

サウンドスコープフォン

浜中雅俊¹

概要：本稿では、頭部方向やハンドジェスチャーで複数パートからなる曲の聴きたいパートを強調して聴くことのできるアプリケーション、サウンドスコープフォンについて述べる。従来提案されていたインタフェースでは頭部方向の検出に、方位センサなどを搭載した特別なヘッドフォンを必要としていた。サウンドスコープフォンは、市販のスマートフォンのインカメラをからの画像によるフェーストラッキング情報と内蔵加速度・ジャイロセンサの情報を統合し頭部方向の検出を可能とした。構築したアプリケーションは SoundScopePhone というアプリ名で Apple App store 上で公開している。

キーワード： SoundScopePhone, Appstore, フェーストラッキング, ハンドジェスチャー

1. はじめに

サウンドスコープフォンは(図1)、曲を聴いているときに、特定のパートの音を、よりはっきりと聴きたいという望みを実現するアプリケーションである。ユーザが右を向けば、右から聞こえていた音が正面で、左を向けば左から聞こえていた音を正面で聴くことができる。ユーザが耳に手のひらを近づける「耳を澄ますようなポーズ」をすると、正面から聞こえていた楽器の音によりフォーカスして聴くことができる。

これまで我々は、ヘッドフォンに顔の向きを検出する方位センサや、手と耳との距離を測定する距離センサにより上記の機能を実現するサウンドスコープヘッドフォンを提案し、日本科学未来館や SIGGRAPH2009 Emerging Technologies 等で展示を行ってきた(図2)[1, 2]。対面での展示が社会情勢により一時的に困難となっていることから、同様の機能をアプリで実装し、スマートフォンを使って多くの人に体験いただけるようアプリの提供を考えた。



図1 サウンドスコープフォンの使用風景

2. 関連研究

頭部の方向や位置を検出するセンサを搭載したヘッドフォンは従来から存在していたが、その目的は、仮想の音源位置を固定することで臨場感を高めることであつたため、本研究のように特定のパートを強調する目的では使われてこなかった [3, 4]。したがって、たとえば聴きたい楽器のすぐ近くに聴きたくない楽器が定位していた場合、聴きたい楽器のみを選択的に聴けるようにすることは困難であつた。

一方、空間音響システム [5, 6] では、聴取者のアバタや各パートの位置をマウスで移動することでミキシングを変更することを可能にしていた。しかし、各パートのミキシングを適切に変更するためには、ソロが始まったパートを近くに配置したり、ソロが終わり伴奏に戻ったパートを遠くに配置しなおすなどの煩雑な操作が必要なため、初心者には困難であつた。



図2 サウンドスコープヘッドフォンの展示風景

¹ 理化学研究所 革新知能統合研究センター 音楽情報知能チーム
Music Information Intelligence Team, Center for Advanced Intelligence
Project (AIP), RIKEN

3. サウンドスコープフォン

複数パートからなる曲を聴くときに、自分が聴きたいパートを探しながら演奏を聴くという音楽鑑賞の形態を提案してきた。

3.1 どのように聴きたいパートを探すのか

「聴きたいパートを探す」ことを可能にするためには、探したパートの音とその他のパートの音が区別できなければならない。本研究では、各パートの定位と音量に違いをもたせることで、探したパートとその他のパートが区別できるようにする。具体的には、探したパートは定位を中央に位置させ音量を大きくする。一方、その他のパートは定位を左右に振り、音量を小さくする。このようにすれば、たとえ楽器の聴きわけが苦手な音楽初心者でも、自分の聴きたい楽器を探し出すことが可能となる。

3.2 どのように動作を検出するのか

従来のサウンドスコープヘッドフォンでは、地磁気センサをヘッドフォンのアーク部に搭載することで、頭部の方向を検出していた(図3)。一方、サウンドスコープフォンでは、スマートフォンのインカメラが取得した画像を処理することで頭部方向の検出を行う(図4上部)。

