

# Verwendung von Audioobjekten im Fahrzeug

Christoph Sladeczek, Mario Seideneck, Wolfgang Lorenz, Anja Chilian, Draško Mašović, Katrin Pursche, Joachim Bös

**Aktuelle Studien zeigen, dass sich der Fahrzeuginnenraum so stark verändern wird wie seit Jahrzehnten nicht mehr. Neue Nutzungskonzepte verlagern den Fokus auf das Innenraumerlebnis. Damit verbunden sind völlig neue Anforderungen an das Audiosystem. Die Verfügbarkeit immersiver Unterhaltungstechnologien, die für neue Komfortfunktionen und mobile Arbeiten genutzt werden, wird essenziell. Durch den veränderten Aufmerksamkeitsfokus wird nicht nur die personalisierte Klanginszenierung selbst, sondern auch deren korrekte räumliche Abbildung eine neue Bedeutung erhalten. Dies gilt für eine Vielzahl von Funktionen wie Komfort, Fahrsicherheit, Wohlbefinden oder Kommunikation. In diesem Zusammenhang wird der Aufwand für die Erstellung von Audioinhalten eine wesentliche Rolle spielen, was eine neue, einheitliche Schnittstelle für die räumliche Darstellung erfordert, um den Produktionsaufwand zu begrenzen. Dieser Beitrag beschreibt, wie objektbasiertes Audio (OBA) als Schnittstellentechnologie genutzt werden kann, um diese Anforderungen zu erfüllen.**

## Einleitung

Der Fahrzeuginnenraum wird sich in naher Zukunft stark verändern. Viele Fahraufgaben werden zunehmend von intelligenten Assistenzsystemen übernommen. Dadurch verlagert sich der Fokus der Insassen von der Straße auf das Erlebnis im Innenraum, und die Schnittstelle zwischen fahrender Person und Fahrzeug erfährt neue Aufmerksamkeit [1]. Heutige Assistenzsysteme konzentrieren sich häufig auf visuelle Informationen und schöpfen das Potenzial des akustischen Kanals nicht voll aus, obwohl das Ohr permanent empfängt und Informationen aus allen Richtungen wahrnimmt [2, 3]. Um den steigenden Ansprüchen an Komfort, Kommunikation und Sicherheit gerecht zu werden, spielt die auditive Schnittstelle eine immer wichtigere Rolle. Der Einsatz immersiver Technologien ist von entscheidender Bedeutung, um das Erlebnis im Fahrzeug neu zu definieren [4]. Damit steigen auch die Anforderungen an die Audiosysteme im Fahrzeug, um sie für neue Anwendungsbereiche nutzbar zu machen. Gleichzeitig besteht seitens der Fahrzeughersteller der Wunsch, alle Audiofunktionen über ein zentrales Soundsystem wiederzugeben, so dass keine zusätzlichen Lautsprecher für spezifische

## Using audio objects in the vehicle

**Current studies show that vehicle interiors will change more than they have for decades. New usage concepts shift the focus to the interior experience. This is associated with completely new demands on the audio system. The availability of immersive entertainment technologies, which are used for new comfort functions and mobile working, becomes essential. The changed focus of attention means that not only the personalized sound staging itself, but also its correct spatial representation will take on a new significance. This applies to a wide range of functions such as comfort, driving safety, well-being, and communication. In this context, the effort required to create audio content will play a significant role, requiring a new, standardized interface for spatial representation in order to limit the production effort. This article describes how object-based audio (OBA) can be used as an interface technology to meet these requirements.**

Funktionen eingebaut werden müssen.

Neue Anwendungen zukünftiger Audiosysteme sind in Abbildung 1 dargestellt. Ziel ist es, interaktive, immersive Hörerlebnisse im Fahrzeug zu schaffen, die nicht nur der Unterhaltung dienen, sondern auch ein akustisches Abbild der Fahrzeugumgebung und der Fahrzeugfunktionen selbst ermöglichen. Durch eine räumlich korrekte Wiedergabe von Warnsignalen kann die Aufmerksamkeit auf potenzielle Gefahrenquellen außerhalb des Fahrzeugs gelenkt werden. Die Wahrnehmung des Fahrzeugverhaltens kann aktiv unterstützt werden, indem künstlich generierte Funktionsklänge ihre räumliche Position fahrdynamisch verändern oder sich immersiv im Fahrzeuginnenraum ausbreiten. Darüber hinaus kann eine Individualisierung des Fahrerlebnisses erreicht werden, indem die gesamte Innenraumakustik im Fahrzeug an persönliche Vorlieben angepasst wird.

