THE JOURNAL OF NAVIGATION (2004), **57**, 407–415. © The Royal Institute of Navigation
DOI: 10.1017/S0373463304002875 Printed in the United Kingdom

An Efficient Neural Network Model for De-noising of MEMS-Based Inertial Data

Ahmed El-Rabbany and Mohammed El-Diasty

(*Ryerson University*, *Toronto*)
(Email: rabbany@ryerson.ca)

Micro-Electro-Mechanical System (MEMS)-based inertial technology has recently evolved. It holds remarkable potential as the future technology for various navigation related applications. This is mainly due to the significant reduction in size, cost, and weight of MEMS sensors. A major drawback of low-cost MEMS-based inertial sensors, however, is that their output signals are contaminated by high-level noise. Unless the high frequency noise component is suppressed, optimizing the pre-filtering methodology cannot be achieved. This paper proposes a neural network-based de-noising model for MEMS-based inertial data. A modular, three-layer feedforward neural network trained using the back-propagation algorithm is used for this purpose. Simulated and real MEMS-based inertial data sets are used to validate the model. It is shown that the model is capable of reducing the noise of the Crossbow's AHRS300CA IMU data by over one order of magnitude without altering the stochastic nature of the original signal. This is of utmost importance in developing a generic stochastic model for MEMS-based inertial data. A comparison between the developed neural network model and the wavelet de-noising method is made to further validate the model. It is shown that achieving the same level of noise suppression with wavelet-based de-noising model changes the stochastic characteristics of original signal.

KEY WORDS

1. Neural network. 2. MEMS. 3. INS.

1. INTRODUCTION. There exist some circumstances where part of the Global Positioning System (GPS) signal may be obstructed to the extent that a GPS receiver may not "see" enough satellites for positioning. An example of those situations is positioning and navigation in urban canyons. In addition, GPS receivers are known to have limited dynamic capabilities. That is, high receiver dynamics can cause temporary signal losses, or cycle slips. Nevertheless, these limitations were successfully overcome by integrating GPS with a relatively environment independent system, the inertial navigation system (INS). Unfortunately however, the size, cost and weight of traditional inertial sensors have limited the use of integrated GPS/INS systems to certain applications (El-Rabbany, 2002). More recently, an emerging new technology, Micro-Electro-Mechanical System (MEMS)-based inertial technology, has been developed which holds remarkable potential as the future technology for various navigation related applications. This is mainly

https://doi.org/10.1017/S0373463304002875 Published online by Cambridge University Press

due to the significant reduction in size, cost, and weight that MEMS sensors hold. Unfortunately however, output signals of low cost, MEMS-based inertial sensors are characterized by their high noise level. Suppressing the high frequency noise component is essential for optimizing the pre-filtering methodology.

This paper proposes a neural network model for de-noising of MEMS-based inertial data. The motivation behind the selection of neural networks lies in their ability to learn from the signal (useful) component of the raw output data only. They cannot learn from the noise component. As such, the neural network solution would mainly contain the signal component; thus suppressing the noise. A modular, three-layer feedforward neural network trained using the back-propagation algorithm is used for this purpose. The design parameters of this model are the number of neural network inputs, outputs, and hidden nodes, as well as the structure of the network itself. Simulated data as well as real data collected at two different rates (1 and 10 Hz) from the Crossbow's AHRS300CA IMU are used to validate the model. It is shown that the model is capable of reducing the noise by over one order of magnitude without altering the stochastic characteristics of the original signal. This is of utmost importance in developing a generic stochastic model for MEMS-based inertial data. A comparison between the developed neural network model and the wavelet de-noising method is also presented.

2. EVOLUTION OF INERTIAL TECHNOLOGY.

Inertial navigation technology, which employed traditional accelerometers and gyroscopes, was used for the first time in World War II (Lawrence, 1998). Inertial technology has since evolved from the complex gimballed system where sensors are mounted in a set of gimbals, to strapdown systems where sensors are mounted directly on the vehicle. A gimballed system uses mechanical gyroscopes with spinning wheels, while a strapdown system uses more advanced inertial sensor technology such as optical gyroscopes. Unfortunately, although the new generation of inertial sensors have many advantages over old technology, their size, cost and weight have limited the use of inertial systems to certain applications.

