FEATURE EXTRACTION IN IMAGES***

Charles A. Harlow*, Jacqueline L. Otto*, David L. Hall**, Gwilym S. Lodwick**
Departments of Electrical Engineering* and Radiology**, University of Missouri-Columbia
Columbia, Missouri

Abstract

The problem considered is the extraction of features from a complicated set of images. In this approach the basic concepts used in the feature extraction process are: 1) an algorithm that enumerates regions, a region being a contiguous set of points whose grey shade varies little; 2) a search for large objects in the picture first and then their details; 3) a description of the type of picture class which guides the feature extraction; 4) "a field of vision" concept: given an area of the picture and description of the objects that comprise this part of the picture, classify all the points in each of the objects. A graphical description is used to convey these concepts. The program is applied to anteroposterior chest x-rays but applies equally to other picture classes.

Introduction

Many potential application areas exist for image processing systems. These application areas include character recognition, fingerprint recognition, aerial reconnaissance photographs, bubble chamber pictures, automatic visual inspection, robot vision, vision for machine tool applications, pathology slides, medical x-rays, and nuclear medicine images. The number of possible application areas, then, is large while the number of present applications is small. However, image analysis techniques have, to some degree, been successful in character recognition, chromosome analysis, bubble chamber pictures, and in some medical applications.

The successes to date in developing image analysis systems have been modest. The major problem is the lack of a paradigm to guide one in developing a system of this kind. Invariably an investigator is faced with the problem that an analysis scheme which solves one problem provides only little insight into the solution of another one. This statement does not imply that an experienced investigator does not have a useful subjective intuition toward image analysis problems, but, rather, that he has no general principles to use as a guide in solving these problems.

In order to discuss image processing systems it is useful to consider the structure given in Figure 1. This figure should be self-explanatory; however, some comments on items not indicated in the figure are appropriate. At the University of Missouri-Columbia our images are scanned with an image dissector camera under control of an SEL 840A computer. The later processing is performed on an IBM 360/50 computer.

In stage 1, the images are digitized by an image dissector camera. This camera records the transmittance of light which passes through the x-ray. Since the human eye views the light intensities as the logarithm of the transmittance, we use a log amplifier to calculate the logarithm of this transmittance. These values are stored on tape as the picture points. Thus, the images stored in the computer are similar to the human visual perception of these images.

In the preprocessing stage, stage 2, one can usually devise methods, such as filter routines and other preprocessing routines, that are adequate, but not optimal, for the later stages of processing. In some of our studies on chest x-rays we have found that we need no preprocessing. In some of the studies in nuclear medicine a simple averaging technique suffices, which reduces the data to a resolution of 32 x 32 approximating the resolutions of the input camera.

The difficult stages in the processing are stage 3, feature extraction, and stage 4, pattern recognition. The pattern recognition phase can usually be done if one has an adequate feature extraction stage. Therefore, we believe that the crux of any image analysis system is the feature extraction stage which is the primary concern of this paper.

Feature Extraction

To repeat: we are interested in developing a generalized feature extraction method that will apply to a large number of difficult picture classes. Specifically, we will consider a standard chest x-ray--AP view although the method applies to other picture classes equally well. Time and space limitations, however, prevent discussion of the latter examples. Figure 2 shows such an image, one of a complicated class of pictures. It is characterized by intersecting and overlapping objects. The objects, such as the heart and lungs, can be of varying sizes and shapes. A high degree of resolution may be needed to detect some of the abnormalities that might occur, such as rib notching, small nodules, and shadows, which are of significance in detecting heart abnormalities. The question then becomes the following: is there any general approach that one can apply to a complicated set of images such as this? In an attempt to answer it, we have been developing a method for feature extraction, whose description follows.

A fundamental concept behind this method is that one must first consider the large objects in an image and, once these objects have been sufficiently recognized, proceed in the analysis of their details. For example, one first locates a

***This research was partially supported by USPHS Grant GM 17729-01 and Grant GM 01738-04.

large object (i.e., lung), then searches further for smaller objects within the lung (i.e., ribs and vascular or bronchial branches). Another concept is that one must have a description, within the program, of the class of input images before one can successfully analyze that class of pictures. Most programs that presently process images for recognition purposes as contrasted with filtering programs, have an implicit description of the class of images to be analyzed contained within the program. The description is often spread throughout the program and not isolated and designated as a description of the images.

