

可聴音を用いた周波数選択に基づく距離推定法の実環境利用に向けた評価

Assessment of distance estimation using audible sound based on spectral selection.

高尾麻衣子¹, 干場功太郎¹, 中臺一博^{1,2}

Maiko TAKAO¹, Kotaro HOSHIBA¹, Kazuhiro NAKADAI^{1,2}

東京工業大学 工学院 システム制御系¹

(株) ホンダ・リサーチ・インスティテュート・ジャパン²

Tokyo Institute of Technology¹

Honda Research Institute Japan Co., Ltd.²

{mtakao,hoshiba,nakadai}@ra.sc.e.titech.ac.jp

Abstract

環境認識は長年にわたって研究されており、自動運転やロボットを利用した災害救助など様々な目的に利用可能である。本稿では、人間に不快感を与えない形で周辺環境認識を行うため、音楽など日常環境に存在する音響信号を用いた距離計測手法を提案する。既存の音響信号を用いた距離計測手法は分解能が不十分、狭帯域信号に対するノイズ耐性が低いといった問題がある。提案手法では周波数ごとに重みを設定することにより、これらの問題の解決を図る。実際に、室内で距離計測実験を行った結果、既存手法と比べ雑音環境下での検出性能で有効性を確認した。

1 はじめに

環境認識は長年にわたって研究されており、自動運転やロボットを利用した災害救助など様々な目的に利用可能である。周辺環境認識技術のうち、音響信号を用いた距離計測は、環境光の影響を受けにくいこと、比較的安価で構成できることなどのメリットがあるため、レーザーやカメラでは計測が難しい暗所やガラス面などでも利用可能である。

音響信号は人間の耳に聴こえない超音波信号と、人間の耳に聴こえる可聴音信号の2つに分けることができる。超音波信号による計測は、音響信号による計測の中では波長が短いため分解能がよく、低周波域に多く存在する環境ノイズ影響を受けにくいいため、広く用いられている。しかし、出力の際に耳障りな立ち上がり音が伴ってしまうことや、実際に人間の耳では超音波を感知することができないため、超音波曝露の危険性 [1] が考えられる。そのため、人間のいる環境において超音波を使用することは不向きといえる。可聴音を信号として使用する場合、送信信号を工夫すれば耳障りでない音響信号を利用した距離

計測が可能になる。また誤って大きな音を出してしまうとすぐに気づくことができ、曝露の危険性も少なくなる。一方で、超音波に比べて周波数が低いことによる分解能の劣化や、環境ノイズに影響を受けてしまうというデメリットがある。例えば、移動ロボットに搭載して足元の状態を確認する際に使用する場合を想定すると、段差を検知できる程度の分解能が必要である。また、実環境で使用することを考えるとノイズ耐性も必要となる。そこで本稿では可聴音を用いた周波数選択に基づく距離推定法を提案する。耳障りでない音として、送信信号に音楽信号を用いることを考える。実環境での実験を通し、既存手法と比較することで提案手法の有用性を評価する。

2 音響信号を用いた距離計測

音響信号を用いた距離計測の手法として、パルス波を用いて伝搬時間を計測する手法 [2]、相互相関により伝搬時間を計測する手法 [3]、定在波を利用する手法 [4]、受信信号の振幅により正規化した相互相関関数を用いた白色化相互相関 (Cross-power Spectrum Phase Analysis : CSP) 法 [5] などが挙げられる。パルス波を用いたパルスエコー法は最も単純な手法であるが、時間方向のパワーが小さいため雑音の影響を受けやすく、計測距離の分解能はその波長に依存するため、可聴領域では超音波領域と比べ分解能が低下してしまう。また、相互相関を用いた場合、時間方向に出力されたエネルギーを圧縮するため、雑音耐性が高くなる。しかし、分解能はサンプリング周期に依存するため、音声の解析などでよく用いられる 16 kHz 程度のサンプリングレートでは十分な分解能が得られるとは言えない。定在波を用いる場合、正確なキャリブレーションが必要となってくるため、ロボットのような移動体への実装が困難であると考えられる。CSP 法は、高い雑音耐性を有しているが、広帯域の雑音がある場合は、性能が劣化してしまう。また基本的には相互相関処理に基づく手法であるため前述のように分解能はサンプリング周波

