HAL
archives-ouvertes.fr

Spatial properties of the DEMAND noise recordings

Joachim Thiemann, Emmanuel Vincent, Steven van de Par

▶ To cite this version:

Joachim Thiemann, Emmanuel Vincent, Steven van de Par. Spatial properties of the DEMAND noise recordings. 40th Annual German Congress on Acoustics (DAGA 2014), Mar 2014, Oldenburg, Germany. hal-00985979

HAL Id: hal-00985979
https://hal.inria.fr/hal-00985979

Submitted on 30 Apr 2014

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Spatial Properties of the DEMAND Noise Recordings

Joachim Thiemann[1], Emmanuel Vincent[2], Steven van de Par[1]
[1] *"Hearing4All" Cluster of Excellence, Acoustics Group, CvO University of Oldenburg, 29126 Oldenburg, Germany*
[2] *Inria, Centre de Nancy - Grand Est, 54600 Villers-lès-Nancy, France*
Email: joachim.thiemann@uni-oldenburg.de

Introduction

"DEMAND" (Diverse Environments Multichannel Acoustic Noise Database) is a set of recordings of environmental noises in both indoor and outdoor settings. The recordings were performed with a 16-channel planar array of microphones. The purpose of the recording is to provide researchers with a large set of freely available noise recordings (licensed under a Creative Commons licence) for use in developing algorithms such as beamforming, noise reduction, and source separation, although anyone may use the data for any purpose they see fit. A more detailed description of the DEMAND recordings can be found in [1].

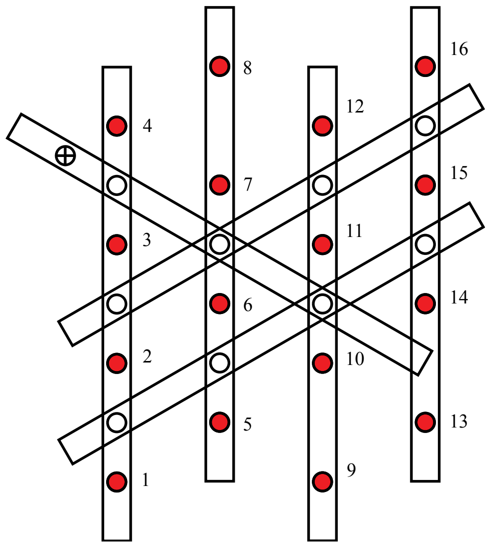

Figure 1: Schematic of the DEMAND microphone array construction. The filled circles indicate microphone positions, with adjacent numbering. The empty circles indicate bolts used to join the support rods, and the circle with cross indicates the tripod attachment.

In this article, we examine some of the spatial properties of the DEMAND recordings, in particular the cross-channel correlations. Notably, the quality of the reverberation characteristics is compared to the theoretical ideal. This property is used as a post-recording calibration of the microphone positions, compared to the design specifications of the array.

Coherence of Signals from Microphones in Diffuse Reverberant Noise

Given an ideal reverberant acoustic space, it is known that the average coherence between two signals, picked up by identical microphones in that space, is dependant only on the frequency, the distance between the microphones, and the speed of sound in that space [2]. This coherence can be expressed as

$$\Gamma_{ij}^{\text{diffuse}}(f) = \operatorname{sinc}\left(\frac{2\pi f d_{ij}}{c}\right), \tag{1}$$

where d_{ij} is the distance between the ith and jth microphone (in m), c is the speed of sound (m/s), f is the frequency (in Hz), and $\operatorname{sinc}(x) \equiv \sin(x)/x$.

Given a set of signals $x_i(f,n)$ in the short-time Fourier transform domain, we calculate the coherence by first computing the average cross-spectral density as

$$\phi_{ij}(f) = \sum_n x_i(f,n)x_j^\star(f,n), \tag{2}$$

where n is the frame index. From these cross-spectral densities the empirical coherence estimate is computed as

$$\Gamma_{ij}(f) = \frac{\phi_{ij}(f)}{\sqrt{\phi_{ii}(f)\phi_{jj}(f)}}. \tag{3}$$

Calibration of Microphone Arrays given Reverberant Signals

Given a set of time-aligned signals from pairs of microphones in a reverberant space, we can estimate the distance between the microphones by fitting the measured coherence, calculated using (3), to the theoretical curve given by (1). Similar to the algorithm for on-line estimation of microphone distances described by McCowan [3], we define an error function

$$\epsilon_{ij}(d) = \sum_{f=0}^{f_s/2} \left| \mathcal{R}\{\Gamma_{ij}(f)\} - \operatorname{sinc}\Big(\frac{2\pi f d}{c}\Big)\right|^2, \tag{4}$$

with f_s being the sampling frequency, and use a nonlinear optimisation algorithm to solve

$$d_{ij} = \operatorname{argmin}_d \epsilon_{ij}(d). \tag{5}$$

As in [3], the Levenberg-Marquart algorithm is used, as provided in MATLAB.

