Digital Filters

Principles and Applications with MATLAB®

FRED J. TAYLOR

IEEE
IEEE PRESS

IEEE SERIES ON
DIGITAL
& MOBILE
COMMUNICATION
John B. Anderson, *Series Editor*

MATLAB
examples

WILEY

DIGITAL FILTERS
Principles and Applications with MATLAB

FRED J. TAYLOR

IEEE SERIES ON
DIGITAL
& MOBILE
COMMUNICATION

John B. Anderson, *Series Editor*

IEEE Series on Digital and Mobile Communication

IEEE

IEEE PRESS

WILEY

A JOHN WILEY & SONS, INC., PUBLICATION

Library of Congress Cataloging-in-Publication Data:

Taylor, Fred J., 1940-
Digital filters : principles and applications with MATLAB / Fred J. Taylor.—1st ed.
p. cm.—(IEEE series on digital & mobile communication ; 30)
Includes index.
ISBN 978-0-470-77039-9 (hardback)
1. Electric filters, Digital. 2. Signal processing–Digital techniques–Mathematics. 3. MATLAB. I. Title.
TK7872.F5T39 2012
621.3815'324–dc23

2011021434

Printed in the United States of America
oBook ISBN: 9781118141151
ePDF ISBN: 9781118141120
ePub ISBN: 9781118141144
eMobi ISBN: 9781118141137

10 9 8 7 6 5 4 3 2 1

To my angel, Lori

PREFACE

The history of digital filters essentially began in the mid-1970s, concurrent with the advent of the field of study called digital signal processing (DSP). Over the ensuing 30 something years, digital filters have become both a facilitating and enabling technology. They serve as analog replacements as well as serving in unique DSP roles in a host of application domains including communications, control, defense, audio, biomedicine, geophysics, radar, entertainment, and others. I have been blessed to be able to witness and participate in all of these phases of digital filter evolution.

A digital filter is a device that can modify the attributes of a signal using digital means. Required filter attributes can be assumed or defined in terms of published standards that specify amplitude and phase behavior as a function of frequency. Besides altering a signal's attributes, digital filters must often meet a host of other constraints such as speed, complexity, power consumption, cost, and other factors. In the pantheon of digital filters, the majority are identified as being finite impulse response (FIR), infinite impulse response (IIR), or multirate systems. The book's primary goal is to provide the needed understanding of both design and analysis strategies as they apply to mainstream digital filters.

In the normal course of an engineer's career, regardless of their disciplinary training, they will be called upon to design or analyze a mainstream filter. Unfortunately, many engineers and technologists have little to no formal digital filter experience. Fortunately, today's workplace is abundant with filter design software packages with various levels of sophistication. One of the leaders in this field is Mathwork's MATLAB™. Today, both practicing engineers and students of engineering exhibit a growing reliance on these tools with MATLAB being a de facto standard. However, after observing how these tools are being used in the workplace and classroom, concerns arise in that users are often overwhelmed with a plethora of filter design options, often developing a filter solution that may not be best for the target application. In addition, users often have insufficient experience or understanding of filter theory to be able to make even minor enhancements to a MATLAB-produced filter outcome. This too is a motivation for developing this book, which elevates the reader's understanding of how to characterize a digital filter, to make proper design choices, and to enhance a computer-generated design into a well-crafted outcome.

In reality, using tools such as MATLAB to design a mainstream digital filter is the easiest step in a solution process that ends with a successfully implemented digital filter. Implementation, whether in software or hardware, is generally the more challenging problem. Tools, such as MATLAB, provide the user with some basic implementation support. Unfortunately, most engineers have no, or only a rudimentary, understanding of the implementation choices offered by MATLAB. This provides additional motivation to develop filter implementation awareness skills, providing content that is generally missing in the current collection of digital filter books and monographs.

The book has been organized to support the stated objectives. The presentation begins with the fundamentals, including sampling, data acquisition, data conversion and quantization, and transforms. Next, the design, implementation, and analysis of an FIR filter are presented. Topics include FIR attributes, types, special cases, and implementation. Following FIRs, the design, implementation, and analysis of an IIR filter are presented. Like FIRs, topics include IIR attributes, types, special cases, and implementation. Additional attention is given to understanding state variables as an IIR architectural description language. Finally, multirate systems are explored, ranging from a discussion of their properties to case studies. In most cases, each topic is supported with MATLAB examples and exhibits.

