

Musikinstrumente im virtuellen Raum.**Strategien zur Digitalisierung von akustischen Kulturgütern am Beispiel der Orgel und
Konzepte eines interaktiven Zugangs zu virtuellen akustischen Objekten**Abstract:

Die Digitalisierung von materiellen und immateriellen Kulturgütern entwickelt zunehmend das Potenzial einer weltumspannenden Vernetzung kulturpraktischen Wissens, auf das mittels virtueller Zugänge audiovisuell oder in datenbasierter Form zugegriffen werden kann. Zu diesen gehören in aktuellen Vorhaben neben architektonischen Bauwerken und rekonstruierten Stadtlandschaften auch Museumsobjekte, deren interaktives Erleben jedoch allein auf die Betrachtung beschränkt bleibt, während ihre Funktion nur über Infotexte oder mediale Verweise vermittelt wird. Eine verbreitete Diskrepanz in der Objektpräsentation wird insbesondere an Musikinstrumenten deutlich, deren Zweck zwar in ihrem Erklängen durch ein physisches Interagieren besteht, die in musealen Kontexten jedoch auch in digitalisierter Form stumm verbleiben. Das Promotionsvorhaben zielt darauf ab, diese akustischen Objekte am Beispiel der Orgel mit evaluierten Methoden in interaktive virtuelle Repräsentationen zu übertragen und sämtliche multimodalen Forschungsdaten offen zur Verfügung zu stellen, die während dieses konsistenten Prozesses generiert werden. Von 21 Orgeln verschiedener Stile und Regionen werden sämtliche Klänge, organologische Objekteigenschaften und raumakustische Parameter erfasst, wofür mittels fotogrammetrischer Methoden auch 3D-Modelle von Objekt und Raum generiert werden. Aus der Analyse von gewonnenen Datensätzen sollen ein spektrales Klassifizierungssystem, Modellierungen des akustischen Verhaltens und organologische Relationskonzepte hervorgehen, wofür auch künstliche neuronale Netze mit Rückführungen auf physikalische Parameter angewandt werden. Diese finden ihren Einsatz auch in der Augmentation und Synthese von Audiodaten zur Optimierung der virtuellen Orgelmodelle, welche wiederum interaktiv sowie in Form von offenen Forschungsdatensätzen über eine systemübergreifende Plattform erreichbar sein werden.

Schlüsselbegriffe:

Organologie; Digitalisierung; Orgel; Audioaufnahme; Fotogrammetrie; virtuelles Instrument; 3D-Audio; akustische Modellierung; Simulation; Auralisierung; künstliche neuronale Netze; offene Forschungsdaten; Digital Humanities; Kulturerbe; Immersion; Interaktivität; Zugang

I. Einleitung

Die digitale Präsentation von akustischen Kulturgütern der öffentlichen und privaten Sammlungen erfolgt aktuell häufig über visuelle Abbildungen, in wenigen Fällen ergänzt von darauf eingespielten Werkaufnahmen, womit lediglich das vorherrschende museale Konzept einer Konservierung und Platzierung akquirierter Sammlungsobjekte repliziert wird. Die zunehmende Einbindung von 3D-Modellen der Objekte bietet hier einen höheren Interaktivitätsgrad, der aktuell jedoch allein auf die Betrachtungsebene beschränkt bleibt und somit kein Erfahren des akustischen Verhaltens eines Objektes ermöglicht. Mit dem technischen Ansatz von samplebasierten virtuellen Instrumenten, die bereits im Feld der Musikproduktion verbreitet sind, lässt sich auch die klangliche Zugangsebene eines Objekts herstellen, während Digitalisierungsvorhaben wie das BMBF-Projekt *TASTEN* [1] demonstrierten, dass auch historische Musikinstrumente nach konservatorischer Einschätzung ohne jegliche Beschädigung aufgenommen werden können. Doch neben den musealen verbleiben auch im öffentlichen Raum verankerte Instrumente zumeist stumm, die jedoch seit mehreren Jahrhunderten kulturelle Handlungen begleiteten und nach wie vor zum Klingen gebracht werden können.

Die Orgel ist ein in Europa weitverbreitetes Musikinstrument, das hervorragend als Untersuchungsgegenstand für die genannten Diskrepanzen bei der Digitalisierung von akustischen Kulturgütern dient. Diese Klasse an Instrumenten vereint materielle und zumeist an prägnante architektonische Bauwerke gekoppelte Objekte mit dem immateriellen Weltkulturerbe des Orgelbaus in Deutschland. Jedes historische und zeitgenössische Instrument dieses Typs stellt eine einzigartige Anfertigung dar und ist zumeist ortsgebunden, während einzelne Instrumente durch den Rückbau von Kirchenräumen und demografischen Binnenwanderungen ebenso gefährdet sind wie durch Naturkatastrophen oder einem zunehmenden Verfall durch ausbleibende Instandhaltungsmaßnahmen. So bietet etwa die Deutsche Stiftung Denkmalschutz eine Übersicht von als gefährdet erkannten Kulturgütern, von denen historische Sakralbauten in strukturschwachen Regionen den größten Teil ausmachen [2]. Doch auch die Aufnahme des Orgelbaus in die repräsentative Liste des immateriellen Kulturerbes der Menschheit durch die UNESCO im Jahre 2017 zeigt auf, dass diese potenziell als mittel- oder langfristig gefährdete Kulturtechnik wahrzunehmen ist, die es zu bewahren und zu fördern gilt [3]. Aktuelle Entwicklungen in Europa wie kriegerische Handlungen lassen in Verbindung mit dem massiv schwindenden Interesse an kirchlichen Einrichtungen mitsamt ihrer Kulturpraktiken eine weiterhin sinkende Bereitschaft zur Erhaltung dieser Instrumente erwarten. Aufgrund der geografischen Verteilung in teils abgelegenen Standorten ist ein Erfahren und gezieltes Beforschen dieser Objekte mit einem recht hohen Reiseaufwand verbunden, nicht zuletzt da diese aufgrund ihrer Größe und akustischen Raumabstimmung nur selten in zentralisierte Sammlungen aufgenommen werden können.

Neben der Entwicklung einer multimodalen Digitalisierungs- und Analysemethodik steht auch die Konzeptbildung eines interaktiven und immersiven Zugangs zu diesen akustischen Kulturgütern und ihren zugehörigen Forschungsdaten im Fokus, welches exemplarisch für die Digitalisierung von komplexen Musikinstrumenten und vergleichbaren Objekten stehen soll. Es entspringt der Intention, die Klänge und das Verhalten dieser Instrumente digital zu sichern und zur nachhaltigen Verfügung zu stellen, worüber hinaus es z. B. möglich sein soll, die virtualisierte Orgel am heimischen Keyboard zu erfahren oder beim Betreten einer Kirche die

speziell darin installierte Orgel mit eigenen Kopfhörern im dreidimensionalen Schallfeld ein beliebiges Werk spielen zu lassen. In gänzlich virtuellen Umgebungen ermöglicht dies bspw. auch Mobilitätsbeeinträchtigten einen barrierefreien Zugang zu bisher unbekanntem Perspektiven, da Orgelpodeste aufgrund der häufig engen Treppenaufgänge i. d. R. kaum erreichbar für diese wären.

Für die Orgel lassen sich in Europa und speziell in Deutschland eine gute Instandhaltung und Betreuung der Instrumente sowie eine gute Dokumentation durch Orgelbauunternehmen mitsamt ihrer Restaurierungsberichte erwarten. Im Laufe des Vorhabens sollen nicht nur jeder einzelne Klang ausgewählter Orgeln aufgenommen, sondern während des Aufnahmevorgangs möglichst viele multimodale Daten generiert und erfasst werden, sodass eine authentische Modellierung des Instruments und weitere z. T. interdisziplinäre Forschungen ermöglicht werden. Ein weiterer zentraler Forschungsansatz des Projekts besteht in der Erkennung von Mustern in den Audiodaten mittels künstlicher neuronaler Netze, die zur Identifizierung von modifizierten Pfeifen und der Kartierung von spektralen Similaritäten innerhalb einer Orgel und zwischen verschiedenen Orgeln eingesetzt werden, die auf physikalische Korrelationen zurückgeführt werden können. Auch die Klassifizierung von Audiodaten der Orgeln soll durch maschinelle Lernverfahren erreicht werden, weiterhin sollen Klänge über neurale Syntheseverfahren rekonstruiert und gänzlich neu generiert werden, die wiederum auf nächstgelegene Dispositionen und spektrale Verwandtschaften untersucht und eingeordnet werden können.

Sämtliche entwickelten Methoden sollen mittels wiederverwendbarer Repositorien und grafischer Oberflächen auch ohne jegliche Expertise sowie mit einfachen technischen Mitteln reproduzierbar sein, insbesondere die Aufnahme und Generierung von virtuellen akustischen Objekten (VAOs). Diese können somit weltweit erzeugt, geteilt und erfahrbar gemacht werden, während die Konsistenz der generierten Datensätze durch den Workflow der Objektaufnahme gewährleistet und damit eine nachhaltige Forschungsgrundlage geboten wird. Es wird Wert darauf gelegt, dass sowohl die Quellcodes öffentlich im GitLab der Universität Leipzig einsehbar sind als auch die Entwicklung der Workflows und das Voranschreiten der Forschung in Form eines Blogs veröffentlicht werden, um eine Plattform des Transfers und aktiven Diskurses über das Projekt zu ermöglichen. Es sollen nur quelloffene und nicht-proprietäre Standards verwendet werden, die auf möglichst vielen Geräten unterstützt werden, darüber hinaus wurden präventive Maßnahmen gegen Datenverluste und aktuell auch in der Wissenschaft zunehmende Verschlüsselungsangriffe (z. B. Ransomware) nach Empfehlungen des BSI vorbereitet [4]. Die virtuellen akustischen Objekte werden auf einer eigens eingerichteten Webplattform zugänglich sein, sowohl das darauf spielbare dreidimensionale Virtualisat als auch die multimodalen Datensätze zu den Instrumenten; auf dieser Plattform können auch externe Objekte von anderen beitragenden Institutionen oder Personen hochgeladen werden, die mit den entwickelten Methoden konsistent generiert wurden, wobei aktuellen und kommenden europäischen Datenschutzstandards entsprochen werden soll. Mit den im Promotionsprojekt geplanten Techniken der Mikrofonierung, Klangsynthese und dreidimensionalen Objekt- und Raumerfassung soll nicht nur ein authentisches Erfahren der Objekte und Objekt-Raum-Relation ermöglicht werden, sondern auch die detaillierte Untersuchung des mechanischen und akustischen Verhaltens. Die entwickelten Tools und Techniken sollen nach Abschluss des Projekts auch auf andere Instrumente übertragen werden können, um die Plattform langfristig um andere Instrumentenarten von Sammlungen, Institutionen und Privatpersonen erweitern zu können, womit sich eine stets erweiternde Quelle für organologische und

akustische Forschungsdaten herausbildet. Im Laufe der Promotion wird sich konsequent darum bemüht, die aus dem Projektfortschritt resultierenden Erkenntnisse in Form von Publikationen zu sichern und zu verbreiten, wofür bereits geeignete Konferenzen herausgearbeitet und Einreichungen zu vorbereiteten Systemen des ersten Arbeitsblocks initiiert wurden.

