Synthesis of Research ■

Medical Image Databases: A Content-based Retrieval Approach

HEMANT D. TAGARE, PHD, C. CARL JAFFE, MD, JAMES DUNCAN, PHD

Abstract   Information contained in medical images differs considerably from that residing in alphanumeric format. The difference can be attributed to four characteristics: (1) the semantics of medical knowledge extractable from images is imprecise; (2) image information contains form and spatial data, which are not expressible in conventional language; (3) a large part of image information is geometric; (4) diagnostic inferences derived from images rest on an incomplete, continuously evolving model of normality. This paper explores the differentiating characteristics of text versus images and their impact on design of a medical image database intended to allow content-based indexing and retrieval. One strategy for implementing medical image databases is presented, which employs object-oriented iconic queries, semantics by association with prototypes, and a generic schema.

■ **J Am Med Inform Assoc.** 1997;4:184–198.

Medical knowledge, examined from an "information" perspective, arises in highly diverse forms. For example, consider two important classes of medical knowledge: anatomy and physiology. Anatomic information rests on visual appearances (e.g., "enlarged heart"). Physiologic information arises from biologic processes, and it may not be visual. It could include data about metabolism, diet, age, environment, exercise, numeric parameters from physiologic tests such as blood pressure, etc. Quite often, anatomic and physiologic information are obtained simultaneously, but only the text-numeric information, such as blood chemistry values, is conveniently stored in a database.[1] Several such conventional text-numeric databases with sophisticated indexing and search mechanisms have been developed, (e.g., the human genome data bank).[2] Intense commercial activity, much of it focused on developing effective search engines for the massive text/numeric repositories of digital files on the World Wide Web (WWW), is being applied to retrieve content-related documents (Lycos, Inktomi, Altavista, Yahoo, Hotbot, etc.). Text indexing by concordances of keywords can imply a massive inverted index table, and weighting functions implemented on metathesauri or text pattern associations are conceptually understandable. But numerical concepts applicable to content-based image indexing—methods not dependent on text-based key words or other alphanumeric identifiers—are often less intuitive.

This article points out some of the unique challenges confronting retrieval engines for medical digital image collections and describes a successful example of a topologic approach devised by the authors that employs geometric properties applicable to tomographic images of body organs. That approach, based on interactively segmented image abstracts, illustrates one tractable problem with a satisfactory solution possible amongst the diverse technologies that give rise to medical images. Medical images arising from photography (e.g., endoscopy, histology, dermatology), radiographic projection (e.g., x-rays, some nuclear medicine), and tomography (e.g., CT, MRI, ultrasound) impose unique, image-dependent restrictions on the nature of features available for abstraction. Automatic, medically useful image abstraction processes

capable of structural or texture pattern retrieval matching have had limited success to date. The motivation for developing new retrieval methods applicable to large image databases rests on the need for disease research on groups of technologically related (e.g., MRIs, CTs) images that share some appearance of diagnostically relevant elements.

Research Problem

The process of determining relevant image features is often complicated by contradictory tensions at work when images are viewed for diagnostic purposes. A duality arises from the simultaneous but cognitively separable processes in which a global gestalt diagnostic impression is formed simultaneously with an awareness of evidentiary sub-element features. For example, a diagnostic conclusion drawn from an image is often greater than, and not merely a result of, an assemblage of small decisions about the existence of particular elemental features (e.g., congestive heart failure diagnosis on X-ray is not a deterministic conclusion from the presence of an enlarged heart and vascular prominence). Thus, diagnostic classifications may be distinct from explanations rationalized from the sum of anatomic features identifiable on an image. Hence, retrieval of groups of images sharing a common feature but perhaps not the same diagnostic classification can be motivated by the intent to better understand the expression of disease. The computational tools applicable to visually perceptible features commonly rest on histograms of hue, saturation and intensity, texture measurements, and edge orientation, as well as on object shape calculated over the whole or some designated local area of an image.

Digital networks have begun to support access to widely distributed sources of medical images as well as related clinical, educational, and research information. The information, however, is voluminous, heterogeneous, dynamic, and geographically distributed. This heterogeneity and geographic spread create a demand for an efficient picture archiving system, but they also generate a rationale for effective image database systems.[3] Without development of the latter, the former would act as a means of communication but would not produce significant new medical knowledge. Picture collections remain an unresolved challenge except for those special class of images adaptable to geographic information systems (GIS), in which conventional geometry and verifiable ground truth are available.

In medicine to date, virtually all picture archive and communication systems (PACS) retrieve images simply by indices based on patient name, technique, or some observer-coded text of diagnostic findings.[4-6] Using conventional database architecture, a user might begin with an image archive (an unorganized collection of images pertaining to a medical theme— e.g., a collection of magnetic resonance cardiac images) and some idea of the type of information needed to be extracted. Fields of text tags, such as patient demographics (age, sex, etc.), diagnostic codes (ICD-9, American College of Radiology diagnostic codes, etc.), image view-plane (saggital, coronal, etc.), and so on usually are the first handles on this process. There are a number of uses for medical image databases, each of which would make different requirements on database organization. For example, an image database designed for teaching might be organized differently than a database designed for clinical investigation. Classification of images into named (e.g., hypernephroma, pulmonary atelectasis, etc.) or coded diagnostic categories (e.g., ICD-9) may suffice for retrieving groups of images for teaching purposes. In the case of text databases, tables of semantic equivalents, such as can be found in a metathesaurus, permit mapping of queries onto specific conventional data fields. Textual descriptors, however, remain imprecise markers that do not intrinsically lend themselves to calculable graded properties. For example, thesaurus entries commonly imply related but nonsynonymic properties, as seen in the terms used to describe variant shapes of the aorta: tortuous, ectatic, deformed, dilated, bulbous, prominent.

This textual approach, however, fails to fully account for quantitative and shape relationships of medically relevant structures within an image that are visible to a trained observer but not codable in conventional database terms. The development of suitable database structures addressing the visual/spatial properties of medical images have lagged. More effective management of the now rapidly emerging large digital image collections motivates a need for development of database methods that incorporate the relationship of diagnostically relevant object shape and geometric properties.[3] However, unlike maps, whose conventional geometric properties make them suitable for graphic information systems, the concepts of content and objects relevant to medical image databases must accommodate the heterogeneity, imprecision, and evolving nature of medical knowledge.

Recently, some investigators have proposed image database structures organized by certain properties of content.[7-10] Most of these techniques are devoted to indexing large conventional collections of photographic images for the purpose of open-ended browsing[11,12] or fixed objects found in industrial parts.[7,13-15] For example, the Query By Image Content (QBIC) sys-

tem rests on color histogram extraction. This permits queries based on color percentages, color distribution, and textures. Automatic color abstraction has advantages because it is easily segmented, but the approach would only be applicable to medical images based on light photography (e.g., dermatology), where color is an inherent feature. Moreover, lacking reasoning procedures, all other metadata is left untapped and unreachable. If the abstraction does not address the image property the query is suited for, the search is hopeless, and retrieval of an complete set of appropriate images residing in the collection is unlikely.

