

Eine Fokusgruppe für die Entwicklung eines Vokabulars zur sensorischen Beurteilung virtueller akustischer Umgebungen

Alexander Lindau¹, Vera Erbes¹,
Steffen Lepa¹, Hans-Joachim Maempel¹, Fabian Brinkmann¹, Stefan Weinzierl¹

¹ *Fachgebiet Audiokommunikation, Technische Universität Berlin, E-Mail: alexander.lindau@tu-berlin.de*

Motivation und Stand der Forschung

Die sensorische Beurteilung virtueller akustischer Umgebungen (*Virtual Acoustic Environments*, VAEs) kann auf Basis integrativer Qualitätskriterien wie der *Plausibilität* [1] oder *Authentizität* [2] vorgenommen werden. Allerdings liefern solche Urteile kaum genauere Hinweise für gezielte technologische Verbesserungen. Für solche Zwecke sind qualitativ differenzierte Vokabulare besser geeignet. Resultate früherer Studien erscheinen jedoch entweder wenig einschlägig [3,4] verwendeten *ad hoc* Verfahren bei der Erstellung der Vokabulare [5,6] oder deckten nicht alle potentiell relevanten Aspekte ab [7]. Die empirisch fundierte Neuentwicklung eines Vokabulars zur differentiellen Diagnose sensorischer Unzulänglichkeiten in VAEs sämtlicher Typen und unter Berücksichtigung potentiell aller praktisch relevanten technischen Aspekte erschien uns daher wünschenswert.

Das Fokusgruppenverfahren

In [8] (S. 43 ff.) diskutieren Bech und Zacharov verschiedene Verfahrensarten für die Erhebung konsensueller sensorischer Vokabulare. Die sprachlich basierten Gruppenverfahren – und unter diesen vor allem das Fokusgruppenverfahren [9], eine nicht stimulusbasierte Expertenmethode – erschienen uns besonders dabei geeignet, da sie direkt verwertbare Ergebnisse liefern. Die Expertenfokusgruppe beruht auf der Annahme, dass ausgeprägte praktische und theoretische Expertise zielführende Diskussionen ermöglichen, ohne auf eine – im Falle von VAEs eventuell nur schwer zu beschaffende – repräsentative Auswahl expliziter Beispielstimuli angewiesen zu sein. Die Partnerschaft in einem größeren Forschungskonsortium zum Thema Virtuelle Akustik (DFG Forschergruppe SEACEN) ermöglichte den Autoren zudem einen besonders einfachen Zugang zu den benötigten Experten.

Methodisch kann eine Expertenfokusgruppe – zunächst bestehend aus einem sowohl fachlich als auch interviewtechnisch geschulten Moderator und einer Gruppe von Experten als Diskutanten – als Kombination aus Leitfadeninterview und Gruppengespräch betrachtet werden. Diese Kombination gilt als ein besonders geeigneter Zugang zu Expertenwissen betrachtet, insofern als letztere es routinemäßig gewohnt sind, zu einem diskursbasierten Erkenntniskonsens zu gelangen [10]. Mit der Erweiterung auf das sog. *dual-moderator* Verfahren können thematische Unvollständigkeit und irrelevante Diskussionen reduziert werden: Dabei übernimmt ein Hauptmoderator die Diskussionsleitung während ein Ko-moderator die Einhaltung der inhaltlichen und verfahrenstechnischen Agenda überwacht. Ergänzend wurde eine aus Mitgliedern der Diskussionsgruppe rekrutierte Beobachtergruppe eingesetzt (*two way Focus Group*). Dabei sollten die Kommentare der distanzierter, und

damit objektiver urteilenden Beobachter Vollständigkeit und Objektivität des Verfahrens weiter befördern.