Die technischen Mittel und infrastrukturellen Ressourcen, welche für eine ideale Realisierung dieses Vorhabens notwendig sind, werden von der Forschungsstelle *DIGITAL ORGANOLGY* am Musikinstrumentenmuseum der Universität Leipzig (MIMUL) und des zugehörigen Universitätsrechenzentrums sowie vom Institut für Systematische Musikwissenschaft der Universität Hamburg zur Verfügung gestellt. Weitere Konzepte zur Sicherung, Distribution und nachhaltigen Verwendung von Forschungsdaten werden mit dem Konsortium für digitalisierte materielle und immaterielle Kulturgüter *NFDI4Culture* [5] ausgearbeitet. Das am Arbeitsort bereits bekannte Umfeld zeichnet sich neben den Musikwissenschaftler:innen, Restaurator:innen und Organolog:innen durch einen stets gewinnbringenden Austausch mit anderen Promovierenden und Forschenden der Fachbereiche Museologie, Digital Humanities, Informatik und der Physik aus. In ebendiesem Umfeld konnten auch wertvolle praktische Erfahrungen und für das Vorhaben hochrelevante Kenntnisse akquiriert werden, so z. B. aus einem vergangenen Digitalisierungs- und einem aktuellen Forschungsprojekt, die vom Bundesministerium für Bildung und Forschung gefördert wurden. Darüberhinausgehend wird auch der Kontakt zu weiteren, für das Vorhaben relevanten Expert:innen gesucht, bspw. zum Organisten und Leipziger Universitätsmusikdirektor Prof. David Timm, mit dem sinnvolle Anspieltechniken für Aufnahmeprozesse evaluiert und objektspezifisches Verhalten antizipiert wurden. Die von ihm betreute Universitätskirche bietet mit der kürzlich eröffneten Schwalbennestorgel zudem ein ideales Objekt für qualitative Untersuchungen über längere Zeiträume hinweg, da in dem neu errichteten Gebäude im Gegensatz zu anderen Kirchen und Konzerthäusern ein konstantes Raumklima gewährleistet wird.

II. Forschungsfrage / Problemstellung / Arbeitshypothese

In der Konzeption des Promotionsprojektes wurden in Rücksprache mit den Betreuern zahlreiche Forschungsfragen auf das Interesse der internationalen Forschungsgemeinschaft und die Realisierbarkeit hin bewertet, woraus sich hohe Erfolgspotenziale zur Beantwortung einiger Kernfragen mit aktuellen Methoden erschlossen. Der erste Komplex widmet sich der Frage nach der idealen Methodik zur dreidimensionalen Virtualisierung von Orgeln sowie vergleichbaren aerophonen Instrumenten und inwiefern diese auf andere akustische Objekte übertragbar ist. Neben der Entwicklung von Strategien zur Klang-, Objekt- und Raumerfassung sind hierbei auch die Generierung und der Umgang mit multimodalen Daten entscheidend, die ausnahmslos konsistent erfasst werden und somit für die Beantwortung der weiteren Forschungsfragen geeignet sind. Die darauffolgende Frage wird von der akustischen Analyse des gewonnenen Materials und den darin identifizierbaren Zusammenhängen bestimmt, insbesondere der Korrelation von spektralen und dispositionalen Klassifikationen mit organologischen und physikalischen Parametern. Eine komparative Analyse innerhalb der Pfeifenreihen einer Orgel und zwischen den Orgeln verschiedener Orgelbaufirmen, Regionen und Bauweisen soll neue Erkenntnisse darüber liefern, inwiefern sich das Material, die Bauart, die Disposition, der akustische Raum oder auch Modifizierungen und Schäden auf spektrale Muster, also auf die klanglichen Produkte einer Orgel

auswirken. Die dritte Forschungsfrage nutzt die hieraus gewonnenen Erkenntnisse, um prognostische und generative Ansätze zu untersuchen und die Ergebnisse in Relation zu realen Orgeldaten zu setzen. Dies betrifft nicht nur die Vorhersage akustischer Verhaltensweisen auf Grundlage der vorliegenden Daten, sondern auch die darauf basierende synthetische Augmentation von Audiodaten zugunsten eines authentischen Virtualisats. Im Speziellen sollen die Initialphasen und Fluktuationen prognostiziert und in den Syntheseprozess implementiert, aber auch modifizierte und beschädigte Pfeifen digital rekonstruiert werden, während experimentelle Ansätze wie Klangfarbentransfers und neurale Synthesetechniken ihre Anwendung finden und die Resultate in die Beantwortung der bestehenden Forschungsfragen eingeordnet werden. Mithilfe der multimodalen Datensätze soll auch evaluiert werden, welche Methoden zum Skalieren von einfachen Stereoaufnahmen auf mehrdimensionale Formate geeignet und welche Ansätze der Dereverberation und Regulation von diffusen und direkten Schallereignissen in den vorliegenden Audiodaten sinnvoll sind. Der vierte und letzte Komplex widmet sich zum einen dem Transfer der virtuellen akustischen Objekte und zugehörigen Datensätze über eine interaktive und partizipative Webplattform, zum anderen der Entwicklung einer App, die das authentische Hören der virtuellen Orgeln in eben jenen Kirchen ermöglicht, in denen das physische Original installiert ist. Diese Anwendungen inkludieren Evaluationsmechanismen und ermöglichen über die Plattform einen fachlichen Diskurs über die zur Verfügung gestellten Datensätze. Somit wird das gesamte Vorhaben auch im Sinne der europäischen Konvention über den Wert des Kulturerbes für die Gesellschaft abgeschlossen, indem digitale Technologien zugunsten eines qualitativ erweiterten, modernen sowie barrierefreien Zugangs zu kulturellen Erzeugnissen weiterentwickelt und eine offene Partizipation ermöglicht werden [6]. Letztendlich soll dieses Projekt auf eine umfassendere Digitalisierung von Kulturgütern mit interaktiven und immersiven Ebenen gerichtet sein, wie aktuell von der EU-Kommission für die Plattform *Europeana* empfohlen, da sich bis November 2021 lediglich 2,47% audiovisuelle und 0,03% dreidimensionale Medien darauf befanden [7].

III. Forschungsstand / eigene Vorarbeiten

Die zielsichere Durchführung des Projektvorhabens wird insbesondere dadurch gestützt, dass für einzelne Teilaspekte bereits umfassende Erkenntnisse vorliegen, welche für die methodische Konzeption, eine wissenschaftlich relevante Zielsetzung und die zeitliche Abschätzung einzelner Aufgaben genutzt wurden. So sind die mit der Orgel zusammenhängenden physikalisch-akustischen Phänomene weitestgehend gut untersucht, insbesondere durch Judit Angster am Fraunhofer-Institut für Bauphysik in Stuttgart [8] und ihren Kollegen [9,10], auch die Kopplungsphänomene wurden ebenso präzise erfasst [11,12] wie die Facetten der Kirchenakustik [13], raumakustische Aspekte der Orgel [14] und Methoden zur Transientenanalyse [15]. Historische und musikästhetische Betrachtungen wurden durch die Arbeiten von Roland Eberlein [16,17] gesichert und werden u. a. durch die Walcker-Stiftung für orgelwissenschaftliche Forschung weitergeführt. Ansätze zur physikalischen Modellierung von Orgelklängen wurden mit konventionellen [18,19] und neuronalen Methoden [20] untersucht, darüber hinaus konnten mittels photogrammetrisch generierter 3D-Modelle und der Finite-Elemente-Methode (FEM) Klänge von historischen Glocken synthetisiert werden [21]. Mit Bezug auf virtuelle Orgeln existieren quelloffene Softwares wie GrandOrgue [22] oder Aeolus [23] sowie proprietäre und kommerzielle Klangbibliotheken wie Hauptwerk [24] oder KONTAKT [25], die jedoch entweder auf der additiven

Synthese oder einfachem Sampling zur Wiedergabe der erfassten Orgelklänge basieren, keine Forschungsdaten zur Verfügung stellen, auf Instrumente z. T. nicht frei zugegriffen werden kann und weitere Diskrepanzen aufweisen. Als wichtige Quelle zur Selektion potenzieller Untersuchungsobjekte dient die Orgeldatenbank *orgbase* [26], welche über 55.000 Orgeln mit Angaben zu Orgelbaufirmen, Standorten, Dispositionen, Mechaniken, Stimmungen und zugehörigen Medien enthält. In der Forschungsstelle Digital Organology des MIMUL bestehen darüber hinaus Zugänge zu weiteren, teils präziseren Orgeldatenbanken, ebenso werden die Daten der von der Forschungsstelle herausgegebenen Plattform *musiXplora* [27] genutzt, um auf verknüpfte Informationen über Orgelbauer, Orgelbaufirmen und weitere relevante Institutionen und Personen zuzugreifen. Hiermit wurden bereits Relationsuntersuchungen durchgeführt und zugehörige Datenbankstrukturen mit Visualisierungskonzepten geschaffen [28]. Es liegen hilfreiche Ansätze für das Datenmanagement [29,30] sowie den nachhaltigen Umgang mit kulturbezogenen und akustischen Forschungsdaten [31] vor, z. B. auch unter Einhaltung der sog. FAIR-Prinzipien [32], ebenso wie Einordnungen und Nutzungskonzepte zu digitalisiertem Kulturerbe [33,34]. Ein interessantes technisches Konzept bot die Einbindung von Plugins auf der EU-geförderten Plattform *PLUGGY* [35], durch die dreidimensional erfasste Klanglandschaften aus Europa hochgeladen, geteilt und dank binauraler Synthese erfahrbar gemacht werden konnten. Die effizientere Integration von virtuellen Instrumenten auf Webseiten wurde mithilfe des WebAssembly-Standards umgesetzt [36] und mündete in einem offenen Framework [37]. Der Fachausschuss für Virtuelle Akustik, Musikalische Akustik und Hörakustik der *Deutsche[n] Gesellschaft für Akustik e.V.* (DEGA) lieferte einen Leitfaden zur Dokumentation von wissenschaftlich orientierten akustischen Aufnahmevorhaben [38] und exemplifizierte das System anhand von Orgelaufnahmen, in deren Kontext eine Stimulus-Datenbank von Werkaufnahmen für virtuelle auditive Umgebungen initiiert wurde [39]. Eine kurze Einordnung der zu beachtenden Aspekte und Probleme des Orgelsamplings wurde von Urban [40] unternommen, das schwedische Projekt *The Organ as memory bank* erweiterte die Dokumentation von Tastendruckparametern auf Orgeln [41].

Relevante Techniken zur Aufnahme von dreidimensionalen Schallfeldern mit dem Ziel einer Bewegung durch virtuelle akustische Umgebungen in 6 Freiheitsgraden (6DoF) wurden u. a. mit Ambisonics-Mikrofonen und einer distanzabhängigen linearen [42] oder einer parametrischen Interpolation [43] erreicht. Auch die instrumentale Klangquellenweite [44] wurde detailliert untersucht, Methoden zur akustischen Lokalisierung wurden probabilistisch [45] oder auf neuronalen Netzen basierend entwickelt [46]. Die Synthese von Abstrahlmustern und raumakustischen Effekten wurde über objektorientierte [47,48,49] Ansätze oder mittels konventioneller Wellenfeldsynthesen [50] realisiert. Zur Erfassung von Raumimpulsantworten (RIR) wurden Sinus-Sweeps optimiert [51], die wiederum über eine generalisierte spektrale Subtraktion [52] oder partielle Resynthese [53] entauscht werden können. Zur Messung diverser Impulsantworten bestehen zahlreiche Methoden, die sphärische Mikrofonarrays einbeziehen [54] und zur Auralisierung des Raums einbezogen werden können [55], während auch Messunsicherheiten wie die des Klarheitsmaßes [56], die Effekte der Raumakustik auf Klassifizierungen [57], Raum-Divergenz-Effekte [58] und Anwendungen maschinellen Lernens zur Berechnung der Raumparameter [59] thematisiert wurden. Zur Evaluation existieren häufig referierte offene Datenbanken, spezieller sind jene, die in Modellen gerasterte Impulsantworten für Schallfeldanalysen [60], räumliche Impulsantworten (SRIR) in 6DoF [61] oder solche zur Validierung akustischer Raumsimulationen [62] beinhalten.

Zu den vielversprechenden Techniken zur Modellierung der Schallausbreitung im Raum gehören die Randlelementmethode (BEM) unter Zuhilfenahme der effizienten Fast-Multipole-Methode [63], die Nutzung von grafischen Prozessoren (GPU) zur Simulation von Raumimpulsantworten über die Image-Source-Methode (ISM) [64] oder die FEM mit Raytracing [65]. Eine perzeptuelle und objektive Evaluation dieser Methoden zur Synthese akustischen Verhaltens in virtuellen 3D-Räumen wurde mit dem Ergebnis durchgeführt, dass Raytracing für die Simulation in großen und ISM in kleinen bis mittleren Räumen geeignet sind [66]. Das Wellenfeld konnte mithilfe direktonaler Raumimpulsantworten (DRIR) analysiert werden [67], in einer komparativen Untersuchung mehrerer Sakralbauten ließen sich jedoch auch konventionelle RIR zielführend anwenden [68]. Es liegen zahlreiche Fallstudien zur akustischen Analyse [69,70,71,72] und Auralisierung [73] von Kirchenräumen vor, in denen auch Simulationen anhand von 3D-Modellen eingesetzt werden [74], die wiederum direkt aus Punktwolken generiert werden können [75]. Versuche einer Dereverberation wurden mittels einer Dekonvolution von Impulsantworten [76], Beamforming und Filtertechniken [77] oder mithilfe von visuell informierten Modellen [78] unternommen.