In recent years, MEMS technology has evolved, which holds a number of unique characteristics. MEMS systems are miniature devices, which comprise one or more of micromachined components or subsystems (Maluf, 2000). In addition to their small size, MEMS systems are inexpensive, light, and consume very low power compared with conventional sensors. As a result, a wide range of applications are expected to benefit from MEMS technology, including the automotive industry, medical instrumentation, telecommunications, environmental monitoring, and others. MEMS systems are commonly fabricated using silicon, which possesses significant electrical and mechanical advantages over other materials. A critical stage in MEMS fabrication is the packaging process, which constitutes about 75% to 95% of the total cost of a MEMS system (Maluf, 2000). So far, as a result of diversity of MEMS applications, no standards exist for MEMS packaging. Large packaging will destroy the small size advantage that MEMS technology possesses (Maluf, 2000).

The focus of MEMS-based inertial sensors' research was on accelerometers and gyroscopes (Tung, 2000). According to Maluf, 2000, micro machined accelerometers were first demonstrated in 1979 at Stanford University. This was followed by a steady increase in unit-volume market, which was accompanied by a decrease

https://doi.org/10.1017/S0373463304002875 Published online by Cambridge University Press

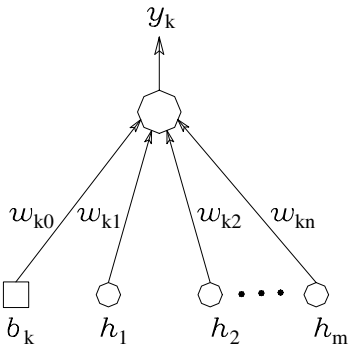


Figure 1. Simple neuron model.

in the unit cost. There exist different designs for silicon accelerometers, including simple pendulum (Lawrence, 1998). Micromachined gyroscopes, on the other hand, operate based on the principle of the Coriolis effect – that is, when an object moves on the surface of a rotating body, it undergoes a Coriolis force. Although not yet matured, MEMS accelerometers and gyroscopes have exceeded expectations in their performance. One of the major challenges that has yet to be overcome is the high noise of the output of MEMS inertial sensors compared with that of traditional technology.

3. NEURAL NETWORK MODELLING. Artificial neural networks (ANN) are computational models that imitate the human brain in performing a particular task. They are capable of solving complex problems through learning, or training, and then generalizing the network outputs for other inputs. A neural network consists of processing elements, or neurons, h_i, that are massively interconnected. Each of the connecting links is characterized by its own weight, w_{ki} (Figure 1). An activation function, e.g., a sigmoid function, is applied to limit the amplitude of the neuron, while an external bias, b_k, is applied to increase or lower the net input of the activation function (Haykin, 1999). The network is trained to find the optimal values for the weights and biases.

Artificial neural networks can be designed in various ways, depending on how the neurons are structured and the learning algorithms, or rules, used. Network architectures may be classified as single-layer feedforward, multi-layer feedforward, and recurrent networks (Haykin, 1999). Initially, we selected the traditional non-modular feedforward structure for our noise suppression model. Unfortunately this structure failed to learn the required mapping. Consequently, a modular three-layer feedforward neural network trained using the back-propagation algorithm was selected (Figure 2). A modular neural network has the capability of dividing a problem into sub-problems and resolving each sub-problem rather well. It consists of a group of back-propagation networks, sometimes referred to as "local experts", each has the same architecture. This group of networks competes to learn the various aspects of the problem, which is then controlled by a "gating network". The back-propagation algorithm uses the gradient descent in weight space to minimize the output error. It converges to a locally optimal solution when an adequate number of input-output pairs of samples are subjected to the network for an adequate number of epochs. The back-propagation algorithm can be summarized in three steps as follows

https://doi.org/10.1017/S0373463304002875 Published online by Cambridge University Press

