A Computational Approach to Understanding Advanced THERMAL BARRIER Coatings’ Performance

PAGE 38
ABOUT DSIAC

PURPOSE
The purpose of DSIAC is to provide information research and analysis for DoD and federal government users to stimulate innovation, foster collaboration, and eliminate redundancy.

MISSION
The mission of DSIAC is to generate, collect, analyze, synthesize, and disseminate scientific and technical information (STI) to DoD and federal government users and industry contractors.

VISION
DSIAC will be the premier information research partner and curator of technology advancements and trends for the defense systems community.

PRODUCTS & SERVICES

TECHNICAL INQUIRIES (TIs)
We offer 4 hours of information research (free to the client) in response to TIs on any of the DSIAC subject areas. TI response efforts generally include literature searches, document requests, answers to technical questions, and expert referrals.

CORE ANALYSIS TASKS (CATs)
We conduct customer-funded CATs that expand the limits of the free TI service. Each CAT is limited to a $1M ceiling and can be incrementally funded.

STI UPLOAD
We collect STI, adding it to the DTIC collection, and ensuring that the body of information is available to share with the DoD community.

TRAINING/EVENTS
We participate, host, and promote key technical conferences and forums to engage and network with the science and technology (S&T) community.

PROMOTION
We publicize events, technologies, information research products, and more to our +75,000 members.

INFORMATION RESEARCH PRODUCTS
We publish the DSIAC Journal, Defense Systems Digest, State-of-the-Art Reports (SOARs), and many other information products available for free electronic download on our website.

HOW TO CONTACT DSIAC

Ted Welsh
DSIAC Director

DSIAC HEADQUARTERS
4695 Millennium Drive
Belcamp, MD 21017-1505
Office: 443.360.4600
Fax: 410.272.6763
Email: contact@dsiac.org

Brian Benesch
DSIAC Technical Project Lead

DSIAC TECHNICAL CONTRACTING OFFICER REPRESENTATIVE (TCOR)
Peggy M. Wagner
704 TG/OL-AC
2700 D Street, Building 1661
Wright-Patterson AFB, OH 45433-7403
Office: 937.255.4126

WPAFB SATELLITE OFFICE
704 TG/OL-AC/DSIAC
2700 D Street, Building 1661
Wright-Patterson AFB, OH 45433-7403
Office: 937.255.3828
DSN: 785.3828
Fax: 937.255.9673

DSIAC COR/PROGRAM MANAGEMENT ANALYST
Emese Horvath
IAC Program Management Office (DTIC-I)
8725 John J. Kingman Road
Fort Belvoir, VA 22060
Office: 571.448.9753

Reference herein to any specific commercial products, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or DSIAC. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or DSIAC and shall not be used for advertising or product endorsement purposes.

Distribution Statement A: Approved for public release; distribution is unlimited.

Copyright © 2020 by the SURVICE Engineering Company.

The DSIAC Journal is a quarterly publication of the Defense Systems Information Analysis Center (DSIAC). DSIAC is a DoD Information Analysis Center (IAC) sponsored by the Defense Technical Information Center (DTIC) with policy oversight provided by the Office of the Under Secretary of Defense (OUSD) for Research and Engineering (R&E). DSIAC is operated by the SURVICE Engineering Company with support from Georgia Tech Research Institute, Texas Research Institute/Austin, and The Johns Hopkins University.

Copyright © 2020 by the SURVICE Engineering Company. This journal was developed by SURVICE under DSIAC contract F8075-14-D-0001. The Government has unlimited free use of and access to this publication and its contents, in both print and electronic versions. Subject to the rights of the Government, this document (print and electronic versions) and the contents contained within it are protected by U.S. copyright law and may not be copied, automated, resold, or redistributed to multiple users or for multiple simultaneous user access to the journal, is desired, please contact DSIAC at 443.360.4600 for written approval.

Download on our website.

Other information products, including reports (SOARs), and many other information products available for free electronic download on our website.

Email: contact@dsiac.org

ISSN 2471-3392 (Print)
ISSN 2471-3406 (Online)
A Computational Approach to Understanding Advanced Thermal Barrier Coatings’ Performance

By Virginia G. DeGiorgi, Edward P. Gorzkowski, Heonjune Ryou, and Stephanie A. Wimmer

Advanced Materials

Many advanced engine system designs result in significant increases in operating temperatures, which, in turn, require new thermal protection coatings (TBCs). In this work, the relationship between material microstructure and performance of ceramic coatings is identified through computational approach. All work used yttria-stabilized zirconia (YSZ) as a model material, but the approach can be applied to many different ceramic materials.
SUMMARY

With the emphasis on lightweighting, composites are being turned to help reduce weight while still maintaining strength and stiffness. However, composites tend to be linear elastic to failure, so there is often no warning of failure (unlike in metallic components). For this research, a pipe section was fabricated from an Inconel 718 liner with a carbon composite overwrap. The pipe was then subjected to increasing internal pressure until failure. The results from this experiment were used to assist in creating and validating a finite-element model of the experiment. The model uses advanced numerical techniques to predict when failure will occur. This article will present the fabrication, testing, and modeling of this effort.

INTRODUCTION

Research into using composites for high-pressure and high-temperature applications has been conducted before [1], but full maturity has not yet been reached. Advances in modeling, testing, and validation are still required before this technology can be fully implemented. One of the issues is that composites can undergo microdamage, leading to variations of the mechanical properties. The microdamage can be the result of high internal pressures and/or exposure to physical trauma. Modeling techniques, such as homogenization, are effective outside these critical regions. However, they break down in regions where the microdamage occurs. In order to account for cracking, the mesh would need to follow the crack, resulting in remeshing and a high computational cost. To address
this issue, the extended finite-element method (XFEM) [2] was developed and has demonstrated a significant advantage over other approaches, such as boundary element methods [3] and remeshing [4]. In XFEM, the finite-element space is enriched with a discontinuous function and near-tip asymptotic functions through the framework of unity partition [5].

A composite tube was made from a candidate high-temperature composite and subjected to a burst test. An initial basic model based on classical laminated plate theory was developed to predict when failure would occur. The results from this experiment were used to assist in creating and validating a finite-element model of the experiment. The model uses XFEM techniques to predict when failure will occur. This article will cover the design, fabrication, burst testing, and modeling of the candidate composite tube.

**HIGH-TEMPERATURE COMPOSITE TUBE**

All recent efforts by Benét Laboratories in composite gun tubes have focused on tank cannons [6]. In that application, the composite is a prestressed jacket over only part of the gun tube, and the temperatures in the composite are significantly lower than in other applications. The materials and fabrication methods used for tank cannons do not directly apply to every situation, as thinner walls cannot support a substantial prestress and the materials themselves cannot handle the operating temperatures seen in every application. Other work at Benét [7, 8] looked at using ceramic liners surrounded by a polymer composite. The need for both hoop and axial prestress coupled with the inherent brittleness of the ceramics made this approach unfeasible.

Material options—high-temperature and high-pressure applications—are carbon/carbon, geopolymer composites, and metal matrix composites (MMC). Both the geopolymer and MMC options needed a thermal barrier coating (TBC) layer to reduce the heat getting into the structural layer and a metallic superalloy liner to act as a gas seal and sliding surface for the round. This concept of a liner, TBC, and structural composite was then patented [9].

For the current effort, the TBC layers were omitted, as this was a structural test and elevated temperatures were not part of the test. This reduced both the cost and manufacturing complexity.

**Tube Fabrication**

Under a Foreign Technology (and Science) Assessment Support program, a composite tube was fabricated at Pyromeral Systems in Barbery, France. (They use their PyroSiC and PyroKarb formulations in a variety of different high-temperature applications, such as Formula 1 exhaust ducts [10].) We selected the PyroKarb resin with IM7 carbon fiber as the reinforcement. The tube was filament wound over an Inconel 718 liner using a wet winding process and cured in an oven. It was then postcured under nitrogen at 704 °C (1300 °F). Figure 1 shows the composite tube being fabricated, and Figure 2 shows it after fabrication.

Originally, the tube was intended for a firing test inside a larger steel tube, so it was required to have a specific inner and outer diameters. This and the limitations of filament winding determined the specific composite layup. The layup selected was [90, (±60, 90), ±60, 90]T. The 90-degree plies are cylindrical windings, with each winding circuit being a single coverage in one direction. Thus, a single down and...
back pass on the winder creates two hoop plies. The ±60-degree layers are helical windings, with each winding layer creating an interwoven double-thickness layer.

**Initial Modeling**

As a starting point, Autodesk Helius Composite 2017 (hereeto referred as Helius) was used to model the layup. Each layer was treated as a unidirectional layer in Helius. The individual layer thickness was determined by physically measuring the tube and its Inconel liner. The composite thickness was divided by 36 to give the average per layer thickness for the model. This resulted in a thickness of 0.686 mm (0.027 in) for the Inconel liner and 0.107 mm (0.004 in) for the composite layers. Exact material properties were not known for the PyroKarb/IM7 material, so IM7/977-2 was used as an analog in Helius. This would provide an overestimation of the tube’s performance but was expected to be in the right ballpark.

Helius used classical laminated plate theory to generate the layup properties and outputs of those properties, as well as the ABD matrix (which relates cross-sectional forces and moments to mid-plane strains and curvatures), which can then be fed into Abaqus for a higher fidelity analysis. The tube/beam analysis option was used in Helius to get an idea of when the tube would fail. The tube’s average measured internal radius of 40.44 mm (1.592 in) was used, and the pressure was increased until first-ply failure was observed based on a maximum stress criterion. Failure was observed at 103 MPa (15 ksi). Helius predicted that the Inconel liner and the first layer would fail at this pressure. The predicted strain at failure can be seen in Figure 3.

**BURST TEST**

**Initial Modeling**

The 1.139-m (44.875-in)-long test specimen, with a smooth Inconel bore surface measuring a nominal 80.5 mm (3.170 in), was provided to the Tube Fatigue Laboratory, Benét Laboratories, Watervliet, NY. The evaluated test specimen had a nominal outer diameter of 90.5 mm (3.563 in) and a nominal wall thickness of 4.4 mm (0.174 in). Four rosette strain gages were placed 90 degrees apart from each other on the exterior surface of the test specimen, centered about its overall length. The gages were Micro-Measurements CEA-06-250UR-350 gages, with grid 3 in the hoop direction.

The entire test fixture assembly was placed on top of riser blocks inside the 3-million-pound press. This test assembly utilized enclosures on the top and bottom of the specimen, with a rubber O-ring and a metal seal in each sealing pocket of the enclosures. The load frame prevents the sealing assemblies and enclosures from exiting the test specimen during pressurization. The combination of the O-ring and metal seal allows sealing between low and high pressures. The O-ring provides low-pressure sealing, as well as the force necessary to drive the metal seal against the enclosure and, in turn, against the sealing pocket of the test specimen. A well-machined surface finish on the metal seal interfaces allows high-pressure sealing. Figure 4 shows the tube in the hydraulic press and a close-up of the strain gage wiring.
Data were collected from the four strain gage rosettes and a pressure transducer. The pressure transducer was calibrated via a deadweight tester for a full-scale pressure of 482 MPa (70 ksi). All 13 channels of data (12 strain and 1 pressure) were recorded by a DeweSoft HD STG-S module. This instrument took raw data at 100 Hz and saved it at 20 Hz.

**Burst Test**

The pressure was ramped up until the specimen was no longer able to hold pressure. There were two audible indications of failure during the test. The first noise was the composite failing just above the midsection, and the second was failure of the liner and composite above it. Upon reviewing the data, it was found that the 12:00 gage had failed during setup, but the other three gages took data throughout the test. The 9:00 strain gage failed at 67 MPa (9,716 psi), and the remaining two gages (6:00 and 3:00) failed at 74 MPa (10,735 psi), roughly 4/10 of a second before the interior Inconel liner failed at 86 MPa (12,483 psi). Figure 5 shows the failed composite tube. It is apparent that it failed along the 60-degree plies and that the hoop plies unwrapped as part of the failure.

### Strain Gage and Pressure Data

The data from the three surviving strain gages and the pressure transducer were analyzed. The strains were corrected for transverse sensitivity, and the principal strains were calculated [11]. The internal pressure can be seen in Figure 6. From the figure, it can be seen that the pressure ramped to about 30 MPa (4.3 ksi), held there for about 15 s, and then rapidly ramped to 86 MPa (12.4 ksi) before final failure. This rapid ramping may have led to premature failure, as it was closer to a dynamic than a static loading. This rapid ramping may be the reason for two failure locations instead of one.

Figures 7–9 show the principal stresses at the three working strain gages. All of the gages failed after the specimen failed. To calculate the maximum principal strains, only data prior to failure were considered. Peak pressure occurred at 27.7 s into the test, so anything after that was considered to be after failure. The time of the maximum principal strain for the 3:00 and 6:00 gages was 27.3 s, with an internal pressure of 74 MPa (10.7 ksi). The 9:00 gage obtained its maximum principal strain slightly earlier at 27.1 s and 67 MPa (9.7 ksi). This most likely corresponded to the first audible noise noticed during the test and would be the initial failure of the composite overwrap. The liner itself then held additional pressure for a short amount of time before failing at 27.7 s. The fact that the 9:00 gage reached its maximum first would imply that failure started near its location, though it could also be the gage coming loose.

Table 1 presents the principal strains for each gage and their angle relative to the axial direction, with pure hoop at 90 degrees. The angles for the maximum principal strains all align with the hoop direction, as would be expected in a cylindrical pressure vessel. Comparing the data in Table 1 to the Helius plot in Figure 4, our measured values are lower in hoop and higher in axial than predicted. We also failed at about 25%
lower in pressure than expected. Given that the Helius data were based on data for a different resin, this could account for the differences.

Table 1 shows basically no strain at the 3:00 gage, as expected, but there are significant shear strains at the 6:00- and 9:00-gage locations. As noted, the maximum values occur earlier at the 9:00 gage, implying that it failed first. Figure 6 shows the failed specimen; it is obvious that it failed along the 60-degree plies. The nonzero shear strain could be indicative of the onset of this failure. The lack of this at the 3:00 location is most likely from it being directly opposite the 9:00 gage.

As the crack propagates, remeshing is required, which will increase computational time and introduce inaccuracies within the solution.

### NUMERICAL MODELING

Using standard finite elements to model scenarios where cracking occurs within a part is typically performed by embedding various crack shapes and sizes into predetermined critical regions of the model. As the crack propagates, remeshing is required, which will increase computational time and introduce inaccuracies within the solution. To address this issue, the extended XFEM [2] was developed and demonstrated a significant advantage over other approaches, such as boundary element methods [3] and remeshing [4]. In XFEM, the finite-element space is enriched through the framework of partition of unity with a discontinuous function and near-tip asymptotic functions [5].

In the XFEM framework, the displacement field in the region around a growing crack is redefined to include terms that account for a crack growing through an element and the stress field seen at the crack tip. This modified displacement field is dubbed, enriched, and applied to nodes within a region \( \Omega_{\text{Enriched}} \), which will evolve with the crack (shown in Figure 10).

The modified displacement field is given in equation (1). The terms \( u_I, a_I, \) and \( b_I \) represent degrees of freedom at node \( I \). \( N_I \) is the shape function for node \( I \). In the equation, the green term highlights the contribution that represents the separation of the element caused from the crack. The blue terms represent the stress field produced by the crack tip.
The H term is a Heaviside function, assigning nodes values of ±1, depending on what side of the crack the node falls on. For this analysis, the simulation was performed using the phantom node approach [12]. In this approach (Figure 11), as the crack passes through the element, a set of phantom nodes (nodes not representing real structure) is initially placed on the elements nodes. The phantom nodes are used to define a cohesive link between two elements coexisting on the original element. As the crack propagates and the cohesive breaks, the element separates. Only the real elements are visible.

In Figure 11(a), the initial crack is shown, with red and green representing the real nodes. In Figure 11(b), the simulation sees two elements—one composed of green and black nodes, where the black nodes are phantom nodes, and one composed on red and magenta nodes, where the magenta nodes are phantom nodes. Figure 11(c) shows what is viewed in the output, with only the green and red nodes visible. The cohesive relationship is shown in Figure 12 [12], where $T_{\text{max}}$ is the maximum traction at the initialization of the crack through the element. For this simulation, the crack initializes when a critical strain is seen at the centroid of the element. $\delta_{\text{max}}$ is the maximum opening of the crack before the element fails.

