WAVELETS

SEEING THE FOREST – AND THE TREES

*O*n November 25, 1998, Walt Disney Pictures and Pixar Animation Studios released a full-length computer-animated feature film called A Bug's Life. *It was the second such collaboration for Disney and Pixar and, like its groundbreaking predecessor* Toy Story *three years earlier, it opened to rave reviews. A Bug's Life, said one reviewer, "teems with beautiful visual inventions…; with intricate details that will keep adults as well as kids bug-eyed from start to finish…; and with colors teased from some new, hitherto-secret pastel spectrum…"*

Only the most computer graphics-savvy moviegoers would have given any thought to the mathematical modeling techniques that made it possible to develop all the characters in the animated ants' story—not to mention their many textures, their myriad expressions, and the way they jumped, flitted, and buzzed around. As it happened, though, a particular type of modeling technique made its debut in the movie, a method of computer animation that makes use of a collection of mathematical procedures called "wavelets."

One way of thinking about wavelets is to consider how our eyes look at the world. In the real world, you can observe a forest like the one shown in the photograph on the next page from many vantage points—in effect, at different scales of resolution. From the window of a cross-country jet, for example, the forest appears to be a solid canopy of green. From the window of an automobile on the ground, the canopy resolves into individual trees, and if

you get out of the car and move closer, you begin to see branches and leaves. If you then pull out a magnifying glass, you might find a drop of dew at the end of a leaf. As you zoom in at smaller and smaller scales, you can find details that you didn't see before. Try to do that with a photograph, however, and you will be disappointed. Enlarge the photograph to get "closer" to a tree and all you'll have is a fuzzier tree; the branch, the leaf, the drop of dew are not to be found. Although our eyes can see the forest at many scales of resolution, the camera can show only one at a time.

Computers do no better than cameras; in fact, their level of resolution is inferior. On a computer screen, the photograph becomes a collection of pixels that are much less sharp than the original.

Soon, however, computers everywhere will be able to do something that photographers have only been able to dream of. They will be able to display an interactive image of a forest in which the viewer can zoom in to get greater detail of the trees, branches, and perhaps even the leaves. They will be able to do this because wavelets make it possible to compress the amount of data used to store an image, allowing a more detailed image to be stored in less space.

Even though as an organized research topic wavelets is less than two decades old, it arises from a constellation of related concepts developed over a period of nearly two

Flik and the other characters in the movie, A Bug's Life, *were generated by a method of computer animation that uses wavelets. (© Disney Enterprises, Inc./ Pixar Animation Studios)*
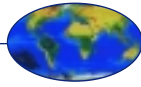
This photo captures the forest at one scale of resolution. Soon computers everywhere will be able to display interactive images in which viewers can zoom in to see greater detail of the trees, branches, and leaves. (Gerry Ellis/Minden Pictures)

centuries, repeatedly rediscovered by scientists who wanted to solve technical problems in their various disciplines. Signal processors were seeking a way to transmit clear messages over telephone wires. Oil prospectors wanted a better way to interpret seismic traces. Yet "wavelets" did not become a household word among scientists until the theory was liberated from the diverse applications in which it arose and was synthesized into a purely mathematical theory. This synthesis, in turn, opened scientists' eyes to new applications. Today, for example, wavelets are not only the workhorse in computer imaging and animation; they also are used by the FBI to encode its data base of 30 million fingerprints. In the future, scientists may put wavelet analysis to work diagnosing breast cancer, looking for heart abnormalities, or predicting the weather.

Transforming Reality

Wavelet analysis allows researchers to isolate and manipulate specific types of patterns hidden in masses of data, in much the same way our eyes can pick out the trees in a forest, or our ears can pick out the flute in a symphony. One approach to understanding how wavelets do this is to start with the difference between two kinds of sounds—a tuning fork and the human voice. Strike a tuning fork and you get a pure tone that lasts for a very long time. In mathematical theory, such a tone is said to be "localized" in *frequency;* that is, it consists of a single note with no higher-frequency overtones. A spoken word, by contrast, lasts for only a second, and thus is "localized" in *time.* It is not localized in frequency because the word is not a single tone but a combination of many different frequencies.

Graphs of the sound waves produced by the tuning fork and human voice highlight the difference, as illustrated on page 3. The vibrations of the tuning fork trace out what mathematicians call a *sine wave,* a smoothly undulating curve that, in theory, could repeat forever. In contrast, the graph of the word "greasy" contains a series of sharp spikes; there are no oscillations.

