

The Breadth and Depth of DSP

Digital Signal Processing is one of the **most powerful technologies** that will **shape science and engineering** in the twenty-first century. Revolutionary changes have already been made in a *broad* range of fields: communications, medical imaging, radar & sonar, high fidelity music reproduction, and oil prospecting, to name just a few. Each of these areas has **developed a deep DSP technology**, with its **own algorithms, mathematics, and specialized techniques**. This **combination of breadth and depth** makes it **impossible for any one individual to master all of the DSP technology** that has been developed. DSP education involves two tasks: **learning general concepts that apply to the field as a whole**, and **learning specialized techniques for your particular area of interest**. This chapter starts our journey into the world of Digital Signal Processing by describing the **dramatic effect** that DSP has made in several diverse fields. **The revolution has begun.**

The Roots of DSP

数据的独特类型

Digital Signal Processing **is distinguished from** other areas in computer science by the unique type of data it uses: **signals**. In most cases, these signals originate as **sensory data from the real world**: seismic vibrations, visual images, sound waves, etc. **DSP** is the **mathematics**, the **algorithms**, and the **techniques** used to **manipulate these signals** after they have been converted into a digital form. This includes a wide variety of goals, such as: **enhancement** of visual images, **recognition and generation** of speech, **compression** of data for storage and transmission, etc. Suppose we attach an analog-to-digital converter to a computer and use it to acquire a chunk of real world data. DSP answers the question: *What next?*

The roots of DSP are in the 1960s and 1970s when digital computers first became available. Computers were expensive during this era, and DSP was limited to only a few critical applications. Pioneering efforts were made in **four key areas**: **radar & sonar**, where national security was at risk; **oil exploration**, where large amounts of money could be made; **space exploration**, where the

data are irreplaceable; and *medical imaging*, where lives could be saved. The personal computer revolution of the 1980s and 1990s caused DSP to explode with new applications. Rather than being motivated by military and government needs, DSP was suddenly *driven by the commercial marketplace*. Anyone who thought they could make money in the rapidly expanding field was suddenly a DSP vendor. DSP reached the public in such products as: mobile telephones, compact disc players, and electronic voice mail. Figure 1-1 illustrates a few of these varied applications.

This technological revolution occurred from the top-down. In the early 1980s, DSP was taught as a *graduate* level course in electrical engineering. A decade later, DSP had become a standard part of the *undergraduate* curriculum. Today, DSP is a *basic skill* needed by scientists and engineers

