

Technische Hochschule Ostwestfalen-Lippe

Studienprogramm Medienproduktion B.A.
Schwerpunkt Audiovisuelle Medien



Bachelorarbeit zum Thema:

Künstliche Intelligenz im Filmtton

Intelligente Mikrofone und deren Anwendungsbereich

Zur Erlangung des Grades: Bachelor of Arts

Erstprüfer: Prof. Dr. Phil. Frank Lechtenberg

Zweitprüfer: M.A. Eike Hollermann

Lizenz: CC BY-SA (4.0)

Vorwort

Audio fasziniert mich seit meiner Kindheit. Als kleiner Junge quetschte ich die ersten schrägen Töne aus der Trompete, spielte später im Schulorchester und erreichte schließlich die erste Trompete der Bigband. Um neue Facetten kennenzulernen, erlernte ich die Bassgitarre als zweites Instrument und experimentierte als DJ und Freelancer in der Veranstaltungsbranche mit synthetischen und digitalen Sounds. Diese vielseitigen Erfahrungen führten mich schließlich zur Faszination für den Filmtone.

Zielgruppe dieser Arbeit sind Personen, welche in der professionellen Filmbranche im Bereich Audio arbeiten. Zugehörig sind hier beispielsweise Tontechniker*innen und Tonmeister*innen, als auch Student*innen, die sich in einer Ausbildung für dieses Arbeitsfeld befinden. Diese Arbeit soll die Leser*innen dazu motivieren, sich mit künstlicher Intelligenz auseinander zu setzen und dessen Potenziale zu erkennen. Die rasante Entwicklung kann herausfordernd sein, dementsprechend soll ebenfalls aus dem lesen dieser Arbeit ein Ausblick in die Zukunft gewagt werden und die Einschüchterung von künstlicher Intelligenz innerhalb der Tonbranche genommen werden.

Abschließend möchte ich mich bei allen Personen bedanken, die mich in der Ausarbeitung meiner Bachelorarbeit unterstützt haben. Mein besonderer Dank gilt Professor Dr. Frank Lechtenberg, dessen Expertise und Inspiration meine Arbeit bereichert hat. Des Weiteren hätte ich ohne die Hilfe von Aimee Tiamson, die als Sängerin die Testaufnahmen durchgeführt hat , nicht zu den Aussagekräftigen Ergebnissen kommen können. Darüberhinaus hat Adele Koopmann mich in der Findung der Interviewpartner*innen maßgeblich unterstützt und mir somit tiefe Einblicke in die aktuelle Arbeitsweise von Tontechniker*innen und -meister*innen ermöglicht.

Mein großer Dank gilt Daniela Cordes, Urs Krüger und Ole Schmidt, die durch ihre Expertise maßgeblich dazu beigetragen haben, diese Arbeit qualitativ hochwertig zu gestalten.

Abschließend danke ich meiner Familie und vor allem meiner Freundin, die während dieser Zeit an meiner Seite standen und stetig ein offenes Ohr hatten.

Vincent Schlagkamp

Detmold, 18.06.2025

Inhaltsverzeichnis

Tabellenverzeichnis	I
Abbildungsverzeichnis	II
Abkürzungsverzeichnis.....	III
1. Einleitung.....	1
1.1 Fragestellung	1
1.2 Methodik und Material.....	1
2. Historie der Tontechnik.....	3
2.1 Historische Meilensteine der Mikrofonentwicklung.....	3
2.2 Traditionelle Mikrofone	4
2.3 Moderne Mikrofontechnologien	4
2.4 Mikrofone in audiovisuellen Medien	6
2.5 Zukunftsperspektiven der Mikrofontechnologie.....	6
3. Intelligente Mikrofonierung und Künstliche Intelligenz	8
3.1 KI in der Audiotechnologie	8
3.2 Intelligente Mikrofone	9
3.2.1 KI gestützte Sprachverarbeitung	9
3.2.2 Flexible piezoelektronische Akustiksensoren (NPAS).....	10
3.4 Anwendungsbeispiele von intelligenten Mikrofonen in der Filmproduktion	10
3.4.1 Automatische Sprechertrennung.....	10
3.4.2 Geräuschreduktion und Nachbearbeitung.....	11
3.4.3 3D-Klanglokalisierung für immersiven Filmtone	11
3.4.4 Ausblick durch intelligente Mikrofone.....	12
3.4.5 Anwendung von KI im Bereich Audio	12
4. Experteninterviews	13

4.1 Durchführung der Experteninterviews	13
4.2 Vorstellung der Interviewpartner*innen	13
4.2.1 Kurzprofil der Expertin Daniela Cordes.....	13
4.2.2 Kurzprofil des Experten Urs Krüger	13
4.2.3 Kurzprofil des Experten Ole Schmidt	14
4.3 Leitfragen zur intelligenten Mikrofonierung	14
4.3.1 Aktuelle Nutzung von KI im Arbeitskontext.....	14
4.3.2 Erwartungen an KI-gestützte Mikrofontechnologien	15
4.3.3 Herausforderungen & Bedenken	18
4.4 Perspektiven auf KI als Assistenzsystem – Chancen und Grenzen.....	19
4.5 Rückschlüsse für die technische Entwicklung und zukünftige Anwendungen	20
4.7 Ergebnisauswertung der Experteninterviews	22
4.7.1 Aktueller Stand der KI-Nutzung im Filmtton.....	22
4.7.2 Erwartungshaltungen an intelligente Mikrofontechnik	22
4.7.3 Herausforderungen und Grenzen.....	23
4.7.4 Ausblick	24
4.6 Zukunftsperspektiven und Forschungspotentiale.....	24
8. Technischer Exkurs Lewitt RAY	26
8.1 Mikrofon Lewitt RAY	26
8.2 Zielsetzung	29
8.3 Methodik	29
8.4 Technische Auswertung.....	32
8.4 Auswertung Bewertungsbögen.....	34
9. Fazit und Ausblick	36
Definitionen	IV
Literaturverzeichnis	VII

AnhangXI

Eidesstattliche Erklärung.....XII

Tabellenverzeichnis

Tabelle 1. - Vergleichende Übersicht der Expert*inneneinschätzungen zur KI im Filmtone.....	23
--	----

Abbildungsverzeichnis

Abbildung 1. - Techgraph Lewitt RAY 5cm Entfernung.....	23
Abbildung 2. - Techgraph Lewitt RAY 30cm Entfernung.....	24
Abbildung 3. - Techgraph Lewitt RAY 50cm Entfernung.....	24
Abbildung 4. - Techgraph Lewitt RAY 100cm Entfernung.....	25
Abbildung 5. - Richtcharakteristik Lewitt RAY.....	26
Abbildung 6. - Klangquellenpositionierung & Mikrofonposition (10cm).....	28
Abbildung 7. - Klangquellenpositionierung & Mikrofonposition (30cm).....	28
Abbildung 8. - Klangquellenpositionierung & Mikrofonposition (50cm).....	28
Abbildung 9. - Klangquellenpositionierung & Mikrofonposition (100cm).....	28
Abbildung 11. - AURA Technology OFF - Bewegung nah/fern/nah - Lewitt RAY.....	29
Abbildung 12. - AURA Technology ON - Bewegung nah/fern/nah - Lewitt RAY.....	29

Abkürzungsverzeichnis

KI.....	Künstliche Intelligenz
CNNs.....	Convolutional Neural Networks
MFCCs.....	mel-frequency cepstral coefficients
t-SNE.....	t-Distributed Stochastic Neighbor Embedding
f-PAS.....	piezoelektronische Akustiksensoren
MEMS.....	Micro-Electro-Mechanical-Systems
LUFs.....	Loudness Units relative to Full Scale
EB.....	elektronische Berichterstattung
dB.....	Dezibel
DR.....	Dynamic Range

1. Einleitung

Die vorliegende Arbeit untersucht zunächst die historische Entwicklung und die heutigen Einsatzgebiete von Mikrofonen. Anschließend wird der aktuelle Stand der Mikrofontechnik vorgestellt und ein Ausblick auf künftige, KI-gestützte Technologien gegeben.

1.1 Fragestellung

In der Filmtone Szene bzw. der Tontechnik wurde stetig daran gearbeitet, die Klangqualität und die Präzision von Mikrofonen zu verbessern. Künstliche Intelligenz birgt auch hier ein großes Potenzial und wird in den kommenden Jahren vermutlich großen Einfluss auf die Entwicklung haben.

Schon jetzt gibt es Mikrofone, welche mit Sensoren versuchen, die Klangqualität zu verbessern. Allerdings finden diese Mikrofone hauptsächlich ihre Anwendungsbereiche in Podcast- oder Gesangsmikrofonen. Wäre diese Technologie auch für die Filmproduktion eine Bereicherung und könnte in Mikrofone etabliert werden? Was müssten intelligente Mikrofone an einem Filmset an technischen Aspekten haben?

1.2 Methodik und Material

Um diese Fragen zu klären, wird zunächst relevante Literatur zu den Themen Geschichte der Tontechnik und Intelligenter Mikrofonierung analysiert und auf den aktuellen Forschungsstand aufmerksam gemacht. Alle unklaren Begriffe in Form von Definitionen geklärt, sodass jeder Leser denselben Wissensstand beim Lesen hat. Basierend auf der Ausarbeitung wird der aktuelle Stand der Mikrofontechnologie dargestellt und eine erste Tendenz der zukünftigen Entwicklungen herausgearbeitet.

Nach der Darstellung historischer Meilensteine erfolgt im zweiten Schritt die empirische Analyse mittels Experteninterviews, in denen mit den Partner*innen darüber diskutiert wird, welche Berührungspunkte sie mit intelligenten Mikrofonen und künstlicher Intelligenz in Bezug auf Audio haben. Die Personen, welche interviewt werden, sind alle in der Audiobranche tätig - ob Tonmeister*innen für Filmproduktionen, elektronische Berichterstattung oder Live-TV-Produktionen. Diese haben dementsprechend alle eine zusammenhängende Branchenherkunft. Ebenfalls wird darüber gesprochen, welche Chancen und Risiken mit der Nutzung von künstlicher Intelligenz bei Audioaufnahmen an Filmsets einhergehen könnten. Anschließend werden diese Interviews transkribiert, analysiert und

ausgewertet, um eine professionelle und vielfältige Meinungsvertretung zu generieren und sich die Forschungslücke zu füllen, da es bis zum jetzigen Zeitpunkt keinerlei Ausarbeitungen in diesem Bereich gibt.

Anhand eines intelligenten Mikrofons (Lewitt RAY), welches mit Hilfe von Sensoren und der Firmeneigenen AURA Technologie arbeitet, werden außerdem Testaufnahmen durchgeführt, um weitere Verbesserungen im Bereich Filmtone zu erkennen. Da Lewitt RAY hauptsächlich als Podcast bzw. Gesangsmikrofon fungiert, werden hierbei nicht direkte Auswirkungen auf die Arbeit im Filmtone untersucht, sondern Hypothesen aufgestellt, falls Filmtone-Mikrofone mit ähnlicher Technologie ausgestattet werden.

Da Literatur zur Historie von Audiotechnik vorhanden ist, allerdings kaum Informationen hinsichtlich intelligenter Mikrofonierung bzw. künstlicher Intelligenz vorhanden sind, wird im Kontext dieser Arbeit versucht, diese Forschungslücke zu erschließen. Durch die Meinungen der Expert*innen wird dementsprechend durch ihre Professionalität und die unterschiedlichen Tätigkeiten im Audibereich ein breitgestreutes Meinungsbild und Zukunftsvisionen generiert.

Schlussendlich wird basierend auf den erlangten Erkenntnissen ein Ausblick und ein Fazit der aktuellen Technologien und deren Potenziale aufgestellt. Ziel der Bachelorarbeit ist es, den aktuellen Stand der Mikrofontechnologie aufzugreifen und darüber hinaus einen Ausblick in die zukünftige Arbeit zu wagen. Mit Hilfe der Testaufnahmen und der Interviews wird das Potenzial von künstlicher Intelligenz und intelligenten Mikrofonen möglichst klar definiert und ein hypothetischer Entwicklungsverlauf definiert.

2. Historie der Tontechnik

2.1 Historische Meilensteine der Mikrofonentwicklung

Die Entwicklung von Mikrofonen hat im späten 19. Jahrhundert einen Sprung von einfachen Modellen hin zu komplexen Systemen gemacht. Ohne die kleinen Anfänge wären Audioaufnahmen und damit einhergehend unser Hörverstehen, wie wir es heutzutage kennen, nicht möglich (Knobloch et. Al., 2020, S. 13-15).

Mit der Erfindung des Kohlemikrofons, welches vor allem für die Telefonie eingesetzt wurde, durch Emile Berliner und Alexander Graham Bell, wurde während des 19. Jahrhunderts eine Mikrofonentwicklung eingeleitet. Die Klangqualität dieser Mikrofone war alles andere als hochwertig, dennoch wurde diese Erfindung ein prägendes Ereignis für die drahtlose Kommunikation (Knobloch et al., 2020, S. 13-15). Dies waren die ersten Schritte, die sich in Zukunft zu einer Qualitätssteigerung der Schallaufnahme führen sollten (Koyama et al., 2014). Die Erfindung des Kohlemikrofons markierte den Beginn für neue Forschungen und Innovationen im Audiobereich.

E.C. Wentes Kondensatormikrofon (1916) etablierte sich schnell durch seine Klangpräzision im Studiobereich. Mit der Einführung wurde ein maßgeblicher Fortschritt in der Mikrofontechnologie erreicht (Manolas & Pauletto, 2009). Darauf folgend wurden im 20. Jahrhundert dynamische Mikrofone populär und stellten ihre Robustheit und Verlässlichkeit für Live-Auftritte und Rundfunkanwendungen unter Beweis. Durch ihre robuste Bauweise waren dynamische Mikrofone den Kondensatormikrofonen in Bezug auf Umwelteinflüsse überlegen, allerdings waren deren Empfindlichkeit und die Klangqualität deutlich besser. Durch die Erfindung dieser beiden Mikrofontypen wurde die Möglichkeit geschaffen, je nach Anwendungsbereich ein passendes Mikrofon zu nutzen und professionellere Aufnahmen in Tonstudios aufzunehmen (Knobloch et al., 2020, S. 13-15).

Um Mikrofonensignale zu verstärken, wurde im 20. Jahrhundert ebenfalls die Röhrentechnologie entwickelt. Durch diese Technik wurde es möglich, schwache oder minimale Signale zu verstärken. Das Grundrauschen wurde dabei kaum erhöht, was diese Technologie zu einem Meilenstein der Signalverarbeitung machte. Durch ihren Einsatz wurden Tonstudios hochwertiger und machten diese somit zu einem populären Ort für Audioaufnahmen (Huscher, R., 2017, S. 1-3).

