

KI in der Filmtone-Postproduktion – Anwendungen aus der Praxis

Eine Veranstaltung von Filmtone Austria und der Filmakademie Wien

Wann: 2. Okt. 2025, 19 bis 21 Uhr

Wo: Arthouse Kino, Future Art Lab, MDW, Anton-von-Webern-Platz 1, 1030 Wien

Teilnehmer:innen: 22 Personen, Mitglieder von FTA, Studierende der MDW

Künstliche Intelligenz ist in alle unsere Lebensbereiche vorgedrungen. Gerade in der Kreativ-Branche wird sie mit unterschiedlichen Gefühlen wahrgenommen - von Euphorie über Skepsis bis hin zu Angst um die Existenz und das eigene künstlerische Schaffen. Was auf jeden Fall hilft, ist darüber zu sprechen, diese neuen Tools sich genau anzusehen und wenn möglich zu überlegen, wie man KI für sich und die eigene Arbeit nützen kann.



In der Filmtone-Postproduktion existieren mittlerweile zahlreiche Tools, die uns die Arbeit erleichtern. Gerade im Bereich der Restauration technisch schwieriger Aufnahmen und auch in der Veränderung und Kreation von Stimmen sind mittlerweile viele Produkte am Markt. Die Frage nach der Verantwortung, die man bei Benutzung solcher Tools trägt – sich selbst, seiner Berufsgruppe, anderen Berufsgruppen und der gesamten Kreativ-Branche gegenüber schwingt immer mit und muss immer wieder neu gestellt werden.

Filmtone Austria, in Kooperation mit der Filmakademie Wien, hat erfahrene Mischtonmeister eingeladen, über KI-Tools zu sprechen, mit denen sie aktuell ihre Arbeit bestreiten.

Manuel Grandpierre und **Manuel Meichsner** sind unserer Einladung gefolgt, Alexander Koller musste leider kurzfristig wegen Erkrankung absagen.

Manuel Meichsner stellte folgende Programme vor.

Hush Pro (<https://hushaudioapp.com/products/hush-pro>)

Clear (<https://www.supertone.ai/en/clear>)

dxRevive (<https://www.accentize.com/dxrevive/>)

Wo früher Waves, Clarity oder der Dialog Isolator von RX eingesetzt wurde, verwendet Manuel derzeit Hush Pro und Clear. Beide Programme arbeiten ähnlich, nicht generativ, im Bereich Noise Reduction und De-Reverbing.

Hush Pro ist ein AAX AudioSuite-Plugin, konzipiert für Pro Tools auf Apple Silicon Macs. Es nutzt maschinelles Lernen, um gesprochene Sprache von Raumreflexionen, Hintergrundgeräuschen und Transienten zu trennen.

- Mit „Hush Mix“ lassen sich Sprache, Hall und Rauschen über separate Fader in Echtzeit neu im Mix justieren.
- Der „Hush Split“ Modus rendert diese drei Komponenten als getrennte Clips, für weiterführende nicht-destruktive Bearbeitung.
- Die Verarbeitung erfolgt offline mit „random access“, also wird das gesamte Audiomaterial analysiert, nicht im Live-Insertbetrieb.

Clear funktioniert sehr ähnlich:

- Es dient zur Rauschunterdrückung („De-Noise“) & Nachhallentfernung („De-Reverb“) plus optionaler Trennung der Stimme vom Hall im Signal.
- Die Audioverarbeitung erfolgt mithilfe eines neuronalen Netzes, das Audiosignale in drei Kanäle aufteilt: *Voice*, *Ambience* und *Voice Reverb*, die über drei Regler zu steuern sind.
- Es läuft in Echtzeit und steht als Plugin in Formaten wie AU, VST, VST3 und AAX zur Verfügung, für macOS und Windows.
- Ein „Eco Mode“ bietet eine geringere CPU-Belastung; nach der Aktivierung ist Offline-Betrieb möglich.