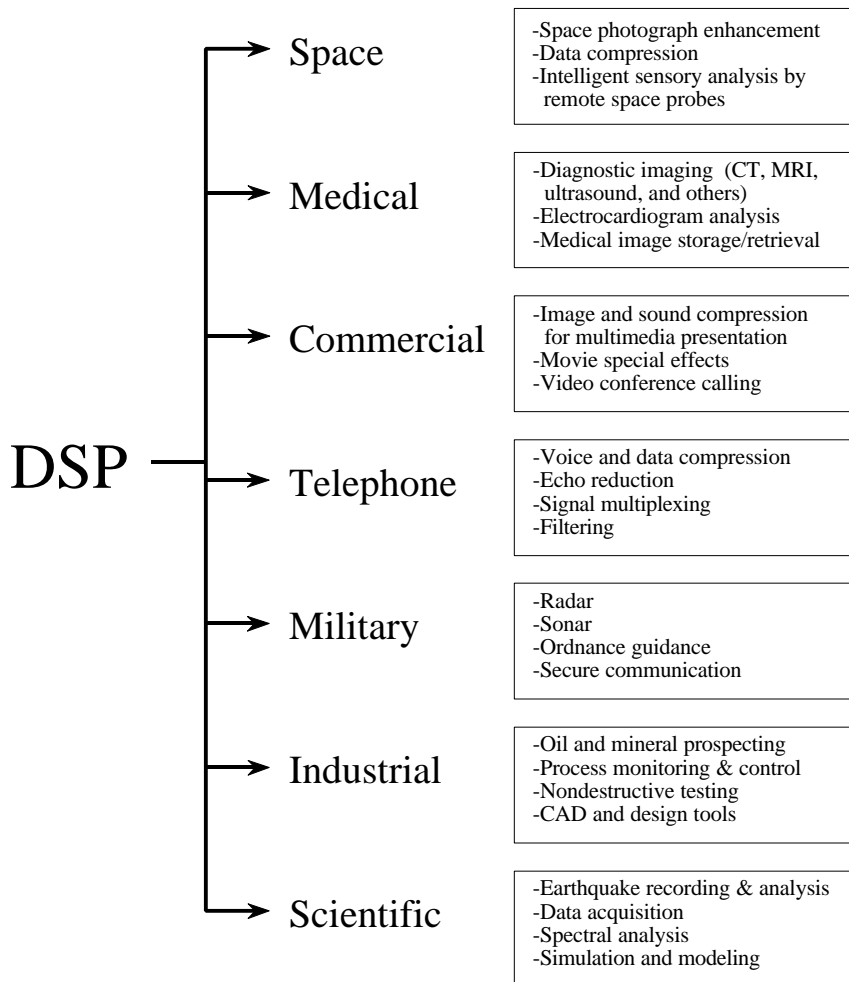


FIGURE 1-1

DSP has revolutionized many areas in science and engineering. A few of these diverse applications are shown here.

in many fields. As an analogy, DSP can be compared to a previous technological revolution: *electronics*. While still the realm of electrical engineering, nearly every scientist and engineer has some background in basic circuit design. Without it, they would be lost in the technological world. DSP has the same future.

This recent history is more than a curiosity; it has a tremendous impact on your ability to learn and use DSP. Suppose you encounter a DSP problem, and turn to textbooks or other publications to find a solution. What you will typically find is page after page of equations, obscure mathematical symbols, and unfamiliar terminology. It's a nightmare! Much of the DSP literature is baffling even to those experienced in the field. It's not that there is anything wrong with this material, it is just intended for a very specialized audience. State-of-the-art researchers need this kind of detailed mathematics to understand the theoretical implications of the work.

A basic premise of this book is that most practical DSP techniques can be learned and used without the traditional barriers of detailed mathematics and theory. *The Scientist and Engineer's Guide to Digital Signal Processing* is written for those who want to use DSP as a tool, not a new career.

The remainder of this chapter illustrates areas where DSP has produced revolutionary changes. As you go through each application, notice that DSP is very interdisciplinary, relying on the technical work in many adjacent fields. As Fig. 1-2 suggests, the borders between DSP and other technical disciplines are not sharp and well defined, but rather fuzzy and overlapping. If you want to specialize in DSP, these are the allied areas you will also need to study.

跨学科

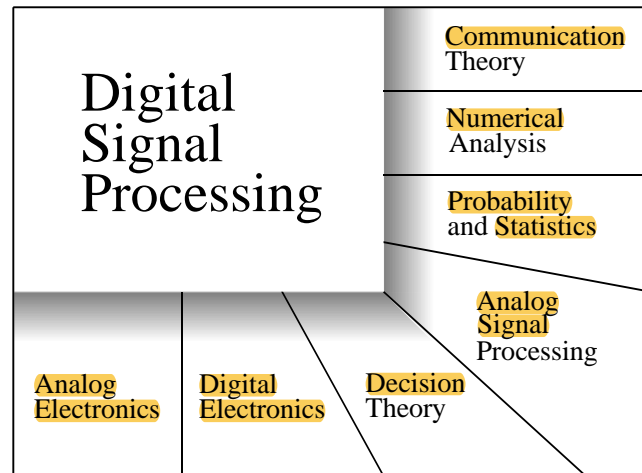


FIGURE 1-2

Digital Signal Processing has fuzzy and overlapping borders with many other areas of science, engineering and mathematics.

