Proceedings 30th International Intelligent Motion Conference, Power Systems World '96.

Advanced Sensors *

Francis L. Merat

Dept. of Electrical Engineering and Applied Physics
Case Western Reserve University
Cleveland, Ohio 44106-7221
Phone: (216)368-4572; FAX: (216)368-2668
E-mail: flm@po.cwru.edu

ABSTRACT

The integration of electronics with sensors is resulting in a revolution in sensors. A new non-contact optical torque sensor which utilizes neural network signal processing is described as an example. This sensor serves as a model for advanced sensors arrays which combine integrated sensors and signal processing for fault tolerance, high noise immunity, and non-linear signal processing. Micromechanical sensors are one example of sensors which can be produced on silicon with the associated signal processing electronics. Already this approach has produced accelerometers, position sensors, pressure gages, and magnetic sensors which may be useful for advanced diagnostic applications.

1. INTRODUCTION

Advances in electronics and data processing have made possible the integration of sensors and pattern recognition systems on a single integrated circuit. One can measure a single quantity, such as temperature, using a single sensor; however, complex measurements such as machine diagnostics may require many sensors measuring a variety of different modalities to correctly predict machine performance, or a multiplicity of sensors measuring the same quantity to provide improved accuracy, linearity, or fault tolerance. In these situations, the output patterns from the sensor arrays must be utilized to estimate the physical quantity of interest. Since the outputs from a sensor array can be viewed as multi-dimensional patterns, this becomes similar to many pattern recognition problems.

Generalized integrated circuit sensing structures have made rapid advancements in recent years due to advancements in integrated circuit technology, i.e., surface and bulk micromachining, and high aspect ratio lithography. In particular, these techniques can be used to selectively etch silicon producing mechanical structures such as beams and diaphragms which can be used as the basis for

* Parts of this work have been supported by the Reliance Electric Division of Rockwell, Inc.; the Cleveland Advanced Manufacturing program through the CWRU Center for Automation and Intelligent Systems Research; the

a wide range of sensors such as pressure, acceleration, and temperature [Wise,1985]. The mechanical structures which form the basis of these sensors are called micro-electromechanical systems (MEMS) and is the basis of considerable development in the sensor community [Bryzek, 1996]. Other techniques have been used to integrate optical, chemical, and biological sensors on silicon substrates which can detect smells or gas composition.

However, such sensors and arrays of sensors are of limited utility without adequate signal processing. Typically, we tend to think of sensor electronics as linearizing or calibrating a sensor. However, we are often interested in fault tolerance as well as combining outputs from a variety of sensors for purposes such as fault tolerance and pattern recognition for such applications as machine diagnostics. In particular, artificial neural networks (ANNs) can be used to handle inaccurate or statistical sensor output patterns and can learn the characteristics of unknown/unmodeled systems. Only training samples (i.e., inputs and the corresponding outputs) are needed to characterize a given system. One does not have to have an analytical system model. In conventional approaches dealing with statistical sensor output patterns, there must be an appropriate mathematical model of the system. Many times, it is not possible to develop an accurate model for the system, or, it is not possible to mathematically solve the model. In this case, conventional systems will either perform very poorly or simply be impossible to implement. It is not difficult to see that ANNs outperform the classical approach using Kalman filters [Park, 1991; Pao, 1991; Lo, 1994].

2. SENSORS

In this section we describe a variety of silicon and optical sensors which may be suitable for machine diagnostics. Their signal processing will be described in the next section.

a. Optical Torque Sensor

Various non-crystalline transparent materials such as certain polymeric plastics are optically isotropic under normal conditions but become birefringent when stress is applied. When polarized light is passed through a stressed (under load) birefringent material an optical phase shift which is proportional to the local strain is produced in the light. The resulting changes in the polarized light reflected from the sensor may be analyzed by a polarization filter, which converts the two-dimensional optical phase function to a two-dimensional intensity function (fringe pattern) such as shown in Figure 1. These optical fringe patterns are a function of the stress on the plastic and are suitable for stress analysis but need human interpretation for proper analysis. However, a neural

U.S .Army under Contract # DAAH04-95-10097, and ARPA under Contract # F49620-94-C-0007.

network may be used to analyze these patterns to produce a shaft torque sensor as will be described below.

