Least Squares Adaptive Processing in Military Applications

Delores M. Etter, Allan O. Steinhardt, and Susan L. Stoner

> *A great idea, having stood the test of time, appears obvious, and indispensable.*
>
> —Anonymous

It is often the case that an idea is clearly obvious once it becomes ubiquitous. This is why it is so difficult to judge the innovation content of a new idea or concept. A creative, but familiar, idea invariably seems less brilliant than something new and esoteric. A theoretically complex notion first heard may stimulate the mind but generally cannot compare in innovation with, say, something as simple, but ultimately culture transforming, as the wheel. Only the test of time can separate the truly great ideas from the merely clever ones. By this measure, Widrow's work on adaptive processing is unambiguously seminal.

Though Dr. Widrow initiated data adaptive least squares processing, the initial optimal noise filtering concept was conceived by Norbert Wiener [3] and Andrey Kolmogorov [4] (independently), both whom developed stochastic least squares theory. These prior efforts, while a tour de force of mathematics—one that provided essential insight and analytical tools—suffered from an assumption of the presence of prior knowledge of time series statistics.

Widrow's Contributions

Reformulating Least Mean Squares in the Data Domain

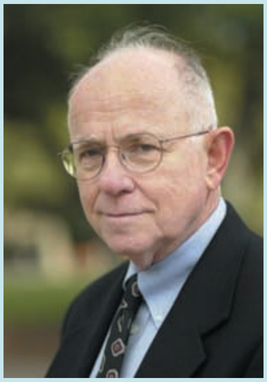
> *Progress in mathematics is often obtained by finding the right notation.*
>
> —Kline

In [5], Kline makes the case that matrix theory arises almost automatically once

Benjamin Franklin Medal

For nearly two centuries, the Franklin Institute has recognized outstanding achievements in science and technology. Award recipients illustrate the diversity of talents necessary to take science from its basic theoretical level to the consumer marketplace. Today the Institute awards 22 medals in a range of science and technology fields with a total endowment of over $16.6 million. Winners are selected by the Committee on Science and the Arts.

The Benjamin Franklin Medal in the field of electrical engineering was awarded in 2001 to Dr. Bernard Widrow for pioneering work in adaptive signal processing as exemplified by the least mean squares algorithm, adaptive filters, adaptive control, adaptive antennas, noise cancellers, artificial neural networks, and directional hearing aids [1]. Dr. Widrow is the author of two books and numerous journal articles on adaptive signal processing and is a Professor of Electrical Engineering at Stanford University.



This article is an extension of a presentation made to celebrate the awarding of the Franklin medal to Dr. Widrow in April 2001 [2]. In it, we reflect on the impact of Dr. Widrow's work on military applications. His least squares adaptive processes have enhanced the performance of a range of systems. Many of these applications have long-term potential for spin-off in civilian arenas, most notably communications and remote sensing.

We hope this article gives readers an appreciation for the general manner in which practical system level constraints give rise to signal processing requirements. Such an appreciation is fundamental to the signal processing community's continued success in areas such as communications, medicine, and space exploitation.

We first summarize least squares adaptation, drawing on citations and graphics from Widrow's early work and contrasting it to previous activities. We next survey deployed military systems that apply his work and discuss briefly possible future defense applications of least squares processing. Finally, we conclude with a retrospective of why this work has been so influential, and what lessons can be derived from it.

the proper notation is established. He shows that determinant theory actually precedes linear algebra. It is only after the notation of a matrix—a single variable describing an entire array of quantities—is introduced that linear algebra, with spectral decomposition and rank, flourishes. Likewise, one can argue that notational innovation in Widrow's 1959 work [6] virtually guaranteed a renaissance in adaptive processing.

Fig. 1, created 42 years ago, formulates the linear prediction problem not as a shift operator applied to a second-order stochastic representation of an ergodic process, but as a circuit! This simple graph changed forever how people view least squares estimation. It cast prediction as a problem of identifying parameters in a circuit, not solving for a Green's function in a stochastic integral. Certainly this figure can be recast in the language of Wiener-Kolmogorov's integral equations. However, it is the ease of thought that this great notation brings, here manifested as a combination of data based and graphical visualization, that is an intellectual force multiplier.