Da die Anwendungen in unterschiedlichen Fachabteilungen der Automobilhersteller entwickelt werden, ist die Art und Weise, wie Audioinhalte für die verschiedenen Fahrzeugplattformen kreiert werden, ein entscheidender Faktor für die Bestimmung des erforderlichen Audioproduktionsaufwands. Gleichzeitig



Abb. 1: Illustrierte Verwendung von Audioobjekten im Fahrzeug für unterschiedliche Anwendungsfälle. Während räumliche Entertainment-inhalte wiedergegeben werden (violett), erfolgt auch die richtungskorrekte Wiedergabe eines akustischen Warnsignals (blau) und eines Funktionsklangs (grün) über das Audiosystem.

führt der Trend immersiver Wiedergabetechnologien zu höheren Qualitätsansprüchen der Nutzerinnen und Nutzer an die Audiosysteme, sodass neben der klanglichen Brillanz auch eine hohe räumliche Wiedergabequalität gefordert wird. Der objektbasierte Audioansatz bietet als interaktive Schnittstelle und Gesamtsystemarchitektur neue Möglichkeiten, diese Anforderungen zu erfüllen.

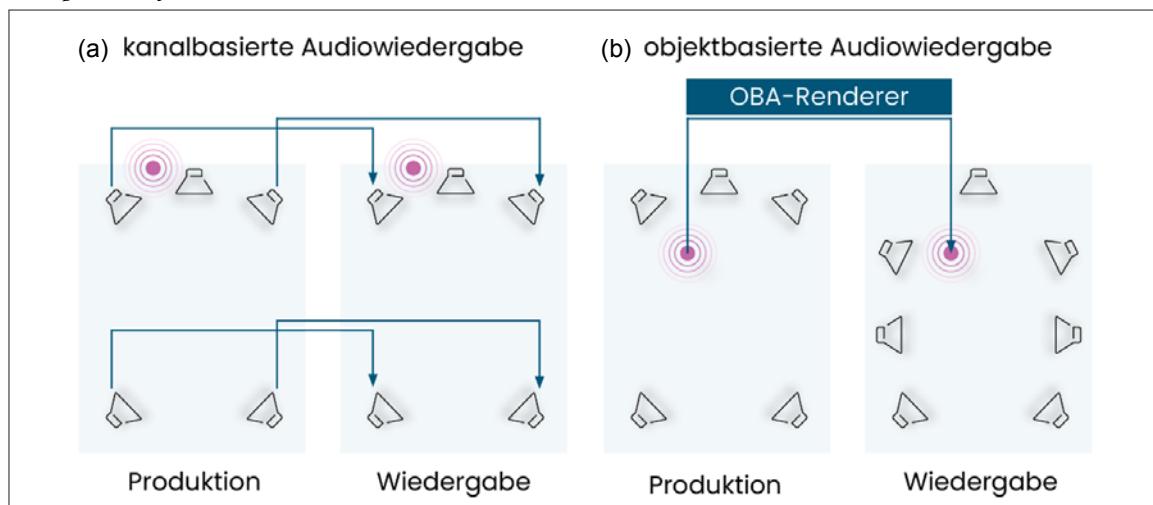
### Objektbasiertes Audio

Bei der klassischen kanalbasierten Audioproduktion werden Lautsprechersignale als Ergebnis der Mischnung gespeichert. Da die räumliche Information des Audiosignals in den Lautsprechersignalen ko-

diziert ist, müssen für eine korrekte räumliche Wiedergabe sowohl die Anzahl als auch die Positionen der Wiedergabelautsprecher identisch mit der Produktions situation sein. Das zeigt Abbildung 2(a).

