1995, *The British Journal of Radiology*, 68, 945–957

Invited review: Computer aids for decision-making in diagnostic radiology—a literature review

P TAYLOR, MSc

Advanced Computation Laboratory, Imperial Cancer Research Fund, 61 Lincoln's Inn Fields, London WC2A 3PX, UK

Abstract
This review looks at a variety of different ways in which computers can be used to assist in the interpretation of radiological images and in radiological decision-making. The issues involved in the design of computerized decision aids are introduced and four criteria proposed for evaluating such aids: need, practicality, veracity and relevance. These criteria are used to assess research into decision aids based on: image databases, numerical methods, expert systems, image processing and image understanding systems. Possible directions for research leading to aids of practical value are discussed in the conclusion.

Computer systems are now used to capture, store, transmit and display radiological images. The use of information technology to provide the framework within which radiologists work opens up the possibility of "computer aided radiology" and, in particular, the prospect of decision support systems for image interpretation. This review looks at a number of areas of computer science within which research has been conducted into decision aids for radiological image interpretation. Since the focus of this review is decision support, attention is restricted to systems which would come into play when or after the image is displayed, to be used by a radiologist seeking assistance. The various ways in which computers can be used to create new kinds of image are not covered: the creation of tomographic images, the reconstruction of three-dimensional (3D) images, the enhancement of digital radiographs, the segmentation of magnetic resonance (MR) images and the registration of images of different modalities. Electronic information sources such as hypertext systems and medical databases are also excluded since, although they could be used to assist in decision making, they are not specifically designed to do so.

Wyatt and Spiegelhalter [1] define medical decision aids as "active knowledge systems which use two or more items of patient data to generate case-specific advice". Such systems include a computerized knowledge source, from which the advice is culled, and a mechanism by which a user may quickly obtain relevant information from this source on the basis of patient data. A computerized knowledge source may take many forms: images, algorithms for manipulating images, medical data, derived probabilities and symbolic representations of medical facts. Different computations and methods of interaction are appropriate to each of these forms of information. Five distinct classes of decision support system are currently the subject of research, they are discussed in turn in the next section. First, some of the issues common to all types of decision support systems are outlined.

The first is the need for a decision aid. In diagnostic radiology, a decision aid is required when interpretation demands specialist expertise, when many images are generated or when interpretation is especially difficult. Interpretation may be problematic because the image is noisy or visually complicated; its relation to imaged anatomy is complex; the potential lesions are subtle; the clinical significance of a lesion depends on a variety of factors or the variety of clinical conditions is large.

A second issue is encountered once a need has been identified: the medical domain to be covered places constraints on the kind of system that will be successful. For example, if a class of images is perceptually difficult, a database of images with known diagnoses may be more useful than a system providing detailed information about possible diagnoses. In addition to restrictions intrinsic to the domain, there may be others which stem from the setting in which the system is to be used: if a large number of images are being interpreted extremely quickly, as in screening situations, a useful system will require only minimal input from the user.

Once the clinical need has been established and the constraints imposed by the medical domain and clinical situation have been evaluated, the problems of designing and building an effective decision aid must be solved: an information source must be created and a mechanism for displaying the information implemented. The information source must be accurate, it must be complete (within limits which are understood by the user) and it must be well constructed, allowing efficient retrieval.

The next issue concerns the mechanism for retrieval from the information source. One of the key issues here is the extent to which the designer has understood how

Received 27 January 1995 and in revised form 18 April 1995, accepted 22 May 1995.

the user will make use of the information in making a decision. Increasingly, decision support systems are being designed as collaborative systems which complement the skills of the decision maker.

The above discussion can be summarized by listing four general criteria necessary for a successful decision support system:

- **NEED**: that there be a clear clinical need for a system.
- **PRACTICALITY**: that the constraints imposed on a system by the medical domain and the clinical setting be understood and the approach taken work within them.
- **VERACITY**: that the information on which the decision support is based be accurate and complete within understood limits.
- **RELEVANCE**: that the system can provide the user with information which is known to improve his or her decision-making.

In the next section five classes of decision support system are considered in turn and the extent to which research in these areas meets the criteria is assessed.

Computers and medical image interpretation

In recent years an increasing number of articles has appeared on computer aids for radiologists, with origins in various strands of research in computer science: databases, numerical methods, expert systems, image processing and image understanding. Each of the following sub-sections gives a brief introduction to these topics and the research in them which is relevant to diagnostic radiology; a representative paper from each section is discussed in detail.

Image databases

Research into image databases began in the late 1970s when work in image interpretation created a need for systems able to store and retrieve large numbers of images. More recently the impetus has come from the database community, where researchers have been looking at image databases in the context of work on multimedia databases, hypermedia systems and picture archiving and communication systems (PACS) for storing large numbers of clinical images [2]. The development of this technology has led to the possibility of using image databases in decision support: if unsure of the significance of an image feature, a radiologist could inspect database images with similar features but with known pathology. This requires the storage of large numbers of images and the capacity to retrieve them quickly on the basis of their visual content.

A great deal of research in image databases in non-medical domains involves attempts to develop "visual query languages" for this purpose, such as that of Chang et al [3] which allows "queries" to be built up from icons, a difficult technique to apply in radiology. One group [4], has developed a system for a database of MR images, in which the user indicates a section of interest in the viewed image to be used as the query in the retrieval of similar images. The designer of such a system faces two problems: first, the process of matching the query with the images must be flexible enough to allow for normal variation between images of similar anatomy, second, the process must be fast enough to be useful. In this system both query and database image are reduced to binary (black and white with no shades of grey) images showing the outline of anatomical features and the binary query is matched with every point in every binary image. In tests the system was able to retrieve similar images (those containing the same view) but the process took 40 s per database image. This is clearly too slow to be used interactively.

Wiederhold et al [5] advocate computing sets of parameters from images, to provide indices for image databases. They implemented a system for computing fetal volumes from ultrasound scans, but no evaluation of an image database indexed in this way has been reported. A more flexible system for the retrieval of images on the basis of content is described by Kofakis and Orphanoudakis [6] who implemented a system in which both an automatic and a user-guided image analysis system are used to derive representations of image content for images in the database. Rules embodying background knowledge are used to check the consistency of this representation. The retrieval is based on a two-phase process in which inappropriate candidates are filtered out on the basis of gross characteristics of image segments and then detailed matches are sought for the midpoints and other characteristic features of polygonal approximations to the image segments making up the query. Experiments with a database of 400 MR and computed tomography (CT) images have demonstrated the efficiency of the system.

The most complete account of an image database designed to provide decision support is Cohn et al's [7] description of AXON. In AXON each database image (chest radiographs) is stored with a set of keywords labelling the lesions and the disease. Associated with the database are "frame" hierarchies representing taxonomies of the keywords. A frame is a way of representing information about classes. Each frame represents the properties defining a class and provides a link to the immediate generalization of the class. This link allows properties to be inherited from more general frames. For example, the disease hierarchy includes a frame for *tuberculosis* which inherits properties from *mycobacterial infection*, which in turn inherits properties from *infectious diseases*. The lowest level in the diseases hierarchy represents the cases known to the system. At this level there are links between the frames of different hierarchies, which represent the occurrences of lesions and connect the images with the appropriate cases.