Für spektrale Klassifikationen wurden Konzepte entwickelt, die auf selbstorganisierenden Karten (SOMs) [79,80], Formanten [81] und dynamischen Formantenfeldern [82], mehrdimensionalen Merkmalsprofilen [83] sowie in neuronale Netze eingespeiste MFCC-Repräsentationen (= Mel-Frequenz-Cepstrum-Koeffizienten) [84] basieren. Multimodale Deskriptoren der Klangfarbe wurden neben berechenbaren spektralen Merkmalen durch Techniken des Music Information Retrieval (MIR) und weiteren physikalischen Korrelaten [85] auch über psychoakustische Kategorien [86] wie der ‚Helligkeit‘ [87,88] und linguistische Untersuchungen [89] eingeordnet. Ein hervorstechender Ansatz zur spektralen Similaritätsanalyse basiert auf der Verwendung eines tiefen neuronalen Netzes, dessen innere Vorgänge jedoch interpretierbar sind und somit Formulierungen mit einer höheren wissenschaftlichen Signifikanz ermöglicht [90]. Die Theorie und Anwendung von maschinellen Lernverfahren in konventionellen akustischen Anwendungen wurde übersichtlich von Bianco [91] zusammengetragen und ausgeführt, mit *L3DAS* (= „Learning 3D Audio Sources“) wurde zudem ein jährlicher Wettbewerb zur Anwendung dieser Verfahren auf die dreidimensionale Audiosignalverarbeitung initiiert [92].

Die neuronalen Klangrepräsentationen wurden allgemein [93] und für spezifische Architekturen [94] sowie anhand von perzeptuellen Dimensionen [95] evaluiert, weitere nutzen u. a. eine lautstärkeinvariante Vektorquantisierung [96] oder solche wie das sog. *Neuralogram* [97]. Es wurden zahlreiche Versuche zu neuronalen Syntheseverfahren ausgeführt, die als Basis entweder latent variable Modelle wie konditionale und variationale Autoencoder (CVAE/VAE) [98,99], Generative Adversarial Networks (GAN) [100,101], rekurrente neuronale Netze (RNN) [102], die differenzierbare Wavetable-Synthese (DWTS) [103], neurales Waveshaping [104] oder die Anwendung der differenzierbaren digitalen Signalverarbeitung (DDSP) [105] verwenden. Die Intensität der auftretenden artifiziellen Verzerrungen in den generierten Klangresultaten kann mithilfe der Fréchet-Audio-Distanz (FAD) [106] quantifiziert werden, darüber hinaus können mittels verschiedener Techniken des Upsamplings inhärente Artefakte geglättet [107] oder niedrigdimensionale Aufnahmen anhand von räumlichen Punktwolken augmentiert werden [108]. Die Möglichkeit zum Transfer von Klangfarben wurde in mehreren Ansätzen mittels neuronaler Netze verschiedener Architekturen [109,110,111] untersucht und konnte bereits in Echtzeit-Anwendungen inklusive grafischer Oberflächen gelingen [112].

Es wurden mehrere quelloffene Toolboxes und Repositorien entwickelt, die ihren Fokus auf die Analyse und Merkmalsextraktion von Audiodaten [113,114] setzen, frequenzanalytische Ansätze erweitern [115], der konsistenten Erhebung und Auswertung akustischer Daten [116] sowie raumakustischen Analysen [117,118] und Simulationen dienen [119]. Auch die Umwandlung und Verwendung von Spektrogrammen für eine direkte Verwendung in neuronalen Netzen ist mittels GPUs gelungen und resultierte in einem offenen Framework [120], während sämtliche modernen Algorithmen für Audioanalysen umfangreich zusammengetragen und theoretisch durchleuchtet wurden [121]. Die hierin aufgeführte Literatur spiegelt den aktuellen Forschungsstand wider und bietet zahlreiche methodische Ansätze, die für das Promotionsprojekt als zielführend erachtet werden und aufgrund ihrer direkten Reproduzierbarkeit ein hohes Erfolgspotenzial bieten.

Die gesamte Planung des Projektes profitierte zudem nicht nur von den Erkenntnissen meiner Bachelorarbeit, in der ich mich der Klangsynthese mittels künstlicher neuronaler Netze (KNN) widmete, sondern auch von den multidisziplinären Erfahrungen der Masterarbeit, in der die akustische Virtualisierung und funktionale Augmentation einer elektropneumatischen Kinoorgel im MIMUL fokussiert wurde. Für diese Arbeit habe ich 44 Register und 31 Effekte der Orgel in zahlreichen Spielvarianten aufgenommen, analysiert und daraus ein virtuell spielbares Instrument mit exakt nachmodellierter Schaltungslogik generiert. Darüber hinaus habe ich einen Zugang zur digitalen Kontrolle des Instruments über MIDI entwickelt, mit dem sich nicht nur automatisierte Aufnahmen und eine kabellose Fernsteuerung über eine App realisieren ließen, sondern auch nicht-invasive Modifikationen der physischen Klangerzeugung durch präzise Eingriffe in die zeitlichen Verhältnisse der Signalgebungen ermöglichte. Im Besonderen führte jedoch das TASTEN-Projekt (BMBF) zum Forschungspfad der Virtualisierung akustischer Instrumente, da mir hierin die mehrjährigen Aufgaben einer Aufnahme von 36 historischen Tasteninstrumenten, der Generierung von virtuell spielbaren Abbildern sowie der Erfassung organologischer Daten und ihrer redaktionellen Einbindung in die lexikalische Plattform *musiXplora* anvertraut wurden. Im Laufe dieses Projekts bildeten sich fundierte Methoden und Strategien zur Aufnahme von tastengebundenen historischen Musikinstrumenten heraus, gleichzeitig wurden diverse Tools entwickelt, welche die Analyseprozesse, den Schnitt von Einzeltönen aus Rohaufnahmen, zahlreiche Klassifizierungen und die Kontrolle der Datenkonsistenz automatisieren; diese Strategien und Tools sollen im Laufe des ersten Projektjahres erweitert, zur Publikationsreife geführt und veröffentlicht werden. Darüber hinaus konnte dank der zahlreichen Transferveranstaltungen dieses Digitalisierungsprojektes eine permanente Reflexion und ein kritischer Diskurs mit Expert:innen erfahren werden, welcher als sehr wertvoll für die Optimierung einzelner Prozesse erachtet wurde. Im DISKOS-Projekt (BMBF) erweiterte ich die entwickelten Strategien um eine dreidimensionale audiovisuelle Objekterfassung, ein offenes sowie konsistenteres Datenmanagement und eine Kompatibilitätsmaximierung digitalisierter Toninformationsträger zur Einbindung in virtualisierte mechanische Abspielgeräte. Hier konnte ich bereits erfolgreich 3D-Modelle generieren und diese mit den daraus gewonnenen akustischen Modellierungen in eigens entwickelte AR-Anwendungen einbinden. Schließlich ermöglichte mir ein sechsmonatiges Exposé-Stipendium zum einen das Studium relevanter Literatur und das Schließen dadurch aufgezeigter Wissenslücken, zum anderen das Gespräch mit Expert:innen über thematische Teilaspekte gefolgt von der finalen Konzeption des Promotionsvorhabens mit mehrfach überdachten Arbeitskonzepten.

IV. Methodik / vorläufige Gliederung

Das gesamte Vorhaben kann in die vier großen Arbeitsblöcke der Datenakquise, Analyse, Augmentation/Synthese und dem Transfer unterteilt werden, wofür jeweils mehrere Arbeitspakete definiert sind. Nach Abschluss einiger zusammenhängender Arbeitspakete erfolgen Verschriftlichungen und Publikationen, um Resultate und Folgeansätze sowohl mit den Betreuern als auch mit der Forschungsgemeinschaft diskutieren zu können. Der erste Arbeitsblock des Promotionsprojekts wird von der Generierung der Forschungsdaten und der parallelen Optimierung von hierfür verwendeten Programmen zur konsistenten Datenerfassung und Datenverwaltung bestimmt. Im ersten Arbeitspaket dieses Blockes wird die Selektion von 21 aufzunehmenden Orgeln nach fest definierten Bewertungskriterien vorgenommen, die sich insbesondere auf das Erfolgspotenzial und die kulturelle Bedeutung eines Objekts ausrichten. Diese kann durch die stadtgeschichtliche Einordnung, die Verbindung zu bekannten Organist:innen, der Bekanntheit des Orgelbauers oder der Feststellung außergewöhnlicher Objekteigenschaften aufgeschlüsselt werden, während der allgemeine Zustand bspw. anhand von Berichten der zurückliegenden Restaurierungsmaßnahmen erkennbar ist. Der zeitliche Aufwand einer Objektaufnahme kann antizipiert werden, wenn die Anzahl der Register und Pfeifen aus Datenbanken abgerufen, der Nachhall des Raums anhand der Gebäudedimensionen geschätzt und die potenzielle Häufigkeit von verkehrs- und baustellenbedingten Außengeräuschen festgestellt werden, zu deren Einschätzung bereits Tools entwickelt wurden. Zur Beurteilung der Qualität vorliegender Objektdaten, die etwa in Datenbanken oder Restaurierungsberichten von Orgelbauern eingesehen werden können, werden zumindest die Lage der Pfeifenreihen im Orgelsystem und die Angabe einiger Messdaten vorausgesetzt, um im Analyseprozess Vergleichswerte für bauliche und physikalische Parameter zu gewährleisten. Eine zuverlässige Betreuung der Instrumente - z. B. durch das Kantorat, Organist:innen oder Orgelbauern – ist ebenso entscheidend, um neben diversen Rückfragen auch feste Termine für die Objekterfassung unter Garantie keiner konkurrierenden Veranstaltungen zu vereinbaren. Es wurden bereits bedeutende Orgeln diverser Orgellandschaften aus Spätgotik, Barock, Romantik und Moderne eingeplant, darunter solche, die von Orgelbauern wie Silbermann, Schnitger, Walcker und Rieger gebaut wurden, in Verbindung zu Persönlichkeiten wie J. S. Bach und Max Reger stehen oder die weltweit ältesten Pfeifen enthalten.

Nach der Selektion der Orgeln erfolgt die Vorbereitung zur Aufnahme, wozu im ersten Schritt nach der Terminvereinbarung die Erstellung eines Aufnahme- und Datenverwaltungsplans anhand der zuvor akquirierten Informationen gehört. Im Aufnahmeplan wird das sequenzielle Vorgehen während des Aufnahmeprozesses festgelegt und als Checkliste vorbereitet, d. h. welche Register, Manuale und Geräusche in welcher Abfolge und Kombination aufgenommen werden. Im Datenverwaltungsplan wird kohärent dazu festgelegt, in welcher Datei-/Ordnerstruktur und unter welcher Benennung die generierten Daten abgelegt werden. In der Vorbereitung einer Aufnahme sollen auch mögliche Probleme antizipiert und für die Planung bedacht werden, wozu etwa eine fehlerhafte Mechanik oder beschädigte Pfeifensysteme gehören. Im letzten Schritt der Vorbereitung wird die Vorlage der digitalen Protokollierung des Aufnahmeprozesses vorbereitet, sodass neben unvorhergesehenen Änderungen im Ablauf auch sämtliche Auffälligkeiten wie Resonanzeffekte, mechanische Probleme, Stimmungsschwankungen oder zusätzlich hinzukommende Klänge protokolliert und in den Datensatz eingebunden oder während des Bearbeitungsprozesses konsultiert werden können.