In der gleichen Zeit trieben die Entwicklungen in der Elektronik und der Miniaturisierung die Verfügbarkeit tragbarer und drahtloser Mikrofone voran. Mit der Erfindung von Transistoren revolutionierten die Nutzungsmöglichkeiten von Mikrofonen. Durch Transistoren wurden Mikrofone leichter und vielfältiger einsetzbar, denn im Vergleich zu Röhren waren diese effizienter und kompakter. Zusätzlich wurde die Qualität der Aufnahmen erhöht und die Mikrofone erhielten durch die Transistoren eine erhöhte Haltbarkeit und waren somit robuster. Miniaturisierte Mikrofone ermöglichten ab diesem Zeitpunkt demnach viele neue Arten der Nutzung in portablen Audiorekordern und für drahtlose Kommunikationssysteme (Kallmeyer, W., 2000, S. 16-18).

Miniaturisierte Mikrofone und die damit einhergehende Flexibilität und Mobilität brachten komplett neue Möglichkeiten mit sich und führten zu einer neuen Art von Mikrofonen und deren Nutzungsmöglichkeiten, wie zum Beispiel in tragbaren Recordern und in drahtlosen Kommunikationssystemen (Hillman & Pauletto, 2014).

Auf dieser Grundlage entstanden neue technische Konzepte, welche im folgenden Abschnitt erläutert werden.

2.2 Traditionelle Mikrofone

Traditionelle Mikrofone sind mit analoger Technik ausgestattet und verarbeiten demnach Schwingungen von Schallwellen in elektrische Signale. Hierbei ist wichtig zu unterscheiden, dass die Klangqualität zusätzlich von der Bauweise, den Komponenten wie zum Beispiel der Membran und den räumlichen Gegebenheiten abhängt. Die Grenzen dieser Mikrofone liegen oft bei geräuschintensiven Situationen oder starken Störgeräuschen, da sie diese nur bedingt filtern (Koyama et al., 2012, S. 1841).

2.3 Moderne Mikrofontechnologien

Die Entwicklung digitaler Mikrofone war ein bedeutender Schritt. Diese Mikrofone wandeln Schallwellen in digitale Signale um und ermöglichen somit eine rauschreduzierte und präzisere Aufnahme. Folglich konnten Audioaufnahmen digital bearbeitet werden, wodurch neue Anwendungsbereiche von Mikrofonen geschaffen wurden (Kallmeyer, W., 2000, S. 16-18). Durch diese Revolution war die Signalverarbeitung durch Algorithmen erstmals umsetzbar.

Heutzutage arbeiten Mikrofone wie zum Beispiel das Shure MXA920 System mit integrierten Microflex Advance Mikrofonarrays mit digitaler Signalverarbeitung, um diese Rauschreduzierung in unkontrollierten Umgebungen zu garantieren. Beim sogenannten Beamforming werden unterschiedliche Mikrofonkapseln angesteuert, sodass eine Richtcharakteristik gezielt für die sprechende Person erzeugt wird. Dadurch werden anderweitige Störgeräusche eliminiert. Ergänzend hierzu unterstützt die IntelliMix DSP von Shure die Echounterdrückung, den Automix, die Rauschreduzierung und die Verstärkungsregelung, um noch präziseren Klang zu gewährleisten (Website Shure). Ein gezielter Anwendungsfall für diese Art von Mikrofon-Systemen wären Konferenzräume.

Eine ähnliche Verarbeitungsweise ist die Wellenfeldsynthese, bei der es sich um ein Verfahren handelt, welches dazu führt, dass eine möglichst genaue Lokalisierung von Klangquellen bei Filmproduktionen sichergestellt wird. Weitere Anwendungsmöglichkeiten finden sich in immersiven Umgebungen wie Virtual Reality (VR). Diese Methode, Tonaufnahmen durchzuführen, trägt dazu bei, das Klangerlebnis möglichst real darzustellen (Tsutsumi & Takada, 2017).

Eine weitere zentrale Entwicklung sind Mikrofonarrays, welche sich durch bestimmte Mikrofonkombinationen und deren Aufnahmen auszeichnen. Zur Verbesserung von Sprachsystemen, welche beispielsweise Anwendung in der Telekommunikation, Sprachassistenten und Konferenzsystemen finden, haben sich Mikrofonarrays durchgesetzt und sind mittlerweile weitverbreitet (Zhang et al., 2023).

Weiterverbreitet sind heutzutage MEMS-Mikrofone (Micro-Electro-Mechanical-Systems), welche unter anderem in Smartphones und Laptops verbaut werden. Diese haben für ihre kleine Bauweise einen vielfältigen Einsatzbereich und eine sehr gute Klangqualität. Hinzu kommt, dass die Bauart äußerst kostengünstig und effizient ist, sodass diese robusten Mikrofone durch ihre stetige Weiterentwicklung ebenfalls die Bereiche Spracherkennung, Augmented Reality und Herakles vorantreiben (Füldner, M., 2012, S. 10-12).

Beispielhaft für diese Art von Mikrofonen ist das Zylia ZM-1, welches durch seine 19 omnidirektional verbauten MEMS-Mikrofone 360° Sound-Szenen abdecken kann und somit für Audioaufnahmen für Dolby Atmos genutzt werden kann. (Zylia Website).

2.4 Mikrofone in audiovisuellen Medien

Die Rolle der Mikrofone hat sich besonders in audiovisuellen Medien weiterentwickelt. Die Nutzung von Surround-Sound-Technologien wie Dolby Atmos hat die räumliche Dimension des Klangs in Filmen erheblich verändert und somit das Klangerlebnis auf eine neue Ebene getragen. Diese Systeme ermöglichen eine präzise Platzierung von Klängen in dreidimensionalen Räumen, was es den Zuschauern erlaubt, tiefer mit ihrer akustischen und emotionalen Wahrnehmung in das Geschehen einzutauchen (Manolas & Pauletto, 2009).

In Virtual-Reality-Umgebungen kommen Mikrofone zum Einsatz, um eine immersive Klanglandschaft zu schaffen. Hierbei werden Mikrofone genutzt, um diese mit den visuellen Elementen synchron interagieren zu lassen. Diese Technologien kombinieren visuelle und akustische Eindrücke und verbessern das Gesamterlebnis (Zhang et al., 2023).

2.5 Zukunftsperspektiven der Mikrofontechnologie

Zukünftige Mikrofontechnologien versprechen weitere Entwicklungen, die die Empfindlichkeit, den Frequenzband und die Richtwirkungen von Mikrofonen weiter verbessern sollen. Hierfür wird an neuen Materialien und Designs geforscht. Hinzu kommt, dass künstliche Intelligenz immer stärker für die Signalverarbeitung in Echtzeit und die Optimierung von Klangqualität in Mikrofonen integriert wird (Busch-Lauer, I., 2023, S. 10-12). Durch diese Schritte ermöglichen sich neue Arbeitsweisen für die Mensch-Maschine-Kommunikation (Dickel, S., 2021, S. 4-6).

Die fortlaufende Entwicklung der Mikrofontechnologie verfolgt in der Geschichte, die Klangqualität und Präzision stetig zu verbessern. Auch zukünftig ist dies weiterhin das Ziel in der Weiterentwicklung von Mikrofonen. Überarbeitete oder neue Technologien könnten eine noch engere Verknüpfung von künstlicher Intelligenz und Machine Learning beinhalten. Dabei könnten Klangquellen automatisch analysiert und optimiert werden. Des Weiteren wird der Nutzungsbereich von Mikrofonen immer weiter ausgeweitet, zum Beispiel für spezielle Anwendungen wie hochsensible Messungen oder den Einsatz in medizinischen Geräten (Koyama et al., 2014).

Die Entwicklung der Mikrofontechnologie hat sowohl in technischer Hinsicht als auch im Hinblick auf deren Anwendungen große Fortschritte gemacht. Von den frühen mechanischen Konstruktionen bis zu heutigen digitalen Systemen hat sich die Rolle des Mikrofons von einem einfachen Übertragungsgerät zu einem präzisen Werkzeug in einer Vielzahl von

Bereichen und jeder Menge Möglichkeiten gewandelt. Zukünftige Entwicklungen, insbesondere im Bereich immersiver Technologien und künstlicher Intelligenz, werden die Bedeutung von Mikrofonen in Wissenschaft, Medien und Kommunikation weiter stärken.

Funkstreifen bzw. drahtlose Mikrofontechnologien werden durch weiterentwickelte Übertragungsstandards und die Verbesserung von Akkus oder Batterien beispielsweise die Reichweite und Laufzeit erhöhen und somit ihre Anwendung für mobile Aufnahmen oder Live-Auftritte verbessern (Knobloch et al., 2020, S. 1-3).

In der Mikrofontechnologie ist die Zukunft durch Innovationen in den Bereichen KI, Digitalisierung, Miniaturisierung und drahtlose Übertragung geprägt. Wie in der Geschichte werden diese Arbeitsweisen stetig weiterentwickelt und die Anwendungsbereiche von Mikrofonen weiter ausweiten und bestimmen. Unsere Kommunikation und die Audioqualität werden sich folglich entwickeln und verändern (Busch-Lauer, I., 2023, S. 1-3).

3. Intelligente Mikrofonierung und Künstliche Intelligenz

3.1 KI in der Audiotechnologie

Künstliche Intelligenz (KI) verändert zunehmend unseren Alltag und wirkt sich auf verschiedene Bereiche aus. Dies betrifft auch die Audiotechnologie, in der KI neue Möglichkeiten für die Erstellung, Bearbeitung und Nutzung von Audioinhalten eröffnet (Stubbe et al., 2019). In diesem Kapitel wird eine Einführung in die KI-gestützte Audiotechnologie gegeben, um die Grundlage für eine detaillierte Auseinandersetzung mit diesem Themenfeld zu schaffen. Dabei werden die grundlegenden Funktionsweisen und Anwendungsbereiche von KI in der Audiotechnologie erläutert.

Die Anwendung von KI in der Audiotechnologie bietet zahlreiche Vorteile. KI-gestützte Tools können Aufgaben schneller, effizienter und präziser erledigen als traditionelle Methoden (Steffek, F., 2023). Sie können große Datenmengen analysieren, Muster erkennen und personalisierte Lösungen anbieten.

Doch auch Herausforderungen und Risiken bleiben nicht aus bei der Nutzung von KI. Ein wichtiges Thema ist der Datenschutz, da KI-Systeme oft große Mengen an persönlichen Daten verarbeiten (Steffek, F., 2023). Auch ethische Fragen spielen eine Rolle, insbesondere im Hinblick auf die Autonomie und Verantwortung von KI-Systemen (Schmohl et al., 2023). Es ist wichtig, dass KI-Systeme transparent und nachvollziehbar sind, damit ihre Entscheidungen überprüft und bewertet werden können.

Trotz dieser Herausforderungen sind die Chancen durch die KI in der Audiotechnologie enorm. KI kann dazu beitragen, die Qualität von Audioinhalten zu verbessern, die Effizienz von Produktionsprozessen zu steigern und neue Möglichkeiten für die Nutzung von Audio zu schaffen (Vecera, E., 2020).

KI erweitert die Möglichkeiten von Mikrofonen durch adaptive Klangoptimierung und automatisierte Nachbearbeitung. Die Integration von künstlicher Intelligenz in Mikrofone bietet Potenziale in der Filmvertonung, sowohl in der Tonqualität als auch zur erheblichen Vereinfachung von Arbeiten während der Postproduktion. Besonders für die Sprachverarbeitung und das Sound Design finden sich nützliche Tools, die Chancen bieten. Im Vergleich zu anderen Audiotechnologien wird durch KI ein dynamischer und adaptiver Arbeitsweg eingeschlagen (Miranda & Williams, 2015, S. 76). Algorithmen, welche auf KI

basieren, reduzieren Hintergrundgeräusche und verbessern die Sprachverständlichkeit und -qualität enorm (Busch-Lauer, I., 2023, S. 10-12).

Im folgenden Kapitel wird ein kurzer Überblick über verschiedene intelligente Mikrofone und ebenfalls über Anwendungsbeispiele dieser Mikrofone und künstlicher Intelligenz gegeben.

3.2 Intelligente Mikrofone

a) Historische Wandler vs. KI-Systeme

Die Audiotechnologie, wie wir sie heute kennen, wäre ohne Mikrofone kaum denkbar, denn sie sind ein elementarer Bestandteil und haben sich seither von akustischen Wandlern bis hin zu KI-gestützten Systemen weiterentwickelt. Wohingegen ältere Mikrofone Schallwellen in elektrische Signale wandeln, arbeiten intelligente Mikrofone mit künstlicher Intelligenz, um die Sprachverarbeitung und die Klangqualität zu verbessern (Jung et al., 2021, S. 2).

b) Deep Learning- und Sensor-basierte Ansätze

Insbesondere intelligente Mikrofone, welche mit beispielsweise Deep Learning oder piezoelektrischen Sensoren arbeiten, werden immer relevanter und überzeugen mit ihren vielfältigen Einsatzmöglichkeiten, wenn es um Klangverbesserung der Stimmen und Rauschunterdrückung während der Aufnahmen geht (Jung et al., 2021, S. 3).

Im folgenden Abschnitt werde ich verschiedene intelligente Mikrofone vorstellen, um ein Grundverständnis zu entwickeln, was ein intelligentes Mikrofon ausmacht und wie vielfältig diese sind.

3.2.1 KI gestützte Sprachverarbeitung

Eine Kombination aus modernster Technik von Mikrofonen und Deep Learning-Algorithmen bietet das Potenzial, Audiosignale aus lauten oder sogar stark störgeräuschbelasteten Szenarien herauszufiltern. Die technische Grundlage sind Convolutional Neural Networks (CNNs). Hierbei handelt es sich um eine Klasse von tiefen neuronalen Netzwerken, welche speziell für Bild- und Audioverarbeitung optimiert wurden, um hierarchische Strukturen zu analysieren. Diese arbeiten gezielt auf Sprachsignale heraus, analysieren diese und filtern auf deren Basis die Hintergrundgeräusche (Ko et al., 2022, S. 2).

Hinzu kommen Deep U-Net Modelle, die durch eine ausgewählte Verarbeitung mehrkanaliger Audiosignale die Klangqualität und somit die Klarheit des Gesprochenen verbessern. Diese

Modelle haben ihren Namen dank der U-förmigen Struktur und wurden vorerst für die Bildsegmentierung genutzt, bis diese sich dann zur Erkennung und Rekonstruktion von Audiomeerkmalen ausweitete (Jung et al., 2021, S. 4).

3.2.2 Flexible piezoelektronische Akustiksensoren (NPAS)

Einen neuen Ansatz von intelligenten Mikrofonen und dessen Technik verspricht man sich von Mikrofonen, in denen piezoelektronische Akustiksensoren (f-PAS) verbaut sind. Diese Sensoren imitieren das menschliche Gehör und sind ausgestattet mit vielen dünnen und flexiblen Membranen, die darauf ausgelegt sind, Sprache besonders breit mit bis zu 8 kHz abzubilden (Jung et al., 2021, S. 3). Hingegen zu MEMS-Mikrofonen (Micro-Electro-Mechanical-Systems), welche mit einer mechanischen Membran versehen sind und eine integrierte Signalverarbeitung haben, sind piezoelektronische Akustiksensoren deutlich sensitiver und somit besser geeignet für umfangreiche Klangkulissen (Ko et al., 2022, S. 3).