Having prompted the reformulation of the ensemble-based least squares work of Weiner-Kolmogorov, Fig. 1 gives rise to considerations of cost minimization in the data domain (see Fig. 2). This leads to the elegant least mean squares (LMS) recursion, which is a gradient implementation of least squares in the data domain

$$\vec{B}_{k+1} = \vec{B}_k + 2\mu\varepsilon_k \vec{x}_k$$

where $\vec{B}_k$ is the vector of filter coefficients, μ is the step size, ε_k is the error, and $\vec{x}_k$ is the data in the filter. The practical value of this simple equation is appropriately summarized:

> Within a half hour of the time that the LMS algorithm was written on the blackboard, [my student] Hoff had it working in hardware. [There was] a large analog computer in the building, right across the hallway. There was nobody in the computer room. We just went in, picked up a plug board, and Hoff wired it together. This happened on a Friday afternoon in the autumn of 1959.
>
> —Widrow [7]

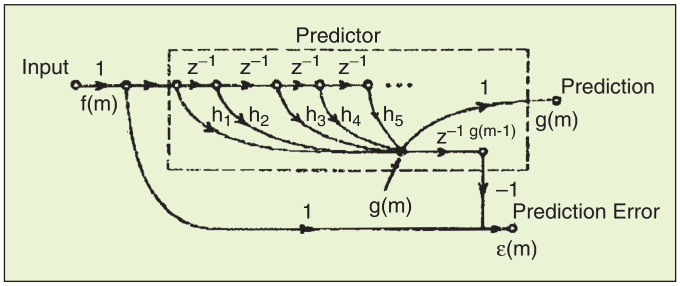
The LMS algorithm is not only simple but also robust to nonstationarity, non-Gaussian input, and model mismatch. In particular, the LMS algorithm can track statistical variations of the noise process, since these variations manifest themselves as changes in each sample's surface whose gradient is being updated. (The first convergence analysis of second-order moments of the LMS algorithm is found in Dr. Senne's thesis dissertation under Dr. Widrow [8].) The LMS algorithm makes no assumption on the noise marginal density and performs well on non-Gaussian processes.

Widrow and his colleagues identified a number of subtle, practical considerations to the algorithm that continue to attract attention. One example is the issue of "targets" in the training set, as depicted in Fig. 3. Here a target refers to the signal of interest. Unlike the Wiener equations, where statistics are known a priori, the LMS scheme assumes data adaptive training. Rarely is the signal distinctive enough that it is isolated in the measurement process. As such, one ends up with the problem that if the target occurs during training (that is within the data set used to adapt the LMS filter), it may appear as noise and is therefore subtracted out. It becomes especially pronounced when there is model mismatch between the target model and reality. For this reason, Widrow devised the concept of a pilot signal, a signal deliberately injected to provide the proper adaptive convergence constraints, to improve training performance. This is illustrated in Fig. 3, where a synthetic signal, formed in the pilot signal generator box, is added at the moment of adaptation to ensure proper signal preservation.

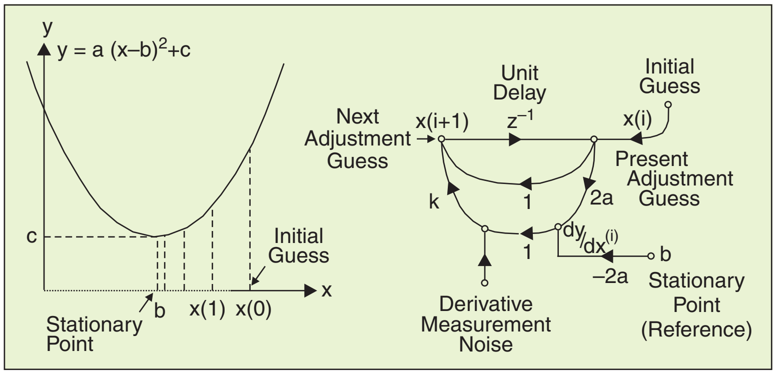
Fig. 4 portrays a second example that considers wideband effects and their associated need for tapped delay lines. Tapped delay lines, finite impulse response filters on each adaptive channel, allow for time and space processing and are referred to today as STAP, space time adaptive processing. Wideband data, wherein a single phase term is insufficient to characterize changes from channel to channel, can be addressed by means of transversal filters on each channel. This innovation allows wideband data to be processed as if it was narrowband data.

Spatial Adaptivity

Most defense-unique applications for least squares adaptivity are in the area of spatial adaptivity. Often, the

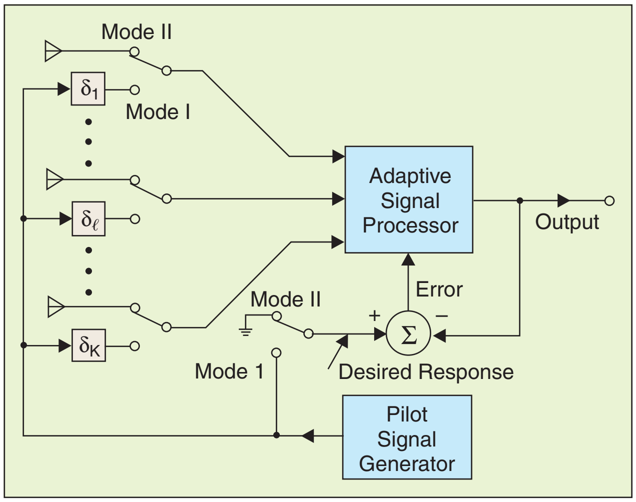


▲ 1. Sample data representation of the linear predictor [6].

