Image Processing and Measurement

John C. Russ
Department of Materials Science and Engineering, College of Engineering, North Carolina State University, Raleigh, North Carolina, U.S.A.

Abstract
Images provide important information in scientific, technical, and forensic situations, in addition to their role in everyday life. Extracting information from images acquired by digital cameras involves image processing to correct colors, reduce noise, and correct for nonuniform illumination or nonplanar views. Enhancement of image details is generally accomplished by reducing the contrast of other information in the image, so that, for example, lines and edges that make measurements of structure are more accessible. The processing steps use a variety of computer algorithms and may be performed on the pixel array, or in a different space, for example, by using a Fourier transform. Some applications, especially forensic ones, require simple comparisons, but for object identification, classification, or correlations, quantitative measurements of color or density, position, size, and shape are needed. Several possible measurement quantities are available for each category, particularly shape, for which a variety of dimensionless ratios, Fourier or wavelet coefficients, and invariant moments may be used. Interpretation of the measurements depends on the nature of the image and of the specimen or scene, for instance, whether it consists of discrete objects on a surface, a section through a complex structure, or a projection through a three-dimensional space.

INTRODUCTION
Humans depend to a high degree on images to gather information about their world, and to organize and understand that information. This dependence extends to scientific, technical, and forensic analysis as well, and to scales that include the microscopic and astronomical, aided by a broad variety of instruments designed to use infrared light, X-rays, radar, sound waves, and so on.

Human vision is not a quantitative tool, and is easily fooled by illusions and distracted by extraneous or random background features. Measurement requires direct comparison to appropriate standards (rulers, protractors, color scales, etc.). Consequently, the design of instruments and computer algorithms that collect, process, and analyze images is a key part of acquiring quantitative data for many scientific, technical, and forensic activities.

Image processing is done for two principal reasons: to improve visual appearance for a human observer, including printing and transmission, and to prepare images for measurement and analysis of the features and structures which they reveal. Image processing methods can be considered in two principal categories: the correction of defects or limitations in acquisition, and the enhancement of important details. Image processing may alter the values or locations of pixels (picture elements) to produce another image. Image analysis, on the other hand, extracts numerical measurement information from the picture.

It is important to understand that the scale of an image (μm, feet, miles, or light years) matters little, as does the type of signal used to form the image. Most processing and measurement tools are equally applicable to a broad variety of images, and may be used in a very wide range of applications.

Correction of Defects 1: Color Adjustment
Digital cameras, and earth-observing satellites, capture color images. Color correction should be the first operation performed if it is required. Compensation for variations in illumination can be made in several ways. The best results require capturing an image of known color standards under the same lighting, or having sufficient independent knowledge of the characteristics of the light source and the physics of the instrumentation.

With standards, a tristimulus matrix can be calculated that corrects for the overlap in the wavelength ranges of the filters used to form the red, green, and blue (RGB) signals that are typically stored. In some cases, a simpler and more approximate approach is used in which neutral gray objects are located and the RGB values adjusted to be equal. This constructs adjustment curves for each color channel which are then applied throughout the image.
Most cameras and computers store and display color images as RGB values for each pixel, but for most processing and measurement purposes other color spaces are more useful. \(L\), \(a\), and \(b\) color coordinates are often used as shown in Fig. 1. In the \(L\), \(a\), \(b\) space, which may be represented as a sphere with orthogonal axes, \(L\) is the luminance, or brightness, while the “\(a\)” and “\(b\)” axes are red–green and blue–yellow.

HSI space is more complicated, with \(H\) or hue represented as an angle on the color wheel from red to yellow, green, cyan, blue, magenta, and back to red, while \(S\) is saturation, or the amount of color (e.g., the difference between gray, pink, and red), and \(I\) (also called \(V\) for value or \(B\) for brightness) is the intensity. This space may be represented as a cylinder, cone, or bi-cone. In the bi-cone shown in the figure, saturation is reduced to zero at the ends. The color saturation at maximum intensity can be increased only by reducing some color contribution, and likewise at the dark end saturation can be increased only by increasing intensity. Conversion from one color space to another is performed in software as necessary.