The study of filters is supported with a number of examples, many involving the use of MATLAB. In an attempt to actively engage the reader, the MATLAB script used to generate the MATLAB examples and graphics are available from John Wiley & Sons Supplemental Book Material site at http://booksupport.wiley.com. The MATLAB scripts can be easily copied into MATLAB's Command Window and reparameterized to reflect the reader's filter applications and needs. Many of the scripts were polished by Mr. Rajneesh Bansal, to whom I owe a great debt.

Fred J. Taylor
IEEE Fellow
Professor Emeritus, University of Florida
Board Chairman and Senior Scientist, The Athena Group Inc.

CONTENTS

CHAPTER 1

INTRODUCTION TO DIGITAL SIGNAL PROCESSING

INTRODUCTION

Signal processing refers to the art and science of creating, modifying, manipulating, analyzing, and displaying signal information and attributes. Since the dawn of time, man has been the quintessential signal processor. Human signal processing was performed using one of the most powerful signal processing engines ever developed: the 25-W human brain that commits about 10 W to information processing. In that context, this biological processor is comparable to the Intel mobile Pentium III processor. As humans evolved, other agents were added to man's signal processing environment and repertoire, such as information coding in the form of intelligible speech, art, and the written word. In time, communication links expanded from local to global, global to galactic. It was, however, the introduction of electronics that enabled the modern information revolution. Analog electronics gave rise to such innovations as the plain old telephone system (POTS), radio, television, radar/sonar, and a host of other inventions that have revolutionized man's life and landscape. With the introduction of digital technologies over a half century ago, man has witnessed a true explosion of innovations that has facilitated the replacement of many existing analog solutions with their digital counterparts. In other instances, digital technology has enabled solutions that previously never existed. Included in this list are digital entertainment systems, digital cameras, digital mobile telephony, and other inventions. In some cases, digital technology has been a disruptive technology, giving rise to products that were impossible to envision prior to the introduction of digital technology. An example of this is the now ubiquitous personal digital computer.

ORIGINS OF DIGITAL SIGNAL PROCESSING (DSP)

Regardless of a signal's source, or the type of machine used to process that information, engineers and scientists have habitually attempted to reduce signals to a set of parameters that can be mathematically manipulated, combined, dissected, analyzed,

Digital Filters: Principles and Applications with MATLAB, First Edition. Fred J. Taylor.

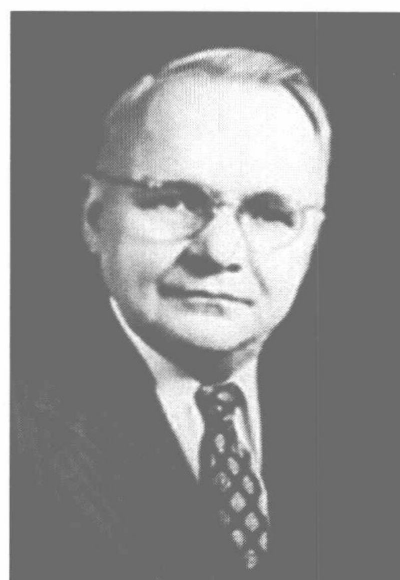
Published 2012 by John Wiley & Sons, Inc.

or archived. This obsession has been fully realized with the advent of the digital computer. One of the consequences of this fusion of man and machine has been the development of a new field of study called *digital signal processing*, or DSP. Some scholars trace the origins of DSP to the invention of iterative computing algorithms discovered by the ancient mathematicians. One early example of a discrete data generator was provided in 1202 by the Italian mathematician Leonardo da Pisa (a.k.a. Fibonacci*). Fibonacci proposed a recursive formula for counting newborn rabbits, assuming that after mating an adult pair would produce another pair of rabbits. The predictive Fibonacci population formula is given by $F_n = F_{n-1} + F_{n-2}$ for the initial conditions $F_0 = 1$, $F_{-1} = 0$, and produces a discrete-time sequence that estimates the rabbit population $\{1, 1, 2, 3, 5, 8, 13, 21, 34, 55, \ldots\}$ as a function of discrete-time events. However, those who promote such action as evidence of DSP are overlooking the missing "D-word." DSP, at some level, must engage digital technology in a signal processing activity.

