

聴覚情報に基づく移動ロボットのナビゲーションに関する 実験的研究

藤井文武 (理工学研究科機械工学専攻) 穂垣暢久 (前・理工学研究科機械工学専攻)

Experimental study on the auditory feedback navigation of mobile robot

Fumitake FUJII (Mechanical Engineering, Graduate School of Science and Engineering)

Nobuhisa HOGAKI (Mechanical Engineering, Graduate School of Science and Engineering)

Abstract This paper discloses the result of the experimental study which explores the feasibility and performance of manual auditory navigation system of a mobile robot. The navigation is based on the sensory capability of a human which is referred to as the sound source localization. The operator is instructed to navigate his mobile robot to the location where a sound source is located, based on the auditory cues given to the operator via three different sound playback devices for three different sound sources. The binaural sound which is fed to the operator is captured with a dummy head microphone mounted on the mobile robot. Experimental results indicate feasibility of the auditory navigation, provided that the interaural differences contained in the binaural signal are properly transmitted to the operator's ears.

Key Words: *Auditory Navigation, Sound Source Localization, Head Related Transfer Function, Binaural/Transaural Audio Playback System*

1 緒言

人間には、視覚という2次元・3次元の空間情報把握に優れ、分解能も高い感覚器が備わっている。人間が外界とのインタラクションを行う場合、外界の状態検出に用いられる感覚器は、通常は視覚となるであろう。一方で、人間の聴覚には、外界に存在する発音体のおおよその位置と方向を両耳で観測される音に含まれる両耳間差から把握する音源定位能という能力が備わっている。聴覚の空間分解能は視覚に及ばず、音源方向に依存して分解能が変化するなど扱いの難しい点もある¹⁾が、視覚にあるような死角が存在しないという優れた特徴も有している。人間はこのように特徴の異なる複数の感覚器を使い分けながら外界の様子を認識しているが、人間が立ち入れないような環境に人間の操縦により立ち入り調査する能力を有する移動ロボットに、人間の聴覚特性を踏まえた音検出能を付与することができれば、未知環境探索に非常に役立つと期待される。

視覚障害を有する方や、消防士、兵士などの一時的に視覚情報が機能しない状況におかれる人のために、音を利用したナビゲーションシステムを構築しようとする研究がいくつか試みられている。Walker and Lindsay²⁾は、発音ビーコンによる経路誘導システム構築に関する検討の一環として、一般に提供されているサウンドカードの頭部伝達関数(HRTF)を利用して空間内に仮想音

像を作成して提示し、仮想空間内を移動させる実験を行い良好な成績を得ている。また、Dunaiら³⁾は、音源からの距離を変えて実測したHRTFを利用し、連続的なクリック音の提示間隔を変更した幾種類かのモノラル音源に、音源の方位角と距離が連続的に変化する状況を再現すべく異なるHRTFを連続的に畳み込んで作成したバイノーラル音を仮想音像として再生提示、受聴者が知覚した音源の位置軌跡を描かせる実験を行い、クリック音の提示間隔を最適化することで距離知覚を改善できることを示した。これらの検討は、提示する音がHRTFを利用して生成された仮想音源である点で共通しており、距離の認知と移動スケールが被験者の感覚と不一致を起こす可能性がある。一方、視覚障害を有する方のための音によるナビゲーションシステム構築のために、実際に人間に対して音による移動ナビゲーションを行った研究もあるが、人間が歩いて移動しながら音の変化を観測する行為と移動ロボットの移動による音変化を観測して音源の位置情報に置き換える行為の間には大きな差があると考えられる。

そこで本研究では、聴覚を利用した環境の認知を移動ロボット操縦に活用するシステム構築に向けた試みとして、聴覚のみによるフィードバックに基づく移動ロボットナビゲーションの実行可能性を探るため、バイノーラル信号を收音可能なダミーヘッドマイクロホンを搭載した聴覚フィードバック移動ロボット操縦システムを構築し、移動ロボットの存在する空間に設置した外部目標音

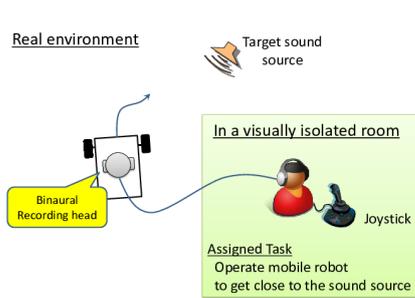


Fig.1 Mobile robot navigation experiment based solely on the binaural auditory feedback

