

Räumliche Darstellung in Telefonkonferenzen: Implementierung eines abwärtskompatiblen Telefonkonferenzdienstes mit 3D-Audio-Funktion (3D-Telko)

Abschlussbericht

Antragsteller	Dennis Guse Quality and Usability Lab Technische Universität Berlin Ernst-Reuter-Platz 7 D-10587 Berlin dennis.guse@tu-berlin.de
Laufzeit	1. Januar 2016 bis 31. Dezember 2016 (12 Monate)
Beteiligte Partner	Technische Universität Berlin, Fachgebiet Quality and Usability Lab – betreut durch Prof. Dr.-Ing. Sebastian Möller Deutsche Telekom AG (ungeförderter Industriepartner) – betreut durch Marlene Gernoth

Hinweis

Das diesem Bericht zugrunde liegende Vorhaben wurde mit Mitteln des Bundesministeriums für Bildung und Forschung unter dem Förderkennzeichen 01IS11556 gefördert. Die Verantwortung für den Inhalt dieser Veröffentlichung liegt beim Autor (siehe Antragsteller).

1. Aufgabenstellung und Motivation

Globale Joint Ventures, umfassende Netzwerke und mobiles Arbeiten machen Telefonkonferenzen im Geschäftsalltag von Unternehmen sowie im wissenschaftlichen Austausch unerlässlich. Bisher ringen diese den Teilnehmern jedoch ein hohes Maß an Aufmerksamkeit ab. Bestehende Konferenzsysteme bieten nur die Möglichkeit einer einkanaligen Übertragung, wobei die Sprachsignale von gleichzeitig sprechenden Teilnehmern übereinandergelegt werden. Das Resultat: Es gestaltet sich schwierig, die einzelnen Sprecher zuzuordnen und klar zu verstehen. Durch die physische Abwesenheit lässt sich ein Sprechwunsch kaum artikulieren. Gespräche werden verzögert, überlappen oder arten in Monologe einzelner Teilnehmer aus. Insgesamt wird den Teilnehmern dadurch eine erhöhte Konzentration abverlangt, die vom Inhalt der Telefonkonferenz ablenken kann.

Für den effizienten, reibungslosen Geschäfts- und Wissensaustausch ist es daher essentiell, diesen Kommunikationskanal zu optimieren. Dafür bietet die State-of-the-Art Audiotechnologie entscheidende Impulse: Die Nutzung von Stereokopfhörern ermöglicht es, Audioquellen räumlich darzustellen und damit auch Telefonkonferenzen natürlicher zu gestalten. Einzelne Teilnehmer werden akustisch im virtuellen Raum verteilt, ähnlich einer Sitzung rund um einen Konferenztisch. Die räumliche Darstellung erleichtert den Teilnehmern die Identifikation einzelner Sprecher.

1.1. Schwerpunkte und Ziele

Zielsetzung des Vorhabens war die Implementierung einer Software, konkret einer Voice-over-IP-Telefonanlage, welche Telefonkonferenzen mit 3D-Audio-Darstellung bereitstellt. Die Neuartigkeit der erstellten Implementierung besteht darin, dass die 3D-Audio-Darstellung mit herkömmlichen Stereo-Endgeräten genutzt werden kann.

Weiterhin war geplant mit der implementierten Softwarelösung Versuche durchzuführen, um den Mehrwert sowie Akzeptanz von 3D-Audio für Telefonkonferenzen zu untersuchen.

1.2. Wissenschaftliche und technische Ziele des Vorhabens

Zur erfolgreichen Umsetzung dieses Ziels war die Beantwortung einer Reihe anwenderorientierter Forschungsaspekte notwendig, welche im Rahmen des Forschungsprojektes erfolgte:

1. Abwärtskompatible und standardkonforme Erweiterung eines bestehenden eines Open-Source Telefonkonferenzsystems für 3D-Audio.
2. Analyse und ggf. Anpassung von Client-Software zur Nutzung, so dass das synthetisierte 3D-Audio-Signal korrekt wiedergegeben werden kann.
3. Untersuchung der Vorteile räumlicher Darstellung.

1.3. Bezug des Vorhabens zu förderpolitischen Zielen / Förderprogramm

Die Hightech-Strategie 2020, an der sich das Bundesministerium für Bildung und Forschung orientiert, will brachliegende Potenziale in Wissenschaft und Wirtschaft gezielt aktivieren, um auch zukünftig eine Führungsposition der Bundesrepublik zu behaupten. Es werden einzelne Technologiefelder zur Lösung wichtiger gesellschaftspolitischer Zielstellungen oder als Innovationstreiber für andere Technologiefelder als wesentliche Voraussetzungen verstanden. Zu diesen Schlüsseltech-

nologien zählt unter anderem der Bereich der *Informations- und Kommunikationstechnologie*. Im Rahmen des Software Campus werden die IT-Führungskräfte von morgen ausgebildet und jungen IT-Experten exzellente Karriereperspektiven in Deutschland eröffnet.