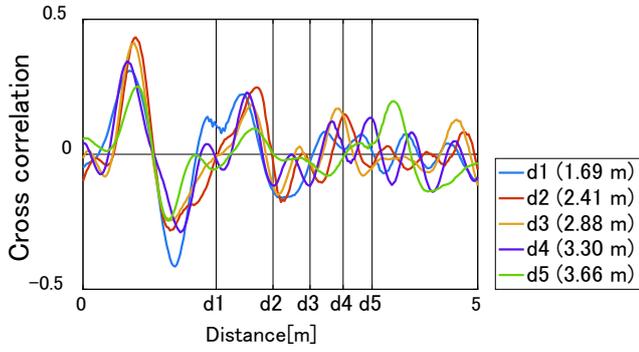


図 1: Distance estimation using cross correlation.

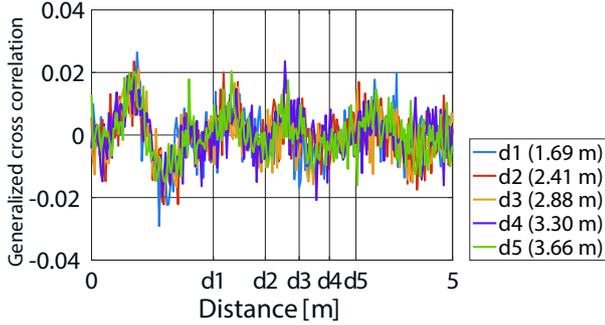


図 2: Distance estimation using CSP.

数に依存する。

実際に、上記の既存手法を利用した距離計測の問題点を確かめるため、2つの実験を試みた。まず、相互相関処理を用いた際の問題点を確認する。SNRが30 dB程度の雑音が少ない環境で、スピーカから部屋の壁に向けて、音楽信号を出力し、その反射音を受信した。スピーカと壁の距離については5パターンの計測を行った。Fig. 1は、送信信号と受信信号の相互相関を計算した結果である。横軸は伝搬時間を距離に換算したもので、レーザー距離計による距離計測値をd1からd5で示す。0.5 m程度の場所に存在するピークは直接波による相互相関値を示し、最大ピークとして観測できる。一方で反射波のピークは雑音に埋もれてしまって、検出が難しいことがわかる。次に、CSP法の問題点を確認する。CSP法は前述の通り、広帯域の雑音がある場合や計測に用いる信号の周波数帯域が狭い場合、性能が劣化してしまう。以下のような実験でこれを確認する。先ほどの実験と同じ環境で、スピーカから部屋の壁に向けて、単一周波数の正弦波を送信し、その反射音を受信した。音源の周波数は3種類、音源と壁の距離は3パターンの計測を行った。距離ごとに計測した3種類の周波数の受信波をPC上で足し合わせ、疑似的な受信波を作成し、これについてCSP法で距離推定を行う。Fig. 2に結果を示す。横軸はFig. 1と同様に伝搬時間を距離に換算した値である。正しく計測ができている場合、d1~d5それぞれの位置にピークが現れるはずであ

るが、Fig. 2ではその周辺にピークを観測することができず、正しい距離推定ができていない。CSP法は全周波数を参照して相関処理を行う手法であるため、本実験のように計測で利用できる周波数が限られている場合には正しく計測できないことがわかる。

3 提案手法

前節で述べたように、既存手法では可聴領域で狭帯域の送信信号に対し雑音耐性を備えた距離計測を行うことは難しい。そこで、これらの問題を解決するため、周波数選択に基づいた距離推定法を提案する。

ターゲットからの反射波を含む受信信号を $s(t)$ 、参照信号を $s_{ref}(t)$ とし、これらの周波数領域での表現を $S(\omega)$ 、 $S_{ref}(\omega)$ とする。 $S(\omega)$ は周波数 ω におけるスペクトルの振幅と位相情報を持った複素信号となるため、以下のように参照信号との位相差 $\phi(\omega)$ を算出する。

$$\phi(\omega) = \text{angle}(S(\omega) \cdot \text{conj}(S_{ref}(\omega))) \quad (1)$$

ここで $\phi(\omega)$ は $0 \leq \phi(\omega) < 2\pi$ であり、各周波数成分の伝搬距離による位相のずれを表す。得られた位相差を、以下のように伝搬距離 $d_n(\omega)$ に変換する。

$$d_n(\omega) = \frac{2n\pi + \phi(\omega)}{2\pi} \lambda, \quad (n \in \mathbb{Z}_+) \quad (2)$$

