University of Pennsylvania
ScholarlyCommons

Departmental Papers (EES)

Department of Earth and Environmental Science

September 2000

Finding the right cross-correlation peak for locating sounds in multipath environments with a fourth-moment function

John L. Spiesberger
University of Pennsylvania, johnsr@sas.upenn.edu

Follow this and additional works at: http://repository.upenn.edu/ees_papers

Recommended Citation

Spiesberger, J. L. (2000). Finding the right cross-correlation peak for locating sounds in multipath environments with a fourth-moment function. Retrieved from http://repository.upenn.edu/ees_papers/4

Copyright ASA. Reprinted from *Journal of the Acoustical Society of America*, Volume 108, Issue 3, September 2000, pages 1349-1352.
Publisher URL: http://dx.doi.org/10.1121/1.1287709

This paper is posted at ScholarlyCommons. http://repository.upenn.edu/ees_papers/4
For more information, please contact libraryrepository@pobox.upenn.edu.

Finding the right cross-correlation peak for locating sounds in multipath environments with a fourth-moment function

Abstract
To locate calling animals in reverberant environments from recordings on widely separated receivers, a fourth-moment "Augmented-Template Correlation Function" (ATCF) helps identify which of many peaks in each cross-correlation function is that corresponding to the difference in travel times for the first arrivals (reference-lag). This peak may not be the largest. The ATCF, by providing an approximate correlation between auto- and cross-correlation functions, can be orders of magnitude more efficient in selecting the reference-lag than the alternative of randomly selecting peaks. The ATCF's efficacy increases with the number of paths and their signal-to-noise ratios.

Comments
Copyright ASA. Reprinted from *Journal of the Acoustical Society of America*, Volume 108, Issue 3, September 2000, pages 1349-1352.
Publisher URL: http://dx.doi.org/10.1121/1.1287709

This journal article is available at ScholarlyCommons: http://repository.upenn.edu/ees_papers/4

Finding the right cross-correlation peak for locating sounds in multipath environments with a fourth-moment function

John L. Spiesberger
Department of Earth and Environmental Science, 240 South 33rd Street, University of Pennsylvania, Philadelphia, Pennsylvania 19104-6316

(Received 11 March 1999; revised 3 April 2000; accepted 23 May 2000)

To locate calling animals in reverberant environments from recordings on widely separated receivers, a fourth-moment ''Augmented-Template Correlation Function'' (ATCF) helps identify which of many peaks in each cross-correlation function is that corresponding to the difference in travel times for the first arrivals (reference-lag). This peak may not be the largest. The ATCF, by providing an approximate correlation between auto- and cross-correlation functions, can be orders of magnitude more efficient in selecting the reference-lag than the alternative of randomly selecting peaks. The ATCF's efficacy increases with the number of paths and their signal-to-noise ratios.
© *2000 Acoustical Society of America.* [S0001-4966(00)00909-7]

PACS numbers: 43.60.Cg, 43.60.Gk [JCB]

I. INTRODUCTION

A method was recently introduced for locating calling animals recorded on widely separated receivers.[1,2] The wide-bandwidth calls traveled to the receivers along multipath. In the rural environments in which this method was applied, locations were not obtained by fitting a forward model to the data[3–9] because multipath reflected from obstacles whose locations were unknown, such as uneven topography, trees, rocks, and houses. The only paths that could be used to locate animals were the ones which arrived first because they were nearly straight.[1] The cross-correlation function contained many contending peaks for the desired lag, called a reference-lag, corresponding to the difference in travel times of the first arrivals. One of the methods used to narrow down the number of candidates for the reference-lag was based on a fourth-moment function called an ''Augmented-Template Correlation Function'' (ATCF). The purpose of this paper is to demonstrate that the ATCF is more effective in selecting the reference-lag than is possible by choosing peaks at random. This is the first quantitative assessment of this function.

