
# Seeing Photons: Progress and Limits of Visible and Infrared Sensor Arrays

ISBN
978-0-309-15304-1

194 pages
7 x 10
PAPERBACK (2010)

Committee on Developments in Detector Technologies; National Research Council

# SEEING PHOTONS

## PROGRESS AND LIMITS OF VISIBLE AND INFRARED SENSOR ARRAYS

Committee on Developments in Detector Technologies

Standing Committee on Technology Insight—Gauge, Evaluate, and Review

Division on Engineering and Physical Sciences

NATIONAL RESEARCH COUNCIL
*OF THE NATIONAL ACADEMIES*

THE NATIONAL ACADEMIES PRESS
Washington, D.C.
**www.nap.edu**

**THE NATIONAL ACADEMIES PRESS 500 Fifth Street, N.W. Washington, DC 20001**

NOTICE: The project that is the subject of this report was approved by the Governing Board of the National Research Council, whose members are drawn from the councils of the National Academy of Sciences, the National Academy of Engineering, and the Institute of Medicine. The members of the committee responsible for the report were chosen for their special competences and with regard for appropriate balance.

This is a report of work supported by contract HHM402-05-D-0011 between the Defense Intelligence Agency and the National Academy of Sciences. Any opinions, findings, conclusions, or recommendations expressed in this publication are those of the author(s) and do not necessarily reflect the view of the organizations or agencies that provided support for the project.

International Standard Book Number-13: 978-0-309-15304-1
International Standard Book Number-10: 0-309-15304-2

*Limited copies are available from*
Division on Engineering and Physical Sciences
National Research Council
500 Fifth Street, N.W.
Washington, DC 20001
(202) 334-3118

*Additional copies are available from*
The National Academies Press
500 Fifth Street, N.W.
Lockbox 285
Washington, DC 2001
(800) 624-6242 or (202) 334-3313
(in the Washington metropolitan area)
http://www.nap.edu



Printed in the United States of America

## COMMITTEE ON DEVELOPMENTS IN DETECTOR TECHNOLOGIES

STEVEN R.J. BRUECK, *Chair*, University of New Mexico
PAUL McMANAMON, *Vice Chair*, Exciting Technology, LLC
STEFAN BAUR, Raytheon Vision Systems
VALERIE BROWNING, ValTech Solutions, LLC
JOHN DEVITT, Georgia Tech Research Institute
THOMAS HARTWICK, Independent Consultant
ANGELA HODGE, Johns Hopkins University, Applied Physics Laboratory
MICHAEL HOPMEIER, Unconventional Concepts, Inc.
STEVEN JOST, BAE Systems
LINDA KATEHI, University of California, Davis
SEETHAMBAL MANI, Sandia National Laboratories
C. KUMAR PATEL, Pranalytica, Inc.
TAMAR PELI, Charles Stark Draper Laboratory
DAVID SHAVER, Massachusetts Institute of Technology, Lincoln Laboratory
JONATHAN SMITH, University of Pennsylvania

*Staff*

CARTER W. FORD, Study Director
NORMAN HALLER, Consultant
SARAH CAPOTE, Research Associate
MARGUERITE SCHNEIDER, Administrative Coordinator
URRIKKA WOODS, Program Associate (to February 26, 2010)

**STANDING COMMITTEE ON TECHNOLOGY INSIGHT—GAUGE, EVALUATE, AND REVIEW**

RUTH A. DAVID, *Chair*, ANSER, Inc.
BRIAN BALLARD, ARES Systems Group
STEVEN R.J. BRUECK, University of New Mexico
ANN N. CAMPBELL, Sandia National Laboratories
ALAN H. EPSTEIN, Pratt and Whitney
JOHN GANNON, BAE Systems
SHARON C. GLOTZER, University of Michigan
CHRISTOPHER C. GREEN, Wayne State University
DIANE E. GRIFFIN, Johns Hopkins Bloomberg School of Public Health
J.C. HERZ, Batchtags, LLC
ALLISON HICKEY, Accenture National Security Services
J. JEROME HOLTON, ARES Systems Group
KENNETH KRESS, KBK Consulting
FREDERICK R. LOPEZ, AMC Incorporated
GILMAN G. LOUIE, Alsop-Louie Partners
STUART PARKIN, IBM Almaden Research Center
JULIE J.C.H. RYAN, George Washington University

*Staff*

MICHAEL A. CLARKE, Lead DEPS Board Director
CARTER W. FORD, Program Officer
DANIEL E.J. TALMAGE, JR., Program Officer
MARGUERITE SCHNEIDER, Administrative Coordinator
KAMARA BROWN, Research Associate
SARAH CAPOTE, Research Associate
SHANNON THOMAS, Program Associate

# Preface

The Department of Defense has recently highlighted intelligence, surveillance, and reconnaissance (ISR) capabilities as a top priority for U.S. warfighters. Contributions provided by ISR assets in the operational theaters in Iraq and Afghanistan have been widely documented in press reporting. While the United States continues to increase investments in ISR capabilities, other nations not friendly to the United States will continue to seek countermeasures to U.S. capabilities.

The Technology Warning Division of the Defense Intelligence Agency's (DIA's) Defense Warning Office (DWO) has the critical responsibility, in collaboration with other components of the intelligence community (IC), for providing U.S. policy makers insight into technological developments that may impact future U.S. warfighting capabilities. To this end, the IC requested that the National Research Council (NRC) investigate and report on key visible and infrared detector technologies, with potential military utility, that are likely to be developed in the next 10-15 years. This study is the eighth in a series sponsored by the DWO and executed under the auspices of the NRC TIGER (Technology Insight—Gauge, Evaluate, and Review) Standing Committee.

A committee of experts in the scientific and technical areas relating to visible and infrared detectors was formed to conduct this study. Faced with a relatively short time frame for completing the study, the committee very much appreciates the timely and informed cooperation of the IC members who sponsored the study, as well as the many government, industry, and university participants who contributed valuable information during the committee's meetings.

We wish to thank all of the committee members for their efforts in producing this report in a very short period of less than four months from first meeting to peer review. In addition, the peer reviewers and monitor provided insightful and useful comments that improved the quality of the report. A sincere thank you is due to the NRC staff including Carter Ford, Sarah Capote, Marguerite Schneider, and Urrikka Woods. The contributions of Norm Haller in providing organizational and technical writing assistance were also of immense value to the committee.

Steven R.J. Brueck, *Chair*
Paul McManamon, *Vice Chair*
Committee on Developments in Detector Technologies

# Acknowledgment of Reviewers

This report has been reviewed in draft form by individuals chosen for their diverse perspectives and technical expertise, in accordance with procedures approved by the National Research Council's (NRC's) Report Review Committee. The purpose of this independent review is to provide candid and critical comments that will assist the institution in making its published report as sound as possible and to ensure that the report meets institutional standards for objectivity, evidence, and responsiveness to the study charge. The review comments and draft manuscript remain confidential to protect the integrity of the deliberative process. We wish to thank the following individuals for their review of this report:

R. Stephen Berry, University of Chicago
James Coleman, University of Illinois
Ruth David, Anser, Inc.
Donald Gaver, Naval Postgraduate School
Anthony Hyder, University of Notre Dame
Kenneth Kress, KBK Consulting, Inc.
Robert Latiff, Science Applications International Corporation
Manijeh Razeghi, Northwestern University.

Although the reviewers listed above have provided many constructive comments and suggestions, they were not asked to endorse the conclusions or recommendations nor did they see the final draft of the report before its release. The review of this report was overseen by Elsa Garmire, Dartmouth College. Appointed by the NRC, she was responsible for making certain that an independent examination of this report was carried out in accordance with institutional procedures and that all review comments were carefully considered. Responsibility for the final content of this report rests entirely with the authoring committee and the institution.

# Contents

APPENDIXES

# Acronyms

| | |
|---|---|
| ABL | airborne laser |
| ADC | analog-to-digital converter |
| AGC | automatic gain control |
| APD | avalanche photodiode |
| APS | active pixel sensor |
| ARGUS-IS | Autonomous Real-time Ground Ubiquitous Surveillance Imaging System |
| ARTEMIS | Advanced Responsive Tactically Effective Military Imaging Spectrometer |
| ASAT | antisatellite (capability) |
| ASIC | application-specific integrated circuit |
| ATC/R | automatic target cuing-recognition |
| | |
| BLIP | background-limited infrared photodetection |
| | |
| CCD | charge-coupled device |
| CDL | Common Data Link |
| CMOS | complementary metal oxide semiconductor |
| CNT | carbon nanotube |
| COP | coefficient of performance |
| COTS | commercial off-the-shelf |
| CPU | central processing unit |
| CTE | charge transfer efficiency |

| | |
|---|---|
| CTIA | capacitor transimpedance amplifier |
| CZT | cadmium zinc telluride |
| | |
| 3-D | three-dimensional |
| DARPA | Defense Advanced Research Projects Agency |
| DCT | discrete cosine transform |
| DIA | Defense Intelligence Agency |
| DOD | Department of Defense |
| DRAM | dynamic random access memory |
| DWELL | (Quantum) Dots in a (quantum) well detector |
| DWO | Defense Warning Office |
| DWT | discrete wavelet transform |
| | |
| EMI | electromagnetic interference |
| EO | electro-optical |
| | |
| FET | field-effect transistor |
| FLIR | forward-looking infrared |
| FOV | field of view |
| FPA | focal plane array |
| FPDP | front-panel data port |
| FPGA | field-programmable gate array |
| | |
| GEO | geosynchronous orbit |
| GM-APD | geiger mode avalanche photodiode |
| GPU | graphics processing unit |
| G-R | generation-recombination |
| | |
| HDMI | high-definition multimedia interface |
| HEO | high Earth orbit |
| | |
| IC | integrated circuit; intelligence community |
| IED | improvised explosive device |
| IR | infrared |
| IRST | infrared search and track system |
| ISR | intelligence, surveillance, and reconnaissance |
| ITAR | International Traffic in Arms Regulations |
| ITO | indium tin oxide |
| | |
| JPEG | Joint Photographic Experts Group |
| JWST | James Webb Space Telescope |

| | |
|---|---|
| LACOSTE | Large Area Coverage Optical Search-while-Track and Engage |
| LADAR | laser detection and ranging |
| LED | light-emitting diode |
| LEO | low Earth orbit |
| LPE | liquid-phase epitaxy |
| LRU | line-replaceable unit |
| LVDS | low-voltage differential signaling |
| LWIR | long-wavelength infrared |
| | |
| MBE | molecular beam epitaxy |
| MCT | mercury cadmium telluride |
| MDE | Multicore Development Environment |
| MEMS | microelectromechanical system |
| MEO | middle Earth orbit |
| MGM | metal-graphene-metal |
| MOS | metal oxide semiconductor |
| MTF | modulation transfer function |
| MWIR | mid-wavelength infrared |
| | |
| NA | numerical aperture |
| NASA | National Aeronautics and Space Administration |
| NATO | North Atlantic Treaty Organization |
| NEP | noise-equivalent power |
| NIR | near infrared |
| NNI | National Nanotechnology Initiative |
| NRC | National Research Council |
| NRDA | National Research and Development Act |
| NTIS | National Technical Information Service |
| NUC | nonuniformity correction |
| | |
| OTCCD | orthogonal transfer CCD |
| | |
| PCR | polymerase chain reaction |
| PnC | phononic crystal |
| | |
| QDIP | quantum-dot IR photodetector |
| QE | quantum efficiency |
| QWIP | quantum-well IR photodetector |
| | |
| R&D | research and development |
| RF | radio frequency |

| | |
|---|---|
| RGB | red, green, blue (color model) |
| ROIC | readout integrated circuit |
| RTG | radioisotope thermoelectric generator |
| RTI | Research Triangle Institute |
| | |
| SERS | surface-enhanced Raman scattering |
| Si PIN | diode with an intrinsic silicon layer between the P- and N-type regions |
| SITP | Shanghai Institute of Technical Physics |
| SLS | strain-layer superlattice |
| SOI | silicon on insulator |
| SOT | statement of task |
| SPD | single-photon detector |
| SPW | surface plasma wave |
| SWaP | size, weight, and power |
| SWIR | short-wavelength infrared |
| SWNT | single-wall carbon nanotube |
| | |
| TDI | time delay-and-integrate |
| TE | thermoelectric |
| TIGER | Standing Committee on Technology Insight—Gauge, Evaluate, and Review |
| TOMBO | Thin Observation Module using Bound Optics |
| | |
| UGS | unattended ground sensor |
| UV | ultraviolet |
| | |
| VLWIR | very long wavelength infrared |

# Summary

## CONTEXT AND TASKING

Detector technologies for both military and civilian applications have evolved over many years to a sophisticated state of current development.[1] Advanced technologies, such as nanoscale-engineered materials, will provide flexibility and functionality in the design and development of future sensor systems and their components. The increasing availability of commercial products will also impact detector-based electro-optical (EO) and infrared (IR) systems and lead to new sensor system-level capabilities.

At the same time, mission needs will change, and sensor system designs and capabilities over the next 10-15 years will need to evolve to match these changed mission needs. New EO-IR sensor system challenges—processing, storing, and communicating—are arising from the enormous increase in data generated as a result of the proliferation of more and ever-higher-pixel-count sensors. Generating data is not the same as providing actionable intelligence; this requires conversion of the data into usable information.

Leveraging of commodity-level developments enables unprecedented capabilities for technologically advanced nation-states and, simultaneously, lowers the barrier to entry for non-state, transnational groups to pose asymmetric threats.

[1]For this report, the committee defines a *detector* as representing a single pixel that receives photons. A *focal plane array* (FPA) is composed of many detectors arranged in a two-dimensional grid and generates an image. A *sensor system* is composed of FPAs plus other components, such as signal processing, data transmission, coolers, optics, and pointing and tracking mechanisms.

Funding, influenced by military needs and commercial market conditions, will drive investments. Commercial funding is expected to be at a greater level than military funding, but this will be restricted to commodity areas with the potential for large-volume manufacturing. Military needs will likely leverage commercial off-the-shelf capabilities in areas such as advanced semiconductor manufacturing tools.

In this overall context, the intelligence community (IC)[2] asked the National Research Council (NRC) to conduct an in-depth technical assessment of detector technologies. Specifically, the NRC was asked to do the following:

- Consider the fundamental, physical limits to optical and infrared detector technologies with potential military utility, with priority on passive imaging systems. Elucidate trade-offs between sensitivity, spectral bandwidth and diversity, dynamic range, polarization sensitivity, operation temperature, and so forth. Compare these limits to the near-term state of the art, identifying the scaling laws and hurdles currently restricting progress.[3]
- Identify key technologies that may help bridge the gaps within a 10-15 year time frame, the implications for future military applications, and any significant indicators of programs to develop such applications. Speculate on technologies and applications of relevance that are high-impact wild cards or have a low probability of feasible deployment within 15 years. Discuss trends in availability and format scalability and in available cooling technologies.
- Consider the pros and cons of implementing each existing or emerging technology, such as noise, dynamic range, processing or bandwidth bottlenecks, hardening, power consumption, weight, et cetera.
- Identify which entities currently lead worldwide funding, research, and development for the key technologies. Highlight the scale, scope, and particular strengths of these R&D efforts, as well as predicted trends, time scales, and commercial drivers.

## THE BOTTOM LINE

### Fundamentals of Visible and Infrared Detectors

There are fundamental limits to detection. To be seen at all, an object must emit or reflect electromagnetic radiation in some spectral band. That radiation

[2]According to Intelligence.gov, the IC is composed of 17 federal agencies. Accessed March 24, 2010.

[3]In several consultations with the committee over a period of months, the sponsor requested that the committee address the imaging spectrum from ultraviolet to very longwave infrared.

must pass though the medium between the object and the sensor. For cross-link, space-based sensor systems, the transmission medium is not a limitation; however, if either the sensor system or the object being viewed is within the atmosphere, atmospheric transmission must be taken into account. Difficult new mission requirements, such as viewing objects below the surface of water or behind strongly scattering media (e.g., foliage), require creative combinations of sensor technologies. Finally, the received electromagnetic information must be transduced into another form, usually an electrical signal, with sufficient signal-to-noise ratio to allow further extraction of information.

Developments in detection have a long history. Both visible and IR detector technologies have undergone significant maturation, and high-performance detectors are available across most spectral bands from 0.2 to 20 μm. The most sensitive IR detectors require cooling to reduce dark current noise and reach background-limited IR photo detection (BLIP), resulting in an increase in size, weight, and power (SWaP), as well as cost, along with a reduction in reliability. Single-photon detectors are available today in the visible and near IR; there are active research efforts to extend this capability throughout the IR. In accordance with the statement of task, this report emphasizes passive sensing; however, developments in active sensing are included as appropriate throughout the report.

In spite of the fact that there are high-end sensor systems capable of close to theoretical sensitivity limits in most bands, significant improvements remain possible for sensor systems by adding functionality, such as multi- and hyperspectral response, polarimetric sensitivity, dynamic resolution, and sensitivity adaptation, as well as reductions in SWaP and cost. Certainly, processing and communications requirements and capabilities will continue to drive improvements in sensor systems. Some of these improvements are fundamental to the detectors or sensor systems, and some are in the ancillary components, such as optics, cooling, pointing and tracking, data handling, and compression.

## Key Current Technologies and Evolutionary Developments

A relatively new technology relates to advances in solid-state detector materials. These advances tend to be used to render immaterial the sources of noise downstream from the detector. Initially, most detectors with gain tended to use linear gain; however, more recently, the significant advantages of uncontrolled avalanche gain, called Geiger mode operation, in which an arbitrarily large number of electrons are released based upon the arrival of a single photon, are creating new imaging modalities. In addition, there remains considerable opportunity to improve other parameters, such as operating temperature, power dissipation, manufacturability, and cost.

Going forward, many of the advances in detector technologies will be in “peripheral" areas. One important area is cooling, particularly for IR sensors. Historically, for mid-wavelength IR (MWIR) and long-wavelength IR (LWIR) detectors, cooling has been a major limitation. There has been a push toward higher-temperature detector operation, based on developing detector technologies that have lower dark current at a given operating temperature. Progress has been slow, however, and considerable room for improvement remains. An alternate approach is to reduce the SWaP requirements of cooling.

The increase in digital processing capabilities, fueled by the semiconductor industry, is a further trend that will continue to have a major effect on sensor systems. Digital processing systems can be adaptable and allow customization for specific applications. Lowering cost can also make a detector technology much more widely available and cause its impact at the systems level to be greatly increased.

Tracking novel adaptations of widely available and inexpensive imagers will continue to be of interest to the IC. One example is the Defense Advanced Research Projects Agency's (DARPA's) Autonomous Real-time Ground Ubiquitous Surveillance-Imaging System (ARGUS-IS) program that involves integrating a large number of cell phone camera chips to provide a revolutionary 1.8 gigapixel imager. Consumer demands for improved, higher-pixel-count, cell phone imagers, which did not even exist until recently, made this revolutionary imaging capability possible. For countries that do not invest significant funding in purpose-built imaging technology, the development of low-cost commodity imagers has significantly lowered the barriers to having a militarily significant imaging capability. The use of large-volume commercial sensors can enable new capabilities for both less advanced asymmetric adversaries and near-peers alike.

**KEY FINDING**
The evolutionary trends are semiconductor detectors characterized by increased pixel pitch and count, higher readout speed, higher operating temperature (especially MWIR), lower power consumption, and decreased sensor thickness. The need for larger fields of regard is a significant driver for larger arrays. Even beyond the diffraction limit of the optical system, oversampling can lead to slightly enhanced resolution.

**KEY FINDING**
The global proliferation of low-cost, commodity imagers, such as cell phone cameras and automobile thermal imagers, enables adversaries to develop sensing systems at relatively low cost, reducing the barrier to achieving limited operational capabilities. As an example, the rapid proliferation of low-cost "night vision technology" is eroding the overwhelming dominance of the United States in nighttime operations, even with the superior performance of advanced systems.

**KEY FINDING**
The availability of very low cost imagers developed for large consumer markets is providing opportunities to develop new sensor systems and architectures, even though the component-level imagers may not have the capabilities typical of high-performance sensors developed specifically for military applications. Additionally, the technology and manufacturing base used to make these low-cost imagers will extend the manufacturing base that can be used for fabricating customized military parts.

**RECOMMENDATION 3-1**
The intelligence community should pay careful attention to the new capabilities inherent in both the proliferation of commodity detector technologies and their integration into novel sensor systems. ARGUS-IS and Gnuradio are examples of how available, low-cost, mature commodity visible focal plane array (FPA) technology (cell phone camera chips) and commercial off-the-shelf (COTS) communications circuitry, through sensor integration, have enabled new, advanced, high-performance imaging capabilities.

**KEY FINDING**
Existing, mature mercury cadmium telluride, indium antimonide, indium gallium arsenide, silicon charge-coupled devices, silicon complementary metal oxide semiconductors, and avalanche photodiode focal plane technologies provide sensors with excellent performance and set a very high barrier to entry for any emerging technology. For some performance parameters, such as detectivity, mature imager technologies already are operating very close to fundamental limits. However, there is still considerable opportunity to improve other parameters such as operating temperature, power dissipation, manufacturability, and cost.

**KEY FINDING**
Rapid progress is being made in the development of closely related single-photon and photon counting detectors and arrays. Single-photon detection and photon counting imagers are key enablers for a wide range of new secure communications, passive sensors, three-dimensional laser detection and ranging, and active optical sensors. Specifically, quantum cryptography relies on the distribution of entangled, single-photon qubits (keys) between the transmitter and receiver; this is inherently a single-photon process. In most cases, these applications involve physical processes in which only a small number of photons are available for detection. These detectors require high quantum efficiencies, low dark count rates, fast recovery times, and capabilities for photon number resolving.

TABLE S-1 Trigger Points of Technical Progress and Their Implications

| Single Photon | 2010 (SOA) | 2015 (TP) | 2020 (TP) | 2025 (TP) |
|---|---|---|---|---|
| Q efficiency | 90% | >90% | >90% | >90% |
| Speed | 10 GHz | 100 GHz | THz | THz |
| Wavelength | Visible | 1.55 μm | 1.55 μm | 1.55 μm |
| Operating temperature | 4 K | 77 K | 300 K | 300 K |
| Application | | QKD | QKD/quantum computer | Quantum computer |

NOTE: QKD = quantum key distribution; SOA = state of the art; TP = trigger point, which indicates a capability that should stimulate the intelligence community to do significant collection and/or analysis.

**RECOMMENDATION 3-2**
The intelligence community should carefully track developments related to single-photon and photon counting detectors across the full spectrum, from the ultraviolet to very long wavelength infrared. Table S-1 lists trigger events that would cause a significant shift in capability and should be carefully monitored by the intelligence community.

**KEY FINDING**
There is significant opportunity to customize image sensor architectures for specific applications that can lead to dramatic improvements in system-level performance, including size, weight, and power. Advanced architectural design, including integration of sensing and processing (in-pixel and on-chip), can have greater system-level impact than making small gains in driving detector performance incrementally closer to fundamental detectivity limits.

**RECOMMENDATION 3-3**
The intelligence community should evaluate and track system capabilities rather than focusing solely on component technical achievements. These include technologies that enable in-pixel and on-chip processing, lower-power operation, and higher operating temperatures, as well as technologies that improve manufacturability.

**KEY FINDING**
For both cryocooler and thermoelectric cooler technologies, there are a number of commercial market drivers, separate from sensor cooling applications, that will drive evolutionary improvements in SWaP. Over the next 10-15 years, it is reasonable to expect that these improvements will achieve overall reductions in SWaP on the order of 20-30 percent.

**Emerging Technologies with Potentially Significant Impacts**

The user always wants more resolution and a wider field of view. Resolution is limited by diffraction, but field of view is not limited in a fundamental manner. Developers continue to try to meet these ever-increasing demands. Users are willing to pay for development to meet these needs, especially an increase in field of regard without sacrificing available resolution, which directly leads to increased pixel count and larger-area arrays.

Emerging thermoelectric, phononic crystal, and laser cooling technologies offer potential for improving sensor systems, because such technologies might be able to replace the cooling furnished by current bulk coolers, with their attendant SwaP penalties.

Another peripheral area that has a major impact is the ability to handle the vast amount of data generated. There are many new sensors coming along that generate large amounts of data. Hyperspectral sensor systems, for example, generate significant data volumes as a result of the additional spectral dimension. Digital developments are not driven by the relatively small number of sensor systems. The gaming industry has a much more significant impact on digital progress, and advances in computation and communication will have a major impact on sensor technology.

Countries around the world are poised to take advantage of nanotechnology to potentially build entirely new sensors and sensor systems. Therefore, international progress in the nanotechnology field constitutes a principal driver for significant advances in sensors.

**KEY FINDING**
Thin-film thermoelectric devices have the potential to substantially reduce size, weight, and power requirements of the active cooling component for room-temperature focal plane arrays. If these devices can meet cost and lifetime metrics, they will displace the currently used bulk coolers. The near-term driver for these developments likely will be in fields such as microelectronics with much larger market potential than detectors.

**RECOMMENDATION 4-1**
The intelligence community should monitor commercial developments in thin-film cooler technology.

**KEY FINDING**
Scaling the data throughput of focal plane sensor systems involves not only the sensor chip but also the detector-processor interface, signal processing and compression, and the communication link (wireless for remote air- and

space-borne missions). Advanced compression and filtering with on-board processing provided by commodity multicore architectures are reducing communications demands.

**RECOMMENDATION 4-2**

Analyses of national capabilities should include consideration of advances in processing technologies for other uses—for example, commercial developments—that could also enhance the use of detectors in future sensor systems.

## The Global Landscape of Detector Technologies

To date, the United States has been the international leader in designing, developing, and implementing detector technologies. An exception is the migration of visible detectors, driven by consumer requirements, to an Asian manufacturing base. Significantly, existing U.S. export control policies have eroded and will continue to erode U.S. advantages in areas of military detector technologies.

Significant detector technology developments will continue to occur in Europe, specifically in the United Kingdom, France, and Germany, as well as in Israel. China, today, is a second-tier nation in designing and fielding detector technologies; however, it is investing substantial resources and is anticipated to emerge as a significant competitor within the 10- to 15-year time frame of this study.

**KEY FINDING**

Current export restrictions will continue to have a significant effect on development and maturation of detector technologies over the next decade. Numerous foreign countries are already developing their own technology base rather than utilizing U.S. technology and often will compete with U.S. technology. U.S. export restrictions are a primary driver creating this competition. U.S. companies invest significant resources in obtaining, funding, and exploiting foreign products so that they can compete in foreign markets without export restrictions.

# 1

# National Security Context of Detector Technologies

## BACKGROUND AND INTRODUCTION

This is the ninth report issued by an ad hoc committee under the general purview of the National Research Council's (NRC's) Standing Committee on Technology Insight—Gauge, Evaluate, and Review (TIGER). The statement of task for this study appears in Box 1-1 and is followed by a synopsis of the committee's approach for conducting the study and a general discussion of uses of detector technologies for future military applications.[1] Finally, a roadmap for the remaining report chapters explains the organization of this report.

Since 2005, the TIGER Standing Committee has assisted the intelligence community[2] (IC) in identifying appropriate areas of study to help that community better understand, assess, and forecast the national security implications of future scientific and technological advances. This particular study, initiated in 2009, focuses primarily on passive visible and infrared (IR) detectors. The Committee on Developments in Detector Technologies was formed to conduct the study. Biographical sketches of the committee members appear in Appendix A.

During the course of the study the committee had several opportunities to in-

[1]For this report, the committee defines a *detector* as representing a single pixel that receives photons. A *focal plane array* (FPA) is comprised of many detectors arranged in a two-dimensional grid and generates an image. A *sensor system* is comprised of FPAs plus other components, such as signal processing, data transmission, coolers, optics, and pointing and tracking mechanisms.

[2]According to Intelligence.gov, the IC is composed of 17 federal agencies. Accessed March 24, 2010.

BOX 1-1
Statement of Task

The NRC will

Consider the fundamental, physical limits to optical and infrared detector technologies with potential military utility, with priority on passive imaging systems. Elucidate trade-offs between sensitivity, spectral bandwidth and diversity, dynamic range, polarization sensitivity, operation temperature, etc. Compare these limits to the near-term state of the art, identifying the scaling laws and hurdles currently restricting progress.

Identify key technologies that may help bridge the gaps within a 10-15 year time frame, the implications for future military applications, and any significant indicators of programs to develop such applications. Speculate on technologies and applications of relevance that are high-impact wild cards or have a low probability of feasible deployment within 15 years. Discuss trends in availability and format scalability and in available cooling technologies.

Consider the pros and cons of implementing each existing or emerging technology, such as noise, dynamic range, processing or bandwidth bottlenecks, hardening, power consumption, weight, etc.

Identify which entities currently lead worldwide funding, research, and development for the key technologies. Highlight the scale, scope, and particular strengths of these R&D efforts, as well as predicted trends, time scales, and commercial drivers.

teract with members of the IC and gain a fuller understanding of the community's needs relating to detectors. In summary, the key points were the following:

- Focus on the underlying science, physics, and fundamental limits.
- Identify where there is room for improvement across the spectrum of possibilities.
- Cover a breadth of topics rather than delve into great depth for a particular topic.
- In accordance with the statement of task, emphasize passive sensing; however, comment on developments in active sensing where appropriate.

## COMMITTEE APPROACH TO STUDY

The committee met three times to receive briefings from government, industry, and university experts in the field of detectors. Mindful of its task, the committee cast a wide net for invited experts to help ensure broad coverage of the relevant aspects of detector science and technology. The committee met a fourth time to finalize its report. Appendix B lists specific meetings and participating organizations.

Recognizing the relatively short period of time available for preparing its report, the committee concentrated on responding expeditiously, yet comprehensively, to the many elements of the statement of task. The committee also took extra steps to supplement the text of its report with numerous references aimed at bolstering the abilities of members of the IC, and other readers as well, to inquire more deeply into particular subjects as necessary.

## GENERAL DISCUSSION OF DETECTOR TECHNOLOGIES FOR FUTURE MILITARY APPLICATIONS

### Overview

The first two paragraphs of the statement of task establish the military context of this study (i.e., the words "potential military utility" and "future military applications"). The statement of task covers a 10-15 year time frame and asks the committee to (1) identify technologies to help bridge gaps and (2) speculate on technologies and applications that are high-impact wild cards.

Important as they are, detectors are only a part of usable military sensor systems, as shown in Figure 1-1, which include optics; coolers; pointing and tracking systems; electronics, communication, processing, and information extraction subsystems; and displays (for detailed information on the fundamentals of detector technologies, see Chapter 2). Thus, it is essential to consider the *combination* of new detector technologies and the demand provided by military customers that drives the resulting sensor system developments.

Many future military sensor system requirements are classified, and the more critical missions, which usually require the most advanced sensor system technology, are highly classified. For the above reasons, the committee used unclassified data and open literature references to (1) describe general categories of military applications and (2) hypothesize possible detector-related military developments 10-15 years in the future. This process is necessarily imperfect, but it captures many of the implications of future sensor systems employing detector technology advances. The following sections are general military applications that, in the committee's estimation, will benefit from future advances in detector technologies.

### Wide-area, Continuous, Airborne Surveillance

There is a strong interest in sensor systems for continuous, wide-area surveillance. One aspect of these types of sensors is the potential ability to hit "rewind" to

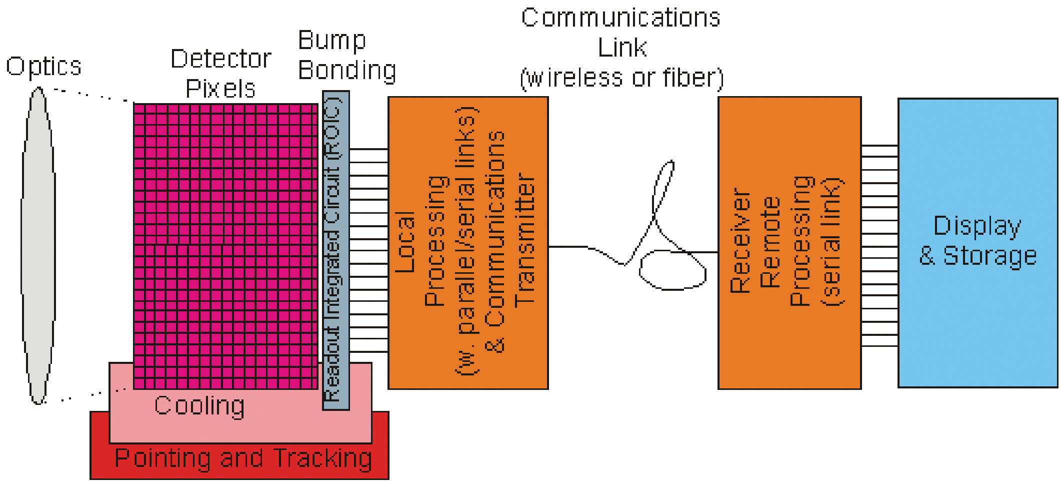


FIGURE 1-1
Schematic representation of an imaging system showing important subsystems.

determine exactly what happened in a particular situation, such as a terrorist attack. An example of how valuable a rewind capability can be comes from a bus bomb event in London, where there are cameras on most street corners. By monitoring many cameras and rewinding the recordings on them, the London police were able to trace the bus bomber back to the house that contained the bomb factory. Certainly in military situations this continuous, wide-area surveillance can have both real-time and postprocessing advantages.

### Inexpensive Airborne Sensors

There is strong interest in having sensors on small drones to obtain close-up, multiple views of objects of interest. These cheap and small sensors can also be used for dangerous tasks, such as explosive removal. Visible- and uncooled IR sensor systems are strong candidates for this type of application. It is not worth spending much time defining exact sensor parameters. Each application will have specifics, but they will vary. Requirements for this application will bend to available technology. If one does not have the resolution, move the sensor system closer.

### Airborne Military Targeting

A targeting sensor's main function is to detect and identify an object as far away as possible. The main identification limitation for a targeting sensor will be the diffraction limit, which relates resolution to wavelength and aperture diameter (i.e., wavelength divided by diameter). Historic aircraft sensors have had about an 8-inch-diameter clear aperture. The First Gulf War modified earlier low-altitude operations, which avoided ground defenses by being low and fast, to higher-altitude operations because the United States and its allies could destroy most ground-based defenses. As a result air operations were conducted mostly above 15,000 feet, an altitude at which long-range target identification became an issue. Size and weight constraints reduced the clear aperture to about 5 or 6 inches in diameter, thus reducing identification ability at constant wavelength.

### Missile Warning Sensors

One requirement for missile warning is field of view to see the approaching missile. Six well-placed sensors with $90 \times 90$ degree fields-of-regard can accomplish full ($4\pi$) situational awareness. For aircraft defense, the highest priority is to cover the rear. The second issue to consider is resolution. A missile has to be discriminated against a potentially cluttered background. The good news is that when missiles are fired they create a bright signature across many wavelengths.

### See-and-avoid Sensors

These sensors can be very similar to the missile warning sensors mentioned above. One military application for this type of sensor is to allow drones to fly under visible flight rules. For visual flight rules an aircraft is supposed to be able to identify another aircraft and maneuver to avoid it. To meet Federal Aviation Administration guidelines, any see-and-avoid sensor must discriminate oncoming aircraft at least as well as a pilot, but this is actually an easy standard since a pilot does not see oncoming aircraft well. See-and-avoid sensors are needed for efficient drone flights in visual-flight-rule conditions, but they will also be used in the future as a supplement for general aviation. It is likely that initial systems will operate in the visible wavelengths.

### Infrared Search and Track Systems

Another type of passive airborne sensor might be an infrared search and track system (IRST), a detection sensor for air-to-air targets. Field of regard is a major requirement. An IRST system should see other aircraft in the forward $2\pi$ of an aircraft (if one has 120 degrees in azimuth and 40 degrees in elevation, most potential threats will be detected). Resolution is important because of the need to detect aircraft at a distance, when they are points or near points; improved resolution helps discriminate against clutter.

### Inexpensive Terrestrial-based Sensors

There is strong interest in having sensors on small, unmanned ground robots and in stationary locations on structures to obtain close-up and multiple, views. These cheap, small sensors can also be used for dangerous tasks, such as explosive removal. Visible and uncooled IR sensor systems are strong candidates for this application. Each application will have specific parameters, and they will vary. Also, requirements for this application will bend to available technology.

### Ground-based Targeting Sensors

Tanks, for example, require targeting sensors. Atmospheric turbulence can be an issue along the ground. A range of a few kilometers is desirable.

### Satellite Platforms

This discussion is intended to characterize the important considerations for satellite sensor system designs, as opposed to an extensive discussion of specific applications and specific designs; specific sensor systems are mentioned only to illustrate the satellite platform considerations. Mission requirements usually start the sensor design process, and the usual "top-down" and "bottom-up" system engineering discipline develops the sensor system. Through the course of this process, the particular platform selected for the mission strongly influences the design choices. Detector materials and ancillary components are selected to optimize the design for a given satellite platform subject to the particular sensing requirements. In turn, expected performance improvements in materials or components influence the overall sensor performance relative to the design.

Several U.S. organizations launch satellite sensors for diverse purposes. By far the largest users of satellite sensors are U.S. government intelligence and military agencies whose missions, payloads, and orbits are usually classified.[3] Most of the early satellite sensors were deployed for strategic intelligence collection but have gradually become an indispensable tool for tactical military missions. On the nonmilitary side, collectors of geophysical data and various mapping and weather observation organizations—for example, the National Aeronautics and Space Administration (NASA) and National Oceanic and Atmospheric Administration—also employ optical, IR, and radio-frequency sensor systems.

Most satellite platforms are intended to gather data on the Earth; hence atmospheric parameters dominate the collection channel. Naturally occurring molecules

[3]A.D. Wheelon. 1997. Corona: The first reconnaissance satellite. *Physics Today* 50(2):24-30. This series of satellite imagery sensors began with the first successful Corona launch in August 1960.

dominate the absorption spectrum (e.g., $CO_2$ and $H_2O$) and limit the clear spectral windows available for satellite imaging. Surface weather producing cloud cover, dust storms, et cetera, further obstructs the collection channel.

Other satellites gather data using the "deep-space" collection channel. NASA observatories in orbit, including the Hubble, are examples. Another example is the Missile Defense Agency's mission of detecting mid-course ballistic-missile payloads with cold space as a background.

The nature of the mission is also important and constrains both the choice of orbit for the platform and the sensor design. Obviously, viewing the Earth's surface is best done in the clear windows listed above. On the other hand, picking out a "target of interest," such as a hot missile or aircraft exhaust, against the thermal Earth background would employ a sensor tailored to that particular emission spectrum.[4] NASA might employ a very narrow spectrum sensor matched to an excited ionic or atomic species in the upper atmosphere to monitor the atmospheric physics. These comments set the stage for the discussion below of orbital platform choices.

*Orbits and Applications*

The selection of the specific orbit is mission dependent. Orbits commonly used are LEO (low Earth orbit, up to approximately 500 km altitude); MEO (middle Earth orbit, up to approximately 8,000 km altitude); GEO (geosynchronous orbit, up to approximately 36,000 km altitude); and HEO (high Earth orbit, with looping, elliptical pattern having apogee at LEO and perigee at GEO over one of the poles). LEO satellites orbit Earth in approximately 1.5 hours; sun-synchronous orbits are roughly 580 km high; and GEO satellites are stationary overhead, with an orbital time of 24 hours, or one day.

Obvious direct consequences of the orbit choice are optical resolution fixed by altitude, wavelength, and aperture and time to view events fixed by the transit time relative to a fixed point on Earth. Another obvious impact is the sun's position relative to the sensor viewing geometry, which influences the choice of wavelength. One indirect consequence is the constraint of the communication channel used to relay collected data to Earth, which affects the on-board data processing and storage requirements; either ground-data nodes must be in view globally or a cross-link to a data relay satellite has to be provided. Yet another indirect consequence is the payload weight, size, and power dependence on the available booster size and overall mission cost.

More subtle consequences include the natural radiation environment that requires shielding of sensors and electronics; particularly sensitive is the MEO

[4]The Defense Support Program has provided early warning of intercontinental ballistic missile launches for several decades.

orbit that transits the Van Allen belts. Design life, fixed by mission and cost for the sensor platform, overarches the above considerations; continuous boost for a LEO satellite to keep it in orbit to achieve the design life has to be traded against the lower available payload weight for a GEO satellite whose life is limited by component failures.

The above criteria for choosing an orbit for a specific mission are the conventional guidelines that are adjusted by advances in launch capability and other satellite technology. Specialized government missions entail the admission of other criteria based on advances in foreign capability. A particularly direct influence relates to foreign threats to U.S. satellites. In 2007 a Chinese antisatellite (ASAT) capability was established,[5] which may impact future orbit choices for satellite imagery.

Planned future applications stretch the boundaries of sensor performance. For example, the James Webb Space Telescope (JWST),[6] in development, is an IR observatory with a planned launch date of 2013. The primary mirror size of 6.5 m stretches the boundaries of precision space construction for sensor systems given the launch weight limitations. A low-cost and short-development-time satellite sensor example is provided by Advanced Responsive Tactically Effective Military Imaging Spectrometer (ARTEMIS), a multispectral sensor configured from existing components launched recently on TacSat 3.[7] Finally, some totally new sensing concepts are being explored by the Defense Advanced Research Projects Agency (DARPA).[8]

### *Fractionated Space Systems*

Fractionated (or networked) space systems offer a novel concept for replacing the large, monolithic systems currently fielded. Such systems divide the functions of a large satellite between many small satellites, which are networked together as a large spacecraft. For example, a networked system could have separate spacecraft for each subsystem, such as power, payload, and navigation, or these subsystems could

[5]Kaufman Marc and Dafna Linzer. 2007. China criticized for anti-satellite missile test. *Washington Post*, p. A01, January 19. Available at http://www.washingtonpost.com/wp-dyn/content/article/2007/01/18/AR2007011801029.html. Last accessed March 24, 2010.

[6]For additional information on the JWST, see http://www.jwst.nasa.gov/index.html. Last accessed March 24, 2010.

[7]See Raytheon press release, June 2009, available at http://raytheon.mediaroom.com/index.php?s=43&item=1289&pagetemplate=release. Accessed March 24, 2010.

[8]Jason C. Eisenreich , Major, United States Air Force. 2009. *The All Seeing Eye: Space-Based Persistent Surveillance in 2030.* Alabama: Maxwell Air Force Base. Available at https://www.afresearch.org/skins/rims/q_mod_be0e99f3-fc56-4ccb-8dfe-670c0822a153/q_act_downloadpaper/q_obj_351d0f8b-02da-4fae-90fc-0daec34a01d9/display.aspx?rs=enginespage. Accessed March 24, 2010.

be hosted on more than one spacecraft. Currently, DARPA is executing Project F6 to investigate the feasibility and value of fractionated space systems. This example emphasizes the need to continually and creatively explore new satellite imaging sensor designs to maximize mission results.

*Implications for Sensor Systems*

Arguably, satellite platforms present the most exacting requirements for intelligence, surveillance, and reconnaissance (ISR) sensors. Satellite payloads are more expensive, typically around $100,000 per pound versus an order of magnitude less for avionic-based platforms, and the lifetime requirements are much more severe since they cannot be easily repaired. Many mission requirements are exacting, special, and innovative to the extent that precision in ranking needed improvements in sensor design is difficult. The general statements below are only intended to provide a context for this class of platform-based applications.

1. *Resolution and field of view (FOV).* Most if not all applications in all orbits profit from increased optical resolution, which means larger apertures for a fixed wavelength.
2. *Two colors.* Most target detection and tracking systems utilize two or more colors to discriminate targets against a static background and clutter. Filter wheels and related technologies have provided the mechanisms in the past. Dual- and multi-wavelength detectors enable a more robust, nonmechanical means to accomplish this.
3. *Data readout and processing.* For most advanced designs, the information handling is crucial to successful mission performance and tied to the data channel to Earth.
4. *Coolers.* Radiation cooling of detector arrays is a tried and true method for satellite-based sensors. This method requires solar thermal shielding using appropriate sunshades, which strongly depend on the orbit choice, and a means for ensuring the radiator always looks to deep space. Phase-change heat pipes are occasionally employed to physically configure the position of radiators on the satellite, and thermal blankets are routine. Auxiliary coolers ease the mechanical design constraints and are essential for achieving particularly low temperatures, approximately 25 K, for cold-object detection. Advanced means for cooling would enhance satellite-based sensors.
5. *Radiation shielding.* Space is a harsh environment, and radiation can severely impact the operation of low-noise electronics, a source of lifetime limitations for many components. In addition to normal environmental radiation, defense missions often have additional nuclear-burst requirements. Shielding is routinely provided by appropriate baffles and electronic

redundancy, and circuitry to sense and handle single-event upsets is necessary. Finally, designers for satellite sensors select electronic chips and other components that are specifically designed for radiation tolerance or have been screened for sustained functionality in the radiation environment.

6. *Designing to thwart threats.* All optical sensors are susceptible to saturation or destruction by in-band, high-power optical radiation.[9] Tactical system designs have included this type of requirement for several years, beginning with eye protection goggles. As laser technology has improved and larger-diameter mirrors have been launched in satellites to collect and focus optical radiation, the need to harden space sensors has increased. Inclusion of spectral rejection filters, baffles, or limiters creates a major new "optical path" design consideration that impacts resolution and detectivity. Elements for hardening usually require cooling, which imposes new physical constraints. Finally, this threat mitigation requirement may well impact the choice of orbit to conduct the mission since LEOs are the lowest altitude and, hence, the most susceptible to intentional disruption.

## POSSIBLE FUTURE DETECTOR-RELATED MILITARY DEVELOPMENTS

In considering the 10-15 year time frame, the statements below reflect content, in the form of presentations or documentation, provided to the committee during its data-gathering stage:

- Second-generation focal plane array (FPA) and forward-looking infrared (FLIR) technology is being globally distributed;[10]
- China is establishing world-class FPA fabrication facilities;[11]
- Single-photon detectors for short-wavelength infrared (SWIR) are under aggressive development;[12]

[9] Jeff Hecht. 2009. Half a century of laser weapons. *Optics and Photonics News* 20(2):14-21.

[10] Zvi Kopolovich. 2009. Status and Trends at Semiconductor Devices—Cooled and Uncooled Detectors. Presentation to the committee on January 21. John Miller. 2009. Future of Imaging. Presentation to the committee on January 21. FPAs and FLIR are discussed in Chapter 2.

[11] John Miller. 2009. Future of Imaging. Presentation to the committee on January 21. Paul Norton. 2009. Georgia Tech trip to China and Korea. Information provided to the committee in June 2009.

[12] Mark Itzler. 2010. Ultimate Sensitivity at Shortwave Infrared Wavelengths Using Single Photon Detection. Presentation to the committee on February 16. Hooman Mohseni. 2010. Novel Nano-injector Detectors: Towards High-resolution Single-photon Imagers at Short-wave Infrared (SWIR). Presentation to the committee on February 16. Bill Farr. 2010. Detectors for Photon-starved Optical Communications: Present and Future Directions. Presentation to the committee on February 17. SWIR has wavelengths of 0.7 to 2.5 microns (see Chapter 2).

- Persistent ISR doctrine will provide real-time information to warfighters;[13]
- Computational imaging work is accelerating;[14] and
- Trends are toward more extensive on-FPA signal processing.[15]

These major statements are embellished in the other chapters of this report. To this list the committee adds factual open knowledge:

- U.S. rules of engagement in battle strive for zero collateral damage and zero unintended casualties;[16]
- Urban warfare scenarios stress sensor system designs that meet ISR needs (namely, detecting improvised explosive devices [IEDs] and suicide bombers);
- Lasers are reaching maturity suitable for battlefield deployment;[17]
- China's new focus on satellite capability challenges U.S. supremacy;[18] and
- Tests of a U.S. airborne laser (ABL) have demonstrated success in shooting down missiles with a high-power laser.[19]

**Deductions, Extrapolations, and Speculations**

From the preceding discussions the future status of evolutionary sensor systems can be inferred (see Chapters 2, 3, and 4 for details of the detector technologies highlighted below):

[13]John Pellegrino. 2009. Emerging Sensor Technologies for Army Applications. Presentation to the committee on December 8. Lyn Brown. 2010. Air Force Research in Detector Technologies. Presentation to the committee on January 20.

[14]John Miller. 2009. Future of Imaging. Presentation to the committee on January 21. John Pellegrino. 2009. Emerging Sensor Technologies for Army Applications. Presentation to the committee on December 8. Nbir Dhar. 2010. Discussion with the committee on February 18.

[15]John Miller. 2009. Future of Imaging. Presentation to the committee on January 21. Vyshnavi Suntharalingam. 2010. Advanced Imager Technology Development at MIT Lincoln Laboratory. Presentation to the committee on January 21. Nbir Dhar. 2010. Discussion with the committee on February 18.

[16]Department of Defense. 2007. *U.S. Army Counterinsurgency Handbook*. New York: Skyhorse Publishing, Inc.

[17]10 kw Fiber Laser; available at www.ipqphotonics.com. Accessed March 24, 2010; numerous *Laser Focus World* issues.

[18]John Miller. 2009. Future of Imaging." Presentation to the committee on January 21. Lyn Brown. 2010. Air Force Research in Detector Technologies. Presentation to the committee on January 20.

[19]U.S. Missile Defense Agency press release; available at http://www.mda.mil/news/10news0002.html. Accessed March 24, 2010.

1. Third-generation IR two-color FPAs will achieve maturity;
2. Single-photon detectors (SPDs) for the SWIR will be ready for sensor system insertion;
3. Quantum-well and quantum-dot IR photo detectors (QWIP and QDIP) and type II strain-layer superlattice (SLS) technologies will achieve a sufficient performance level for specialized system needs;
4. Establishment of on-FPA analog-to-digital converters and increased onboard digital data processing will occur;
5. Improved filters, coolers, lightweight optics, et cetera, will mature; and
6. More capable strategic satellite sensors will be deployable having higher resolution, multispectral capability, and data fusion signal processing.

To the above sensor system status statements can be added potential revolutionary or wild-card achievements during the next 10-15 years:

1. Single-photon detectors using mid-wavelength infrared (MWIR, at 2.5 to 7.0 μm) and long-wavelength infrared (LWIR, at 7.0 to 12 μm) sensor systems via wavelength translation into the visible; a 10-times signal-to-noise ratio is possible for reduced cooling;
2. All digital processing capability directly on-FPA will emerge to provide improved data compression, feature extraction, and lowered overall data system complexity; and
3. Computational imaging application to "conformal imaging," "speckle imaging," and "hyperspectral" for new airborne platforms; these techniques feature system configuration advantages to compensate for atmospheric turbulence, airfoil boundary-layer effects, and optical train optimization.

These specific detector advancements could then be used to establish new or extended sensor system performance for whichever entity were to develop them. The committee notes that an evolutionary or wild-card detector technology for system insertion cannot just be an academic demonstration; the technology has to be practical and producible at a reasonable cost. The committee believes the above listing conforms to these criteria. (A counterexample would be single-photon superconductor sensors operating at 4 K, which are deemed impractical because of the cooling requirement; preliminary experiments suggest that materials having higher critical temperatures will not be suitable for this application.)

All of the above developments provide the design tools for the sensor system architect to respond to operational system requirements of the next decade:

- Active imaging to extract three-dimensional information for greater target or adversary location capability in the battlefield (using SPDs and laser radar);
- Passive or active imaging sensor systems merging all available data via new digital data processing, sensor fusion, and global positioning system coordinates to provide a total ISR solution for the battlefield (uses the technology listed above plus two- and multicolor FPA, active-pixel complementary metal oxide semiconductors, decision-making algorithms, and other evolutionary improvements);
- Addition of laser jamming to deny the enemy the use of second-generation FLIRs and maintain total night vision superiority—selected searchlight laser illumination may be employed to provide greater target discrimination at longer ranges (via latest fiber laser, diode-pumped solid-state lasers, and other anti-surface-to-air-missile technology); and
- All of the above items in a airborne platform with sensing and laser jamming power sufficient to deny LEO satellite ISR over U.S. territory. If a space conflict were to develop in the next decade the first phase may well involve sensor jamming or destruction in an ever-escalating series of steps (via conformal imaging, segmented mirrors, speckle imaging, etc., and high-power laser technology).