In dem hier dargelegten Vorhaben wurde ein abwärtskompatibles Telefonkonferenzsystem mit räumlicher Darstellung implementiert. Dieses Vorhaben fällt damit genau in den Bereich einer der wichtigen Schlüsseltechnologien.

1.4. Wissenschaftlich-technische Umsetzung des Vorhabens

Im Rahmen des Vorhabens wurde die Open-Source Telefonieservers Asterisk¹, welche auch Telefonkonferenzen bereitstellen kann, erweitert. Die Erweiterung umfasst die Verarbeitung von zweikanaligen Audiosignalen und Implementierung von Binauralsynthese. Letztere ermöglicht das Rendering einer virtuellen Umgebung in 3D-Audio. Hierzu wurden zuerst die technischen Möglichkeiten, auch abseits von Asterisk, zur Umsetzung evaluiert, um daraufhin zielgerichtet die Implementierung umzusetzen. Die ausgewählte Möglichkeit der Implementierung wurde aufgrund der Zukunftstauglichkeit (zukünftiger Wartungsaufwand erfolgt teilweise durch Open-Source-Community) sowie bestehender Vorkenntnisse beim Antragsteller ausgewählt.

Bei der Implementierung wurde eng mit der Open-Source Community hinter Asterisk zusammengearbeitet, damit die finale Lösung erfolgreich integriert werden konnte.

¹ Projektwebseite von Asterisk: [HTTP://WWW.ASTERISK.ORG](http://www.asterisk.org)

2. Ausführliche Beschreibung des Arbeitsplans

2.1. Arbeitspaket 1: Projektmanagement

Arbeitspaket Nummer	AP1	Zeitliche Einordnung:	M1-M12
---------------------	-----	-----------------------	--------

Ziel

In diesem Arbeitspaket wurden Entwürfe für Schnittstellen und Entwicklungspläne für die anderen Arbeitspakete entwickelt und parallel laufende Tätigkeiten koordiniert. Dabei wurde für eine termingerechte Abgabe der lieferbaren Ergebnisse gesorgt, die finanztechnische Abrechnung sichergestellt und das Projekt-Controlling durchgeführt. Die Ergebnisse wurden durch adäquate Maßnahmen nach außen sichtbar gemacht. Ebenfalls wurden die Personalbesetzung sowie die Personalführung in diesem Arbeitspaket durchgeführt.

Detailbeschreibung (Tasks)

- **T1.1: Administratives Management und Finanzen**
- **T1.2: Wissenschaftliches Management**
- **T1.3: Technisches Management**

Ergebnisse:

Alle Ziele des Arbeitspakets wurden erfolgreich umgesetzt. Insbesondere konnte sich der Projektleiter in den Umgang mit einem eigenen Team ausprobieren. Dies umfasste sowohl Personalführung als auch die erfolgreiche Kommunikation mit externen Stakeholdern.

Weiterhin wurde in diesem Arbeitspaket der Transfer der Projektergebnisse (technischer als auch praktischer Natur) durchgeführt. Dies umfasste auch die Planung der Projektwebseite.

2.2. Arbeitspaket 2: Implementierung des Telefonkonferenzservers

Arbeitspaket Nummer	AP2	Zeitliche Einordnung:	M1-M6
---------------------	-----	-----------------------	-------

Ziel

In diesem Arbeitspaket wurde der Open-Source-Telefonieservers (Asterisk) derart angepasst und erweitert, so dass dieser mittels Binauralsynthese eine räumliche Darstellung für jeden Telefonkonferenzteilnehmer bereitstellen kann.

Dies erforderte zuerst die Erweiterung von Asterisk, so dass dieser mehrkanalige Signale verarbeiten kann. In einem zweiten Schritt wurde eine Konferenzbrücke implementiert, welche Mono-Signale einzelner Teilnehmer mittels Binauralsynthese rendert, so dass ein räumlicher Höreindruck entsteht.

Detailbeschreibung (Tasks)

- **T2.1: Erweiterung Telefonkonferenzservers zur Mehrkanalfähigkeit**
- **T2.2: Implementierung der Konferenzbrücke mit Binauralsynthese**
- **T2.3: Automatisierte Tests für Konferenzbrücke mit räumlicher Darstellung**

Ergebnisse:

Dieses Arbeitspaket wurde vollständig, in Time und in Budget erfüllt. Besonders die Nutzung eines Freelancers hat sich ausgezahlt, da so der richtige Experte für die Implementierung rekrutiert werden konnte. Alle Deliverables wurden pünktlich und qualitativ hochwertig geliefert.