ただし、 λ は周波数 ω の波長、 \mathbb{Z}_+ は0以上の整数である。このように、周期が λ の周期関数として $d_n(\omega)$ を定義することができ、これをFig. 3(a)に示す。横軸は距離で、青、赤、黄色の点はそれぞれ異なる周波数 $\omega_1, \omega_2, \omega_3$ の成分から求められた伝搬距離 $d_n(\omega_1), d_n(\omega_2), d_n(\omega_3)$ を示す。これは、上記に挙げた、位相差を用いる場合には伝搬距離が波長以内でないと距離を一意に定めることができないという問題に対応するため、位相の周回を考慮した伝搬距離候補を挙げるという意味を持つ。理想環境では、送信信号に含まれるすべての周波数に共通する $d_n(\omega)$ が存在し、その値が伝搬距離となる。Fig. 3(a)では3つの周波数の $d_n(\omega)$ が真値の位置で一致している。しかし実環境では雑音の影響で誤差が発生する。Fig. 4(a)に、計測に誤差が発生してしまった場合の計測における、Fig. 3(a)と同様の図を示す。ここでは計測誤差によって推定距離が真値と一致していないことがわかる。そこで、得られた $d_n(\omega)$ に対し、ガウス分布を畳み込んだ周期関数であると仮定して、伝搬距離尤度を以下のように定義する。

$$p_\omega(x) = \sum_{\omega} \frac{1}{\sqrt{2\pi\sigma^2(\omega)}} \exp\left(-\frac{(x - d_n(\omega))^2}{2\sigma^2(\omega)}\right) \quad (3)$$

ここで、 $\sigma^2(\omega)$ はガウス分布の分散、 x は距離である。このようにすることで、雑音による計測誤差を考慮した推定ができる。これにより、尤度を持った伝搬距離の候補を定義できる。Fig. 3(b)では理想環境での計測に対して伝

搬距離尤度を求めた結果を、Fig. 4(b) では誤差のある計測に対して伝搬距離尤度を求めた結果を示す。横軸が伝搬距離、縦軸が尤度であり、3つの周波数それぞれで求めた結果を異なる色で示す。Fig. 3(b) ではノイズによる計測誤差がないため、全ての $p_\omega(x)$ が真値の位置にピークを持っている。Fig. 4(b) では誤差が発生しているため、ピークが真値と少しずれている。しかし、伝搬距離尤度としてガウス分布を畳み込むことで、どの周波数でも真値の位置に十分大きな値を持っている。さらに、各周波数で得られる $p_\omega(x)$ に対し周波数重みを定義し、全周波数を統合したブロードバンド尤度関数を以下のように定義する。

$$p(x) = \sum_{\omega} p_\omega(x) \cdot W(\omega) \quad (4)$$

ここで、 $W(\omega)$ は各周波数に対する重み関数である。重みを定義することにより、参照信号の持たない周波数を除外したり、計測しやすい周波数の影響を強く受けるように設定できるため、より雑音の影響を受けにくい距離推定が可能となる。このようにして求められたブロードバンド尤度関数に対し、伝搬距離 D を以下のように算出する。

$$D = \underset{x}{\operatorname{argmax}} p(x) \quad (5)$$

つまり、ブロードバンド尤度関数の最大ピーク位置が求めたい伝搬距離 D である。Fig. 3(c) では理想環境での計測に対して求めたブロードバンド尤度関数、Fig. 4(c) では誤差のある計測に対して求めたブロードバンド尤度関数を示す。横軸が伝搬距離、縦軸が尤度を示す。Fig. 3(c) では理想環境での計測なので、真値の位置に問題なく最大ピークが観測できる。また、Fig. 4(c) でも同様に真値付近で最大ピークを観測することができる。伝搬距離尤度を求める際にガウス分布を想定したことにより、この時最大ピークのほかに大きなピークが見られるが、サンプルに使用する周波数を増やせばピーク以外の部分がランダム分布となり平滑化されるので、最大ピークがより強調される。重みを調整することでも、計測に関係のないピークを抑制することができる。

本手法の特徴は、 $\sigma(\omega)$ や $W(\omega)$ を状況に応じて適切に定めることができる点であり、広帯域にわたる特徴的なノイズの中であっても雑音耐性を持った距離推定が可能であることが期待できる。例えば、広帯域のノイズに対し一部の周波数域しか持たない狭帯域の信号を送信する場合、送信信号に含まれる周波数に対する重みを $W(\omega) = 1$ とし、それ以外の周波数に対する重みを0とすれば、計測に必要な周波数のみを使用できるので、雑音の影響を受けにくい距離推定を行うことができる。