The $F$ functions are four terms representing the asymptotic solution of the crack tip stress field. For the simulation in this article, these terms are not included—only the separation enrichment terms. The crack tip stress field terms will be implemented in future simulations.

**Burst Test Simulation**

The finite-element simulation was performed using a two-dimensional representation of a quarter cross section of the tube (shown in Figure 13). The simulation was modeled as a static plane strain model using a nonlinear response. The composite was modeled as an orthotropic material using material properties generated from the Autodesk Helius Composite 2017 software. It is assumed there is a rough surface connection between the composite and Inconel. The bottom and left edges of the quarter tube are considered symmetric. The pressure linearly increases on the interior of the tube until failure occurs. To initialize cracking, a critical principal strain on the centroid of the element is set for 3500 με. For the cohesive relationship, the $\delta_{\text{max}}$ is set to 0.41 mm (0.016 in).

As the simulation progresses, cracking begins in the composite at 94.4 MPa (13.7 ksi) (as seen in Figure 14). The crack continues to propagate through the composite until reaching a critical pressure at 100.7 MPa (14.6 ksi), in which case, the simulation fails.
(as seen in Figure 15). The pressure in simulation corresponds with the pressures in the experiment and the Helius software. The principal strains, seen as the simulation failure on the outer surface of the composite, are 3100 με (shown in Figure 16). This is about 11.4% off from the maximum principal strains seen in Table 1 at the 3 and 9 positions. Given the ideal situation the model represents, this error is within reason. The compressive strain along the axis is not captured, as the model is performed using plane strain.

CONCLUSIONS
A composite tube section was fabricated from an Inconel 718 liner with a high-temperature, carbon-fiber composite overwrap. The tube was then subjected to increasing internal pressure until failure. External strain gages and video were used to monitor the test. An initial basic model based on classical laminated plate theory was developed to predict when failure would occur. The results from this experiment were used to assist in creating and validating a finite-element model of the experiment. The model uses advanced numerical techniques to predict when failure will occur. The results from the initial predictions using Helius, the experimentation, and the finite-element simulation using the extended finite-element method all show critical failure between 83 and 103 MPa.
REFERENCES


BIOGRAPHIES

ANDREW LITTLEFIELD is the lead composite engineer at CCDC AC Benét Laboratories, Watervliet, NY, where he works on composite gun tubes and shrouds, electromagnetic railgun launchers, gun barrel vibration absorbers, and simulated proof testing of mortar base plates. Prior to joining CCDC AC, he worked for the U.S. Air Force Research Laboratory on applying composites to spacecraft structures. He received the 2010 Army Science Conference Best Paper Lethality Award and several research and development awards. He has authored several journal and conference papers and technical reports. He has patents for a self-powering prognostic gun tag, an electromagnetic gun launcher, and composite mortar base plate and tube. Dr. Littlefield holds a Ph.D. in mechanical engineering from Rensselaer Polytechnic Institute (RPI).

LUCAS SMITH is the Tube Fatigue Laboratory leader at Benét, where he researches, designs, develops, manufactures, tests, and supports fatigue testing and conducts applied research and development support for all new, modified, and legacy thick-wall pressure vessel assemblies. He has published multiple government reports. He received a patent related to the gun shield of an M60 main battle tank and applied for patents related to his high-pressure work and fatigue-related testing. Mr. Smith holds a B.S. in mechanical engineering from RPI and an M.S. in mechanical engineering from Union Graduate College.

MICHAEL MACRI is a mechanical engineer at Benét, where he leads several research projects in developing advanced numerical algorithms for analysis of artillery systems. He is also the lead instructor for the Armament Graduate School’s finite-element and numerical methods courses. Prior to his work at Benét, he worked for the U.S. Army Research Laboratory as a National Research Council post-doc. His research expertise includes meshless methods, multiscale modeling techniques, numerical fatigue and fracture analysis, and partition of unity enrichment methods, such as XFEM. Dr. Macri holds a Ph.D. in mechanical engineering from RPI.

JOSHUA ROOT is a mechanical engineer at Benét, where he designs, builds, and tests composite gun tubes and associated hardware. His expertise is in working with composite materials and managing complicated systems, designing parts to take advantage of the unique properties of composites, researching nontraditional composite materials, performing various types of analyses on composite materials, testing composite materials at the coupon and part levels, and performing forensic analysis on composite parts that have seen field service. He holds two patents and coauthored the second-place outstanding paper for SAMPE 2004. Mr. Root holds a degree in mechanical engineering from Cedarville University and an M.S. in mechanical engineering from Union Graduate College.
INTRODUCTION

For almost two centuries, the optical community has been using the same method of producing refractive lenses and optically-transparent windows. They start with an optical blank that is cut, ground, and then polished/lapped to form the final optic. While this process is well suited to the simplest optics, those with a combination of planar and spherical curved surfaces, modern demands for high-definition optics are pushing the limitations of this centuries-old technology. Spherical lenses suffer from spherical aberration, which is a kind of blur caused by imperfect focusing.

Modern refractive lenses have aspheric surfaces to correct spherical and other aberrations in the lens system—the surface curvature is not constant across the aperture, as is the case for spherical surfaces. This is usually defined mathematically as the sum of a first-order curvature having a constant radius (i.e., a spherical surface) and a series of aspheric terms (i.e., intentional deviations from the spherical surface). The more aspheric terms in a surface, the higher order the optic. This involves more manufacturing complexity and
costly modern methods of manufacture, such as computer-controlled precision polishing, single-point diamond turning (SPDT), or precision glass molding. Even these methods are somewhat limited in that they require the surfaces have rotational or translational symmetry and are not suitable for asymmetric, or freeform, surfaces. The old traditional means of producing optics cannot support the requirements for new emerging technologies that require higher resolution. These limitations are imposed by multiple factors, such as the cost for higher-order optics and the inability to improve on the existing manufacturing technology [1].

The state of the art in current lens manufacturing technology is a 7th-order lens [2]. Order of an optical lens refers to the complexity of the shape of the lens represented by a polynomial of certain order. The more complex a lens shape is, the higher the order of its aberration correction. Aberration is defined as the degradation of the performance of an optical lens, when the light from one point of an object does not converge into a single point after transmission through the lens. A higher-order aberration is the distortion acquired by a wavefront of light as it passes through the lens with irregularities or inhomogeneities of its refractive index and imperfections in its geometrical shape.

To correct these various imperfections/aberrations in the current technology, various techniques of refracting light are employed. For example, chromatic aberration, or color blur, is imparted by the dispersion or wavelength dependence of refractive index in the material. Similar to how a prism separates light into its constituent colors, a simple lens with chromatic aberration focuses light of different colors at different spots (Figure 1, left), resulting in color blur at the image plane (left inset). A common solution to this problem is the achromatic doublet, which uses two different types of glass—crown (low dispersion, positive focusing power) and flint (high dispersion, negative focusing power) to bring the primary colors into the same focus (Figure 1, right). The refracted spot size for the red-green-blue components is almost identical in size and location in the XYZ coordinate system (right inset).

This type of aberration corrections comes at cost, where different components, materials, and adhesives have to be used to bind them. The first penalty of these corrections is the reduction in optical power transmitted through the optics. The second is the size and weight. Additionally, the use of optical adhesives limits the amount optical power that the optics can transmit. Another major limitation of the current technology is that as the size of the optics grows, the fabrication of the component becomes more difficult.

Within the last decade, additive manufacturing (AM) has been successfully implemented in producing products from polymers, ceramic powders, and metals. AM is capable of producing products with intricate details on the micrometer (µm) scale. To harness the power of this new manufacturing technique, the topic...
Within the last decade, AM has been successfully implemented in producing products from polymers, ceramic powders, and metals.

“Additive Manufacturing of Inorganic Transparent Materials for Advanced Optics” was proposed and funded by the U.S. Navy Small Business Innovation Research/Small Business Technology Transfer (SBIR/STTR) Office to address the current limitation of processing optical components [1].

The sweeping intentions of this AM topic are twofold. One is to develop a technique to perform front surface repair on existing optical components. The other is to develop a manufacturing technique for producing high-definition optical components that can get beyond the 7th-order limitation of current technology. The combined goal is to produce self-contained solutions that can solve the Navy’s needs of extending the life of existing optical products and rapidly develop new optical components with higher fidelity in controlling the optical wavefronts.

Naval air operations have a broad array of weapon and surveillance systems that utilize high-performance optics. Many of these applications require greater wavelength transmission range, hardness, and temperature stability. The ability to print glass lenses using AM will provide the Navy the ability to (1) deposit net-shape or near net-shape, free-form optics; (2) locally adjust the index of refraction and other optical properties, such as dispersion; (3) create high-precision, low-thermal expansion meteorological frames which can form the basis for refractive optics; and (4) repair existing optical systems [1].

The benefits of AM are widely realized for structural systems. However, the work on printing optical components is still in its comparative nascency. The majority of the AM effort in optics has primarily focused on polymers. Processes that have been demonstrated for printing optically-transparent polymers include inkjet printing [3] with/without in-situ ultraviolet (UV) curing [4] and multiphoton stereolithography to directly polymerize resin [5, 6]. These techniques have been used for rapid prototyping of nonimaging optics using polymethyl methacrylate and similar plastics [5].

This inorganic AM topic uses a 3-inch achromatic lens as its demonstration vehicle for producing an AM process that can be used to manufacture an array of naval optical components with varying geometries. Demonstration of the AM process should include the effectiveness of fabricating fully-densified optical components with precision control of the component’s geometry and smooth surface quality. One of the key metrics that will be used to track the progress of different AM processes is the variation in the index across the surface [%n] of the lens. Additionally, standard optical metrological processes will be used to examine the finished achromat lens [1].

AM FOR REPAIRS AND REFURBISHMENT IN THE FIELD

In order to extend the life of deployed, fielded instruments, an AM process is needed to repair damages to the front optics of these systems. Damages to these optical systems could be scratches, fractured windows, pitting, etc., to the front surface optics. These damages are most prevalent in instruments used by dismounted Warfighters exposed to harsh environments. Instrument reliability and replacement costs are issues commonly raised by product managers of Soldier-borne systems. Additionally, complex optical systems installed on ground-based mobile platforms, surface and subsurface naval vehicles, and fixed- and rotary-wing aircraft are subjected to diverse environmental hazards, including sand and wind erosion, salt spray, mechanical and thermal shock, and vibrations. Damaged or inoperable optical trains currently need the entire sensor module replaced. A robust optical AM platform would enable rapid refurbishment and redeployment of sensors while reducing the logistical burden of maintaining replacement part inventories.

One promising technique in development under this effort is the filament-fed, laser-heated (FFLH) AM process. FFLH is a layer-by-layer technique where a rod or filament of the lens material is fed to the deposition zone, where it is softened by a focused laser beam and fused to the layer below as it is printed (Figure 2, left). This was developed by researchers at Missouri University of Science and Technology (MS&T) [7]. The filament diameter (currently between 100 µm and 3 mm) limits the overall resolution of the deposition, which can be compensated somewhat with careful design of the process parameters and toolpath mapping. FFLH has been used to demonstrate homogeneous glass plates with good transparency and lenses with spherical surfaces and diameters between 4 mm and 10 mm (Figure 2, right). In addition to manufacturing lenses with freeform...
surfaces, FFLH is a promising technique for repairing glass parts with physical damage (scratch/dig) if the resolution can be improved and suitable filaments are developed.

Product Innovations and Engineering (PINE, Inc.) of St. James, MO, creates innovative solutions to transform the AM industry and is partnering with Dr. E. C. Kinzel at the University of Notre Dame to further improve the FFLH process. They have shown that the FFLH process is capable of creating fully-dense, optically-transparent objects, including lenses, freely-supported structures, waveguides, and microfluidic networks.

**AM FOR ADVANCED OPTICS**

A multimaterial, precision AM optical platform would be a game changer for advanced refractive optical components based on freeform surfaces and graded index (GRIN) optics. These technologies bring increased control over the transmission and manipulation of the optical wavefront. They offer the potential for very high-performance optical designs (e.g., high image quality, wide field of view, etc.) in compact forms that would benefit platforms where size and weight are a premium, such as space-based platforms, small unmanned aerial vehicles, and the dismounted Warfighter. Optical systems based on these technologies have been historically difficult to implement with current technologies [8] and, thus, their potential has not yet been fully realized.

Conventional refractive optics with spherical and even aspherical surfaces have rotational symmetry that make them compatible with modern optical surface generating methods (e.g., SPDT). Freeform optics, which lack rotational or translational symmetry, are recently being used in developing high-quality optical systems. The spherical or low-order aspheric surfaces in conventional lenses and mirrors have a simple shape and thus are limited in their potential light-beam paths. Lenses and mirrors with more complex freeform surfaces, which are mathematically described by Zernike polynomials, are needed for many advanced applications.

Traditional 2 degrees of freedom manufacturing processes, such as grinding, polishing, and ultraprecision turning, are incapable of manufacturing freeform surfaces. Advanced methods (ion-beam figuring and five-axis diamond milling) are required to realize these surfaces. A typical lens element contributes up to 15 aberrations that must be corrected in the optical assembly by additional lens elements and aperture stops. Using a freeform shape for the surface offers an opportunity to more directly correct the limiting aberrations and reduce the number of lens elements. The additive generation of three-dimensional (3-D) optical surfaces in glass via AM is an opportunity to bolster the manufacturing and realization of freeform optics for Navy applications.

So far, we have only considered the role of optical surfaces in manipulating the optical wavefront, as the body of a homogeneous lens (ideally) does not alter its trajectory. In addition to refraction at the front and rear surfaces, GRIN optics have an internal refractive index distribution that bends light within the body of the lens. Distributing the internal GRIN profile offers the optical designer an additional degree of freedom roughly equivalent to having an additional surface.

GRIN lenses have been commercially available for some time, although the GRIN profiles have been restricted in size and shape of the index profile to those easily generated by diffusion processes. Ion-exchange diffusion through a glass rod has been used to generate purely radial gradients (e.g., SELFOC by GoFoton). Diffusion between stacked glass plates has been used to generate purely axial gradients (Gradium by LightPath). But the long time for diffusion, typically measured in months, limits the diameters of radial...
GRIN optics to several millimeters, with roughly parabolic index profiles. The limiting factors for the widespread adoption of GRIN in advanced optical systems are manufacturing difficulty and the constraints on the GRIN profiles that can be obtained by current methods [8].

Based on inkjet printing technology, the Volumetric Index of Refraction Gradient Optics (VIRGO) process from Voxtel, Inc., is capable of 3-D printing of lenses using two or more “inks” having different optical properties (refractive index and dispersion). This process opens the possibility to creating GRIN optics with gradients unrestrained by previous techniques. It utilizes inks comprised of optical material nanoparticles deposited onto a substrate and dispersed in a solvent voxel by voxel, layer by layer [8]. The solvent is subsequently removed, and the optic is densified under appropriate conditions to form a solid, transparent lens free of voids or pores. By overlapping multiple inks with different optical properties and allowing small-scale diffusion, optics with arbitrary, volumetric GRIN profiles can be realized, can compensate for geometric and chromatic aberrations, and can provide optical power.

If we think of GRIN from a design perspective as having an additional optical surface, 3-D GRIN represents an additional high-order, freeform surface inside the optic. Additionally, since the diffusion distances are small (~100 µm) compared to those used in other methods, the processing time is fast and scalable. Currently, the medium-scale VIRGO platform can produce organic/inorganic optics within the 1 m x 1.5 m build area, including homogenous optics, GRIN lenses, phase corrector plates, and lens array plates (as shown in Figure 3). The process is compatible with other important wavelength regimes, including UV and infrared (IR), by exploiting the appropriate materials for inks.

Voxtel and the Institute for Optics at the University of Rochester (UR) are developing an AM method to create glass gradient refractive index (GRIN) lenses. This research is an extension of Voxtel’s existing organic/inorganic VIRGO AM platform (Figure 3, left), which can simultaneously print six different inks in a 1 m x 1.5 m build area, up to 60 mm tall, at a production rate of 100 optics (25-mm diameter) per hour. The process has been used to demonstrate 4-inch-diameter GRIN lenses and phase corrector plates (Figure 3, center) as well as lenslet arrays (Figure 3, right). The UR glass melt technology will melt green state Voxtel lens samples to deliver high-performance optics.

