HIGH PERFORMANCE SUPERVISED TIME-DELAY ESTIMATION USING NEURAL NETWORKS

Ludwig Houégnigan[1], Pooyan Safari[2], Climent Nadeu[2], Mike van der Schaar[1], Marta Solé[1], Michel André[1]

1 Laboratory of Applied Bioacoustics (LAB), Polytechnic University of Catalonia, UPC Barcelona Tech, Spain
2. TALP Research Center - Dept. TSC, Polytechnic University of Catalonia, UPC Barcelona Tech, Spain

ABSTRACT

Time-delay estimation is an essential building block of many signal processing applications. This paper follows up on earlier work for acoustic source localization and time delay estimation using pattern recognition techniques; it presents high performance results obtained with supervised training of neural networks which challenge the state of the art and compares its performance to that of well-known methods such as the Generalized Cross-Correlation or Adaptive Eigenvalue Decomposition.

Index Terms— time-delay estimation, neural networks, source localization, acoustics

1. INTRODUCTION

Time-delay estimation (TDE) is a task as fundamental as spectral estimation and a key step for many popular applications such as sonar and radar direction finding, seismology, biomedicine, satellite navigation or acoustic source localization.

Recent advances in machine learning invite us to revisit classical signal processing problems. Over the past years, the authors of this paper have proposed original approaches [1,2] using machine learning and data-specific modelling to improve TDE in the context of both air and underwater acoustic source localization (biological sources such as cetaceans, or artificial ones such as pingers, ships, navy sonar, etc).

Comprehensive studies on TDE such as [3] were published but none of them, to our knowledge, has ever included supervised learning. On the other hand, little has yet been published on time-delay estimation using supervised learning besides benchmark papers by Shaltaf et al. [4, 5] that were working only in a very limited set of conditions which we successfully extended.

2. MODELS AND METHODS FOR TIME-DELAY ESTIMATION

In the proposed approach - contrary to correlation-based estimators [6, 7], minimum entropy [8] or Eigenvalue Decomposition (AED) [9] - no particular assumption is made with regard to modelling. The capabilities of neural networks for system identification and interpolation permit to construct a system that minimizes the error between its output and an ideal response represented by a peak at the localization of the correct time-delay. It is proposed here to provide as target the dirac delta function: $\delta(n-\tau_{12})$.

3. IMPLEMENTATION AND TEST

Eight datasets, each containing 400,000 chirp signals were constructed. Each signal featured random duration (with 10 to hundreds of samples at a sampling rate of 16 kHz, this number being randomly selected from uniform distributions) and varying noise. In each of seven first dataset the variance $\sigma_S{}^2$ of the signal of interest is related to the noise variance $\sigma_N{}^2$ by a factor of respectively 0, 0.2, 0.4, 0.5, 0.6, 0.8, 1. The last dataset contains 400000 samples where that ratio is drawn uniformly between 0.2 and 1. These noisy datasets aim at mimicking adverse conditions which typically cause failures in time-delay estimators as will be shown in section 4.

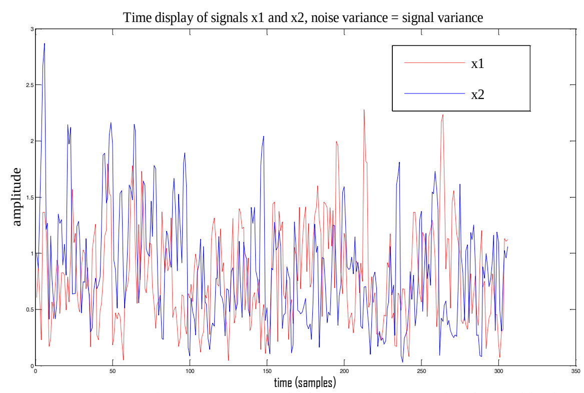

Fig.1 signals x1 and x2 are composed of noisy chirp signals of variable duration.

3.2 Neural network parameters

Multilayer perceptrons ("mlps") architectures including a single hidden layer and 30 hidden units were used. Sigmoid and linear activation functions were respectively used for the hidden and output units. The training procedure was conducted using a standard backpropagation algorithm with a fixed mini-batch size of 100 and 100 epochs. The fixed momentum and weight decay for all the systems were respectively set to 0.9 and 10^{-7}. A sparsity target of 0.05 and a sparsity penalty of 10^{-4} were used for all the networks. L2 regularization norm was set to 10^{-3}.

For each dataset, 19 neural nets were trained with varying learning rates. Among those 19 nets, for the sake of concision, only 4 were selected and are displayed here, namely the nets providing respectively the best (lowest) mean error, the worst (highest) mean error, the best (lowest) variance and the worst (highest) variance, referred to respectively as MLPA, MLPB, MLPC, MLPD. Those nets are then compared to 4 other estimators: the generalized cross-correlation with SCOT and PHAT filters (GCC-SCOT and GCC-PHAT), standard unbiased cross-correlation (XCOR) and the Adaptive Eigenvalue Decomposition (AED).

4. RESULTS

Figure 1. represents the output of cross-correlation estimator (XCOR), the output of an “mlp” and the ideal response (target) when the noise in the training data is variable. It can be observed that cross-correlation is performing poorly at estimating the nominal delay whereas the neural network closely matches the target. Overall shape of the distribution is also slightly affected by noise as shown by the measurement of the Kullback-Leibler divergence. Q_{KL} . the neural network, although it performs much closer to the target than cross-correlation does, is noisier than previously and has more leakage and ripples. This is adequately reflected by the Q_{KL} measures: $Q_{KL}(\text{Target})=0$, $Q_{KL}(\text{MLP}) = 0.1895$, $Q_{KL}(\text{XCOR}) = 14.30$.