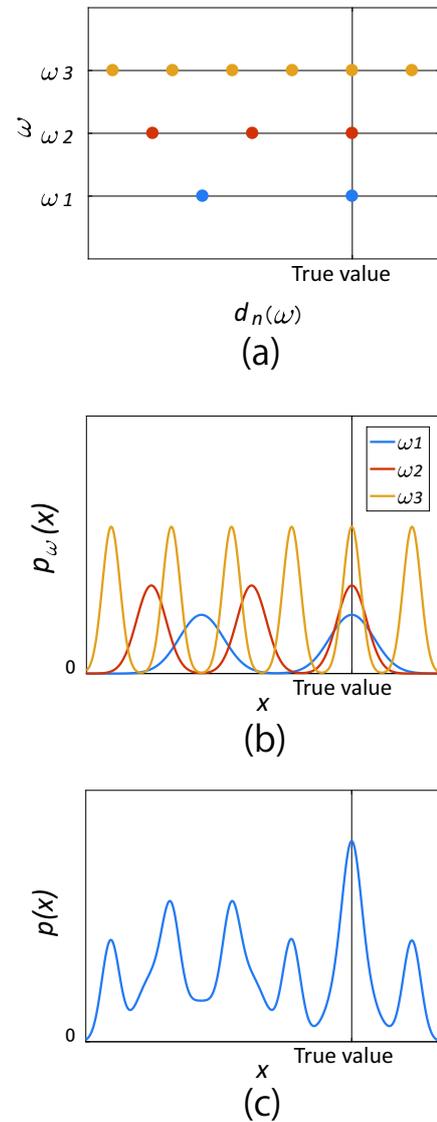


図 3: Algorithm of proposed method in noiseless environment. (a) $d_n(\omega)$, (b) $p_\omega(x)$, (c) $p(x)$.

4 評価実験

4.1 実験状況

提案手法の有用性を確認するため、評価実験を行った。実験状況を Fig. 5 に、計測機器の正面写真を Fig. 6 に示す。送波器 1 から信号を送波し、計測対象である壁からの反射波を受波器 1 にて受信する。また、参照信号 s_{ref} として、送波器 1 の近傍に設置した受波器 2 にて送波器 1 からの直接波を受信する。これは、送波器、受波器の周波数特性が計測に影響を与えないようにするためである。送波器 1 から受波器 1 への直接波を低減させるため、送波器 1 と受波器 2 の間に吸音材を挟み、送波器 1 をダンボールの覆いで囲った。計測は SNR が 30 dB 程度の雑音の少ない環境において、対象までの距離を 5 通りに変化させて実験を行う。送波器 1 には GENELEC 8010APM、受波器には MEMS マイクロホンを使用した。また、送信信号

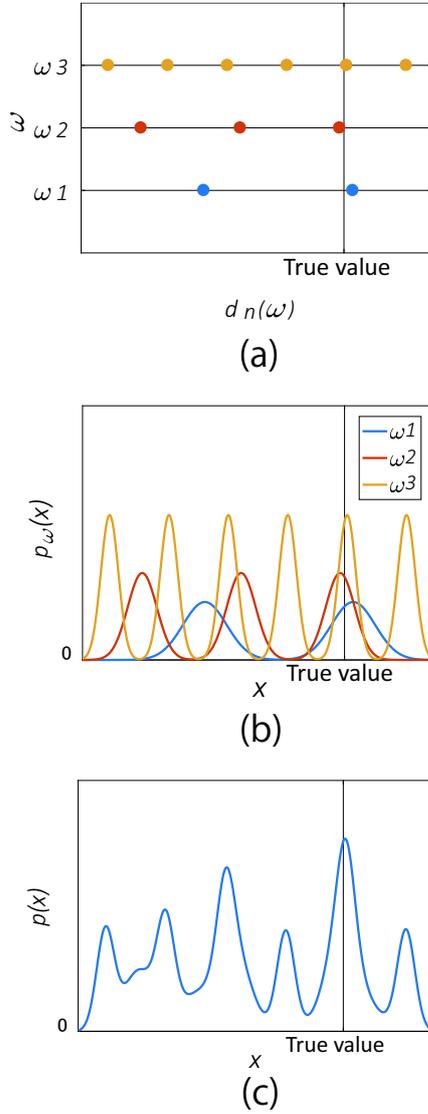


図 4: Algorithm of proposed method in noisy environment. (a) $d_n(\omega)$, (b) $p_\omega(x)$, (c) $p(x)$.