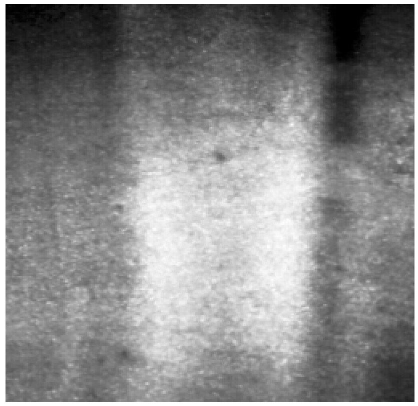

Figure 1. Typical Intensity Pattern produced by Optical Torque Sensor

An experimental optical torque sensor is shown in Figure 2. A cylinder of polycarbonate plastic (birefringent) was epoxied to aluminum mounts which were then attached to a motor shaft (see Figure 2) using collets. The plastic cylinder was illuminated by polarized light which passes through the plastic, was reflected from a coating of reflective epoxy on the inner surface of the

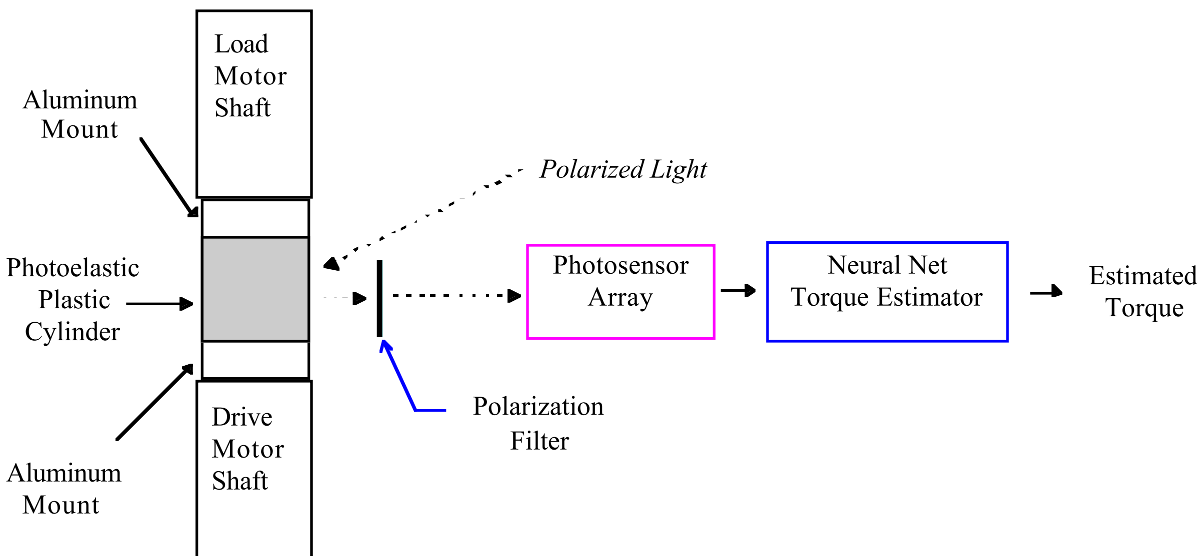

Figure 2. Experimental Torque Sensor

plastic cylinder, and then passed through the plastic a second time. The light traveling through the plastic cylinder incurs an optical phase shift which is dependent upon the photo-induced strain (a function of the shaft torque) at that point as it travels through the sensor's plastic cylinder. This optical phase shift is a two-dimensional function of position on the sensor surface and loading (torque) on the motor shaft.

As shown in Figure 2 a neural net was used to process the sensor data. The motivation for this was that neural nets are excellent functional estimators and should be able to learn the functional mapping from intensity patterns to shaft torque. Shaft intensity patterns such as shown in Figure 1 can be recorded using a conventional CCD camera. However, a conventional camera image has far too many pixels to directly input pixel information into a neural net and would be far too expensive for a low-cost torque sensor. In this experimental sensor, image preprocessing was used to reduce the number of inputs to the neural net and simulate a photodiode array as might be found in a commercial version of this torque sensor. Figure 3 shows this horizontal strip of 32 virtual sensor cells.

