DIGITAL IMAGE ANALYSIS

OF A CLASS OF GASTRIC RADIOGRAMS

Shigehiro Fukushima

January 1983

DIGITAL IMAGE ANALYSIS OF A CLASS OF GASTRIC RADIOGRAMS

Shigehiro Fukushima

January 1983

Department of Electrical Engineering
Kyoto Institute of Technology
Matsugasaki, Sakyo, Kyoto 606

Submitted to Kyoto University in partial fulfillment
of the requirements for the degree of Doctor of Engineering

ABSTRACT

This dissertation presents a number of approaches to computerizing the interpretation of a class of gastric radiograms. This study was motivated by the need of such facility in mass screening which is conducted throughout Japan to find people diseased in the stomach and to reduce the high mortality due to gastric cancer. Since a large number of gastric radiograms are taken routinely in gastric mass screening, an increased emphasis has been placed on their automatic interpretation. Contents of gastric radiograms depend on the dose of contrast material and the position of the subject, and therefore more than six kinds of radiograms are taken to obtain complementary information. Among these images, the most important image class is of the standing position-anteroposterior (SPAP) image. This is because an SPAP image gives fundamental information in the gastric radiography. In the remaining part of this dissertation, attention is focused on this image class.

Firstly, a conceptual scheme of diagnostic digital image analysis is introduced. An image processing technique was developed in order to detect the gastric contour by computer, since its detection is a most fundamental task in computerized interpretation of SPAP images. This technique is mainly based on a dynamic threshold concept as well as other methods of binary image processing such as labeling, distance transformation, thinning, and inverse distance transformation. A software system was designed for ease of modification by users by installing newly acquired pieces of knowledge and was implemented both on a large computer system and on a small one.

Secondly, this dissertation discusses the feature extraction and the pattern classification of SPAP images. For this purpose, two different approaches were examined, *i.e.*, the descriptive method and the quantitative method. But, the quantitative method is discussed in detail. The discriminant analysis with the feature selection has proved usefulness of some geometrical measurements and Fourier descriptors which were obtained from the gastric contour. Combination of several features attaines best discrimination with 7.5 percent error according to the leaving-one-out method of the error rate estimation. The discriminant analysis was also applied to the problem of discriminating abnormal pattern of the apex region, a portion where diseases are frequently detected. In this experiment, features were extracted by two-dimensional orthogonal transforms. Several features were selected by the discriminant analysis, resulting best discrimination with 6.4 percent error by the leaving-one-out estimation. Distribution of the samples and the similarity of the features were also inspected by the clustering. The results of this method supported those of the discriminant analysis. The two experiments denied the validity of the principal component analysis in the feature selection, or the dimensionality reduction of the feature space, for the pattern classification purpose.

In conclusion, a promising perspective was obtained by this study for the computerized analysis of SPAP radiograms of the stomach.

ACKNOWLEDGEMENT

This study has been supervised by deceased Professor Takashi Soma from its beginning. The author is very grateful to him for his constant guidance and encouragement. The author expresses his sincere gratitude to Professor Michiyoshi Kuwahara for his warm and invaluable advice. Particularly, the author is indebted to him in preparing the manuscript. The author also expresses his many thanks to Associate Professor Sigeru Eiho for his helpful discussion. The author's many thanks are also attributed to collaboration of Professor Keiichi Kawai and his gastro-intestinal research group from the Kyoto Prefectural University of Medicine, particularly, Dr. Kazunori Ida (currently with Gifu Dental University,) Dr. Yuzo Akasaka, Dr. Fumio Misaki, and Mr. Kyohei Hayashi. The author must express his special gratitude to Professor Bunji Kondo for his generosity. This study was conducted partly at his laboratory at the beginning and has been encouraged by him.

CONTENTS

CHAPTER 1 INTRODUCTION

1.1 Pattern Recognition in Biomedical Image Analysis

It has been more than two decades since the advent of the concept of pattern recognition by computer with particular interest of its application in science and technology. In image pattern analysis, character recognition was a major subject in the early days. Today, techniques have grown up to handle a variety of classes of image patterns such as images of industrial parts, human faces, and finger prints, visual scenes, weather maps, areal photographs, satellite images, microscopic images, and biomedical images. Approaches to analyzing most of these images are rather different from one to character recognition. In character recognition problems, an *a priori* set of distinct symbols exists for the objective language. These are source symbols in the sense of Shanon's communication model. On the contrary, in natural image recognition problems as listed above, the property of the information source is not well-defined in general. It is even doubtful sometimes whether there is any definite communication source. For example, in medical image recognition, a major concern is to identify whether photographed organs or tissue is normal or diseased. However, usually, consistent definition about normal patterns and diseased ones is not given in advance. There may be even cases that the distinct boundary between normal patterns and diseased ones is difficult to define or does not exist. Therefore, it is necessary to analyze the concerned class of images in order to obtain knowledge for designing a dedicated system for pattern recognition. The analysis includes inspection of assumable prototypes and features prescribing them.

