Sensor-Based Cognitive State Assessment in a Mobile Environment

Santosh Mathan, Natalia Mazaeva, Stephen Whitlow

Honeywell Laboratories
3660 Technology Dr,
Minneapolis, MN 55418

Andre Adami, Deniz Erdogmus, Tian Lan, Misha Pavel

Bio Medical Engineering Department
Oregon Health Sciences University,
Beaverton, OR 97006

Abstract

Inferring cognitive state from non invasive neurophysiological sensors is a challenging task even in pristine laboratory environments. Artifacts ranging from eye blinks, to muscle artifacts and electrical line noise can mask electrical signals associated with cognitive functions. These concerns are particularly pronounced in the context of the Honeywell team's ongoing efforts to realize neurophysiologically driven adaptive automation for the dismounted ambulatory soldier. Besides the typical sources of signal contamination, the Honeywell team has to deal with the effects of artifacts induced by shock, rubbing cables and gross muscle movement. This paper presents the Honeywell team's efforts to make reliable sensor based cognitive state assessments given the constraints just cited. Cognitive state classification results suggest that it is feasible to classify cognitive state in ambulatory, military-relevant task contexts.

1 Introduction

DARPA's (Defense Advanced Research Projects Agency) Augmented Cognition program is an effort aimed at tailoring computer based assistance to a user's cognitive state. Technologies that have matured under this program promise to foster a fundamentally new type of human computer interaction in complex task domains. Currently, computer based assistance in many challenging contexts takes the form of rigidly automated systems (Sarter, Woods, & Billings, 1997). These systems largely relegate the human to the role of a passive observer and occasionally force the human to take over in extremely demanding task conditions. Computer based assistance conceived by the AugCog program is of a more compliant nature, where the human is actively engaged in the task at all times. Automation merely serves to help users cope with the most difficult of circumstances (c.f. Norman, 1990). This type of mixed initiative interaction, offers the promise of realizing the best attributes of both humans and machines in the service of performing complex tasks.

This paper describes Honeywell's efforts in conjunction with the Augmented Cognition program. The Honeywell Augmented Cognition team focuses on the dismounted Future Force Warrior (FFW). FFW is a component of the US Army's Advanced Technology Development (ATD) program. A critical element of the FFW program is a reliance on networked communications and high density information exchange. Such an infrastructure is expected to increase situational awareness at every level of the operational hierarchy. It is hoped that an information technology based transformation of the military will facilitate better individual and collaborative decision making at every level. However, effective decision making on the basis of broad access to mission relevant information is constrained by the limits of the human information processing system. The goal of the Honeywell Augmented Cognition team is to use physiological and neurophysiological sensors to detect occasions when cognitive resources may be inadequate to cope with mission relevant demands. Efforts of the Honeywell team focus on ways to leverage automation to effectively manage information under difficult task conditions. To the best of our knowledge, the Honeywell team's efforts represent one of the first attempts to create a wearable cognitive state classification system in the context of a fully ambulatory individual. A major factor limiting such applications is the potential for artifacts induced by gross body movements to overwhelm task related neurophysiological signals.

Realizing the vision of the AugCog program in the context of an ambulatory soldier is constrained by several challenges. First, as Schmorrow and Kruse (2002) have noted, processing and analysis of neurophysiological data is largely conducted off-line by researchers and practitioners. However, in order for Augmented Cognition technologies to work in practical settings, effective and computationally efficient artifact reduction and signal processing solutions are necessary. Second, inferring the cognitive state of users demands pattern recognition solutions that are robust to noise and the inherent non stationarity in neurophysiological signals. Third, it requires the development of means to collect reliable neurophysiological data outside the laboratory. Hence, compact and robust form factors associated with neurophysiological sensors and processors are a matter of critical concern. Users should be able to move around freely.

In the following pages we describe a system designed to facilitate cognitive state classification in mobile environments. We describe a hardware configuration that allows neurophysiological data to be collected and processed in a body-worn wireless platform. We provide an overview of software components used for signal processing and artifact reduction. We highlight our classification approach. Additionally, we present results that show it is feasible to discriminate among workload levels on the basis of neurophysiological sensors in ambulatory contexts.