Telecommunications

Telecommunications is about transferring information from one location to another. This includes many forms of information: telephone conversations, television signals, computer files, and other types of data. To transfer the information, you need a channel between the two locations. This may be a wire pair, radio signal, optical fiber, etc. Telecommunications companies receive payment for transferring their customer's information, while they must pay to establish and maintain the channel. The financial bottom line is simple: the more information they can pass through a single channel, the more money they make. DSP has revolutionized the telecommunications industry in many areas: signaling tone generation and detection, frequency band shifting, filtering to remove power line hum, etc. Three specific examples from the telephone network will be discussed here: multiplexing, compression, and echo control.

Multiplexing

There are approximately one billion telephones in the world. At the press of a few buttons, switching networks allow any one of these to be connected to any other in only a few seconds. The immensity of this task is mind boggling! Until the 1960s, a connection between two telephones required passing the analog voice signals through mechanical switches and amplifiers. One connection required one pair of wires. In comparison, DSP converts audio signals into a stream of serial digital data. Since bits can be easily intertwined and later separated, many telephone conversations can be transmitted on a single channel. For example, a telephone standard known as the *T-carrier system* can simultaneously transmit 24 voice signals. Each voice signal is sampled 8000 times per second using an 8 bit companded (logarithmic compressed) analog-to-digital conversion. This results in each voice signal being represented as 64,000 bits/sec, and all 24 channels being contained in 1.544 megabits/sec. This signal can be transmitted about 6000 feet using ordinary telephone lines of 22 gauge copper wire, a typical interconnection distance. The financial advantage of digital transmission is enormous. Wire and analog switches are expensive; digital logic gates are cheap.

Compression

When a voice signal is digitized at 8000 samples/sec, most of the digital information is redundant. That is, the information carried by any one sample is largely duplicated by the neighboring samples. Dozens of DSP algorithms have been developed to convert digitized voice signals into data streams that require fewer bits/sec. These are called data compression algorithms. Matching uncompression algorithms are used to restore the signal to its original form. These algorithms vary in the amount of compression achieved and the resulting sound quality. In general, reducing the data rate from 64 kilobits/sec to 32 kilobits/sec results in no loss of sound quality. When compressed to a data rate of 8 kilobits/sec, the sound is noticeably affected, but still usable for long distance telephone networks. The highest achievable compression is about 2 kilobits/sec, resulting in

sound that is highly distorted, but usable for some applications such as military and undersea communications.

Echo control

Echoes are a serious problem in long distance telephone connections. When you speak into a telephone, a signal representing your voice travels to the connecting receiver, where a portion of it returns as an echo. If the connection is within a few hundred miles, the elapsed time for receiving the echo is only a few milliseconds. The human ear is accustomed to hearing echoes with these small time delays, and the connection sounds quite normal. As the distance becomes larger, the echo becomes increasingly noticeable and irritating. The delay can be several hundred milliseconds for intercontinental communications, and is particularly objectionable. Digital Signal Processing attacks this type of problem by measuring the returned signal and generating an appropriate *antisignal* to cancel the offending echo. This same technique allows speakerphone users to hear and speak at the same time without fighting audio feedback (squealing). It can also be used to reduce environmental noise by canceling it with digitally generated *antinoise*.

Audio Processing

The two principal human senses are **vision and hearing**. Correspondingly, much of DSP is related to image and audio processing. People listen to both *music* and *speech*. DSP has made revolutionary changes in both these areas.

Music

The path leading from the musician's microphone to the audiophile's speaker is remarkably long. Digital data representation is important to prevent the degradation commonly associated with analog storage and manipulation. This is very familiar to anyone who has compared the musical quality of cassette tapes with compact disks. In a typical scenario, a musical piece is recorded in a sound studio on multiple channels or tracks. In some cases, this even involves recording individual instruments and singers separately. This is done to give the sound engineer greater flexibility in creating the final product. The complex process of combining the individual tracks into a final product is called *mix down*. DSP can provide several important functions during mix down, including: filtering, signal addition and subtraction, signal editing, etc.