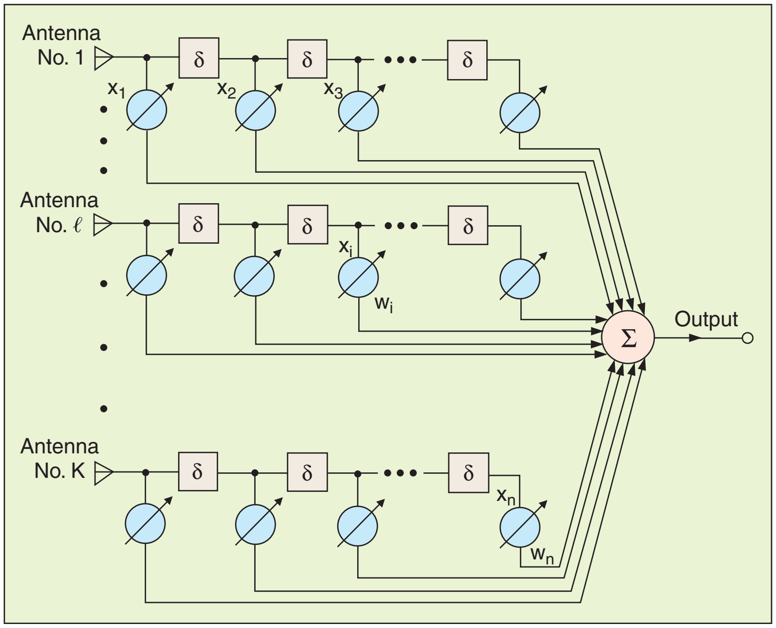


▲ 2. Fast objective function minimization: (a) error surface and (b) update algorithm [6].

side lobes of an array are not sufficient to reject signals that can overcome deterministic attenuation. Such signals can be deliberate jamming, natural, or unintentional human artifacts. (Examples are described in the following sections.) Normally, the precise antenna pattern in the side lobe region, i.e., away from where the antenna is pointing, is unknown, but the least squares approach does not require this knowledge. In addition to side lobe suppression, the least squares process can address main beam nulling, in which the interferer is separated by less than a quiescent pattern beam width. Fig. 5 displays the first openly published schematic of an adaptive spatial pattern that is the plot of gain of the adaptive system, after convergence, as a function of the arrival of a signal impinging on the array.



▲ *3. Pilot signal generation for improved adaptive convergence [17].*



▲ *4. Tapped delay line beamformer [17].*

Least Squares Applications in Deployed Military Systems

> *Emerging military requirements:*
>
> *A robust multisensor information grid providing dominant awareness of the battle space...*
>
> *A joint communications grid with adequate capacity, resilience, and network-management capabilities...*
>
> —Joint Vision 2020

Joint Vision 2020 [9], a strategy plan developed by the Joint Chiefs of Staff to guide the transformation of the armed forces for the new and emerging threats of this new century, gives a description of emerging requirements in the Department of Defense (DoD). Situational awareness, the knowledge of the current state of affairs, and robust communications are crucial to conflict avoidance and to decisive victory. As shown in the following sections, adaptive least squares processing is a truly enabling technology for both of these application venues.

Not all research in the area of spatial adaptivity yields results that can be applied in deployed systems. That which has been theoretically studied and demonstrated in the laboratory often does not survive the rigors of testing during DoD operational evaluation. Some examples of signal processing technologies that have not yet significantly transitioned into military systems include chaos theory (the theory of nonlineary systems exhibiting self similarity), cumulants (the theory of higher order moments in statistical models), and singular value/eigendecomposition (spectral theory of linear operators).