2 Hardware Configuration

The wireless sensor suite employed by Honeywell is assembled using a variety of off-the-shelf components. EEG data is collected using the BioSemi Active Two system using a 32 channel EEG cap and eye electrodes. This system integrates an amplifier with an Ag-AgCl electrode – this affords extremely low noise measurements without any skin preparation. The system also incorporates a wearable Arousal Meter. The Arousal Meter, developed by Clemson University, senses a subject's ECG signals and outputs inter-beat interval data in conjunction with a derived measure of a subject's cognitive arousal. Information regarding physical context is obtained using a combination of a Dead Reckoning Module (DRM) manufactured by Point Park Research and an Inertia Cube manufactured by InterSense. The DRM unit is a self contained navigation component that fuses information from several internal sensors to determine displacement from a specific geographical position. The internal sensors consist of a thermometer, barometer, magnetometer, accelerometer, gyroscope, and GPS receiver. The system is specifically designed to work with intermittent GPS signals. The InertiaCube provides information about head orientation about the head's pitch, roll, and yaw axes.

Figure 1: Body worn sensor suite and signal processing system

Information from the sensors described above is processed on a body worn laptop. The sensors are connected to the laptop via a combination USB, serial port and Bluetooth interfaces. The sensor electronics and the laptop are mounted in a backpack worn by the subject (Figure 1). Sensor data is collected and processed on the laptop computer during the experiment. A base station computer controls the experiment and communicates with the body-worn laptop computer via an 802.11 wireless network.

3 Signal Processing Software

The cognitive state classification efforts reported here rely primarily on EEG data. As mentioned earlier, the sensor monitoring equipment consists of a BioSemi Active Two EEG system with 32 electrodes. Vertical and horizontal eye movements and blinks are recorded with electrodes below and lateral to the left eye. All channels reference the right mastoid. EEG is sampled at 256Hz from 7 channels (CZ, P3, P4, PZ, O2, P04, F7) while the subject is performing tasks. These sites were selected based on a saliency analysis on EEG collected from various subjects performing cognitive test battery tasks (Russell & Gustafson, 2001). EEG signals are pre-processed to remove eye blinks using an adaptive linear filter based on the Widrow-Hoff training rule (Widrow & Hoff, 1960). Information from the VEOGLB ocular reference channel was used as the noise reference source for the adaptive ocular filter. DC drifts were removed using high pass