Requirements for medical image databases, however, differ substantially from those applicable to general commercial image collections (commonly referred to as "stock house" photo collections). To appreciate the difference, we can categorize databases along three dimensions: (1) The extent to which the database schema can understand and reason about its content. We will call this the "content understanding" axis. (2) The ease with which the database query mechanism allows the user to specify what the user wants. If the

database does not allow easy and intuitive translation of users' common queries, then it cannot guarantee all relevant data have been retrieved from the database. This can be called the "query completion" axis. (3) In addition, there is the extent of interaction required by an image librarian at data entry or by the end user at retrieval. We will call this the "user interaction axis." Figure 1 shows a conceptual model of the *content understanding–query completion–user interaction* space, plotting the location of text databases, commercial image browsing databases, and medical image databases.

Most commercial text databases lack implementation of mechanisms for reasoning on elements of their content. Aside from databases employing domain-specific semantic nets, conventional databases operating on strings do not present the user with a reasoning environment for data retrieval. Data are treated either as numbers or as strings. Data-entry requires significant interaction with the database to specify a complete set of semantic relationships. However, having carefully defined these semantic relationships, text databases behave deterministically and guarantee full query completeness. That is, the database guarantees that all data satisfying the query are successfully retrieved. Thus, text databases are located in the corner of the space characterized by low content understanding, high user interaction (at least at data entry), and high query completion (all relevant items successfully retrieved).

Commercial image databases intended for (photograph) browsing are equipped with only a rudimentary understanding of content.[16] Most of these databases do not distinguish between foreground and background (i.e., important and unimportant features) or between multiple objects in the image. Images are often indexed by features that characterize the entire image rather than by unique objects present in the image. For example, an image indexing scheme for stock-house advertising photographs, like QBIC[12] and others,[17,18] can index by dominant color or texture properties as well as by keywords, so "red sunsets" may be retrieved. Consequently, these schemes have low data-entry costs. The database can automatically compute the image index as the image is entered into the database. But these databases cannot guarantee query completion—a mechanism by which all images in the collection that would satisfy a query are guaranteed to be retrieved. Query completion is possible only if the user can successfully adapt his or her query to properties allowed by the precomputed image features. Consequently, there can be no guarantee that a complete sequential examination of the collection might not uncover additional images that should have
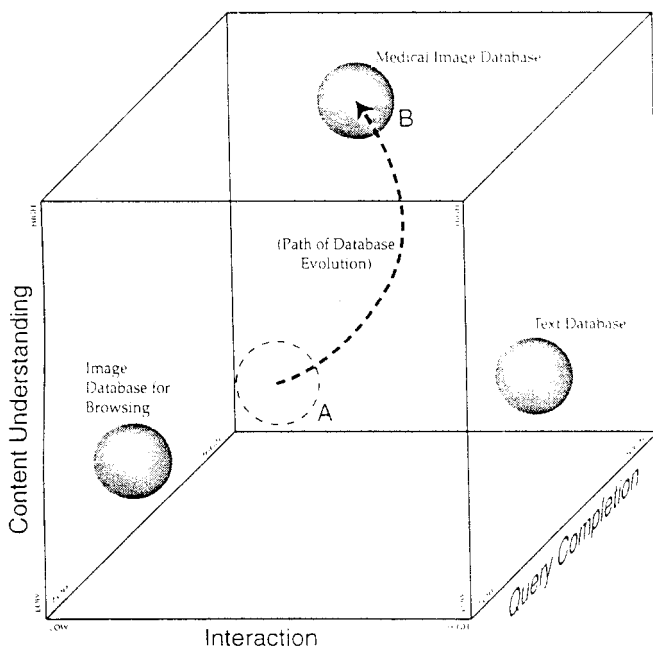

Figure 1 A conceptual model of the *content understanding–query completion–interaction* space, plotting the location of text databases, commercial image browsing databases, and medical image databases. Note the scale designations from low to high. The lowest degrees of each property are located in the lower left hand corner and the highest lie in the farther right top corner. The evolving property of content-based medical image databases is depicted by the first location at its original implementation, point "A," but evolving over time to point "B."

satisfied the query. Thus, image browsing databases are located close to the "LLL" corner of the space—low content understanding, low demand on user interaction at entry, and low guarantee of complete retrieval of appropriate images in response to a query.

Medical image databases, however occupy a distinct location in the *content understanding–query completion–user interaction space*. These databases demand a moderate-to-high degree of content understanding. To be useful they need to account for the elemental structures within images because organs, their relative locations, and other distinct features are likely properties intended for retrieval. The database also has a moderate-to-high requirement for query completion. While moderate query completion might be satisfactory when the database is used in a browsing mode, high query completion is required when the database is used for comparative diagnosis or medical research. Clinical users would find it disconcerting if a nearly complete set of relevant cases were not retrieved in response to a well-formed query. Consequently, if an image database can provide content understanding at an organ level and can guarantee a high level of query completion, the user may be willing to invest moderate effort in the entry interaction process at the time each image is added to the collection or at retrieval time with regard to the effort needed to specify an effective query.

Medical image databases developed for content-based retrieval have one more unique characteristic that distinguishes them even from other standard relational database management systems that require schema evolution. In the case of image databases, their location in the *content understanding–query completion–interaction* space evolves in a more complex way over time. As argued below, medical image understanding is imprecise, and even expert diagnosticians cannot, at outset, indicate how to convert what they perceive as image information into purely quantitative properties. In our experience, the act of developing the database itself serves to further refine the concepts, features, and necessary image processing. Over the lifetime of such a database, the structure and effectiveness typically evolve from location A in Figure 1. That initial location is characterized by moderate content understanding, moderate user interaction, and low query completion, since at outset the users' requirements have not yet been satisfactorily translated into mathematical features. As the database evolves, it typically follows the trajectory to point B, where, after iterative redefinition of concepts and features, it should settle into acceptable performance at high levels of query completion and image understanding. A key point here is that an evolutionary capacity is an essential requirement, and therefore appropriate tools
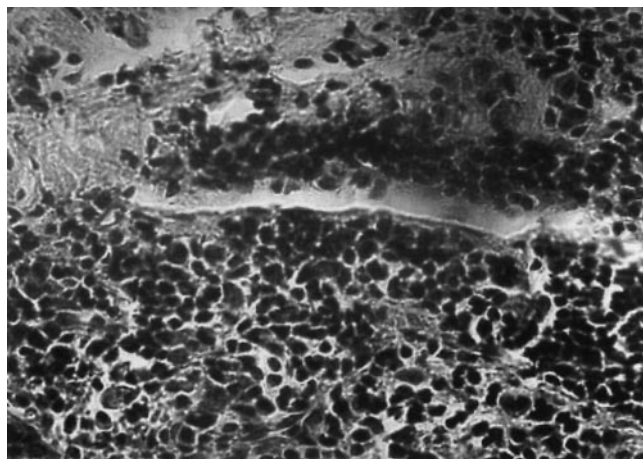

Figure 2 Medical images created by diagnostic instruments can result in large digital collections. Microscopic histology images possess unique color signatures and cell textures that might be used as indexing methods for grouping slides that share similar staining techniques. Thus, a database indexing scheme could take into account color hue as an indexing feature. A query structure could therefore be devised that would make it possible for a user to retrieve images that share a common staining technique.

must be designed at the outset to allow the expected evolution of the database. A medical image database design that permitted only fixed field structure mechanisms would result in a brittle application that would be discarded with each advance in medical image understanding.

