AN IMAGE PROCESSING SYSTEM FOR AUTOMATED SCREENING OF CHEST X-RAYS

J.H. Kulick, T.C. Challis, K.B. Martin, D.C. Haig, and C.M. Haig
Queen's University, Kingston, Ontario

ABSTRACT

Over the past 4½ years, the Department of Computing and Information Science at Queen's University has developed an interactive image processing laboratory. This laboratory is being used to develop an automated system for the screening of chest X-rays.

The minicomputer-based system includes a colour raster image display system, a TV camera, and an image digitizer. Supporting software includes: device drivers, image display and manipulation routines, image processing operations, and virtual image handling.

The approach taken has been to develop a model of a "normal" chest X-ray. This model includes various forms of knowledge: characteristics of X-ray films, structural knowledge concerning location and shape of various anatomical components, and radiologist-based information concerning the distinguishing features of abnormal X-rays. Recent results using statistical classifiers have been very encouraging, and will be discussed.

UN SYSTÈME DE TRAITEMENT D'IMAGES POUR LA RECONNAISSANCE DE RAYONS-X DE LA POITRINE

RÉSUMÉ

Durant les quatre et demi années dérrières le département d'informatique à l'Université Queen's a developé un laboratoire pour le traitement interactif des images. Ce laboratoire est présentement utilisé pour la conception d'un système de reconnaissance automatique de rayons-X.

Le système utilise un mini-ordinateur et comprend des commandes d'unités, des programmes pour la manipulation et l'affichage des images, des opérations de traitement d'images ainsi que la manutention des images virtuelles.

L'approche adoptée a été de développer un modèle pour le rayon-X d'une poitrine "normale". Ce modèle comprend plusieurs formes d'informations: des caractéristiques de films de rayons-X, des connaissances structurales concernant la location et la forme de différentes composantes anatomiques, ainsi que des informations basées sur les connaissances de radiologues concernant les caracteristiques distinguant les rayons-X anormaux. De récentes expériences utilisant des classificateurs statistiques ont été très encourageantes et seront discutées.

1. Introduction

There are several reasons contributing to the desirability of an automated chest x-ray "screening" system:

(1) several hundred million chest films are taken every year in North America,

(2) the diagnostic agreement in a screening environment by radiologists is about 70% [1],

(3) and the number of abnormal films in a screening environment is usually low, about 7%.

At Queen's University, a computer-based system is being developed to operate in a mass "screening" environment. The system's function is to detect "abnormal" films so that they may be analysed more thoroughly by trained radiologists.

The major goals of the project are:

(1) developing hardware and software for acquiring, processing, and displaying imagery data

(2) developing computer algorithms for analysing chest x-rays

(3) evaluating the system's performance.

2. Hardware

The Image Processing Laboratory hardware now consists of the following (see Fig. 1):

(1) Norpak RGP-3000 colour raster display - 480 x 640 x 8 bits of resolution and colour and grey level lookup table

(2) Norpak VID Video Image Digitizer - 480 x 640 x 8 bits of resolution with 2 second image acquisition time

(3) Sierra Scientific LSV-1 black and white TV camera

(4) Conrac QQA B/W TV monitor

(5) Electrohome 19" colour TV monitor

(6) Tektronix 4006 storage tube graphic display terminal

The above hardware is interfaced to the department's PDP 11/45 minicomputer which runs under the Unix timeshared operating system. Resources include: 104K words of core memory, 28 million bytes of disk storage, floating point processor, and the usual peripherals (magnetic tape, user terminals, printers, etc.).