The nanoparticle electrospray laser deposition (NELD) technology developed by researchers at the Center for Research and Education in Optics and Lasers (CREOL) at the University of Central Florida (UCF) [9] is also designed to 3-D print an optic layer by layer using suspended nanoparticles. The NELD process is a droplet-on-demand delivery system where a syringe pump delivers microdroplets of a suspension containing optical nanoparticles onto a substrate while a laser beam focuses onto the droplet, boiling the delivery fluid and fusing the nanoparticles together (Figure 4). A hollow laser beam is
focused onto the deposition site to boil away the delivery fluid (typically water) and fuse the nanoparticles to each other and the layer below, in-situ as the part is printed (Figure 4, inset). The spherical shape of the microdroplet focuses laser energy as it passes through the hollow laser beam, thereby generating heating within the microdroplet before it impacts the surface. So far, the process has been shown to precisely deposit subwavelength optical structures and is adaptable to different lens materials through carefully selecting the nanoparticle species, delivery fluid, and laser wavelength.

IRflex Corporation and the CREOL are teaming up to design and develop the NELD AM process and demonstrate its capabilities with an athermalized achromatic (achrothermic) doublet lens. IRflex has considerable knowledge and expertise in specialized glass fabrication and the scientific and modeling capabilities to design the achothermic doublet with the desired specifications. The NELD approach will be modified by CREOL to fabricate high-performance optics based on silica glass. While the NELD process is being developed using a single ink, in principle, it extends to printing multiple inks for printing 3-D GRIN optics and freeform surfaces.

CONCLUSIONS

Extending these optical AM methods to infrared systems, particularly those used for thermal imaging in the mid-wave infrared (3–5 µm) and long-wave infrared (8–12 µm) wavebands, is promising but poses another set of challenges. The materials do not have significant IR transmission beyond about 1 µm, and IR refractive materials must be adapted for use in AM. IR lenses are typically crystals (e.g., silicon, CaF2, germanium, ZnSe, ZnS, and GaAs) or chalcogenide glasses (e.g., amorphous material transmitting infrared radiation and Schott/Vitron infrared glasses series). Deposition and densification of IR crystalline nanoparticles into high-quality, transparent optical ceramics is a topic of ongoing research and potentially compatible with the AM methods discussed here. Chalcogenide glasses offer a wide range of optical properties that can be controlled via the quantity of their chemical constituents [10]. However, when heated to processing temperatures, their chemistries and, hence, optical properties can change due to high vapor pressures and reactivity with water and oxygen.

Advanced infrared optics via AM, particularly IR-GRIN and freeform surfaces, can significantly reduce the size and weight of imaging and threat-warning sensors, enabling dual-band common aperture sensors and sensors with conformal apertures. In the IR wavebands, where there is little diversity in existing lens materials, chromatic...
The potential of AM for field repair of optical systems on demand will dramatically enhance the logistics and maintenance of Navy platforms and associated sensors across all Services. Aberration is a key challenge and often requires numerous lens elements. This is compounded when the system must operate in multiple wavebands.

The primary advantage of GRIN for IR is to utilize the additional degree of freedom it provides to correct chromatic aberration, thereby reducing the number of lens elements, overall system size, and shifting some of the system complexity to the lens element level for relaxed assembly tolerances and improved reliability [2, 8, 10]. To date, GRIN profiles in infrared materials have been limited to axial and hybrid axial-radial configurations due to diffusion-based fabrication techniques. A true AM approach would enable gradients with 3-D precision and advanced capabilities to platforms for all the Services.

Optical quality AM of inorganic glasses opens up new opportunities for engineering optics with volumetrically varying properties, such as GRIN lenses, as well as homogeneous lenses with advances freeform surfaces, unconstrained by the limitations of conventional lens manufacturing methods. The potential of AM for field repair of optical systems on demand will dramatically enhance the logistics and maintenance of Navy platforms and associated sensors across all Services.

ACKNOWLEDGMENTS

In writing this article, many people have made this effort significantly easier. I want to thank Dan, my co-author, who stepped in the last minute and added significant clarity to this article and my understanding of freeform optics. I want to also thank IRFlex, Voxel, PINE, and their associated universities (UCF, UR, and Notre Dame) for their contributions; Kishan Goel (retired) in creating this STTR; Petra Braonthoover for her tireless efforts in working with me to get this topic out; and Naval Air Systems Command’s SBIR/STTR Office for allowing me to expand the Navy’s portfolio in optical technology.

On a personal note, I would like to thank Jill Tyler Hurst for donating her heart to me last September. She was an amazing young woman who saved the lives of five people with her generous gifts. I want to let her loved ones know that I am eternally grateful for her gift. I want to thank the Jacksonville Mayo Clinic cardiac team for saving my life. They are truly skilled, compassionate, and caring professionals. I would also like to thank my wife, my kids, my mom, my brothers, and family and friends for getting me through this challenging period.

REFERENCES


BIographies

CHANDRAIKA (JOHN) SUGRIM is an electronic engineer for the U.S. Naval Air Warfare Commander Fleet Readiness Center in Jacksonville, FL. His research interests include electro-optical/IR systems design, composite laser-induced fluorescence detection, and polarization switcher optics. He has received several Engineer of the Year awards from the Navy; written manuals and publications; and presented at symposiums. He has four granted patents and five nonprovisional patent applications and was awarded a Professional Engineering license. Mr. Sugrims holds a bachelor’s degree in electrical engineering and an M.S. in physics from State University of New York at Stony Brook, an M.S. in risk management from Florida State University, and an M.S. in electrical engineering (electronic warfare) from the Navy Post Graduate School.

DANIEL GIBSON is a research scientist in the Optical Materials and Devices Branch at the U.S. Naval Research Laboratory in Washington, DC, where he has been developing infrared optical materials and fibers since 2003. He invented diffusion-based GRIN IR glass lenses and thermally-bonded IR doublet lenses that obviate the need for IR-transparent adhesives. He also developed technologies to fabricate novel infrared optical fibers, including hollow waveguides, single crystal fibers, heavy metal oxide and chalcogenide glass fibers with solid core, photonic crystal, photonic bandgap, and negative curvature structures. Dr. Gibson holds over a dozen patents on his technologies, several of which have been licensed to industry, and has served as principle investigator on multiple internally- and externally-funded research and development programs. Dr. Gibson holds a Ph.D. in ceramic and materials engineering from Rutgers University.
CAN COMPRESSIVE SENSING Solve Your Sensor and Measurement Problems?

By Michael Don

SUMMARY

Compressive sensing (CS) is a relatively new field that has generated a great deal of excitement in the signal-processing community. Research has applied CS to many forms of measurement, including radio detection and ranging (RADAR), light detection and ranging (LIDAR), magnetic resonance imaging (MRI), hyperspectral imaging, high-speed imaging, X-ray tomography, and electron microscopy. Benefits range from increased resolution and measurement speed to decreased power consumption and memory usage. CS has received mixed reviews in commercial and government circles. Some have touted CS as a cure-all that can be “thrown” at any sensor problem. Others consider CS all hype—just a rebranding of old theories. Who is right?

In order to answer this question, an overview of CS is presented, clarifying common misconceptions. Case studies are brought to illustrate the advantages and disadvantages of applying CS to various sensor problems. Guidelines are extracted from these case studies, allowing the readers to answer for themselves, “Can CS solve my sensor and measurement problems?”
INTRODUCTION

At first look, it appears that CS is widespread and revolutionizing a variety of sensor systems. Four new Institute of Electrical and Electronics Engineers (IEEE) paper classification categories have been created specifically for CS. Thousands of research papers have been published [1], representing significant academic funding by government and industry. The Wikipedia entry on CS [2] declares that “compressed sensing is used in a mobile phone camera sensor” and “commercial shortwave-infrared cameras based upon compressed sensing are available.” The infrared (IR) camera refers to InView’s single-pixel camera, a highly-publicized, real-world example of compressive sensing in action [3]. The MIT Technology Review article “Why Compressive Sensing Will Change the World” [4] explains how CS has supplanted the Nyquist-Shannon sampling theorem, a foundation in signal processing during the last century, and that CS is “going to have big implications for all kinds of measurements” [4].

A closer look, however, reveals some doubts about CS. There are anonymous comments, such as “Most of it seems to be linear interpolation, rebranded” [5] or “Compressed sensing... was overhyped” [6]. There are researchers like Yoram Bresler, a professor at the University of Illinois, that claim that CS is not really new. He asks, “Would a rose by any other name smell as sweet?,” claiming that CS is just a new name for earlier techniques, such as image compression on the fly and blind spectrum sampling [7].

Some CS researchers acknowledge shortcomings. Thomas Strohmer, a professor and the University of California Davis, asks, “Is compressive sensing overrated?” and notes that “the construction of compressive sensing based hardware is still a great challenge” [8]. Simon Foucart, a professor of mathematics at Texas A&M University, describes how “projects to build practical systems foundered...” and that “...compressed sensing has not had the technological impact that its strongest proponents anticipated” [9].

When asked what was holding CS back from imaging applications, Mark Neifeld, a professor at the University of Arizona, answered that “we haven’t discovered the ‘killer’ application yet” [10].

There are other scholars that are openly critical of CS. Leonid Yaroslavsky, a professor at Tel Aviv University and an Optical Society fellow, writes, “Assertions that CS methods enable large reduction in the sampling costs and surpass the traditional limits of sampling theory are quite exaggerated, misleading and grounded in misinterpretation of the sampling theory” [11]. In a section of his website titled “Fads and Fallacies in Image Processing,” Kieran Larkin, an independent researcher with 4,274 paper citations, declares that “everyone knew that the single-pixel camera research was a failure” [12]. He is referring to the InView single-pixel camera previously mentioned that was purported to be a successful application of CS. Who is right? Will CS bring about a revolution in sensing and measurement or is it really just all hype?

There are many tutorials [13–15], review papers [16–18], and articles [4, 19, 20] on CS, but they tend to be too technical or general for many readers. The technical sources are inaccessible to those without a signal processing background, while the nontechnical sources are too vague to give an intelligent perspective on CS. They do not address specific criticisms, leaving the readers on their own to judge between the supporters and detractors of CS. They are also generally written by CS researchers that may justly or unjustly be suspected of bias.

This article seeks to provide an accessible explanation of CS that gives enough background to examine claims and criticisms. In an effort to make CS understandable to the layman, concepts have been simplified, details have been glossed over, and equations have been replaced by intuitive explanations. For a more in-depth treatment of CS, several tutorials provide a good starting point [13–15].

TRADITIONAL SAMPLING

Overview

CS is sometimes referred to as compressive sampling since it is the sampling process that lies at the heart of CS. Sampling transforms continuous analog signals into discrete digital values that can be processed by a computer. In this age of low-cost computing, virtually all sensor systems sample signals, from commercial audio and video equipment to specialized medical and military systems. The speed or resolution at which a signal is sampled is called the sampling rate. Samples are typically taken at regular intervals, and the sampling rate determines the size of the features that can be identified in a signal. The more samples taken, i.e., the higher the sampling rate, the smaller the features that can be identified. When designing any system that samples data, the sampling rate must be carefully considered. If it is too high,
the extra samples can increase power consumption, memory usage, computing complexity, and cost for little or no gain in performance. If it is too low, important information is lost, degrading performance, or even making the system unusable.

**Nyquist-Shannon Sampling Theorem**

In order to appreciate CS, we must first explain traditional sampling in some detail. The top left plot of Figure 1 shows a 1-Hz sinusoidal signal, i.e., there is one cycle per second. Such a signal could represent many different types of physical phenomenon, such as a voltage oscillating over time. The middle left plot shows a 10-Hz signal, i.e., 10 cycles per second. The bottom left plot shows the summation of these two signals. The top plot on the right shows this summed signal sampled at 10 Hz, i.e., 10 samples per second. This is an example of traditional sampling.

An analog-to-digital converter (ADC) would sample the signal at regular intervals 30x over the 3-s period to produce the samples marked as red dots. These samples are connected by a green line in an attempt to reconstruct the original signal. The reconstruction completely misses the 10-Hz signal, creating a waveform similar to the original 1-Hz signal. Clearly, the 10-Hz sample rate is too slow to detect the 10-Hz signal. The middle right plot shows the same signal sampled at 20 Hz, successfully capturing the 10-Hz signal.

The Nyquist-Shannon sampling theorem states that the sampling rate must be at least twice the highest frequency of a signal. This intuitively makes sense. In order to detect the important features of a signal, there needs to be at least one sample per feature. The important features of a sinusoid can be viewed as the valleys and peaks of each cycle.

Therefore, two samples are needed per cycle. In this case, a 20-Hz sampling rate allows us to sample at the valley and peak of each cycle of the 10-Hz signal. The bottom right plot shows the boxed detail from the plot above it. The green lines connecting the samples give a good approximation of the signal but do not perfectly match the original waveform in blue. However, we will see that the signal can be perfectly reconstructed using sampling theory.

The idea of representing signals in different bases will be an important concept in CS.

The signal in Figure 1 is relatively simple, just a combination of two sinusoids. A plot of the spectrum of this signal is shown in the top left plot of Figure 2. Instead of viewing the signal as a voltage oscillating in the time domain, the plot shows the amplitudes of the sinusoids that make up the signal in the frequency domain. The frequency domain is commonly referred to as the Fourier basis. The normal domain in which we typically view the signal, in this case, the time domain, is commonly called the standard basis. The idea of representing signals in different bases will be an important concept in CS that will be revisited later. The bottom left plot shows a more complicated spectrum with 14 nonzero values, corresponding to 14 sinusoids in the time domain that are summed to produce the signal in the top right plot. This signal is sampled above the Nyquist rate and reconstructed using sampling theory as the dashed red line in the bottom right plot, perfectly matching the original signal in blue. This is a remarkable result. A continuous analog signal composed of many frequencies that is sampled at the Nyquist rate can be perfectly reconstructed.
Two-Dimensional (2-D) Sampling

Figure 3 gives an example of 2-D sampling. In the case of one-dimensional (1-D) signals, one sensor (the ADC) samples at regular time intervals. In imaging, multiple sensors are typically spaced at regular intervals to create a 2-D sensor array. Just as the sampling rate determined the smallest detectable feature for the 1-D signal, the resolution of the 2-D array determines the smallest detectable features in the image. The relatively high-resolution image on the left is 1024 x 1024 pixels. The middle image shows the magnified detail of an airplane, with an ARL logo clearly recognizable. The right image has a resolution of 512 x 512, where the logo is now unrecognizable. The line width (i.e., feature size) of the letters was 1 pixel for the 1024 x 1024 array. When we decrease the resolution to less than 1 pixel per feature, those features become unrecognizable.

A Scanning Single-Pixel Camera

Images are typically sampled using a 2-D sensor array, but there are other methods as well. A single sensor could scan the scene, pixel by pixel, row by row, to acquire a complete 1024 x 1024 image. Figure 4 shows an example of such a system. A lens projects a scene onto a digital micromirror device (DMD), a 2-D array of tiny mirrors. Each mirror can be independently controlled to reflect light toward or away from a single sensor. The DMD steps through all of its mirrors, reflecting the light from one mirror toward the detector, while all of the other mirrors reflect light away from the detector. At each step, only the part of the scene reflected by the single mirror is seen by the sensor, capturing the single pixel of the image corresponding to that mirror. After all the pixels have been collected, they can be arranged to form a 2-D image of the scene. The image will have the same resolution as the DMD, e.g., a 1024 x 1024 DMD can create a 1024 x 1024 pixel image.

In essence, the traditional imager with a 2-D sensor array has been replaced by a 2-D mirror array. This process may seem excessively complicated for the visible spectrum, where high-resolution 2-D sensors arrays are inexpensive; but for expensive IR sensors, this system might be viable. If the high-resolution DMD is less expensive than a high-resolution IR sensor array, a low-cost IR camera can be created using a DMD and a cheap, single-pixel IR sensor.

Compressive Sensing, Single-Pixel Camera

There is another important factor besides the relative cost of the DMD that will determine the practicality of this

![1024x1024](image1.png)
![1024x1024](image2.png)
![512x512](image3.png)

Figure 3: A 2-D Sampling Example (Source: CCDC ARL).
single-pixel camera—the measurement speed. A 2-D sensor array can capture an entire image in one snapshot. A single-pixel camera with a 1024 x 1024 DMD has to step through all 1 million mirrors to take a picture. This will take some time, even if the mirrors can move very fast. Is there any way to speed up the measurements without significantly affecting the image quality? A 512 x 512 array would have 4X less pixels and be 4x faster; but as we saw in Figure 3, this will degrade the image quality. This is based on the Nyquist-Shannon sampling theorem, which, in essence, says that at least one measurement is needed per feature. Using traditional sampling, there is no way to avoid the sampling theorem, but we can circumvent it using compressive sampling.