The ATCF also plays a role in the blind solution of the impulse response of the channel at each receiver.[2,10] Blindly estimating the impulse response of a channel has applications in acoustic and electromagnetic communication systems, as well as tomographic imaging systems of environmental fields.[1,10,11]

II. AUGMENTED-TEMPLATE CORRELATION FUNCTIONS

The ATCFs are defined as

$$F_{ij}(T) \equiv \int_0^{\tau_{\max}} \breve{R}_{ii}(\tau)\breve{R}_{ij}(\tau - T)\,d\tau, \tag{1}$$

$$B_{ij}(T) \equiv \int_0^{\tau_{\max}} \breve{R}_{jj}(\tau)\breve{R}_{ji}(\tau - T)\,d\tau, \tag{2}$$

where the peak-selection function is[2]

$$\breve{R}_{ij}(\tau) \equiv \begin{cases} 0; & \text{if there is no peak at lag } \tau \text{ in } R_{ij}(\tau) \\ 1; & \text{if there is a peak at lag } \tau \text{ in } R_{ij}(\tau), \end{cases} \tag{3}$$

and the auto- and cross-correlation functions are

$$R_{ij}(p) \equiv \frac{1}{K}\sum_{k=1+p}^{K} r_i(k) r_j(k-p); \quad p \geqslant 0; \quad 1 \leqslant i < j \leqslant \mathcal{R}. \tag{4}$$

For $i \neq j$, the indices on k go from 1 to $K+p$ for p less than zero. The kth time sample at receiver i is $r_i(k)$. There are $\mathcal{R}$ receivers. For auto-correlation functions, $i=j$. The effectiveness of ATCFs comes from the fact that the relative travel times of paths at any receiver appear as similar patterns in the auto- and cross-correlation functions. When the auto- and cross-correlation functions are lined up properly, the ATCF tends toward large values. A complete discussion of this function appears on pp. 304–305 and Fig. 2 of Ref. 2.

An objective function used to estimate the efficacy of ATCFs in selecting the reference-lag is

$$O_{ij}(T) \equiv F_{ij}(-T) + B_{ij}(T), \tag{5}$$

where candidates for the reference-lag, $T = T_{FB_{ij}}$, are taken from the intersections of three sets,

$$\begin{aligned} T_{FB_{ij}} \in & \{T_{F_{ij}}(1), T_{F_{ij}}(2), \ldots T_{F_{ij}}(P_{F_{ij}})\} \\ & \cap \{T_{B_{ij}}(1), T_{B_{ij}}(2), \ldots T_{B_{ij}}(P_{B_{ij}})\} \\ & \cap \{\tau_{ij}(1), \tau_{ij}(2), \ldots \tau_{ij}(P_{\text{data}})\}. \end{aligned} \tag{6}$$

The number of candidate reference-lags from the F- and B-ATCFs are $P_{F_{ij}}$ and $P_{B_{ij}}$, respectively. The third set contains the P_{data} peaks occurring in the cross-correlation function $i-j$. The intersection with the third set guarantees that candidate reference-lags are taken from the set of picked peaks in the cross-correlation function. Thus, the reference-lag cannot be found in this study unless it is picked in the cross-correlation function. Since ATCFs tend to have large values when T is the reference-lag,[2] ranked candidates for reference-lags, $T_{FB_{ij}}$, are taken by ordering values of

$O_{ij}(T_{FB_{ij}})$ in monotonically decreasing order.

The peak-selection function utilizes peaks that exceed a certain signal-to-noise ratio. In this paper, a statistical peak-selection function is introduced which suppresses smaller peaks that are more likely due to noise than signal. Once the superiority of the statistical peak-selection function is established, further exploration of the efficacy of ATCFs will be based only on the statistical peak-selection function. The efficacy of the ATCFs will be investigated as a function of the number of paths and their signal-to-noise ratios.

III. STATISTICAL PEAK-SELECTION FUNCTION

Assume there are K data at receiver i. For simplicity, assume that the greatest lag p considered in the cross-correlation function, $R_{ij}(p)$, obeys $|p| \ll K$. Assuming noise samples at the receiver are mutually uncorrelated, then $R_{ij}(p)$ is, asymptotically with K for $|p| \ll K$, a mutually uncorrelated Gaussian random variable with mean zero and variance σ_{ij}^2 except for lag zero of the auto-correlation function which is not treated here. The Gaussian nature of the correlation values is a result of the central limit theorem.[12] The variance, σ_{ij}^2, is taken to be a constant, which is a good approximation when $p \ll K$. So, the probability density function of $R_{ij}(p)$ is

$$f_{R_{ij}}(R_{ij}(p)) = \frac{1}{\sqrt{2\pi\sigma_{ij}^2}} \exp\left[-\frac{R_{ij}^2(p)}{2\sigma_{ij}^2}\right]; \quad p>0. \tag{7}$$

One of many possible statistical peak-selection functions is

$$\breve{R}_{ij}(p) = 1 - \exp\left[-\frac{R_{ij}^2(p)}{2\sigma_{ij}^2}\right], \quad \text{if } R_{ij}(p) \geqslant R_0, \tag{8}$$

and is zero if $R_{ij}(p) < R_0$. The peak-selection function has values from zero, for small peaks, to one, for big peaks.