One of the most interesting DSP applications in music preparation is *artificial reverberation*. If the individual channels are simply added together, the resulting piece sounds frail and diluted, much as if the musicians were playing outdoors. This is because listeners are greatly influenced by the echo or reverberation content of the music, which is usually minimized in the sound studio. DSP allows artificial echoes and reverberation to be added during mix down to simulate various ideal listening environments. Echoes with delays of a few hundred milliseconds give the impression of cathedral like

locations. Adding echoes with delays of 10-20 milliseconds provide the perception of more modest size listening rooms.

Speech generation

Speech generation and recognition are used to communicate between humans and machines. Rather than using your hands and eyes, you use your mouth and ears. This is very convenient when your hands and eyes should be doing something else, such as: driving a car, performing surgery, or (unfortunately) firing your weapons at the enemy. Two approaches are used for computer generated speech: *digital recording* and *vocal tract simulation*. In digital recording, the voice of a human speaker is digitized and stored, usually in a compressed form. During playback, the stored data are uncompressed and converted back into an analog signal. An entire hour of recorded speech requires only about three megabytes of storage, well within the capabilities of even small computer systems. This is the most common method of digital speech generation used today.

Vocal tract simulators are more complicated, trying to mimic the physical mechanisms by which humans create speech. The human vocal tract is an acoustic cavity with resonant frequencies determined by the size and shape of the chambers. Sound originates in the vocal tract in one of two basic ways, called *voiced* and *fricative* sounds. With voiced sounds, vocal cord vibration produces near periodic pulses of air into the vocal cavities. In comparison, fricative sounds originate from the noisy air turbulence at narrow constrictions, such as the teeth and lips. Vocal tract simulators operate by generating digital signals that resemble these two types of excitation. The characteristics of the resonate chamber are simulated by passing the excitation signal through a digital filter with similar resonances. This approach was used in one of the very early DSP success stories, the *Speak & Spell*, a widely sold electronic learning aid for children.

Speech recognition

The automated recognition of human speech is immensely more difficult than speech generation. Speech recognition is a classic example of things that the human brain does well, but digital computers do poorly. Digital computers can store and recall vast amounts of data, perform mathematical calculations at blazing speeds, and do repetitive tasks without becoming bored or inefficient. Unfortunately, present day computers perform very poorly when faced with raw sensory data. Teaching a computer to send you a monthly electric bill is easy. Teaching the same computer to understand your voice is a major undertaking.

Digital Signal Processing generally approaches the problem of voice recognition in two steps: *feature extraction* followed by *feature matching*. Each word in the incoming audio signal is isolated and then analyzed to identify the type of excitation and resonate frequencies. These parameters are then compared with previous examples of spoken words to identify the closest match. Often, these systems are limited to only a few hundred words; can only accept speech with distinct pauses between words; and must be retrained for each individual speaker. While this is adequate for many commercial

applications, these limitations are humbling when compared to the abilities of human hearing. There is a great deal of work to be done in this area, with tremendous financial rewards for those that produce successful commercial products.

Echo Location

A common method of obtaining information about a remote object is to bounce a *wave* off of it. For example, radar operates by transmitting pulses of radio waves, and examining the received signal for echoes from aircraft. In sonar, sound waves are transmitted through the water to detect submarines and other submerged objects. Geophysicists have long probed the earth by setting off explosions and listening for the echoes from deeply buried layers of rock. While these applications have a common thread, each has its own specific problems and needs. Digital Signal Processing has produced revolutionary changes in all three areas.

Radar

Radar is an acronym for *RADio Detection And Ranging*. In the simplest radar system, a radio transmitter produces a pulse of radio frequency energy a few microseconds long. This pulse is fed into a highly directional antenna, where the resulting radio wave propagates away at the speed of light. Aircraft in the path of this wave will reflect a small portion of the energy back toward a receiving antenna, situated near the transmission site. The distance to the object is calculated from the elapsed time between the transmitted pulse and the received echo. The direction to the object is found more simply; you know *where* you pointed the directional antenna when the echo was received.