In the nineteenth century, mathematicians perfected what might be called the "tuning fork" version of reality, a theory known as Fourier analysis. Jean Baptiste Joseph Fourier, a French mathematician, claimed in 1807 that any repeating waveform (or periodic function), like the tuning fork sound wave, can be expressed as an infinite sum of sine waves and cosine waves of various frequencies. (A cosine wave is a sine wave shifted forward a quarter cycle.)

A familiar demonstration of Fourier's theory occurs in music. When a musician plays a note, he or she creates an irregularly shaped sound wave. The same shape repeats itself for as long as the musician holds the note. Therefore, according to Fourier, the note can be separated into a sum of sine and cosine waves. The lowest-frequency wave is called the fundamental frequency of the note, and the higher-frequency ones are called overtones. For example, the note A, played on a violin or a flute, has a fundamental frequency of 440 cycles per second and overtones with frequencies of 880, 1320, and so on. Even if a violin and a flute are playing the same note, they will sound different because their overtones have different strengths or "amplitudes." As music synthesizers demonstrated in the 1960s, a very convincing imitation of a violin or a flute can be obtained by re-combining pure sine waves with the appropriate amplitudes. That, of course, is exactly what Fourier predicted back in 1807.

Mathematicians later extended Fourier's idea to non-periodic functions (or waves) that change over time, rather than repeating in the same shape forever. Most real-world waves are of this type: say, the sound of a motor that speeds up, slows down, and hiccups now and then. In images, too, the distinction between repeating and non-repeating patterns is important. A repeating pattern may be seen as a texture or background, while a non-repeating one is picked out by the eye as an object. Periodic or repeating waves composed of a discrete series of overtones can be used to represent repeating (background) patterns in an image. Non-periodic features can be resolved into a much more complex spectrum of frequencies, called the "Fourier transform," just as sunlight can be separated into a spectrum of colors.
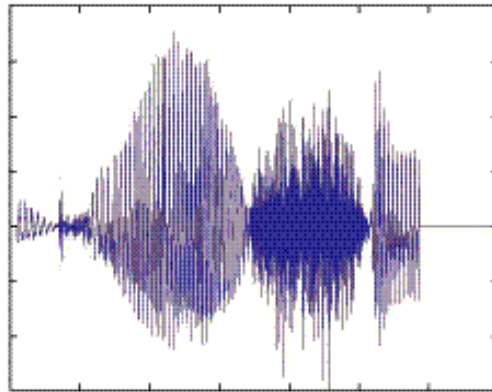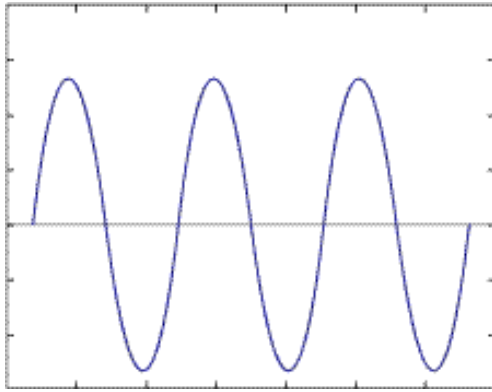
Graphs of the sound-waves produced by a tuning fork (left) and the spoken word "greasy" (right) illustrate the difference between a tone localized in frequency and one localized in time. The tuning fork produces a simple "sine wave." (Courtesy of Ofer Levi, Stanford University)

The Fourier transform portrays the structure of a periodic wave in a much more revealing and concentrated form than a traditional graph of a wave would. For example, a rattle in a motor will show up as a peak at an unusual frequency in the Fourier transform.

Fourier transforms have been a hit. During the nineteenth century they solved many problems in physics and engineering. This prominence led scientists and engineers to think of them as the preferred way to analyze phenomena of all kinds. This ubiquity forced a close examination of the method. As a result, throughout the twentieth century, mathematicians, physicists, and engineers came to realize a drawback of the Fourier transform: they have trouble reproducing transient signals or signals with abrupt changes, such as the spoken word or the rap of a snare drum. Music synthesizers, as good as they are, still do not match the sound of concert violinists, because the playing of a violinist contains transient features—such as the contact of the bow on the string—that are poorly imitated by representations based on sine waves.