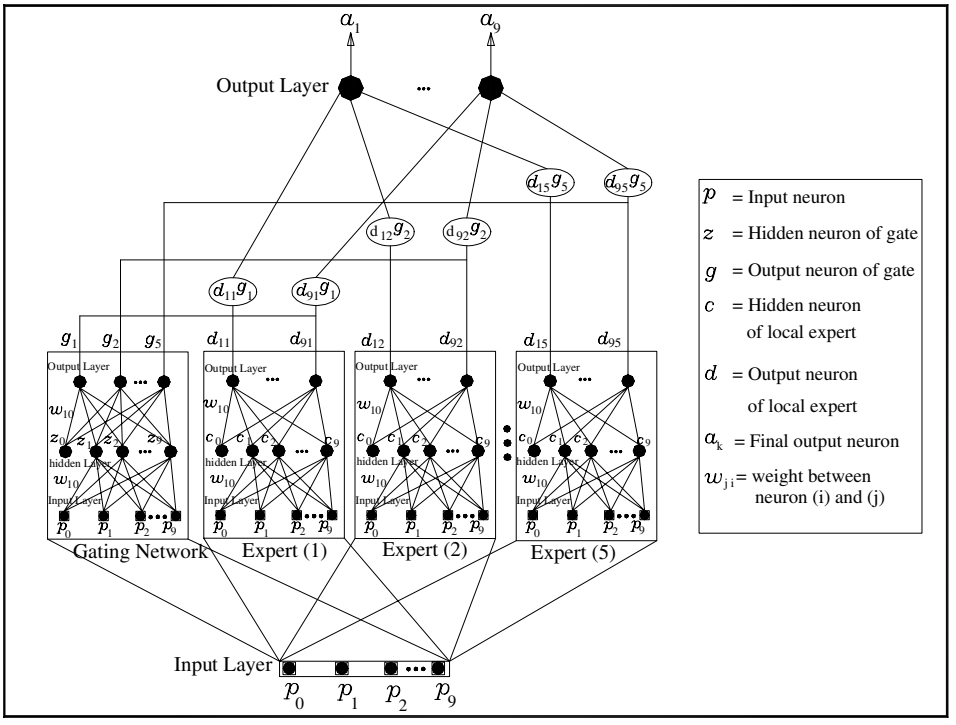


Figure 2. Modular, three-layer feedforward neural network with the structure [9-9-5-9].

(see El-Rabbany and El-Diasty, 2002 for more details). The first step is to propagate the input forward through the network. The neurons in the first layer receive external inputs:

$$\boldsymbol{a}^0 = \boldsymbol{p} \tag{1}$$

$$\boldsymbol{a}^{m+1} = \boldsymbol{f}^{m+1}(\boldsymbol{W}^{m+1}\boldsymbol{a}^m + \boldsymbol{b}^{m+1}), \quad \text{for } m = 0, 1, \ldots, M-1 \tag{2}$$

where M is the number of layers in the network. Equation (1) provides the starting point for equation (2), and the output of the layer m is the input to the layer $m+1$. The outputs of the neurons in the last layer are considered as the network outputs:

$$\boldsymbol{a} = \boldsymbol{a}^M \tag{3}$$

The next step is to calculate the output error and *back-propagate* it to calculate the sensitivities ($\boldsymbol{s}$ values).

$$\boldsymbol{s}^M = -2\dot{\boldsymbol{F}}^M(\boldsymbol{n}^M)(\boldsymbol{t} - \boldsymbol{a}) \tag{4}$$

$$\boldsymbol{s}^m = \dot{\boldsymbol{F}}^m(\boldsymbol{n}^m)(\boldsymbol{W}^{m+1})\boldsymbol{s}^{m+1}, \quad \text{for } m = M-1, \ldots, 2, 1 \tag{5}$$

Finally, the weights and biases are updated, based on the calculated sensitivities, using the approximate steepest descent rule:

$$\boldsymbol{W}^m(k+1) = \boldsymbol{W}^m(k) - \alpha \boldsymbol{s}^m(\boldsymbol{a}^{m-1})^T \tag{6}$$

https://doi.org/10.1017/S0373463304002875 Published online by Cambridge University Press