## Areas Not Considered or Considered Superficially

Again, in accord with the statement of task and due to the relatively short time line of the study, some important areas for sensor systems and sensor mission applications were not considered in detail sufficient to include in this report, including (1) short-range sensor systems and applications possible with cell phone cameras, medical imagers, perimeter-intrusion detection, netted arrays of short-range sensors, and pollution sensors; (2) coherent receivers to receive or eavesdrop on all laser communication links that relate to entangled photon-encryption channels; (3) image feature extraction for terrain, automatic target recognition, change detection, and facial recognition; (4) specific military missions, such as warning detection, missile seekers, missile fusing, and star trackers; and (5) satellite sensor system trades for classified missions.

Sensor system designs involve intricate trades impacting the choice of imaging sensor and a host of ancillary components affecting size, weight, and power. For example, satellite mission requirements critically determine booster throw-weight, choice of orbit, data transmission method, mission life, and cost. In turn, all sensor parameters, such as collector aperture, cooling method, needed power, and radiation shielding, have to be chosen to conform. This is an intricate process, and improvements in mundane engineering materials or processes can have a major impact on overall sensor system performance.

## REPORT ORGANIZATION

The report is structured as follows to correspond to the four main paragraphs of the statement of task. Chapter 2, "Fundamentals of Ultraviolet, Visible, and Infrared Detectors," covers the first paragraph. The report addresses the second and third paragraphs in the statement of task according to whether the important technologies are considered evolutionary or emerging. Thus, Chapter 3, "Key Current Technologies and Evolutionary Developments," addresses the elements of both the second and the third paragraphs for existing technologies expected to undergo significant evolution in the next 10-15 years. Chapter 4, "Emerging Technologies with Potentially Significant Impacts," does the same for potentially "game-changing" technologies that might emerge and have a major impact during the same time frame. With respect to the fourth paragraph in the statement of task, Chapter 5, "The Global Landscape of Detector Technologies," discusses the international scope of work in the detector, FPA, and sensory system fields.

# 2

# Fundamentals of Ultraviolet, Visible, and Infrared Detectors

## INTRODUCTION

Electro-optical detectors are used to measure or sense the radiation emitted or reflected by objects within the detector's optical field of view (FOV). Passive systems operate without any illumination of the object by the observer, relying on either self-luminosity (for example, a hot rocket exhaust) or reflection-transmission of ambient light. In active systems, the observation is associated with irradiation of the scene (as in a camera flash) in the spectral region of interest. A detector converts incident radiation to an electrical signal that is often proportional to the incoming intensity. This electrical signal is processed, usually digitally, transmitted, and/or stored. A two-dimensional array of detectors, called a focal plane array (FPA), is often placed at the focal plane of an optical system so that the spatial variation of the incident intensity is recorded as an image. There are many excellent texts at both introductory and advanced levels that deal with the fundamentals and applications of ultraviolet (UV), visible, and infrared detectors.[1,2] The committee's intent is to provide a brief introduction to facilitate reading the material that follows.

[1]E.L. Dereniak and G.D. Boreman. 1996. *Infrared Detectors and Systems.* Hoboken, N.J.: John Wiley and Sons.

[2]S. Donati. 2000. *Photodetectors: Devices, Circuits and Applications.* Saddle River, N.J.: Prentice-Hall.

## SOURCES

Sources include self-emission from hot objects that generally follow a blackbody radiation curve depending on the temperature of the source, modified by the spectral emissivity of the object. Alternatively for passive sensors, the reflection or transmission modification of ambient sources can be detected. During daytime, the dominant source in the visible is the sun. There is a significant night glow in the spectral region around 1.5 μm that makes short-wavelength infrared (SWIR) imaging an alternative to visible image intensifier night vision goggles for some night vision applications.[3,4] The semiconductor absorbance ranges that enable passive night vision and the nightglow irradiance spectrum are illustrated in Figure 2-1. The peak of the room temperature blackbody curve is at about 10 μm in the infrared.

## TRANSMISSION

### Spectral Regions

Over the years a number of designations for spectral regions have become somewhat standard, but there is significant overlap and it is useful to define the regions used in this report to assist the reader (see Table 2-1). The transitions between these regions are not sharply defined and the designations are to be interpreted loosely; the detection mechanisms, the transmission, and the dominant noise sources all vary across these bands. These definitions help in discussing those variations cohesively.

Electromagnetic sensors cover the entire range from 200 nm to 20 μm and beyond; this taxonomy is intended merely to provide a nomenclature for the most frequently used bands for long-range imaging.

### Atmospheric Transmission

Atmospheric transmission is an important aspect of any terrestrial remote sensing application. Figure 2-2 shows the atmospheric transmission across the 0.2-20 μm region (~1 km horizontal path length at sea level, temperature = 15°C, with 46 percent relative humidity) along with the wavelength bands defined above.

While Figure 2-2 is representative, the transmission curve will vary with atmospheric conditions, as well as the path taken through the atmosphere; for example,

[3] T.R. Hoelter and B.B. Barton. 2003. Extended short wavelength spectral response from InGaAs focal plane arrays. *Proceedings of SPIE* 5074:481-490.

[4] Available at http://www.sensorsinc.com/downloads/paper_HighSpeedSWIRImagingAndRangeGating.pdf. Last accessed March 25, 2010.

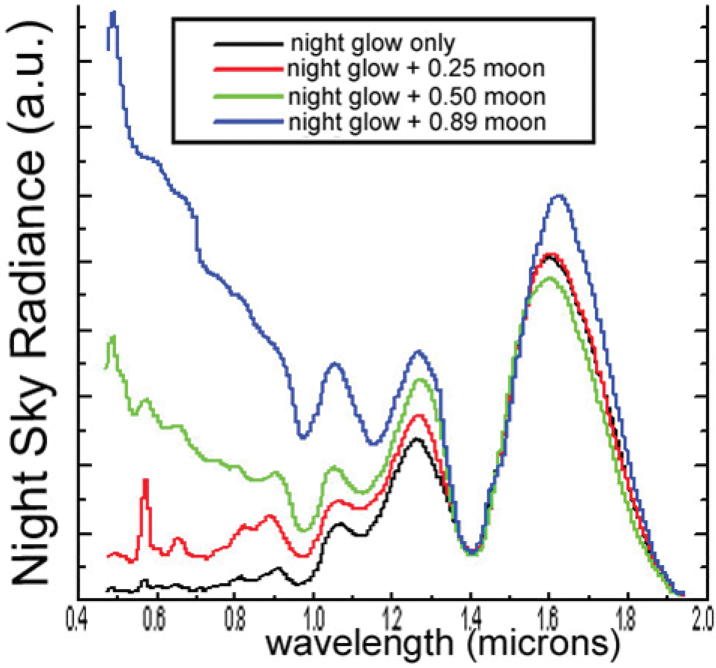


FIGURE 2-1
Nightglow irradiance spectrum under different moonlight conditions. SOURCE: Vatsia, L.M. 1972. Atmospheric optical environment. Research and Development Technical Report ECOM-7023. Prepared for the Army Night Vision Lab, Fort Belvoir, Va.

TABLE 2-1 Definition of Spectral Regions with Long-range Atmospheric Transmission

| Designation | Wavelength Band (μm) | Physical Significance and Comments |
|---|---|---|
| Solar blind UV | 0.2-0.28 | Solar radiation in this band is blocked by the Earth's ozone layer, so any radiation in this region is likely man-made. Once under the ozone layer, the atmosphere is transparent to wavelengths as short as ~200 nm where oxygen absorption limits the transmission |
| UV | 0.28-0.4 | Atmosphere is transparent |
| Visible | 0.4-0.7 | Peak of solar spectrum |
| Near infrared | 0.7-1.0 | Long-wavelength cutoff defined by silicon detector response |
| SWIR | 1.1-2.7 | Overlaps with telecommunications wavelengths; large commercial infrastructure available at 1.3 and 1.55 μm |
| MWIR | 2.7-6.2 | Atmospheric transmission window, molecular vibrational absorptions |
| LWIR | 6.2-15.0 | Atmospheric transmission window, molecular vibrational absorptions |
| VLWIR | 15.0-20.0 | Molecular vibrational absorptions |

NOTE: LWIR = long-wavelength infrared; MWIR = mid-wavelength infrared; VLWIR = very long wavelength infrared.

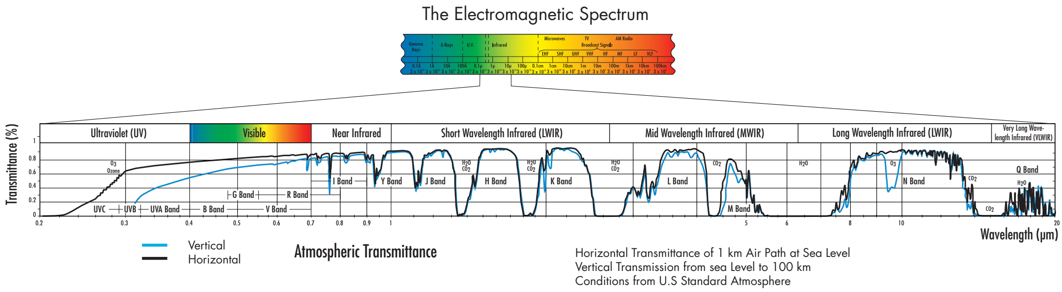


FIGURE 2-2
Display of the atmospheric transmittance levels. SOURCE: Data from the Santa Barbara Research Institute, a subsidiary of Hughes, and OMEGA Engineering, Inc. Available at http://www.coseti.org/atmosphe.htm. Accessed March 29, 2010.

looking through the atmosphere from a space-based platform will differ in the details.

**FINDING 2-1**
For any sensor application, the relevant spectral range is set by the overlaps of the spectral signature of the target and the pass bands of the transmission medium between the target and the detector.

## DETECTION

In general, detectors are divided into two classes: thermal and photon (or quantum).[5] Thermal detectors operate by the absorption of incoming radiation causing a change in temperature of the detector and by the sensitivity of some measurable parameter—for example, resistance—to that temperature. Thermal detectors are typically sensitive across a wide range of incident wavelengths. Quantum detectors depend on the direct interaction of the incoming light with the detector materials, resulting, for example, in electron-hole pair creation in a semiconductor. Photogenerated carriers can be measured by directly measuring charge collected during an integration period, by measuring photocurrent, by a change in resistance (photoconductive), or by voltage generation across a junction (photovoltaic).

[5] R.L. Petritz. 1959. Fundamentals of infrared detectors. *Proceedings of IRE* 47(9):1458-1467.

## Thermal Detection

In thermal detectors, photon absorption leads to a small temperature rise of the detector, which is sensed by a temperature-dependent property of the material such as a pyroelectric effect or a temperature-dependent resistance. An advantage of using thermal detectors is that they typically are very broadband; a disadvantage is that it is a challenge to make a structure that has measurable temperature rise for low power signals.[6] It is also possible to track thermoelectric effects using thermocouples and thermopiles or with the aid of Golay cells that can track thermal expansion in a system. In general, there is a trade-off between the response speed of a thermal detector and its sensitivity. Thermal isolation allows longer integration times to detect weaker signals, but this means that the detector response time is necessarily increased.

## Quantum Detection

Quantum or photon detectors, typically semiconductors with bandgaps matched to the photon energy, operate by the generation of electron-hole pairs by the absorption of a photon. There are two major classes of photon detectors: photoconductive and photovoltaic.

### *Photoconductors*

In a photoconductor, the excited carriers are detected through the change in resistance induced by the photoexcited carriers. Often the mobilities of electrons and holes are quite different in the semiconductor, with the consequence that the faster carrier can transit the detector several times before the carriers recombine. This provides a gain mechanism.

### *Photovoltaic Detectors*

In a photovoltaic device, the photoexcited electron and hole are separated by the built-in field associated with a p-n junction and collected. Particularly for indirect bandgap semiconductors, such as silicon, the absorption region has to be extended to ensure a good quantum efficiency leading to p-i-n designs. There is usually a trade-off between extending the absorption region for high probability

[6] B. Cole, R. Homing, B. Johnson, K. Nguyen, P.W. Kruse, and M. C. Foote. 1994. High performance infrared detector arrays using thin film microstructures. *Proceedings of the Ninth IEEE International Symposium on Applications of Ferroelectrics* 653-656.

of absorbing a photon and shortening it to ensure that recombination mechanisms do not impact the collection efficiency.

*Avalanche Photodiodes*

Avalanche photodiodes incorporate high-field regions that lead to carrier multiplication to increase signal levels above the characteristic noise sources downstream in the electronics. The carrier multiplication is accomplished by imparting sufficient kinetic energy to a carrier for it to create an additional electron-hole pair by impact ionization. There is always some excess noise associated with the multiplication, but this can be minimized by designs that allow primarily one carrier to be multiplied while suppressing the multiplication of the oppositely charged carrier.

## INFORMATION ENCODED BY PHOTONS

A photon is the quantum mechanical element of all electromagnetic radiation. Photon energy is given by

$$E_{ph} = h\nu = \frac{hc}{\lambda} = \frac{1.986 \times 10^{-19}}{\lambda} \mathrm{J},$$

where $h$ is Planck's constant, $c$ is the speed of light, and $\lambda$ is the wavelength of the infrared photon in micrometers. By collecting photons, measurements can be made of light's intensity, temporal variations in intensity, spectrum, polarization, electric field phase, incident angle, and photon time of arrival. These types of measurements will now be defined in greater detail.

### Intensity

Intensity, the incident power per unit area, is the most commonly used optical imaging signal. The variation of signal intensity across the focal plane is recorded as a gray-scale image.

### Spectrum

Images can be panchromatic, monochromatic, multispectral (including three-color traditional RGB [red, green, blue]), or hyperspectral (multiple spectral bands across the wavelength range of interest). Spectral information can be obtained in several ways, including dispersion into different pixels (using diffraction or refraction), temporal modulation of spectral filters, use of on-chip absorptive filters, measuring the size of a charge packet (for example, for X-ray energy spectroscopy), or varying the bandgaps of multiple photon absorbing regions.

### Polarization

Imagers have been developed to measure the complete polarization states of the electromagnetic field described by the Stokes parameters or, more commonly, the linear polarization components.[7]

### Dynamics

Time scales can range from still imaging, to video rates, to fast (e.g., kilohertz) amplitude fluctuations due to target phenomenology, to high-speed imaging (e.g., megahertz), to acquiring sub-ns (nanosecond) range information from single photons for active LADAR (laser detection and ranging) imaging.

### Time Delay

The time delay from an excitation to the reception of a photon provides a measure of distance to the object in the same way as in a radar receiver. This is an active sensor application and is beyond the scope of this study. However, it is worthwhile to note that advances in both ultrafast sources and high-speed photon counting detectors will make available in the visible and near-infrared (NIR) spectral regions many of the advanced radar concepts, such as chirped pulses and synthetic imaging concepts, that have been so successful in longer-wavelength spectral regions.

Imagers are available with many different designs and architectures to exploit these different characteristics of optical signals, but it is difficult to design a single imager that is optimized for simultaneously measuring all of these attributes.

### Phase and Incidence Angle

Imagers can perform heterodyne or other types of carrier-phase detection. High-speed detectors can allow detection of temporal-phase variation by measuring the beat frequency between a local oscillator and a return signal. Alternatively, wavefront sensors are used to measure spatial-phase variation, allowing analysis of atmospheric wavefront distortions for adaptive optical correction. These sensors enable measurement of a small phase distortion in optical waves under significant

[7]For additional detailed information on Stokes polarization parameters, please see http://spie.org/x32376.xml. Last accessed on May 6, 2010.

background wavefront aberrations. A Shack-Hartmann sensor is a frequently used wavefront sensor that consists of a microlens array in front of a multiple element detector or focal plane array to register the local wavefront tilt in the position of the imaged spots from each microlens on the sensor array.[8]

**FINDING 2-2**
While the spatial variation of signal intensity is most often the quantity evaluated to produce an image, spectral distributions, polarization, phase, and temporal characteristics are additional information channels that can be exploited in some applications.

## THE LIMITS IMPOSED BY DIFFRACTION

### Spatial Resolution

Optical imaging systems can have limitations in resolution caused by imperfections in the lenses or by their misalignment, which results in defects of the image and is often referred to as an optical aberration. In addition, for transmission through the atmosphere, variations of the index of refraction due to air currents and temperature variations also cause changes in the image. Aberrations describe the amount by which a geometrically traced ray misses a specific location in the image.[9] If all of these aberrations are dealt with, diffraction is the ultimate limit of optical focusing. For an aberration-free optical system with uniform illumination of a circular input aperture, the result at the focus is a bright central disk surrounded by a series of concentric rings of rapidly diminishing amplitude. This is known as an Airy pattern (shown in Figure 2-3), and the diameter of the central disk is given by ~$1.22\lambda/NA$, where $\lambda$ is the wavelength and $NA$ the numerical aperture of the optical system (the half-angle of the light acceptance cone).[10] Mathematically, the intensity versus position in the Airy pattern is given by

$$I(r)=\left(\frac{NA}{\lambda}\right)^2\left(\frac{2J(m)}{m}\right)^2 \text{ with } m=\frac{2\pi NAr}{\lambda},$$

[8]J. Schwiegerling and D.R. Neal. 1994. Historical development of the Shack-Hartmann wavefront sensor. *J Opt Soc Am A* 11:1949-1957.

[9]Harold Rothbart and Thomas H. Brown. 2006. *Mechanical Design Handbook: Measurement, Analysis, and Control of Dynamic Systems*, Second Edition. New York: McGraw-Hill Companies, Inc.

[10]$NA$ is related to the $F/\#$ notation commonly used to describe the light acceptance cone in photography. $NA = \sin\theta$, where light incident on the lens at angles up to $\theta$ is imaged onto the focal spot. In terms of the diameter of the lens, $D$, and the focal length, $f$, $\tan\theta = D/2f$. $F/\# = f/D$, so for small angles where $\sin\theta \approx \tan\theta \approx \theta$, $NA \approx 1/[2F/\#]$.

FIGURE 2-3
The Airy disk. SOURCE: Figure courtesy of Cambridge in Colour. Available at http://www.cambridgeincolour.com/tutorials/diffraction-photography.htm. Accessed on March 29, 2010.

where $r$ is the radial coordinate and $J_1$ is the first-order Bessel function, with a first zero at $m = 0.61$.

As is very well known, the minimum focal spot diameter also sets the separation distance at which two point objects can be resolved as distinct. The Rayleigh resolution criterion is obtained by setting the minimum separation of two objects equal to the radius of the Airy disk,[11]

$$R_{min} = 0.6\lambda / NA.$$

Detector optical systems capable of producing images with angular resolutions that are as good as the instrument's theoretical limit are said to be diffraction limited.

For an ideal circular aperture, the two-dimensional diffraction pattern, the Airy disk, is used to define the theoretical maximum resolution for the optical system. When the diameter of the disk's central peak becomes large with respect to the size of the pixel in the FPA, it begins to have a visual impact on the image.

## OPTICAL SYSTEMS

### Numerical Aperture and Field of View

The numerical aperture, $NA = \sin\theta \leq 1$, describes the light collection power of the optical system. A larger $NA$ results in higher resolution (see equation above) and, therefore, requires more pixels in the focal plane array if the same area is to

[11] J.W. Strutt (III Lord Rayleigh). 1879. Investigations in optics, with special reference to the spectroscope. *Monthly Notices of the Royal Astronomical Society* 40:254.

be imaged at this higher resolution. The field of view is the extent of the imaged region on the focal plane array referred back to the objects being imaged.

### Curved Focal Planes

Everyone is familiar with one optical system that uses a curved focal plane array, namely the human eye. Nature chooses this curvature because it makes the optics much simpler. In contrast, the many optical elements in, for example, a standard camera lens are required to faithfully reproduce the image on the flat focal plane of the camera. Our materials technology, which relies on epitaxial crystal growth and the accompanying device fabrication technologies and has largely derived from planar silicon integrated circuit technology, make curved focal plane arrays a difficult option. Recently there has been significant work, particularly in visible systems based on silicon materials to adapt to curved focal surfaces.[12] A flat surface can conformally map to a cylinder, but it cannot map to a sphere without deforming. Practical curved focal plane technologies are making a significant difference in image capture and in the size and weight of optical systems.

## DETECTIVITY

The noise-equivalent power (*NEP*) is the input power at which a detector exhibits a signal-to-noise ratio of unity. The detectivity, *D*, is the inverse of *NEP*; this of course depends on the detector area (*A*) and the detection bandwidth (*BW*). For observation of an extended object, the signal scales as the area, while the noise associated with the dark current scales as $\sqrt{A}$ ; the noise also scales as $\sqrt{BW}$ . These simple extrinsic parameters can be eliminated with a simple normalization; the resulting parameter is $D^* = \sqrt{A \times BW} \,/\, NEP$ which is more characteristic of detector material performance.

### Quantum Efficiency

The signal level at the detector is directly proportional to the probability that an incident photon results in an electrical signal; this is known as the quantum efficiency (*QE*). The external quantum efficiency includes effects such as reflection from optical surfaces that can be addressed with additional engineering (for example, antireflection coatings), while the internal quantum efficiency is more characteristic of the detector material and device geometry.

The prerequisite for quantum efficiency is absorption of a photon leading

[12]R. Dinyari, S-B. Rim, K. Huang, P.B. Catrysse, and P. Peumans. 2008. Curving monolithic silicon for non-planar focal plane array applications. *Applied Physics Letters* 92:091114.

to some, typically electronic, change in the material such as the creation of an electron-hole pair in a semiconductor. A high absorption coefficient allows thinner material, which facilitates the second component of the quantum efficiency—sensing the electron-hole pair. In a photovoltaic detector this is accomplished by separating the carriers across a p-n junction resulting in a voltage proportional to the number of carriers. This process can be disrupted by recombination, either radiative or nonradiative, before the carriers diffuse into the junction region. In a photoconductive detector, the carriers are sensed as change in the conductance, which is measured as the current for a fixed voltage applied across the device. If the carriers cycle more than once through the detector before recombination, there is a gain associated with the detection that can make it easier to overwhelm noise further downstream in the electronics.

## Noise

There are many noise sources whose relative importance varies with the detector material properties, the ambient temperature, the detector operating temperature, the device design, the readout electronics, and other variables. Some of the most important sources are catalogued here. Since these noise sources are in general uncorrelated, the total noise is proportional to the square root of the sum of the squares of the individual noise sources.

### *Photon Statistics and Background-limited Infrared Detection*

There is noise associated with the signal itself. Since photodetection is a discrete process, and most natural sources exhibit Poisson statistics in the fluctuations of the signal level, this noise scales as the square root of the signal level. Photon noise is unavoidable for natural signals and sets a fundamental noise floor. For an extended source (image structure large compared to an individual pixel) the current scales as the pixel area, so the noise is $n_{photon} = \sqrt{i_{photon}(BW \times A)}$.

For engineered sources, it is possible to reduce the shot noise at the expense of increased phase fluctuations, and vice versa. Collectively these are known as squeezed states and have been investigated for communications applications.

Any background photons impinging on the detector also contribute to the noise. While the background is usually not an issue in the UV and visible, in the infrared there is substantial background flux associated with blackbody emission from a room-temperature scene. The peak of the 300 K blackbody emission is in the middle of the LWIR at 10 μm. For cooled infrared detectors (discussed below) this dark current associated with the background radiation and the accompanying noise levels often set the detection limit. This is known as background-limited infrared photodetection (BLIP). Many current infrared systems are close to BLIP;

thus, further improvements in detector dark currents will have little impact on performance. Of course there are many scenarios other than looking at a terrestrial scene, and these have other, often more sensitive, BLIP limits. For example, looking up, a cold sky has a lower BLIP limit, requiring lower detector noise, and space-based cross-link applications have very low backgrounds. There is increasing interest in multispectral and hyperspectral sensing. The spectral filtration inherent in these concepts also reduces the background contribution to the noise.

*Dark Current*

Both photovoltaic and photoconductive detectors are biased under operating conditions and exhibit some dark current even in the absence of illumination. This dark current is usually proportional to the pixel area. Since the dark current is carried by discrete charges (electrons and holes), there is shot noise scaling as $n_{dark} = \sqrt{i_{dark}(BW \times A)}$ associated with this dark current. This dark current noise is the dominant limitation in many detectors. The sources of dark current include Johnson noise and generation-recombination noise.

***Johnson Noise*** For a photoconductive detector, one component of the dark current is associated with the dark resistance of the detector. This is known as Johnson or Nyquist noise (kTC noise); the dark current is given by

$$n_{Johnson} = \sqrt{\frac{4kT(BW \times A)}{R_0 A}},$$

where $k$ is Boltzman's constant, $T$ is the device temperature, $R_0$ is the detector resistance (slope of the $I$-$V$ curve for a photovoltaic device), $A$ the detector area, and $BW$ the electrical bandwidth. The noise current is written in this form since the factor $\sqrt{BW \times A}$ is eliminated in evaluating its contribution to $D^*$.

***Generation and Recombination Noise*** For infrared detectors, the fluctuations in thermally generated carrier densities in the active region also contribute to the noise; this noise source is significant if thermal energies ($kT$) are comparable to the semiconductor bandgap. This noise source is generally negligible for UV and visible detection. Cooling the detector also eliminates this noise source, but it can be the dominant noise source for uncooled devices. A relatively new device concept, which relies on bandgap engineering concepts available in III-V epitaxial growth, is an nBn (and variants including pMp) geometry that incorporates a barrier layer for the majority carrier in place of the traditional p-n junction carrier separation region, but does not impede the conduction of minority carriers. The result is to block generation-recombination (G-R) noise currents.[13,14] In principle, this can dramatically reduce or eliminate G-R noise and allow improved detector performance including higher-temperature operation.

*Readout Noise*

New very low power monolithic high-speed analog-to-digital converters have advanced the state-of-the-art noise performance in IR sensors. This can allow BLIP-limited performance to be achieved over a broad range of operational conditions (which means that performance purely from a sensitivity standpoint has plateaued for these conditions). True 14-bit performance can be achieved at pixel rates of more than 20 megapixels per second per channel and noise floors of 0.325 count (approximately 50 $\mu V_{rms}$), all while exposed to full-EMI (electromagnetic interference) environments. High-speed ADCs (analog-to-digital converters) allow for oversampling techniques that were previously not possible in 14-bit resolution. ADCs are now available in Quad and Octal packages with high-speed serialized outputs, eliminating hundreds of wires and field-programmable gate array (FPGA) pins when connected to large-format FPAs, further increasing integration and reducing power. Single-board designs with 64 video channels of high-speed, 65 million samples per second, ADCs allow 2K $\times$ 2K arrays to run at 30 Hz video rates achieving an overall processing bandwidth of more than 125 megapixels per second. These digitized video data can then be transmitted at 5 gigabits per second using a common LVDS (low-voltage differential signaling) interface over 15 m of cable.

Multiple standard video interface protocols are currently supported such as FPDP (front-panel data port), Camerlink, Ethernet, Hotlinks, and LVDS. Additionally, newer standards are coming into favor for IR sensors such as HDMI (high-definition multimedia interface), Gigabit Ethernet, FireWire, and USB 3.0. These interfaces allow for easier system integration and can support large-format arrays.

*Other Sources of Noise*

There are many other sources of noise of varying degrees of importance in specific situations. Some of these include low-frequency (1/*f*) noise, temperature

[13]S. Maimon and G.W. Wicks. 2006. nBn detector, an infrared detector with reduced dark current and higher operating temperature. *Applied Physics Letters* 89:151109.

[14]Binh-Minh Nguyen, Siamak Abdollahi Pour, Simeon Bogdanov, and Manijeh Razeghi. 2010. Minority electron unipolar photodetectors based on type II InAs/GaSb/AlSb superlattices for very long wavelength infrared detection. *Proc SPIE* 7608:760825-1.

fluctuations that change the parameters such as dark current of the device, microphonics, and impurity ionization associated noise (Barkhuesen noise).

## BRIEF SURVEY OF DETECTORS BY SPECTRAL REGION

### Ultraviolet

For applications where both UV and visible responses are desired, silicon photodiodes are very good UV detectors. For wavelengths below 360 nm, silicon exhibits a direct bandgap resulting in a very strong absorption. Traditional vertical p-n junction devices can be inefficient in this wavelength region if the absorption occurs in the heavily doped contact layers, before the photons can penetrate to the junction region. In-plane devices such as Schottky barrier detectors, in which the transport is parallel to the semiconductor surface, provide an alternative geometry that also has the advantage of very high speed as a result of the low capacitance of the structure.

#### *Solar Blind*

Because of the low natural background in the solar blind region ($\lambda < 280$ nm), photodetectors and focal plane array imagers operating in this range allow for a number of unique applications—generally any terrestrial 280 nm radiation can be assumed to arise from man-made sources. Currently, most solar blind imaging is performed with either a photocathode and microchannel plate combination or a UV-enhanced silicon photodiode with a band-pass filter. Neither of these options is ideal: the photocathode and microchannel plate combination is a fragile vacuum tube device requiring a high-voltage power supply, while the silicon photodiode is not intrinsically solar blind and suffers from increased size and complexity and decreased efficiency due to the optical filtering requirement. Technological and scientific advances in high-aluminum-composition AlGaN-based semiconductor materials have led to the development of visible blind p-i-n photodiode FPA cameras[15,16,17,18] and a renewed interest in the development of intrinsically solar

[15]J.D. Brown, Zhonghai Yu, J. Matthews, S. Harney, J. Boney, J. F. Schetzina, J. D. Benson , K. W. Dang, C. Terrill , Thomas Nohava, Wei Yang, and Subash Krishnankutty. 1999. Visible-blind UV digital camera based on a 32 × 32 array of GaN/AlGaN p-i-n photodiodes. *MRS Internet Journal of Nitride Semiconductor Research* 4(9):1-6.

[16]J.D. Brown, J. Matthews, S. Harney, and J. Boney. 1999. High-sensitivity visible-blind AlGaN photodiodes and photodiode arrays. *MRS Internet Journal of Nitride Semiconductor Research* 5S1(W1.9).

[17]B. Yang, K. Heng, T. Li, C. J. Collins, S. Wang, R. D. Dupuis, J. C. Campbell, M. J. Schurman, and I. T. Ferguson.2000. 32×32 Ultraviolet $Al_{0.1}Ga_{0.9}N$/GaN p-i-n photodetector array. *Quantum Electronics Letter*, 36(11):1229.

[18]J.D. Brown, J. Boney, J. Matthews, P. Srinivasan, and J.F. Schetzina.2000. UV-Specific (320-365

blind FPA cameras. The first solar blind FPA camera was reported by BAE systems in 2001.[19] The first images from a solar blind FPA camera were published in 2002, but the quality was lacking and the FPA did not provide full frame imaging.[20] The first 320 × 256 imaging was reported in 2005.[21] The only recent reports of solar blind FPAs are from a Chinese group.[22,23]

## Visible

Visible detectors are broadly divided into charge-coupled device (CCD) imagers and complementary metal oxide semiconductor (CMOS) imagers. Prior to discussing each of these in greater detail, the performance and operation of CCD and CMOS technologies will be contrasted. In addition, avalanche photodetectors are important for low-light-level applications.

In general, CCD and CMOS technologies can share much of the same processing equipment, and both have benefited from Moore's law scaling advances (see Figure 2-4); however, the detailed process flows for CCDs and CMOS evolved separately and with different requirements. During the evolution of CMOS technology it became possible to implement reasonable-quality imagers that had some advantages and disadvantages over the already mainstream CCD imaging technology.

Initial claims were that the increasing performance of CMOS imagers, driven by rapid Moore's law progress, as well as the level of integration available (which enabled single-chip solutions) and the low cost of commodity CMOS processes, would rapidly make CCD imaging technologies obsolete. CCDs were perceived as requiring "specialized" processes, implemented in low volumes on less cost-

nm) digital camera based on a 128×128 focal plane array of GaN/AlGaN p-i-n photodiodes. *MRS Internet Journal of Nitride Semiconductor Research* 5(6).

[19] P. Lamarre, A. Hairston, S.P. Tobin, K.K. Wong, A.K. Sood, M.B. Reine, M. Pophristic, R. Birkham, I.T. Ferguson, R. Singh, C.R. Eddy, Jr., U. Chowdhury, M.M. Wong, R.D. Dupuis, P. Kozodoy, and E.J. Tarsa. 2001. AlGaN UV Focal Plane Arrays. *Physica Status Solidi (A)*, 188(1):289.

[20] J.P. Long, S. Varadaraajan, J. Matthews, and J.F. Schetzina. 2002. UV detectors and focal plane array imagers based on AlGaN p-i-n photodiodes. *Opto-Electronics Review* 10(4):251.

[21] R. McClintock, K. Mayes, A. Yasan, D. Shiell, P. Kung, and M. Razeghi. 2005. 320x256 solar-blind focal plane arrays based on $Al_xGa_{1-x}N$ *Applied Physics Letter*, 86(1):011117.

[22] Yuan YongGang, Zhang Yan, Chu KaiHui, Li XiangYang, Zhao DeGang, and Yang Hui. 2008. Development of solar-blind AlGaN 128x128 ultraviolet focal plane arrays. *Science in China Series E: Technological Sciences* 51(6):820.

[23] Yongang Yuan, Yan Zhang, Dafu Liu, Kaihui Chu, Ling Wang, and Xiangyang Li. 2009. Performance of 128×128 solar-blind AlGaN ultraviolet focal plane arrays. *Proceedings of the SPIE,* 7381:73810I.

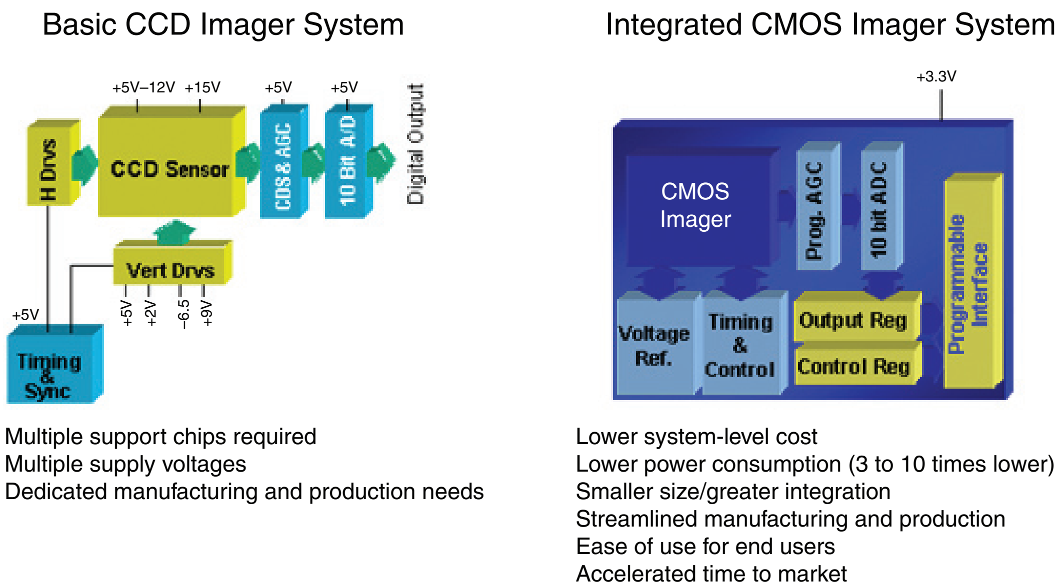


FIGURE 2-4
A depiction of the CCD and CMOS imager systems. The photoresponsive elements are the CCD sensor and the CMOS imager, respectively. Available at http://www.siliconimaging.com/ARTICLES/CMOS%20PRIMER.htm#imagesensors. Accessed March 29, 2010.

effective dedicated process lines. The reality has proven different from these early and simplistic predictions.[24]

CCDs have continued to dominate the market for most high-performance imaging applications. Secondly, CMOS processes used for imagers have increasingly become specialized, enabling the higher performance needed to compete with CCD imagers. CCD processes have also become more complex and CCD-CMOS processes have emerged that allow CCD imagers to be monolithically integrated with CMOS electronics used for control, analog-to-digital conversion, and other on-chip processing.[25,26] To understand some of the performance issues, a short description of some key features of CCD and CMOS technologies is provided here.

[24] S. Paurazas, J. Geist, F. Pink, M. Hoen, and H. Steiman. 2000. Comparison of diagnostic accuracy of digital imaging by using CCD and CMOS-APS sensors with E-speed film in the detection of periapical bony lesions. *Oral Surgery, Oral Medicine, Oral Pathology, Oral Radiology and Endodontology* 89(3):356-362.

[25] Craig L. Keast and Charles G. Sodini. 1993. A CCD/CMOS-based imager with integrated focal plane signal processing. *IEEE Journal of Solid-State Circuits* 28(4):431-438.

[26] V. Suntharalingam, B.E. Burke, J.A. Burns, M.J. Cooper, and C. L. Keast. 2000. Merged CCD/SOI-CMOS technology. *Proc SPIE* 3965:246-253.

*Charge-coupled Device Imagers*

CCD imagers typically collect and store charge (photoelectrons or holes) generated by incident light under collection electrodes and then sequentially shift the stored charge packets to a readout amplifier to produce a time-dependent output signal containing the image information. To isolate charge packets in adjacent pixels from one another, multiple electrode phases are used.[27]

Typically either three or four electrode phases are employed, which has generally resulted in a need for two, three, or even four layers of polysilicon gate material to define the clock phases needed for the electrodes; this is in contrast to most standard CMOS processes, which employ one only layer of polysilicon to provide the gates for the nMOS and pMOS transistors.

In a CCD, charge packets are swept from pixel to pixel along a CCD register by varying the potential applied to each clock phase. Since several thousand transfers may be required to reach the edge of the chip, the charge transfer efficiency must be kept quite large, with losses of only $10^{-5}$ to $10^{-6}$ per transfer, both to ensure that the amplitude of the charge packet is preserved during transfer and to prevent charge smearing during readout. Although a number of factors can cause charge loss or trapping during transfers, one important factor is to ensure that charge transfer is not blocked by potential barriers between phase electrodes.

To help minimize the formation of potential barriers between phases, it is advantageous to have very small gaps (less than 100-200 nm) between electrodes defining adjacent clock phases. A decade ago, defining such small gaps using lithography was not practical, especially given the requirement for high yields (no shorts allowed between phases) and the length of the region that must remain defect free (on the order of 10 cm or more for a large-format CCD.) The alternative is to define the electrodes for one clock phase in a layer of polysilicon, thermally oxidize that polysilicon, and then cover it with another layer of polysilicon in which the next clock phase is defined.[28] Using this method, small gaps can be defined between clock phases with high yield and without using aggressively scaled lithography, but at the expense of a process that often uses three layers of polysilicon.

It is worth noting that the ability to avoid the need to pattern extremely small features in the CCD process has allowed the use of older-generation lithography systems such as 1:1 reduction scanning slit lithography systems (for example, the Perkin-Elmer/SVGL Micralign systems). Since these systems can print a field as large as a full 150 mm diameter wafer, very large format CCD imagers can be de-

[27]See, for example, James R. Janesick. 2001. *Scientific Charge-Coupled Devices*. Bellingham, WA: SPIE-The Society of Photo-Optical Instrumentation Engineers, or J.D.E. Beynon and D.R. Lamb. 1980. *Charged-Coupled Devices and Their Applications*. McGrawHill.

[28]See, for example James R. Janesick. 2001. *Scientific Charge-Coupled Devices*. Bellingham, WA: SPIE-The Society of Photo-Optical Instrumentation Engineers, or J.D.E. Beynon and D.R. Lamb. 1980. *Charged-Coupled Devices and Their Applications*. London; New York: McGrawHill.

fined cost-effectively and without field stitching techniques, but with the limitation that printed features sizes must generally be on the order of 1 μm or larger. This should be contrasted to the lithography techniques needed for CMOS imagers employing aggressive transistor scaling (say 180 nm or below), where devices are usually patterned using high numerical aperture deep UV steppers or step-and-scan systems. Because of field size limitations of those lithography systems, CMOS imager chip sizes must currently be limited to standard lithographic field sizes of less than 33 × 22 mm, unless field stitching methods are used.

There are both advantages and disadvantages of shifting charge packets across a CCD imager to move charge to the output amplifier.[29] One advantage is that the very high fidelity of the charge transfer process, combined with the simple pixel design (a capacitor), results in very low fixed pattern noise.[30] A second advantage is that highly optimized readout amplifiers can be implemented, enabling very low readout noise, and it is also straightforward to implement correlated double sampling to eliminate the kTC noise associated with the gate capacitance of the readout circuit. To date scientific users such as astronomers, requiring the lowest possible noise levels, continue to use CCD imagers over CMOS imagers. Another benefit of being able to move charge packets from pixel to pixel is that charge packets can be moved during integration of the image, to allow a charge packet to "follow" a moving image. This feature is typically exploited for time delay-and-integrate (TDI) imagers that move the charge packets at a steady rate in one direction to precisely match the motion of an image slewing across the focal plane.

Applications for TDI imagers range from machine vision applications such as imaging parts in motion on a high-speed conveyor belt, to airborne imaging systems where large swaths of data are imaged in "pushbroom" fashion. A more recent development is a two-dimensional TDI capability available from the orthogonal transfer CCD (OTCCD), which allows charge shifting in both the *x*- and the *y*-axes, permitting left, right, up, or down shifting of the charge during integration.[31,32,33] The OTCCD can provide electronic image stabilization during image integration,

[29] Eric R. Fossum. 1991. Wire transfer of charge packets using a CCD-BBD structure for charge-domain signal processing. *IEEE Transactions on Electron Devices* 38(2):291-298.

[30] Eric Fossum, Sunetra K. Mendis, Bedabrata Pain, Robert H. Nixon, and Zhimin Zhou. California Institute of Technology, assignee. February 1, 2000. Active pixel sensor having intra-pixel charge transfer with analog-to-digital converter. U.S. Patent 6,021,172.

[31] B.E. Burke, R.K. Reich, E.D. Savoye, and J.L. Tonry. 1994. An orthogonal-transfer CCD imager. *IEEE Transactions on Electron Devices* 41(12):2482-2484.

[32] John Tonry and Barry E. Burke. 1998. The orthogonal transfer CCD. *Experimental Astronomy* 8(1):77-87.

[33] John Tonry, Barry Burke, and Paul Scheter. 1997. The orthogonal transfer CCD. *Publications of the Astronomical Society of the Pacific* 109:1154-1164.

allowing low light imaging with long integration times when images require stabilization to correct for atmospheric turbulence or platform jitter.

The need to shift charges over long distances in a CCD does have some disadvantages. For satellite applications, radiation damage of the silicon can result in defects that generate minority carriers under bias (bright spots) and trapping sites (which increase charge transfer inefficiency [CTI]), and shifting charge packets through a long path across the chip increases the probability that a charge packet will come into contact with one or more damage sites.[34,35,36] Thus, large CCDs are susceptible to displacement damage, whereas CMOS sensors, where photocharge is read immediately at the pixel, are generally much more tolerant of radiation damage. Another disadvantage of shifting charge packets is the power dissipation associated with the repetitive clocking of the phases. Since the electrodes are capacitors, the energy dissipated per phase clock cycle is ~ $CV^2$, where $C$ is the capacitance of the electrodes attached to the phase being clocked and $V$ is the voltage swing of the clock. Many scientific CCDs use relatively high voltages (say 10-15 V) compared to the voltages used in CMOS imagers (1-3 V); this can contribute significantly to power requirements. However, it should be noted that CCDs have been designed and fabricated to operate at low voltages, so that the voltage difference is not fundamental; in some cases the choice of a higher voltage is driven by application requirements such as the need to obtain deep depletion or a higher full-well charge capacity, and not by some inherent limitation of CCD technology. It is also worth noting that resonant energy recovery techniques can be applied to reduce power consumption from CCD clocks.

Finally, it should be noted that CCDs can be, and have been, successfully integrated monolithically with CMOS transistors, allowing on-chip control, analog-to-digital conversion, and processing functions. For cases where the cost of a custom CCD-CMOS process may be undesirable, clever three-dimensional packaging techniques provide an alternative way of placing separately fabricated CMOS on CCD chips.[37]

---

[34]V.A.J. Van Lint. 1987. The physics of radiation damage in particle detectors. *Nuclear Instrumentation Methods, Physics Research* A253:453-459.

[35]Albert J.P. Theuwissen. 2007. Influence of terrestrial cosmic rays on the reliability of CCD image sensors—Part 1: Experiments at room temperature. *IEEE Transactions on Electron Devices* 54(12):3260-3266.

[36]Albert J.P. Theuwissen. 2008. Influence of terrestrial cosmic rays on the reliability of CCD image sensors—Part 2: Experiments at elevated temperature. *IEEE Transactions on Electron Devices* 55(9):2324-2328.

[37]J.Y. Yang, A. Taddiken, and Y.C. Kao. 1991. Monolithic integration of GaAs LED array/Si CMOS LOGIC. *Technical Digest for the Gallium Arsenide Integrated Circuit (GaAs IC) Symposium* 301-304.

*Complementary Metal Oxide–Semiconductor Imagers*

The distinguishing feature of most CMOS imagers is that transistors are placed in each pixel, typically to allow resetting and readout of the detector within that pixel.[38] Figure 2-4 shows some examples of possible pixel electronics. Designing a very simple pixel with a single transistor per pixel is certainly possible and is analogous to the design of a DRAM (dynamic random access memory) cell. In fact, early DRAMs were occasionally used as imagers by experimentalists. When the pixel's transistor is switched on, any accumulated photocharge in the pixel is dumped onto a column line, allowing the pixel to be read out and reset to a reference voltage. When the switch is off, charge integrates on the pixel from photocurrent. This pixel design is not generally used for high-performance imaging, because the large column capacitance leads to low responsivity and high-input-referred noise.

More typical pixel designs employ three transistors.[39,40] One transistor is used to reset the pixel to a reference voltage, the second provides a source follower to buffer the charge integrated on the pixel and drive a voltage onto a column line, and the third transistor provides a selection transistor that allows only that pixel to drive the column when the pixel's row in the imager is selected for readout.

More complex pixels can include additional transistors to provide switches for snapshot shutters or correlated double sampling, or more sophisticated amplifiers such as capacitor transimpedance amplifiers (CTIAs).[41,42] It is also possible to place analog-to-digital converters in each pixel, resulting in direct digital outputs from each pixel. Another very different pixel design could be the readout electronics needed for a Geiger-mode[43] avalanche photodiode (APD) array, where digital outputs are generated when single photons are detected; depending on the readout design, a count of the number of detected photons could be kept in the pixel, or

---

[38]Zeljko Ignjatovic, Yang Zhang, and Mark Bocko. 2008. CMOS image sensor readout employing in-pixel transistor current sensing. In *Proceedings IEEE International Symposium on Circuits and Systems*, May.

[39]Bedabrata Pain, Thomas Cunningham, Shouleh Nikzad, Michael Hoenk, Todd Jones, Bruce Hancock, and Chris Wrigley. 2005. A back-illuminated megapixel CMOS image sensor. Jet Propulsion Laboratory, NASA.

[40]Vyshnavi Suntharalingam, Dennis D. Rathman, Gregory Prigozhin, Steven Kissel, and Mark Bautz. 2007. Back-Illuminated three-dimensionally integrated CMOS image sensors for scientific applications. *Proceedings of SPIE* 6690:6690009-9.

[41]X. Liu, B. Fowler, S. Onishi, P. Vu, D. D. Wen, H. Do, and S. Horn. 2005. CCD/CMOS hybrid FPA for low light level imaging. *Proceedings of SPIE* 5881:58810C.

[42]Haluk Kulah and Tayfun Akin. 1999. A current mirroring integration based readout circuit for high performance infrared FPA applications. *IEEE Transactions on Circuits and Systems—II: Analog and Digital Signal Processing* 50(4):181-186.

[43]In Ggeiger mode operation of an avalanche photodiode, the bias is sufficiently large that a single incident photon causes an uncontrolled discharge that is not self-limiting. Instead additional circuitry is supplied to remove the bias to reset the detector.

the time of arrival of the photon (time stamp) could be stored, as is desirable for LADAR applications.

CMOS pixel electronics can be integrated monolithically with silicon detectors, or detector arrays can be hybridized with CMOS readout integrated circuits (ICs).[44] A very typical hybridization arrangement is to bump-bond (usually with indium bumps) a detector array on top of a silicon CMOS readout IC.[45] One compelling reason for the hybridized arrangement is that detectors can potentially occupy 100 percent of the pixel area, since detector area does not compete with the transistors for real estate. Another great advantage of this technique is that the detector arrays can be fabricated in a different fabrication facility, using a different process, from the CMOS. Thus, special processes can be used to fabricate deep-depletion low-dark-current silicon p-i-n diodes, or higher-voltage devices such as APDs, without altering the CMOS foundry processes. Using bump-bonding can make small pixel pitches (below 10 μm) difficult, especially when high yield and 100 percent pixel operability are required. Recently three-dimensional integration processes using wafer bonding have been developed, and even disparate materials such as silicon and InP have been monolithically integrated, with pixel size down to 6 μm.[46,47,48,49]

A great advantage of monolithic CMOS imagers has been the ability to integrate a complete imaging system, including pixel electronics, addressing and control circuitry, analog-to-digital conversion, and even some signal processing into a single chip that has relatively simple digital interfacing requirements and does not require the user to design analog readout electronics (which often have

[44] A.G. Andreou, P.O. Pouliquen, and C.G. Rizk. 2009. Noise analysis and comparison of analog and digital readout integrated circuits for infrared focal plane arrays. Pp. 695-699 in *Proceedings of the 43rd Annual Conference on Information Sciences and Systems (CISS09)*, Baltimore, Md., March 18-20.

[45] Kun-Sik Park, Tae-Woo Kim, Yong-Sun Yoon, Jong-Moon Park, Jin-Yeong Kang, Jin-Gun Koo, Bo-Woo Kim, J. Kosonen, and Kwang-Soo No. 2007. Fabrication of a direct-type silicon pixel detector for a large area hybrid X-ray imaging device. *IEEE Nuclear Science Symposium Conference Record* M18-194-M18-197.

[46] S. Das, A. Chandrakasan, and R. Reif. 2003. Design tools for 3-D integrated circuits. Pp. 53-56 in *IEEE Proceedings of the Asia and South Pacific Design Automation Conference*.

[47] K. Banerjee, S. Souri, P. Kapur, and K. Saraswat. 2001. 3D ICs: A novel chip design for improving deep-submicrometer interconnect performance and systems-on-chip integration. *Proceedings of IEEE* 89(5):602-633.

[48] Steven E. Steen, Douglas LaTulipe, Anna W. Topol, David J. Frank, Kevin Belote, and Dominick Posillico. 2007. Overlay as the key to drive wafer scale 3D integration. *Microelectronic Engineering* 84(5-8):1412-1415.