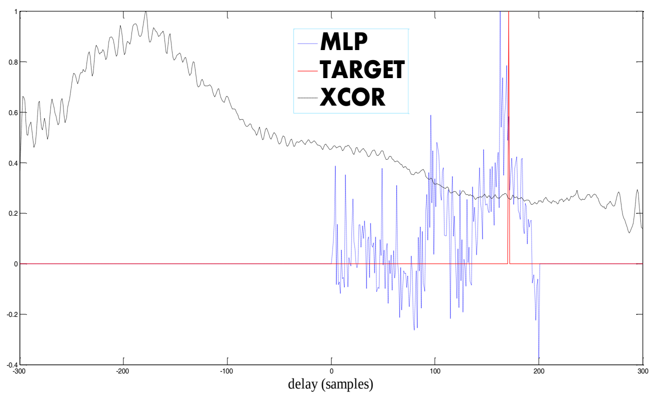

Figure 1. mlp and xcor estimators against target. Variance is variable (dataset 8).

Table 1 summarizes the evolution of the mean of the error for the tested estimators as variance changes.

Table 1. Mean of the error of various estimators for current datasets

	MLPA	MLPB	MLPC	MLPD	SCOT	PHAT	XCOR	AED
Noise Var. 0	0.98	5.56	0.98	5.56	0.40	0.16	0	43.29
N.V.0.2	2.13	3.46	2.19	3.19	196.74	196.38	222.51	207.95
N.V.0.4	4.86	6.88	4.86	6.69	181.46	181.09	271.72	200.99
N.V.0.5	6.60	9.23	8.55	9.05	171.46	171.02	292.26	197.93
N.V.0.6	8.55	12.12	11.15	11.43	161.99	161.57	304.85	195.04
N.V.0.8	13.09	17.66	16.49	16.57	147.86	147.38	313.32	191.80
N.V. 1	18.74	23.27	18.95	19.49	138.76	138.22	312.02	191.41
Variable N.V.	8.84	11.62	8.84	11.62	171.77	171.35	274.01	199.54

As variance increases, the neural solutions prove to perform consistently better than any of the other methods at stake. Even in high noise it is found that the neural solution remains satisfactory.
Standard boxplots *(figure 2)* provide us additionally with a compact understanding of the performance of the various estimators and some additional statistics. It can be observed that all trained “mlps” systematically outperform all non-supervised methods when noise is present.

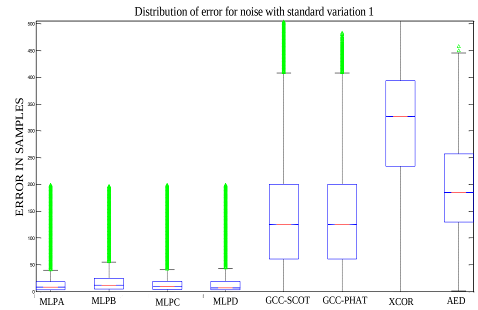

Figure 2. boxplot representation of the error distribution of various estimators when noise variance equals signal variance.

5. CONCLUSIONS

In this paper supervised neural networks were used for a successful time-delay estimation and proved to outperform benchmark methods both for the nominal estimation of time-delay and in approximating an ideal time-delay response. As an entry for localization this robust time-delay estimates would produce drastically more consistent location estimates. The integration of these improved time-delay estimators both in underwater and in room acoustics is the object of ongoing research projects.

6. REFERENCES

[1]André, M., et al. "Localising Cetacean Sounds for the Real-Time Mitigation and Long-Term Acoustic Monitoring of Noise, Advances in Sound Localization." *InTech* (2011).

[2] Houegnigan, Ludwig, et al. "Neural networks for the localization of biological and anthropogenic source at neutrino deep sea telescope." *OCEANS 2015-Genova*. IEEE, 2015.

[3] Chen, Jingdong, Jacob Benesty, and Yiteng Huang. "Time delay estimation in room acoustic environments: an overview." *EURASIP Journal on applied signal processing* 2006 (2006): 170

[4]Shaltaf, Samir. "Neural-network-based time-delay estimation." *EURASIP Journal on Applied Signal Processing* 2004 (2004): 378-385.

[5]Shaltaf, Samir J., and Ahmad A. Mohammad. "Neural networks based time-delay estimation using DCT coefficients." *American Journal of Applied Sciences* 6.4 (2009): 703.

[6] Carter, G. C. "Time delay estimation for passive sonar signal processing." *IEEE Transactions on Acoustics, Speech, and Signal Processing* 29.3 (1981): 463-470.

[7] Knapp, Charles, and Glifford Carter. "The generalized correlation method for estimation of time delay." *IEEE Transactions on Acoustics, Speech, and Signal Processing* 24.4 (1976): 320-327.

[8]Benesty, Jacob, Yiteng Huang, and Jingdong Chen. "Time delay estimation via minimum entropy." *IEEE Signal Processing Letters* 14.3 (2007): 157-160.

[9] Benesty, Jacob. "Adaptive eigenvalue decomposition algorithm for passive acoustic source localization." *The Journal of the Acoustical Society of America* 107.1 (2000): 384-391.