として, RWC 研究用音楽データベース (RWC-MDB-P-2001-M01NO.11) の男声楽曲を利用した. 送波器 1, 受波器 1, 2 は, システムインフロンティア社製の多チャンネル音響信号収録装置 RASP-24¹ に接続され, 同期収録を行う. 音響信号は, サンプル周波数 16 kHz, 量子化ビット数 16 bit で収録される. 収録した音響信号は, PC にて処理を行う. 受信信号の 0~0.5 kHz, 4~8 kHz の周波数帯域に対して, 中心値が $\pi/4$, 標準偏差が $\pi/6$ のガウスノイズに置き換えた. これは, 0.5~4 kHz 以外の周波数を参照すると, 方向性のある雑音として大きなピークが生じるように, 疑似的に設定したものである. 特に CSP 法を利用した場合, すべての周波数帯を参照するので, このノイズによる影響を強く受け, 計測とは関係のないピークによって正しい値が計測できないような状況に

¹http://www.sifi.co.jp/system/modules/pico2/index.php?content_id=4

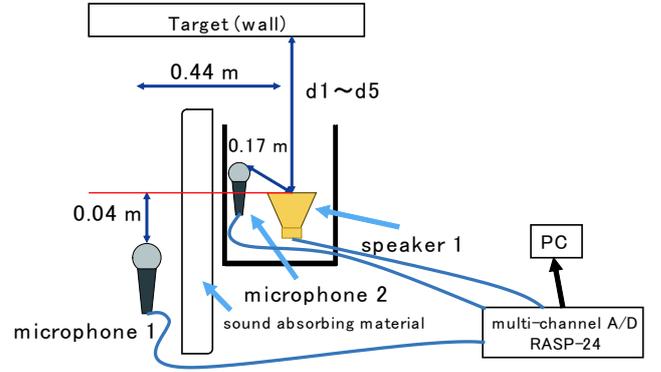


図 5: Experimental configuration.

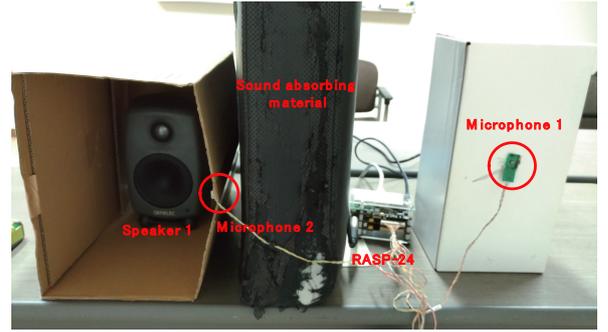


図 6: Measurement system for proposed method.

なると考えられる.

4.2 実験結果

実験を行った結果について述べる. 受波器 2 にて受信した信号の周波数スペクトルの例を Fig. 7 に示す. 周波数スペクトルから, 送信信号の周波数領域がほぼ 3 kHz 以下におさまっていることが確認できる. また, 500 Hz 以下の低周波において, 回り込みの影響で直接波が受波器 1 に到達しやすくなることがわかっている. これらを加味し, 以下のように $\sigma, W(\omega)$ を設定した.

$$\sigma = \frac{30}{\omega}, W(\omega) = \begin{cases} \sqrt{A} \cdot \frac{\sqrt{2\pi\sigma^2}}{\omega} & (0.5 < \omega < 3) \\ 0 & (\text{otherwise}) \end{cases} \quad (6)$$

ここで, A は ω における送信信号の周波数スペクトルの大きさを表す. 周波数が高くなるほど波長は短くなるので, σ は周波数に反比例する値とした. ガウス分布の特性上, 周波数ごとに伝搬尤度 $p_\omega(x)$ のピークの値が変わってしまうので, ピークの値を統一するために $W(\omega)$ として $\sqrt{2\pi\sigma^2}$ をかける. また, 周波数が低いと波長が長くなるため $d_n(\omega)$ の間隔が大きくなる. おおまかな距離推定のためには $d_n(\omega)$ の間隔が大きいもの, 詳細な距離推定のためには $d_n(\omega)$ の間隔が小さいものが有効である. そのため $1/\omega$ をかけることで, 低い周波数になるほど評価関数に強く影響するように周波数に反比例する値とした. さらに \sqrt{A} をかけることにより送信信号がより大きなパ