**Correction of Defects 2: Noise Reduction**

Noise is generally any part of the image that does not represent the actual scene, but arises from other sources. These may include the statistics of charge production in the detector, thermal or electronic noise in the amplifier and digitization process, electrical interference in transmission, vibration of the camera or flickering of the light source, and so on. The two principal kinds of noise are random and periodic; they are treated in different ways, under the assumption that they can be distinguished from the important details. Random or speckle noise usually appears as fluctuations in the brightness of neighboring pixels and is treated in the spatial domain of the pixels, while periodic noise involves larger-scale variations and is best dealt with using the Fourier transform of the image.

Fig. 2 shows an image with significant random noise, visible as variations in the pixels in the uniform background above the cat’s head. It arises primarily from the amplification required, because the photo was taken in dim light. The most common, but generally poor, approach used for random noise reduction is a Gaussian blur, which replaces each pixel value with the weighted average of the pixels in a small neighborhood. This reduces the noise as shown, but also blurs detail and shifts edges. It is identical to a low-pass filter in Fourier space that keeps low frequencies and reduces the high frequencies (variations over a short distance) that constitute the pixel-to-pixel noise variations, but which are also needed to define edges, lines, and boundaries. Extensions of the Gaussian model may adjust the weights applied to neighboring pixels based on their difference in value or the direction of the local brightness gradient.

![Fig. 1 Color spaces: (A) cubic RGB; (B) spherical \(L\), \(a\), \(b\); (C) biconic HSI.](image1)

![Fig. 2 Random noise reduction: (A) original; (B) Gaussian smooth; (C) median filter; (D) nonlocal means filter.](image2)
Median filters replace each pixel with the mean value found by ranking the pixel values in the neighborhood according to brightness (all of the examples in the figure use a neighborhood with a radius of three pixels). The median filter is a nonlinear operation that has no equivalent in Fourier space. This filter, and variations that combine partial results from multiple neighborhoods, or use vectors for color images, are widely used and do a better job of preserving details such as lines and edges while reducing random noise. More computationally complex filters such as the nonlocal means filter\[1\] produce even better results. This works by replacing each pixel with a weighted average of all pixels in the image, based on the similarity of their neighborhoods.

Fig. 3 shows an example of periodic noise. In the Fourier transform, this appears as “spikes” at radii corresponding to the frequency (inverse of the spacing of the lines) and at angles that correspond to their orientation. Removal of the spikes and calculating the inverse Fourier transform restores the image with the noise removed but all other information, which is composed of different frequencies and orientations, intact.

**Correction of Defects 3: Nonuniform Illumination**

A key assumption behind most methods for selecting features for measurement is that an object should have the same color and brightness wherever it happens to lie in the field of view. In some controlled situations, such as microscopy and laboratory setups, uniform illumination can be achieved. In real world imagery, including crime scene photos and satellite imaging of a curved planet, it may be difficult or impossible to do so. There are several ways to adjust the resulting image to correct the nonuniformity.

Fig. 4 shows the preferred approach—recording an image of the background or substrate with the objects of interest removed. This background can then be subtracted or divided into the original to remove the variation. The choice of subtraction or division depends on how the camera recorded the brightness, as explained in the section Detail Enhancement 7: Image Combinations. When recording a background image is not practical, it may be possible to model the background by fitting a smooth function, typically a polynomial, to multiple points in the image that are known or assumed to be the same, or in some cases to calculate a background based on independent knowledge of the circumstances (such as the lighting of a spherical planet by the sun).

In other cases, it may be possible to “remove” the objects of interest by a morphological procedure called an opening. As shown in Fig. 5, replacing each pixel by its brightest neighbor, and repeating the operation until the dark letters are removed, and then reversing the operation and replacing each pixel by its darkest neighbor to restore the position of the edges and creases, produces a background image that can be subtracted.
Correction of Defects 4: Geometric Distortion

Measurements are most straightforwardly performed when the image shows the subjects of interest in a normal view of a flat surface. Transforming an image taken at an angle, or of a curved surface, requires knowing the geometry and performing a correction as shown in Fig. 6. Including rulers in images, and locating fiducial marks, is a critical step to enable this procedure and is standard practice for forensic imaging. Pixel values are interpolated from those in the original to generate the corrected image.

Enhancement 1: Histogram Adjustments

After the corrective steps shown earlier, it is often useful to make adjustments to contrast and brightness. This is done by referring to the image histogram, a plot showing the number of pixels as a function of brightness. For color images, there may be a histogram for each channel, but adjustments, like all of the enhancement operations, should be performed on the brightness, luminance, or intensity values leaving the color information unchanged. Attempting to make adjustments to the RGB channels, for example, would alter the relative amounts producing new and strange colors in the resulting image.

