Detecting Acoustic Reflectors using a Robot's Ego-Noise

PAPER ID : 3788

Usama Saqib*, Antoine Deleforge† and Jesper Rindom Jensen*
*Audio Analysis Lab, CREATE, Aalborg University, Denmark
† Universite de Lorraine, CNRS, Inria, LORIA, F-54000 Nancy, France

AALBORG UNIVERSITY DENMARK

1. Introduction

- Ego-noise is considered as a source of problem in many robotic applications, as it corrupts audio recordings captured by microphones [Nakadai, 2000][Okuno, 2015]
- Ego-noise reduction is an active area of research that plays an important role in many autonomous systems [Wake, 2019], and has enabled applications such as speech recognition for HRI or acoustic scene analysis
- Instead of treating ego-noise as a source of problem, we propose utilizing ego-noise constructively for acoustic reflector, e.g., wall, localization
- We propose an estimator for acoustic reflectors based on the *time difference of echo* (TDOE) which exploits the comb-filtering effect that emerges from the direct-path component of the sound source mixing with its delayed version, due to the presence of the acoustic reflector

2. Problem Statement

Consider a setup with a single microphone that records both the ego-noise generated by the rotors of a drone, $x[n]$, and a background noise from the environment $v[n]$. The signal model is then:

$$y[n] = (h * s)[n] + v[n] = x[n] + v[n], \quad (1)$$

Furthermore, we can rewrite (1) in a compact expression by separating $x[n]$ into the direct-path and early reflection components as:

$$y[n] = x_d[n] + x_r[n] + v'[n], \quad (2)$$

where $x_d[n] = g_1 s[n-\tau_1]$ and $x_r[n] = \sum_{q=2}^{R} g_q s[n-\tau_q]$ contains all the early reflections. If we vectorize (2) and express it in terms of the gains and delays:

$$\mathbf{y}[n] \approx g_d \mathbf{D}_{\tau_d} s[n] + g_r \mathbf{D}_{\tau_r} s[n] + \mathbf{v}'[n], \quad (3)$$

$$\mathbf{y}[n] = \begin{bmatrix} y[n] & y[n+1] & \cdots & y[n+N-1] \end{bmatrix}^T,$$

where $\mathbf{D}_\tau$ is a cyclic shift register. Assuming $R = 2$, the signal $\mathbf{x}_r[n]$ is then a delayed version of the direct-path component, (3) can be expressed as shown:

$$\mathbf{y}[n] = \mathbf{x}_d[n] + \frac{g_r}{g_d}\mathbf{D}_{\Delta\tau}\mathbf{x}_d[n] + \mathbf{v}'[n], \quad (4)$$

$$= (\mathbf{I} + \alpha \mathbf{D}_{\Delta\tau})\,\mathbf{x}_d[n] + \mathbf{v}'[n], \quad (5)$$

where $\Delta\tau$ is the TDOE of the observed signal, such that $\Delta\tau = \tau_r - \tau_d$ and $\alpha = \frac{g_r}{g_d}$, while $\mathbf{I}$ is the identity matrix.

3. Ego-noise-based acoustic reflector estimation and detection

$$\{\widehat{\Delta\tau}, \widehat{\alpha}\} = \underset{\Delta\tau,\alpha}{\arg\min} \|\mathbf{y} - (\mathbf{I} + \alpha\mathbf{D}_{\Delta\tau})\,\mathbf{x}_d\|^2 \quad (6)$$

$$= \underset{\Delta\tau,\alpha}{\arg\min}\, J(\Delta\tau, \alpha). \quad (7)$$

By zeroing the derivative with respect to α we get:

$$\frac{\delta J}{\delta \alpha} = -\mathbf{y}^T\mathbf{D}_{\Delta\tau}\mathbf{x}_d - \mathbf{x}_d^T\mathbf{D}_{\Delta\tau}^T\mathbf{y} + \mathbf{x}_d^T\mathbf{D}_{\Delta\tau}^T\mathbf{x}_d + \mathbf{x}_d^T\mathbf{D}_{\Delta\tau}\mathbf{x}_d + 2\alpha\mathbf{x}_d^T\mathbf{D}_{\Delta\tau}^T\mathbf{D}_{\Delta\tau}\mathbf{x}_d = 0. \quad (8)$$

By observing that $\mathbf{D}_{\Delta\tau}^T\mathbf{D}_{\Delta\tau} = \mathbf{I}$, this becomes:

$$\widehat{\alpha}(\Delta\tau) = \frac{(\mathbf{y} - \mathbf{x}_d)^T\mathbf{D}_{\Delta\tau}\mathbf{x}_d}{\|\mathbf{x}_d\|^2}. \quad (9)$$

$$\widehat{\Delta\tau} = \underset{\Delta\tau}{\arg\max}\ \widehat{\alpha}(\Delta\tau)^2. \quad (10)$$

