

音源定位・分離ソフトウェア HARK

本田技研工業株式会社

知的財産・標準化統括部 先端技術知財課 紺社明彦

音源定位・分離ソフト『HARK』とは

音源定位・音源分離とは

人の声や鳥の鳴き声などの音が**どの方向から聞こえてくるのか**を特定することを「**音源定位**」といいます。また、聞こえてくる様々な音の中から**特定の音を拾うこと**を「**音源分離**」といいます。

HARK (*) とは

マイクで拾った音の聞こえてくる方向の**特定（音源定位）**と、**特定音のみを拾う（音源分離）**ことを可能にするソフトウェアです。

ロボットと人が声でコミュニケーションを行うためのツールとして開発されました。

(*) “HRI-JP Audition for Robots with Kyoto Univ” の各ワードの頭文字からの名称です。Hondaと京都大学との共同研究より生まれました。



HARKはASIMOに搭載されました。
左の写真は3人が同時に話した内容をASIMOが聞き取るデモンストレーション実行時のものです。
ASIMOの頭部に取り付けられた複数の小型マイクで收音し、ASIMOに搭載されたHARKを使って音源分離処理を行い、話した内容を認識しています。

デモンストレーションの様子は以下のページの動画でもご覧いただけます。
<https://www.honda.co.jp/ASIMO/about/>

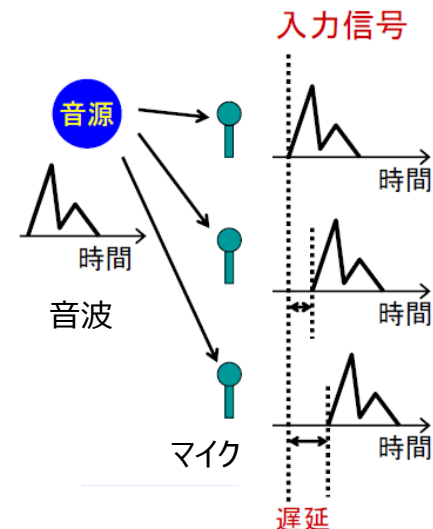


音源定位の原理

人は音が聞こえてくる方向が分かります。人は二つの耳により音を聞き取りますが、両耳に聞こえる音の時間差によって方向を認識しています。

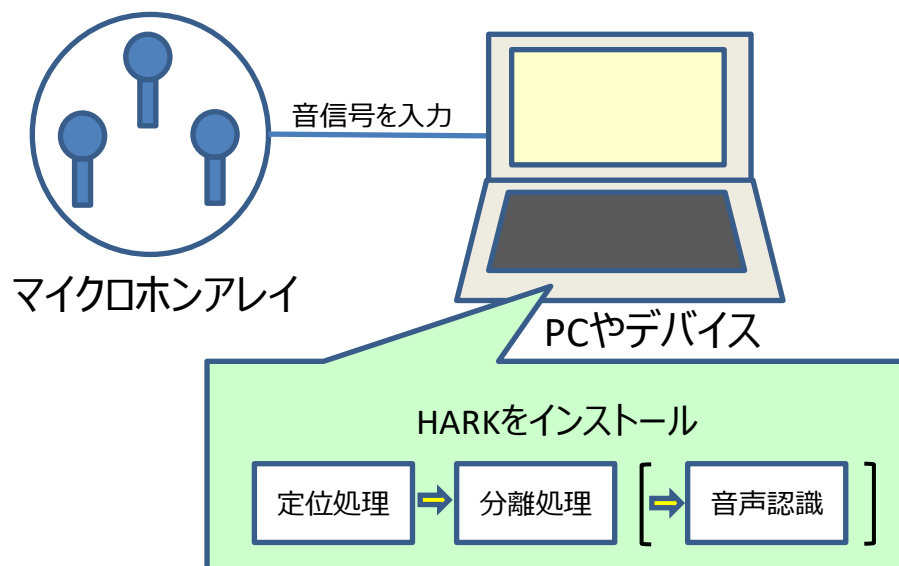
音源定位も同様に、複数のマイク（マイクロホンアレイ）によって收音し、マイク間の音の時間差を利用して音の方向を推定します。