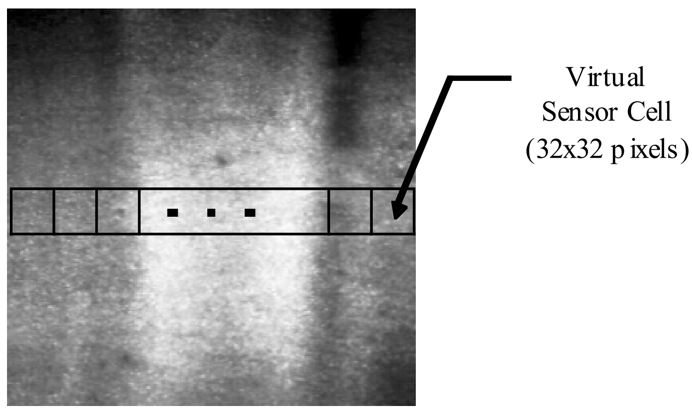

Figure 3. Pre-processing of 256x256x8 image

b. MEMS SENSORS

Silicon is an excellent mechanical material which can also support integrated electronic fabrication. The important mechanical property of silicon for most sensor applications is its flexibility - it is very elastic below the breaking point. Aside from its mechanical advantages silicon based sensors also benefit from batch fabrication techniques and generous support from the integrated circuit industry. These latter advantages result in a well established industry which can produce low-cost sensors.

i. silicon micromachined accelerometers

Silicon based accelerometers have been produced by bulk micromachining of silicon to produce a large proof mass supported by a thin flexure area (a cantilever arm) as shown in Figure 4. A complete survey of the field may be found in Yun [Yun, 1991]. Sensors of the type shown in Figure 4 are produced by bulk micromachining which etches along crystal planes in the silicon to delineate the features. Other, typically much smaller sensors, are produced by surface micromachining in which sacrificial oxide layers are deposited on top of a silicon crystal, followed by additional layer(s) of polysilicon. The oxide is then selectively etched away to produce thin, free-standing silicon sensing elements. Such devices are typically limited to several microns thickness and cannot produce large masses (such as shown in Figure 4) which are needed for low-g sensing applications. The accelerometer in Figure 4 works by measuring the flex of the seismic mass through the cantilever arm. Figure 4 shows a piezoresistive sensor placed at the cantilever arm to measure the acceleration although many other techniques such as hall effect, capacitive and resonant beam techniques may be used for electronic readout [Bicking,1993]. The nominal range of commercial silicon accelerometers is ±2g to ±100 g although higher and lower sensitivities can be achieved [Kubler, 1995].

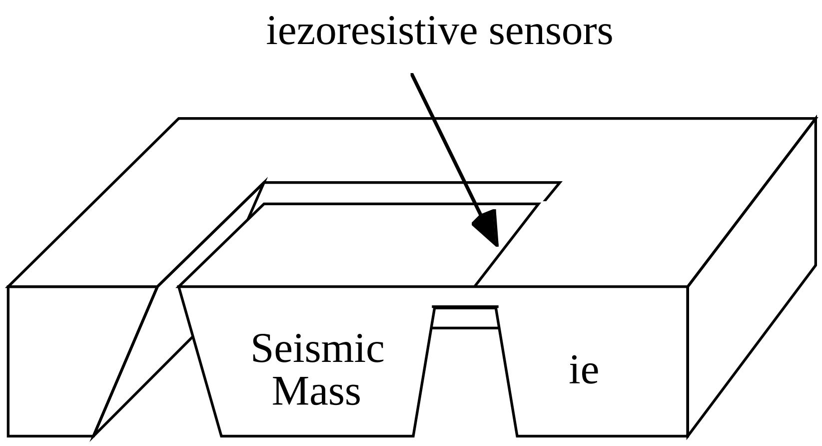

Figure 4. Cross-section of silicon piezoresistive sensor

ii. silicon micromachined pressure sensors

The same micromachining techniques that can be used to produce accelerometers can also be used to produce silicon pressure transducers. Figure 5 shows a silicon pressure sensor where a thin diaphragm has been etched into a silicon substrate. The etched silicon chip is then bonded (sealed) to a second wafer which then forms a sealed compartment containing a standard reference pressure. The readout mechanism is capacitive sensing of the displacement of the diaphragm as a function of external pressure variations. Because of the capacitive sensing mechanism these sensors typically need relatively sophisticated on-chip electronics to detect the small signals.