The principle underlying this problem can be illustrated by what is known as the Heisenberg Indeterminacy Principle. In 1927, the physicist Werner Heisenberg stated that the position and the velocity of an object cannot both be measured exactly at the same time even in theory. In signal processing terms, this means it is impossible to know simultaneously the exact frequency and the exact time of occurrence of this frequency in a signal. In order to know its frequency, the signal must be spread in time, or vice versa. In musical terms, the trade-off means that any signal with a short duration must have a complicated frequency spectrum made of a rich variety of sine waves, whereas any signal made from a simple combination of a few sine waves must have a complicated appearance in the time domain. Thus, we can't expect to reproduce the sound of a drum with an orchestra of tuning forks.

An Idea with No Name

Over the course of the twentieth century, scientists in different fields struggled to get around these limitations, in order to allow representations of the data to adapt to the nature of the information. In essence, they wanted to capture both the low-resolution forest—the repeating background signal—and the high-resolution trees—the individual, localized variations in the background. Although the scientists were each trying to solve the problems particular to their respective fields, they began to arrive at the same conclusion—namely, that Fourier transforms themselves were to blame. They also arrived at essentially the same solution: Perhaps by splitting a signal up into components that were not pure sine waves, it would be possible to condense the information in both the time and frequency domains. This is the idea that would ultimately be known as wavelets.

The first entrant in the wavelet derby was a Hungarian mathematician named Alfred Haar, who introduced in 1909 the functions that are now called "Haar wavelets." These functions consist simply of a short positive pulse followed by a short negative pulse. An example is shown on page 5. Although the short pulses of Haar wavelets are excellent for teaching wavelet theory, they are less useful for most applications because they yield jagged lines instead of smooth curves. For example, an image reconstructed with Haar wavelets looks like a cheap calculator display, and a Haar wavelet reconstruction of the sound of a flute is too harsh.

From time to time over the next several decades, other precursors of wavelet theory arose. In the 1930s, the English mathematicians John Littlewood and R.E.A.C. Paley developed a method of grouping frequencies by octaves, thereby creating a signal that is well localized in frequency (its spectrum lies within

one octave) and also relatively well localized in time. In 1946, Dennis Gabor, a British-Hungarian physicist, introduced the Gabor transform, analogous to the Fourier transform, which separates a wave into "time-frequency packets" or "coherent states" that have the greatest possible simultaneous localization in both time and frequency. And in the 1970s and 1980s, the signal processing and image processing communities introduced their own versions of wavelet analysis, going by such names as "subband coding," "quadrature mirror filters," and the "pyramidal algorithm."

While not precisely identical, all of these techniques had similar features. They decomposed or transformed signals into pieces that could be localized to any time interval and could also be dilated or contracted to analyze the signal at different scales of resolution. These precursors of wavelets had one other thing in common: No one knew about them beyond individual specialized communities. But in 1984, wavelet theory finally came into its own.

The Great Synthesis

Jean Morlet didn't plan to start a scientific revolution. He was merely trying to give geologists a better way to search for oil.

Petroleum geologists usually locate underground oil deposits by making loud noises. Because sound waves travel through different materials at different speeds, geologists can infer what kind of material lies under the surface by sending seismic waves into the ground and measuring how quickly they rebound. If the waves propagate especially quickly through one layer, it may be a salt dome, which can trap a layer of oil underneath.

Figuring out just how the geology translates into a sound wave (or vice versa) is a tricky mathematical problem, and one that engineers traditionally solve with Fourier analysis. Unfortunately, seismic signals contain lots of transients—abrupt changes in the wave as it passes from one rock layer to another. These transients contain exactly the information the geologist is looking for, namely the location of the rock layers, but Fourier analysis spreads that spatial information out all over the place.

Morlet, an engineer for Elf-Aquitaine, developed his own way of analyzing the seismic signals to create components that were localized in space, which he called "wavelets of constant shape." Later, they would be known as "Morlet wavelets." Whether the components are dilated, compressed, or shifted in time, they maintain the same shape. Other families of wavelets can be built by taking a different shape, called a mother wavelet, and dilating, compressing, or shifting it in time. Researchers would find that the exact shape of the mother wavelet strongly affects the accuracy and compression properties of the approximation. Many of the

Timeline

Wavelets have had an unusual scientific history, marked by many independent discoveries and rediscoveries. The most rapid progress has come since the early 1980s, when a coherent mathematical theory of wavelets finally emerged.