When the brightness range captured in the image does not cover the full available dynamic range, a linear stretch of the values can be applied (Fig. 7b). It is important not to push pixel values beyond the black and white limits, causing them to be clipped to those values and data to be irretrievably lost. After the contrast expansion, there are just as many possible brightness values that have no pixels, as shown by the gaps in the histogram, but they are uniformly distributed across the brightness range rather than being collected at one or both ends of the histogram. Linear stretching is not the only possibility. Fig. 7 shows several other possibilities, with the resulting histogram shown for each case.

Adjusting the “gamma” value (Fig. 7d) changes the mid-gray point in the histogram and can expand the contrast for either the bright or dark portion of the image by compressing the values at the opposite end of the range. Rather than this manual adjustment, applying histogram equalization (Fig. 7e) adjusts values so that the histogram is as nearly uniform as possible, and all levels of brightness are represented by equal areas of the image. This is shown in the cumulative histogram, shown in Fig. 7e, which becomes a straight line. Equalization is often useful for comparing images taken under different lighting conditions. A more computationally intensive approach is the homomorphic transformation, which is applied in Fourier space by adjusting the amplitudes of dominant frequencies. In Fig. 7f, details in both the bright and dark regions are clearly evident.

Enhancement 2: Sharpening Detail

Human vision locates lines and edges in images as places where the brightness changes abruptly, and from these forms a mental sketch of the scene. Increasing the local contrast at steps, or narrowing the distance over which the change occurs, makes the image appear sharper. The simplest approach to this is the Laplacian filter, which calculates the difference between each pixel and the average value of its neighbors. A more flexible routine, called the unsharp mask and implemented in many programs, subtracts a Gaussian smoothed copy of the image from the original. This is a high-pass filter (it removes low frequencies or gradual variations in brightness, and “passes” or keeps the high frequencies) and may equivalently be performed using the Fourier transform. The results of all
these “detail extracting” routines are typically added back to the original for viewing.

The most flexible such approach is the difference of Gaussians or DoG filter, which calculates the difference between two copies of the image which have been smoothed with Gaussians having different radii. This is a band-pass filter that selects a range of frequencies and can enhance detail while suppressing high frequency noise as well as low frequency variations. It is shown in Fig. 8b. Using similar logic but calculating the difference between median values in different size neighborhoods (Fig. 8c) requires more computation but is superior in its ability to avoid haloes around edges. Local equalization (Fig. 8d) performs histogram equalization within a local neighborhood and keeps the new value only for each central pixel; this emphasizes fine detail by increasing the difference, either positive or negative, between each pixel and local neighbors.

**Detail Enhancement 3: Defining edges**

In addition to the visual enhancement of images, edges and boundaries are important for the measurement of features. Defining their position may be performed using several different approaches. The most common, the Sobel filter, replaces the value of each pixel with the magnitude of the local gradient of pixel brightness, as shown in Fig. 9b. A different approach, the variance filter (Fig. 9c), calculates the statistical variance of pixel values in a neighborhood, responding strongly to local changes. Both of these produce broad lines because of the size of the neighborhood used for the calculation. The Canny filter (Fig. 9d) begins with the gradient but keeps only those pixels with the maximum value in the gradient direction, producing single-pixel-wide lines that mark the most probable location of the boundary.

In addition to marking the location of boundaries, the brightness gradient vector has a direction that can be used to measure the orientation of edges. Fig. 10 shows the use of the Sobel gradient vector to mark cellulose fibers used in papermaking with gray values proportional to the local angle. A histogram of values, shown as a compass plot, indicates the nonisotropic distribution of fiber orientations.

**Detail Enhancement 4: Revealing Texture**

Features in images are not always distinguished by differences in brightness or color, or by outlined boundaries.
Another criterion can be texture, which can be understood as a local variation in brightness or color. Fig. 11 shows an example: the curds in the cheese do not have a distinct brightness, but have a “smooth” appearance while the surrounding matrix is highly textured.