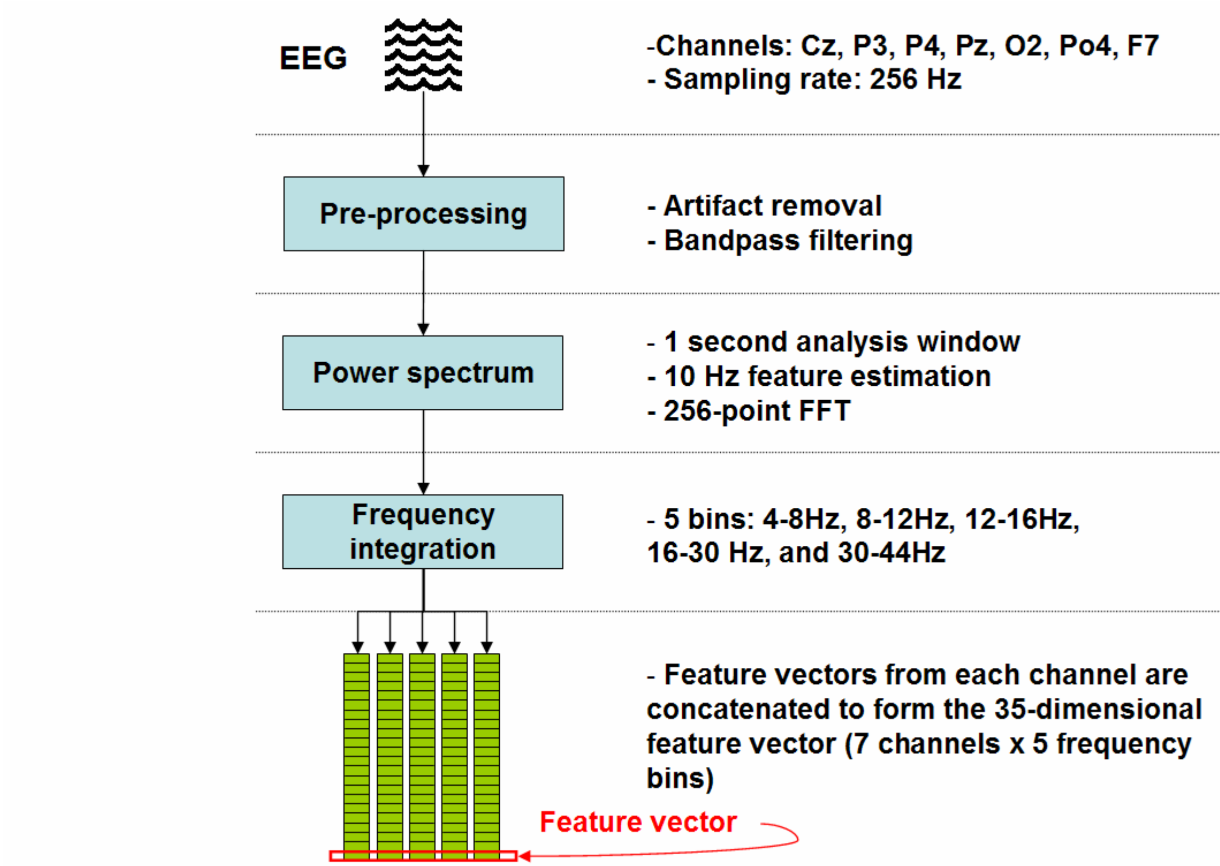

Figure 2: Signal processing system

filters (0.5Hz cut-off). A band pass filter (between 2Hz and 50Hz) is also employed, as this interval is generally associated with cognitive activity. The power spectral density (PSD) of the EEG signals is estimated using the Welch method (Welch, 1967). The PSD process uses 1-second sliding windows with 50% overlap. PSD estimates are integrated over five frequency bands:

4-8Hz (theta), 8-12Hz (alpha), 12-16Hz (low beta), 16-30Hz (high beta), 30-44Hz (gamma). These bands, sampled every 0.1 seconds, are used as the basic input features for cognitive classification. The particular selection of the frequency bands is based on well-established interpretations of EEG signals in prior cognitive and clinical (e.g. Gevins, Smith, McEvoy & Yu, 1997) contexts. The overall schematic diagram of the signal processing system is shown in Figure 2.

4 Cognitive State Classification System

Estimates of spectral power form the input features to a pattern classification system. The classification system uses parametric and non parametric techniques to asses the likely cognitive state on the basis of spectral features; i.e. estimate p(cognitive state | spectral features). The classification process relies on probability density estimates derived from a set of spectral samples. These spectral samples are gathered in conjunction with tasks representative of the eventual task environment. It is assumed that these sample patterns are representative of the population of spectral patterns one would expect in the performance environment. The classification system uses three distinct classification approaches: K nearest neighbor (KNN), Parzen Windows, and Gaussian Mixture Models (Figure 3). We describe each of these components next.

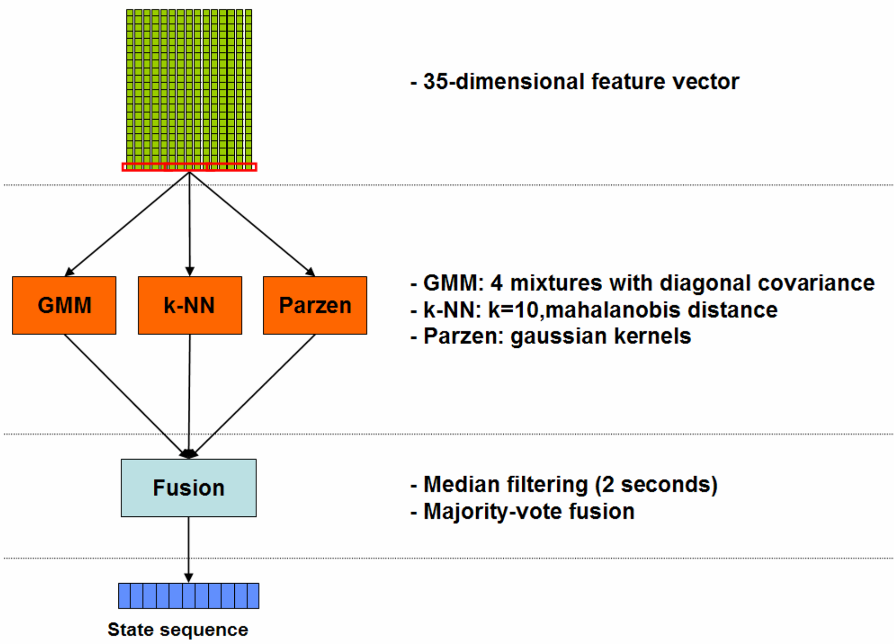

Figure 3: Classification system

4.1 Gaussian Mixture Models

Gaussian Mixture models provide a way to model the probability density functions of spectral features associated with each cognitive state. This is accomplished using a superposition of Gaussian kernels. The unknown probability density associated with each class or cognitive state is approximated by a weighted linear combination of Gaussian density components. Given, an appropriate number of Gaussian components, and appropriately chosen component parameters (mean and covariance matrix associated with each component), a Gaussian mixture model can model any probability density to an arbitrary degree of precision.