1807
Jean Baptiste Joseph Fourier claims that any periodic function, or wave, can be expressed as an infinite sum of sine and cosine waves of various frequencies. Because of serious doubts over the correctness of his arguments, his paper is not published until 15 years later.

1930
John Littlewood and R.A.E.C. Paley, of Cambridge University, show that local information about a wave, such as the timing of a pulse of energy, can be retrieved by grouping the terms of its Fourier series into "octaves."

1976
IBM physicists Claude Galand and Daniel Esteban discover subband coding, a way of encoding digital transmissions for the telephone.

1984
Joint paper by Morlet and Grossmann brings the word "wavelet" into the mathematical lexicon for the first time.

1909
Alfred Haar, a Hungarian mathematician, discovers a "basis" of functions that are now recognized as the first wavelets. They consist of a short positive pulse followed by a short negative pulse.

1946
Dennis (Denes) Gabor, a British-Hungarian physicist who invented holography, decomposes signals into "time-frequency packets" or "Gabor chirps."

1981
Petroleum engineer Jean Morlet of Elf-Aquitaine finds a way to decompose seismic signals into what he calls "wavelets of constant shape." He turns to quantum physicist Alex Grossmann for help in proving that the method works.
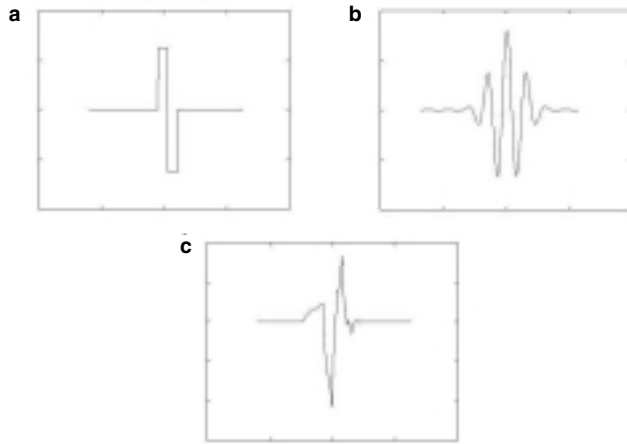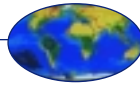
Graphs of several different types of wavelets. (a) Haar Wavelet, (b) Morlet Wavelet (c) Daubechies Wavelet. (Courtesy of Ofer Levi, Stanford University)

differences between earlier versions of wavelets simply amounted to different choices for the mother wavelet.

Morlet's method wasn't in the books, but it seemed to work. On his personal computer, he could separate a wave into its wavelet components, and then reassemble them into the original wave. But he wasn't satisfied with this empirical proof, and began asking other scientists if the method was mathematically sound.

Morlet found the answer he wanted from Alex Grossmann, a physicist at the Centre de Physique Théorique in Marseilles. Grossmann worked with

Morlet for a year to confirm that waves could be reconstructed from their wavelet decompositions. In fact, wavelet transforms turned out to work better than Fourier transforms, because they are much less sensitive to small errors in the computation. An error or an unwise truncation of the Fourier coefficients can turn a smooth signal into a jumpy one or vice versa; wavelets avoid such disastrous consequences.

Morlet and Grossmann's paper, the first to use the word "wavelet," was published in 1984. Yves Meyer, currently at the École Normale Supérieure de Cachan, widely acknowledged as one of the founders of wavelet theory, heard about their work in the fall of the same year. He was the first to realize the connection between Morlet's wavelets and earlier mathematical wavelets, such as those in the work of Littlewood and Paley. (Indeed, Meyer has counted 16 separate rediscoveries of the wavelet concept before Morlet and Grossmann's paper.)

Meyer went on to discover a new kind of wavelet, with a mathematical property called orthogonality that made the wavelet transform as easy to work with and manipulate as a Fourier transform. ("Orthogonality" means that the information captured by one wavelet is completely independent of the information captured by another.) Perhaps most importantly, he became the nexus of the emerging wavelet community.

In 1986, Stéphane Mallat, a former student of Meyer's who was working on a doctorate in computer

1985
Yves Meyer of the University of Paris discovers the first smooth orthogonal wavelets.

1986
Stéphane Mallat, then at the University of Pennsylvania, shows that the Haar basis, the Littlewood-Paley octaves, the Gabor chirps, and the subband filters of Galand and Esteban are all related to wavelet-based algorithms.