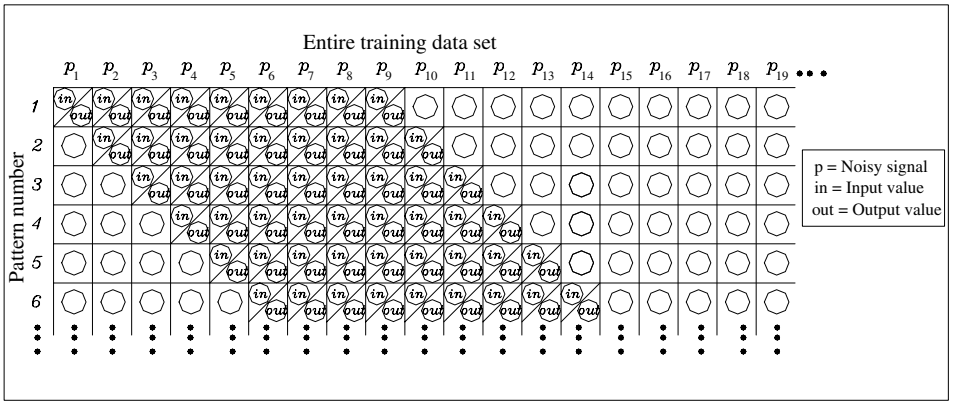


Figure 3. Training patterns for the proposed neural network model.

$$\boldsymbol{b}^m(k+1) = \boldsymbol{b}^m(k) - \alpha \boldsymbol{s}^m \tag{7}$$

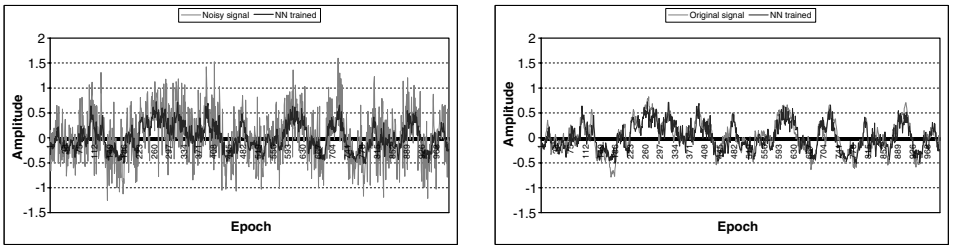
where α is the step size. The three steps are repeated until the error reaches a minimum; hence, the calculated weights of the connections would represent the network solution.

Our neural network de-noising model attempts to train the network to reproduce the input signal. As such, it uses a sliding window with the same number of input and output processing elements (Figure 3). The sliding window is suggested to be three times the duration of most noise (NeuralWare, 2001). In our case, a window containing 9 elements was selected. Of the 9 output processing elements, only the centre element of each window is considered as the filtered record. A critical stage to this approach was the stopping criterion for training the network. In our case, the training was stopped based on examining the performance of the neural network when the root-mean-square (RMS) settles at a minimal value.

4. WAVELET ANALYSIS. A wavelet is a waveform of finite interval and zero mean (Mathworks, 2002). Wavelet analysis is a relatively new way of modelling and processing signals, which have traditionally been done by Fourier analysis. While Fourier analysis breaks up a signal into sine and cosine functions, wavelet analysis breaks up a signal into translated (i.e., shifted) and scaled versions of the original wavelet. Translating a wavelet means shifting it forward (or backward) in time. Scaling a wavelet, on the other hand, means stretching (or compressing) it to obtain low and high frequency wavelets. Smaller scale factors correspond to more compressed (or high frequency) wavelets and vice versa. There are some advantages of wavelet analysis over Fourier analysis, including the ability of the former to analyse non-stationary signals and signals with more localized features (Boggess and Narcowich, 2001; Mathworks, 2002).

A wavelet-based de-noising is accomplished by first decomposing the noisy signal to obtain the wavelet coefficients, both the approximations and details. The approximations constitute the low-frequency constituents of the signal, while the details constitute the high-frequency constituents. It should be pointed out that a suitable decomposition level must be used, which would depend on the signal characteristics

https://doi.org/10.1017/S0373463304002875 Published online by Cambridge University Press



(a) (b)

Figure 4. Simulated first-order Gauss-Markov data versus neural network de-noised data. (a) with noise; (b) without noise.