3.4 Anwendungsbeispiele von intelligenten Mikrofonen in der Filmproduktion

Um intelligente Mikrofone zu verstehen und deren Potenzial zu erkennen, werden im folgenden Abschnitt unterschiedliche Technologien vorgestellt und die Stärken in deren Anwendungsbereichen aufgezeigt.

3.4.1 Automatische Sprechererkennung

Durch Multi-Kanal-Arrays und die Unterstützung durch künstliche Intelligenz können intelligente Mikrofone mehrere Sprecher*innen erkennen, analysieren und trennen. Aktuell wird in Forschungen aufgezeigt, dass noise-robuste, flexible piezoelektrische akustische Sensoren (NPAS) in Echtzeit Stimmen von Personen trennen können. Dies ist auch in geräuschintensiven Umgebungen der Fall. Durch eine unabhängige Vektoranalyse (IVA) werden diese Stimmen in ihrer Frequenz untersucht und von anderen Geräuschen abgekapselt. Dieses Verfahren hilft vor allem bei mehreren Sprecher*innen (Jung et al., 2021, S. 16).

Durch die Analyse von mel-frequency cepstral coefficients (MFCCs), welche das aufgenommene Audiosignal auf der Mel-Skala abbilden, sind diese Ergebnisse besonders effizient für Sprachverarbeitung. Hierbei werden akustische Merkmale von den einzelnen Sprecher*innen identifiziert und unterschieden. Zusätzlich wird durch eine vereinfachte Visualisierung durch t-Distributed Stochastic Neighbor Embedding (t-SNE) durchgeführt, bei dem Datenpunkte ausgewertet werden und somit Stimmen klar voneinander getrennt werden

können (Jung et al., 2021, S. 4). Eine automatische Sprechertrennung wäre demnach besonders sinnvoll in Szenarien, in denen mehr als eine Person einen großen Anteil für Dialogaufnahmen hätte und ein möglichst klares Klangbild gewünscht ist.

3.4.2 Geräuschreduktion und Nachbearbeitung

An nahezu jedem Filmset bzw. bei jeder Audioaufnahme kommt es immer wieder zu unbeeinflussbaren Faktoren wie Wind, Verkehr oder unbeteiligten Personen, welche ungewünschte Klangquellen sind. Intelligente Mikrofone, welche mit CNNs ausgestattet sind, können diese Klangquellen erkennen und eliminieren. *„Das Deep U-Net basierte Speech Enhancement (DEEP-SEA)-Modell reduziert Hintergrundgeräusche signifikant und verbessert die Sprachverständlichkeit um bis zu 109%“* (Jung et al., 2021, S. 17).

Stellt man intelligente Mikrofone, ausgestattet mit NPAS gegenüber herkömmlichen MEMS, gegenüber, kommt man zu dem Ergebnis, dass die mit NPAS ausgestatteten Mikrofone eine deutlich bessere Performance als die herkömmlichen Mikrofone aufweisen. Dies würde dazu führen, dass die Audioaufnahmen in Situationen mit vielen Klangquellen und sogar Störgeräuschen deutlich klarer und somit besser für eine Nachbearbeitung während der Postproduktion sind (Jung et al., 2021, S. 13).

3.4.3 3D-Klanglokalisierung für immersiven Filmtton

Der Einsatz von mehrkanaligen Mikrofonarrays mit KI gestützter Schallquellenlokalisierung kann Klangquellen an einer Filmlocation orten und insbesondere für immersive Film- und Soundformate wie beispielsweise Dolby Atmos aufbereiten (Ko et al., 2022, S. 4). Hierfür werden Deep Neural Networks (DNNs) für die Lokalisierung genutzt, um die Position der Schallquellen im dreidimensionalen Raum zu erfassen. Dies geschieht durch eine Analyse von Richtung und Entfernung der Klangquelle und führt dazu, dass diese konkret ermittelt werden kann (Ko et al., 2022, S. 2). In dem Artikel *„Real-Time Sound Source Localization for Low-Power IoT Devices Based on Multi-Stream CNN“* schreiben die Autoren *„Ein Deep Neural Network (DNN) basiertes System analysiert Multi-Kanal-Spektren, um die Richtung und Entfernung der Klangquellen mit hoher Präzision zu bestimmen“* (Ko et al., 2022, S. 7). Folglich hiervon wäre, nach der Verarbeitung der Audiospuren in der Postproduktion, ein intensiveres Filmerlebnis für das Publikum.

3.4.4 Ausblick durch intelligente Mikrofone

Anhand von automatisierter Sprechertrennung, Geräuschreduktion und Nachverarbeitung und 3D-Klanglokalisierung lässt sich in Zusammenhang mit intelligenten Mikrofonen erkennen, dass die Möglichkeiten für Filmproduktionen eine Erleichterung sind und vor allem größere Präzision der Audioaufnahme ermöglichen. Die Autoren des Artikels „*Deep Learning-based Noise Robust Flexible Piezoelectric Acoustic Sensors for Speech Processing*“ schreiben zu diesem Thema „*die zukünftige Entwicklung intelligenter Mikrofone könnte adaptive Algorithmen integrieren, die sich kontinuierlich an verschiedene akustische Umgebungen anpassen und so die Effizienz der Sprach- und Geräuschtrennung weiter steigern*“ (Jung et al., 2021, S. 10). Hierbei ist wahrzunehmen, dass einerseits Nutzer der Technik einen praktischen Vorteil bekommen, als auch Konsumenten der erstellten Aufnahmen eine eindrucksvollere Klangwelt genießen können.

3.4.5 Anwendung von KI im Bereich Audio

In diesem Kapitel geht es um einen kurzen Exkurs in die Anwendungsbeispiele von KI im Bereich Audio anhand des Beispiels „KI in der Musikkomposition und -analyse“.

Bei der computergestützten Musikkomposition ist KI mittlerweile weit verbreitet. Die Anfänge waren hierbei regelbasierte Systeme, welche dahingehend programmiert wurden, Musikstile zu imitieren. Basierend auf diesen Systemen wurden neue KI-gestützte Systeme entwickelt, die ebenso Daten analysieren, allerdings deutlich umfangreicher und anhand dessen neue Kompositionen generieren. Zu unterscheiden ist hier zwischen zwei unterschiedlich arbeitenden Systemen - symbolischen und biologischen KI-Modellen. Während symbolische Systeme sich an Notation und Harmonieregeln orientieren, kreieren biologische Systeme musikalische Muster durch einen Lernprozess (Miranda & Williams, 2015, S. 78).

4. Experteninterviews

4.1 Durchführung der Experteninterviews

Um die technischen Ergebnisse meiner Testaufnahmen durch einen praktischen Branchenbezug zu ergänzen, wurden im Rahmen dieser Arbeit Interviews mit drei Expert*innen aus der professionellen Audioproduktion durchgeführt. Ziel war es, die Perspektive von erfahrenen Anwendern auf die Nutzung von KI im Filmtone und intelligenten Mikrofonen – insbesondere adaptive Systeme wie das Lewitt RAY – einzuholen.

Die Interviewpartner verfügen über langjährige Erfahrung in den Bereichen Tonaufnahme, Live-Sound und elektronischer Berichterstattung und konnten wertvolle Einschätzungen zu Anwendungsszenarien, Qualitätskriterien sowie Chancen und Grenzen intelligenter Mikrofone geben. Die Gespräche wurden halbstrukturiert geführt, wobei der Fokus auf offen formulierten Fragen lag, um individuelle Einschätzungen zu ermöglichen. Die Interviews fanden im Zeitraum vom 24.04.2025 bis 18.06.2025 online über die Plattform Zoom statt und dauerten jeweils etwa 30 bis 45 Minuten. Zugehörige Audiodateien finden sich im Anhang.

Die gewonnenen Aussagen wurden anschließend inhaltlich kategorisiert und in Bezug zur zentralen Forschungsfrage dieser Arbeit gesetzt, um fundierte Rückschlüsse für die Einordnung adaptiver Mikrofontechnologie im praktischen Einsatz zu ziehen.

4.2 Vorstellung der Interviewpartner*innen

4.2.1 Kurzprofil der Expertin Daniela Cordes

Daniela Cordes ist eine erfahrene Toningenieurin aus dem Bereich TV- und Liveproduktion. Ihre Expertise, Meinungen und Antworten werden unter den verschiedenen Aspekten dieser Arbeit ausgewertet und lassen somit Einblicke in die Nutzung von KI im Filmtone, die technischen Herausforderungen und die Zukunftseinschätzung von Cordes zu. Im Gespräch sollte insbesondere geprüft werden, wie relevant KI-Technologien im Live-Tonbereich bereits heute sind und welches Entwicklungspotenzial in der praktischen Anwendung gesehen wird.

4.2.2 Kurzprofil des Experten Urs Krüger

Urs Krüger, Set-Tonmeister mit über zehn Jahren Erfahrung in szenischer Filmproduktion, gab im Kontext eines Experteninterviews Einblicke in seine Nutzung und Erwartungen bezüglich KI-gestützter Mikrofontechnik. Die Auswertung erfolgt entlang zentraler Teilfragen

der Arbeit, die sich auf die aktuellen Anwendungen, Erwartungen und Herausforderungen von KI-gestützter Mikrofontechnik beziehen. Die Aussagen wurden thematisch gebündelt und mit Zeitstempeln aus dem Interview belegt.

4.2.3 Kurzprofil des Experten Ole Schmidt

Das Experteninterview mit Ole Schmidt, einem freiberuflichen Tontechniker mit jahrelanger Erfahrung in der elektronischen Berichterstattung sowie der Postproduktion, orientiert sich ebenfalls an zentralen Leitfragen. Ziel des Interviews war auch hier, qualitative Einschätzungen zur Rolle Künstlicher Intelligenz (KI) und zu intelligenten Mikrofonlösungen im Tonbereich zu erhalten.

4.3 Leitfragen zur intelligenten Mikrofonierung

4.3.1 Aktuelle Nutzung von KI im Arbeitskontext

In den folgenden Kapiteln werden die Ergebnisse und Antworten der unterschiedlichen Interviews zusammengetragen und reflektiert.

Für Daniela Cordes spielt aktuell KI im Tonbereich keine präsenste Rolle ihrer Arbeitsweise, verdeutlicht sie und ergänzt, dass allerdings aus ihrer Sicht Equalizer, Autotune und Kompressoren eine ähnliche Verarbeitungsweise vorweisen.

„Autotunen und alles diese Geschichten sind ja schon sehr nah an der KI dran.“ [00:01:49]

Nach Cordes Aussagen zur Folge sind KI gestützte Tools in der Bildverarbeitung deutlich präsenter und etablierter als in der Tonverarbeitung, geschweige denn Tonaufnahmen. Software wie RX10 von iZotope sei ihr grundsätzlich bekannt, Berührungspunkte hat sie dennoch aktuell keine [00:03:08]. Eine Entwicklung in Richtung Integration von KI in bestehende Tools und Etablierung von KI gestützten Plug-Ins schließt Sie nicht aus, im Gegenteil - laut ihr ist der Prozess in vollem Gange und merkbar.

Urs Krüger nutzt künstliche Intelligenz im Bereich Filmtone derzeit primär in der Postproduktion. Dort wird KI insbesondere zur Störgeräuschreduzierung, Dialogbearbeitung und Sprachisolierung eingesetzt:

„Gerade wenn es um Sprache isolieren geht oder bereinigen von Störgeräuschen, sind auf jeden Fall schon KIs [...] mehrheitlich in Benutzung.“

[00:01:09]

Am Filmset selbst kommt KI derzeit nicht aktiv zum Einsatz, was Krüger auf technologische Hürden und fehlende Geräte zurückführt. Zudem fehle es an nutzerseitiger Erfahrung und Routine im Umgang mit KI-Systemen.

„Aktuell gibt es einfach keine Mikrofone im professionellen Bereich und auch noch zu wenig Wissen von den Nutzern.“

[00:02:54]

Auch intelligente Mikrofone, wie etwa solche mit Abstandsmessung und automatischer Pegelanpassung, sieht Krüger noch nicht im professionellen Setzeinsatz, wenngleich sie im semiprofessionellen Umfeld zunehmend denkbar seien, z. B. bei VJs, Influencern oder kleinen Doku-Teams.

„Gerade bei Dokus oder kleinen Produktionen [...] kann das eine Unterstützung sein.“

[00:09:45]

Schmidt nutzt KI-basierte Werkzeuge im Filmtoneinsatz bislang primär in der Postproduktion. Insbesondere nennt er Tools wie iZotope RX, die für Rauschunterdrückung und Dialogbearbeitung zum Einsatz kommen:

"Ich benutze iZotope RX regelmäßig. [...] Das hilft bei vielen Dingen, wo früher viel mehr Handarbeit nötig war."

[00:04:21]

Am Filmset selbst kommt KI dagegen noch kaum zur Anwendung, was Schmidt sowohl auf fehlende geeignete Mikrofone als auch auf Skepsis im Team zurückführt. Der Fokus liege weiterhin auf klassischer Mikrofontechnik, wobei die Nachbearbeitung durch KI unterstützend agiert.

"Am Set geht's eher klassisch zu. Ich arbeite da lieber mit dem, was ich kontrollieren kann."

[00:06:13]

4.3.2 Erwartungen an KI-gestützte Mikrofontechnologien

Große Potenziale sieht Cordes für KI in der Aufbereitung und Überarbeitung von Aufnahmen, in denen Hintergrundgeräusche oder allgemeine Störgeräusche wie z.B. Schuhquietschen oder einen Kammfilter-Effekt unterdrückt werden müssen.

„Dann muss ich kein Moosgummi mehr unter die Schuhe kleben.“ [00:04:35]

Zukunftsorientiert ergänzt sie, dass ein Mikrofon, welches es zulässt, seine Richtcharakteristik zu wählen und den voreingestellten Parametern als intelligent gelten würde und somit eine enorme Hilfestellung in herausfordernden Situationen wäre.

„Ich stelle das hin und gebe die Parameter ein und sage: Mach mir das mal hier schön das Interview im Schwimmbad.“ [00:10:19]

Darauf aufbauend wäre eine automatische Pegelanpassung in Kombination der einstellbaren Richtcharakteristik fortschrittlich, um Tonaufnahmen während stressiger Produktionen und ebenfalls nicht geschulter Sänger*innen zu kompensieren und somit optimale Aufnahmeergebnisse zu gewährleisten [00:13:39].

Ein intelligentes Mikrofon könnte für Sie ebenfalls ausgestattet sein mit Sensoren für die Entfernungsmessung, sodass die Klangquelle erkannt und die Aufnahme direkt situationsbedingt verarbeitet werden kann [00:13:39].