Hush Pro ist vor allem im Bereich De-Reverbing sehr gut und im Unterschied zu Clear werden bei Hush die Atmer und stimmliche Reaktionen besser erhalten.

dxRevive ist ein generativ arbeitendes Plug-in zur Sprachrestauration mit folgenden Kernfunktionen:

- Es entfernt Rauschen, reduziert Hall und soll Artefakte aus Kompressionen beseitigen.
- Es reintegriert fehlende Frequenzen und rekonstruiert Anteile, die durch Bandbegrenzung oder schlechte Aufnahme verloren gingen.
- Es repariert geclippte Audiosignale und ersetzt fehlende Samples bei Aussetzern.
- Es arbeitet in Echtzeit als Insert-Effekt mit algorithmischen Modi (Studio, Retain etc.)
- In der Pro-Version kommen zusätzliche Algorithmen hinzu (z. B. „EQ Restore“, „Studio 3“)

Manuel setzt es in der Sprachbearbeitung nicht allgemein ein, sondern nur für kurze Passagen oder einzelne Sätze. In den unteren Frequenzen ist das oft kritisch, aber in den hohen Frequenzen arbeitet es sehr gut. Im Bereich 2 bis 3 kHz können manchmal Artefakte entstehen. Die drei von Manuel vorgestellten Tools sind auch unter den Kolleg:innen weit verbreitet.

Manuel Grandpierre beschäftigte sich bei einem Filmprojekt, bei dem es aus produktionstechnischen Gründen nötig war, die Stimme einer Protagonistin neu zu generieren mit den Tools von **Eleven Labs**, einer generativen KI. (<https://elevenlabs.io/>)
Zum Beispiel: Text-to-Speech, Voice Changer und Voice Cloning.

Text-to-Speech (TTS) wandelt eingegebenen Text in ausdrucksvolle, emotionale und mehrsprachige Sprachaufnahmen um – mit Steuerung von Sprechweise, Pausen und Intonation.

Der **Voice Changer** (Stimmenveränderung) modifiziert eine existierende Audioaufnahme so, dass sie wie eine andere Stimme klingt, indem Tonhöhe, Timbre oder Klangcharakter angepasst werden (z. B. zur Dramatisierung, Charakterisierung).

Die Tools funktionieren vor allem in der englischen Sprache durchaus beeindruckend. Teilweise aber auch schon mit deutschen Stimmen.

Achtung, diese Stimmen dürfen aus rechtlichen Gründen nicht einfach in filmischen Werken verwendet werden! Hier ist es unbedingt notwendig, sich mit den Geschäftsbedingungen vertraut zu machen!

Voice Cloning erzeugt aus Sprachproben eine KI-Stimme, die neue Texte in der Stimme einer realen Person spricht (als „digitale Kopie“).

ElevenLabs bietet „Instant Voice Cloning“ für wenige Minuten Tonmaterial und „Professional Voice Cloning“ für hochwertige, personalisierte Kopien mit höherer Qualität.

Die Tools werden mittlerweile auch im Bildschnitt eingesetzt. Wo früher noch Untertitel zur Information über notwendige ADR-Aufnahmen standen, werden heute oft schon als Layout Sätze im Schneiderraum gesprochen und mit Hilfe von Voice Cloning der Schauspieler:in oder dem Schauspieler in den Mund gelegt. Die rechtlichen Bedenken werden dabei außer Acht gelassen. Bernd Dormayer erzählt, dass die Schauspieler:innen, die dann zur ADR ins Tonstudio kommen und diese Layouts hören, oft sehr ungehalten über diese Praxis sind, obwohl das Tonstudio diese gar nicht hergestellt hat. Bernd regt eine Veranstaltung gemeinsam mit den Kolleg:innen vom Schnitt an, um auf diese Problematik hinzuweisen.

Bernd Dormayer stellt noch **Orion SFX** (<https://www.orion-sfx.com/>) vor.