Processing the image to replace each pixel value with the result from calculating various statistical properties of the local neighborhood can convert the image to one in which the regions have a unique brightness and can be isolated for measurement. The most commonly used properties are the range (difference between the brightest and darkest value) or the variance of the pixel values. In the figure, the fractal dimension has been calculated; this is a more complex calculation that fits the slope (on log-log axes) of the variance as a function of the size of the neighborhood. The resulting difference in brightness allows outlining the boundaries of the curds, so that their volume fraction and surface area can be determined using stereological relationships as explained in the section Measurements 2: Stereology.

**Detail Enhancement 5: Principal Components**

RGB color images, and satellite images covering multiple wavelengths, may be processed using principal components analysis (also known as the Hotelling or Karhunen–Loève transform) to obtain one or more new color channels as a combination of the existing ones, which can provide optimum contrast for the details in a particular image.

This can be visualized as a rotation of the color coordinate axes as shown in Fig. 12. The original image is a fingerprint on a check which has an imprinted texture pattern. In the original RGB channels, the minutiae in the print are difficult to discern. Plotting each pixel’s RGB values in a three-dimensional graph shows correlation, and fitting a plane to the data produces the maximum dispersion of the values and hence the greatest contrast.

Using the position of each pixel’s point along the new principal components axes results in the images shown that separate the fingerprint from the printed background pattern. The third axis, which is perpendicular to the plane, generates an image with little contrast, containing primarily the random noise in the original image.

**Detail Enhancement 7: Image Combinations**

The example of subtracting a recorded background image is shown in the section Correction of Defects 3: Nonuniform Illumination. There are other situations in which two or more images of the same scene may be acquired, for instance using different wavelength bands, or different lighting, or different camera focus. Processing an image may also produce an additional representation (e.g., the Gaussian blurred copy that is subtracted to produce the unsharp mask result).

Arithmetic operations between images are performed pixel-by-pixel, with scaling and offset applied to keep the resulting values within the permitted range (for single-byte images this is 0...255, but some programs accommodate many different bit depths and normalize all of them to 0...1 using real numbers rather than integers).

Either subtraction or division is used for removing background, depending on whether the acquisition device responds logarithmically (like film and vidicon cameras) or linearly (solid state detectors, but the electronics may convert the result to logarithmic in order to mimic film). Division is used to ratio one wavelength band to another, compensating for variations in illumination and, for example, the curvature of the earth. Addition may be used to superimpose difference-of-Gaussian or edge-delineation results on the original image for visual enhancement. Multiplication is less often employed, but is used in graphics applications, for example, to superimpose texture on smooth regions. In addition, mathematical operations include keeping whichever pixel value is greater or smaller, and for
black and white or “binary” thresholded images the various Boolean operations (AND, OR, Exclusive-OR, and their combinations) are useful for combining various selections and information.

When a series of images acquired with different focal planes are captured, they can be combined to keep whichever pixel value at each location gives the sharpest focus, resulting in an extended focal depth. The pixel value selected may be the one with the highest local contrast or variance in its neighborhood. Fig. 13 shows an example, with a map indicating from which original image each pixel in the composite was selected.

**DETAIL ENHANCEMENT 8: DECONVOLUTION**

When the Hubble telescope was first launched, a fabrication error in the curvature of the primary mirror caused the images to be out-of-focus. Several years later a replacement secondary mirror was installed that compensated for the incorrect primary curvature, restored the focal sharpness, and increased the amount of light directed to the instrument package. But in the interim, sharp images were obtained by deconvolution using computer software. If the point-spread-function (PSF) of the optics, which is simply the recorded image produced by point of light like a star, can be either calculated or measured, it can be used to remove much of the blur introduced in image capture, either due to the optics or to motion.

Fig. 14 shows an example. The process is usually performed in Fourier space, with the most basic algorithm (Wiener deconvolution) dividing the transform of the blurred image by that of the PSF, plus a small scalar constant that depends on the amount of noise present. Other methods include iterative techniques that may try to determine the PSF from the image itself (e.g., Lucy-Richardson deconvolution). The results are never as good as a perfectly focused original image, because the noise is increased and not all of the blur can be removed. But the improvement over the original blurred image can be great, and for images such as forensic evidence may be critical.