Krüger sieht Potenzial insbesondere in Mikrofontechnologien, die Sensorik und adaptive Algorithmen nutzen. Er verweist auf Forschungsansätze zu membranlosen, sensorbasierten Mikrofonen, die den gesamten Raum erfassen und nachträglich differenzierbar machen:

„Dass man sagt: Ich lege das Mikrofon in den Raum [...] und kann dann im Nachhinein sagen, ich will nur die Person hören, die da langläuft.“ [00:30:00]

Technischer Exkurs:

In Bezug auf die vorherigen Aussagen von Cordes und Krüger liefert das Lewitt LCT 640 TS spannende Ausblicke - durch die zwei XLR-Ausgänge, einen für das Front-, einen für das Back-Signal der Kapsel, erlaubt es die flexible Mischung verschiedener Polar Patterns und Richtwirkungen nach der Aufnahme mittels der POLARIZER-Software. Dies entspricht exakt der Vorstellung, die die beiden Expert*innen beschrieben haben: eine Möglichkeit, die Aufnahmeperspektive und Raumcharakteristik im Nachgang präzise zu steuern. Auch wenn das LCT 640 TS nicht für den Einsatz an Filmsets geeignet ist, lassen sich so zukunftsorientierte Visionen für Mikrofone für genau diesen Anwendungsfall erahnen.

Als wichtige Entwicklung nennt Krüger zudem Systeme, die automatisch Frequenzen für beispielsweise Funkmikrofone koordinieren, sich selbstständig auf saubere Kanäle umstellen und den Aufnahmezustand überwachen:

„Dass da eine KI im Hintergrund eigenständig arbeitet und immer dafür sorgt, dass man volle Netzabdeckung hat.“

[00:10:58]

Gleichzeitig hebt er hervor, dass eine zunehmende KI-Automatisierung zu einer Verschiebung im Berufsbild führen könnte – hin zu einer Rolle als „Tonregisseur“, der mit der KI kommuniziert und Ergebnisse kuratiert:

„Vielleicht macht es dann auch gar kein Sounddesigner mehr, sondern der Regisseur selber.“

[01:00:58]

Für Schmidt liegt das größte Potenzial intelligenter Mikrofone in der automatischen Ausrichtung und Richtcharakteristik, etwa für dynamischen Szenen oder bei beengten Drehorten. Er betont, dass adaptive Systeme insbesondere bei Drehs mit wenig Personal hilfreich sein könnten:

"Wenn ich nur ein Mikrofon habe, das aber mitdenkt und sich auf den Sprecher richtet - das wäre genial."

[00:08:02]

Zudem sieht er große Chancen in der automatischen Signaltrennung, also der Fähigkeit, Sprache, Geräusche und Raumanteile getrennt zu erfassen:

"Wenn ein Mikro von sich aus entscheidet, dass es den Hintergrund nicht braucht - das wäre sehr nützlich."

[00:09:45]

4.3.3 Herausforderungen & Bedenken

Cordes betont zwei zentrale Herausforderungen:

1. Technisch: Um KI komplett in die Tonverarbeitung voll und ganz zu integrieren, würde eine hohe Rechenleistung gefragt sein, welche möglicherweise nicht in jeder Situation gewährleistet sein kann. Außerdem könnte eine direkte Signalverarbeitung zu Verlust und Verfälschung von guten Tonsignalen führen [00:06:19].

2. Kreativ: Die Intuition von erfahrenen Toningenieur*innen fehlt als wichtiger Aspekt während der Tonaufnahme. Die Entscheidungsgewalt darüber, ob ein Geräusch gewollt oder ungewollt ist, scheint eine KI laut Cordes noch nicht tragen zu können.

„Es gibt ja Geräusche, die gewollt sind [...] und andere hat man sich damit eingefangen.“
[00:06:47]

Zusätzlich hierzu kommt das Problem, dass eine visuell-akustische Absprache herrschen muss. Die Location und das Gespielte bzw. das Geschehen muss logisch abgebildet werden und ist somit nicht mit einer einfachen Signalüberarbeitung getan [00:07:30].

Daniela Cordes ist es zudem wichtig, dass die Arbeit einer KI ausgiebig getestet werden muss, um in spontanen und nicht planbaren Situationen handeln zu können. Vor allem im Live-Betrieb scheint ihr das Risiko zu hoch zu sein, um etwas dem Zufall zu überlassen [00:19:54].

Als zentrale technische Herausforderung nennt Krüger die Integration in bestehende Workflows und zusätzliche Komplexität am Set, z.B. durch hochleistungsfähige Rechentechnik:

„Das ist einfach die Masse an Technik am Set, die sich massiv erhöht.“

[00:20:00]

Im kreativen Bereich sieht er Grenzen der KI insbesondere bei sprachlicher Feinarbeit, etwa wenn eine Schauspielerleistung nicht stimmig wirkt. Hier sei Erfahrung, Intuition und das Gespür für Sprache entscheidend:

„Das sind Sachen, da werden KIs [...] Probleme haben. [...] Weil da gibt's nicht unbedingt richtig oder falsch.“

[00:22:26]

Weiterhin weist Krüger auf urheberrechtliche Unsicherheiten hin. Inhalte, die von KI-Stimmen generiert wurden, könnten ggf. nicht geschützt oder unklar lizenziert sein, was insbesondere für Sender und Produktionsfirmen ein Problem darstellt:

„Wenn jetzt eine KI-Stimme hingehört und einen Satz nachspricht [...] ist der Inhalt [...] nicht urheberrechtlich schützbar.“

[00:18:13]

Schmidt nennt eine weitere Kombination aus Hürden:

1. Technische Unberechenbarkeit: KI-gestützte Systeme können unvorhersehbar reagieren, was in der Tonaufnahme kritisch sei:

"Ich will wissen, was das Mikrofon tut. Wenn es anfängt, für sich zu entscheiden, verliere ich die Kontrolle.“

[00:10:57]

2. Mangelnde Transparenz der Algorithmen: Er bemängelt, dass viele KI-Systeme wie eine "Black Box" funktionieren - die Entscheidungskriterien seien oft nicht nachvollziehbar.

"Es wäre wichtig, dass man nachvollziehen kann, warum ein Mikro etwas macht.“

[00:11:34]

Zudem merkt er an, dass Ton oft nicht getrennt vom Bild gedacht werden kann - eine zu starke Automatisierung könne zu einer Entfremdung zwischen Tonbild und Raumwirkung führen.

4.4 Perspektiven auf KI als Assistenzsystem – Chancen und Grenzen

Das Interview mit Daniela Cordes verdeutlicht eine grundsätzlich offene, aber vorsichtige Haltung gegenüber KI-Technologien. Besonders im Live-Kontext sieht sie praktische Anwendungsmöglichkeiten, wenn KI zur Unterstützung dient – nicht jedoch als Ersatz menschlicher Expertise.

Die Forderung nach intelligenter Richtcharakteristik, automatischer Pegelregelung und kontextbezogener Selektierung des Audiosignals spiegelt eine klare Vision für die Mikrofontechnik der Zukunft. Gleichzeitig bleiben Bedenken hinsichtlich ungewollter Signalveränderung, Klangverfälschung und fehlender Intuition bestehen.

Die Analyse zeigt, dass KI im professionellen Filmbereich derzeit primär als postproduktionsunterstützendes Werkzeug eingesetzt wird. Am Set ist sie noch nicht verankert, was sowohl auf technische als auch auf workflow-bezogene Hindernisse zurückzuführen ist. In der Praxis könnten intelligente Mikrofone dann sinnvoll werden, wenn sie Funktionen wie automatische Richtcharakteristikanpassung, Signalüberwachung und adaptive Pegelregelung übernehmen, ohne den kreativen Prozess zu behindern.

Gleichzeitig entstehen durch KI neue Aufgabenfelder (z. B. Tonregie, Qualitätssicherung synthetisierter Inhalte) und ethische sowie rechtliche Fragestellungen, die derzeit noch unzureichend reguliert sind. Krügers Skepsis hinsichtlich der kreativen Intuition und der zwischenmenschlichen Abstimmung am Set betont die Notwendigkeit eines hybriden Ansatzes, in dem KI als Assistenzsystem, nicht aber als Ersatz fungiert.

Schmidts Aussagen spiegeln eine kritisch-reflektierte, aber interessierte Haltung gegenüber KI im Filmbereich. Während er in der Postproduktion klare Vorteile durch KI-Tools sieht, ist er bei Echtzeit-Anwendungen am Set deutlich vorsichtiger. Sein Fokus liegt auf Kontrolle, Transparenz und Verlässlichkeit, was intelligente Mikrofone erfüllen müssten, um akzeptiert zu werden.

Die zentralen Potenziale liegen in automatischer Quellenerkennung, anpassungsfähiger Mikrofontechnik und Signalauftrennung, doch diese müssten nutzergesteuert und nachvollziehbar bleiben. Gleichzeitig zeigt Schmidt auf, dass in komplexen Aufnahmesituationen (z. B. One-Take-Shots) eine KI-Unterstützung enorme Arbeitserleichterung bringen könnte.

4.5 Rückschlüsse für die technische Entwicklung und zukünftige Anwendungen

Basierend auf den gesammelten Daten lassen sich folgende praxisnahe Rückschlüsse ziehen:

- Sensorik-basierte Mikrofone mit Entfernungserkennung könnten Mikrofonplatzierungen optimieren – insbesondere im Live- oder EB-Bereich. Die Kombination aus Sensorik (Bewegung/Ort) und KI-Logik könnte außerdem die Nachbearbeitung von Mikrofonplatzierungen vereinfachen.
- Eingrenzbare Aufnahmemodi (z. B. durch Voreinstellungen: „Interview“, „Theater“, „Musikaufnahme“) könnten künftig via App gesteuert werden.

- KI-gestützte Mikrofone mit adaptiver Richtwirkung könnten zentrale Tools für kleinere Teams oder hektische Produktionsumgebungen werden. Mikrofone mit Raumdatenanalyse und nachträglicher Quellengewichtung könnten zusätzlich klassische Richtcharakteristiken langfristig ersetzen. Dies erfordert jedoch neue Produktionsstandards und Aufnahmemethoden. Zudem könnte selektive Raumabbildung eine wichtige Funktion werden, um Atmosphäre gezielt zu steuern – ohne zusätzliche Stereo-Setups für eine Atmo.
- In Low-Budget-Produktionen kann KI-gestützte Technik zur Personalkompensation beitragen – mit qualitativen Abstrichen, aber ökonomischem Vorteil.
- KI-Unterstützung in Echtzeit ist primär dort relevant, wo Live-Übertragungen oder komplexe Multikanal-Setups (z. B. Dolby Atmos) zum Einsatz kommen. KI kann hier Laufzeitunterschiede, Übersprechen und Kanalzuweisungen automatisiert steuern.
[01:52:17]

Ein transparenter Modus, der Nutzer*innen über KI-Entscheidungen informiert, wäre für die Akzeptanz all dieser Dienste essenziell. Diese Entwicklungen erfordern allerdings neue Schnittstellen für die manuelle Steuerung und müssen nahtlos in bestehende Workflows integrierbar sein.

4.7 Ergebnisauswertung der Experteninterviews

Dieses Kapitel wertet die Interviewergebnisse vergleichend aus und gibt eine Zusammenfassung der Ergebnisse. Die durchgeführten Interviews und die Expertise der drei Teilnehmer*innen ermöglichen einen tiefen Einblick in den aktuellen Stand von Künstlicher Intelligenz im Bereich Filmton und darüber hinaus lassen sich Perspektiven für die Zukunft erkennen - insbesondere für das Thema intelligente Mikrofonierung. Erkennbar ist, dass der Einsatz von KI aktuell eher in der Postproduktion zu finden ist, es allerdings Ideen und Konzepte für die Audioaufnahme am Set oder während Live-Produktionen gibt. Diese sind gegenwärtig nicht etabliert und dienen als Konzept oder zukünftige Erwartungen.

4.7.1 Aktueller Stand der KI-Nutzung im Filmton

Während der Postproduktion sind allen drei Interviewpartner*innen - Cordes, Krüger und Schmidt - das Nutzen von KI gestützter Software wie iZotope RX bekannt. Allerdings ist der Anwendungsfall durch die praktischen Arbeitsfelder (Live-Produktion, Setmeister, EB) der drei Kandidatinnen nicht präsent während ihrer Tätigkeiten.

Dementsprechend bleibt das Arbeitsfeld, sowohl bei der Live-Produktion als auch am Filmset, konventionell. Dies kann an fehlenden markttauglichen Produkten und begrenzter Rechenleistung liegen. Außerdem fehlen die Erfahrungswerte, sodass nicht gewährleistet sein kann, dass die KI richtig trainiert ist und dementsprechend zuverlässig funktioniert.

Aspekt	Cordes	Krüger	Schmidt
KI im Einsatz	Nein (außer EQ/Comp)	Ja, Postproduktion	Ja, Postproduktion
Potenzial für Mikrofone	Ja, Live/TV	Ja, Doku/Fiktion	Ja, EB/VJs
Hauptsorge	Verlust von Kontrolle	Technischer Overload	Blackbox Verhalten
Ethikbedenken	Kaum genannt	Ja, Rechtliche Grauzone	Ja, Rechtliche Grauzone

Tabelle 1. - Vergleichende Übersicht der Expert*inneneinschätzungen zur KI im Filmton

4.7.2 Erwartungshaltungen an intelligente Mikrofontechnik

In vielen Punkten sind sich alle einig- intelligente Mikrofone könnten Audioaufnahmen erleichtern und unterstützen. Dennoch gibt es Unterschiede in den Anwendungsfeldern.

Für Videograf*innen, welche als einzelne Personen einen Dreh durchführen, wäre ein intelligentes Mikrofon durchaus hilfreich. Ein Tool, welches als Assistenzsystem dient und die

Tätigkeit vereinfacht. Außerdem wäre hierfür interessant, „Allzweck-Mikrofone“ zu haben. Diese könnten intelligent ihre Richtcharakteristik anpassen, eine automatische Pegelanpassung und Silbentrennung haben und sensorgestützt sein. Cordes, Krüger und Schmidt sehen das Potenzial vor allem in Anwendungsfeldern wie Doku, Vlogs oder kleine EB-Teams.

4.7.3 Herausforderungen und Grenzen

Auf die Frage, welche Herausforderungen oder Risiken die drei sehen, lassen sich bei allen drei Stimmen klare Zusammenhänge erkennen.

Technische Herausforderungen wie zum Beispiel die hohe Rechenleistung oder die Echtzeitfähigkeit der KI könnten Probleme an Filmsets erzeugen. Zusätzlich hierzu kommt es darauf an, wie transparent die KI arbeitet - im Kontext bedeutet dies, dass Entscheidungen oder Befehle der KI klar nachvollziehbar sein müssen. Zur Folge hätte dies, dass Fehler bzw. Fehlentscheidungen gefunden und gelöst werden könnten und im besten Fall nicht mehr vorkommen, nach erneutem Trainieren der KI.