We believe that recognizing first the large objects, then their details has several advantages. First, one should have a much greater possibility of locating the large objects reliably. For example, in Figure 2, if one were to first begin searching for the ribs, one would most likely encounter severe problems, caused by the intersecting lines. These problems would be compounded if the lungs had increased vascularity, because some of the ribs would blend into the vascularity. What we are really saying is this: we believe that it is not feasible to analyze this type of picture by using local algorithms which perform calculations on only a small area of the picture (such as line identification or contour trace programs) because any rule that will allow faint objects (such as the rib outlines) to be present will also allow a vast array of other lines to come through the process. Moreover, it is probably not possible to develop clever enough local logic to sort the ribs out of a picture of this type. In other words, it is difficult to develop logic to recognize a small object in an image without first recognizing the region which encompasses it. However, if one proceeds with a search for large, then smaller objects, one has a chance to identify the finer details of the large object. In Figure 2 one might first look for the large objects such as the lungs and heart which one has a good chance to find, probably at a coarse resolution which should speed processing time. Once the large objects are located, one can repeat the process if one is interested in details of the object, but confine the search to the area of the object, i.e., the lung, and discard the rest of the picture data. One could of course examine its details at a higher resolution than the one which searched for the object, because the area of search is much restricted.

As we have previously stated, if one is to analyze pictures, one must have some description of the pictures in the program. Since we want a program that applies to different classes of pictures, we have to have a way of describing many different classes of pictures to it. We will have more to say about this later.

Before we proceed with the details of our method, however, we might mention the following problem. Any feature extraction algorithm will have input parameters, and yet the correct parameter settings for identifying objects will vary because of uncontrollable factors such as lighting or poor adjustment in the photographic process. This means that no one setting of the parameters will suffice. Therefore, one must build into any useful program the automatic adjustment of all the parameters.

To summarize: we believe that one needs feature extraction algorithms which embody the following concepts: a) Proceed with a search from large to small objects. That is, the search is vertical. b) Utilize a structural description of the pictures to guide the search for objects. c) Automatically adjust all parameters. The following are details of a system that has these properties.

One idea that we have decided to use is that of a 'field of vision'. By this we mean that a picture is composed of objects (large) that completely fill the space in the picture. These items fit together in jig-saw fashion and comprise the gross objects in the image. For example, in a chest x-ray one might divide the chest into five gross objects (Figure 3): a) heart; b) right lung; c) left lung; d) right arm; and e) left arm. We believe that this concept will aid in the problem of automatically adjusting parameters. For example, if one is attempting to identify the points located in an object, such as the left lung, the question is how one decides automatically where this object terminates.

One solution is to find that part of each basic region (e.g., lung or heart) where the grey level remains constant within that given area, but leave the boundary points unclassified since the relationship between its grey level and the region's grey level is high in variance. Later after the average grey level of each region has been adjusted because of the addition of new points (as not all region points are classified the first time), boundary points are placed in that region which matches their grey level best. Since each object has a definite boundary in which all of its visible points are contained, we can say that one object terminates where another one begins. Visible points are those points which are not hidden behind an overlapping object, i.e., the overlapping object conceals what is behind it. For example, the left lung is surrounded by the heart on the right and left arm on the left. Part of the left lung is hidden by the heart. We identify the left lung as that region bounded on the right by the heart and on the left by the left arm. Thus, we can say that one object terminates where another object begins.

