Application of Adaptive and Neural Network Computational Techniques to Traffic Volume and Classification Monitoring

W. C. MEAD, H. N. FISHER, R. D. JONES, K. R. BISSET, AND L. A. LEE

A traffic volume and classification monitoring (TVCM) system based on adaptive and neural network computational techniques is being developed. The value of neural networks in this application lies in their ability to learn from data and to form a mapping of arbitrary topology. The piezoelectric strip and magnetic loop sensors typically used for TVCM provide signals that are complicated and variable and that correspond in indirect ways with the desired FHWA 13-class classification system. Furthermore, the wide variety of vehicle configurations adds to the complexity of the classification task. The goal is to provide a TVCM system featuring high accuracy, adaptability to wide sensor and environmental variations, and continuous fault detection. The authors have instrumented an experimental TVCM site, developed personal computer-based on-line data acquisition software, collected a large data base of vehicles' signals together with accurate ground truth determination, and analyzed the data off-line with a neural net classification system that can distinguish between class 2 (automobiles) and class 3 (utility vehicles) vehicles with better than 90 percent accuracy. The neural network used, called the connectionist hyperprism classification network, features simple basis functions; rapid, linear training algorithms for basis function amplitudes and widths; and basis function elimination that enhances network speed and accuracy. Work is in progress to extend the system to other classes, to quantify the system's adaptability, and to develop automatic fault detection techniques.

The FHWA 13-class classification scheme (*1*) divides vehicles largely according to application or axle configuration. Standard traffic volume and classification monitoring (TVCM) practice typically combines one or more piezoelectric strip sensors with one or more magnetic loop sensors in a road-embedded sensor group that provides signals to a commercial electronics package. Although these systems appear to be quite simple, they are in reality quite complicated and possess performance characteristics that can significantly degrade the reliability of the vehicle-monitoring information provided. For example, piezoelectric strip sensors vary greatly in output pulse characteristics from one sensor to another and also from one event to another, even for interactions with similar vehicles, making the process of "simply" counting axles an error-prone task. Furthermore, the sensor installations and electronics typically drift with changes in environmental conditions and with installation aging. These characteristics can lead to unacceptable classification inaccuracies.

Nonlinear adaptive network computing has progressed greatly in the past decade and has demonstrated capabilities and opened new applications to high-speed digital computers. Many traditional computer applications are preprogrammed; that is, algorithms are specifically designed to implement a known numerical solution for an application. Adaptive algorithms, however, offer somewhat greater generality: the adaptive algorithm "learns" by adjusting modeling coefficients to optimize a fit to the available data or to maximize some performance criteria. Artificial neural networks combine many simple, individual processing units interconnected to perform prediction, control, and classification tasks via machine learning. The "neurons" or "nodes" are usually simple nonlinear transfer functions. The training consists of adjusting weights (basis function parameters or interconnection strengths) to best match a training set or to minimize an energy function. Artificial neural networks and adaptive cellular automata show interesting and useful behaviors. Capabilities already demonstrated by existing adaptive computing systems include machine learning (*2–4*), self-organization (*2–4*) bidirectional associative memories (*3,4*), feature detection and pattern recognition and classification (*2–4*), signal processing and noise reduction (*3,4*), processing of speech, handwriting, and natural language (*2–4*), modeling of multidimensional nonlinear and chaotic functions (*5–9*), prediction of physical dynamical processes (*7–9*), and providing new solutions to control (*7–11*) and classification (*12*) problems.

Our goal in the project described here is to harness the capabilities of adaptive and neural network computational techniques to the TVCM application to obtain high classification accuracy, adaptability to a fairly wide range of sensor and environmental conditions, and automatic detection of faults when the adaptive range is exceeded. In addition to these beneficial performance objectives, neural networks have certain other advantages for various applications, including TVCM. They learn inductively from data and can be quite versatile and robust. Using on-line learning, neural networks can predict, control, or classify in drifting systems. Their implementation in software permits low-cost development, whereas implementation in special-purpose large-scale integration (LSI) hardware provides low-cost replication with high-speed performance.

CLASSIFICATION SCHEME

A slight extension of the FHWA 13-class system (Table 1) is used here. The major class boundaries agree with the FHWA scheme; subclasses have been added that are expected to be distinguishable, for example, to separate vehicles that are towing trailers from those

Applied Theoretical Physics Division, Los Alamos National Laboratory, P.O. Box 1663, Los Alamos, N.M. 87545.