[49] P. Leduc, F. de Crecy, M. Fayolle, B. Charlet, T. Enot, M. Zussy, B. Jones, J.-C. Barbe, N. Kernevez, N. Sillon, S. Maitrejean, and D. Louisa. 2007. Challenges for 3D IC integration: bonding quality and thermal management. Pp. 210-212 in *Proceedings of the IEEE International Interconnect Technology Conference.*

high standby power requirements).[50] A principal disadvantage is that the CMOS support electronics consume real estate on the chip. In the case of the per-pixel electronics, the area available in the pixel for the detector is reduced, so fill factor is often limited to 30 to 60 percent. This limits low-light performance of the devices, unless microlenses are used to improve the fill factor. Unfortunately microlenses are less effective when used in low-*F*/# (*F*-number) imaging systems and may not be appropriate for all applications. The presence of other electronics outside the pixel, such as banks of analog-to-digital converters, can also be an issue if the chip is going to be used in a four-side abutted (tiled) arrangement as part of a large mosaic focal plane. In some cases, as much as 50 percent of the die area may not be used for the actual image sensing, making it more difficult to array the chips without losing large parts of the imaging field, throwing away light, or having duplicate imaging systems (optics.)

The ability to pack a number of transistors into a small area of a pixel is improved if more deeply scaled CMOS processes are used.[51,52] However, more deeply scaled processes, especially those optimized for digital applications, often have limitations that can adversely affect imager performance. Maximum voltages are more limited, often to 1 to 2 V for deeply scaled (45 to 180 nm) processes, which reduces the dynamic range of the imager. If transistors are fabricated on thin epitaxial layers, or if the doping levels in the substrate are increased, the thickness of the silicon region able to collect photoelectrons is decreased, greatly limiting the red and VNIR responsivity. Microlens arrays may be needed to focus incident light onto small photosites (poor inherent fill factor), which may preclude effective use of the device with low-*F*/# optics.[53] In some cases, back-side illumination is now being used on CMOS devices to improve the fill factor and the spectral response, but the increased complexity makes these devices increasingly similar to the CCDs made by "specialized processes."[54,55]

---

[50]Fayçal Saffih and Richard Hornsey. 2007. Reduced human perception of FPN noise of the pyramidal readout CMOS image sensor. *IEEE Transactions on Circuits and Systems for Video Technology* 17(7):924-930.

[51]Michael Aquilino. 2006. Development of a Deep-Submicron CMOS Process for Fabrication of High Performance 0.25 mm Transistors. M.S. thesis, Rochester Institute of Technology, Rochester, NY. Available at https://ritdml.rit.edu/bitstream/handle/1850/5193/2006_Michael_Aquilino.pdf?sequence=1. Last accessed March 29, 2010.

[52]L. Wilson ed. 1997. *The National Technology Roadmap for Semiconductors*. San Jose, Calif.: Semiconductor Industry Association.

[53]E.A. Watson, W.E. Whitaker, C.D. Brewer, and S.R. Harris. 2002. Implementing optical phased array beam steering with cascaded microlens arrays. *Proceedings of IEEE Aerospace Conference* 3: 1429-1436.

[54]Tommy A. Kwa, Pasqualina M. Sarro, and Reinoud F. Wolffenbuttel. 1997. Backside-illuminated silicon photodiode array for an integrated spectrometer. *IEEE Transactions on Electron Devices* 44(5):761-765.

[55]A.G. Golenkov, F.F. Sizov, Z.F. Tsybrii, and L.A. Darchuk. 2006. Spectral sensitivity dependencies of backside illuminated planar HgCdTe photodiodes. *Infrared Physics and Technology* 47(3):213-219.

It should be noted that CMOS imagers shine in certain applications and that it is relatively straightforward to implement region-of-interest readout, or even random accessing of pixels, which can be very valuable for certain tracking functions. Moreover, fast shuttering is easier to implement in these imagers than in their CCD counterparts.

Because CMOS imagers are often targeted at low-cost applications and because deeply scaled CMOS processes are comparatively expensive on a cost-per-square-centimeter basis, CMOS imager manufacturers are incentivized to deliver the maximum number of pixels in the smallest possible die area. This is especially true for very cost-sensitive markets such as cell phone cameras, which are also the high-volume drivers for the market but have small profit margins. This has led to a trend to very small pixel sizes, which keeps the chip costs low and also allows chips with multiple megapixels to fit within lithographic stepper field sizes and to be produced with high yield. Unfortunately, these pixel sizes may not be well matched to certain optical systems and may lead to unfavorable trade-offs in dynamic range, low light sensitivity, and other parameters.

## Avalanche Photodiodes

Avalanche photodiodes, mentioned briefly above, are highly sensitive semiconductor electronic devices that use amplification by avalanche processes to enhance the sensitivity for low light levels (see Figure 2-5). The ideal APD would be low cost and would have background-limited dark noise, broad spectral and frequency response, no excess noise, and a gain that ranges from 1 to $10^6$ or more.

APDs achieve gain by accelerating either photoexcited carrier to energies above the bandgap where it can create an additional electron-hole pair by an inverse Auger process.

One specific example of an APD is characterized by a large active area (~180 μm), wide depletion region (~30 μm), a photon detection probability in excess of 50 percent, and a good timing response that is less than 300 ps.[56] Among the disadvantages of this device is that it has a high 300 V breakdown voltage, which is not compatible with silicon electronics; in addition, the process required to thin the device structure to allow the depletion region to reach through the entire device is proprietary.

The shortcomings associated with the reach-through APD were addressed with

[56]Don Phelan and Alan P. Morrison. 2008. Geiger-mode avalanche photodiodes for high time resolution astrophysics. Pp. 291-310 in *High Timing Resolution Astrophysics*, Don Phelan, Oliver Ryan, and Andrew Shearer, eds. New York: Springer.

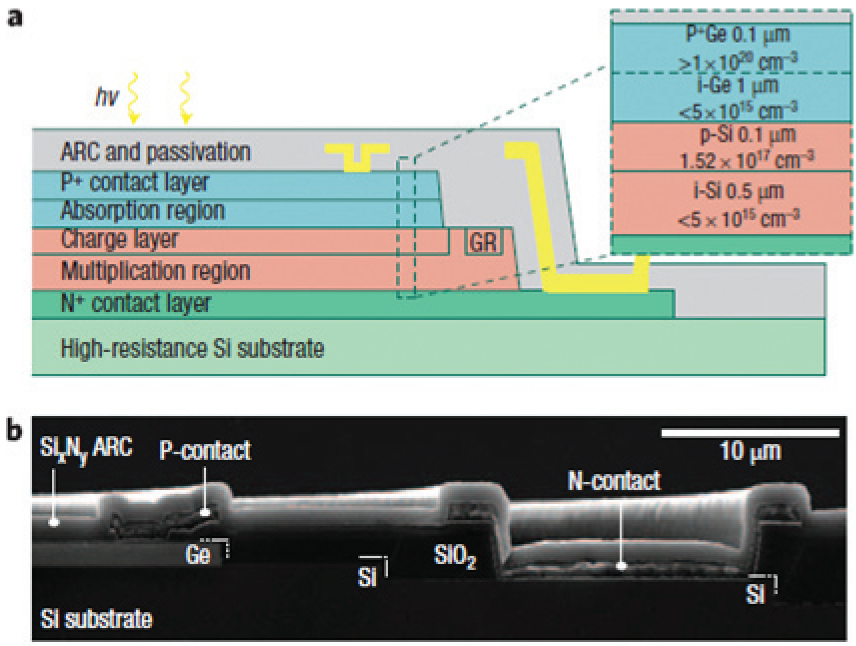


FIGURE 2-5
Schematic (a) and SEM (b) cross sections of a germanium/silicon APD. Doping concentrations and layer thicknesses were confirmed by secondary ion mass spectrometry (SIMS). The floating guard ring design labeled GR in cross section (a) was used to prevent premature breakdown along the device perimeter. ARC, anti-reflection coating. (Reprinted by permission from Macmillan Publishers Ltd: Nature Photonics. Yimin Kang, Han-Din Liu, Mike Morse, Mario J. Paniccia, Moshe Zadka, Stas Litski, Gadi Sarid, Alexandre Pauchard, Ying-Hao Kuo, Hui-Wen Chen, Wissem Sfar Zaoui, John E. Bowers, Andreas Beling, Dion C. McIntosh, Xiaoguang Zheng, and Joe C. Campbell. 2009. Monolithic germanium/silicon avalanche photodiodes with 340 GHz gain–bandwidth product. *Nature Photonics* 3:59-63. Copyright 2008.) These devices are widely deployed in long-wavelength and high-bit-rate optical transmission systems (see J.C. Campbell and H. Nie. 2000. High speed, low noise avalanche photodiodes. *Proceedings of the Device and Research Conference* 23:458-461).

the development of the Geiger mode avalanche photodiode (GM-APD).[57,58,59,60] In fact the steps used to fabricate this device are compatible with CMOS processing. The disadvantages, however, are that these devices have a limited active area on the order of approximately 50 μm. As the active area diameter increases, there is a rapid increase in dark count noise caused by the presence of process-induced defects, which act as carrier generation centers within the device.[61] In addition, this APD has a reduced quantum efficiency due to the smaller detection volume that is limited by the depletion region whose width is typically less than 1 μm.

In considering the schematic cross section for typical APDs a few basic structural elements are observed; these include an absorption region and a multiplication region. In the presence of light, an electric field that separates the photo-generated holes and electrons is present across the absorption region.[62] This field sweeps one carrier toward the multiplication region, which is designed to support a large electric field to provide internal photo-current gain by impact ionization.

The APD gain region must be wide enough to provide a gain of at least 100 for silicon APDs or 10-40 for germanium or InGaAs APDs. In addition to this, it is expected that the multiplying electric field profile must enable an effective gain at a field strength below the breakdown field of the diode.

If the reverse bias voltage is less than the breakdown voltage, the avalanche dies down due to losses. When this happens a single photon will generate hundreds or even thousands of electrons. Above the breakdown voltage, the acceleration of the current carriers is great enough to keep the avalanche alive even without additional external stimulus. Thus, a single photon is sufficient to generate a constant current that can be measured using external electronic equipment. This current is calculated as

[57] A.M. Moloney, A.P. Morrison, C.J. Jackson, A. Mathewson, and P.J. Murphy. 2002. Large-area geiger-mode avalanche photodiodes for short-haul plastic optical fiber communication. *Proceedings of SPIE, Opto Ireland, Optoelectronic and Photonic Devices* 4876:438-445.

[58] A.M. Moloney, A.P. Morrison, J.C. Jackson, A. Mathewson, and P.J. Murphy. 2002. Geiger mode avalanche photodiode with CMOS transimpedance amplifier receiver for optical data link applications. *IT&T Annual Conference, Transmission Technologies.*

[59] A.P. Morrison, V.S. Sinnis, A. Mathewson, F. Zappa, L. Variscoand M. Ghioni, and S. Cova. 1997. Single-photon avalanche detectors for low-light level imaging. *Proceedings of SPIE, EUV, X-Ray, and Gamma-Ray Instrumentation for Astronomy VIII* 3114:333-340.

[60] In Geiger mode operation of an avalanche photodiode, the bias is sufficiently large that a single incident photon causes an uncontrolled discharge that is not self-limiting. Instead additional circuitry is supplied to remove the bias to reset the detector.

[61] E.A. Dauler, P.I. Hopman, K.A. McIntosh, J.P. Donnelly, E.K. Duerr, R.J. Magliocco, L.J. Mahoney, K.M. Molvar, A. Napoleone, D.C. Oakley, and F.J. ODonnell. 2006. Scaling of dark count rate with active area in 1.06 μm photon-counting InGaAsP/InP avalanche photodiodes. *Applied Physics Letters* 89(11):111102.

[62] Available at http://www.perkinelmer.com/. Last accessed March 29, 2010.

$$I_S = M \cdot R_0(\lambda) \cdot P_S,$$

where $R_0(\lambda)$ is the spectral responsivity of the APD, $M$ is the internal gain, and $P_s$ (watts) is the incident optical power. The gain of the APD is dependent on the applied reverse bias voltage.

APDs are recommended for very high bandwidth applications or where internal gain is needed to overcome secondary amplifier noise. The devices are typically used for low-light detection, laser radar systems, optical data transmission, bar-code scanners, or biomedical equipment.[63] They have found their way into military, medical, and communications applications and include positron emission tomography and particle physics.

APDs are photodetectors that provide first-stage gain through avalanche multiplication. APDs show an internal current gain effect of around 100 due to impact ionization (avalanche effect) and can employ doping techniques that allow greater voltage to be applied before breakdown is reached and therefore allow for a greater operating gain (>1,000).

Avalanche photodiode arrays generate digital outputs when a single photon is detected. These devices are capable of storing a count of the number of photons detected and the time of arrival of the photon on a single pixel. These photodiodes exhibit photoelectron gains of up to about 50, but issues associated with excess noise and nonuniformity have precluded widespread use of this phenomenon for FPAs employed in low-light-level applications. Attempts to move to longer wavelengths significantly decrease the yields due to the deviation from material lattice mismatch to the indium phosphide substrate, and the added defect densities contribute to excess dark current. This technology is being incorporated into higher-end short-wavelength infrared (SWIR) imaging and hyperspectral sensors but is currently too expensive for high-volume soldier sensor applications.

As is true of most detectors, the utility of APDs depends on many parameters such as quantum efficiency and total leakage current (the sum of the dark current, photocurrent, and noise). Knowledge of these parameters is important to fully characterize and efficiently operate avalanche photodiodes in Geiger mode.[64]

Thus, it is necessary to quantify the dark count rate, the excess reverse bias voltage, the optimum operating temperature, the photon detection probability, the after-pulsing probability, and the hold-off time. Because many of these parameters are interdependent, it is necessary to perform trade-offs between the variables to

[63]More information on avalanche photodiodes is available at http://www.lasercomponents.com/fileadmin/user_upload/home/Datasheets/lc/applikationsreport/avalanche-photodiodes.pdf. Accessed March 29, 2010.

[64]Geiger mode is an avalanche mode in which an unlimited avalanche occurs based on detection of a single photon or any number of photons. The avalanche is quenched by reducing the bias voltage.

achieve optimum performance for specific applications. For example, it is desirable to have a low dark count rate because this will maximize signal-to-noise ratio and minimize statistical uncertainty. A decrease of the dark count rate occurs exponentially with decreased temperatures; however, operating at low *T* increases the after-pulsing probability. Thus, changing one parameter, while it will improve performance in a specific area, will affect other parameters as well.[65]

## Near Infrared

### *Silicon*

Silicon sensors are sensitive throughout the visible to wavelengths as long as the silicon bandgap of ~1.1 μm. Just the opposite of the UV situation discussed above, the very long absorption length associated with the indirect bandgap of silicon requires very different optimization of the device structure for quantum efficiency and carrier collection.

### *Intensifiers*

In many fields, it is common to use image intensifiers in front of a camera tube because they permit cameras to work at the lowest light levels possible. These are electron optic systems that are made up of an input phosphor-photocathode screen that converts incoming radiation into a beam of electrons, electrodes to control the movement of electrons, and an output phosphor screen that produces the output image.[66] They convert spectral radiation to a visible light image, which after additional processing can be displayed on a monitor. Most commercially available image intensifiers have axial symmetry; however, some nonaxisymmetrial intensifiers have recently been designed.[67] Intensifiers work utilizing an avalanche or Geiger mode gain in back of a photocathode. Thus, extremely small photon fluxes are multiplied several thousandfold, allowing viewing under extremely low light conditions.

[65] B.S. Robinson, D.O. Caplan, M.L. Stevens, R.J. Barron, E.A. Dauler, S.A.Hamilton, K.A. McIntosh, J.P. Donnelly, E.K. Duerr, and S. Verghese 2005. High-sensitivity photon-counting communications using Geiger-mode avalanche photodiodes. *Proceedings of the IEEE Lasers and Electro-Optics Society* 559-560.

[66] K.G. Vosburgh, R.K. Swank, and J.M.J. Houston. 1997. X-ray image intensifiers. *Advances in Electronics and Electron Physics* 43:205-244.

[67] N.W. Adamiak, J. Dabrowski, and A. Fenster. 1996. Design of nonaxisymmetrical image intensifiers. *IEEE Transactions on Industry Applications* 32(1):93-99.

## Short-wavelength Infrared

InGaAs detector technology is quite well developed as a result of its dominant use in fiber-optic telecommunications at ~1.3 to 1.7 μm. By varying the composition, the bandgap can be shifted to as long as 2.6 μm.

## Mid-, Long-, and Very Long Wavelength Infrared

*Brief History of Infrared Detection*

Thallium sulfide and lead sulfide (or galena) were among the first infrared detector materials, developed during the 1930s. Many other materials have been investigated for application to infrared detection. Lead-salt detectors are polycrystalline and are produced using vacuum evaporation and a chemical deposition process from solution followed by post-growth sensitization.[68] Reproducibility has historically been poor, but well defined, although somewhat empirical recipes were eventually found.

Significant improvement in detector manufacture occurred with the discovery of the transistor, which stimulated growth and material purification techniques. This resulted in novel techniques for detector production from single crystals.

High-performance detectors were initially based on the use of germanium with the introduction of controlled impurities. Development of high-performance visible and NIR detectors based on silicon began to occur in the 1970s after the invention of the CCD. This resulted in the development of sophisticated readout schemes that allowed both detection and readout to occur on one common silicon chip.

In the 1950s, there was extensive investigation of III-V semiconductors. As a result of its small bandgap (5 μm at 77 K), InSb showed promise as a material for MWIR detection, and indeed vastly improved InSb FPA arrays remain a mainstay of cooled MWIR imaging.

Shortly thereafter, in 1959, HgCdTe (mercury cadmium telluride, or MCT) was found to exhibit semiconducting properties over much of its composition range. The alloy's bandgap was variable from 0.0 to 1.605 eV. Later, long-wavelength photoconductivity was demonstrated in HgCdTe, leading the way to development of infrared detectors.

A shift occurred in the mid-1960s toward using the PbSnTe alloy because of production and storage problems associated with HgCdTe. However, limitations in the speed of response for PbSnTe detectors and the better suitability of HgCdTe for

[68] A. Rogalski, and J. Pitrowski. 1988. Intrinsic infrared detectors. *Progress in Quantum Electronics* 12:287-289.

infrared imaging device production, as well as improvements in the technology of the material, once again shifted the focus back to HgCdTe at the beginning of the 1970s. In the mid- to late 1980s, HgCdTe remained the most promising narrow-gap semiconductor for infrared detector arrays. Today, after many years of intensive development, photovoltaic HgCdTe is widely used across all infrared bands for high-performance IR FPAs.

### *Indium Antimonide*

InSb MWIR detectors have been developed continuously since the 1950s and are a quite mature technology. InSb has a bandgap of about 5.4 μm at 77 K, making this material a good choice for MWIR detection. InSb detectors are based on bulk materials rather than epitaxy, and processing involves impurity diffusion or ion implantation. Relatively large wafers, ~3 to 4 inches (100 mm), are available.

### *Mercury Cadmium Telluride*

HgCdTe is a ternary compound whose bandgap can be adjusted by varying the relative proportions of mercury and cadmium. HgCdTe is a pseudobinary alloy between HgTe and CdTe, written as $Hg_{1-x}Cd_xTe$. The composition range $0.21 < x < 0.26$ covers the LWIR regime. Nearly all of today's LWIR HgCdTe is grown epitaxially in thin layers by molecular beam epitaxy (MBE) or liquid-phase epitaxy (LPE). Both p-type and n-type doping can be reproducibly accomplished. The composition can be varied during growth, allowing the formation of heterojunctions, barriers, and multiband devices. The most commonly used substrate is CdZnTe, but there is extensive work on using both silicon and GaAs substrates, although the large lattice mismatch has resulted in slow progress. Development of photovoltaic arrays began more than 30 years ago, and a high level of technology readiness has been achieved.

Recently, Tennant presented an empirical result, known as Rule 07 (the 07 refers to the year this result was obtained and was used to stress its transient status),[69] that provides a characterization of dark current as a function of bandgap and temperature across a wide range of HgCdTe materials. Figure 2-6 presents the measured and semiempirical model. Tennant revisited this characterization and found that it remains a reliable guide.[70] This rule is relevant to high-quantum-efficiency devices,

[69]W.E. Tennant, D. Lee, M. Zandian, E. Piquette, and M. Carmody. 2008. MBE HgCdTe technology: a very general solution to IR detection, described by "Rule 07," a very convenient heuristic. *Journal of Electronic Materials* 37(9):1406-1410.

[70]W.E. Tennant. 2010. "Rule 07" revisited: still a good heuristic predictor of *p/n* HgCdTe photodiode performance? *Journal of Electronic Materials* DOI:10.1007/s11664-010-1084-9.

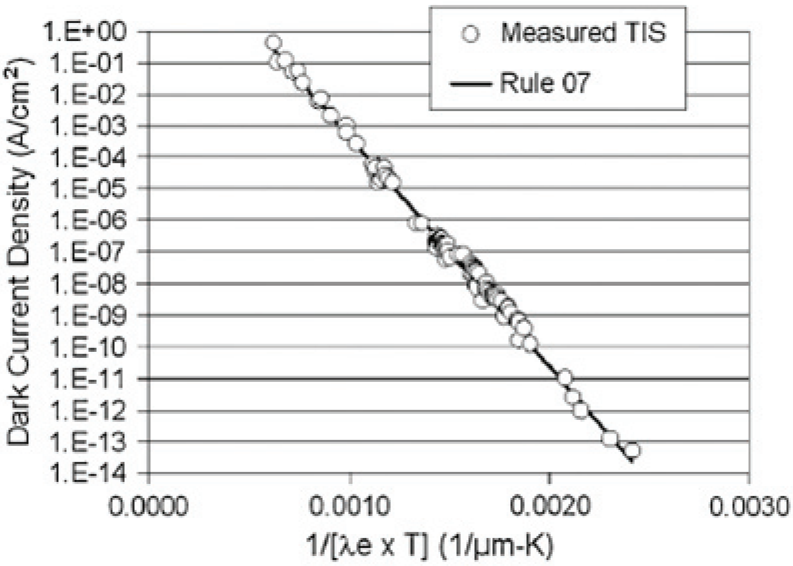


FIGURE 2-6
Dark current density for HgCdTe as a function of the cutoff wavelength × temperature product. NOTE: TIS = Teledyne Imaging Sensors. SOURCE: W.E. Tennant, 2010. "Rule 07" revisited: still a good performance heuristic predictor of *p*/*n* HgCdTe photodiode performance. *Journal of Electronic Materials*. DOI: 10.1007/s11664-010-1084-9.

and the analysis suggests that Auger recombination (Auger 1) in the n-type HgCdTe is the dominant limiting mechanism.[71] This latest paper also compared both the theoretical and the experimental results for various strain-layer superlattice (SLS) structures against the HgCdTe results. The best SLS results are approaching the Rule 07 limits, while the theoretical work shows that significant improvement remains possible pending improvements in materials quality and processing.

Tennant then went on to compare this current density with the background dark current for a $4\pi$ steradian background at the operating temperature of the device (e.g., surrounded by a cold shield). This result is shown in Figure 2-7. For 77 K operation and a MWIR cutoff, the device is close (within a factor of ~5) to this background limit, while for both SWIR and LWIR operation, the dark currents are substantially above the BLIP limit for this very low background situation.

[71]Auger processes in semiconductors are three-body interactions in which an electron-hole pair recombines without emission of a photon but rather with excitation of a second carrier to a higher-energy state. Auger processes are essentially the inverse of impact ionization in which an energetic carrier relaxes by creating an electron-hole pair.

Tennant ascribes the LWIR result to Auger recombination processes, while at short wavelengths the increased noise results from a deviation between the optical and electrical bandgaps. Substantial efforts have already been made to reduce the Auger recombination, and it does not appear likely that much further improvement is available in present material and device configurations. Thus, the low-temperature "external radiative limit" remains an elusive goal that does not appear approachable within the constraints of current technology.

It is important to recognize that the BLIP current for LWIR tactical applications, looking at a 300 K background with *F*/1 optics, is much higher than this "external radiative" limit. At 77 K operation temperature and with a 12 μm LWIR cutoff, the dark current is ~ $10^{-4}$ A/cm$^2$ (see Figure 2-7). From the blackbody irradiance (300 K, *F*/1 optics), the dark current is ~ 0.18 A/cm$^2$, orders of magnitude greater than the intrinsic device dark current.

**FINDING 2-3**

MWIR and LWIR detectors are already close to fundamental BLIP for terrestrial operations that look at a 300 K background. Future innovations will focus on device and system optimization for specific applications.

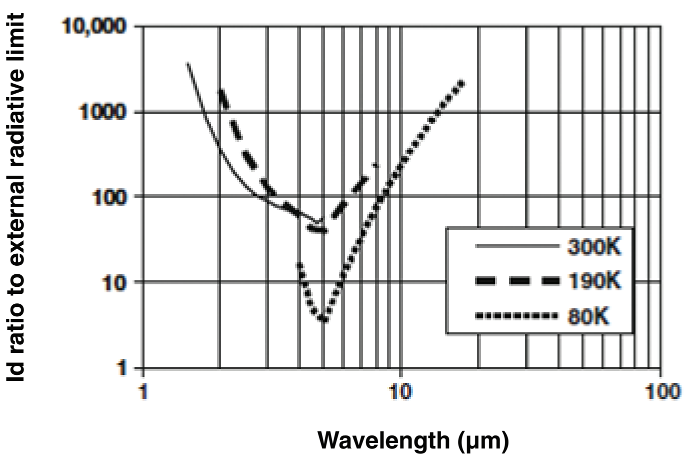


FIGURE 2-7
HgCdTe dark current from Rule 07 relative to the external radiative limit, corresponding to a cold shield at the device temperature. At MWIR there is relatively little room for improvement at the lowest temperatures; however at other wavelengths there is substantial excess dark current, which limits the detector performance for low-background situations. SOURCE: W.E. Tennant, D. Lee, M. Zandian, E. Piquette, and M. Carmody. 2008. MBE HgCdTe technology: a very general solution to IR detection, described by "Rule 07," a very convenient heuristic. *Journal of Electronic Materials* 37(9):1406-1410.

*Strained-layer Superlattice*

Strained-layer superlattice material consists of alternating thin layers of InAs and GaSb. A typical LWIR example of the superlattice structure has a period consisting of 4.4 nm of InAs and 2.1 nm of GaSb. This pair is repeated 300 times or more to form the IR-absorbing region. Because of the type II band offset between the two constituent materials, in which the conduction band of InAs is below the valence band of GaSb, the structure exhibits new bands for holes and electrons, which are separated by an energy difference that is smaller than either of the bandgaps of the InAs or GaSb themselves and is adjustable by varying the thicknesses of the layers. This small effective bandgap is suitable for absorbing IR photons. The structure is grown by MBE and the commonly used substrate is GaSb. Both *n*- and *p*-type doping have been demonstrated. Heterojunctions are typically formed by growing contacting regions adjacent to the absorbing region, having a shorter superlattice period and a wider effective bandgap than the absorber. SLS materials are based on very well developed III-V materials, and the vast experience in bandgap engineering in these and related systems holds promise for continuing developments. Several variants have been and continue to be introduced including "W" and "M" structures.[72,73] This remains an active research area with significant potential for dramatic advances.

Additionally, the basic SLS structure can be modified by inserting a very thin (a few angstroms) layer of AlSb as a barrier for the majority carrier electrons. This opens up a wide parameter space for bandgap engineering, to enable specialized barriers as well as various heterojunction designs. Devices incorporating these structures have been grown and tested recently with the aim of reducing the dark current.[74]

The SLS LWIR technology development effort has received substantial funding recently because of its potential as a future, low-cost, III-V compatible replacement for HgCdTe in some applications. Although progress has been made the device performance, the dark currents remain significantly greater than those of HgCdTe as shown in Figure 2-8. These are laboratory studies, the technology readiness level of SLS material is well behind that of HgCdTe.

[72]B.M. Nguyen, D. Hoffman, P.Y. Delaunay, and M. Razeghi. 2007. Dark current suppression in type II InAs/GaSb superlattice long wavelength infrared photodiodes with M-structure barrier. *Applied Physics Letters* 91:63511-1.

[73]E.H. Aifer, J.G. Tischler, J.H. Warner, I. Vurgaftman, W.W. Bewley, J.R. Meyer, J.C. Kim, L.J. Whitman, C.L. Canedy, and E.M. Jackson. 2006. W-structured type-II superlattice long-wave infrared photodiodes with high quantum efficiency. *Applied Physics Letters* 89:053519-1.

[74]D.Z. Ting, C.J. Hill, A. Soibel, S.A. Keo, J.M. Mumolo, J. Nguyen, and S.D. Gunapala. 2009. A high-performance long wavelength superlattice complementary barrier infrared detector. *Applied Physics Letters* 95:023508.

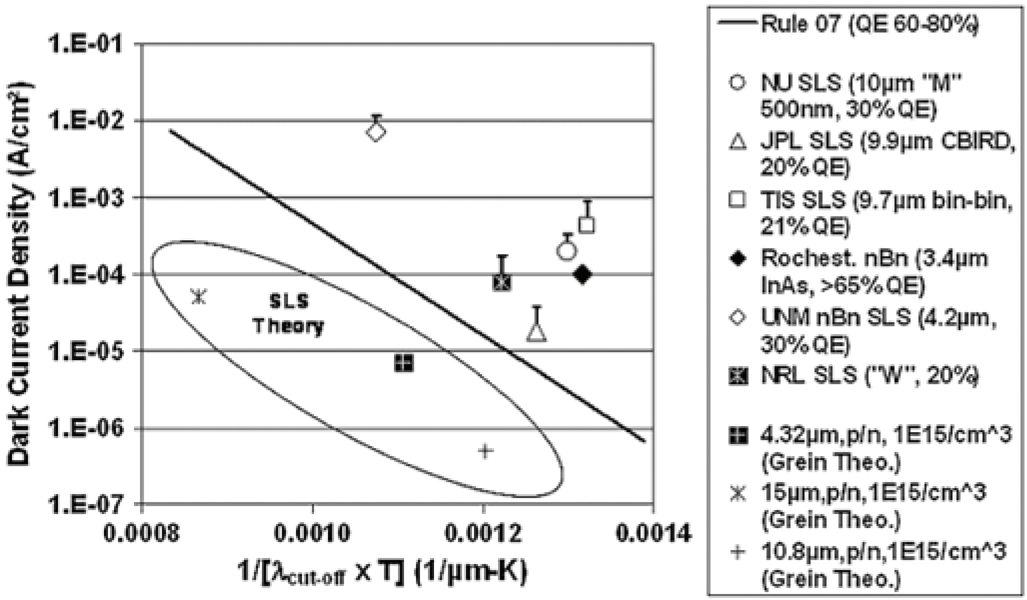


FIGURE 2-8
Comparison of theoretical (*inside circled region*) dark currents of SLS devices with the HgCdTe Rule 07 metric (*solid line*). The experimental dark currents are above those achieved in HgCdTe, while the theory shows a potential advantage of SLS pending better materials and device processes. SOURCE: Tennant, W.E., D. Lee, M. Zandian, E. Piquette, and M. Carmody. 2008. MBE HgCdTe technology: a very general solution to IR detection, described by "Rule 07," a very convenient heuristic. "Rule 07" revisited: still a good performance heuristic predictor of *p/n* HgCdTe photodiode performance. *Journal of Electronic Materials* 37(9):1406-1410 DOI: 10.1007/s11664-010-1084-9.

### *Quantum-well Infrared Photodetectors and Quantum-dot Infrared Photodetectors*

Quantum-well infrared photodetectors (QWIPs) and quantum-dot infrared photodetectors (QDIPs) are unipolar photoconductive devices based on intraband absorption between electronic levels defined by quantum confinement in traditional III-V semiconductors, principally GaAs and InP. The promise is that the III-V growth and processing technology is quite mature, substrates are readily available, and scaling to large arrays should be simpler (and have higher yield) than for HgCdTe-based devices. The issues are related to the relatively weak absorption associated with the quantum confined structures.

For the case of intrasubband transitions in III-V QWs, selection rules forbid the absorption of normally incident light requiring a grating or other optical element to scatter the incident light into the QW plane. Due to the weak absorption, associated with the small fill factor of the QWs relative to the wavelength, this can lead to cross-talk issues and constraints on decreasing the pixel size. QDIPs, as a result of the three-dimensional confinement, eliminate this selection rule allowing normal incidence detection, but the absorption is still quite weak, about 2 to 3 percent for a single pass through a typical active layer, which results in poor quantum efficiency. This is somewhat alleviated in the detectivity by the low dark currents, which also depend on the total volume of quantum dots. Recently, there has been quite a bit of activity in adding nanostructures such as plasmonics to QDIPs; this is discussed more fully in Chapter 4. Rogolski has recently reviewed progress in both HgCdTe and QWIPs-QDIPs for FPAs.[75,76,77]

*Very Long Wavelength Infrared*

Many of the same detector technologies being developed for the LWIR also can be optimized for VLWIR operation beyond 12 μm. Historically, bulk doped semiconductors have dominated in this spectral region, which is potentially important for missile detection against a cold (space) background. Mercury cadmium telluride detectors suffer from increasing noise due to thermally generated carriers and Auger processes at these land wavelengths. Both type II superlattice and QDIP-QWIP detectors have shown promise for this spectral region. This remains an active area of investigation.

**FINDING 2-4**
Continued detector advancement requires improved growth and processing of low defect density compound semiconductor materials. The 30-year trend has been improvements in existing materials along with the incorporation of nanoscale structures in one, two, and three dimensions.

## FABRICATION OF DETECTORS AND FOCAL PLANE ARRAYS

### Detectors

Each material system brings its own unique set of fabrication issues to maintain high performance. Overall the dimensional scale of even visible pixels is large

[75] A. Rogalski. 2006. Competitive technologies of third generation infrared photon detectors. *Opto-Electronics Review* 14(1):87-101.
[76] P. Martyniuk and A. Rogalski. 2008. Quantum-dot infrared photodetectors: status andoutlook. *Progress in Quantum Electronics 32*(3-4):89.
[77] P. Martyniuk, S. Krishna, and A. Rogalski. 2008. Assessment of quantum dot infrared photodetectors for high temperature operation. *Journal of Applied Physics 104*(3):034314-1.

compared to the minimum feature size of current lithographic tools (which are following a Moore's law curve with current manufacturing at the 45 nm node).

## Focal Plane Arrays

A focal plane array is created by arranging individual detector elements in a lattice-like array. Individual detectors in an array are often referred to as pixels, short for picture elements. However, the process of developing an integrated array of detectors is significantly more challenging than fabricating an individual detector element. The overall scheme of silicon-based visible detectors is discussed in the sections on CMOS and CCDs. As a result of the advanced state of silicon technology, these imaging chips integrate to some extent both the detection and the electronics. For infrared detectors, in contrast, the signals have to be moved from the detector material to silicon circuitry, called the readout integrated circuit (ROIC); this is usually accomplished by bonding each pixel to a silicon readout circuit using a myriad of indium bump bonds. The number of bonds scales as the number of pixels, and for very large arrays this is a difficult manufacturing step. Typically each pixel has one independent contact and shares the second contact with other pixels in the array. The distribution of the common contacts impacts electrical cross-talk and readout speed.

A fundamental limitation in the development of arrays of detectors is that light is easily coupled to neighboring pixels in an array, which leads to the development of false counts, or cross-talk.[78] There are approaches that may be exercised to mitigate this limitation, but they add additional complexity to the manufacturing.[79] In addition, the array fabrication process becomes even more complicated by the requirement to maintain low leakage current in the individual pixels, making the fabrication process even more unwieldy.[80] The progress in arrays has been steady and has paralleled the development of dense electronic structures such as DRAMs.

[78]Don Phelan and Alan P. Morrison. 2008. Geiger-mode avalanche photodiodes for high time resolution astrophysics. Pp. 291-310 in *High Timing Resolution Astrophysics*, Don Phelan, Oliver Ryan, and Andrew Shearer, eds. New York: Springer.

[79]J. Ziegler, M. Bruder, M. Finck, R. Kruger, P. Menger, T. Simon, and R. Wollrab. 2002. Advanced sensor technologies for high performance infrared detectors. *Infrared Physics and Technology* 43(3-5):239-243.

[80]Alexis Rochas, Alexandre R. Pauchard, Pierre-A. Besse, Dragan Pantic, Zoran Prijic, and Rade S. Popovic. 2002. Low-noise silicon avalanche photodiodes fabricated in conventional CMOS technologies. *IEEE Transactions on Electron Devices* 49:387-394.

### Manufacturing Infrastructure

The manufacturing infrastructure for large array fabrication is discussed in Chapter 5. At this point it suffices to recognize that the manufacturing tools are largely those developed by the integrated circuit industry and adapted for FPA manufacturing. The FPA industry is not sufficiently large to support the development of a complete set of unique tools. As the evolution of the silicon industry is driven by a different set of goals, this can lead to divergence and to gaps in the FPA tool set. One simple example is that the silicon industry has standardized on a field size of $22 \times 33$ mm$^2$ for its lithography tools. The drive to larger pixel counts for FPAs often requires much larger overall FPA sizes, which can only be accomplished by abutting multiple fields, requiring special considerations in the design of the focal plane arrays.

**FINDING 2-5**
An advanced equipment set is required for manufacturing large-pixel-count detector arrays. Equipment availability is dependent on leveraging silicon CMOS developments. The detector market is not in itself sufficiently large to drive equipment development.

## CONCLUDING THOUGHTS

Detection and imaging of electromagnetic radiation across the UV, visible, and infrared spectrum has a long history. As a result of its very advanced stage of technological development, silicon is now, and undoubtedly will continue as, the dominant material for visible sensors. One exception is the need for solar blind detectors that are insensitive to the solar spectrum after it is filtered bypassing through the ozone layer surrounding the Earth. Large-bandgap materials such as AlGaN are being actively developed for this application. First-generation night vision systems used intensified (amplified) visible detection. Increased interest is now being placed on SWIR detection using InGaAs and related materials technology. Much of the progress at these longer wavelengths was catalyzed by the needs of the telecommunications industry for fiber-optic receivers. Infrared detectors have been under development for many years, primarily for military applications. The traditional material systems for cooled detectors are InSb for MWIR and HgCdTe for both MWIR and LWIR. Emerging material systems include SLS antimonides and intersubband transition QWIPs and QDIPs in the AlGaAs system. Both of these have the advantage of epitaxial growth on GaAs and possibly silicon substrates and of leveraging off of the mature GaAs technology developed for electronics and photonics. However, they are at a much earlier stage of development and technology readiness. Figure 2-9 shows the material systems relevant for different wavelength regimes.

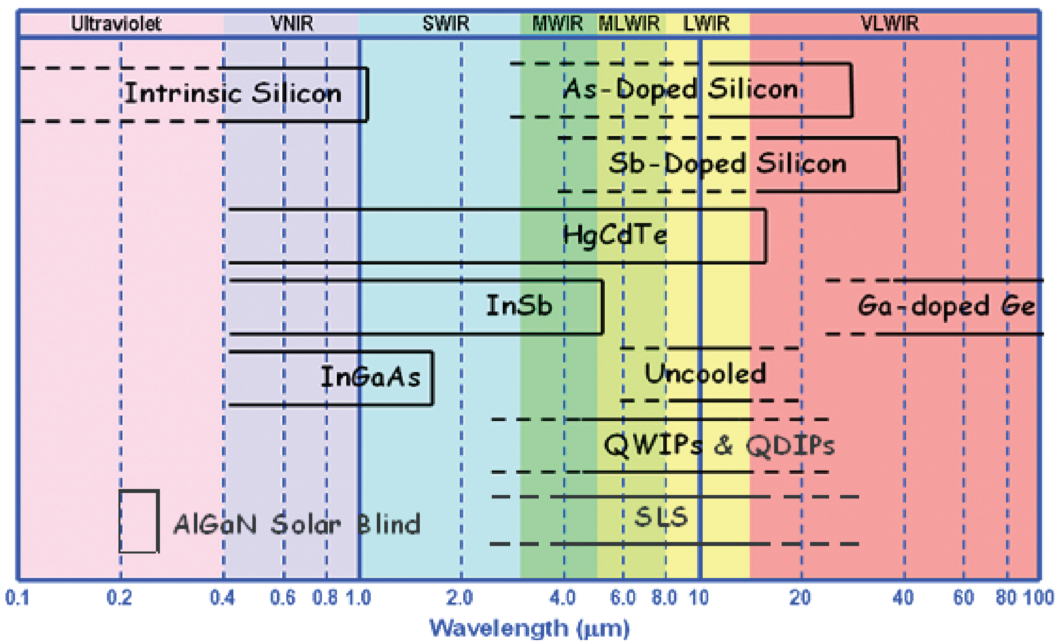


FIGURE 2-9
Material systems for UV-visible-infrared detection. Except for the bottom two entries, these material systems have been known and developed for decades. SOURCE: Presented to the committee by Dr. "Dutch" Stapelbroek, University of Arizona.

# 3

# Key Current Technologies and Evolutionary Developments

## INTRODUCTION

Advanced-technology sensors, coupled with data processing and fusion and networked communications, have enabled new approaches to warfighting and have been essential for reducing combat losses, minimizing collateral damage, and allowing a smaller but more effective military force. U.S. military strategy is expected to remain dependent on maintaining technological superiority, including the development of increasingly sophisticated sensor systems and new sensor types that exploit novel threat signatures. While the possibility exists that radically new sensor technologies will be invented, there are many opportunities to customize and improve the performance of existing, and in some cases relatively mature, focal plane technologies. There is an active research community pursuing these goals on a global scale. The focus of this chapter is to describe some of the likely near-term developments to existing visible and infrared (IR) detector technologies.

The proper definition of a sensing problem drives optimization of the design features of the focal plane to lead to the best system-level performance. In many cases, new sensor systems will be enabled not just by driving closer to "physical limits" but by tailoring designs and performance to be exceptionally well matched to a specific application. Clever design architectures, adding powerful on-focal-plane processing features, lowering power dissipation, increasing detector operating temperature, and dramatically reducing costs are examples of seemingly evolutionary changes that could enable revolutionary capabilities. Innovation will likely not be driven exclusively by moving closer to the physical limits of detector performance.

A number of advances, such as putting more processing into a pixel or making a smaller pixel, depend on continued improvements in silicon complementary metal oxide semiconductor (CMOS) process technology, driven by the semiconductor industry's push to stay on the Moore's law scaling curve. It is worth noting that device scaling is becoming increasingly technically challenging and expensive and that physical limits are becoming real roadblocks to CMOS scaling. Many process technologies developed for 65 nm, 45 nm, or 32 nm, while excellent in terms of digital circuit performance, are less than ideal for the analog portions of a pixel. They may have high transistor leakage levels and very limited dynamic range due to smaller voltage swings, thus limiting performance. Accordingly, device scaling is likely to be exploited by adding digital functionality to pixels. As more advanced CMOS process technologies are used, the cost per transistor drops, but the cost per area of CMOS chips increases. This, in turn, carries negative implications for focal planes that require physically large areas.

Advanced processes are also costly and have large nonrecurring expenses associated with design, mask generation, and initial design and debugging, making it expensive to develop custom devices needed only in small volumes. It is also worth noting that changes in the lithography processes used at smaller feature sizes require specialty techniques, such as field stitching, to produce large-area devices that exceed the field size of modern lithography tools. In short, while CMOS scaling will dramatically shape the design options for advanced focal planes, significant learning and innovation will be required to apply these advanced technologies optimally. In many cases, visible sensor applications do not benefit from the high-volume manufacturing imperatives that both drive and allow amortization of the increasing tool cost that accompanies CMOS scaling.

A number of areas in which near-term technology advances are expected include ultralarge-format arrays; mosaic tiling technologies; pixel size reduction; smarter pixels and on-focal plane processing; improved three-dimensional (3-D) integration and hybridization; higher-operating-temperature devices; multicolor pixels; improved short-wavelength infrared (SWIR) arrays; photon counting technologies and lower readout noise; curved focal surfaces; lower-power operation; radiation hardening; cost reduction; and improved cooler technology.

## KEY TECHNOLOGIES EXPECTED TO DRIVE ADVANCEMENTS IN EXISTING DETECTOR TECHNOLOGIES OVER THE NEXT 10-15 YEARS

With respect to areas in which near-term technology advances are expected, this section examines each of the advances and addresses benefits, risks or drawbacks, impact on system performance, and implications for military applications of the expected advances.

## Ultralarge-format Focal Plane Arrays

Progress is expected to continue in developing increasingly large arrays for the visible, SWIR, mid-wavelength infrared (MWIR), long-wavelength infrared (LWIR), and very long wavelength (VLWIR) portions of the spectrum. While both visible and IR arrays depend on the development of larger CMOS readout integrated circuits (ROICs), IR arrays have additional challenges—for example, yielding large detector arrays in difficult material systems such as mercury cadmium telluride (MCT), availability of large substrates, and developing high-yielding and small pixel-pitch bump bonding.

Low-cost and high-performance (by some metrics) CMOS active pixel sensor (APS) multi-megapixel arrays are commercially available for the visible spectrum. Charge-coupled devices (CCDs) for scientific applications are routinely made with pixel counts exceeding 20 megapixels, and the trend to ever-larger single chips will continue. Using four-side buttable tiling, CCD focal planes of 1.4 gigapixels have been made and are in use by astronomers. In addition to monolithic silicon active-pixel sensors and CCDs, hybrid megapixel visible arrays will also become more available, providing sensitivity advantages over the traditional commercial products. Visible 50-megapixel arrays are now available with digital outputs having ROIC noise levels of less than 10 electrons and offering a sensitivity advantage over consumer products.

In the IR, highly sensitive, multi-megapixel infrared focal plane arrays (FPAs), up to 4 megapixels, are also widely available for the MWIR and SWIR spectrum. Recently, 16-megapixel MWIR FPAs have been demonstrated. As shown in Figure 3-1, detector array pixel count has paralleled the exponential growth of silicon dynamic random access memory (DRAM) bit capacity.[1] Visible arrays with close to 100 megapixels offer the largest formats. In the IR, 16-megapixel arrays are now available.

System-level benefits of large FPAs are generally related to providing a large instantaneous field of view (FOV). As focal planes become less expensive per pixel, it increasingly makes sense to eliminate costly, power-hungry, heavy, and unreliable mechanical scanning or optical pointing systems and replace these with a fully electronic selection of the FOV by reading out a region of interest from a larger FPA. Large FPAs allow monitoring of large areas and are important for persistent surveillance applications. These FPAs enable important applications, such as high-resolution, wide-area airborne persistent surveillance and distributed aperture systems providing full spherical coverage of platforms.

The single 8-inch ROIC wafer shown in Figure 3-2 contains 2K × 2K, 2K × 4K,

[1] Paul Norton. 2006. Third-generation sensors for night vision. *Opto-Electronics Review* 14(1):1-10. Available at http://www.springerlink.com/content/h126r11q13747524/fulltext.pdf. Accessed March 24, 2010.

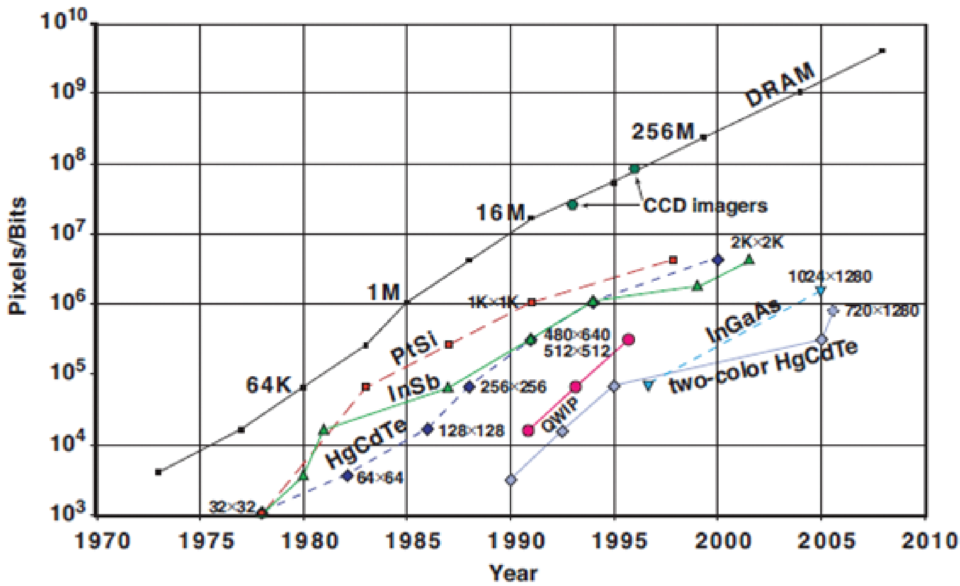


FIGURE 3-1
Detector array pixel count has paralleled the exponential growth of silicon DRAM bit capacity. SOURCE: Paul Norton. 2006. Third-generation sensors for night vision. *Infrared Photoelectronics, SPIE Proceedings Vol. 5957.*

and 4K × 4K (4-, 8-, and 16-megapixel) die. Scaling up to the 16-megapixel FPA provides larger sensor FOV and improved full-Earth coverage of the ballistic missile theater. Individual larger arrays are advantageous over tiling multiple smaller FPAs and result in 100 percent coverage without the additional effort required to account for the gaps between tiled arrays.

**FINDING 3-1**
The evolutionary trends are semiconductor detectors characterized by increased pixel pitch and count, higher readout speed, higher operating temperature (especially MWIR), lower power consumption, and decreased sensor thickness. The need for larger fields of regard is a significant driver for larger arrays. Even beyond the diffraction limit of the optical system, oversampling can lead to slightly enhanced resolution.

## Mosaic Tiling Technologies

For a number of years, visible-band CCD imagers have been built in three-side buttable and four-side buttable formats, allowing tiling of large focal planes from

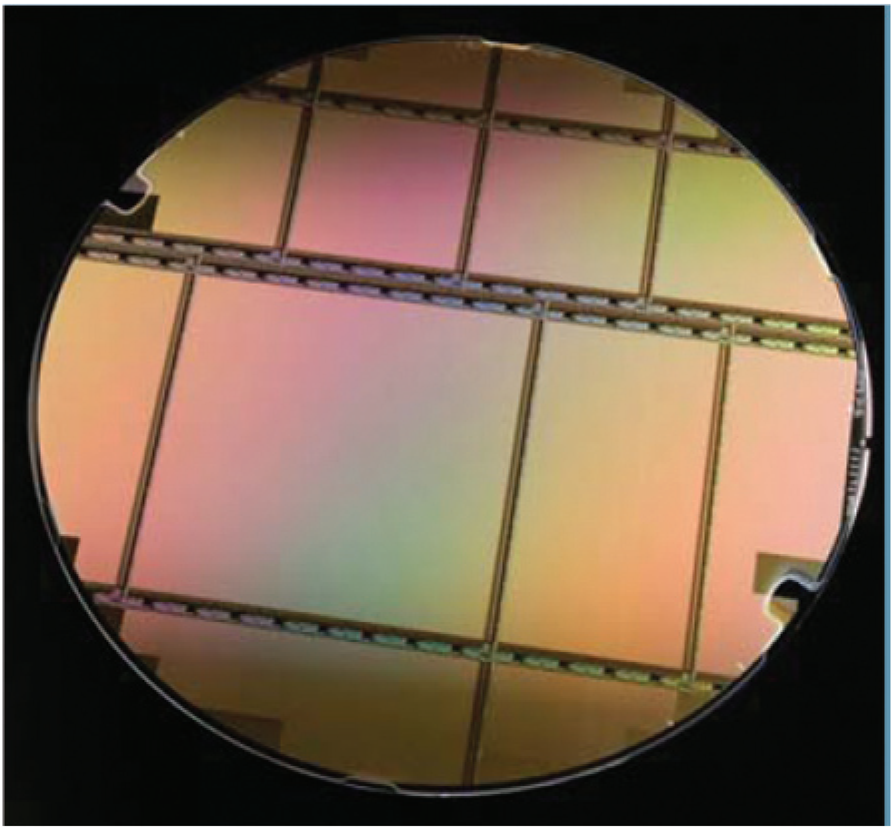

FIGURE 3-2
A single 8-inch ROIC wafer from 2007 Raytheon industry research and development. SOURCE: Angelo Scotty Gilmore, Stefan Baur, and James Bangs. 2008. High-definition infrared focal plane arrays enhance and simplify space surveillance sensors. *Raytheon Technology Today* 1:5-8. Available at http://www.raytheon.com/newsroom/rtnwcm/groups/public/documents/content/rtn08_tech_sensing_pdf2.pdf. Accessed March 26, 2010.

smaller, higher-yielding chips, with only modest gaps between the chips. For example, the 1.4-gigapixel orthogonal transfer array CCD imager used in PanSTARRS is composed of 60 chips, each of 22 megapixels. An array of this size could not be made monolithically since it exceeds the size of the largest silicon wafers used by the IC (integrated circuit) industry. Currently, and for the better part of the next decade, silicon wafer sizes are not expected to exceed the 300 mm diameter currently in use by leading-edge semiconductor facilities, although the silicon IC industry is actively exploring a transition to 450-mm-diameter wafers. Many imaging chips are made using process technologies being run on 200-mm-diameter wafers, for both cost and technological reasons.