The simplest way of interacting with the system is to enter a keyword and a condition. The system will match the keyword with its frame and then search down the hierarchy retrieving images matching the condition. For example, all images of neoplastic disease showing pulmonary lymphadenopathy could be retrieved by indicating that the system should search for cases in frames underneath the *neoplastic disease* frame and check for connections to the *pulmonary lymphadenopathy* frame before retrieving the images for the case.

This procedure has been made more sophisticated by using four "axes of clinical relevance": radiographic findings, underlying aetiology, clinical findings and imaging modality. The stored images are grouped according to these four axes. Figure 1 shows eight images retrieved from a search for pulmonary infiltrates in Hodgkin's disease and related disease states. These have been ranked along the aetiology axis so the first cases where the infiltrates were caused by Hodgkin's disease come first. In the later cases the infiltrates were caused by an infection commonly found in Hodgkin's disease.

In addition, domain knowledge has been captured in the form of search heuristics which have an application condition and a distance measure (an indication of similarity along an axis). One heuristic states that if, when searching along the aetiology axis, fewer than five cases are found, the immediate generalization of the current aetiology should be used. When a user enters a query, the system first carries out a keyword search and gathers together all the heuristics whose application condition has been met. The heuristics are then applied in order of distance measure and the retrieved images are presented in order of merit.

AXON meets some of the four criteria outlined in the introduction:

- **NEED**: it has been developed for an area in which there is a clear clinical need; studies show that error rates as high as 30% are found in the interpretation of chest radiographs [8].

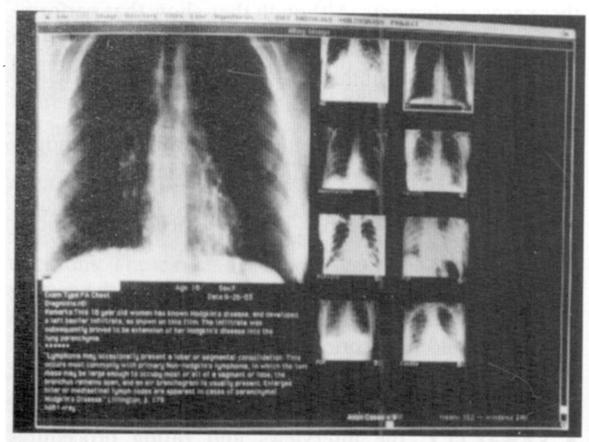

Figure 1. The first eight cases retrieved from an image database by a search for "pulmonary infiltrates in Hodgkin's Disease" using Cohn et al's AXON [7]. The number next to each image indicates the order in which it was retrieved. The large picture on the left is an expanded picture of case number two (each of the eight cases may be enlarged).

- **PRACTICALITY**: the system is provided as part of an image display system in which both AXON and other kinds of information source are available on request.
- **VERACITY**: the prototype contains only 60 cases and no mention is made of how an adequate coverage of image features and disease processes could be ensured.
- **RELEVANCE**: if the retrieval technique is successful the system would provide the user with information he or she could use in decision making.

The method for assisting in content-based retrieval seems appropriate for a decision support system, in that it is based on an understanding of clinicians' ideas of relevance—although the paper does not explain where this understanding comes from. It is, however, not clear that the simple four-dimensional classification will be adequate to partition a search space much larger than 60 images, or that understanding what clinicians consider to be relevant could inform a sufficiently rich classification. The value of the heuristics is unclear. They are described as embodying knowledge about the domain, the domain in this case is as much "how to search a large database of medical images", as it is "medical images" and the heuristics given don't suggest that there is much to be said about this domain.

The focus in this section has been on the technical problems which must be solved before image databases can be used as decision support tools. Most of the work in this area is concerned with the problem of retrieving a relevant image and other problems have received much less attention: for example, ensuring that an image database provides adequate coverage of a medical domain.

Atlases of radiological images exist as conventional textbooks, *e.g.* Tabar and Dean [9], and are used in teaching but these may not provide an adequate analogy for a decision support tool. Textbooks also exist which provide lists of all the possible diagnoses associated with each radiological finding for every class of images [10], but no attempt has been made to wed this to an image database. The analogy between textbooks and image databases also supports the idea that image databases are an appropriate way of providing decision support, but it isn't known how often clinicians consult such resources, nor in what circumstances nor to what effect. Research into these questions could help inform the design of future image databases.

Decision support systems based on numerical methods

The earliest group of computer aids to medical decision-making were based on numerical methods. Ledley and Lusted [11] characterized diagnosis as a logical activity of reasoning from symptoms to causes on the basis of "probabilistic" or "uncertain" information. They advocated the use of Bayes' rule, a mathematical equation describing how probabilities should be combined, to improve decision-making. This initiative led to the development of computerized decision aids which use Bayes' rule to combine, for example, information about the frequency of a disease and the frequency with which signs are associated with that disease, to compute the probability that a patient showing these signs has the disease. A review of eight such systems [12], developed between 1971 and 1979 as aids to image interpretation, concluded that where the scope and range of the problem is narrowly defined, computer aids can perform better than humans. The optimistic tone of the review is at odds with the impression that this was not a flourishing area of research. Some authors [13] have argued that such systems can be fitted into a collaborative model of diagnostic problem solving since clinicians are much better than computers at identifying the key symptoms and reducing the number of diagnoses to a manageable subset, while computers are much better at performing the numerical calculations which identify the most likely option in that subset. However, Good [14] gives an account of an experiment in which half the radiologists ignored the output of a numerical decision support system when rating mammograms for malignancy.

One of the most successful attempts to use numerical methods in a decision aid for radiology is that of D'Orsi et al [15] which uses data derived from radiologists' judgments of mammograms. The authors interviewed five expert readers to establish a set of 50 features indicative of malignancy. These experts were also asked to rate 24 images for similarity in an exhaustive series of pairwise comparisons. A computerized statistical technique (multidimensional scaling analysis) was used to derive a set of abstract dimensions that would explain the rating data. The experts then matched these statistically derived dimensions to real world features. The original and derived sets of features were then combined by the radiologists at a meeting where they agreed on a set of 23 important features. 150 images were then rated for the presence of each feature by the same experts. This information was used to identify the 12 most significant features and their predictive values. These predictive values formed the basis for the decision aid. The decision aid consisted of two components: a set of rating scales for radiologists to indicate their degree of certainty about each feature's presence and a computer program which combined the 12 ratings with their established predictive value to produce an estimate of malignancy. The system was tested by six other radiologists, all of whom were experienced but not specialists in mammography. They were asked to rate the probability of each mammogram being malignant and the pooled data was used to construct receiver operating characteristic (ROC) curves, which showed improvement with the use of the decision aid. At a chosen threshold the decline in false negatives was found to be significant. ROC curves drawn up from data using only the 51 most difficult cases (difficulty being measured by the diversity of readers' judgments) showed a clearer improvement, both in false positive and false negative rates, as shown in Figure 2.

