Image Analysis: Application to Automated Medical Diagnosis

The potential of optical scanning equipment and digital computers for assisting or replacing human judgment in medical diagnosis has been recognized for some time. Efforts have been made, with varying degrees of success, to develop techniques for recognizing and classifying blood cells, for typing chromosomes, and for identifying cancer cells. These image analysis techniques have now been extended to radiology, and in one of the first such applications doctors and engineers at the University of Missouri at Columbia have demonstrated the feasibility of diagnosing rheumatic heart damage with an automatic system for analyzing chest x-ray images. Trials with the image analysis system showed that its diagnosis was more accurate, on the average, than that of a panel of radiologists.

The basic technique involves the conversion of an image—whether a photomicrograph of a cell or an x-ray film—to a two-dimensional array of numbers which represents the image, and the analysis, with pattern recognition methods and computers, of the digitized image. For many applications the goal is not just the enhancement of visual information—as is the case, for example, with the digital techniques used in the space program to improve the quality of satellite photographs—but rather the extraction of nonvisual information from the image.

Image analysis techniques are still experimental in many medical applications, and they are often still too costly for widespread use. Both university scientists and several industrial companies are actively working on improving these techniques. Image analysis is also being applied to many nonmedical problems as well. Much of the early research in image analysis, for example, was sponsored by the Air Force, and work on applications to industrial quality control and military intelligence is also going on, in some cases in the same laboratories that are working on medical applications. Use of image analysis techniques for some nonmedical applications may in fact, because of the greater resources that are available, become widespread long before those in medicine.

In their application to rheumatic heart disease, a University of Missouri team consisting of Gwilym Lodwick, Samuel Dwyer III, Charles Harlow, and others has refined the image analysis process to three major steps. First the image is converted to digital information, then the pertinent features of the image are extracted and measured, and a diagnosis is made according to probabilistic logic built into the computer program.

In the first step, the x-ray images are scanned with a specialized camera that produces an output voltage whose average value is proportional to the amount of light transmitted through a particular point on the x-ray. The optical image is thus converted point by point into a two-dimensional array of numbers, each of which represents the "grayness" of the image—that is, the gradation of light and dark on the film—at a particular location.

The camera can distinguish 64 shades of gray, and is capable of resolving an image into an array as large as 1024 by 1024, although for the rheumatic heart application only 256 by 256 data points are taken. The amount of data to be processed and the cost of analyzing increase rapidly with greater resolution. The camera is essentially an electron tube on which the optical image is converted to an electrical image. The electrical image is then sampled, and the resulting analog output signal from the camera is filtered to eliminate background noise by integrating the signal over a period of about 1 millisecond. The signal is then converted to a digital value and stored on magnetic tape; because the response of the human eye is proportional to the logarithm of the intensity of light, it is this logarithmic value that is stored. The system can sample at the rate of 1000 points per second.

Once the x-ray image has been converted to a series of grayness values, the data are processed with a computer for the purpose of identifying and extracting the features of the image. This is the most difficult part of the image analysis process, since the chest x-ray typically contains interfering overlapping objects, such as the heart, lungs, and rib cage, which can be of varying sizes and shapes. The Missouri team has evolved several empirical techniques to locate the heart within the image and to measure its size and shape. Heart dimensions are important because, in rheumatic heart disease, the organ is characteristically enlarged or disfigured.

The computer first searches the data to establish the approximate position of all the major features (that is, the heart, lungs, arms) by scanning for regions whose intensities are roughly continuous and whose size and average gray values are close to those expected. An alternative approach to locating the heart involves summing the gray values of the image by row and column; analysis of the resulting simplified "signature" of the image is used to help identify the heart. The tentative identification is tested by the computer by ascertaining the physiological reality of the relative positions of the regions—that is, whether the right lung is to the right of the heart. Once the major features have been crudely identified, the computer then searches for details of the feature by examining points bordering the initial features and assigning the points appropriately. Overlapping of the different regions within the image determines the boundaries of the major features and thus the outline, for example, of the heart (see Fig. 1). The process continues until all data points have been assigned to a definite feature.

Once the heart has been located and its outline has been determined, the computer extracts information on which diagnostic decisions can be made. Curves (fourth order polynomials) are fitted to the left and right boundaries of the heart, horizontal and diagonal dimensions are measured, and the ratio of heart to lung area is determined. A total of 13 measurements are made.

Rheumatic heart disease shows up primarily in adults. Usually the result of a streptococcus infection in combination with rheumatic fever, the disease attacks the valves and muscles of the heart, causing scarring and calcification of the valves and dilation of the heart. The disease no longer presents as large a medical problem as, for example, congenital heart damage, and both the diagnosis and treatment of the