1987
Ingrid Daubechies constructs the first smooth orthogonal wavelets with compact support. Her wavelets turn the theory into a practical tool that can be easily programmed and used by any scientist with a minimum of mathematical training.

1990
David Donoho and Iain Johnstone, at Stanford University, use wavelets to "denoise" images, making them even sharper than the originals.

1992
The FBI chooses a wavelet method developed by Tom Hopper of the FBI's Criminal Justice Information Services Division and Jonathan Bradley and Chris Brislawn from Los Alamos National Laboratory, to compress its enormous database of fingerprints.

1995
Pixar Studios releases the movie *Toy Story*, the first fully computer-animated cartoon. In the sequel, *Toy Story 2,* some shapes are rendered by subdivision surfaces, a technique mathematically related to wavelets.

1999
The International Standards Organization approves a new standard for digital picture compression, called JPEG-2000. The new standard uses wavelets to compress image files by 1:200 ratios with no visible loss in image quality. Web browsers are expected to support the new standard by 2001.

vision, linked the theory of wavelets to the existing literature on subband coding and quadrature mirror filters, which are the image processing community's versions of wavelets. The idea of multiresolution analysis—that is, looking at signals at different scales of resolution—was already familiar to experts in image processing. Mallat, collaborating with Meyer, showed that wavelets are implicit in the process of multiresolution analysis.

Thanks to Mallat's work, wavelets became much easier. One could now do a wavelet analysis without knowing the formula for a mother wavelet. The process was reduced to simple operations of averaging groups of pixels together and taking their differences, over and over. The language of wavelets also became more comfortable to electrical engineers, who embraced familiar terms such as "filters," "high frequencies," and "low frequencies."

The final great salvo in the wavelet revolution was fired in 1987, when Ingrid Daubechies, while visiting the Courant Institute at New York University and later during her appointment at AT&T Bell Laboratories, discovered a whole new class of wavelets, which were not only orthogonal (like Meyer's) but which could be implemented using simple digital filtering ideas, in fact, using short digital filters. The new wavelets were almost as simple to program and use as Haar wavelets, but they were smooth without the jumps of Haar wavelets. Signal processors now had a dream tool: a way to break up digital data into contributions of various scales. Combining Daubechies and Mallat's ideas, there was a simple, orthogonal transform that could be

rapidly computed on modern digital computers.

The Daubechies wavelets have surprising features—such as intimate connections with the theory of fractals. If their graph is viewed under magnification, characteristic jagged wiggles can be seen, no matter how strong the magnification. This exquisite complexity of detail means there is no simple formula for these wavelets. They are ungainly and asymmetric; nineteenth-century mathematicians would have recoiled from them in horror. But like the Model-T Ford, they are beautiful because they work. The Daubechies wavelets turn the theory into a practical tool that can be easily programmed and used by any scientist with a minimum of mathematical training.

How Do Wavelets Work?

So far, the "killer app" for wavelets has been digital image compression. They are central to the new JPEG-2000 digital image standard and the WSQ (wavelet scalar quantization) method that the FBI uses to compress its fingerprint database. In this context, wavelets can be thought of as the building blocks of images. An image of a forest can be made from the broadest wavelets: a big swath of green for the forest, a splash of blue for the sky. More detailed, sharper wavelets can help distinguish one tree from another. Branches and needles can be added to the image with even finer wavelets. Like an individual brush stroke in a painting, each wavelet is not itself an image, but many wavelets together can recreate anything. Unlike a brush stroke in a painting, a wavelet can be made arbitrarily small: A wavelet has no physical size limitations because it is simply a series of 0s and 1s stored in a computer's memory.

Contrary to popular belief, wavelets themselves do not compress an image: Their job is to *make compression possible*. To understand why, suppose that an image is encoded as a series of spatially arranged numbers, such as 1, 3, 7, 9, 8, 8, 6, 2. If each number represents the darkness of a pixel, with 0 being white and 15 being black, then this string represents some kind of gray object (the 7s, 8s, and 9s) against a light background (the 1s, 2s, and 3s).