Image Classes as Database Candidates

Medical images created by diagnostic instruments offer digital collections of substantial size, although they do not represent the complete spectrum of images for which image databases might be desirable. Microscopic histology (Fig. 2) or even macrophotography used in dermatology could suggest different database organizational ideas because of their inherent properties of color, shading, and resolution. For example, those two latter examples suggest a database indexing scheme that could take into account color hue as an indexing feature.[18] Thus, a query structure could be devised to retrieve images sharing a common staining technique. Ultrasound images of large organs with relatively uniform tissue such as spleen or liver (Fig. 3) present relatively homogeneous image patterns. These might be indexable by mathematical image processing approaches that characterize global image texture (e.g., the Fourier or Laplace transforms).

Diagnostic images from computed tomography, magnetic resonance, nuclear medicine, and radiography

Figure 3 Ultrasound images of large organs (here the liver) appear to be dominated by textures and may lend themselves to indexing schemas that emphasize extracting a global property rather than local features.

are increasingly being created in digital form by instruments connected to high-bandwidth networks and the Internet. Certain distinguishing features of these disparate collections, beyond the usual distinctions of resolution and dynamic range, have implications for image database design. It is of foremost importance to consider whether the image arises from a projection technique such as conventional radiography (Fig. 4) or from a tomographic technique such as magnetic resonance imaging (Fig. 5). This distinction governs whether the image possesses distinguishable, individually bounded anatomic objects (as in the MRI) versus overlapping structures and patterns (as in the radiographic display of the lung markings). Tomographic images grouped by acquisition from individual subjects also have the unique virtue of retaining the data required for unambiguous, three-dimensional reconstruction of tissue structures, thus offering further computational opportunities for novel, shape-matching similarity operations. Patterns and superimposed overlapping structures require a different image database approach than do image collections composed of tomographic slices of organs and cavities. Images whose dominant features are patterns of overlapping structures might lend themselves to computational indexing by global image processing parameters (Figs. 2, 3, and 4). These computational approaches could perhaps concentrate on reducing, say, the image intensity profile into its nonvisible frequency components, such as edge boundary orientation.[19–21] An alternative approach might be ab-

straction by mathematical morphology[22] to density "blobs," which can then be compared and ranked mathematically.

Tomographic images, on the other hand, are more tractable and lend themselves to organizational schemes that take into account the multiple organ boundaries whose configuration and relationships can be mathematically compared.[23,24] For example, tomographic images of the heart seem particularly attractive for the application of topologic tools as a means of indexing image subfeatures. The arguments later in this paper are illustrated with examples from work with a magnetic resonance cardiac image collection.

Content-based Image Database Search Strategies

Computer-aided diagnostic schemes are currently under development in several research institutions to assist the physician and improve diagnostic accuracy by reducing the number of missed diagnoses. Using pattern recognition and feature-extraction techniques, a number of computer-aided diagnosis applications are being developed: computerized detection of pulmonary nodules and mammography microcalcifications[25] automated analysis of heart sizes; automated sizing of stenotic lesions and tracking of vessels in angiographic[26] images; and detection and characterization of interstitial disease.[27] To the degree that mathemat-

Figure 4 Chest x-ray images are projections of many overlapping structures (for example lung tissues). Indexing procedures might be constructed to address textures rather than organ boundaries because of the difficulty of isolating organs without overlap.

ical distinctions might be made between images,[28] implementation of these tools as well as new knowledge-based tools are certain to be developed over the next few years and will need to interface with medical imaging databases.

Similarly, three-dimensional visualization technology has made rapid advances over the past few years so that all manner of display and visualization of human anatomy are now possible. Virtual reality is a term that describes procedures for interacting with these three-dimensional representations in a realistic way. For example, surgical procedures of the future can be envisioned in which the surgeon, using special gloves and a head-mounted monitor, rehearses surgical procedures before the actual event. Both these possibilities require prior segmentation of images as mathematical representations for which an image database structure can be implied.[29]

Given the limited implementation capacity of image collection organization presently available, most current clinical image search behavior is driven by the simple desire to retrieve images from a specific patient or by a named (or coded) diagnosis for which a conventional relational database structure might suffice. These simplistic efforts do not take advantage of exploring image collections on the basis of images that possess "similar appearance" or contain a given structure with a special spatial relationship to another. Perhaps this mode is not familiar to clinicians because of the present lack of graphic and feature-based search
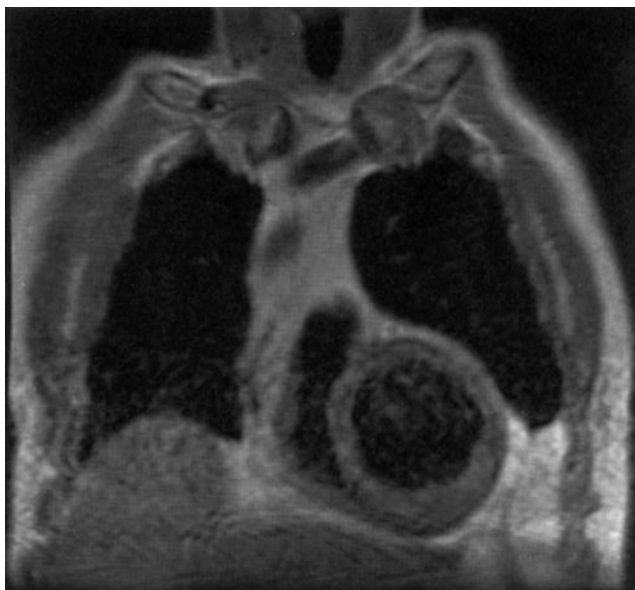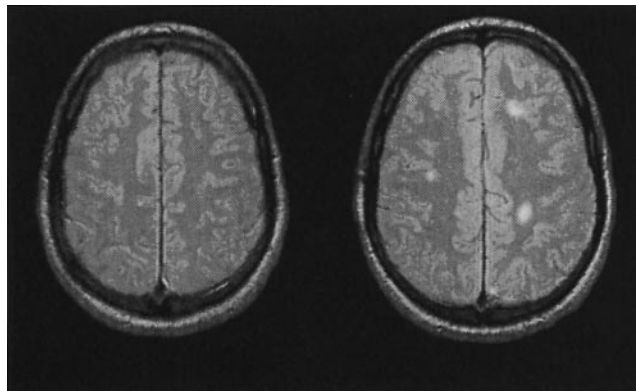

Figure 6 Axial MRI sections of the brain. The left image is normal. The right image shows high-intensity lesions typical of multiple sclerosis. An image database would need to provide an ability to retrieve groups of images whose lesion sizes, shapes or clustering would bear some notion of similarity.

mechanisms. For example, a clinician may observe an image with multiple discrete abnormalities (e.g., an MRI of the brain in patients with multiple sclerosis often has multiple, discrete, high-intensity signals) (Fig. 6). To better understand the disease, a clinician might wish to search a large image collection of MRIs of the brain and retrieve only those images (cases) with a similar viewplane and brain tomographic level that contain abnormalities that appear to be of similar size, number, and location.[3] Once the collection of images is larger than 50 or so, satisfying this task by sequential visual inspection of an image collection becomes very unwieldy and motivates indexing by some mathematic schema.