The compressed sample is a randomly weighted sum of the entire signal. In this case, instead of using one mirror at a time, we can use a random pattern of multiple mirrors simultaneously to reflect a random pattern of the scene onto the sensor (illustrated in Figure 5). Each column represents the process of taking one compressed sample. The top row is the scene, in this case, a simple “L” that remains the same for each measurement. The low-resolution, 8 x 8 image and DMD are only used here for illustration; real applications would use higher resolutions. The next row shows the weights produced by the DMD, which changes for each measurement. The white squares represent the mirrors reflecting light toward the sensor, effectively weighting (i.e., multiplying) the light by 1. The black squares represent the mirrors reflecting light away from the sensor, effectively weighting the light by 0. The “L” of the image is outlined to illustrate how the weights overlap the scene. The next row shows the weighted scene. This is the light seen by the sensor, which is the product of the scene and the DMD pattern.

The bottom plot shows the compressed sample produced by each column. For each measurement, the light of the weighted scene is focused on the single sensor, which sums all of the light and measures its intensity, producing a compressed sample. If the value of each black and white pixel of the weighted scene is 0 and 1, respectively, then the value of the compressed samples will be the number of white pixels in the weighted scene. The figure shows five example DMD patterns that produce five compressed samples. The samples themselves do not resemble the scene. But once enough samples have been taken, they can be processed...
by an optimization algorithm to recover an image of the scene.

Figure 6 shows the image from Figure 3 that used 512 x 512 = 262,144 pixels on the left compared to a simulated CS recovered image using 262,144 compressed samples on the right. This used the same method of taking compressed samples, as shown in Figure 5, except a 1024 x 1024 DMD was set to 262,144 different random patterns, producing 262,144 compressed samples. An algorithm processed these samples to produce a 1024 x 1024 image, with the “ARL” clearly visible. We have achieved a resolution of 1024 x 1024 using 4x less measurements, beating the Nyquist-Shannon sampling rate! This example illustrates the essential components of CS. Compressed samples are formed as randomly-weighted sums of the signal, typically requiring some form of specialized hardware. These samples are then postprocessed to reconstruct the original signal using fewer samples than predicted by traditional sampling theory.

**Sparsity**

The Nyquist-Shannon sampling theorem states that the sampling rate must be at least twice the highest frequency of the signal—equivalently, the pixel width cannot be larger than the smallest feature of the scene. The sampling rate depends on the highest frequency of the signal. In CS, the number of measurements required depends on the sparsity of the signal, not its highest frequency. The sparsity of a signal is inversely proportional to its number of nonzero elements. Sparse signals are mostly close to 0, with a relatively few number of larger nonzero values. Remarkably, CS works well as long as the signal is sparse in any basis, not just the standard basis. For example, the 1 + 10 Hz signal in Figure 1 shown in the time domain is not sparse, i.e., most of the values are not 0. However, when the signal is represented in the Fourier basis (frequency domain), as in the top left of Figure 2, there are only two nonzero values, making it very sparse. Many naturally-occurring signals are sparse in some basis, allowing CS to be applied to a variety of applications. CS worked in the airport image because images are generally sparse in bases such as the 2-D Fourier or 2-D wavelet.

The left plot of Figure 7 shows a normal image in the standard basis. The middle plot shows this image in the wavelet basis, with the amplitude shown in a log scale for emphasis. The right plot compares the sparsity of the image in the standard and wavelet. For this 256 x 256 image, all 65,536 pixel values were arranged in descending order to create the line labeled “Original.” The same was done for the image in the wavelet basis and labeled “Wavelet.” Just like the 1 + 10 Hz signal, even though the image is not sparse in the standard basis, it is sparse in the wavelet basis. The important point is that the sparser the signal, the better CS works, i.e., the signal can be reconstructed with less measurements. Higher-dimensional signals, such as three-dimensional (3-D) images, are typically very sparse, making them ideal candidates for CS.

**Incoherence**

In the single-pixel camera example, the DMD was used to produce randomly-weighted sums of the data. The weights do not have to be random; they just have to be incoherent with the sparse basis of the data. Incoherence can be thought of as maximally different. For example, when using a Fourier basis, the pattern of weights should be as different as possible from the sinusoids of the Fourier basis. It happens to be that completely random weights are incoherent to any basis, but pseudorandom or structured weights can also be used in CS. Using CS theory, these weights can be optimized.
to achieve maximum performance [21], but they also have to be realizable in hardware. For example, a DMD cannot produce arbitrary valued weights. The mirrors can only point two directions, away or toward the sensor, resulting in weights of 0 or 1. Depending on the hardware used to implement the CS weights, the weight values may have limitations that affect CS performance.

**Data Reconstruction**

A detailed description of CS reconstruction algorithms is outside the scope of this article, but there are a few relevant points to mention. In traditional sampling, a signal can be perfectly reconstructed if it is sampled above the Nyquist rate. Similarly, in CS, a signal can be perfectly reconstructed if there are enough measurements relative to its sparsity. Real signals are not perfectly sparse, i.e., many of the values will be close to 0 but not actually 0. If these values are small enough, however, they will have minimal impact on CS performance.

The distinction between signal noise and measurement noise is important in CS. Measurement noise is created in the measurement process, e.g., electronic noise in the sensor. Signal noise is present in the signal being measured before it reaches the sensor, e.g., external interference. Signal noise is frequently measured as a signal-to-noise ratio (SNR), where a low SNR indicates a noisy signal. CS performs well in the presence of measurement noise. Reconstruction remains stable, with the quality of the reconstruction proportional to the noise level. CS performs poorly, however, when the signal has a low SNR [22]. Unfortunately, the CS process itself can degrade the SNR. In order to reconstruct the original signal, the random weights used to make the compressed measurements must be known. For example, the states of the DMD in the single-pixel camera must be known for each measurement and used in the reconstruction algorithm. In an ideal case, this information is known perfectly and is not a source of error. In practice, movement or miscalibration will introduce error, decreasing the signal’s SNR and hampering CS.

The number of measurements needed to produce an accurate reconstruction is proportional to the sparsity of the data. However, the exact sparsity of the data is not known a priori. Therefore, practical systems have to be designed for worst-case scenarios, increasing the required number of measurements. Another important consideration is that the optimization algorithms that reconstruct the data are very computationally intensive. Progress has been made to accelerate these algorithms [23], but they can still hamper real-time applications or systems with limited computing resources.

**Compressive Sensing vs. Compression**

CS is related to compression. When data is compressed, a large quantity of data is represented by a smaller amount of data. Typically, the full, uncompressed data is acquired first before a compression algorithm reduces it to a manageable size. For example, in traditional imaging, high-resolution imagers capture every pixel of an image and then throw away most of the data as it is compressed into a JPEG format. On the other hand, CS uses specialized hardware to compress the data at the time of measurement. Only the compressed measurements are saved; nothing is thrown away. From this viewpoint, CS is more efficient than...
typical imaging practices. CS can be used directly on data as a compression algorithm, independent of any sensor system. But traditional compression algorithms generally perform better than CS when the full data set is already available.

**Compressive Sensing vs. Inpainting**

In some cases, CS can be confused with inpainting [24]. Typically, inpainting refers to filling in the missing samples of an image, but it can also refer to filling in the missing samples of other types of data. Inpainting works with traditional samples, it does not use weighted sums of the data. Missing samples can occur accidently due to occlusions, noise, or damage; or nonuniform sampling can be done on purpose. Inpainting is confused with CS because CS reconstruction algorithms can also be used for inpainting. The main result of CS—that sparse data can be perfectly reconstructed using a small number of compressed samples—does not apply to inpainting. This is why CS is a much more powerful tool than nonuniform sampling for enhancing sensor systems.

**CASE STUDIES**

**Single-Pixel Camera**

The single-pixel camera can be viewed as an application of CS to increase measurement speed or, alternatively, image resolution. A scanning camera that measures 1 pixel at a time would require 1024 x 1024 measurements to create a 1024 x 1024 image. A CS architecture using 4x less measurements would decrease the acquisition time by a factor of 4. Alternatively, a scanning camera using a 512 x 512 DMD would take the same amount of time to acquire an image as a 1024 x 1024 CS camera. However, the CS camera would have a resolution 4x higher than the scanning camera. Although CS can significantly increase the measurement speed or resolution of a scanning camera, there are several reasons why the single-pixel camera might not be a commercial success. They are as follows:

- Even if CS can increase measurement speed, it may not increase it enough to be practical for many applications.
- Given the long measurement time, camera or subject motion may introduce noise that impacts the reconstruction results.
- The effectiveness of CS is related to the sparsity of the data. Even though 2-D images are generally sparse, they may not be sparse enough to make this application practical.
- The cost of traditional short-wave IR (SWIR) cameras has been decreasing [25]. In addition, the DMD mirrors are limited to the near IR and SWIR range, preventing application to mid-wave IR, long-wave IR (LWIR), and far IR (FIR) imaging. These factors limit the marketability of a DMD-based solution.

InView has been developing technology to address some of these problems, such as using multiple sensors to decrease measurement time [26] and hyperspectral cameras that target sparser 3-D data sets [27]. But at this time, it appears that InView has not made large inroads into the IR imaging market.

**Cell Phone Camera**

Another CS imaging application is in low-powered, complementary metal-oxide semiconductor (CMOS) imagers [28]. Instead of using CS for measurement speed or resolution, this application focuses on reducing power. Each pixel in a typical 2-D imaging array is made up of a light sensor that produces a voltage and an ADC that converts the analog voltage to a digital value. This is illustrated in Figure 8 on the left with an example 4 x 4 imager. For a 1024 x 1024 sensor, each image requires about 1 million analog to digital (A/D) conversions. Multiplying that by the number of images needed for a video results in significant power usage, especially for mobile devices with limited battery life.

CS can be used to reduce power consumption by reducing the number of A/D conversions required for each image. Compressed samples are produced by connecting each ADC to a random pattern of sensors, requiring less A/D conversions per image (as illustrated in Figure 8 on the right). The problem with this approach is the image reconstruction. CS image reconstruction takes time and computing resources, probably using more power than initially saved. This CS imager would only be

![Figure 8: Comparison of a Typical CMOS Imager and CS Low-Power CMOS Imager (Source: CCDC ARL).](Image)
useful in a niche application where the video is not needed in real time and can be reconstructed in postprocessing using powerful computers. It would certainly be undesirable as a cell phone camera.

**Magnetic Resonance Imaging**

The only truly successful commercial application of CS is possibly the MRI [29, 30]. MRI detects radio frequency emissions from tissue excited by magnetic fields. Due to the physics of the system, measurement takes place in the Fourier basis, requiring a basis change back to the standard basis to retrieve the image. Figure 9 shows a simple example of a 2-D MRI, which is similar to the $1 + 10$ Hz, 1-D signals in Figures 1 and 2. The first plot on the left in Figure 9 shows an image in the standard basis composed of low- and high-frequency 2-D sinusoids. A real MRI might depict an image of a brain. The next plot shows the image in the Fourier basis. The white dot in the lower left represents the low-frequency 2-D sinusoid shown in the third plot, while the grey dot in the upper left represents the high-frequency sinusoid in the last plot. A typical MRI system would scan through the Fourier basis, acquiring all of the points at the desired sampling rate. Once all of the Fourier samples are taken, they can be transformed to the standard basis to retrieve the image.

The amplitude of the points in the Fourier basis corresponds to the correlation between the sinusoid represented by that point and the image. This is calculated by the sum of the image multiplied by that sinusoid. For example, the amplitude of the white point is the sum of the image multiplied by the sinusoid in the third plot. Thus, each traditional MRI sample in the Fourier basis is really a compressed sample—a weighted sum of the image, where the sinusoids act as the weights. This means that CS can be applied to MRI without any hardware changes. Using CS theory, a fraction of the full number of samples typically required for MRI can be used to reconstruct an image, significantly reducing measurement time. MRI CS has a number of advantages [31]. They are as follows:

1. MRI is often used to produce 3-D data (e.g., Figure 10). Higher-dimensional data is typically sparser than lower-dimensional data, enabling 3-D MRI to benefit more from CS than 2-D imaging applications.

2. MRI imaging is performed in a high-SNR laboratory environment.

3. No new hardware is required to produce the compressed samples. Flexible hardware that can be programmed with arbitrary weights would be ideal, but MRI that uses Fourier weights has worked well in practice [32].

4. There is substantial motivation for increasing the speed of MRI. Patients are required to remain still for long periods of time, which can be difficult in many instances. In addition, MRI equipment is very expensive. Increasing the throughput of an MRI machine will decrease overall cost.

5. Image reconstruction can easily be done using powerful computers in postprocessing. Real-time image reconstruction is not required.

**GUIDELINES**

The following five positive aspects of applying CS to MRI can be formulated as general conditions for the successful application of CS.

1. The data should be very sparse; higher-dimensional data is ideal.
2. The SNR should be high; a laboratory environment is ideal.

3. The hardware that produces the incoherent weights needed for CS should be easily available.

4. There must be substantial motivation for adopting a CS strategy so the benefits outweigh any disadvantages.

5. The application must be able to accommodate the long reconstruction time and high computing costs of CS.

It is possible to use CS in a case that does not satisfy these conditions, but it will be more difficult to produce a practical system. We will try applying these guidelines to two test cases to determine their suitability for a CS implementation. The first application is an IR imager for spinning munitions (shown in Figure 11) [33]. The scene is projected through the coded aperture onto the sensors to create weighted samples. The coded aperture is a randomly-patterned mask that blocks random sections of light from reaching the sensor. The coded aperture pattern cannot be changed like the mirrors of a DMD, but the rotation of the aperture relative to the scene via the natural rotation of the munition can produce the different weights for each compressed sample. The coded aperture shown in Figure 11 has a relatively low resolution for illustration purposes. In practical applications, the resolution of the coded aperture is much higher than the sensor array, increasing the resolution of the sensor array the same way the DMD increases the resolution of a single sensor. We apply the guidelines as follows:

1. Two-dimensional data is only moderately sparse.

2. A high-dynamic munition flying through the sky will not have a high SNR.

3. The rotating, coded aperture produces semistructured weights not ideal for CS.

4. High-resolution IR imagers are expensive, motivating the use of CS to increase the resolution of inexpensive low-resolution imagers.

5. A munition imager for target recognition must operate in real time, although there may be ways to use compressed CS data without reconstruction [34].

Only the fourth guideline is encouraging. The cost benefits of CS must be weighed against the cost of the CS hardware and other disadvantages associated with it. Clearly applying CS in this case would be challenging.

Another possible CS application is antenna pattern measurement [35]. Figure 12 shows an example antenna pattern as a blue line around an antenna under test (AUT). A traditional measurement system is shown on the left. The circle of dots around the AUT represent test antennas that transmit one at a time (green), while the others are inactive (red). The sensitivity of the AUT is measured relative to each test antenna, creating an antenna pattern. The middle plot shows a CS implementation where random patterns of test antennas transmit simultaneously.

The transmissions are summed in the AUT, creating a compressed sample. The CS implementation reduces the number of measurements required, accelerating the measurement process. We can show that the following five conditions apply, making this a promising application of CS:

1. Antenna patterns are often 3-D (azimuth, elevation, and radio frequency) and therefore very sparse.

2. Measurement is typically performed in a controlled laboratory environment with high SNR.

3. There are already systems that use multiple test antennas placed in a ring around the AUT to measure antenna patterns. These systems could be easily modified to activate random combinations of these antennas instead of using them one at a time. The on/off patterns of test antennas are similar to the on/off patterns created by a DMD, which work well for CS.

4. Antenna pattern measurement is typically an expensive, slow

![Figure 11: Architecture of a CS IR Imager for Spinning Munitions (Source: CCDC ARL).](image-url)
procedure, giving substantial motivation to use CS to accelerate measurement by reducing the number of samples.

5. There is no problem postprocessing the data; the data are not needed in real time.

**CONCLUSIONS**

Now that we know what CS is and have examined some case studies, we can answer the questions raised in the Introduction.

- Many unsubstantiated comments about CS are simply untrue. CS is not “linear interpolation, rebranded.” The example of signal reconstruction using traditional sampling in Figure 2 is a form of interpolation. CS uses an optimization algorithm to identify signal coefficients in a sparse basis from compressed samples. This is not interpolation.

