

リアルとバーチャルの音を融合する音響XR技術

耳を塞がないオープンイヤー型イヤホンの普及とともに、周囲のリアルな音とイヤホンから聴こえるバーチャルの音を組み合わせる聴くことによる新たな体験が提案されています。NTTでは、このようなオープン型イヤホンが実現するリアルとバーチャルの融合を音響XRと呼び、サービス実現に向けた技術開発を進めています。本稿では、取り組んできたトライアルを中心に、技術や今後の展望について解説します。

キーワード：#XR, #PSZ, #オープンイヤー型イヤホン

のぐち けんいち ちば ひろのぶ
野口 賢一 / 千葉 大将
かこ たつや こづか しほり
加古 達也 / 小塚 詩穂里
くろかわ よしあき わたなべ ゆき
黒川 義昭 / 渡邊 悠希
なかやま あきら
中山 彰

NTTコンピュータ&データサイエンス研究所

オープン型イヤホンが実現するリアルとバーチャルの融合

NTTは、究極のプライベート音空間であるパーソナライズドサウンドゾーン(PSZ)の実現に向けて、研究開発を行っており、耳を塞ぐことなく利用者にしか聴こえないイヤホンの設計技術を開発⁽¹⁾しました。耳を塞がないオープンイヤー型イヤホンは、さまざまな製品が登場し、近年急速に普及しています。NTTでは、周囲で発生する音を自然に聴くことが可能であるオープンイヤー型イヤホンの特徴に着目し、イヤホンから聴こえるバーチャルの音と、直接耳で聴くリアルな音を融合させて聴く音響XR技術を提唱し、研究開発を開始しました。

音響XR技術を利用することで、周囲の音を聴きながら、イヤホンからの付加音により、聴こえる音を拡張できます。例えば、観劇やコンサートでは、通常P.A.スピーカ

から再生された音を客席で聴取をしますが、観客の近傍で発生する音響再現、音の方向や距離感の制御、広がりのある音響の表現はまだ課題があります。P.A.スピーカ音とイヤホンからの音を融合させて聴取する音響XR技術により、従来困難であったさまざまな音響表現や観客個人に最適化された音響提示を行うことが可能となります。また、スタジアムでのスポーツ観戦では、周囲の歓声を感じながら、イヤホンから解説音声を楽しむことが可能となります。多言語で行われる国際会議では、話し相手の発話のニュアンスを感じると同時に、イヤホンから翻訳音声を聴くことができます。

音響XR技術には、図1に示すように、大きく2つの技術課題があります。「オープンイヤー型イヤホン対応立体音響」と「バーチャル音空間レンダリング」です。通常、イヤホンから再生する音は、頭内に音像が定位しますが、「オープンイヤー型イヤホン対応立体音響」は、頭外の空間上

の任意の位置、例えばリアルなオブジェクトの位置からあたかも発生するような音を、イヤホンを介してバーチャルの音として再現することにより、さまざまな音響表現を可能にします。その際、イヤホンの装着位置のずれや、個人の耳の形状の違いにも対応し、オープンイヤー型イヤホンによる立体音響提示を実現します。リアルな音とバーチャルの音を単純に組み合わせるのではなく、「バーチャル音空間レンダリング」は、人の認知特性に合わせて、イヤホン再生音を制御し、直接耳で聴くリアルな音とイヤホンから聴くバーチャルの音を違和感なく融合して聴くことを実現します。