Other examples illustrate a variety of search strategies on the part of the user. In certain cases, there may be a need for a precise retrieval (the user needs an exact match). This is a common objective when one knows that a collection contains a specific needed image but immediate access to that image is obscured by the size of the collection and a failure to recall the desired image's text tag. Under these circumstances, there is need for a query mechanism that allows the user to create a sketch of the important feature, which can be used for a geometric match.[30,31]

Alternatively, there might be circumstances under which the user might accept a narrow group of fundamentally interchangeable but individually distinct images. Here the search motivation would be satisfied by merely finding a "proximity match." It can be further expected that a mathematically based image database could also allow a much looser match to allow exploring a series of content-similar images with

Figure 5 A coronal MRI tomographic section of the chest and heart. Tomographic images readily permit non-overlapping, geometrically bounded organs and tissues to be identified as a collection of individual features.

greater variations in the content features. Here a much broader subset of retrieved images would be acceptable. This search strategy would characterize a scientific effort motivated by hypothesis checking in which the user desires to explore a theme and variations of images from a more broadly defined concept of structural similarity. An example of such a query might be the user who wishes to retrieve a set of coronal cardiac MRI images that are candidate examples of left ventricular aneurysm.[32,33] The challenges of formalizing that geometrically based conception and creating an effective query are discussed at greater length below. Finally, there is the most open-ended search strategy, best characterized by the term "browsing." Here the database user has a less well-formed idea of which images would be desired for retrieval and is therefore willing to inspect a larger, more wide-ranging retrieved subset that fits relatively loose match criteria.

Relevant Questions

Having created the context within which image databases capable of content-based indexing and retrieval are discussed, there are now a variety of relevant questions that database designers should consider: What constitutes a collection? Should it require some text presorting? Would simply a semantic net as a data model be sufficient? What is the data model, and who defines it? What is a geometric (mathematical) abstract, and need it include all objects in the image? What is an index? What is similarity? How can similarity, defined on the global image, relate to similarity metrics of each component feature within the image? What are relevant metrics of similarity? Are the similarity–distance metrics discrete or continuous? What does it mean to index a shape? What statistics should be gathered? Should indices be precomputed or calculated on the fly? What does it mean to "evolve" a database? Must queries be in some way restricted? How does one specify a query, and how are the results to be displayed? What is implied by the placement sequence in a retrieved image subset? How should images indexed as equivalent (e.g., arising from the same "bin") be displayed? What are the implications of image entry into the database that require domain knowledge (how do you insert domain knowledge expertise into an image database)? How are new images added to the collection? What are database maintenance mechanisms? How do you validate the search retrieval engine? How does the database structure deal with data inconsistencies and conflicts? Is there a role for multimodality registration (that is, images of the same anatomy acquired by different techniques such as magnetic resonance imaging and computed tomography)?

To define what image indexing means, it could be stated that images should contain features that are mathematical and generalizable; the features should be organized for fast retrieval; the search mechanisms must be provably complete (not merely statistical); and it should structure the collection into small, examinable subcollections sharing similarity. In general, indexing can be described as the search for an element of the database on the basis of reduced information. Reduced information in the setting of a text database is (in descending order of reduction): the set of keywords; the abstract of the text; the text itself. The analogy for image indexing should, in this view, be operating on some abstract of the images in the database. In practice, this abstract, if not a text string describing the image, may be a strongly reduced version of the image, such as an icon or a cartoon. What may be concluded from such a discussion is that an index is not the collection itself and that the process of image indexing should not be mistaken for "image understanding."

Research by Others

Particular approaches to image databases have been made by other investigators, some of whom have proposed shape, texture, and geometric descriptors as indexing mechanisms.[7,28,30,34] To date, however, the image database techniques so far developed (e.g., for collections of faces using an averaged "eigenface" template as a model; animal outline forms analyzed as binary images) would not satisfy the complex demands created by medical imaging.[35]

Medical Knowledge

As stated above, medical knowledge is heterogeneous, often imprecise and ill-defined, and particularly difficult to obtain from images in an automatic fashion. Diagnostic images of any complexity seldom lend themselves to observational findings that can be agreed upon by all observers. Intra- and interobserver variation is common. This impediment cannot be conveniently circumvented nor even accounted for in structured text/numeric reporting schemes. There are often, however, significant portions of image content that can be agreed upon. When an "image librarian" entering a new image into a larger collection interactively defines an object (organ or tissue) boundary, these features can act as convenient data entry elements of geometric indexing schemas. Further elaboration on the influence of medical knowledge heterogeneity and imprecision in the design of an image database is discussed more fully in the following paragraphs.

Figure 7 Cardiac ventriculograms illustrating the difficulty of precisely textually defining the term "ventricular aneurysm." Two exemplary angiograms from different individuals are shown on the top. Below these are boundary drawings pairs (systole and diastole) illustrating contraction patterns of other patients who are candidates for being labeled as having one form of aneurysm or another. Separation into meaningful visual subgroups and the threshold for such classification are subject to considerable debate.
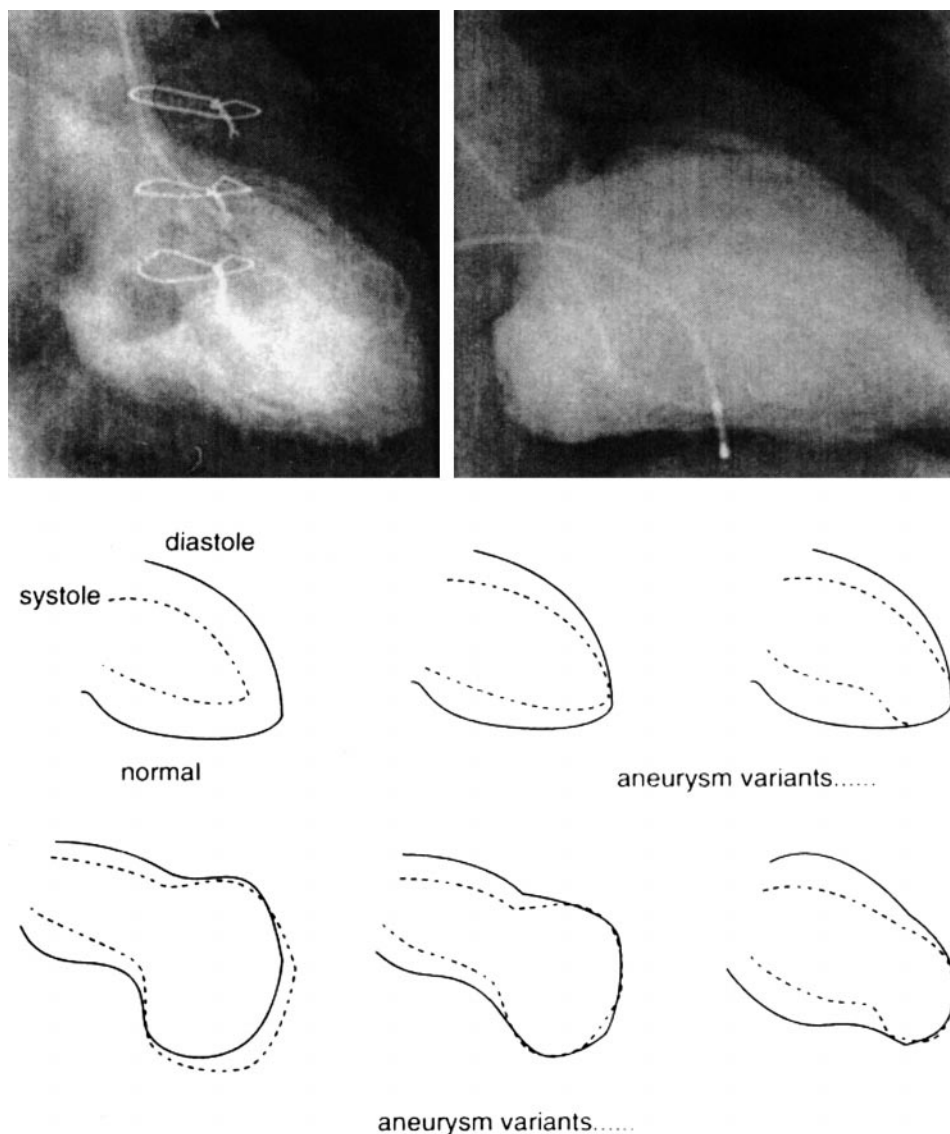
Heterogeneity