The simplest kind of multiresolution analysis filters the image by averaging each pair of adjacent pixels. In the above example, this results in the string 2, 8, 8, 4: a lower-resolution image that still shows a grayish object against a light background. If we wanted to reconstruct a degraded version of the original image from this, we could do so by repeating each number
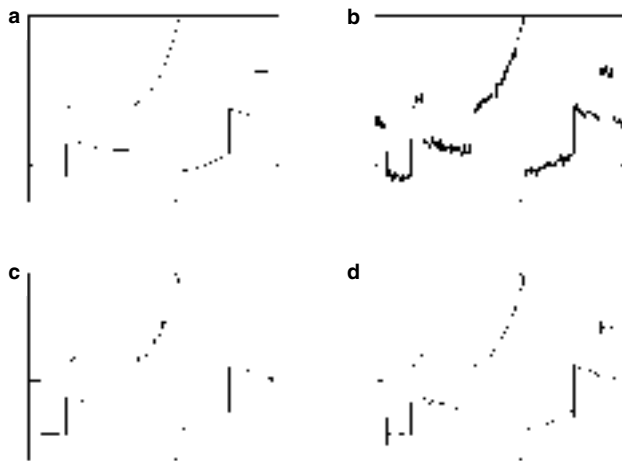
Wavelets can be used to filter out a signal from noise. The top images show the original signal exhibiting jumps and smooth pieces (a) and a noisy version of the signal, which one would like to "denoise" (b). On the bottom, the result of denoising using Haar wavelets yields a jagged line instead of a smooth curve (c), in contrast, using Daubechies wavelets results in smoother curve (d). (Courtesy of Ofer Levi, Stanford University)
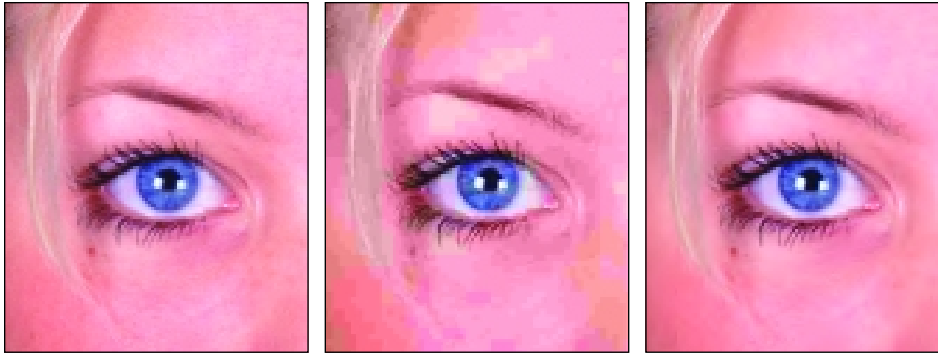
Wavelets make it possible to compress images with little image quality degradation. From left to right are the original image, the image compressed at a 220:1 ratio using standard JPEG technology and the image compressed to the same ratio using JPEG-2000, a method that uses wavelets. (Image courtesy of ImageState; graphic manipulated by Aware, Inc.)

in the string: 2, 2, 8, 8, 8, 8, 4, 4.

Suppose, however, that we wanted to get back the original image perfectly. To do this, we would have to save some additional information in the first step, namely a set of numbers that can be added to or subtracted from the low-resolution signal to obtain the high-resolution signal. In the example, those numbers are -1, -1, 0, and 2. (For example: Adding -1 to the first pixel of the degraded image, 2, gives 1, the first pixel of the original image; subtracting -1 from it gives 3, the second pixel of the original image.)

Thus the first level of the multiresolution analysis splits the original signal up into a low-resolution part (2, 8, 8, 4) and a high-frequency or "detail" part (-1, -1, 0, 2). The high-frequency details are also called the Haar wavelet coefficients. In fact, this whole procedure is the multiresolution version of the wavelet transform Haar discovered in 1909.

It might not seem that the first step of the wavelet transform has gained anything. There were eight numbers in the original signal, and there are still eight numbers in the transform. But in a typical digital image, most pixels will be very much like their neighbors: Sky pixels will occur next to sky pixels, forest pixels next to forest pixels. This means that the averages of nearby pixels will be almost the same as the original pixels, and so most of the detail coefficients will either be zero or very close to zero. If we simply round those coefficients off to zero, then the only information we need to keep is the low-resolution image plus a smattering of detail coefficients that did not get rounded off to zero. Thus, the amount of data required to store the image has been compressed by a factor of almost 2. The process of rounding high-precision numbers into lower precision numbers with fewer digits is called quantization (the "Q" in "WSQ"). An example is the process of rounding a number to two significant figures.

