbeitung digitaler Audio- und Videostreams.

Der verwendete Befehl war:

```
ffprobe movie.mov -hide_banner -select_streams d -show_streams -show_data
```

Zusammengefasstes Ergebnis:

- QuickTime Movie Container Format (mov)
- Zusätzlicher Datentrack mit Timecode
 - codec_tag_string=tmcd
 - codec_tag=0x64636d74
- Timecode Track hat dieselbe Bildrate (24 fps) und Bildanzahl (768 frames) wie der Videotrack
 - time_base=1/24000
 - nb_frames=542
- Timecode ist 32 Bit pro Frame (bei 24 fps, 768 frames)
 - bit_rate=768
- Zusätzlich wird der Start Timecode in einem Metadatenfeld angegeben
 - TAG:timecode=00:04:16:05

Mit dem folgenden Befehl lesen wir noch testweise die eingebetteten Timecode-Daten aus:

```
ffprobe movie.mov -hide_banner -select_streams d -show_streams -show_data -show_packets
```

Das Ergebnis ist ein 32 Bit Hex Wert pro Frame, hier nur das Ergebnis des ersten und letzten Frames:

- Erster Frame
 - 0000 1805 (hexadezimal)
 - entspricht 6149 (dezimal)
 - entspricht 00:04:16:05 (bei 24 fps)
- Letzter Frame
 - 0000 1a22 (hexadezimal)
 - entspricht 6690 (dezimal)
 - entspricht 00:04:38:18 (bei 24 fps)

Das entspricht (erwartungsgemäß) genau den Angaben in Premiere Pro, wichtig ist aber, dass ein fortlaufender Timecode zu jedem Einzelbild separat gespeichert wird. Nur durch diese Eigenschaft hat der Timecode einen Wert für die Verwendung in Research Video.

Anmerkung: Dieses Timecode Format ist beschrieben im Abschnitt "Timecode Media" von in der Spezifikation zu Apple QuickTime (https://developer.apple.com/library/archive/documentation/QuickTime/QTFF/QTFFChap3/qtff3.html#//apple_ref/doc/uid/TP40000939-CH205-69831)

Nun ist noch zu klären ob die Timecode-Spur im geschnittenen und aus Premiere Pro exportierten Video enthalten ist. Dazu haben wir in zwei unterschiedlichen Formaten exportiert und die erhaltenen Dateien wieder mit ffprobe auf Timecode tracks untersucht. Das Ergebnis kurz zusammengefasst:

- Format: H.264 (MP4 Container)
 - Kein Timecode Track (nur Audio + Video)
 - Nur Start Timecode kann exportiert werden (Option: "Set Start Timecode")
- Format: Quicktime (ProRes 422)
 - Nur Start Timecode im Ergebnis (als Metadata Tag sowie Timecode Track mit einem einzelnen Frame)

Obwohl beim Quicktime Format ein Timecode Track exportiert wird, bleibt nur der Timecode des ersten Frames erhalten. Das ist ausreichend für die Synchronisation mehrerer Videos, aber ungeeignet für den vorliegenden Anwendungsfall mit Research Video.

Fazit

- Beim Rendern eines neuen Videos aus Premiere wird ein neuer, fortlaufender Timecode erzeugt
- Nur der Beginn-Timecode (Media Start) kann manuell gesetzt werden
- Informationen zum Schnittvorgang gehen verloren
- Diese Form des Timecodes ist daher ungeeignet für Research Video

3. EDL

Die "Edit Decision List" (EDL) ist ein relativ altes Verfahren, entstanden als noch hauptsächlich Videobänder zur Aufzeichnung von TV-Produktionen verwendet wurden.

Prinzipiell ist eine EDL ein Textdokument, das in Listenform die Zusammenstellung eines finalen Videos aus Teilen von Quellvideos beschreibt. Da die EDL die Schnitte eines bearbeiteten Videos ("Edit Decisions") strukturiert beschreibt, könnte sie dazu benutzt werden den Schnittvorgang innerhalb von Research Video nachzuvollziehen und die Annotationen entsprechend zu transformieren.

Nachfolgend werden die Möglichkeiten des EDL-Exports in Adobe Premiere Pro und die Nutzbarkeit in Research Video untersucht.

EDL ist kein einheitlicher Standard, es gibt mehrere hersteller-spezifische Formate u. A. von CMX Systems, Grass Valley und Sony. Die meisten EDL-Formate sind textbasiert und menschenlesbar. Eine kurze Übersicht über einige EDL-Formate findet sich unter:

http://www.edlmax.com/EdlMaxHelp/Edl/Edl_Overview.htm

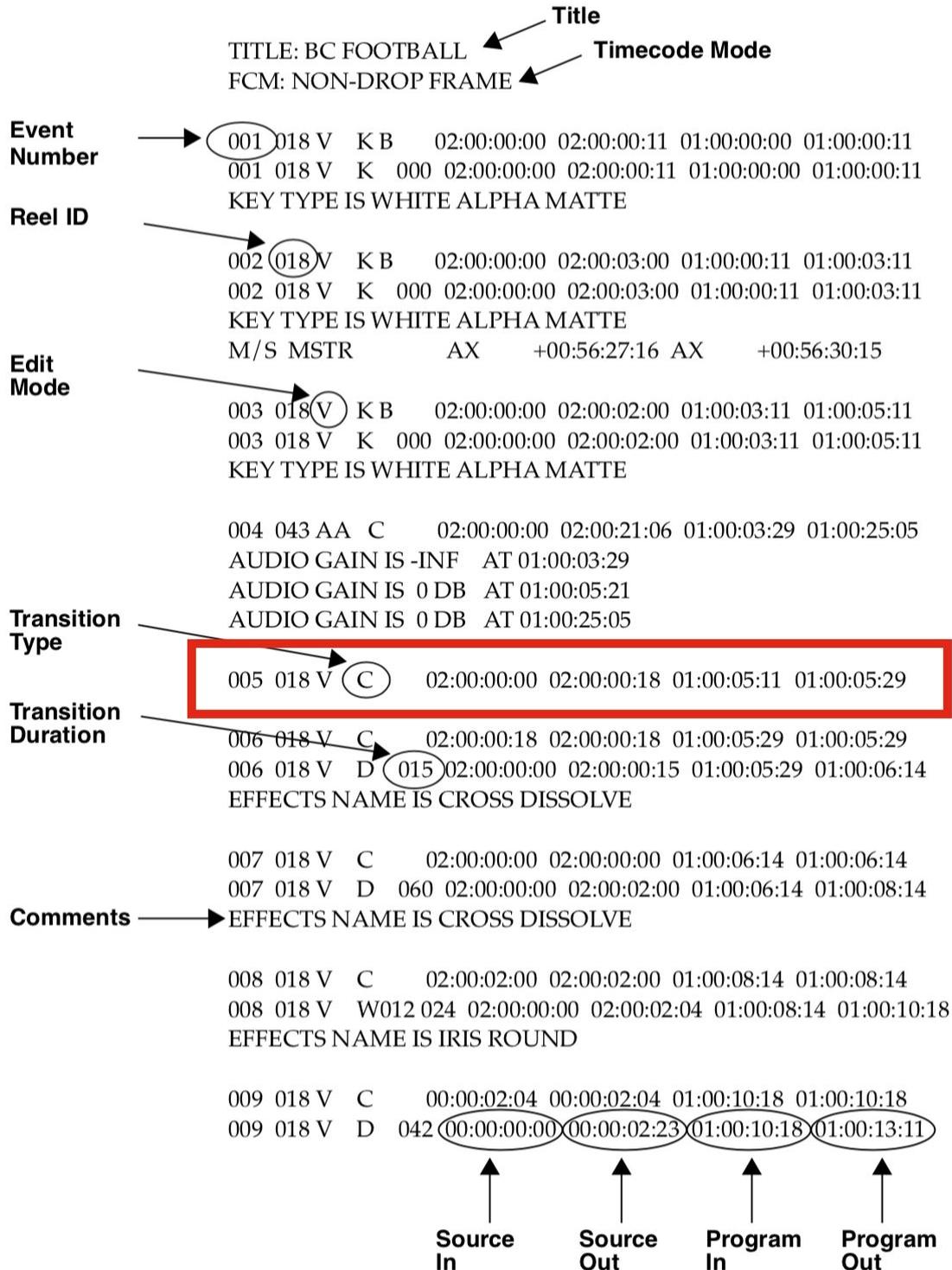
Premiere Pro unterstützt das Format CMX3600, wie folgender Auszug aus dem Premiere Pro User Guide (<https://helpx.adobe.com/premiere-pro/using/exporting-projects-applications.html>) zeigt:

With Premiere Pro, you can export your project as an edit decision list (EDL) in the CMX3600 format. This format is the most widely accepted and most robust of the EDL formats.

Die CMX-3600 der Firma CMS Systems war ein populäres Videoband Schnittgerät in den 1980er Jahren. Eine Spezifikation des CMX 3600 EDL Formats von 1988 ist unter <http://xml.biz/EDL-X/CMX3600.pdf> verfügbar und beginnt wie folgt:

The CMX Edit Decision List is composed of edit decision statements which are intended to work as a data base for the assembly of a video tape master.

Hier ein Beispiel einer CMX-3600 EDL mit Hinweisen zur Struktur:



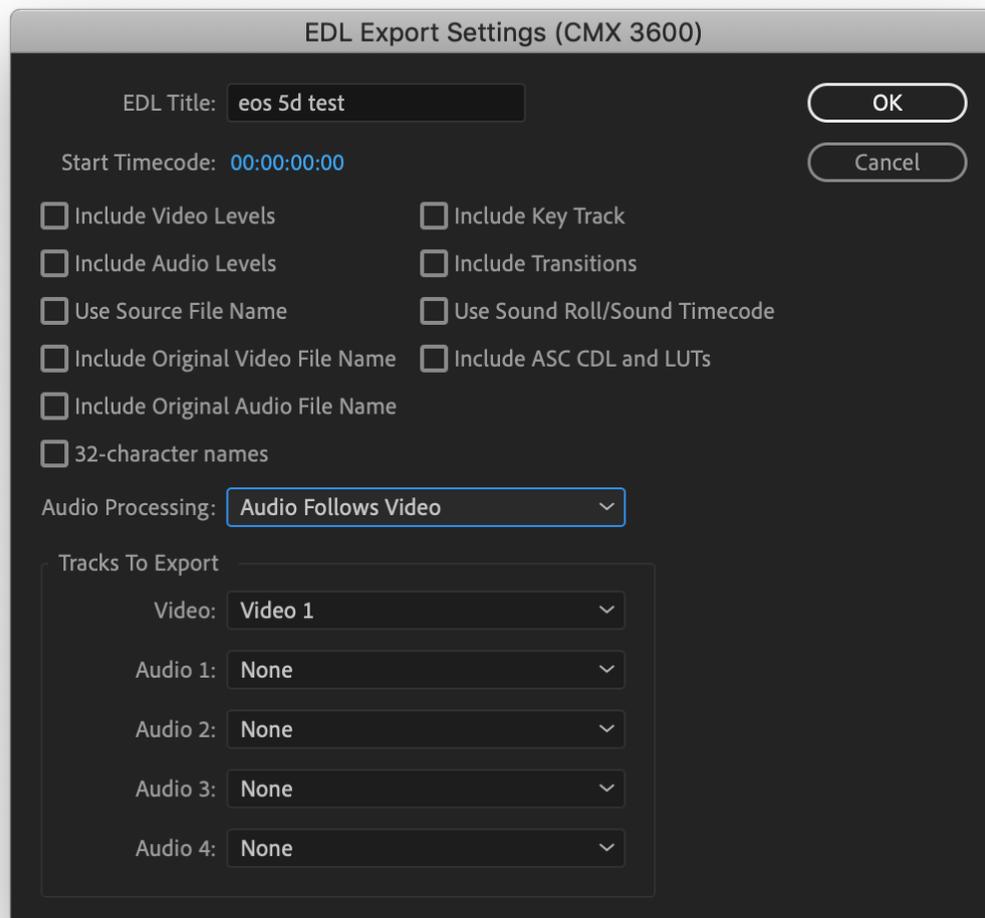
Die rot markierte Zeile beschreibt einen einfachen Schnitt. Es handelt sich um Event Nummer 5 (laufende Nummer der Einträge) und bezieht sich auf Reel ID 18. Der Begriff "Reel" bezeichnet die betreffende "Filmrolle", im Falle von digitalen Videoclips oft nicht relevant. (In Premiere Pro wird der betreffende Videoclip in einem der Zeile folgenden Kommentar genannt) "V" deutet auf eine Bearbeitung der (einzigen) der Videospur und "C", dass es

sich um einen einfachen Schnitt ("Cut") handelt. Die nachfolgenden vier Timecodes beschreiben welcher Teil des Quellvideos ("Source") an welche Stelle der Ausgabe ("Program") kopiert wird.

Aufgrund von Alter und Herkunft des CMX-3600 EDL Formats, besitzt es einige wichtige Einschränkungen:

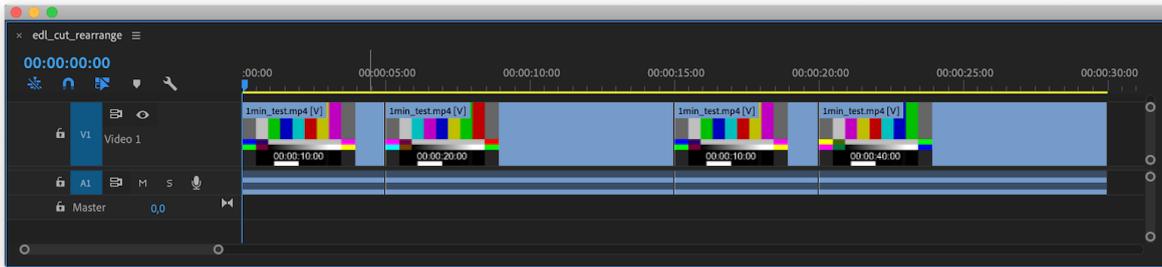
- Limitiert auf 999 "Events" d.h. "Schnittereignisse"
- Nur eine Videospur wird unterstützt
- Vier Audiospuren werden unterstützt
- Keine Unterstützung für verschachtelte Clips
- Keine Unterstützung für variable Geschwindigkeitsänderung von Clips

Der "EDL Export-Settings" Dialog in Premiere Pro bietet eine Reihe an Optionen, die Benutzer beachten muss. Hier wird auch die Einschränkung auf eine Videospur klar, denn man muss unter "Tracks To Export" den Videotrack auswählen der exportiert werden soll:



Wir haben drei Tests mit dem EDL Export durchgeführt, bei denen jeweils unterschiedliche Quellvideos und Schnittsituationen vorliegen und dann die resultierende EDL-Datei analysiert wird:

Test 1: Ein Originalvideo. Teile schneiden, duplizieren und umsortieren



Resultierende EDL:

```

TITLE: edl_cut_rearrange
FCM: NON-DROP FRAME

000001 AX      AA/V C      00:00:10:00 00:00:15:00 00:00:00:00 00:00:05:00
* FROM CLIP NAME: 1min_test.mp4

000002 AX      AA/V C      00:00:20:00 00:00:30:00 00:00:05:00 00:00:15:00
* FROM CLIP NAME: 1min_test.mp4

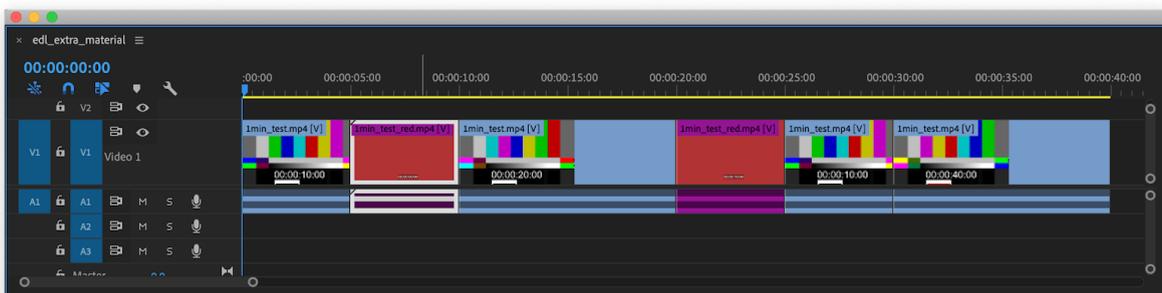
000003 AX      AA/V C      00:00:10:00 00:00:15:00 00:00:15:00 00:00:20:00
* FROM CLIP NAME: 1min_test.mp4

000004 AX      AA/V C      00:00:40:00 00:00:50:00 00:00:20:00 00:00:30:00
* FROM CLIP NAME: 1min_test.mp4