This methodology was used to develop a system for assisting radiologists in combining evidence from different imaging modalities [16]. Feature lists and computerized decision aids were developed for breast X-rays and breast diaphanography. Tests of the decision aids showed no great improvement in the case of mammography, but considerable improvement in the case of diaphanography and significant improvement when used as an aid to decisions based on inspection of the two images. The authors claim that their tool allows users to merge information from the two tests at the "feature level" and that radiologists inspecting images from different modalities otherwise make separate assessments and only combine these gross judgments. Although they show that performance on the combined result is better with the system than without it, they do not show that that combined judgment is better than that based on the more effective of the two tests.

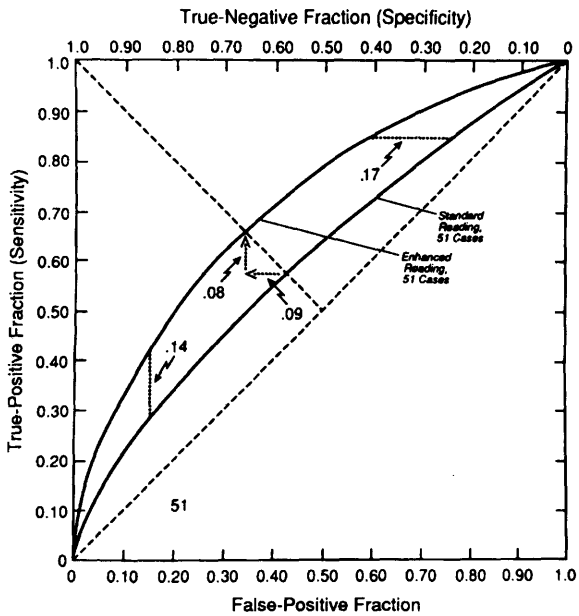

Figure 2. ROC curves for the interpretation of 51 difficult mammography cases with (enhanced) and without (standard) D'Orsi et al's [15] numerical decision support aid. The dashed diagonal line from bottom left to top right indicates chance performance, the dotted lines and arrows indicate differences in specificity and sensitivity.

The approach seems to fare quite well when judged on the four previously mentioned criteria.

- **NEED**: both systems are designed for difficult problems.
- **PRACTICALITY**: nothing is said about the setting in which these aids might be used or how an interface to them would function. It seems likely that rating each image for 12 properties would be considered a heavy additional burden by many radiologists.
- **VERACITY**: the interviews and rating procedures ought to ensure that the information used is sufficient and correct.
- **RELEVANCE**: the ROC tests, although on a small sample, demonstrate that some improvement is obtained through using these systems.

The methodology employed by D'Orsi et al [15] in the creation of their information source seems rigorous, but the basis of it is still radiologists' judgments and not, for example, statistical data about what was actually found on biopsy for mammograms rated in different ways. Such data are becoming easier to obtain; for example, Kuhns et al [17] have compiled tables of incidence and mortality rates for radiologically diagnosed diseases, together with information about the risks and successes of treatments and the reliability of diagnostic tests. Their aim is to allow a rough calculation to be made, first of how certain one should be of a diagnosis before ordering treatment and second, of the impact which a positive test result would have on that certainty.

It should be noted that numerical aids have been around a long time, in areas other than image interpretation, and have not been adopted with enthusiasm by the medical community. There are many possible reasons for this (see [18] for a discussion), some of them to do with social and political problems, others with the general difficulties of technology transfer. De Dombal et al [19], discuss Bayesian decision systems and argue that one of the principal problems is ensuring that the physician is able accurately to assess the significance of the computer's decision. They advocate presenting overview data for what happened to patients with similar assessments, rather than a diagnostic prediction. There are advantages to using computers in decision-making: computers are less variable than people and can deal with a much greater volume of information. However, where the decision-making process must remain under human control, computers must provide appropriate input. It would have been interesting to know to what extent the improved performance found by D'Orsi et al was due to the use of the checklist of 12 significant features and to what extent it was the calculation of probability that improved performance.

Expert systems

Another class of decision support systems is the expert system. In expert systems the information source is a representation of clinical knowledge. This contrasts with the numerical decision support systems described above in which the information source is made up of statistical data derived from past cases or radiologists' judgments. Developers of expert systems compile sets of rules which represent the knowledge used by clinicians in making decisions. These rules are used, together with information supplied by the user, to generate the inferences which form the basis of the decision support.

There are two common control strategies used to guide inference: forward and backward chaining. In forward chaining, inference is driven by the case data and rules are "fired" if the conditions match the data. Firing a rule adds the conclusion to the data, which may enable more rules to fire. In backward chaining the inference process is an attempt to prove a conclusion by searching for rules with the required conclusion and inspecting their conditions to see if they match either the data or the conclusions of other rules whose conditions may be inspected. Backward chaining is a much more tightly controlled process and is useful in cases where a small set of possible inferences are known in advance.

Other terms commonly found in descriptions of expert systems refer to methods of organizing the representation of domain knowledge. Expert system developers commonly make use of hierarchies of "frames" to describe classes of objects. As described earlier, a frame is a set of propositions which defines a class. A refinement of the frame idea is that of the "object". Objects have the additional properties of allowing the developer to include rules and parts of programs in the object definition. An "object oriented" approach allows a computer system to be considered as a network of communicating objects.

Expert systems have been developed for many areas since the 1960s, coming into more widespread use in the 1980s. A number of medical domains have been tackled with apparent success but, as far as the author is aware, no systems have entered widespread clinical use. Expert systems have also been developed as decision aids for diagnostic radiology; of these a handful have been developed to provide radiologists with assistance in the application of computerized image processing techniques [20–23]. The weakness of these systems is the absence of formalized knowledge on the applicability of image processing techniques. Many more systems have been developed to provide diagnostic advice on the basis of information entered by the radiologist about what he or she sees on the image. Published accounts of such systems fall into three categories:

(1) Small scale systems.

(2) Expert system components of larger systems.

(3) Novel developments of expert system ideas.

Small scale systems are generally developed using a commercial expert system shell which provides the interface and inference mechanism. The system developer then works with a radiologist to construct a set of rules which provides the knowledge base or information source. Cook and Fox [24] developed such a system to distinguish between benign and malignant mammographic anomalies. The system contained a number of rules (the number is not given) distinguishing 16 common mammographic manifestations. The system did not improve the diagnostic accuracy of radiologists but residents and even students with no training in radiology were able to perform as well as experts when using the system. The problem with such systems is that they are of little value to a specialist and yet too specialized to merit the attention of a generalist. They may be of some value to students of a speciality but suffer through being designed as decision rather than teaching aids.

One of the problems with these systems is that the users must enter accounts of the lesions they see on the image in order to drive the process of generating advice. This may involve a lengthy interaction and the introduction of errors where the terminology used by the designers is not understood by the users (see [25] for a discussion of this problem). Some authors [26–28] have published accounts of expert systems in which it is intended that the input should ultimately come from an image processing system able to detect lesions automatically and to generate the needed description. However, none of these systems has actually been completed. Developing the required image processing is no easy task and it may be that a cautious estimate of the kind of results image processing can provide should precede the construction of the expert system.