Da dies im Fahrzeug in der Regel nicht gewährleistet werden kann, wird seit jeher versucht, durch Remix-, Downmix- und Upmix-Algorithmen die räumliche Information in den gespeicherten Lautsprechersignalen so zu verändern, dass die neu generierten Ausgangssignale an die Anzahl und Positionierung der Lautsprecher angepasst sind [5]. Hierzu werden einfache Algorithmen, wie die gewichtete Addition der Ausgangskanäle, bis hin zu komplexen Algorithmen zur Quellentrennung verwendet. Bei stark begrenzten

Abb. 2: Kanalbasierter und objektbasierter Audiowiedergabeansatz. Während bei der kanalbasierten Audiowiedergabe (a) das Lautsprecher setup identisch bei Produktion und Wiedergabe sein muss, damit Audioinhalte korrekt räumlich wiedergegeben werden können, erlaubt die objektbasierte Wiedergabe (b) die Verwendung unterschiedlicher Lautsprecher aufbauten.



Prozessorressourcen ist eine solche Anpassung immer kompromissbehaftet. Ein wesentlicher Nachteil ist der hohe Produktionsaufwand, der sich aus der Kombination von fehlender Flexibilität und gleichzeitiger Variantenvielfalt der Fahrzeugplattformen ergibt, insbesondere durch unterschiedliche Lautsprechersetups in diversen Ausstattungsvarianten. Hinzu kommt, dass die Abbildung von Quellbewegungen nicht möglich ist, da die beschriebenen Algorithmen keine interaktive Anpassung der räumlichen Audioinhalte erlauben. Im Gegensatz zur kanalbasierten Audioproduktion liefert eine objektbasierte Produktion keine fertigen Lautsprechersignale. Objektbasiertes Audio ist ein Konzept für die Speicherung, Übertragung und Wiedergabe von Audioinhalten, bei der Audioobjekte durch ein Audiosignal und zugehörige Metadaten definiert sind [6]. Die Metadaten beschreiben Eigenschaften der Audioobjekte wie Position, Typ oder Verstärkung. Eine Audioszene besteht aus mehreren Audioobjekten, die statische oder dynamische Eigenschaften haben können. Der Vorteil gegenüber kanalbasierten Methoden besteht darin, dass die so definierten Audioszenen unabhängig von der Lautsprecherkonfiguration sind, da die Erzeugung der Lautsprechersignale erst unmittelbar im Wiedergabeprozess stattfindet. Der Prozess der Berechnung von Wiedergabesignalen wird als Rendering bezeichnet. So müssen im Idealfall Inhalte nur einmal produziert werden, da sie durch eine einfache Konfiguration des Renderings auf einem nahezu beliebigen Lautsprechersetup wiedergegeben werden können. Dies zeigt schematisch Abbildung 2(b). Für die Wiedergabe von Audioobjekten über Lautsprecher stehen verschiedene Methoden zur Verfügung, die von einfachen Stereofonieverfahren bis hin zu komplexen Schallfeldrekonstruktionsverfahren reichen [7, 8]. Mathematisch kann das objektbasierte Rendering im Frequenzbereich durch

$$S(\vec{r}, \omega) = \sum_l T_l(\omega) Q_l(\vec{r}_0, \omega) F_l(\omega) G_l(\vec{r}|\vec{r}_0, \omega)$$

beschrieben werden. Das Schallfeld  $S(\vec{r}, \omega)$  eines Audioobjekts mit der Kreisfrequenz  $\omega = 2\pi f$  an einer beliebigen Hörposition  $\vec{r}$  innerhalb eines Hörbereichs ergibt sich durch die Superposition der Schallfelder jedes beteiligten Lautsprechers  $l$  an der Position  $\vec{r}_0$ , der mit der Funktion

$$Q_l(\vec{r}_0, \omega) = L_l(\vec{r}_0) C_l(\omega) D_l(\omega) A(\omega)$$

angesteuert wird. Die in der Theorie häufig verwendete Greensche-Funktion  $G_l(\vec{r}|\vec{r}_0, \omega)$  einer Monopolschallquelle gilt nur unter Freifeldbedingungen. Reale Übertragungseigenschaften wie Richtcharakteristik und Raumakustik sind in  $F_l(\omega)$  zusammengefasst.