TABLE 1 FHWA 13-Class Vehicle Classification Scheme Extended To Specify Distinguishable Subclasses

Class #	Subclass	Configuration	Sub-Configuration
1		motorcycle	
	a		w/o trailer
	e		w/ trailer
2		passenger car	
	a		subcompact
	b		compact
	c		full-sized
	d		jumbo
	e		w/ trailer
3		2 axle, 4 tire single unit,utility	
	a		small
	b		medium
	c		large
	d		jumbo
	e		w/ trailer
4		bus	
	a		2 axle,short wheelbase
	b		2 axle,long wheelbase
	c		3+axle
	e		w/ trailer
5		2 axle, 4-6 tire large single unit	
	a	4-tire	w/o trailer
	b	6-tire	w/o trailer
	e	4- or 6-tire	w/ trailer
6		3 axle single unit	
	a		w/o trailer
	e		w/ trailer
7		4-5+ axle single unit	
	a		4 axle
	b		5+ axle
8		3-4 axle single trailer	
	a		3 axle
	b		4 axle
9		5 axle single trailer	
	a		long-tongue trailer
	b		standard semi
10		6-7+ axle single trailer	
	a		6 axle
	b		7+ axle
11		5 axle multiple trailer	
12		6 axle multiple trailer	
13		7+ axle multiple trailer	
14		other	

that are not. In some cases the class is subdivided according to vehicle size. Work to date and the present paper deal exclusively with classes 2a to 2d and 3a to 3d (cars without trailers and utility vehicles without trailers) but neglect subclass information. These classes were initially focused on for two reasons: (a) they cover about 98 percent of the vehicles at our first sensor test site (STS1), and (b) this is a fairly subtle class boundary, which serves well to test the adaptive/neural network approach.

EXPERIMENTAL SENSOR TEST SITE

STS1 was designed to provide a conveniently accessible experimental site with good traffic flow. It is located on the State Route 4 Truck Route about 15 min from Los Alamos National Laboratory. The sensor layout, shown in Figure 1, was designed to offer redundant measurements to permit internal cross-validation and multiple sensor subgroupings that can simulate several different monitoring

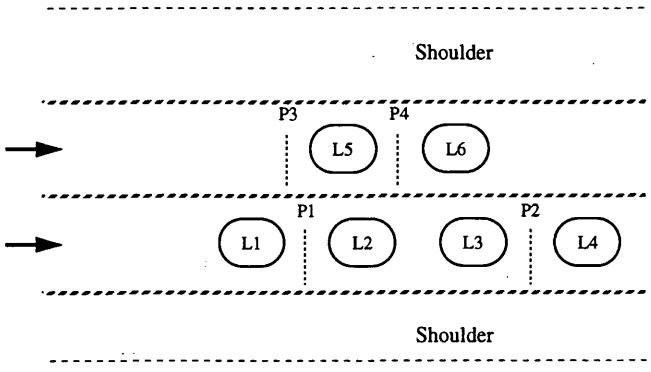

FIGURE 1 Schematic layout of STS1.

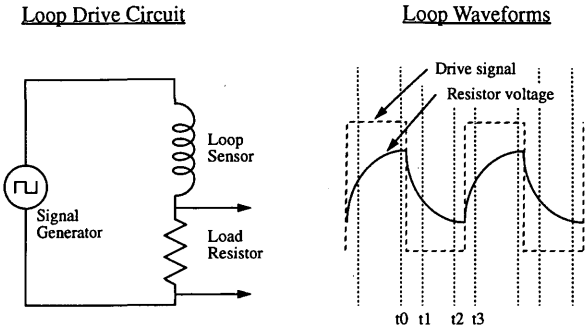

FIGURE 2 Loop readout scheme that is simple, direct, and fast and that uses versatile hardware.

configurations. Both lanes of traffic are monitored with on-track sensors only. The data presented here were obtained at STS1 over the period from November 1992 through June 1993.

DATA ACQUISITION SYSTEM

A key design criterion for the data acquisition system was hardware flexibility combined with off-the-shelf availability. Thus, a personal computer (PC)-based acquisition system was chosen and a 16-channel analog-to-digital conversion (ADC) board for input and a 20-channel counter-timer-board for output were used. This hardware is expected to be readily adaptable to any current TVCM or weigh-in-motion (WIM) sensors and to most future sensor types.

A loop readout scheme was chosen that, although unconventional, is simple, direct, and fast. The scheme is illustrated in Figure 2. A square-wave drive signal is applied to a resistive-inductive (R-L) circuit containing the loop. The voltage across the load resistor is read at four times per square-wave cycle: twice near the maximum current and twice when the current is in the exponential decay following the square wave's voltage transitions. The difference between the peak and decaying readings is related simply to the inductance of the loop, which in turn varies according to the characteristics of a proximate vehicle.

A schematic of the loop data acquisition system is shown in Figure 3. The 5 to 10-kHz square waves for driving up to six loop sensors (about 1 ohm of impedance) are generated by the counter-timer board with sequentially delayed phases. The timer outputs are individually cleaned up and amplified and are then applied to the R-L circuits containing the six respective loop sensors. Most data have been acquired using high-quality audio amplifiers as loop drivers, although a custom-designed seven-channel instrumentation amplifier has been used as well. The loop circuits were typically driven with a 0.4-V *p-p* square wave and the loop current signal amplified by a gain-of-10 amplifier on the ADC board.