The most interesting class is the third, where researchers have attempted to extend the conventional expert system technology to tackle the problems which beset the less ambitious systems. One such system tackles the problem of diagnosis of lung disease from chest radiographs [29]. The knowledge base contains rules and frames and also a text generator. The user supplies clinical data, findings and a proposed diagnosis. The rules are then used to make further inferences from the supplied data and these, together with that data, are used to generate a short text which provides a "critique" of the user's proposed diagnosis. A later paper [30] describes a version of the system adapted for mammography.

The author's own work is on the development of decision aids for image interpretation based on a model of human decision making. A prototype has been developed as a decision aid for radiologists reading mammograms taken as a result of referrals from the breast screening programme [31]. Decision-making is viewed as a set of four activities: proposing candidates, generating arguments for and against, establishing relations between the candidates and evaluating their relative merits. These activities can be represented using a set of rules which may be used, together with knowledge about different kinds of decision, to select facts from a knowledge base relevant to a particular case and to construct a graphical display of those facts. The model is applied to three decisions in mammography: the classification of abnormalities, diagnosis and the selection of a course of action.

An on-going project described by Keravnou et al [32, 33] is developing an ambitious expert system to assist in the diagnosis of skeletal dysplasia on the basis of information entered by radiologists. The system is based on a model of diagnostic reasoning comprising three activities: the triggering, differentiation and evaluation of hypotheses. At the beginning of a consultation the user inspects the image and enters some findings into the system. If a finding is known to be strongly suggestive of a dysplasia, it is triggered and becomes an active hypothesis. The differentiation module selects the most plausible active hypotheses, divides them into clusters and then investigates each cluster in turn, a process which involves requesting information that distinguishes between hypotheses, and which may trigger new hypotheses. Less plausible hypotheses can be added to clusters which contain promising hypotheses. To evaluate hypotheses the system compares the findings entered and the findings which would be expected given the hypothesis, both to determine how many of the entered findings are explained and how many of the expected findings have been entered. In addition to knowledge about dysplasias, the system contains two types of "background knowledge": foundational knowledge related to diagnostic findings (*e.g.* "if a part of a component is abnormal then a component is abnormal") and a temporal model which supports reasoning with time-related facts about the case and about dysplasias. No evaluation of the described model has yet been performed although a limited evaluation of an earlier prototype showed a marked improvement over "manual diagnosis".

Again, the system meets some but not all of the criteria:

- **NEED**: although collectively dysplasias affect 1% of the population in the UK, there are about 2000 syndromes which individually occur infrequently and are regularly misdiagnosed.

- **PRACTICALITY**: published reports give no account of the interface or how it might meet the demands of the clinical setting.

- **VERACITY**: the bulk of the effort in the project has been devoted to developing a well designed representation of the domain and also of some of the less formalized "background knowledge".

- **RELEVANCE**: the question of how the information supplied by the system might be incorporated into the user's decision-making process has not been addressed.

Expert systems have, in recent years, been applied in a number of fields. They have tended to be most successful in clearly delimited domains where a large body of formalized knowledge exists and most useful where there is an identifiable class of users with a frequent need to consult this expertise. Expert systems for image interpretation have not yet been developed successfully because there is no real demand for the simpler systems and a number of research problems need to be solved before larger systems can be useful. Users need to be able rapidly to enter unambiguous descriptions of images and receive succinct descriptions of the likely diagnoses. This requires sophisticated interfaces for user input, novel techniques for knowledge representation and inference mechanisms which are designed as collaborative problem solvers.

Image processing systems

An image is represented digitally as an array of numbers or pixels, each number representing the brightness of a point in the image. The term "image processing" can be applied to any mathematical transformation of this array. The term covers the numerical treatment of digital images for quantitative measurements, image enhancement, object recognition, segmentation, 3D reconstruction and tomography. Work in all of these areas could be considered relevant to decision support; however, very little of it was explicitly intended to be. This section concentrates on the detection and classification of imaged objects, since image processing with these goals has been considered as a way of providing clinicians with decision aids.

Image processing techniques for object recognition or classification can be divided into two classes:

(1) Feature vector classification.

(2) Fitting models to photometry.

The easiest objects to recognize and identify are those which give rise to pixels of a distinctive brightness. If images are considered as three-dimensional distributions of pixels, with the dimension of brightness, or pixel value, being added to the two dimensions of the image, then setting thresholds on the brightness dimension is sufficient to identify such objects. In practice, this technique can be used only where some additional subtlety is employed in setting the thresholds, Parker et al [34] have developed a technique which detects microcalcifications in mammograms on the basis of significant peaks in the distribution of grey levels.

In feature vector classification systems, instead of using bands of the single dimension of pixel value to identify objects, they are matched with regions of a multidimensional space the dimensions of which correspond to photometry-based measures such as area or mean brightness. Object recognition is then performed by identifying the region of interest on the image, perhaps on the basis of brightness, then computing the feature vector (the values of the various photometry-based measures) for that region and classifying it accordingly. Systems of this type are often developed by feeding the photometry based measures for regions into software which is able to combine them with varying weights until a combination is found which produces a reliable classification, for example, Kegelmeyer and Allmen [35] used 46 features of regions identified as microcalcifications as input to a "Binary Decision Tree" in order to generate a classifier which would filter out false positives.

Other systems model objects in ways which can be matched directly with the image. Matsumoto et al [36] describe a system in which an idealized image of a projected 9 mm lung nodule is used as the pattern in a system for the detection of lung nodules. Since similar objects can give rise to widely different image patterns some flexibility must be built into the process. In model-based object recognition, the flexibility is generally built into the function which attempts to match the model to the image pattern, although some systems build the flexibility into the model. The standard example of a flexible matching function is the Generalised Hough Transform. This is used to match templates to image data in cases where the templates may be described using a set of parameters. For example, a circular template would be described by three parameters: its radius and the x and y co-ordinates of its centre. Patterns in the image data "vote" in parameter space (which in this case has three dimensions, one for each parameter), thereby identifying the preferred parameter values. This method has been used to detect stellate lesions [37] in mammograms.

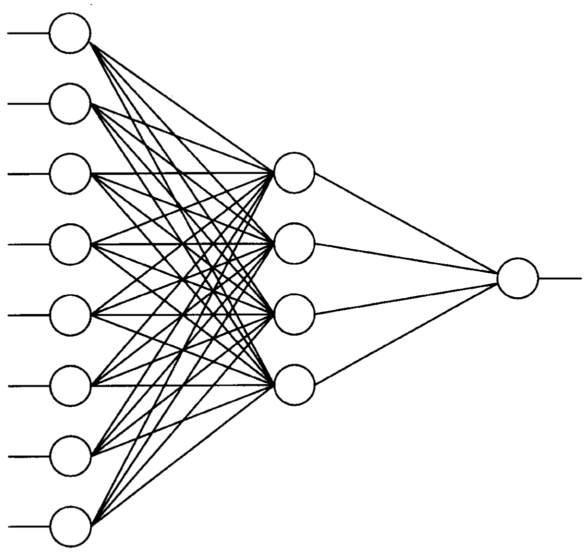

Figure 3. The nodes and connections in a typical back-propagation neural network with input and output layers and one hidden layer.