- The claim that CS is overhyped is probably true. Even CS researchers acknowledge that “...compressed sensing has not had the technological impact that its strongest proponents anticipated” [9]. When reading about all of the applications of CS, it may seem that CS is revolutionizing many sensor technologies. However, most of these applications are experimental systems that have not overtaken traditional techniques. For example, when Wikipedia states that “compressed sensing is used in a mobile phone camera sensor” [2], it is not talking about all phone cameras. It is referring to the experimental low-power CMOS imager in the case study that was never actually used in any commercial cell phone.

- Open criticism of CS from researchers such as Leonid Yaroslavsky and Kieran Larkin [11, 12] partially stem from the overhyped publicity of CS. Their comments can be generally understood to say that CS will not provide the best performance in all cases, and, in many practical situations, more traditional sampling or compression strategies will outperform CS. This is certainly true. On the other hand, CS has achieved significant results that will increase performance for certain applications. CS does deserve recognition, as well as further research and funding, even if it is not the panacea some have claimed.

- An additional point is that the growth in CS has also advanced topics, such as the properties of random matrices, sparse representation, and optimization algorithms applicable to areas outside sensing and measurement. Even if CS sensor hardware has been slow to develop, there are many related fields benefiting from CS research.

Can CS solve your sensor and measurement problems? As with most significant questions, there is not an easy yes or no answer. CS is definitely not a cure-all that can be used in every situation. The fact that Mark Neifeld declared that “we haven’t discovered the ‘killer’ application yet” [10] when accessing the potential of CS imagers with leaders in the defense sector should give one pause if they think that their application is the “killer” application. The five guidelines listed are a good starting point. CS should realistically be compared to other alternatives, whether with other sampling techniques such as basis scanning [38] or alternate sensor technologies. CS results that seem promising in idealized settings might not perform well in more realistic scenarios. CS is still developing. An application that is not practical in the near term might still deserve longer-term research. Government and industry leaders should approach CS with their eyes open. They should be aware of the advantages and disadvantages of CS and know if the funded research is practical and short...
term or more theoretical and long term. In 1956, there was a surge in optimism about information theory much like the hype CS is experiencing. The words of Claude Shannon, of the Nyquist-Shannon sampling theorem, ring true today as much as they did then [39].

Information theory has, in the last few years, become something of a scientific bandwagon… What can be done to inject a note of moderation in this situation? In the first place, workers in other fields should realize that the basic results of the subject are aimed in a very specific direction… A thorough understanding of the mathematical foundation and its …application is surely a prerequisite to other applications. The subject of information theory has certainly been sold, if not oversold. We should now turn our attention to the business of research and development at the highest scientific plane we can maintain… A few first-rate research papers are preferable to a large number that are poorly conceived or half finished.

Whether CS can solve your problem or not, there is a final lesson to learn from CS methodology. The hardware and software aspects of sensor systems should not be designed independently, rather there should be an interdisciplinary codesign resulting in an optimal solution [8].

REFERENCES


BIOGRAPHY

MICHAEL DON is an electrical engineer for ARL, where he specializes in high-performance embedded computing, signal processing, and wireless communication. He began his career as an intern at Digital Equipment Corporation, performing integrated circuit design for their next-generation Alpha processor, the fastest processor in the world at that time. After graduating, he worked for Bell Labs, where he engaged in the mixed-signal design of read channels for their mass storage group. Mr. Don holds a bachelor’s degree in electrical engineering from Cornell University and is currently pursuing a Ph.D. in electrical engineering at the University of Delaware, where he is researching compressive sensing applications for guided munitions.
ADDITIVE MANUFACTURING

HIGH-PERFORMANCE POLYMERS

for Space and Aerospace

By Hilmar Koerner

(Photo Source: U.S. Air Force)
INTRODUCTION

Additive manufacturing (AM) has come a long way. However, there are still a number of problems to solve for AM of polymer matrix composites (PMCs) [1], especially for future aerospace in U.S. Department of Defense (DoD) applications. The largest benefits for the aerospace industry and DoD-related challenges are complexity-enabled capability, low-volume/low-cost manufacturing, part reduction, and service life; prototyping and form fit function; and toolless part manufacturing and lightweighting interior hardware (ducts, seat framework, wall panels, and avionic bays).

The U.S. Air Force Resource Laboratory (AFRL) is completing the first steps toward complex load-bearing structures in composites, focusing on continuous reinforcement three-dimensional printing (3-DP). In addition, composite tooling, an area for noncritical parts, provides lead time and cost reduction. Sometimes it does not make sense to print an article in 3-D, but it does save a lot of time and cost to print the mold that makes the part. Replacing metal parts in jet engines or low-cost aircraft technologies is another area in which the ability to print something that cannot be made conventionally has a potential for lightweighting and part reduction.

Further opportunities for AM composites include on-demand printing parts that do not have a digital or hardcopy blueprint (B-52) for ground-based and nonstructural replacements. Based on demand signals and technology readiness, we look at a roadmap in AM with the following near-mid and far-term goals (including metal and polymer-based processes):

- Near/midterm – thermal management (heat exchangers, nozzles, and ducting), embedded functionality
- Far term – critical and load-bearing structures, canopy frames, wing spars, tunable radio frequency, integrated power, and embedded sensors.

These goals are addressed primarily with metals, polymers, soft matter, multicomponent, and reinforced composites. Composites via AM are still in the early stages and face stringent certification and qualification processes—an area for which tools must still be developed. For example, measuring the mechanical properties of a printed specimen vs. a conventionally-manufactured specimen is like comparing apples to oranges (i.e., one peels = orange, while the other breaks = apple). The current AM field can also be described as creating parts with well-ordered defects due to the step-by-step process. Understanding the effects of these regarding the overall structure and performance is key to success.

For the U.S. Air Force (AF), AM is a pervasive processing technology and spans many domains, including everything from soft-bio sensors to extreme temperature metals and ceramics and flexible electronics to survivable and tailored munitions.

AFRL is completing the first steps toward complex load-bearing structures in composites, focusing on continuous reinforcement three-dimensional printing.

(antennae, electronics, and wiring), smart tooling, fairing mounts, and small engine parts (brackets, fixtures, and shrouds).

The work described in this article was implemented by a team very familiar with composite manufacturing high-temperature thermosetting resins and an expert in selective laser sintering (SLS) to successfully process these resins into AF-relevant articles. The Materials and Manufacturing Directorate (RX) at Wright-Patterson Air Force Base (WPAFB), Dayton, OH, has been driving the field of high-temperature polyimide thermosets for several decades in conventional composite manufacturing. These thermosets, originally developed at the National Aeronautics and Space Administration (NASA) and refined and optimized at AFRL during the 90s, are used in advanced AF PMC applications [2].

Six years ago, the Composites Branch at AFRL decided to seriously investigate opportunities and limitations of AM of composites focusing on thermosetting resins. The community was heavily involved; very little was carried out on thermosets (almost nonexistent for reinforced composites). The initial target was additively-manufactured components for jet engine propulsion that required higher service temperatures reaching in excess of 350 °C. The research recently expanded into lower-cost materials for limited life technologies [3, 4].

The very first experiments in high-temperature resins produced feedstock materials for fused deposition modeling and allowed researchers at AFRL to print coupons in 3-D, with service temperatures approaching 300 °C [5]. Recognizing that not all resins amend to extrusion into filaments, the team established a partnership with NASA to work with their resin-/transfer molding resins of the same polyimides family. Rather than create a filament, which can be a daunting process, the idea was to simply create a powder with the right particle size distribution, use well-known
selective laser sintering AM processes to print articles from the powders, and add chopped carbon fibers for reinforcement in subsequent studies (Figure 1, top).

Figure 1 shows the schematic overview of the project from synthesis, to mixing with milled carbon fibers, to powder, and SLS. This process produces a small bracket for jet propulsion AF applications or NASA door panels with AM-enabled noise reduction, primarily focusing on enhancements in mechanical properties and applying space for exquisite platforms such as the F-35.

**SELECTIVE LASER SINTERING**

Understanding the effects of processing on microstructure and void formation and the influence on mechanical properties in service life is poorly understood in any type of AM, but specifically for manufacturing composites. This is because of the complex multimaterial process that requires reducing defects and enabling well-formed interfaces between reinforcement constituents and the matrix. It is further complicated by the highly-nonequilibrium process for which timing of several materials attributes is essential, such as timing of hardening, crystallization, and interdiffusion.

Ultimately, these poorly-understood phenomena will affect fatigue and aging of 3-D printed parts and components. Therefore, we focus on nonstructural and tooling for complex critical parts. Standardization and the American Society for Testing and Materials protocols for 3-DP must still be matured. Three-dimensional printing will not replace conventional manufacturing but rather complement it in function, lightweighting, and complexity. For example, a wing could be a hybrid of conventional and 3-DP, with a topology optimized lightweight and mechanically adaptable, responsive core structure and a conventional composite skin.

While scientists at AFRL/RX have designed filament feedstock for fused deposition modeling that creates articles with service temperatures above 350 °C, some materials in this class of thermosets cannot be extruded into filament feedstock. In this case, SLS is the better processing route because it only requires the material to be in a well-defined powder form, with a narrow size distribution of particles.

Figure 1 demonstrates the overall approach. A polyimide is synthesized via polymerization using well-established organic synthesis routes. The raw material is then powderized in special equipment that will lead to particle sizes of 50–100 µm (Figure 2). This is necessary for the SLS process.

Figure 2 shows a one-pot reaction and vial, with the final product in powder form. A composite panel manufactured via resin transfer molding from the polyimide and a carbon fabric is shown.
on the right. During this process, a roller runs across a bed of the powder to evenly distribute the powder. A high-power laser is programmed to fuse these particles along a trace given by a computer-aided design (CAD) program containing the underlying article geometry. The laser power, speed, and bed temperature have to be just right to melt the particles and fuse them to get a coherent structure. A subsequent layer of powder is then spread across the bed. The laser continues fusing particles in discrete patterns dictated by the CAD design until a full article is printed. Finally, the remaining unfused powder is removed, and the rough surface of the part is polished.

**Neat Resin**

Efforts to additively manufacture coupons for mechanical testing via SLS are extremely challenging because the resin system has a very broad processing window with low viscosity, specifically for resin-transfer molding from ~150 °C to ~320 °C. While this large temperature window is required for the resin to transfer and flow into a carbon fiber fabric preform and fill the entire mold, it is detrimental for the SLS process. During SLS, the laser pulse may enable fusion but appears insufficient for crosslinking to occur. Typical laser speeds are 10 m/s (400 in/s). The crosslinking is necessary because a part that has been melted and fused still needs to be postcured into a fully crosslinked functional part. Uncured, laser consolidated material is still very brittle and cannot be handled for postprocessing.

The crosslinking process leads to an exotherm during which the uncured object uncontrollably collapses into a puddle (Figure 3). Figure 3 shows selective laser-sintered coupons on the top (array of different processing parameters) and the same coupons after postprocessing on the bottom. None of the coupons keep the original shape or melt or are too brittle to handle.

There are several parameters that can be changed to influence the material's behavior during processing; however, none of these were successful. While the power of the laser increased, the speed lowered, and the bed temperature raised, the total time at temperature was never enough to induce any crosslinking in the resulting melt pool to increase the viscosity. Furthermore, the difference between melt temperature of ~160 °C to the crosslinking initiation temperature of 320 °C was too broad. This would not allow sufficient crosslinking for self-support of the printed material during subsequent postprocessing to obtain a fully crosslinked object without losing shape. The viscosity of the material within this temperature range is like maple syrup (30 poise). Further
Efforts to additively manufacture coupons for mechanical testing via SLS are extremely challenging because the resin system has a very broad processing window with low viscosity.

Trials were carried out with precured (b-staged) material.

Figure 4 (left) shows differential scanning calorimetry (DSC) of the raw, unprocessed polyimide. The material melts at around 160 °C and starts to crosslink, with an onset of 320 °C and an exotherm peak at 372 °C (typical for these types of thermosets). The very low viscosity of the material between 160 °C and 320 °C is the bottleneck for processing via SLS.

Figure 4 (right) shows material that was cured at 300 °C for 2.5 hr. A filament is formed during separation of the rheometer plates, indicating an increase in molecular weight. The result is that this material can be drawn into a filament in the melt due to increase in molecular weight of the resin while still exhibiting melt flow. Essentially, the powder was heated systematically to a temperature that would promote crosslinking to a small percentage (10%–20%). If crosslinking does not proceed beyond a 30%–40% conversion of the crosslink moieties, the material should still melt during laser sintering. The molecular weight is now increased, and the material keeps its shape during subsequent postprocessing steps. However, even this approach was unsuccessful, despite a viscosity approaching values for materials that can be successfully sintered.

**Carbon Fiber-Filled Resin**

After restarting, the team recognized that the energy from the laser to the resin had to increase so that partial crosslinking was enabled while the laser scanned across the powder bed. There are examples from the literature that discuss adding carbon black into polyamide to establish better thermal conductivity and energy transfer from the laser to the resin material [6]. The reinforcing carbon fibers planned for reinforcement of the resins in the second phase of the program to improve mechanical performance at temperature are perfect laser energy absorbers. They can transfer the energy as heat to the matrix and therefore promote crosslinking.

The resin was mixed with 35% carbon fibers (~60 µm long, [see top of Figure 1]) and dry blended for printing composite coupons. As expected, the thermal conductivity of the carbon fiber-filled resin was greater by a factor of 3, compared to the neat resin. In addition, the penetration depth of the laser energy increased substantially and enabled fusing thicker layers.

With further optimization of increased laser power and layer thickness, the team could finally print coupons for mechanical testing after postcure steps (Figure 5). Figure 5 shows laser scanning across the powder bed at 10 m/s (left) and printed mechanical testing specimens before postprocessing (right). The rough surface is from
agglomerating milled carbon fibers and b-staged resin particle sizes exceeding original particle size of 50–100 µm.

**Postprocessing**

Three-dimensional printing thermoplastic materials does not require a postprocess step, such as nylon, unless it is a finishing step. This is different for thermoset resins described here. While the material has the desired shape, it is not yet fully crosslinked and will be very brittle. It is essentially a fused powder of a semicrystalline polymer and will easily fall apart like a compressed powder. A postcure step must be introduced to crosslink the material at elevated temperatures without “melting” the material and thereby losing the printed shape. This can be accomplished by a systematic stepwise increase in temperature while holding the material at each step for a given time (e.g., 1 hr).

DSC and rheology measurements provided information on the state of the material. A key parameter is the degree of cure or the conversion of crosslinkable moieties into crosslinks within the resin. It is well understood for resins of this type that when a degree of cure of 0.5 is reached (0 being not crosslinked at all and 1 being completely crosslinked), the material starts to gel. When heated above the softening temperature, the material does not melt anymore and behaves like a rubber. A gel may turn soft with temperature, but it will not lose its shape. Hence, once the gel regime is reached via the time at temperature, the material can be ramped up quickly in temperature to fully cure based on DSC measurements (for the resin in this study, 370 °C).

Resulting cure programs, also called cure cycles, were very similar to what needs to be done with the same resins in conventional composite processes. These programs can be accomplished within about 10 hr for 3-D printed parts. It is obvious that AM or 3-DP for higher temperature composite applications at AFRL will not be high throughput; however, the technology will enable unprecedented part complexity and part consolidation.

**Performance**

Once small demo articles such as brackets were printed, several mechanical testing specimens were printed to evaluate the material’s performance. After microscopy of the printed samples and density measurements (gas pycnometer, 15%–20%), it was apparent that the printed specimens were not perfect and that additional work needs to be conducted to obtain void-free articles (Figure 6).

Figure 6 shows a fracture in center of gauge section (left) and a fracture surface indicating porosity of the specimens (right). All specimens in the test fractured in the center gauge section, indicating proper sample clamping and homogeneity in the specimens. Irrespective of the voids, these first-printed specimens still showed exceedingly promising mechanical properties at elevated temperature, which is key to many of the application’s challenges of these materials. The results of the selective laser-sintered specimens were compared to results of another common resin used in conventional composite manufacturing, a toughened epoxy (5320-1) with a high-performance carbon fiber (IM7).

Table 1 summarizes the initial results and demonstrates that the performance of these selective laser trials with high-temperature thermosetting polyimides is very promising. The comparison epoxy sample would decompose long before the measurement temperature of the high-temperature polyimide. A fairer comparison includes values of mechanical properties at reduced temperatures—in this case, at a temperature relative to the softening
energy transfer from the laser to the resin. Considering these preliminary results on 3-D printed articles and specimens, the resulting mechanical properties at high temperatures were outstanding.