Matching the DEMAND data to the theory

Applying the above optimisation to the DEMAND recordings, we find a great variation in the residual error between each individual recording item. In Table 1, the RMS error is given for each item in DEMAND, computed as

$$E = \frac{1}{C} \sum_{c=1}^{C} \epsilon_c(d_c), \tag{6}$$

where C is the number of unique pairs of i and j, c enumerating the pairs. All items are analysed using a sampling frequency of 16 kHz, a 2048 sample FFT with a Hanning window, and a frame advance of 1024 samples.

Table 1: Sum of the residual errors after optimisation of d_{ij} over all microphone pairs, for each item in DEMAND.

Item	E	Item	E
DKITCHEN	25.95	PCAFETER	3.30
DLIVING	92.06	PRESTO	4.52
DWASHING	54.10	PSTATION	3.85
NFIELD	100.84	SPSQUARE	16.23
NPARK	24.09	STRAFFIC	10.47
NRIVER	24.19	SCAFE	15.39
OHALLWAY	99.70	TBUS	4.58
OMEETING	62.58	TCAR	51.09
OOFFICE	48.26	TMETRO	5.02

It can be seen that there is a large variation in how well the data fits (1), and this variation can be interpreted as an indication of the relative "diffuseness" of the noise in the recordings.

Calibration of the array using the "best" noise

Given a DEMAND recording which fits the theoretical model of diffuse noise reasonably well, we can infer the geometry of the array from the intermicrophone distance estimates using multidimensional scaling (MDS). In Figures 2 and 3, we show the results of using MDS to find a three dimensional geometry from the distances estimated from the interchannel coherences taken from the PSTATION recording. In Fig. 2, the black dots indicate the positions of the microphones as designed, while the red crosses show the estimated positions from the data.

In Fig. 3 we show the estimated vertical position of the microphones. Note that the array design is for all microphones to be aligned to the plane.

We note that the differences in geometry are small in the XY-plane (max 4.1 mm, average 2.1 mm), while the differences in the vertical axis are considerably larger (max 1.2 cm, average 6.1 mm). The magnitude of this difference suggest it is primarily an artifact of the estimation procedure.

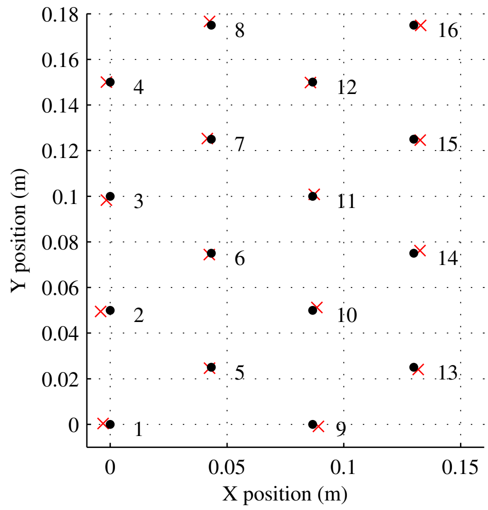

Figure 2: Comparison of the locations of the microphones as designed (dots) versus locations estimated from the audio signals (crosses) in the XY-plane.

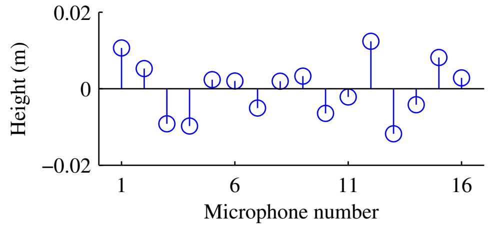

Figure 3: Estimated vertical displacement of microphones.

Discussion

The DEMAND recordings provide researchers with a large amount of data for use in the development of multichannel algorithms. In this study, we examine some of the spatial properties of the data in terms of the interchannel coherence, and compare it with respect to the theoretical prediction for diffuse reverberant noise. We find that there is considerable variation between the items in DEMAND. Noting that fitting the empirical coherence estimate to the theoretical coherence can be used to estimate intermicrophone distances, we use MDS to infer the geometry of the array from the recordings. We find a close match in the XY plane, while the vertical positions are less well estimated. This demonstrates that DEMAND can be used to explore multichannel algorithms, and gives a glimpse at interesting properties that one could still extract from the database.

References

[1] J. Thiemann, N. Ito, and E. Vincent, "The DEMAND database of multichannel environmental noise recordings," in *Proc. Int. Cong. Acous. 2013 (ICA2013)*, Montreal, Canada, 2013.

[2] R. Cook, R. Waterhouse, R. Berendt, S. Edelman, and M.C. Thompson, Jr, "Measurement of correlation coefficients in reverberant sound fields," *J.Acoust. Soc. Amer.*, vol. 27, pp. 1072–1077, 1955

[3] I. McCowan, M. Lincoln, and I. Himawan, "Microphone Array Shape Calibration in Diffuse Noise Fileds," *IEEE Trans. Audio, Speech, and Language Proc.*, vol. 16, no. 3, March 2008