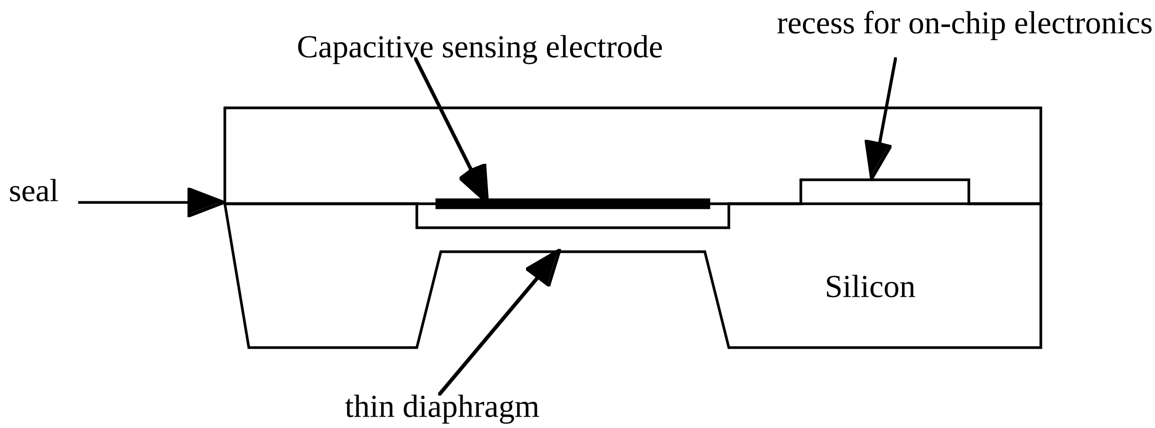

Figure 5. Silicon capacitive pressure sensor

iii. magnetic sensors

Magnetic sensors are particularly important for motor diagnostics. The Hall effect is the most commonly utilized transduction scheme for magnetic field sensing. Such devices detect the Hall potential resulting from the Lorentz force acting on moving carriers and are sensitive to the magnetic field in a particular direction. Micromachining can be used to produce two-dimensional structures which can be used for 2-D sensing of in-plane magnetic fields [Paranjape, 1994; Paranjape, 1995]. However, Hall effect sensors are suitable for high magnetic field strength applications (i.e., 1 mT). Fluxgate sensors based upon the non-linear B-H curve of ferromagnetic materials (i.e., nickel-iron) are also compatible with silicon processing and can provide two orders of magnitude better resolution. Fluxgate sensors use ferromagnetic cores surrounded by several coils. An excitation coil(s) drives the sensor core into saturation periodically. Pick-up coils then measure the resulting changes in magnetic flux via the induced voltage. Gottfried-Gottfried describes a microfabricated fluxgate sensor in which the coils are fabricated by interconnecting two metal layers and the sensor is provided by a 0.46 µm thick nickel-iron layer [Gottfried-Gottfried, 1995]. The signal processing for this sensor is quite complex requiring detection of signals at even harmonics of the signal yet was able to be put onto a single CMOS ASIC [Gottfried-Gottfried, 1995]

iv. “smart” sensors

The future of smart sensors is to incorporate micromechanical sensing elements on the same substrate as the signal conditioning and processing electronics. This electronics will facilitate communications between sensors in arrays or complex machinery, will provide analog or digital signals as needed, and will use signal processing to correct for non-ideal sensor characteristics such as temperature drift and non-linearities. Such smart electronics will provide on-chip calibration, eliminating the need for laser trimming and similar expensive calibration techniques.

3. Sensor Signal Processing

On-chip signal processing is essential for many applications providing stability, accuracy and ruggedness. Commercial accelerometers for automotive air-bag deployment are among the best examples of sensor/electronic integration. Similar requirements are also found in industrial pressure sensors [Moore, 1995; Matthews, 1996].

a. Neural networks

Neural nets can compute any computable function [Funahashi, 1989; Hornik, 1989]. Anything that can be represented as a mapping between vector spaces can be approximated to arbitrary precision. In practice, neural nets are useful for mapping problems which are tolerant of errors, have example data available, but to which hard and fast rules can not easily be applied. They are particularly useful for the non-linear mappings such are found in sensor calibration and compensation.