The data acquisition system for the piezoelectric strips is illustrated in Figure 4. Since the piezoelectric strips are active sensors, no drive signal is needed, and the acquisition system is simpler. A

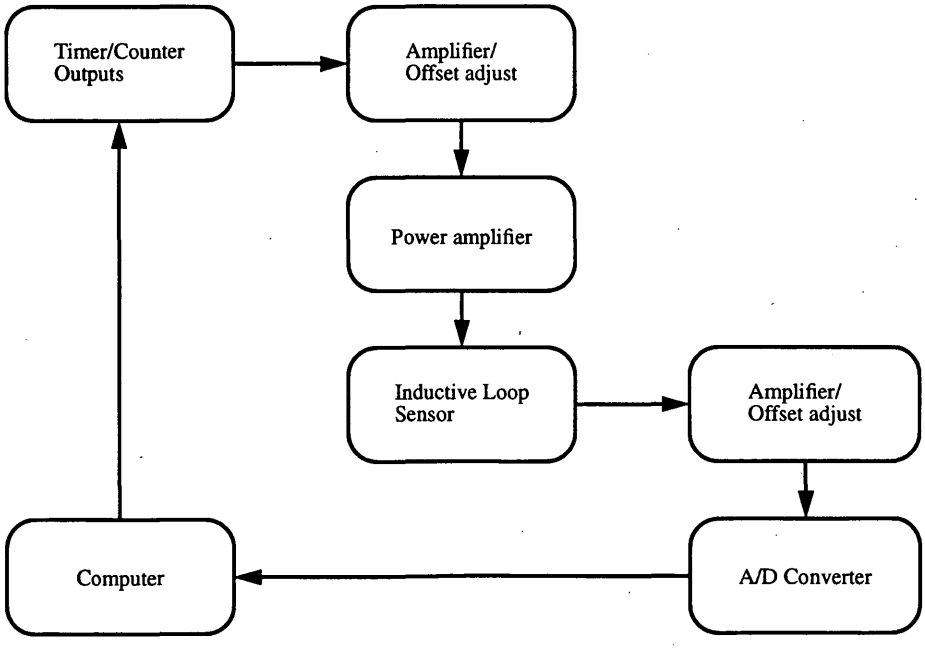

FIGURE 3 Schematic of inductive loop sensor data acquisition system.

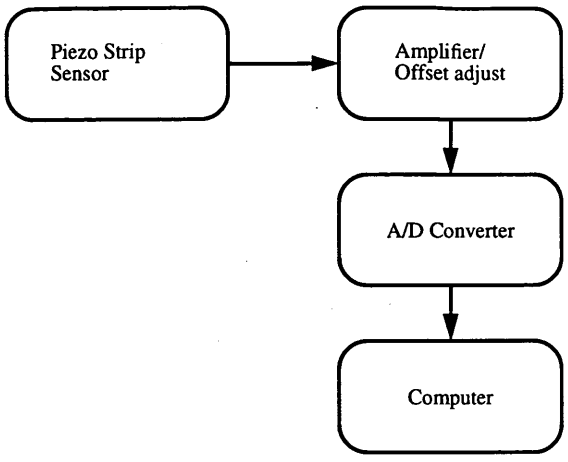

FIGURE 4 Schematic of piezoelectric strip sensor data acquisition system.

seven-channel instrumentation amplifier is generally used with high input impedance at voltage gains of 1 to 10 to match the piezoelectric outputs to the ADC input requirements.

Data from up to six loops and up to six piezoelectric strips are acquired cyclically. The total data rate is about 50 to 100 kilosamples/sec, providing ample time for resolution of individual sensors' outputs. Amplitude resolution is about 1 part in 4,096.

Phase I data were collected in one of two modes: first a "spooling" mode was used: the data stream was directly recorded on 90-MB high-speed, removable disks for later analysis. Later the capability of collecting event mode data was developed and employed: PC-based, on-line analysis reduced the incoming data stream to a list of sensor-activation events, thus obviating the need to record large amounts of quiescent data.

GROUND TRUTH DETERMINATION

A two-channel video system, illustrated in Figure 5, was used to acquire the data needed to determine ground truth. The two cameras viewed the test site from widely differing angles on opposite sides of the roadway. The primary camera view was perpendicular to the road at an angle of about 40 degrees above the horizontal. Data were recorded on videotapes, together with a signal light that indicated the timing of the PC sensor-data acquisition runs.