源に向けた移動ロボットの手動誘導実験を行うこととした。このような移動ロボットにおいては、聴覚センサは視覚の補助として位置づけるべきものであると考えられるが、本稿では人間の音源定位能に基づく聴覚によるナビゲーションの特性を明らかにし、より効果的な人間への聴覚情報提示のあり方を検討するという観点から、視覚情報を排した実験を行ったので、以下にその結果を報告する。

2 実験システムのハードウェア構成

聴覚ナビゲーション移動ロボットの実験システムの全体図を図1に示す。システムは、移動ロボット制御系に関する部分と、移動ロボット走行環境で収録された音を別室の被験者に伝送する部分に分かれ、両者は信号伝達経路としては独立している。ここでは、今回の実験システムの構成について説明する。

2.1 マニュアルナビゲーションのための制御系の構成

人間による移動ロボットナビゲーションに関する信号の流れを図2に示す。実験システムの制御は、MATLAB がインストールされたノート PC により行う。MATLAB 上で動作するスクリプトは、操縦者がジョイスティックを操作することで与えられるロボットに対する並進速度と姿勢角速度の指令値を、bluetooth 無線により移動ロボットに伝送する。

移動ロボット側に組み込まれたプログラムは、この指令値を受信して、次章で説明する制御則に従って左右の動輪に取り付けられたモータへの制御入力を計算し、出力する。その際の指令入力値や左右のモータユニットに組み込まれているエンコーダのパルスカウント値など、制御によって値が変化していく変数の値をロギング用変数としてノート PC 側に bluetooth 無線を介して返送す

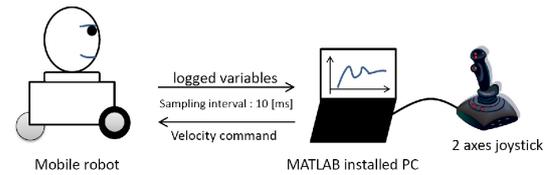


Fig.2 Control signal connections

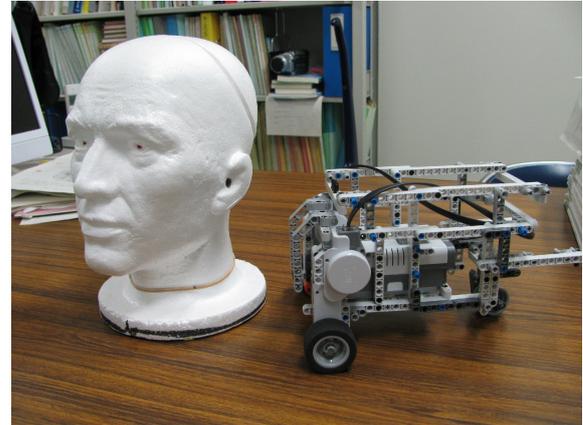


Fig.3 Dummy head microphone for recording binaural environmental sounds and the mobile robot

る。以上の一連の処理が1制御周期で実行され、これを繰り返すことで移動ロボットの動作が手動制御される。

今回、移動ロボットとしてはLEGO社のMindstorms NXT 2.0 キットを、図3右に示すような2動輪1キャスト型の移動ロボットとして組み上げて利用した。このキットには制御ロジックのプログラム用にARM7マイコンが搭載されている。Mindstorms キットには様々なソフトウェア開発環境が提供されているが、今回のシステム構成には実時間のタスク実行管理や複数タスク間の優先度に基づく実行制御の機能を持つリアルタイム OS である nrtOSEK⁴⁾ を利用した。前述の制御のメインタスクの実行周期は 10[ms] に設定した。

移動ロボットより返送されたログ変数の値は、MATLAB スクリプト中で受信しメモリに格納すると同時に、MATLAB の plot コマンドを利用してログ変数値の時間変化を画面のグラフでモニタリングできるようにした。これにより、走行中のロボットの状況をリアルタイムに把握することが可能となる。

2.2 音信号の収録と伝達および再生

図3右に示す移動ロボットの上には、マイクロホンを設置する。被験者の音源定位能により音源の位置を認知させるには、被験者の両耳にバイノーラル信号と呼ばれる人の耳に入る音に相当する信号を提示する必要が