Medical concepts of health and disease commonly rest on knowledge of basic biologic or biochemical processes. For example, hypercalcemia may lead to the appearance of dystrophic calcifications on x-rays. In these images, calcific deposits may be responsible for extraneous image densities. It is the disease-concept understanding of the trained clinician that permits distinction of these densities from other normal calcific structures by integrating the visual observation with associated physiologic/microscopic knowledge.

Imprecision

Imprecision has at least three components: semantic imprecision, feature imprecision, and signal impreci-

sion. Semantic imprecision is revealed in medically image-based knowledge by its inability to precisely articulate concepts such as (in the case of cardiology) "left ventricular aneurysm" [32,33] (Fig. 7). Lacking agreement by all observers, nonuniversal unique dictionaries (custom semantics) must be devised for each user.

Items in a conventional text database are commonly considered a fixed asset, prospectively defined at the time of entry (field definitions may be text, number, calculation, etc.). Search mechanisms implying a higher level of abstraction than defined at the time of database inception may be complicated or impossible. Although medical imaging experts usually recognize diverse anatomic features from an image and use
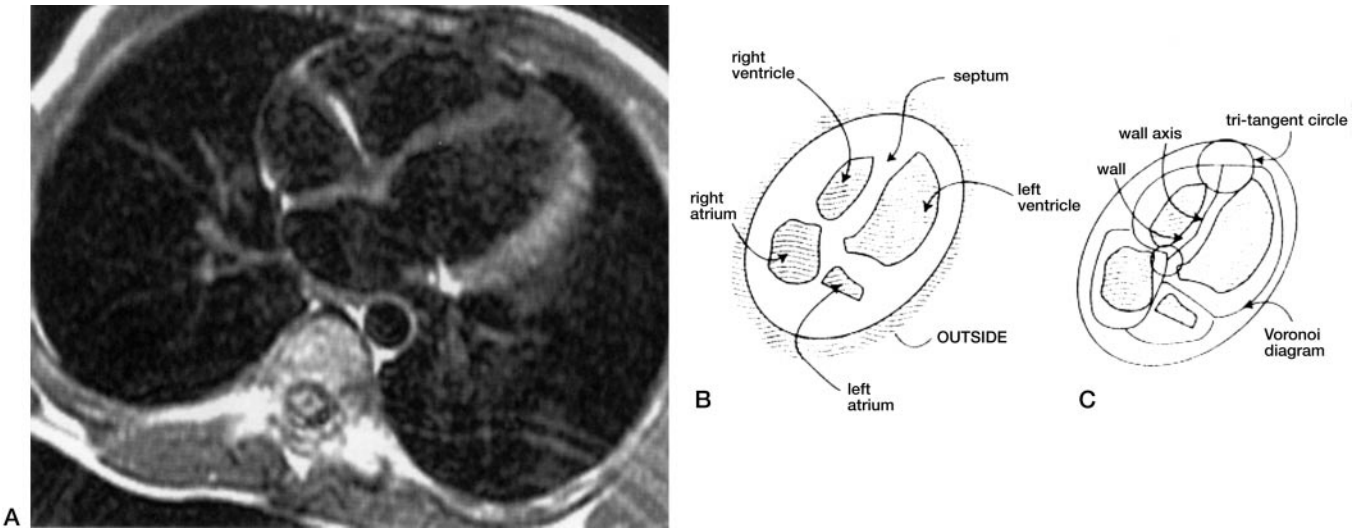
Figure 8 A geometric schema for organizing the arrangement and properties of component features of an image. The left image is a "cardiac four-chamber" MRI section displaying the cardiac chambers. The middle image represents an interactively generated abstract of the main image features. The right image shows the topologic operator (Voronoi diagram) that uniquely creates an indexable mathematical value derived from these segments. This operation takes into account both the explicitly identified objects and implicitly calculated features such as the shape and properties of a wall between features (here, the interventricular septum).

them to infer disease, image features, as well as the categories into which they are placed, are often ill defined. Considerable effort may be required to formalize even seemingly simple clinical terms in an objective, reproducible way so as to convey statistically reliable and anatomically meaningful information.

Many industrial applications of image databases do not share this difficulty. Common industrial objects, particularly those resulting from computer-aided design (CAD) possess ground-truth knowable and modelable configurational geometry.[14] The notion of "object," such as an industrial fastener, incurs little uncertainty. Often the main challenge to industrial part categorization is inferring three-dimensional objects from two-dimensional views.

For an example of complexity in the medical domain, consider the notion of "thickness" of a wall (the myocardium) between two cardiac chambers: There are at least two different ways of defining what it means. It might mean the average separation between neighboring boundaries, or it might mean the maximum value of separation between neighboring boundaries (Fig. 8). Both formalizations of the term are meaningful in certain contexts. The first is meaningful when one seeks a representative measure of a systematic change in the configuration of the wall. The second is meaningful when one seeks a simple measure of the amount of deposit of a substance (e.g., myoma) in the wall.

Another example is the formalization of the term "axis of a bone." In a recent study,[23] two different formalizations were compared: the choice of the "principal axis" of the bone versus simply a line connecting two points on the bone that are farthest apart. Numerical results of the two methods produce substantially different data. Either alternative is reasonable. Surprisingly, as far as reliably representing a radiologist's intuitive notion of "axis," the second choice was in fact preferable.

It is equally hard to obtain precise formalization of semantic categories (such as "large ventricles" or "tortuous aorta") used by cardiac imagers. Although the number of semantic categories for medical images is often small (the American College of Radiology includes only a few hundred terms to categorize images), there is considerable variation in the physician's interpretation of what constitutes these categories. In the absence of objective image processing computation, nominal lists or categorical scores/scales are common approaches (such as small, medium, large; or 1+, 2+, 3+), much of which is poorly tractable statistically. Clinically meaningful image databases would be collections of images too large to be examinable or processed at once. Therefore, retrieval mechanisms should be at least supported by data structures amenable to robust statistical operations.

Not only are the boundaries of observational categories often fuzzy, but there is also variation in what a

prototypical member of the category might be. One well-documented example of this is the term "left ventricular aneurysm"[32,33] (Fig. 7). The term applies to observations about the size, as well as shape of the cardiac left ventricle and is sufficiently vague that there is considerable interobserver variation. Definitional imprecision and observer variation are fundamental facts regarding medical images, and must be considered in the intelligent design of image databases. Recognizing the inevitability of imprecision at item entry implies that proper image database design correspondingly should allow a degree of imprecision in the manner in which a retrieval query is specified. When such a database incorporates geometry, partial matching of iconic or hand-drawn shapes should be permitted.