Um eine möglichst vergleichbare Abstrahlfunktion aller Lautsprecher zu erreichen, wird die Korrekturfunktion  $T_l(\omega)$  im Tuningprozess bestimmt. Sie berücksichtigt Eigenschaften wie Laufzeit, Lautstärke, Kennlinie und Abstrahlverhalten des Lautsprechers, die in der Regel durch akustische Messungen ermittelt werden. Abhängig von der gewählten Schallfeldrekonstruktionsmethode kann  $Q_l(\vec{r}_0, \omega)$  verschiedene Terme aufweisen. Im Beispiel ergeben sich bei einer Metadatenänderung Verstärkungs-  $C_l$  und Verzögerungswerte  $D_l$ , die auf das Audioeingangssignal  $A$  pro Lautsprecher angewendet werden. Da übliche Lautsprecherkonfigurationen in Fahrzeugen aus einigen wenigen bis zu mehreren Dutzend Lautsprechern bestehen können, die eine dreidimensionale Verteilung aufweisen, beinhalten objektbasierte Rendering-Technologien häufig eine Funktion  $L(\vec{r}_0)$  zur Auswahl relevanter Lautsprecher in Abhängigkeit von der Position des Audioobjekts. Hierbei ist es wichtig, dass die Gesamtlautstärke im Wiedergaberaum unabhängig von der Position der virtuellen Schallquelle ist. Das bedeutet, dass je nach Anzahl der Lautsprecher, die an der Wiedergabe eines einzelnen Objekts beteiligt sind, eine Lautstärkekorrektur vorgenommen werden muss. Dies ist insbesondere bei irregulären Lautsprecheranordnungen der Fall, d.h. wenn die Lautsprecher des Wiedergabesystems unregelmäßig im Raum verteilt sind und eventuell sogar unterschiedliche Abstände zur Raummitte haben.

Die Ansätze der existierenden OBA-Algorithmen zur Berechnung von  $Q_l(\vec{r}_0, \omega)$  unterscheiden sich zum Teil sehr. Gemein ist ihnen, dass zur Konfiguration des Renderings oftmals die realen Positionen der Lautsprecher im Wiedergaberaum verwendet werden. Die Metadaten der Audioobjekte, die für die Berechnung der genannten Koeffizienten herangezogen werden, sind üblicherweise die Positionen und die Lautstärke des wiederzugebenden Signals. Einige Algorithmen erlauben die Benutzung zusätzlicher Objekteigenschaften, wie die Abstrahlcharakteristik und die räumliche Ausdehnung.

## Erzeugung von Metadaten für objektbasiertes Rendering im Fahrzeug

Der Einsatz des objektbasierten Audioansatzes als Schnittstellentechnologie im Fahrzeug ist in Abbildung 3 dargestellt. Der objektbasierte Audiorenderer ist eine Software, die interaktiv die Lautsprechersignale berechnet. Das im Fahrzeug verbaute Lautsprechersetup wird dem Audiorenderer über eine Konfiguration bekannt gemacht. Darin finden sich üblicherweise Informationen zu Position und Typ der verwendeten Lautsprecher. Um aus den Audioobjekten Lautsprechersignale für ein definiertes Lautsprechersetup