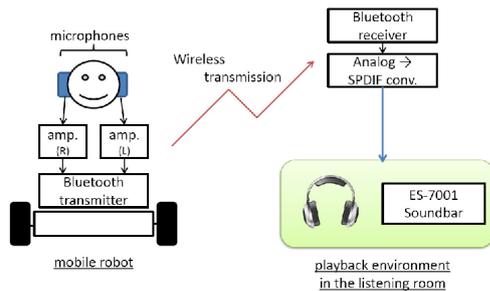


Fig.4 Sound signal transmission

ある。

バイノーラル信号の收音のためには人の頭および耳介の形状を模したダミーヘッドマイクロホンが用いられる。商用のコンテンツを収録するために用いられるダミーヘッドマイクロホンは質量が大きく高価でもあるため、今回はカツラなどのディスプレイスタンドとして市販されている発泡スチロール製の安価な頭部モデルを利用して簡易なダミーヘッドマイクロホンを作成した。作成したマイクロホンを図3左に示す。成人の外耳道の平均的な長さは24[mm]程度であるといわれている⁵⁾。作成した簡易ダミーヘッドでも鼓膜に相当する部分に設置するECMの收音面が外耳道入口から24[mm]程度となるようにしている。

図1に示すように、移動ロボットの走行環境内に設置されたラウドスピーカにより提示される目標信号音は、このダミーヘッドのマイクロホンを介して收音され、移動ロボットの走行環境に関する視覚情報が排除された別室にいる被験者のもとに届けられる。マイクロホンで採取された音は、マイクのアンプ出力口でBluetooth無線送信機により無線送信され、実験室内に設置された受信機で受信される。Bluetooth受信機の出力はステレオアナログ音声信号を光SPDIFデジタル信号に変換する変換機へ接続され、この変換機によって光信号へと変換の上、Tos-linkケーブルにより別室の被験者まで伝達される。

被験者のために用意される音の再生装置は、ステレオヘッドホンと、マランツ社のホームシアターシステムであるES-7001である。ES-7001には、入力バイノーラル信号が両耳にそのまま届くようにスピーカーから左右耳までの音響伝達関数が最適となるOPSODISスピーカーと、伝達関数の細部を補正するためのDSPが備わっており、入力バイノーラル信号が持つ情報を損なうことなく左右の耳に届けることが可能である。

同様の効果をステレオスピーカー再生システムの前段

フィルタとして畳み込むような実装を行う場合、そのような逆フィルタは特定の周波数においてハイゲインとなり再生系のダイナミックレンジを劣化させることがある⁷⁾が、ES-7001はそれを避け、かつ音響伝達関数の誤差が逆フィルタに与える悪影響も低減化される工夫がなされている。今回はこのES-7001をヘッドホン再生の場合とのナビゲーション性能比較のために利用した。なお、ES-7001には、3Dオーディオ再生のための信号処理に加えてステレオ方式をデジタル信号処理で模擬し、一般のステレオスピーカーでステレオ信号を再生するのと同様の状態を再現する再生モードも有している。これも比較のために利用し、計3種類の再生装置に対して評価実験を行った。

3 ロボット運動制御系の構築

ここでは、図3に示す移動ロボットに実装した制御系の構成について説明する。

3.1 モータユニットのモデリング

今回の移動ロボットで動力源として利用する、LEGO NXTのモーターユニットに関する情報を表2にまとめた。力学的特性も含めたより詳細な情報はweb site⁶⁾で確認することができる。ユーザーは、NXTのファームウェアが提供する機能を通じてモータの制御を行うことになる。この際の変更可能なパラメータは、入力PWMのduty比で、0～100[%]の整数値がとりうる値となる。逆回転を指令する場合は与える数値に負号を付すので、回転方向の制御とあわせ、モータへの制御入力は-100～100の間で指定することになる。