Analysis of the video data was largely done by playing the tape into a workstation-based image digitizer, using a software image comparator to select only frames taken during a run and with a vehicle present. Digitized frames were classified by a human analyst. Use of the workstation to preselect the relevant images increased the classification rate from about 1/10 real time to about 1/3 real time. Generally, one camera's digitized image was adequate for classification and lane determination. This process was expensive, but not prohibitively so. The cost of ground truth determination was mitigated by reusing the data for multiple sensor groupings and by repeatedly analyzing the sensor signals as adaptive algorithms were developed.

Formal quality control procedures were observed to evaluate and maintain the accuracy of ground truth determination. At least 10 percent (randomly selected) of the data runs were reclassified and the accuracy of the ground truth determination was found to be better than 99 percent in vehicle volume, major class determination, and lane determination (to the nearest lane). This accuracy is more than adequate for training and testing the sensor-based classification system. The ground truth data base from STS1 currently contains 2,216 vehicles, mostly of classes 2a to 2d and 3a to 3d.

SIGNAL PREPROCESSING

Given the complex character of the sensor signals, signal preprocessing plays a crucial role in preparing the data for input to the

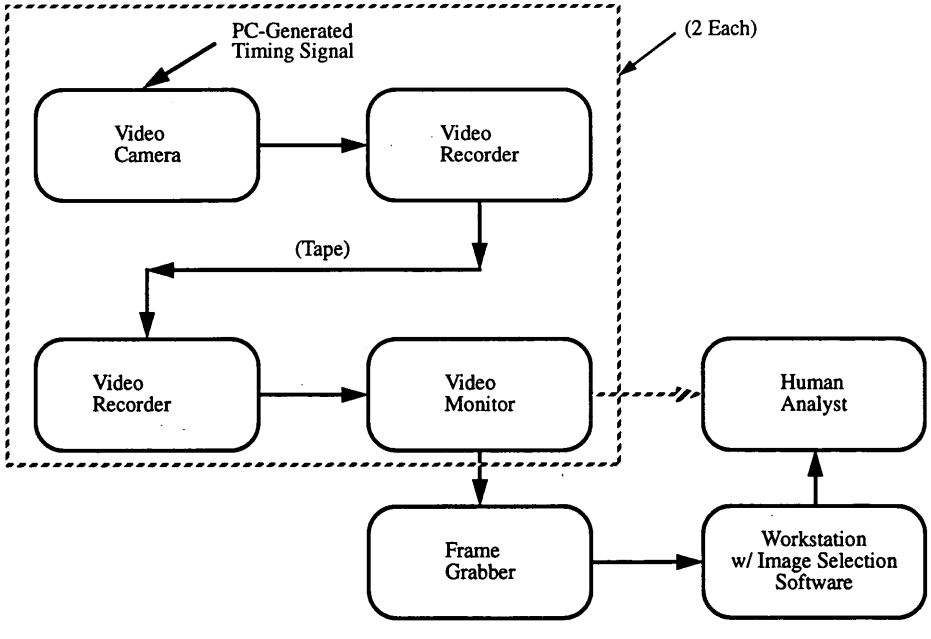

FIGURE 5 Schematic of video data acquisition system.

neural networks. Signal preprocessing steps are discussed in two subgroups here: adaptive signal conditioning steps and data representation steps.

Adaptive Signal Conditioning and Pulse Extraction

The importance of adaptive signal conditioning can be seen by considering the properties of loop and piezoelectric sensor signals. The most difficult aspects of loop signal processing are poor signal-to-noise ratio, significant systematic errors (e.g., asymmetric results of measurements on the up- and down-transition signals), and high ratio of drift to signal. The piezoelectric strips, on the other hand, give processing challenges from three different characteristics: large dynamic range (~1,000:1), large sensor-to-sensor variations (~10:1), and large variations in response due to particular interaction details (~10:1).

By using a sequence of signal conditioning steps, reliable detection and data reduction of the sensor signals are achieved: (a) even-odd correction removes sampling asymmetries of the loop signals; (b) low-pass filtering improves the signal-to-noise ratio, particularly for loops; finally, adaptive adjustments to signal (c) amplitude and (d) noise levels compensate for sensor-to-sensor and event-to-event signal variations. To remain within PC processing capabilities, adaptive (but non-neural-network) signal processing techniques are used in these initial signal-conditioning steps.

The extraction of active signal pulses is performed by an adaptive triggering algorithm. The algorithm distinguishes between baseline (quiescent) and active signal behavior by using a pulse height spectrum for each data channel. The trigger's sensitivity is varied adaptively, using moving averages that account for recent average noise and signal levels. The trigger also incorporates a slope-sensitive term and logic that helps to prevent multiple triggering during a single piezoelectric event.