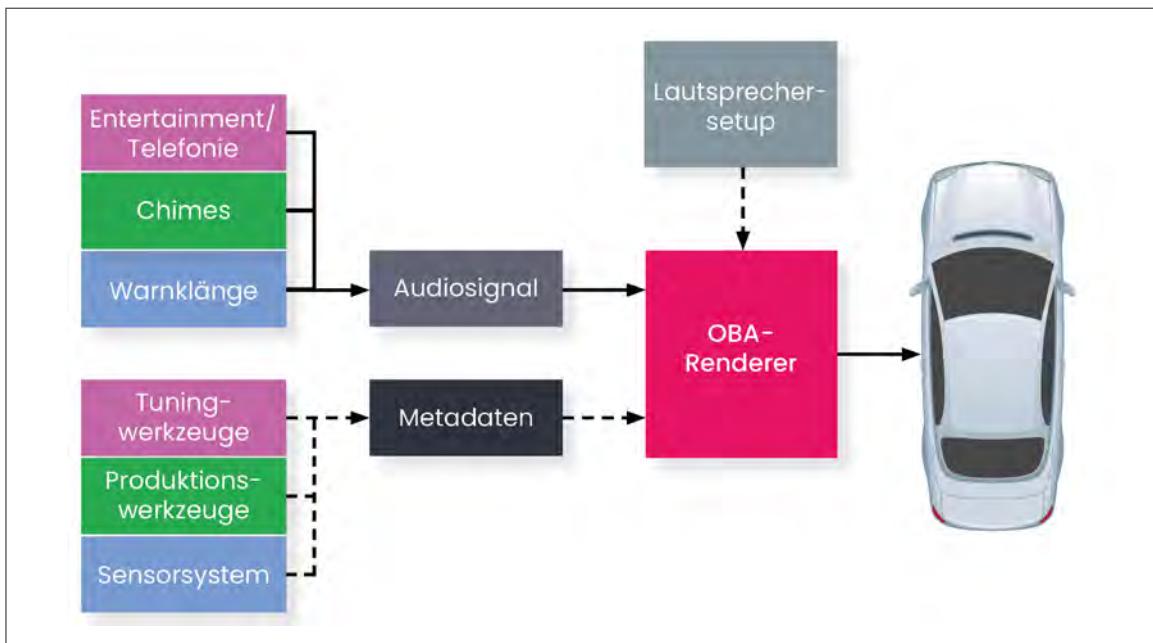


Abb. 3: Konzept zur Verwendung von Audioobjekten im Fahrzeug. Der objektbasierte Audiorenderer kennt das konkrete Lautsprechersetup im Fahrzeug, wie z. B. Position und Typ der Lautsprecher. Anwendungsabhängig erfolgt die interaktive Zuspielung von Audiosignalen synchronisiert mit den zugehörigen Metadaten. Sowohl die Metadaten als auch die Audiodaten können auf unterschiedliche Art und Weise generiert werden.

zu generieren, müssen dem Renderer Audio- und Metadaten synchronisiert zur Verfügung gestellt werden. Der Audiorenderer benötigt hingegen keinerlei Informationen darüber, aus welcher Quelle der Audioinhalt und die Metadaten stammen. Dies ermöglicht eine unabhängige Verwendung für verschiedene Anwendungen und eine einfache Entwicklung. Zur Realisierung interaktiver Audiosysteme wird eine pufferbasierte Audiosignalverarbeitung verwendet. Dabei werden die Audiodaten in kleine Segmente, sogenannte Puffer, unterteilt, die anschließend nacheinander bearbeitet werden. Um die Synchronisation zwischen Audio- und Metadaten zu ermöglichen, muss das Verarbeitungssystem sicherstellen, dass für die Bearbeitung eines Audiopuffers immer der relevante Metadatensatz zur Verfügung steht.

Abbildung 3 zeigt die objektbasierte Audioschnittstelle beispielhaft für verschiedene Automotive-Anwendungen. Während bei Telefonie und Entertainment die Audiosignale live in das Audiosystem gestreamt werden, handelt es sich bei den Chimes (z. B. Willkommens- oder Interaktionsklänge) und Warnklängen um vorproduzierte Inhalte, die zum Wunschzeitpunkt abgespielt werden. Die zugehörigen Metadaten können verschiedenen Quellen entstammen und unterschiedliche zeitliche Auflösungen haben. Es wird zwischen statischen, automatisierten und interaktiven Audioobjekten unterschieden [9].

#### Statische Audioobjekte:

Für Anwendungen im Bereich Entertainment oder Telefonie können Audioobjekte verwendet werden, deren Eigenschaften, wie z. B. Position und Typ, zeitlich konstant sind. Mithilfe eines Produktionswerkzeugs wird eine Audioszene erstellt, die durch einen statischen Satz an Metadaten beschrieben wird. Dadurch können beispielsweise kanalbasierte Audiodaten auf beliebigen Lautsprecheranordnungen wiedergegeben werden, indem statische Audioobjekte als virtuelle Lautsprecher genutzt werden. Dies zeigt Abbildung 4(a) auf der folgenden Seite.