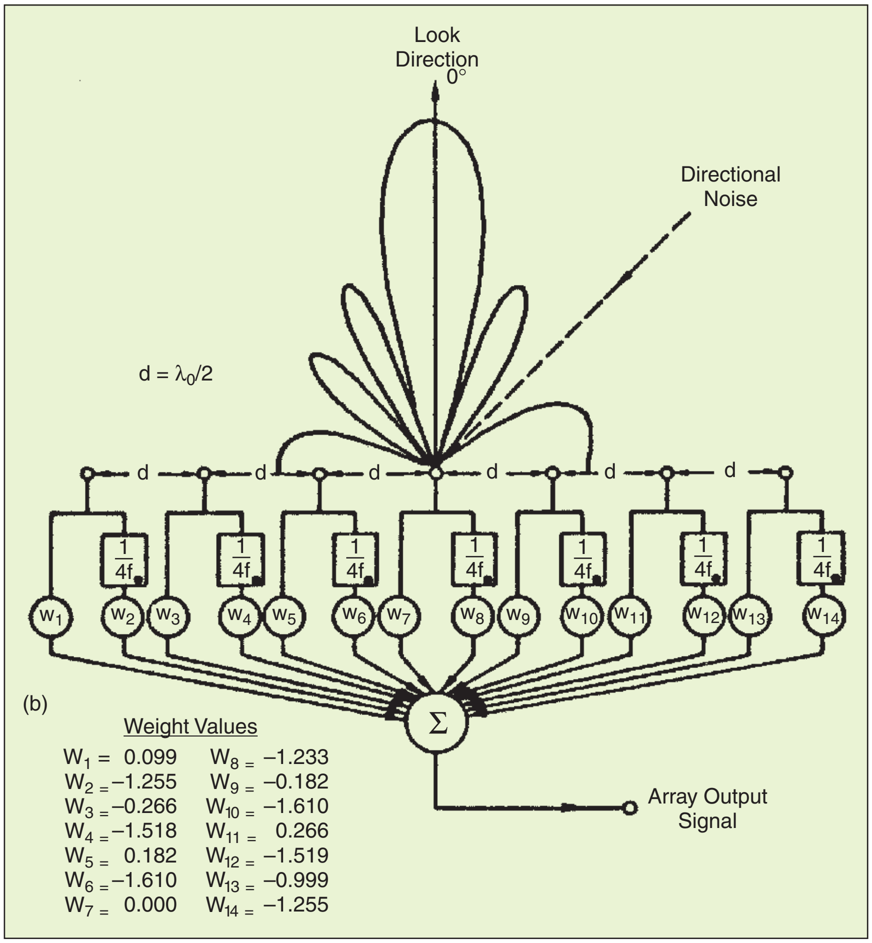
Sensing Applications

Submarine Detection

Submarines played a significant role in both World Wars. German U boats, for example, were a substantial threat to convoys ferrying supplies and military personnel across the Atlantic Ocean. The heavy loss of convoys and surface combatants motivated both militaries (Allied forces and Axis powers) to develop capabilities to address issues associated with submarine detection. The advent of the nuclear submarine, with its substantially reduced acoustic signature, increased dramatically the difficulty of submarine detection. Furthermore, nuclear submarines offered reduced opportunities for surface contact as reoxygenation of diesel batteries was no longer necessary.

Motivation for detection of these quiet nuclear submarines further increased when these vehicles emerged as the third leg in the Cold War triad, joining land-based and airborne components of the nuclear arsenal. For decades, the difficulty of detecting ballistic missile submarines was a key factor in the balance of power between the Cold War superpowers. Any significant breakthrough in detection capability could have tipped the balance of power, and hence this capability was of enormous strategic importance. For this reason, the U.S. Navy generously funded signal processing research in the 1960s-1980s with the hopes of advancing our ability to detect underwater acoustic signatures (Widrow's early work on adaptive beam forming was funded by the U.S. Navy [10]).

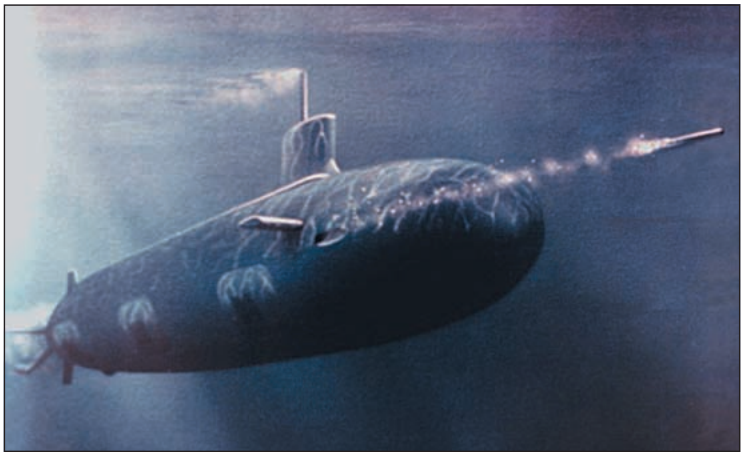
It was during this period that Widrow introduced his least squares processing ideas, and work commenced to put it to practical use. Soon, his work was being used to tackle the problem of growing concern for antisubmarine warfare. Since nuclear submarines could threaten large cities from deep off the continental shelf, there was a need for long-range detection capabilities. Low-frequency energy generally propagates further than high frequency emissions. Hence, low-frequency sensors were needed to address the increased standoff ranges. As submarines became quieter, the range at which they could be detected became shorter (for a fixed sensor suite), so the need for improved range offered by low frequency became doubly important. The need for long- range, low-frequency detection translated into nearly guaranteed main beam nulling requirements; ships or marine life many miles away would be collocated in the same bearing channel as a submerged target. This meant hardware (i.e., physical aperture or frequency) could not be used to separate signals from noise, hence adaptive noise filtering had to be applied. Mainbeam nulling requires careful placement of a precise, scenario dependent null to suppress interference. Unlike sidelobe interference, which can often be adequately removed with deterministic precomputed weights, mainbeam suppression is intrinsically an adaptive proposition.