Implications of Medical Knowledge Imprecision

Recognition of the above considerations imply that, as a developer designs a database and interactively queries it to extract information, his or her intended database schema must have the capacity to evolve as more refined image features are developed. A fixed database structure incapable of dynamic redefinition would freeze indexing methods so that image retrieval could not occur if there were new developments in knowledge of the structure of disease. For example, in the case of the previously mentioned left ventricular aneurysm, geometric image data that fails to distinguish cavity perimeter points of the endocardium from those of the valve plane will lack the flexibility to future index for those separate properties. The extent of database evolution needs to be far greater in medical image databases than in most others, and effective management of database schema evolution should be a primary consideration in design. In our approach, an initial schema, called a "generic schema," is provided to help the user organize the database and roughly express geometrically the image features he or she is interested in.[24] The generic schema is used to interact with the database. It helps formulate hypotheses about possible refinement of the database schema and allows testing these on increasingly larger samples of images by sequentially enlarging the field of view of the database and by using object-oriented queries.

The Generic Schema

Database evolution is a bootstrap process whereby the user customizes the schema for his or her own ends. Thus, the user needs a basic schema to work from that provides an initial framework. A robust solution to the issues involved in the design of a generic schema remains unresolved. There seem to be several approaches depending on the domain of application. Jagadish et al.[36] provide a generic schema called the thin line code that uses curves and curve segments as basic entities. In contrast, we have developed a generic schema that uses point sets in Euclidean space as the basic entity.[24] Thus, points, curves, and regions can be entities in this scheme. Several generic attributes are also being developed for use in this schema. These attributes are topologic, differential geometric, and mathematical morphologic features of the point sets. An example of representing the anatomy of the heart in this schema is shown as follows.[37]

Figure 8A is an MRI of the heart in what is called the "four chamber view." Five point sets are outlined in the image: the left ventricle, the right ventricle, the left atrium, the right atrium, and the "outside." The "outside" region is all of the space that is not part of the heart. An instance of the schema is obtained from the image by finding the point sets (this is the image interpretation process). A number of other point sets are defined in the schema and are computed from the above five point sets. Figure 8C shows the computation of the implicitly defined point set, "wall." The Voronoi diagram (a topologic construct that specifies the relationship of all objects in the plane) of the initial point sets is computed. From this computation, the tritangent circles at the vertices of the diagram are obtained, and the walls are obtained as point sets defined by tangent radii and the boundaries of the original point sets. This definition of "wall" depends only on the Euclidean nature of space and is context independent—therefore, it is generic. The main aim of the generic schema is to exploit the point set structure of image entities without using domain specific knowledge. It is likely that in specific instances the user might wish to substitute a more refined definition of "wall," and will use the schema evolution tools in the database to do so. However, the generic wall is a good starting point from which the user can develop his concept of "wall."

Field of View

As the user seeks possible hypotheses for formalizing image features and tests different formalizations, a powerful means of controlling the complexity is to change the user's field of view of the database. This implies executing test query operations on smaller subsets of the image collection. The initial part of the formalization process is exploratory. The user interacts with the database and forms a series of hypoth-

eses. Rather than retrieve large sets of images from the database, the user navigates through the database, collecting images that are "interesting"—i.e., the ones that suggest possible formalizations. Typical interactions at this stage include extrapolation along a vector in the database ("show me an image with a left ventricle larger than this") and interpolation ("show me an image of a left ventricle that possesses a shape intermediate between these two sample left ventricles"). The field of view is narrow, and the search depends on the user's intuition rather than on precisely formed concepts.

Once a feasible hypothesis about formalization has been formed, the field of view is enlarged, and a larger set of images is obtained on which to try out the formalization. Again, the key here is to control the field of view and enlarge it sequentially as the user iteratively refines his or her formalization. The control of the field of view is obtained through queries that use the generic database schema. For example, suppose a user wishes to formalize the notion of wall thickness and presumes it would be revealed preferentially in one particular tomographic section of the heart. If the generic database schema allows the database to be indexed with respect to sections of the heart, the user can access the set of images in the preferred section and try out the formalization. Once a satisfactory formalization is achieved for these images, the field of view may be further enlarged by including a few other sections and the new formalization can be tried. By iterative mechanisms, the user finally settles on a formalization that is general and reliable enough and incorporates it into the database schema.

Object-oriented Schemas

Object-oriented queries are needed to support the iterative refinement process. These queries contain within them procedures that further process the set of retrieved images. Candidate procedures would be the user's tentative formalizations of image features. This ability is missing in commercially available database query languages, and it appears that objects are the desirable mechanism for creating it. It should be emphasized that the above process of controlling the field of view and using object-oriented queries can be used for other purposes as well. For example, once the database schema has evolved to a point where statistical hypothesis testing can begin, initial hypothesis formulation might be aided by interactively changing the field of view, and procedures that compute statistics may be included in the object oriented queries.

Semantics by Prototypes

As mentioned above, image semantics and categories are also ill defined and can be expected to vary from user to user. The user requires tools to create customized semantics and categories. Research on categorization indicates that mental categories are not defined in terms of necessary and sufficient features, but they are instead defined in terms of closeness to prototypes. A prototype is a member of a category that has the most features in common with other members of the category and is most differentiated from members of other categories. Given an object, its category is determined by measuring its similarity (and dissimilarity) to prototypes or, in the case of medical images, to a visual mental model. This implies the need for developing iconic queries and categorization through iconic association. Iconic queries are queries that use pictorial examples.[30] Thus, instead of asking for a set of images that are examples of "tortuous aorta" or "left ventricular aneurysm," where such terms are ill defined, the user sketches his or her prototypes of tortuous aorta and left ventricular aneurysm (or uses images that contain prototypes) and uses these as examples of what he or she wishes to retrieve. A set of user-defined features (shape, size, etc.) and similarity measures can be used to retrieve the necessary images. Similarly, category formation is achieved through identification of prototypes and by means of measuring the similarity of a given image with the prototypes. For example, Figure 9 shows the formation of the mental category "tortuous aorta." A number of images that contain typical tortuous aorta, and a number of images that contain aortas that are not tortuous, are pooled together in defining the semantic category along with a means of defining similarity with these images. (A comparison of curvatures of the aorta is a possible similarity measure.) The use of icons and associations with prototypes appears to be a consistent way of providing the user with customized semantics.

Recapitulation

To recapitulate, it is clear that medical images represent a particularly unique class of problem for the design of databases.