In many cases, the size of a silicon chip is not limited by the wafer size, but it is limited by the ability to yield working chips, which falls off rapidly with increasing chip size. Achieving high yields on large chips can be expected to be most challenging for extremely dense and complex designs using the most advanced process technologies. Tiling large arrays from smaller chips addresses the practical and economic limits of making larger monolithic chips.

While tiled arrays are already in common use in high-end scientific and military applications, there are a number of areas in which improvements can be expected. Gaps between chips can be reduced from hundreds of micrometers to as little as a few tens of micrometers, especially for monolithic technologies. Greater use of four-side buttable designs is expected. Techniques will be developed to simplify interconnections to the tiles and to lower the cost of tiled arrays. Improved manufacturability and repairability are also active areas of research.

Application areas for large-format tiled arrays with minimal seam loss are likely to include wide-FOV telescope systems with large optics and long focal lengths for such diverse applications as astronomy, space surveillance, and persistent surveillance.

## Pixel Size Reduction

A general trend has been to reduce pixel sizes, and this trend is expected to continue. Several reasons exist for reducing pixel size, and the desirability of reducing pixel size is dependent on details of the application and the operating wavelength. In general, systems operating at shorter wavelengths (e.g., visible and ultraviolet) are more likely to benefit from small pixel sizes because of the smaller diffraction-limited spot size.

Diffraction-limited optics with low *F*-numbers (e.g., *F*/1) could benefit from pixels on the order of one wavelength across, as small as about 0.5 μm in the visible or about 10 μm in the LWIR. Oversampling the diffractive spot may provide some additional resolution for smaller pixels, but this saturates quickly as the pixel size is decreased.

Commercial CMOS imagers and CMOS chips have been demonstrated with pixels sized in the 1 to 2 μm range, with some examples less than 1 μm. On the silicon CMOS imager side, much of the interest in pixel size reduction has been driven by the desire to deliver a large number of megapixels to a consumer while keeping the silicon area used by the chip as small as possible to minimize cost for consumer applications, such as cell phone cameras. Interestingly, many of these cameras have low-performance optics, leading to much poorer resolution than might be expected based on the megapixel count for the imager. Visible-band small-pixel imagers are useful with small optics and can find application in unattended ground sensors as well as other systems requiring small surveillance cameras.

In the IR there remains a steady emphasis on improving uncooled microbolometers that will continue to mature with smaller detector sizes and larger formats. Current products utilize a 17 μm pitch and are available in high-definition formats (640 and 1280) primarily in the LWIR band. MWIR arrays have been fabricated, but they are limited by detector noise due to the lower MWIR photon flux.

In the near future, uncooled 10-12 μm detector pitch arrays will be available in high-definition format (1920 × 1080). This reduction in pitch will enable a reduction in optics size allowing increased range capability without an increase in weight for man-portable applications.

For wide-area persistent surveillance, programs such as ARGUS-IS,[2] as shown in Boxes 3-1 and 3-2, are already working on extremely large mosaics of visible FPAs to enable constant surveillance of large city areas.[3] With the smaller-pitch FPAs becoming available, these types of systems should be more affordable and smaller. The technology challenge is to maintain sensitivity and reduce the thermal time constant as the pitch size is decreased. Uncooled IR technologies also will be developed that exploit piezoelectric and other temperature response mechanisms, such as bimetallic microelectromechanical systems (MEMS) structures, diode forward voltage changes, or capacitance changes.

**FINDING 3-2**

The global proliferation of low-cost, commodity imagers, such as cell phone cameras and automobile thermal imagers, enables adversaries to develop sensing systems at relatively low cost, reducing the barrier to achieving limited operational capabilities. As an example, the rapid proliferation of low-cost "night vision technology" is eroding the overwhelming dominance of the United States in nighttime operations, even with the superior performance of advanced systems.

**FINDING 3-3**

The availability of very low cost imagers developed for large consumer markets is providing opportunities to develop new sensor systems and architectures, even though the component-level imagers may not have the capabilities typical of high-performance sensors developed specifically for military applications. Additionally, the technology and manufacturing base used to make these low-cost imagers will extend the manufacturing base that can be used for fabricating customized military parts.

---

[2]Brian Leininger, Jonathan Edwards, John Antoniades, David Chester, Dan Haas, Eric Liu, Mark Stevens, Charlie Gershfield, Mike Braun, James D. Targove, Steve Wein, Paul Brewer, Donald G. Madden, and Khurram Hassan Shafique. 2008. Autonomous Real-time Ground Ubiquitous Surveillance—Imaging System (ARGUS-IS) Defense Transformation and Net-Centric Systems. *Proceedings of SPIE* 6981:69810H.

[3]The data management challenges posed by a system such as ARGUS-IS are addressed in Chapter 4.

**RECOMMENDATION 3-1**
The intelligence community should pay careful attention to the new capabilities inherent in both the proliferation of commodity detector technologies and their integration into novel sensor systems. ARGUS-IS and Gnuradio are examples of how available, low-cost, mature commodity visible FPA technology (cell phone camera chips) and commercial off-the-shelf (COTS) communications circuitry, through sensor integration, have enabled new, advanced, high-performance imaging capabilities.

### Smarter Pixels and On-focal-plane Processing

A trend has been to take advantage of progress in electronics scaling to integrate as many functions as possible on a single chip. Visible-band CMOS active pixel sensors are particularly well developed examples of this trend. Usually the desire is to achieve cost reductions for high-volume applications by having a single-chip solution, but monolithic integration can also have other substantial benefits, such as power and noise reduction and enabling new interconnection-rich processing architectures that would not be feasible using off-chip inputs and outputs.

In the IR and high-performance visible imaging area, fully digital focal planes are just entering the market. In some cases their electro-optical (EO) performance exceeds that of traditional analog focal planes coupled to discrete electronics. High-performance IR scanners and large-area staring visible digital focal planes pixels are currently being demonstrated and show size, weight, and power (SWaP) advantages. On-ROIC digital logic enables future digital signal processing such as nonuniformity correction (NUC), image stabilization, and compression. These focal planes will enable much smaller systems (microsystems). This technology is poised to go into large-area cryogenic infrared sensors over the next few years.

Sending data off-chip requires substantial power and sending those data through communication links—for example, a radio-frequency (RF) link for an unattended ground sensor—can be even more energy intensive. There is increasing recognition of the value of trying to identify and transmit only small amounts of high-level and actionable information rather than large numbers of raw data bits. This is particularly important for systems with severe power or communication bandwidth constraints. This is leading to specialized imaging chips that have on-chip processing tuned to a specific application. Such customization can lead to dramatic improvements in system performance, but it has the drawback that longer design and fabrication cycles and large nonrecurring costs may be required to make application-specific chips.

The amount of processing that can be placed right at the pixel has generally been limited by the modest numbers of transistors that can be placed within a small pixel. Moore's law device scaling has steadily been improving that situation, and focal planes are now being developed with reasonably small pixels and significant

BOX 3-1
Case Study: DARPA ARGUS-IS

The DARPA ARGUS-IS unmanned aerial system is designed for persistent surveillance and incorporates a 1.8-gigapixel composite visible sensor system composed of 184 FPA pairs, for a total of 368 FPAs. ARGUS-IS is demonstrating on-board real-time processing of wide-area video imagery. Figure 3-1-1 depicts the components of the system. Figure 3-1-2 outlines the specific on-board processing solution, details of which are discussed in Chapter 4, and Figure 3-1-3 provides a sample output image.

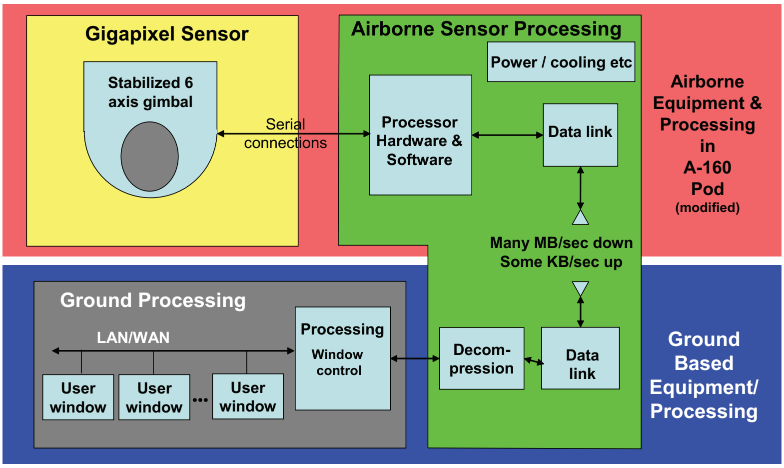


FIGURE 3-1-1
ARGUS-IS block diagram. SOURCE: Leininger, Brian, Jonathan Edwards, John Antoniades, David Chester, Dan Haas, Eric Liu, Mark Stevens, Charlie Gershfield, Mike Braun, James D. Targove, Steve Wein, Paul Brewer, Donald G. Madden, and Khurram Hassan Shafique. 2008. Autonomous Real-time Ground Ubiquitous Surveillance-Imaging System (ARGUS-IS). *Proceedings of the SPIE* 6981:69810H-1.

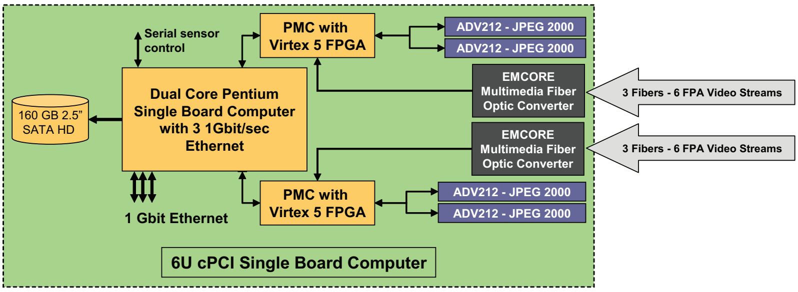


FIGURE 3-1-2
ARGUS-IS airborne processing module. SOURCE: Leininger, Brian, Jonathan Edwards, John Antoniades, David Chester, Dan Haas, Eric Liu, Mark Stevens, Charlie Gershfield, Mike Braun, James D. Targove, Steve Wein, Paul Brewer, Donald G. Madden, and Khurram Hassan Shafique. 2008. Autonomous Real-time Ground Ubiquitous Surveillance-Imaging System (ARGUS-IS). *Proceedings of the SPIE* 6981:69810H-1.

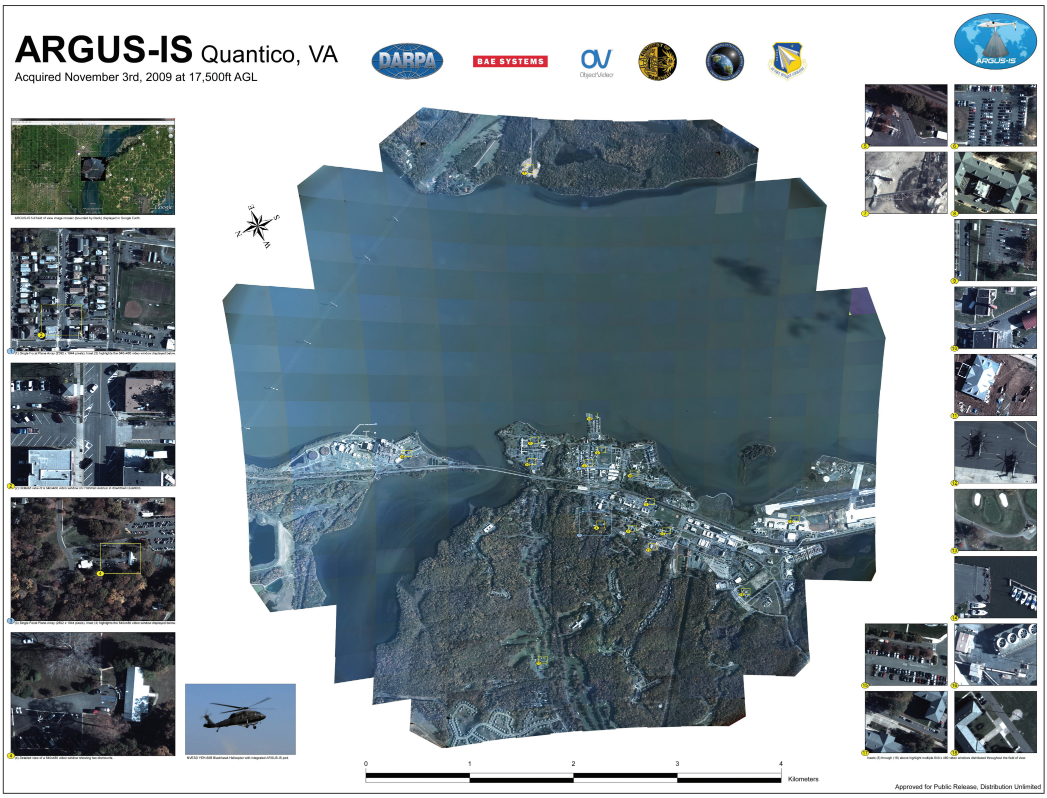


FIGURE 3-1-3
Sample of ARGUS-IS imagery. Mounted under a YEH-60B helicopter at 17,500 feet over Quantico, Va., Argus-IS images an area more than 4 km wide and provides multiple 640 × 480-pixel real-time video windows. SOURCE: Image courtesy of BAE Systems.

BOX 3-2
Impact of Commodity Components

One observation to highlight is the use of readily available commodity components in the ARGUS-IS system; the ARGUS-IS designers state:

> Though many other processing solutions exist, the rapid deployment schedule implied use of low-risk COTS [commercial off-the-shelf], or close to COTS, processing hardware. Hence, other processing elements were considered, but rejected for various reasons including: power consumption (multicore central processor units (CPUs) and graphics processor units (GPUs)), lack of configurability and development time (custom ASICs), or lack of proven integration tools or processing margin.

Multicore processors and GPUs have become COTS since 2007 and enjoy certain performance and power benefits. More importantly, the observation about rapid deployment schedules using commodity components has not been lost on other design teams (e.g., at the Iran University of Science and Technology,[a] Xidian University,[b] and Nanjing University of Science and Technology).[c]

Additional examples of commercial developments relevant to the communication architectures of remote imaging systems include the Gnuradio architecture,[d] with its signal processing architecture partitioned between a universal software radio peripheral with selectable daughtercards and a host processor with more or less arbitrary signal processing performance. A third example is the use of multiple-input multiple-output (MIMO)[e,f] in radio systems, where multiple antennas and significant signal processing are used to overcome localized radio-frequency challenges such as multipath.

[a]B. Zamanlooy, V.H. Vaghef, S. Mirzakuchaki, A.S. Bakhtiari, and R.E. Atani. 2007. A real time infrared imaging system based on DSP & FPGA. Pp. 16-23 in *Lecture Notes in Computer Science.* Berlin: Springer-Verlag.

[b]Huixin Zhou, Shangqian Liu, Rui Lai, Dabao Wang, and Yubao Cheng. 2005. Solution for the nonuniformity correction of infrared focal plane arrays. *Applied Optics* 44(15):2928-2932.

[c]J. Zhang, Y. Qian, B. Chang, S. Xing, and L. Sun. 2005. Signal processing of microbolometer infrared focal-plane arrays. Pp. 137-145 in *Infrared Components and Their Applications.* Bellingham, Wash: SPIE.

[d]Available at http://gnuradio.org/redmine/wiki/gnuradio.

[e]Gregory G. Raleigh and John M. Cioffi. 1998. Spatio-temporal coding for wireless communication. *IEEE Transactions on Communications* 46(3):357-366.

[f]G.J. Foschini. 1996. Layered spacetime architecture for wireless communication in a fading environment when using multiple antennas. *Bell Labs Syst Tech J* 1:4159.

in-pixel processing. As an example, digital FPAs with per-pixel analog-to-digital conversion, high-dynamic-range digital integration, and simple but powerful signal processing primitives have been demonstrated.

### 3-D Integration and Improved Hybridization Technology

Most IR and some visible technologies require hybridization of a detector array with a silicon CMOS readout IC. The detector pixels are generally connected to the per-pixel electronics through indium bump bonds. While bump-bond pitches of 14 μm or larger are relatively common, pitches of less than 8 μm offer significant challenges in terms of array yield, pixel operability, and cost. It can be expected that developments will continue to extend bump-bonding technology to smaller pixels, as well as to improve manufacturability and reduce cost at all pixel sizes.

One research area has been in the development of 3-D integration technologies providing alternatives to bump bonding. For example, oxide-to-oxide wafer bonding and silicon on insulator (SOI)-based 3-D integration have been used to demonstrate SWIR arrays integrated to CMOS with pitches down to 6 μm.

In the past, much of the processing done on a focal plane was confined to the two-dimensional real estate directly under a given pixel. Recently as many as three layers of CMOS have been stacked and vertically interconnected, offering the potential to increase the amount of processing available within a pixel footprint. 3-D integration technology has been applied to photon counting readouts for laser detection and ranging (LADAR) as well as to visible-band passive imagers. This stacked approach allows additional processing real estate in layers under the traditional sensor chip assembly.

A number of process approaches are being explored, including methods using through-wafer vias, wafer-bonded SOI electronics, and thin detector layers attached with epoxy. This technology will allow for higher-performance analog detector amplifiers and will enable on-chip and in-pixel processing of digital video and image data on large-area staring arrays.

### Devices Able to Perform at Higher Temperatures

Increasing the operating temperature is of particular concern for high-performance IR detectors, since the power needed for cooling increases significantly as the operating temperature is dropped. The power savings that result from reduced cooling requirements are particularly important for space systems, for sensors used by dismounted soldiers, and for power-constrained unattended ground sensors. While high operating temperature work is generally focused on the MWIR and LWIR, even silicon visible sensors used in applications requiring long integration times must be cooled to reduce dark current, and improvements in dark current at higher temperatures have enabled the elimination of thermoelectric coolers or a reduction in their power consumption.

High-temperature MWIR detectors will simplify future space systems by eliminating the cold FPA stage. Current space-based MWIR detectors must be cooled to temperatures in the 70 K range in order to reduce noise produced by the detectors and enable background-limited IR photodetection (BLIP). Raising the operating temperature to that of the optical bench will eliminate second-stage cryogens or mechanical coolers. Several competing technologies hold some promise in this area.

As the performance of MCT MWIR continues toward higher temperatures, new detector technologies, such as strained-layer superlattices (SLS), have some promise at high temperatures as well. The band structure of indium arsenide-gallium antimonide allows optimization of the carrier effective mass that theoretically enables detectors to have higher operating temperatures and longer cutoff wavelengths. Because the SLS detectors behave as direct bandgap devices, they avoid the quantum efficiency problems that have affected other III-V detector approaches such as quantum-well infrared photodetectors (QWIPs) and quantum-dot infrared photodetectors (QDIPs).

The literature on SLS detectors is quite active, and limiting behavior has yet to be established. The results to date are quite dependent on the epitaxial growth and processing-induced defects, much as the situation was for MCT a number of years ago. A recent review[4] provided the snapshot in Table 3-1. While there is much work to be done, it appears that SLS detectors will at least provide some competition for MCT FPAs because of the more mature III-V technology and the possibilities of bandgap engineering. One example of the bandgap engineering possibilities is the nBn (or n-type/n-barrier/n-type device [and related pMp]) structure that promises to reduce generation-recombination dark currents.[5] The pMp structure has shown some advantage for LWIR[6] and may be extensible to the MWIR. These developments are just being reported by research groups, with varying success. This work is likely to mature in the next 10 to 15 years and, if its early promise is fulfilled, may find its way first into military systems and then to lower-cost commercial applications.

[4]H.S. Kim, E. Plis, A. Khoshakhlagh, S. Myers, N. Gautam, Y.D. Sharma, L.R. Dawson, S. Krishna, S.J. Lee, and S.K. Noh. 2010. Performance improvement of InAs/GaSb strained layer superlattice detectors by reducing surface leakage currents with SU-8 passivation. *Applied Physics Letters* 96(3):033502.

[5]S. Maimon and G.W. Wicks. 2006. nBn detector, an infrared detector with reduced dark current and higher operating temperature. *Applied Physics Letters* 89(15):151109.

[6]B.-M. Nguyen, S. Bogdanov, S. Abdollahi Pour, and M. Razeghi. 2009. Minority electron unipolar photodetectors based on type II InAs/GaSb/AlSb superlattices for very long wavelength infrared detection. *Applied Physics Letters* 95(18):183502-1.

TABLE 3-1 Comparison of State-of-the-Art Type II Strained-layer Superlattice Photodiodes and nBn Photodetectors for MWIR Detection at Elevated Temperatures

| Parameter | Kim et al. (2008)[a] | Plis et al. (2007)[b] | Wei et al. (2006)[c] | Razeghi et al. (2010)[d] |
|---|---|---|---|---|
| Device | nBn | nBn | pin SLS | pin SLS |
| Cutoff wavelength, μm | 4.2 | 4.5 | 4.9 | 4.2 |
| Dark current density, A cm$^{-2}$ (77 K) | $1.0 \times 10^{-7}$ | $5.0 \times 10^{-7}$ | $4.0 \times 10^{-8}$ | $3.3 \times 10^{-9}$ |
| Dark current density, A cm$^{-2}$ (high T) | NA | 0.15 (240 K) | 0.2 (240 K) | 0.05 (240 K) |
| Responsivity, A/W (77 K) | 1.6 | 0.7 | 1.0 | NA |
| D* Jones (77 K) | $6.7 \times 10^{11}$ | $2.0 \times 10^{12}$ | $1.5 \times 10^{13}$ | $3 \times 10^{13}$ |
| D* Jones (high T) | NA | $2.0 \times 10^{9}$ (240 K) | $1.0 \times 10^{9}$ (300 K) | $1.2 \times 10^{10}$ (240 K) |

NOTE: NA = not available.

[a]H.S. Kim, E. Plis, J.B. Rodriguez, G.D. Bishop, Y.D. Sharma, L.R. Dawson, S. Krishna, J. Bundas, R. Cook, D. Burrows, R. Dennis, K. Patnaude, A. Reisinger, and M. Sundaram. 2008. Mid-IR focal plane array based on type-II InAs/GaSb strain layer superlattice detector with nBn design. *Applied Physics Letters* 92(18):183502.

[b]E. Plis, J.B. Rodriguez, H.S. Kim, G. Bishop, Y.D. Sharma, L.R. Dawson, and S. Krishna. 2007. Type II InAs/GaSb strain layer superlattice detectors with p-on-polarity. *Applied Physics Letters* 91(13):133512.

[c]Y. Wei, A. Hood, H. Yau, A. Gin, M. Razeghi, M.Z. Tidrow, and V. Nathan. 2005. Uncooled operation of type-II InAs/GaSb superlattice photodiodes in the midwavelength infrared range. *Applied Physics Letters* 86(23):233106.

[d]M. Razeghi, B.M. Nguyen, P.Y. Delaunay, S.A. Pour, E.K. Huang, P. Manukar, S. Bognadov, and G. Chen. 2010. High operating temperature MWIR photon detectors based on type II InAs/GaSb superlattice. *Proceedings of SPIE* 7608:76081Q.

## Multicolor Pixels

Multiband detectors provide independent sensing of different spectral bands within individual pixels.[7] The multiband feature provides the ability to optimize detection and identification functions of the sensor system. Multiband detectors enable increased mission robustness due to their ability to provide optimum imagery over a wide range of atmospheric and battlefield scenarios. In addition, when combined with advanced signal processing and fusion algorithms, multiband FPAs provide enhanced target detection and discrimination capability.

Multiband FPAs are currently available in small and standard television formats. These will be further developed and will be available in more spectral regions.

[7]Donald F. King, Jason S. Graham, Adam M. Kennedy, Richard N. Mullins, Jeffrey C. McQuitty, William A. Radford, Thomas J. Kostrzewa, Elizabeth A. Patten, Thomas F. Mc Ewan, James G. Vodicka, and John J. Wootan. 2008. 3rd-generation MW/LWIR sensor engine for advanced tactical systems. *Proceedings of SPIE* 6940:69402R.

The FPAs currently operate mainly in the MWIR and LWIR bands. These will be expanded to offer SWIR and possibly visible capabilities in conjunction with MWIR and LWIR. A wider spectrum choice for multiband operation will enable new and more robust systems that provide better target recognition and identification. Figure 3-3 illustrates the architecture and morphology of single-bump, two-color detectors.

Hyperspectral detectors, separating and sensing hundreds of bands at once, will have a significant effect on materials identification, including gases and volatile materials (ranging from pollutants to explosives), and provide the ultimate passive countermeasure to camouflage.[8] Hyperspectral sensors will be developed in all sizes, from handheld or wall-mounted field units to remote sensing instruments aboard aircraft or spacecraft.

While the MCT material system has been the dominant material used for dual-band MWIR and LWIR focal planes, work is going on with other materials and device types.

QWIPs generally have lower quantum efficiency (QE) and limited bandwidth, remain a niche U.S. technology, and are used more often in long-wave applications, especially for the international military market. The low QE arises as a result of the limited thickness of the multi-quantum-well absorbers and the requirement of a grating or other optical device, since the absorption requires an electric field directed across the quantum well, which is not available for normal incidence radiation. The Department of Defense (DOD) has not made significant use of this approach because of these limitations. For certain multicolor applications, narrow-bandwidth QWIPs may be appropriate. In addition, a corrugated QWIP (C-QWIP) approach has the potential to improve performance for both QE and bandwidth, if ever implemented (see additional discussion on nanophotonics in Chapter 4).

QDIPs are an alternate emerging technology that presents some promise at the research stage but remains unproved for commercial and military applications.[9] The many degrees of freedom offered by a quantum dots in a well (DWELL) detec-

[8]"*Multispectral* deals with several images at discrete and somewhat narrow bands. The "discrete and somewhat narrow" is what distinguishes multispectral in the visible from color photography. A multispectral sensor may have many bands covering the spectrum from the visible to the longwave infrared. Multispectral images do not produce the "spectrum" of an object. Landsat is an excellent example. *Hyperspectral* deals with imaging narrow spectral bands over a contiguous spectral range and produces the spectra of all pixels in the scene. So a sensor with only 20 bands can also be hyperspectral when it covers the range from 500 to 700 nm with 20 10 nm wide bands (while a sensor with 20 discrete bands covering the visible, near infrared, SWIR, MWIR, and LWIR would be considered multispectral). SOURCE: http://en.wikipedia.org/wiki/Hyperspectral_imaging. Last accessed on June 17, 2010.

[9]A.V. Barve, S.J. Lee, S.K. Noh, and S.K. Krishna. In press. Review of current progress in quantum dot infrared photodetectors. *Laser & Photon Rev* DOI 10.1002/lpor.200900031.

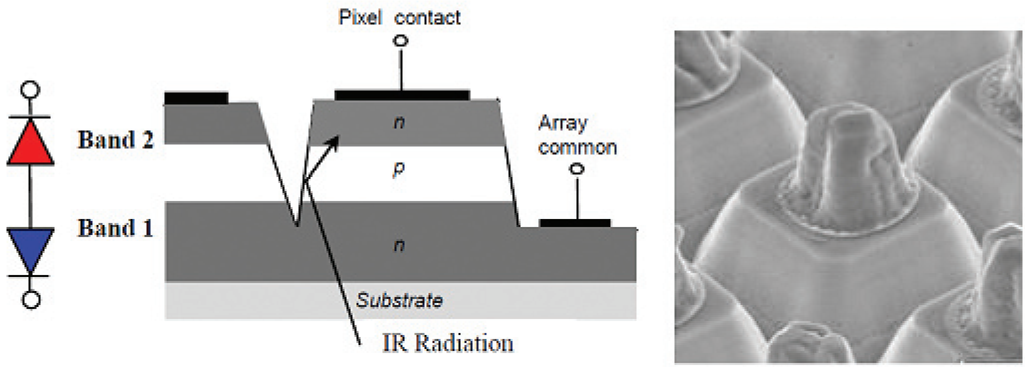


FIGURE 3-3
A cross section diagram (*left*) and scanning electron microscope image (*right*) illustrate the architecture and morphology of single-bump two-color detectors. The lower junction (Band 1) responds to shorter-wavelength radiation, while the upper junction (Band 2) responds to longer wavelengths. SOURCE: King, Donald F., Jason S. Graham, Adam M. Kennedy, Richard N. Mullins, Jeffrey C. McQuitty, William A. Radford, Thomas J. Kostrzewa, Elizabeth A. Patten, Thomas F. Mc Ewan, James G. Vodicka, and John J. Wootan. 2008. 3rd-generation MW/LWIR sensor engine for advanced tactical systems. *Proceedings of SPIE* 6940:69402R.

tor allows multicolor operation in a single device by accessing intradot transitions (VLWIR), QD-QW transitions (LWIR), and QD-continuum transitions (MWIR). These can be accessed with different bias voltage regimes. A significant advantage of QDIPs is that, in contrast to QWIPs, they are sensitive to normal incidence radiation, but they still suffer from generally low quantum efficiency as a result of the small absorption volume. This can be mitigated with plasmonic structures, as discussed in Chapter 4. Table 3-2 provides a snapshot of the current state of development of LWIR detectors across all of the various material systems.

Visible imaging sensors often use integrated color filters to achieve red, green, blue (RGB) color model capability, though high-performance systems may use dichroic filters to make optimum use of photons. There is room for innovation in the visible sensor area to make multispectral and hyperspectral sensors, especially ones that use photons efficiently.

**FINDING 3-4**
Existing, mature mercury cadmium telluride, indium antimonide, indium gallium arsenide, silicon charge-coupled devices, silicon complementary metal oxide semiconductors, and avalanche photodiode focal plane technologies provide sensors with excellent performance and set a very high barrier to entry for any emerging technology. For some performance parameters, such

TABLE 3-2 Comparison of LWIR Existing State-of-the-art Device Systems for LWIR Detectors

| | Bolometer | HgCdTe | Type II SLs | QWIP | QDIP/QDWIP |
|---|---|---|---|---|---|
| Maturity | TRL 9 | TRL 9 | TRL 2-3 | TRL 8 | TRL 1-2 |
| Status | Material of choice for application requiring medium to low performance | Material of choice for application requiring high performance | Research and development | Commercial | Research and Development |
| Military System Examples | Weapon sight, night vision goggles, missile seekers, small UAV sensors, unattended ground sensors | Missile intercept, tactical ground and airborne imaging, hyperspectral, missile seeker, missile tracking, space-based sensing | Being developed in universities and evaluated in industry research environment | Being evaluated for some military applications | Very early stages of development at universities |
| Limitations | Low sensitivity and long time constraints | Performance susceptible to manufacturing variations. Difficult to extend to >14 micron cut-off | Requires a significant, >$100 million, investment and fundamental material breakthrough to mature | Narrow bandwith and low sensitivity | Narrow bandwith and low sensitivity |
| Advantages | Low cost and requires no active cooling. Leverages standard Si manufacturing equipment | Near theoretical performance. Will remain material of choice for at least the next 10-15 years | Theoretically better then HgCdTe at >14 micron cut-off. Leverages commercial III-V fabrication techniques | Low-cost applications. Leverages commercial manufacturing processes. Very uniform material | Not sufficient data to characterize material advantages |

NOTE: TRL = technology readiness level.

as detectivity, mature imager technologies already are operating very close to fundamental limits. However, there is still considerable opportunity to improve other parameters such as operating temperature, power dissipation, manufacturability, and cost.

## Improved SWIR Arrays

Access to the SWIR band provides the tactical advantage of being able to see in a band that has more night illumination than the visible and near-IR bands and that sees signals from all current laser designators, pointers, and range finders. The SWIR-equipped soldier can see adversaries equipped with night vision goggles without being detected, since the SWIR devices are totally passive, in contrast to image intensifiers that radiate as well as detect. A low-cost, low-power, high-resolution SWIR technology could replace the U.S. inventory of night vision goggles and, until fielded by the competition, would provide a significant advantage under low-light conditions. Despite potential advantages, SWIR technology is still not in wide use, compared to image-intensifier-based night vision goggles. SWIR FPAs will be available with increased sensitivity, lower power, smaller pixels, and larger formats.

Currently, the 320 × 240 format is commonly available, with 640 × 480 entering the market. High-definition formats are being developed and will be available for fielding in the near future. Improvements in readout application-specific integrated circuits and detector materials have resulted in SWIR FPAs having noise levels of a few electrons. This will result in high-definition formats capable of matching light levels of image intensifiers but with extended spectral response for night glow applications and a digital interface supporting advanced processing, such as multiband fusion. Technologies that are capable of counting individual photons are currently in the early stages of development and can be expected to yield usable focal planes within 10 years. Low-power, very small cameras will transmit digital SWIR images for weeks, enabled by a combination of lower-power electronics, room-temperature detector operation, and improved batteries. The technology will also be available in 50-megapixel formats for other applications, such as airborne surveillance.

The principal game changers in SWIR will be technologies that can significantly reduce the cost of FPA fabrication, currently a material or process yield issue, facilitate finer-pitch FPAs (less than 10 µm), or exhibit dark current densities that best InGaAs at any cutoff wavelength or operating temperature.

**FINDING 3-5**
Short-wave infrared (SWIR), due to an atmospheric phenomenon called night glow, is emerging as a next-generation tactical imaging technology because of its covertness and the similarity between SWIR and visible imagery. As it matures, SWIR will provide an alternative to intensified visible imaging (night vision goggles).

### Photon Counting Technologies and Lower Readout Noise

Three different technological paths will result in lower readout noise levels or even single-photon sensitivity:

1. For visible imagers, improvements in analog readout technology will continue, increasing the range of devices capable of sub-single-electron noise levels.
2. Innovative approaches to add modest amounts (5-100×) of linear gain to the photon detection process (such as linear-mode avalanche gain) will be pursued, to reduce the input-referred noise from the readout.

3. Significant numbers of new focal planes capable of direct photon-to-digital conversion will appear. For example, Geiger-mode detector arrays that have been developed for UV, visible, and SWIR, and are already used in active imaging systems such as direct detection LADAR, will be further developed and applied to passive photon counting imaging applications in increasingly large array sizes.

Photon counters have a number of important applications, including some low-light imaging applications, hyperspectral sensors, high-speed imaging, 3-D LADAR, and dual-mode active-passive pixels.

**FINDING 3-6**

Rapid progress is being made in the development of closely related single-photon and photon counting detectors and arrays. Single-photon detection and photon counting imagers are key enablers for a wide range of new secure communications, passive sensors, 3-D LADAR, and active optical sensors. Specifically, quantum cryptography relies on the distribution of entangled, single-photon qubits (keys) between the transmitter and receiver; this is inherently a single-photon process. In most cases, these applications involve physical processes in which only a small number of photons are available for detection. These detectors require high quantum efficiencies, low dark count rates, fast recovery times, and capabilities for photon number resolving.

**RECOMMENDATION 3-2**

The intelligence community should carefully track developments related to single-photon and photon counting detectors across the full spectrum from UV to VLWIR. Table 3-3 lists trigger events that would cause a significant shift in capability and should be carefully monitored by the intelligence community.

TABLE 3-3 Trigger Points of Technical Progress and Their Implications

| Single Photon | 2010 (SOA) | 2015 (TP) | 2020 (TP) | 2025 (TP) |
|---|---|---|---|---|
| Q efficiency | 90% | >90% | >90% | >90% |
| Speed | 10 GHz | 100 GHz | THz | THz |
| Wavelength | Visible | 1.55 μm | 1.55 μm | 1.55 μm |
| Operating temperature | 4 K | 77 K | 300 K | 300 K |
| Application | | QKD | QKD/quantum computer | Quantum computer |

NOTE: QKD = quantum key distribution; SOA = state of the art; TP = trigger point, which indicates a capability that should stimulate the intelligence community to do significant collection and/or analysis.

### Curved Focal Surfaces

Over the last decade, the ability to make cylindrically or spherically curved silicon imagers has been demonstrated, and recently a focal plane comprised of spherically curved silicon CCD imagers has been developed for a ground-based telescope. It is likely that these technologies will be refined and extended to other wavelengths (IR) and other detector materials systems. Curved focal planes can simplify the design of wide-FOV optics and allow lighter-weight solutions for high-performance imaging systems on size- and weight-constrained platforms, such as small unmanned air systems. The utility of curved focal planes becomes evident when one realizes that the human retina is curved, because this dramatically simplifies the design and complexity of the lens.

### Lower Power

Moore's law scaling of CMOS technology has led to steady reductions in the power consumed, especially for digital logic operations. Clever circuit design and architectural approaches also lead to dramatic power reductions. It is expected that further reductions in power can be achieved over the coming decade. Since many military systems operate with severe power and cooling constraints, these power reductions will result in the increased feasibility of many systems. In the case of cooled IR systems, power dissipated on the ROIC must be removed through the cooler, incurring another large power penalty, so lower-power ROIC designs are particularly desirable.

### Radiation Hardening

Space systems are of critical importance for defense, and they drive the requirements for radiation-hard imagers.[10] The hardness level required depends on the orbit and lifetime. Although beyond the scope of what can be discussed here, shortfalls and opportunities for improvement do exist for certain imaging technologies in certain environments, and important improvements can be expected over the coming decade. A general discussion of radiation hardness is provided in Appendix C.

[10]John E Hubbs et al. 2010. Radiometric characterization of long wavelength type II strained layer superlattice infrared focal plane array developed by Naval Research Laboratory and Teledyne Imaging Systems. *SENSIAC MSS Symposium*.

## Cost Reduction

Visible sensors, such as interline-transfer CCDs and CMOS active-pixel sensors, have seen dramatic cost reductions as large commercial markets, such as consumer cameras and cell phone cameras, have driven large production volumes. These parts are high performance in the sense that they represent state-of-the-art technology and are highly optimized for their intended markets. These parts may also be very low performance compared to parts optimized for a specific DOD application. While these low-cost, consumer-driven parts are not expected to replace all high-performance custom parts, they are opening up new approaches to DOD sensors because of their low cost and high performance-price ratio. One example is the recent use of large numbers of multi-megapixel cell phone CMOS imagers in wide-FOV video surveillance applications and airborne persistent surveillance systems (see above and Chapter 4 for additional information on this system). The proliferation of low-cost, visible sensors offers opportunities for DOD if it is quick to exploit this technology; however, such technologies are proliferated globally and are also accessible to current and potential adversaries.[11]

As discussed above, in the IR, alternatives to MCT such as InAs-GaSb SLS detectors may provide both a large decrease in cost and an increase in key performance capabilities. This detector technology will exploit the established III-V technology and manufacturing base and may also avoid some of the inherent producibility challenges associated with II-VI IR detectors. These materials can be grown by more traditional III-V molecular beam epitaxy (MBE) wafer vendors that can provide these wafers to the IR industry very much as bulk InSb wafers are supplied today. If successful, these material systems would enable an evolutionary cost reduction in sensors due to simpler technology requirements once it has been established. SLS devices are predicted theoretically[12] to have lower dark currents than HgCdTe; SLS is limited by Shockley-Read generation mechanisms that produce current via point defects in the as-grown material.[13,14,15] Future advancements in SLS technology will require a significant effort and investments to reduce the defect densities.

[11]There is a fuller discussion of these trends in Chapter 4. The ready availability of capable, off-the-shelf imagers will have a dramatic impact on the availability of future imaging systems for friends and foes alike.

[12]M.E. Flatte and C.H. Grein. 2009. Theory and modeling of type-II strained-layer superlattice detectors. *Proc. SPIE* 7222:72220Q.

[13]J. Pellegrino and R. DeWames. 2009. Minority carrier lifetime characteristics in type II InAs/GaSb LWIR superlattice $n^+\pi p^+$ photodiodes. *Proc. SPIE* 7298:72981U.

[14]D.R. Rhiger, A. Gerrish, and C.J. Hill. 2008. Estimation of carrier lifetimes from I-V curve fitting for InAs/GaSb and HgCdTe LWIR diodes. *Proc. MSS Parallel Meeting.*

[15]M.E. Flatte and C.H. Grein. 2009. Theory and modeling of type-II strained-layer superlattice detectors. *Proc. SPIE* 7222:72220Q.

Both cost reduction and the multiple open-source supplier structure of SLS detectors could have significant impacts in the widespread proliferation of large-format IR FPA technology.

**FINDING 3-7**
There is significant opportunity to customize image sensor architectures for specific applications that can lead to dramatic improvements in system-level performance, including size, weight, and power. Advanced architectural design, including integration of sensing and processing (in-pixel and on-chip), can have greater system-level impact than making small gains in driving detector performance incrementally closer to fundamental detectivity limits.

**RECOMMENDATION 3-3**
The intelligence community should evaluate and track system capabilities rather than focusing solely on component technical achievements. These include technologies that enable in-pixel and on-chip processing, lower-power operation, and higher operating temperatures, as well as technologies that improve manufacturability.

## Improved Cooler Technologies

The availability of improved cooler technology can have a huge impact on the system-level performance of detector systems and may make or break whether a given focal plane technology can be considered for a given application. At least four dimensions for improvements in coolers can be identified—each dimension results in different system impacts—including the following:

1. *Development of higher-coefficient-of-performance coolers:* Visible and SWIR systems using thermoelectric coolers, as well as IR systems using refrigeration cycles, would benefit from more efficient coolers, because cooling power often dominates total sensor power consumption. This is particularly important for power-constrained systems, such as unattended ground sensors, and systems that are used only intermittently but must continuously be in a ready-to-operate state.
2. *Reliability improvements and space qualification:* A number of space systems would benefit from the availability of highly reliable and fully qualified cryogenic coolers.
3. *Lower cost:* For systems that must be produced in large quantities—for example, for individual soldiers—and especially for "disposable" systems, low cost is critical.
4. *More compact form factor:* The availability of compact and lightweight

coolers is a prerequisite for a number of applications. An example would include deployment of an imaging system on a micro-air vehicle.

Temperature control and stability are crucial parameters in determining the ultimate resolution, signal to noise, and other performance attributes of various IR and visible EO detectors. For this reason, detectors are often categorized as being either cooled or uncooled. Cooled detectors refer to platforms that require cryogenic temperatures in order to operate. Typical temperatures of operation for cooled sensors range from <10 to 150 K or slightly higher. Uncooled detectors, despite their title, typically incorporate some degree of temperature control near or slightly below room temperature (~250-300 K) to minimize noise, optimize resolution, and maintain stable operating conditions. The two technologies currently available for addressing the cooling requirements of IR and visible detectors are closed-cycle refrigerators and thermoelectric coolers. Closed-cycle refrigerators can achieve the cryogenic temperatures required for cooled IR sensors, while thermoelectric coolers are generally the preferred approach to temperature control for uncooled visible and IR sensors.

Although uncooled sensors offer significant advantages in terms of cost, lifetime, and SWaP, cooled sensors offer significantly enhanced range, resolution, and sensitivity as a result of the reduced dark current and, therefore, lower-noise operation. For this reason, cooled IR sensors have been the technology of choice for many military operations where performance, not cost, is the prime driver. Conversely, for many commercial applications, uncooled sensors are the preferred technology because of their lower cost and higher reliability. In this section, the current status of various cooling technologies is reviewed with respect to their present and near-term applicability to IR and visible detectors.

Thermoelectric coolers and closed-cycle or mechanical refrigerators both involve the use of a working "fluid" to transfer heat from a thermal source to a thermal sink. The major difference between the thermoelectric and mechanical cryocoolers is the nature of the working fluid. A thermoelectric cooler is a solid-state device that uses charge carriers (electrons or holes) as a working fluid, whereas mechanical cryocoolers use a gas such as helium as the working fluid. While each cooling technology has its advantages and disadvantages, limitations in these cooling technologies translate directly to performance limitations in the IR and visible detectors to which they are applied. Conversely, any improvements and breakthroughs achieved in cooling technologies could provide enhancements to detector applications in SWaP metrics and in overall detector performance.

*Cryocoolers*

As mentioned previously, mechanical cryocoolers or refrigerators represent the only currently available technology that can reach the cryogenic temperatures required for many IR and visible sensors. As shown in Figure 3-4, cooling of IR sensors represents one of a number of commercial and research applications for cryocoolers. The major commercial applications for mid- to large-scale cooling include cryopumps for semiconductor fabrication facilities, magnetic resonance imaging (MRI) magnet cooling, and gas separation and liquefaction. For low-power applications, IR sensors represent the largest single application for cryocoolers.

The most common types of cryocoolers can be classified as either recuperative or regenerative. In recuperative systems, gas flows in a single direction. The gas is compressed at ambient fixed temperature and pressure and allowed to expand through an orifice to the desired cryogenic fixed temperature and pressure. The Joule-Thompson and Brayton cycle refrigerators are examples of recuperative systems. In a regenerative system, the gas flow oscillates back and forth between hot and cold regions driven by a piston, diaphragm, or compressor, with the gas being compressed at the hot end and expanded on the cold end. Stirling, Gifford-McMahon, and Pulse Tube cryocoolers are the most common types of regenerative cryocooler systems.

An important parameter in measuring the performance of a cryocooler is the coefficient of performance (COP). COP is a measure of efficiency and is defined as the ratio of cooling power achieved at a particular temperature to total electrical input power to the cryocooler. Often, the COP is given as a fraction of the ideal

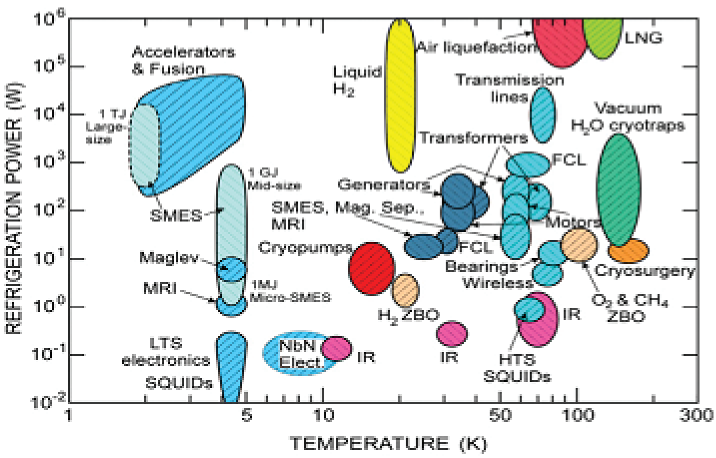


FIGURE 3-4
Cooling of IR sensors represents one of several commercial and research applications for cryocoolers. NOTE: FCL = freon coolant line; HTS = high-temperature superconductivity; LNG = liquid natural gas; LTS = low-temperature superconductivity; MRI = magnetic resonance imaging; SMES = superconducting magnetic energy storage; SQUID = superconducting quantum interference device; ZBO = zero boil-off. SOURCE: Radebaugh, Ray. 2009. Cryocoolers: the state of the art and recent developments. *Journal of Physics: Condensed Matter* 21:164219.

or Carnot efficiency, which is defined as the ratio of $T_c/(T_c - T_h)$, where $T_c$ and $T_h$ refer to cold and hot temperatures, respectively, of operation. In general, recuperative systems have advantages in terms of reduced noise and vibration, whereas regenerative systems tend to obtain higher efficiencies and greater reliability at the temperatures of interest for many IR detector applications. The relative performances of the different technologies as a fraction of the limiting Carnot efficiency are shown in Figure 3-5.

Significant developments in cryocooler technologies during the past several decades have facilitated their increased use in commercial markets. One major area of improvement has been reliability. Because impurities in the working gases can freeze and either clog or damage the various internal components of a cryocooler responsible for gas flow, maintaining a high-purity gas over the lifetime of the cryocooler is a key challenge. The development of novel adsorber materials and designs, as well as improved sealing techniques for maintaining the high pressures required

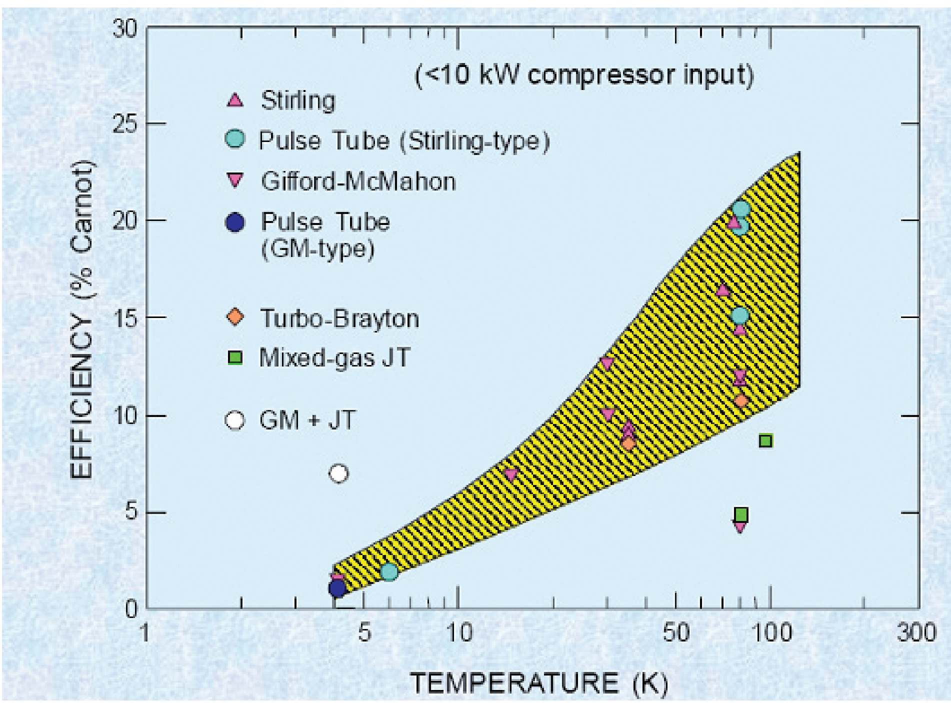


FIGURE 3-5
Relative performances of the different technologies as a fraction of the limiting Carnot efficiency. SOURCE: Radebaugh, Ray. 2009. Cryocoolers: the state of the art and recent developments. *Journal of Physics: Condensed Matter* 21:164219.

for cryocooler operation, has significantly contributed to the improved lifetimes now achievable in many cryocooler systems (5,000-10,000 hours). Improvements in heat exchanger and recuperator designs and materials have yielded absolute COP values for cryocoolers in the range of 0.01 to 0.1 for IR detector applications. Despite these advances, the cryocooler remains a major failure point for cooled IR detectors, as well as more than doubling the weight and power requirements of the integrated IR sensor system.