Such a neural network was used to interpret the intensity patterns produced by the optical torque sensor described in Section 2a. For the optical torque sensor the neural net was used to learn the mapping between the optical fringe pattern as seen by the virtual sensor array and the shaft torque, as measured by a conventional strain gauge torque sensor. There was no known analytical relationship between the given input (optical intensity pattern) and the measured output (shaft torque) for the sensor making this sensor array ideal for neural net signal processing.

Before training a neural network, the connection strengths (weights) between nodes are assigned to small random numbers. Training using the standard backpropagation technique then proceeds by presenting the training pairs to the network and adjusting the weights until a satisfactory level of performance is reached [Rumelhart, 1986].

As an example of typical neural network processing a backpropagation network with 32 nodes in the input layer, 12 nodes in a single hidden layer, and 1 output node was used for torque estimation for the sensor described in section 2a. Backpropagation computes the partial derivatives of error with respect to the neural weights. With these partial derivatives, it is possible to do gradient descent in weight space. If small steps are taken in the direction of the gradient, the error is guaranteed to reach a local minimum. This local minimum has been empirically accepted as a good enough solution for most purposes, although it is a very slow, time consuming process.

Although neural networks are useful in prototype experiments they have some important drawbacks for mass-produced sensors: (1) the time required to train the neural network, and (2) accommodation of manufacturing variations in the sensors, The first drawback is significant for all neural network applications. As an example, the neural networks used for the optical torque sensor typically took several hours to train on a UNIX workstation. Such times are not acceptable for mass production of sensors, even if the device has the on-board electronics to perform such calibration. The second drawback is each and every manufactured sensor must be trained and, even if the electronics is on the chip, will dramatically slow the manufacturing process.

There are solutions which can be used to address each of these points. For a practical sensor it is desirable to train the neural net in the shortest possible time. This requires taking the largest possible steps in the direction of the gradient without overshooting the minimum error solution. A set of partial derivatives collected at a single point does not have enough information for deciding step size. If the higher-order derivatives (the curvature of the error function) are available, it is possible to choose better, or larger, step sizes which will result in faster learning. One such algorithm called the quickprop algorithm [Fahlman, 1988] assumes each weight in the neural net has a quadratic error curve. Each weight in the network is assumed to affect the error independently of the others. The quadratic calculation is then approximated using a difference of gradients between the current and the previous epoch. Special provisions must be made for starting the algorithm, what to do when last weight changes are zero, and what to do when the current derivative is greater than the previous derivative. The quickprop technique works well in practice, and considerably shortens the training time compared to plain backpropagation techniques.

Other techniques such as random vector enhancement have dramatically reduced the training time for these networks and have increased the generalization (interpolation) ability of neural networks by changing the neural network architecture. These techniques eliminate the hidden layer by random vector enhancement of the input, i.e. computing additional functions of the inputs [Chung, 1995; Pao, 1994] .

For the optical torque sensor described in 2a, a random vector enhanced phasor neural network [Chung, 1995] was trained to estimate the applied torque from the supplied fringe pattern. The training pairs thus consisted of 32 element sensor array intensity vectors corresponding to an experimentally obtained fringe pattern from the birefringent sensor, and a corresponding torque value as measured by the strain gauge shaft torque sensor. Typical performance of this neural network processing is shown in Figures 6 and 7 for a statically loaded shaft. Figure 6 shows the training of the neural network for forty-one torque values (70 to 150 pound-inches in 2 pound-inch increments). The neural net output (represented by x's in Figure 6) is superimposed on the linear training curve (represented by a solid line which can barely be seen) and the difference (error) shown by the thick solid line as a percentage of the actual torque. The average training error was 0.16%.