The operating range of a radar system is determined by two parameters: how much energy is in the initial pulse, and the noise level of the radio receiver. Unfortunately, increasing the energy in the pulse usually requires making the pulse *longer*. In turn, the longer pulse reduces the accuracy and precision of the elapsed time measurement. This results in a conflict between two important parameters: the ability to detect objects at long range, and the ability to accurately determine an object's distance.

DSP has revolutionized radar in three areas, all of which relate to this basic problem. First, DSP can *compress* the pulse after it is received, providing better distance determination without reducing the operating range. Second, DSP can filter the received signal to decrease the noise. This increases the range, without degrading the distance determination. Third, DSP enables the rapid selection and generation of different pulse shapes and lengths. Among other things, this allows the pulse to be optimized for a particular detection problem. Now the impressive part: much of this is done at a sampling rate comparable to the radio frequency used, as high as several hundred megahertz! When it comes to radar, DSP is as much about high-speed hardware design as it is about algorithms.

Sonar

Sonar is an acronym for *SOund NAVigation and Ranging*. It is divided into two categories, *active* and *passive*. In active sonar, sound pulses between 2 kHz and 40 kHz are transmitted into the water, and the resulting echoes detected and analyzed. Uses of active sonar include: detection & localization of undersea bodies, navigation, communication, and mapping the sea floor. A maximum operating range of 10 to 100 kilometers is typical. In comparison, passive sonar simply *listens* to underwater sounds, which includes: natural turbulence, marine life, and mechanical sounds from submarines and surface vessels. Since passive sonar emits no energy, it is ideal for covert operations. You want to detect *the other guy*, without him detecting *you*. The most important application of passive sonar is in military surveillance systems that detect and track submarines. Passive sonar typically uses lower frequencies than active sonar because they propagate through the water with less absorption. Detection ranges can be thousands of kilometers.

DSP has revolutionized sonar in many of the same areas as radar: pulse generation, pulse compression, and filtering of detected signals. In one view, sonar is *simpler* than radar because of the lower frequencies involved. In another view, sonar is more *difficult* than radar because the environment is much less uniform and stable. Sonar systems usually employ extensive arrays of transmitting and receiving elements, rather than just a single channel. By properly controlling and mixing the signals in these many elements, the sonar system can steer the emitted pulse to the desired location and determine the direction that echoes are received from. To handle these multiple channels, sonar systems require the same massive DSP computing power as radar.

Reflection seismology

As early as the 1920s, geophysicists discovered that the structure of the earth's crust could be probed with sound. Prospectors could set off an explosion and record the echoes from boundary layers more than ten kilometers below the surface. These echo seismograms were interpreted by the raw eye to map the subsurface structure. The reflection seismic method rapidly became the primary method for locating petroleum and mineral deposits, and remains so today.

In the ideal case, a sound pulse sent into the ground produces a single echo for each boundary layer the pulse passes through. Unfortunately, the situation is not usually this simple. Each echo returning to the surface must pass through all the other boundary layers above where it originated. This can result in the echo bouncing between layers, giving rise to *echoes of echoes* being detected at the surface. These secondary echoes can make the detected signal very complicated and difficult to interpret. Digital Signal Processing has been widely used since the 1960s to isolate the primary from the secondary echoes in reflection seismograms. How did the early geophysicists manage without DSP? The answer is simple: they looked in *easy* places, where multiple reflections were minimized. DSP allows oil to be found in *difficult* locations, such as under the ocean.

Image Processing

Images are signals with special characteristics. First, they are a measure of a parameter over *space* (distance), while most signals are a measure of a parameter over *time*. Second, they contain a great deal of information. For example, more than 10 megabytes can be required to store one second of television video. This is more than a thousand times greater than for a similar length voice signal. Third, the final judge of quality is often a subjective human evaluation, rather than an objective criterion. These special characteristics have made image processing a distinct subgroup within DSP.