Data Representation

One additional group of tasks must be accomplished before presenting the data to the neural net for solution: choosing a representation for the data. The representation determines what information the neural network must process. If too much extraneous information is presented to the network, the neural net can be overwhelmed. This is analogous to a signal-to-noise problem. On the other hand, if the representation chosen does not include the information to be processed, the neural net can be underinformed, that is, the network can be reduced to the status of fortune teller.

For the TVCM application representation requires signal reduction, screening, parsing, and subselecting the signal data. The signals are reduced by extracting simple signal statistics from the detailed signal profiles, for example, peak amplitude, full widths at half- and quarter-maximum, time of peak amplitude, and integral between the half-maximum points. The screening steps apply known physical constraints to remove extraneous pulses. Constraints currently used include minimum and maximum sensor pulse widths and amplitudes, implied vehicle speed greater than 0, and implied axle separation distances greater than 0.6 m (2 ft). One constraint, although not absolutely physically defensible, appears to be helpful, namely, restricting the dynamic range of signals corresponding to one vehicle to a factor of 8:1. Parsing associates subgroupings of the signals into vehicle events. The vehicle parser is built on a few observations that appear to be quite accurate, at least within the authors' current experience: that one vehicle usually generates a single magnetic loop pulse, that an implied gap of more than 24.4 m (80 ft) between axles always signifies a separate vehicle, and that piezoelectric signals usually belong to the vehicle with the nearest (in-time) magnetic loop signal. The signal parser is the most complex nonadaptive part of the classification system. Only time will tell whether it is fully adequate to all traffic and site conditions.

Data subselection is the final task, and that is discussed separately in the next paragraph since subselection issues are expected to be different for different vehicle classes. Also the data subselection issues are closely interwoven with the definitions of the classification scheme and the overall architecture chosen for the classification system (Figure 6). The basic idea is to divide the classification parameter space according to the number of axles detected for a vehicle and then to choose among the classes having constant numbers of axles by using neural networks trained to the task. There are not enough data on the universe of vehicle and installation types to determine whether the initial axle count determination can be performed accurately enough to support this classification system architecture or not. At present the architecture is serving well.

Returning to the data subselection issue, specifically for classes 2a to 2d and 3a to 3d, by a combination of observational, deductive, and statistical analyses, it was determined that the data containing most of the information for distinguishing the boundary between these two classes is the peak amplitude of the magnetic loop data. The reason for this is believed to be that most class 2 vehicles have lower ground clearances than most class 3 vehicles. Therefore, most class 2 vehicles yield larger changes in loop inductance and greater peak loop signal amplitude. This analysis indicates that the practical limit to the accuracy of the separation of these two classes is about 90 percent on a vehicle-by-vehicle basis. That is, given typical measurement errors, traffic behavior, and vehicle configurations, about 10 percent of class 2 vehicles will appear to be class 3 vehicles and about 10 percent of class 3 vehicles will appear to be class 2 vehicles. A large part of the crossover of class 2 vehicles into class 3 appears to be due to off-track events, which reduce the change in loop inductance because of the lateral offset of the vehicle over the loop. A large part of the crossover from class 2 into class 3 appears to be caused by the fact that some class 3 (utility) vehicles are actually built on car chassis and thus do not have larger ground clearances than most cars. To some extent these two

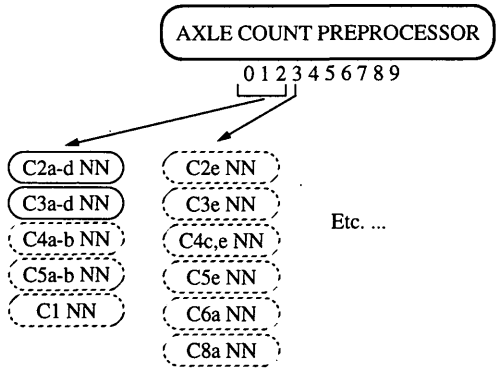

FIGURE 6 Overview of classification system architecture.

crossover effects largely cancel, and it is possible, in an average sense, to end up with greater than 90 percent overall accuracy for an ensemble of vehicles.

One other important issue of adaptation arose in connection with using the loop peak amplitude to distinguish the class 2–3-class boundary. It was found that for unknown reasons the absolute scaling of the peak loop amplitude shifts from one data acquisition run to another; this shift has not been traced to a reparable cause. However, the problem has apparently been solved by observing that there exists a relatively stable vehicle population that gives a fixed maximum signal, and this provides an on-line calibration factor.

Having resolved the crucial issues of data conditioning and representation, the data are ready to be delivered to a classification neural net to obtain the completed classification solution. Note that by "completed," it is meant fully functioning for classes 2a to 2d and 3a to 3d rather than complete in the sense of having solved the entire classification problem.