Let us consider the following two hypotheses:

$$\mathcal{H}_0 : \mathbf{y}[n] = \mathbf{x}_d[n] + \mathbf{v}[n], \quad (11)$$

$$\mathcal{H}_1 : \mathbf{y}[n] = \mathbf{x}_r[n] + \mathbf{x}_d[n] + \mathbf{v}[n], \quad (12)$$

the generalized likelihood ratio test (GLRT) is given as:

$$\mathcal{L}(n) = \frac{p(\mathbf{y}; \mathbf{x}_r[n], \mathcal{H}_1)}{p(\mathbf{y}; \mathcal{H}_0)} > \gamma. \quad (13)$$

The probability density functions (PDFs) for the two hypotheses are given as shown:

$$\mathcal{N}(\mathbf{x}_d[n] + \mathbf{v}[n], \sigma^2), \quad (14)$$

$$\mathcal{N}(\mathbf{x}_r[n] + \mathbf{x}_d[n] + \mathbf{v}[n], \sigma^2), \quad (15)$$

where σ_v^2 is the variance of the background noise, $v[n]$.

$$\ln \mathcal{L}(\mathbf{x}) = \ln \frac{p(\mathbf{y}; \mathbf{x}_r[n], \mathcal{H}_1)}{p(\mathbf{y}; \mathcal{H}_0)} = (\mathbf{y}[n] - \mathbf{x}_d[n])^T(\mathbf{y}[n] - \mathbf{x}_d[n]) > 2\sigma_v^2 \ln\gamma. \quad (16)$$

Hence, the criterion to detect an acoustic reflector is:

$$T(\mathbf{y}) = \|\mathbf{y}[n] - \mathbf{x}_d[n]\|^2 \underset{H_0}{\overset{H_1}{\gtrless}} 2\sigma_v^2 \ln\gamma. \quad (17)$$