また音源分離は、音源定位処理で得た方向の音を強調することで、特定の音を拾うものです。



第18回HARK講習会テキストより

音源定位・分離システム構成



左図のように、マイクロホンアレイ（市販品など）をHARKをインストールしたデバイスやPCに接続することで音源定位・分離処理を行うことができます。

※利用にはマイクロホンアレイをご準備いただく必要があります。

マイクロホンアレイ例

- ・システムインフロンティア社 Tamago
- ・その他 ALSA、ASIO、DirectSoundベースデバイス

1. 独自の信号処理アルゴリズム

・音源定位 GEVD-MUSIC法 および GSVD-MUSIC法

雑音に比較的ロバストなMultiple Signal Classification (MUSIC) 法をベースに、さらに改良を加えたGEVD-MUSIC法 および GSVD-MUSIC法を採用し、より雑音にロバストなアルゴリズムとなっています。

・音源分離 GHDSS-AS法

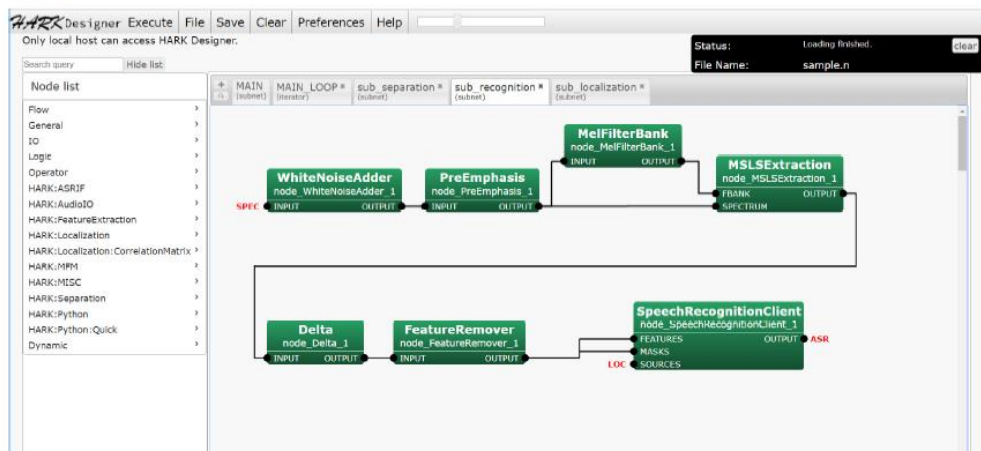
音源分離方法として定位情報を使ったGeometric High-order Decorrelation Source Separation (GHDSS法) に改良を加え、音響環境が変化しても追従できるようにしたGHDSS-AS法を採用し、分離性能を向上させています。

2. GUIプログラミング環境 HARK Designer

音源定位・分離に使う各機能ごとに集約したモジュール（ノード）をブラウザ上で接続することで視覚的にプログラミングを行うことができます。パラメータの条件設定もHARK Designer上で行うことができます。



HARK Designerアイコン



3. 各種サポート・パッケージ

・他言語やライブラリとのインターフェース

Python、ROS（Robot Operating System）やOpenCVをサポートするために、HARK-Python、HARK-ROS、HARK-OpenCVというインターフェースを用意しています。

・音声認識

分離音声の音声認識を行うため、オープンソースの音声認識エンジン（Kaldi）をサポートしています。

・多数の音源分離手法

音源分離手法をGHDSS-ASを含め代表的な分離手法を11種用意し、HARK-SSSにパッケージしました。音源分離はそれぞれの手法ごとに特徴があるため、ケースによってはGHDSS-AS法以外の手法が有利なことも考えられます。そのようなケースに対して他手法も用意しました。

4. ソフトウェアが公開されており、かんたんに入手することができます。

HARKはホームページ（<https://hark.jp>）で公開されており、このページからダウンロードすることができます。
研究・試験段階では無償で利用することができます。
（Linux[Ubuntu]用とWindows用）



応用用途

[応用例]