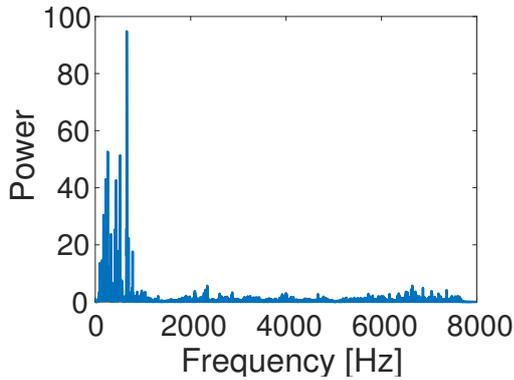


図 7: Spectrum of reference signal.

ワーを持つ周波数を重点的に見ることができる。

以上のパラメータを用い、CSP 法によって距離を算出した結果を Fig. 8(a) に示す。同様の実験について、実際に評価関数を算出した結果を Fig. 8(b) に示す。横軸は伝搬距離を示す。レーザー距離計によって計測した伝搬距離は 1.69 m, 2.41 m, 2.88 m, 3.30 m, 3.66 m であり、それぞれ d1, d2, d3, d4, d5 と表す。伝搬距離が異なる 5 通りの計測の結果を異なる色で示す。次に、同様の実験を同じ環境下で 20 回行った際の、レーザー距離計での計測結果との誤差の分布を算出した結果を Fig. 9 に示す。CSP 法では最大ピークが $x = 0$ 付近であり CSP 法そのものとの精度比較ができないため、 $x = 0.5$ 以降の結果からピーク探査し、伝搬推定距離 D を求める。それぞれの実験環境における誤差の大きさを縦軸に示す。赤が提案手法、青が CSP 法での結果である。

4.3 考察

Fig. 8(a) に示された、位相の偏ったノイズをかけたときの CSP 変換の結果を見ると、推定される伝搬距離 D となる最大ピークは $x = 0$ 付近に存在する。これはノイズ部分に疑似的に指定した位相差によって生じるものであり、正しい計測の弊害となることがわかる。計測距離付近にはピークが見られるものの、相対的には小さくなってしまっているので、ノイズによるピークの位置によっては計測がより困難になるといえる。したがって、参照信号が狭帯域の場合、その他の周波数帯域のノイズの影響により性能が劣化することが読み取れる。一方、Fig. 8(b) に示された提案手法での結果によると、それぞれレーザー距離計での計測距離と同じ位置に最大ピーク D が確認できる。CSP 法で発生したようなノイズ部分の影響はここでは見られないため、この条件の下で提案手法はより有効な計測となっているといえる。

また、Fig. 9 より、どの距離においてもレーザー距離計での計測値より近い結果を示しているのは提案手法である。CSP 法の分解能は 0.04 m 程度であり、提案手法では推定結果の分散が大きいですが、分解能は 0.01 m 程度と十

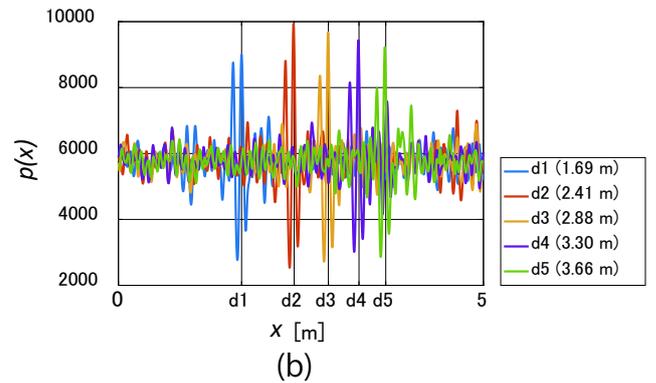
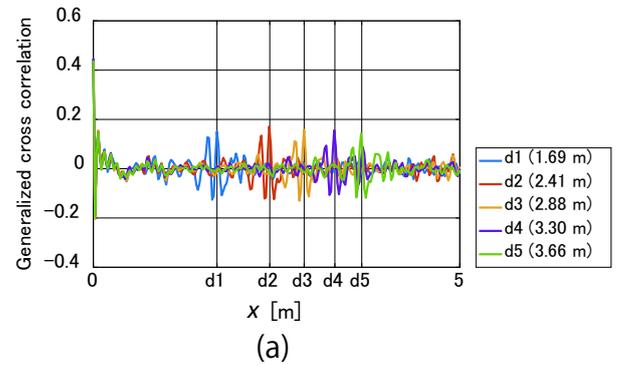


図 8: Distance estimation result using (a) CSP, (b) proposed method.