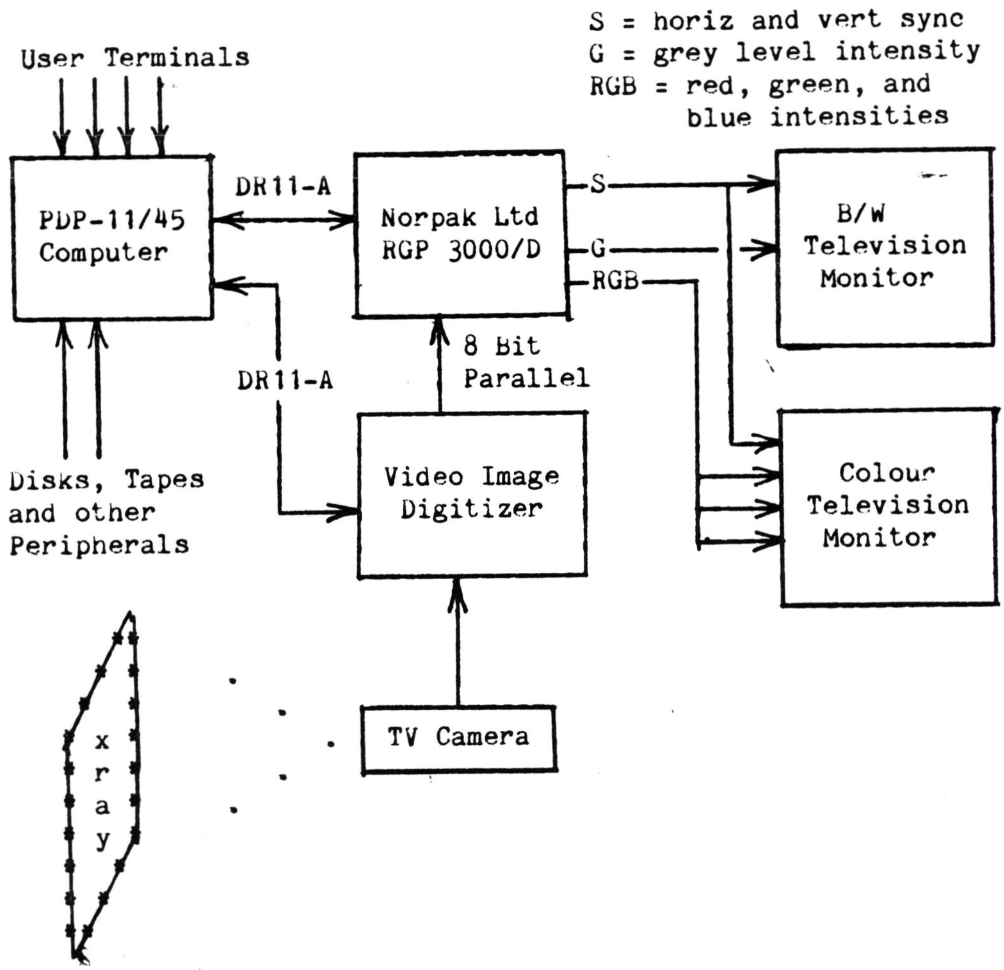

Figure 1. Image Processing Lab Hardware Configuration

The Image Processing Lab also has a parallel processor system, called Atoms, under development. This system, when finished, will allow up to 16 computers to share resources while working on an image processing problem. Currently, the required operating system support is being developed to allow cooperating parallel processes.

3. Software Support

The first level of software was developed to provide a reasonably high-level interface to the image processing laboratory hardware. This includes routines for: (1) digitization of images of varying size and resolution, (2) reading these images from the raster display memory back to the computer and vice versa, (3) colour and grey level lookup table definition allowing such operations as thresholding, image enhancement through "pseudo-colouring" (grey level intensities are mapped into different colours) (4) vector and character generation for graphic overlay of imagery data, and (5) user coordinate input using a tracking cursor.

Another group of routines are used for the storage and retrieval of large images. The major function of these routines is to provide a means for processing images larger than a program's available workspace (e.g. 1024 x 1024 images in a 32K byte workspace). The user may also define "virtual" images, in his own coordinate system, in terms of the original image.

Finally, a collection of routines for performing common image processing tasks have been implemented. These include routines for (1) histogram computation and modification, (2) image thresholding, (3) edge detection, (4) "fast" Fourier and Hadamard transforms, (5) image smoothing, sharpening, and filtering, and (6) output of images to hard copy printers.

4. Analysis of Chest X-Rays

Our basic approach has been to develop algorithms which (a) segment the digitized x-ray image into its major components, (b) measure and classify the heart as either normal or abnormal, and (c) measure and classify the lung fields as either normal or abnormal.

The latter two steps are examples of the general problem called patten classification. In general, a classifier is first "trained" on a set of known patterns, so that it may "learn" how the features vary for each class. Once a classifier has been sufficiently trained, it may be reliably applied to the classification of unkown patterns.