CONNECTIONIST HYPERPRISM CLASSIFICATION NEURAL NETWORK

The classification network used in this work, the connectionist hyperprism classification (CHC) network, is designed to recognize multidimensional patterns presented by the various vehicle signatures produced by the TVCM sensors and to produce corresponding classification outputs. The CHC network has architectural features (Figure 7) that are well matched to the needs of the intended TVCM applications.

The CHC network operates on roughly the same principle as clustering algorithms (*13*), whereas it draws most of its numerical approach from typical neural network methods. Two data sets are imagined, one for training and another for testing, that each contains a number of anonymous samples (labeled 0) plus some number of tagged samples (labeled 1) that are representative of a single class (e.g., several signal sets that correspond to the same vehicle class). Each sample vector (p) consists of an N-dimensional set of inputs x_{pi}, together with the desired output, o_p, equal in this case to 0. or 1. It is assumed that there are M class 1 members of the training set, and initially a network was chosen that contains M nodes or basis functions, each centered at one of the unique training datum points.

The CHC network uses N-dimensional hyperprism basis functions (Figure 8). Each basis function has a vector center, $\boldsymbol{x}_c$, in the input parameter space, and produces nonzero output in a connected region about its center, having width $\boldsymbol{b}_a$ above the center and $\boldsymbol{b}_b$ below it. At the center the basis function's output value is equal to the weight f_c. When, for dimension i, the datum point coordinates satisfy $x_{cji} - b_{bji} < x_{pi} < x_{cji} + b_{aji}$, then the ith component of the jth basis function is

$$f_{ji}(x) = f_{cji} \cdot \{1. - [(1. - m_e) \cdot |x_{pi} - x_{cji}|/b_{wji}]\} \tag{1}$$

where b_{wji} is the appropriate width parameter, either b_{aji} or b_{bji}. Within the active domain the basis function's value is the product of these N components. Outside its active domain the basis function is zero. The constant m_e is the same for all basis functions in the network, and in practice we often use m_e equal to 1., so the basis function components are simply N-dimensional top-hat functions. The network output is

$$g(\boldsymbol{x}) = \Sigma f_j(\boldsymbol{x}) \tag{2}$$

The network training algorithm contains two main parts: one part adjusts the basis function central amplitudes f_c based on only the class 1 data, whereas the other adjusts the widths $\boldsymbol{b}_a$ and $\boldsymbol{b}_b$ based on all of the training data. The f_c's are adjusted to minimize the root mean square error in the network's calculation of the class 1 datum points. The widths are adjusted according to a self-organizing algorithm (*14*), adjusting the domain of each basis function to regulate the number of class 0 datum points that fall within the active region and resetting the width when basis function overlap occurs. The training algorithms are iterative and on-line in the sense that the basis function amplitudes and widths need not be static in time but can grow or shrink to reflect the currently appropriate training conditions.

The training algorithm for the jth node's central amplitude f_{cj} weights each datum point by a manually adjustable parameter, w_1, according to its target output value o_p:

$$w_p = (w_1 \cdot o_p) + [(1. - w_1) \cdot (1. - o_p)] \tag{3}$$

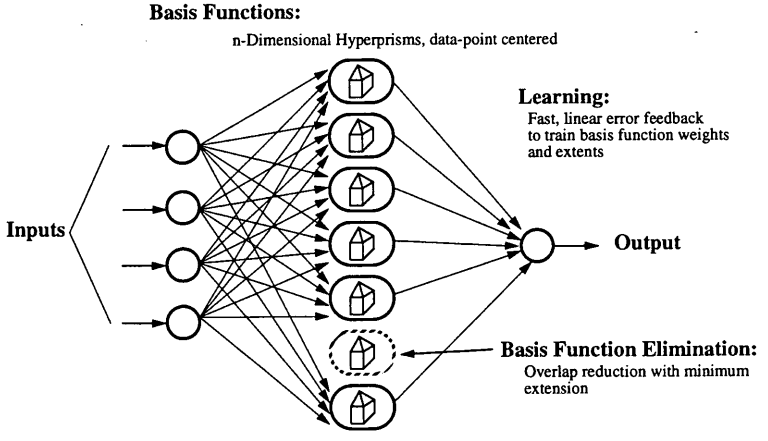

FIGURE 7 Architecture of CHC neural network.

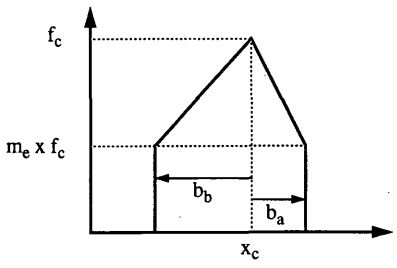

FIGURE 8 Basis function of CHC neural network.