The foundations of DSP were laid, in fact, in the first half of the 20th century. Two agents of change were Claude Shannon and Harry Nyquist. They both formulated the now celebrated sampling theorem that described how a continuous-time signal can be represented by a set of sample values. Such representations were found to be so mathematically perfect that the original signal could be reconstructed from a set of sparsely distributed samples. Nyquist conjectured the sampling theorem in 1928, which was later mathematically demonstrated by Shannon in 1949. Their work

Claude Shannon (1916–2001)

Harry Nyquist (1889–1976)



* Fibonacci is short for *filius Bonacci*, son of Bonacci, whose family name means "good stupid fellow."

provided the motivation and framework to convert signals from a continuous-time domain to and from the discrete-time domain. The sampling theorem, while being critically important to the establishment of DSP, was actually developed prior to the general existence of digital technology and computing agents. Nevertheless, it was the sampling theorem that permanently fused together the analog and discrete-time sample domain, enabling what is now called DSP.

During the 1950s, and into the 1960s, digital computers first began to make their initial appearance on the technology scene. These early computing machines were considered to be far too costly and valuable to be used in the mundane role of signal analysis, or as a laboratory support tool by lowly engineers. In 1965, Cooley and Tukey introduced an algorithm that is now known as the fast Fourier transform (FFT) that changed this equation. The FFT was indeed a breakthrough in that it recognized both the strengths and weaknesses of the classic von Neumann general-purpose digital computer architecture of the day, and used this knowledge to craft an efficient code for computing Fourier transforms. The FFT was cleverly designed to distribute data efficiently within conventional memory architectures and perform computation in a sequential manner. Nevertheless, early adopters of the FFT would not necessarily have considered themselves to be DSP engineers since the field of DSP had yet to exist.

Since the introduction of the FFT, digital computing has witnessed a continuous growth, synergistically benefiting from the increasing computing power and decreasing cost of digital technologies in accordance with Moore's law.* The digital systems available in the 1970s, such as the general-purpose minicomputer, were capable of running programs that processed signals in an off-line manner. This process was often expensive, time-consuming, required considerable programming skills, and generally remained compute bound, limiting the type of applications that could be considered. During this epoch, early attempts witnessed the use of dedicated digital logic to build rudimentary digital filters and radar correlators for national defense purposes. These activities caused engineers and scientists to recognize, for the first time, the potential of DSP even though there was no formal field of study called DSP at that time. All this, however, was about to change.

In 1979, a true (albeit quiet) revolution began with the introduction of the first-generation DSP microprocessor (DSP μp) in the form of the Intel 2920, a device called an "analog signal processor" for marketing reasons. The 2920 contained on-chip analog-to-digital converter (ADC)/digital-to-analog converter (DAC), and a strengthened arithmetic unit that was able to execute any instruction in 200 μs. While initiating a fundamentally important chain of events that led to the modern DSP μp, by itself, the 2920 was a marketplace disappointment appearing in a few 300 b/s modems. It was, nevertheless, warmly embraced by a small but active group of digital audio experimenters. With the second generation of DSP μp (e.g., Texas Instruments TMS320C10), DSP technology exposed its potential value in a host of new applications. For the first time, products with embedded DSP capabilities became a practical reality establishing DSP as an enabling technology. The field,

* Moore's law $(Nt/A)(t_1) = (Nt/a)(t_0) \times 1.58^{(t_1-t_0)}$ predicts that semiconductor density will double every 18 months.

now called DSP, rapidly developed in the form of academic programs, journals, and societies, and developing infrastructure technology. These beginnings swiftly gave way to a third and fourth generation of general-purpose DSP μp as well as custom DSP devices. Even though DSP remains a relatively young science, being only a few decades old, it has become both a major economic and technological force. DSP solutions are now routinely developed using commercial off-the-shelf (COTS) software and DSP μps and field programmable gate arrays (FPGAs), along with application-specific integrated circuits (ASICs).* There is now an abundance of DSP software design and development tools that serve this industry. Through the intelligent use of these resources, DSP has become an enabling technology for high-speed, low-cost data communications (modems), digital controllers, wireless solutions including cellular telephony and other personal communications services, video compression, multimedia and entertainment (audio and video solutions), plus a host of other applications. At the core of this revolution are the tens of thousands of scholars and technologists who now refer to themselves as DSP engineers and scientists. These engineers are hybrids in that they need to have competence in the application area they serve, to possess strong computer hardware and/or software skills, plus to have an understanding of the theory and practice of DSP. They, like DSP technology itself, are still in the formative stage. All that can be accurately predicted at this time is that DSP will be one of the principal technological driving forces of the 21st century economy.