The parameters associated with component Gaussians are iteratively determined using the Expectation Maximization algorithm (Dempster, Laird, and Rubin, 1977). Once the Gaussian parameters have been initialized, the system iterates through a two step procedure for each sample associated with each class. In the first step (expectation step), the system computes the probability of a particular training sample belonging to a particular class based on current model

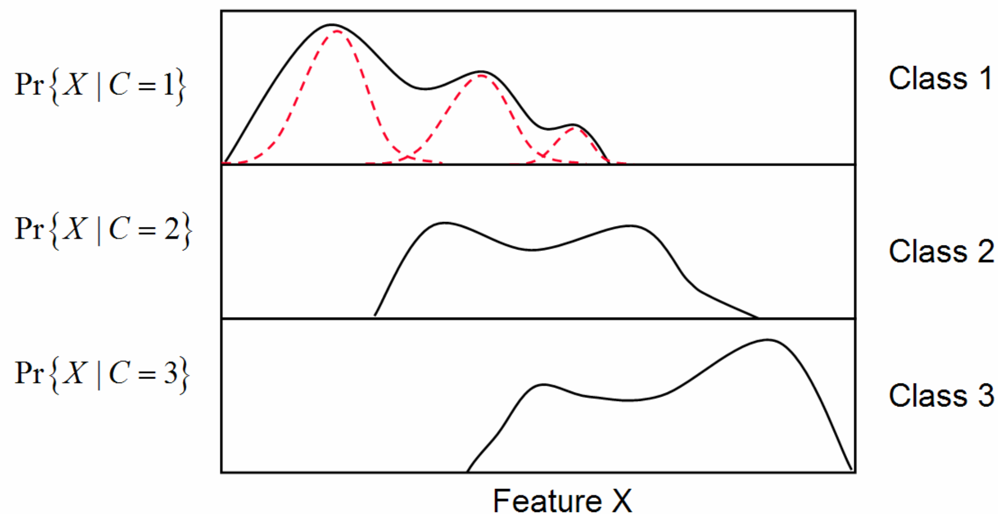

Figure 4: Gaussian mixture models

parameters (posteriori probability). In the maximization step, the model parameters are adjusted in the direction of increasing the class membership likelihood.

Once probability density functions associated with each cognitive state have been generated, it becomes possible to classify individual spectral samples. Each spectral vector is attributed to a class that has the highest posterior probability of representing it. Posterior probabilities are computed using Bayes rule. For example, Figure 4 shows the probability density functions associated with three distinct classes. These probability densities are estimated using three Gaussians. For example, very high values of the data point x are most likely to come from class 3, very low values of x are most likely to come from Class 1.

4.2 K Nearest Neighbor

The K-nearest neighbor approach is a non parametric technique that makes no assumption about the form of the probability densities underlying a particular set of data. Given a particular sample

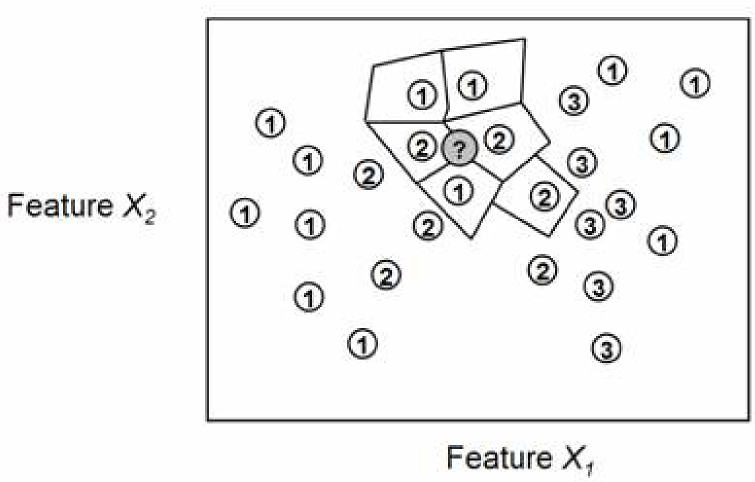

Figure 5: K nearest neighbor

x, the classification process identifies k samples whose features come closest (as assessed by Euclidian or Mahalanobis distance metrics) to the features represented in x. The sample x would be assigned the modal class of the nearest k neighbors. For example, consider the data point represented by the question mark in Figure 5. Based on k = 5, it would be assigned the label associated with the most common class category of it's 5 nearest neighbors: 1. It can be shown that if k is large, but the overall cell small, that the classifier will approach the best possible classification (Bayes rate) (Duda, Hart, & Stork, 2000)