Das zweite Arbeitspaket, die dreidimensionale Objekt- und Raumerfassung, wird für jede der ausgewählten Orgeln anhand der zuvor festgelegten Aufnahmestrategie durchgeführt. Im ersten Schritt wird hierbei die Mikrofonierung mit softwaregestützten Analysen vorgenommen, um ein möglichst ausgewogenes spektrales Klangbild ohne Phasenauslöschungen außerhalb von Wellenknoten zu gewährleisten. Zur Erfassung eines umfassenden 3D-Schallfeldes im Sinne des Projektes, durch das sich letztendlich in sechs Freiheitsgraden (6DoF) bewegt werden kann, werden im Kirchenraum fünf autarke kugelförmige Mikrofon-Arrays mit jeweils 19 MEMS-Kapseln nach einem Quincunx-Schema verteilt. Darüber hinaus sollen ein Paar an kardioiden Stereomikrofonen möglichst im direkten Schallfeld der Orgel sowie ein ebenfalls kardioides Paar an der Spielposition eingerichtet werden, die beide einer ORTF-Systematik folgen. An jedem einzelnen Mikrofon werden nun die Abstände zu den Wänden sowie zu Decken und Boden mittels eines einfachen Laser-Entfernungsmessers protokolliert, um die Messung im 3D-Modell zu validieren. Die Mikrofon-Arrays werden direkt über USB mit einer jeweiligen Auflösung von 48kHz/24bit an einen Lüfterlosen Rechner im Auditorium angeschlossen, während die vier Stereomikrofone über eine USB-Audioschnittstelle mit 384kHz/32bit an ein passiv gekühltes Notebook angeschlossen werden; sämtliche Einstellungen erfolgen somit digital und die Geräte werden zur Vermeidung stromnetzbedingter Schwankungen im Audiosignal an eine Stromversorgung mit Spannungsstabilisator angeschlossen, welche alternativ durch eine sichere akkubasierte Lösung ersetzt werden kann. Beide Rechner werden vor jeder Aufnahme über einen gemeinsamen Kommunikationspunkt synchronisiert, um gleichgeschaltete Zeitstempel in den Audiodaten zu garantieren. Für die spätere Messung der Impulsantworten wird anschließend ein linearer Lautsprecher am Orgelpodest platziert, der die verschiedenen beschaffenen abstrahlenden Messsignale über das Interface vom Notebook erhält.

Im nächsten Schritt erfolgt die visuelle Erfassung des akustischen Raums mittels Fotogrammetrie und einem LiDAR-Sensor auf zwei Höhenstufen, um in einem späteren Arbeitsblock ein texturiertes 3D-Abbild des Kircheninnenraums zu generieren. Es soll hierbei darauf geachtet werden, dass die Positionen sämtlicher Mikrofone im dreidimensionalen Raum verortbar sind, was präzisere akustische Modellierungen ermöglicht. Anschließend wird die Orgel selbst fotogrammetrisch mit hohem Detailgrad erfasst, insbesondere die Pfeifen und der Spieltisch der Orgel, um exemplarische Messdatenvergleiche vorzunehmen und das interaktive virtuelle akustische Objekt auch visuell korrekt nachbilden zu können. Wo es möglich ist, sollen auch Pfeifenreihen im Inneren erfasst werden, die Methoden hierfür müssen jedoch noch bewertet werden, wobei auch die Zusammenführung endoskopischer Videografien für unzugängliche Stellen in Betracht kommt, etwa um akustisch ermittelte Verdachte auf Modifikationen visuell zu fundieren.

Nach Abschluss der visuellen Raum- und Objekterfassung erfolgt die Messung der akustischen Impulsantworten des Raumes mithilfe des linearen Lautsprechers und sämtlicher Mikrofone, wobei auch direktionale und ggf. binaurale Raumimpulsantworten (BRIR) generiert werden. Anschließend folgt die Kalibrierung eines münzgroßen Gyrosensors, welcher vor dem Anspielen auf einer Taste platziert wird, um mechanische Interaktionsdaten wie die Geschwindigkeit und Tiefe des Tastenanschlags zu gegebenen Zeitpunkten während der Aufnahme zu erfassen. Dieser Sensor wird über einen USB-Mikroprozessor an den zweiten Rechner angeschlossen und zeitlich mit den Audiodaten abgestimmt, um Fehlinterpretationen von Korrelationen im Datensatz auszuschließen. Im Auditorium werden ein digitales Thermometer, Barometer und Hygrometer über einen

weiteren Mikroprozessor an den ersten Rechner angeschlossen, um die klimatischen Bedingungen der Schallereignisse in vordefinierten Zeitschritten zu erfassen.

Mit der Kalibrierung des Aufnahmesystems und dem Öffnen der Protokolleingabe kann die eigentliche Tonaufnahme der Orgel nach dem entsprechenden Plan beginnen. Während einer konstant laufenden Aufnahme wird hierbei jeder Ton mit einem Druck auf den Sensor angeschlagen und gewartet, bis der Klang vollständig und ohne auftretende Störgeräusche verklungen ist, was in Echtzeit anhand eines spektralen Monitoring-Systems kontrolliert wird. In diesem System wird unterschieden zwischen Außen-, Innen-, Objekt- und Körpergeräuschen; ob eine Aufnahme wiederholt werden muss, wird je nach Intensität, Frequenzbereich und zeitlichem Auftreten während des Klingens oder Ausklingens beurteilt.

Je nach Höhe und Stabilität des Tons variiert zudem die Haltezeit, um einen ausreichenden Informationsgehalt zu erfassen. Darüber hinaus werden, je nach zeitlicher Planung nur an exemplarischen Pfeifenreihen, weitere Spieltechniken wie Staccati gespielt, um unter Zuhilfenahme der mechanischen Interaktionsdaten Erkenntnisse über das Verhalten eines Objekts wie die An- und Absprache einer Pfeife zu gewinnen. Nach jeder Pfeifenreihe wird die konstant abgelaufene Aufnahme planmäßig gesichert und anschließend eine neue begonnen, bis alle Pfeifen und Effekte der Orgel aufgenommen wurden. Im Anschluss daran werden sämtliche Geräusche des Instruments mit den Mikrofonen an der Spielposition erfasst, wozu u. a. die Druck- und Abfallgeräusche der Tasten aller Manuale und der Pedale, die Bedienung der Registerzüge und Koppeln sowie die Aktivierung und Deaktivierung des Motors gehören. Sofern die Orgel über eine MIDI-Schnittstelle verfügt, werden Einspielungen mit laufenden Aufnahmen ausgeführt, um die MIDI-Datei in einem späteren Arbeitspaket in die virtuelle Orgel einzuspeisen und das synthetisierte Resultat mit der korrespondierenden Aufnahme zu vergleichen. Nach Abschluss des Aufnahmevorgangs wird die Vollständigkeit und Konsistenz der erfassten Daten geprüft, wofür ein Abgleich des Datenplans mit den gesicherten Daten durchgeführt wird, gefolgt von einer automatisierten Analyse des Audiomaterials mittels eines Tools, um potenziell übersehene Störgeräusche zu identifizieren und eine Neuaufnahme der betroffenen Töne zu initiieren. Anschließend werden sämtliche Daten planmäßig in Backups gesichert und indiziert, wofür sowohl ein Cloudspeicher als auch gespiegelte Festplatten zur Verfügung stehen. Die Objekterfassung wird in drei kumulative Phasen aufgeteilt, sodass die aus jeder Phase hervorgehenden Daten analysiert, evaluiert und angewandte Erfassungsprozesse ggf. optimiert werden können. Die Erfassung der Schwalbennestorgel in der Leipziger Universitätskirche soll mit höherer Detailstärke erfolgen, um Annahmen mit qualitativen Belegen stützen und verifizieren zu können, die aus Datenatzanalysen anderer Orgeln getroffen werden. Die Daten dieses Objektes werden für die geplanten Ansätze der Simulation, Augmentation und Synthese primär herangezogen, wonach die anderen Orgeln mit absteigender Datenmenge folgen.

Im zweiten großen Arbeitsblock werden im Wesentlichen die zuvor generierten Datensätze genutzt, um die Hypothesen anhand umfassender Analysen zu einem empirischen Beweis zu führen. Die Vorarbeit besteht hierbei zunächst in der Subtraktion der jeweiligen Latenzen und in der Extraktion von sämtlichen verwertbaren Parametern aus dem Audiomaterial, wozu etwa die ermittelte Fundamentalfrequenz, Spektral- und Lautstärke-daten, Phaseninformationen, Partialtonfluktuationen oder die Dekomposition der Transienten gehören. Im darauffolgenden Schritt werden die erzeugten 3D-Raummodelle zur weiteren Verwendung bearbeitet,

reduziert und mit dem Orgelmodell zusammengeführt. Die Modellgenerierung aus fotogrammetrischen und LiDAR-Daten sowie deren Bearbeitung erfolgen über die quelloffenen Softwares *Meshroom* [122] und *Blender* [123]. Anschließend werden die damit erweiterten Datensätze auf ihre Korrektheit und Konsistenz hin geprüft, sodass keine falschen Zuweisungen auftreten oder verfälschte Daten interpretiert werden. Sämtliche aus diesen Schritten hervorgehenden Datensätze können als Objektdatei durchsucht oder hieraus in gewünschtem Umfang im JSON-Datenformat exportiert werden, sodass einzelne Informationen zu einem Instrument über vordefinierte Klassenattribute abgerufen werden.

Nach dieser Datensatzvorbereitung sollen im zweiten Arbeitspaket des Blocks ideale Methoden evaluiert werden, um die spektrale Similarität von Pfeifen bestimmen und kartografieren zu können, wofür zunächst die Direktschall-Audiodaten herangezogen werden. Hierbei sollen neben konventionellen MIR-Methoden auch SOMs und weitere KNNs zur Mustererkennung eingesetzt werden, deren Ausgabe sowohl anhand der korrespondierenden Bau- und Messdaten als auch auditiv evaluiert werden. In diesem Vorgang sollen auftretende Divergenzen über die Gewichtung der einzubeziehenden Parameter aufgelöst und anhand von Testdaten geprüft werden, doch auch von KNNs generierte Repräsentationen der instrumentalen Audiodaten sollen hierbei evaluiert werden. Für die Erstellung von Testdatensätzen werden klassifizierte Orgelklänge genutzt, die aus frei verfügbaren Quellen bezogen und in Richtung der eigenen Datensatzstruktur konsistent vorverarbeitet wurden. Ziel dieser Analyse ist die Erstellung einer neuen taxonomischen Systematik für Klangfarben von Orgeln, die von einer reproduzierbaren Methodik und empirischen Datenauswertung gestützt wird, um künftige materielle und digitale Erzeugnisse klassifizieren und relational einordnen zu können. Ein weiteres Ziel dieser Methodik liegt darin, modifizierte, beschädigte oder reparierte Pfeifen anhand ihres Klanges zu identifizieren, während die Verwendung von SOMs neue Aufschlüsse über die akustischen und organologischen Relationen verschiedener Orgelbaufirmen und geografischen Regionen bieten kann.

Im dritten Arbeitspaket des Analyseblocks soll mithilfe der Aufnahmen der Mikrofon-Arrays, der Impulsantworten und der 3D-Modelle die Schallausbreitung untersucht werden, um Hörpositionen im virtuellen akustischen Raum und die Audioausgabe für definierte Positionen im Raum mit der Modellvorhersage vergleichen zu können. Dieser Schritt ist auch für die spätere Auralisierung von augmentierten Audiodaten von hohem Nutzen: die in diesen Audiodaten abgebildeten Raumeffekte müssen den räumlichen Dimensionen entsprechend logisch verlaufen, um psychoakustische Divergenzeffekte und eine Hemmung der auditiven Immersion zu minimieren.

Das vierte Arbeitspaket dient der Vorbereitung von Datensätzen zur Simulation von Schwankungen in einem späteren Arbeitsblock. Die vorliegenden Audiodaten werden hierfür zunächst visuell, auditiv und datenbasiert auf Muster untersucht, an denen diese erkennbar wären, wofür zuerst die Pfeifenklänge herangezogen werden, die bereits während der Aufnahme aufgefallen waren und entsprechend protokolliert wurden. Es soll ein automatisches Erkennungsverfahren entwickelt werden, mit dem sich Fluktuationen in den Audiodaten erkennen, klassifizieren und in ihrem Umfang beschreiben lassen.