SIGNAL DOMAINS

Signals are abundant in both the natural and artificial worlds. Nature, in particular, is rich in signals, from cosmic rays, bird trills, and the proverbial tree falling in the woods. Signals found in the natural world are produced by a variety of mechanisms. An example is a biological electrocardiogram (EKG) signal illustrated in Figure 1.1.

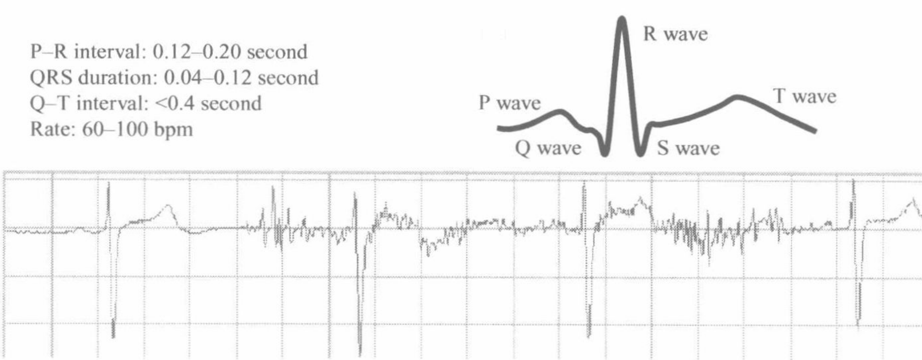


Figure 1.1 Typical EKG model (top) showing the P, Q, R, S, and T phases. Shown also is a typical EKG recording (bottom).

* ASICs are defined by the solution provider, not the manufacturer. A variation on this theme is the application-specific standard parts (ASSPs), which are essentially ASICs developed for high-volume commercial sale.

Figure 1.2 Two-dimensional image "Lena" as a $512 \times 512 \times 8$ bit/pixel image on the left and JPEG compressed image on the right.

Others include man-made signals such as those generated by musical instruments or the human voice. Artificial signals can be created without a natural production mechanism but using algorithms and electronic sound reproduction equipment. Signals from these domains can be classified as being one-, two-, and M-dimensional. A simple one-dimensional sinusoid $x(t) = \cos(\omega t)$ that can be expressed as an amplitude versus time trajectory. Images are often classified as being two-dimensional signals that can be expressed as a function of two spatial parameters, such as $f(x_1,x_2) = \cos(\omega x_1 + \omega x_2)$. A simple black and white image can be represented as an array of image values $i(x,y)$, where $i(x,y)$ is the image intensity at coordinates (x,y) (see Fig. 1.2). Signals of higher dimension also exist. The Dow Jones industrial average, for example, is a function of 30 economic variables and represents a multidimensional signal.

Causality is also an important signal property. Causal signals are produced by causal systems (nonanticipative systems) where the output signal output $y(t)$, at some specific instant t_0, depends only on the system input $x(t)$ for $t \leq t_0$. Signals that are not causal are called noncausal, anticausal signals, or anticipatory.* While noncausal signals are not products of a physically realizable signal generation mechanism, they will play a significant role in the mathematical study of signals and systems or in performing off-line simulations. For example, the signal $x(t) = \cos(\omega_0 t)$ technically persists for all time (i.e., $t \in (-\infty, \infty)$) and is therefore noncausal. It is also recognized that $x(t) = \cos(\omega_0 t)$ represents a mathematically important signal even though it could not have been created by a physical signal generator.

SIGNAL TAXONOMY

Any meaningful signal taxonomy needs to recognize that signals can live in different domains. Specifically, there are three important signal domains, and they are developed below.

* He who answers before listening—that is his folly and his shame—Proverbs 18:13.

Continuous-Time Signals

Continuous-time or analog signals $x(t)$ are defined as a continuum of points in both independent and dependent variables. Both the signal's amplitude $x(t)$ and time instance t are numbers known with infinite precision. Continuous-time signals can be further portioned into differentiable, analytic, piecewise differentiable, continuous, piecewise constant, as well as others.