Medical

In 1895, Wilhelm Conrad Röntgen discovered that x-rays could pass through substantial amounts of matter. Medicine was revolutionized by the ability to look inside the living human body. Medical x-ray systems spread throughout the world in only a few years. In spite of its obvious success, medical x-ray imaging was limited by four problems until DSP and related techniques came along in the 1970s. First, overlapping structures in the body can hide behind each other. For example, portions of the heart might not be visible behind the ribs. Second, it is not always possible to distinguish between similar tissues. For example, it may be able to separate bone from soft tissue, but not distinguish a tumor from the liver. Third, x-ray images show *anatomy*, the body's structure, and not *physiology*, the body's operation. The x-ray image of a living person looks exactly like the x-ray image of a dead one! Fourth, x-ray exposure can cause cancer, requiring it to be used sparingly and only with proper justification.

The problem of overlapping structures was solved in 1971 with the introduction of the first **computed tomography** scanner (formerly called computed axial tomography, or *CAT* scanner). Computed tomography (CT) is a classic example of Digital Signal Processing. X-rays from many directions are passed through the section of the patient's body being examined. Instead of simply forming images with the detected x-rays, the signals are converted into digital data and stored in a computer. The information is then used to *calculate* images that appear to be *slices through the body*. These images show much greater detail than conventional techniques, allowing significantly better diagnosis and treatment. The impact of CT was nearly as large as the original introduction of x-ray imaging itself. Within only a few years, every major hospital in the world had access to a CT scanner. In 1979, two of CT's principle contributors, Godfrey N. Hounsfield and Allan M. Cormack, shared the Nobel Prize in Medicine. *That's good DSP!*

The last three x-ray problems have been solved by using penetrating energy other than x-rays, such as radio and sound waves. DSP plays a key role in all these techniques. For example, Magnetic Resonance Imaging (MRI) uses magnetic fields in conjunction with radio waves to probe the interior of the human body. Properly adjusting the strength and frequency of the fields cause the atomic nuclei in a localized region of the body to resonate between quantum energy states. This resonance results in the emission of a secondary radio

wave, detected with an antenna placed near the body. The strength and other characteristics of this detected signal provide information about the localized region in resonance. Adjustment of the magnetic field allows the resonance region to be scanned throughout the body, mapping the internal structure. This information is usually presented as images, just as in computed tomography. Besides providing excellent discrimination between different types of soft tissue, MRI can provide information about physiology, such as blood flow through arteries. MRI relies totally on Digital Signal Processing techniques, and could not be implemented without them.

Space

Sometimes, you just have to make the most out of a bad picture. This is frequently the case with images taken from unmanned satellites and space exploration vehicles. No one is going to send a repairman to Mars just to tweak the knobs on a camera! DSP can improve the quality of images taken under extremely unfavorable conditions in several ways: brightness and contrast adjustment, edge detection, noise reduction, focus adjustment, motion blur reduction, etc. Images that have spatial distortion, such as encountered when a flat image is taken of a spherical planet, can also be *warped* into a correct representation. Many individual images can also be combined into a single database, allowing the information to be displayed in unique ways. For example, a video sequence simulating an aerial flight over the surface of a distant planet.

Commercial Imaging Products

The large information content in images is a problem for systems sold in mass quantity to the general public. Commercial systems must be *cheap*, and this doesn't mesh well with large memories and high data transfer rates. One answer to this dilemma is *image compression*. Just as with voice signals, images contain a tremendous amount of redundant information, and can be run through algorithms that reduce the number of bits needed to represent them. Television and other moving pictures are especially suitable for compression, since most of the image remain the same from frame-to-frame. Commercial imaging products that take advantage of this technology include: video telephones, computer programs that display moving pictures, and digital television.