#### Automatisierte Audioobjekte:

Sollen bestimmte Klangereignisse mit verschiedenen Funktionen im Fahrzeug verknüpft werden, können automatisierte Audioobjekte zum Einsatz kommen. Die Wiedergabe der vorproduzierten Audio- und Metadaten wird durch ein bestimmtes Ereignis ausgelöst. Dies erfordert ein Produktionswerkzeug, in dem man Positions- und Eigenschaftsänderungen der Audioobjekte über einen bestimmten Zeitraum definiert. Die so generierten Metadaten können als Wertefolgen abgespeichert werden. Gestartet durch einen Auslöser werden diese Wertefolgen an den Audiorenderer gesendet. Eine mögliche Anwendung für automatisierte Audioobjekte ist das Abspielen eines individuellen, sich dreidimensional um die fahrende Person bewegenden Willkommensklangs beim Einstieg in das Fahrzeug. In Abbildung 4(b) (siehe folgende Seite) wird dies beispielhaft gezeigt.

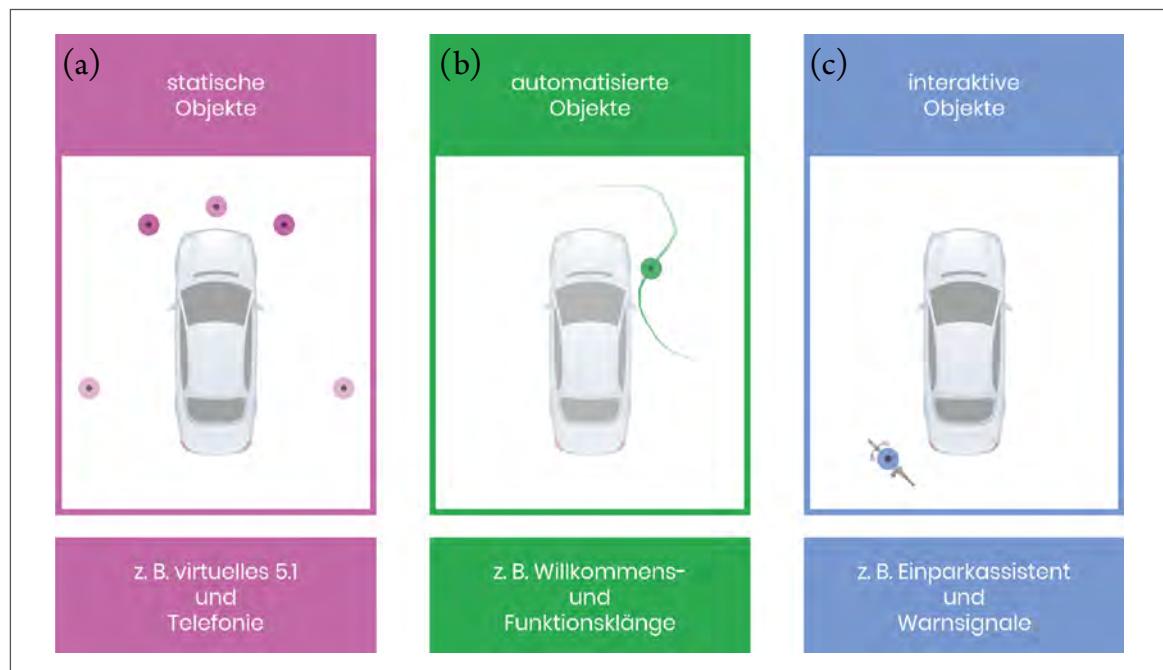
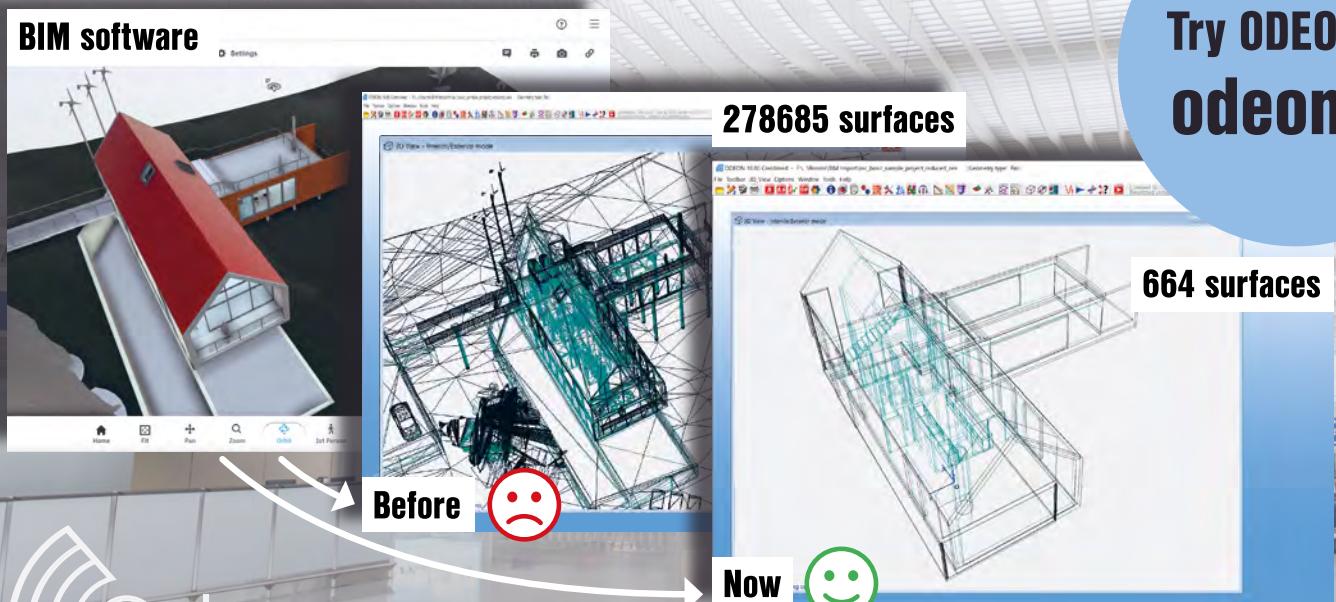