Discrete-Time Signals

A discrete-time signal $x[k]$ is obtained by sampling a continuous-time signal $x(t)$ with an ideal sampler. Specifically, if T_s denotes the sample period, then $f_s = 1/T_s$ is called the sample rate or sample frequency and is measured in sample per second, denoted "Sa/s," although "Hz" is commonly used interchangeably with Sa/s. The sample value at the sample instance $t = kT_s$ is denoted $x(t = kT_s) = x[k]$. A collection of such sample values is called a time series. A discrete-time series consists of sample values that are continuously resolved along the dependent axis (amplitude) and discretely resolved along the independent axis (time). Whereas the sample instances are discrete (in time), the sample value $x[k]$ is an infinite precision real or complex number. Discrete-time signals can be physically created by passing a continuous signal through an electronic device called an impulse sampler or ideal sampler. The sampled value $x[k]$ can be used to construct a continuous-time signal $y(t)$ using an inverse process called interpolation. An example interpolator is a called sample-and-hold (S/H) circuit, as shown in Figure 1.3. S/H circuits are commonly found in the design of DACs. Other forms of interpolation are possible.

Computing algorithms, such as those studied in discrete mathematics, can be used to produce discrete-time signals. In addition, discrete-time series also arise in the fields of economics, biology, calculus, statistics, physics, plus others. The engineering importance of discrete-time signals can be traced back to a post-World War II era in the form of sampled data control systems and telephony. During the early days of the Cold War, strategic bombers were flying missions having long time durations while attempting to navigate with a high degree of accuracy. This was a challenging problem for the day's analog control systems. Small drifts in the control surface signal values could accumulate over time, resulting in large positional errors. What was required was a more precise autopilot technology. It was discovered that if the control signal was sampled and modulated, drift-free alternating current (AC) amplifiers could replace the troublesome drift prone direct current (DC) amplifiers. The simplest modulation scheme is an alternating sign periodic analog pulse train using a device that was called a "chopper." The modulated signal was then transmit-

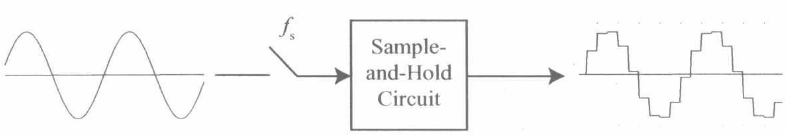


Figure 1.3 Discrete-time signal sampler and sample-and-hold circuit.

ted to a receiver where it was demodulated, returning all the samples to their original sign. This gave rise to a technology called sampled-data control, which had a strong following in the 1950s and 1960s.

The most enduring technology emerging from this era is found in telephony. It was discovered that a number of distinct discrete-time time series could be interlaced (i.e., time-division multiplexed) onto a common channel, thereby increasing the channel's capacity in terms of the number of subscribers per line per unit time. The result was that the telephone company could bill multiple clients for using a single copper line. Claude Shannon developed the mathematical framework by which these signals can be time multiplexed, transmitted along a common line, and reconstructed at each individual receiver. Shannon's innovation, known as Shannon's sampling theorem, has been a driving force behind most DSP techniques and methodologies.

Digital Signals

Digital signals are discrete-time signals that are also quantized along the dependent axis (amplitude). Digital signals can be produced by a digital computer using finite precision arithmetic or by passing an analog signal $x(t)$ through an ADC or A/D, also producing a finite precision approximation of a discrete-time signal. In either case, quantizing the amplitude of the original signal introduces an uncertainty called quantization error. Controlling and managing such errors is often critical to the successful design of a DSP solution.

A general signal taxonomy is presented in Figure 1.4. Contemporary signal processing systems typically contain a mix of analog, discrete, and digital signals and systems. A signal's original point of origin is often the continuous-time or analog domain. Digital signals, however, are becoming increasingly dominant in this mix. Applications that were once considered to be exclusively analog, such as sound

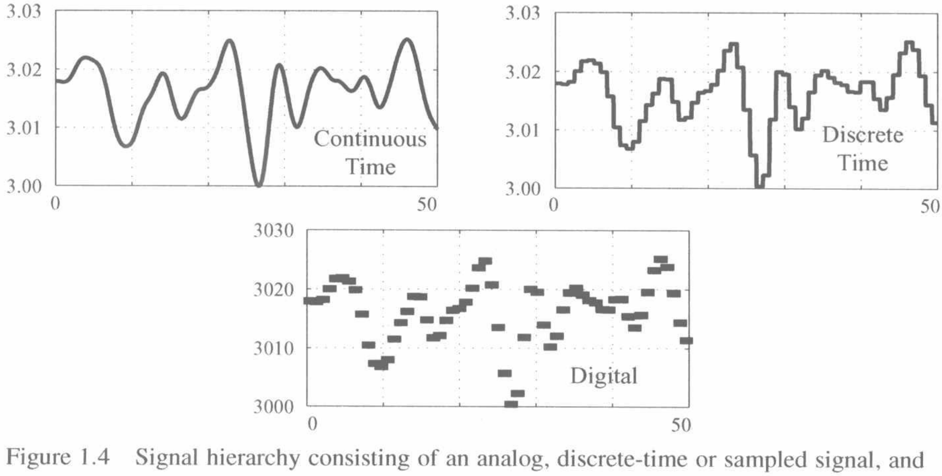


Figure 1.4 Signal hierarchy consisting of an analog, discrete-time or sampled signal, and digital or quantized signal process.