Various image measurements are made, pertaining to the characteristics of normal hearts and lung fields. These features are used for both training and testing phases. Once the classifiers have been trained, unknown x-rays are presented to the system for analysis, and results compared to radiologist's diagnosis.

Reliability in correctly classifying abnormal x-rays is imperative, On the other hand, misclassifying normals as abnormal is not nearly so serious. Various internal decisions made by the system have been designed to reflect this.

Figure 2 depicts the overall flow of control and data in the analysis of a single chest x-ray. The individual components are discussed later, in greater detail.

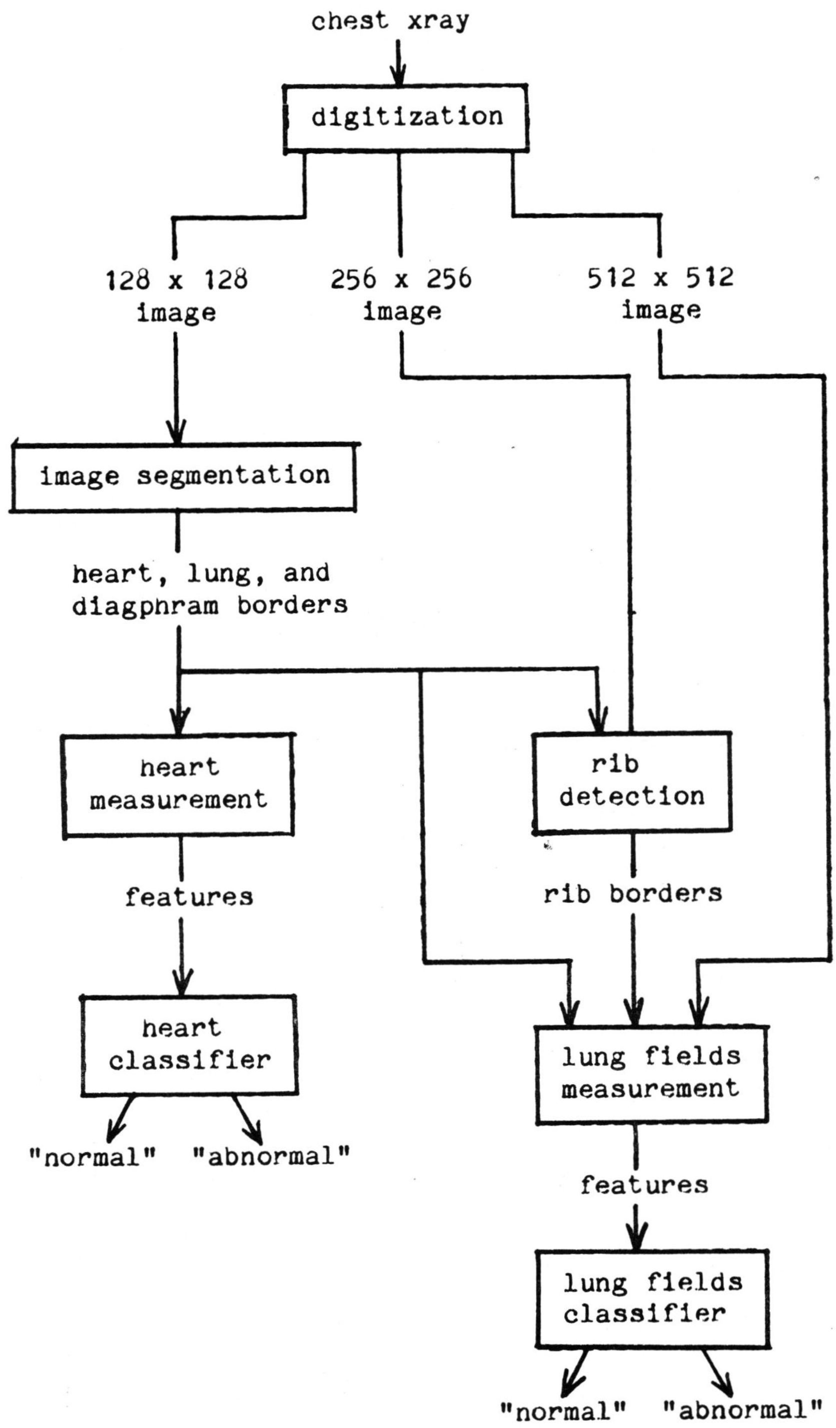

Figure 2. Overall Flow of Control and Data

Image Segmentation

Image segmentation is performed on a low resolution (128 x 128) version of the chest x-ray film being analysed.