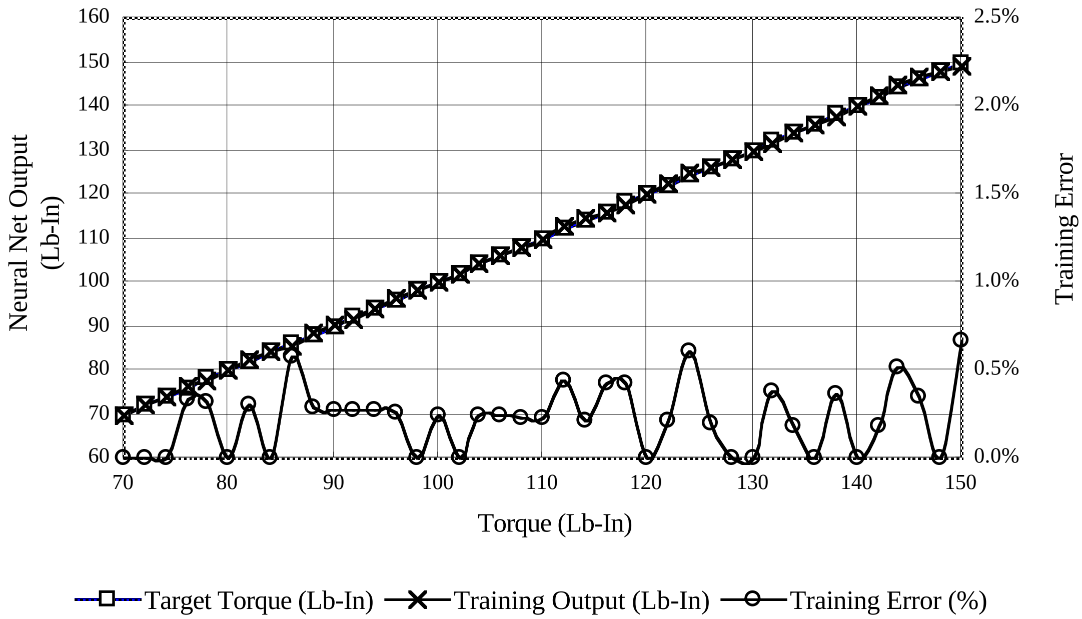

Figure 6. Performance of Neural Net Torque Estimator (Training)

The performance of a single hidden layer neural network torque estimator is shown in Figure 7. The average sensor error was less than 0.4% for stationary shafts, and increased as the shaft rotational speed increased, up to 4% for shaft speeds less than 20 rpm.

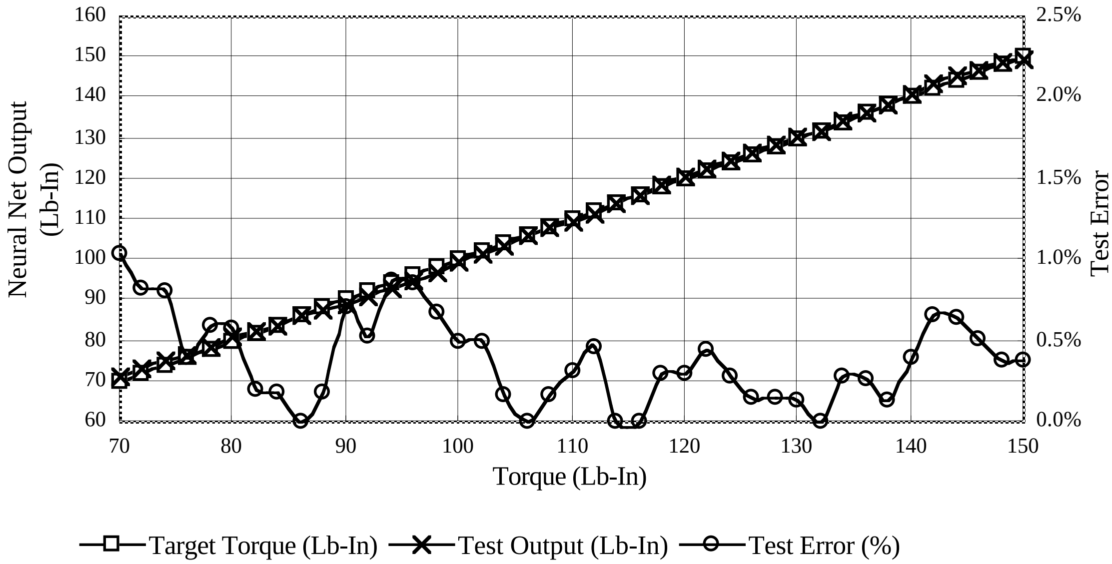

Figure 7. Performance of Neural Net Torque Estimator (Test)