Im fünften Arbeitspaket erfolgt eine organologische Untersuchung der jeweiligen Objekte, indem die Netzwerke und Relationen der Orgelbaufirmen, Ereignisdaten, Bezugsregionen und Wirkungsorte unter Zuhilfenahme der musiXplora-API analysiert werden. So soll bspw. in Erfahrung gebracht werden, ob und in welchen

Graden die erfassten Objekte miteinander in Beziehung stehen könnten, etwa über konvergente Ausbildungsstätten, Mitarbeitende, beauftragte Materialwerkstätten oder Verwandtschaftsverhältnisse. Die Ergebnisse dieser Analyse sollen visualisiert werden und dem Zweck dienen, akustische Similaritäten der Orgeln auf gemeinsame Ereigniskreise zurückzuführen und die visuelle Objektkartierung zu erweitern. Auf Grundlage der muSiXplora-Datenbank soll ein Tool zur Relationsanalyse und Kartografierung entwickelt werden, sodass auch künftig erfasste Objekte auf der Plattform für virtuelle akustische Objekte automatisch in Beziehung zueinander gesetzt und somit neue organologische Untersuchungen angeregt werden.

Der dritte Arbeitsblock wird von generativen Ansätzen bestimmt, zu denen insbesondere die Augmentation der erzeugten Daten gehört, um fehlende Orgelklänge entweder zu rekonstruieren oder gänzlich neue Klänge zu erschaffen, wofür sich der Einsatz von KNNs als äußerst sinnvoll erwiesen hat. Zum ersten Arbeitspaket dieses Blockes gehört die Evaluation geeigneter Netzwerktopologien für verschiedene generative Aufgaben durch die Erprobung und komparative Analyse bestehender Techniken. In der ersten Anwendung sollen für jede aufgenommene Pfeife neue Initialgeräusche generiert werden, wofür rasch eine hohe Menge an Eingabedaten akquiriert und Initialgeräusche mithilfe adaptiver Dekompositionstechniken extrahiert werden können. Sofern die Resultate insgesamt erfolgreich auditiv und spektralanalytisch evaluiert wurden, werden die räumlichen Reflexionen der Ausgabedaten über die entsprechende Impulsantwort als Faltungshall berechnet und das Ergebnis an die Halteschwingungen der korrespondierenden Pfeifen gekoppelt. Die Evaluation soll anhand von Testdaten durchgeführt werden, in denen weitere aufgenommene Initialgeräusche derselben Pfeifen mit den synthetisch generierten verglichen werden, insbesondere in Bezug auf die Lautstärkeverteilung und ihre zeitlichen Verläufe. Der gesamte methodische Komplex soll anschließend auf die Audiodaten der übrigen Mikrofone angewandt werden, speziell der Mikrofon-Arrays unter Einbezug der direktionalen Impulsantworten. Dieses Vorhaben dient dem Zweck, die Ansprache der Orgel in einem virtuellen akustischen Objekt auf einer Grundlage von mehrdimensionalen Mustererkennungen nahezu unendlich häufig nachahmen zu können, da an einer realen Orgel aufgrund der chaotischen Klanganregung keine aufeinanderfolgende Ansprache in gleicher Beschaffenheit erklingt.

Im zweiten Arbeitspaket sollen Fluktuationenmuster in den Halteschwingungen anhand des Audiomaterials und der Modelle identifiziert und prognostiziert werden, wofür die im Analyseblock ermittelten Korrelationen zur Generierung von Eingabe- und Testdaten verwendet werden. Für diese Aufgabe müssen Ereignisse in einer Zeitserie in ihrer Dauer und Intensität bzw. spektralen Auswirkung erkannt und vorhergesagt werden. Das prädiktive Modell soll zunächst auf Basis einer multivariaten Regression mithilfe eines rekurrenten neuronalen Netzes (RNN) gebildet werden, wofür in den Trainingsdaten pro Orgelklang die jeweiligen Pfeifenmaße, die Distanz zu anderen Pfeifen, die Eintrittszeitpunkte und Zeitfenster der Fluktuationen sowie die korrespondierenden Audiodaten mitsamt der extrahierten Spektraldaten vorliegen. Die Testdaten bestehen aus dem Netzwerk unbekanntem verkürzten Klängen, deren weiterer Verlauf mitsamt den auftretenden Schwankungen vorliegen und die es zu prognostizieren gilt. Ziel ist es, Fluktuationen über die Aufnahmedauer eines Samples hinaus nachvollziehbar zu simulieren, indem allein die in den Audiodaten klassifizierten Ereignisse zur Extrapolation genutzt werden.

Im dritten Arbeitspaket steht die systematische Augmentation der vorhandenen Audiodaten im Fokus, indem neue Orgelklänge mithilfe der neuronalen Klangsynthese generiert und durch die im zweiten Arbeitsblock entwickelten Analysemethoden eingeordnet werden sollen. Dies dient zum einen einer Prüfung der spektralorientierten Systematik und ihrer Rückführung auf physikalische Parameter, im Wesentlichen sollen jedoch neue Methoden evaluiert werden, um auf Grundlage einer mehrdimensionalen Mustererkennung in instrumentalen Audiodaten verwandte Klänge generieren und in ein stimmiges Gesamtinstrument integrieren zu können. Es soll auch erforscht werden, ob und inwiefern sich die Ergebnisse an bestimmte spektrale Muster einer Pfeifenreihe annähern lassen, um neue Daten für die mögliche Rekonstruktion fehlender oder modifizierter Klänge zu erhalten. Eine bereits bekannte Herausforderung dieses Arbeitspaketes liegt darin, sowohl die logische Verteilung von Phaseninformationen als auch Diskontinuitäten im spektralen Verlauf auszugleichen, wofür Ansätze einer Resynthese auf Basis der Ausgabedaten evaluiert werden sollen.

Im vierten Arbeitspaket sollen mit den zuvor gewonnenen Erkenntnissen fehlende, beschädigte oder modifizierte Orgelklänge mit generativen Methoden rekonstruiert und in bestehende Instrumente integriert werden. Hierbei besteht nicht die Behauptung, im Produkt einen physikalisch korrekt rekonstruierten Klang generieren zu können, sondern Orgelklänge mit einer Methodik zu ergänzen, die nicht auf simplen Tonhöhenverschiebungen der Nachbartöne basiert und an deren Resultat sich anhand der entwickelten Analysemethoden dennoch eine sinnvolle Integration erkennen lässt. Sofern geeignete Messdaten vorliegen, sollen auch Konzepte einer Synthese durch die physikalische Modellierung angewandt und verglichen werden.

Der vierte große Arbeitsblock behandelt den öffentlichen Transfer der Datensätze und VAOs sowie die Umsetzung der Zugangskonzepte. Die Programmierung der virtuellen Instrumente vollzieht sich im Wesentlichen über einen hybriden Ansatz auf Grundlage des Samplings und der Simulation akustischen Verhaltens. Hierfür wird das quelloffene Framework *JUCE* [124] verwendet und erweitert, weiterhin soll ein offenes Format für VAOs entwickelt und zur Erweiterung freigegeben werden, in dem sämtliche Datensätze integriert und abgerufen werden können. Zu diesen zählen neben den Audiodaten selbst die organologischen und akustischen Informationen, ebenso die zuvor berechneten Simulationenmuster und ggf. reduzierte 3D-Modelle. Zur direkten Nutzung dieses Formats erfolgt die Programmierung einer einfachen Oberfläche, mit der sich VAOs importieren, spielen und sämtliche darin kodierten Informationen auslesen lassen. Dieses kann auf gängigen Betriebssystemen sowie in App- oder Webumgebungen genutzt werden, da die Objekte systemunabhängig dekodiert und mit dem WebAssembly-Standard bei hoher Performanz eingebettet werden können. Hierin können auch Einspielungen im MIDI-Format auf dem VAO abgespielt werden, eine solche Datei kann jedoch auf anderen Orgeln eingespielt oder digital angefertigt worden sein und somit nicht zur vorliegenden Disposition passen. Zur Lösung soll ein Konverter greifen, der die ursprünglich kodierten Pfeifenreihen anhand der neuen Systematik auf die als nächstverwandt klassifizierten Register überträgt, gleichsam werden auch die Effekte und Umfänge der Manuale in den Konvertierungsprozess einbezogen. Grundsätzlich können sowohl die Oberfläche als auch die zuvor generierten virtuellen Instrumente in andere quelloffene sowie in proprietäre Standards exportiert und heruntergeladen werden, was bspw. aufgrund der verbreiteten Verwendung von VST-Instrumenten als sinnvoll erachtet wird. Innerhalb des Programms kann die vorliegende Stimmung einer Orgel oder einzelner Pfeifenreihen nach Belieben geändert werden, wofür eine eigene Datenbank als Referenz

eingerrichtet wird, die einen Großteil aller bekannten Stimmungssysteme enthält. Mit der Entwicklung dieser Abspieloberfläche und des Formats kann auf sämtliche VAOs über die meisten Endgeräte überall und jederzeit interaktiv oder datenbasiert zugegriffen werden.

Im zweiten Arbeitspaket soll schließlich die Webplattform entworfen werden, auf der sich die virtuellen Instrumente hochladen, spielen, herunterladen und zugehörige Datensätze einsehen lassen. Diese Plattform wird auch für andere kooperierende Institutionen und Personen geöffnet, die ihre Sammlungen oder einzelne Instrumente nach den entwickelten Standards virtuell zur Verfügung stellen möchten. Die Erreichbarkeit von Objekten und Daten auf dieser Plattform wird durch ein persistentes Strukturierungsprinzip gewährleistet, welches auf der Vergabe von Identifikatoren für Institutionen oder Personen, den von ihnen hochgeladenen Objekten und sämtlichen darin enthaltenen Objektdatensätzen basiert. Jede beitragende Instanz erhält mit dem Kürzel eine eigene repräsentative Seite, auf der die von ihnen hochgeladenen Objekte übersichtlich erreichbar sind. Auch jedes Objekt erhält eine eigene Seite, auf der das virtuelle Instrument in gewünschtem Umfang geladen und gespielt werden kann, während weitere Schaltflächen auf die Datensätze, die zugehörigen Seiten, das Abspielen von digitalisierten Toninformationsträgern und auf die Downloads des virtuellen Instruments in sämtlichen verfügbaren Formaten verweisen. Zu weiteren Seiten gehören solche, die eine Diskussion über das Objekt oder Anfragen zu weiteren Daten, die Verlinkung weiterer Medien, das Anzeigen von Beiträgen zum Objekt oder die Auflistung von Veranstaltungen ermöglichen, in denen es gespielt, präsentiert oder im Fokus eines Vortrages steht. Das virtuelle Instrument soll hierbei je nach Bandbreite und Rechenleistung in verschiedenen Qualitätsstufen spielbar sein, wozu auch die erfassten 3D-Schallfelder gehören, die über den offenen Standard MPEG-H 3D Audio und der binauralen Synthese erklingen können [125]. Für diese Plattform sollen zunächst ausschließlich die kooperierenden Institutionen freigeschaltet werden, wobei bestimmte Evaluationsmechanismen eingebunden sind; nach Abschluss des Projekts soll die Plattform auch für andere Beitragende geöffnet und um andere Instrumententypen erweitert werden. Die Idee hinter dieser Plattform ist weiterhin, dass diese auch für künftige Digitalisierungs- und Forschungsprojekte genutzt werden kann.

Das dritte Arbeitspaket des Transferblocks beinhaltet die Programmierung einer mobilen App, über die alle VAOs an ihren physischen Standorten über einfache Kopfhörer mit vorliegenden Werken bespielbar sind. Dies soll dank der Synthese von 3D-Schallfeldern mit einem überaus realistischen Klang ermöglicht werden, die eine zunehmend verbreitete Anwendung finden, da beinahe alle aktuellen Kopfhörer oder mobilen Betriebssysteme bereits diverse 3D-Audiostandards unterstützen. Auch hier soll der Standard MPEG-H 3D Audio verwendet werden, während die Einspielungen im MIDI-Format vorliegen und vorgespicherte oder direkt herunterladbare VAOs adressieren. Ziel dieser Anwendung ist es, zumeist still verbleibende akustische Kulturgüter im öffentlichen Raum authentisch zum Klingen zu bringen, sodass etwa beim Betreten einer Kirche während einer Stadtbesichtigung auch die dort installierte Orgel auf Wunsch und in vollem Umfang über eine Smartphone-Anwendung virtuell erklingen kann.