The iterative adjustment to the central amplitudes is then given by

$$f_{cj}^{n+1} = f_{cj}^{n}\{1 + [\alpha\, w_p\, (g(\boldsymbol{x}_p) - o_p)]\} \tag{4}$$

where α is a learning rate parameter and n is the iteration count. For most of this work α equal to 0.1 and w_1 equal to 0.7 were used.

The training algorithm that adjusts the basis function widths is slightly more elaborate. It operates in two successive stages—one that adjusts widths on the basis of datum point inclusion and another that revises the widths to remove basis function overlap.

The datum point-based width adjustment increases the appropriate widths as class 1 points are encountered within the basis function's active domain and decreases the widths as included class 0 points are found. The choice of what widths to adjust is made by calculating the distance (d_{pji}) of each point included within the active domain to each dimension's nearest basis function edge and choosing the smallest. The data-dependent width adjustment is

$$b_{wji}^{n+1} = (b_{wji}^{n}) \cdot \{1. + [\beta\, (b_{wji}^{n}/b_{\text{avg}})(b_1\, o_p + b_0\, (1. - o_p))]\} \tag{5}$$

where β is an overall width learning rate (0.01 − 1.0, for this work), and b_1 (~0.1) and b_0(~−0.2) are the width adjustment factors associated with class 1 and class 0 datum points, respectively. A maximum basis function half-width is enforced ($b_{\max}$ = 0.2 here).

The overlap-based width adjustment eliminates basis function overlap. This part of the training operates as a logical constraint that shifts basis function boundaries by the least amount that removes overlap. Basis functions are allowed to overlap only if removal of the overlap would reduce a basis function below a set minimum half-width ($b_{\min}$ = 0.005 here).

Finally, the training algorithm includes a basis function elimination scheme that removes basis functions whose widths in any dimension have become less than an elimination threshold value (b_{elim} = 0.01, typically) and whose elimination would not permanently orphan any class 1 datum points. This algorithm is not useful for all classification problems, but if it is used it decreases the size of the network required and increases training and testing speeds. In the present application the basis function elimination works very well. Generally, the network size can be reduced from 30 to 40 nodes to 5 to 10 nodes, whereas the performance of the classifier is either constant or actually improves slightly.

CLASSIFICATION RESULTS

A sample of the training data and the corresponding fit produced by a network trained to identify class 2a to 2d vehicles are shown in Figure 9. The CHC network solution shown uses very little basis function elimination, and thus has more nodes than the minimum required to obtain good performance. The results of applying this network to a test data set are illustrated in Figure 10.

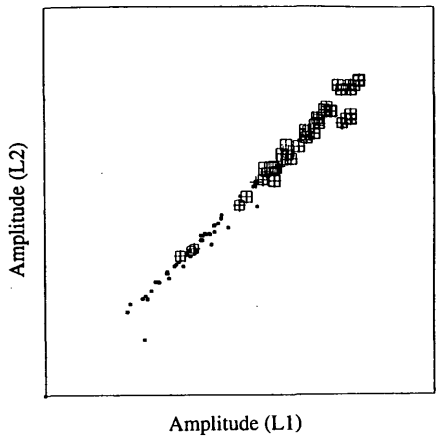

FIGURE 9 Training data (small + is class 0 datum point; large + is class 1 datum point) and network fit (shown by boxes, scaled the same way as the datum point pluses) for sample class 2a to 2d classification task.

Four networks have been similarly trained to handle class 2 and class 3 identifications for six sensor groupings. One advantage of using one network per class is that an estimate can be obtained of the classification error by comparing the sum of the networks' outputs with the actual, known total vehicle count. Since the errors here are dominated by overlap of the two classes, the estimated error is given by the difference between the sum of the network's outputs and the actual vehicle count. Table 2 shows the results of classifying a data set consisting of STS1 measurements for which ground truth is known. The results are summed over six independently processed sensor groupings. The CHC networks used here were

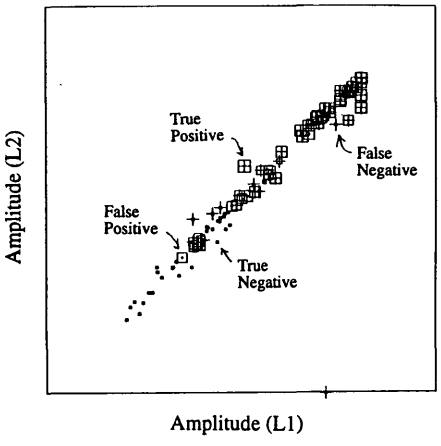

FIGURE 10 Sample test data and network prediction for same network shown in Figure 9. The four kinds of test prediction outcomes are labeled.

TABLE 2 Composite Results of Off-Line Classification Tests on All STS1 Data Processed to Ground Truth Status

Category	NN Pred.	Ground Truth	Act. Error
Volume	6407	6438	-31
Class 2	3955	3828	127
Class 3	2452	2562	-110

trained using about 10 percent of the STS1 data set. The overall volume accuracy obtained is better than 99 percent. The inferred memberships of classes 2 and 3 are accurate to better than 95 to 98 percent in net population count when the classification neural networks have been tuned to near-optimum performance.