Agenda

Den Leitfaden der Gruppendiskussionen bildete die geplante Erstellung eines konsensuellen, deskriptiven Vokabulars, bzw. eines Fragebogens für die sensorische Beurteilung aller Entwicklungsarten und -stufen von VAEs. Dabei sollte dieses Vokabular alle auditiv wahrnehmbaren Unterschiedsqualitäten umfassen, die bei Vergleichen von VAEs zur – real präsentierten oder auch nur vorgestellten – akustischen Realität, oder zu anderen VAEs auftreten könnten. Dabei wurden die Experten instruiert, ein ausgewogenes Verhältnis von Vollständigkeit und praktischer Relevanz anzustreben. Schließlich sollten die finalen Qualitätsbezeichner semantisch möglichst eindeutig und wechselseitig exklusiv formuliert und für andere Experten selbsterklärend sein. Nur wenn unbedingt nötig, sollte es möglich sein, noch kurze ergänzende Präzisierungen, Beispiele für Skalenlabel (Tonhöhe: ‚hoch – tief‘) oder sogar die Erstellung illustrativer Klangbeispiele zu vereinbaren.

Diskussionsverlauf

Nach Empfehlungen in [11] wurden die Teilnehmer zur Steigerung von Motivation und inhaltlicher Fokussierung in alle wesentliche Entscheidungsprozesse einbezogen (Abstimmungen mit Mehrheitsentschied). Es zeigte sich jedoch, dass nach gründlichen (und zuweilen langen) Diskussionen, meist Einstimmigkeit erreicht werden konnte. Nach dem Bekanntgeben der Diskussionsregeln, begannen die ersten beiden Diskussionsrunden mit je einem 20-minütigen Brainstorming. Diese lieferten 62 vorläufige Begriffe als Ausgangsbasis der folgenden Diskussionen. Eine Diskussionsrunde dauerte 3,5 Std. Pro Tag wurden maximal 2 Runden durchgeführt. Insgesamt wurden während mehrerer Gelegenheiten in Berlin und Aachen 16 Diskussionsrunden (56 Std.) durchgeführt. An den Diskussionen nahmen jeweils ca. 10-15 teils wechselnde Experten aus einer Gesamtgruppe von 21 Personen teil. Jede neue Runde begann mit einer randomisierten Aufteilung in Diskutanten und Beobachter (Verhältnis im Mittel etwa 70:30). Danach begann eine erste 90-minütige Diskussion. Nach einer 20-minütigen Konsolidierungspause kehrten die Beobachter in die Gruppe zurück und diskutierten ihre Eindrücke innerhalb weiterer 90 Minuten. Alle Beobachterkommentare wurden festgehalten, bis sie innerhalb der Gruppe zufriedenstellend ausdiskutiert werden konnten. Die letzte Runde eines Treffens endete jeweils mit einer Feedbackrunde.

Fertigstellung

Die Gruppendiskussionen lieferten ein vorläufiges Vokabular, von 49 Begriffen welche wie folgt final bearbeitet wurden: Zunächst wurde das konsensuelle inhaltliche Verständnis aller Begriffe überprüft. Dazu glichen die 12 Experten mit der

höchsten Teilnahmeghäufigkeit ihre je individuell verschriftlichten Umschreibungen aller Begriffe miteinander ab. Dabei wurden auch – wie vereinbart – illustrative Klangbeispiele für drei der Begriffe erzeugt. Danach wurde die Verständlichkeit des Vokabulars zusätzlich von fünf externen Experten beurteilt. Dazu lieferten diese wiederum Umschreibungen ihres Verständnisses aller Begriffe und Kommentare zur Einschlägigkeit der Klangbeispiele. Diese etwa 250 Aussagen wurden auf semantische Übereinstimmung mit den vorherigen Gruppensätzen überprüft. Dabei wurden für etwa ein Drittel der Begriffe kleinere Missverständnisse identifiziert. Diese wurden in Gruppensitzungen mit fünf der ursprünglichen Experten behoben. Dabei wurde auch ein Begriff als obsolet bestätigt, passende Skalenlabel für alle Begriffe vereinbart und die finalen Umschreibungen aller Qualitäten in das Vokabular übernommen.