In practical situations, the noise and reverberation may be correlated. The results in this paper may not be sensitive to the amount of correlation because the variance of the noise and reverberation, σ_{ij}^2, can be empirically estimated as a function of lag.

R_0 is chosen by specifying the probability that selected peaks are due to noise. Since $R_{ij}(p)$ is a random variable, so is $\breve{R}_{ij}(p)$. The probability that a noisy peak exceeds R_0 is

$$P(\breve{R}_{ij}(p) \geqslant R_0) = 2\,\mathrm{erfc}(\sqrt{-2\ln(1-R_0)}), \tag{9}$$

where the complementary error function is

$$\mathrm{erfc}(x) \equiv \frac{1}{\sqrt{2\pi}} \int_x^\infty \exp(-t^2/2)\,dt.$$

Given $P(\breve{R}_{ij}(p) \geqslant R_0)$, R_0 is found numerically.

IV. CANDIDATE REFERENCE-LAGS CHOSEN AT RANDOM

The probability of randomly selecting the reference-lags among the candidates at $\mathcal{R}$ receivers is computed next. The number of cross-correlation functions that can be formed from $\mathcal{R}$ receivers is $\mathcal{R}(\mathcal{R}-1)/2$. For the cross-correlation between receivers i and j, suppose there are n_{ij} candidates for the reference-lag and one of these is correct. All physically possible reference-lags must have values between $-d_{ij}/c$ and d_{ij}/c, where d_{ij} is the distance between the receivers and c is the slowest speed at which sound propagates. If a peak is picked at random without replacement from each cross-correlation function, what is the probability, $p_f(f)$, that all the reference-lags are picked after making f random selections from each of the $\mathcal{R}(\mathcal{R}-1)/2$ cross-correlation functions? For example, $p_f(f=1)$ is the probability that all $\mathcal{R}(\mathcal{R}-1)/2$ reference-lags are chosen from the first selection.

To find this probability function, suppose for example there are three receivers. Then,

$$p_f(f=1) = \frac{1}{n_{12}} \frac{1}{n_{13}} \frac{1}{n_{23}}.$$

The probability that the reference-lags are picked among the first two selections is

$$p_f(f=2) = \frac{\min(2,n_{12})}{n_{12}} \frac{\min(2,n_{13})}{n_{13}} \frac{\min(2,n_{23})}{n_{23}},$$

where $\min(a,b)$ is the minimum value of a and b. So the general formula for $\mathcal{R}$ receivers is

$$p_f(f) = \prod_{i=1}^{j-1} \prod_{j=2}^{\mathcal{R}} \frac{\min(f,n_{ij})}{n_{ij}}. \tag{10}$$

Monte Carlo simulations will be used to measure the efficacy of ATCFs in comparison with this probability. So there is a need to define an average probability of $p_f(f)$ over many simulations,

$$\overline{p_f(f)} \equiv \frac{1}{Q(f)} \sum_{q=1}^{Q(f)} p_{f,q}(f), \tag{11}$$

where $p_{f,q}(f)$ is $p_f(f)$ for simulation q and $Q(f)$ is the number of simulations which have f selections.

V. SIMULATIONS

Three receivers are located on the circumference of a circle of radius 15 m. A source is placed at its center. Each receiver picks up five paths with the difference in arrival time between the last and first being 0.4 s. Arrival times are selected using a uniform random number generator. The speed of sound is 330 m/s. The signal-to-noise ratio of the first path correlated with itself, R_{ii}, in each auto-correlation function is set to 15 dB, i.e.,

$$15\,(\mathrm{dB}) = 10\log_{10}\frac{R_{ii}^2}{\sigma_{ii}^2}. \tag{12}$$

The amplitudes of the remaining paths decay following a spherical spreading of energy. The signal-to-noise ratio is set so that there is a high probability that the first arrivals will be picked in the cross-correlation functions, but not so high so that the later arrivals will all be picked. The value for R_0 is chosen so that there is a 0.2 probability that a selected peak is due to noise only. This corresponds to picking peaks with signal-to-noise ratios exceeding 2.2 dB. The correct reference lag appears among each set of picked peaks from each