```

Dieser Test bestätigt die grundsätzliche Eignung des EDL Exports für den Anwendungsfall in Research Video. Aus den Timestamps können die Schnitte nachvollzogen und auf die Annotationen übertragen werden. Wie beim Videomaterial selbst können Annotationen gelöscht, verschoben oder dupliziert werden. Da es nur eine einzige beteiligte Videodatei (1min_test.mp4) gibt, besteht kein Zweifel über die Zuordnung des Videomaterials.

Test 2: Ein Originalvideo + Extramaterial (rot)



Resultierende EDL:

```
TITLE: edl_extra_material
FCM: NON-DROP FRAME

000001 AX      AA/V C      00:00:10:00 00:00:15:00 00:00:00:00 00:00:05:00
* FROM CLIP NAME: 1min_test.mp4

000002 AX      AA/V C      00:00:00:00 00:00:05:00 00:00:05:00 00:00:10:00
* FROM CLIP NAME: 1min_test_red.mp4

000003 AX      AA/V C      00:00:20:00 00:00:30:00 00:00:10:00 00:00:20:00
* FROM CLIP NAME: 1min_test.mp4

000004 AX      AA/V C      00:00:10:00 00:00:15:00 00:00:20:00 00:00:25:00
* FROM CLIP NAME: 1min_test_red.mp4

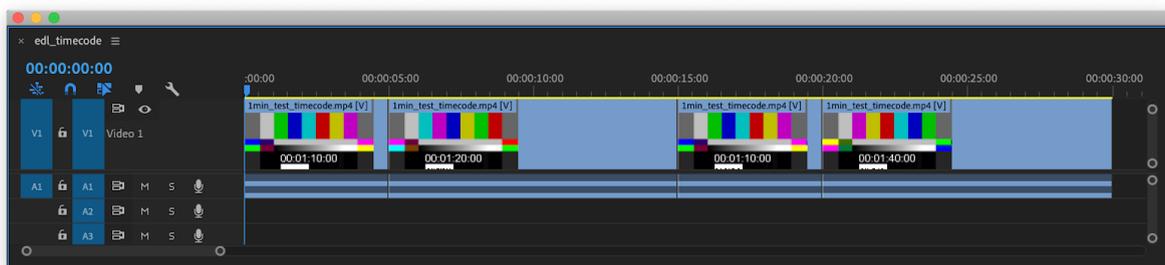
000005 AX      AA/V C      00:00:10:00 00:00:15:00 00:00:25:00 00:00:30:00
* FROM CLIP NAME: 1min_test.mp4

000006 AX      AA/V C      00:00:40:00 00:00:50:00 00:00:30:00 00:00:40:00
* FROM CLIP NAME: 1min_test.mp4
```

In diesem Test gibt es zwei beteiligte Videofiles: Zum Originalvideo (1min_test.mp4), das bereits in Research Video vorhanden ist und annotiert wurde, wird durch Schneiden an mehreren Stellen neues Material aus einer anderen Videodatei (1min_test_red.mp4, in Rot dargestellt) hinzugefügt. Wie im vorigen Test sind die Ausschnitte und Platzierungen der beiden Videodateien aus den Timestamps ablesbar.

Die Zuordnung zu den beiden Videodateien ist über den Timestamps nachfolgende Zeilen auszulesen, in denen die Dateinamen der Clips angegeben sind. Hierbei ist wichtig, dass Research Video den Dateinamen des Originalvideos kennen muss, denn nur diese Schnitte sind für Research Video wichtig und alle anderen müssen ignoriert werden.

Test 3: Originalvideo mit Timecode (beginnt bei 00:01:00:00)



Resultierende EDL:

```
TITLE: edl_timecode
FCM: NON-DROP FRAME

000001 AX      AA/V C      00:01:10:00 00:01:15:00 00:00:00:00 00:00:05:00
* FROM CLIP NAME: 1min_test_timecode.mp4

000002 AX      AA/V C      00:01:20:00 00:01:30:00 00:00:05:00 00:00:15:00
* FROM CLIP NAME: 1min_test_timecode.mp4
```

```
000003 AX      AA/V C      00:01:10:00 00:01:15:00 00:00:15:00 00:00:20:00
* FROM CLIP NAME: 1min_test_timecode.mp4

000004 AX      AA/V C      00:01:40:00 00:01:50:00 00:00:20:00 00:00:30:00
* FROM CLIP NAME: 1min_test_timecode.mp4
```

In diesem Fall wird getestet, wie sich der EDL-Export verhält, wenn das Originalvideo einen eingebetteten Timecode enthält. Es wird eine Videodatei verwendet, die bei 00:01:00:00 beginnt, also einen Offset von 1 min hat. Es zeigt sich, dass die resultierende EDL Datei den Timecode des Originalvideos beachtet. Die angegebenen Zeiten beziehen sich somit nicht auf den Beginn der Videodatei sondern auf den Timecode Offset.