As we have stated before, we need a way to describe the class of pictures that we are considering. In addition, we want to proceed in the analysis by means of a vertical search. We have decided to use a graphical method in this description. So that the same program might apply to different picture classes, the graphical descriptions will be read in as data. If one uses a tree structure, then one gets a natural ordering of the picture from gross objects to their details. For example, consider Figure 3, which is a schematic drawing of the chest, and Figure 4, which gives a vertical ordering for the objects in the chest. It should be noted that the vertical ordering may correspond to natural objects, the anatomy, but the real criterion is a system that is computationally natural. Therefore, we see that the graphical description implies the vertical ordering. We will now describe the manner in which we use the concept of a 'field of vision' and a graphical description of the class of pictures to guide in the extraction of the features at a particular level in the tree. For concreteness we shall refer to the class of chest x-rays and the ordering given in Figure 4. We will consider the level containing the lungs, arms, etc.

Restated, the problem is: Given an area of a picture and a description of the objects that comprise this part of the picture, find a way to identify each object by finding all its points. Our method is as follows. Each node on the graph represents an object in the picture. Attached to each node is a list of attributes for that object. The list of attributes describes the object identified with the node. This description might include facts as: a) the average grey level or average local histogram; b) average number of points; and c) shape descriptors such as higher order moments.

Consider the node which contains the attributes for the left lung. The two items on this list are the expected average grey level and expected average number of points. Both attributes are calculated prior to the running of the algorithm, and the calculation is based on the resolution of the picture. For example, the larger the resolution, the more points in each region and the more accurate the average grey level. Of course, the attribute list is altered throughout the algorithm. For example, the grey level of this lung might be set at 40, but after the first pass the average might be calculated to be 38. A constant update on the average grey level and number of points found is kept.

If one is to identify the points in an object, one needs a method for examining the picture data points for this purpose. We use a method that enumerates the points in a region by starting with one initial point SOR (start of region point) and then examines the neighboring points to see if they share a common property P. In our algorithm the property P is defined as similar grey level. If the neighboring point has the property P, then it is placed in the region with SOR. The process is then repeated until all the contiguous points starting from SOR which have the property P are placed in the region. In our present program, whether or not a point has the property P is determined by an edge detector. The value of the average grey level, called INT, of the region is recomputed every time a new point has been added to this region. As a new point x = PIC(I,J) is considered as a candidate for the region, a local histogram is computed about x and then a number VAR is computed which is a measure of the deviation of the histogram from INT. If the absolute value of VAR is less than DELTA, then x has the property P. DELTA is an input parameter to the program that enumerates the region points and the property P holds if point x is not on a transition boundary of the object. We call programs of this type REGIONEN programs. It should be noted that the larger DELTA is, the more points will be included in the region. Thus a problem is how to adjust DELTA to get the exact points that lie in the region. It should be noted that we need the exact boundaries of objects because, in order to determine medical diagnosis, we must not only decide where the heart, for example, lies, but also determine its exact size and shape. In order to make these adjustments of DELTA, we associate a parameter list with each node of the graph. This parameter list contains the following items:

1. NI1 - Minimum value allowed for I in searching for the SOR point in the search for an object identified with node.

2. NI2 - Maximum value allowed for I in searching for the SOR point in the search for an object identified with node.

3. NJ1 - Minimum value allowed for J in searching for the SOR point in the search for an object identified with node.

4. NJ2 - Maximum value allowed for J in searching for the SOR point in the search for an object identified with node.

5. DELTA1 - The first setting of DELTA.

6. DELTA2 - A parameter used to speed the search for SOR.

7. CHDELTA - The amount by which DELTA is incremented.

8. KMAX - The maximum allowed incrementation of DELTA.

9. FC - A counter giving the present level for DELTA.

10. LOOKD - Governs the rate of spread of a region linked to node.

As an example of a parameter list, let us use the node associated with the heart. In order to find the heart one has to know where to look for it. The first four items in the parameter list describe the rectangular area of the chest x-ray where the search for the heart is to begin. Of course, the heart is not confined to this rectangle; it is just a starting place and the search for more heart points expands past these bounds. The first point placed in the heart region is critical. Since the grey level is not constant over the whole object, a higher or lower average grey level can result if the wrong start point is chosen. For example, the heart has grey level of 2 or 3 near the diaphragm and 14 or 15 in the mediastinum (upper portion of heart). If the start point has grey level of 14, then probably that portion near the diaphragm would not be accepted as part of heart region. DELTA2 controls the acceptance of a point as being the start point. DELTA1 holds the first value for DELTA. DELTA controls the acceptance of a point as a member of that region. For example, let DELTA be 3 and average intensity of the heart retion be 5. If a point under consideration has a grey level of 6, then the variance of the average grey level and point would be 1 and thus, the point would be accepted since the variance is less than DELTA. CHDELTA is a fixed amount used to increment DELTA. DELTA is incremented so that more points may be accepted into the region. The reasons for accepting more points are two. First, in the initial linkage a link between node and region could be rejected because of too few points found. Second, our goal is to classify each point as member of a region. During the initial linking of a region with node, KMAX is the maximum number of times DELTA may be incremented. The purpose of KMAX is to provide an exit in case the object with the given attributes can not be found. Finally, LOOKD is used in the structure tests. It is the maximum distance in some direction one looks to find a particular grey level.

In addition to the attributes and parameters associated with each node of the graph, we also have structure tests. These are in the form of propositional calculus statements and give information as to the structure of the class of pictures. These tests would give information such as a) the right lung is to the right of the heart or b) the right arm is to the right of the right lung. This completes the information associated with the graph which describes the class of pictures being studied. For a different class of pictures the graph would change. The graph should be considered as data for a picture analysis supervision. We will now describe the present operation of the picture supervisor.

The feature extraction program is divided into three stages of operation. The first stage is given the task of making an initial linkage of the objects in the picture with the graph nodes. It does this in the following manner. The nodes on a given level have consecutive node numbers associated with them, NODEN1 holding the low node number and NODEN2 the highest. Each of these nodes is interrogated in turn beginning with NODEN1. Node is the current node under consideration. The parameters of NODE are given to a REGIONEN program. These parameters include an initial DELTA and an area of search for the object. All of the regions within the area are enumerated. The attributes (for example, number of points and average grey level) of each of these regions are tested against the attributes of NODE. The region, called REGION, whose attributes best match those of NODE is tentatively linked to NODE. The match is tested to see if it is close. It is possible to have a best match and still reject it because the number of points in the region is too small or grey level is not close enough. If the match is not close, the link is removed and FC is either increased or decreased, which effectively increases or decreases DELTA, depending upon how the number of points found in REGION compares with that expected for NODE. Another attempt will then be made to link a region with NODE. It should be noted that at this stage we are seeking only the coarsest links. For example, we are trying to merely distinguish the lung from the heart, not get the exact boundaries. Each node from NODEN1 to NODEN2 is considered, until a tentative link is made to some region in the input picture. DELTA1 will normally be set so that one is likely to make a link from REGION to NODE where REGION has fewer points than actually should be in the region within the picture which accurately reflects the object associated with NODE. We do this because we intend to have the objects grow toward each other to get the accurate boundaries between the objects. It might also be noted that in the first stage of processing a check is made to determine if any of the linked regions intersect. If they do, then the links are removed and the parameter FC for each affected node is decreased.

Stage 2 of the processing consists of increasing FC in steps of 1 until KMAX for the node is achieved or until the REGION linked to NODE intersects some other region. When either of these events occurs, NODE is placed on an inactive list. When every node is inactive, then stage 2 is completed. Stage 2 therefore consists of expanding as much as possible the regions initially linked to the nodes. Upon completion of stage 2, each REGION will be enlarged; however, they will not, in general, fill the entire area under consideration, which is one of the criteria that we have established. A third stage of processing completes the algorithm.

In stage 3 each point in the area of interest, say x, is examined. The node regions that are visible from this point are determined. A node region NODE is visible from x if no other region is between x and the NODE region. The point x is placed in one of the visible regions whose grey level attribute most closely matches that of x. In the present algorithm we use the same program that we already used to verify the property P in the original enumeration of the regions. Namely, we compute the variance of x for each region using a local histogram about x and the average intensity of each region. The region whose variance is the smallest is selected as a candidate region that point x might be placed in, called NODE for reference. At this point the structure test for NODE must also be passed before x is placed in the region NODE.