4.3 Parzen Windows

Parzen windows (Parzen, 1967) are a generalization of the *k*-nearest neighbor technique. Instead of choosing the nearest neighbors and assigning a sample x with the label associated with the modal class of its neighbors, one can weight each vote by using a kernel function. With Gaussian kernels, the weight decreases exponentially with the square of the distance. As a consequence, far away points become insignificant. Kernel volumes constrain the region within which neighbors are considered. Consequently, Parzen windows may be a better choice when there are large differences in the variability associated with each class. The data point shown in Figure 4, will be assigned to the dominant class in it's immediate vicinity.

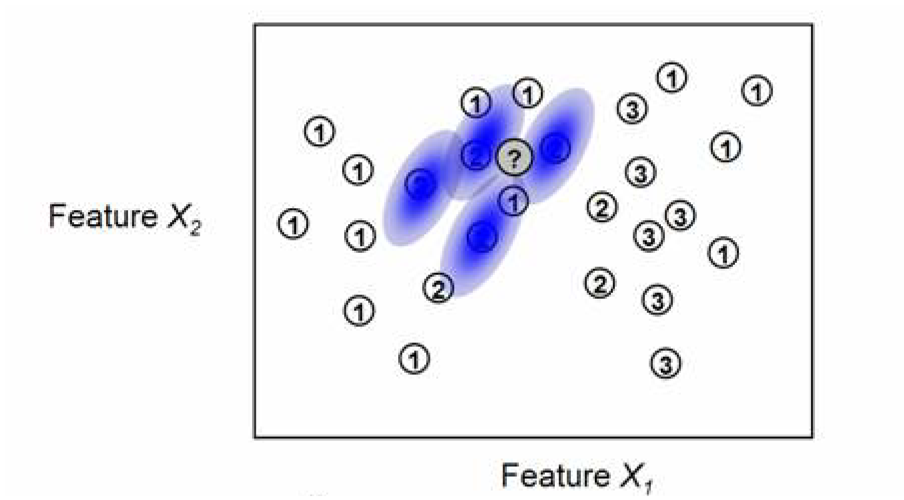

Figure 6: Parzen windows

4.4 Composite Classifier

These statistical classification techniques were chosen over multi-layer neural networks because they require minimal training time. KNN and Parzen Windows require no training, whereas the EM algorithm used to generate GMMs, converges relatively quickly. KNN and Parzen Window approaches require all training patterns to be held in memory. Every new feature vector has to be compared to each of these patterns. However, despite the computational cost of these comparisons at run time, the system was able to output classification decisions well within real-time constraints.

The composite classification system regards the output from each classifier as a vote for the likely cognitive state. The majority vote of the three component classifiers forms the output of the composite classifier. When there is no majority agreement, the Parzen window decision is selected. A classification decision is output at a rate of 10Hz. Outputs from the composite classifier are passed through a modal filter before an assessment of cognitive state is output by the classification system. Modal filtering serves to make the cognitive state assessment process more robust to undesirable fluctuations in the underlying EEG signal. Modal filtering is done over a sliding 2 second window with the assumption that cognitive state remains stable over that period of time.

5 Results

The system described here was empirically assessed. The experiment compared classification accuracy across three workload levels in two mobility conditions: stationary and walking. The tasks in the stationary case were: *relaxed* (waiting for orders), *communicate* (getting orders from base via radio communication), and *count* (starting from 100 and decreasing by 7). Tasks in the mobile case were: *navigate* (walking to a designated target), *navigate and visual search* (walking while looking for snipers), and *navigate and communicate* (receiving and giving mission status reports). The subject wore the sensor suite described earlier in this paper in both mobility conditions. EEG was collected as subjects performed each of the tasks mentioned above.

After the preprocessing and PSD feature extraction stages, approximately 3000 samples were obtained. One third of this data was used for training the classifiers, and the remaining two-thirds were used for testing. Classification results for both stationary and mobile cases are presented in the confusion matrix shown in Figure-7. As the diagonals associated with each confusion matrix indicate, classification accuracy was well over 90%. The results presented here are representative of outcomes replicated with a large number of independent data sets and cognitive tasks.