Further work is required to remove the voids during SLS, as well as increase carbon fiber content. While there is great potential in this technology, there will be limitations going forward to explore applications in which complex-shaped polymer matrix composites might be beneficial (e.g., for replacing metal parts on jet engines). One of the disadvantages of this process is that it will not be feasible for continuous carbon fiber composite 3-DP. Technologies, such as automated fiber or tape placement, will soon merge with 3-DP technologies to provide engineering solutions for such applications.

Further work is required to remove the voids during SLS, as well as increase carbon fiber content.
By Virginia G. DeGiorgi, Edward P. Gorzkowski, Heonjune Ryou, and Stephanie A. Wimmer

SUMMARY

Many advanced engine system designs result in significant increases in operating temperatures, which, in turn, require new thermal protection coatings (TBCs). In addition to providing a much higher level of thermal protection, TBCs must have sufficient strength at these higher operating temperatures to avoid spallation or other structural failure. Advanced TBCs are often made from ceramics because of its high melting temperature, durability, and relative ease of application. In this work, the relationship between material microstructure and performance of ceramic coatings is identified through computational approach. All work used yttria-stabilized zirconia (YSZ) as a model material, but the approach can be applied to many different ceramic materials.
INTRODUCTION

Advances in engine design have led to the need for additional thermal protection for engine components. It is not unreasonable to expect that the next generation of TBCs must provide at least 1 to 2 orders of magnitude increase in thermal protection over existing coating technologies. To design an effective coating, it is important to realize that TBCs are not single-purpose materials. In addition to providing thermal protection, a TBC must withstand high temperatures and structural loads without failure. Structural failure of TBCs is often simply not acceptable. For instance, coating spallation on turbine blades can result in severe engine damage from TBC coating debris entering the engine.

Obtaining increased thermal protection and sufficient strength at higher temperatures can be problematic. Increased porosity, a commonly used feature to decrease thermal conductivity, is also known to decrease mechanical strength, increasing the risk of structural failure. Reducing grain size, however, decreases thermal conductivity and increases mechanical strength. Computational modeling can be used to design ceramic coatings that balance features, such as porosity and grain size, by quickly evaluating thermal performance and structural strength.

In this article, we review use of computational simulations to evaluate using porosity in a ceramic TBC for sufficient thermal and structural performance. We will first introduce the material characteristics of interest and our modeling process. We will then present initial model validation through direct comparison with experimental results. Next, trends in performance identified via a series of computational simulations will be presented. Finally, we end with a few words addressing future plans.

MATERIAL CHARACTERISTICS

In order for computational modeling to be meaningful, the model must include a reasonably-accurate constitutive response and accurate geometric representation of pertinent features. In the current work, the baseline material used for all studies is YSZ, a well-characterized, ceramic TBC. YSZ bulk material structural properties are readily available in the literature.

Capturing appropriately-sized features is more challenging. In this case, the limitation is not the smallest dimension that can be incorporated into a computational model but rather the smallest grain size that can be fabricated. This smallest grain size will dictate the size of porosity that we can incorporate into the TBC. Developing sintering processes minimizing grain growth is a major challenge for ceramics since grain growth in ceramics tends to occur at the same temperatures required for sintering.

The U.S. Naval Research Laboratory (NRL) has the ability to create nanograin ceramic materials. Gorzkowski et al. [1] successfully limited grain growth in a ceramic through a two-step sintering process, which paired a short-duration, high-temperature heating stage to isolate any porosity with a long time step at a lower temperature to achieve final densification. Thermal diffusivity decreased ~30% when this process was applied to nanosized YSZ powder (100-nm grain size) with random nanoscale pores [1, 2].

Now that we know we can create a ceramic with a small-scale porosity, we need to develop an understanding of the geometric and spatial characteristics of porosity. Nakamura et al. [3] examined porosity by using a regular array of defects with random size, shape, and orientation identifying the relationship between measured strength and defect size. Wang et al. [4] and Kulkarni et al. [5] quantified real porosity patterns from scanning electron microscope images for coatings generated by various techniques. Amsellem et al. [6, 7] focused on image reconstruction highlighting the importance of three-dimensional (3-D) features in calculating effective moduli. Cronin et al. [8] used 3-D images to identify and include grain boundaries in their finite-element (FE) models as critical features relating to thermal performance. In contrast to using full 3-D characteristics of porosity, Wei et al. [9] demonstrated through simplified modeling that the largest defect was the most important factor in determining the effective thermal conductivity.

All of these studies provide insight, but many lacked details on porosity characteristics which could be used to further evaluate the characteristic-performance relationship. However, Wang et al. [4] provided geometric and statistical information on porosity characteristics. Porosity was defined by three types of defects—voids, pores,
and cracks. Each defect category can be represented by a two-dimensional (2-D) ellipse or a 3-D ellipsoid. In two dimensions, voids are spheres (equal minor and major axes), pores are oblate ellipses, and cracks are flattened ellipses (minor axes are much less than major axes). We used this data to define the baseline porosity condition examined and included variations in defect shape, size, and orientation. We also created a 3-D representation of porosity by defining the third axis as equal to the 2-D minor axis. Our matrix of porosity definitions allowed us to develop an understanding of how defect shape, size, orientation, and the amount of porosity present impact TBC performance.

**COMPUTATIONAL MODELS**

Thermal and structural performance was evaluated by FE techniques for a range of porosity level, defect sizes, and spatial orientations. All of the porosity geometries considered were feasible based on NRL's successful fabrication of nanograin ceramics. The baseline material was fully-dense YSZ, with a thermal conductivity of 2.5 W/mK and Young's modulus of 200 GPa. Linear material response was utilized in all analyses.

A commercial FE code [10] was used for all analyses, which were run on U.S. Department of Defense (DoD) high-performance computing (HPC) facilities. The 2-D model represented a surface area coating of 30 x 50 µm and unit thickness. The 3-D models had a thickness of 5 µm, resulting in a volume of 30 x 50 x 5 µm. This thickness was chosen to allow 3-D effects while creating a model that was not too large to solve on available computer resources. Two-dimensional models consisted of ~300,000 elements and ran in less than an hour. The 3-D models had between 7 and 9 million elements, depending on the porosity level, and ran in less than 10 hr.

**Two-Dimensional Models**

Multiple 2-D models were created to evaluate defect size, number, and orientation [11]. Defect dimensions were based on measured data [3]. The baseline 2-D model had 600 defects distributed between voids, pores, and cracks (as shown in Table 1). In addition, defect size and orientation within the plane varied. A typical defect pattern is shown in Figure 1.

Performance is evaluated by effective thermal and strength properties as defined by equations (1) and (2):

\[
K_{\text{eff}} = \frac{QL}{W \Delta L},
\]

and

\[
E_{\text{eff}} = \frac{FL}{W \Delta L},
\]

where \(K_{\text{eff}}\) and \(E_{\text{eff}}\) are the effective thermal conductivity and effective elastic modulus. For vertical loading (either temperature difference or displacement), \(L\) is the width, \(W\) is the height, \(\Delta T\) is the temperature difference between the top and bottom edges. \(\Delta L\) is the displacement of the top edge, \(Q\) is the sum of reaction heat flux at the node points along the bottom edge, and \(F\) is the sum of reaction forces at the node points along the bottom edge. For horizontal loading (horizontal temperature difference or displacement), the \(L\) and \(W\) change places in the equations, the \(\Delta T\) and \(\Delta L\) are measured horizontally, and \(Q\) and \(F\) are summed along the left-hand side.

**Three-Dimensional Models**

While performance trends and other information can be inferred from 2-D simulations, results have to be critically evaluated since the 3-D nature of the material is not incorporated into the analysis. The 3-D computational modeling is addressed by Wimmer et al. [12, 13]. A typical 3-D porosity model is shown in Figure 2, as well as representative porosity patterns through the thickness. Examination of the variation in porosity at any cross section highlights the inherent 3-D nature of the material.

---

**Table 1: Nominal Radii and Standard Deviations**

<table>
<thead>
<tr>
<th>DEFECT TYPE/ELLIPtical REPRESENTATION</th>
<th>MAJOR AXIS (µm)</th>
<th>MINOR AXIS (µm)</th>
<th>STANDARD DEVIATION</th>
</tr>
</thead>
<tbody>
<tr>
<td>Void/Equal Axis</td>
<td>5.8</td>
<td>0.687</td>
<td>0.04</td>
</tr>
<tr>
<td>Pores/Oblate</td>
<td>31.7</td>
<td>0.528</td>
<td>0.106</td>
</tr>
<tr>
<td>Crack/Flat</td>
<td>62.5</td>
<td>0.528</td>
<td>0.053</td>
</tr>
</tbody>
</table>

Figure 1: Geometry for the 2-D Model With 3.5% Porosity (Source: NRL).
Typical variations in localized porosity are shown in Table 2. Large differences in local porosity may occur; for instance, a TBC with 3.54% porosity may have regions with localized porosity level ranging from 1.93% to 6.91%. Calculated effective properties will vary based on the local porosity level.

Effective material properties for the 3-D models are calculated by including the thickness \( D \) in equations (3) and (4):

\[
K_{\text{eff}} = \frac{QL}{WD \Delta L}, \quad \text{(3)}
\]

and

\[
E_{\text{eff}} = \frac{FL}{WD \Delta L}. \quad \text{(4)}
\]

**MODEL VALIDATION**

Model validation is a critical step in any computational effort. The fundamental question in the validation process is how well does the model represent reality? Once confidence is gained through validation, additional “what if” scenarios can be undertaken.

NRL successfully fabricated YSZ based ceramics with specific porosity levels (Figure 3), as fabricated ceramics were close to single-defect porosity conditions [13]. Thermal diffusivity and heat capacity were measured, and effective thermal conductivity was calculated. Effective elastic moduli were calculated from measured indentation hardness values.

Validation results used all 2-D and 3-D calculated values. In addition to models with multiple types and sizes
of defects, we examined the impact of porosity made up of only one type of defect—spherical voids [13]. This defect geometry was created specifically to mimic the creation of a TBC using spherical sacrificial material (i.e., material which will disintegrate during the sintering process).

Effective thermal conductivity vs. porosity values are shown in Figure 4. As expected, the effective conductivity decreases with increasing porosity. The 3-D model shows effective values ~10% greater than 2-D values. The measured conductivity is between the calculated effective conductivity from the 3-D and 2-D models showing a variation of less than 10% between computational and experimental results. Single-defect porosity calculated values fall within the general performance trend defined by all calculated values.

Figure 5 shows the effective elastic moduli for vertical loading. As expected, the effective moduli decreases with increasing porosity. Measured moduli are higher than computational values, with the 3-D results closer to measured values. Once again, single-defect porosity computational results follow the same performance trends as all other computational porosity patterns.

Initial comparisons are very good. While additional work is needed to improve computational-experimental validation metrics, this is a good start, and it is reasonable to discuss computational trends based on this level of agreement.

**PERFORMANCE TRENDS**

Computational modeling provides insight on how specific porosity characteristics affect performance. Horizontal chaining of heat flux is desirable in TBCs to block heat transfer across the coating. However, chaining stress values is undesirable since it leads to connecting regions of localized failure and creates percolation of material separation across the thickness. Ultimately, this can lead to percolation of material separation and spallation along high-stress seams. The general trend is for the effective structural modulus to decrease with increasing porosity. However, variations in structural and
Computational modeling provides insight on how specific porosity characteristics affect performance.

Thermal performance can be observed, with variations in defect geometries. Variations in porosity resulted in the greatest changes in effective performance metrics. Variations in thermal and structure results due to variations in pore lengths are shown in Figure 6.

Three-dimensional FE studies [12, 13] show trends consistent with the 2-D analysis results. There are, however, significant and subtle differences. General trends are similar between 2-D and 3-D computational results. However, the interactions between individual defects appear remarkably different between 2-D and 3-D models, even at the same porosity level as seen in Figure 7 for structural response. The differences in defect interactions can be seen directly when we duplicated the porosity pattern found at a slice point in a 3-D model and in a 2-D model.

Figure 8 shows the von Mises stress pattern at three slice locations in a 3-D model: (a) 3-D model at Slice 1, (b) 3-D model at Slice 2, (c) 3-D model at Slice 3, (d) 2-D duplication of porosity pattern found at 3-D model at Slice 1, (e) 2-D duplication of porosity pattern found at 3-D model Slice 2, and (f) 2-D duplication of porosity pattern found at 3-D model at Slice 3. In general, the differences in maximum stress values based on localized porosity levels can be clearly seen in both 2-D and 3-D calculated results. Similar results are seen for thermal properties.

While 2-D representations of microstructures, including porosity characterization, have a place in identifying performance trends, the need to understand the variations through the thickness (i.e., the 3-D effect) cannot be understated. As we move forward in designing TBCs, it is important to consider how porosity varies in the third dimension.

**CONCLUSIONS**

In this article, we have given an overview of NRL’s computational modeling effort to understand thermal and structural characteristics of ceramic TBCs. Initial
computational model validation required fabricating nanostructured ceramics with specific porosity levels. This is itself a challenge. NRL has successfully fabricated nanostructured YSZ ceramic with four different levels of porosity ranging from <0.01% to 9%. There is still additional validation work to complete, but this initial effort provides enough to discuss general trends.

In general, computational results indicate that geometric characteristics...
3. Additional measures of thermal and structural performance. Currently-used effective properties represent averaged response and obscure localized performance. The metrics identified will have to amend to computational and experimental measurements. These changes will provide improved computational models, which we can exercise to explore the connection between performance and microstructure, ultimately assisting in developing the next generation of ceramic TBCs.

ACKNOWLEDGMENTS

Funding for this project was provided by the Office of Naval Research through the NRL’s Basic Research Program. The FE analysis was supported by the DoD HPC Modernization Program using the U.S. Air Force Research Laboratory’s Major Shared Resource Center under project 416, subproject 231.

REFERENCES


BIOGRAPHIES

VIRGINIA G. DEGIORGI is the head of the Multifunctional Materials Branch in the Materials Science and Technology Division at NRL. Prior to joining NRL, she worked at Westinghouse Electric Corp., where she became the first woman and youngest employee at the time to be awarded the prestigious corporate B. G. Lamme Scholarship. Her research interests include applying computational modeling techniques to diverse fields of study, such as electrochemical corrosion and ceramic thermal barrier coatings. She is a Fellow of ASME and an American Business Women’s Association Top Ten. Dr. DeGiorgi holds a B.S. and M.E. in civil engineering from the University of Louisville and a Ph.D. in engineering mechanics from Southern Methodist University.

EDWARD P. GORZKOWSKI is a materials research engineer and head of the Ceramics and Rapid Prototyping Section at NRL. His research interests include piezoelectric materials for sensor and actuator applications; processing dielectric and ferroelectric materials; unique processing methods to create bulk nanostructured ceramics; and 3-D microstructure characterization and modeling for improved ceramics, such as thermal barrier coatings and capacitor materials. He received the 2015 Du-Co Young Professionals Award and a 2018 Best Paper Award for the Journal of American Ceramics Society. Dr. Gorzkowski holds a B.S. and Ph.D. in materials science and engineering from Lehigh University.

HEONJUNE RYU is a staff scientist at the Multifunctional Materials Branch at NRL. His multidisciplinary research focuses include interfaces between biofunctionalizing organisms, additively manufactured metals, and nanocrystalline ceramics. Dr. Ryu holds a Ph.D. in mechanical engineering from the University of Maryland, Baltimore County.

STEPHANIE A. WIMMER is a mechanical engineer at the Multifunctional Materials Branch at NRL. She has more than 20 years of experience in computational modeling and simulation of structures involving multiphysics, fracture, corrosion, porosity, and nonlinear materials. Dr. Wimmer holds a B.S. in ocean engineering from Texas A&M University and a Ph.D. from the Naval Architecture and Marine Engineering Department at the University of Michigan.
THE IMPORTANCE OF EARLY PROTOTYPING in Defense Research, Engineering, Acquisition, and Sustainment

By Eric Spero, Zeke Topolosky, and Karl Kappra
SUMMARY

The U.S. Department of Defense (DoD) operates across the entire life cycle of capability development, including exploratory science, research and development (R&D), production and deployment, operations and sustainment, and disposal. The processes used by the DoD to acquire systems are based on the ability to plan and budget for long-term, stable mass production of solutions addressing variable and predictable scenarios. However, current global political conditions, the rate of technological change, and the application of emerging technologies create uncertainty in the future security environment.