**Detail Enhancement 8: Cross-Correlation**

Cross-correlation is used to align images, and also to locate a target in a scene. It is often used for aerial surveillance, machine and robotics vision, and finding faces in images. It is frequently carried out using Fourier transforms, but for small targets may be applied in the spatial or pixel domain. It is easy to visualize the process as having the target image on a transparent film and sliding it across all locations in the scene image to find a match. The result is another image in which each pixel records a measure of the similarity of that location to the target. Fig. 15 shows an example. Searching for the target particle shape finds all of the occurrences with high matching scores, in spite of the different contrast for single particles versus those in groups, while
ignoring the background texture of the filter and objects present with other sizes or shapes.

**Binary Images 1: Thresholding (Automatic)**

Except for manual measurements on images, in which a human marks points using a mouse and the computer reports distances, most image measurements are performed after thresholding or segmentation to delineate the objects, structures, or other features of interest. Manual measurements are generally suspect, because of nonreproducibility and the possible influence of expectation or desire. For the same reason, manual thresholding, although often used, is not a preferred approach.

Thresholding selects pixels based on some defining characteristics as belonging to the features of interest. The process may identify all pixels at once as part of one or another of several classes of structure, or simply erase as background those which are not part of the structure of current interest. The simplest of all types of thresholding is based on the brightness histogram of the image, as shown in Fig. 16a.

A peak in the histogram indicates that many pixels have similar brightnesses, which may indicate that they represent the same type of structure. Placing thresholds “between peaks” may distinguish the features of current interest. In the example in the figure, the bright peak corresponds to the paper but there is no dark peak representing the ink. Instead, a statistical test is used to select the threshold value (marked with an arrow) that is used to (hopefully) isolate the printed characters for measurement and ultimately identification.

The test illustrated is one of the most widely used, producing often satisfactory and at least reproducible results. It uses the Student’s t-test to compare the values of pixels above and below each possible threshold setting and selects the one that produces the greatest value of t. This indicates that the two groups are most different and distinguishable. However, the statistical test makes the tacit assumption that the two populations have Gaussian or normal distributions, which is rarely the case. There are a variety of other statistical tests, which use entropy, fuzzy weighting of values, and other means, and which produce somewhat different threshold settings.

A different approach to automatic threshold setting uses not only the value of the pixels, but also those of their immediate neighbors. The logic behind the test is that pixels within features, or within background, should be similar to their neighbors, while those along borders should not. A co-occurrence matrix that counts the number of pixels with each value along one direction and the number with each average neighbor value along the other is used, as indicated schematically in Fig. 16b. In the figure, the areas marked A and C are the more uniform regions, corresponding to features and background in which pixels are similar to their neighbors. Those marked B and D represent the borders where they are different. Iteratively adjusting thresholds to maximize the total counts or the entropy in A and C and minimize those in B and D is a more computer intensive approach, but is superior in performance.

Another powerful method is indicated in Fig. 16c. The k-means procedure is particularly appropriate for color or multichannel images (the figure shows just two dimensions, but the method generalizes directly to any number). The values of all pixels are plotted and the method searches for clusters. An initial set of k locations are selected arbitrarily, and all pixel points that are closest to each location are temporarily given that class identity. The mean of each class is then used as the next proposed cluster center, and the procedure is repeated. This causes some points to change identity, and the cluster boundaries and cluster means to change. The procedure continues until no further changes take place.
**Binary Images 2: Thresholding (Interactive)**

Other approaches to thresholding or segmentation are sometimes useful, and may involve some degree of human interaction. Fig. 17a shows the use of an edge-marking procedure as described in the section Detail Enhancement 3: Defining Edges to outline each object. Since there are also some lines and edges drawn within the objects (e.g., the pumpkin ridges and stems), it becomes necessary for the operator to select those lines that are the object boundaries, but which have been located automatically.

Human selection also operates in the seed-fill or region-growing approach, shown in Fig. 17b. Marking an initial point (indicated by an asterisk in the figure) begins the process. Then every neighboring point is examined and ones that are similar are added to the growing region. This continues until no further neighboring points are added. The resulting selection is outlined in the figure. The test for similarity can be a fixed range of color or brightness, or it may be based on the statistics of the growing region, or may be weighted toward the values of the pixels near the local expanding boundary. The most common problem with region growing is that it may “escape” from the feature and become free to spread across background or other objects.