To summarize: all the neighboring regions of point x are examined. Out of these regions the one that most closely matches the grey level properties of x is selected. If the structure tests are passed, x is placed in the region. In our example of chest x-ray pictures, the structure tests for the left arm region, for example, would require that the left lung lie always to the left of the left arm. If point x is considered and NODE is the left arm node, then x must be to the right of the left lung or this test fails and x is not placed in the left arm region. The structure tests are easy to write and are very valuable. In our example pictures, the grey shades of the left lung will tend to be typically about 45 and those of the heart about 10, while those of the left arm will be around 30. If one did not have structure tests of the sort we described, the right arm regions would tend to wrap around the lung into the heart and mediastinum. The structure test effectively prevents this. Stage 3 is repeated until every point in the area of interest is placed in one of the regions.

Some examples of this program are given in Figures 5, 6. We believe that they demonstrate the feasibility of the program as a method for identifying the objects in a complicated class of pictures. However, at this stage of development the program is by no means perfect. An analysis of Figure 6 shows that the heart is fairly well outlined. One problem is that the region including the left arm protrudes too far into the left lung. This problem, we believe, is caused by the fact that in Stage 3 of the processing, the local properties of a point x are compared against the average intensities of the neighboring regions before x is placed in one of these regions. The left lung region is characterized by grey levels in the upper lobe that are around 50 and decrease to around 20 at the tip of the lower lobe. The grey levels along the lower lobe will therefore look more like those of the left arm region. A better method would probably be not to use the average grey level for the selection of the region to put x into, but, rather, consider the characteristics in the lower lobe for the classification in this area. Note that in the upper lobe we have a good boundary for the left lung.

Conclusion

We have presented a concept that, we believe, will form a basis for identifying objects in complex images. The concept is that the search for objects should be a top-down or vertical search. First, one must find large objects, then their details. The search should be guided by a structural description of the class of pictures. In addition, we believe that the concept of a 'field of vision' is an aid in object identification. That is, an entire area of the picture is completely filled by objects that must be identified. An example of a working program that embodies these concepts is discussed. The program demonstrates feasibility and is a good vehicle for continuing research in image analysis problems.

Bibliography

1. J. A. Hawkins; Image Processing: A Review and Projection; Automatic Interpretation and Classification of Images, A. Grasselli (Ed.); Academic Press, 1969.

2. Azriel Rosenfeld; Picture Processing by Computer; Academic Press, 1969.

3. Azriel Rosenfeld, E. B. Troy; Visual Texture Analysis; Proceedings of the UMR-Mervin J. Kelly Communications Conference, Rolla, Missouri, October 1970.

4. Azriel Rosenfeld; Figure Extraction; Automatic Interpretation and Classification of Images, A. Grasselli (Ed.); Academic Press, 1969.

5. Craig Cook; Size Detectors; Technical Report 70-127; University of Maryland Computer Science Center, August 1970.

6. Franco P. Preparata and Sylvian R. Ray; An Approach to Artificial Nonsymbolic Cognition; Coordinated Sciences Laboratory Technical Report, UILC-ENG 70-223, July 1970; University of Illinois; Urbana, Illinois.

7. M. B. Clowes; Transformational Grammars and the Organization of Pictures; Automatic Interpretation and Classification of Images, A. Grasselli (Ed.); Academic Press, 1969.

8. J. L. Muerle, D. C. Allen; Experimental Evaluation of Techniques for Automatic Segmentation of Objects in a Complex Scene; Pictorial Pattern Recognition, Cheng, Ledley, Pollock, Rosenfeld (Eds.); Thompson Book Company, 1968.

9. L. Uhr, Ed.; Pattern Recognition; Wiley, 1966.

10. M. K. Hu; Visual Pattern Recognitions by Moment Invariants; IRE Trans. Information Theory; February 1962.

11. Nils J. Nilsson; Learning Machines; McGraw Hill, 1965.

12. G. S. Lodwick, C. L. Hann, W. E. Smith, R. F. Keller, E. Robertson; Computer Diagnosis of Primary Bone Tumors; Radiology 80, 273, 1963.