4. Experiments

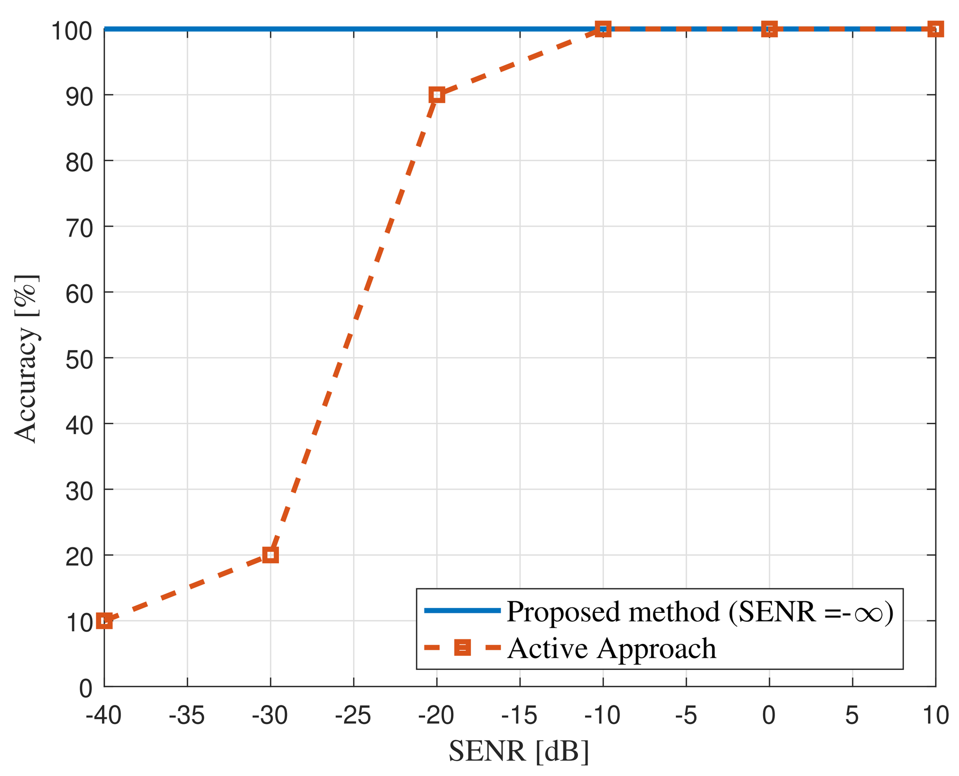

Fig. 1

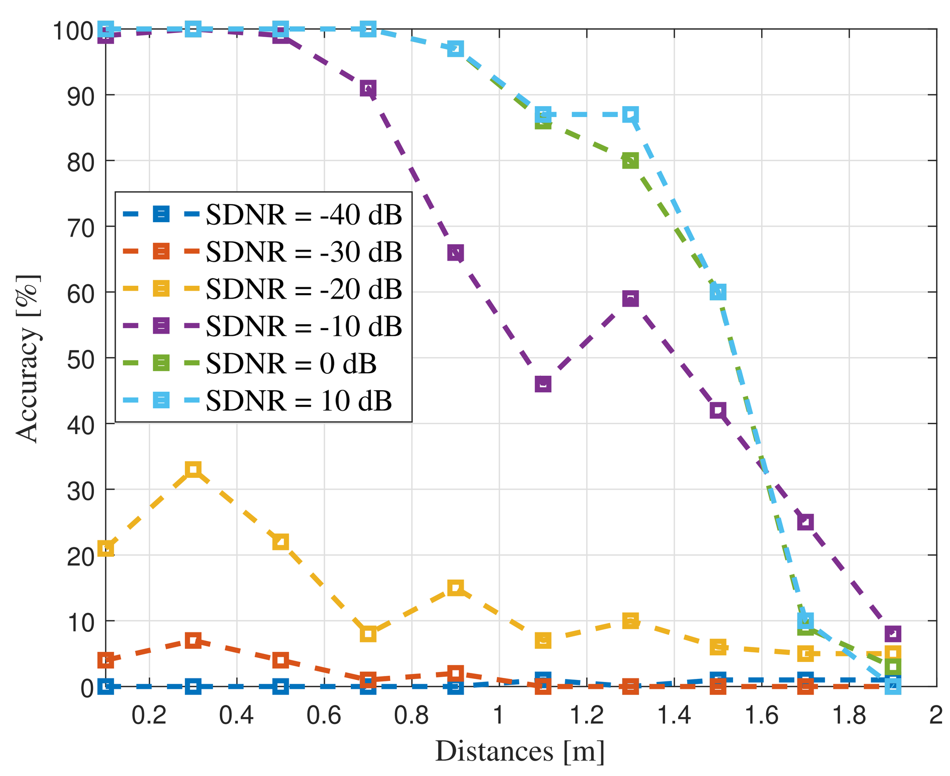

Fig. 2

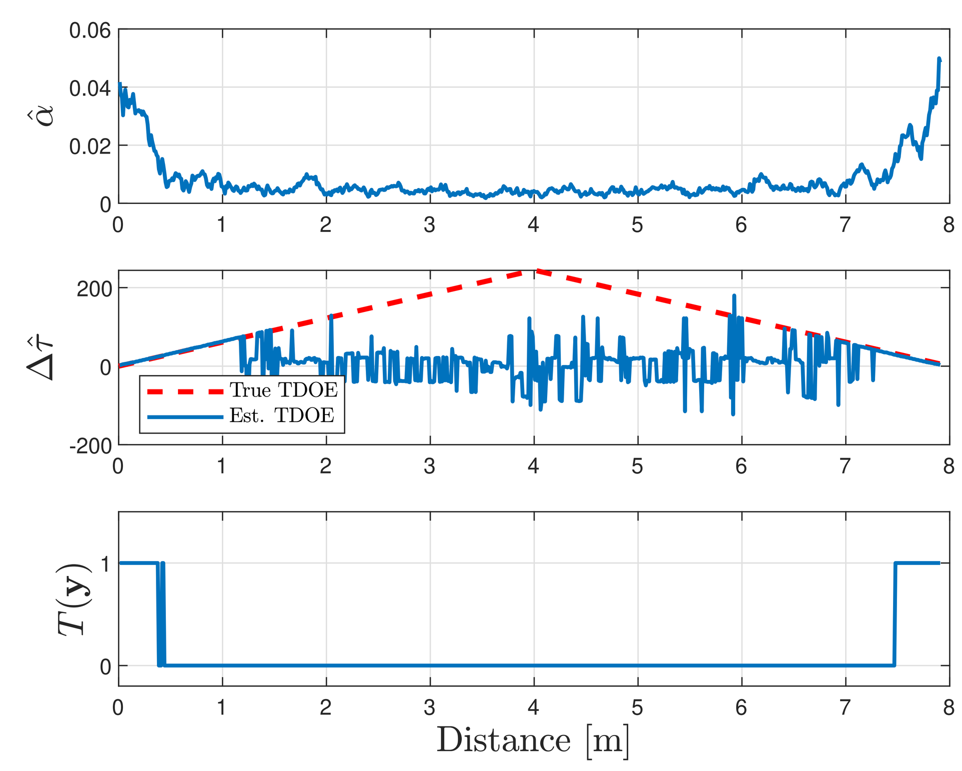

Fig. 3

5. Conclusion

- In this paper, we proposed a TDOE estimator and an echo detector to estimate the proximity of an acoustic reflector.
- The proposed method could lead to the development of new sound-based collision avoidance systems for, e.g., drones.
- The proposed method is robust against active/intrusive approach proposed by [Saqib, 2019] as shown in Fig. 1.
- The proposed method detects an acoustic reflector up to a distance of 1 m under low SDNRs as shown in Fig. 2