Resultate

Das finale Vokabular wurde *Spatial Audio Quality Inventory* (SAQI) benannt. Es beinhaltet 48 Begriffe für auditive Unterschiedsqualitäten, die sich acht Kategorien gruppieren lassen (s. Tabelle 1).

Tabelle 1: Begriffe des deutschen SAQI, [12,14]

Kategorie	Qualitätsbezeichner („Wahrnehmung eines Unterschieds bzgl. ...“)
Klangfarbe (8)	<i>Klangfarbe hell-/dunkel, Klangfarbliche Ausprägung im Höhen-/Mitten-/Tiefenbereich, Schärfe, Rauigkeit, Kammfilterartigkeit, Metall, Klangfarbe</i>
Tonalität (3)	<i>Tonhaltigkeit, Tonhöhe, Dopplereffekt</i>
Geometrie (10)	<i>Richtung Azimut/Elevation, Vorn-Hinten-Lage, Entfernung, Tiefen-, Breiten-, Höhenausdehnung, Externalisierungsgrad, Lokalisierbarkeit, Räumliches Zerfallen</i>
Raum (3)	<i>Nachhallstärke, Nachhalldauer, Nachhallumhüllung</i>
Zeitverhalten (7)	<i>Vor-, Nachechos, Zeitliches Zerfallen, Knackigkeit, Wiedergabegeschwindigkeit, Szenenablauf, Reaktionsschnelligkeit</i>
Dynamik (3)	<i>Lautheit, Dynamik, Kompressor-effekte</i>
Artefakt (7)	<i>Tonhaltiges/Impulshaftes/Rauschhaftes Fremdgeräusch, Fremdquelle, Geisterquelle, Verzerrungen, Vibration</i>
Allgemeines (7)	<i>Unterschied, Klarheit, Sprachverständlichkeit, Natürlichkeit, Präsenz, Gefallen, Sonstiges</i>

Folgend einem *bottom-up* Verständnis auditiver Wahrnehmung zeigen manche dieser Attribute starke Bezüge zu basalen zeitlichen oder spektralen Signaleigenschaften. Andere wiederum beziehen sich offensichtlich auf spezifisch-technologische Aspekte von VAEs. Wiederum andere, eher wertende, integrative oder einstellungsbezogene Aspekte entsprechen eher *top-down* Prozessen der Urteilsbildung. Das vollständige Vokabular ergänzt jeden Begriff um die erwähnte inhaltliche Umschreibung, um geeignete Endlabel für Ratingskalen und, in einigen Fällen, um illustrative Klangbeispiele. Die Unterschiedsqualität ‚Sonstiges‘ ist enthalten, um potentiell übersehene oder zukünftig neu entstehende Qualitäten abfangen zu können¹. Weiterhin – und wie in [12,14] näher ausgeführt – wurde ein prototypisches geschlossenes System von Bezugsobjekten vorgeschlagen, um wahrgenommene Unterschiede typischen Teilelementen von VAEs zuordnen zu können (Teil der simulierten Szene: Vordergrundquellen / Hintergrundquellen / raumakustische Umgebung, nicht Teil der simulierten Szene: Übertragungssystem / Laborumgebung). Weiterhin wurde es als praktisch

relevant empfunden, wahrgenommene Unterschiede nach Zeitvarianz (konstant / periodisch-regelhaft / nicht regelhaft, stetig / unstetig) und Ursächlichkeit (nutzerverursacht / szenenverursacht / unabhängig) weiter differenzieren zu können. Vollständige deutsche und englische Versionen des SAQI sind in [12] dokumentiert. Um die Anwendung zu erleichtern, wurde das SAQI als weitgehend parametrierbarer Hörversuch in die freie Matlab® Hörversuchssoftware whisPER (vs. 1.8.0) integriert [13].