Ein weiterer wichtiger Faktor ist die Komplexität am Set - viele Geräte und deren Infrastruktur könnten den Produktionsprozess belasten. Das technische Know-How müsste mit Schulungen unterstützt werden. Außerdem muss der Betrieb der Geräte bei einem Drehtag von zehn Stunden gewährleistet sein, was folglich auf zusätzliche Akkus und/oder Batterien hinausläuft.

Als abschließenden Punkt ist die rechtliche Nutzung von KI zu klären. Aktuell gibt es kaum Regulation, welche den Nutzen von KI für das Urheberrecht definiert. Ein klar definiertes Gesetz für die Nutzung von KI generierter als auch modellierter Stimme muss hierfür gegeben sein.

Zusammenfassend lässt sich erkennen, dass es Schwierigkeiten in der Nutzung von KI in der praktischen Audioaufnahme gibt, welche vor dem kommerziellen Einsatz geklärt werden müssen. KI dient demnach aktuell als Unterstützung und Helfer - nicht als Ersatz der ausführenden Fachkräfte am Filmset.

4.7.4 Ausblick

Auf Basis der Interviewergebnisse zeigt sich, dass Künstliche Intelligenz im Bereich Filmtone noch in den Anfängen steckt - zwar werden viele leistungsstarke Tools bereits in der Postproduktion eingesetzt, im Bereich Setton bestehen noch Herausforderungen. Konzepte und Ideen sind in der Mikrofontechnik zu finden und lassen mit ambitionierten Vorstellungen in eine innovative Zukunft blicken.

Daniela Cordes' Perspektive ist besonders aufschlussreich, da sie nicht aus der klassischen Filmproduktion, sondern aus dem Bereich TV und Live-Audio kommt. Dadurch liefert sie wertvolle Impulse für zukünftige Forschungen jenseits des Spielfilms.

Urs Krügers' Einschätzungen bieten praxisnahe Perspektiven auf die aktuellen und zukünftigen Entwicklungen intelligenter Mikrofontechnik. Seine langjährige Erfahrung am Set erlaubt es, technische Potenziale ebenso wie Risiken realistisch einzuordnen. Für weitergehende Forschung wären vergleichende Interviews mit Postproduktions-Tonmeister*innen, Sounddesigner*innen sowie Hersteller*innen von Mikrofontechnik aufschlussreich.

Ole Schmidt liefert wertvolle Einsichten aus der praktischen elektronischen Berichterstattung. Seine Aussagen unterstreichen die hohe Relevanz praxisorientierter Forschung und zeigen, dass KI nicht nur technisches, sondern auch laut ihm ebenfalls gestalterisches und ethisches Neuland im Filmtonebereich betrifft.

Langfristig könnte sich das Berufsbild im Ton verändern – weg vom technischen Operator, hin zum Kurator von Klangrealitäten, der mit KI-Systemen arbeitet, diese aber auch reflektiert und kontrolliert. Wichtig bleibt: KI darf den kreativen Anspruch nicht untergraben, sondern sollte menschliche Stärken durch technische Entlastung fördern. Langfristig könnten selbstlernende Mikrofonarrays zur Norm werden, die sich automatisch an Umgebungsbedingungen anpassen.

4.6 Zukunftsperspektiven und Forschungspotentiale

Wichtige Aspekte für eine zukunftsorientierte Arbeit mit neuen KI-Technologien und Potenziale für weiterführende Forschung:

- Vergleichende Interviews mit Postproduktionstechniker*innen zur Einschätzung von Zeitersparnis und Automatisierung.

- Tests intelligenter Mikrofone in Live- und EB-Settings, z. B. mit Sensorik, KI-basiertem Signalrouting oder dynamischer Richtwirkung.
- Usability-Tests intelligenter Mikrofone mit Fokus auf Steuerbarkeit, Transparenz und Fehleranfälligkeit.
- Entwicklung von Erklärbarer KI (Explainable AI) speziell für kreative Arbeitsfelder wie den Filmtone.

Zudem wären quantitative Studien denkbar, die z. B. den tatsächlichen Einsatzgrad KI-gestützter Tools in der Filmtoneproduktion erfassen. Ein weiteres Forschungspotenzial liegt in der Ethik synthetisierter Stimmen, besonders im Spannungsfeld zwischen kreativer Kontrolle, Persönlichkeitsrechten und technischen Möglichkeiten.

8. Technischer Exkurs Lewitt RAY

8.1 Mikrofon Lewitt RAY

Im Rahmen der Ausarbeitung zum Themas „künstliche Intelligenz im Filmtone“ hat die Firma Lewitt, welche in der Herstellung von Mikrofonen für Audioaufnahmen tätig ist, ein Lewitt RAY für die Testaufnahmen zur Verfügung gestellt. Das Mikrofon ist ausgestattet mit einer 1“ Echkondensatorkapsel, besitzt eine Nierencharakteristik mit Rückdämpfung und dient dazu, Podcast- und Gesangsaufnahmen zu erstellen. Laut der Firma kann das Mikrofon „deinen Pegel und Sound perfekt ausbalancieren, weil der AURA Sensor deinen Abstand zum Mikrofon kennt und deinen Sound entsprechend anpasst.“ (Website Lewitt). Die Technologie arbeitet hier mit einem time-of-flight (ToF) Sensor, welcher den Abstand zur Klangquelle misst (im Bereich von 5cm bis 100cm) und daraufhin das Signal und die Dynamik anpasst (Website Lewitt). Ziel ist es hierbei, der Person vor dem Mikrofon die Möglichkeit zu geben, sich frei zu bewegen, ohne dass sich der Klangcharakter verändert. Mit den integrierten Sensoren und der dazugehörigen AURA Technology fällt das Lewitt RAY in die Kategorie „intelligenter Mikrofone“ und liefert somit triftige Ergebnisse für diese Ausarbeitung.

In den nachfolgenden Techgraphen wird aufgezeigt, wie die unterschiedlichen Frequenzgänge des Lewitt RAY bei ein- oder ausgeschalteter AURA Technology sind. Auf der y-Achse wird der Dynamikumfang in Form von Dezibel (dB) und auf der x-Achse der Frequenzband von 20Hz - 20000Hz dargestellt. Die grün gestrichelte Linie zeigt dabei jeweils den Frequenzgang

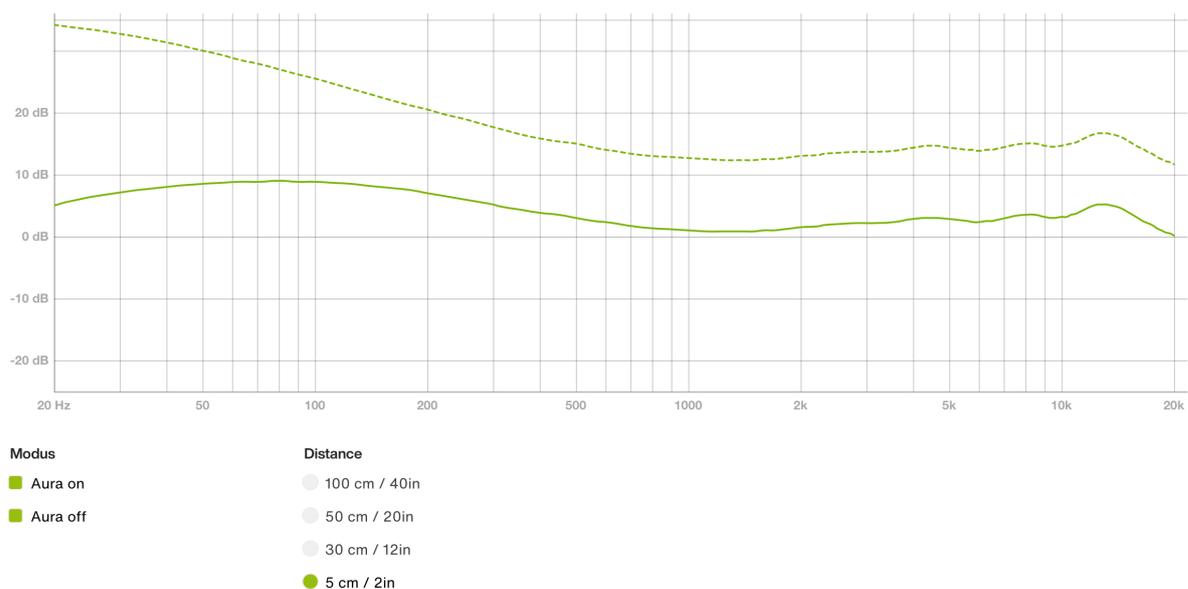


Abbildung 1. - Techgraph Lewitt RAY 5cm Entfernung

bei ausgeschalteter AURA Technology und die grün durchgezogene Linie dementsprechend bei eingeschalteter AURA Technology. Die Legende ist für die Abbildungen 1, 2, 3 und 4 exakt in diesem Schema identisch. Beginnend bei Abbildung 1 lässt sich erkennen, dass in den Frequenzen 20Hz - 200Hz eine Bassabsenkung durchgeführt wird, sodass der Nahbesprechungseffekt bei einer Entfernung von 5cm kompensiert wird. Das Klangbild wird demnach von der AURA Technology modelliert und dadurch ein ausgewogener Klang erzeugt.

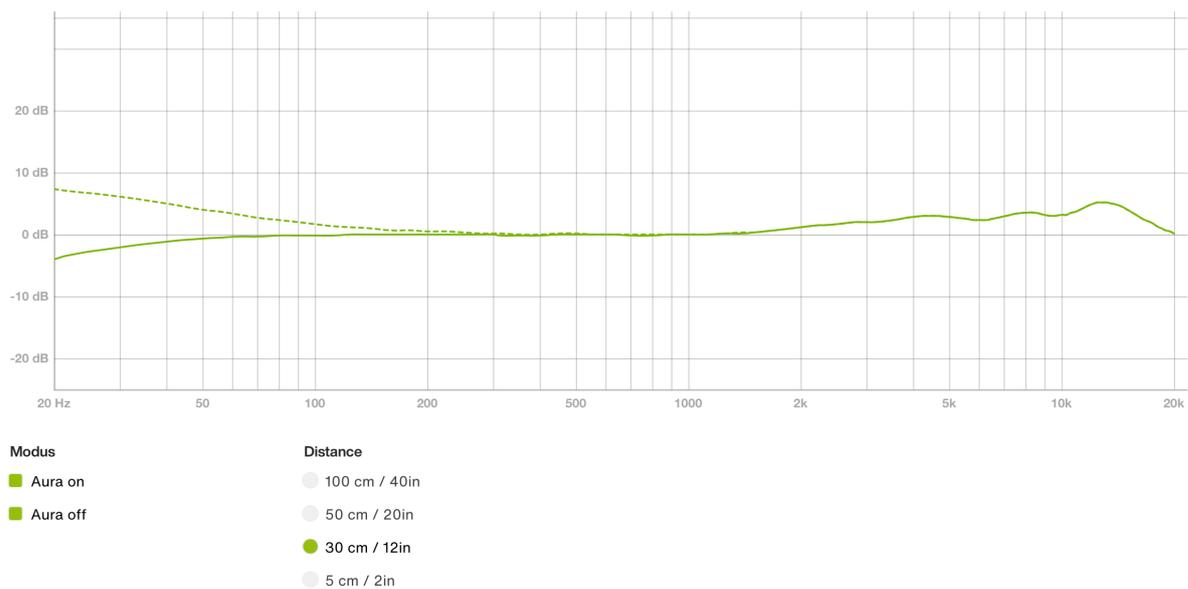


Abbildung 2. - Techgraph Lewitt RAY 30cm Entfernung

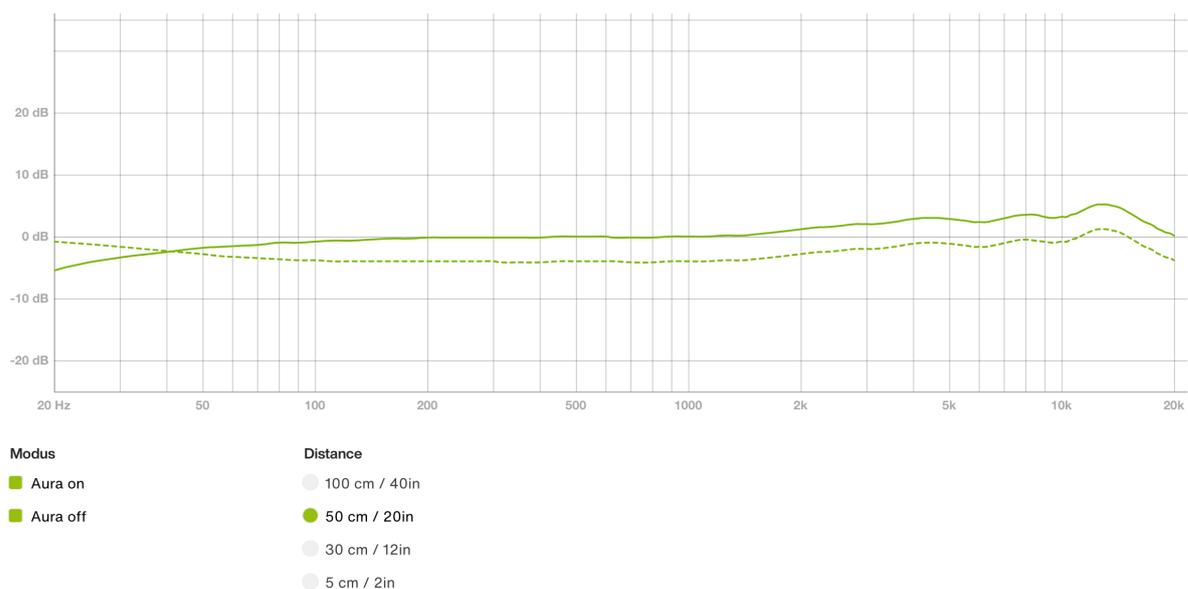


Abbildung 3. - Techgraph Lewitt RAY 50cm Entfernung

Fortlaufend zeigen Abbildung 2 und Abbildung 3 ein konsistentes Klangbild, bei dem ebenfalls eine Bassabsenkung zu erkennen ist. Wohingegen die grün gestrichelte in den Frequenzen von 20Hz - 100Hz abfällt, ist die grün durchgezogene Linie abgesenkt. Grund dafür ist auch hier die Kontrolle eines Nahbesprechungseffekts durch die AURA Technology. Aufgrund der größeren Entfernungen von 30cm und 50cm fällt diese Absenkung allerdings nicht so stark aus (ca. 10dB und 5dB).