このようなハード構成であるので、制御の前提となるモータユニットの伝達特性としては、duty比入力から出力軸角度間の伝達特性のモデルを求めておくことが好都合である。そこで、モーターユニットに対して10, 40, 70[%]の3種類のduty比入力をステップ状に与えた場合の出力角度応答を測定した。その結果が図5である。これより、duty比入力 $u(t)$ から出力角度 $\theta(t)$ の間のモータの伝達関数 $M(s)$ は一次遅れと積分器からなる

$$M(s) = \frac{K}{s(Ts + 1)} \quad (1)$$

と表すことができることを確認した。以降の設計では、パラメータ K は0.094、 T は0.5の値を利用した。

3.2 ロボット速度・姿勢角速度制御系の構成

先の実験系の装置構成で述べたように、ロボットの制御系は聴覚情報に基づいて被験者が操縦したジョイスティックの入力量に対応する並進速度および姿勢角速度指令値に追従して運動するよう構成されなければならない。式 (1) より、直接観測はできない左右のモータの回転角速度を $\omega_R(t)$, $\omega_L(t)$, 左右モータに対する duty 比入力をそれぞれ $u_R(t)$, $u_L(t)$ とすれば

$$\frac{\Omega_R(s)}{U_R(s)} = \frac{0.094}{0.5s + 1} \triangleq P_R(s) \left(= \frac{\Omega_L(s)}{U_L(s)} \triangleq P_L(s) \right) \quad (2)$$

なる関係式が得られる。ここでモータユニット特性の個体差は無視している。すると、動輪の半径を R , 動輪間距離を L として、制御量 $V(s), \Omega(s)$ と操作量 $U_R(s), U_L(s)$ の間には

$$\begin{bmatrix} V(s) \\ \Omega(s) \end{bmatrix} = \begin{bmatrix} R/2 & R/2 \\ R/L & -R/L \end{bmatrix} \times \begin{bmatrix} P_R(s) & 0 \\ 0 & P_L(s) \end{bmatrix} \begin{bmatrix} U_R(s) \\ U_L(s) \end{bmatrix} \quad (3)$$

なる関係式が成り立つ。そこで、移動ロボットの並進速度および姿勢角速度を制御する入力 $U_V(s)$ および $U_\Omega(s)$ を考え、左右モータへの指令値を $U_R(s)$ および $U_L(s)$ を用いて

$$\begin{bmatrix} U_R(s) \\ U_L(s) \end{bmatrix} = \begin{bmatrix} R/2 & R/2 \\ R/L & -R/L \end{bmatrix}^{-1} \begin{bmatrix} U_V(s) \\ U_\Omega(s) \end{bmatrix} \quad (4)$$

にて生成することにすれば、 $P_R(s) = P_L(s) = P(s)$ のとき、

$$\begin{bmatrix} V(s) \\ \Omega(s) \end{bmatrix} = \begin{bmatrix} P(s) & 0 \\ 0 & P(s) \end{bmatrix} \begin{bmatrix} U_V(s) \\ U_\Omega(s) \end{bmatrix} \quad (5)$$

が得られ非干渉化できる。今回は速度および姿勢角速度の制御則に

$$\begin{aligned} U_V(s) &= K_{P_V} \left(1 + \frac{1}{T_{I_V} s} + T_d s \right) (V_R(s) - V(s)) \\ U_\Omega(s) &= K_{P_\Omega} \left(1 + \frac{1}{T_{I_\Omega} s} + T_d s \right) (\Omega_R(s) - \Omega(s)) \end{aligned} \quad (6)$$

で与えられる PID 制御則を採用した。ただし、 $V_R(s)$ および $\Omega_R(s)$ は、被験者のジョイスティック操作により与えられる並進速度と姿勢角速度の指令値である。結果として構成された制御系のブロック線図を図 6 に示す。図中に存在する観測器は、各モータの回転角速度を推定するもので、式 (1) に対して設計し、各動輪の回転角速度を推定する。式 (6) で与えられる制御則の計算で必要と

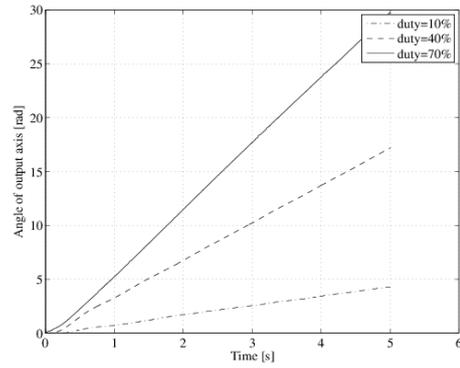


Fig.5 Step reponses for 3 different duty ratio inputs

Table 2 Hardware features of LEGO NXT motor unit

Element	Feature
control input	PWM duty ratio with direction sign (-100 to 100 integer)
gear ratio	48:1
encoder	12 [PPR] with two phased outputs (connected to motor with 10:32 gear resulting in 1[deg] resolution of output)