3人が同時に発する言葉を聞き分けるロボット

3人の声をそれぞれ分離し音声認識で何を話しているか聞き分けます。



関連動画 <https://www.honda.co.jp/ASIMO/about/>

[応用例]

ドローンによる災害時の要救助者探索

ドローンに付けたマイクロホンアレイで要救助者の声を拾い、どこにいるのかを特定します。飛行中の風切り音などの雑音がある環境でも声を拾うことができます。



東京工業大学ホームページ
<https://www.titech.ac.jp/news/2017/040001.html>



[応用例]

会議での会話の記録

会議中の会話を録音し、各参加者の声を分離することができます。会議中の発言者にカメラを向けることにも利用できます。



[応用例]

野生生物の観測

森林での野生生物の鳴き声（鳥や動物）を録音し、声の方向や分布の経時変化から生物の生態を観測することができます。環境アセスメント調査での活用が考えられます。



名古屋大学 大学院情報学研究科 鈴木麗聖准教授のグループではHARKを利用した野鳥観測用ソフトウェア「HARKBird」を作成され、ご研究に活用されています。「HARKBird」はHARKBirdのホームページから取得することが可能です。
 (鈴木准教授のHP) <http://www.alife.cs.i.nagoya-u.ac.jp/~reiji/>
 (HARKBirdのHP) <https://sites.google.com/view/alcorsuzuki/home/harkbird>

HARKBirdを使った観測の写真

2019 IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP 2019 – Proceedings P.8246-8250

音源定位・分離技術では、線形代数などを使ってモデルを構築します。特許内容もそれに従います。

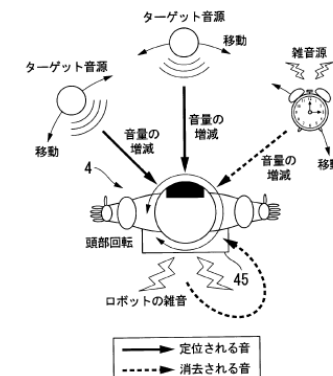
■ 主な関連特許

①特許第5663201号 音源方向推定装置及び音源方向推定方法

②特許第5952692号 音源方向推定装置、音響処理システム、音源方向推定方法及び音源方向推定プログラム

【内容】

GEVD-MUSIC法関連。従来MUSIC法では対象音が雑音に比べて強いことが条件であったが、GEVD（一般固有値展開）を基づく方法を採用することで雑音を除外し音源定位を行う。



③特許第5967571号 音響信号処理装置、音響信号処理方法、及び音響信号処理プログラム(*)

【内容】

GSVD-MUSIC法関連。GSVD（一般特異値展開）を基づく方法を採用することで方向性のある雑音に対し、その方向から到来する雑音を抑制するフィルタが算出され、雑音影響を低減できる。

④特許第5706782号 音源分離装置及び音源分離方法

【内容】

GHDSS法関連。音源の変動（音源の位置の変化、音のパワー）が生じて音源信号を分離できる。音の変動状態を検出し、それに基づき分離行列を算出して音源分離に用いる。

(*) 国立大学法人熊本大学との共同権利です。

HARKについて上記内容の特許を保有し、保護を図っています。

想定ケース1：HARKソフトウェアをダウンロードしお使いいただく。

HARKソフトウェアは、HARKホームページ（<https://hark.jp>）からダウンロードいただけます。
HARKを貴社の製品やサービスにご活用ください。
研究や、商用化のための試験でのご利用は無償で行っていただけます。
貴社の製品・サービスを商用化される際に、有償のライセンス契約を締結させていただきます。

想定ケース2：貴社における新規ソフトウェア開発（特許実施許諾のみ）

HARKに利用されている技術の詳細は、弊社研究者らによる学術論文や特許文献により公開されています。
HARKソフトウェアそのものをご利用いただくかわりに、特許文献の記載の技術をご活用いただいて、貴社独自のソフトウェアを作成いただくことも可能です。
その場合、有償の特許実施許諾契約を締結させていただきます。

ライセンスのお問合せや学術論文や特許リストなど各種お問合せにつきましては、
hm_hark_licensing_admin@hm.honda.co.jp
で承ります。お気軽にお問合せください。

ご清聴ありがとうございました