Im vierten Arbeitspaket soll exemplarisch geprüft werden, inwiefern sich die audiovisuellen VAOs für die Integration in einer Augmented-/Virtual-Reality-Umgebung (AR/VR) eignen. Da die 3D-Raum-/Objektdaten bereits im hierfür verwendbaren Format vorliegen und das Verhalten der VAOs in der Sprache C++ beschreibbar ist, können die im Projekt erhaltenen Daten problemlos eingebunden werden.

Als VR-Entwicklungsumgebung soll die quelloffene *Open 3D Engine* (O3DE) [126] verwendet werden, während für die Implementierung und die AR-Entwicklung der ebenfalls quelloffene Standard *OpenXR* [127] genutzt werden soll. Mit der über bereits bekannte Schnittstellen realisierbaren AR-Anwendung soll das eigene Keyboard oder E-Piano über visuelle Marker und der Kamera eines Smartphones/Tablets oder einer AR-Brille dahingehend erfasst werden, dass es in das 3D-Modell eines Orgelspieltisches hineinintegriert bzw. mit Überblendungen versehen wird und somit den audiovisuellen Eindruck des originalen Spieltisches beim Spielen simuliert.

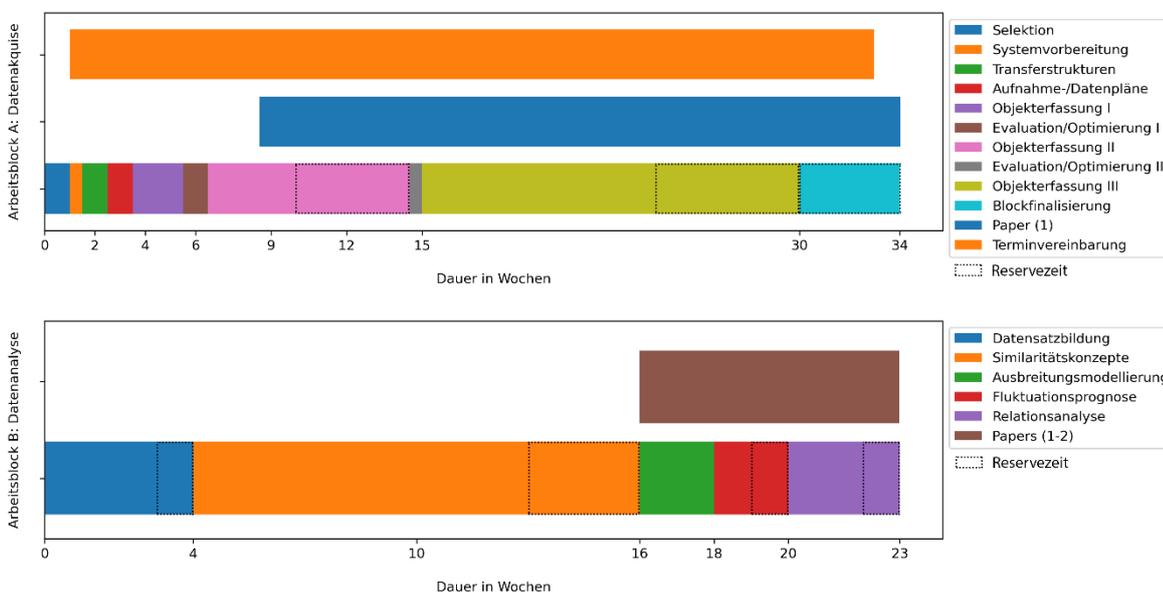
Im letzten Arbeitsblock erfolgt zunächst die Rücksprache mit den Betreuern und kooperierenden Institutionen, gefolgt von der Einordnung daraus hervorgehender Optimierungsvorschläge. Im zweiten Arbeitspaket erfolgt schließlich die Zusammenführung, Korrektur und Finalisierung der Dissertation, in die nur im dringenden Falle ggf. neu aufkommende Inhalte einfließen sollen. Diese soll in englischer Sprache verfasst und der internationalen Forschungsgemeinschaft nach ihrer Bewertung in einem ansprechenden Layout zur Verfügung gestellt werden.

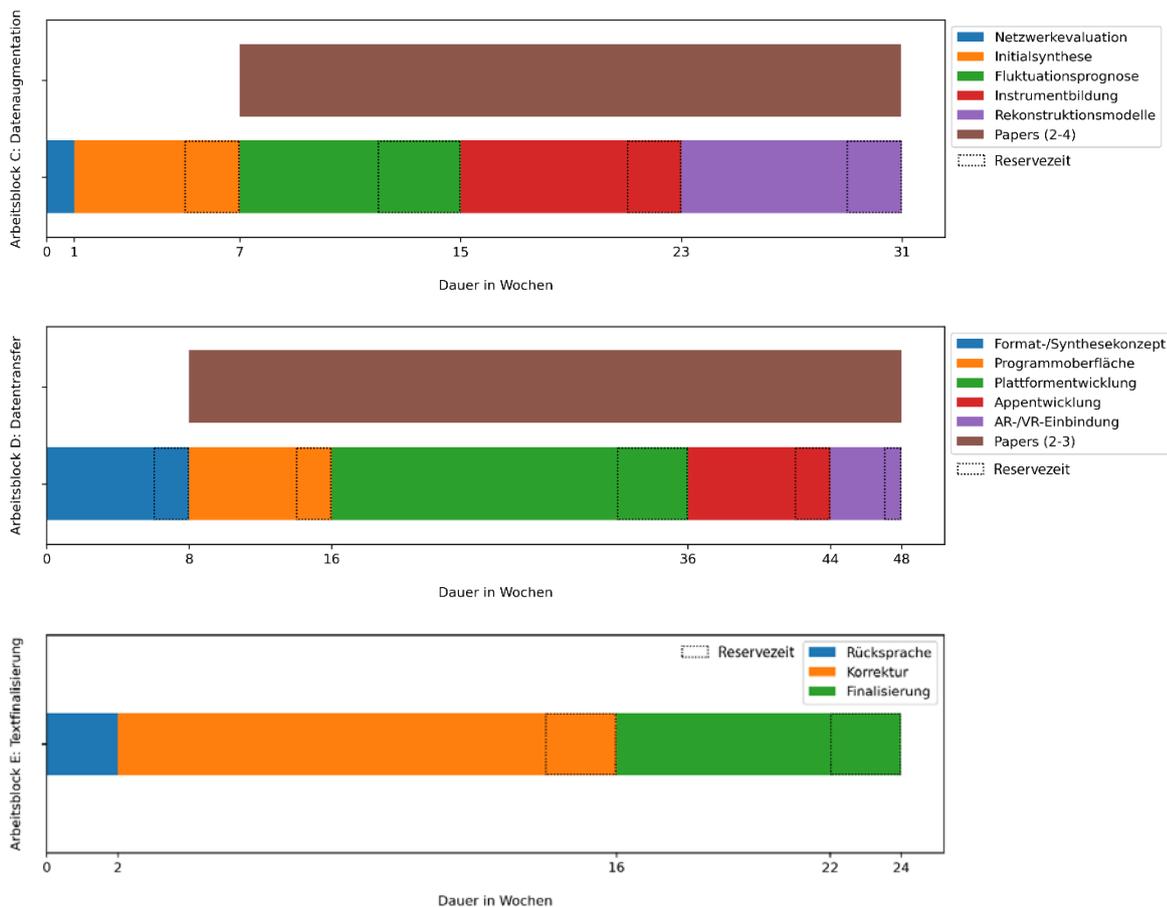
Für die Durchführung des Promotionsvorhabens wurden begleitende Mechanismen entwickelt, die eine wöchentliche Evaluation des Arbeits- und Forschungsfortschritts sowie der eigenen Auslastung gewährleisten. Diese ermöglichen nicht nur eine präzise Einschätzung des Fortschritts im Projekt, sondern tragen auch zu einer beständigen Leistungsfähigkeit bei und dienen der Vorbeugung von stressbedingten Fehlern oder Unachtsamkeiten. Anhand dieser Evaluationsmechanismen wird auch entschieden, ob die fest im Zeitplan integrierten Pufferzonen am Ende eines Arbeitspaketes aktiviert werden, etwa falls eine länger andauernde Unzufriedenheit bzgl. des Arbeitsfortschrittes auftritt. Somit können Problemfelder frühzeitig erkannt, gezielt bearbeitet oder gänzlich umgangen werden, bspw. falls die Bearbeitung eines zeitintensiven Problems für die Beantwortung einer Forschungsfrage von vergleichsweise geringer Relevanz ist oder mittels eines reduzierten Datenaufwands erfolgen könnte. Die Auslastung wurde mit Blick auf die Zeitplanung zwar so weit wie möglich antizipiert, in der aktuellen pandemischen Situation können jedoch unvorhergesehene Beschränkungen eintreten, weshalb die Arbeitspakete modular gestaltet wurden und bereits nach Abschluss von wenigen Objekterfassungen jederzeit umorganisiert werden können. Um die einzelnen Aufgaben der Arbeitspakete, die Arbeitsstände und das progressive Fortschreiten im Projekt übersichtlich verfolgen zu können wird die Plattform *Asana* [128] zum Projektmanagement verwendet, die sich bereits ideal bewähren konnte. Die permanente Beobachtung von neuen, relevanten Forschungsarbeiten im thematischen Umfeld des Promotionsprojektes erfolgt neben konventionellen Recherchen über Feeds und Kanäle ausgewählter Forscher:innen oder Institutionen sowie über die App *Researcher* [129], welche mit der Literaturverwaltungssoftware *Mendeley* [130] synchronisierbar ist, die wiederum mit einem Cloudspeicher verbunden ist und somit einen geräteübergreifenden Zugang auf sämtliche relevante Literatur ermöglicht.

V. Zeit- und Arbeitsplan

Der konzipierte Arbeitsplan beruht, wie zuvor im methodischen Abschnitt geschildert, auf einer Einteilung des Projekts in vier umfassende Arbeitsblöcke mit fest definierten Arbeitspaketen und Aufgaben. Diese Arbeitsblöcke bauen wesentlich aufeinander auf, da sie als Akquirierung, Analyse, Synthese/Augmentation und

Transfer von Forschungsdaten klassifizierbar sind, worauf die Korrektur und Finalisierung der Verschriftlichung als eigener Arbeitsblock folgt. Die zeitliche Aufteilung der Arbeitspakete erfolgte über ein Prinzip, nach dem zunächst für jedes Arbeitspaket ein Prioritätswert aus einer quantifizierten Einschätzung der Komplexität, des Verzögerungspotenzials, der eigenen Erfahrung und des Automatisierungspotenzials (z. B. hoch bei Metadatengenerierung, niedrig bei Objekterfassung) auf einer Skala von 1-5 berechnet wurde. Anschließend wurden angemessene Zeiträume und zeitliche Reserven eingeschätzt, die einkalkulierten potenziellen Verzögerungen entsprechen und bei einer idealen Einhaltung des Zeitplans nicht angebrochen werden sollen, jedoch stets eine realistische Einhaltung des Projektplans ermöglichen können. Sofern es aus unvorhersehbaren Gründen auch zu einer Überschreitung dieser Zeitreserven kommen sollte, so kann dem betroffenen Arbeitspaket die für äußerst niedrigpriorisierte Aufgaben eingeplante Zeit neu zugewiesen werden. Zu diesem Zwecke wurden die als vergleichsweise am wenigsten forschungsrelevant eingeschätzten Aufgaben vorwiegend an das Ende des Projektvorhabens positioniert, sodass eine solche Zuweisung ohne vorherige zeitliche Verluste einhergehen würde. Trotz der konkreten zeitlichen Einteilung der Arbeitspakete bleibt die Zeitplanung stets dynamisch, da wöchentliche Fortschritte registriert werden und die verfrühte oder verspätete Fertigstellung eines Arbeitspakets automatisch zu einer Umverteilung von Zeit zugunsten der ausstehenden Arbeitspakete mit absteigender Priorität führt. Eine frühere Fertigstellung des Arbeitspakets *Objekterfassung II* würde somit die ausstehende Zeit der *Objekterfassung III* zugewiesen, umgekehrt würde eine verspätete Fertigstellung von *Objekterfassung II* zu einem entsprechenden Zeitabzug des Arbeitspaketes *AR-/VR-Einbindung* führen. Zusätzlich zur genannten Zeit- und Arbeitsplanstruktur sind wöchentliche Routinen eingeplant, welche die Registrierung und Evaluierung des Arbeitsfortschritts, das Formulieren von Tages- oder Wochenzielen, das Integrieren von derweil publizierter Literatur sowie die Bewertung der eigenen Belastung und Zufriedenheit betreffen, was aufgrund der vorgefertigten Formulare wenig Zeit in Anspruch nimmt und einen großen Nutzen in sich birgt. Die Arbeitszeit wird mit mind. 45-50 Stunden pro Woche berechnet, wovon 3 Stunden für diverse Gesprächs- oder Vortragstermine und für die Routinen sowie Transfertätigkeiten 1 Stunde einkalkuliert werden. An außereuropäischen Fachkonferenzen soll aus zeitlichen und umweltschonenden Gründen möglichst virtuell teilgenommen werden, während für den jährlichen Urlaub 30 Tage eingeplant sind.