Das hat zur Folge, dass Research Video ebenfalls einen gegebenenfalls vorhandenen Timecode in der Videodatei berücksichtigen muss, damit es nicht zur falschen Transformation von Annotationen kommt. In diesem Beispiel würde Research Video alle Annotationen um eine Minute nach hinten verschoben platzieren, würde der Timecode Offset nicht berücksichtigt. Das ist wichtig, da der Offset prinzipiell beliebig ist und auch Stunden betragen kann, womit Annotationen gänzlich verloren gehen könnten.

Premiere Pro bietet eine alternative Exportmöglichkeit als "Final Cut Pro XML"-Datei. Dieses enthält wie EDL Informationen zum Schnittvorgang und ist aufgrund des XML-Formats ebenfalls sowohl maschinen- als auch menschenlesbar. Es ist neuer als EDL, enthält wesentlich mehr Informationen und hat nicht dessen Einschränkungen der Event- und Spuranzahl. Beispielsweise sind auch die Timecodes-Offsets der verwendeten Videos enthalten, was das Problem aus Test 3 lösen würde. Nachteile dieses Formats sind die aufgrund der Komplexität aufwändigere Verarbeitung und eventuelle Probleme mit neueren Versionen. Im Unterschied zu EDL muss bei Final Cut Pro XML damit gerechnet werden, dass sich das Format in Zukunft ändern könnte. Normalerweise legen große Softwarehersteller wie Adobe (Premiere Pro) und Apple (Final Cut Pro) Wert auf eine gewisse Formatstabilität, eine Garantie gibt es aber nicht.

Fazit

Möglicher Workflow in Research Video

- (Kein Export aus Research Video nötig)
- Video schneiden im Schnittprogramm
 - Es müssen gewissen Einschränkungen beachtet werden (Nur eine Videospur, keine Effekte verwenden)
- EDL Export aus Schnittprogramm
 - Resultiert in einer kleinen Text oder XML Datei
- Import in Research Video
 - Das geschnittene Video sowie die EDL Datei werden geladen
 - Anhand der EDL Datei können bestehende Annotationen angepasst werden

Vorteile

- Kompaktes, textbasiertes, menschenlesbares Format
- Kein zusätzlicher Datentransfer (z.B. Annotationen) von Research Video zum Schnittprogramm nötig.

Nachteile

- Nur eine Videospur wird berücksichtigt.
- Dateiname des Originalvideos muss bekannt sein, falls mehr als eine Videodatei geschnitten wird
- Timecode-Offset des Originalvideos muss ausgelesen werden, da sich EDL darauf bezieht
- Effekte wie Blenden und Geschwindigkeitsänderungen müssen speziell beachtet werden

4. Audiospur

Die Grundidee in diesem Abschnitt ist die Einbettung von Timing-Informationen in einer zusätzlichen Audiospur. Der Hauptvorteil dabei ist, dass jedes Schnittprogramm Audiospuren unterstützt und diese auch meist automatisch zusammen mit der zugehörigen Videospur geschnitten werden. Der Benutzer ist somit nicht auf ein bestimmtes Schnittprogramm eingeschränkt und hat kaum Zusatzaufwand beim Schneiden.

Enthält das zu schneidende Video frame-spezifische Timing-Information eingebettet in einen zusätzlichen Audiokanal, und wird dieser zusammen mit der Videospur geschnitten, lässt sich der Schnittprozess durch Analyse der zusätzlichen Audiospur im fertigen Video nachvollziehen. Research Video könnte somit bereits vorhandene Annotationen entsprechend transformieren.

Technisch ist diese Methode besonders anspruchsvoll, was hauptsächlich daran liegt, dass die Audiospur entgegen ihres eigentlichen Zweckes verwendet wird. Eine wichtige Anforderung ist das exakte Synchronisieren der Timing-Information im Audiokanal mit der Bildrate des Videos. Außerdem muss die Timing-Information so eingebettet werden, dass sie nicht durch gängige Audio-Bearbeitungen (z.B. Lautstärkeanpassung, Änderung der Samplerate) zerstört wird, die Schnittprogramme mitunter automatisch durchführen.

Es gibt einen alten Industriestandard für Timecodes in Form einer Audiospur, der die genannten Bedingungen erfüllt und auch heute noch im semiprofessionellen und professionellen Bereich zum Einsatz kommt: Der Linear Timecode (LTC) nach SMPTE-12M Standard, zuletzt herausgegeben 2008 von der Society of Motion Picture and Television Engineers (SMPTE). Diese Variante eine "Audio-Timecode" und deren mögliche Anwendung in Research Video wird nachfolgend untersucht.

In der professionellen Videoproduktion gibt es eigene Timecode Generatoren, kleine Geräte die den angesprochenen Linear Timecode ausgeben und sich untereinander und mit Kameras bzw. Audiorecordern synchronisieren lassen. Als Beispiele seien hier Tentacle Sync <https://tentaclesync.com> und Mozegear <https://www.mozegear.com> genannt. Professionelle Kameras (z.B. Arri, Canon) und Audioaufnahmegeräte (z.B. Tascam, Zoom) bieten eigene Timecode Ein- bzw. Ausgänge.