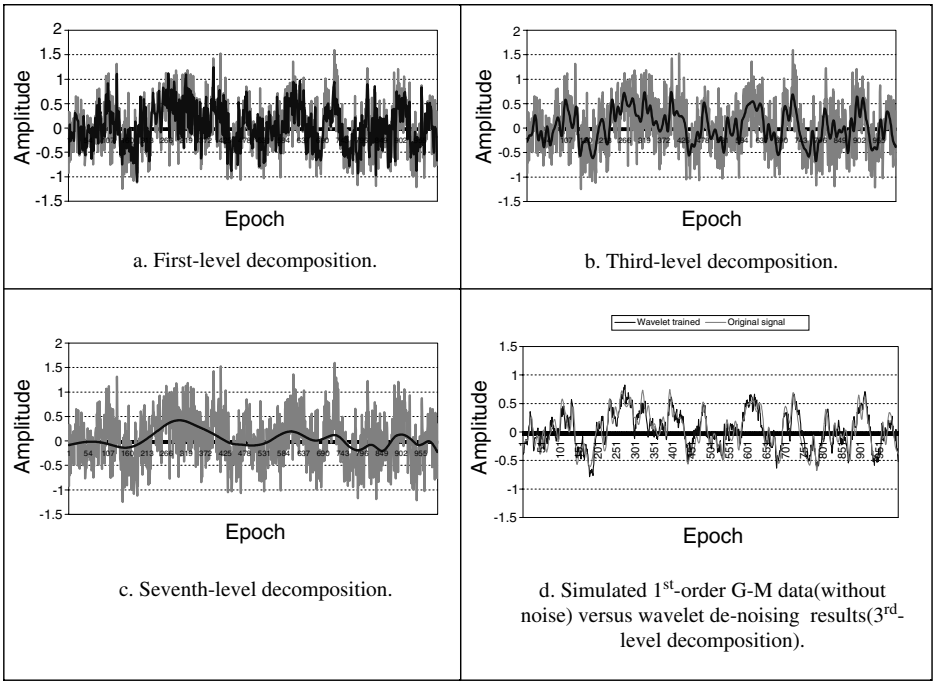
(Mathworks, 2002). Once the wavelet coefficients are obtained, the unwanted coefficients (i.e., details in our case) are removed or modified. The last step is to re-construct the signal using the approximations coefficients and, possibly, the modified details coefficients.

5. RESULTS AND DISCUSSION. As mentioned above, simulated data as well as real data collected at two different rates (1 and 10 Hz) from the Crossbow's AHRS300CA IMU were used to assess the performance of the proposed de-noising model. Two sets of simulated data were generated, namely a random walk and a first-order Gauss-Markov process, both were contaminated by white sequence noise.

5.1. *Results for simulated data.* A modular neural network with the structure [9-9-5-9] – 9 input neurons, 9 hidden neurons, 5 local experts, and 9 outputs – was trained using the back-propagation algorithm (Figures 2 and 3). Each of the data series was divided into three subsets: training, testing and validation subsets. The training was stopped when the value of the normalized RMS became 0·225. Figure 4 shows the results of the neural network model when the simulated data was used. It can be seen that the model is capable of recovering the original signal (i.e., without noise added). Another interesting outcome is that the stochastic nature of the original signal remained unchanged after applying the model. This is of utmost importance in developing a generic stochastic model for the MEMS-based inertial data. To further validate the developed neural network model, a wavelet-based de-noising model, with various decomposition levels, was used for de-noising the same simulated data (Figure 5). Inspecting the autocorrelation functions of the de-noised data shows that, similar to the neural network model, the wavelet-based de-noising model with third level decomposition is capable of reducing the noise without altering the stochastic nature of the original data. However, examining the variance of the de-noised data shows that the noise level is 48% higher when the wavelet de-noising model is used. To reach the same level of noise reduction, the fifth level of decomposition must be applied, which, unfortunately, alters the stochastic nature of the original data as a result of over-smoothing. Another drawback of the wavelet de-noising model is its dependency on the selection of the threshold value.