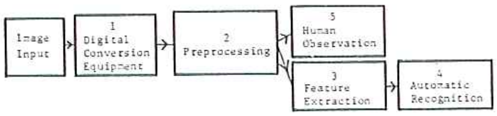
13. G. S. Lodwick, Arch Turner; Computer Aided Diagnosis of Primary Bone Tumors; Department of Radiology, University of Missouri; December 1965.

14. Cheng, Ledley, Pollack, Rosenfeld (Eds.); Pictorial Pattern Recognition; Thompson Book Company; 1968.

15. Lipkin and Rosenfeld; Picture Processing and Psychopictorics; Academic Press, 1970.

16. Tippett, Berkowitz, Clapp, Koester, Vanderburg; Optical and Electro-optical Information Processing; MIT Press, 1965.

17. P. Naur (Ed.); Revised Algol Report; Comm. ACM; January, 1963.



1. Digital conversion equipment--The camera and associated analogue to digital conversion equipment required to convert the image into an n x n array of points whose entries give the brightness value of the image.
2. Preprocessing--Programs used to suppress the noise put into the image data by the digital conversion equipment and also programs used to emphasize certain properties in the image that are more important than others for the later stages of analysis.
3. Feature Extraction--The extraction from the image data of information to be used in the automatic recognition of objects in the image. The data emitted from stage 3 of processing should be greatly reduced from that of stage 2 of the processing.
4. Automatic Recognition--The automatic recognition of the important objects in the image by the computer.
5. Human Observation--A human interpretation of an improved image.

FIGURE 1.

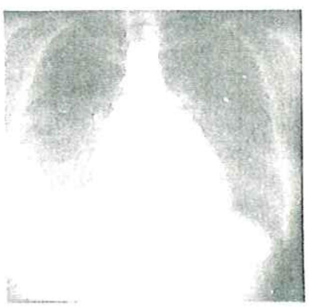


FIG. 2A--ORIGINAL CHEST X-RAY

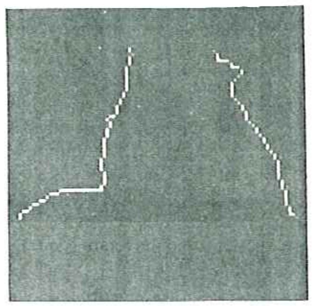


FIG. 2B--TYPICAL CONTOUR TRACE OF THE HEART

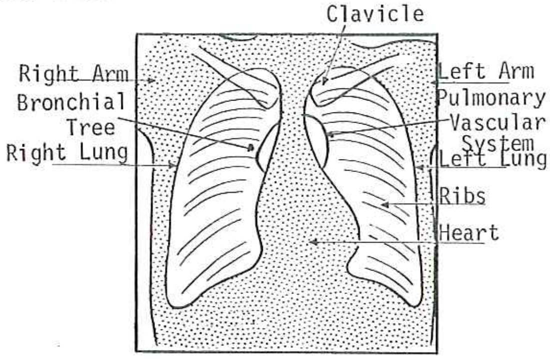
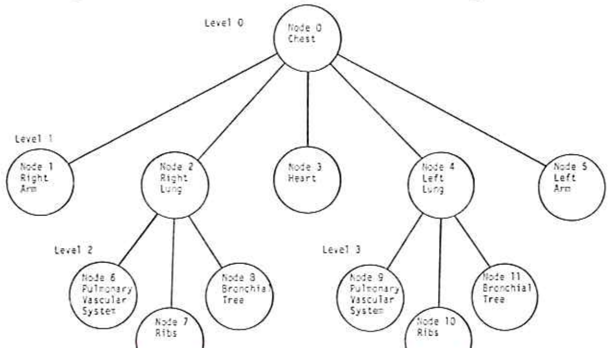


FIG. 3--ANTEROPOSTERIOR VIEW OF CHEST (The Bronchial Tree and Pulmonary Vascular System branch throughout each lung.)