Der als Audiosignal codierte SMPTE Timecode identifiziert (wie auch in Abschnitt 2 beschrieben) jedes Einzelbild in der Form HHMMSSFF (HH ... Stunde, MM ... Minute, SS ... Sekunde, FF ... Frame). Zusätzlich können jedem Frame 32 bit freie Daten hinzugefügt werden.

Um den das LTC Timecode Format praktisch zu untersuchen, haben wir mit einem Online Generator (<http://elteeseee.pehrhovey.net>) eine Beispieldatei generiert. Folgende Einstellungen wurden verwendet:

- 25 fps

- 44.1 kHz
- 16bit
- Start Time 00:00:00:00
- Duration 1 min

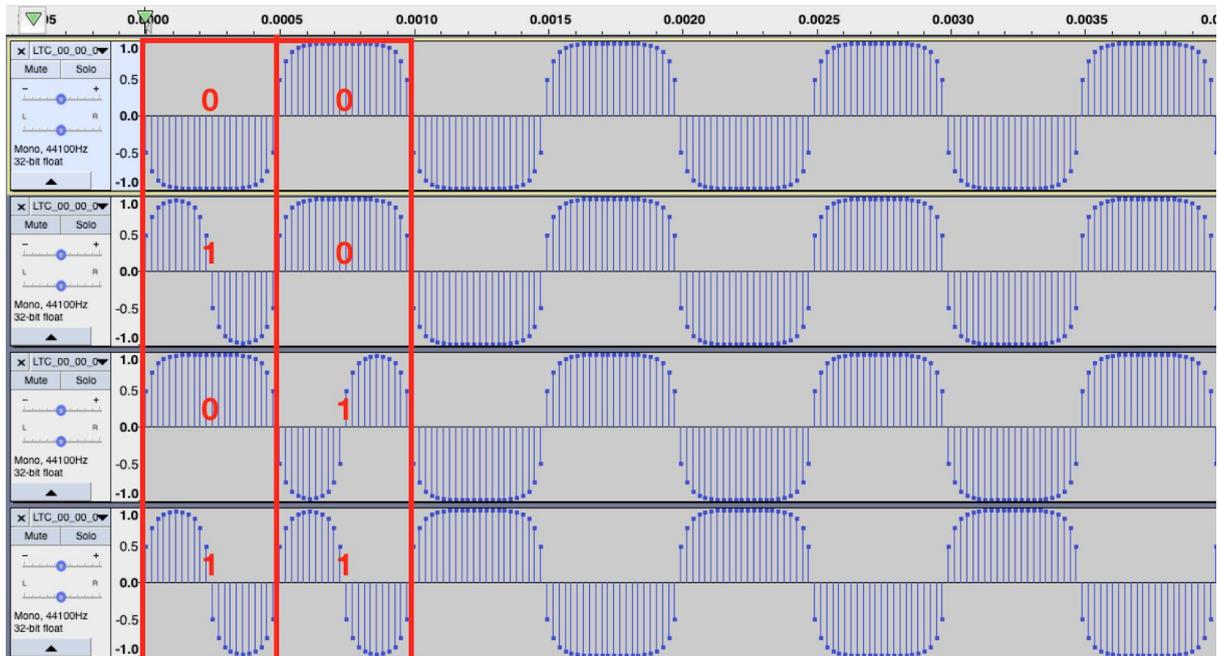
Das resultierende File ist eine Standard Audiodatei im WAV Format die wir mit dem freien Audio-Editor Audacity (<https://www.audacityteam.org>) untersucht haben.

Im folgenden Screenshot ist untereinander das generierte Signal für ersten vier Videoframes dargestellt (jeweils 40 ms Dauer):



Rot markiert ist der Bereich in dem die Framenummer (FF) codiert ist, blau der Synchronisationsbereich, eine eindeutige Sequenz, die aufeinanderfolgende Frames trennt.

Der nächste Screenshot zeigt im Detail die Codierung der Framenummern der ersten vier Einzelbilder:



Rot markiert sind jeweils die ersten beiden Bits der vier aufeinanderfolgenden Frames. Hier ist die Codierung mittels des sog. Biphas Mark Code zu erkennen. Ein "1-Bit" wird durch einen zusätzlichen Übergang in der Mitte codiert, während ein "0-Bit" diesen nicht aufweist.

Weiters können wir die korrekte Nummerierung der ersten vier Einzelbilder verifizieren von Frame 0 bis Frame 3 verifizieren (Binäre Codierung, Least Significant Bit first):

- 00 = 0
- 10 = 1
- 01 = 2
- 11 = 3

Fazit

Möglicher Workflow in Research Video

- Export in Research Video
 - Beim Export wird in das Video eine zusätzliche Audiospur mit fortlaufendem LTC Timecode eingebettet
- Schneiden des Videos
 - Der Timecode Audiotrack wird analog zum Bild "mitgeschnitten" (Standardeinstellung in den Schnittprogrammen)
 - Export des geschnittenen Videos
- Import in Research Video
 - Beim Import wird der geschnittene Timecodetrack identifiziert und ausgelesen. Anhand der Timecodes kann die neue Position jedes Frames bestimmt und die Annotationen entsprechend angepasst werden

Vorteile

- Völlig unabhängig vom Schnittprogramm (jedes Schnittprogramm unterstützt Audio)