Diskussion und Ausblick

Vergleicht man unser Endergebnis mit den Resultaten oben genannter Vorstudien, zeigt sich, dass das SAQI auch alle bereits zuvor genannten relevanten Aspekte enthält. Dieses Ergebnis stützt unseren Eindruck, mit der Expertenfokusgruppe ein effektives Verfahren zur Erstellung fachspezifischer konsensueller deskriptiver Vokabulare gefunden zu haben. Die Objektivität des Verfahrens wurde durch mehrfache Mechanismen der Selbstkontrolle überwacht. Das SAQI wird derzeit in verschiedenen experimentellen Studien des SEACEN-Konsortiums eingesetzt. Erste Ergebnisse ermöglichen die Beurteilung der praktischen Anwendbarkeit und der Erfüllung einschlägiger Testgütekriterien. Zukünftige Anwendungsfelder des SAQI bilden die gezielte technische (Weiter-)Entwicklung von VAES, deren perzeptiv motivierte Aufwandsreduktion oder ein qualitativ differenziertes Benchmarking.

Danksagung

Die Durchführung dieser Studie wurde durch die Deutsche Forschungsgemeinschaft gefördert (DFG WE 4057/1-3). Die Autoren möchten zudem allen Teilnehmern dieser Studie (s. [14]) ihren herzlichen Dank ausdrücken.

Literatur

- [1] Lindau, A.; Weinzierl, S. (2012): "Assessing the Plausibility of ...", in: *Acta Acustica united with Acustica*, 98(55), 804-810
- [2] Brinkmann, F.; Lindau, A.; Vrhovnik, M.; Weinzierl, S. (2014): "Assessing the Authenticity of ...", in: *Proc. of the EAA Joint Symposium on Auralization and Ambisonics*, Berlin
- [3] Berg, J.; Rumsey, F. (1999): "Spatial Attribute Identification ...", in: *Proc. of the 16th International AES Conference*, Rovaniemi
- [4] Lorho, G. (2005): "Individual Vocabulary Profiling ...", in: *Proc. of the 119th AES Convention*, New York, preprint no. 6629
- [5] Pellegrini, R. S. (2001): "Quality Assessment of ...", in: *Proc. of ICAD 2001*, Espoo
- [6] Lokki, T.; Järveläinen, H. (2001): "Subjective evaluation of ...", in: *Proc. of ICAD 2001*, Espoo
- [7] Silzle, Andreas (2007): "Quality Taxonomies for ...", in: *Proc. of the 122nd AES Convention*, Vienna, preprint no. 6993
- [8] Bech, S.; Zacharov, N. (2006): *Perceptual Audio Evaluation: Theory, Method and Application*, Chichester: Wiley
- [9] Stewart, D. W.; Shamsdasani, P. N.; Rook, D. W. (1990): *Focus groups: Theory and practice*, Newbury Park, CA: Sage
- [10] Bogner, A.; Leuthold M. (2005): "Was ich dazu noch sagen wollte...". Die Moderation von Experten-Fokusgruppen", in: Bogner, A.; Littig, B.; Menz, W. (eds.): *Das Experteninterview*, Wiesbaden: VS Verlag für Sozialwissenschaften
- [11] Dürrenberger, G.; Behringer, J. (1999): *Die Fokusgruppe in Theorie und Anwendung*, Stuttgart: Rudolph-Sophien-Stift gGmbH
- [12] Lindau, A. (2014): *SAQI Test Manual*, v1.0, <http://dx.doi.org/10.14279/depositonce-1>
- [13] Ciba, S.; Wlodarski, A.; Maempel, H.-J. (2009): "WhisPER – A new tool for performing listening tests", in: *Proc. of the 126th AES Conv.*, Munich, preprint 7749, <http://www.ak.tu-berlin.de/whisper>
- [14] Lindau, A. et al. (2014): "A Spatial Audio Quality Inventory for Virtual Acoustic Environments (SAQI)", in: *Proc. of the EAA Joint Symposium on Auralization and Ambisonics*, Berlin

¹ Alle Anwender sind herzlich eingeladen, dem Erstautor ihre Erfahrungen mit dieser Kategorie mitzuteilen.