▲ *5. Spatially adaptive beam former [17].*

Incorporating adaptive noise suppression into a system design is exceptionally complicated; the danger of target nulling exists, adaptive losses must be dealt with, computational demands are heavy, and characterizing post-nulled stochastic signals is very challenging. Low frequency, coupled with substantial standoff, ensured that adaptive beam forming would have to be used for submarine detection, even with the advent of long arrays of hydrophones replacing the sonobuoy.

The U.S. Navy deployed the sound surveillance system (SOSUS) array in the 1950s. SOSUS comprised a system of arrays of hydrophones mounted on the bottom of coastal regions along the continental shelf in both the Atlantic and Pacific Oceans. In the 1990s, SOSUS was converted to a programmable commercial off the shelf (COTS)-based system so that algorithm upgrades could be performed merely by changing software code. Adaptive beam forming was then implemented to enhance hydrophone signal clarity. With the diminished need for antisubmarine warfare, dual use applications of SOSUS have emerged in recent years, in-



▲ *6. Seawolf class attack submarine.*

cluding seismic monitoring, marine mammal monitoring, and thermal thermometry.

Los Angeles and Seawolf class submarines (see Fig. 6) are equipped for antisubmarine warfare. Since 1975, towed sonar arrays have been used on these boats for passive acoustic detection of submarines and surface ships. In 1997, a rapid COTS insertion program included the incorporation of a programmable signal processing system on these towed arrays. Adaptive beam forming was part of the suite of algorithms applied to the arrays. Currently towed arrays of three lengths, all incorporating adaptive beamforming, are used.

Adaptive processing for active bow-mounted sonar has been explored in the laboratory, but the quiescent narrow main beam associated with higher frequency activation renders the relative gain of adaptivity less dramatic; hence such systems have not yet been systematically deployed.

Seismic Monitoring

Adaptive least-squares processing is used to reduce surface noise from long-range, low-frequency seismic arrays which are employed to detect nuclear detonations. The DoD has seismic systems located worldwide for the purpose of monitoring nuclear tests. Generally, the natural side lobe rejection from quiescent patterns suffices for the elimination of unwanted noises, but at times, nearby strong acoustic sources (trucks driving by, for example) need to be filtered out adaptively. The impact of adaptive least squares processing in this venue has been less dramatic than in sonar, partly because the propagation physics is more complex, but also because the signal of interest is impulsive and thereby easier to extract from pseudostationary interference by conventional methods [11].

Surface Moving Target Detection

The use of radar to detect moving ground targets (called ground moving target indication or GMTI) is relatively recent. The challenge of GMTI is the rejection of clutter, that is, of ground returns. Ground-to-air radars do not have to contend with clutter. Air-to-air radars only encounter clutter in a look-down mode. Fortunately, these radars can separate targets from clutter by means of waveform selection because the targets, whether aircraft or missiles, are moving very fast, so the apparent velocity of the ground can be filtered out using linear (Doppler) filters. But GMTI targets move slowly and cannot be separated from clutter with standard waveform processes. The motivation for GMTI is to provide beyond line-of-sight knowledge of enemy troop activity. The Defense Advanced Research Projects Agency (DARPA) explored GMTI in the 1960s with the Assault Breaker program, which later transitioned to the Air Force as the Joint Surveillance Target Attack Radar System (initially known as JSTARS and subsequently designated Joint STARS). Assault Breaker and Joint STARS, air-to-ground surveillance systems, were developed during the Cold War, when the major threat was an attack by the Soviets on Western Europe. The improved situational awareness provided by these systems was critical during this period to offset the Soviet's overwhelming numerical advantage.

Joint STARS (see Fig. 7) was used with great success against Iraqi forces in Desert Storm and later in the Balkans in Bosnia and Kosovo. During Desert Storm, Joint STARS provided U.S. air power with enemy ground force targets (tanks and reconnaissance and supply vehicles) enabling devastating Iraqi losses prior to initiation of a quick and decisive ground war. Joint STARS suppresses clutter using a technique known as displaced phase center array (DPCA), a deterministic clutter null placement scheme. DPCA exactly cancels, rather than nulls, clutter. This implies higher signal losses and thermal noise gains than can be achieved with adaptive nulling; the null is placed just deep enough and narrow enough to be effective while minimally distorting the antenna pattern.

▲ *7. E-8C: Airborne platform for the joint surveillance target attack radar system.*

▲ *8. Global Hawk unmanned aerial vehicle.*

GMTI has recently migrated to unmanned aerial vehicles (UAVs). One such aircraft is the Global Hawk. The Global Hawk UAV, shown in Fig. 8, is a high-altitude, long-endurance reconnaissance system. With an endurance of approximately 30 hours, Global Hawk can provide near real-time imagery of large geographic areas. Combined with satellite connectivity, its endurance and range allow it to be controlled from, and to relay data back to, a ground site in the United States from virtually anywhere in the world.