The cognitive state estimator described here was assessed in the context of a real-time, closed-loop, adaptive performance enhancement system. The system optimally scheduled communication traffic to the subject based on cognitive state assessments. Experiments conducted demonstrate that the assistance offered by this interface improves task-related performance greatly. For instance, the scheduling of communication based on the cognitive load assessment resulted in 100% improvement in message comprehension and 125% improvement in situation awareness.

STATIONARY

true class \ classified as	Relaxed	Communicate	Count
Relaxed	1.000	0.021	0.000
Communicate	0.000	0.979	0.098
Count	0.000	0.000	0.902

MOBILE

true class \ classified as	Navigate	Search	Nav & Comm
Navigate	0.959	0.019	0.000
Search	0.003	0.981	0.047
Nav & Comm	0.038	0.000	0.953

Figure 7: Probability of classifying test patterns correctly. Higher numbers on the diagonal of each matrix correspond to better performance.

6 Conclusion

The ability to detect and to classify the cognitive state of the operator is a prerequisite to successful augmentation of a user's task performance. However, there are numerous technical challenges that limit classification accuracy in the context of an ambulatory soldier. This paper describes hardware and software components that were used to gather, filter, and classify neurophysiological signals. Classification results show that it is indeed feasible to accurately classify the cognitive state of an ambulatory individual.

While these results are encouraging, it is important to emphasize that these results were obtained in conjunction with training and testing data obtained from the same experimental session. We have observed classification accuracy in the 60% to 70% range when training and testing data are drawn from different experimental sessions. Long term non stationarity in EEG limits classification accuracy across experimental sessions. However, it is important to note that EEG is only one component in the cognitive state assessment suite that the Honeywell team is developing. We expect information from sources such as fNIR (functional near infra red imaging), accelerometers, and context modeling to complement EEG based cognitive state assessments. We expect the complementary use of these components to compensate for the variability and unpredictability in operational environments.

Acknowledgments

The authors would like to acknowledge the technical contributions of Jim Carciofini, Michael Dorneich, and Jeff Rye. This paper/research was supported by contract number DAAD-16-03-C-0054 funded through DARPA and the U.S. Army Natick Soldier Center. LCDR Dylan Schmorrow serves as the Program Manager of the DARPA Augmented Cognition program and Mr. Henry Girolamo is the DARPA Agent. The opinions expressed herein are those of the authors and do not necessarily reflect the views of DARPA or the U.S. Army Natick Soldier Center.

References:

C.A. Russell, S.G. Gustafson, "Selecting Salient Features of Psychophysiological Measures," Air Force Research Laboratory Technical Report (AFRL-HE-WP-TR-2001-0136), 2001.

A.P. Dempster, N.M. Laird, D.B. Rubin, "Maximum Likelihood from Incomplete Data via the EM Algorithm," Journal of the Royal Statistical Society, vol. 39, pp. 1-38, 1977.

R.O. Duda, P.E. Hart, D.G. Stork, Pattern Classification, 2nd ed., Wiley, 2000.

A. Gevins, M.E. Smith, L.McEvoy, D. Yu, "High Resolution EEG Mapping of Cortical Activation Related to Working Memory: Effects of Task Difficulty, Type of Processing, and Practice," Cerebral Cortex, vol. 7, pp. 374-385, 1997.

Norman DA. The problem with "automation": inappropriate feedback and interaction, not "over-automation". Philosophical Transaction of the Royal Society of London, 1990; B327: 585-593.

E. Parzen, “On Estimation of a Probability Density Function and Mode”, in Time Series Analysis Papers, Holden-Day, Inc., San Diego, California, 1967.

Sarter, N.B., Woods, D.D., and Billings, C.E. (1997). Automation Surprises. In G. Salvendy (Ed.), Handbook of Human Factors and Ergonomics (2nd edition) (pp. 1926-1943). New York, NY: Wiley.

Schmorrow, D.D., & Kruse, A.A., 2002. Improving human performance throughadvanced cognitive system technology. In: Proceedings of the Interservice/Industry Training, Simulation and Education Conference (I/ITSEC'02), Orlando, FL.

B. Widrow and M. E. Hoff, "Adaptive switching circuits," in IRE WESCON Convention Record, 1960, pp. 96--104.

P. Welch, “The Use of Fast Fourier Transform for the Estimation of Power Spectra: A Method Based on Time Averaging Over Short Modified Periodograms,” IEEE Transactions on Audio and Electroacoustics, vol. 15, no. 2, pp. 70-73, 1967.