5.2. *Results for real data.* Figure 6 shows the results of the neural network model when the Crossbow's AHRS300CA IMU data was used (only the accelerometer data

https://doi.org/10.1017/S0373463304002875 Published online by Cambridge University Press



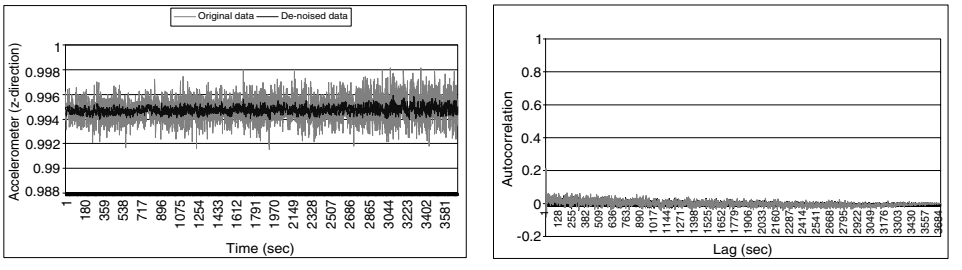
a. First-level decomposition.

b. Third-level decomposition.

c. Seventh-level decomposition.

d. Simulated 1st-order G-M data(without noise) versus wavelet de-noising results(3rd-level decomposition).

Figure 5. Simulated first-order Gauss-Markov data (with noise) versus wavelet de-noising with various decomposition levels.



(a) (b)

Figure 6. (a) Raw accelerometer measurements in z-direction vs. neural network de-noising results; (b) Corresponding autocorrelation functions.

in the z-direction is shown). The model was capable of reducing the noise by a factor of 12·5 without altering the stochastic nature of the original data. Figure 7 and 8 show the results of wavelet de-noising model. Similar to the results of the simulated data, the neural network model was superior to the wavelet de-noising model. Inspecting the variance of the de-noised inertial data shows that the noise level is 62% higher when the wavelet de-noising model is used. As discussed above, further noise reduction, for example through the use of higher decomposition level, would alter the stochastic characteristics of the original data.

https://doi.org/10.1017/S0373463304002875 Published online by Cambridge University Press

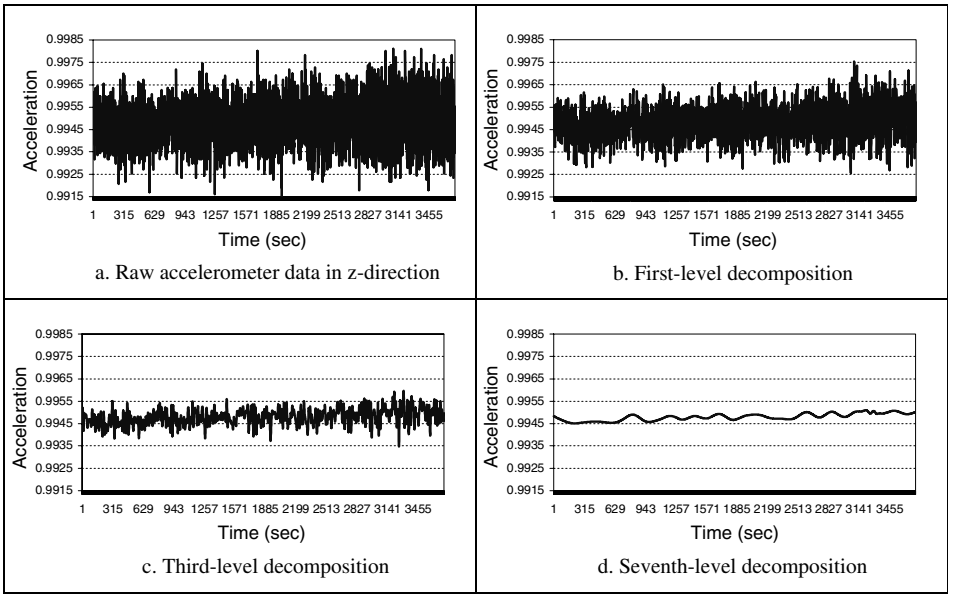


Figure 7. Results for wavelet de-noising method.