- Relativ unabhängig vom Audioformat bei geeigneter Wahl der Codierung (z.B. SMPTE Timecode)
 - Robust gegenüber Änderungen von Samplerate und Bittiefe, z.B. bei Umwandlung des Audioformats im Schnittprogramm.
 - NICHT robust gegenüber Änderung der Bildrate des Videos

Nachteile

- Hoher Umsetzungsaufwand
 - Beim Export aus Research Video muss das Timecodesignal speziell generiert werden.
 - Beim Import in Research Video muss das Timecodesignal identifiziert, gelesen und interpretiert werden.
 - Zusätzlicher Speicherbedarf (je nach Länge des Videos)

5. Untertitelspur

Ähnlich wie im vorigen Abschnitt soll bei dieser Methode eine zusätzliche Spur verwendet werden die als Träger von Timing-Informationen dient und im Schnittprozess mitgeschnitten und mit der finalen Videodatei exportiert wird. Wie bei den anderen Methoden könnte Research Video so die Schnitte nachvollziehen und die Annotationen entsprechend anpassen.

Im Unterschied zur Audiospur eignet sich eine Untertitelspur naturgemäß besser zur Einbettung von Timing-Daten, da diese einfach als Text dargestellt werden können. Andererseits ist die Unterstützung von Untertiteln in Schnittprogrammen wesentlich schlechter als jene für Audioformate, und es gibt eine Vielzahl konkurrierender Standards.

Nachfolgend wird die Unterstützung von Untertiteln in Adobe Premiere Pro und deren Anwendbarkeit in Research Video evaluiert.

Untertitel fallen in Premiere Pro unter den Begriff "Captioning". Dazu der entsprechende Auszug "Captioning in Premiere Pro" aus der Onlinehilfe (<https://helpx.adobe.com/premiere-pro/using/working-with-captions.html>):

Premiere Pro provides a full set of captioning features that let you edit captions, create your own, and export them for display for all supported formats.

You can import and display closed captions and edit the text, color, background, and the timing quickly and easily. Once you're done editing, you can export the closed captions files as a "sidecar" file, embed them in a QuickTime movie or MXF file, or burn the captions into a video.

Untertitel können nach verschiedenen Kriterien eingeteilt werden. Die wichtigste Unterscheidung ist jene zwischen "open" und "closed" Captions:

- Closed Captions: Können ein- und ausgeblendet werden. Sind zunächst "geschlossen" ("closed") und müssen erst "geöffnet" werden.
- Open Captions: Sind immer sichtbar ("offen"). Da sie Teil des Videos sind, werden sie auch "burnt in" ("eingebrannt") oder "hard" Subtitles genannt.

Für die Anwendung in Research Video sind nur Closed Captions interessant, da Open Captions Teil des Videobildes sind und daher nicht ausgelesen werden können. Closed Captions können weiter nach der Art ihrer Speicherung unterschieden werden:

- Embedded: Einbettung der Untertitel-Daten in der Videodatei selbst, als eigene Untertitelspur.
- Sidecar File: Die Untertitel werden in einem separaten File gespeichert.

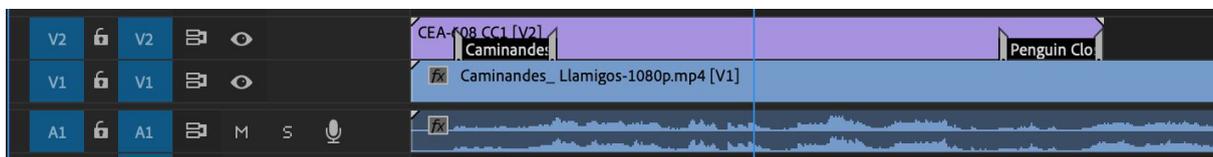
Für Research Video sind grundsätzlich beide Varianten denkbar, die Möglichkeit als eigene Datei (Sidecar File) ist unabhängig von Videocontainerformat und deshalb wahrscheinlich leichter zu verarbeiten.

Ein erster Test in Premiere Pro hat zum Ziel herauszufinden, wie sich eine Untertitelspur beim Schneiden verhält. Beim Anlegen eines neuen Untertitel-Elements (mit File / New... / Captions...) muss ein "Standard" ausgewählt werden:

- CEA-608 (Standardeinstellung)
- CEA-708
- Teletext
- Open Subtitling
- Australian
- Open Captions

Die Recherche ergibt, dass sich die Optionen hauptsächlich durch die Gestaltungsmöglichkeiten des Erscheinungsbildes (Schrift, Farbe, Größe usw.) unterscheiden. Es ist daher, abgesehen von der ungeeigneten Option "Open Captions", für uns egal welche Option gewählt wird und wir bleiben bei der Standardeinstellung CEA-608.

Die Untertitel scheinen nicht automatisch zusammen mit dem Video in der Zeitleiste auf. Sie sind ein eigenes Element im Media-Panel, zusammen mit den verwendeten Videoclips und müssen manuell in die Zeitleiste gezogen werden, wo sie als zusätzliche Videospur aufscheinen (V2 im folgenden Screenshot):



Es ist zunächst wichtig darauf zu achten, dass die Untertitel synchron mit der Videospur beginnen, da es andernfalls zu einer Verschiebung der Annotationen kommen würde.