Orion SFX ist ein Tool zur intelligenten SFX-Bibliothekssuche mit folgenden Funktionen:

- **Sonic Similarity Search** – Man kann beliebige Sounds anklicken und Orion zeigt andere Sounds mit ähnlichem Ton, Timbre, Hüllkurve, Tonhöhe und Rhythmus.
- **Golden Nuggets** – Die Software zerlegt Audiodateien in Abschnitte („Chunks“), analysiert diese separat und hilft so, passende Teilstücke aus komplexen Aufnahmen zu finden.
- **Search by Meaning** – Man kann Suchbegriffe (z. B. „sad monster“) eingeben, und Orion sucht Geräusche, die semantisch und atmosphärisch dazu passen – ähnlich wie es ein Sounddesigner intuitiv machen würde.
- **Filter by Sonic Characteristic** – Ergebnisse lassen sich weiter eingrenzen und sortieren nach Eigenschaften wie Hoch-/Tiefertonanteil, Verzerrung, Geräuschcharakter etc.
- **Live Input Search** – Man kann direkt ins Programm hineinsprechen oder Audio einspeisen und Orion sucht dann passende Sounds in der Bibliothek.

Der große Nachteil ist die enorme Zeitdauer, bis das Programm die eigene Sound Library „gescannt“ hat. Bei großen Archiven stößt es hier an seine Grenzen.

Weitere Tools, die von den Anwesenden kurz diskutiert wurden:

Chameleon von Accentize (<https://www.accentize.com/chameleon/>)

Chameleon ist ein intelligentes Reverb-Plugin, das die Nachhallcharakteristik einer Referenzaufnahme analysiert und modelliert.

In Sekundenschnelle erzeugt es ein Impuls-Reverb-Profil (mit Einstellungen wie Dry/Wet, Pre-Delay, Stereo-Breite), das auf trockene Signale angewendet werden kann. So lässt sich ADR oder Foley akustisch an den Raumton von Originalaufnahmen sehr gut angleichen.

Eine Surround-Version (Chameleon Surround) erweitert das Prinzip auf mehrkanalige Wiedergabe (z. B. 5.1, 7.1, Atmos) durch 3D-Raummodellierung.

Chameleon ist derzeit das beste Produkt für die Angleichung von ADR und Foley an den Originalton.

Supertone Air (<https://www.supertone.ai/en/air>)

Supertone Air ist ein Plug-in zur Reverb- und EQ-Anpassung (Dialogue Match). So wie Chameleon analysiert es eine bestehende Dialogaufnahme und extrahiert deren Raumcharakter (Nachhall & Frequenzgang), um diese auf trockene ADR-Aufnahmen zu übertragen.

Mit einer Reverb-Steuerung plus spezieller Parameter kann das „Air Profile“ feinjustiert werden.

Air ist optimiert für Pro Tools AudioSuite sowie DAWs über VST3/AU/AAX und unterstützt Offline-Workflows und Profil-Export/Import.

In einem anschließenden Gespräch wurde noch einmal darauf hingewiesen, dass die Verwendung von KI natürlich rechtlich Fragen aufwirft. Auch die Achtung vor dem eigenen Schaffen sowie der Kreativität anderer Berufsgruppen muss bewahrt bleiben. Eine Verwendung von KI muss mit Verstand und Überlegung passieren.

Abschließend formulierte Manuel Grandpierre den eindrücklichen Appell an alle Kolleg:innen, dass das Verständnis unseres eigenen künstlerischen Schaffens bestimmen wird, ob unsere Arbeit einmal durch eine KI ersetzt werden wird oder nicht. Wir müssen unsere eigene Arbeit als Quelle von Kreativität verstehen, die immer Neues hervorbringt. Wenn wir selber nur noch den einfachen Weg der Reproduktion anstelle der Neuerschaffung gehen, dann wird unser Tun ersetzbar. Eine KI kann nur bereits Hervorgebrachtes reproduzieren, bestenfalls kombinieren, nur der Mensch kann wirklich Neues erschaffen.

Das sollte das Ziel unserer Arbeit sein.

Herzlichen Dank an die beiden Referenten und an alle Kolleg:innen, die sich in das Gespräch eingebracht haben!