VI. Literaturverzeichnis

- [1] Musikinstrumentenmuseum der Universität Leipzig, „TASTEN“, <https://mfim.uni-leipzig.de/dt/Forschung/Tastenprojekt.php> [03.02.2022]
- [2] Deutsche Stiftung Denkmalschutz, „Denkmale in Not“, <https://www.monumente-online.de/de/denkmale-in-gefahr/> [03.02.2022]
- [3] UNESCO, „Organ craftsmanship and music“, <https://ich.unesco.org/en/RL/organ-craftsmanship-and-music-01277> [03.02.2022]
- [4] BSI, „Ransomware. Bedrohungslage, Prävention & Reaktion 2021“, <https://www.bsi.bund.de/SharedDocs/Downloads/DE/BSI/Cyber-Sicherheit/Themen/Ransomware.html> [03.02.2022]
- [5] Akademie der Wissenschaften Mainz, „NFDI4Culture“, <https://nfdi4culture.de/> [03.02.2022]
- [6] Europarat, „Council of Europe Framework Convention on the Value of Cultural Heritage for Society“, in: *CETS*, No. 199, <https://rm.coe.int/1680083746>, angefertigt am 27.10.2005
- [7] Europäische Kommission, „Empfehlung (EU) 2021/1970 der Kommission vom 10. November 2021 für einen gemeinsamen europäischen Datenraum für das Kulturerbe“, in: *C/2021/7953*, <http://data.europa.eu/eli/reco/2021/1970/oj>, angefertigt am 10.11.2021
- [8] Angster, J. u. a. (2019), „25 years applied pipe organ research at Fraunhofer IBP in Stuttgart“, in: *Proc. ISMA* (Sept. 2019), Pap. 16: S. 1–15
- [9] Rucz, P. (2009), „Determination of organ pipes’ acoustic parameters by means of numerical techniques“, in: *Acoust. Rev.*, Bd. 9(5): S. 1–9
- [10] Miklos, A. u. a., „Properties of the Sound of Flue Organ Pipes“, in: *Acta Acust.*, Bd. 86(4): S. 611–622
- [11] Fischer, J. L. u. a. (2016), „Aeroacoustical coupling and synchronization of organ pipes“, in: *JASA*, Bd. 140(4): S. 2344–2351
- [12] Okada, M. u. a. (2019), „Acoustic measurements of the infinite-simal phase response curve from a sounding organ pipe“, in: *Phys. Lett. Sect. A*, Bd. 383(15): S. 1733–1741
- [13] Girón, S. u. a. (2017), „Church acoustics: A state-of-the-art review after several decades of research“, in: *J. Sound Vib.*, Bd. 411: S. 378–408
- [14] Meyer, J. (2003), „Wechselbeziehungen zwischen Orgel und Raumakustik“, in: *DAGA* (2003), S. 518–519
- [15] Rioux, V. u. a. (1999), „Noise quality of transient sounds: Perception of ‚hiss‘ and ‚cough‘ in a flue organ pipe“, in: *JASA*, Bd. 105(2): S. 1002
- [16] Eberlein, Roland (2011), *Die Geschichte der Orgel*, Köln 2011
- [17] Eberlein, Roland (2008), *Orgelregister, ihre Namen und ihre Geschichte*, Köln 2008
- [18] Czystewski, A. u. a. (1995), „New Approach to the Synthesis of Organ Pipe Sound“, in: *Proc. 98th Conv. AES*, Pap. 3957: S. 1–13
- [19] Smith, J. O. (1992), „Physical Modeling Using Digital Waveguides“, in: *Comput. Music J.*, Bd. 16(4): S. 74
- [20] Gabrielli, L. u. a. (2019), „A multi-stage algorithm for acoustic physical model parameters estimation“, in: *IEEE/ACM Trans. Audio Speech Lang. Process.*, Bd. 27(8): S. 1229–1240
- [21] Carvalho, M. u. a. (2021), „Physical modelling techniques for the dynamical characterization and sound synthesis of historical bells“, in: *Herit. Sci.*, Bd. 9(1): S. 157
- [22] GrandOrgue, „GrandOrgue“, <https://github.com/GrandOrgue/grandorgue> [03.02.2022]
- [23] Fugal, H., „Aeolus“, <https://github.com/fugalh/aeolus> [03.02.2022]
- [24] Milan Digital Audio, „Hauptwerk – Virtual Pipe Organ“, <https://www.hauptwerk.com/> [03.02.2022]
- [25] Native Instruments, „Kontakt 6“, <https://www.native-instruments.com/de/products/komplete/samplers/kontakt-6/> [03.02.2022]
- [26] Bron, P., „Organ Database“, <http://www.orgbase.nl/> [03.02.2022]

- [27] Focht, J. (Hrsg.), „musiXplora“, <https://musixplora.de/> [03.02.2022]
- [28] Khulusi, R. u. a. (2020), „MusiXplora: Visual analysis of a musicological encyclopedia“, in: *Proc. 15th Int. Conf. VISIGRAPP2020*, Bd. 3, S. 76–87
- [29] Balint, J. u. a. (2021), „Research data management across disciplines in AUDICTIVE“, in: *DAGA (2021)*, S. 1348–1351
- [30] Universität Konstanz, „Forschungsdaten und Forschungsdatenmanagement“, <https://www.forschungsdaten.info/> [03.02.2022]
- [31] Spors, S. u. a. (2017), „Towards Open Science in Acoustics: Foundations and Best Practices The Scientific Method“, in: *DAGA (2017)*, S. 218–221
- [32] Wilkinson, M. D. u. a. (2016), „Comment: The FAIR Guiding Principles for scientific data management and stewardship“, in: *Sci. Data*, Bd. 3(160018): S. 1–9
- [33] Ronchi, A. M. (2019), „e-Culture: On Culture in the Digital Age“, in: *e-Services*, S. 115–193
- [34] Aliprantis, J. u. a. (2018), „Linked Open Data as Universal Markers for Mobile Augmented Reality Applications in Cultural Heritage“, in: *Digit. Cult. Herit.*, S. 79–90
- [35] Comunità, M. u. a. (2019), „Web-based binaural audio and sonic narratives for cultural heritage“, in: *Conf. Immersive Interact. Audio* (März 2019), Pap. 14: S. 1–11
- [36] Kleimola, J. u. a. (2015), „Web audio modules“, in: *Proc. 12th Int. Conf. SMC*, S. 249–256
- [37] Larkin, O. u. a. (2018), „iPlug 2: Desktop Plug-in Framework Meets Web Audio Modules“, in: *Proc. Int. Web Audio Conf.*, S. 1–6
- [38] DEGA (2018), „Memorandum zur Durchführung und Dokumentation von Audio-Produktionen für wissenschaftliche Anwendungen in der Akustik“, in: *DEGA VA 12* (Nov 2018)
- [39] Epe, C. u. a. (2018), „DEGA-Audiodatenbank zur Evaluierung von virtuellen auditiven Umgebungen: Erstellung und Dokumentation von Kirchenorgelaufnahmen zur Verwendung als Stimuli in virtuellen Umgebungen Einleitung“, in: *DAGA (2018)*, S. 106–108
- [40] Urban, O. (2006), „Introduction to Problems of In Situ Pipe Organ Sampling“, in: *Proc. 33rd Int. Acoust. Conf.*, S. 86–91
- [41] Asutay, E. u. a. (2012), „Development of methodology for documentation of key action properties and haptic sensation of pipe organ playing“, in: *Acoust. Bull.*, Bd. 5(37), S. 42–44
- [42] Patricio, E. u. a. (2019), „Toward six degrees of freedom audio recording and playback using multiple ambisonics sound fields“, in: *146th Int. Conv. AES*, Pap. 10141
- [43] Tylka, J. G. u. a. (2019), „Domains of practical applicability for parametric interpolation methods for virtual sound field navigation“, in: *JAES*, Bd. 67(11): S. 882–893
- [44] Ziemer, T. (2017), „Source Width in Music Production. Methods in Stereo, Ambisonics, and Wave Field Synthesis“, in: *Source Width Music Prod. Methods Stereo, Ambisonics, Wave F. Synth.*, hrsg. v. A. Schneider, Cham 2017, S. 299–340
- [45] Braun, S. u. a. (2019), „Acoustic localization using spatial probability in noisy and reverberant environments“, in: *IEEE Work. Appl. Signal Process. to Audio Acoust.* (Okt 2019), S. 353–357
- [46] Grumiaux, P.-A. u. a. (2021), „A Survey of Sound Source Localization with Deep Learning Methods“, in: *arXiv:2109.03465*
- [47] Jot, J. M. u. a. (2021), „Rendering Spatial Sound for Interoperable Experiences in the Audio Metaverse“, in: *arXiv:2109.12471*
- [48] Sz wajcowski, A. (2021), „Objective-Oriented Method for Uniformation of Various Directivity Representations“, in: *151st Conv. AES* (Okt 2021), Pap. 10521: S. 1–11
- [49] Yeoward, C. u. a. (2021), „Real-time binaural room modelling for augmented reality applications“, in: *J. Audio Eng. Soc.*, Bd. 69(11): S. 818–833
- [50] Ziemer, T. u. a. (2017), „Psychoacoustic sound field synthesis for musical instrument radiation characteristics“, in: *JAES*, Bd. 65(6): S. 482–496
- [51] Richard, A. u. a. (2020), „Sine sweep optimization for room impulse response measurements“, in: *Proc. Forum Acusticum*, S. 147–154
- [52] Chen, M. u. a. (2021), „De-noising process in room impulse response with generalized spectral subtraction“, in: *Appl. Sci.*, Bd. 11(6858): S. 1–20
- [53] Massé, P. u. a. (2020), „Measurement, Analysis, and Denoising of Directional Room Impulse Responses in Complex Spaces“, in: *Proc. Forum Acusticum*, S. 137–144
- [54] Barré, S. u. a. (2014), „Room impulse response measurement with a spherical microphone array, application to room and building acoustics“, in: *43rd Int. Congr. INTERNOISE*, Pap. 840, S. 1–6
- [55] Alary, B. u. a. (2021), „A Method for Capturing and Reproducing Directional Reverberation in Six Degrees of Freedom“, in: *arxiv:2110.04082*
- [56] Kliesch, D. u. a. (2020), „Die Messunsicherheit des Klarheitsmaßes bei raumakustischen Messungen“, in: *DAGA (2020)*, S. 846–849
- [57] Emmanouilidou, D. u. a. (2019), „The effect of room acoustics on audio event classification“, in: *Proc. Int. Congr. Acoust.* (Sept. 2019), S. 102–109
- [58] Werner, S. u. a. (2016), „A summary on acoustic room divergence and its effect on externalization of auditory events“, in: *8th Int. Conf. QoMEX*, S. 1–6
- [59] Falcón Pérez, R. u. a. (2019), „Machine-learning-based estimation of reverberation time using room geometry for room effect rendering“, in: *Proc. 23rd Int. Congr. Acoust.*, S. 7258–7265
- [60] Koyama, S. u. a. (2021), „MeshRIR: A Dataset of Room Impulse Responses on Meshed Grid Points For Evaluating Sound Field Analysis and Synthesis Methods“, in: *arXiv:2106.10801*
- [61] McKenzie, T. u. a. (2021), „Dataset of Spatial Room Impulse Responses in a Variable Acoustics Room for Six Degrees-of-Freedom Rendering and Analysis“, in: *arXiv:2111.11882*
- [62] Aspöck, L. u. a. (2018), „Eine Datenbank zur Validierung von akustischen Simulationen“, in: *DAGA (2018)*, S. 1337
- [63] Gumerov, N. A. u. a. (2021), „Fast multipole accelerated boundary element methods for room acoustics“, in: *JASA*, Bd. 150(3): S. 1707–1720
- [64] Diaz-Guerra, D. u. a. (2021), „gpuRIR: A python library for room impulse response simulation with GPU acceleration“, in: *Multimed. Tools Appl.*, Bd. 80(4): S. 5653–5671
- [65] Kimmich, J. M. u. a. (2018), „Berechnung des Schallfeldes in der Deutschen Oper Berlin mit Raytracing und der Finiten-Elemente-Methode Einleitung“, in: *DAGA (2018)*, S. 1097–1099
- [66] Garí, S. V. A. u. a. (2019), „Evaluation of real-time sound propagation engines in a virtual reality framework“, in: *Conf. Immersive Interact. Audio*, Pap. 37, S. 1–10
- [67] Ziesemer, S. u. a. (2021), „High Resolution Sound Field Analysis with Directional Room Impulse Responses“, in: *DAGA (2021)*, S. 820–823
- [68] Błaszczak, P. u. a. (2021), „Analysis of sound field distribution in architecturally diverse temples“, in: *Arch. Acoust.*, Bd. 46(1): S. 121–133
- [69] Tronchin, L. u. a. (2020), „Evaluation of acoustic similarities in two italian churches honored to S. Dominic“, in: *Appl. Sci.*, Bd. 10(20): S. 1–16
- [70] Ciaburro, G. u. a. (2021), „The acoustic characteristics of the “Dives in Misericordia” Church in Rome“, in: *Build. Acoust.*, Bd. 28(2): S. 197–206
- [71] Lorenz-Kierakiewitz, K.-H. u. a. (2021), „Mehrdimensionale visuelle Clusteranalyse der raumakustischen Parameterwerte von Wiener Sälen, Kirchen und Auditorien“, in: *DAGA (2021)*, S. 1121–1123
- [72] Alvarez-Morales, L. u. a. (2019), „Mapping the acoustics of Ripon cathedral“, in: *Proc. 23rd Int. Congr. Acoust.*, S. 2335–2342
- [73] Autio, H. u. a. (2021), „Historically based room acoustic analysis and auralization of a church in the 1470s“, in: *Appl. Sci.*, Bd. 11(4): S. 1–25
- [74] Boren, B. B. (2021), „Acoustic simulation of J.S. Bach’s Thomaskirche in 1723 and 1539“, in: *Acta Acust.*, Bd. 5(14): S. 1–11
- [75] Wahbeh, W. (2021), „Parametric modelling approach to reconstructing architectural indoor spaces from point clouds“, in: *Int. Arch. ISPRS*, Bd. 43(B4): S. 251–257