Die Projektergebnisse dieses Arbeitspakets wurden in das Asterisk-Projekt (Upstream) integriert. Dies erforderte allerdings einen nicht geplanten Mehraufwand, denn die funktionierende Implementierung entsprach nicht den Anforderungen der Open-Source-Community. Dies wurde bei der Projektplanung übersehen, so dass der Mehraufwand für Kommunikation im Arbeitspaket 1 geleistet worden ist sowie für die praktische Umsetzung Ressourcen aus dem Arbeitspaket 4 zur Unterstützung genutzt werden mussten.

2.3. Arbeitspaket 3: Implementierung eines mehrkanaligen Telefonie-Clients

Arbeitspaket Nummer	AP3	Zeitliche Einordnung:	M2-M7
---------------------	-----	-----------------------	-------

Ziel

In diesem Arbeitspaket wurde ein Telefonieclient umgesetzt, welcher binaural-gerenderte Signale darstellen kann. Dieser Client sollte möglichst plattformunabhängig sein als auch einfach zu warten sein.

Detailbeschreibung (Tasks)

- **T3.1: Analyse verfügbarer Telefonieclients auf Mehrkanalfähigkeit**
- **T3.2: Implementierung geeigneten mehrkanalfähigen Telefonieclients**

Ergebnisse:

Im Rahmen dieses Arbeitspakets wurde nach umfassender Analyse erfolgreich ein Client unter Nutzung der neuartigen Technologie WebRTC implementiert. WebRTC ist neuer Web-Standard und ermöglicht die Nutzung von Telefonie im Webbrowser. Die Nutzung dieser standardisierten Technologien ermöglichte hier eine techn. einfach wartbare Lösung und auch einfache Installation. Dies erforderte allerdings einen personellen Mehraufwand, welcher vom mit allokierten Ressourcen aus Arbeitspaket 4 realisiert worden ist.

2.4. Arbeitspaket 4: Durchführung von Nutzerstudien

Arbeitspaket Nummer	AP4	Zeitliche Einordnung:	M5-M12
---------------------	-----	-----------------------	--------

Ziel

In diesem Arbeitspaket sollte das Telefonkonferenzsystems (AP2) und der Telefonieclients (AP3) in ein- sowie mehrtägigen Nutzerstudien eingesetzt werden. Hier sollte insbesondere der praktische Mehrwert und auch Einschränkungen der räumlichen Darstellung untersucht werden.

Im Rahmen dieses Arbeitspakets wurde auch die techn. Performanz des Telefonkonferenzsystems evaluiert.

Detailbeschreibung (Tasks)

- **T4.1: Durchführung von Nutzerstudien unter Laborbedingungen**
- **T4.2: Durchführung von mehrtägigen Nutzerstudien unter Feldbedingungen**

Ergebnisse:

Dieses Arbeitspaket wurde nur reduziert bearbeitet, um den ungeplanten Mehraufwand in den Arbeitspaketen 2 und 3 zu erfüllen. Daher wurde auf mehrtägige Versuche komplett verzichtet, damit sehr wahrscheinlich erfolglosen Arbeiten (durch den reduzierten verfügbaren Ressourcen für dieses Arbeitspaket) zu vermeiden. Insgesamt wurde eine Nutzerstudien in Laborsituation zur Untersuchung der Verbesserung der Sprecheridentifikation in einer Telefonkonferenz bei Bereitstellung von 3D-Audio erfolgreich durchgeführt.

Weiterhin wurde in diesem Arbeitspaket das implementierte System getestet und keine Einschränkung bei bis zu 100 Teilnehmern pro Telefonkonferenz auf Standard-Hardware festgestellt.

3. Mittel- und Zeitplanung

Das Projekt wurde im geplanten Zeitraum durchgeführt. In diesem wurde das Hauptziel, die Implementierung des Telefonkonferenzsystems mit 3D-Audio-Funktionalität, erfolgreich umgesetzt. Durch den Mehraufwand für die Integration des Quellcodes nach Upstream waren weniger wurden die geplanten Nutzerstudien nur teilweise umgesetzt.

Die Zeitplanung wurde für alle Arbeitspakete eingehalten.

Die Mittelplanung wurde eingehalten. Lediglich die Mittel für Nutzerstudien wurden entsprechend der nicht durchgeführten Nutzerstudien nur anteilig abgerufen.