Various anatomical knowledge concerning location of the major organs is used to drive a variety of histogram analysis, threshold selection, and smoothing procedures in obtaining reliable approximations to the lung, diaghram, and heart border outlines.

The output of the segmentor is a data structure describing the various borders and their intersection points. This data is used extensively by subsequent components of the x-ray analysis system.

Heart Analysis

Based on radiologist's suggestions, various measurements are made of heart shape and size. These measurements are computed using the image segmentation data. Examples of measurements taken are heart/lung area ratio, heart/thorax width ratio, and polynomial coefficients describing curvature of left heart border. A total of 15 features (7 ratios, 4 shape coefficients, and 4 shape approximation residuals) were obtained.

The classifier we have initially implemented and used is called Fisher's Linear Discriminant Function [2]. For the two class case, patterns are classified by first projecting them onto a line and then comparing with a single decision value. The orientation of the line and the decision value are obtained by maximizing the ratio of between-class scatter and within-class scatter. The dimensionality of the line is equal to the number of features being used in the classification process.

Features and pairs of features of were used to build individual classifiers.

The performance of the individual classifiers was measured using leave-one-out jack-knife testing. This involves training on all but one of the patterns in the data set, then presenting the removed pattern to the classifier as an "unknown", and recording its classification. This process is repeated with a second pattern, and so on until all patterns have been classified in this way. The total number of errors (i.e. misclassifications) is a very good measure of the classifier's performance.

In order to achieve high success rates in detecting abnormal x-rays, it was decided to use a sequential classifier, consisting of several individual classifiers. An heart would be classified normal only if classified normal by all of the individual classifiers.

Heart analysis has been performed on two data sets. First, on a rather small training set (35 films), and then on a second much larger set (250 films). The results are presented below as confusion matrices, where N = normal and A = abnormal.

35 films

		Computer Findings N	Computer Findings A	
Radiological Findings	N	12	3	82%
	A	0	19	100%
				91%

250 films

	N	A	
N	49	26	68%
A	2	173	99%
			89%

Rib Detection

We believe that the accurate detection of ribs is a necessary step in the automated analysis of chest x-rays. The reasons are two-fold: Firstly, lesions along rib margins can only be determined by accurately locating the ribs. Secondly, a key part of our approach to lung field evaluation requires texture measurements to be made separately along ribs and within inter-costal spaces (i.e. the areas between ribs).

Previous algorithms have been very time-consuming, requiring pre-processing of the entire image to enhance rib borders.

Our approach has been to use previously derived image segmentation data, as well as information concerning the number of nature of ribs, to greatly reduce both time and cost.

The following are the major steps in the location and description of the ribs:

(1) Enhance a number of vertical slices through the lung field, spaced at equal intervals between the heart and outer lung borders, by simple spatial averaging

(2) Process each slice by a technique known as "hysteresis" smoothing [2], varying the smoothing window height until the number of major peaks detected approaches the number of ribs being sought

(3) Estimate rib borders from the peaks of the smoothed curve

(4) Use anatomical information to reject implausable border point pairs from further consideration

(5) Perform global linking of all detected rib border points into a set of rib point lists using a constrained minimal spanning algorithm

(6) Fit 2nd order polynomial curves to rib point lists

(7) Verify validity of final set of ribs.

Preliminary results obtained on a set of 35 films indicate further work is required. Ribs which are clearly visible are detected very reliably, while highly transparent ribs, which are barely visible, are sometimes missed. Also, when large areas of the lung are diseased, one or more ribs may be obscured, and as a result, go undetected.

Lung Field Analysis

The analysis of the lung fields is, of course, a major aspect of the screening of chest x-rays. Our approach has been to identify the areas utilized by radiologists, such as inter-costal spaces, and to process them by using relationships similar to those used by radiologists. In particular, we are examining the inter-costal spaces in small disjoint zones of the lung and comparing them both bilaterally and within the same lung. These regions are formed by dividing each lung into upper and lower halves, and each half into inner, middle, and outer thirds, for a total of twelve individual zones.