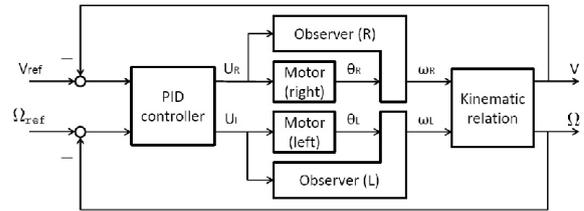


Fig.6 Block diagram of the PI control system for mobile robot

なる $V(s)$ および $\Omega(s)$ は、この観測器の出力を用いて計算したものをを用いている。式 (6) 中のゲインは、ジョイスティック操作に対する応答特性を観察しながらチューニングを行った。実験で利用した値は

$$K_{P_V} = K_{P_\Omega} = 5.0, T_{I_V} = T_{I_\Omega} = 0.2, T_d = 0.8$$

となっている。

4 実験内容

以上説明した実験装置を利用して、聴覚情報のみのフィードバックによる移動ロボットのナビゲーション実験を行った。実験の被験者となったのは、22歳～25歳の研究室男子学生4名で、全員聴覚に異常はなく、バイノーラル音による音像定位実験の被験者となった経験を有している。移動ロボットの走行環境は、大学キャンパ

ス内のラウンジの一角に図7に示すように設定した。実験開始時における移動ロボットの初期位置および姿勢は固定とし、その点を中心として半径3m, 中心角120°の扇形を描く。この扇形の円弧上の点について、ロボットの正面である方位角0°の点と、そこから左右に30°ごとに2点ずつの計5点を選び、ロボットナビゲーションにおける誘導目標となるスピーカークの設置位置とした。走行環境のうち2面が壁面に近接しているが、残る2面の壁は走行領域から遠く離れている。

実験では、被験者1名につき以下の手順で行った。まず、ナビゲーションの目標となるスピーカークの設置位置を図7の5か所の中から1か所ランダムに選択し、スピーカークを設置する。ついで「ホワイトノイズ」「ピンクノイズ」「1kHzの純音」の順に目標音を提示していく。1つの試行は音色の種類によらず最長90秒間で、音量一定の状態を持続音が提示される。別室にいる被験者は、目標音の再生開始と同時に移動ロボットの操縦を開始する。

被験者へは、スピーカークの正面にロボットを誘導するように指示しているが、スピーカークに十分近い状態ではスピーカークが障害となってその場旋回して姿勢角を修正することができない場合もあるため、被験者の操縦するロボットがスピーカークの十分近くに到達した時点で目標到達と実験の終了を被験者に知らせ、それ以上の操縦を行わせないようにした。一方、90秒間の試行時間の間に上記の意味で目標に到達できなかった場合はそこでその試行を打ち切るようにしている。

各被験者とも、再生装置は、ヘッドホン(Pioneer製SE-MJ31), ES-7001(3Dモード), ES-7001(ステレオモード)の順番で実験を行った。この実験を行うに当たっては、被験者が自身のジョイスティック操作に対して移動ロボットがどのように応答するかを事前に把握しておくことが必要である。そこで、聴覚ナビゲーション実験の実施前に移動ロボット操縦の練習時間を設け、すべての被験者に操縦を体験してもらったうえで実験を開始した。このとき、各被験者はロボットが走行する様子を自身の視覚で確認している。また、各被験者には、聴覚ナビゲーション実験の開始前に、図7に示すロボットの走行環境を目で見て確認してもらい、その後実験を開始した。

5 実験結果と考察

ここでは、前章で説明した実験手順に従って行った実験の結果を示し、考察を加える。まず、表3に走行実験の結果をまとめる。記号の意味は表下部で説明したとおりである。なお、図8にナビゲーションに成功した試行