NTTは、2023年に音響XR技術を用いた複数のトライアルを行ってきましたので、詳細を紹介します。

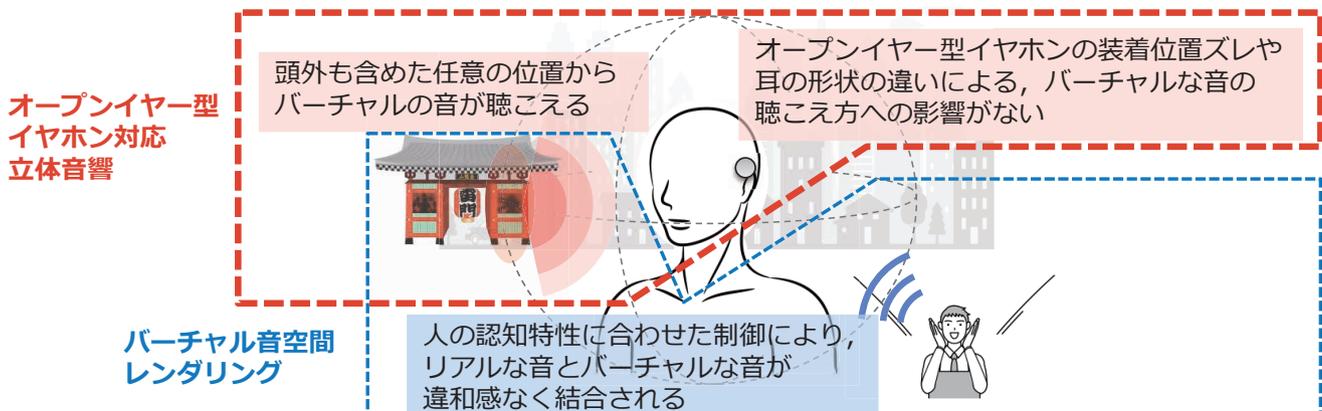


図1 音響XR技術の課題

トライアル1：超歌舞伎（ニコニコ超会議2023）

超歌舞伎では、歌舞伎とNTTの最新テクノロジーを組み合わせ、公演を行ってきました。2023年4月29～30日に幕張メッセで開催された「ニコニコ超会議2023」内で上演された「超歌舞伎 Powered by NTT『御伽草紙戀姿絵（おとぎぞうしこいのすがたえ）』」では、会場のリアルな音と耳元に流れる音がクロスオーバーする空間音響演出に取り組みました⁽²⁾。通常、観客は、P.A.スピーカから再生される、役者の音声、音楽、さまざまな効果音、周囲の客席から聴こえる大向う（歌舞伎の掛け声）を会場のリアルな音として、直接耳で聴きます。トライアルでは、ステージ前方客席の約180席限定で、無線受信機に接続したオープンイヤー型イヤホンを配布し、会場のリアルな音に加えて、イヤホンからの音を聴く体験を行っていただきました。イヤホンからは演目に合わせ、効果音を再生しました。例えば、馬が駆ける音、矢が飛び交う音、風の音等です。P.A.スピーカの音と組み合わせることで、聴取者の近傍を馬が左から右に走り抜ける音、聴取者の上方を矢が飛び交う音を再現でき、臨場感ある空間音響演出を可能としました。

オープンイヤー型イヤホンを用いた立体音響では、筐体自体の音響特性やイヤホンの装着位置から外耳道入り口までの音の伝達特性、耳表面の形状などの影響により、音像が上方に定位してしまうという問題があります。この問題に対し、上方定位の特性を打ち消す補正フィルタを開発し、適用

させました（図2）。これにより、コンテンツ制作者の意図する空間音響再生を可能とすることができました。

トライアル2：サテライト配信公演における音声ガイド

2023年11月2日に開催された「第180回NTT東日本N響コンサート」の開催会場である東京オペラシティ コンサートホール（東京都新宿区）でのオーケストラの演奏を、サテライト会場の北斎ホール（長野県小布施町）に高音質配信するサテライト配信公演が開催されました。北斎ホールでは、大型スクリーンでの映像再生と、5.1ch音響機器による高音質再生を行うとともに、手元のタブレット、スマートフォンで、好きなアングルで演奏を見ることができるマルチアングル配信を実施しました。また、NTT技術による音漏れの少ないオープンイヤー型イヤホンを用いて、演奏される楽曲の解説を配信するトライアルを実施しました。音漏れが少ないため、周囲の人の妨げにならず、耳を塞がないことで、オーケストラの演奏と解説音声を同時に聴くことができます。

一般的にオープンイヤー型イヤホンを利用する場合、周囲のリアルな音とイヤホンからの音が重なることで、音が聴き取りにくくなる問題があります。そこで、イヤホンから提示する音像位置を制御し、周囲のリアルな音が被らない位置から、あたかも聴こえるように提示することで、周囲の音が耳に聴こえる状況下でも音による情報提示を可能にしました⁽³⁾（図3）。今回のサテ

ライト配信公演では、本技術を用い、イヤホンからの解説音声の聴こえる位置をユーザの上方に制御することで、スピーカから再生されるオーケストラの音と分離し、解説音声の聴き取りやすさの向上を図りました。

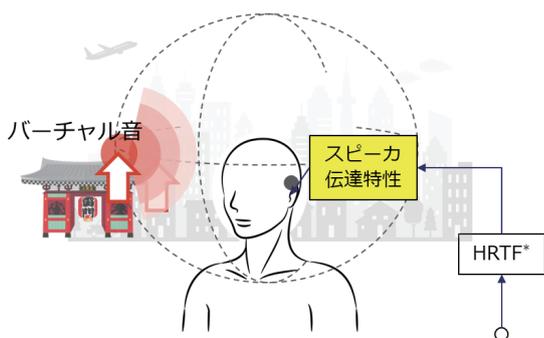
トライアル3：NTT技術史料館の音声ガイド

2023年11月14～17日に行われた「NTT R&D FORUM 2023 — IOWN ACCELERATION」において、NTT武蔵野研究開発センタ内にあるNTT技術史料館にて、オープンイヤー型イヤホンを用いた音声ガイドのトライアルを行いました。本トライアルでは、スマートフォンのセンサによりユーザの位置を推定し、展示物に近づくと自動的に設定された音声ガイドがスマートフォンに接続されたオープンイヤー型イヤホンから再生されるシステムを来場者に配布しました。美術館や博物館等の施設で、これまでもイヤホンを用いたガイドサービスは行われていますが、本トライアルでは、以下の点をポイントとしています。

