Mitsuo Ohta, Yoshifumi Fujita

Inter-subjective relationship of higher-order among spatial-temporal wavy environmental factors – a methodological trial based on a standpoint of "Relationism-First"

APPENDIX IV

EXPERIMENTAL SITUATION RELATED TO TIME SERIES REGRESSION MODEL IN SOUND ENVIRONMENT AND EVALUATION OF PREDICTION ERROR

Figure A4-1 is the system diagram for the measurement in the sound experiment laboratory [1–3] (data 1). The input signal is a white noise of the center frequency 250 Hz with one-third octave band. Due to the reverberation characteristics of the room, there naturally exists the correlation among time series values of signal observed at different times, especially at the low frequency region. The acquired data are transformed into a discrete form by the A-D converter, at the sampling rate of 0.1 s (called case 1) or 0.09 s (called case 2).

The sampling interval should be essentially set in relation to the frequency bandwidth to be restored. In this experiment, two different sampling intervals are used due to the following reasons. The proposed method is applied to the discrete signal with any sampling interval. In the time region, the correlation between random process at time k and $k-1$ is decreased (increased) when the sampling interval is increased (decreased). In other words, the order to be used in the time-series model depends on the sampling interval, even for the same signal.

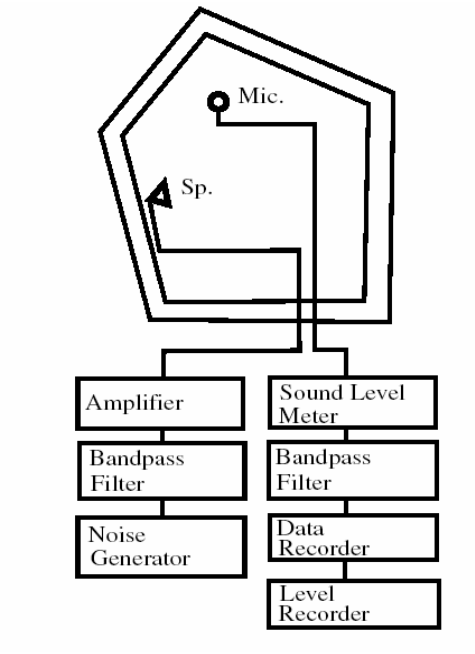

Figure A4-1.

Block diagram of the acoustical experiment

The sound waveform for the washer (data 2) is recorded once on a tape recorder at the measurement site. The replayed data are transformed by the A-D converter in the sound experiment laboratory at the sampling interval of 0.08 s. For each case, 100 non-stationary data are recorded and are divided into training data and prediction data for the locally stationary interval. In other words, the autoregression parameters (expansion coefficients A_{mn} in Eqs. (18) and (27)) are calculated from the training data, and the sequential prediction of the time series is attempted using the obtained expansion coefficients for the prediction data. Equation (17) is used for case 1 of data 1. Equation (26) is used for case 2 of data 1 and data 2.

Next, the parameters of the autoregression model are estimated using the training data. More precisely, A_{mn} are estimated by the moment method following the definitions of Eqs. (18) and (27) (called estimation 1). The parameters also estimated based on the infinite series expansion of Eqs. (17) and (26) for the autoregression model. Considering that there are only a finite number of data in the actual case and only a finite number of expansion coefficients can be utilized, the evaluation criterion that $J = \langle \varepsilon_k^2 \rangle$ should be the minimum is introduced and the autoregression parameters are estimated from external observations (called estimation 2).

In the case of Eq. (10), for example, the autoregression model

$$x_k = c_0 + c_a \left[\sum_{N=0}^{M-1} \sum_{n_1+n_2+\cdots+N_p=N} A_{1\mathbf{n}} \psi_{\mathbf{n}}^{(2)}\left(\mathbf{x}_{k-1}\right) \Big/ \sum_{N=0}^{M} \sum_{n_1+n_2+\cdots+N_p=N} A_{0\mathbf{n}} \psi_{\mathbf{n}}^{(2)}\left(\mathbf{x}_{k-1}\right) \right] + \varepsilon_k \qquad \text{(A4-1)}$$

is used considering that the number of expansion coefficients is finite. Hereupon, $A_{1\mathbf{n}}$ is determined so that

$$J = \left\langle \left(x_k - c_0 - c_a \left[\sum_{N=0}^{M-1} \sum_{n_1+n_2+\cdots n_p=N} A_{1\mathbf{n}} \psi_{\mathbf{n}}^{(2)}\left(\mathbf{x}_{k-1}\right) \Big/ \sum_{N=0}^{M} \sum_{n_1+n_2+\cdots n_p=N} A_{0\mathbf{n}} \psi_{\mathbf{n}}^{(2)}\left(\mathbf{x}_{k-1}\right) \right] \right)^2 \right\rangle \qquad \text{(A4-2)}$$

is minimized(least-squares method). In this case, $A_{0\mathbf{n}}$ denotes the statistical data of only $\mathbf{x}_{k-1}$ and are estimated beforehand by the moment method.