An autoassociative memory using neural networks can be used for sensor failure detection and correction and can also correct for sensor variations eliminating many calibration requirements. Conventional systems for failure analysis and detection rely upon complex models of physical systems and have traditionally been utilized in high-cost applications such as jet aircraft and spacecraft [Chow,1984]. However, as neural networks can be used to represent systems through training for which mathematical models can not be formulated they can be applied to sensor systems and can be used to predict sensor outputs. The autoassociative memory learns a 1:1 mapping for a noise-free set of sensor outputs. Once this mapping is learned, differences between measured sensor outputs and sensor outputs estimated by the autoassociative memory can be used to identify faulty sensors. Furthermore, the autoassociative memory can also be used to substitute correct sensor output values for faulty values. In practice, this autoassociative memory is placed between the sensor outputs and before any additional sensor signal processing to function as a noise-eliminating, fault correcting filter [Chung,1995].

4. Discussion and Conclusion

The integration of electronics and sensors is well underway. However, the advent of MEMS and neural network signal processing is leading to new classes of “smarter” sensors which can correct for sensor noise, drift, failure and manufacturing tolerances. A neural network autoassociative memory uses redundant information from a sensor array to provided corrected sensor outputs. A second neural network combines the sensor information to estimate the required system parameters. This neural network approach avoids the use of complex analytical models and exploits the information redundancy of a sensor array.

In order to detect sensor failures, outputs from the sensors are initially processed by the autoassociative memory and sensor residuals are calculated. Conventional residual generation is typically based on analytical knowledge of the system. However, an autoassociative memory can represent the system through learning and can be used to describe non-linear systems. The residuals can be processed using methods like thresholding or statistical decision theory to identify a particular sensor failure.

We have tested this approach on a one dimensional (i.e., linear) optical sensor array. An autoassociative memory was implemented for a 32 element optical sensor array using real data for functional estimation. With no sensor noise the system was able to estimate torque (the sensed parameter) to an accuracy of about 1%. Computer simulations were done adding independent Gaussian noise to the sensor outputs. Without using autoassociative memory signal processing, the estimation error dropped to 17% for a sensor S/N of 0.75dB. With the autoassociative memory the estimation error was only about 4% under the same conditions. The neural net based birefringent torque sensor has shown its ability to accurately measure shaft torque values over limited torque ranges. The neural net torque estimator has exhibited less than 0.4% average estimation error over a 70 - 150 lb-inch static torque range. The excellent torque measurement accuracy achieved in these experiments shows the potential of accurately measuring torques by means of non-contacting optical sensor arrays. The images used in these experiments are identical to those from a linear array of optical sensors which would detect an average optical intensity corresponding to a limited spatial region. This could be equivalently performed by an array of photodetectors with a considerable reduction in system complexity.

This neural network approach has several advantages which are important to sensor arrays such as might be implemented using MEMS technology: (1) it dramatically reduces the effects of individual sensor noise; (2) it accommodates sensor-to-sensor variation in arrays by treating the variation as noise; (3) it should be capable of multi-sensor fusion.

We have simulated a MEMS array of thermal sensors and are in the process of collecting data from an actual array of MEMS sensors for comparative evaluation. We are also fabricating a specialized silicon chip which will incorporate thermal heaters and sensors. Our simulations show that neural network signal processing can be used to isolate faulty sensors. Simulations and experiments in noise, sensor failure and sensor tolerance for MEMS sensor arrays are in progress.

The actual development of such MEMS based systems is underway in many laboratories; however, the sensor signal processing and especially the reliability and accuracy of the sensor array is very important to critical applications such as diagnostic monitoring and machine maintenance. Such applications seem to be especially well suited to neural network signal processing of the type described. The principle difference is that envisioned MEMS systems are often two-dimensional. To test the applicability of neural network sensor processing and control in MEMS systems we are developing a test chip which will consist of a two dimensional array of semiconductor heaters and thermal sensors on a silicon wafer. The thermal sensors will be interleaved between the resistive heaters as shown in Figure 6. The long-term goal of this work is to use the heater/sensor array with a 2-D feedback control algorithm to precisely generate two-dimensional thermal profiles. While this chip is being designed and fabricated we have developed simple computer models to test our signal processing and control algorithms.

REFERENCES

Bicking, R.E., "*Automotive Accelerometers for vehicle ride and comfort applications,*" in Automotive Sensory Systems, edited by Christopher O. Nwagboso, Chapman & Hall, London (1993), pp. 125-140.