recording and reproduction, have become digital. The wireless communications industry are replacing analog components in radios, as well as back-end audio and signal decoding sections, with digital devices. Images and video signals are now routinely coded and decoded as digital signals. Discrete-time systems, as defined, are rarely found in use today except as part of the sampling subsystems (samplers) found in ADCs. The reason for this paradigm shift from analog signal processing to DSP is primarily due to two breakthroughs. The first is the sampling theorem and the second is the product of the fruitful digital semiconductor industry. Once this bridge was crossed, it became logical to replace everything possible with digital technology.

DSP: A DISCIPLINE

Signal processing is a gift of the sampling theorem and formidable armada of companion theories, methodologies, and tools, such as the celebrated FFT. Initially DSP was envisioned simply as an analog replacement technology. It is now clearly apparent to many that DSP will move into new areas and become the dominant signal processing technology in the 21st century. DSP has matured to the point where it can claim to be an academic discipline replete with a rich infrastructure industry. DSP continues to gain semiconductor market share mainly because it can deliver solutions. The DSP advantage is summarized below.

Digital Advantages

The attributes of a digital solution are as follows:

- Both analog and digital systems can generally be fabricated as highly integrated semiconductor systems. Compared with analog circuitry, digital devices can take full advantage of submicron technologies and are generally more electronically dense, resulting in both economic and performance advantages.
- As semiconductor technologies shrink (deep submicron) and signal voltages continue to decline (1.25 V and lower), the intrinsic signal-to-noise ratio found at the transistor level decreases. Digital systems are far more tolerant of such internal noise. These devices, however, are essentially useless as an analog system (e.g., equivalent 3-bit precision per transistor).
- Digital systems can operate at extremely low frequencies, which would require unrealistically large capacitor and resistor values if implemented as an analog solution.
- Digital systems can be designed with increased precision with only an incremental increase in cost, whereas the precision of an analog system precision is physically limited (10 bits ~ 60-dB dynamic range typical).
- Digital systems can be easily programmed to change their function whereas reprogramming analog systems is extremely difficult.

- Digital signals can be easily delayed and/or compressed, an effect which is difficult to achieve with analog signals.
- Digital systems require no external alignment, while analog systems need periodic adjustment (due to temperature drift, aging, etc.).
- Digital systems do not have impedance-matching requirements, while analog systems do.
- Digital systems, compared with analog devices, are less sensitive to additive noise as a general rule.

Analog Advantages

There are, however, a few analog attributes that resist the digital challenge. They are as follows:

- Analog systems can operate at extremely high frequencies (e.g., microwave and optical frequencies) that exceed the maximum clock rate of a digital device or ADC.
- Analog solutions are sometimes more cost effective (e.g., first-order resistor-capacitor [RC] filter) compared with solutions fashioned with digital components (e.g., ADC, digital filter, plus DAC).

Driven by advancements in semiconductors, software, and algorithms, DSP will be a principal enabling and facilitating technology in the following areas:

Audio
Audio-video receivers
Computing
Digital radio
Home audio
Flat panel displays
Internet audio
Pro audio
Speech
Toys

Transportation
Chassis sensors
Power train
Driver displays
Security systems
Safety systems

Broadband
Wireless local area network (LAN)
Cable
Digital subscriber line (DSL)
Voice-over Internet protocol (VoIP)

Control
Digital power supplies
Embedded sensors
Industrial drives
Motors
Instrumentation

Medical
Automated external defibrillators
Monitoring
Hearing aids
Imaging
Prosthetics

Military
Avionics
Countermeasures
Imaging
Munitions
Navigation
Radar/sonar

Wireless
Handsets
Infrastructure
Radiofrequency tagging
Security
Biometrics
Smart sensors

Telecom
High-frequency radios
Infrastructure