Neural networks are a special kind of flexible model. They consist of networks of elements, each element or node computes a weighted sum of its input and applies a function of some kind to generate an output, hence the analogy with neurons. Nodes are arranged in layers. The input layer receives input which may be raw image data, processed image data or information about the features of an interpreted image. The output layer provides the response of the system. Most neural nets used in object recognition systems include a single "hidden" layer between the input and output, as shown in Figure 3. To use the systems for object recognition they must first be trained. Training consists of providing the system with examples of the possible input and allowing a control loop (*e.g.* feedback or feed-forward) to adjust the weights of the system's nodes to produce the required output. The system is considered a success if it is able to generalize what it has learnt and classify new examples correctly.

Tourassi and Floyd [38] used neural nets to detect cold lesions in simulated noisy single photon emission tomography (SPECT) data, shown in Figure 4b. A network with 256 input nodes, eight hidden nodes and one output node was trained on neighbourhoods of 16×16 pixels. 220 patterns were used in training the net: 80 lesion free neighbourhoods and 10 each of seven different sizes of lesions viewed at two different counts levels.

The network was tested on 64×64 pixel images, again of simulated SPECT data. The 16×16 pixel input region of the neural net was moved over the SPECT images, pixel by pixel, and the response at each point was recorded in an image of the neural net's output. The image was filtered using a noise reduction technique and then regions of the output image where 10 or more adjacent positive responses coalesced were considered as

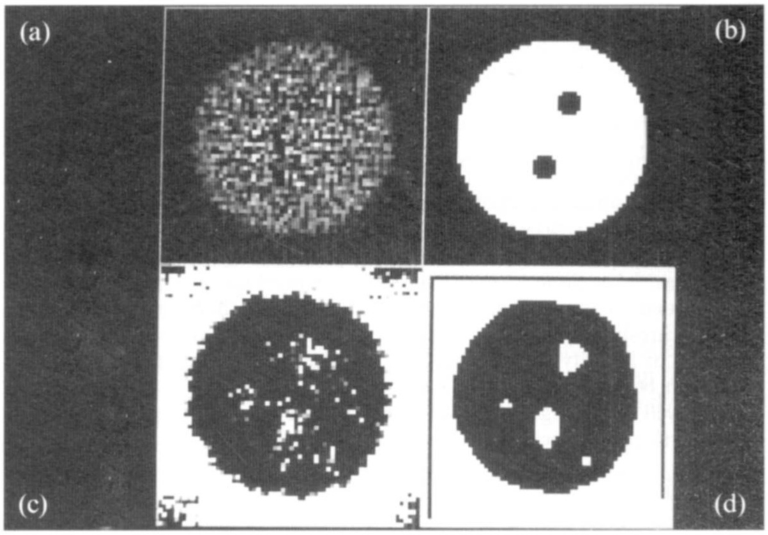

Figure 4. The simulated lesions used by Tourassi and Floyd [38] are shown: (a) after the addition of noise and (b) in their original form. The output of the neural net is shown in (c) and the final output of the system in (d).

representing detected lesions. Tested on 600 images, the network has 100% sensitivity on larger lesions while 80% sensitivity was found at a threshold of 50% specificity for 1 cm lesions.

Much of the work reported here represents attempts to improve techniques that may become useful in the future, rather than to demonstrate their value today. Even techniques that perform better than human observers on a particular detection task will not be useful if they are only slightly better than humans, if their operation is poorly understood or if they are attempting to answer only one of many questions involved in the interpretation of an image. Until recently few papers published on object recognition techniques in radiology described how the system could be used to support a clinical user. One paper which does is that of Astley et al [39]. The authors propose that research be focused on areas in which human interpreters need assistance and that we need to understand how well these techniques work and how they can be used effectively. These proposals correspond roughly to the criteria of NEED and PRACTICALITY used in this review.

Astley et al [39] have concentrated their work on the problems and challenges created by the UK's National Breast Screening Programme. Analysis of the errors made by radiologists can be used to provide evidence about the areas in which assistance is required. Techniques for detecting the different classes of lesions are described, but the authors accept that these are inadequate either for an automated sort prior to screening by radiologists or to serve as a second reader. The authors believe that existing object recognition systems may be most valuable in providing prompts to guide the radiologists' search, and they report research which attempts to establish how this can be achieved. They studied the effects of prompting on radiologists' ability to detect microcalcifications [40]. The information source in this case is the image processing algorithm, which was adequate as a detector of microcalcifications but ignored other abnormalities, hence the VERACITY criterion is only partially met. Detection was significantly better with prompts even with a moderate false positive rate, so the RELEVANCE criterion is met.

Research in object recognition has developed a battery of more or less effective algorithms for certain kinds of object recognition and classification. One problem which makes it difficult to assess work in this field, is that most researchers are developing and testing algorithms using images supplied by their clinical collaborators and which other researchers have no access to, hence different algorithms cannot readily be compared. There have recently been initiatives to make available databases of images on which researchers can test their programs, *e.g.* The Mammographic Image Analysis Society Digital Mammogram Database [41]. However, even where such resources are taken up, different researchers often present results in ways which prevent direct comparison: one researcher might report the percentage of microcalcifications successfully detected in each image, another may present an ROC analysis.

Even with effective techniques, without insight into how they can be employed to improve the performance of human radiologists, they remain unused in clinical practice. The challenge is to design a framework based both on an acceptance of the limitations of the image processing technology and on an awareness of the needs of clinicians. Many important questions remain unanswered. These questions are closely related to the criteria drawn up in the introduction to this review. Mammography is identified as an area where there is a need for support and the constraints which screening puts on how that is provided are understood, meeting our first two criteria. Meeting the VERACITY criterion would require detectors for a useful range of abnormalities. It is not clear if this has been met by research reviewed here, since only some abnormalities can be reliably detected. Gale et al [42] argue that strong individual differences in error rates on certain types of abnormality mean that this is enough to be useful. However, it seems unlikely that software which addresses only the problem of detecting one abnormality will be taken up. Nor is it clear that favourable results found when prompting for one abnormality will transfer to prompting for all abnormalities. Little is known about how prompting serves to improve performance and how its ability to improve performance is affected by parameters such as the false positive rate of the detection software.

Image understanding systems

The final class of systems to be considered is that of image understanding systems, systems which use image processing or analysis to produce a symbolic representation of the image. The major difficulty to be overcome in developing such systems is segmentation: the division of an image into regions which correspond to distinct objects in the imaged scene or, in radiology, anatomy. Once an image has been segmented the identified regions can then be analysed using knowledge about, for example, the imaged anatomy, to generate a symbolic representation of the image. Systems have been developed for the segmentation of CT scans [43–48] and the delineation of blood vessels on angiograms [49–51]. Image understanding is, however, still a research topic and none of these systems has been developed to the point of being useful as a decision aid.