Bryzek, J., "The MEMS Industry in Silicon Valley and Its Impact on Sensor Technology," *Sensors*, vol. 13, no. 7, July (1996).

Chow, E. and A. S. Willsky, "Analytical Redundancy and the Design of Robust Failure Detection Systems", in *IEEE Transactions on Automatic Control,* vol. AC-29, no. 7, pp. 603-614, July 1984.

Chung, D. and F. Merat , "Associative Memory Using Random Vector Enhanced Phasor Neural Network, Ohio Aerospace Institute Neural Network Symposium and Workshop 95, August 1995.

Chung,D. and F. Merat , "Sensor Failure Detection/Correction Using Autoassociative Memory," Ohio Aerospace Institute Neural Network Symposium and Workshop 95, August 1995.

Fahlman, S. E., "Faster Learning Variations on Back Propagation: An Empirical Study," in *Proceedings of the 1988 Connectionist Models Summer School,* San Mateo, CA, Morgan Kaufmann Publishers, 1988.

Funahashi, K., "On the Approximation Realization of Continuous Mappings by Neural Networks," in *Neural Networks*, vol. 2, pp. 183-192, 1989.

Hornik, K., M. Stinchombe, and H. White, "Multilayer Feedforward Networks are Universal Approximators," in *Neural Networks*, vol. 2, pp. 359-369, 1989.

Gottfried-Gottfried, R., W. Budde, R. Jahne, B. Sauer, S. Ulbricht and U. Wende, "*A Miniaturized Magnetic Field Sensor System Consisting of a Planar Fluxgate Sensor and a CMOS Readout Circuitry,*" in Transducers '95 - Eurosensors IX, Stockholm, Sweden (1995), pp. 229-232.

Kubler, J.M. , "Introduction to Variable Capacitance Silicon Microsensors," *Sensors*, vol. 12, no. 8, August (1995).

Lo, J. T., "Synthetic Approach to Optimal Filtering," *IEEE Transactions on Neural Networks*, vol. 5, no. 5, September (1994).

Moore, J., “Smart Pressure Sensors for Industrial Applications,” *Sensors*, vol. 12, no. 6, June (1996).

Matthews, C.A., “Intelligent Pressure Measurement in Multiple Sensor Arrays,” *Sensors*, vol. 13, no. 3, March (1996).

Pao, Y. H., G. H. Park, and D. J. Sobajic, “System Identification and Noise Cancellation: A Quantitative Comparative Study of Kalman Filtering and Neural Net Approaches,” in *Proceedings of the American Control Conference, Boston, MA*, vol. 2, pp. 1408-1411 (1991).

Pao, Y. H., G. H. Park and D. J. Sobajic, “Learning and Generalization Characteristics of the Random Vector Functional-Link Net,”, in *Neuro Computing*, vol. 6, pp. 163-180 (1994).

Paranjape, M, L. Ristic, “*A Micromachined Vertical Hall Effect Magnetic Field Sensor in Standard CMOS Technology,*” Applied Physics Letters, Vol. 60, No. 25 (1992)

Paranjape, M, L. Landsberger, and M. Kahrizi, “*A 2-D Vertical Hall Magnetic Field Sensor Using Active Carrier Confinement and Micromachining Techniques*,” in Transducers ‘95 - Eurosensors IX, Stockholm, Sweden (1995), pp. 253-256.

Park, G. H., “*System Identification and Noise Cancellation using Neural Networks*,” Master Thesis, Dept. of Electrical Engineering and Applied Physics, Case Western Reserve University (1991).

Rumelhart, D. E., Hinton, G. E., and Williams, R. J., “Learning Internal Representations by Error Propagation,” in Rumelhart, D. E. and McClelland, J. L. (editor), *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*, Chapter 8, MIT Press (1986).

Wise, K.D., “*Silicon Micromachining and its Application to High Performance Integrated Sensors,*” in Micromachining and Micropackaging of Transducers, edited by C.D. Fung, P.W. Cheung, W.H. Ko, and D.G. Fleming, Elsevier Science Publishers B.V., Amsterdam (1985), pp. 3-18.

Yun, K.D., and R. Howe “*Silicon Microfabricated accelerometers: a perspective on recent developments,*” in Proceedings of Sensors Expo 1991 (1991), pp. 1-8.