CONCLUSIONS AND FUTURE DIRECTIONS

An end-to-end vehicle classification system, based on adaptive and neural network techniques, has been successfully developed and demonstrated that achieves quite good classification of vehicles in classes 2a to 2d and 3a to 3d. Current net volume accuracy is about 99 percent, and classification accuracy is better than 95 percent for the two classes handled by the system when the neural networks are specifically trained for the installation being used to collect the classification data. These accuracies significantly exceed those of an off-the-shelf commercial unit tested under the same circumstances.

In the near future the authors intend to extend the classification system to other vehicle classes (which requires acquisition and analysis of additional data) and to implement some of the neural network-based fault detection ideas.

ACKNOWLEDGMENTS

The authors thank David Albright of the Alliance for Transportation Research for valuable support and discussions throughout this work. The authors thank L. Blair, D. Wilson, K. Lee, and C. Barnes for helpful comments and support as well.

This research was supported by the U.S. Department of Transportation. The cooperation and various kinds of assistance by the U.S. Department of Energy and the New Mexico State Highway and Transportation Department were very important as well.

REFERENCES

1. *Traffic Monitoring Guide.* FHWA, U.S. Department of Transportation, June 1985, pp. 4-3-1 ff.
2. McClelland, J. L., D. E. Rumelhart, and the PDP Research Group. *Parallel Distributed Processing,* Vol. 1 and 2. Massachusetts Institute of Technology, Cambridge, 1986.
3. Wasserman, P. D. *Neural Computing, Theory and Practice.* Van Nostrand Reinhold, New York, NY 1989.
4. Hecht-Nielsen, R. *Neurocomputing.* Addison-Wesley, New York, 1990.
5. Farmer, J. D. Predicting Chaotic Time Series. *Physical Review Letters,* Vol. 59, 1987, p. 845.
6. Lapedes, A. S., and R. Farber. Nonlinear Signal Processing Using Neural Networks: Prediction and System Modeling. Report LA-UR-87-2662. Los Alamos National Laboratory, 1987.
7. Jones, R. D., Y. C. Lee, C. W. Barnes, G. W. Flake, K. Lee, P. S. Lewis, and S. Qian. Function Approximation and Time Series Prediction with Neural Networks. Report LA-UR-90-21. Los Alamos National Laboratory, 1990.
8. Jones, R. D., C. W. Barnes, G. W. Flake, K. Lee, Y. C. Lee, P. S. Lewis, M. K. O'Rourke, and S. Qian. Prediction and Control of Chaotic Processes Using Nonlinear Adaptive Networks. Report LA-UR-90-2755. Los Alamos National Laboratory, 1990.
9. Jones, R. D., Y. C. Lee, S. Qian, C. W. Barnes, K. R. Bisset, G. M. Bruce, G. W. Flake, K. Lee, L. A. Lee, W. C. Mead, M. K. O'Rourke, I. Poli, and L. E. Thode. Nonlinear Adaptive Networks: A Little Theory, a Few Applications. Report LA-UR-91-273. Los Alamos National Laboratory, 1991.
10. Lee, Y. C., S. Qian, R. D. Jones, C. W. Barnes, G. W. Flake, M. K. O'Rourke, K. Lee, H. H. Chen, G. Z. Sun, Y. Q. Zhang, D. Chen, and C. L. Giles. Adaptive Stochastic Cellular Automata: Theory. *Physica,* Vol. D45, 1990, p. 159; S. Qian, Y. C. Lee, R. D. Jones, C. W. Barnes, et al. Adaptive Stochastic Cellular Automata: Applications. *Physica,* Vol. D45, 1990, p. 181.
11. Mead, W. C., P. S. Bowling, S. K. Brown, R. D. Jones, et al. Optimization and Control of a Small-Angle Negative Ion Source Using an On-Line Adaptive Controller Based on the Connectionist Normalized Local Spline Network. *Nuclear Instruments and Methods,* Vol. B72, 1992, p. 271.
12. Bell, S. E., W. C. Mead, G. A. Eiceman, R. D. Jones, and R. E. Ewing. Connectionist Hyperprism Neural Network for the Analysis of Ion Mobility Spectra: An Empirical Evaluation. *Journal of Chemical Information and Computational Science,* in press.
13. Fukunaga, K. *Introduction to Statistical Pattern Recognition.* Academic Press, Boston, 1990.
14. Kohonen, T. *Self-Organization and Associative Memory,* 2d ed. Springer-Verlag, New York, 1988.

Publication of this paper sponsored by Committee on Transportation Data and Information Systems.