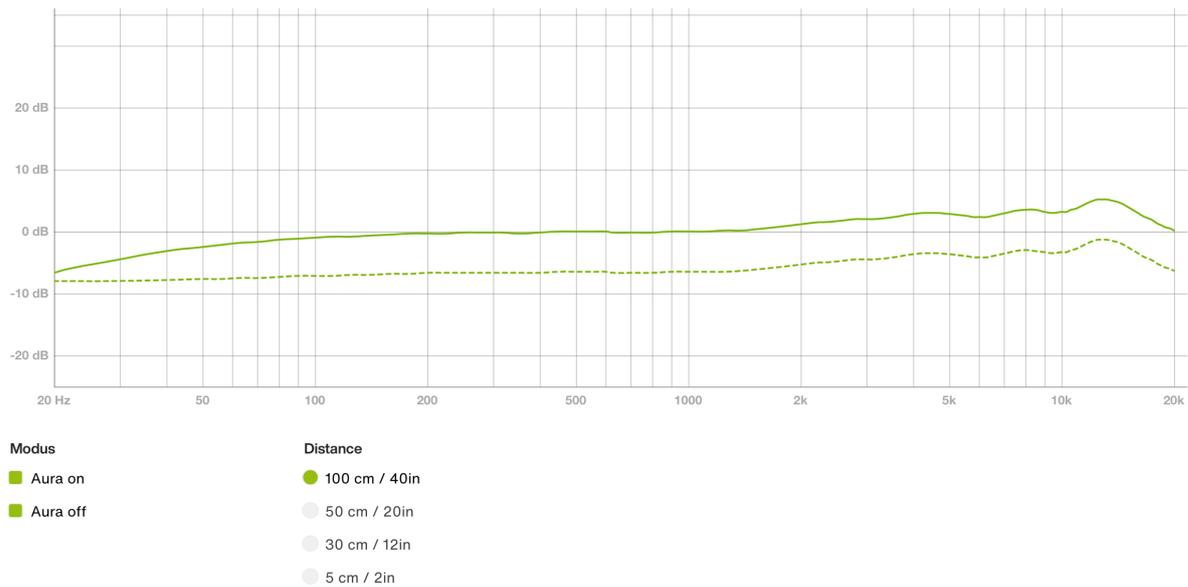


Abbildung 4. - Techgraph Lewitt RAY 100cm Entfernung

Während bei den vorherigen Abbildungen versucht wurde, ein konsistentes und ausgewogenes Klangbild durch eine Bassabsenkung zu schaffen, wird in Abbildung 4 deutlich, dass eine durchgehende Anhebung des Frequenzbandes ab einer Entfernung von 100cm von der Klangquelle zum Mikrofon stattfindet.

Abbildung 5 dient hier zur Ergänzung der Techgraphen und enthält die Richtcharakteristiken für die einzelnen Frequenzen des Lewitt RAY.

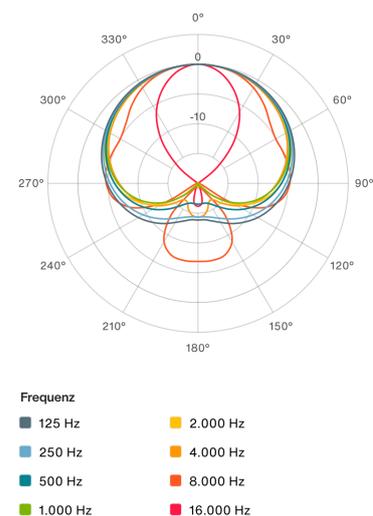


Abbildung 5. -

Richtcharakteristik Lewitt RAY

8.2 Zielsetzung

Die Ergebnisse der durchgeführten Aufnahmen sollen die Funktionsweise und die Eigenschaften des Lewitt RAY aufzeigen. Hierbei werden die Aufnahmen sowohl in kontrollierten Studioumgebungen als auch in herkömmlichen Testumgebungen wie in einem Wohnzimmer untersucht. Ziel war hierbei, die Wirkung der AURA Technology und darüber hinaus das allgemeine Klangbild unter verschiedenen Aspekten zu testen. Anschließend wurden diese Aufnahmen mit Hilfe eines Bewertungsbogens ausgewertet und analysiert.

8.3 Methodik

Für die praktische Untersuchung wurde ein strukturierter Testansatz entwickelt, der in zwei unterschiedlichen akustischen Umgebungen – einem Tonstudio sowie einem typischen Wohnzimmer – durchgeführt wurde (siehe Anhang). Ziel war es, die Leistungsfähigkeit des Lewitt RAY Mikrofons unter kontrollierten Bedingungen zu bewerten, insbesondere im Hinblick auf die Wirksamkeit der integrierten Adaptive Sound Technology (AST). Die Tests umfassten systematisch variierte Parameter wie Mikrofonabstand (10 cm bis 1 m), Dynamikverhalten sowie Bewegung der Klangquelle im Raum.

Um die Vielseitigkeit des Lewitt RAY zu testen, wurden drei verschiedene Klangquellen getestet. Die aufgenommenen Daten sollten möglichst aussagekräftig und vergleichbar sein, sodass versucht wurde, stetig den gleichen Dynamikumfang zu generieren. Für die Gesangsaufnahmen wurde ein Studio als kontrollierte Testumgebung gewählt, um eine möglichst trockene Aufnahme zu generieren. Die Sprachaufnahmen wurden mit der gleichen Technik unter unkontrollierten Voraussetzungen in einem Wohnzimmer aufgenommen. Die Parameter für die einzelnen Aufnahmen wurden wie folgt variiert und dokumentiert:

- **Entfernung zur Schallquelle:** 10 cm, 30 cm, 50cm, 1 m
- **Signalquellen:** Gesang, Sprache, Akustikgitarre
- **AURA Technology:** aktiviert / deaktiviert
- **Bewegungsverhalten:** statisch vs. sich bewegende Schallquelle

Durch die gewonnenen Daten lässt sich die Leistungsfähigkeit des Lewitt RAY spezifisch analysieren. Zur qualitativen Auswertung der Aufnahmen wurde ein standardisierter Bewertungsbogen (siehe Anhang) entwickelt, der gezielt Aspekte wie Raumhall, Klangneutralität, Klarheit, Plosiv- und Sibilantenkontrolle sowie adaptive Reaktionen des Vincent Schlagkamp

Mikrofons erfasst. Der Fragebogen erlaubt sowohl eine objektive Beurteilung durch Vergleich der Aufnahmen als auch eine subjektive Einschätzung durch Testpersonen. Die Fragestellungen sind mehrstufig skaliert, um differenzierte Rückmeldungen zu ermöglichen und gezielt Unterschiede zwischen AST-an und AST-aus-Szenarien zu erfassen. Sowohl zugehörige Audioaufnahmen als auch die Bewertungsbögen finden sich im Anhang.

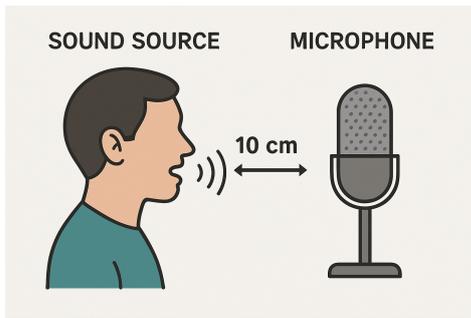


Abbildung 6. -
Klangquellenpositionierung &
Mikrofonposition (10cm)

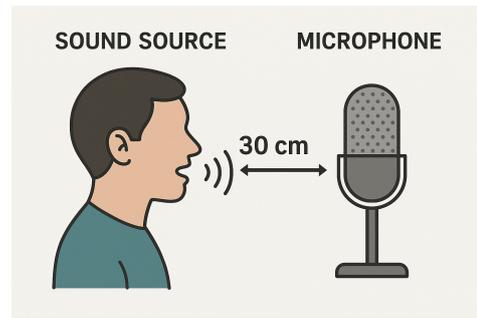


Abbildung 7. -
Klangquellenpositionierung &
Mikrofonposition (30cm)

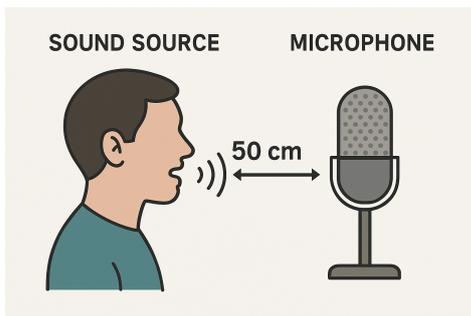


Abbildung 8. -
Klangquellenpositionierung &
Mikrofonposition (50cm)



Abbildung 9. -
Klangquellenpositionierung &
Mikrofonposition (100cm)

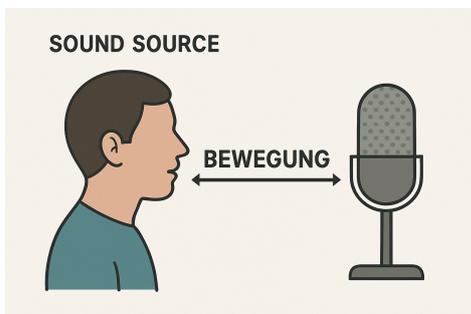


Abbildung 10. - Bewegliche
Klangquellenpositionierung & Mikrofon

Die dargestellten Abbildungen 6 bis 10 visualisieren die räumliche Beziehung zwischen einer Klangquelle (in Form einer sprechenden Person) und dem Lewitt RAY-Mikrofon. Sie zeigen verschiedene Abstände zwischen Mund und Mikrofon: 10 cm, 30 cm, 50 cm, 100 cm sowie eine abstrahierte Darstellung mit der Beschriftung "Bewegung" anstelle einer konkreten Zentimeterangabe.

8.4 Technische Auswertung

Um eine technische Auswertung durchzuführen, wurde das Metering Plug-In von Youlean genutzt. In den folgenden Abbildungen sind jeweils die LUFs (Loudness Units relative to Full Scale), die Dynamic Range (DR) und der True Peak (TP) der dazugehörigen Aufnahmen abzulesen. Die zugehörigen Audioaufnahmen und weitere Youlean-Auswertungen finden sich im Anhang. Für die Analyse wurden diese zwei Audioaufnahmen ausgewählt, da diese die triftigsten Ergebnisse lieferten.

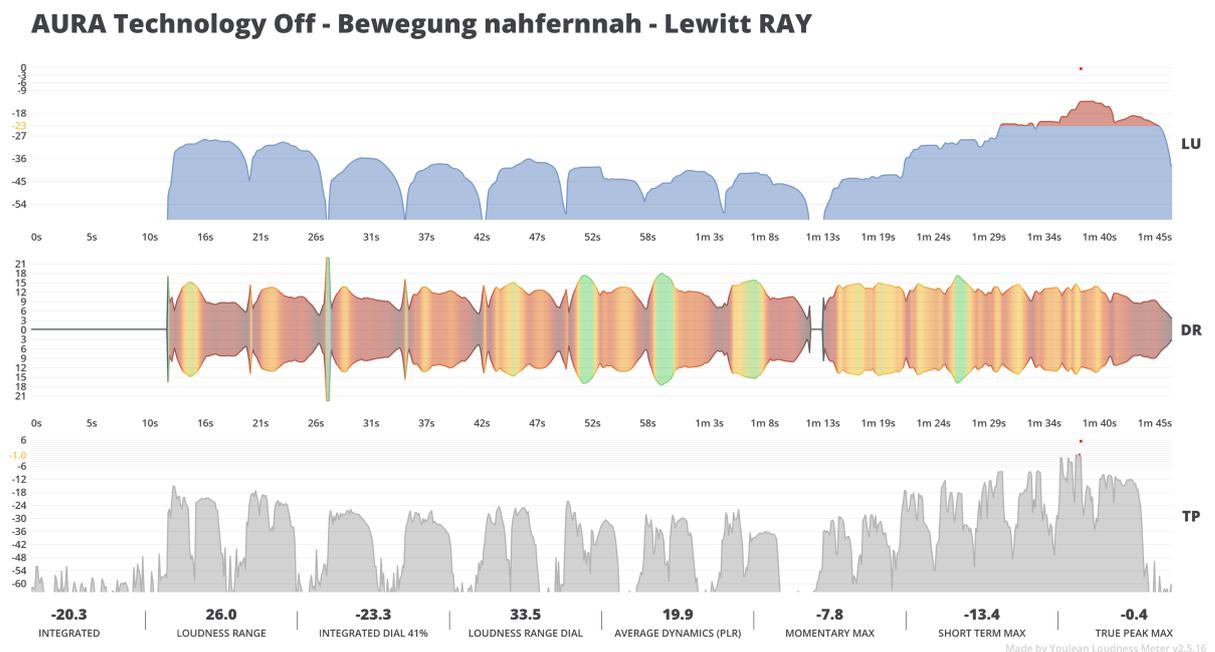


Abbildung 11. - AURA Technology OFF - Bewegung nah/fern/nah - Lewitt RAY

In Abbildung 11. ist eine Grafik zu sehen, welche den Verlauf einer Audioaufnahme einer Sängerin unter Anbetracht der vorher genannten Parameter aufzeigt. Während der Aufnahme war die AURA Technology des Lewitt RAY ausgeschaltet. Erkennbar ist, dass zu Beginn der Aufnahme, bei ca. Sekunde 12, die Aufnahme -27 LUFs und einen True Peak von -30 dB aufweist. Im Verlauf der Zeit schwachen diese Werte ab, bis sie bei 1 m 8 s ihr Minimum von -46 LUFs und -36 dB erreichen. Zum Ende der Aufnahme lässt sich nun wieder im Graph erkennen, dass die beiden Werte ansteigen auf ihr Maximum von -20 LUFs und -6 dB.

Während der Audioaufnahme hat sich die Sängerin zunächst in einem Abstand von 10cm vom Mikrofon befunden, hat sich dann bis auf 1 m vom Mikrofon entfernt und schließlich wieder angenähert bis auf 10cm. In der Aufnahme ist zu hören, dass zu Beginn der Aufnahme die

Lautstärke der Stimme zunächst gering ist und die Sängerin ihre Stimme soft und bewusst einsetzt. Gegen Ende jedoch wird ihre Stimme kraftvoller und die Lautstärke hebt sich.