- [76] Liu, J. u. a. (2019), „Dereverberation based on deep neural networks with directional feature from spherical microphone array recordings“, in: *Proc. 23rd Int. Congr. Acoust.*, S. 4932–4939
- [77] Bai, M. R. u. a. (2019), „Deep Learning Applied to Dereverberation and Sound Event Classification in Reverberant Environments“, in: *Proc. 23rd Int. Congr. Acoust.*, S. 4917–4923
- [78] Chen, C. u. a. (2021), „Learning Audio-Visual Dereverberation“, in: *arXiv:2106.07732*
- [79] Plath, N. u. a. (2021), „Piano Timbre Development Analysis using Machine Learning“, in: *arXiv:2112.03214*
- [80] Bader, R. u. a. (2021), „Computational timbre and tonal system similarity analysis of the music of Northern Myanmar-based Kachin compared to Xinjiang-based Uyghur ethnic groups“, in: *arXiv:2103.08203*
- [81] Reuter, C. u. a. (2018), „Instrumentale Formantnähe und Klangfarbenähnlichkeit aus menschlicher und rechnerische Perspektive“, in: *DAGA (2018)*, S. 1711–1714
- [82] Siddiq, S. u. a. (2021), „Formanten bringen Farbe ins Spiel Möglichkeiten der Klangfarbenbeschreibung im dynamischen Formantenfeld“, in: *DAGA (2021)*, S. 520–523
- [83] Reymore, L. (2021), „Characterizing prototypical musical instrument timbres with Timbre Trait Profiles“, in: *Musica. Sci.* (Apr 2021), S. 1–27
- [84] Meng, Z. (2021), „Research on timbre classification based on BP neural network and MFCC“, in: *J. Phys. Conf. Ser.*, Bd. 1856(2021)
- [85] Siddiq, S. u. a. (2018), „Die physikalischen Korrelate von Instrumentalklangfarben“, in: *DAGA (2018)*, S. 1695–1698
- [86] Kazazis, S. u. a. (2021), „Ordinal scaling of timbre-related spectral audio descriptors“, in: *JASA*, Bd. 149(6): S. 3785–3796
- [87] Saitis, C. u. a. (2019), „The role of attack transients in timbral brightness perception“, in: *Proc. 23rd Int. Congr. Acoust.*, Pap. 5506
- [88] Saitis, C. u. a. (2020), „Brightness perception for musical instrument sounds: Relation to timbre dissimilarity and source-cause categories“, in: *JASA*, Bd. 148(4): S. 2256–2266
- [89] Wallmark, Z. u. a. (2021), „Describing Sound“, in: *Oxford Handbook of Timbre*, S. 577–608
- [90] Zinemanas, P. u. a. (2021), „An interpretable deep learning model for automatic sound classification“, in: *Electronics*, Bd. 10(7):850
- [91] Bianco, M. J. u. a. (2019), „Machine learning in acoustics: Theory and applications“, in: *JASA*, Bd. 146(5): S. 3590–3628
- [92] Guizzo, E. u. a. (2021), „L3DAS21 Challenge: Machine Learning for 3D Audio Signal Processing“, in: *arXiv:2104.05499*
- [93] Haider, D. u. a. (2021), „Zeit-Frequenz Darstellungen und Deep Learning“, in: *DAGA (2021)*, S. 29–32
- [94] Natsiou, A. u. a. (2022), „Audio representations for deep learning in sound synthesis: A review“, in: *arXiv:2201.02490*
- [95] Esling, P. u. a. (2018), „Bridging audio analysis, perception and synthesis with perceptually-regularized variational timbre spaces“, in: *Proc. 19th ISMIR*, S. 175–181
- [96] Bitton, A. u. a. (2020), „Vector-Quantized Timbre Representation“, in: *arXiv:2007.06349*
- [97] Verma, P. u. a. (2019), „Neuralogram: A Deep Neural Network Based Representation for Audio Signals“, in: *arXiv:1904.05073*
- [98] Subramani, K. u. a. (2020), „Vapar Synth - A Variational Parametric Model for Audio Synthesis“, in: *ICASSP Proceedings (Mai 2020)*, S. 796–800
- [99] Caillon, A. u. a. (2021), „RAVE: A variational autoencoder for fast and high-quality neural audio synthesis“, in: *arXiv:2111.05011*
- [100] Marafioti, A. u. a. (2020), „Time-Frequency Analysis for Neural Synthesis of Audio“, in: *DAGA (2020)*, S. 130–133
- [101] Cho, Y. u. a. (2021), „Efficient adversarial audio synthesis via progressive upsampling“, in: *ICASSP Proceedings (Jun 2021)*, S. 3410–3414
- [102] Mehri, S. u. a. (2016), „SampleRNN: An Unconditional End-to-End Neural Audio Generation Model“, in: *arXiv:1612.07837*
- [103] Shan, S. u. a. (2021), „Differentiable Wavetable Synthesis“, in: *arXiv:2111.10003*
- [104] Hayes, B. u. a. (2021), „Neural Waveshaping Synthesis“, in: *Proc. 22nd ISMIR*, S. 254–261
- [105] Engel, J. u. a. (2020), „DDSP: Differentiable Digital Signal Processing“, in: *arXiv:2001.04643*
- [106] Kilgour, K. u. a. (2019), „Fréchet Audio Distance: A Reference-Free Metric for Evaluating Music Enhancement Algorithms“, in: *Proc. Interspeech (Sept. 2019)*, S. 2350–2354
- [107] Pons, J. u. a. (2021), „Upsampling Artifacts in Neural Audio Synthesis“, in: *ICASSP Proceedings (Jun 2021)*, S. 3005–3009
- [108] Lluís, F. u. a. (2021), „Points2Sound: From mono to binaural audio using 3D point cloud scenes“, in: *arXiv:2104.12462*
- [109] Gabrielli, L. u. a. (2018), „Deep Learning for Timbre Modification and Transfer: an Evaluation Study“, in: *144th Conv. AES*, Pap. 9996
- [110] Huang, S. u. a. (2019), „Timbretron: A wavenet(CycleGAN(CqT(Audio))) pipeline for musical timbre transfer“, in: *arXiv:1811.09620*
- [111] Bonnici, R. S. u. a. (2021), „Timbre Transfer with Variational Auto Encoding and Cycle-Consistent Adversarial Networks“, in: *arXiv:2109.02096*
- [112] Ganis, F. u. a. (2021), „Real-time Timbre Transfer and Sound Synthesis using DDSP“, in: *arXiv:2103.07220*
- [113] McFee, B. u. a. (2015), „librosa: Audio and Music Signal Analysis in Python“, in: *Proc. 14th Python Sci. Conf.*, S. 18–24
- [114] Bogdanov, D. u. a. (2013), „Essentia: An audio analysis library for music information retrieval“, in: *Proc. 14th ISMIR*, S. 493–498
- [115] Hollomey, C. u. a. (2021), „The Large Time Frequency Analysis Toolbox“, in: *DAGA (2021)*, S. 26–28
- [116] Brinkmann, F. u. a. (2017), „AKtools-eine offene Software zur Erhebung, Verarbeitung und Inspektion akustischer Signale“, in: *DAGA (2017)*, S. 210–213
- [117] Fonseca, W. u. a. (2019), „PyTTA: Open source toolbox for acoustic measurement and signal processing“, in: *Proc. 23rd Int. Congr. Acoust.*, S. 4983–4990
- [118] Gomes, O. C. u. a. (2020), „Spatial Energy Analysis for Room Acoustical Evaluation“, in: *DAGA (2020)*, S. 834–837
- [119] Scheibler, R. u. a. (2018), „Pyroomacoustics: A Python Package for Audio Room Simulation and Array Processing Algorithms“, in: *ICASSP Proceedings (Apr 2018)*, S. 351–355
- [120] Cheuk, K. W. u. a. (2020), „NnAudio: An on-the-Fly GPU Audio to Spectrogram Conversion Toolbox Using 1D Convolutional Neural Networks“, in: *IEEE Access*, Bd. 8, S. 161981–162003
- [121] Sharma, G. u. a. (2020), „Trends in audio signal feature extraction methods“, in: *Appl. Acoust.*, Bd. 158(107020)
- [122] AliceVision, „AliceVision. Photogrammetric Computer Vision Framework“, <https://alicevision.org/> [03.02.2022]
- [123] Blender Foundation, „Blender. Free and Open 3D Creation Software“, <https://www.blender.org/> [03.02.2022]
- [124] JUCE, „JUCE“, <https://juce.com/> [03.02.2022]
- [125] Herre, J. u. a. (2015), „MPEG-H 3D Audio - The New Standard for Coding of Immersive Spatial Audio“, in: *IEEE J. Sel. Top. Signal Process.*, Bd. 9(5): S. 770–779
- [126] Open 3D Engine Contributors, „Open 3D Engine“, <https://www.o3de.org/> [03.02.2022]
- [127] Khronos Group Inc., „OpenXR Overview - The Khronos Group Inc“, <https://www.khronos.org/openxr/> [03.02.2022]
- [128] Asana Inc., „Asana“, <https://asana.com/> [03.02.2022]
- [129] Imperial College London, „Researcher | An App For Academics“, <https://www.researcher-app.com/> [03.02.2022]
- [130] Mendeley Ltd., „Mendeley“, <https://www.mendeley.com/> [03.02.2022]