Tables A4-1 to A4-4 show the evaluation of the error (mean-square error between the actual value x_k and the predicted value $\hat{x}_k$) given by

$$e \triangleq \left\{ \frac{1}{n} \sum_{k=1}^{n} \left(x_k - \hat{x}_k \right)^2 \right\}^{\frac{1}{2}} \left(n \text{ is the number of data} \right) \qquad \text{(A4-3)}$$

for each datum, each case, and each estimation.

Table A4-1. Evaluation of error of predicted value to measured value (case 1 of data 1).

Estimation error	Prediction error [dB]			
	AR Model	Proposed time series regression model		
		M=2	M=3	M=4
1	1.0012	1.0012	0.9980	0.9918
2	1.0012	1.0012	0.9980	0.9963

Table A4-2. Evaluation of error of predicted value to measured value (case 2 of data 1).

Estimation error	Prediction error [dB]			
	AR Model	Proposed time series regression model		
		M=2	M=3	M=4
1	1.0077	1.0077	1.0005	0.9904
2	1.0077	1.0077	1.0005	0.9974

Table A4-3. Evaluation of error of predicted value to measured value
(estimation method 1 for data 2)

p	Prediction error [dB]			
	AR Model	Proposed time series regression model		
		M=2	*M*=3	*M*=4
1	—	1.3478	1.3283	1.3250
2	1.3335	1.3335	1.3051	1.3229

Table A4-4. Evaluation of error of predicted value to measured value
(estimation method 2 for data 2)

p	Prediction error [dB]			
	AR Model	Proposed time series regression model		
		M=2	*M*=3	*M*=4
1	—	1.3448	1.3315	1.3257
2	1.3335	1.3335	1.3045	1.2524

It can be seen in particular from Tables A4-3 and A4-4 that the prediction performance of the proposed method is better for p=1(although it is smaller than the optimal order p=2) if the higher-order expansion terms are considered than the AR model for p=2. Thus, it is verified also experimentally that the proposed method is a generalization, including the well-known (linear type) AR model as the special case.

Figure A4-2 shows the cumulative distribution functions (abbr. cdf) for the data1 (derived from the prediction data), once after the above correlation analysis based on the idea of "Relationism-First". To determine the extent of non-Gaussian property of the data, the figure also shows the result when the statistical Hermite expansion distribution expression is used, which is applicable to any type of non-Gaussian distribution with the Gaussian distribution in the first expansion term.

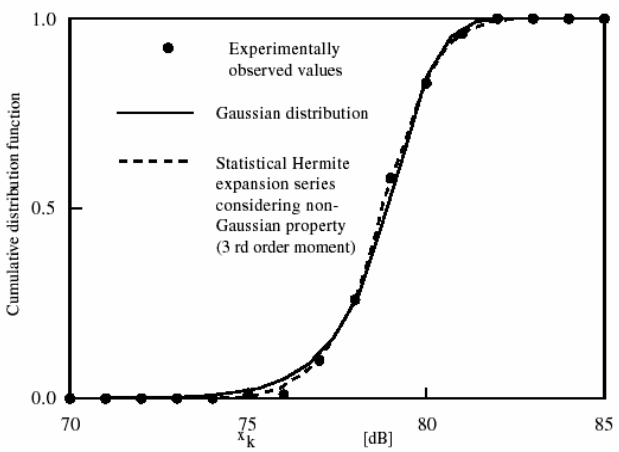

Figure A4-2.

Cumulative distribution function of data 1

It can be seen from Figure A4-2 that data 1 are considerably close to the Gaussian distribution. However, the result is closer to the actual value if the expansion coefficient reflecting the third-order moment also is considered.

REFERENCES

1. M. Ohta. Several dynamical state estimation methods for environmental system of acoustic noise and vibration type. I. Methodological style and establishment of a general theory (in Japanese). Japanese society of automatic control: System and control (explanation), vol. 23, N°10, 27–33, 1979. II. State estimation theory of energy system and its application of living environment (in Japanese). Japanese society of automatic control: System and control (explanation), vol. 23, N°12, 26–33, 1979.
2. M. Ohta, K. Hatakeyama. Stochastic system theory of acoustic environment and its digital signal processing. Japanese society of automatic control (explanation, special issue), vol. 31, N°11, 823–831, 1987 (in Japanese).
3. M. Ohta, A. Ikuta. Stochastic signal information processing – Bayes' type digital filter using higher order mutual correlation and its application to state estimation. Journal of Acoustic Society of Japan (specific explanation), vol. 50, N°2, 140–148, 1994 (in Japanese)