The process of transforming and quantizing can be repeated as many times as desired, each time decreasing the bits of information by a factor of almost 2 and slightly degrading the quality of the image. Depending on the needs of the user, the process can be stopped before the lower resolution starts to become apparent, or it can be continued to obtain a very low-resolution "thumbnail" image with layers of increasingly accurate details. With the JPEG-2000 standard, one can achieve compression ratios of 200:1 without a perceptible difference in the quality of the image. Such wavelet decompositions are obtained by averaging more than two nearby pixels at a time. The simplest Daubechies wavelet transform, for instance, combines groups of four pixels, and smoother ones combine six, eight, or more.

One fascinating property of wavelets is that they automatically pick out the same features our eyes do. The wavelet coefficients that are still left after quantization correspond to pixels that are very different from their neighbors—at the edge of the objects in an image. Thus, wavelets recreate an image mostly by drawing edges—which is exactly what humans do when they sketch a picture. Indeed, some researchers have suggested that the analogy between wavelet transforms and human vision is no accident, and that our neurons filter visual signals in a similar way to wavelets.

Wavelets in the Future

With the foundations of wavelet theory securely in place, the field has grown rapidly over the last decade. A distribution list on wavelets that began with 40 names in 1990 is now an online newsletter with more than 17,000 subscribers. Moreover, it has continued to evolve through a healthy mix of theory and practice. Engineers are always trying new applications, and for mathematicians, there are still important theoretical questions to be answered.

Although wavelets are best known for image com-

pression, many researchers are interested in using wavelets for pattern recognition. In weather forecasting, for example, they might slim down the data-bloated computer models that are now in use. Traditionally, such models sample the barometric pressure (for instance) at an enormous number of grid points and use this information to predict how the data will evolve. However, this approach uses up a lot of computer memory. A model of the atmosphere that uses a 1000-by-1000-by-1000 grid requires a billion data points—and it's still a fairly crude model.

However, most of the data in the grid are redundant. The barometric pressure in your town is probably about the same as the barometric pressure a mile down the road. If the weather models used wavelets, they could view the data the same way weather forecasters do, concentrating on the places where abrupt changes occur—warm fronts, cold fronts and the like. Other problems in fluid dynamics have been tackled the same way. At Los Alamos National Laboratory, for example, wavelets are used to study the shock waves produced by a bomb explosion.

As demonstrated by the recent spate of full-length computer-animated films, wavelets also have a promising future in the movies. Because the wavelet transform is a reversible process, it is just as easy to *synthesize* an image (build it up out of wavelets) as it is to *analyze* it (break it down into wavelet components). This idea is related to a new computer animation method called *subdivision surfaces,* basically a multiresolution analysis run in reverse. To draw a cartoon character, the animator only has to specify where a few key points go, creating a low-resolution version of the character. The computer can then do a reverse multiresolution analysis, making the character look like a real person and not a stick figure.

Subdivision surfaces debuted in the 1998 movie *A Bug's Life,* replacing a more clumsy method called NURBs (for non-uniform rational B splines) that had been used in the first *Toy Story* movie in 1995. Interestingly, NURBs and subdivision methods coexisted in 1999's *Toy Story 2,* where the characters that appeared in the first *Toy Story* remained NURBs, but where the new characters were based on the subdivision method. The next frontier for subdivision surfaces may be the video game industry, where they could eliminate the blocky look of today's graphics.

Meanwhile, on the theoretical side, mathematicians are still looking for better kinds of wavelets for two- and three-dimensional images. Although the standard wavelet methods are good at picking up edges, they do it one pixel at a time—an inefficient way of representing something that may be a very simple curve or line. David Donoho and Emmanuel Candès of Stanford University have proposed a new class of wavelets called "ridgelets," which are specifically designed to detect discontinuities along a line. Other researchers are studying "multiwavelets," which can be used to encode multiple signals traveling through the same line, such as color images in which three color values (red, green, and blue) have to be transmitted at once.

When asked to justify the value of mathematics, mathematicians often point out that ideas developed to solve a pure mathematical problem can lead to unexpected applications years later. But the story of wavelets paints a more complicated and somewhat more interesting picture. In this case, specific applied research led to a new theoretical synthesis, which in turn opened scientists' eyes to new applications. Perhaps the broader lesson of wavelets is that we should not view basic and applied sciences as separate endeavors: Good science requires us to see both the theoretical forest and the practical trees.