Various ideas are common to many image understanding systems. A key one is that of processing at different levels. A succession of processes are applied to the image, each producing a new representation of the image which can form the input to the process at the next level. Early processing is heavily numerical and pixel-based, later processing is more symbolic and based on derived facts. Variations of this idea include the application of higher-level knowledge to control low-level processes and the ability to revise low-level processing on the basis of high-level results. A third variation is "blackboard processing": an attempt to allow the different processes applied in image interpretation to co-operate more closely. In a blackboard system, instead of each process taking input from a lower-level process and delivering output to a higher-level process, all processes read from, and write to, a blackboard which represents the current state of the interpretation.

One image understanding system is presented as a decision support tool [52]. The same technology has been applied in an early system for analysing sequences of scintigraphic images of the heart [53] and more recently for MR-based diagnosis of the knee [54]. The system is intended to be capable of deriving a symbolic description of the medical evidence contained in an image, without any human input whatsoever. The authors believe that such a system will help physicians make diagnoses and free them from the need to handle routine cases. In the scintigraphic system, the input consists of sequences of between 12 and 32 images of 64×64 pixels. Each one is smoothed and then a simple set of heuristics is used to identify a point which is certain to be within the left ventricle. The image is then reconstructed in polar co-ordinate space around this point and an edge detection program used to identify the ventricle boundary. Direction and expected ranges of occurrence are used to divide the edge up into the four ventricular regions, the results are shown in Figure 5. Subjective evaluation suggests that these low-level processes successfully analyse 93% of 420 test images. The knowledge-based processing is query-driven: the user asks if a particular concept represented in the knowledge base is true of the sequence under analysis and the system gives a likelihood derived for that concept. The query must match exactly a concept represented in the knowledge base. The answer is computed by breaking down the concept into its components to generate a set of primitive concepts, which correspond to the results of the low-level processing.

The knowledge representation has two components: a declarative component comprising an is-a and a part-of hierarchy and a procedural component which contains rules for computing areas, lengths, etc., for deriving descriptions of heart motion from area changes over the sequence, and for combining medical evidence. The system was evaluated on nine sequences for which a doctor gave a diagnosis. In each case the system suggested a likelihood estimate for that diagnosis which was higher than for the other possible diagnoses. It is an impressive system with reliable image processing at the early stages and a large knowledge base at the later stages. The evaluation, although no less thorough than many, is only on a limited series of images. Assessing it against the criteria:

- **NEED**: the need for such a system is assumed rather than argued.

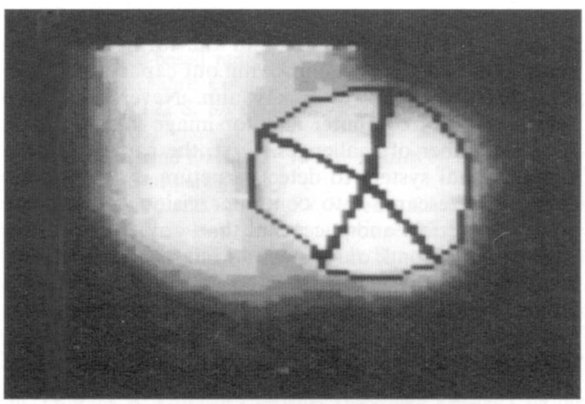

Figure 5. An anatomically meaningful segmentation of the left ventricle, generated automatically by Niemann et al's knowledge-based image understanding system [53]. The four sectors correspond to the spetal, inferioapical, posterolateral and basal regions of the left ventricle (© IEEE 1985).

• **PRACTICALITY**: no account is given of the demands of the clinical setting.

• **VERACITY**: the knowledge representation used in the system is derived from an analysis of the objects and concepts that make up the domain.

• **RELEVANCE**: the described interface allows the user to ask any question and receive a reply, but the basis for the reply must inevitably remain obscure in so large a system.

In the majority of papers describing image understanding systems the architecture of the system is the key element of interest: the way processing is divided up into levels, or into sub-systems, and the way in which the interaction of different components is controlled. It is assumed that the difficulty in image interpretation is that of managing the way the processing is carried out. One of the consequences of this is that relatively little is said about the first two criteria, or about the medical knowledge required by the VERACITY criterion.

The other component of the information source in these systems is the image processing. The recurring problem in image interpretation is segmentation: the problem of dividing up an image into regions which correspond to planes or to objects in the imaged scene. There is no definitive solution to the problem and all systems include *ad hoc* approaches to dealing with less than perfect segmentations. This is more of a problem on more complicated images.

The major weakness of these systems is revealed when the fourth criterion is considered. A system which performs the whole of the image interpretation task leaves no clear role for the user. It seems that there is some way to go before image interpretation systems can be developed to provide useful assistance in difficult areas of image interpretation.

Conclusion

Progress in medical imaging has been as rapid as that in other technologically advanced fields and our capacity to generate and display data in the form of medical images is now such that improving our capacity to interpret this data must be a key aim. Nevertheless, the development of computer aids for image interpretation faces a number of challenges. First, the capacity of the human visual system to detect structure is still, despite decades of research into computer vision, much more powerful, flexible and successful than any technological alternative. Second, our capacity to design tools to augment, rather than replace, human interpretative skills is limited by our understanding of how the human perceptual system works.

The research reviewed here meets these two challenges with varying degrees of success. The use of image databases for decision support is an appealing strategy since images are a source from which the human visual system can effectively extract information. If a computer is to be used to provide decision support from an image database, a system must be devised which allows the user to retrieve an image on the basis of its content. Progress is being made both on ways of analysing images into their components [6] and with regard to methods of indexing images relevant to clinicians' decision problems [7]. Numerical methods and expert systems also leave the interpretation of images to the eye, simply making available information to help in assessing the significance of image features. As the computerized collection of medical data becomes more developed, the potential for systems which make numerical data available to help in decision-making will increase, and so will the need for research into the most effective ways of presenting statistical information. Expert systems in medicine suffer from the problems of representing and handling large amounts of knowledge. The challenge is to develop a system which is able to handle a domain complex enough for such systems to be required. One of the key problems in these systems is in constructing a knowledge base that allows users to describe the features of visual images adequately, another is in developing a style of interaction which ensures that the expert system is complementing the user's expertise.

The use of image processing techniques is problematic because the human visual system is still better than computers at distinguishing between normal and abnormal structure in a previously unseen image. There is, however, an increasing amount of evidence to suggest that these techniques can be applied in certain cases and do improve radiologists' decision-making. The trick here is to detect certain classes of abnormality on images which humans find difficult. Even then, current algorithms are sufficiently sensitive to be useful only when a high false positive rate is permitted. It has been shown that, even if there are two false positives per image, using the positive responses from an abnormality detector as prompts does improve radiologists' detection rates. However, if a battery of detectors is to be provided for images where a range of abnormalities may be detected, the false positive rates must be kept much lower if the total number of prompts is to be held at an acceptable level. Meanwhile, the prospect for image understanding systems able to provide an interpretation of complex medical images is still remote. The central problem here is that of segmentation; the human eye is still the best system for identifying anatomically meaningful boundaries on noisy images.