■ The database schema can (and must) evolve considerably over the lifetime of the database. The changeability of the schema seems to be the single most important aspect of medical image databases, and much design effort must be focused on management of this change.
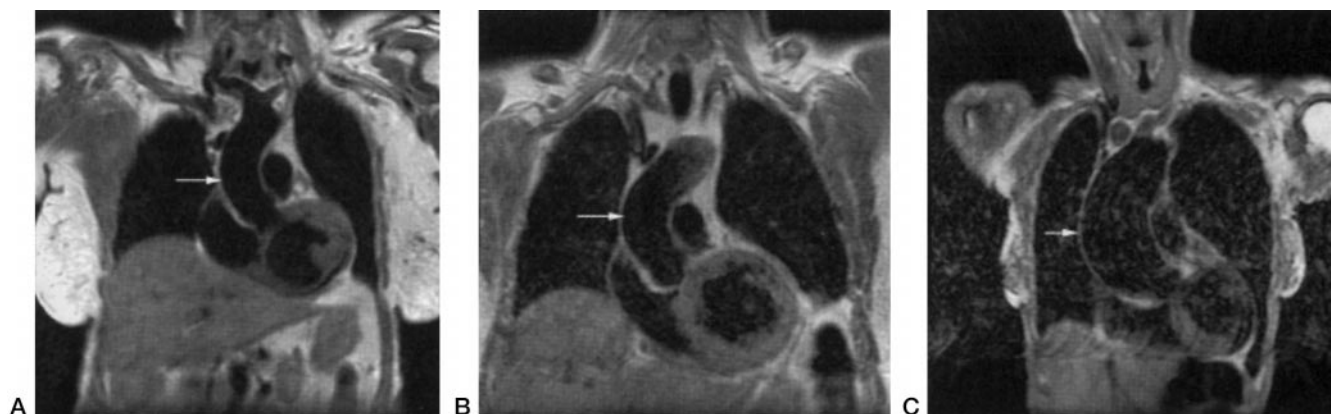
Figure 9 Coronal MRIs of three different individuals exemplifying variations in the shape and configuration of the ascending aorta (arrow). The image on the left is normal while the middle image might be textually described as "tortuous" and the right image might be variably called "ectatic" or "aneurysmal." In a large collection, shape based on geometry might serve to better index the collection.

- Object oriented queries and generic schemas to control the field of view provide mechanisms to manage the evolution of the database schema.

- Medical image interpretation is a complex and poorly understood process. There is a need to decouple the database activity from the interpretation activity, and the database schema appears to be one mechanism where this can be achieved. This implies that entry of an image's feature geometry should be as objective as possible and not be influenced by knowledge bias arising from a gestalt-driven diagnostic interpretive process.

- Where image entry into collections (particularly where entry points are distributed over a network) is conducted by different catalogers, objectivity in the methods of feature selection entry is critical for predictable retrieval.

A medical imaging database requires tools not required by traditional textual databases. These include:

- **Non-textual indexing.** There must be a means for nontextual indexing in addition to textual indexing, with links between the two kinds of data. For example, geometric information can be obtained by analyzing the outlines of organs and tumors. Physiologic information, however, will come from such sources as laboratory results and other parts of the patient record. Links between the two classes will be important to understanding the clinical implications of a particular therapeutic or diagnostic decision.

- **Customized schema.** The clinical researcher will require tools that allow for end-user designed ad hoc customized schema for retrieval and search that can be edited, modified and adapted to new queries. Queries to an image database by different users may present vastly different demands on the query language. For example, a medical oncologist may want to generate complex queries about an image that relate to the functionality and/or structure of organs in the image. A database to be used for teaching, on the other hand, may require a means of accessing images that all exhibit a particular morphological characteristic.

- **Dynamic.** Users must be able to generate queries of a set of medical images that are changing and dynamic. A patient who undergoes a CT scan to delineate a primary tumor of the lung may subsequently be discovered to have liver metastases. Thus, the physician accessing the database will want to incorporate new knowledge about liver metastases into the database and have the capability to develop relevant questions about the patient's condition, both in the past and the present, that pertain to the new information.

- **Similarity modules.** As the user develops possible hypotheses for exploring a database, having the ability to navigate through the database collecting images that are "interesting" will suggest new formalizations. Thus, tools that allow for "show me one like this that is larger" or "show me one between these two" may provide the user with powerful means of developing new conceptualizations and knowledge. This "changing the field of view"

approach is perceived as an important attribute of a medical imaging database.

- **Comparison modules.** Another query structure would show the user ten "normal" examples of a feature for purposes of comparing them with one under investigation to determine whether it falls within normal limits. Special tools that allow for flexible shape matching will be needed.

- **Iconic queries.** Iconic queries are queries that use examples that are user generated. These may be sketches of features that are important or they may be prototypes. The use of icons and associations with prototypes will provide the user with a means of developing customized semantics. (e.g., "This is what I choose to define as a left ventricular aneurysm.") Generic schema will be needed to develop a starting point in the schema evolution so that a user can define which relations and which similarity measures are appropriate for the problem under consideration. A variety of objects at different levels of abstraction will be required by users to support the iconic queries and the customized schema. A pictorial-based query language will be essential to the full utilization of a medical imaging database.

- **Descriptive language.** New descriptors of image features will lead to new knowledge and new categories for staging of disease. For example, a large neuroblastoma extending across the midline and encasing the celiac axis is inoperable; while a neuroblastoma that extends across midline but does not encase the celiac axis is operable. Both are pathologically called "Stage III." New tools need to be developed that allow for the evolution of new definitions that distinguish between such two possibilities.

- **Multi-modality registration.** Accurate registration of images from different imaging modalities adds new knowledge to the decision-making process,[29] but tools for the rapid implementation of this procedure are still lacking.

- **Image manipulation.** As intrinsic operating environments, imaging databases need to incorporate many of the already existing tools used for manipulating images: zoom, pan, rotation, contrast enhancement, region-of-interest contours; pattern recognition tools, such as edge detection, similarity retrieval; three-dimensional display features, complete with surface rendering and texture discrimination; movie loops that display multiple images, possibly from several different studies, in rapid sequence on the same screen; automatic segmentation of features of interest; ability to electronically "mark" on the images as is done on film; and customized user-defined functions.

Validation

Meaningful proof of adequacy of implementation of a medical image database should incorporate a rational test by which operation of the instrument can be judged to be successful. In the case of commercial image databases devised for browsing where poorly formulated queries are expected and incomplete recovery of image subsets is not problematic, validation might be impossible and unnecessary. Medical image databases, however, impose more stringent justification criteria and cannot be satisfied by merely acclamation. Comparison of performance in the context of a realistic clinical goal is mandatory and the search mechanism must be provably complete (not merely statistical). The database's response to a well-formed query should be compared with the trained clinician's capacity to find a similar set of images by exhaustive search. Since human judgment is involved, duplicate trials with several observers is necessary to compensate for expected observer variation. Appropriate statistical tests such as the *kappa* measure of agreement should be considered, although alternative statistical methods may be appropriate. Although it is intended that image databases are designed to make accessible very large image collections, testing procedures validated by humans conducting exhaustive search must necessarily be limited to reasonable but statistically valid size collections.

Conclusion

Aside from considerations that apply to an exact match, the critical consideration in image retrieval by similarity stems from the answer to the question, What makes two images similar? Similarity from a medical perspective is predominantly context dependent. Design of medical image databases imposes requirements that differ from those of other domains. The complexity and context-relatedness of medical image content should dismiss false hopes that image indexing can occur fully automatically or that there exists some universal primitive.

Image abstracts, by nature, are simplifications of complexity. By defining abstractions for images, and distance metrics which allow the comparison of abstractions, the computational burden can be greatly reduced. Rapid comparisons can be done on abstract representations because they require fewer calculations. Alternatively, under conditions of great computational speed such as can occur with parallel pro-

cessing, queries constructed as custom processed, operator-directed image processing may proceed to examine each image in the database sequentially at query run-time without the benefit of a precomputed abstract for indexing; nevertheless, these provide an effective search mechanism.