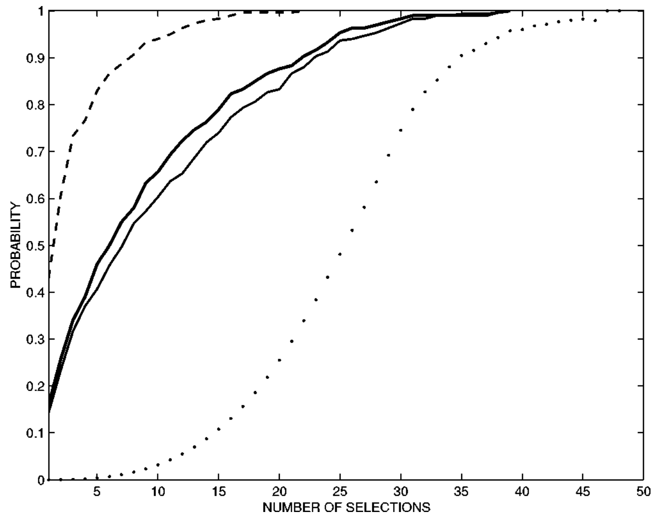

FIG. 1. Probability of choosing all three correct reference-lags as a function of the number of times one lag has been selected, without replacement, from each of the three pools of candidate reference-lags from each of the cross-correlation functions. The dotted line refers to random selection. The probability of obtaining the correct reference-lags using the Augmented-Template Correlation Functions (ATCFs) is shown as a function of the indicated number of selections using the statistical (thick line) and nonstatistical (thin line) peak-selection functions. Selections from the ATCFs are taken from the set $T_{FB_{ij}}$ [Eq. (6)] in the order of decreasing size of $O_{ij}(T_{FB_{ij}})$. The dashed line shows the efficacy of the ATCFs when the statistical peak-selection function is used with simulations having higher signal-to-noise ratios.

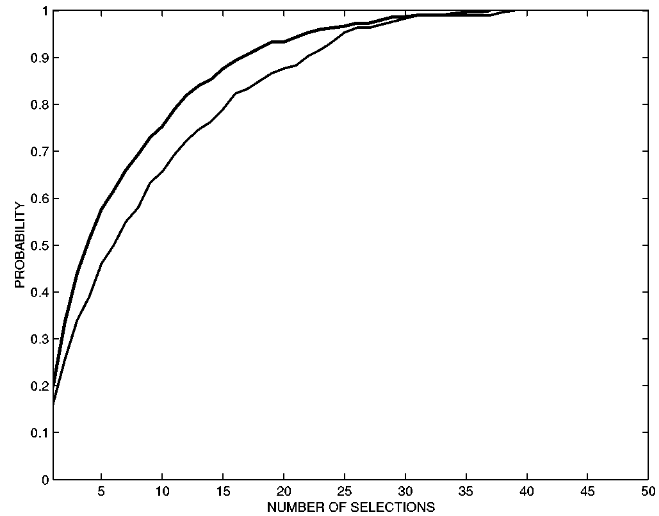

FIG. 2. Probability of selecting all three correct reference-lags using the ATCFs as in Fig. 1 as a function of the number of selections made from the candidate pools of reference-lags from the three cross-correlation functions. The thick and thin lines are based on simulations using ten and five paths, respectively, at each receiver.

cross-correlation function. Each simulation below uses 100 Monte Carlo runs.

ATCFs based on the statistical peak-selection function yield higher probabilities of selecting the reference-lags than the nonstatistical peak-selection function (thick and thin lines, Fig. 1).

The average number of candidate reference-lags found at each receiver for the statistically-based peak-selection function is 33. The reference-lags are found with much higher probabilities from the ATCFs than by random selection (thick and dotted lines, Fig. 1). For example, when the probability of selecting all three reference-lags using ATCFs is 0.8, the probability of randomly selecting all three reference lags is 0.1. The gain in probability is 0.8/0.1=8.

The efficacy of ATCFs increases with the signal-to-noise ratio. For example, when the signal-to-noise ratio in Eq. (12) is raised to 40 dB, the probabilities of selecting the reference-lags increase significantly (dashed and thick solid lines, Fig. 1). When the probability of selecting all three reference-lags using ATCFs is 0.8, the probability of randomly selecting all three reference-lags is 0.004. The corresponding gain in probability is 0.8/0.004=200.

ATCFs do better at larger signal-to-noise ratios because there are more signal-related peaks in the correlation functions, offering a more complete template of signal-related patterns to be used by the ATCFs. In fact, the average number of signal-related peaks that are picked from three auto- and three cross-correlation functions is 36 for the 15 dB simulation and 93 for the 40 dB simulation. The total number of signal-related lags in all the correlation functions is about 105, so less than half of these are picked for the 15 dB case. The improved efficacy of ATCFs is unrelated to the number of noisy peaks, because both the 15 and 40 dB simulations contain the same average number of noisy peaks, about 760, in all the correlation functions.