This review has highlighted a number of problems in the development of decision aids for image interpretation. First, the choice of domain. There are many areas in which the images are difficult to interpret or the expertise required is in short supply or the number of images generated stretches the available resources: areas in which computer aids could be useful. The problem is that these are not necessarily the areas for which it is easiest to develop a decision aid. In a number of cases it seems that research has been driven more by what is possible than by what is desirable. Second, there are the problems relating to how the system is to be used. A successful system must be designed for the setting in which it is to be used. It is less obvious what the standard of usability should be for a research prototype, especially when one considers the pace at which digital technology is changing the nature of radiology. Three points are clear: a system which requires any lengthy interaction is only going to be used as a measure of last resort; any system which is integrated into the normal routine of the clinician and performs some clerical tasks is more likely to gain acceptance and, finally, clinicians will be unwilling to learn how to use many different decision support systems.

The different sections of this review have surveyed decision support systems based on different kinds of computerized information source. No one kind of knowledge is likely to prove pre-eminent and yet the proliferation of systems would mean redundancy and be confusing and daunting for users. There is therefore a strong argument for attempting to develop an integrated decision aid which is able to provide information of different kinds on request, or in response to different problems. Greenes [55] proposes one solution to this problem: that the developers should consider how their system could be connected to other systems and a common interface developed for a set of "building block" decision aids. Swett et al [29] argue for the development of a radiologists' workstation, which suggests a closer knit integration. They outline a model of human vision as consisting of a pre-attentive phase in which recognition may occur automatically, followed by an attentive phase in which conscious consideration is given to identified features and finally a decision phase in which additional information may be sought. Computer aids complementing the radiologists' skills could be developed for each of these phases.

Few medical decision support systems, and none of the early systems, have been developed as collaborative problem solvers. Their makers assumed that it would be possible to develop a system that performed demonstrably better than a human decision-maker. Such a system could then be installed and would be consulted by the decision maker when he or she required special assistance. Few such systems entered widespread use. One of the reasons for this failure is undoubtedly the fact that the designers were more concerned with ensuring that the systems out-performed clinicians than with improving the performance of clinicians. This leads to weaknesses because it means that the systems haven't built on the users' own expertise. Perhaps more importantly, it leads to a failure of user acceptance, since it is not clear how the user should behave when he or she feels that the system is in error. The responsibility for the decision clearly lies with the clinician and although the machine is supposedly more reliable than the clinician, it is certainly not infallible.

Computer technology, of the kind reviewed here, may be useful both as a mechanism for independently identifying relevant image features and in making information available to help in decision-making. Further development in computer technology, notably in image processing, is required before a radiologists' workstation becomes possible, but much has been achieved and much of the remaining challenge lies in understanding how best to apply technology to meet the needs of radiologists.

Acknowledgments

The author is supported by the Imperial Cancer Research Fund, London and the Institut Bergonié, Bordeaux. Professor Fox, Professor Todd-Pokropek, Ms Hajnal, Dr Dicks-Mireaux, Dr Wooding and Ms McNicol all read drafts of this review, their help and advice is gratefully acknowledged, as are the comments of the anonymous referees. Dr Cohn, Dr Floyd, Professor Niemann and Dr D'Orsi very kindly supplied the figures used in this article.

References

1. WYATT, J and SPIEGELHALTER, D, Field trials of medical decision-aids: potential problems and solutions. In *Proceedings of Fifteenth Annual Conference on Computer Applications in Medical Care*, ed. by P Clayton (American Medical Informatics Association, Bethesda, USA), pp. 3–7 (1991).
2. ALLISON, D J, MARTIN, N J, REYNOLDS, R A and STRICKLAND, N H, On clinical aspects of PACS, *Proc. 18th Int. Congress on Radiol.*, 23–28 (1994).
3. CHANG, S K, YAN, C W, DIMITROFF, D C ET AL, An intelligent image database system, *IEEE Trans. on Software Engineering*, *14*, 681–688 (1988).
4. GEVERS, T and SMEULDERS, A W M, Indexing of images by pictorial information. In *Visual Database Systems II*, ed. by E Knuth and L M Wegner (Elsevier Science Publishers), pp. 93–100 (1992).
5. WIEDERHOLD, G, BRINKLEY, J, SAMADAMI, R and CLAUER, C R, Model-driven image analysis to augment databases. In *Visual Database Systems: Proceedings of IFIP TC2/WG 2.6*, ed. by T I Kunii (Elsevier Science, Amsterdam), pp. 159–180 (1989).
6. KOFAKIS, P and ORPHANOUDAKIS, S, Image indexing by context. In *A Second Generation PACS Concept*, ed. by M Osteaux, A R Bakker, D Bell et al (Springer-Verlag), pp. 250–293 (1992).
7. COHN, A I, MILLER, P L, FISHER, P R ET AL, Knowledge-based radiologic image retrieval using axes of clinical relevance, *Comp. Biomed. Res.*, *23*, 199–221 (1990).
8. SCHWARTZ, S, *Medical Thinking* (Springer-Verlag, New York) (1986).
9. TABAR, L and DEAN, P, *Teaching Atlas of Mammography* (Thieme, Stuttgart) (1983).
10. REEDER, M and BRADLEY, W (eds), *Reeder and Felson's Gamuts in Radiology (3rd edn)* (Springer-Verlag) (1992).
11. LEDLEY, R S and LUSTED, L B, Reasoning foundations of medical diagnosis, *Science*, *130*, 9–21 (1959).
12. LODWICK, G S, Computers in radiologic diagnosis, *Appl. Radiol.*, *Jan/Feb*, 61–65 (1986).
13. BLOIS, M S, Clinical judgement and computers, *New Engl. J. Med.*, *303*, 192–197 (1980).
14. GOOD, B, *Effect of Computer Generated Feedback on the Performance of Radiologists (PhD thesis)* (University of Pittsburgh, USA) (1992).
15. D'ORSI, C, GETTY, D J, SWETS, R ET AL, Reading and decision aids for improved accuracy and standardization of mammographic diagnosis, *Radiology*, *184*, 619–622 (1992).