AURA Technology ON - Bewegung nahfer/nah - Lewitt RAY

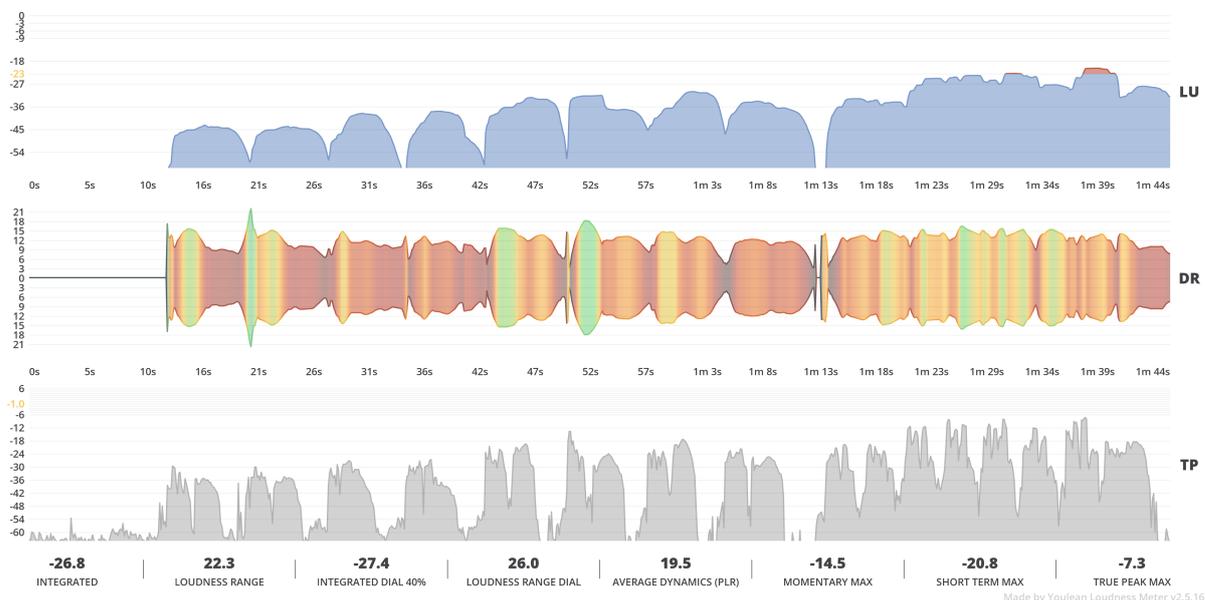


Abbildung 12. - AURA Technology ON - Bewegung nah/fern/nah - Lewitt RAY

Im Vergleich zu der vorherigen Grafik ist in Abbildung 12. zu sehen, wie der Verlauf einer Audioaufnahme sich unter eingeschalteter AURA Technology unter denselben Parametern verhält.

Die dargestellten Kurven zeigen den Verlauf der Lautheit in LUFS sowie der True Peak-Werte über die Dauer der Aufnahme. Zu Beginn der Aufnahme, bei ca. 12 Sekunden, liegt der Lautheitswert bei -44 LUFs und der True Peak bei -30 dB. Trotz der sich im Verlauf ändernden Mikrofonposition bleiben die Pegel insgesamt konstanter und ausgeglichener als bei deaktivierter AURA-Technologie.

Die Sängerin bewegt sich während der Aufnahme erneut vom Mikrofon weg – beginnend bei einem Abstand von 10 cm, sich entfernend auf etwa 1 m, und nähert sich zum Ende wieder an. Diese Bewegung ist in der Grafik dieses Mal kaum erkennbar. Die Lautheits- und Pegelschwankungen fallen ebenfalls geringer aus als in der vorherigen Messung. Allgemein lässt sich somit ein kontinuierlicher Lautstärkeanstieg erkennen, welcher mit der Dynamiksteigerung der Sängerin in Verbindung steht und von dem Lewitt RAY klar abgebildet wird.

Die AURA Technology gleicht die Pegelunterschiede, die durch den variierenden Mikrofonabstand entstehen, automatisch aus. Dies wird in der Grafik daran deutlich, dass die Differenz zwischen Minimal- und Maximalwerten von -44 LUFS und -25 LUFS deutlich geringer ausfällt. Auch der True Peak Max bleibt innerhalb eines kontrollierten Bereichs von -7,3 dB, was für eine gleichmäßigere Wahrnehmung sorgt.

In der auditiven Wahrnehmung äußert sich dies darin, dass die Stimme während der gesamten Aufnahme präsenter und konsistenter klingt – unabhängig von der tatsächlichen Position der Sängerin vor dem Mikrofon.

8.4 Auswertung Bewertungsbögen

Die Klangeigenschaften des Lewitt RAY lassen sich nach der Auswertung der Bewertungsbögen und unter Anbetracht der getesteten Umgebungen drei Faktoren erkennen, welche eine große Rolle spielen - die Entfernung zur Klangquelle, ob die AURA Technology ein- oder ausgeschaltet ist und um was für eine Klangquelle es sich handelt. Insbesondere wird hier unter Berücksichtigung der Testaufnahmen unterschieden von Sprach-, Gesangs- oder Instrumentalaufnahmen.

Wenn es um das Verhalten des Raumhalls in Bezug auf die Klangeigenschaften des Lewitt RAY geht, wird schnell klar, dass sowohl bei Aufnahmen in Studioumgebung als auch im unkontrollierten Wohnzimmer dieser stark zunimmt ab einer Entfernung von 50cm. Im Vergleich zu den Aufnahmen, welche in einer Entfernung von 10-30cm aufgenommen wurden. Hier lässt sich kaum Raumhall wahrnehmen und das Klangbild erscheint natürlich. Reflektierend kontrolliert die AURA Technology demnach den Raumhall bis hin zu einer Entfernung von 50cm effektiv, verliert daraufhin allerdings ihre Wirksamkeit.

In Faktor Natürlichkeit der Aufnahmen punktet das Lewitt RAY, denn sowohl während der kleinen, als auch der großen Entfernungen der Klangquelle zum Mikrofon wirken die Aufnahmen warm und präsent. Ausschließlich bei vereinzelt aufgenommenen Aufnahmen kommt es zu geringerer Klangqualität, was durchaus auf die Art und Weise der Beschallung zurückzuführen sein kann.

Durch die integrierten Sensoren und die damit einhergehende AURA Technology kompensiert das Lewitt RAY beeindruckend Dynamikschwankungen der Klangquelle. Durch einmaliges Einstellen des Pegels wurde gewährleistet, dass eine Messbarkeit der Aufnahmen in Bezug

auf die Pegelkompensation umsetzbar ist. Betrachten wir die optischen Darstellungen und die technischen Auswertungen des Bewertungsbogen lässt sich erkennen, dass es kaum zu Lautstärkeschwankungen kommt. Außerdem gibt es bei der Pegelkompensation keine Auffälligkeiten und demnach beeindruckt das Lewitt RAY vor allem bei den Entfernungen 10-50cm.

Bei der Beurteilung der Klangdetails und Klarheit fällt auch hier der Bewertungsbogen gut aus. Bei steigender Entfernung differenzieren sich auch hier manche Aufnahmen, doch ebenfalls bei den Bewegungsaufnahmen lassen sich kaum Veränderungen wahrnehmen und die Aufnahmen bleiben klar verständlich. Wie bei vorherigen Parametern schon angemerkt, ist ausschließlich ein deutlicher Unterschied bei ausgeschalteter AURA Technology merkbar.

Um weitere Unterschiede zu herkömmlichen Mikrofonen herauszuarbeiten, wurde zusätzlich ein Plosiv- und Sibilantentest durchgeführt. Wohingegen die Sibilanten gut kontrolliert werden und keinen ausgeprägten Unterschied zu traditionellen Mikrofonen haben, sind Plosivlaute präsent in den Aufnahmen. Die Kontrolle konnte hierbei nur über die Entfernung und nicht über ein- oder ausschalten der AURA Technology garantiert werden. Schlussfolgernd arbeitet das Lewitt RAY in diesem Fall nicht innovativer als andere Mikrofone.

Wie zuvor kurz aufgegriffen arbeiten die Entfernungssensoren des Lewitt RAY ausgesprochen gut. Sowohl bei Bewegungen zum Mikrofon hin der auch weg Arbeiter die pegelkompensation funktional und bietet somit eine Absicherung der Aufnahmequalität für gesprochene Inhalte und auch Gesangsaufnahmen. Obwohl die Bewegungen für Sänger*innen unnatürlich wirken, hilft dies kontrollierte Aufnahmen zu gewährleisten. Die gleichen Aufnahmen ohne die AURA Technology wirken verglichen dünn und der Dynamikumfang nimmt merklich ab, bei Steigerung der Entfernung. Dies führt dazu, dass bestimmte Aufnahmen nicht brauchbar wären, um eine professionelle und hochwertige Postproduktion zu garantieren.

Unter Anbetracht der ausgewerteten Ergebnisse ist das Lewitt RAY ein professionelles Mikrofon um hochwertige Aufnahmen durchzuführen. Bis zu einer Entfernung von 50cm arbeitet das Mikrofon zudem tadellos und bildet ein schönes und natürliches Klangbild ab. Die oben genannten Faktoren haben dementsprechend maßgeblich um messbare und vor allem auch direkt hörbare Auswirkungen der Aufnahmequalität.

9. Fazit und Ausblick

Im formfolgenden Kapitel werden die wesentlichen Erkenntnisse zusammengefasst und ein Ausblick auf zukünftige Forschungsfragen gegeben.

Diese Arbeit hat die historische Entwicklung, den aktuellen Stand und die zukünftigen Potenziale von Künstlicher Intelligenz im Filmtone umfassend untersucht. Zunächst konnte gezeigt werden, wie sich Mikrofon- und Tonaufnahmetechnologien von den mechanischen Anfängen bis hin zu digitalisierten Systemen gewandelt haben. Darauf aufbauend wurde der Einsatz moderner KI-Algorithmen – insbesondere in Form neuronaler Netze zur Klangoptimierung und automatischen Geräuschunterdrückung – analysiert. Anhand der Auswertung von Experteninterviews und der Fallstudie zum Lewitt RAY-Mikrofon ließ sich belegen, dass KI-gestützte Verfahren bereits heute zu einer deutlich gesteigerten Klangqualität am Filmset beitragen könnten.

Die Interviews mit Expert*innen aus der Branche haben verdeutlicht, dass technische Innovationen wie Adaptive Noise-Cancellation-Module oder Echtzeit-Klangmodellierungen nicht nur die Produktionsabläufe vereinfachen, sondern auch neue kreative Freiräume schaffen. Gleichzeitig wurde deutlich, dass der Einsatz von KI nicht frei von Herausforderungen ist: Fragen der Zuverlässigkeit unter extremen Aufnahmebedingungen, ethische Aspekte beim Einsatz adaptiver Lernsysteme sowie die Notwendigkeit klarer Standards für Evaluation und Zertifizierung sind zentrale Punkte, die in zukünftigen Entwicklungszyklen adressiert werden müssen.

Abschließend lässt sich festhalten, dass Künstliche Intelligenz im Filmtone ein enormes Potenzial besitzt, die Klangqualität und Effizienz in der audiovisuellen Produktion nachhaltig zu verbessern. Um dieses Potenzial voll auszuschöpfen, bedarf es jedoch einer engen Verzahnung von Forschung und Praxis, transparenter Kriterien zur Beurteilung KI-basierter Tools sowie weiterführender Studien zur ergonomischen und ethischen Dimension der KI-Nutzung. In zukünftigen Arbeiten sollten neben der technischen Weiterentwicklung insbesondere Nutzer*innenerfahrungen und langfristige Auswirkungen auf Produktionsprozesse systematisch untersucht werden. Damit schafft die vorliegende Arbeit eine solide Grundlage für die weitere Integration intelligenter Audiosysteme in die professionelle Filmproduktion.

Definitionen

Künstliche Intelligenz

KI bezeichnet ein interdisziplinäres Forschungsfeld, in dem „intelligente Agenten“ entwickelt werden – also Systeme, die ihre Umwelt wahrnehmen, aus diesen Wahrnehmungen lernen, eigenständig Ziele verfolgen und basierend darauf Entscheidungen treffen und handeln.

Bartneck, C. et al. (2021) ‘What Is AI?’, in C. Bartneck et al. (eds) An Introduction to Ethics in Robotics and AI. Cham: Springer International Publishing, pp. 5–16. Available at: https://doi.org/10.1007/978-3-030-51110-4_2.

Filmton

Filmton umfasst in der Postproduktion sämtliche Audioaspekte eines Films oder Videos: Dialog, Umgebungsgeräusche, Effekte, Foley-Arbeit und Musik. Die Klanggestaltung erfolgt synchron zum visuellen Bild, mit dem Ziel, narrative und emotionale Tiefe zu erzeugen. Qualitative Grundlagen und Methoden werden in standardisierten Produktionsabläufen erläutert

Stavrides, S.C. (2019) ‘Film Sound Design & Effects | Cyber Film School’, 30 June. Available at: <https://cyberfilmschool.com/film-sound-design/> (Accessed: 15 June 2025).

Wellenfeldsynthese

WFS ist ein räumliches Audiowiedergabe-Verfahren, bei dem durch steuert betriebene Lautsprecherarrays künstlich erzeugte Schallwellenfronten aufgebaut werden, sodass virtuelle Klangquellen räumlich unabhängig von der Position der Zuhörer im Raum wahrgenommen werden.

Brandenburg, K., Brix, S. and Sporer, T. (no date) ‘WAVE FIELD SYNTHESIS: FROM RESEARCH TO APPLICATIONS’. Available at: https://www.eurasip.org/Proceedings/Eusipco/Eusipco2004/defevent/papers/cr1938.pdf?utm_source=chatgpt.com.

Convolutional Neural Networks (CNNs)

Ein CNN ist ein mehrschichtiges neuronales Netzwerk, das für die automatische Hierarchie-Extraktion von Merkmalen aus Gitterdaten (z. B. Bilder oder spektrale Audiodaten) entwickelt wurde.

Was sind konvolutionale neuronale Netze? | IBM (2021). Available at: <https://www.ibm.com/de-de/think/topics/convolutional-neural-networks> (Accessed: 15 June 2025).

Deep U-Net Modelle

Deep U-Net-Modelle sind spezielle neural-netzwerk-basierte Architekturen im Bereich des Deep Learning, die ursprünglich für präzise Bildsegmentierungsaufgaben entwickelt wurden.

Ronneberger, O., Fischer, P. and Brox, T. (2015) 'U-Net: Convolutional Networks for Biomedical Image Segmentation'. arXiv. Available at: <https://doi.org/10.48550/arXiv.1505.04597>.

piezoelektronische Akustiksensoren (f-PAS)

Piezoelektronische Akustiksensoren basieren auf dem piezoelektrischen Effekt, bei dem mechanischer Druck auf ein piezoelektrisches Medium (z. B. Keramiken wie PZT oder ZnO) eine proportionale elektrische Ladung bzw. Spannung erzeugt. Diese Sensoren wandeln akustischen Druck direkt in ein elektrisches Signal um.

Campo-Valera, M. et al. (2023) 'Characterization of a Piezoelectric Acoustic Sensor Fabricated for Low-Frequency Applications: A Comparative Study of Three Methods', Sensors, 23(5), p. 2742. Available at: <https://doi.org/10.3390/s23052742>.

MEMS-Mikrofonen (Micro-Electro-Mechanical-Systems)

MEMS-Mikrofone sind winzige (typischerweise < 1 mm) Mikrofonbauteile, die mittels mikro-elektromechanischer Systeme integriert akustischen Druck in elektrische Signale umwandeln.