Reference 2 hypothesized that the efficacy of ATCFs should increase with the number of paths at each receiver. With more paths, the probability that noise could look like the more complicated templates of signal-related lags should decrease. Indeed, when the number of paths at each receiver is increased from five to ten, the probability of selecting all the reference-lags increases (Fig. 2).

VI. CONCLUSIONS

The difference in the travel times of the first arrivals at two receivers is useful for locating a naturally occurring sound such as a calling animal when the corresponding paths are known to be nearly straight.[1] Cross-correlation functions may not tell us what this difference is when there are many peaks of similar amplitudes due to multipath. A quantitative assessment of a fourth-moment function, called the Augmented-Template Correlation Function (ATCF), shows its effectiveness in identifying the desired difference compared to alternative methods. For example, simulations show that the ATCF is able to identify the difference in travel time with probabilities that are between 8 and 200 times that found by randomly selecting candidates for the reference-lags from the peaks in the cross-correlation functions. The performance of the ATCFs improves with the signal-to-noise ratio and the number of paths. These results are the first quantitative measure of the efficacy of this function.

It is important to emphasize that the ATCF alone may not provide a reliable enough estimate for the desired difference called the ''reference-lag.'' It is part of the algorithm used to blindly estimate the reference-lag. The reliability of the algorithm in simulations and field experiments comes also from the use of the lag and amplitude equations which link the auto- and cross-correlation functions.[1,2] The ATCF is incorporated in two ways. First, it helps reduce the computational burden of that algorithm by providing guidance for reference-lag candidates. Second, it is an objective function used for selecting reference-lags.

ACKNOWLEDGMENTS

This work was supported by ONR Grant No. N00014-97-1-0613. During the review process of Ref. 2, an anonymous reviewer mentioned the possibility of letting the peak-selection function take on a continuum of values due to the presence of noise. This idea motivated the author to explore this idea. I thank the reviewers for the comments.

[1] J. L. Spiesberger, ''Locating animals from their sounds and tomography of the atmosphere: Experimental demonstration,'' J. Acoust. Soc. Am. **106**, 837–846 (1999).

[2] J. L. Spiesberger, ''Linking auto- and cross-correlation functions with correlation equations: Application to estimating the relative travel times and amplitudes of multipath,'' J. Acoust. Soc. Am. **104**, 300–312 (1998).

[3] W. S. Burdic, *Underwater Acoustic System Analysis*, 2nd ed. (Prentice-Hall, Englewood Cliffs, NJ, 1991).

[4] H. P. Bucker, ''Use of calculated sound fields and matched-field detection to locate sound sources in shallow water,'' J. Acoust. Soc. Am. **59**, 368–373 (1976).

[5] M. D. Collins and W. A. Kuperman, ''Focalization: Environmental focusing and source localization,'' J. Acoust. Soc. Am. **90**, 1410–1422 (1991).

[6] J. L. Krolik, ''Matched-field minimum variance beamforming in a random ocean channel,'' J. Acoust. Soc. Am. **92**, 1408–1419 (1992).

[7] S. Li and P. M. Schultheiss, ''Depth measurement of remote sources using multipath propagation,'' IEEE J. Ocean Eng. **18**, 379–387 (1993).

[8] S. Narasimhan and J. L. Krolik, ''Fundamental limits on acoustic source range estimation performance in uncertain ocean channels,'' J. Acoust. Soc. Am. **97**, 215–226 (1995).

[9] A. Kristensen, A. Caiti, F. Ingenito, M. D. Max, J. M. Berkson, M. D. Collins, L. T. Fialkowski, N. C. Makris, E. McDonald, J. S. Perkins, and W. A. Kuperman, ''Environmental confirmation of focalization,'' in *Full Field Inversion Methods in Ocean and Seismo-Acoustics*, edited by O. Diachok *et al.* (Kluwer Academic, Dordrecht, 1995), pp. 153–158.

[10] ''Special issue on blind systems identification and estimation,'' Proc. IEEE **86**, 1903–2089 (1998).

[11] J. L. Spiesberger and K. M. Fristrup, ''Passive localization of calling animals and sensing of their acoustic environment using acoustic tomography,'' Am. Nat. **135**, 107–153 (1990).

[12] R. J. Serfling, *Approximation Theorems of Mathematical Statistics* (Wiley, New York, 1980).