Attribute list for each node consists of

a) grey shade (level) of node
b) expected number of points for node

Parameter for each node consists of

a) vicinity to look for object
b) first DELTA setting
c) value used to speed the search for start point
d) maximum number of increments for DELTA
e) counter for number of times DELTA has been incremented
f) look distance for structure test

FIGURE 4. PARTIAL VERTICAL ORDERING OF A CHEST X-RAY PICTURE

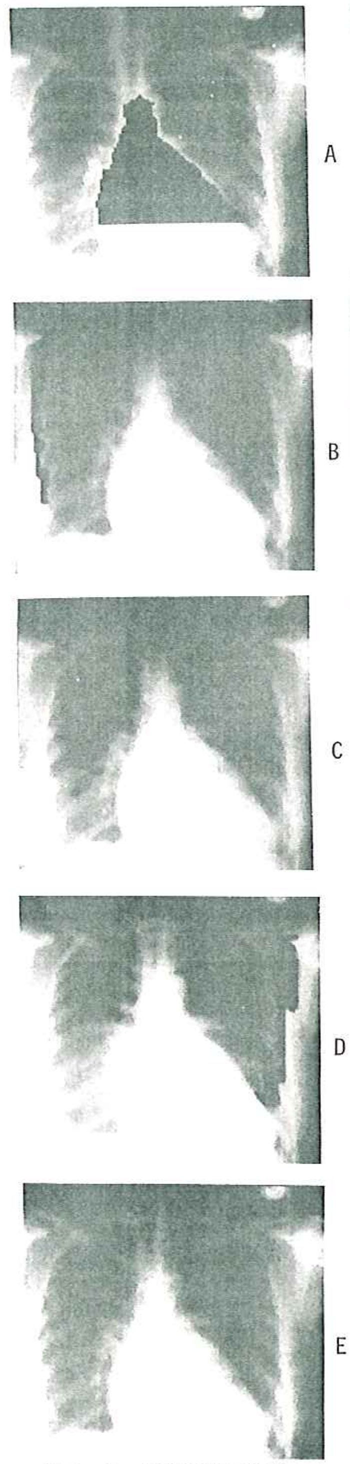


FIG. 5--INTERMEDIATE LOCATION OF:
A. HEART REGION
B. RIGHT ARM REGION
C. RIGHT LUNG REGION
D. LEFT ARM REGION
E. LEFT LUNG REGION

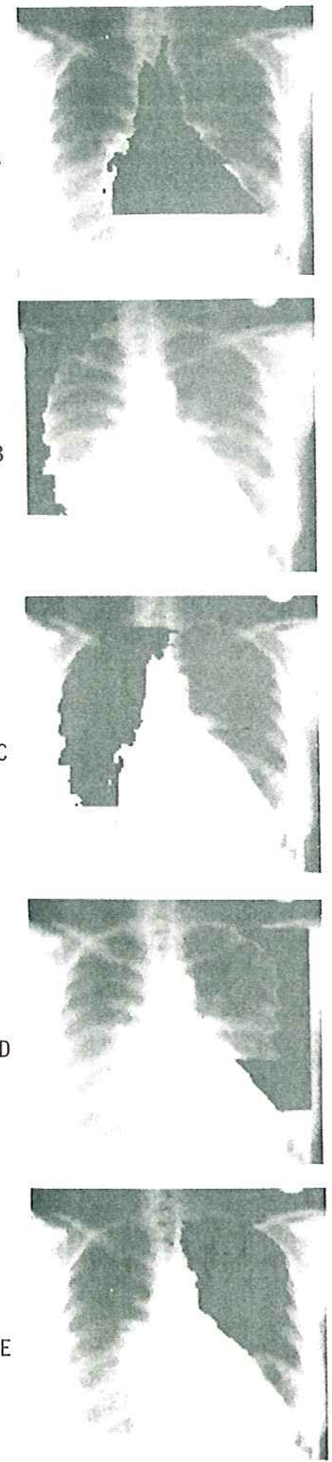


FIG. 6--FINAL PROCESSING OF:
A. HEART
B. RIGHT ARM
C. RIGHT LUNG
D. LEFT ARM
E. LEFT LUNG

1971 IEEE SYSTEMS, MAN AND CYBERNETICS GROUP

ANNUAL SYMPOSIUM RECORD

OCTOBER 25-27, 1971
ANAHEIM, CALIFORNIA

IEEE CATALOG NO. 71 C 46-SMC