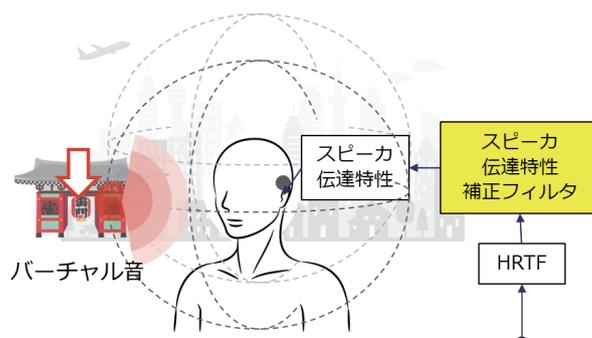
- (1) 周囲と調和するオープンイヤー型イヤホンでの鑑賞

耳を塞がないことで、周囲の音を自然に聴くことができるため、周囲の様子を把握することができます。これにより、ガイドを聴きながらも、周囲の混雑状況を感じ取ることができ、かつ、イヤホンの音漏れは少なく、他の来場者の迷惑となる心配もありません。また、館内のスピーカや展示物自体が発する音があれば、その音も自然に

バーチャル音が上方から聴こえてしまう



上方定位問題を打ち消すようなフィルタを設計



* HRTF (Head-Related Transfer Function: 頭部伝達関数)

図2 オープンイヤー型イヤホン着用時の定位位置の補正フィルタ適用イメージ



図3 空間的な特性を利用してイヤホンからの音を聞き取りやすくする音源配置技術

聴くことができます。環境によっては、一緒にいる友人等と会話を交わしながら鑑賞するといった体験も可能となります。

(2) あたかも展示物から聴こえる立体音響

ステレオの音響再生を活かし、展示物からユーザの両耳までの音の伝達特性を模擬する信号を提示することで、あたかも展示物から聴こえる立体音響を可能としました。このとき、音像が上方に定位してしまうというオープンイヤー型イヤホンの特性の補正技術を導入しています。また、今回、ユーザの位置をリアルタイムに推定しており、ユーザが移動しても、同じ位置にある展示物からあたかも聴こえるよう立体音響を実現しています。具体的には、静態展示している昔の通信機の周囲を動き回るとき、通信機を模擬した効果音をユーザ位置に応じて立体音響処理してイヤホンで提示し、あたかも通信機から音が聴こえるような音響演出を実現しました。

(3) クロスリンガル音声合成技術による多言語ガイド

インバウンド需要の増加に伴い、多言語による音声ガイドの提供が求められています。NTTが開発したクロスリンガル音声合成技術は、日本語のみの音声データから、同じ声質を保ちつつ、英語、中国語等の異なる言語による音声合成を可能とする技術です。今回は、声優の日本語音声データを基に、クロスリンガル音声合成技術による日本語と英語の音声ガイドを用意し、アプ

リ上で切り替え可能としました。

今後の展開

本稿では、イヤホンから聴こえるバーチャルの音と、直接耳で聴くリアルな音を融合させて聴く音響XR技術とその取り組みについて紹介しました。オープンイヤー型イヤホンは、周囲の音を自然に聴くことができる特徴から、観光・エンタテインメントといった分野以外にも、ビジネスや日常生活のさまざまな場面での活用が考えられます。例えば、視覚障がい者の方へ音声ガイドを行うことも有用なアプリケーションです。実際に視覚障がい者の方にオープンイヤー型イヤホンを装着し、音声ガイドを聴きながら、敷地内を歩く体験をしていただきました。体験後のヒアリングでは、「外でイヤホンをつけることに抵抗があったが、外の音の聴こえ方があまり変わらなかった」と、コメントをいただきました。また、耳を塞がないため、長時間装着していても疲れにくい特徴から、一日中ユーザがオープンイヤー型イヤホンを装着しながら、ユーザの状況に応じたパーソナライズ化された複数の音響XRサービスを楽しむといったことも考えられます。今後、想定される利用場面での技術課題を解決し、サービスの実現に向けて、取り組んでいきます。

参考文献

- (1) <https://group.ntt.jp/newsrelease/2022/11/09/221109a.html>
- (2) <https://group.ntt.jp/newsrelease/2023/04/29/230429b.html>
- (3) <https://group.ntt.jp/newsrelease/2023/11/10/231110c.html>



(後列左から) 中山 彰 / 野口 賢一 / 小塚 詩穂里 / 黒川 義昭 (右上)

(前列左から) 渡邊 悠希 / 千葉 大将 / 加古 達也

音響技術を核として、リアルとバーチャルを融合させることによる新たな体験を創出しています。今後もさまざまなトライアルを繰り返しながら、価値あるサービスの実現に向けて取り組んでいきます。

◆問い合わせ先

NTTコンピュータ&データサイエンス研究所
企画担当
TEL 046-859-4003
FAX 046-855-1149
E-mail cd-koho-ml@ntt.com