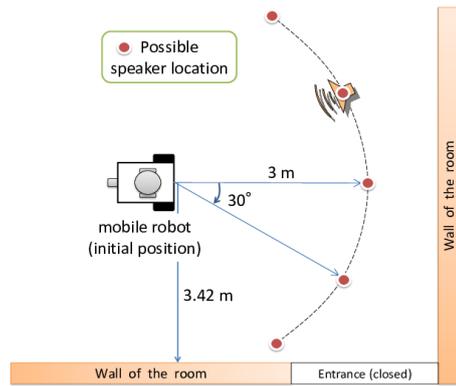


Fig.7 Setup of the Experiment

Table 3 Result of the experiment

	white	pink	pure tone
Headphone	○○×○	○△○○	○△××
ES-7001 (3D mode)	○○△○	○○○○	○×△○
ES-7001 (Stereo)	○○××	△××○	×△×○

What symbols in the table mean.

- ... Successful attempt
- △ ... Heading to correct direction, but ended an attempt not sufficiently close to target. Or mobile robot is close to target but operator lost direction.
- × ... Lost and/or ended an attempt too far from target

から被験者による操作入力とロボットの応答の一例を、図9に対応するロボットの走行軌跡を、図10に同一被験者が誘導に失敗した場合の操作軌跡を示した。図8および図9はヘッドホン再生のピンクノイズ利用、図10はES-7001(stereo)再生のホワイトノイズ利用の場合である。この失敗の試行においてスピーカークはもっとも右の位置に設置されているが、60秒を過ぎるあたりまで左の方に向けて走行していたことが方位角のグラフから見取れる。

提示音のうち1kHzの純音については別段の考察が必要のため、ホワイトノイズとピンクノイズの結果をもとに、再生装置がナビゲーション性能に与える影響について検証する。ヘッドホン再生とES-7001(3D)については両者の結果に差はない。複数の被験者から、主観的な意見としてES-7001(3D)の方がヘッドホンより距離感を感じやすかったという感想を得ている。バイノーラル音のヘッドホン再生では頭内定位が生じやすいとされているが、そのことを反映した感想であると考えられる。ヘッドホン再生の場合でも音源方位をとらえることは可

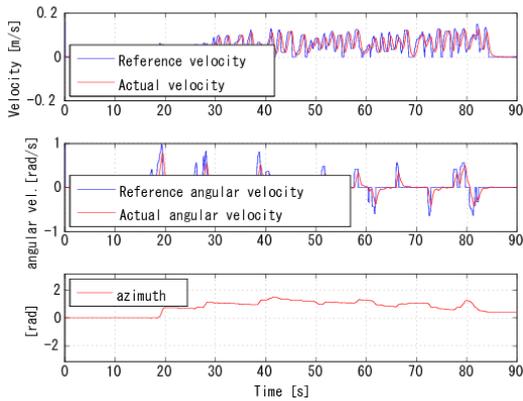


Fig.8 An example of navigation input and the robot response : Successful attempt

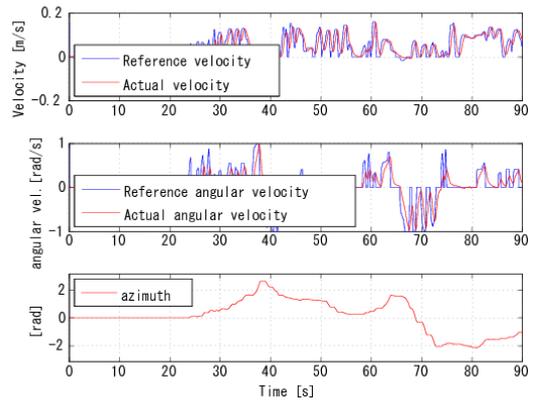


Fig.10 An example of navigation input and robot response : Unsuccessful attempt

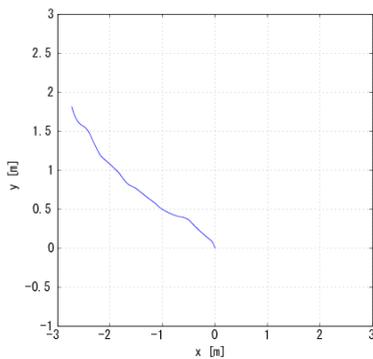


Fig.9 (x, y) trajectory corresponding to Fig.8

能であり、正しい方向に前進していれば再生音圧の増加により音源への接近を把握することは可能である。頭外定位の場合と比べ直感的に操作することはできないが、誘導は可能であったものと分析する。