4. Fortschreibung des Verwertungsplans

4.1. Erfindungen, Schutzrechtanmeldungen und erteilte Schutzrechte

Es wurden keine Schutzrechte geplant zu beantragen, beantragt oder erteilt.

4.2. Wirtschaftliche Erfolgsaussichte nach Projektende

Die im Rahmen des Projektes erstellte Implementierung auf Basis von Asterisk wurde erfolgreiche in das Open-Source Projekt integriert. Dies vereinfacht eine zukünftige wirtschaftliche Verwertung, da die notwendige Instandhaltung der Software zu großen Teilen durch die Open-Source-Community hinter Asterisk durchgeführt wird.

4.3. Wissenschaftlich-technische Erfolgsaussichten nach Projektende

Die Ergebnisse des Vorhabens können nun für wissenschaftliche Forschung sowie Industrieforschung eingesetzt werden. Die Wartung durch die Asterisk-Community ist insbesondere für wissenschaftliche Forschung vorteilhaft, da in diesem Rahmen häufig keine Ressourcen zur zukunfts-sicheren Implementierung von benötigten Softwarelösungen zur Verfügung stehen.

4.4. Wissenschaftliche und wirtschaftliche Anschlussfähigkeit

Eine geplante wirtschaftliche Nutzung durch den Antragsteller oder andere Projektbeteiligte ist aktuell nicht vorgesehen. Da das techn. Projektergebnis als Open-Source verfügbar ist, kann eine wirtschaftliche Nutzung der Ergebnisse durch Dritte erfolgen. Über solche Aktivitäten hat der Autor zu diesem Zeitpunkt keine Kenntnis.

Das gleiche gilt für die wissenschaftliche Anschlussfähigkeit.

4.5. Erfolglose Arbeiten

Alle in diesem Vorhaben durchgeführten Arbeiten waren erfolgreich. Insbesondere die Neuordnung der Ressourcen des Arbeitspakets 4 (Nutzerstudien) für Arbeitspaket 2 und 3 (beide Implementierung) hat es ermöglicht, die ungeplante notwendige Mehrarbeit zur Integration der erstellten Software in die Open-Source Community zu ermöglichen. Ansonsten wäre die Wartung der Softwarelösung nach Projektende nicht gegeben und als Teil des Projektergebnisses als potentiell erfolglose Arbeit zu werten.

4.6. Während der Durchführung des Vorhabens bekannt gewordener Fortschritt auf dem Gebiet des Vorhabens bei anderer Stelle

Der Autor hat keine neuen Erkenntnisse über Fortschritte Dritter während der Durchführung des Vorhabens erhalten.

4.7. Präsentationsmöglichkeiten für mögliche Nutzer

Die im Rahmen des Projektes erstellte Webseite ([HTTP://WWW.STEAKCONFERENCING.DE](http://www.steakconferencing.de)) enthält eine umfangreiche Beschreibung der im diesen Projekt entstandenen Software und der theoretischen Grundlagen.

Eine implementierte Demo wird Mangels Ressourcen über die Projektlaufzeit nicht bereitgestellt. Die vollständige Dokumentation für das Setup der Demo findet sich im Github-Repository ([HTTPS://GITHUB.COM/STEAKCONFERENCING/DEMO](https://github.com/STEAKCONFERENCING/DEMO)). Diese erlaubt es potentiellen Nutzer das umgesetzte Demo-Setup zu reproduzieren.

4.8. Veröffentlichungen (geplant und erfolgt)

4.8.1. Erfolgte Veröffentlichungen

- Guse D., Haase F., „STEAK: Backward-Compatible Spatial Telephone Conferencing for Asterisk“, Audio Engineering Society Convention 141, September 2016, Los Angeles, CA, USA.
- Guse D., Haase F., „Presentation about binaural (positional) audio support“, AstriDevCon, September 2016, Phoenix, AZ, USA.
- Spur M., Guse D., Skowronek J.: „Influence of Packet Loss and Double-Talk on the Perceived Quality of Multi-Party Telephone Conferencing with Binaurally Presented Spatial Audio Reproduction“, März 2016, DAGA, Aachen, Deutschland.
- Projektwebseite: [HTTP://WWW.STEAKCONFERENCING.DE](http://www.steakconferencing.de)
- Quellcode: [HTTPS://GITHUB.COM/STEAKCONFERENCING](https://github.com/STEAKCONFERENCING)

4.8.2. Geplante Veröffentlichungen

- Guse D., Haase F., „TheTelephone: A Software-based Flexible Speech Telephony System for Conversational User Studies“, SoftwareX, Elsevier BV.