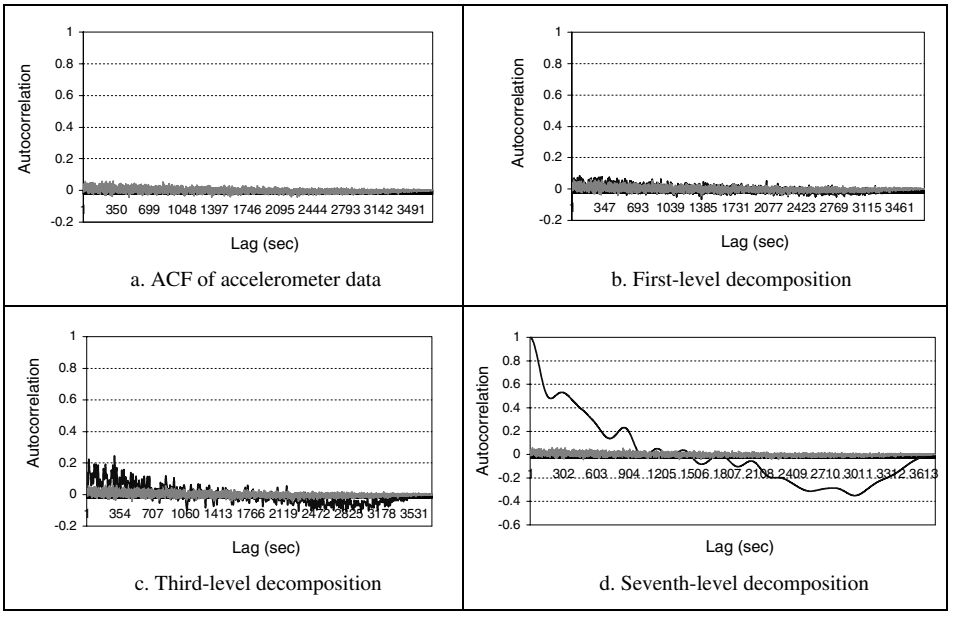


Figure 8. Autocorrelation functions of raw accelerometer measurements in z-direction and wavelet-based de-noised data.

6. CONCLUSIONS. This paper developed a modular neural network-based model for noise suppression. It has been shown that the model can effectively be used to suppress the noise of MEMS-based inertial data as the network learns from the signal component only. Results of simulated and real MEMS-based inertial

https://doi.org/10.1017/S0373463304002875 Published online by Cambridge University Press

data sets showed that the neural network de-noising model is capable of suppressing the noise of the inertial data by more than one order of magnitude without altering the stochastic characteristics of the original signal. In contrast, as a result of over-smoothing the data, achieving the same level of noise suppression with wavelet-based de-noising model changes the characteristics of original signal.

ACKNOWLEDGEMENTS

This research was supported in part by the Geomatics for Informed Decision (GEOIDE) NCE. The Crossbow's AHRS300CA IMU data was provided to us by Dr. Naser El-Sheimy of the University of Calgary.

REFERENCES

Boggess, A. and F. J. Narcowich (2001). *A First Course in Wavelets with Fourier Analysis*. Prentice Hall, USA.

El-Rabbany, A. (2002). *Introduction to GPS: The Global Positioning System*. Artech House Publishers, Boston, USA.

El-Rabbany, A. and M. El-Diasty (2003). A New Approach to Tidal Prediction. This *Journal* **56**, 305–314.

Haykin, S. (1999). *Neural Networks: A Comprehensive Foundation*. Second Edition. Prentice Hall, USA.

Lawrence, A. (1998). *Modern Inertial Technology*. Second Edition. Springer-Verlag New York, Inc.

Maluf, N. (2000). *An Introduction to Microelectromechanical Systems Engineering*. Artech House Publishers, Boston, USA.

Mathworks, Inc. (2002). Wavelet Toolbox User's Guide. Version 2, Natick, MA, USA.

NeuralWare (2001). NeuralWorks Professional II/Plus Reference Guide, PA, USA.

Tung, S. (2000). Position Paper: An Overview of MEMS Inertial Sensors. Department of Mechanical Engineering, University of Arkansas, Fayetteville, Arkansas 72701, USA.

https://doi.org/10.1017/S0373463304002875 Published online by Cambridge University Press