Fig. 17c illustrates the active contour approach. It begins with a manually drawn outline, which then contracts until it is stopped by the borders of the object (active contours that expand from an inner outline can also be used). The stopping may be based on color or brightness, gradient, or other criterion. Active contours can bridge over gaps where the border is indistinct because the shrinking criterion seeks to minimize the energy in the boundary, based on its length and curvature. Active contours may be called “snakes,” and when applied in three dimensions (3D) are referred to as “balloons.”

These are not the only approaches used for thresholding and segmentation. Top-down split and merge segmentation examines the histogram for the image, and if it is not uniform by some statistical test it divides the area into parts. Each of these is examined similarly and divided, and the process continues. At each iteration, adjacent regions with different previous parents are compared and joined if they are similar. The final result reaches the level of individual pixels and produces a set of regions. Other computer-intensive methods include fuzzy approaches to cluster analysis that weight pixels by how different they are from the cluster mean, and neural net approaches which begin with the entire array of pixel values as input.

**Binary Images 3: Morphological Processing**

Thresholded images are often imperfect delineation of the features or structures of interest. Random variations in pixel values may cause some individual errors, boundaries may be poorly defined if the finite size of pixels straddle them and have intermediate values, and some pixels may have values that are the same as those within the structures of interest. These flaws are usually small in dimension (often single pixels) and are dealt with by morphological operations of erosion and dilation, which remove or add pixels according to the identity of their neighbors.

Dilation in its simplest form adds background pixels that are adjacent to a feature boundary, and erosion removes feature pixels that are adjacent to background. Since each of these changes the size of the object, they are usually used in combination. Fig. 18 shows an example, in which a closing (the sequence of dilation followed by erosion) is able to fill internal gaps without changing the external dimensions of the fibers. The opposite sequence, erosion followed by dilation, is called an opening and is used to remove background noise or speckle.

Continued erosion with a rule that a pixel may not be removed if it causes an object to divide into two parts generates the feature skeleton. An alternative method assigns to each pixel within a feature a value that measures its straight line distance to the nearest background point. The ridges in this Euclidean distance map (EDM) define the skeleton and their values form the medial axis transform, which is often useful for measurement purposes.

In the example in Fig. 19, the number of end points in the skeleton (pixels with only one neighbor) identifies the number of teeth in the gear. In other cases, the number of node points (pixels with more than two neighbors) measure network connectivity. Euler’s rule.

---

**Fig. 17** Additional segmentation methods: (A) edge delineation; (B) region growing; (C) active contours.

**Fig. 18** Applying a closing: (A) original image showing cross-sections of glass fibers; (B) thresholded image showing cracks; (C) filling the cracks with a closing.
for the topology of skeletons in two-dimensional images is (Number of Loops – Number of Segments + Number of Ends + Number of Nodes = 1).

The EDM is also used to separate touching features, as shown in Fig. 20. The watershed segmentation method considers “rain” falling on the EDM and proceeds downhill from the peaks to locate points that would receive runoff from more than one initial peak. These locations mark watershed boundaries and are removed, leaving separated features for measurement. The method works for mostly convex features that have only a single peak in their EDM, with overlaps less than their radii.

**Measurements 1: Photogrammetry**

Dimensions and spatial arrangements of objects in 3D scenes can be determined from measurements on images. In some cases, such as accident reconstruction, image measurements are used to construct detailed 3D models. Sometimes measurement is based on multiple images taken from different positions, for example, stereo pair images, employing trigonometry. But even single images often can be accurately interpreted to determine 3D information.

For example, knowing the location and lens specification of a surveillance camera makes it possible to determine the height of a person from the image. This can be done trigonometrically, but a scaled drawing of the geometry also provides a solution and is easier to explain, for instance to a nontechnical jury. An even simpler method, called “reverse projection,” requires taking a suitable measuring ruler to the scene and recording its image using the same camera and geometry, and then superimposing the two images as shown in Fig. 21 so that the height or other dimension can be read directly.

Another forensic example is the measurement of a blood spatter pattern. The elongation of each droplet gives the angle and direction from which it arrived at the surface (a wall, floor, table, etc.). The intersection point of lines projected back in the indicated directions locates the point in space where the droplets originated, which is the exit wound from a gunshot and hence determines the location of the victim when shot.