While the needs of commercial and military cryocooler applications will continue to drive improvements in cost, reliability, and efficiency, it is likely that these improvements will be incremental over the next 10-15 years. Areas for improvement include the ability to operate regenerative systems at higher frequency and the ability to operate using gases at higher pressures.[16,17] Operation at higher frequency enables the use of smaller, lighter-weight compressors that are often the dominant volume in regenerative cryocoolers. Similarly, increasing the gas pressure can enable larger COP by increasing the effective thermal capacitance per cycle of operation. Designs and concepts currently being considered for operation at higher frequency and pressure, as well as ongoing material and process developments, will ultimately lead to improved seals and structural performance for cryocooler components and enclosures.[18] There will likely continue to be a trade-off between cost and performance as commercial and military demands for improvements in SWaP drive the technology to more demanding operating conditions where maintaining lifetimes in excess of 10,000 hours will continue to be a challenge.

### *Thermoelectric Coolers*

Thermoelectric materials can be used to convert thermal energy to electricity or to use electricity to pump heat. In a generator configuration, thermoelectric devices exploit the Seebeck effect—the voltage created between two dissimilar conductors in the presence of a temperature difference. Thermoelectric coolers work by exploiting the Peltier effect, which refers to the creation of heat flux at the junction of two dissimilar conductors in the presence of current flow. In either configuration, the optimal performance of a given set of thermoelectric materials in a device is determined by the dimensionless figure of merit $ZT = S^2\sigma/\kappa$, where $S$ = Seebeck coefficient (thermopower), $\sigma$ = electrical conductivity, $\kappa$ = thermal conductivity, and $T$ is the average temperature $(T_1 + T_2)/2$. Due to performance

[16] S. Vanapalli, M. Lewis, Z. Gan, and R. Radebaugh. 2007. 120 Hz pulse tube cryocooler for fast cooldown at 50 K. *Appl. Phys. Lett.* 90:072504.

[17] S.L. Zhu G.Y. Yu, W. Dai, E.C. Luo, Z.H. Wu, X.D. Zhang. 2009. Characterization of a 300 Hz thermoacoustically-driven pulse tube cooler. *Cryogenics* 49:51.

[18] Ray Radebaugh. 2009. Cryocoolers: the state of the art and recent developments. *Journal of Physics: Condensed Matter* 21(16):164219-164228.

limitations, thermoelectric generators and coolers have found only a few niche commercial applications. As generators, these applications include radioisotope thermoelectric generators (RTGs) for space applications and waste heat conversion to power small electric devices such as fans, lights, or battery chargers. Commercial applications of thermoelectric coolers include (1) climate-controlled seating for automobiles; (2) temperature control and stability for bolometric and ferroelectric detectors, laser diodes, and ink jet printers; (3) dark-current reduction in mid-wave IR detectors; and (4) noise reduction in CCD arrays.

Thermoelectric coolers are essentially solid-state heat pumps where the flow of thermal energy is determined by the polarity of the applied current. The coefficient of performance for a thermoelectric cooler is a function of $ZT$ as well as the overall temperature difference between the hot side and the cold side of the cooler. At maximum temperature difference, $\Delta T$, the COP of a thermoelectric generator goes to zero (i.e., no heat can be removed). Conversely, at zero $\Delta T$, a thermoelectric cooler achieves maximum heat pumping capacity. The typical load profile for a thermoelectric cooler is shown in Figure 3-6. Therefore, for most thermal management applications, including temperature stabilization for IR and visible

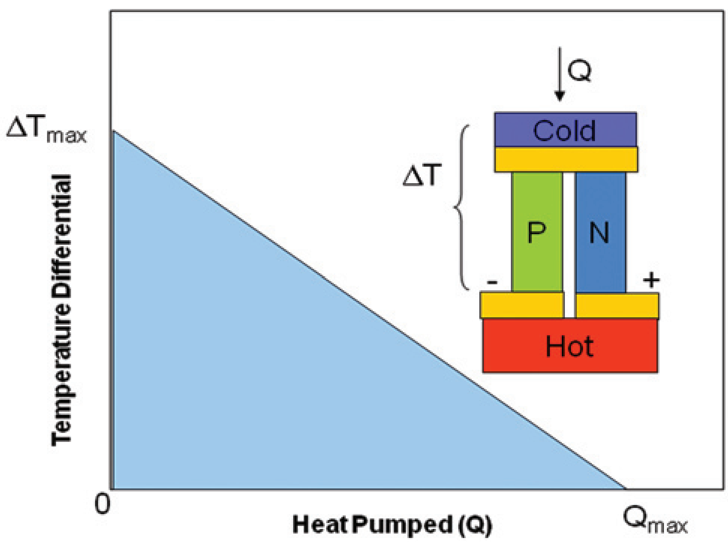


FIGURE 3-6
Typical load profile for a thermoelectric cooler. SOURCE: Rama Venkatasubramanian and colleagues, Research Triangle Institute, Research Triangle Park, N.C.

"uncooled" detectors, thermoelectric coolers typically operate at the minimal $\Delta T$ that provides acceptable detector performance.

Historically, the maximum achievable material $ZT$ stood at ~1 and saw little improvement despite many decades of research starting in the 1960s through the early 2000s. However, in the early 1990s, DOD began investing significant research-and-development funding to explore the potential for achieving higher $ZT$ values for both cooling and power generation applications. These investments paid off with the reports of breakthrough $ZT$ values in low-dimensional thin-film thermoelectric materials. The enhancement in $ZT$ achieved via nanostructuring in thin-film materials has prompted similar investigations in bulk materials. Recent reports have indicated that bulk thermoelectric performance in the $ZT$ range of about 1.5 near room temperature and approaching 2.0 at higher temperatures is achievable.[19]

The bulk material discoveries have prompted the formation of several commercial new start companies—for example, GMZ and ZT Plus—and it is likely that these activities will transition directly to improved thermoelectric devices within the next 5-10 years. The potential impact of the thin-film materials is less certain since reproducibility, scalability, and fabrication costs remain significant challenges. For this reason, they are discussed in detail in Chapter 4. This section focuses on what is currently achievable with existing bulk commercial thermoelectric materials and what developments are likely to occur over the next 10-15 years.[20]

Current commercially available thermoelectric coolers are based on alloys of bismuth telluride and antimony telluride materials. These materials exhibit $ZT$ values close to 1 unity, but in a device configuration achieve values closer to $ZT$ = 0.7. In a single-stage device, this equates to a maximum $\Delta T$ of about 70 K as measured from an ambient temperature of ~ 293 K. An approximation of maximum achievable $\Delta T$ can be obtained using the relationship $\Delta T_{max} \sim 1/2ZT_c^2$. However, as previously mentioned, at maximum $\Delta T$, a thermoelectric cooler cannot dissipate any heat. Therefore, in order to achieve $\Delta T$ of 70 K or more and still have some cooling capacity, thermoelectric coolers are stacked in "stages." The idea is to have each stage operating at less than $\Delta T_{max}$, with its "hot" side starting at the "cold"

[19]Poudel, Bed, Qing Hao, Yi Ma, Yucheng Lan, Austin Minnich, Bo Yu, Xiao Yan, Dezhi Wang, Andrew Muto, Daryoosh Vashaee, Xiaoyuan Chen, Junming Liu, Mildred S. Dresselhaus, Gang Chen, and Zhifeng Ren. 2008. High-thermoelectric performance of nanostructured bismuth antimony telluride bulk alloys. *Zhifeng Science* 320:634.

[20]A good literature summary of the current status of TE technology is found in Lon E. Bell. 2008. Cooling, heating, generating power, and recovering waste heat with thermoelectric systems. *Science* 321:1457-1461. Available at http://www.sciencemag.org/cgi/reprint/321/5895/1457.pdf?maxtoshow=&HITS=10&hits=10&RESULTFORMAT=&author1=bell&titleabstract=an+information-maximization+approach+to+blind+separation+and+blind+deconvolution&searchid=1&FIRSTINDEX=170&fdate=//&tdate=//&resourcetype=HWCIT. Accessed March 24, 2010.

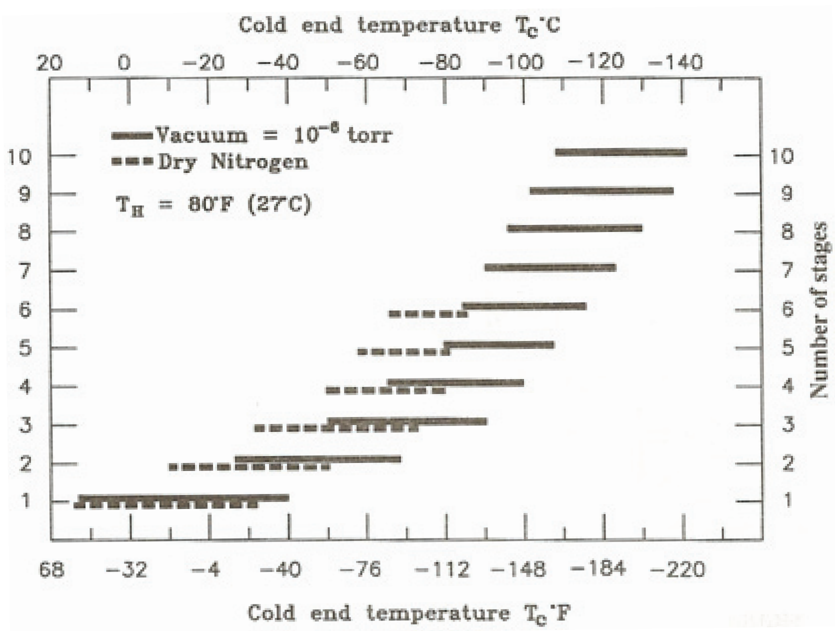


FIGURE 3-7
Theoretical range for maximum $\Delta T$ achievable using existing commercial materials as a function of the number of stages. SOURCE: Rowe, D.M., ed. 1995. *CRC Handbook of Thermoelectrics.* Boca Raton, Fla.: CRC Press, Inc.

side of the prior stage. The theoretical range for $\Delta T_{max}$ achievable using existing commercial materials is plotted in Figure 3-7 as a function of the number of stages. Although the figure shows values for up to a 10-stage device, commercially available coolers typically do not go beyond 6 stages.

If recently reported $ZT$ advances in bulk bismuth telluride can be transitioned to commercial products, it is reasonable to assume that device-level $ZT$s will see some modest improvements, perhaps as much as a 50 percent improvement in $\Delta T_{max}$ for a single-stage device. In terms of efficiency, existing thermoelectric coolers can achieve COPs of ~1.0 (i.e., 1 W of cooling for 1 W of electrical input) at $\Delta T$ of ~30 K. If the bulk materials available for commercial coolers were to achieve $ZT$ values closer to 1.5, then the achievable COPs would be anticipated to approach 1.5 (1.5 W of cooling for 1 W of electrical input) at similar 30 K temperature differentials. There is some evidence that modest improvements in bulk materials are transitioning to commercial thermoelectric (TE) products.[21] The reduction in material usage is attributable to anticipated improvements in material properties.

There are a number of commercial factors driving improvements in TE materials. In fact, the efforts to commercialize new, advanced bulk materials, coupled with industry efforts to improve cooler design and minimize parasitic losses, make it quite likely that these projected improvements will be realized within a 10-15 year time frame. For example, based on available data, the current market for TE coolers is relatively small, but it appears to be growing at an average rate of between 10 and 40 percent per year across a number of applications. One recent estimate puts the total TE cooler market at approximately $200,000,000.[22] These estimates include TE coolers for a number of applications including climate control, EO cooling, personal coolers, and biomedical refrigeration. Figure 3-8 shows an average growth of roughly 40 percent per year between 2001 and 2008. In contrast, a *Wall Street Journal* 2004 article reported that Igloo products estimated the personal cooler (small coolers for refrigerating drinks and food) market at $50 million and predicted a 10 percent annual growth. Given these estimates it is reasonable to assume that a significant portion of the commercial drivers for improved TE cooler technology are coming from nondetector applications.

In summary, both cryocoolers and thermoelectric coolers are poised to achieve modest improvements in performance over the next 10-15 years in terms of efficiency. In addition, cryocooler costs and reliability at today's performance levels are likely to improve due to commercial market drivers associated with their use in the semiconductor (cryopumps), medical (MRI, cryosurgery), and gas liquefaction industries. Reliability at today's SWaP requirements will likely increase to exceed 10,000 hours, while costs will likely be driven lower as these markets continue to grow. For niche applications, such as high-end detectors and large staring arrays, advances in SWaP will be critical. The trend toward larger-footprint arrays is particularly challenging because it requires greater thermal stability and minimal vibration across a larger area. Although concepts for operating at higher frequency and higher pressure are likely to result in improved SWaP for cryocoolers, these advances will probably be obtained, at least initially, at the expense of reliability and cost. In addition, the higher frequency of operation may induce additional unwanted vibrational noise and susceptibility to electromagnetic interference.

**FINDING 3-8**

Low-cost uncooled infrared focal plane arrays are approaching the perfor-

[21] Lon Bell, Amerigon, Inc., personal communication with the committee on March 3, 2010. Amerigon anticipates material usage in its products to decrease linearly over the next four years from 16 g to 11 g.

[22] Available at http://www.its.org/node/5263. Accessed March 24, 2010.

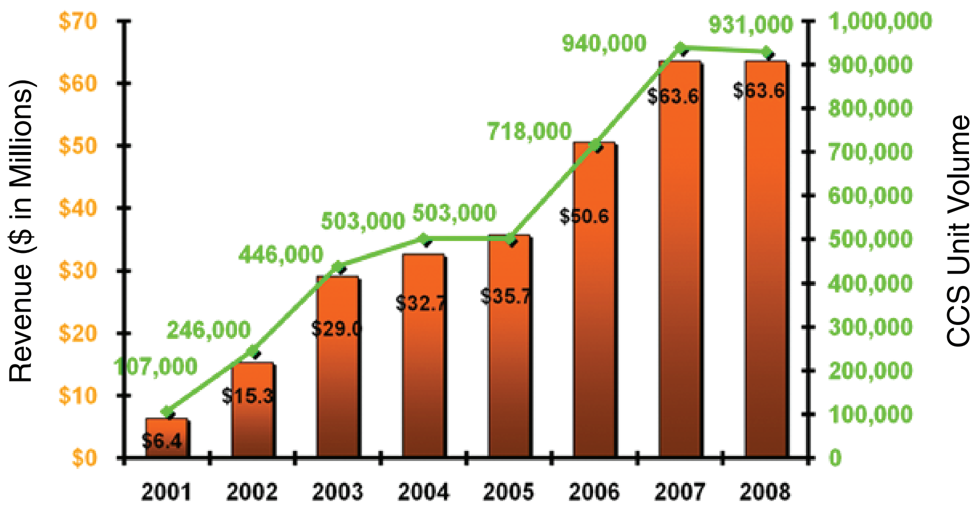


FIGURE 3-8
Estimates of the TE cooler market. SOURCE: Courtesy of Amerigon CCS.

mance needed for applications that have traditionally relied on expensive cooled devices.

**FINDING 3-9**
For both cryocooler and TE cooler technologies, there are a number of commercial market drivers, separate from sensor cooling applications, that will drive evolutionary improvements in SWaP. Over the next 10-15 years, it is reasonable to expect that these improvements will achieve overall reductions in SWaP on the order of 20-30 percent.

## CONCLUDING THOUGHTS

In addition to a proliferation of sensor numbers and types, new sensors must be developed to exploit unique target and background phenomenologies and be capable of processing significantly larger volumes of data. Advanced technology sensors have the potential to be used in novel sensing situations—for example, the vetting of potential enemies or the identification of combatants and noncombatants in a counterinsurgency operations. IR sensors with high-performance imaging FPAs have that capability and have long been critical to U.S. relative military superiority. Maintaining this superiority requires continual improvements in the technologies required for advanced ROICs and detector materials growth and continual awareness and incorporation of advances originating from both domestic and foreign developments.

# 4

# Emerging Technologies with Potentially Significant Impacts

## INTRODUCTION

The ultimate performance of a particular detector system is dependent on the integration of the various component technologies. Chapter 3 discusses the current and anticipated 10-15 year status of the various detector component technologies together with their likely impact on overall system performance. In contrast, this chapter focuses on technology breakthroughs that are more speculative in nature but, if achieved, could represent "game changing" improvements in system-level detector performance. Technologies enabling (1) advanced detection, (2) innovative optics, (3) improved coolers, and (4) enhanced signal processing are discussed in detail.

## ADVANCED DETECTION TECHNOLOGIES

### Epitaxial Growth Approaches

Epitaxial growth techniques are used to produce the active material in most long-wavelength infrared (LWIR) photon detectors. The detector material generally consists of two or more thin layers grown in succession on a substrate. Epitaxial growth implies that the crystal structure of the layers is aligned with that of the substrate, a necessary requirement for good material quality and appropriate electrical characteristics. The two most common families of epitaxial materials for LWIR applications are mercury cadmium telluride (MCT) and antimonide-based III-Vs.

For MCT, the technique of liquid-phase epitaxy (LPE) was demonstrated in the early 1980s and has matured to become a workhorse of the industry. Elements to form the layers are first dissolved in a melt of mercury or tellurium. The substrate is immersed in the melt and the temperature is ramped down, causing the elements to crystallize and form a layer. A second melt is used to form a second layer. N-type and p-type dopants, respectively, are included in the melts, so that the interface between layers becomes a p-n junction. The substrate for nearly all LPE growth is cadmium zinc telluride (CZT), which is chemically and physically compatible with MCT and is transparent in the IR.

Advantages of LPE are that (1) it occurs close to thermodynamic equilibrium, typically near 500°C, causing it to be relatively forgiving of defects; (2) dopants can be incorporated in a very controllable manner; and (3) excellent material quality is routinely achieved. The disadvantages are that detector structures requiring more than two layers are impractical, and it is not possible to maintain sharp interfaces between layers because of interdiffusion during growth. Also, LPE growth cannot be performed on alternative substrates such as silicon.

Molecular beam epitaxy (MBE) has become the preferred growth method for more advanced MCT device structures, such as two-color arrays for third-generation sensors, as well as avalanche photodiodes. It also enables the growth of MCT on silicon and GaAs substrates that are larger and cheaper than CZT. MBE growth is performed in an ultrahigh-vacuum chamber with the elements being emitted by hot effusion cells and depositing on the substrate, which is held at about 200°C. Sharp interfaces can be formed because the molecular beams can be turned on and off abruptly and because interdiffusion is negligible at the low growth temperature. In most cases the substrate is CZT, silicon, or GaAs.

MBE is more challenging than LPE because it is less tolerant of growth defects, and it requires very tight control of the substrate temperature and the beam pressures of the species arriving at the substrate. MBE equipment is more expensive to acquire and maintain than that of LPE. However, MBE technology has matured to the point that multilayer epitaxial structures, in which the MCT alloy composition and the doping are controllably changed several times during the growth run, are produced on a regular basis. This has enabled complex device structures in MCT that would have otherwise not been possible. Also, the ability to grow MCT on silicon has enabled the fabrication of very large focal plane arrays (FPAs). There is a large lattice mismatch between HgCdTe and both GaAs and silicon, which has severe implications for the growth process, in particular the formation of dislocations and other growth defects. Significant progress has been made in learning how to accommodate the lattice mismatch, particularly in the HgCdTe:GaAs system as discussed in Chapter 3. The availability of larger and cheaper substrates for epitaxial growth will have a major impact on the performance and cost of future HgCdTe FPAs.

The antimonide-based III-V materials, including strain-layer superlattice (SLS), quantum-well infrared photodetector (QWIP), and quantum-dot infrared photodetector (QDIP) structures, are grown by MBE in most cases. The technique is similar to that for MCT, except that the substrate temperature is higher (about 400 to 500°C). MBE technology for III-V materials is relatively mature, but some additional development has been required to control the composition of the atomic monolayers at the interfaces between the indium arsenide and the gallium antimonide components of the SLS in order to control the strain. Further improvements in the MBE technique are needed to minimize the populations of point defects that limit the carrier lifetimes in SLS material. The large experience base in the growth and design of electronic and photonic devices using bandgap engineering—the incorporation of multiple functional layers into device structures—opens new avenues for optimizing device performance. Examples include barrier layers to reduce dark current and amplification layers to extend the concepts of avalanche gain and Geiger mode detection further into the IR. Most of these efforts are at an early stage of research development and are likely to bear significant fruit within the next 15 years. This is an area to pay attention to for further improvements in infrared detection.

## Nanophotonics

Over the recent past, the global scientific community gradually developed technologies that could structure materials on a nanometer scale and the field of nanotechnology was developed. Arguably, nanotechnology has its roots in the challenge of Professor Richard Feynman in 1959 to build the world's smallest motor.[1] In 1996, a federal interagency working group was formed to consider the creation of a National Nanotechnology Initiative (NNI) to focus U.S. research and development (R&D) efforts,[2] and in 2000 the NNI became a formal government program. In 2003, the 21st Century Nanotechnology Research and Development Act (NRDA) gave the NNI the legislative backing needed to establish a management structure and funding.[3] The National Science Foundation has played a key role, leading and coordinating the various agencies involved, including the Department of Defense, the Department of Energy, the Environmental Protection Agency, the National Institutes of Heath, the Department of Commerce, the U.S. Department of Agriculture, the National Aeronautics and Space Administration, and many

---

[1]R.P. Feynman. 2002. *The Pleasure of Finding Things Out and the Meaning of It All.* New York: Perseus.

[2]National Research Council. 2002. *Small Wonders, Endless Frontiers: A Review of the National Nanotechnology Initiative.* Washington, D.C.: The National Academies Press.

[3]National Research Council. 2006. *A Matter of Size: Triennial Review of the National Nanotechnology Initiative.* Washington, D.C.: The National Academies Press.

others. In the past several years, the annual funding for the NNI has been at about \$1.5 billion.[4] Table 4-1 shows the distribution among the various funding sources within the U.S. government.

A similar historical evolution has occurred in many foreign countries.[5] The intent in this chapter is not to provide exhaustive historical detail; in short, Japan, Europe, and other Asian countries have kept pace with the United States with comparable government expenditures and even more significant commercial funding. Much of the work up to the past half dozen years has been oriented toward fundamental materials, establishing methods for creating nanometer structures and measuring properties. More recently, efforts have matured for building entirely new device structures using these material fabrication techniques. Quantum dots, nanotubes, and layered carbon graphene structures can be cited as examples of specific materials from which entirely new devices and applications will arise. Countries around the world are poised to take advantage of nanotechnology to potentially build entirely new sensors and sensor systems. Therefore, foreign progress in the nanotechnology field constitutes a principal driver for significant advances in sensors.

Of the many nanoscale material systems being explored, simple carbon structures are undoubtedly the most studied; perhaps these carbon structures will be, in fact, the "silicon" for nanotechnology. With several different morphologies (graphene two-dimensional [2-D] or three-dimensional [3-D] layered structures, spherical fullerene ["Bucky Balls"], and single- and multiwall nanotubes [CNT] of different chiralities) and outstanding physical properties (>30× the strength of steel and approaching the conductivity of copper), carbon is being studied for a host of applications. All of the III-V and II-VI compound classes are also extensively being explored for basic science attributes and applications; hence, techniques have emerged for reproducible quality device prototypes using the full power of modern electronic material and chip fabrication methods. It would distract us in this short introduction to list the many other types of materials such as biology-based building blocks that might have some bearing on new techniques for sensors; instead it suits our purpose here to focus on the most significant potential that nanomaterials can provide to advance the sensor state of the art. The impact of nanotechnology on future designs of sensors and sensor systems can be anticipated along the following lines.

Graphene is a material that has recently been a subject of intense study for its potential application in high-frequency electronics and photonics. Graphene consists of a monolayer of carbon atoms arranged in a 2-D hexagonal lattice.

[4] Available at www.nano.gov/NNI_FY09_budget_summary.pdf. Accessed March 25, 2010.

[5] Woodhouse, E.J., ed. 2004. Special Issue on Nanotechnology. Prepared as a Publication of the IEEE Society on Social Implications of Technology. *IEEE Technology and Society Magazine* 23(4).

TABLE 4-1 Funding of Nanophotonics by Federal Agency

| | 2009 Actual | 2009 Recovery[a] | 2010 Estimated | 2010 Proposed |
|---|---|---|---|---|
| DOE[b] | 332.6 | 293.2 | 372.9 | 423.9 |
| NSF | 408.6 | 101.2 | 417.7 | 401.3 |
| HHS-NIH | 342.8 | 73.4 | 360.6 | 382.4 |
| DOD[c] | 459.0 | 0.0 | 436.4 | 348.5 |
| DOC-NIST | 93.4 | 43.4 | 114.4 | 108.0 |
| EPA | 11.6 | 0.0 | 17.7 | 20.0 |
| HHS-NIOSH | 6.7 | 0.0 | 9.5 | 16.5 |
| NASA | 13.7 | 0.0 | 13.7 | 15.8 |
| HHS-FDA | 6.5 | 0.0 | 7.3 | 15.0 |
| DHS | 9.1 | 0.0 | 11.7 | 11.7 |
| USDA-NIFA | 9.9 | 0.0 | 10.4 | 8.9 |
| USDA-FS | 5.4 | 0.0 | 5.4 | 5.4 |
| CPSC | 0.2 | 0.0 | 0.2 | 2.2 |
| DOT-FHWA | 0.9 | 0.0 | 3.2 | 2.0 |
| DOJ | 1.2 | 0.0 | 0.0 | 0.0 |
| TOTAL[d] | 1,701.5 | 511.3 | 1,781.1 | 1,761.6 |

NOTE: CPSC = Consumer Product Safety Commission; DHS = Department of Homeland Security; DOC = Deparment of Commerce; DOD = Department of Defense; DOE = Department of Energy; DOJ = Department of Justice; DOT = Department of Transportation; EPA = Environmental Protection Agency; FDA = Food and Drug Administration; FHWA = Federal Highway Administration; HHS = Department of Health and Human Services; NASA = National Aeronautics and Space Administraion; NIFA = National Institute of Food and Agriculture; NIH = National Institutes of Health; NIOSH = National Institute for Occupational Safety and Health; NIST = National Institute of Standards and Technology; NSF = National Science Foundation; USDA = Department of Agriculture.

[a]Based on allocations of the American Recovery and Reinvestment Act (ARRA) of 2009 (P.L. 111-5) appropriations. Agencies may report additional ARRA funding for small business innovative research (SBIR) and small business technology transfer (STTR) projects later, when 2009 SBIR-STTR data become available.

[b]Includes the Office of Science, the Office of Energy Efficiency and Renewable Energy, the Office of Fossil Energy, the Office of Nuclear Energy, and the Advanced Research Projects Agency-Energy.

[c]The 2009 and 2010 DOD figures include congressionally directed funding that is outside the NNI plan ($117 million for 2009).

[d]Totals may not add, due to rounding.

SOURCE: Data from http://www.nano.gov/html/about/funding.html. Accessed May 2, 2010.

Electronically, it behaves as a zero bandgap semiconductor with extraordinary carrier mobility, even at room temperature. Graphene has also demonstrated strong photocurrent responses near graphene-metal interfaces. The combination of graphene's attractive electronic and photonic properties holds great promise for visible detector applications. In fact, recent results have demonstrated the use of graphene detectors in a 10 gigabit per second optical link with an external photoresponsivity of 6.1 mA/W at 1.55 μm wavelength.[6] The same group also reports

[6]T. Mueller, F. Xia, and P. Avouris. 2010. Graphene photodetectors for high-speed optical communications. *Nature Photonics* 4:297-301.

having demonstrated a strong photoresponse in a metal-graphene-metal (MGM) based photodetector at 514 nm, 633 nm, and 2.4 μm. Graphene's high switching speed combined with a broadband photoresponse underscores its potential to have a disruptive impact on future detector performance. The promise of a material that may surpass the performance of silicon for many electronic applications has focused a significant body of research on graphene because the mechanisms of transport in this material are not fully understood. As these research efforts mature over the next several years and new techniques for processing and fabricating graphene devices are developed, the true potential for graphene in electronic and photonic devices will become better clarified and quantified.

*Photonic Structures*

Nanostructures can be built through bottom-up self-assembly processes taking advantage of both organic and inorganic routes and top-down approaches applying lithographic techniques. Integrated circuit scales are approaching transverse dimensions of ~10 nm scale,[7] and an important trend is the merging of top-down processes, which offer long-range order and complex hierarchical structure, and bottom-up self-assembly, which offers nanometer-scale capabilities and below, with "directed self-assembly."[8] Since the first demonstration[9] of a photonic crystal in 1989, detailed work has accelerated quickly and been extended from microwave to optical frequencies. Again, not to dwell on explanations that can be found in textbooks,[10] the concepts use the precision attendant to nanostructure construction to form periodic one-, two-, and three-dimensional subwavelength structures for controlling optical radiation. Analogous to electrons in semiconductors, light propagates through periodic structures with pass bands or stop bands depending on the wavelength. All of the well-known components familiar to microwave engineers can therefore be constructed for light—for example, wavelength pass-rejection filters, resonators, isolators, circulators, and bends. Embedding absorbing or emitting optical elements in these structures permits tailoring of features such as spontaneous emission probability through a lower density of radiation states. In addition to potential advantages in designing more compact optical trains trans-

[7]Data derived from the *International Technology Roadmap for Semiconductors*, available at http://www.itrs.net. Accessed March 25, 2010.

[8]J.A. Liddle, Y. Cui, and P. Alivasatos. 2004. Lithographically directed self-assembly of nanostructures. *Journal of Vacuum Science and Technology B* 22(6):3409-3414.

[9]E.Yablonovitch and T.J. Gmitter. 1989. Photonic band structure: the face-centered cubic case. *Physics Review Letters* 63:1950-1953.

[10]J.D. Joannapoulos, S.G. Johnson, J.N. Winn, and R.D. Meade. 2008. *Photonic Crystals: Molding the Flow of Light.* Princeton, N.J.: Princeton University Press.

porting light from the collection aperture to the detector element, some tailoring of the thermal noise background is possible.[11]

These structures should not be confused with composite dielectrics, which are composed of two or more materials interspersed on a subwavelength scale without any consideration for ordering. An example is perfectly black carbon surfaces comprising "steel wool-like" features or a mixture of low- and high-index material to achieve a particular index of refraction. While subwavelength surface absorbing elements do imply the possibility of building sensors with subwavelength pixel size, diffraction effects limit the minimum pixel sizes independent of the length scale of the absorber as discussed in Chapter 2. Metamaterials are an emerging class of materials with wholly new properties such as a negative index of refraction that offer additional possibilities for managing and directing optical paths in nonclassical ways.[12] Additionally, plasmonics takes advantage of the very large (and negative) dielectric constant of metals, to compress the wavelength and enhance electromagnetic fields in the vicinity of metal conductors. This has been referred to as "ultraviolet wavelengths at optical frequencies"[13] and is the basis of many well-studied phenomena such as surface-enhanced Raman scattering (SERS) and surface plasma wave chemical-biological sensors.[14] Additional discussion of the application to infrared detectors is presented in the following section.

### *Electronics*

The broad applicability of nanotechnology to electronics is obvious; for example, the use of cathodic electron field emission from an assemblage of nanotubes for high-power microwave transmitters[15] and other vacuum electronic applications offers copious production of electrons; this particular technology may find immediate application in fielded systems. On an individual scale, single-wall carbon nanotubes (SWNTs) can be isolated with adequate properties[16] to demonstrate transistor action for microelectronic circuits. Techniques for generating and manipulating individual SWNTs have been perfected to the point that metal-metal,

[11] S-Y. Lin, J.G. Fleming, E. Chow, J. Bur, K.K. Choi, and A. Goldberg. 2000. Enhancement and suppression of thermal emission by a three-dimensional photonic crystal. *Physical Review B* 62(4): R2243-R2246.

[12] W. Cai and V. Shalaev. 2010. *Optical Metamaterials: Fundamentals and Applications.* Berlin: Springer-Verlag.

[13] M. Dragoman and D. Dragoman. 2008. Plasmonics: applications to nanoscale terahertz and optical devices. *Progress in Quantum Electronics* 32:1-4.

[14] J. Homola, S.S. Lee, and G. Gauglitz. 1999. Surface plasmon resonance sensors: review. *Sensors and Actuators B: Chemical* 54:3-15.

[15] K.L. Averett, J.E. Van Nostrand, J.D. Albrecht, Y.S. Chen, and C.C. Yang. 2007. Epitaxial overgrowth of GaN nanocolumns. *Journal of Vacuum Science and Technology B* 25(3):964-968.

[16] Sang N. Kim, Zhifeng Kuang, James G. Grote, Barry L. Farmer, and Rajesh R. Naik. 2008. Enrichment of (6,5) single wall carbon nanotubes using genomic DNA. *Nano Letters* 8(12):4415-4420.

metal-semiconductor, and semiconductor-semiconductor junctions can be reproducibly formed and the I-V curves measured;[17] however, scaling this to the densities and defect levels already reached for complementary metal oxide semiconductor (CMOS) applications remains an open question. It can be noted that this last reference is from the Indian Institute of Science in Bangalore, illustrating the global sweep of this important technology. SWNTs have been assembled into electronic circuits in elementary "chips" with >20,000 elements,[18] and field effect transistors (FETs) have been reproducibly constructed to build a 10 FET ring oscillator.[19] Competitive electronic applications are many years behind the level of sophistication needed to contemplate actual integration of a nanoscale readout integrated circuit (ROIC) into operating sensors. Still the ultimate payoff of a fully integrated sensor element with nanoscale processing requires ongoing monitoring of global improvement and activity.

*Sensor Elements*

Quantum sensor elements receive an incoming photon and free a bound electron(s) for amplification and signal processing. The nanomaterial necessarily must have well-defined optical and electronic properties. One such material is a quantum dot wherein the physical dimensions are reduced to the point that electron states are no longer defined by an infinite crystal lattice; rather, the dot's physical dimension fixes the permissible energy bands, very much a man-made atomic system. Quantum-well structures also tailor bands, and QWIP sensor elements are discussed in another section of this report along with pixel-sized antennas to guide incoming radiation into the element. At this stage of nanotechnology detector elements, QWIP and QDIP structures are the most studied, but entirely new configurations might be possible and literature should be appropriately scanned.

*Plasmonic Enhancement of Detectors*

The dielectric properties of metals are often described by a free-carrier Drüde model given by

$$\varepsilon(\omega)=1-\frac{\omega^2}{\omega(\omega+i\nu)},$$

[17]C.N.R. Rao, R. Voggu, and A Govindara. 2009. Selective generation of single-walled carbon nanotubes with metallic, semiconducting, and other unique electronic properties. *Nanscale* 1:96-105.

[18]C.W. Zhou, J. Kong, E. Yenilmez, and H. J. Dai. 2000. Modulated chemical doping of individual carbon nanotubes. *Science* 290:1552-1555.

[19]P. Avouris. 2009. Carbon nanotube electronics and photonics. *Physics Today* 62(1):34-40.

where

$$\omega_p = \sqrt{\frac{4\pi e^2 N}{\kappa \varepsilon_0 m^*}}$$

is the plasma frequency in the metal, with $e$ the electronic charge, $N$ the carrier concentration, $\kappa$ the relative dielectric constant arising from the bound electrons, and $m^*$ the electron effective mass. For single-electron per atom metals such as gold and silver, $\omega_p$ is in the ultraviolet spectral region. Here $\nu$ is the electron collision frequency that is typically in the terahertz regime. At radio frequencies, $\omega/\nu << 1$, and the metal response is large and imaginary (out of phase with the driving electric field). Throughout the infrared, $\omega/\nu >> 1$, and the metal response is large and negative with a smaller imaginary part. This is the plasmonic regime. At visible frequencies, additional losses due to bound electron transitions become more important and the dielectric function is both lossier and not as negative. For most metals except gold, silver, and aluminum, the dielectric function is positive across the visible.

Some of the implications of this dielectric function have been recognized for more than 100 years. Sommerfeld was the first to recognize the existence of a bound surface mode at the interface between a lossless dielectric and a lossy metal in his analysis of Marconi's wireless transmission experiments (there the loss was associated with currents in Earth's surface).[20,21,22] More recently, interest in plasmonics has been rekindled with the discovery of anomalous transmission through a metal slab perforated with a 2-D array of holes.[23,24] This transmission is associated with resonances involving the coupling of the incident radiation to surface plasma waves (SPW) localized to a metal-dielectric interface (a thin metal film has two such SPWs one on either side of the film) and the localized resonances associated with the holes (or other unit-cell structures such as annuli[25]).

[20] A. Sommerfeld. 1909. Über die Ausbreitung der Wellen in der drahtlosen Telegraphie. *Annalen der Physik* 28:665-737.

[21] A. Baños. 1966. *Dipole Radiation in the Presence of a Conducting Half-Space.* Oxford: Pergammon Press.

[22] S.R.J. Brueck. 2000. Radiation from a dipole embedded in a dielectric slab. *IEEE Journal of Selected Topics in Quantum Electronics* 6:899-910.

[23] T.W. Ebbesen, H.J. Lezec, H.F. Ghaemi, T. Thio, and P.A. Wolff. 1998. Extraordinary optical transmission through sub-wavelength hole arrays. *Nature* 391:667-669.

[24] For a more recent review, see C. Genetand T.W. Ebbesen. 2007. Light in tiny holes. *Nature* 445:39-46.

[25] W. Fan, S. Zhang, B. Minhas, K.J. Malloy, and S.R.J. Brueck. 2005. Enhanced infrared transmission through subwavelength coaxial metallic arrays. *Physics Review Letters* 94:033902.

The implications for detectors were recognized quite some time ago[26] and rediscovered soon after the discovery of the anomalous transmission.[27] In addition to the distributed SPW coupling, work has also been reported using a shaped plasmonic lens structure to funnel all of the incident radiation to a single small detector at the center of a bull's-eye pattern.[28] This particular experiment was for a SWIR detector for integration with silicon integrated circuits, and the goal was reduced capacitance for higher-speed operation. For the infrared, this approach is aimed mainly at reducing the detector volume and, consequently, thermal noise sources such as generation-recombination dark current in high-operating-temperature MWIR detectors.[29] The difficulty is finding the appropriate combination of SPW coupling, hole transmission, and angular and spectral bandwidth while still retaining the ability to collect the photo- or plasmon-generated carriers.

Very recently, a 30× enhancement in detectivity was obtained for a SPW coupled QDIP detector[30] by coupling using a similar transmission metal grating. This is very early work, still in the exploratory research stage, and far from ready for integration into commercial focal planes, but it does offer the potential of a multispectral, MWIR-LWIR focal plane array with high quantum efficiency, polarization, and spectral selectivity and some degree of electrical tuning based on the quantum-confined Stark effect in QDIP detectors. Ultimately, based on previous results on coupling to SPWs, quantum efficiencies near unity should be possible. These arrays could be a new direction in IR FPAs; however, the road to a fielded product is long and success is not guaranteed. The difficulty is finding the appropriate combination of SPW coupling, hole transmission, and angular and spectral bandwidth while still retaining the ability to collect the photo- or plasmon-generated carriers.

**FINDING 4-1**
An emerging trend in focal plane array technologies is multispectral band sensing enabling enhanced system capability through a single aperture. Spectral information is an added discriminant for enhanced detection selectivity and material identification.

[26]S.R.J. Brueck, V. Diadiuk, T. Jones, and W. Lenth. 1985. Enhanced quantum efficiency internal photoemission detectors by grating coupling to surface plasma waves. *Applied Physics Letters* 46:915-917.

[27]Z. Yu, G. Veronis, S. Fan, and M.L. Brongersma, Design of mid-infrared photodetectors enhanced by surface plasmons on grating structures. *Applied Physics Letters* 89, 151116 (2006).

[28]T. Ishi, J. Fujikata, K. Makita, T. Baba, and K. Ohashi. 2005. Si nano-photodiode with a surface plasmon antenna. *Japanese Journal of Applied Physics* 44: L364-L366.

[29]R.D.R. Bhat, N.C. Panoiu, S.R.J. Brueck, and R.M. Osgood. 2008. Enhancing the signal-to-noise ratio of an infrared photodetector with a circular metal grating. *Optics Express* 16(7):4588-4596.

[30]S.C. Lee, S. Krishna, and S.R.J. Brueck. 2009. Quantum dot infrared photodetector enhanced by surface plasma wave excitation. *Optics Express* 17(25):23160-23168.

**FINDING 4-2**
By manipulating fields at the subwavelength scale, nanophotonics offers a potential for enhanced detector functionality, particularly in adding wavelength and/or polarization selectivity along with enhanced detectivity at the pixel scale.

## Antennas

An antenna is a transduction device that is used to transmit or receive electromagnetic radiation. Traditionally, antennas have been associated with transmit or receive applications in the radio-frequency (RF) spectrum. Examples of RF antennas can be found in many aspects of everyday life since they are a ubiquitous component in television, radio, voice, data, radar, and other communication networks. However, within the past decade, the potential for developing antennas that work in the visible and IR spectra has been explored with increasing interest. In this section, recent research in IR and optical antennas is reviewed and assessed in terms of relevance to visible and IR detector technologies.

Antennas are composed of various conducting elements arranged in a pattern designed for optimum performance for a given application. Typical performance parameters cited are gain, bandwidth, and efficiency and these are all functionally related to both the size of the antenna elements and the conducting (electrical and magnetic) properties of the materials that comprise the antenna. While the basic equations that govern antenna performance in the RF spectrum have analogues in the IR and optical spectra, the physical realization of elements that constitute a functional antenna are vastly different. The two major reasons for these differences include the fact that antenna size scales with wavelength and that material losses increase significantly in conductors with increased frequency. The scaling issue requires that antennas working in the IR and optical spectra be typically on the order of microns or smaller, and the loss issue presents significant challenges in designing antennas that achieve reasonable efficiencies at IR and/or optical wavelengths.

Despite the challenges in developing antennas for optical and/or IR applications, there has been a significant body of research focused on exploring this topic. The primary drivers for this research are the need for achieving large absorption cross section together with high field localization and/or enhancement for applications in nanoscale imaging and spectroscopy, solar energy conversion, and coherent control of light emission and/or absorption.[31] It is difficult to speculate on the

[31] An overview of the current research on optical antennas as well as a discussion of their potential applications can be found in Palash Bharadwaj, Bradley Deutsch, and Lukas Novotny. 2009. Optical antennas. *Advances in Optics and Photonics* 1(3):438-483 along with the article's associated references. Additional useful summaries of research that is representative of the fundamental nature of current optical and IR antenna research are found in P. Mühlschlegel, H.-J. Eisler, O.J.F. Martin, B.Hecht, and D.W. Pohl. 2005. Resonant optical antennas. *Science* 308(5728):1607-1609; J. Greffet. 2005. Applied

potential future impact of optical-IR antennas on detector systems at this time due to the fundamental and immature status of the research. It is worth considering, however, how an optical and/or IR antenna might be employed in a detector system and what performance benefits it might enable.

One example is to exploit the wavelength and polarization selectivity of antennas to enable detectors with electronically tunable spectral responses. A brief discussion of this concept follows. An antenna-coupled IR sensor receives flux with an antenna that is tuned to optical wavelengths, and only photons in a selective waveband pass into the detector. Typically the antenna is used to receive radiation and is coupled to an infrared or thermal detector to capture the signal power. Such devices have been used for infrared and millimeter-wave sensing. One advantage of an antenna-coupled IR sensor is that spectral and polarization responsivities can be set by antenna size and orientation. In addition, the spectral response of these devices can be tuned electronically.[32] Typical devices are fabricated on top of a dielectrically coated metallic ground plane as shown in Figure 4-1(a) and 4-1(b), in cross section and top views, respectively. Figure 4-1(c) shows an equivalent circuit model that was developed for a device that incorporated a metal oxide semiconductor (MOS) capacitor for spectral tuning. The gold antenna arms, the $SiO_2$, and the silicon substrate form the MOS-cap. Applying bias to the backside of the silicon changes the depletion width of the capacitor.

The MOS-capacitor pair acts as a varactor ($C_{mos}$), in series with the antenna capacitance $C_{antenna}$. Their equivalent capacitance $C_{eq}$, is in parallel with the antenna inductance $L_{antenna}$ and the fringe-field capacitance $C_{farfield}$. When the MOS-capacitor is biased in the depletion region, $W_{eff}$ continues to expand until the device enters into the inversion region. Figure 4-2 displays both the measured and modeled results from an IR MOS tuner.[33] It shows actual tuning of a nanofabricated antenna, coupled with a conventional photodetector. These measured results were taken with a single antenna that was hand-fabricated by etching with a mask made by simple contact photolithographic masking. Higher precision would be expected in production, with the use of more sophisticated fabrication and measurement

physics: nanoantennas for light emission. *Science* 308(5728):1561-1563; and C. Soukoulis, S. Linden, and M. Wegener. 2007. Negative refractive index at optical wavelengths. *Science* 315(5808):47-49. A survey of these references and other literature indicates that the practically achievable efficiencies for absorbing or emitting radiation in the IR or visible range using antenna structures are quite low (1 percent or less) unless the antenna and source are maintained at distances considerably smaller than the radiation wavelength.

[32]M.A. Gritz. 2003. Fabrication of Infrared Antennas. Ph.D. dissertation. University of Central Florida, Orlando.

[33]M.A. Gritz. 2003. Fabrication of Infrared Antennas. Ph.D. dissertation. University of Central Florida, Orlando.

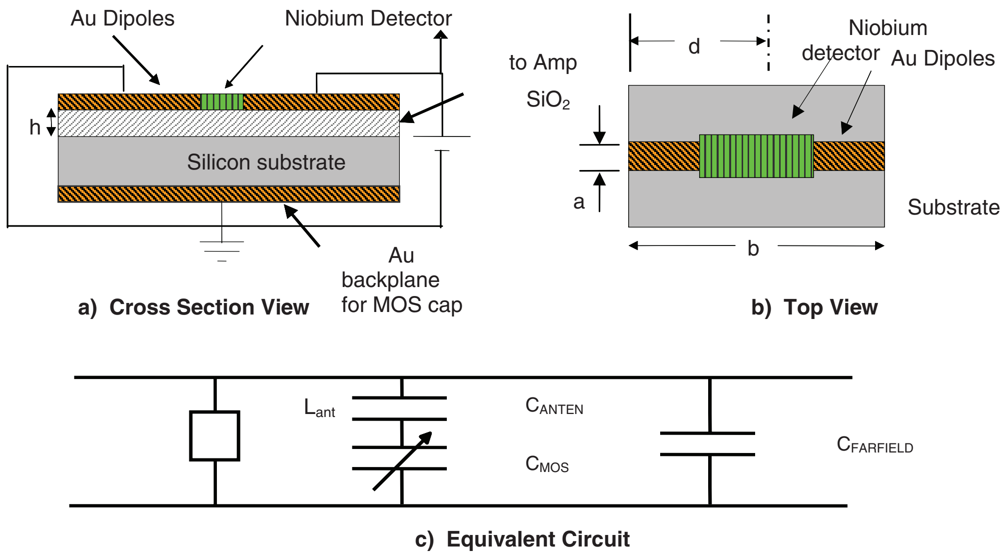


FIGURE 4-1
Typical dipole antenna-coupled detector.

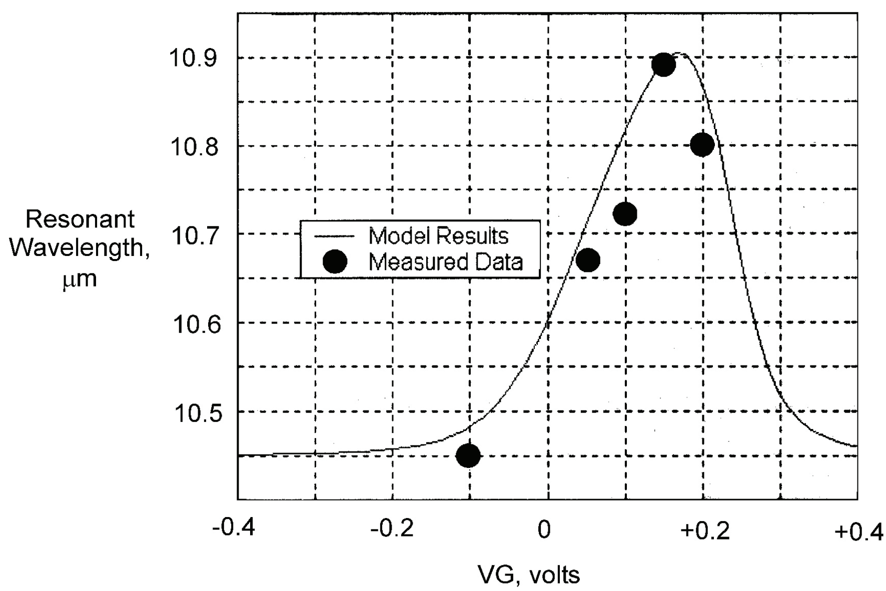


FIGURE 4-2
Modeled and measured data for the wavelength tuning. SOURCE: Gritz, M.A. 2003. Fabrication of Infrared Antennas. Ph.D. dissertation. University of Central Florida, Orlando.

equipment. Nonetheless, the tunability of this antenna structure is evident, as is the accuracy of the predictive model, showing practicality of design and fabrication.

**FINDING 4-3**
IR and/or optical antennas for detectors are theoretically capable of surpassing current uncooled bolometer detector sensitivity, specifically in the MWIR. However without significant investments in materials, manufacturing, and fabrication, systems based on antenna technology are not likely to achieve a maturity level that would enable their practical use in a 10-15 year time frame.

## Wavelength Up-conversion

In principle, wavelength conversion is very attractive. Visible detectors are inexpensive, sensitive, and uncooled; have huge formats; and have very small pixel pitch. Visible detectors are superior to all subbands of infrared detectors in almost all aspects. The concept of wavelength conversion from an infrared wavelength to the visible would make these inexpensive, superior, focal planes available to detect radiation at infrared wavelengths. Unfortunately up-conversion has not achieved significant success to date. Visible photons contain more energy than infrared photons. Energy must be added to convert IR photons to visible photons. A pump will be required that adds energy, preferably without adding significant noise. An alternate way to approach up-conversion could be to have IR photons driving the emission of visible photons.[34] Major issues for any up-conversion approach are (1) achieving high conversion efficiency and (2) minimizing noise added by the up-conversion process. Background-limited infrared photodetection (BLIP) performance, or near-BLIP, performance is desired.

A variety of physical mechanisms have been attempted for up-conversion without achieving practical device-level success. Some recent schemes have focused on variants to achieving room-temperature detectors as an alternative to microbolometers: (1) free carriers in silicon using reemission from silicon clusters,[35] and (2) optical refractive index change with absorbed radiation.[36] These sensor concepts may find use in specialized applications but have yet to provide a competitive level of performance. In this class of technologies, a film or material of some sort is used with a bandgap that is too large to absorb infrared radiation. A pump laser, or potentially some other energy source, is used to keep the material in a metastable

[34]Hui Chun Liu. 2006. Photon upconversion devices for long wavelength imaging. *SPIE Newsroom.* Available at http://spie.org/x8627.xml?highlight=x2408&ArticleID=x8627. Last accessed March 25, 2010.

[35]R. Kipper, D. Arbel, E. Baskin, A. Fayer, A. Epstein, N. Shuall, A. Saguy, D. Veksler, B. Spektor, D. Ben-Aharon, and V. Garber. 2009. The roadmap for low price-high performance IR detector based on LWIR to NIR light up-conversion approach. *Proceedings of the SPIE* 7298:72980J-72980J-5.

[36]Available at http://www.redshiftsytems.com. Last accessed March 25, 2010.

excited state where an infrared photon can provide the required energy to affect the photoionization. Several companies have pursued these types (and related) technologies including Sirica in Israel and Red Shift Technologies in the United States. The low manufacturing cost of these alternative uncooled technologies could lead them to be widely adopted if technology breakthroughs are achieved. The excited state needs to be decoupled from the thermal energy of an uncooled conversion layer. A second area of recent interest has been for quantum communication in the 1.3 to 1.5 μm band, up-converting to be able to use silicon detector arrays. This appears to be an active area of research.