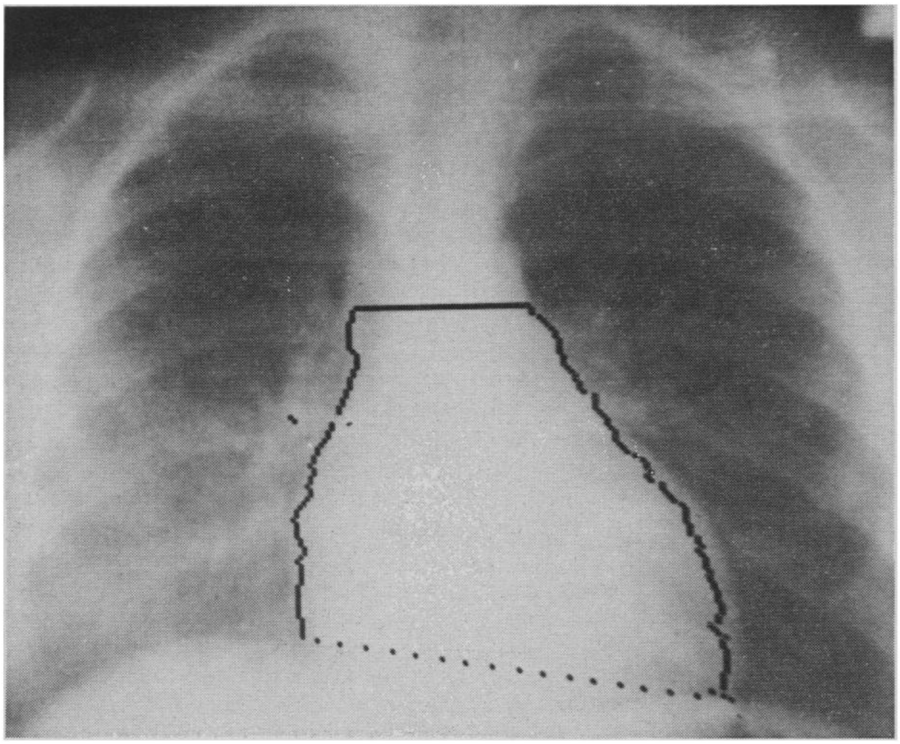

Fig. 1. X-ray of a rheumatic heart whose shape has been determined by an automatic image analysis system. [University of Missouri]

disease are well understood. The Missouri team chose rheumatic heart disease for the initial application of their techniques because it was well understood and because an adequate number of cases was available.

In a comparison of chest x-rays of both normal and damaged hearts measurements were tested to see which were most effective as predictive variables. The effectiveness of a particular measurement and the optimum combination of factors was thus empirically determined. In practice, diagnosis by the computer is made with a Bayesian probabilistic method. Analysis of x-rays from some 281 patients with clinically confirmed symptoms resulted in a table of probabilities that gave the likelihood of a particular diagnosis as a function of the measured quantities. In a particular case, the computer obtains the measurements from an x-ray image, decides on the basis of the probability table whether the heart is normal or abnormal, and, if abnormal, selects one or more of four different types of rheumatic abnormality as the most probable diagnosis.

The image analysis system was tested against a panel of ten radiologists on a series of 135 patients for which the correct diagnosis was known. The radiologists, using both front and lateral x-rays, made individual diagnoses, whereas the computer used only the frontal x-ray. The overall accuracy of correct diagnosis by the radiologists was 62 percent as compared to an accuracy of about 73 percent for the computer. The computer's accuracy was also higher than that of any individual radiologist. In determining abnormal hearts, the radiologists and the computer achieved a comparable accuracy, but the radiologists had a greater tendency to make false positive identifications, and the computer was more accurate in placing abnormal cases into the correct class of rheumatic heart disease. The performance of the image analysis system and the accuracy of the fully automated diagnosis thus established, in Lodwick's view, the feasibility of using image analysis for this medical application. Clinical evaluation of the system is now underway at the University of Missouri.

Although rheumatic heart disease is not the most pressing of medical problems, image analysis may be of practical importance. In rural medical practice, for example, a central facility equipped with a computer could process and evaluate x-rays from areas where trained radiologists are scarce. Image analysis techniques may eventually provide a consistent standard of comparison for what is often widely varying clinical judgement, and thus lead to means of monitoring and supplementing medical diagnostic practices. The cost of the system is still somewhat high (the analysis of each x-ray requires about 3 minutes of computer time on an IBM 360/50), but as part of a network serving large numbers of users the system may become economically feasible.

The Missouri scientists have begun work on other radiological applications of image analysis, including an attempt to diagnose congenital heart damage from chest x-rays and efforts to identify and classify bone tumors from x-ray films. They are also trying to develop means of evaluating coal miners' pneumoconiosis (black lung disease).

Elsewhere other medical applications are also being pursued, including methods to automate differential blood counts, to characterize chromosomes for karyotype analysis and for research in cytogenetics, and to identify cancer cells. The feasibility of image analysis for blood cell recognition was established several years ago when a group headed by P. Mendelsohn of the University of Pennsylvania showed that white blood cells could be accurately classified with an automatic system. Several commercial machines are now available for classifying blood cells, although development of new methods is continuing. Image analysis systems for typing chromosomes, such as that developed by R. S. Ledley of Georgetown University in Washington, D.C., are also available. Other groups are developing these techniques for research purposes—Mendelsohn's group, for example, is analyzing the DNA content of chromosomes.

A group headed by George Wied of the University of Chicago and Peter Bartels of the University of Arizona is trying to automate the examination of endometrial cells for cervical cancer; they hope to improve the reliability of diagnosis in early malignant and premalignant tumor cells, and, by modeling the evolution of the disease, to make objective comparisons of different therapies.

The image analysis techniques used in all these applications are basically similar to those that have now been applied to rheumatic heart disease. Perhaps the most significant feature of these efforts is their interdisciplinary nature, which, in the case of the Missouri group, involved the sustained cooperation of electrical engineers and radiologists over a period of 5 to 6 years—a cooperation that may eventually improve the practice of medicine and make its benefits more widely available.—ALLEN L. HAMMOND