16. SELTZER, S E, MCNEIL, B J, D'ORSI, C J D ET AL, Combining evidence from multiple imaging modalities, a feature-analysis method, *Comp. Med. Imaging Graphics, 16*, 373–380 (1992).
17. KUHNS, L R, THORNBURY, J R and FRYBACK, D G, *Decision Making in Imaging* (Year Book Medical Publishers, Chicago) (1989).
18. HEATHFIELD, H and WYATT, J, Philosophies for the design and development of clinical decision-support systems, *Methods Info. Med., 32*, 1–9 (1993).
19. DE DOMBAL, F T, BARNES, S, DALLOS, V ET AL, How should computer-aided decision-support systems present their predications to the practising clinician, *Theor. Surg., 7*, 111–116 (1992).
20. TANAKA, T and SUEDA, N, Knowledge acquisition in image processing expert system 'Explain', *Proc. Artificial Intell. Indust. Appl.*, 267–272 (1988).
21. HU, Z P, PUN, T and PELLEGRINI, C, An expert system for guiding image segmentation, *Comp. Med. Imaging Graph., 14*, 13–24 (1990).
22. DE ROSIS, F, GIANCASPRO, A, PIZZUTILO, P and PASQUEARIELLO, G, A consultation system for medical image analysis, *Expert Systems Appl., 5*, 35–42 (1992).
23. MATSUYAMA, T, Expert systems for image processing, analysis, and recognition for computer vision, *Advances Electronics Electron Phys., 86*, 81–171 (1993).
24. COOK, H M and FOX, M C, Application of expert systems to mammographic image analysis, *Am. J. Physiol. Imag., 4*, 16–22 (1989).
25. PIRAINO, D W, RICHMOND, B J, UETANI, M ET AL, Problems in applying expert system technology to radiographic image, *J. Digital Imag., 2*, 21–26 (1989).
26. ENGELMANN, U, MEINZER, H P, SCHEPPELMAN, D and SPRINGUB, A, Knowledge-based image analysis on the basis of predicate logic, *Med. Inform. Eur., '90*, 763–767 (1990).
27. STRAUSS, A, *ESPOIR: un système multi-expert d'aide à la décision devant des mammographies presentant des microcalcifications (PhD thesis)* (Center Anti Cancereux de Montpelier) (1989).
28. STEWART, B D, CAIRNS, A Y and RICKETTS, I W, The analysis of breast abnormalities by computer. In *Digital Mammography*, ed. by A Gale, S Astley, D Dance and A Cairns (Elsevier Science, Amsterdam), pp. 241–251 (1994).
29. SWETT, H A and MILLER, P L, ICON: a computer-based approach to differential diagnosis in radiology, *Radiology, 163*, 555–558 (1987).
30. SWETT, H A, GIGER, M L and DOI, K, Computer vision and decision support. In *The Perception of Visual Information*, ed. by W R Hendee and P N Temple (Springer–Verlag, Berlin), pp. 274–315 (1993).
31. TAYLOR, P, Decision support for image interpretation: a mammography workstation. In *Information Processing in Medical Imaging*, ed. by Y Bizais, C Barillot and R Di Paola (Kluwer Academic Publishers, Dordrecht), pp. 227–238 (1995).
32. KERAVNOU, E T, DAMS, F, WASHBROOK, J ET AL, Background knowledge in diagnosis, *Artific. Intell. Med., 4*, 263–279 (1992).
33. KERAVNOU, E T, DAMS, F, WASHBROOK, J ET AL, *Modelling Diagnostic Skills (Technical report)* (Department of Computer Science, University College London) (1993).
34. PARKER, J, DANCE, D R and SUCKLING, J, The hysteresis technique for detection of calcifications in digital mammograms. In *Digital Mammography*, ed. by A Gale, S Astley, D Dance and A Cairns (Elsevier Science, Amsterdam), pp. 41–51 (1994).
35. KEGELMEYER, W P and ALLMEN, M C, Dense feature maps for the detection of microcalcifications. In *Digital Mammography*, ed. by A Gale, S Astley, D Dance and A Cairns (Elsevier Science, Amsterdam), pp. 3–13 (1994).
36. MATSUMOTO, T, YOSHIMURA, H, DOI, K ET AL, Potential usefulness of computerized nodule detection in screening programmes for lung cancer, *Invest. Radiol., 27*, 471–475 (1992).
37. ASTLEY, S, TAYLOR, C, BOGGIS, C ET AL, Cue generation and combination for mammographic screening. In *VISUAL SEARCH II*, ed. by D Brogan, pp. 207–223 (1993).
38. TOURASSI, G D and FLOYD, C E, Artificial neural networks for single photon emission computed tomography, *Invest. Radiol., 28*, 671–678 (1993).
39. ASTLEY, S, HUTT, I, ADAMSON, S ET AL, Automation in mammography: Computer vision and human perception, *Proc. Biomed. Image Processing Biomed. Visualiz.*, 716–730 (1993).
40. HUTT, I W, ASTLEY, S M and BOGGIS, C R M, Prompting as an aid to diagnosis in mammography. In *Digital Mammography*, ed. by A Gale, S Astley, D Dance and A Cairns (Elsevier Science, Amsterdam), pp. 389–398 (1994).
41. SUCKLING, J, PARKER, J, DANCE, D R ET AL, The Mammographic Image Analysis Society Digital Mammogram Database. In *Digital Mammography*, ed. by A Gale, S Astley, D Dance and A Cairns (Elsevier Science, Amsterdam), pp. 375–378 (1994).
42. GALE, A G, WILSON, A R M, and ROEBUCK, E J, Mammographic screening: Radiological performance as a precursor to image processing, *Proc. Biomed. Image Processing Biomed. Visualiz.*, 458–464 (1993).
43. LIN, W C, WENG, Y T and CHEN, C T, Expert vision systems integrating image and segmentation and recognition processes, *Eng. Appl. Artificial Intell., 1*, 230–249 (1988).
44. KARSSEMEIJER, N, VAN ERNING, L J T O and EIJKMAN, E G J, Recognition of organs in CT image sequences: a model guided approach, *Comp. Biomed. Res., 21*, 434–448 (1988).
45. DHAWAN, A P and JUVVADI, S, Knowledged-based analysis and understanding of medical images, *Comp. Methods Prog. Biomed., 33*, 221–239 (1990).
46. NATARAJAN, K, CAWLEY, M G and NEWELL, J A, A knowledge-based system paradigm for automatic interpretation of CT scans, *Medical Informatics, 16*, 167–183 (1991).
47. DELLEPIANE, S, VENTURI, G and VERNAZZA, G, Model generation and model matching of real images by a fuzzy approach, *Pattern Recog., 25*, 115–137 (1992).
48. LI, H, DE KLERCK, R and CORNELIS, J, Integration of multiple knowledge sources in a system for brain CT-scan interpretation based on the blackboard model, *Second Int. Conf. Practical Appl. Prolog.*, 336–343 (1994).
49. STANSFIELD, S A, ANGY—a rule-based expert system for automatic segmentation of coronary arteries vessels from digitally subtracted angiograms, *IEEE Trans. Pattern Analysis Machine Intell., 8*, 188–199 (1986).
50. CATROS, J Y and MISCHLER, D, An artificial intelligence approach for medical picture analysis, *Pattern Recog. Letters, 8*, 123–130 (1988).
51. SMETS, C, VAN DE WERF, F, SUETENS, P and OOSTERLINCK, A, An expert system for the labeling and 3D reconstruction of the coronary arteries from two projections, *Int. J. Card. Imaging, 5*, 145–154 (1990).

52. NIEMANN, H, SAGERER, G, SCHRODER, S and KUMMERT, F, ERNEST: A semantic network system for pattern understanding, *IEEE Trans. Pattern Analysis Machine Intell.*, *12*, 883–906 (1990).
53. NIEMANN, H, BUNKE, H, HOFMANN, I ET AL, A knowledge based system for analysis of gated blood pool studies, *IEEE Trans. Pattern Analysis Machine Intell.*, 7, 246–259 (1985).
54. WEIERICH, P, WETZEL, D and NIEMANN, H, Audigon—a medical expert system with new judgement functions for the fuzzy rulebase, *Proc. First Europ. Congress Fuzzy Intell. Technol.*, 472–477 (1993).
55. GREENES, R A, A building block approach to application development for education and decision support in radiology, *J. Digital Imaging*, *4*, 213–225 (1991).