Future directions for wavelength up-conversion sensors may utilize the pass-stop band effects of engineered materials to advantage. Simultaneous impingement of a strong pump signal and an emitted weak detected signal are necessary requirements. Photonic materials provide this effect and might help isolate extraneous pump signal noise and improve the overall quantum efficiency.

**FINDING 4-4**

Efficient, image wavelength conversion from IR to optical would have a high impact due to the low cost, low inherent noise, and technological maturity of visible imaging. To date the low efficiency and added noise of wavelength conversion approaches have not made this an attractive alternative to direct IR detection. The technologies reviewed for up-conversion do not show an obvious path to reach maturity within 10-15 years.

### MEMS Bi-morph Cantilevers

Microelectomechanical systems (MEMS) bi-morph cantilever devices are of interest for the night vision area. As an alternative to resistive bolometers, there has been a recent interest in trying to optimize mechanical cantilever systems. In cantilever sensing, IR energy is absorbed by a thermally isolated paddle and then transferred to a mechanical actuator comprised of two materials with dissimilar thermal expansion coefficients (e.g., a metal and an insulator). The mechanical actuator physically displaces the paddle based on the temperature change. The thermal isolation of both the paddle and the actuator help to control the trade-off between the amount of displacement (sensitivity) and the relaxation time (speed) of the pixel. The position of the pixel can be read out either electronically (as capacitance) or optically. Optical readout removes the need for a complex and expensive ROIC by using a visible light source (such as a laser) to transducer cantilever position. Readout can be realized using a visible camera or eye depending on the system. Visible cameras can leverage the cost reductions, power efficiency improvements, and volumes of the market. An imager built using this principle would be very power efficient using the strong market drivers for the visible camera. The key challenge is to build an array with sufficient uniformity to yield an acceptable image and to address or eliminate the pixel-by-pixel corrections used in other uncooled IR imagers.[37,38]

**FINDING 4-5**
If the material combination, growth, and deposition uniformity challenges can be overcome and if pixel-specific correction is developed at the unit cell level, the MEMs bi-morph cantilever technology is attractive for uncooled LWIR arrays. MEMs fabrication and infrastructure are already in place, and the commercial pull for lower cost and reduced power would rapidly drive this transition.

## Optomechanical Devices

Photonic devices without electronic components have the potential to be detectors in environments where traditional electronics has failed such as high temperature, hostile electromagnetic environments, et cetera. Optical sensing and actuation incorporating mechanical components could be explored as detectors as this technology matures further.[39,40,41,42,43,44,45]

---

[37] C.D.W., Jones, C.A. Bolle, R. Ryf, M.E. Simon, F. Pardo, V.A. Aksyuk, W.Y.-C. Lai, J.E. Bower, J.F. Miner, F.P. Klemens, R.A. Cirelli, T.W. Sorsch, E.J. Ferry, L.A. Fetter, C.-S. Pai, J.A. Taylor, B. Vyas, G.P. Watson, B. Stekas, M.R. Baker, A.R. Papazian, N.R. Basavanhally, W.M. Mansfield, A. Kornblit, R.C. Keller, J.V. Gates, and A.P. Ramirez. 2009. MEMS thermal imager with optical readout. *Sensors and Actuators A* 155(1):47-57.

[38] Dong Feng-Liang, Zhang Qing-Chuan, Chen Da-Peng, Miao Zheng-Yu, Xiong Zhi-Ming, Guo Zhe-Ying, Li Chao-Bo, Jiao Bin-Bin, and Wu Xiao-Ping. 2007. Optimized optomechanical micro cantilever array for uncooled infrared imaging. *Chinese Physical Letters* 24(12):3362-3364.

[39] P.T. Rakich, Miloš A. Popovic, and Zheng Wang. 2009. General treatment of optical forces and potentials in mechanically variable photonic systems. *Optics Express* 17(20):18116-18135.

[40] P.T. Rakich, Miloš A. Popovic, Marin Soljaęič, and Erich P. Ippen. 2007. Trapping, corralling and spectral bonding of optical resonances through optically induced potentials. *Nature Photonics* 1:658-665.

[41] Amit Mizrahiand Levi Schächter. 2007. Two-slab all-optical spring. *Optics Letters* 32(6): 692-694.

[42] Matt Eichenfield, Christopher P. Michael, Raviv Perahia, and Oskar Painter. 2007. Actuation of micro-optomechanical systems via cavity enhanced optical dipole forces. *Nature Photonics* 1:416-422.

[43] Michelle Povinelli, Steven Johnson, Marko Lonèar, Mihai Ibanescu, Elizabeth Smythe, Federico Capasso, and J. Joannopoulos. 2005. High-Q enhancement of attractive and repulsive optical forces between coupled whispering-gallery mode resonators. *Optics Express* 13(20):8286-8295.

[44] Michelle Povinelli, Marko Lonęar, Mihai Ibanescu, Elizabeth J. Smythe, Steven G. Johnson, Federico Capasso, and John D. Joannopoulos. 2005. Evanescent-wave bonding between optical waveguides. *Optics Letters* 30(22):3042-3044.

[45] Masaya Notomi, Hideaki Taniyama, Satoshi Mitsugi, and Eiichi Kuramochi. 2006. Optomechanical wavelength and energy conversion in high-Q double layer cavities of photonic crystal slabs. *Physical Review Letters* 97(2):023903.

## Bioinspired Detection

It has long been recognized that nature, after many millennia of evolution, has achieved an ability to control the optical properties of complex and hierarchical biological surface structures in ways that, to date, cannot be replicated via artificial means. Examples include the iridescent color of butterfly wings and the reflecting structures in squid that are used for camouflage.[46] It has only been in recent years that the research community has been able to identify the protein structures used for coloration and to characterize and reproduce the complex, hierarchical reflective structures used for tuning the optical reflectance.

Understanding how nature achieves such a high degree of optical functionality at minimal size, weight, and power is the focus of a number of government-sponsored fundamental research programs. In particular, the Defense Advanced Research Projects Agency (DARPA) has recently initiated a Bioinspired Photonics program, and the Air Force Office of Scientific Research is funding a Multidisciplinary University Research Initiative (MURI) on Optical Effects Through Nature's Photonic Control. These efforts are considered fundamental research and are generally representative of the state of maturity of the field. It is expected that these efforts, with significant university involvement, will assist in cross-training individuals in both biological and traditional engineering; this will be required for progress in the area. Progress in this field can best be monitored by following these programs and periodic review of relevant articles in research journals. The following background material gives an overall picture of the current research in bioinspired photonics:

- Memis, Omer Gokalp, and Hooman Mohseni. 2008. A single photon detector inspired by the human eye. *SPIE Newsroom.* Available at http://spie.org/x19173.xml?ArticleID=x19173. Last accessed March 25, 2010.
- Memis, Omer Gokalp, and Hooman Mohseni. 2007. Long-wave infrared detectors: inspired by nature, IR detector targets long-wavelength applications. Available at http://www.optoiq.com/index/photonics-technologies-applications/lfw-display/lfw-article-display/289406/articles/laser-focus-world/volume-43/issue-4/features/long-wave-infrared-detectors-inspired-by-nature-ir-detector-targets-long-wavelength-applications.html. Last accessed March 25, 2010. This article reports on the single-photon detectors developed by BISOL.
- Wu, Wei, Alex Katsnelson, Omer G. Memis, and Hooman Mohseni. 2007. A deep sub-wavelength process for formation of highly uniform arrays of

[46] Peter Forbes. 2009. *Dazzled and Deceived: Mimicry and Camouflage.* New Haven: Yale University Press.

nano-holes and nano-pillars. *Nanotechnology* 18(48):485301. This has been rated among the most popular articles in *Nanotechnology* journal, exceeding 250 downloads in two months.

- Memis, Omer Gokalp, Alex Katsnelson, Soon-Cheol Kong, Hooman Mohseni, Minjun Yan, Shuang Zhang, Tim Hossain, Niu Jin, and Ilesanmi Adesida. 2007. A photon detector with very high gain at low bias and at room temperature. *Applied Physics Letters* 91:171112.
- Gelfand, Ryan M., Lukas Bruderer, and Hooman Mohseni. 2009. Nanocavity plasmonic device for ultrabroadband single molecule sensing. *Optics Letters* 34(7):1087-1089.
- Rizk, C.G., P.O. Pouliquen, and A.G. Andreou. 2010. Flexible readout and integration sensor (FRIS): new class of imaging sensor arrays optimized for air and missile defense. *John Hopkins APL Technical Digest* 28(3):252-253. Available at http://techdigest.jhuapl.edu/td2803/33RizkFlexible.pdf. Last accessed March 25, 2010.
- Andreou, A.G., P.O. Pouliquen, and C.G. Rizk. 2009. Noise analysis and comparison of analog and digital readout integrated circuits for infrared focal plane arrays. Pp. 695-699 in *Proceedings of the 43rd Annual Conference on Information Sciences and Systems* (CISS09), Baltimore, Md.
- Kohoutek, John, Ivy Yoke Leng Wan, and Hooman Mohseni. 2010. Dynamic measurement and modeling of the Casimir force at the nanometer scale. *Applied Physics Letters* 96:063106.
- Wu, Wei, Dibyendu Dey, Omer G. Memis, and Hooman Mohseni. 2009. Modeling and fabrication of electrically tunable quantum dot intersubband devices. *Applied Physics Letters* 94:193113.
- Wu, Wei, Dibyendu Dey, Alex Katsnelson, Omer G. Memis, and Hooman Mohseni. 2008. Large areas of periodic nanoholes perforated in multistacked films produced by lift-off. *Journal of Vacuum Science and Technology B* 26(5):1745-1747.

## EMERGING INNOVATIVE OPTICAL TECHNOLOGIES

### Microlenses

Microlenses have historically been widely used in visible cameras to increase the light coupling efficiency. Typically, the unit cell area in a focal plane array does not have 100 percent fill factor due to the peripheral electronics that are needed to store the signal and to communicate with external electronics. As shown in Figure 4-3, in order to maximize the optical efficiency, microlens arrays are used above the pixel to focus light into the sensitive part of the cell and away from any other non-photo-sensitive circuitry.

Recently, this technology has been extended to IR FPAs. This can include both

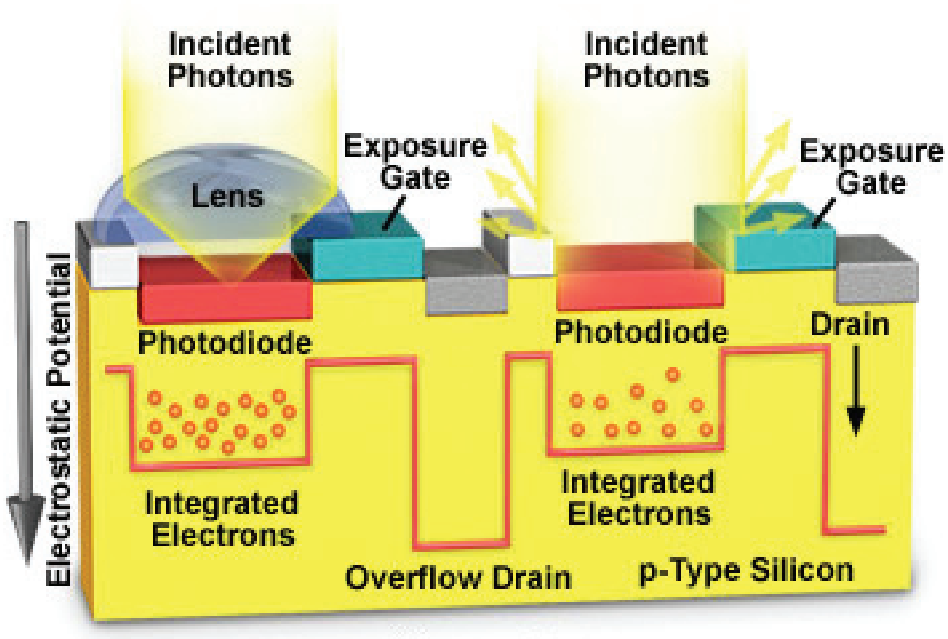


FIGURE 4-3
Microlens array architecture. SOURCE: Available from http://micro.magnet.fsu.edu/primer/digitalimaging/concepts/microlensarray.html. Accessed August 2, 2010.

traditional microlenses similar to the visible counterpart and microwave guide devices as per Figure 4-3, which can also be used to guide light into the photoactive region and attain higher effective fill factors. Several domestic IR FPA manufacturers use various embodiments of this technology.

Recent data, including presentations from the 2010 SPIE DSS Conference, indicate that foreign IR FPA manufacturers (specifically Sofradir of France) are also incorporating microlens technology for fill factor enhancements. This is an important technological enhancer because it allows greater flexibility in pixel unit cell layout, without compromise in optical absorption sensitivity. As such, it is an enabler of both higher IR FPA sensitivity and advanced unit cell electronic features. Manufacturing maturity in this area also indicates the potential for next steps into other novel micro-optical features, such as on-board polarizers, as discussed below.

### Integration of Optics with Focal Plane Arrays

Novel optical configurations, specifically configurations in which optical components are integrated with the detector, can provide new and improved levels of performance for optical detectors for certain applications, as shown in Figure 4-4. For example, in Thin Observation Module using Bound Optics (TOMBO) lenslets, under the DARPA Montage program, an array of multiple small lenses is used to replace the single macrolens typically used in imaging systems. Although on-detector microlenses have been used extensively in industry for many years, these are generally just light concentrators behind traditional optics. In the TOMBO configuration, the minilenses are the imaging element. In a typical example, a focal plane array of detectors (for example 1280 × 1024) is broken up into 20 squares of 256 × 256 pixels each in a 5 × 4 grid. Twenty lenslets form 20 independent images of the scene on the detector, each slightly offset from the next. Superresolution reconstruction is used to put together a full resolution image of the scene.

A variant of the TOMBO configuration allows a simple multispectral and/or polarimetric imager to be formed. In a typical multispectral or polarimetric imager, adjacent pixels with pixel-level polarizers having alternating polarization orientation are used. Difficulties arise in fabricating pixel-sized (~20 μm) polarizers with high quality. Also, diffraction effects impact the polarization extinction ratio. In a TOMBO configuration, of the 20 available images in the example above, a few can be dedicated to polarimetric and/or spectral image products. Since a large block of pixels is used, the fabrication becomes simpler, and the diffraction from the larger

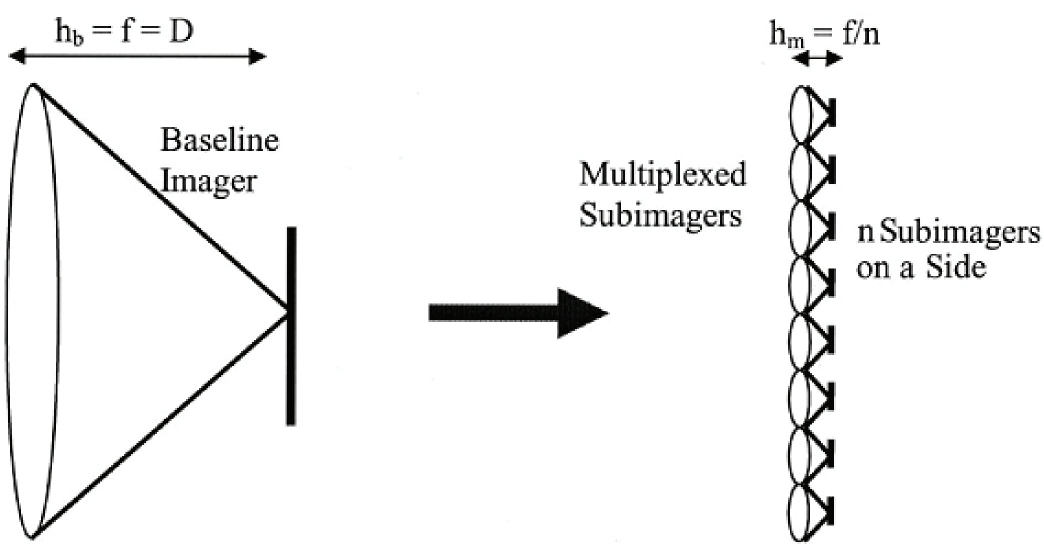


FIGURE 4-4
A substantial size, weight, and power (SWaP) advantage is possible by replacing traditional optics with a focusing lenslet array at the expense of additional processing to combine the multiple images.

aperture is essentially negligible. In the example below, each of several bands and polarizations is imaged at 256 × 256 resolution, still leaving 1280 × 768 pixels for a monochrome main image of HD resolution. As a final point, the small amount of parallax between subimages can generate some depth or range information as well, as shown in Figure 4-5. Developments in this arena could allow a wide diversity of image products to be obtained from a single compact imaging system.

## Compressive Sensing

The idea of compressive sensing is to measure only the data you will keep.[47,48] The concepts behind compressive sensing began only a few years ago.[49,50,51,52] A precursor to compressed sensing was seen in the 1970s, when seismologists constructed images of reflective layers within the Earth based on data that did not seem to satisfy the sampling criterion.[53] Compressive sensing could have a major effect on detector use in systems because a smaller number of detectors could image over a large area. While the size of detector arrays is ever increasing, if it were possible to use the same array and obtain an image similar to that obtained from a much larger array, this would have major impact at the systems level.

People have for years taken large images, but then data handling and communications limitations reduced the amount of data to be transmitted, in either a lossless, or a lossy, manner. Converting a bit-mapped image into a Joint Photographic Experts Group (JPEG) image is an example of what is done on a regular basis. Most digital cameras today store pictures in a JPEG format, but the picture is actually taken in a fully bit-mapped manner. In concept, compressive sensing is very simple where the image is compressed directly, reducing the sampling below the Nyquist-Shannon limit. The ability to accomplish this can vastly multiply the effective size of a detector array and have a major impact on visible and IR sensors. Compressive sensing does not follow sampling theory. Sparse sampling can be used,

[47]Emmanuel Candès and Michael Wakin. 2008. An introduction to compressive sampling. *IEEE Signal Processing Magazine* 25(2):21-30.

[48]Justin Romberg. 2008. Imaging via compressive sampling. *IEEE Signal Processing Magazine* 25(2): 14-20.

[49]More information is available at http://en.wikipedia.org/wiki/Compressed_sensing. Last accessed March 25, 2010.

[50]Emmanuel Candès. 2006. Compressive sampling. *Int. Congress of Mathematics* 3:1433-1452, Madrid, Spain.

[51]David Donoho. 2006. Compressed sensing. *IEEE Trans. on Information Theory* 52(4):1289-1306.

[52]Richard Baraniuk. 2007. Compressive sensing. *IEEE Signal Processing Magazine* 24(4):118-121.

[53]The sampling theorem, also called the Nyquist-Shannon criterion, states that a function that is band limited to frequencies (spatial frequencies in imaging) less than $B$ is completely determined with a series of samples of separation $1/2B$.

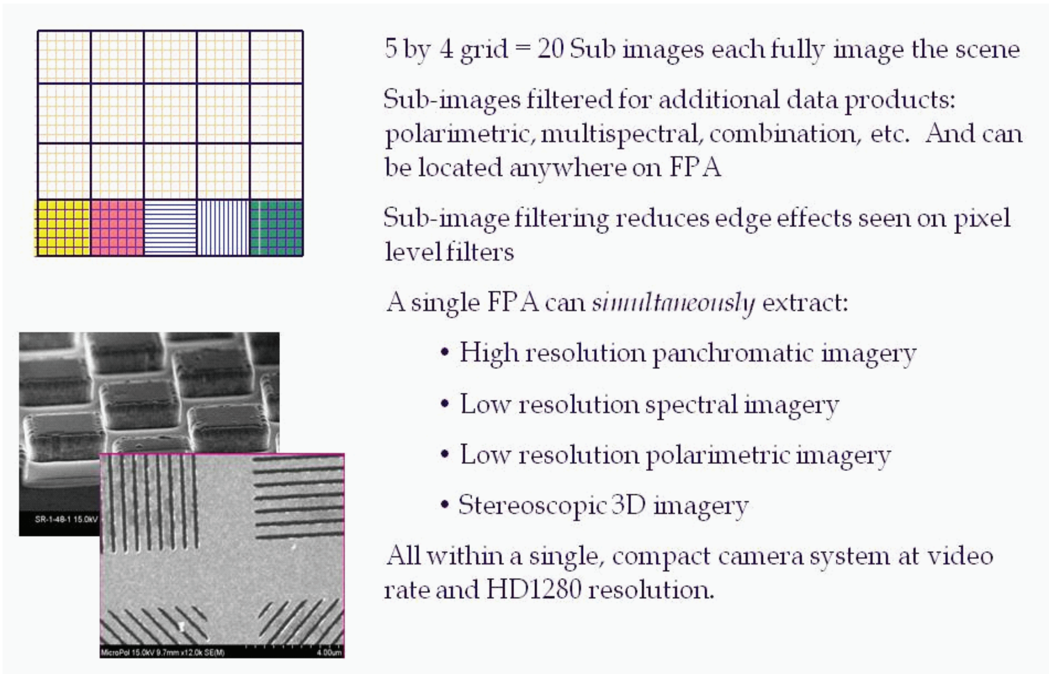


FIGURE 4-5
Description of an optical diversity imager.

but the full image can be recreated under certain constraints.[54] This is an active area of research, as a natural extension of data compression approaches.

Because of this interest, DARPA has initiated a program called Large Area Coverage Optical Search-while-Track and Engage (LACOSTE). The idea was to develop extremely wide field-of-view coded aperture imaging technology to support single-sensor day-night persistent tactical surveillance of all moving vehicles in a large urban battlefield. LACOSTE-coded aperture imaging technology focused on achieving a very wide instantaneous field of regard using multiple simultane-

[54]Emmanuel Candès, Justin Romberg, and Terence Tao. 2006. Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information. *IEEE Transactions on Information Theory* 52(2):489-509.

ous wide field-of-view (FOV) images.[55,56,57,58,59,60,61,62,63,64,65,66] In coded apertures, a structured mask of pinhole cameras is created such that the image from each individual pinhole falls across a common FPA. With a large-area mask centered above a small FPA, the pinhole camera structure opens as a lens in the desired look direction while the remainder of the mask remains opaque. Knowing the mask structure, the multiple images on the same focal plane are digitally deconvolved to form an image. This concept provides several unique and enabling features. Orthogonally coded masks allow multiple simultaneous images at arbitrary pointing angles. Grouping pinholes enables electronically switchable resolution. The coded aperture replaces optical gain with digital processing gain.

A critical enabling technology for LACOSTE was the adaptive pinhole mask

[55] A. Mahalanobis, C. Reyner, H. Patel, T. Haberfelde, David Brady, Mark Neifeld, B.V.K. Vijaya Kumar, and Stanley Rogers. 2007. IR performance study of an adaptive coded aperture “diffractive imaging” system employing MEMS “eyelid shutter” technologies. *Proceedings of the SPIE* 6714:67140D.

[56] A. Mahalanobis, C. Reyner, T. Haberfelde, M. Neifeld, and B.V.K. Vijaya Kumar. 2008. Recent developments in coded aperture multiplexed imaging systems. *Proceedings of the SPIE* 6978:6978G.

[57] A. Mahalanobis, M. Neifeld, B.V.K. Vijaya Kumar, and T. Haberfelde. 2008. Design and analysis of a coded aperture imaging system with engineered PSFs for wide field of view imaging. *Proceedings of the SPIE* 7096:7096C.

[58] A. Mahalanobis, M. Neifeld, B.V.K. Vijaya Kumar T. Haberfelde, and D. Brady. 2009. Off-axis sparse aperture imaging using phase optimization techniques for application in wide-area imaging systems. *Applied Optics* 48(28):5212-5224.

[59] C. Slinger, N. Gordon, K. Lewis, G. McDonald, M. McNie, D. Payne, K. Ridley, M. Strens, G. De Villiers, R. Wilson, and M. Eismann 2007. An investigation of the potential for the use of a high-resolution adaptive coded aperture system in the mid-wave infrared. *Proc SPIE* 6714:671408.

[60] M.E. McNie, D.K. Combes, G.W. Smith, N. Price, K.D. Ridley, K.M. Brunson, K.L. Lewis, C.W. Slinger, and S. Rogers. 2007. Reconfigurable mask for adaptive coded aperture imaging (ACAI) based on an addressable MOEMS microshutter array. *Proc SPIE* 6714:67140B.

[61] K. Lewis 2007. Challenges in the evolution of advanced imaging systems. *Proc SPIE* 6714:671402.

[62] M.E. McNie, D.O. King, N. Price, D.J. Combes, G.W. Smith, A.G. Brown, N.T. Gordon, S.M. Stone, K.M. Brunson, K.L. Lewis, C.W. Slinger, and S. Rogers. 2008. A large-area reconfigurable MOEMS microshutter array for a coded aperture imaging systems. *Proc SPIE* 7096:70960D.

[63] Chris W Slinger, Kevin Gilholm, Neil Gordon, Mark McNie, Doug Payne, Kevin Ridley, Malcolm Strens, Mike Todd, Geoff De Villiers, Philip Watson, Rebecca Wilson, Gavin Dyer, Mike Eismann, Joe Meola, and Stanley Rogers. 2008. Adaptive coded aperture imaging in the infrared: towards a practical implementation. *Proc SPIE* 7096:709609.

[64] G.D. de Villiers, N.T. Gordon, D.A. Payne, I.K. Proudler, I.D. Skidmore, K.D. Ridley, C.R. Bennett, R.A. Wilson, and C.W. Slinger. 2009. Subpixel superresolution by decoding frames from a reconfigurable coded-aperture camera: theory and experimental verification. *Proc SPIE* 7478:846806.

[65] K.D. Ridley, G.D. de Villiers, D.A. Payne, R.A. Wilson, and C.W. Slinger. 2009. Visible band lens-free imaging using coded aperture techniques. *Proc SPIE* 7478:746809.

[66] M.E. McNie, D.O. King, G.W. Smith, S.M. Stone, A.G. Brown, N.T. Gordon, C.W. Slinger, K. Cannon, S. Riches, and S. Rogers. 2009. A scalable multichip architecture to realise large-format microshutter arrays for coded aperture applications. *Proc SPIE* 7468:74780E.

where each pixel is physically small and addressable (either independently or in spatially distributed orthogonal groupings) and has millisecond switching speed, sufficient MWIR transmittance, low power consumption, large-area manufacturability, and extremely low cost per pixel. Another issue is compensating for dispersion effects while still using the pinholes, since there is a desire to use this "camera" in a broadband imager.

## Lensless Imaging

Most sensors detect an image by using optics to focus light onto a detector array. This is referred to as detecting in the image plane because you form an image in this plane conjugate to the object you are viewing. A lens can be used alternatively to take a Fourier transform of the input light.[67] This converts the input from the pupil plane to the image plane. If the phase and amplitude of the light wave are known in the pupil plane, they can be digitally converted to the image plane by taking a Fourier transform. This is the essence of so-called lensless imaging. The major difficulty with detecting in the pupil plane, and digitally converting to the image plane, is that detectors are sensitive to the intensity of the impinging light. The carrier frequency of light is orders-of-magnitude too high a bandwidth for existing detectors to follow. The fact that optical detectors do not detect the phase of the carrier can be mitigated by guessing the phase and using an image sharpness metric to iteratively find values of phase across the pupil plane that will provide the sharpest image.[68] For passive visible and IR sensors, this is the only option to obtain phases for lensless imaging because of the broadband nature of these imagers. Spatial variation in phase can be estimated using wavefront measurement devices, such as the Shack-Hartman device discussed above. Measurements by such a device can assist the initial estimate of phase.

To measure the phase of an optical wave, a second, coherent optical wave can be directed at the detector at the same time as the first wave. This is called heterodyne detection. If this is done, then phase and amplitude can be measured, so a Fourier transform can be taken. To measure the phase of the detected signal, the detector bandwidth must be high enough to measure the beat frequency between the local oscillator and the signal to be detected. This cannot be accomplished for most passive sensors because the bandwidth of the detected signal is much broader than the bandwidth of the detector. For narrow-band active sensors, heterodyne detection can be used. There is a technique called spatial heterodyne that allows the use of low

[67]Joseph Goodman. 2004. *Introduction to Fourier Optics.* Greenwood Village, Colo.: Roberts and Company Publishers.

[68]J.R. Fienup. 2006. Lensless coherent imaging by phase retrieval with an illumination pattern constraint. *Optics Express* 14(2):498-508.

bandwidth with a spatially offset local oscillator.[69] Spatial heterodyne measures the spatial variation of phase across an aperture. Temporal phase is not measured.

Lensless imaging may not be completely lensless. If the detector focal plane array is the size of the aperture being used, no optics will be required. It is however likely that for many applications the receive aperture will be larger than available FPAs. If the receive aperture is not the same size as the FPA to be used, a telescope will be required to adjust the size of the pupil plane image prior to detection.

Signal-to-noise and sampling considerations are different in the pupil plane than in an image plane. Bright point objects in the image plane are distributed across the full aperture in the pupil plane. For a diffraction-limited set of optics, the detector size in the image plane is set to optical diffraction limit. Oversampling occurs when the detector size is smaller than the diffraction limit. The field-of-view of a traditional image plane sensor is set by multiplying the pixel linear dimension by the number of detectors in that dimension. In the pupil plane, detector sampling is opposite. The maximum sensor FOV is set by the size of the detector, while the maximum resolution to be sampled is set by the full array width, the product of detector linear dimension times the number of detectors.

## IMPROVED COOLERS

### Thermoelectrics

Current commercial thermoelectric (TE) coolers can achieve a maximum $\Delta T$ of ~100-130 K below room temperature. An example of a current commercially available multistage TE cooler that achieves a maximum $\Delta T$ of ~133 K, as measured from room temperature, is Marlow Industries' MI6030-01BC shown in Figure 4-6. At this extreme temperature difference, however, the module cannot dissipate any heat (i.e., the coefficient of performance [COP] is essentially zero). Maximum COP for this device at DT = 0, as reported by Marlow Industries, is ~0.03.

In contrast, single-stage thermoelectric devices can achieve COP values approaching 0.6-1.0; however, the maximum $\Delta T$ for these devices is only ~70 K. Again, at $\Delta T_{max}$ no heat can be pumped, so TE coolers typically operate at cold side temperatures that are only a fraction of their maximum achievable $\Delta T$.

Existing commercial TE coolers have a market in providing temperature stability and noise reduction or dark current reduction for detector systems that operate at or near room temperature. Current available TE coolers do not have sufficient cooling capacity to be used for systems that require operation at temperatures below ~200 K. However, several recent developments in thermoelectric materials

[69] J.C. Marron and R.L. Kendrick. 2007. Distributed aperture active imaging. *Proceedings of the SPIE* 6550:65500A.

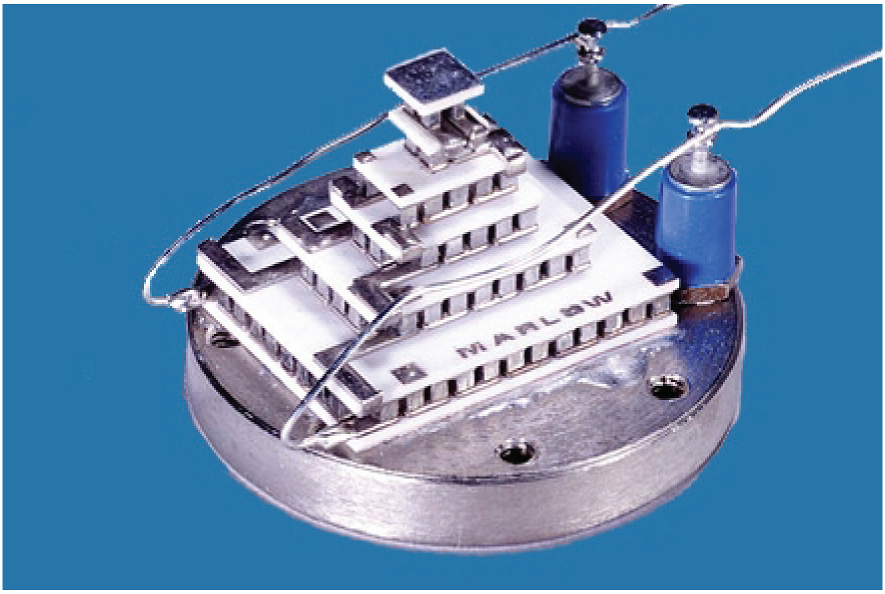


FIGURE 4-6
Marlow Industries' MI6030-01BC. NOTE: The base dimensions are 21.84 × 28.19 mm. The module height is 20.73 mm. SOURCE: Marlow Industries, Inc., Subsidiary of II-VI Incorporated, 10451 Vista Park Rd, Dallas, TX 75238. Image available at http://www.marlow.com/thermoelectric-modules/six-stage/mi6030.html. Accessed March 25, 2010.

research suggest the potential for breakthroughs that may ultimately transform the use of TE coolers in detector systems. These developments include the demonstration of thin-film thermoelectrics with performance that is two to three times greater than conventional bulk materials at or near room temperature and new understandings of how nanostructurer can improve thermoelectric performance at cryogenic temperatures.

In 2001, Rama Venkatasubramanian and his colleagues at Research Triangle Institute (RTI) reported a significant breakthrough in the achievable figure of merit, $ZT$, in a thin-film superlattice of alternating bismuth telluride and antimony telluride materials. Based on the performance demonstrated, it is projected (see Figure 4-7) that the achievable COP for a TE cooler could approach or even exceed that of small-scale mechanical systems. In addition, the thin-film format allows for a much smaller-profile device for comparable performance compared to existing TE coolers (see Figure 4-8).

The potential for thin-film TE coolers for use in current and future detectors that operate near room temperature remains speculative. There are challenges with manufacturing, reliability, cost, and scale-up that still need to be overcome.

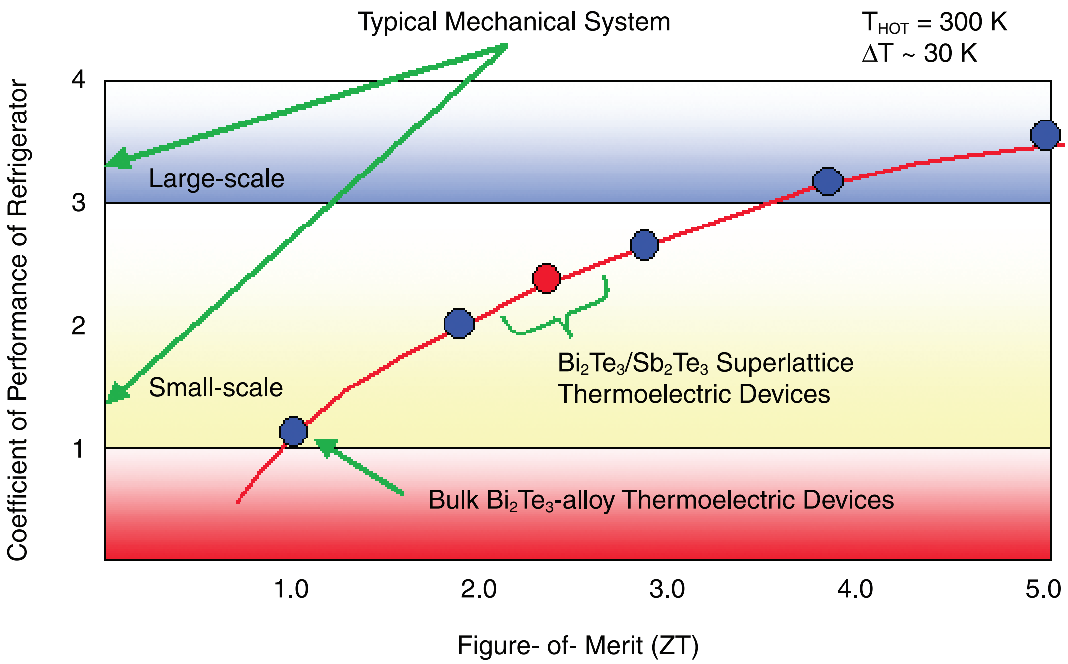


FIGURE 4-7
Achievable figure of merit, *ZT*, in a thin-film superlattice of alternating bismuth telluride-antimony telluride materials. SOURCE: Rama Venkatasubramanian, and colleagues, RTI International, Research Triangle Park, N.C.

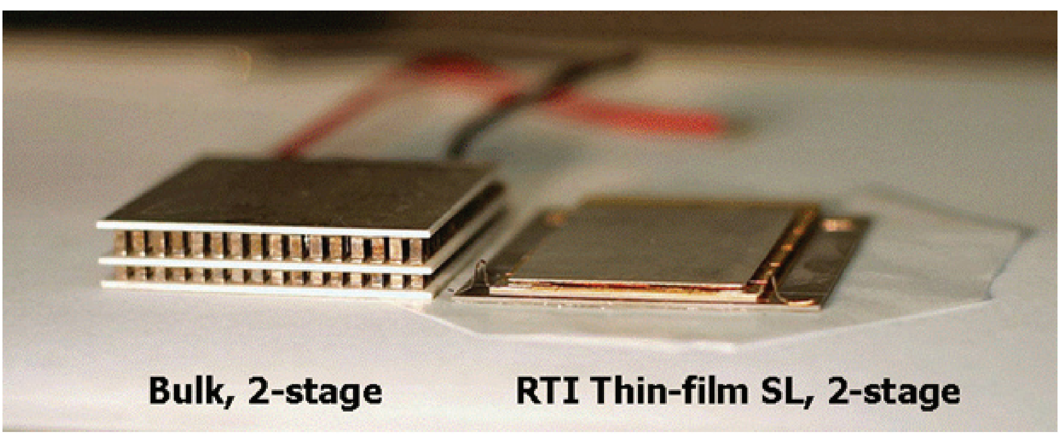


FIGURE 4-8
Advanced thin-film superlattice thermoelectric modules for FPA applications. NOTE: The bulk two-stage module—50 × 50 × 8 mm—versus the thin-film two-stage module—50 × 50 × 2.5 mm. The thin-film module has a factor of three advantage in module profile thickness and is significantly more lightweight, by a factor of 100 or more, for similar performance of heat pumping capacity. SOURCE: Rama Venkatasubramanian, and colleagues, RTI International, Research Triangle Park, N.C.

However, RTI has established a spinoff company, Nextreme Solutions, with the goal of commercializing its thin-film thermoelectric technology for a number of thermal management applications including hot spot management of microelectronic devices. According to Nextreme's website (www.nextreme.com), it is able to provide samples of its thin-film thermoelectric coolers for evaluation by potential customers. Target applications include laser diodes, photodetectors, polymerase chain reaction (PCR), and light-emitting diodes (LEDs). While it does not appear that these coolers have currently achieved large-volume commodity status, Nextreme does report commercial shipments of some of its cooler products as having been initiated in 2009. The next several years will be critical in determining the long-term viability of this technology. Nextreme's success or failure will likely determine the near-term opportunities for achieving significant breakthroughs in reduced size, weight, and power for the active cooling component of detectors that operate near room temperature.

A recent development reported up to a 100 percent enhancement in lead telluride compounds at 773 K by doping to create "resonant states."[70] This approach may be ideal for enhancing the thermoelectric performance of materials at cryogenic temperatures. If thermoelectric materials could be realized at cryogenic temperatures with *ZT* values in the range of 1-2, then it is conceivable that solid-state refrigerators could replace current mechanical cryocoolers. This could easily lead to an order-of-magnitude or more reduction in overall system weight and/or power requirements. The results are very preliminary, and there is comparatively little (compared to funding for TE power generation) funded research to specifically address thermoelectric development for cryogenic temperatures.

**FINDING 4-6**

Thin-film thermoelectric devices have the potential to substantially reduce size, weight, and power requirements of the active cooling component for room-temperature focal plane arrays. If these devices can meet cost and lifetime metrics, they will displace the currently used bulk coolers. The near-term driver for these developments likely will be in fields such as microelectronics with much larger market potential than detectors.

**RECOMMENDATION 4-1**

The intelligence community should monitor commercial developments in thin-film cooler technology.

[70] Josef Heremans, Vladimir Jovovic, Eric S. Toberer, Ali Saramat, Ken Kurosaki, Anek Charoenphakdee, Shinsuke Yamanaka, and G. Jeffrey Snyder. 2008. Enhancement of thermoelectric efficiency in PbTe by distortion of the electronic density of states. *Science* 321:554-557.

## Phononic Crystals for Cooling

Work is under way in the area of silicon phononic crystal-based thermoelectric active cooling modules. The device concept is based on planar phononic crystal topologies that have the potential to surpass silicon nanowires in thermoelectric performance from both practical and scientific points of view. An approach that would permit the hybrid use of nano-patterning, quantum confinement, and coherent and incoherent scattering mechanisms for phonons in a mass producible setting, while at the same time leaving the electrical properties of silicon unaltered, can be achieved via the so-called phononic crystal (PnC) lattice. A PnC is a periodic arrangement of Mie scattering centers in a host matrix. 2-D planar phononic crystals have some significant advantages over nanowire topologies. The flexible nature of the 2-D phononic crystal topology could allow one to uniquely tailor the electronic structure of the silicon to facilitate improved Seebeck coefficients over those of nanowire topologies. 2-D PnC topologies have the potential to surpass silicon nanowires in thermoelectric performance from both practical and scientific points of view. Due to their improved mechanical strength, their larger contactable areas, reduced phononic thermal transport, and somewhat tailorable Seebeck coefficient, a 2-D PnC approach has a number of distinct advantages for creating thermoelectrics in a silicon platform. In addition to nanometer-geometry boundary scattering, PnC topologies will exhibit coherent Bragg scattering and reductions in group velocity that reduce phonon mobility beyond boundary scattering alone. Additionally, the flexible nature of the 2-D phononic crystal topology could allow one to uniquely tailor the electronic structure of the silicon to facilitate improved Seebeck coefficients over those of nanowire topologies.[71,72]

## Laser Cooling

Laser cooling refers to one of several mechanisms by which laser light interacts with matter in a way that results in a net reduction in temperature. The reduction in temperature is associated with absorption and emission processes that result in a net loss of momentum at the atomic or molecular level. The most common form of laser cooling is Doppler cooling. Doppler cooling is a technique used to cool, low-density elemental gases (Cs, Rb, Na, etc.) to extremely low temperatures. The original motivation for developing techniques for cooling gases to temperatures of

[71]Patrick E. Hopkins, Peter T. Rakich, Roy H. Olsson, Ihab F. El-kady, and Leslie M. Phinney. 2009. Origin of reduction in phonon thermal conductivity of microporous solids. *Applied Physics Letters* 95:161902.

[72]Roy Olsson and Ihab El-kady, Sandia National Laboratories, Albuquerque, N.M., personal communications with the committee on March 9, 2010.

hundreds of microkelvins or less was to improve the accuracy of atomic clocks as well as to explore the quantum mechanical nature of atomic physics.

Doppler cooling is observed when an atom is moving towards a laser light source. Under certain conditions, the atom can preferentially absorb photons with an energy that is slightly below a particular atomic resonance. In absorbing these photons, conservation of momentum requires the atom to slow its momentum in the direction opposite to the incident photon. As the atom relaxes back to its lower state, a photon is emitted in a random direction. The randomization of the atom's momentum results in a net slowing of the atom's motion and, therefore, a lowering of temperature.

The demonstration that Doppler cooling could successfully be used to achieve sub-millikelvin temperatures in elemental gases was acknowledged with a Nobel Prize in 1997 that was shared among Steven Chu, Claude Cohen-Tannoudji, and William D. Phillips.

Doppler cooling continues to be a useful research tool for exploring quantum physics and has found practical use in advanced inertial navigational systems. However, it is unlikely that this technology would be of value for sensor applications. It is a technique best suited to cooling gases as opposed to solids, and the achievable temperatures are well below the operating temperatures required for detector systems.

More recently, laser cooling based on the physics of anti-Stokes fluorescence has been demonstrated to result in substantial cooling of certain rare earth materials. In anti-Stokes fluorescence cooling, a solid absorbs photons at a particular frequency and then reemits at a higher frequency. The frequency shift results in a net loss of energy in the form of heat from lattice vibrations (phonons). This effect was first reported in 1995 by researchers from Los Alamos who observed a 0.3 K reduction in temperature in $Yb^{3+}$-doped fluorozirconate glass.[73] More recently, Seletskiy et al. have reported a record 155 K temperature reduction in ytterbium-doped $LiYF_4$ using laser cooling.[74] This exceeds the $\Delta T$ achievable using current state-of-the-art solid-state thermoelectric cooling technologies and represents a promising breakthrough for achieving a practical laser cooling capability.

This research is still at a very early stage, and there are few data on which to base a prediction of achievable coefficient of performance in a practical device. The technology would have the benefit of being solid state (no moving parts), so it could be of benefit to future detector systems where noise due to vibration is of concern. However, understanding the potential for achieving greater $\Delta T$ and reasonable coefficient of performance remains a significant challenge for this technology, which is still a very new and immature field of research.

[73] R. Epstein et al. 1995. *Nature* 377, 500-503.

[74] D.V. Seletskiy, S.D. Melgaard, S. Bigotta, A. Di Lieto, M. Tonelli, and M. Sheik-Bahae. 2010. Laser cooling of solids to cryogenic temperatures. *Nature Photonics* 4:161-164.

## ENHANCED SIGNAL PROCESSING

Key technology advancements in electronics and electronic devices have allowed for a previously unseen level of systems integration and miniaturization in IR sensors. Advancements in analog-to-digital converters, field-programmable gate arrays, and very small outline packaging have allowed IR systems to go from large multicard line-replacable units (LRUs), a complex component of a vehicle that can be replaced quickly at the organizational level, to battery-operated helmet-mounted devices over the last 15 years.

As shown in Figure 4-9, processing technologies that support both data and information transmission, as well as information and knowledge extraction, will be vital to future detector technologies. When a technological concept is reduced to practice, various design factors are traded against each other to achieve a new capability within constraints imposed by other elements of the system or to achieve a new capability as a consequence of a breakthrough removing a constraint elsewhere in the system. As a specific example of a constraint, remote sensing architectures will inevitably be constrained by the communications capabilities available to move data from the sensor system to processing or archival sites. On the other hand, the emergence of single-chip multicore processor architectures may allow significant local pre-processing of sensor data before they are transmitted from the remote sensor, reducing the bandwidth required and allowing the use of sensors with improved precision (e.g., the number of pixels in a focal plane array) or higher scan rates, each of which may be valuable in some applications.

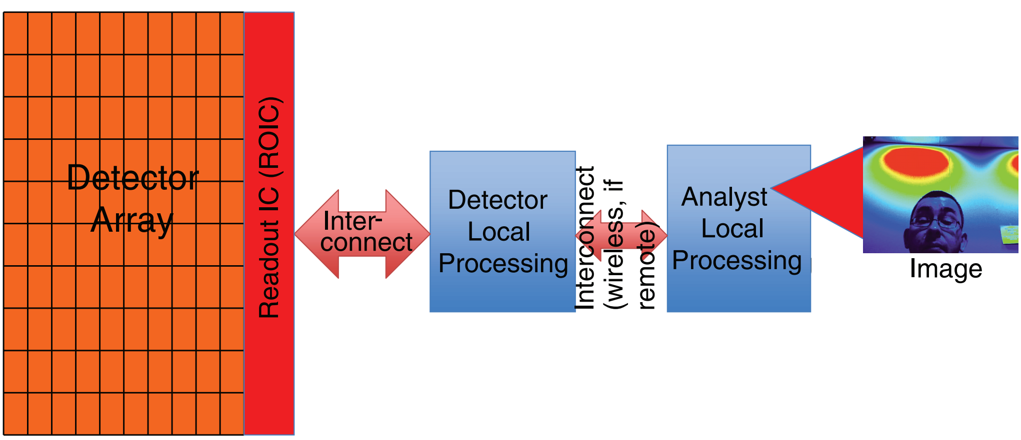


FIGURE 4-9
A high-level model that demonstrates the methodology and constraints involved with data and information transmission and information and knowledge extraction.

## Data and Information Transmission

The major challenges facing data and information transmission are primarily at remote collector or sensing platforms. These challenges are distinct for close-in platforms, such as unattended ground sensors (UGS) and remote platforms (such as airborne and space). Both close-in and remote platforms have limited communication bandwidth; however, close-in platforms have additional severe constraints on power consumption. Once the raw data or derived information have reached the ground processing station, there are fewer limitations in terms of power requirements, computer power, and communication bandwidth. The technologies that address data transmission limitations on communication bandwidth belong to two major classes: (1) data compression and (2) data screening.

### *Data Compression*

***Lossless Compression Techniques*** A number of methods are aimed at achieving lossless compression;[75,76,77,78,79] however, lossless compression can only a achieve a modest degree of compression. For example, an ARGUS-IS-like system can produce up to 770 gigabits per second. The use of a Common Data Link (CDL) operating at 274 megabits per second would require compression ratios on the order of 2,800, far beyond the capabilities of lossless compression techniques.

***Lossy Compression Techniques*** Most lossy compression[80] methods include three major steps: (1) signal decomposition-transformation, (2) quantization-thresholding, and (3) encoding. The first step is focused on transforming the signal into a representation that can be compressed more efficiently by reducing dynamic range, removing redundancy, among other things. This step is also taken by many lossless techniques and is reversible. Commonly-used transforms include discrete cosine transform (DCT)[81] and discrete wavelet transform (DWT).[82] The quantization step that follows the signal transformation is the step that reduces the number of output symbols and is the source of information loss and also encoding-

[75]Available at http://en.wikipedia.org/wiki/Run-length_encoding.

[76]Available at http://en.wikipedia.org/wiki/DPCM.

[77]Available at http://en.wikipedia.org/wiki/ Predictive_Coding.

[78]Available at http://en.wikipedia.org/wiki/Entropy_encoding.

[79]Available at http://en.wikipedia.org/wiki/LZW.

[80]Lossy compression reduces the size of the file required to store information but does not save all of the original information.

[81]N. Ahmed, T. Natarajan, and K.R. Rao. 1974. Discrete cosine transform. *IEEE Transactions on Computers* C-23:90-93.

[82]M. Vetterli and J. Kovacevic. 1995. *J. Wavelets and Subband Coding.* Englewood Cliffs, N.J.: Prentice Hall.

compression efficiency. The last step of encoding is aimed at achieving rates that approach the entropy of the quantized values or symbols.[83]

***Emerging Compression Techniques*** Compressive sensing or sampling is an emerging technology that is surfacing in various implementations ranging from a direct application of image compression to new sensors that embed the concept of compressed sensing at the analog-optical layer.[84] Compressive sensing or sampling relies on two basic principles, sparsity and incoherence. The basic assertion is that many signals are sparse, meaning that they have a very compact representation if the basis functions are chosen properly. Incoherence is an extended case of the duality concept between time and frequency domains, sparse in time and spread in frequency. If the signal has a compact representation in some basis, the sampling-sensing waveforms have a very dense representation. The implication is that we can design efficient sampling that captures the useful information embedded in the signal and transforms it into a small amount of data. Large upfront data reduction directly translates to a reduced requirement for communication bandwidth. The simplicity of the sampling stage is obtained at the cost of a complex reconstruction stage that requires the application of computing-intensive linear programming techniques.

The same concepts are also being explored in the design of a new generation of imagers that employ compressive sensing at the optical-analog layer. In this new concept of sensing a smaller number of measurements is acquired, and each measurement corresponds to a quantity in the transformed space that is optically accomplished as the inner product of the scene with the basis functions. Because of the significant reduction of sensor components and elements (without sacrifice in performance), these new imagers likely will be deployed on many more platforms than currently feasible.