Abb. 4: Anwendungsabhängige Erzeugung von Metadaten für die objektbasierte Audiowiedergabe.

- (a) zeigt die Verwendung von statischen Metadaten für Audioobjekte zur Erzeugung virtueller Lautsprecheranordnungen.
- (b) symbolisiert eine vorproduzierte Bewegungsbahn eines automatisierten Audioobjekts, das durch eine Zeitreihe von Metadatenwerten beschrieben und ab Zeitpunkt der Auslösung »abgespielt« werden kann.
- (c) zeigt interaktive Metadaten, die z. B. durch einen Außensensor zur Hinderniserkennung direkt an das objektbasierte Audiosystem übertragen werden.

# IFC and BIM import made possible!

Enjoy great new features with the latest version of ODEON,  
including automated import from REVIT and ArchiCAD



Try ODEON 18  
[odeon.dk](http://odeon.dk)

### Interaktive Audioobjekte:

Für Anwendungen, die Informationen über bestimmte Ereignisse in der Umgebung des Fahrzeugs liefern, können interaktive Audioobjekte zum Einsatz kommen. Dies wird in Abbildung 4(c) dargestellt. Hierbei interagiert z. B. ein externes Sensorsystem direkt mit dem Audiosystem im Fahrzeug. Die bereitgestellten Sensordaten lösen nicht nur die Audiowiedergabe aus, sondern beeinflussen gleichzeitig die Eigenschaften des Audioobjekts, indem sie interaktiv die Metadaten generieren. So kann die Position der Audiosignale realen Ereignissen in der Fahrzeugumgebung zugeordnet werden. Hierfür müssen die eingehenden Sensordaten wahrnehmungsbezogen umgerechnet und auf vordefinierte Grenzwerte (z. B. für die Verstärkung) beschränkt werden. Automatisierte Audioobjekte, die räumlich eindeutig im Fahrzeug platziert werden, können insbesondere für Infotainment-Anwendungen zum Einsatz kommen. Beispiele hierfür sind die akustische Warnung vor Hindernissen bei einem Einparkassistenten oder Warnsignale, die die Aufmerksamkeit auf die Position gefährdet Verkehrsteilnehmender lenken. So ermöglichen interaktive Audioobjekte eine intuitive 360°-Wahrnehmung der Fahrzeugumgebung, was zur Verbesserung der Fahrsicherheit beitragen kann [2].