The first airborne demonstration of GMTI on the Global Hawk (also a DARPA effort) was in December 2000. Shortly thereafter, maritime target tracking was demonstrated as well. Unlike Joint STARS, Global Hawk's GMTI rejects clutter by using least squares processing, mainly covariance based space-time adaptive processing. Global Hawk's high endurance mandates that it have a small airframe, which, in turn, requires a small radar antenna aperture. This smaller aperture results in a reduced ability to deterministically reject clutter (as required by DPCA); hence, adaptive processing is required.

Unmanned reconnaissance is increasingly demanded by the combination of deeper look requirements to enable the deployment of weapons from safe, stand-off locations, and increased sensor standoff range requirements (due to the improved range of counter surface-to-air weapons). Global Hawk is perhaps the first transition of least squares processing which is driven by post Cold War mission requirements—uncertain theater of conflict, reduced manning requirements, and increasing intolerance for human casualty.

There is growing interest in putting least squares adaptive processing on the E-2C Hawkeye aircraft to allow this carrier-based air defense radar to be converted from a maritime air surveillance system to a littoral GMTI surveillance system. (Littoral surveillance is the examination of a coastal region, especially the shore zone between high and low watermarks.) The GMTI capability would provide surveillance in support of deployment of land forces from coastal regions. Such efforts, however, lack mature definition and have not yet been funded.

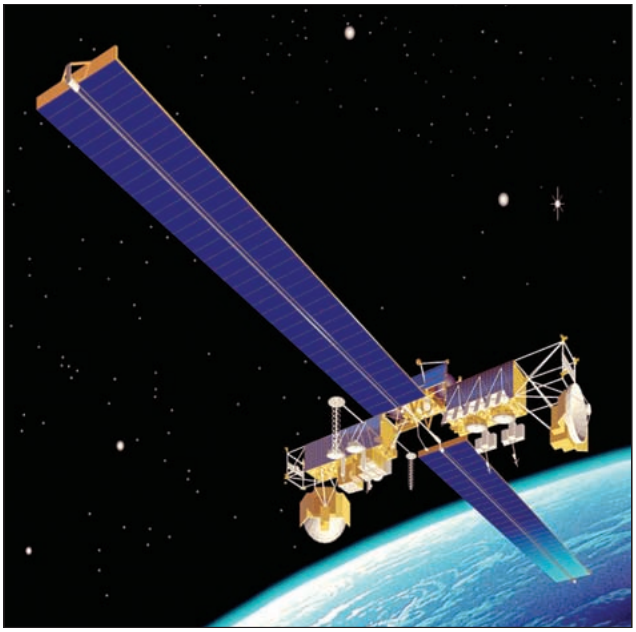
Robust Communications

Military uses place unique demands on communications. Intelligence, surveillance, and reconnaissance requirements, such as the need to track enemy movements across the battle space (airborne, ground, maritime), will produce massive amounts of data that must be broadly disseminated in a timely fashion. In addition, other operational, administrative, and logistical information must be rapidly conveyed and updated. Since the location of future conflicts is uncertain, the communications grid must be global. Furthermore, military communications systems must be secure against electronic attacks in the forms of jamming and intrusion ("hacking").

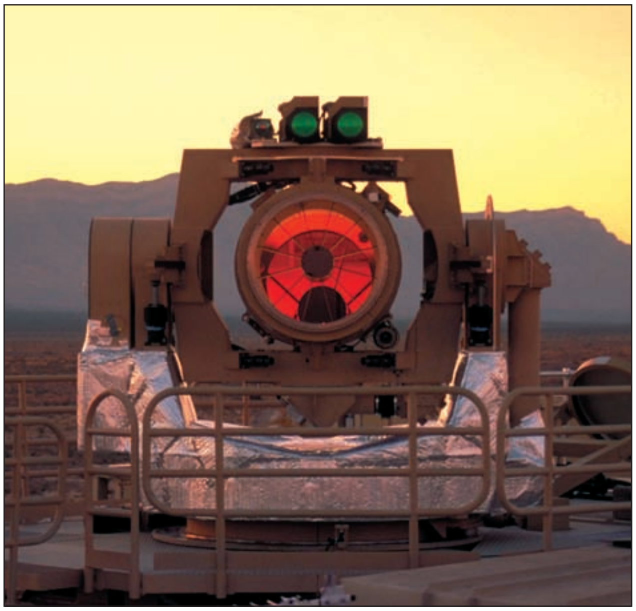
Milstar

Milstar (see Fig. 9) is a joint service satellite communications system which enables encrypted voice, data, teletype, and facsimile communications. It is designed to provide national authorities the means to instantaneously communicate with their forces worldwide, secure from eavesdropping, and protected from enemy electronic jamming. Four satellites are required for the Milstar constellation, three of which are presently in orbit. Two of these are Milstar 1 satellites, which are capable only of low data rate communications. Upgraded Milstar 2 satellites, the first of which was launched in February 2000, will make up the remainder of the constellation.