[83]Entropy is considered a standard measure of complexity. It is a property of a distribution over a discrete set of symbols. For a sequence $\{i\}$ of symbols $x$ drawn from an alphabet with a probability $p(i)$, the entropy $H(i)$ of the random variable $I$ is given *by* $H(i) = S\,p(i)\log 2\,p(i)$. The units of entropy and entropy rate defined above are bits and bits per symbol. The entropy of a sequence is the length of shortest binary description of the states of the random variable that generates the sequence, so it is the size of the most compressed description of the sequence. For additional information, see S. Lloyd. 1990. Physical measures of complexity. In E. Jen, ed. *1989 Lectures in Complex Systems, SFI Studies in the Sciences of Complexity*, Vol. II, pp. 67-73. Addison-Wesley; Claude E. Shannon. 1948. A mathematical theory of communication, *Bell System Technical Journal* 27:379-423, 623-656. There is a related measure of complexity, called Kolmogorov complexity, that measures the size of the smallest program necessary to produce an output. For additional information, see http://en.wikipedia.org/wiki/Entropy_(information_theory. Last accessed June 21, 2010.

[84]Emmanuel J. Candès and Michael B. Wakin. 2008. An introduction to compressive sampling. *IEEE Signal Processing Magazine* 13(March).

*Data Screening Techniques*

The basic principle behind the application of data screening methods is to filter (screen) the raw data and extract only the data segments that are of interest. This is called relevance filtering in other contexts. Example screening techniques include automatic target detection and/or recognition and material identification, among others. In principle, the use of data screening techniques as a compression mechanism amounts to shifting the specific mission processing algorithms from the receiving ground station to the collection platform. A number of approaches can be taken once the image regions of interest have been identified. One approach is to transmit only the data associated with the regions of interest (with no loss of fidelity) and discard the remaining data. However one might choose to further reduce the volume of data either through lossy compression of the selected data or the use of derived attributes such as the identification and location of the detected targets. This compression method is typically used in unmanned ground sensors, as well as on some airborne platforms.

## Application-specific Processing

The signal processing and exploitation chain consists of two major components; (1) conversion of sensor input into physical meaningful values that then can be further processed by (2) mission- or application-specific algorithms (e.g., missile detection and tracking, target detection, automatic target recognition, and material identification). In this report the primary focus is on the core signal processing chain that is accomplished on the collection platform and is essential for high-quality image acquisition, as shown in Figure 4-10.

In the IR and high-performance visible imaging area, fully digital focal planes are just entering the market. In some cases their performance exceeds that of traditional analog focal planes coupled to discrete electronics. On-ROIC digital logic enables future digital signal processing such as nonuniformity correction (NUC), image stabilization, and compression leading to much smaller systems (microsystems). This technology is poised to go into large-area cryogenic infrared sensors over the next few years.

## Local Processing

Sending data off-chip requires substantial power, and sending those data through communication links (for example, an RF link for an unattended ground sensor) is even more energy intensive. As described above, most of the basic functionality depicted in Figure 4-10 such as AGC (automatic gain control) and TDI time delay integration is currently embedded and implemented as an integral part of the ROIC functionality. There is increasing recognition of the value of trying to

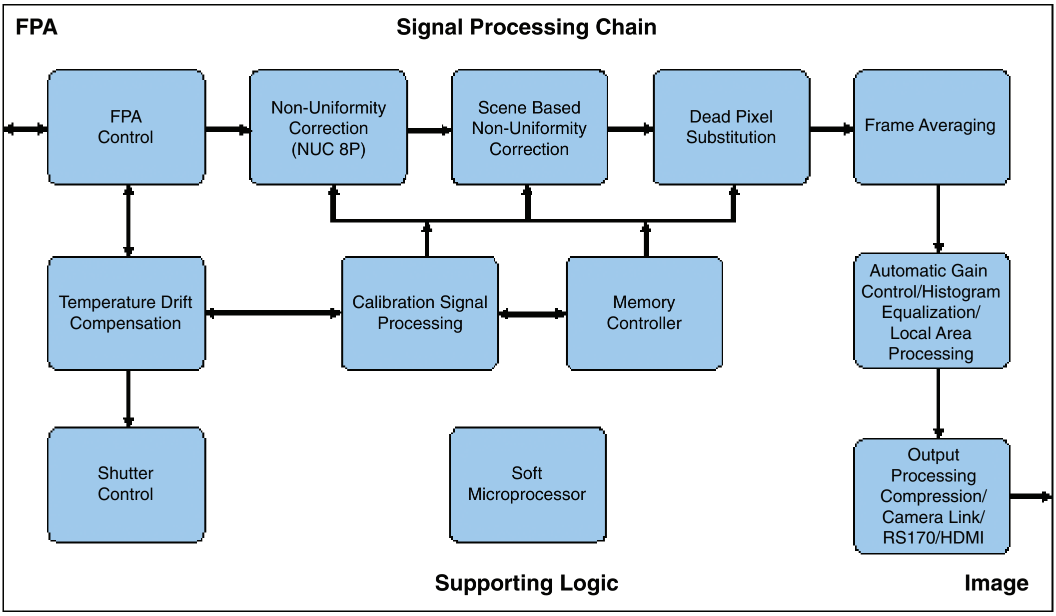


FIGURE 4-10
Components of a signal processing architecture.

identify and transmit only small amounts of high-level "actionable information" rather than large numbers of raw data bits.

Architectural bottlenecks occur when steps in digital signal processing are mismatched in their performance. The goal of later processing steps should be to extract as much information as possible from this sensing infrastructure. In practice this means processing to remove noise, to obtain true information, and to lose as little information as possible. A bottleneck in a system can be viewed as a lossy filter. In the case of a parallel front end, it will have a number of bits of attributes, such as amplitudes, but perhaps frequencies and phases as well, available to late processing stages. For example, for a 10-megapixel array, each pixel might produce 32 bits of information per sample at a sampling rate of 30 frames per second, leading to an aggregate "bit rate" of $9.6 \times 10^9$, or about 1.2 gigabytes, per second. More information per sample, more pixels, or a greater scan rate has a significant effect on the processing and communication demands of such a sensor.

As a specific example, the DARPA ARGUS-IS unmanned aerial system described in Chapter 3 (Boxes 3-1 and 3-2) contains 368 visible FPAs.[85] At data rates

[85]Brian Leininger, Jonathan Edwards, John Antoniades, David Chester, Dan Haas, Eric Liu, Mark

of 96 megapixels per second per FPA, 12 bits per pixel, and 368 FPAs, the total data rate from the sensor is about 424 gigabits per second. This data rate is beyond the capacities of conventional processing elements. In addition, ARGUS-IS uses a spread-spectrum jam-resistant CDL wireless data link of 274 megabit per second capacity. If the wireless data link is fully utilized, the on-board systems must achieve a data rate reduction of 423,936/274 or more than 1,500, which is difficult to achieve with compression technologies alone.

ARGUS-IS approaches the data management challenge with a novel on-board processing architecture, characterized by parallel interconnects. Each FPA pair feeds a field-programmable gate array (FPGA) that multiplexes the data from the two FPAs (a total of 2.3 gigabits per second), time tags these data, and interleaves the data onto a fiber-optic transceiver. The transceiver is a commercial off-the-shelf (COTS) device operating at 3.3 gigabits per second. Sixteen 12-fiber ribbon cables connect the 184 FPGA pairs to the ARGUS-IS airborne processor system, a multiprocessor system illustrated in Box 3-1, which consists of 32 processor modules. Each processor module can handle 6 fibers, or about 20 gigabits per second of data, and consists of two Xilinx Virtex 5 FPGAs and a low-power Intel Pentium general-purpose central processing unit (CPU).

The ARGUS-IS designers believe that the processor modules can provide more than 500 billion operations per second each, for a total processing capacity for the 32 processor modules in excess of 16 trillion operations per second. To overcome some of the data rate limitations of the CDL downlink, JPEG 2000 compression is done using application-specific integrated circuits (ASICs) to provide hardware assist. The ARGUS-IS designers note the severe limitations of the 200+ megabit per second data link and propose moving target tracking into the on-board software. ARGUS-IS illustrates many of the system architecture trade-offs discussed in this chapter.

As the demand is increasing for on-board processing functionality that mirrors what traditionally has been accomplished on the ground, the need for compute power that meets size, weight, and power (SWaP) constraints is continuously growing. The amount of processing that can be placed right at the pixel has generally been limited by the modest numbers of transistors that can be placed within a small pixel.

Significant advances have been made in computer architectures and are currently referred to as high-performance computing. Commercial applications such as video gaming, increased cell phone functionality, and so forth, have been pushing significant advances in small, low-power, high-performance computing platforms

---

Stevens, Charlie Gershfield, Mike Braun, James D. Targove, Steve Wein, Paul Brewer, Donald G. Madden, and Khurram Hassan Shafique. 2008. Autonomous Real-time Ground Ubiquitous Surveillance—Imaging System (ARGUS-IS). *Proceedings of the SPIE* 6981:69810H-1.

that are available on the commercial market (for example, multicore processors and multicore graphics processing units [GPUs]) and have no International Traffic in Arms Regulations (ITAR) restrictions. Figure 4-11 depicts a Texas Instruments system-on-the-chip solution for cell phone applications. The significant advances in high-performance computing over the last decade are making real-time on-board processing a reachable reality.

However, the complexity of programming these powerful processors and achieving their potential computer power has significantly increased—both the exploitation of the available parallelism and the memory organization of the computation are subtle and require significant effort. A key enabler for exploiting this emerging computational power is the new sets of software tools that enable rapid porting and debugging of existing algorithms into multicore computing platforms. As discussed in the following two examples, extensive sets of dedicated software tools are emerging in support of the hardware platforms.

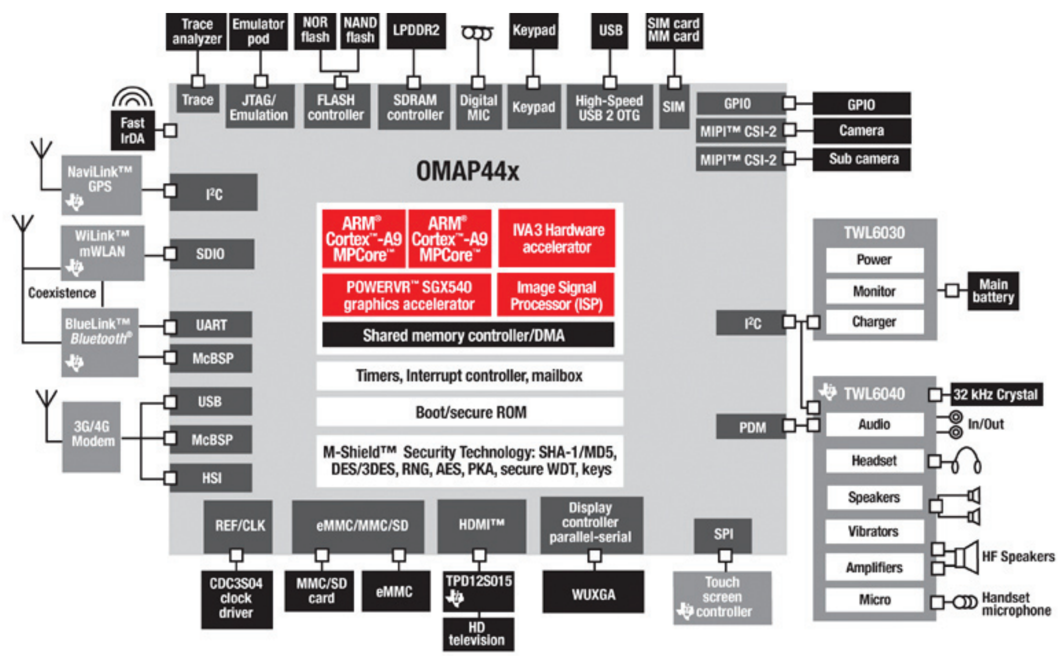


FIGURE 4-11
Texas Instruments OMAP 4430 system on a chip. SOURCE: Courtesy Texas Instruments. Available at http://focus.ti.com/general/docs/wtbu/wtbuproductcontent.tsp?templateId=6123&navigationId=12843&contentId=53243. Last Accessed March 25, 2010.

NVIDIA multicore technology is one example of commercially available high-performance computing technologies. In 2010, NVIDIA announced the launch of its next generation Tegra, a multiprocessing system focused on the mobile web.[86] The Tegra has eight independent processors, including a dual-core CPU for mobile applications. These processors are used together or independently to optimize power usage.

At another point on the performance scale, NVIDIA has also introduced the capability to perform petascale computing with teraflop processors. The NVIDIA Tesla C1060 computing processor provides energy-efficient parallel computing power by bringing the performance of a small cluster to a workstation environment. The CUDA programming environment ("C" programming language, enhanced with thread management primitives to inform the runtime for the GPU of what can be executed concurrently) simplifies the development of applications for its multiple 240-core processors.

Another example that has a significantly lower power profile is the series of multicore chips and boards provided by Tilera. The first generation Tilera TILE64 commercial chip is laid out as an 8 tile × 8 tile interconnect using a pipe-lined and programmable two-dimensional, proprietary, high-performance, low-latency mesh. The mesh can transport streams of data, memory blocks, or scalar values, adding to the flexibility of the programming models available for the chip. In addition, this approach can support arbitrary numbers of tiles, so that 8 × 8 in this initial chip is not a fixed configuration. This makes the architecture particularly suitable for radiation hard applications where a different number of tiles per chip may be required for power dissipation and several other technical reasons. The switch engine in each tile completely offloads the tile processing engine from iMesh™ network routing and protocol handling, and provides buffering and flow control so that tiles can perform processing in an asynchronous manner. Each network link is full duplex. The dynamic networks are routed in a tile-layout fashion (*x*-direction first, then *y*-direction). A RadHard by design chip (MAESTRO) is currently being developed under the National Reconnaissance Office-funded OPERA program.

The third generation of Tilera Corporation's multicore processor is aimed at delivering the highest available general-purpose compute at the lowest power consumption. The TILE100 will provide a 4× to 8× increase in performance over Tilera's current TILE*Pro*64 processor and will double the performance-per-watt metric. Table 4-2 provides a comparison between the Tile64, the Tile100, and the MAESTRO processors.

[86] 2010 International Consumer Electronics Show (CES), Las Vegas, Nevada.

TABLE 4-2 Comparison Between the Tile64, the Tile100, and the RadHard by Design Processors

| Performance | TILE*Pro*64 | Tile100 (Greylock) | Maestro (RHBD) |
|---|---|---|---|
| Number of cores | 64 | 100 | 49 |
| Temperature range | 0°C-70°C | 0°C-70°C | –55°C-125°C |
| Foundry | TSMC | TSMC | IBM |
| Feature size | 90 nm | 40 nm | 90 nm |
| On-chip cache (MB) | 5.6 | 32 | 4.3 |
| Floating point operations (GFLOPS) | ~10 (software emulated) | 29 (FPU accelerator) | 22 (IEE 754 FPU per each core) |
| On-chip bandwidth (Tbit/s) | 38 | 232 | 16 |
| Clock speed (MHz) | 700, 866 | 1250, 1750 | 300 |
| Typical power (W) | 27 | <50 (estimated) | 18 (estimated) |
| Total I/O bandwidth (Gbps) | 40 | 44 | 40 |
| Ethernet bandwidth | 2 XAUI, 2GbE | 2 XAUI, 2GbE | 4 XAUI, 4GbE |

NOTE: FPU = floating point unit; I/O = input-output.
SOURCE: Data derived from http://www.tilera.com/products/TILEPro64.php; http://www.tilera.com/products/TILE-Gx.php; http://nepp.nasa.gov/mapld_2009/talks/083109_Monday/03_Malone_Michael_mapld09_pres_1.pdf. Accessed March 25, 2010.

Tilera provides a Multicore Development Environment (MDE). The MDE tool kit provides a complete Integrated Development Environment (IDE).[87]

**FINDING 4-7**

Scaling the data throughput of focal plane sensor systems involves not only the sensor chip but also the detector-processor interface, signal processing and compression, and the communication link (wireless for remote air- and space-borne missions). Advanced compression and filtering with on-board processing provided by commodity multicore architectures are reducing communications demands.

[87]Standard Eclipse-based IDE; ANSI standard C compiler (see Section 6.0) and C++ compiler; Multi-tile cycle-accurate simulator; Whole chip debug and performance analysis; Complete SMP Linux 2.6 environment—standard runtime environment and command line tools; ILib library for efficient intertile communications; PCIe hardware development platform support; Linux and Windows host environments.

**RECOMMENDATION 4-2**
Analyses of national capabilities should include consideration of advances in processing technologies for other uses—for example, commercial developments—that could also enhance the use of detectors in future sensor systems.

## Multisensor Data Fusion

Performance enhancements can be achieved when combining data collected with IR sensors with additional sensor modalities. Multisensor data fusion[88,89,90] has been an evolving set of architectures and algorithms for drawing inferences from a multiplicity of sensors used in combination. Any given multisensory system will have an architecture that defines its ultimate capabilities, and then a set of algorithms will be used to draw the necessary inferences from the fused data. Algorithms for sensor fusion include techniques such as Bayesian inference,[91] Dempster-Shafer [92,93] evidential reasoning, and voting. The specific techniques depend on both the mission or application and the specific sensor modalities. Sample applications include image enhancement for improved navigation in low-light conditions, target detection, and target recognition. An example is the fusion of data from a pulsed radar and an IR sensor, as shown in Figure 4-12.[94] The radar can determine range, but not angular direction, while the forward-looking infrared (FLIR) can determine angular direction, but not range. By fusing the two sensor modalities, both range and angular direction can be determined.

Three specific methods for sensor fusion include raw data fusion, feature-level fusion, and decision-level fusion. The highest level of fusion occurs when the multiple images are combined into a single multivalue image that is then exploited. This requires accurate alignment of sensor measurements across the multiple sensors resulting in a vector of multimodal measurements associated with a common ground location. The combined multivalue image is then processed by an application-specific algorithm, such as automatic target cuing-recognition

[88]David L. Hall and James Llinas. 1997. An introduction to multisensory data fusion. *Proc. IEEE* 85(1).

[89]David L. Hall and Sonya A.H. McMullen. 2004. *Mathematical Techniques in Multisensor Data Fusion*, 2nd Edition. Artech, ISBN 978-15805333355.

[90]Lawrence A. Klein. 2004. *Sensor and Data Fusion: A Tool for Information Assessment and Decision Making*. SPIE ISBN 978-0819454355.

[91]Lawrence A. Klein. 2004. *Sensor and Data Fusion: A Tool for Information Assessment and Decision Making*. SPIE ISBN 978-0819454355.

[92]A.P. Dempster 1968. Generalization of Bayesian inference. *J. Royal Statistical Society* 30:205-247.

[93]G. Shafer. 1976. *A Mathematical Theory of Evidence*. Princeton, N.J.: Princeton University Press.

[94]David L. Hall and James Llinas. 1997. An introduction to multisensory data fusion. *Proc. IEEE* 85(1).

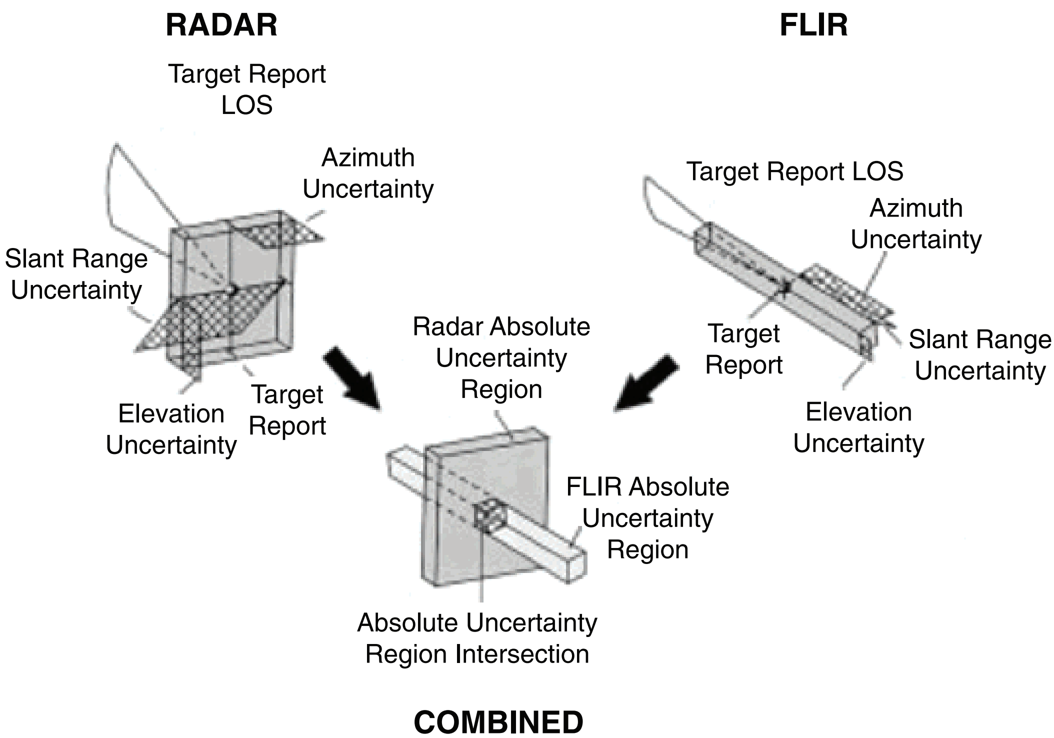


FIGURE 4-12
Fusion of data from a pulsed radar and an IR sensor. SOURCE: David L. Hall and James Llinas. 1997. An Introduction to multisensory data fusion. *Proc. IEEE* 85(1).

(ATC/R), that simultaneously operates on the vector of measurements. Fusion techniques that operate in this mode are known as raw data fusion as well as centralized data fusion methods, and they typically assume a common image projection plane for the multiple sensors. Multi- and hyperspectral sensors are well matched for this type of fusion approach.[95,96] Centralized data fusion is not typically applied to sensors that do not share a common imaging plane such as synthetic aperture radar (SAR).[97]

[95]G. Shafer. 1976. *A Mathematical Theory of Evidence.* Princeton, N.J.: Princeton University Press.

[96]Tamar Peli, Ken Ellis, Robert Stahl, and Eli Peli. 1999. Integrated color coding and monochrome multi-spectral fusion. In *Detection and Ccountermeasures: Infrared Detection and Detectors Conference.*

[97]M. Aguilar, D.A. Fay, W.D. Ross, A.M. Waxman, D.B. Ireland, and J.P. Racamato. 1998. Real-time fusion of low-light CCD and uncooled IR imagery for color night vision. *SPIE* 3364.

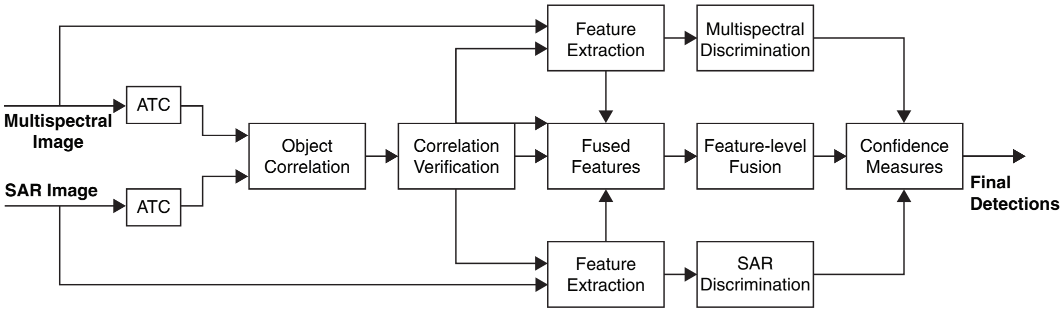


FIGURE 4-13
Improved ATC through feature-level fusion of SAR and multispectral imagery. SOURCE: T. Peli, M. Young, R. Knox, Ellis, K., and F. Bennett. 1999. Feature level fusion. *Proc. SPIE Sensor Fusion: Architecture, Algorithms and Applications III*, Vol. 3719. Orlando, Fla.

In contrast to raw data fusion, feature- and decision-level fusion methods do not require precise alignment. In a decision-level fusion, also known as distributed data fusion, an independent decision is made based on a single data modality and the decisions are passed to the fusion node where a global decision is made using a variety of algorithms including Bayesian inferencing. Feature-level fusion is a hybrid between raw data fusion and decision-level fusion. In feature-level fusion,[98] each sensor output is processed independently, and attributes associated with events or entities of interest that have been extracted in each sensor domain are combined and a decision is made based on the joint feature set. Figure 4-13 depicts improved ATC/R through feature level fusion of SAR and multispectral imagery.

## CONCLUDING THOUGHTS

Emerging technologies could enable new capabilities, with examples being the ability of some advanced detector technologies to enable multispectral sensing through a single aperture and the potential advantages in selectivity accruing from nanophotonics. Advances in non-detector-specific technologies, such as communications technology and signal processing technologies, have a direct bearing on the ability to turn detector data into useful information. As an example, the effects of miniaturization and parallel processing have made commodity components

[98]A.M. Waxman, M. Aguilar, R.A. Baxter, D.A. Fay, D.B. Ireland, J.P. Racamoto, and W.D. Ross. 1998. Opponent-color fusion of multi-sensor imagery: visible, IR and SAR. *Proceedings of the 1998 Conference of the IRIS Specialty Group on Passive Sensors.*

sufficiently powerful to provide onboard processing capabilities for a gigapixel capability.

Both the basic science underlying advances in detectors, optics, coolers, and algorithms and the commodity processing capabilities enabling new trades in signal processing are available worldwide. Emerging detector and related system-level technologies have significant potential for advancing sensor systems and deserve attention from the intelligence community in assessing present and future global sensor system capabilities.

# 5

# The Global Landscape of Detector Technologies

## INTRODUCTION

Previous chapters have established the military context and technical bases for addressing evolutionary and emerging, perhaps "breakthrough," technologies relating to detectors as well as the collection of components and subsystems that constitute a full sensor system. This chapter takes a global view of detectors, sensors, and sensor systems and addresses forces that drive detector development and tend to encourage development more in certain areas of the world than in others. Specific topics are worldwide leaders: government roles, markets, and scale; U.S. export restrictions; and supply-chain bottlenecks. Additional considerations and concluding remarks complete the chapter.

The following observations provide a context for the discussions. First, although evidence may emerge of an advance in some aspect of detector- or sensor system-related technologies, an entity's ability to mature a new technology to producible and deployable states is the ultimate determinant of the utility of that advance. Second, global commercial competition involving detector technologies and sensor systems has become significant.

## WORLDWIDE LEADERS

Several countries are actively involved in the development of infrared (IR) detector technologies. Below is a graphical depiction of open-source publications by country of origin and decade from 1980 to 2010. The data were obtained through a search using Compendex and National Technical Information Service (NTIS) databases. The information contained in Figure 5-1 shows trends in technology interests by country and geographical region. The identification is by the institutional affiliations of the authors. By comparing the number of publications that have emerged from various countries over the last 30 years, one can see acceleration in research reporting by most of the countries shown in Figure 5-1. While the United States maintained dominance, the contributions from the People's Republic of China more than doubled during the last decade, to about 12 percent of the total.

The Web of Science has a different list of open-source publications from which to draw. Web of Science is a citation database with multidisciplinary coverage of more than 10,000 high-impact journals in the sciences, social sciences, and arts and humanities, as well as coverage of international proceedings for more than 120,000 conferences. Figures 5.2 through 5.5 reflect the countries whose papers are drawn from this database. The figures comprise both a 30-year and a 10-year look back and were generated using the search criteria "infrared + detect*." Figure 5-2 is the 30-year compilation. Figures 5-3 through 5-5 are broken down in 10-year periods to show evolving trends.

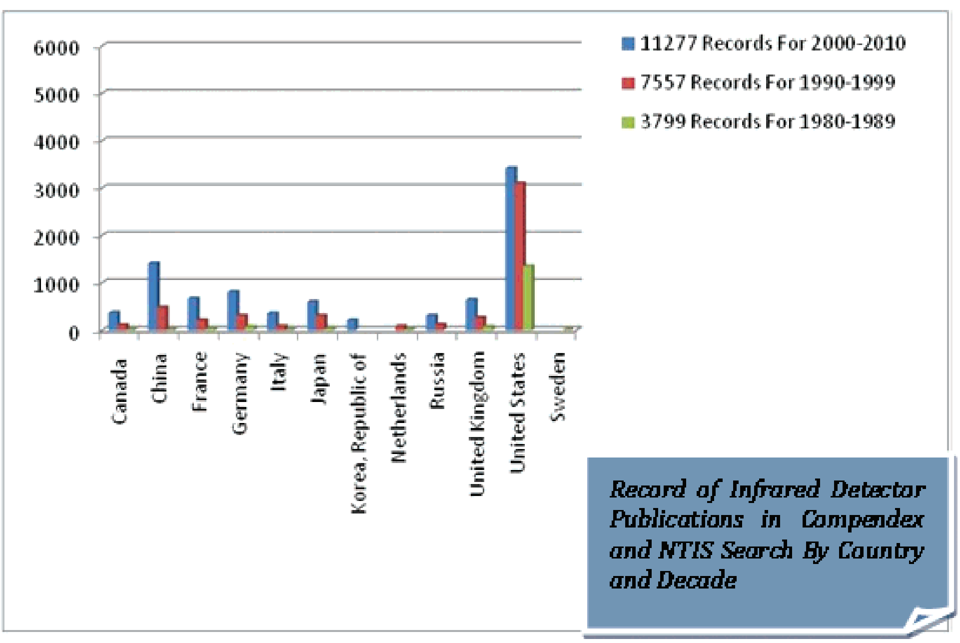


FIGURE 5-1
Illustrative global infrared detection publication activities.

| View Records / Exclude Records | Field: Country/Territory | Record Count | % of 52903 | Bar Chart | Save Analysis Data to File |
|---|---|---|---|---|---|
| ☐ | USA | 20318 | 38.4061 % | | |
| ☐ | GERMANY | 6102 | 11.5343 % | | |
| ☐ | FRANCE | 4378 | 8.2755 % | | |
| ☐ | JAPAN | 4189 | 7.9183 % | | |
| ☐ | PEOPLES R CHINA | 3921 | 7.4117 % | | |
| ☐ | ENGLAND | 3690 | 6.9750 % | | |
| ☐ | ITALY | 2845 | 5.3778 % | | |
| ☐ | SPAIN | 2348 | 4.4383 % | | |
| ☐ | CANADA | 2233 | 4.2209 % | | |
| ☐ | NETHERLANDS | 1621 | 3.0641 % | | |
| View Records / Exclude Records | Field: Country/Territory | Record Count | % of 52903 | Bar Chart | Save Analysis Data to File |

(134 Country/Territory value(s) outside display options.)

FIGURE 5-2
Publications on IR detector technologies from 1980 to 2010 (52,903 results).

| View Records / Exclude Records | Field: Country/Territory | Record Count | % of 719 | Bar Chart | Save Analysis Data to File |
|---|---|---|---|---|---|
| ☐ | USA | 364 | 50.6259 % | | |
| ☐ | FED REP GER | 60 | 8.3449 % | | |
| ☐ | ENGLAND | 56 | 7.7886 % | | |
| ☐ | JAPAN | 55 | 7.6495 % | | |
| ☐ | FRANCE | 40 | 5.5633 % | | |
| ☐ | CANADA | 35 | 4.8679 % | | |
| ☐ | USSR | 22 | 3.0598 % | | |
| ☐ | AUSTRALIA | 14 | 1.9471 % | | |
| ☐ | INDIA | 11 | 1.5299 % | | |
| ☐ | POLAND | 10 | 1.3908 % | | |
| View Records / Exclude Records | Field: Country/Territory | Record Count | % of 719 | Bar Chart | Save Analysis Data to File |

(26 Country/Territory value(s) outside display options.)
(29 records (4.0334%) do not contain data in the field being analyzed.)

FIGURE 5-3
Publications on IR detector technologies from 1980 to 1989 (719 results). Conference proceedings are included only from 1990 to the present. Therefore, results in this decade are significantly lower because coverage does not include conference proceedings.

| View Records / Exclude Records | Field: Country/Territory | Record Count | % of 16620 | Bar Chart | Save Analysis Data to File |
|---|---|---|---|---|---|
| ☐ | USA | 6807 | 40.9567 % | | |
| ☐ | GERMANY | 2036 | 12.2503 % | | |
| ☐ | FRANCE | 1381 | 8.3093 % | | |
| ☐ | JAPAN | 1345 | 8.0927 % | | |
| ☐ | ENGLAND | 1188 | 7.1480 % | | |
| ☐ | ITALY | 771 | 4.6390 % | | |
| ☐ | CANADA | 752 | 4.5247 % | | |
| ☐ | SPAIN | 728 | 4.3803 % | | |
| ☐ | NETHERLANDS | 547 | 3.2912 % | | |
| ☐ | PEOPLES R CHINA | 453 | 2.7256 % | | |
| View Records / Exclude Records | Field: Country/Territory | Record Count | % of 16620 | Bar Chart | Save Analysis Data to File |

(99 Country/Territory value(s) outside display options.)

FIGURE 5-4
1990-1999 Web of Science publications. A total of 16,620 results are found.

| View Records / Exclude Records | Field: Country/Territory | Record Count | % of 35565 | Bar Chart | Save Analysis Data to File |
|---|---|---|---|---|---|
| ☐ | USA | 13148 | 36.9689 % | | |
| ☐ | GERMANY | 4066 | 11.4326 % | | |
| ☐ | PEOPLES R CHINA | 3465 | 9.7427 % | | |
| ☐ | FRANCE | 2957 | 8.3144 % | | |
| ☐ | JAPAN | 2789 | 7.8420 % | | |
| ☐ | ENGLAND | 2446 | 6.8775 % | | |
| ☐ | ITALY | 2068 | 5.8147 % | | |
| ☐ | SPAIN | 1619 | 4.5522 % | | |
| ☐ | CANADA | 1446 | 4.0658 % | | |
| ☐ | NETHERLANDS | 1073 | 3.0170 % | | |
| View Records / Exclude Records | Field: Country/Territory | Record Count | % of 35565 | Bar Chart | Save Analysis Data to File |

(121 Country/Territory value(s) outside display options.)

FIGURE 5-5
2000-2010 Web of Science publications. A total of 35,565 results are found. In looking at the data by decade, two noticeable trends are the decrease in the percentage lead of the United States and the dramatic increase of publications originating from the People's Republic of China, from not being in the top ten to a strong third place.

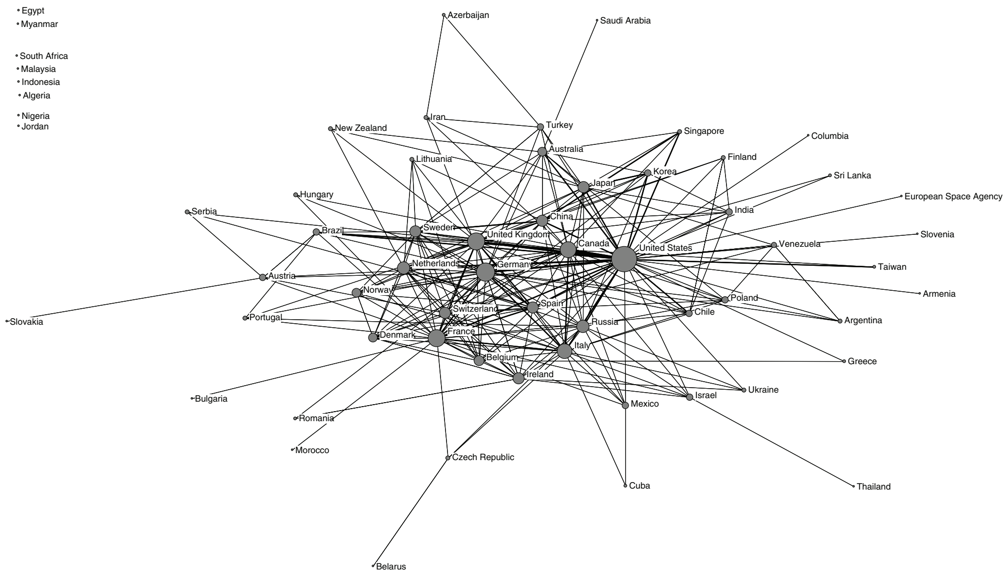


FIGURE 5-6
A representation of joint publication activity on infrared detectors from 2000 to 2010. The size of each country circle represents the number of joint publications from that country, and the weight of the interconnecting lines represents the number of joint publications with joint authorship affiliations from the two countries.

There is substantial international collaboration at the research level. Of the publications listed above, 12 percent had authors from multiple countries and 2 percent had authors from three countries or more. This interconnectivity is represented in Figure 5-6, in which the size of the circle is a representation of the number of joint publications with authors from that country and the weight of the interconnecting lines represents the number of joint publications between authors from the two countries. The data are from the Compendex database covering the years 2000-2010.

## GOVERNMENT ROLES, MARKETS, AND SCALE

In some nations, commercial products may be subsidized by governments through direct funding (or indirect assistance, such as tax breaks) for research and development as well as for the infrastructure to manufacture or distribute a product. The incentive "boost" from government funding can accelerate development and deployment of applications that would otherwise take much longer. Government investments in the commercial arena can have a significant effect on the development and maturation trend for detector and sensor system-related technologies. These developments can be applied to military applications as well as commercial. Some examples of government assistance follow:

- Declarations of local, regional, or national goals related to detector and sensor technologies;
- Large investments in major enabling technologies, such as industrial plants and fabrication facilities;
- Opening of university research centers focused on key technology issues;
- Efforts to attract major manufacturing players (i.e., analogous to courting a foreign automaker to open a new plant in a particular U.S. state); and
- Efforts to attract technical leaders, managers, and financial investments in specific technical areas to particular regions.

Creation of a government incentive is merely the first step, albeit a large one, in influencing certain technologies. Nonetheless, the decision to create an incentive should be taken as an indicator of where to focus attention and resources when assessing the research, development, and maturation of various classes of detector and sensor technology. One example of apparent foreign government investment is mentioned in Chapter 3 (Box 3-2), suggesting Chinese and Iranian interests in rapid deployments using commodity components.

An example of substantial foreign government investment is that of the Chinese government. Sensor and detector laboratory complexes are located in Shanghai, at the Shanghai Institute of Technical Physics (SITP), and at Wuhan, which the Chinese refer to as their "optics valley." Reports by visitors indicate the following:

1. Active quantum-well IR photodetectors (QWIP) and type II strained superlattice (SLS) programs;
2. III-V materials and InGaAs material-based sensors;
3. HgCdTe material growth and sensor array development;
4. Low Earth orbit and geosynchronous orbit (LEO and GEO) satellite sensor fabrication for weather and Earth usage;
5. Sensor payload fabrication using both HgCdTe and silicon detectors made in-house;
6. Both linear and two-dimensional array formats including readouts designed at SITP; and
7. Formats as large as $512 \times 1024$.[1]

[1]Paul Norton. 2009. Georgia Tech Trip to China and Korea. Santa Barbara.

These activities, in addition to technical papers presented at various international conferences, indicate a major national effort with the intent to become a leading source of quality sensors, maybe even a dominant source. There was no report of extraordinary activity by China in the commoditized IR sensor markets; hence it is possible these activities will find primary use in military or intelligence missions. Close scrutiny of Chinese activities in high-yield, high-quality production of these sensors classes will determine whether China will soon become a major player in military detector technology.

**FINDING 5-1**
An emerging foreign force in sensor technology is the set of newly established government-sponsored institutes in China. Extensive new laboratory facilities are known to be producing quality materials and sensor arrays.

**RECOMMENDATION 5-1**
The intelligence community should closely monitor Chinese activities for signs that an operational capability is being established for manufacturing high-quality sensor arrays.

Visible and IR sensor systems tend to vary significantly in the way they are applied in end products and also in their end customers. Visible sensors have a broad base of interest and a vast array of customers, ranging from the global civilian world (e.g., commercial, industrial, scientific, academic, and civilian government agencies) to the military and the intelligence community. Their broad application to civil uses is the primary driver for their development and represents the majority of their applications. An example would be proliferation of cell phones, especially in areas with less established landline telephone infrastructure than in the United States. These cell phones usually come with visible cameras. Alternatively, IR sensors have a much smaller application space, primarily either in national security applications or in specialized niche areas, such as scientific research, medicine, process control, and instrumentation. This smaller customer base results in a better-focused effort, but also a smaller and less diverse source of funding, principally from the government.

**FINDING 5-2**
Visible sensor technologies are more strongly driven by commercial markets, especially overseas, than by national security requirements. In contrast, IR and thermal sensors are more strongly influenced by national security requirements.

The differences between these widely varying applications and customers, coupled to differences in the business and funding models, lead to differences in the signatures and key indicators of the visible and IR product lines. Below are examples of the way different factors have affected the research, development, and manufacture of sensor technologies. The examples could also provide some insight as to how these factors may shape the future of sensor technologies.

Foreign governments have military-related activities under way to use sensor systems to provide intelligence, surveillance, and reconnaissance (ISR) information. Foreign governments also use sensor systems for internal law enforcement and civil concerns, such as pollution monitoring or tracking vehicular traffic.

Among companies there is constant competition to be first to market. In contrast to lengthy government cycles, times for commercial sensor system technologies may be as short as six months or less, but seldom more than a few years. Also, commercial sales for some items can be larger than sales for the government.[2] Even though commercial sales can exceed those for government, the life cycle from development to fielding to retirement may be a fraction of that for government systems. This ratio could result in commercial systems being backward-integrated into, or augmenting, government capabilities. In other words, commercial products may end up driving government products.[3]

Often unit-cost reductions are obtained as a result of economies of scale in the production of large numbers of sensors, such as cell phone cameras. However, a high-performance sensor system, which foreign governments might fly on aircraft or satellites, would be costly. For near-peer competitors of the United States, such as China or Russia, costs for what is deemed to be an essential national security sensor system would probably not be a major prohibition to its development and deployment. For other nations, however, such costs could be a major driver in deciding whether or not to develop and deploy a costly, high-performance sensor system.

[2]One example from the FLIR Systems Inc. Annual Report for 2009, page 63, revenue from external customers (in million dollars).

| Year | 2009 | 2008 | 2007 |
|---|---|---|---|
| Government Systems | 655 | 569 | 382 |
| Commercial Systems | 492 | 508 | 397 |

Note: For this data the term "Commercial Systems" represents sales by the Thermography division plus the Commercial Vision Systems division because FLIR is combining both divisions into a single Commercial Systems division in 2010 (page 2 of report). From http://files.shareholder.com/downloads/FLIR/913990835x0x353521/81C2AFD8-637C-4E0B-99CA-039BCCAB36A9/Form_10_K_typeset.pdf. Last accessed on May 24, 2010.

[3]Consider that the time frame of 10-15 years represents only about one generation (at most, two generations) of technologies and systems in use by the U.S. military today. That same time frame may represent perhaps 5-30 generations of commercial technology. Contrast the military deployment cycle to the two- to three-year pace of Moore's law shrinking of integrated circuits.

The United States has long enjoyed relative dominance with respect to night operations due to its superior tactical IR sensors. This situation, however, is changing. As technology advances have occurred, from starlight goggles to cooled IR to uncooled microbridges, the result has been lower cost and wider use in areas such as law enforcement, environmental surveillance, border surveillance, and even sport hunting. Earlier-generation technology has diffused into the marketplace (as an example, a Google search on night vision goggles pulls up hundreds of competing commercial sources).

The U.S. IR sensor industry has gradually transformed over the past two decades from Department of Defense (DOD) prime-contractor dominance controlling the marketplace, to smaller research and development establishments and commercial system suppliers having a larger share. Foreign defense companies have also begun to vertically integrate their system products. For example, Thales,[4] in France, formed SOFRADIR,[5] in 1986, to produce mercury cadmium telluride (MCT) detectors for insertion into tactical systems; SOFRADIR now produces QWIPs, and it acquired Electrophysics[6] in the United States in 2008. Further, Finmeccanica,[7] an Italian conglomerate, acquired the key U.S. tactical sensor supplier DRS Technologies in 2008.[8] Fortunately these acquisitions involve North Atlantic Treaty Organizations (NATO) countries, and "firewalls" are set up to prevent the diffusion of classified work out of the United States. Nevertheless, this trend to foreign defense company ownership would seem to dilute the U.S. supremacy in tactical IR sensors.

Strategic IR sensors with higher performance are closely controlled by security classification. Often, cues to progress can be obtained from knowledge of what is happening in research and development facilities, but actual system deployment requires production capability. Experimental demonstrations are not sufficient for integration into militarily useful sensor systems. Hence, the critical transition to industrial production becomes an important tip-off for closer inspection to signal an escalation of capability by a foreign power or a foreign-owned company.

In contrast to IR sensors, visible sensor developments are historically more accessible. Detector elements, sensors, and readout devices can be commercially procured with good capability. System configurations (e.g., multispectral or hyperspectral) are generally application specific, and they may be classified since they depend on mission requirements.

---

[4]For additional information on Thales, please see http://www.thalesgroup.com/.

[5]For additional information on SOFRADIR, please see http://www.sofradir.com/.

[6]For additional information on Electrophysics, please see http://www.electrophysics.com/.

[7]For additional information on Fenmeccanica, please see http://www.finmeccanica.it/Holding/IT/index.sdo.

[8]For additional information on DRS Technologies, please see http://www.drs.com/.

**FINDING 5-3**
There is a significant difference in infrastructure requirements and development paths between "strategic" (low volume, high unit value and capability) and "commodity" (high volume, lower unit value and capability) sensors and systems. Commodity sensors tend to have short (i.e., months or a small number of years) life cycles and to be closely tied to short-term market requirements.

Satellites, which contain some of the most costly, complex, and technically advanced imaging platforms known to man, demand the highest possible quality and capability from their sensor systems. Historically the domain of only the most technically advanced and prosperous governments, satellites are rapidly becoming the purview of commercial industry, primarily for communications and broadcast applications.

Importantly, high-resolution commercial satellites, such as the GeoEye-1 with 41 cm resolution are challenging military capabilities. Not only does commercial satellite imaging technology represent capabilities on a par with military systems, but the market size for commercial applications is much larger as well. The end result is greater interest and concomitant diversity of manufacturers, leading to increased competition and greater research and development in these areas.[9]

**FINDING 5-4**
Lowering of barriers to commercial satellite systems is expanding the market for "strategic-class" imaging technology. The next decade will see greater investment in these capabilities.

## U.S. EXPORT RESTRICTIONS

U.S. companies are constrained from passing controlled information to foreign interests by numerous regulations and laws, including the International Traffic in Arms Regulations (ITAR), which is managed by the Department of State, but governs DOD operations.

ITAR is intended to control, or at a minimum maintain knowledge of, the spread and dissemination of technology to minimize or mitigate the threat that advanced technology could pose to U.S. national security interests. The threat can manifest itself in many ways, but the two primary areas of interest are use of U.S. technologies against the United States (i.e., adversaries having U.S. capabilities, or nearly so) and adversarial development of countermeasures to negate or mitigate the advantage of U.S. technologies (i.e., counters to U.S. systems). Unfortunately ITAR appears to have exacerbated some of these problems, not mitigated them.

[9]See http://www.sciencedaily.com/releases/2009/05/090526183858.htm.

For example, with respect to U.S. night vision technology, foreign capability—uncontrolled and at times unknown (to the United States)—has been growing and providing a wide array of technologies to numerous customers.

When a nation limits export of technology, it reduces the size of that nation's potential market and consequently the number of, and profit associated with, items that are sold. This has a direct impact by reducing potential benefit from manufacturing learning and economies of scale and also reducing incentives to improve or evolve new technologies. Reductions in profits reduce available capital and thereby diminish industry's ability to invest in research and development. There could also be a stifling effect on basic research at academic institutions. The cascading effect ultimately leads to slowed rates of growth and innovation, generally supporting a trend to drive innovation overseas and possibly to slow U.S. enhancement of current technologies.

Where foreign markets exist, U.S. manufacturers will seek to compete if they believe they can benefit. One way this may be reconciled with ITAR restrictions is through purchase and development of foreign-made sensors, therefore maintaining the ability to export the product. Consequently, investment that might be made in improving U.S. technologies is diverted into research and development for improving unrestricted foreign technologies.

ITAR ultimately results in smaller market size and reduced international share for U.S. firms. In addition to consequences (e.g., with respect to night vision technology), the U.S. government is forced to provide a larger portion of funding for research and development. Aside from costing more, this also limits innovation because research and development will be focused almost exclusively on meeting DOD needs, making it more difficult to pursue high-risk, high-payoff avenues. The area in which ITAR restrictions are most significant is where there exists a large commercial market, as well as a military market.

**FINDING 5-5**

Current export restrictions will continue to have a significant effect on development and maturation of detector technologies over the next decade. Numerous foreign countries are already developing their own technology base rather than utilizing U.S. technology and often will compete with U.S. technology. U.S. export restrictions are a primary driver creating this competition. U.S. companies invest significant resources in obtaining, funding, and exploiting foreign products so that they can compete in foreign markets without export restrictions.

As highlighted in Box 5-1, control of sensitive technology was recently cited in the *Quadrennial Defense Review* that included Presidential direction for a "comprehensive review" to identify reforms in the current export system.

BOX 5-1
Control of Defense-related Technologies

The global economy has changed, with many countries now possessing advanced research, development, and manufacturing capabilities. Moreover, many advanced technologies are no longer developed predominantly for military applications with eventual transition to commercial uses, but follow the exact opposite course. Yet, in the name of controlling the technologies used in the production of advanced conventional weapons, our system continues to place checks on many that are widely available and remains designed to control such items as if Cold War economic and military-to-commercial models continued to apply.

The U.S. export system itself poses a potential national security risk. Its structure is overly complicated, contains too many redundancies, and tries to protect too much. Today's export control system encourages foreign customers to seek foreign suppliers and U.S. companies to seek foreign partners not subject to U.S. export controls. Furthermore, the U.S. government is not adequately focused on protecting those key technologies and items that should be protected and ensuring that potential adversaries do not obtain technical data crucial for the production of sophisticated weapons systems.

These deficiencies can be solved only through fundamental reform. The President has therefore directed a comprehensive review tasked with identifying reforms to enhance U.S. national security, foreign policy, and economic security interests. Reform efforts must reflect an inherently interagency process as current export control authorities rest with other departments. Similarly, meaningful reforms will not be possible without congressional involvement throughout the process. The Department of Defense has a vital stake in fundamental reform of export controls and will work with our interagency partners and Congress to ensure that a new system fully addresses the threats that the United States will face in the future.

SOURCE: *2010 Quadrennial Defense Review.* http://www.defense.gov/qdr/. Last accessed on August 29, 2010.

As reform efforts are considered, three issues should be taken into account: (1) technology export controls may encourage proliferation; (2) U.S. markets could be reduced due to export controls; and (3) U.S. manufacturers might make more use of foreign technologies as U.S. technology enhancement slows. Also, it would be useful to keep in mind that the overall goal of regulating the dissemination of technology should be to minimize the threat posed by new and emerging technologies.

## SUPPLY CHAIN BOTTLENECKS

As new technologies develop there may be only limited ability to supply early adopters. Until sufficient production becomes available, bottlenecks could arise. When there exists a single-supplier bottleneck of key components, several problems as well as vulnerabilities arise:

1. Failure of the single supplier can have significant deleterious effect on research, development, and manufacture of sensor technologies and systems.
2. The single supplier is subject to natural disasters, intentional attacks, or coercion, each of which can have significant impacts on organizations that depend on single-supplier products.
3. The single supplier may affect U.S. domestic efforts, but some foreign companies may pursue technologies different from the United States. It is conceivable that situations may arise in which U.S. research, development, and manufacture may be affected, while foreign efforts might not.