インカメラを用いて頭部の方向の検出を行う場合、カメラから見て頭部方向が90度を超えると、顔の前側の半分以上が見えなくなり、検出が困難となる。そこでサウンドスコープフォンでは、スマートフォンに搭載された加速度・ジャイロセンサが取得したスマートフォン本体の角度情報と、インカメラからの頭部角度情報を統合することで、360度すべての方向での頭部方向検出を可能とする。

すなわち、スマートフォンを手に持ち後ろを振り返る場合、首から上だけで大きく振り返ることは困難で上半身も連動して動くことになる。このときスマートフォンを持った手もある程度連動して動くため、インカメラの映像に頭部の前面が捉えやすくなり、これにより、頭部を90度以上回転させても頭部方向の検出が可能となる。

従来のサウンドスコープヘッドフォンでは、右スピーカの外側に距離センサを搭載することで手が耳に近づいていることを検出していた。一方、サウンドスコープフォンではインカメラからの映像に対して画像処理を行うことで、手が耳に近づいていることを検出している。



図3 ヘッドフォンに搭載したセンサ

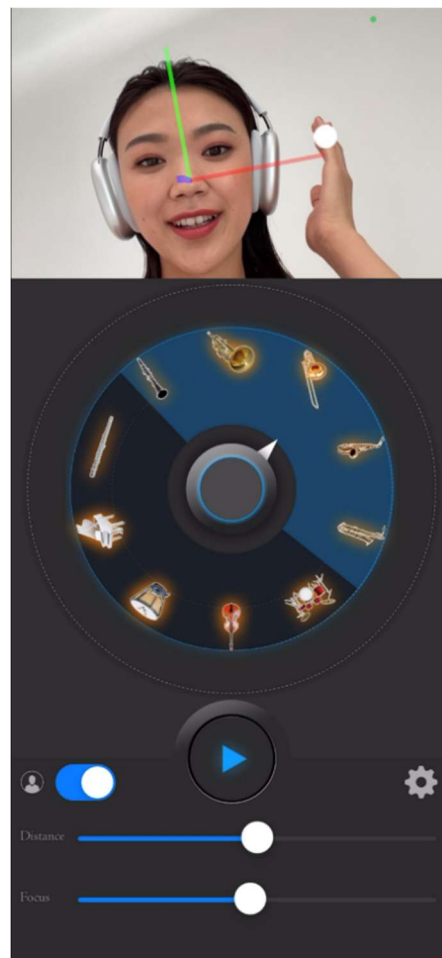


図4 サウンドスコープフォンのスナップショット

3.3 どのように強調するのか

従来のサウンドスコープヘッドフォンでは、検出した動作とオーディオミキサーの操作とを関連づけることで、聴きたいパートの音が正面で強調されて聞こえるようにしていた。一方、サウンドスコープフォンでは、3D音響ライブラリ OpenAL[7]で定義される3D空間上に n 個のパートを配置することで、頭部の方向に配置されたパートの音を正面で聴くことができる(図4中部)。このとき、聴者のアバタから n 個のパートそれぞれとの作る角度を θ_n とする。

頭部の方位に応じて次式で増幅率 h_n^θ ($0 \leq h_n^\theta \leq 1$) を変化させることにより、正面に近い角度にあるパートの音をより強調して聴くことができる。以下、頭部の正面方向とパートの配置されている位置とのつくる角度を θ ($-\pi \leq \theta < \pi$)、手と耳との距離を δ ($0 \leq \delta \leq 1$)とする。なお、実装上での手と耳の距離は、手の小指の先端がインカメラの撮影範囲の10%内側の位置するときを1、インカメラの撮影範囲の中心に位置するときを0として、その間を正規化した値である。インカメラの撮影範囲の外側から10%は、認識対象外とすることで動作を安定させている。

$$h_n^\theta = \begin{cases} 0 & \widetilde{h}_n^\theta < 0 \\ \widetilde{h}_n^\theta & 0 \leq \widetilde{h}_n^\theta \end{cases} \quad (1)$$