Im weiteren Schnittprozess muss die Untertitelspur mit der Videospur "mitgeschnitten" werden d.h. jeder Schnitt, Verschiebung, Duplizierung etc. auf der Videospur, muss genauso auch auf der Untertitelspur gemacht werden. Zu diesem Zweck können die beiden Spuren verbunden ("linked") werden, der Prozess wird dann automatisch. Die Verbindung der Spuren muss aber jedes Mal manuell gemacht werden.

Premiere Pro unterstützt eine Reihe von Untertitel-Formaten:

- Embedded: MOV, DNxHD MXF Op1a, MXF Op1a
- Sidecar Files: SRT, SCC, MCC, XML (W3C TTML, SMPTE-TT, EBU-TT), STL

Wir testen zunächst einige Formate beim Import:

- SubRip (.srt): Ein sehr einfaches und daher leicht zu verarbeitendes Format. Export aus Research Video wird bereits unterstützt. Wird

standardmäßig als 'Open Captions' importiert und kann folglich nur 'burnt-in' exportiert werden, was für unsere Zwecke ungeeignet ist.

- EBU N19 File (.stl): Wird standardmäßig als 'Teletext' importiert. Kann anschließend als Sidecar File exportiert werden.
- Die Import Settings für .srt and .stl können geändert werden, allerdings nicht während des Imports, sondern nur davor, in einem relativ versteckten Einstellungsdialog.
- Die anderen Formaten (.scc, .mcc, .xml) scheinen die Art des Imports selbst festzulegen. Es gibt keine Standardeinstellungen für sie.

Anschließend Testen wir den Export von Untertiteln. Die Schritte sind wie folgt:

- Sequenz auswählen
- File / Export / Media... (Cmd-M)
- Export Options
 - None
 - Create Sidecar File -> auswählen
 - Burn Captions Into Video
- File Formats
 - Scenarist Closed Caption File (.scc)
 - MacCaption VANC File (.mcc)
 - W3C/SMPTE/EBU Timed Text File (.xml)
 - SubRip Subtitle Format (.srt)

Der Test zeigt, dass die Timecodes in den Untertitel beim Schneiden korrekt übertragen werden, selbst wenn Lücken im Untertitel-Track sind. Das Format der Wahl ist W3C/SMPTE/EBU Timed Text File (.xml). Es ist ein W3C-Standard (<https://www.w3.org/TR/2018/REC-ttml1-20181108>) und daher relativ zukunftssicher. Es wird in Premiere Pro korrekt importiert (Drag & Drop), ohne dass zuvor Standardeinstellungen geändert werden müssen. Die Zeitstempel können auch in Millisekunden angegeben, womit die Angabe unabhängig von der Framerate des Videos wird.

Fazit

Möglicher Workflow in Research Video

- Video und Untertitel aus Research Video exportieren
- Video und Untertitel im Schnittprogramm importieren
- Untertitel auf die Zeitleiste hinzufügen und an der Videospur ausrichten
- Untertitel- mit Videospur verbinden
- Beim Schneiden darauf achten, dass Untertitel immer zusammen mit Video bearbeitet werden
- Export der Untertitel mit den korrekten Einstellungen

Vorteile

- Nutzung eines standardbasierten XML Formats

Nachteile

- Sehr spezieller Workflow
- Hoher Aufwand für den Benutzer

Ergebnis

Im Folgenden fassen wir die Ergebnisse der verschiedenen getesteten Methoden zusammen. Dabei bewerten wir jeweils die Erfolgsaussichten und den voraussichtlichen Aufwand zur Umsetzung bzw. für den Endanwender.

Erfolgsaussichten

- Wie gut wird das Ziel "Schneiden von Annotationen" voraussichtlich erfüllt?

Umsetzungsaufwand

- Ist die Methode realistisch umsetzbar?
- Wie viel Aufwand ist die Umsetzung der Methode?

Anwenderaufwand

- Wie praktikabel ist die Methode für den Anwender von Research Video?

Methoden	Erfolgsaussichten	Umsetzungsaufwand	Anwenderaufwand
1. Trimfunktion	nicht geeignet	sehr hoch	mittel
2. Timecode	nicht geeignet	nicht geeignet	nicht geeignet
3. EDL	gut	mittel bis hoch	mittel
4. Audiospur	sehr gut	hoch	niedrig
5. Untertitelspur	gut	mittel bis hoch	hoch

Fazit und Empfehlung

Die ersten beiden untersuchten Methoden (Trimfunktion, Timecode) eignen sich nach jetzigem Stand nicht für den vorliegenden Anwendungsfall und fallen damit aus der weiteren Betrachtung.

Die Methode Untertitelspur kann zwar theoretisch das vorliegende Problem lösen, allerdings erfordert sie hohe Disziplin vom Endanwender und ist daher praktisch eher ungeeignet.

Die Methode Audiospur wäre aus Anwendersicht ideal, da kaum Zusatzaufwand besteht. Audiotracks sind integraler Teil jedes Schnittprogramms und damit universell unterstützt, sowohl beim Schneiden als auch beim Export. Großer Nachteil ist der hohe Umsetzungsaufwand.

Die beste Bilanz zeigt die Methode EDL. Sie leistet von ihrer Intention her am ehesten das Nötige, nämlich eine Abbildung des Schnittvorgangs in strukturierter Form und kann mit vertretbarem Aufwand implementiert werden. Unsere Empfehlung ist daher, diesen Ansatz weiter zu verfolgen, und zunächst die Unterstützung in anderen Schnittprogrammen zu klären, besonders für das besser geeignete Final Cut Pro XML Format.