It seems evident that cataloging of individual images at the time of entry has numerous advantages, except for the time demand of the entry process. Manual methods can benefit from computer assistance in these tasks such as image processing operations like boundary-finding and region growing. Adding images to a collection, much like the acquisitions process of a conventional library, requires effort. Geometric feature abstracts and their implications for user queries must be designed together. In this process, one must define a useful set of specifiers and design a graphic user interface to set up specifications. Once this is done, there must be defined a set of resemblance (similarity) functions with tolerance functionals.

In the development of future proposed medical image databases, the following issues are important:

- The management of evolution of the database schema over its lifetime, in particular the use of the database itself as a means of refining the schema.

- The development of generic schemas that may be used as a starting point in the schema evolution.

- The development of data models that permit the existence of entities in the database, and the computation of relations and attributes by user defined procedures.

- The need to address database models that incorporate the following: index an imaging database using image features; support spatial relations for queries that can detect change, such as by shape and size, but are robust enough to adjust for deformations; develop object-oriented solutions that can handle levels of uncertainty in identifying objects with fuzzy boundaries.

- It is necessary to develop queries that permit icons, pictorial examples, and procedures.

- There must be shielding of the database activities from image interpretation, so that the image feature sets are constructed as objectively as possible.

The goal of medical imaging databases is to provide a means for organizing large collections of heterogeneous, changing, pictorial, and symbolic data. This must reside in a structured environment that can be synthesized, classified, and presented in an organized and efficient manner to facilitate optimal decision making in a health care environment. Approaches taken to the design of future image databases will likely differ and result in applications that take into account technologically related domains (e.g., tomographic imaging, such as CT and ultrasound) or are customized by connection with particular diagnostic knowledge domains (e.g., organ configurations). In all, a properly organized imaging database can compensate for obvious human visual memory limitations and provide a basis for improved patient care, research, and education.

References ■

1. Bohern BF, Hanley EN Jr., Extracting Knowledge from Large Medical Databases: an Automated Approach. Computers and Biomedical Research, 1995. 28:191–210.
2. Altschul SF, Miller W et al., Basic local alignment search tool. J Mol Biol. 1990. 215:403–10.
3. Zink S, Jaffe C. Medical Imaging Databases: an NIH Workshop. Investig Radiol. 1993. 28:366–72.
4. Dayhoff RE, Kuymak PM, Shepard B., Integrating medical images into hospital information systems. J Digital Imaging. 1991. 4:87–93.
5. Huang HK. Picture archiving and communications systems. Comput Med Imaging Graph. 15, 1991.
6. Kim Y, Haynor DR. Requirements for PACS workstations. in 2nd Int. Conference on Image Management and Communication in Patient Care. 1991. Kyoto: IEEE Computer Society Press.
7. Chang SK, Hsu A. Smart Image Design for Large Image Databases. J. Visual Languages and Computing. 3, 1992.
8. Grosky W. Toward a data model for integrated pictorial databases. Computer Vision, Graphics, and Image Processing, 1984. 25:371–82.
9. Grosky WI. Image database management. IEEE Computer, 1989. 22:7–8.
10. Iyengar SS. Guest editors' introduction: Special section on image databases. IEEE Transactions on Software Engineering, 1988. 14(5):608–9.
11. Flickner M. Query by Image and Video Content: the QBIC System. IEEE Computer, 1995. 28(9).
12. Niblack W. Query by Image and Video Content: the QBIC System. IEEE Computer, 1995. 28(9):23–32.
13. Califano A. Multidimensional Indexing for Recognizing Visual Shapes. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1994. 16(4):373–92.
14. Mehrotra R, Grosky WI. Industrial part recognition using a component-index. Image and Vision Computing, 1990. 8(3): 225–232.
15. Ogle VE, Chabot SM. Retrieval from a Relational Database of Images. IEEE Computer, 1995. 28(9).
16. Tamura H, A Management System for an Integrated Database of Pictures and Alphanumeric Data. Computer Graphics and Image Processing, 1981. 16:270–86.
17. Liu J. Multiresolution Color Image Segmentation. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1994. 16(7):689–700.
18. Swain MJ. Color Indexing. Int. J. of Computer Vision, 1991. 7(1).
19. Chang SK. Picture Indexing and Abstraction Techniques for Pictorial Databases. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1984. 6(4).

20. Gevers T. Enigma; An Image Retrieval System. in Proc. 11th IAPR. 1992. The Hague, Netherlands: IEEE Computer Press.

21. Gevers T. Indexing of Images by Pictorial Information. in Proc. of Visual Database Systems II. Budapest, Hungary: Elsevier, 1992.

22. Serra, J. Image Analysis and Mathematical Morphology. London: Academic Press, 1982.

23. Tagare HD, Stoner DM, Viegas SF, Hillman GR. Location and Geometric Description of Carpal Bones in CT Images. Annals of Biomedical Engineering, 1993. 21:715–26.

24. Tagare HD, Jaffe C, Duncan JS. Arrangement: A Spatial Relation Between Parts for Evaluating Similarity of Tomographic Section. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1995. 17(9):880.

25. Chan HP, Vyborny C, Schmidt RA, et al. Improvements in radiologists detection of clustered microcalcifications on mammograms. Invest. Radiol. 1990. 25:1102–10.

26. Pinciroli F, Pozzi G. ARCADIA: a System for the Integration of Angiographic Data and Images by Object-Oriented DBMS. Computers and Biomedical Research, 1995. 28:5–23.

27. MacMahon H, Chan H-P, et al. Computer-aided diagnosis in chest radiology. J. Thoracic Imag 1990. 5:67–74.

28. Grosky WI. Iconic Indexing Using Generalized Pattern Matching Techniques. Computer Vision, Graphics, and Image Processing, 1986. 35:383–403.

29. Pelezzari A, Halpern H, Chien CT, Cooper MD. Three dimensional correlation of PET, CT and MRI images. J Nuc Med. 1987. 28:682–89.

30. Joseph T. A high level query language for pictorial database management. IEEE Trans. on Software Engineering, 1988. 14:630–8.

31. Cardenas A, Tiara RK, Barker R, Breant CM. The Knowledge-Based Object-Oriented PICQUERY Language. IEEE Transactions on Knowledge and Data Engineering, 1993. 5: 644–57.

32. Meizlish JL, Plankey M, Erico D, Levy W, Zaret B. Functional Left Ventricular Aneurism Formation After Acute Anterior Transmural Myocardial Infarction. N Engl J Med. 1984. 311:1001–6.

33. Hamer HH. Redefining true left ventricular aneurysm. Am J Card. 1989. 64:1192.

34. Saund, E. Distributed Symbolic Representation of Visual Shape. Neural Computation. 1990. 2:138–51.

35. Moghaddam B, Pentland A. Bayesian Face Recognition using Deformable Intensity Surfaces, in IEEE Conference on Computer Vision & Pattern Recognition. San Francisco, CA: IEEE Computer Press, 1996.

36. Jagadish HV. An object model for image reconstruction. IEEE Computer, 1989. 22:33–42.

37. Tagare HD, Duncan JS, Jaffe CC. A Geometric Indexing scheme for an Image Library. In CAR '91, Computer Assisted Radiology. Berlin: Springer–Verlag, 1991.