Prototyping ideas throughout a system’s life cycle offers an opportunity to continuously keep pace with change and uncertainty. A survey of the literature reveals a desire to prototype early in the development process as a means of generating knowledge while reducing risk. However, budget and legal authorities pin prototyping to technology maturity assessment and as a pathway to accelerate delivery of near-final, full-up systems. Missing is a complementary strategy for early prototyping. This article presents an overview of prototyping in the DoD and promotes consideration of, and funding for, physical and virtual prototyping throughout the life cycle as an effective way to proactively test hypotheses, engage stakeholders early, learn, prune decision paths, and ultimately deliver the right capability faster to Warfighters.

INTRODUCTION

Adversaries of the United States are competent, quick, and effective at fielding emerging technologies, which presents an impediment to maintaining overmatch both now and in the future [1–5]. The processes used by the DoD to develop and acquire its capabilities are all based on the ability to plan and budget for long-term, stable mass production of solutions addressing variable and predictable scenarios [6]. The result is defense systems designed to requirements based on long-term accuracy and certainty that take years to design and build because of their inherent complicatedness. However, current global political conditions, the rate of technological change, and the application of emerging technologies create uncertainty in the future security environment; “future warfare will feature constant myriad technological advances that come at a tempo that disallows mass production” [7]. The decision-making processes that worked for mass-producing solutions to accurately predictable conditions may be incompatible with the conditions under which defense systems are envisioned to operate in the future and the speed with which these systems must be developed and integrated.

An increasingly effective path to ensuring faster fielding of the right solution is through an approach that engages technologists with users to learn their pain points, puts form to function early on in the development process, tests hypotheses through research experiments, quickly gets hardware into the hands of a user, and collects feedback to assess whether the product or service solution is on the path to meeting user needs [8–10]. This approach can occur as early as the ideation stage, and the early form-to-function effort can look like a sketch, a computer code with merely basic functionality, a hardware mock-up, a controlled scientific experiment, or some other mechanism to tactilely communicate an idea or concept before proceeding into the next phase of development. Colloquially, this process is called prototyping, and the operand in the process is called a prototype. Prototyping generates information that supports a decision; if faster decision-making is the end, then prototyping offers the way and prototypes are the means. A classic example of using nonfunctional prototype hardware to gain information is when the Industrial Design team at Apple made wooden iPhone mock-ups to determine the optimum size of icons for a human finger [11].

Issues surrounding DoD prototyping activities have recently arisen—definitions are inconsistent and budgets for prototyping are largely constrained to the later stages of technology maturity, revealing risk aversion in the acquisition system. Further, although there now exists a middle acquisition tier in which prototyping takes center stage to accelerate delivery of capability to the Warfighter, the Government Accountability Office (GAO) recently found that “[the] DoD has yet to fully determine how it will oversee middle-tier acquisition programs, including what information should be required to ensure informed decisions about program selection and how to measure program performance” [12].

A review of the John S. McCain National Defense Authorization Act (NDAA) for fiscal year 2019 (FY19) [13] shows that only 2 of 11 prototyping programs have been authorized as applied research—$10M to accelerate Army railgun development and prototyping and $160M for Innovative Naval Prototypes (INPs) applied research, where INPs “comprise potentially game-changing or disruptive technologies...
developed around anticipated Naval needs rather than in response to established requirements” [14]. All other such FY19 program authorizations ($678M) fall under advanced technology development and advanced component development and prototype budget activities. Overall, there is disagreement across the research and acquisition spectrum on what even constitutes a “prototype.”

WHAT IS A PROTOTYPE ANYWAY?

The word “prototype” originates from the Greek “prōtotupos,” which literally means first impression, mold, or pattern or the first from which all subsequent copies will derive [15]. Within the DoD alone, a multitude of definitions exists, each influenced by a product’s life cycle stage, the level of system hierarchy for which the prototype is built, the type of knowledge the prototype aims to uncover, and the final disposition of the prototype [16–19].

The Defense Acquisition University (DAU), the go-to source for defense acquisition professional training, defines a prototype as “a preliminary type, form, or instance of a system or system element that serves as a model for what comes later. They can be at the system level or can focus on subsystems or components” [20–24].

The DAU’s continuous learning module on Prototyping and Experimentation (CLE 082) uses the definition of a prototype model from DAU Acquipedia [20]:

A physical or virtual model used to evaluate the technical or manufacturing feasibility or military utility of a particular technology or process, concept, end item, or system. Prototype models have various types depending [on] the phase of the system life cycle. Prototype models range from the early development phase to production ready as from breadboard, brassboard, engineering development model to production repetitive models.

Note that a breadboard refers to an experimental device used in prototyping electronic circuits where components are “plugged” into the board. A breadboard enables temporary prototypes and circuit design experimentation in the laboratory [22]. A brassboard refers to an experimental or demonstration test model intended for field testing outside the laboratory environment. A brassboard follows the breadboard prototyping stage and contains both the functionality and approximate physical configuration of the final operational product [21].

The DoD Prototyping Guidebook, a living document that consolidates prototyping approaches, best practices, and recommendations into a single guide, offers this definition: “a model (e.g., physical, digital, conceptual, and analytical) built to evaluate and inform its feasibility or usefulness” [25].

With the Other Transaction (OT) Authority contracting mechanism becoming the preferred approach for certain eligible R&D efforts, the Office of the Under Secretary of Defense for Acquisition and Sustainment (OUSD[A&S]) defines a prototype project in the context of an OT in their Other Transactions Guide [26]:

A prototype project addresses a proof of concept, model, reverse engineering to address obsolescence, pilot, novel application of commercial technologies for defense purposes, agile development activity, creation, design, development, demonstration of technical or operational utility, or combinations of the foregoing. A process, including a business process, may be the subject of a prototype project. A prototype may be physical, virtual, or conceptual in nature.

The MITRE Systems Engineering Guide, which is based on MITRE’s application of systems engineering across the federally-funded research and development centers (FFRDCs) it operates for the U.S. government, offers a similarly broad definition spanning the life cycle [27]:

Prototyping is a practice in which an early sample or model of a system, capability, or process is built to answer specific questions about, give insight into, or reduce uncertainty or risk in many diverse areas. This includes exploring alternative concepts, technology maturity assessments, requirements discovery or refinement, design alternative assessments, and performance or suitability issues.

Collectively, these definitions cover a wide range of system hierarchy, element fidelity, product realization, and life cycle phase. Yet the conventional mindset in DoD research, development, test, and evaluation (RDT&E) of a “prototype” is not one of permission to proactively assess the realm of the possible. Rather, it is one of either a full-scale working model that closely resembles the final product that will be mass produced or a low-volume, reactive solution to a specific capability need. The root cause of these differing perspectives is unclear.
A search of DoD RDT&E regulations reveals further inconsistency in perspectives. The U.S. Army Regulation on Test and Evaluation Policy defines a prototype as “an article in final form employing standard parts and representative of articles to be produced on a production line with production tooling” [28].

The U.S. Air Force Test and Evaluation Guide defines a prototype as “a model suitable for evaluation of design, performance, and production potential. Note: The Air Force uses prototypes during development of a technology project or acquisition program for verification or demonstration of technical feasibility. Prototypes are not usually representative of the final production item” [29]. While not applicable to science and technology (S&T) programs that operate pre-Milestone A, the guide recommends that S&T activities follow its intent as much as possible and tailor applying its principles.

Neither the Navy’s Operational Test Director’s Manual [30] nor the DoD Test and Evaluation Management Guide [31], a generic reference written for all personnel involved in DoD acquisition management, provide a definition.

The Marine Corp’s Integrated Test and Evaluation Handbook also does not provide a definition but refers to using prototypes in pre-production qualification and production prove-out tests [32]. So, across Service RDT&E regulations, prototypes are considered both as near-final articles and a proactive means of evaluating feasibility.

After its independent review of the literature, the OUSD Research and Engineering (R&E) Emerging Capability and Prototyping (EC&P) concluded that the seemingly endless reasons for why S&T, R&D, and acquisition professionals conduct prototyping efforts boil down to generating information that supports a specific decision [25].

**HISTORY AND LAW**

When trying to understand the scope and complexity of an ubiquitous activity such as prototyping, it is important to map out the origin and history.

In 2009, the Weapon Systems Acquisition Reform Act (WSARA) [33] included a requirement that the Acquisition Strategy of each Major Defense Acquisition Program (MDAP) provide competitive prototyping prior to Milestone B approval. If a waiver for competitive prototyping was approved, then the program was required to produce a prototype prior to Milestone B approval—only if the life cycle’s benefit exceeded the cost; if full system prototyping was not feasible, then prototyping at the system or critical subsystem level may be required. With WSARA, near-final system prototypes were preferred, but the door was opened to build and evaluate prototypes at the subsystem level.

The prototyping requirements in WSARA initiated discussion on the role of prototyping in defense acquisition. Prototypes were then considered to be a valid means of solving a long-standing problem with the defense acquisition system (e.g., too slow, too cumbersome, too much risk, and too little information) and were acknowledged to have roles across the spectrum of concept generation, technology development, system integration, and test [34]. Further, prototypes were acknowledged to provide multiple opportunities, such as reducing risk early, fostering innovation, inspiring a new generation of designers and engineers, recruiting and retaining technical leaders, and increasing public interest [35].

The competitive prototyping requirements for MDAPs mandated by WSARA and codified as amended at 10 U.S.C. § 2430 were repealed by the NDAA for FY16 [36] and replaced with language referring to prototyping as one of several alternative risk management and mitigation approaches. The revised language is codified as amended at 10 U.S.C. § 2431b, where a comprehensive approach to technical, cost, and schedule risk management and mitigation in an acquisition strategy starts with the technique of “prototyping (including prototyping at the system, subsystem, or component level and competitive prototyping, where appropriate).”

The same public law (NDAA for FY16 [36]) that repealed portions of WSARA also produced Section 804—Middle Tier of Acquisition for Rapid Prototyping and Rapid Fielding—codified as amended at 10 U.S.C. § 2302, as well as Section 815—authority of the DoD to carry out certain prototype projects—codified as amended at 10 U.S.C. § 2371b. The acquisition-focused nature of these authorities requires that system prototypes be demonstrated in an operational environment within 5 years and potentially provide residual operational capability. However, 10 U.S.C. § 2371b gives the Director of the Defense Advanced Research Projects Agency (DARPA), the secretary of a military department, or any other official designated by the Secretary of Defense the authority to carry out prototype projects “that are directly relevant to enhancing the mission effectiveness of military personnel and the supporting platforms, systems, components, or materials proposed to be acquired or developed by the Department of Defense, or to improvement of platforms, systems, components, or materials in use by the armed forces.”
Although no specific guidance is available on how the research, innovation, disruption, and entrepreneur communities of the DoD should plan, fund, or implement prototyping efforts, 10 U.S.C. § 2371b enables application of agency funding to such efforts prior to the initiation of an acquisition program. (For the most relevant current directions and authorities pertaining to prototype projects, see the DoD’s Prototyping Guidebook [25].)

THE ROLE OF PROTOTYPING

The act of prototyping has historically been greatly beneficial in terms of risk reduction and concept demonstration prior to system development. Specifically, prototyping has advanced new technologies, enhanced industry workforce skills between major acquisitions, and dissuaded adversaries by showcasing new capabilities [3, 18, 35, 37].

A common perception of the engineering development process is that ideas flow linearly from the laboratory into prototyping, then through engineering and development, and eventually into production and sustainment. This linear model is incompatible with the current global environment, where the desire to emerge from competition as the victor drives both industry and government to a risk-tolerant, “fail early, fail often” approach in every phase of development [38–41].

Not all implementations of prototyping need to be 7-year, $700M [42] or 4-year, $938M [43] efforts performed by teams of defense contractors. There are organizations who employ prototyping as a tool and a strategy while operating outside of, and even independent of, defense acquisition, such as academia, internal R&D departments across private industry, FFRDCs, university-affiliated research centers, DARPA, and corporate laboratories within the Services. It is important to recognize that these organizations are not ineligible to create prototypes or prototype their ideas before making long-term investment decisions. To the contrary, this strategy is desired. Scientists and engineers routinely prototype their ideas to test hypotheses, reduce uncertainty, and explore feasibility [44, 45]. Whether hardware or software, these assume various names and forms, such as mock-ups, wire frames, breadboards, brassboards, proofs of concept, subscale models, and digital representations.

The DoD Prototyping Guidebook emphasizes prototyping as an enabler across all communities involved in system development, including exploratory S&T, R&D, and acquisition, regardless of whether the prototyping activity is occurring inside of, in support of, or completely independent of a program of record. The DAU course on prototyping and experimentation summarizes benefits and applications that span the development spectrum from exploration, to engineering, to acquisition, and even includes emerging capability shortfalls given the speed at which policies are adapting to the current national security climate. Prototyping is happening across the system’s life cycle and helps all communities understand the problem, come up with alternative solutions, assess the alternatives, learn (through success and failure), and make informed decisions.

Enabling rapid, but disciplined, progression from idea to prototype allows early and continual testing of ideas to screen for promising concepts and iterate to a validated solution before committing to the pursuit of operational viability and full, operational capability.

The purposes of prototyping are well documented. They are as follows:

- Generates information that supports specific decisions [46].
- Helps justify subsequent investments made in technology and technology maturation [25].
- Creates a preliminary version of something to resolve risk and explore operational potential [16].
- Ensures that new, innovative, and disruptive technologies are available to include in potential future systems and demonstrates the value of new technologies or systems [18].
- Increases user buy-in and participation, develops a better understanding of the product and its requirements, and reduces risk [47].
- Gains practical, operational knowledge and experience shared across defense industrial base, appeals to public interest in new technology, and inspires future innovators [35, 48].
- Helps bridge the gap between research and applications by enabling individuals and organizations with different technical backgrounds to exchange ideas in a common, intuitive, and understandable form [49].

Lauff’s in-depth study of prototyping reveals that prototypes are static objects until they are given meaning through the socially-constructed contexts and environments in which they are being used; they are a form of design language that enables communication, aids in learning, and informs decision making [50].

Horning et al. [51] promote prototyping as an operational strategy to reduce
development time and continually deliver mission-custom solutions by blending mission engineering, digital engineering, early synthetic prototyping, and advanced manufacturing.

Mullenburg and Gundo [52] promote prototypes as a means of quickly assessing design feasibility through trying out ideas. The authors present merits of a design-by-prototype process, applied to small, high-risk projects, using three case studies. In each case, designers and decision makers interacted with low-fidelity, tangible hardware mock-ups that informed final product decisions and led to a successful outcome [52].

Prototyping in research may not have anything to do with a user—neither a product nor a market may yet exist—and will instead focus on testing the feasibility of an idea and reducing uncertainty through a learning process [53]. An example is the transistor prototype created by Bell Telephone Laboratories researchers in late 1946 [54]. After an underpinning theory was verified through experimentation, researchers built a rough prototype device and demonstrated its functionality in the laboratory. A group of top engineers then spent the next 6 months considering applications of the technology. The transistor, a replacement for vacuum tube technology, is considered one of the most important technologies of the twentieth century.

BREAKING THE PARADIGM

Despite the documented flexibility in time and purpose of prototyping, the prevailing mental model in defense systems acquisition appears to be tied to a near-final, full-up system. One reason for this may be the use of the term in the Technology Readiness Level (TRL) scale [55] definitions and in the DoD Financial Management Regulation (FMR) [56] that defines an RDT&E budget activity (BA).

The TRL scale, conceived by the National Aeronautics and Space Administration (NASA) and tailored by each organization that implements the concept, is a measurement system that supports assessments of readiness, or maturity, of a technology on a scale of 1 (least mature) to 9 (most mature). Practical examples of each level of the TRL scale are provided by Grudo [57].

The RDT&E budget activities span from Activity 1 (basic research; systematic study tailored toward greater knowledge or understanding of a scientific principle) to Activity 7 (operational system development and upgrades).

The first appearance of the term “system prototype” in the TRL scale is at TRL 6, while the first appearance of the terms “prototype system” and “high-fidelity operating environment” in the FMR is at BA 6.4. On the surface, this presents a problem for the S&T community when they want to perform prototyping to reduce risk, gain knowledge, and prune decision paths early because S&T funding ends with BA 6.3. Though the TRL scale is the most widely used measure of technology maturity, organizations such as DoD, NASA, and the Department of Energy routinely tailor the definitions to suit their application and serve as a common language [58].

The current DoD paradigm of the relationships between technology maturity, funding, technology description [55], environment, and prototype category that should be reconsidered, in part because of its assumption of linear and sequential development, is summarized in Table 1. The linear model of innovation, where development progresses from basic research, to applied research, to development, and to production originated in the early 20th century and evolved over the next several decades as natural scientists, R&D business industry scholars, and economists built on the original taxonomy [59].