There is considerable benefit to be gained by maintaining cognizance of these single-point suppliers as they emerge with new technologies and evolve with the technology base.

In the United States, responsibility for supply chain issues in government programs falls to the system program office and prime contractor in a shared responsibility. Normally, the supply of piece parts from foreign or domestic suppliers is ensured by requiring at least two sources or by stockpiling parts. This is the guarantee against disruption of system operations or future deliveries. The DOD bureaucracy acknowledges this possible disruption and has established an office to monitor this issue.

Most DOD ISR systems pay particular attention to key technologies that are the heart of the sensor system concept. Considerable attention has been paid to parts such as application-specific integrated circuits (ASICs) that increasingly are made in off-shore production facilities. Most of the silicon foundries are located off shore, with concerns about continuity of supply and security of classified chips. Establishment of a U.S. "trusted foundry" and increased use of FPGAs (field-programmable gate arrays), which can be system-configured in controlled fashion, have mitigated this worry.

**FINDING 5-6**
Evolution of new technologies often generates single-supplier bottlenecks. These can have significant, though transient, impacts on research, development, and manufacturing.

**RECOMMENDATION 5-2**
The intelligence community should be aware of the development and status of single-supplier bottlenecks.

## ADDITIONAL CONSIDERATIONS

Identifying trends is complicated by many different factors involved in making a deployable sensor system. Important macro trends include (1) growth of foreign businesses selling first- and second-generation sensor systems for military and civilian applications and (2) roughly fixed rates of U.S. research and development for sensors, primarily sponsored by government.

Much ancillary equipment, such as optics and electronics, can be produced by several nations. For second-generation and commodity sensors, the focal plane array (FPA) is a small part of the system cost and is a mass-produced competitive item. The trend is market driven. For high-value sensors that use cutting-edge sensor technology—with chip fabrication that is proprietary or classified, low in volume, derived from U.S. government research and development, and has special processing features—the trend is performance driven.

As foreign businesses equip much of the world with night vision capability, the once-prominent U.S. lead fades. The trend is for foreign production to focus on large quantities of sensors to gain the cost advantage associated with increased yield. Similar statements can be made about visible active pixel sensors for cell phones, except that market motivation in this case is for consumer products.

The trend for high-performance sensors, such as high-sensitivity charge-coupled devices (CCDs) and two-color, large-pixel-count MCTs, is different. These third-generation sensors are exclusively fabricated in the United States. No current foreign commercial business profit motive exists to drive the efforts; hence, a different motivation is involved. Certainly strategic-level satellite sensor systems constitute a foreign motivation, and the ~500 Chinese satellites could employ an advanced-performance type of sensor to advantage. Therefore, a general trend by the Chinese to fabricate advanced sensor types is anticipated. This is supported by advanced research and development laboratories, particularly for MCT and single-layer superlattice (SLS) structures. Reports of heightened laboratory investment in Wuhan support this conclusion (see prior discussion).

Although detectors and sensors may have considerable capability, this may not be fully exploited due to limitations in other portions of the system. For example, given the existence of high-resolution, broad-band imagers, the data throughput needed to successfully exploit their full capabilities may not exist (see discussion in Chapter 4).

Although much of this report addresses technology developments directly related to detectors and sensors, they must be integrated into an overall system. To make these technologies viable, a vast array of support, interface, and associated technologies must also mature and be integrated into a functional system. In many cases the limiting factor of a capability is not the detector or sensor technology but rather an ancillary technology. For example, a sensor may work effectively and with great capability in a laboratory, but because of the large amount of data generated it may be unable to function effectively in a system due to cooling or electronic limitations. Further, even without such limitations, other problems (e.g., manufacturability, usability, and sustainability in the field) could conspire to limit ultimate system effectiveness.

Consider the problem of replacing 100,000 military night vision systems. Even though a generational leap in technology may be available now, it may take many years to obtain funding and put in place a supply chain sufficient to replace the system worldwide.

## CONCLUDING THOUGHTS

The ability of the U.S. military to operate at night is no longer unique: Focused foreign investment and the reluctance of the United States to share leading-edge technology with allied nations have resulted in the proliferation of sensor-system component manufacturing to Europe and Israel with equivalent levels of performance. To offset costs associated with maintaining high-technology detector production, U.S. allies have exported critical components and technology because diplomacy does not always trump economic incentive. Thus, the technology is migrating to Asia, where American influence has little bearing on export control.

From the U.S. perspective, two principal threats from the proliferation of imaging sensor systems are (1) organizations that resort to clever application of available technology—for example, improvised explosive devices—and (2) developed countries with imaging systems that approach parity with the United States. The difference in threat level between some night vision capability and none is more significant than the difference between first- and third-generation imaging IR systems.

As low-cost thermal imagers become more readily available, such as in some luxury automobiles, it remains only a matter of time before the U.S. military faces adversaries with IR rifle scopes and even night sights for man-portable missile launchers on current battlefields. This technology escalation should come as no surprise, and continued U.S. research and development of sensors and technologies to detect and counter such threats will be increasingly significant.

Persistent surveillance is emerging as the centerpiece of network-centric warfare for the asymmetric engagements the United States has encountered in Afghanistan and Iraq. The United States hopes to become the omnipresent "eye-in-the-sky" and continuously monitor everything of interest. The United States now has the capability of doing this in daylight, but the mountains of data that are generated to survey a relatively small area require significant resources to generate actionable information. As larger-format IR sensors become more readily available, persistent surveillance will quickly become a data bandwidth and signal-processing nightmare and will not be truly practical without significant innovations in autonomous signal processing, video compression, and low-power, high-data-rate, radio-frequency links.

With the current focus on the Mideast, it is easy to ignore the potential threat of more sophisticated sensor technology from the developed nations. Here, any technology advantage the United States may have once enjoyed has eroded to the point where the warfighters are essentially at parity. The deployment of third-generation IR sensor technology will provide some level of overmatch for the user, but gaining a significant operational advantage will mandate the development of a new generation of active or passive IR sensors that provide the user with higher-resolution imagery at longer standoff range through FPAs with gain, advanced signal processing, and covert illumination sources.

In addition to the United States, the French (SOFRADIR), the Germans (AIM), and the United Kingdom (SELEX) are investing in the development of sensors based on detectors with "noiseless" gain that can amplify low-intensity signals to overcome system noise limits at wavelengths ranging from 1 μm through the long-wavelength infrared (LWIR). In conjunction with new solid-state laser illumination sources and recent developments in silicon complementary metal oxide semiconductor (CMOS) electronics, these prototype systems can provide three-dimensional images that significantly enhance identification capability and can employ time of flight to eliminate background at ranges other than the target. Examples of a range of competitive foreign detector technologies are shown in Tables 5-1 and 5-2.

TABLE 5-1 Competitive Non-U.S. Cryogenic IR Sensor Technology

| Country | Company | MCT | InSb | Superlattice | SWIR | Optical Materials |
|---|---|---|---|---|---|---|
| China | Various | LPE/MBE | | R&D | ? | Dominant position |
| France | SOFRADIR | LPE/MBE | | R&D | InGaAs | R&D |
| Germany | AIM/Fraunhofer | LPE/MBE | | Production | MCT | R&D |
| Israel | SCD | LPE | Dominant position | R&D | ? | R&D |
| Japan | Various | LPE/MBE | | ? | InGaAs | Dominant position |
| Russia | Various | Bulk/LPE | Small arrays | Theory | ? | Dominant position |
| United Kingdom | SELEX | LPE/MOCVD | | | InGaAs | R&D |

NOTE: LPE = liquid-phase epitaxy; MBE = molecular beam epitaxy; MOCVD = molecular organic chemical vapor deposition.

TABLE 5-2 Competitive Non-U.S. Uncooled IR FPA Technology

| Country | Company | Microbolometers | MWIR |
|---|---|---|---|
| China | Alpha Si | MCT R&D | |
| Germany | AIM/Fraunhofer | ? | MCT R&D |
| Israel | SCD | $VO_x$ | nBn R&D |
| Various | Buy French | MCT R&D | |
| France | SOFRADIR | | |
| Japan | Various | Si diodes | ? |
| Russia | Buy French | ? | MCT R&D |
| United Kingdom | SELEX | Buy ? | MCT R&D |

Gains in sensor performance and cost reduction can be attributed to advances in detector materials and devices and silicon CMOS technology. The commercial investment in CMOS has resulted in "smarter" FPAs and more capable signal processors. State-of-the-art CMOS technology is readily available worldwide and has certainly contributed to leveling the playing field for both consumer and military electronics. Custom integrated circuits, such as for FPA readouts, can be fabricated in any number of foundries, so performance and functionality are only constrained by the talent of the designer.

Imaging systems will continue to deliver improvements in resolution and provide coverage of larger areas. The fundamental limits to sensitivity for passive broadband, Earth-viewing sensors are being approached. Leap-ahead sensor capabilities could occur related to active systems, where sensors will incorporate illumination sources that enable new imaging modes that enhance recognition range by generating three-dimensional renditions of a scene. An additional example could be non-imaging sensors, where spectral and temporal signatures will enable identification of specific materials within a given field of view. Also, it is anticipated that a large gain in surveillance capability will come from parallel signal processing and improved video compression for data transmission.

The globalization of state-of-the-art CMOS foundries and production capabilities will tend to level the playing field for semiconductor electronics. The ability to develop semiconductor technology is no longer restricted to Western high-technology centers as is evidenced by the number of non-U.S. graduate students enrolled in world-class graduate science and engineering programs. As a result of this trend in globalization, sensor proliferation will accelerate; U.S. ability to maintain a leadership technology position will require more focus with shorter development cycles.

The evolution of low-cost reliable coolers will facilitate higher-performance sensor systems. Already, this trend is being seen in some hand-held soldier systems where the cooler and FPA cost-power trades render this solution more acceptable than larger-aperture, uncooled imagers. A new generation of minicoolers is becoming available and, with the emergence of higher-operating-temperature detector technology, could supplant the move to uncooled thermal sensors.

As manufacturing costs continue to fall, commercialization of imaging sensors will make less capable night vision technology more readily available in the marketplace. Subsequently, nonaligned adversaries will adapt these sensors to low-cost weapons systems. Advanced processors, such as those available in video games, are being adapted to sensor image processing tasks; it no longer requires an expensive, custom real-time processor to sift video data, and knowledgeable engineers are also becoming available globally.

### Tracking Developments

How does one track sensor development, and what does one look for? Generally, observers fail to recognize breakthrough technologies when they are featured at a technical conference, show up in a journal article, or are touted by some start-up on a website. Generally, few of these innovations will survive the journey to market and, on occasion, a promising new technical development will be kept under wraps.

For asymmetric engagements, low-cost commercial products, creative talent, and black-market availability will determine the sophistication of sensor systems. These can be tracked, and engineering tiger teams can postulate the possibilities of employing a BMW forward-looking infrared (FLIR) as a night site for a crew-served weapon or a man-portable missile launcher and ways to detect and counter such weapons systems. Innovative commercial cameras with autofocus or image stabilization technology may lend themselves to the development of low-end drones. The United States must be prepared for surprise and respond quickly to neutralize any short-term advantage presented by our adversaries.

Silicon CMOS technology is the manufacturing platform upon which most modern sensors are based, from readout integrated circuits to sophisticated signal processors. State-of-the-art CMOS is globally available and makes establishment of a significant advantage in sensor technology potentially short-lived. The developments that should be tracked include the heterogeneous integration of other technologies onto the CMOS platform. Examples include the incorporation of germanium into the CMOS platform for higher-speed electronics and for detectors and microelectromechanical (MEMs) devices for accelerometers or uncooled microbolometer-based thermal detectors.

The integration of compound semiconductors with CMOS creates the potential for monolithic FPAs that span the entire electromagnetic spectrum. As the technology continues to evolve, one could envision FPAs with integral signal processors that are layered into the CMOS chip. Here, only specified intelligence would be the output, significantly reducing the data bandwidth for the sensor. As heterogeneous integration becomes a manufacturing reality, the developers of this technology will need to be monitored to avoid surprise at its initial introduction into sensor systems.

# Appendixes

# Appendix A

# Biographical Sketches of Committee Members

**Steven R.J. Brueck**, *Chair*, is the director of the Center for High Technology Materials (CHTM) and is a distinguished professor of electrical and computer engineering, physics, and astronomy at the University of New Mexico. As CHTM director, he manages research and education at the boundaries of two disciplines. The first, optoelectronics, unites optics and electronics and is found in CHTM's emphasis on semiconductor laser sources, optical modulators, detectors, and optical fibers. The second, microelectronics, applies semiconductor technology to the fabrication of electronic and optoelectronic devices for information and control applications. Examples of these unifying themes at work are silicon-based optoelectronics and optoelectronics for silicon manufacturing sensors. He is also a former research staff member of the Massachusetts Institute of Technology (MIT) Lincoln Laboratory. He is a member of the American Physical Society and the Materials Research Society and is a fellow of the Institute of Electrical and Electronics Engineers (IEEE), the Optical Society of America (OSA), and the American Association for the Advancement of Science. Dr. Brueck is a member of the National Research Council's (NRC's) TIGER (Technology Insight—Gauge, Evaluate, and Review) Standing Committee and was a member of the NRC's Committee on Nanophotonics Accessibility and Applicability and Committee on Emerging Micro- and Nanotechnologies.

**Paul McManamon**, *Vice Chair*, is an independent consultant and works half-time as the technical director of the Ladar and Optical Communications institute, LOCI, at the University of Dayton. Until May of 2008 he was chief scientist for the Sensors

Directorate, Air Force Research Laboratory (AFRL), Air Force Materiel Command, Wright-Patterson Air Force Base, Ohio. The Sensors Directorate consists of about 1,250 people responsible for developing new sensor technology for the Air Force. Dr McManamon was responsible for the technical portfolio of the Sensors Directorate, including radio-frequency (RF) sensors and countermeasures, electro-optical (EO) sensors and countermeasures, and automatic object recognition. He has developed multidiscriminate electro-optical sensors, including multifunction laser radar, novel EO countermeasure systems, and optical phased-array beam steering. Dr. McManamon has participated in three Air Force Scientific Advisory Board (AFSAB) summer studies: New World Vistas (1995); A Roadmap for a 21st Century Aerospace Force (1998); and Sensors for Difficult Targets (2001). Dr. McManamon was instrumental in the development of laser flash imaging, initiating the ERASER program as a method to enhance EO target recognition range by a factor of 4 or 5. Dr. McManamon is widely recognized in the electro-optical community. Dr. McManamon was the 2006 president of SPIE. He was on the SPIE board of directors for seven years and on the SPIE Executive Committee from 2003 through 2007. Dr. McManamon serves on the executive committee for the Military Sensing Symposia (MSS). He is a fellow of SPIE, IEEE, OSA, AFRL, and MSS.

**Stefan Baur** is director of Technology and Advanced Programs at Raytheon Vision Systems. His responsibilities include establishing critical technology roadmaps, developing Raytheon's focal plane array (FPA) intellectual property portfolio, and managing an approximately 30 development programs. Prior to this position, he spent 15 years at the Raytheon El Segundo developing advanced EO systems. Highlights of his career include leading the technical direction of the Raytheon Land Warrior program, program manager of Thermal Weapon Sight (TWS), program manager of the core staring imaging model for Raytheon Airborne EO systems, program manager of Navy Shipboard Long Range IRST (Infrared Search and Track System), and manager of FPA test and design center. Mr. Baur serves on several internal Raytheon and U.S. government-sponsored committees that help shape the direction of future EO systems. He is a graduate of the University of California in Los Angeles.

**Valerie Browning** is an independent consultant and subject matter expert for ValTech Solutions, LLC. She serves as a subject matter expert for a number of Department of Defense (DOD) and other government activities in the areas of advanced materials and alternative energy. Prior to forming ValTech Solutions, LLC, in December 2007, Dr. Browning served as a program manager in the Defense Sciences Office (DSO) at the Defense Advanced Research Projects Agency (DARPA). During her tenure at DARPA, she assumed full responsibility for the strategic planning, operating management, leadership, and development of multiple DOD research and development programs providing innovative technologies in power and energy, radar, telecommunications, and biotechnology for diagnostics, therapeutics, and chemical-biological warfare defense. Specific programs managed by Dr. Browning include the Metamaterials, Palm Power, Direct Thermal to Electric Conversion, Negative Index Materials, Robust Portable Power Systems, and Bio-Magnetic Interfacing Concepts Programs. She also served as the DARPA liaison to the DOD Integrated Product Team (IPT) on energy security and served as acting DSO director prior to her departure from government service. In addition to her time at DARPA, Dr. Browning spent 16 of her 24 years of government service as a research physicist at the Naval Research Laboratory. Her primary areas of research were thermoelectric materials, high-temperature superconductors, and magnetic oxide materials. Upon leaving her government position, Dr. Browning was awarded the Secretary of Defense Award for Outstanding Public Service. She is active in a number of professional organizations including the American Physical Society, the Materials Research Society (MRS), and Sigma Xi. Most recently, Dr. Browning served as co-chair for a 2007 MRS Symposium on magnetic materials and was the Technical Program Committee chair for the 2008 Fuel Cell Seminar. She continues to serve on the Technical Program Committee for the Fuel Cell Seminar and Exposition and was recently appointed as a member of the National Materials Advisory Board. She has a Ph.D. in physics from the Catholic University of America.

**John Devitt** is division chief of GTRI-Electro Optics Sciences Lab's remote sensing group at Georgia Tech. He's the former manager of the Systems Analysis and Test group in the Infrared Products Engineering organization at L-3 Cincinnati Electronics (CE). He has more than 20 years' experience in advanced technology projects in EO-IR, optics, sensors, and related areas, including direct experience in leading both large-format FPA development and novel detector programs. He has led major research and development, manufacturing technology, and other advanced technology developmental projects with broad technical and economic scope at both L-3 CE and, previously, GE Global Research Center. This has involved coordinating multiple cross-technical and organizational groups within businesses, managing university and national laboratory subcontracts, and being a key interface with the major government agencies. In 1993, Mr. Devitt was awarded the GE Sanford Moss Award for Most Outstanding Test and Measurement Program. He has 12 U.S. patents and numerous publications. Mr. Devitt is a certified Six Sigma Green Belt and has an M.S. in physics from Ohio State University. He currently serves as chairman of the MSS Passive Sensors Committee and is on the SPIE IR Technology Committee.

**Thomas Hartwick** is a technology and business adviser with long experience in technical management for commercial and government activities with emphasis on research and development. With a specialty in electronics and optics, his background includes R&D and management positions at Hughes Aircraft Company, the Aerospace Corporation, and TRW. In addition to participation in key business strategies of more than a dozen companies, he has served on numerous national advisory panels, including the Advisory Group on Electron Devices for DOD, review panels for the Defense Science Board and AFSAB, and various boards and committees of the National Research Council. Hartwick currently serves on five corporate boards and committees.

**Angela Hodge** joined the Johns Hopkins University Applied Physics Laboratory (APL) in November 2005 and is a member of the Senior Professional Staff. Her work focuses on sensor fusion and the investigation, development, and design of algorithms and models for target detection and tracking of air and space systems. This work includes the analysis of infrared sensors, radar, and flight test data. It involves the application of systems engineering and digital signal processing principles. She has researched, implemented, and demonstrated performance for several designs. Prior to joining APL, Dr. Hodge worked as a civil servant and electrical engineer at various U.S. national laboratories, including the National Institute of Standards and Technology, the Naval Research Laboratory, and the U.S. Army Night Vision and Electronic Sensors Directorate. She has published more than a dozen electrical engineering articles on sensor technologies for various journals and conference proceedings. The majority of that work focuses on development of sensors for detection of specific biological agents in fluids. On January 2, 2007 patent 7,158,067, "Analog to digital converter using sawtooth voltage signals with differential comparator," for which she is co-inventor, was issued.

**Michael Hopmeier** is the founder and president of Unconventional Concepts, Inc. He serves as technical adviser and operational consultant to numerous government agencies and commercial firms. Some of his current project areas include homeland security, chemical-biological incident response, combat casualty care, crisis response and management, operational medical support, unconventional pathogen countermeasure programs, space station technology and exploration assessment, federal agency protective measures, counterterrorism, integrated federal-civilian disaster response, suicide terrorism, terrorist motivation and societal analysis, training and preparedness, and Special Operations technology support. Mr. Hopmeier previously served as technical adviser and/or operational liaison for the U.S. Army Medical Research and Materiel Command, U.S. Air Force (USAF) Wright Laboratory, USAF 59th Medical Wing, U.S. Navy Amphibious Warfare Program, explosive ordinance disposal program, DARPA, and the Federal Emergency Management Agency. He has memberships in several organizations including the World Health Organization (WHO) Working Group on Stockpile for Countering Radiological Disasters and the Editorial Board of the International Society of Disaster Medicine. He is the first U.S. participant, Israel Home Front Command Search and Rescue Course, Israel, for which he received special commendation, Air Force Materiel Command and Air Mobility Command, support to Joint Chiefs of Staff Program. Mr. Hopmeier received M.S. in mechanical engineering from the University of Florida.

**Steven Jost** is director of photonics, advanced systems, and technology at BAE Systems Electronics and Integrated Solutions. He has more than 35 years' experience in research, development, and production of electro-optical components and systems. Currently, he is responsible for the development of photonic integrated circuits, infrared focal planes, and EO components for military and commercial applications. With support from DARPA's Electronic Photonic Integrated Circuit (EPIC) program, BAE Systems has been developing a complementary metal oxide semiconductor (CMOS)-based photonic integrated circuit capability at its state-of-the-art foundry in Manassas, Virginia. He has worked in numerous infrared materials including mercury cadmium telluride (MCT), InSb, GaSb-InAs strained-layer superlattice, GaAs-GaAlAs quantum-well infrared photodetector (QWIP), and both single crystal and polycrystalline lead salts, and he specialized in detector physics-surface passivation. Dr. Jost led the first U.S. technical team to successfully implement CdTe heterostructure surface passivation on long-wavelength infrared (LWIR) MCT production FPAs and developed a radiation-tolerant surface passivation for InSb detectors. Currently his team is transitioning a novel heterostructure surface passivation for resonant cavity lead salt detectors to threat warning sensor applications. Dr. Jost received his Ph.D. in solid-state device physics from Princeton University.

**Linda Katehi** is chancellor of the University of California, Davis. As chief executive officer (CEO), she oversees all aspects of the university's teaching, research, and public service mission; she also holds faculty appointments in electrical and computer engineering and in women and gender studies. Dr. Katehi chairs the President's Committee for the National Medal of Science and is chair of the Secretary of Commerce's Committee for the National Medal of Technology and Innovation. She is a fellow and board member of the American Association for the Advancement of Science and a member of many other national boards and committees. Previously, Professor Katehi served as provost and vice chancellor for academic affairs at the University of Illinois at Urbana-Champaign; the John A. Edwardson Dean of Engineering and professor of electrical and computer engineering at Purdue University; and associate dean for academic affairs and graduate education in the College of Engineering and professor of electrical engineering and computer science at the University of Michigan. Professor Katehi is an expert in the following areas: development and characterization (theoretical and experimental) of microwave, millimeter-wave printed circuits; computer-aided design of VLSI (very large scale integration) interconnects; development and characterization of micromachined circuits for microwave, millimeter-wave, and submillimeter-wave applications including microelectromechanical system (MEMS) switches, high-Q evanescent mode filters, and MEMS devices for circuit reconfigurability; development of low-loss lines for submillimeter-wave and terahertz frequency applications; theoretical and experimental study of uniplanar circuits for hybrid-monolithic and monolithic oscillator, amplifier, and mixer applications; and theoretical and experimental characterization of photonic bandgap materials. Some of her research projects that have created new directions in high-frequency frequency design include W-band power cube; novel packaging approaches for high-density three-dimensional integrated circuits (ICs); device and circuit approaches for next-generation wireless communications; MEMS for microwave and millimeter-wave applications; study of photonic bandgap substrates for use in frequency-selective structures; silicon-based on-wafer packaging for high isolation in high-density circuits; high-Q micromachined resonators for RF filters-diplexers; and MEMS transfer switches. Her work in electronic circuit design has led to numerous national and international awards both as a technical leader and as an educator, 16 U.S. patents, and an additional six U.S. patent applications. She is the author or coauthor of 10 book chapters and about 600 refereed publications in journals and symposia proceedings. She earned her Ph.D. in electrical engineering from the University of California at Los Angeles (UCLA).

**Seethambal Mani** is a professional member of the technical staff in the Sensors and Analog System Department at Sandia National Laboratories. She has experience in manufacturing and assembly of focal plane arrays. She has been involved in the hypertemporal focal plane array project involving tiled arrays. She has coordinated work toward integrating the tiles on a motherboard while developing and optimizing processes between Ziptronix and Sandia microfabrication groups specifically in the area of precision singulation. She also carried out physical and electrical characterization work of the tiled assemblies received from Ziptronix. She continues working on this project interfacing with Ziptronix for the next batch of assemblies scheduled for later this summer. On the BTB project she worked closely with the TIS (Teledyne imaging Systems) group getting familiar with and understanding their process for assembly and detectors. She worked with the staff at Teledyne modifying their processes and procedures to improve yield for the large-area focal plane arrays their assembly process. She also worked with the Teledyne detector design and fabrication groups while reducing risk in their detector fabrication process using specially designed PECs (process evaluation coupons). She led reliability and failure analysis efforts for the BTB 2K and 8K focal plane arrays. She has worked closely with Infrared Radiation Effects Laboratory at AFRL supporting the radiation testing of the BTB 2K and 8K arrays. Currently she is the principal investigator for the nanobump Laboratory Directed Research and Development project. This project is looking into developing a process for hybridization for smaller and smaller features and pixels in order to support large-area focal plane arrays. Prior to joining this group she has experience in the microsystems area in developing processes, technologies, and integrating components. She has worked closely with DRS pushing the state of the art developing hybridization processes. She has a Ph.D. in materials engineering from Rensselaer Polytechnic Institute.

**C. Kumar Patel** is a professor of physics and astronomy, chemistry, and electrical engineering at the University of California, Los Angeles. He is also the founder and CEO of Pranalytica, Inc., a Santa Monica, California based company that carries out R&D and manufacturing and sells trace gas sensors for in situ detection of chemical warfare agents and explosives, systems for standoff detection of explosives (IEDs), and high-power mid-wave infared and long-wave infrared quantum cascade lasers for applications in defense, homeland security, and commercial systems. He served as vice chancellor for research at UCLA from 1993 to 1999. Prior to joining UCLA in March 1993, he was the executive director of the Research, Materials Science, Engineering and Academic Affairs Division at AT&T Bell Laboratories, Murray Hill, New Jersey. He joined Bell Laboratories in 1961 where he began his career by carrying out research in the field of gas lasers. He is the inventor of the carbon dioxide laser, which is one of the most widely used lasers in industry. Dr. Patel received his Ph.D. in electrical engineering from Stanford University in 1961. In 1988, he was awarded an honorary doctor of science degree from the New Jersey Institute of Technology. In 1996, Dr. Patel was awarded the National Medal of Science by the President of the United States of America.

**Tamar Peli** is currently associate director for Special Operations Programs and Signal Processing and Exploitation for the Draper Corporation. She has more than 25 years' experience in signal and image processing with application to the Department of Defense and the intelligence community (IC). Ms. Peli has developed technologies for a wide range of platforms and sensor modalities including imaging and non-imaging radar from X band to low frequencies, IR-IRST-FLIR (forward-looking infrared), video EO-IR, multispectral-hyperspectral, and measurement and signature intelligence. Ms. Peli is leading Draper's White Space Initiative in exploitation technologies. As the lead for the exploitation initiative, Ms. Peli is responsible for developing the technology roadmap, the business strategy, for expanding Draper's role and business opportunities within the IC DoD community and its execution. Ms. Peli is a MSS NSDF (National Symposium on Data Fusion) committee member co-chairing the sensitive compartmented information session.

Ms. Peli holds an M.S. in electrical engineering from Technion-Israel Institute of Technology. She was a member of the NRC's Committee on Sensing and Communications Capabilities for Special Operations Forces and is also a member of the Panel on Digitization and Communications Science of the NRC's Army Research Laboratory Technical Assessment Committee.

**David Shaver** is head of the Solid State Division at MIT Lincoln Laboratory where he oversees research in such varied areas as high-performance imaging sensors, deeply scaled silicon microelectronics, solid-state lasers, optoelectronics, photonics, superconductive devices, quantum computing, and biological agent sensors. His current personal technical interests have included development of photon counting sensors and three-dimensional integrated imagers, and technology related to trusted electronics. Before his present position, Dr. Shaver was responsible for bringing online a Class 10 silicon fabrication facility. He also led the Submicrometer Technology Group, which pioneered the development of 193 nm wavelength optical lithography and served as the technical champion for 193 nm on the Silicon Arsenide Lithography Technical Working Group. Before that he served as chief scientist and director of research for Micrion Corporation (now part of FEI) where he was involved in the development of focused ion-beam and laser-beam microchemistry systems for photomask, microcircuit, and flat-panel display repair and modification. He received his Ph.D. in electrical engineering from the Massachusetts Institute of Technology in 1981. He also attended the Harvard JFK School of Government program for Senior Executives in National and International Security. Dr. Shaver is a fellow of the IEEE.

**Jonathan Smith** is the Olga and Alberico Pompa Professor of Engineering and Applied Science at the University of Pennsylvania, to which he recently returned after almost three years at DARPA. He was elected IEEE Fellow in the Class of 2001 for "contributions to the technology of high-speed networking." He was previously at Bell Telephone Laboratories and Bellcore, which he joined at the AT&T divestiture. His current research interests range from programmable network infrastructures and cognitive radios to architectures for computer-augmented immune response. Dr. Smith serves on the President's Council of Advisors on Science and Technology Network and Information Technology Technical Advisory Group. He was a member of the NRC Committee on Sensing and Communications Capabilities for Special Operations Forces and is a current member of the NRC Board on Army Science and Technology.

# Appendix B

# Meetings and Participating Organizations

## MEETING 1

**December 7-8, 2009**
**The Keck Center of the National Academies**
**Washington, D.C.**

**U.S. Infrared Focal Plane Array (IRFPA) Revitalization**
A. Fenner Milton, Director, Night Vision & Electronic Sensors Directorate, U.S. Army

**Advanced Detector Arrays for Navy Sensors**
Mel Kruer, Senior Scientist, Naval Research Laboratory

**Antenna Coupled IR Detectors**
Glenn Boreman, Trustee Chair Professor of Optics and Electrical Engineering, University of Central Florida

**Recent Progress in HgCdTe IR Detectors at DRS Technologies**
Pradip Mitra, Director, Advanced Development Programs

**Sponsor Presentation**
Kurt M., Science and Technology Analyst, Intelligence Community
Elliott Lehman, Science and Technology Analyst, Defense Intelligence Agency

**Emerging Sensor Technologies for Army Applications**
John Pellegrino, Director, Sensors and Electron Devices Directorate, Army Research Laboratory

**Type II InAs-GaSb Superlattices: A Developing Material System vs. Mercury Cadmium Telluride: The State-of-the-art Infrared Detection Technology**
Manijeh Razeghi, Walter P. Murphy Professor and Director of the Center for Quantum Devices, Northwestern University

**MEETING 2**

**January 20-21, 2010**
**The Keck Center of the National Academies**
**Washington, D.C.**

**Air Force Research in Detector Technologies**
Lyn Brown, Program Manager, Air Force Research Laboratory

**Technology Developments**
Terence Haran, Electro-Optical Systems Laboratory, Georgia Tech Research Institute

**Status and Trends at Semiconductor Devices—Cooled and Uncooled Detectors**
Zvi Kopolovich, Vice President for Programs and Product Lines Management

**Future of Imaging**
John Miller, Chief Technical Officer for FLIR (Forward-looking Infrared) Government Systems Division, FLIR Systems

**Advanced Imager Technology Development at MIT Lincoln Laboratory**
Vyshnavi Suntharalingam, Senior Technical Staff, Advanced Imaging Technology, Lincoln Laboratory

**MEETING 3**

**February 16-18, 2010**
**Albuquerque, N. Mex.**

**Ultimate Sensitivity at Shortwave Infrared Wavelengths Using Single-Photon Detection**
Mark Itzler, Chief Technical Officer, Princeton Lightwave, Inc.

**Insights on Global R&D and Markets**
James “Spider” Marks (Major General, U.S. Army, retired), Managing Partner, Ergo

**Passive Infrared Optimal Waveband Selection and Time Series Processing**
Paul Oglesby, Senior Engineering Fellow, Raytheon Missile Systems

**Novel Nano-injector Detectors: Towards High-resolution Single-photon Imagers at Shortwave Infrared (SWIR)**
Hooman Mohseni, Assistant Professor, Northwestern University

**Infrared Focal Plane Technologies with Emphasis on Multi- and Hyperspectral**
Paul LeVan, Branch Technical Advisor, Air Force Research Laboratory

**Advanced Focal Plane Technology for National Intelligence Missions**
Kurt Lanes, Senior Engineer, Space Missions Program Office, Sandia National Laboratories

**Performance Update on New Generation of Hybrid Silicon, Visible Focal Plane Arrays**
John Hubbs, Chief Scientist, Infrared Radiation Effects Laboratory, Air Force Research Laboratory

**Imaging Infrared Detector Arrays**
Dutch Staplebroek, Research Professor, University of Arizona

**Identifying Key Technologies from a Systems Integration Perspective**
Neil Siegel, Sector Vice President and Chief Engineer, Northrop Grumman Mission Systems

**Detectors for Photon-starved Optical Communications: Present and Future Directions**
Bill Farr, Optical Communications Technology Manager, Optical Communications Group, Jet Propulsion Laboratory

**CMOS (Complementary Metal Oxide Semiconductor) Active Pixel Image Sensors: Status and Future Direction**
Eric Fossum, President, ImageSensors, Inc.

**Next-generation Infrared Detectors: Evaluating Type II Superlattices and Quantum Dots**
Sanjay Krishna, Associate Professor, University of New Mexico

## MEETING 4

**March 9-12, 2009**
**The Beckman Center of the National Academies**
**Irvine, California**

Writing meeting.

# Appendix C

# Background Information on Radiation Hardening for Detectors

## SPACE RADIATION ENVIRONMENT

All focal plane arrays suffer risk of radiation damage in the space environment. The need for radiation hardening of sensors and detectors has become crucial with the strategic goal of creating autonomous spacecraft which rely on on-board information processing. Most satellites are deployed in orbit around the earth. The lowest radiation doses occur in low Earth orbit (LEO), less than 500 km from the surface. Only a few heavy ions penetrate the magnetic fields to this level. In the Polar Regions, there are slightly increased levels due to the Van Allen belts which will allow more heavy ions to penetrate. At geosynchronous orbit, doses are somewhat higher, but still low compared to interplanetary space due to geomagnetic shielding. However, more heavy ions can penetrate to this level than for LEO. As ions penetrate the skin of the space craft, they can emit X-rays. These X-rays will enter the detector and semiconductor materials, and cause the different layers to ionize. This can be temporary, such as corrupting the contents of a memory cell, or permanent, when the ionization triggers latchup in the device. The charge injected into the device will collect at a circuit node and cause data in the device to change.

Electrons and X-rays produce electron-hole pairs which are normally collected at the power supply nodes. The ROIC and the detector respond differently to the ionization. Ionization can cause eventual shifts in MOS transistor thresholds, which causes changes in the device characteristics. When there is no bias on the transistor, almost all of the electron-hole pairs immediately recombine. When there is a positive bias, however, the electrons are swept away while the holes migrate slowly to the negative channel, and become trapped. This causes N-channel enhancement transistors to become easier to turn on, while P-channel transistors are harder to turn on adversely impacting the ROIC performance.

There are four basic categories of radiation vulnerabilities for an integrated circuit. These four effects are Neutron, Total Ionizing Dose (TID), Transient Dose, and Single Event Effect (SEE).

1. **Neutron Effects:** When neutrons strike a semiconductor chip, they displace atoms within the crystal lattice structure. The minority carrier lifetime is reduced because of the increased density of recombination centers. Silicon devices begin exhibiting changes in their electrical characteristics at levels of $1\times10^{10}$ to $1\times10^{11}$ neutrons/cm$^2$. Because bipolar components are minority carrier type devices, neutron radiation affects them at lower doses than for MOS devices. In bipolar integrated circuits, the base transit time and width are the main physical parameters affected. Neutron radiation significantly reduces gain in bipolar devices. MOS devices aren't normally affected until levels of $1\times10^{15}$ neutrons/cm$^2$ are reached.
2. **Total Ionizing Dose Effects:** Total ionizing dose is the accumulation of ionizing radiation over time, typically measured in rads. Slow, steady accumulation of ionization over the life of an integrated circuit causes performance parameters to degrade. Eventually, the device fails. The total dose creates a number of electron-hole pairs in the silicon dioxide layers of MOS devices. As these recombine, they create photocurrents and changes in the threshold voltage that make n-channel devices easier to turn on and p-channel devices more difficult to turn on. Even though some recovery and self-healing takes place in the device, the change is essentially permanent. Some holes created during ionizing pulses are trapped at defect centers near the silicon/silicon oxide interface. Charges induced in the device create a field across the gate oxide sufficiently high to cause the gate oxide to fail, or sufficient carriers are generated in the gate oxide itself to cause failure.
3. **Transient Dose Effects:** A transient dose is a high-level pulse of radiation, typical in a nuclear burst, which generates photocurrents in all semiconductor regions. This pulse creates sudden, immediate effects such as changes in logic states, corruption of a memory cell's content, or circuit ringing. If the pulse is large enough, permanent damage may occur. Transient doses can also cause junction breakdown or trigger latch-up, destroying the device.
4. **Single Event Effect (SEE):** Single event effects typically only affect digital devices significantly. A SEE occurs when a single high-energy particle strikes a device, leaving behind an ionized track. The ionization along the path of the impinging particle collects at a circuit node. If the charge is high

enough, it can create a soft error single event upset (SEU), such as a bit flip, a change in state that causes a momentary glitch in the device output, or corruption of the data in a storage element. A SEE can possibly trigger a device latch-up and burnout. Latch-up occurs when sufficient current is induced in part of the device that it causes the device to latch into a fixed state regardless of circuit input. Burnout occurs when the radiation induces sufficient power dissipation to cause catastrophic device failure. Burnout often occurs as a result of latch-up. SEEs can wreak havoc on satellites, spacecraft, and aircraft as well. Therefore, circuits used in aerospace controls systems must be protected from potentially disastrous SEEs.

a. **Single Event Upsets (SEU):** These are also known as soft errors that occur due to either the deposition of depletion of charge by a single ion at a circuit node, causing a change of state in a memory cell. In very sensitive devices, a single ion hit can also cause multiple-bit upsets (MBUs) in adjacent memory cells. This type of event causes no permanent damage and the device can be reprogrammed for correct function after such an event has occurred.
b. **Single Event Latch-up (SEL):** This can occur in any semiconductor device which has a parasitic n-p-n-p path. A single heavy ion or high energy proton passing through either the base emitter junction of the parasitic n-p-n transistor, or the emitter-base junction of the p-n-p transistor can initiate regenerative action. This leads to excessive power supply current and loss of device functionality. The device can burnout unless the current is limited or the power to the device is reset. SEL is the most concern in bulk CMOS devices.
c. **Single Event Snapback:** This is also a regenerative current mechanism similar to SEL, but a device does not need to have a p-n-p structure. It can be triggered in a n-channel MOS transistor with large currents, such as IC output driver devices, by a single event hit-induced avalanche multiplication near the drain junction of the device.
d. **Single Event-Induced Burnout (SEB):** This event may occur in power MOSFETs when the passage of a single heavy ion forward biases the thin body region under the source of the device. If the drain-to-source voltage of the device exceeds the local breakdown voltage of the parasitic bipolar, the device can burn out due to large currents and high local power dissipation.
e. **Single Event Gate Rupture (SEGR):** This has been observed due to heavy ion hits in power MOSFETs when a large bias is applied to the gate, leading to thermal breakdown and destruction of the gate oxide. It can also occur in nonvolatile memories such as EEPROM

during write or erase operations, the time when high voltage is applied to the gate.

During the past several decades, several companies have developed manufacturing processes to produce a range of rad-hard electronic products. These processes are somewhat different from the ones used in commercial foundries because they include a few modified process steps that produce circuits with greater radiation resistance. These parts are more expensive than their commercial counterparts and have lagged several generations behind in terms of processing speed, power, and size. Moreover, many companies that were in the business of supplying rad-hard components a decade ago have dropped out of the market. Only two foundries remain active today—Honeywell and Sandia National Laboratories.

The high cost of maintaining dedicated foundries to create space electronics has motivated an exploration of alternatives for next-generation space systems. One strategy in particular has been gaining popularity in recent years. Known as radiation hardening by design (RHBD), this approach relies solely on circuit design techniques to mitigate the damage, functional upsets, and data loss caused by space radiation.[1]

Different aspects of this approach have been in use for some time, but most frequently in combination with dedicated rad-hard manufacturing facilities. More recently, a number of research institutions and corporations have demonstrated the basic feasibility of RHBD using standard commercial foundries; however, to satisfy the military's need for a wide range of part types and hardness levels, a self-sustaining RHBD infrastructure must be established, and the RHBD approach must be proven robust enough to use without some degree of fabrication process control.[2]

The manufacturing processes used to build commercial electronic components in the 1970s and 1980s were severely inadequate to meet the needs of the space community. But as commercial CMOS processes have advanced, the inherent radiation resistance of these devices has improved—and thus, the RHBD approach has become more feasible.[3] For example, the current that flows through CMOS transistors is governed by a low-voltage gate over each device, isolated by a layer of oxide. These insulating layers can develop a charge after long exposure to ionizing radiation, and this charge can affect the flow of current through the device; but

[1]D.R. Alexander, D.G. Mavis, C.P. Brothers, and J.R. Chavez, "Design Issues for Radiation Tolerant Microcircuits in Space," *1996 IEEE Nuclear and Space Radiation Effects Conference* (NSREC) Short Course, V-1 (1996).

[2]Ibid.

[3]Ibid.

as circuits have shrunk, the thicknesses of these insulating layers have decreased, presenting a decreased opportunity for charge buildup.

The radiation-induced increases in leakage current result in unregulated current flowing across unintended areas of the semiconductor.[4,5,6] When leakage current bypasses the transistor's isolated regions, it degrades the ability to distinguish the transistor's "on" and "off" states. Leakage also increases the circuit's background current, or the amount of current flowing when the device is in a quiescent state. Such an increase, multiplied by the tens of millions of switches in each circuit, can drive up power consumption, increasing heat-dissipation needs and prematurely draining the power source of the satellite. In an extreme case, the isolation between discrete components can also be lost, rendering the circuit useless.[7,8,9]

There are interface areas which are prone to leakage in a radiation environment such as edges of the transistors where the thin gate oxide abuts with the much thicker field oxide. The process traditionally used to manufacture the transistor borders can induce significant material stress, which may facilitate the increase in leakage current when exposed to radiation. The newest isolation-oxide manufacturing processes impart less stress and seem to have achieved a greater inherent radiation resistance.

It has been shown that total-dose effects can be mitigated by designing transistors in an enclosed shape, thereby eliminating the edges that can trigger current leakage along the borders of conventional transistors. Current flows from the center to the outside of these devices, making them immune to edge leakage current, but requiring a larger area for each transistor. Furthermore, transistor-to-transistor leakage can be reduced by incorporating guard bands around individual transistors

[4]Ibid.

[5]G. Anelli, M. Campbell, M. Delmastro, F. Faccio, S. Florian, A. Giraldo, E. Heijne, P. Jarron, K. Kloukinas, A. Marchioro, P. Moreira, and W. Snoeys, "Radiation Tolerant VLSI Circuits in Standard Deep Submicron CMOS Technologies for the LHC Experiments: Practical Design Aspects," *IEEE Transactions on Nuclear Science*, Vol. 46, pp. 1690-1696 (1999).

[6]T. Calin, M. Nicolaidis, and R. Valazco, "Upset Hardened Memory Design for Submicron CMOS Technology," *IEEE Transactions on Nuclear Science*, Vol. 43, pp. 2874-2878 (1996).

[7]D.R. Alexander, D.G. Mavis, C.P. Brothers, and J.R. Chavez, "Design Issues for Radiation Tolerant Microcircuits in Space," *1996 IEEE Nuclear and Space Radiation Effects Conference* (NSREC) Short Course, V-1 (1996).

[8]G. Anelli, M. Campbell, M. Delmastro, F. Faccio, S. Florian, A. Giraldo, E. Heijne, P. Jarron, K. Kloukinas, A. Marchioro, P. Moreira, and W. Snoeys, "Radiation Tolerant VLSI Circuits in Standard Deep Submicron CMOS Technologies for the LHC Experiments: Practical Design Aspects," *IEEE Transactions on Nuclear Science*, Vol. 46, pp. 1690-1696 (1999).

[9]T. Calin, M. Nicolaidis, and R. Valazco, "Upset Hardened Memory Design for Submicron CMOS Technology," *IEEE Transactions on Nuclear Science*, Vol. 43, pp. 2874-2878 (1996).

or groups of transistors.[10] Other novel techniques are being applied to conventional transistor switches to boost their immunity to total ionizing dose radiation. These techniques consume area in the design, thereby reducing the total number of transistors available for a given circuit function and increasing the capacitance, and thus the power consumption, of the circuit. The trade-off may be worthwhile: Using RHBD, several manufacturers have demonstrated radiation hardness in excess of 20 Megarads using commercial CMOS foundries, making them suitable for use in nuclear reactors as well as severe space environments.[11,12] Companies such as Raytheon Vision systems, Teledyne Imaging Systems and BAE are all experienced in manufacturing RHBD products for space applications.

Single-event upsets require different mitigation techniques. Single-event upsets occur when energetic particles deposit charge into memory circuits, causing stored data to change state (from a "1" to a "0," for example). As circuits shrink and transistor volumes become smaller, the total charge needed to cause an upset in a circuit element decreases. Thus, even protons moving through the circuit may deposit sufficient charge to disrupt sensitive locations. Susceptibility to single-event upsets can be reduced by increasing the amount of charge needed to trigger a bit flip or by providing feedback resistors that give the circuit time to recover from a particle strike. Perhaps the most common approach is to use redundant information storage or error-checking circuitry. For example, a technique known as "voting logic" can be used to catch and correct potential errors in latches. With this technique, a single latch does not affect a change in bit state; rather, several identical latches are queried, and the state will only change if the majority of latches are in agreement. Thus, a single latch error will be "voted away" by the others.

Another technique useful for overcoming single-event upsets is known as "error detection and correction." In this technique, the system architecture provides extra check bits in each stored word in memory; when these extra bits are read and interrogated, errors become apparent and can be corrected. Perhaps the simplest approach would be to insert a single bit that denotes whether the content of a word has an even or odd parity; this requires minimal overhead, but does not automatically identify the location of any observed errors. On the other hand, to uniquely

[10] G. Anelli, M. Campbell, M. Delmastro, F. Faccio, S. Florian, A. Giraldo, E. Heijne, P. Jarron, K. Kloukinas, A. Marchioro, P. Moreira, and W. Snoeys, "Radiation Tolerant VLSI Circuits in Standard Deep Submicron CMOS Technologies for the LHC Experiments: Practical Design Aspects," *IEEE Transactions on Nuclear Science*, Vol. 46, pp. 1690-1696 (1999).

[11] R.C. Lacoe, J.V. Osborn, R. Koga, S. Brown, and D.C. Mayer, "Application of Hardness-By-Design Methodology to Radiation-Tolerant ASIC Technologies," *IEEE Transactions on Nuclear Science*, Vol. 47, pp. 2334-2341 (2000).

[12] J.V. Osborn, R.C. Lacoe, D.C. Mayer, and G. Yabiku, "Total-Dose Hardness of Three Commercial CMOS Microelectronics Foundries," *IEEE Transactions on Nuclear Science*, Vol. 45, pp. 1458-1463 (1998).

detect and correct a single error in a 16-bit word using the common "Hamming code" method requires the insertion of six additional bits. Thus, the error detection and correction technique requires a significantly greater number of memory bits to store a given amount of information.

Circuit designers use computer-aided design tools to define and verify the final circuit layout, to perform logical simulation of the design, to identify potential failure modes, and to perform static and dynamic timing simulations. These tools use so-called "cell libraries" to simplify the design process as much as possible. Each library is a collection of individual circuit elements that includes functional and performance information about each element. Effective use of RHBD requires that knowledge of the behavior of the circuits in the space environment be incorporated into the computer-aided design tools. For instance, the programs would need to simulate the electrical behavior of the transistor switch in a radiation environment based on the structure of the device and the physics of the radiation interactions.[13]

Rad-hard cell libraries are developed and maintained such that they will include provisions for reliable operation in harsh environments. A number of cell libraries will probably be needed for each CMOS generation to meet the needs of a range of space programs operating in various orbits, and with a range of reliability, survivability, and cost requirements. Funding for libraries with the most stringent requirements—and thus the smallest markets—must be generated by the customer community.[14]

Commercial foundries typically provide the starting material for all electronic components manufactured in their processing facilities; however, nonstandard starting materials incorporating epitaxial layers or insulating substrates, for example, may enhance radiation immunity. The part supplier and the selected foundry may agree to substitute appropriate starting materials to provide additional levels of radiation hardness.

Each foundry typically uses proprietary procedures developed over many years; however, nonstandard processing steps involving, for example, novel implants or modifications of layer thicknesses may help enhance radiation immunity. In this approach the RHBD part supplier and the selected foundry may agree to substitute or augment appropriate manufacturing steps to provide additional levels of radiation hardness.

NASA has been employing design-hardening concepts in various projects. The Europa satellite, for example, will be exposed to more than 6 megarads over the life of the mission. To meet this high total-dose requirement, NASA is using rad-hard

[13]Available at http://www.aero.org/publications/crosslink/summer2003/index.html. Last accessed on March 25, 2010.

[14]Ibid.

processors along with several digital and analog circuits designed using redundancy and other RHBD techniques.

For the ROIC radiation hardness by design has quickly evolved from a concept to a strategy that may well redefine the way electronic components are procured for defense space systems. Companies have demonstrated that RHBD techniques can provide immunity from total-dose and single-event effects in commercially produced circuits. CAD tools that can model these radiation effects and cell libraries that use a range of these techniques have been developed at a number of government agencies, universities, and private companies during the past several years, culminating in the commercial production of RHBD memories, microprocessors, and application-specific integrated circuits that are being specified. The infrastructure needed to make RHBD a mainstream procurement approach is gradually being developed.[15]

In addition to the ROIC performance, the detector performance is also impacted by radiation due to different layers ionizing, increase in dark currents due to charge carrier generation etc. The detector response depends on numerous factors such as detector material type, growth process, detector design, detector fabrication, and defects arising thru these different steps. The detector response is not characterized in isolation but together with the ROIC as a FPA. An example is observed by Hubbs et al where they discuss changes in the lateral diffusion length in HgCdTe detectors in a proton environment for LWIR detectors. They found that the non-ionizing energy loss (NIEL) in HgCdTe provides a frame work for estimating responsivity degradation in LWIR MCT detectors due to on-orbit exposure from protons. They found that their comparison of the responsivity degradation at different proton energies suggested that the atomic columbic interaction of the protons with the MCT detector is likely the primary mechanism responsible for the degradation in the responsivity at proton energies below 30 MeV.[16]

[15]Available at http://www.aero.org/publications/crosslink/summer2003/index.html. Last accessed on March 25, 2010.

[16]John E. Hubbs et al., "Lateral Diffusion Length Changes in HgCdTe Detectors in a Proton Environment," *IEEE Transactions on Nuclear Science*, Vol. 54, no. 6 (2007), p. 2435.