We are currenlty using two types of digital texture measures: first order and second order. First order measures tend to characterize average tendencies, the most familiar being average density. Second order measures characterize how pairs of pixel vary. One can compute any order texture measure, but psychophysical experiments indicate that the human eye can only detect first and second order differences reliably [3].

The first order texture measures used were the mean and variance of grey levels over the specified zones.

The second order measures are based on the grey level difference statistics [4]. These measures are related to the joint probability distributions of pairs of pixels at various distances apart and orientations. Four different measures were taken in four directions (0, 45, 90, and 135 degrees) and for two distances (1 and 2 pixels apart). This yields a total of 32 second order features for each zone. Together with the 2 first order we now have a total of 34 features per zone.

The twelve zones of the lung fields are examined both individually and in pairs to determine which measures give better classification results. In addition to the 12 sets of features for individual zones, the following sets of features were computed: bilateral inter-lung ratios, horizontal intra-lung ratios, and vertical intra-lung ratios.

Fisher's Linear Discriminant Function classifiers were trained, one for each zone. The lungs were classified as "abnormal" if any of its zones were so classified, otherwise it was classified as "normal".

For the purposes of preliminary study, 50 films were used to train and test the system, again using leave-one-out testing. The over-all film classification rates for the various types of features are as follows:

Radiological Findings	Computer Findings: N	Computer Findings: A	
N	16	12	57%
A	1	22	96%
			75%

Individual Zones

	N	A	
N	12	16	43%
A	0	23	100%
			69%

Bilateral Pairs

	N	A	
N	10	18	36%
A	0	23	100%
			65%

Horizontal Pairs

	N	A	
N	11	17	39%
A	2	21	91%
			63%

Vertical Pairs

The overall results for abnormals are quite good; admittedly, at the expense of poorer results for the normals. Since the goal is a reliable screening system, this is tolerable.

5. Future Plans

Much improvement is needed, especially in the rib detection and lung field analysis components of the system. The performance of the lung field classifier on normal xrays is not satisfactory.

One of the difficulties of testing such a system is obtaining test data. Currently, we are developing a significant-sized (1000 films) library of chest x-ray films with corresponding diagnoses from five radiologists.

The radiologists are recording their diagnoses on specially designed forms, which depict regions of interest and possible diagnosis categories. These forms will then be stored in an on-line database using our image processing system.

The database will be used to record computer results, so that overall system performance may be measured. Also, when training the system, it will be possible to build training data sets consisting only of x-rays on which there is a high consensus (80%) of agreement among the five radiologists.

6. References

[1] Ziskin et al., "X-Ray Image processing and Its Evaluation", Journal of the Association of Advanced Medical Instrumentation, 1972, pp. 283-287.

[2] Duda, R. O. and Hart, P. E. Pattern Classification and Scene Analysis. New York, Wiley, 1973

[3] Julesz, B., "Experiments in the Visual Perception of Texture", Scientific American, Vol. 232, No. 4, April 1975, pp. 34-43.

[4] Weszka, J. S., "A Comparative Study of Texture Measures for Terrain Classification", IEEE Trans. on Systems, Man, and Cybernetics, Vol. SMC-6, No. 4, April 1976, pp. 269-385.

[5] Haig, D. C, Kulick, J. H., Challis, T. W., Haig, C. M., "Automated Screening of Chest X-Rays: Image Segmenation and Cardiac Silhouette Analysis", Proceedings of the Fourth International Joint Conference on Pattern Recognition, Kyoto, Japan, November 1978.

[6] Brace, C. M., Kulick, J. H., Challis, T. W., "Automated Detection of Ribs in Chest X-Rays", Proceedings of the International Joint Conference on Artificial Intelligence, M. I. T., August 1977.

[7] Haig, C. M, "Automated Texture Analysis of Lung Fields", Master's Thesis, Department of Computing and Information Science, Queen's University, Kingston, April, 1978.

[8] Challis, T. W., Haig, C. M., Haig, D. C., Kulick, J. H, "Automated Analysis of Chest X-Rays: Texture Analysis of Lung Fields", presented at the SPIE Conference, Toronto, March 1979.