### **Zusammenfassung und Ausblick**

Aktuelle Entwicklungstrends in der Automobilindustrie führen dazu, dass immer mehr Funktionen über das eingebaute Audiosystem abgebildet werden müssen. Gleichzeitig steigt die Erwartungshaltung der nutzenden Personen an die dargebotene räumliche Klangqualität. Durch die angebotene Variantenvielfalt der Fahrzeuge ist eine Steigerung der Aufwände zur Audioinhalterstellung und -anpassung zu erwarten. Eine Umstellung von der kanalbasierten Audioproduktion hin zur objektbasierten Audioproduktion bietet das Potenzial, den gestiegenen Anforderungen an das Audiosystem hinsichtlich Immersion und Interaktivität gerecht zu werden und gleichzeitig die Produktionsaufwände signifikant zu reduzieren. Die ersten Fahrzeuge mit integrierter objektbasierter Audioplattform gehen ab dem Jahr 2025 in Serie.

### **Literatur**

- [1] Möller, T.; Schneiderbauer, T.; Garms, F.; Gläfke, A.; Köster, N.; Stegmüller, S.; Kern, M.; Werner, M.; Bobka, K.: The Future of Interior in Automotive. 12. November 2021. <https://www.mckinsey.com/industries/automotive-and-assembly/our-insights/the-future-of-interior-in-automotive>
- [2] Yost, W.: Fundamentals of Hearing: An Introduction. Amsterdam, Elsevier Science Publishing Co Inc, 2006. <https://doi.org/10.1121/1.1398047>
- [3] Laack, A.; Torschmied, A.; Tuzar, G.-D.: Immersive Audio HMI to Improve Situational Awareness. In: Mensch und Computer 2015 Workshopband, Stuttgart, S. 501–508, 2015. <https://doi.org/10.1515/9783110443905-071>
- [4] Kovačević, J.; Kaprocki, N.; Popović, A.: Review of Automotive Audio Technologies: Immersive Audio Case Study. In: 2019 Zooming Innovation in Consumer Technologies Conference (ZINC), S. 98–99, 2019. <https://doi.org/10.1109/ZINC.2019.8769396>
- [5] Bai, M.R.; Shih, G.: Upmixing and Downmixing Two-channel Stereo Audio for Consumer Electronics. In: IEEE Transactions on Consumer Electronics. Vol. 53, no. 3, S. 1011–1019, August 2007. <https://doi.org/10.1109/TCE.2007.4341580>
- [6] Gasull Ruiz, A.; Sladeczek, C.; Sporer, T.: A Description of an Object-Based Audio Workflow for Media Productions. In: 57th International Conference of the Audio Engineering Society (AES), Hollywood, CA, USA, 2015.
- [7] Brix, S.; Sladeczek, C.; Franck, A.; Zhykhar, A.; Clausen, C.; Gleim, P.: Wave Field Synthesis Based Concept Car for High-Quality Automotive Sound. In: 48th Conference of the Audio Engineering Society (AES), Munich, Germany, 2012.
- [8] Roginska, A.; Geluso, P.: Immersive Sound: The Art and Science of Binaural and Multi-Channel Audio. New York, Taylor & Francis, 2018. <https://doi.org/10.4324/9781315707525>
- [9] Sladeczek, C.; Seideneck, M.; Lorenz, W.; Pursche, K.; Schneider, B.: Object-Based Audio as Platform Technology in Vehicles. In: 2022 International Conference on Automotive Audio, Dearborn, Michigan, USA, 2022. ■

### **Christoph Sladeczek**

*Fraunhofer-Institut  
für Digitale Medien-  
technologie IDMT,  
Ilmenau;  
transfer function  
GmbH, Ilmenau*

### **Mario Seideneck**

*Fraunhofer IDMT,  
Ilmenau;  
transfer function  
GmbH, Ilmenau*

### **Wolfgang Lorenz**

*Fraunhofer IDMT,  
Ilmenau;  
transfer function  
GmbH, Ilmenau*

### **Anja Chilian**

*Fraunhofer IDMT,  
Ilmenau*

### **Dr. rer. nat. Dr.**

**phil. Draško  
Mašović**  
*Fraunhofer IDMT,  
Ilmenau*

### **Katrin Pursche**

*Fraunhofer IDMT,  
Ilmenau*

### **Prof. Dr.-Ing.**

**Joachim Bös,  
M.S. /SUNY**  
*Fraunhofer IDMT,  
Ilmenau;  
Technische Uni-  
versität Ilmenau,  
FG Industrielle  
Anwendungen von  
Medientechnologien  
IAM*