**Measurements 2: Stereology**

Sections through 3D samples are typically imaged in various kinds of light and electron microscopes, and are also produced by tomographic imaging using light, X-rays, sound, neutrons, and many other signals. The features revealed in these section images do not directly show the size or even the number of objects present in the space, because the sampling plane may pass through any portion of the object, not necessarily showing its full extent. However, it is possible using rules derived from geometric probability to infer many important structural parameters including the volume fraction, surface area, length, curvature, number, and connectivity of the objects.

This field is known as stereology (from the Greek for study of three-dimensional space). Many of the rules and procedures are simple to apply and involve counting of “events”—the intersection of the structure(s) of...
interest with properly designed grids of lines or points—rather than the measurement of dimensions. The key to using stereological relationships is understanding that a section plane intersects a volume to produce an area, intersects a surface to generate a line, and intersects a linear feature producing points. In all cases, the dimension of the structure of interest is one greater than the evidence found in the image.

For example, the volume fraction of a structure is measured by the fraction of points in a regular or random grid that fall on the structure. The surface area per unit volume is equal to two times the number of intersections that a line grid makes with the surface, divided by the total length of the line, or to \((4\pi/3)\) times the length of the boundary line divided by the image area. The length of a linear structure per unit volume is two times the number of intersection points divided by the image area. In all cases, care is needed in the design of grids and the sectioning techniques used in order to produce unbiased results. This somewhat specialized topic is well covered in texts such as Baddeley and Vedel Jensen.[13]

Measurements 3: Feature Brightness, Size, and Location

The measurements of individual features in images fall generally into four groups: brightness or color, location, size, and shape. It is also important in many cases to count the number of features present. Fig. 22 shows an image of rice grains captured using a desktop flatbed scanner. Some of the rice grains intersect the edges of the image, indicating that this is a sample of a larger field of objects. One unbiased procedure for counting the number per unit area is to count as one-half those grains that intersect the edges, since the other “half” count would be obtained if the adjacent field of view was measured.

For measurement purposes, the edge-intersecting grains cannot be used, as their dimension is unknown. Since large objects are more likely to intersect an edge, the bias in a measured size distribution such as the one shown in the figure can be compensated by counting each measurable grain with a weighting function equal to \((W_x-F_x)(W_y-F_y)/(W_x\cdot W_y)\), where \(W_x\) and \(W_y\) are the dimensions of the image in the \(x\) and \(y\) directions, and \(F_x\) and \(F_y\) are the projected or box dimensions of each object in those directions. For very small features, this weight is nearly 1, but for large features it is greater than one to compensate for other similar size objects that would have intersected the borders of the image and been excluded from the measurements.

The distribution of the length of the rice grains is used, for example, to determine that the sampled rice has a small percentage of short grains and can be sold as “long grain” rice. There are many other useful measures of size, such as area (which may or may not include internal holes and peripheral indentations), the radii of the maximum inscribed and minimum circumscribed circles, and the perimeter.

Perimeter is the most difficult measurement to determine properly. It may be calculated using the center-to-center path through the boundary pixels, or along their outer edges, or by fitting smooth curves, and these all give slightly different results. More important, the perimeter depends on the pixel size and resolution of the original image, and in many cases as magnification increases the resolution reveals more and more irregularities, so that the perimeter is not a well-defined concept. Indeed, the rate at which perimeter varies with resolution is one of the ways to determine the fractal dimension of a shape[14]

Pixel brightness values can be calibrated to measure density and other object parameters, but the values recorded in the RGB channels cannot be used to measure color in the sense of a spectrophotometer. This is because the filters used in cameras cover ranges of wavelengths so that different combinations of intensity and wavelength can produce identical results. This is also true for satellite images, which record many bands with each one covering a range of visible or infrared wavelengths.

The location of objects can be determined as their centroids, which may be weighted by density determined from the pixel values. Location may also be based on the center of the circumscribed or inscribed circles in some cases; the latter location is the only one guaranteed to lie within the boundary of the object. One use of location data for a collection of objects is determining whether the objects are clustered, randomly arranged, or self-avoiding. Cacti in the desert are...
Measurements 4: Feature Shape

Shape is a difficult concept to describe, and humans generally resort to nouns rather than adjectives (“... shaped like a ...”). “Round” may mean “like a circle” (or a sphere or cylinder) but might also mean without indentations and sharp corners. “Skinny” and “bent” generally have meaning only by comparison to other forms. Putting numbers to shape description is complicated as well. The simplest and most widely used approach to measuring shape uses dimensionless ratios of size measurements. Table 1 lists a few as examples, but it should be understood that various names are assigned to these relationships with no consistency, and that it is possible to have shapes that are visually entirely different that share values for one or several of these ratios.