Milstar 2 satellites feature medium data rate (in addition to low data rate) communications that enable the



▲ *9. Milstar military communications satellite.*



▲ *10. Tactical high-energy laser.*

near-instantaneous transmission of regular voice communications, imagery, intelligence, and air tasking orders. Milstar 2 spacecraft also incorporate a nulling feature, a multichannel antijam adaptive beam former, that will neutralize close-in enemy electronic jamming capabilities [12].

A communications satellite has a large field of view and a predictable orbit making it easy prey to jamming from ground-based emitters. To offset this possibility, beam-forming techniques provide robust electronic counter-countermeasures effectiveness since they do not rely on bandwidth diversity (spread spectrum), quiescent side lobes, or power management. The beam former weights must be computed in space due to the relay nature of space communications. The simplicity of the LMS is a key enabler for the confined form factor (in size, weight, and power) for space-based implementations. Milstar 2 is the first system in space to implement space-time adaptive processing.

High-Energy Lasers

The use of lasers to attack incoming missiles was first given serious consideration as part of the Strategic Defense Initiative, referred to as "Star Wars" in the 1980s. Despite the end of the Cold War, the threat from theater, intercontinental, and cruise missiles has not diminished. The Tactical High-Energy Laser (THEL) system, shown in Fig. 10, has recently demonstrated the ability to engage and defeat short-range artillery rockets. Since June of 2000, THEL has successfully shot down 23 rockets. The THEL demonstrator is being developed in a joint U.S.-Israeli program. The weapon design has been driven, in part, by Israel's requirements for an air defense system to protect communities located along country's northern border from terrorist rocket attacks. The demonstrator system uses deuterium-fluoride chemical laser technologies. In the future, a smaller, mobile version of the THEL may employ a solid-state laser. The THEL system contains a fully digital control system with a 10-kHz sample rate. One of the challenges of the control system was to stabilize the laser beam with two 6-in beryllium mirrors. In the design of the control system, a slightly modified adaptive noise cancellation filter was required to remove a destabilizing harmonic signal.

Emerging Applications

While we have focused in this article on deployed military systems, emerging applications of least squares adaptivity hold great promise for the DoD as well. Adaptive optics, introduced in the THEL application to correct for atmospheric turbulence, has great growth potential for both energy weapons and wideband communications. Very wideband clutter rejection, through the use of Widrow's LMS technique, is under consideration for imaging moving vehicles to improve situational awareness under dynamic conditions [13]. A robust, survivable use of the global positioning system geolocation capability [14] is another growth area for least squares adaptive processing. Here, the intent is to render GPS immune to jamming by means of adaptive interference rejection.

Hyperspectral imagery has the potential for significantly enhancing intelligence collection and analysis. Key hyperspectral applications include terrain characterization, anomaly detection and material identification, and atmospheric characterization. In this type of imaging, noise must be suppressed to extract targets. The steering vector is the temporal spectral structure, so in a sense, this application presents the mathematical twin of space-time processing first envisioned by Widrow.

Lastly, the vision of global access and near continuous tracking of targets, in all weather, anywhere, anytime, lures one to entertain notions of a constellation of GMTI and maritime moving target indication-capable space-based radars. The utility of such a system is so compelling that it is a matter of when, not if, it will be realized. Such a system, once orbiting, will surely be enabled by clutter suppression by means of space time least squares adaptive filtering.

Conclusions

A few insights can be derived from the above survey. First, it is often hard to predict the scope of applications of a radical new idea. This suggests a need to invest in basic research and development, without necessarily knowing how the research will be applied in the future. This does not imply, however, that an applied focus is not useful. Widrow's early work was motivated by neural science, specifically mimicking biological processes for artificial intelligence [6], [7], [15], [16]. Working towards an application (and almost any application will do) provides valuable focus. The impact of real-world effects (in least squares these include nonstationarity, model mismatch, non-Gaussianity, channel decorrelation, etc.) are difficult to predict a priori but are often omnipresent across application domains.

Next, a novel idea can transform entirely the applications it touches. In addition to enhancing effectiveness, the entire complexion of the system can evolve. Without LMS theory, modern DoD surveillance, communications, and weapon systems would look very different. As an example, the aperture of many surveillance systems would have to be much larger—increasing overall system size and hence decreasing military effectiveness—if robust main beam nulling was not available.

A stunning aspect of both the LMS algorithm and the sampled data view of stochastic estimation furnished by Widrow are their extreme simplicity. This is worth noting; complex, esoteric solutions have a habit of falling prey to the corroding influence of real-world considerations.

We have surveyed the use of LMS and LMS-like applications in various military systems. These include radar surveillance (air and space-based sensing of moving surface targets), sonar surveillance (from towed and bottom mounted arrays), seismo-acoustic surveillance (stationary geophones), robust antijam communications, and energy directed weapons. The demonstration of the significance of the LMS algorithm in military systems is a tribute to the elegant, simple, and robust legacy in the adaptive least squares processing work of Dr. Bernard Widrow.