分小さくなっている。実環境において、例えば移動ロボットが足元の状態を確認するために使用する場合に必要となる分解能を考える。移動ロボットは壁や段差を検知し、進行方向を変えるなどの選択が必要である。特に段差の検知を見落としてしまうと、ロボットの転倒や故障の恐れがある。分解能が 0.04 m の場合、階段のような大きな段差は検知できるが、路肩のような段差を検知することはできない。一方、分解能が 0.01 m の場合、そのような段差でも検知することができる。したがって、この場合 0.01 m の分解能を持つ提案手法がより優れているといえる。

一方で、提案手法では各周波数についてそれぞれ処理を行うため、計算に時間がかかる。また、使用信号に低周波が含まれていない場合に計測が難しいなどの欠点もあり、実用化には改良の必要がある。

5 おわりに

本稿では、将来的なロボット聴覚への応用を目標に、可聴音領域の音楽音響信号をもとに距離を計測する技術について述べた。従来から用いられる手法や超音波計測で用いられる手法をそのまま適用しただけでは、可聴音領域信号での距離計測には問題があることを示し、これを解決する手法として、周波数重みを用いる手法を提案した。参照信号や環境ノイズ、周波数特性に応じた適切な周波数重みを用いることで、参照信号が狭帯域であっても口

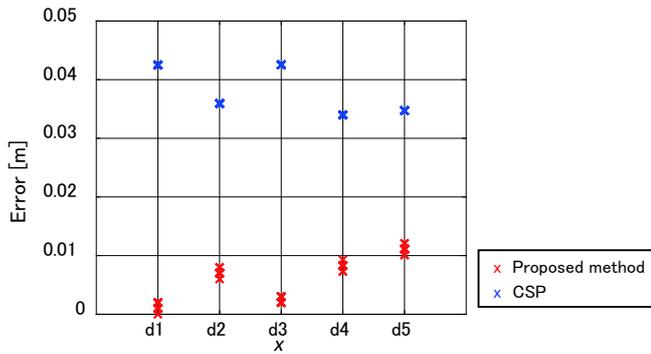


図 9: Error distribution.

バストに、かつ高分解能で距離推定が可能であることを示した。

今後は距離の計測だけではなく、二次元的、あるいは三次元的な地図の作成、システム（送受信器）が移動する場合や周囲の環境が動的に変化する場合への対応などの課題への取り組みを通じて、実用的な技術の構築を目指したい。

謝辞

本研究は、JSPS 科研費 16H02884, 16K00294, 17K00365 および、JST ImPACT タフロボティクスチャレンジの助成をうけた。

参考文献

- [1] Wesley L. Nyborg. Biological effects of ultrasound: development of safety guidelines: Part i: personal histories. *Ultrasound in Medicine & Biology*, Vol. 26, No. 6, pp. 911–964, 2000.
- [2] 松野保久, 山中有一. 2 球体の超音波 (50khz) 反射パルス波形情報利用のための実験に基づく検討. 日本水産学会誌, Vol. 56, No. 8, pp. 1219–1224, 1990.
- [3] 松尾成光, 平野憲雄, 片尾浩, 安藤雅孝. 超音波を利用した精密音響測距装置の開発. 東京大学地震研究所技術研究報告, No. 5, pp. 65–73. 東京大学地震研究所, 1999.
- [4] Hanabusa Shimpei, Uebo Tetsuji, Tsuchida Yuuta, Shinohara Toshihiro, and Nakasako Noboru. Distance estimation based on interference of audible linear chirp signal. *IEEJ Transactions on Electronics, Information and Systems*, Vol. 129, No. 11, pp. 2027–2033, 2009.
- [5] M Omologo and P Svaizer. Use of the crosspower-spectrum phase in acoustic event location. *IEEE Transactions on Speech and Audio Processing*, Vol. 5, No. 3, pp. 288–292, 1997.