ただし,

$$\widetilde{h}_n^\theta = \begin{cases} 0 & \delta = 0 \\ 1 - (\alpha \cdot |\hat{\theta}_n|) / (\pi \cdot \delta) & \delta > 0 \end{cases} \quad (2)$$

α ($0 \leq \alpha \leq 1$) は調節可能なパラメータで, $\delta < 1$ の場合の増幅率の変化を設定する. $\alpha = 0$ の場合には, 耳に手を近づけても各パートの増幅率に変化はないが, $\alpha > 0$ の場合には, 耳に手が近づくにつれて増幅率が減少する. このとき聴者が向いている方向よりも, 向いていない方向の増幅率の減少のほうが大きいので, 正面の音が相対的に大きな音で聞こえるようになる.

3.4 ブーストモード・ミキシングコンソールの実装

3.3 節の実装のみでは, 耳に手を近づけると聴者の向いていない方向の増幅率が減少する. これにより正面のパートの音は聴きやすくなるが, すべてのパートの増幅率の合計の値は小さくなるため, 曲としての迫力に欠けた状態になりやすい. そこで, 増幅率の合計が一定となるよう正規化したブーストモードを実装した. ブーストモードでは一つのパートの音量が大きくなりすぎる場合があるため, 増幅率に閾値を設け, それ以上大きくならないようにするブーストリミッターを設けた (図 5 右端).

さらに, 各楽器の音を個別に調整可能とするため, ミキシングコンソールを実装した.



図 5 ブーストモードとミキシングコンソールの実装

4. おわりに

本稿では, 頭部方向やハンドジェスチャーで複数パートからなる曲の聴きたいパートを強調して聴くことのできるアプリケーション, サウンドスコープフォンについて述べた. サウンドスコープフォンは, アプリ名「SoundScopePhone」として, Apple Appstore 上で公開している[8]. また, 利用方法を理解するための紹介動画を用意している[9] (図 6). アプリを 30 回利用するとアンケートを依頼するページの案内が表示される. ご協力いただければ幸いである.

iPhone 版のサウンドスコープフォンに続き, 今後演奏者の演奏映像も表示される iPad 版も構築し公開していく予定としている.



図 6 サウンドスコープフォンの紹介動画

参考文献

- [1] Masatoshi Hamanaka, SeungHee Lee. Sound Scope Headphones. ACM Siggraph2009 Talks TK-201 /Emerging Technologies ET-201, August 2009.
- [2] 浜中雅俊, 李 昇姫. サウンドスコープヘッドフォン. 日本バーチャルリアリティ学会論文誌, 2007, vol. 12, no. 3, p. 295-304.
- [3] 佐藤光一. デジタルコードレスサラウンドヘッドフォンの開発. PIONEER 技術情報誌, 2004, Vol. 14, no. 2, p. 66-73.
- [4] 久木元 伸如, Ewe C. Huat, 竹田 仰. プロジェクション 型没入ディスプレイにおける 3 次元音場生成のための実用的検討. 電子情報通信学会 技術研究報告, 1999, EA99-32, p. 29-36.
- [5] 鈴木陽一, 西村 竜一. 超臨場感音響の展開. 電子情報通信学会誌, 2010, vol. 93, no. 5, p. 392-396.
- [6] 安藤彰男. 高臨場感音響技術とその理論. 電子情報通信学会 基礎・境界ソサイエティ, 2009, vol. 3, no. 4, p. 33-46.
- [7] “OpenAL. “. <https://www.openal.org/>, (参照 2021-05-18).
- [8] “SoundScopePhone. “. <https://apps.apple.com/jp/app/soundscopephone/id1559407378>, (参照 2021-05-18).
- [9] “SoundScopePhone. “. <https://gttm.jp/hamanaka/soundscopephone/>, (参照 2021-05-18).