Delores M. Etter received her B.S. and M.S. degrees in mathematics from Wright State University and her Ph.D. degree in electrical engineering from the University of New Mexico. She was with the Electrical and Computer Engineering faculty at the University of New Mexico from 1979-1989 and the Electrical and Computer Engineering faculty at the University of Colorado from 1990-1998. From June 1998 through July 2001, she was the Deputy Under Secretary of Defense for Science and Technology and had oversight responsibility for the Department of Defense S&T program. She is currently a member of the Electrical Engineering faculty at the United States Naval Academy and holds the Office of Naval Research Distinguished Chair in Science and Technology. She is a former President of the IEEE Acoustics, Speech, and Signal Processing Society. She is also a member of the National Academy of Engineering. She is a Fellow of the IEEE.

Allan O. Steinhardt received a Ph.D. degree from the University of Colorado, Boulder, in electrical engineering. From 1984 to 1987, and 1993 to 1996 he was a member of the technical staff at the Lincoln Laboratory, Massachusetts Institute of Technology, Cambridge, where he conducted research in adaptive array processing. From 1987 to 1993, he was on the faculty at Cornell University, Ithaca, NY, where he conducted research in adaptive array processing. He is currently chief scientist for signal processing at the Defense Advanced Research Projects Agency, overseeing the sensor programs in the information exploitation office. He received the 1986 IEEE Signal Processing Society Paper Award and the 1990 Best Professor Award from the Cornell student chapter of the IEEE. In 1998, he received the outstanding achievement award from the United States Office of the Secretary of Defense. He is a Senior Member of the IEEE.

Susan L. Stoner received her B.S. and M.S. degrees in mechanical engineering from New Mexico State University and University of California at Davis, respectively. She has been with Lawrence Livermore National Laboratory since 1985 where she has had responsibilities associated with national and international security, materials engineering, and engineering management. From 1994-1998, she worked in Naka, Japan, on an international team designing an experimental magnetic fusion reactor. Since August 2000, she has been on assignment in the Department of Defense as a science advisor for nuclear matters and as assistant to the Deputy Under Secretary of Defense for Science and Technology.

References

[1] Franklin Institute Online: http://sln.fi.edu/tfi/exhibits/bower/

[2] D. Etter, "Adaptive signal processing in DoD," presented at the Franklin Institute award ceremony, Villanova, PA, Apr. 25, 2001.

[3] N. Wiener, *The Extrapolation Interpolation and Smoothing of Stationary Time Series*. New York: Wiley, 1949.

[4] A. Kolmogorov, "Interpolation and extrapolation," *Bull. Acad. Sci., USSR Ser. Math*., vol. 5, pp. 3-14, 1941.

[5] M. Kline, *Mathematics from Ancient to Modern Times*. Oxford, U.K.: Oxford Univ. Press, 1972.

[6] B. Widrow and M. Hoff, "Adaptive switching circuits" in *IRE WESCON Conven. Rec.*, IEE 1960.

[7] B. Widrow in *Talking Nets: An Oral History of Neural Networks*, J.A. Anderson and E. Rosenfeld, Eds. Cambridge, MA: MIT Press, 2000, pp. 52-53.

[8] K.D. Senne, "Adaptive linear discrete-time estimation," Stanford Univ. Center for Systems Research, Tech. Report 6778-5, June 1968.

[9] Joint Vision, America's Military Preparing for Tomorrow, 2000, http://www.dtic.mil/jv2020/jvpub2.htm

[10] L. Griffiths, "Early history of the LMS algorithm," presented at the Franklin Institute award ceremony, Villanova, PA, Apr. 25, 2001.

[11] R. Lacoss, MIT Lincoln Laboratory, private correspondence, 2001.

[12] USAF Space Command: http://www.space.com/missionlaunches/launches/fl_milstar_010227.html

[13] Affordable Surface Target Engagement, DARPA, 2001, http://www.darpa.mil/spo/programs/affordmovingsurfacetarget.htm

[14] GPS anti-jamming program, DARPA, 2001, http://www.darpa.mil/spo/Programs/gpsguidancepackage.htm

[15] B. Widrow and M. Lehr, "30 years of adaptive neural networks: Perceptron, madaline, and backpropagation" *Proc. IEEE* , vol. 78, pp. 1415-1442, Oct. 1990.

[16] B. Widrow, R.G. Winter, and R.A. Baxter, "Learning phenomena in layered neural networks," in *Proc. 1st. Int. Conf. Neural Networks*, vol. II, 1987, pp. 411-429.

[17] B. Widrow, P.E. Mantey, L.J. Griffiths, and B.B. Goode, "Adaptive antenna systems," *Proc. IEEE*, vol. 55, pp. 2143-2159, Dec. 1967.