Kumar, A. et al. (2022) 'Recent development and futuristic applications of MEMS based piezoelectric microphones', Sensors and Actuators A: Physical, 347, p. 113887. Available at: <https://doi.org/10.1016/j.sna.2022.113887>.

mel-frequency cepstral coefficients (MFCCs)

MFCCs sind kompakte Merkmalsvektoren, die aus einer logarithmierten spektralen Analyse eines Audiosignals auf einer psychoakustisch gewichteten Mel-Skala gewonnen werden. Sie beschreiben den spektralen Verlauf (Envelope) – häufig genutzt zur Sprach- und Musikerkennung

Mel Frequency Cepstral Coefficient - an overview | ScienceDirect Topics (no date). Available at: https://www.sciencedirect.com/topics/computer-science/mel-frequency-cepstral-coefficient?utm_source=chatgpt.com (Accessed: 15 June 2025).

t-Distributed Stochastic Neighbor Embedding (t-SNE)

Ein nicht-linearer Algorithmus zur Dimensionsreduktion, der hochdimensionale Daten in 2D oder 3D abbildet. Er erhält lokale Nachbarschaftsstrukturen durch Minimierung der Kullback–Leibler-Divergenz zwischen Wahrscheinlichkeitsverteilungen über Original- und Zielraum

Jung, S. et al. (2024) 'A Review of t-SNE', Image Processing On Line, 14, pp. 250–270. Available at: <https://doi.org/10.5201/ipol.2024.528>.

LUFs (Loudness Units relative to Full Scale)

Ein genormtes Maß zur objektiven Bestimmung der wahrgenommenen Lautheit digitaler Audiosignale historisch analog zu dBFS. LUFs berücksichtigt psychoakustische Gewichtungen undangaben der Lautheit gemäß EBU R128 für einheitliche Lautstärkelevel bei Sendungen und Produktionen

What Are LUFs? The Complete Guide (no date). Available at: <https://www.izotope.com/en/learn/what-are-lufs> (Accessed: 15 June 2025).

Literaturverzeichnis

AURA Technology - How it works (no date) LEWITT. Available at: <https://www.lewitt-audio.com/blog/aura-technology-how-it-works> (Accessed: 10 June 2025).

Bartneck, C. *et al.* (2021) ‘What Is AI?’, in C. Bartneck *et al.* (eds) *An Introduction to Ethics in Robotics and AI*. Cham: Springer International Publishing, pp. 5–16. Available at: https://doi.org/10.1007/978-3-030-51110-4_2.

Boulianne, G. *et al.* (2003) ‘Automatic segmentation of film dialogues into phonemes and graphemes’, in *Proc. Eurospeech 2003*, pp. 1241–1244. Available at: <https://doi.org/10.21437/Eurospeech.2003-397>.

Brandenburg, K., Brix, S. and Sporer, T. (no date) ‘WAVE FIELD SYNTHESIS: FROM RESEARCH TO APPLICATIONS’. Available at: https://www.eurasip.org/Proceedings/Eusipco/Eusipco2004/defevent/papers/cr1938.pdf?utm_source=chatgpt.com.

Campo-Valera, M. *et al.* (2023) ‘Characterization of a Piezoelectric Acoustic Sensor Fabricated for Low-Frequency Applications: A Comparative Study of Three Methods’, *Sensors*, 23(5), p. 2742. Available at: <https://doi.org/10.3390/s23052742>.

Chattopadhyay, B. (ed.) (2023) ‘Direct Sound and Early Talkies’, in *Sound in Indian Film and Audiovisual Media: History, Practices and Production*. Amsterdam University Press, pp. 69–82. Available at: <https://doi.org/10.1017/9789048551668.003>.

Crossref (no date) *Künstliche Intelligenz in der Hochschulbildung*. Available at: <https://record.crossref.org/> (Accessed: 16 June 2025).

Cutting, J.E. (2019) ‘Sequences in popular cinema generate inconsistent event segmentation’, *Attention, Perception, & Psychophysics*, 81(6), pp. 2014–2025. Available at: <https://doi.org/10.3758/s13414-019-01757-w>.

Gomes Lobato, T.H. and Sottek, R. (2023) *Super-Resolution of Sound Source Radiation Using Microphone Arrays and Artificial Intelligence*. Available at: <https://doi.org/10.4271/2023-01-1142>.

Harwardt, M. and Köhler, M. (2023) ‘Künstliche Intelligenz’, in M. Harwardt and M. Köhler (eds) *Künstliche Intelligenz entlang der Customer Journey: Einsatzpotenziale von KI im E-Commerce*. Wiesbaden: Springer Fachmedien, pp. 21–29. Available at: https://doi.org/10.1007/978-3-658-39109-6_3.

Hessler, M. and Liggieri, K. (eds) (2020) *Technikanthropologie: Handbuch für Wissenschaft und Studium*. 1. Auflage. Baden-Baden: Nomos, Edition Sigma, in der Nomos Verlagsgesellschaft. Available at: <https://doi.org/10.5771/9783845287959>.

Hillman, N. and Pauletto, S. (2014) ‘The Craftsman: The use of sound design to elicit emotions’, *The Soundtrack*, 7. Available at: https://doi.org/10.1386/st.7.1.5_1.

Jung, S. *et al.* (2024) ‘A Review of t-SNE’, *Image Processing On Line*, 14, pp. 250–270. Available at: <https://doi.org/10.5201/ipol.2024.528>.

Jung, Y.H. *et al.* (2021) ‘Deep Learning-based Noise-Robust Flexible Piezoelectric Acoustic Sensors for Speech Processing’. Research Square. Available at: <https://doi.org/10.21203/rs.3.rs-799114/v1>.

Knaus, T. (2024) ‘Warum KI kein Hype ist und die Medienpädagogik sich damit befassen sollte’, *merz | medien + erziehung*, 68(3), pp. 21–30. Available at: <https://doi.org/10.21240/merz/2024.3.10>.

Kotti, M. *et al.* (2008) ‘Audio-Assisted Movie Dialogue Detection’, *IEEE Trans. Circuits Syst. Video Techn.*, 18, pp. 1618–1627. Available at: <https://doi.org/10.1109/TCSVT.2008.2005613>.

Koyama, S. *et al.* (2014) ‘Real-Time Sound Field Transmission System by Using Wave Field Reconstruction Filter and Its Evaluation’, *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences*, E97.A, pp. 1840–1848. Available at: <https://doi.org/10.1587/transfun.E97.A.1840>.

Kumar, A. *et al.* (2022) ‘Recent development and futuristic applications of MEMS based piezoelectric microphones’, *Sensors and Actuators A: Physical*, 347, p. 113887. Available at: <https://doi.org/10.1016/j.sna.2022.113887>.

Li, Y. *et al.* (2024) ‘The Influence of Music Psychology on the Consistency of Mood in Movies: A Review of Historical Works and Empirical Research’, *Herança*, 7(1), pp. 44–57. Available at: <https://doi.org/10.52152/heranca.v7i1.837>.

Liu, Y. *et al.* (2020) ‘Research on Sound Source Localization Algorithm of Spatial Distributed Microphone Array Based on PHAT Model’, in Q. Liang *et al.* (eds) *Communications, Signal Processing, and Systems*. Singapore: Springer, pp. 1443–1446. Available at: https://doi.org/10.1007/978-981-13-6504-1_171.

Ma, R. *et al.* (2017) *Smart microphone array design for speech enhancement in financial VR and AR*, p. 3. Available at: <https://doi.org/10.1109/ICSENS.2017.8234178>.

Manolas, C. and Pauletto, S. (2009) ‘Enlarging the Diegetic Space: Uses of the Multichannel Soundtrack in Cinematic Narrative’, *The Soundtrack*, 2, pp. 39–55. Available at: https://doi.org/10.1386/st.2.1.39_1.

Mel Frequency Cepstral Coefficient - an overview | ScienceDirect Topics (no date a). Available at: https://www.sciencedirect.com/topics/computer-science/mel-frequency-cepstral-coefficient?utm_source=chatgpt.com#definition (Accessed: 15 June 2025).

Microflex® Advance™ Mikrofonarrays (no date). Available at: <https://www.shure.com/de-de/loesungen/microflex-advance> (Accessed: 16 June 2025).

Miranda, E.R. and Williams, D. (2015) ‘Artificial Intelligence in Organised Sound’, *Organised Sound*, 20(1), pp. 76–81. Available at: <https://doi.org/10.1017/S1355771814000454>.

Modern Sound Recording Technologies | Semantic Scholar (no date). Available at: https://www.semanticscholar.org/paper/Modern-Sound-Recording-Technologies-Yunyk-Riazantsev/f76a61adb8cad9911cc838e7010663749180e712?utm_source=consensus (Accessed: 2 December 2024).

‘Movie Sound, Part 2: Preference and Attribute Ratings of Six Listening Environments’ (no date) *AES*. Available at: <https://aes2.org/publications/elibrary-page/> (Accessed: 2 December 2024).

Prather, W.E., Frazier, W.G. and Gilliland, C. (2023) ‘Development of an 128 element array for demonstrating Artificial Intelligence techniques in an Acoustic Sensor System’, *The Journal of the Acoustical Society of America*, 154(4_supplement), pp. A329–A330. Available at: <https://doi.org/10.1121/10.0023702>.

RAY - *Autofokus für deine Stimme*. (no date) LEWITT. Available at: <https://www.lewitt-audio.com/de/ray> (Accessed: 8 June 2025).

Reed, J. and Woolworth, D. s. (2022) ‘Modern movie sound: reality and simulated reality’, *The Journal of the Acoustical Society of America*, 152(4_Supplement), p. A181. Available at: <https://doi.org/10.1121/10.0015961>.

Ronneberger, O., Fischer, P. and Brox, T. (2015) ‘U-Net: Convolutional Networks for Biomedical Image Segmentation’. arXiv. Available at: <https://doi.org/10.48550/arXiv.1505.04597>.

S, S.D.C. and S, R.K. (2023) ‘EFFECTS OF SOUND AND HOW AUDIENCES PERCEIVE’. A SPECIFIC REFERENCE STUDY USING THE MALAYALAM FILM "JALLIKATTU”, *ShodhKosh: Journal of Visual and Performing Arts*, 4(1SE), pp. 31–34. Available at: <https://doi.org/10.29121/shodhkosh.v4.i1SE.2023.408>.

Stavrides, S.C. (2019) ‘Film Sound Design & Effects | Cyber Film School’, 30 June. Available at: <https://cyberfilmschool.com/film-sound-design/> (Accessed: 15 June 2025).

Steffek, F. (2023) ‘DIGITALISIERUNG. Die Veränderung der Konfliktlösung durch künstliche Intelligenz – Teil 2’, *Zeitschrift für Konfliktmanagement*, 26(4), pp. 121–126. Available at: <https://doi.org/10.9785/zkm-2023-260404>.

Stubbe, J., Wessels, J. and Zinke, G. (2019) ‘Neue Intelligenz, neue Ethik?’, in V. Wittpahl (ed.) *Künstliche Intelligenz: Technologie | Anwendung | Gesellschaft*. Berlin, Heidelberg: Springer, pp. 239–254. Available at: https://doi.org/10.1007/978-3-662-58042-4_15.

Teich, I. (2020) ‘Meilensteine der Entwicklung Künstlicher Intelligenz’, *Informatik Spektrum*, 43(4), pp. 276–284. Available at: <https://doi.org/10.1007/s00287-020-01280-5>.

Tsutsumi, K. and Takada, H. (2017) ‘Powerful Sound Effects at Audience Seats by Wave Field Synthesis’, *NTT Technical Review*, 15(12), pp. 25–31. Available at: <https://doi.org/10.53829/ntr201712fa5>.

Varela, M., Nalla, R. and Wirth, W.-D. (2022) ‘Combining remote sensing and artificial intelligence to locate sounds’, *The Journal of the Acoustical Society of America*, 151, pp. A36–A36. Available at: <https://doi.org/10.1121/10.0010575>.

Vecera, E. (2020) ‘Künstliche Intelligenz in Bibliotheken’, *Information - Wissenschaft & Praxis*, 71(1), pp. 49–52. Available at: <https://doi.org/10.1515/iwp-2019-2053>.

Was sind konvolutionale neuronale Netze? | IBM (2021). Available at: <https://www.ibm.com/de-de/think/topics/convolutional-neural-networks> (Accessed: 15 June 2025).

What Are LUFs? The Complete Guide (no date). Available at: <https://www.izotope.com/en/learn/what-are-lufs> (Accessed: 15 June 2025).

XLR-Mikrofon mit Audiofokus: Lewitt Ray im Test | heise online (no date). Available at: <https://www.heise.de/tests/Lewitt-Ray-XLR-Mikrofon-mit-Audiofokus-im-Test-9702041.html> (Accessed: 2 December 2024).

Yamashita, R. *et al.* (2018) ‘Convolutional neural networks: an overview and application in radiology’, *Insights into Imaging*, 9(4), pp. 611–629. Available at: <https://doi.org/10.1007/s13244-018-0639-9>.

Zhang, L., Li, Z. and Gahb, J.S. (2023) ‘Research and analysis of interaction in virtual reality films in the context of new media’, *SHS Web of Conferences*, 167, p. 01020. Available at: <https://doi.org/10.1051/shsconf/202316701020>.

Zhang, Y. (2021) ‘Analysis on the Sound of the Movie The Man Who Knew Too Much’, in. *2nd International Conference on Language, Art and Cultural Exchange (ICLACE 2021)*, Atlantis Press, pp. 510–513. Available at: <https://doi.org/10.2991/assehr.k.210609.101>.

ZYLIA PRO Have It All! - 3rd Order Ambisonics Microphone & Software (no date) Zylia e-shop. Available at: <https://www.zylia.co/shop/zylia-pro-have-it-all-zylia-pro-have-it-all-216> (Accessed: 16 June 2025).

Anhang

Der Anhang zu dieser Arbeit ist über den im Dokument aufgeführten QR-Code digital abrufbar. Aufgrund der Vielzahl an ergänzenden Materialien – darunter Audiodateien, Videodateien sowie PDF-Dokumente – wurde auf eine physische Beilage verzichtet. Um die Inhalte dennoch übersichtlich und kompakt bereitzustellen, wurde ein strukturierter digitaler Ordner angelegt. Dieser ermöglicht einen gezielten Zugriff auf die jeweiligen Dateien und fördert eine effiziente Nutzung der ergänzenden Materialien.



QR - Code führt zum Anhang

Eidesstattliche Erklärung

Hiermit erkläre ich, Vincent Schlagkamp, dass ich die vorliegende Arbeit eigenständig und ohne fremde Hilfe angefertigt habe. Textpassagen, die wörtlich oder dem Sinn nach auf Publikationen oder Vorträgen anderer Autoren beruhen, sind als solche kenntlich gemacht. Die Arbeit wurde bisher keiner anderen Prüfungsbehörde vorgelegt und auch noch nicht veröffentlicht.

Mettingen, 18.06.2025



Vincent Schlagkamp