The prototype category column of Table 1 derives from the course Prototyping and Experimentation to Improve Acquisition [17], which subsequently formed the basis for an update to DAU CLE 082 [16] and resulted in the following descriptions of each of the three types of prototypes to clarify their roles in defense research, engineering, acquisition, and sustainment. Of note is the introduction of the conceptual prototype category, which aligns with the conventional budget and technology maturity categories found in S&T.

**Conceptual.** Demonstrate the art of the possible, provide evidence of overcoming specific technical risks and barriers, and evaluate S&T with a DoD corporate focus. Conceptual prototypes can also be used to support the analysis of a proof of concept or demonstrate feasibility prior to Milestone A. Conceptual prototype models are often breadboards and may be ready to demonstrate or prove in a laboratory environment.

**Developmental.** Validate the technical feasibility and explore the operational value of a capability that has already been proven in laboratory and relevant environments during the Technology Maturation & Risk Reduction phase. A developmental prototype will define the form, fit, function, and “ilities” of the system or technology.
Operational. Develop the technology or system so that Warfighters can use it in the field after it has been demonstrated in a realistic operational environment during the Production & Deployment or Operation & Support phases. An operational prototype can rapidly provide a needed capability to the field.

GAO recommends in numerous reports that technologies be demonstrated in a realistic (i.e., operational) environment prior to starting development. However, DoD permits an MDAP to proceed with development once the milestone decision authority certifies that the technology has been demonstrated in a relevant environment. This difference in demonstrated maturity level, commonly referred to as a technology “valley of death,” results in residual risk that a program of record is neither inclined nor budgeted to further reduce. (“Valley of death” is the colloquial phrase for when a technology is unable to cross ownership boundaries, e.g., from S&T into development, development into production, or production into sustainment. Some reasons include higher-than-acceptable residual risk, inadequate funding, and misaligned technological capability.)

From Table 2 [17], it is clear that prototyping has value across live, virtual, and constructive experimentation venues, even when technology maturity is low, primarily benefitting from gaining feedback on user needs.

The DoD Emerging Capability & Prototyping Office promotes a strategy that spans nearly the entire defense acquisition system and TRL scale, as shown in Figure 1 [60]. There are similarities between Proof of Principle, Pre-Engineering & Manufacturing Development (EMD), and Fieldable prototypes in Figure 1 and Conceptual, Developmental, and Operational prototypes in Table 2, but interchangeability should not be assumed. Instead, the take-away should be that prototyping has a valid role across the acquisition spectrum—from pre-conceptual feasibility testing through operational qualification and assessment—and helps to build up an understanding of system-level impacts.

**MOVING FORWARD WITH A STRATEGY**

With the concept of breaking the linear innovation and technology maturation paradigm now broached, a strategy is needed for how to implement and execute prototyping across the system’s development life cycle. Studies conclude that prototyping should be implemented as both a strategy and a tool to complement experimentation, aid in innovation, and help cross the S&T valley of death [1, 3, 4]. Choosing a strategy to implement prototyping practices in early S&T is not trivial, but the literature provides alternatives for classifying strategies based on purpose, motivation, and expected learning.

Lauff [50] compiles a comprehensive review of literature on prototype frameworks, taxonomies, and strategies and performs three case studies to understand state of the practice and document findings. A research outcome

<table>
<thead>
<tr>
<th>TRL</th>
<th>BA</th>
<th>TECHNOLOGY DESCRIPTION</th>
<th>ENVIRONMENT</th>
<th>PROTOTYPE CATEGORY</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>6.1</td>
<td>Basic principles and properties observed</td>
<td>Laboratory</td>
<td>Conceptual; demonstrate art of the possible and technical feasibility</td>
</tr>
<tr>
<td>2</td>
<td>6.1, 6.2</td>
<td>Speculation on potential applications</td>
<td>Relevant</td>
<td>Developmental; validate technical feasibility and explore operational value</td>
</tr>
<tr>
<td>3</td>
<td>6.2</td>
<td>Proof of concept, hypothesis testing through analysis and experimentation</td>
<td>Operational</td>
<td>Operational; develop for operational use</td>
</tr>
<tr>
<td>4</td>
<td>6.2, 6.3</td>
<td>Component, breadboard of low fidelity; improvised integration</td>
<td>Operational</td>
<td>Operational; develop for operational use</td>
</tr>
<tr>
<td>5</td>
<td>6.3</td>
<td>Component, brassboard of higher fidelity; elegant integration</td>
<td>Operational</td>
<td>Operational; develop for operational use</td>
</tr>
<tr>
<td>6</td>
<td>6.3</td>
<td>Prototype system closely resembling desired final configuration</td>
<td>Operational</td>
<td>Operational; develop for operational use</td>
</tr>
<tr>
<td>7</td>
<td>6.4</td>
<td>Prototype of actual final system</td>
<td>Operational</td>
<td>Operational; develop for operational use</td>
</tr>
<tr>
<td>8</td>
<td>6.5, 6.7</td>
<td>Actual final system qualified to function as expected</td>
<td>Operational</td>
<td>Operational; develop for operational use</td>
</tr>
<tr>
<td>9</td>
<td>6.5, 6.7</td>
<td>Actual final system proven through successful mission operations</td>
<td>Operational</td>
<td>Operational; develop for operational use</td>
</tr>
</tbody>
</table>
is a prototyping canvas that guides users to build the simplest prototype possible to quickly gather feedback through a problem statement, assumptions and questions, stakeholder interactions, a testing plan, resource identification, and an approach to a “minimum viable prototype.” With such a tool, the community can navigate ambiguity and reduce wasted resources [61, 62].

In a seminal work on prototyping, Drezner [63] stresses the importance of understanding timing, level in the system integration spectrum, and goals. The portion of Drezner’s proposed taxonomy that best suits prototyping during R&D starts with the purpose of technology viability, which focuses on generating information to reduce technological risk. Technology viability can be assessed outside the normal weapon system acquisition program structure and done without a specified military mission. The two objectives associated with this purpose in the taxonomy are experimentation to demonstrate a new idea, a new technology, or an existing technology in a new application and exploration to evaluate the possible performance envelope.

Table 2: Types of Prototypes and Experiments

<table>
<thead>
<tr>
<th>TYPES OF PROTOTYPES</th>
<th>TYPES OF EXPERIMENTS</th>
<th>Source: DASD (EC&amp;P) RRTO; Prototyping and Experimentation to Improve Acquisition</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Operational Level (OL) (Fieldable)</strong>&lt;br&gt;TRLs 8-9&lt;br&gt;BAs 5 &amp; 7</td>
<td><strong>Live</strong>&lt;br&gt;(Actual People/Systems in a Realistic Physical Environment)&lt;br&gt;<strong>What:</strong> Actual OL prototype.&lt;br&gt;<strong>Why:</strong> Most realism. Also, no technical, cost, safety, or environmental constraints.&lt;br&gt;<strong>Warfighters:</strong> If involved, trained on the OL prototype.&lt;br&gt;<strong>Benefit:</strong> Validate OL prototype benefits in an operational environment.</td>
<td></td>
</tr>
<tr>
<td><strong>Developmental (Usable in a Pinch)</strong>&lt;br&gt;TRLs 5-7&lt;br&gt;BAs 3-4</td>
<td><strong>Virtual</strong>&lt;br&gt;(Actual People/Systems Interacting With Mock-ups and a CG Environment)&lt;br&gt;<strong>What:</strong> Mock-up of OL prototype.&lt;br&gt;<strong>Why:</strong> Cost, environmental, safety, or other constraints prohibit use of actual OL prototype.&lt;br&gt;<strong>Warfighters:</strong> If involved, trained on the mock-up.&lt;br&gt;<strong>Benefit:</strong> Validate (to a lesser degree) benefits in an operational environment.</td>
<td></td>
</tr>
<tr>
<td><strong>Conceptual (Not Fieldable)</strong>&lt;br&gt;TRLs 1-4&lt;br&gt;BAs 1-3</td>
<td><strong>Constructive</strong>&lt;br&gt;(CG People/Systems in a CG Environment)&lt;br&gt;<strong>What:</strong> CG representation (e.g., model) of OL prototype.&lt;br&gt;<strong>Why:</strong> Cost, environmental, safety, or other constraints prohibit use of actual or mock-up of OL prototype.&lt;br&gt;<strong>Warfighters:</strong> If involved, trained on the CG version.&lt;br&gt;<strong>Benefit:</strong> Validate (to a lesser degree) benefits in an operational environment.</td>
<td></td>
</tr>
</tbody>
</table>

---

**BA** = Budget Activity  
**CG** = Computer Generated  
**OL** = Operational Level  
**TRL** = Technology Readiness Level
Carr and Verner [53] present a framework based on their experience in software development and promote prototypes as instruments used strategically throughout the development process. Three approaches are described—exploratory, experimental, and evolutionary. Exploratory prototyping is used to engage with customers early to discover and clarify requirements, elicit value and preferences, and understand how the technology will be used. At this level, a “quick-and-dirty” or “throw-it-away” prototype can serve as a partial implementation of the system and may not look anything like the anticipated final system; the objective is to learn.

Experimental prototyping uses a breadboard approach to assess feasibility of new ideas and features. At this level, the prototype varies from partial implementation to a mock-up of the anticipated final solution.

Evolutionary prototyping is used in environments of high uncertainty, e.g., when requirements are not all known ahead of time and when the user may neither know how the technology will be used nor the environment in which it is intended to operate. The evolutionary approach enables adaptation to continuous learning of user needs, external factors, and requirements.

Beynon-Davies et al. [47] develop a taxonomy of information system prototype practices and further
promote prototypes as a tool and a technique. Three forms of prototype are described—throwaway, incremental, and evolutionary. Throwaway prototypes are used to gain a specific piece of knowledge and then discarded (exploratory) or continuously test the feasibility of some feature (experimental) offline. An incremental prototype is refined gradually and becomes either a part of the final delivered system or the delivered system itself (e.g., block 1 has limited functionality, block 2 has increased functionality, etc.). An evolutionary prototype is part of a system that is to be delivered in increments (e.g., new features, modules, or plug-ins added over time until full functional capability has been achieved). Incremental and evolutionary prototypes are intended for eventual operational use, whereas throwaway prototypes are intended for offline learning.

Menold et al. [64] compile an extensive literature review related to prototyping frameworks and strategies. Their objective was to help bridge the gap between research and practice by providing designers with a structured set of methods for prototyping activities. The resulting framework is based on three phases—frame, build, and test. In this way, implementers can focus on testing assumptions for gaining knowledge.

Lichter et al. [65] identify a difference between prototyping in practice and prototyping in theory and use case studies to validate a framework that describes prototypes using kinds (e.g., presentation, prototype proper, breadboard, and pilot system), goals (e.g., exploratory, experimental, and evolutionary), construction techniques (e.g., horizontal and vertical), and relationship to application system (e.g., building block, throwaway, and problem clarification). The framework helps describe prototypes and promote them as a basis for discussion via experimentation.

Implementing prototyping as a tool and a strategy shifts the focus from near-final design validation to a purposeful and methodical approach of a manageable number of unknowns and hypotheses across the capability development span. In this way, it becomes a vehicle for early and continuous learning that can guide investments in technologies to close threat-based gaps in a rapidly-innovating global environment.

CONCLUSIONS AND RECOMMENDATIONS

In the current global environment, the aim is to deliver the right capabilities to the Warfighter, quickly and efficiently. There is increased emphasis within the DoD on using prototyping and experimentation to explore new capabilities and reduce technical, cost, and schedule risk prior to entering systems acquisition [16, 18, 67]. Further, prototyping is seen as an enabler for innovation [4, 41], which is important for the DoD as it seeks innovation across operational concepts, organizational structure, business processes, and technology [66].

Prototyping offers an innovative approach to solving technical challenges and a scientific approach to answering research questions, potentially through invention. However, the focus of prototyping in acquisition reform has been on integrating existing technologies to form a new or enhanced system-level capability that intends to be fielded from the outset. The entire defense development community can benefit by treating prototyping as a license to explore and test hypotheses, reduce risk, and learn on an object that may not resemble the anticipated final system in form, fit, or function. While acquisition and sustainment considers prototyping as part of acquisition agility [68], perhaps R&E can think of prototyping as “knowledge agility.”

The conditions are present for prototyping and experimentation to provide early and enduring benefit, from knowledge generation all the way through product sustainment. For example, in the pursuit of informing future concepts vis-à-vis a novel propulsion system, an organization might test basic hypotheses of bearings, lubricants, and transmissions on live, virtual, or constructive mock-ups as opposed to building an entire rotorcraft vehicle. In this case, prototyping at the subsystem and component levels will help reveal where innovation can address a gap and where a scientific breakthrough might be needed to address a lagging subsystem capability, such as power density or mass.

The DoD should consider reexamining the paradigm of prototypes aligned with budget activity and assessments of technology maturity. This way, technology maturity, prototype maturity, and integrated system maturity can take place independently and potentially complement each other instead of competing with each other. As part of a formal technology readiness assessment for critical technologies, the TRL scale is used to provide a consistent maturity evaluation standard, not to assess the readiness of any technology based on prototype test results at increasingly higher levels of system hierarchy.
Perhaps the TRL scale can be modified to include a reference to prototypes at every level, 1 through 9.

The anticipated time and cost benefits of prototyping will only be realized through a targeted and intentional strategy based on learning objectives. Various frameworks, taxonomies, and strategies are documented in the literature and readily available for tailoring.

Research shows the importance of prototyping across the research and technology development spectrum and reveals that various communities use prototypes as leverage to generate knowledge independent of technology maturity levels and budget activity. A recommendation is made for relaxing definitions and policy constraints so that all communities can benefit. A common goal is generating knowledge and unlocking decision paths to get a product into the hands of the Warfighter. Prototyping offers a way to intelligently and methodically evaluate feasibility, resolve risks, refine requirements, gain stakeholder buy-in, and assess military utility. Efficiency and effectiveness in acquisition is the goal, and prototyping ideas early is an enabling solution.

DSIAC was subsequently requested to conduct a technical peer review of a position paper authored by ARL on the span of prototyping in defense research, engineering, acquisition, and sustainment. That discussion paper evolved into the article presented in this journal.

REFERENCES


NOTE FROM THE EDITOR

DSIAC regularly provides information research capabilities, technical expertise, and a network of subject matter experts to U.S. Army Research Laboratory (ARL) engineers and scientists to enhance their technical efforts. Last year, ARL asked DSIAC to identify documented examples in support of their research effort to use early prototyping for science concepts and ideas in hardware and software.
Are you ready
to shape the future of the world’s greatest military?

Our service members
are counting on you to field cutting-edge technologies
to maintain their edge on the battlefield.

Let Defense Technical Information Center (DTIC)
help you get started. For almost 75 years, DTIC has served the information needs of
the defense communities through the collection
and preservation of the DoD’s multi-billion dollar
annual investment in science and technology.
We have over 4.5 million S&T documents in our
collection—technical research, budget reports,
conference proceedings,
grant awards, and international agreements
—to enable eligible users from the DoD labs, academia, federally-funded research and
development centers, and industry partners to
better align their resources, test novel concepts,
and develop new capabilities. Our tools and re-
sources provide insights on research outcomes,
technology development, and maturity; enable
tracking of DoD investments and capability gaps;
and help identify partnership opportunities.

Give us a try.
We’re ready for you.
Visit: https://go.usa.gov/xd622 today.
DSIAC ONLINE

www.dsiac.org

DSIAC PRODUCTS AND SERVICES INCLUDE:

• Performing literature searches.
• Providing requested documents.
• Answering technical questions.
• Providing referrals to subject matter experts (SMEs).
• Collecting, electronically cataloging, preserving, and disseminating Defense Systems scientific and technical information (STI) to qualified users.
• Developing and deploying products, tools, and training based on the needs of the Defense Systems community.
• Fostering and supporting the DSIAC technical Communities of Practice.
• Participating in key DoD conferences and forums to engage and network with the S&T community.
• Performing customer-funded Core Analysis Tasks (CATs) under pre-competed IDIQ Delivery Orders.

DSIAC SCOPE AREAS INCLUDE:

• Advanced Materials
• Autonomous Systems
• Directed Energy
• Energetics
• Military Sensing
• Non-Lethal Weapons
• Reliability, Maintainability, Quality, Supportability, and Interoperability (RMQSI)
• Survivability and Vulnerability
• Weapon Systems

CONNECT WITH US ON SOCIAL MEDIA!