ES-7001 のステレオエミュレーションモードでは、成績が悪化していることも当初の想定のとおりである。なお、表中○としている試行においても、ゴールの直前で方向を見失いその場旋回を行っていた試行が複数あった。

一方、1kHz の純音についてであるが、今回の実験で 1kHz の音を再生すると、実験場所とした部屋空間内に定在波が形成され、室内が音源からの距離とは無関係に音圧の高いところと低いところに分かれた。この音圧の空間分布は、スピーカーからの直接波と壁面や天井、床面などで反射した反射波の干渉により形成されたものと考えられる。実際、現実には人間がいる室内で純音を再生した場合、全く音源と異なる方向に定位することが多々ある。結果として、移動ロボットが室内を直進走行するだけでも、その走行経路によっては音源からの距離に関係なく音圧が上下することになり、HRTF で定量説明される音源方位角毎の両耳間レベル差を観測することはできない。文献²⁾では HRTF が畳み込まれた短い持続時

間の純音を、空白時間において繰り返し再生する音を提示することで良好なナビゲーション成績を得ているが、この検証は仮想空間内でのものであり、今回の実験結果は実空間では別の要因が結果を左右する可能性があることを示すものである。

6 まとめと今後の課題

本研究では、移動ロボットの活動環境で収録された音の情報に基づく移動ロボットのナビゲーションシステム構築に向けた検討の第一歩として、聴覚情報のみによる人間の音源定位能を利用した移動ロボットのナビゲーション実験システムを構築し、複数の被験者によるナビゲーション実験を行った。その結果、聴覚フィードバックがナビゲーションの手掛かりとして十分機能することを示した。

今後に関しては、まずは、このシステムにおいて音源定位精度の向上に資すると考えられる手法の適用について検討し、確認を行う必要がある。これには、收音マイクの頭部伝達関数と被験者のそれが異なることによる定位誤差の問題も含まれる。

今回の試行では、図 8 に示すように被験者がロボットを比較的小刻みに動かそうとする傾向が散見された。これには、距離感の認知が難しい状況下で、命綱である方位を正確に把握しようとする被験者の意図が影響しているものと考えられる。音源に対する距離変化の事実を音で知覚させる手法については検討の必要がある。

最終的には、視覚情報が得られる場合に視覚との効果的な協調を図る方法を確立するという観点から、システムの完成度を高めていきたいと考えている。

Acknowledgement

The authors would like to thank Mr. Allan Waddilove working for the Institute of Sound and Vibration Research, University of Southampton, United Kingdom for his work and suggestions to build up the experimental system.

参考文献

- 1) 日本音響学会編, 飯田一博, 森本政之編著, “空間音響学”, コロナ社, 2010
- 2) Bruce N. Walker and Jeffrey Lindsay, “Navigation Performance With a Virtual Auditory Display: Effects of Beacon Sound, Capture Radius, and Practice,” *HUMAN FACTORS*, Vol.48, No.2, Summer, pp. 265-278, 2006
- 3) L. Dunai *et.al.*, “The Influence of the Inter-Click Interval on Moving Sound Source Localization for Navigation Systems,” *Acoustical Physics*, Vol.56, No.3, pp.348-353, 2010
- 4) NXT OSEK/JSP, ANSI C/C++ with OSEK/ μ ITRON RTOS for LEGO Mindstorms NXT, <http://lejos-osek.sourceforge.net/>
- 5) Jens Blauert, “Spatial Hearing - The Psychophysics of Human Sound Localization -,” The MIT Press, 1996
- 6) NXTTM motor internals - Philo's Home Page, <http://www.philohome.com/nxtmotor/nxtmotor.htm>
- 7) T. Takeuchi and P. A. Nelson, “Optimal Source Distribution for Virtual Acoustic Imaging,” ISVR Technical Report No. 288, University of Southampton, 2000
- 8) B. Gardnew and K. Martin, “HRTF Measurements of a KEMAR Dummy-Head Microphone,” MIT Media Lab Perceptual Computing - Technical Report No.280, <http://sound.media.mit.edu/resources/KEMAR.html>, 1994
- 9) D.R.Begault, “3-D Sound for Virutal Reality and Multimedia,” Academic Press Professional Inc.

(平成27年3月10日受理)