To illustrate the use of dimensionless ratios, a collection of leaves from various trees were used. Fig. 23 shows representative examples (not at the same scale), with a plot of the values for three of the shape factors that are able to identify the various species based on shape alone. The regions occupied by the points in each class are irregular, and improved results can be obtained by using linear discriminant analysis to calculate canonical variables, which are linear combinations of the measured parameters. This produces the plot shown in the figure, in which each class is represented by a spherical region centered on the mean value with a radius of two standard deviations.

Other methods for shape description can also distinguish all of these classes. The principal ones in use are harmonic coefficients and moments. The former is based on the periphery of the feature, for example, expressing the point coordinates along the boundary in complex form (x + iy). A Fourier transform of the boundary then represents the shape as a series of terms, and the amplitudes can be used as numeric shape descriptors. Instead of a Fourier transform, a wavelet transform may also be used.

Moments, on the other hand, use all of the interior pixel coordinates as well, which can be an advantage if the boundary is poorly defined, or when the shape consists of multiple parts (e.g., an animal paw print). There are invariant moments that may be used to describe shape. Both the harmonic coefficients and the moment values can be used in subsequent statistical analysis for comparison and correlation.

MEASUREMENTS 5: DATA ANALYSIS

Measurements on objects and structures obtained from images are typically used for descriptive statistics and classification, and for correlation with object history or function. The common statistical parameters (mean, standard deviation, etc.) are convenient but make the tacit assumption that the values are normally distributed, which is not always the case (especially rarely so for shape parameters).

Nonparametric comparisons between data sets using Mann–Whitney or Kolmogorov–Smirnov statistics are preferred, as they yield meaningful probabilities whether the data are normal or not. Likewise, correlation based on rank order (Spearman correlation) is preferred over the usual Pearson correlation if relationships may be nonlinear. The interpretation of the r-squared value is the same in both cases.

Classification based on measurements such as those shown in Fig. 23 may use linear discriminant analysis, neural nets, fuzzy cluster analysis, or k-nearest neighbor tests. These are standard tools for treating data, not limited to measurements from images, and are well covered in most statistics texts.

A particular interest for image analysis is database searching. Landmark methods, such as the Automated Fingerprint Identification System, work by using the relative location of multiple points. For fingerprints, these are minutiae such as the gaps, ends, and bifurcations of the ridge lines in the print. A list of 12–16 such landmarks can call up the 10 or so most similar fingerprints on file, for a human to compare. Similar use of human judgment of a small number of “most like” selections found by automatic search algorithms is used in medical

<table>
<thead>
<tr>
<th>Table 1 A few dimensionless ratios that may be used to describe shape</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \text{Radius Ratio} = \frac{\text{Inscribed Diameter}}{\text{Circumscribed Diameter}} )</td>
</tr>
<tr>
<td>( \text{Roundness} = \frac{4 \cdot \text{Area}}{\pi \cdot \text{Max Diameter}^2} )</td>
</tr>
<tr>
<td>( \text{Formfactor} = \frac{4\pi \cdot \text{Area}}{\text{Perimeter}^2} )</td>
</tr>
<tr>
<td>( \text{Aspect Ratio} = \frac{\text{Max Caliper Dimension}}{\text{Min Caliper Dimension}} )</td>
</tr>
<tr>
<td>( \text{Solidity} = \frac{\text{Area}}{\text{Convex Area}} )</td>
</tr>
</tbody>
</table>
diagnosis, such as the analysis of Pap smears and mammograms.

An elusive goal for image analysis is “query by example” in which the presentation of an image is used to locate other images of similar objects. The problem is that with a few exceptions such as finding paintings with the same predominant color(s), it is not easy for computer algorithms to decide what it is that the presenter believes to be the important characteristics of the example image. Online Internet searches for images work using the words in accompanying text, not the contents of the images themselves.

ACKNOWLEDGMENT

The explanations and topics covered, and the examples shown, are from The Image Processing Handbook (John C. Russ, CRC Press, 2011). More detailed information, additional examples and comparisons of algorithms, and extended references are available there.

REFERENCES