Although some kinds of knowledge are known through the analysis conventional in the related field or understood heuristically, they must be translated into a form of representation suitable for manipulation by computer: for example, quantification and symbolic description are required in the statistical analysis and in the syntactic analysis, respectively. In order to find an appropriate form of representation, properties of objective images must be investigated. The study on natural pattern recognition inevitably emphasizes the analysis and understanding of specific properties of the class of objective images.

As for pattern recognition of biomedical images, more problems are involved in image analysis:

1. It is difficult to obtain sufficient number of samples for developing methods. It is particularly obstructive in the study of pattern recognition that verification of the diagnosis given to each sample is difficult or impossible. This is particularly claimed for normal samples because live bodies must be tested.
2. Some of the samples may be singular because the samples are usually obtained at hospitals.
3. The knowledge invoked by physicians in the diagnosing process is almost unbounded with various degree of importance. Methods must be developed to install such diverse knowledge.
4. Mass screening is a major front of clinical application of pattern recognition. However, the diagnosis in mass screening depends on the policy determined by the estimated proportion of the diseases and the cost of screening and precise examination.

These aspects cause uncertainties in making the framework of each problem. Multiple phases of building-and-testing process are required

to clearly establish the problems by assuming prototypes and features.

1.2 Background and History of the Research

In the last decade, there has been a considerable growth in a wide range of biomedical image processing[1,2]. Three advantages may be gained by digital methods in this field:

1. When the quality of an image is poor, computer can restore a fine image for better observation.
2. Computer can readily present information which is difficult to obtain or to measure manually.
3. In mass screening or in examination, computer may be a substitute for a human worker as an automatic interpreter of images.

Biomedical images and methods for their analysis can be classified in a number of ways:

1) Subject organs or tissue: lung, heart, stomach, brain, bone, eye-fundus, cell, liver, kidney, chromosome, bloodcell, etc.
2) Imaging: X-ray, gamma-ray (RI), microscope, ultrasound, etc.
3) Content: morphological information or functional one.
4) Mono frame or multi frames.
5) Still object or moving object.
6) Textural differences or those due to contrast.
7) Purpose: automatic interpretation or interactive analysis.
8) Approach: data acquisition and display; signal processing for image restoration and enhancement; or pattern recognition.

The research for the computerized analysis of gastric radiograms has been started in Japan because of the high mortality due to gastric cancer[3]. Because the most efficient means to reduce the mortality is

the early diagnosis and the early treatment, mass screening is conducted throughout the country to find diseased people. Although several examination techniques exist to detect gastric diseases, the most powerful and convenient aid in mass screening is indirect radiograms taken through the fluoroscope. In the usual routine of the indirect radiography, several images are taken of each individual in various positions. Combination of more than six different kinds of images is recommended by the Japanese Society of Gastric Mass Survey, a promoting society of gastric mass screening[4]. In addition, more than three million people receive gastric mass screening every year. Therefore, considerable work load is required of physicians in interpreting the images, increasingly emphasizing the computerized diagnosis.

From early seventies, attention has been mainly focused on the standing position-anteroposterior (SPAP) image because of the most frequent detection of symptomatic changes and because of the simpler densitometric structure. The gastric contour retains most of the diagnostic information in this kind of images.

The first concern has been the extraction of the gastric contour by image processing. A strategy of the computerized interpretation of SPAP images was conceptualized first by Soma *et al.*[5,6]. They studied the contour extraction in a limited portion, the apex, and the diagnostic feature extraction from manually traced contour samples.

Rather optimistic steps were taken earlier. It was believed that dedicated hardware would be developed soon on reflection of character recognition machine. A class of approaches to designing hardware was the application of a spiral reader, a flying spot scanner controlled to track the object boundary by detecting abrupt change of image intensity.

Miyawaki[7] extracted the lesser curvature of a stomach by this method and depicted it on a CRT. Akatsuka *et al.*[8] designed a sophisticated control circuits for a flying spot scanner. This scanner was controlled by a programmed small computer PDP-12 equipped with a small core memory. He reported that less than one-kiloword (12 bits/word) storage was enough to memorize the outline of a stomach with additive information such as gray levels, if a proper coding like Freeman's coding and packaging is used.

Mori *et al.*[9] took another approach to flying spot scanner design. Their machine incorporated two scanning modes. The main scanning, a common manner of flying spot scanning, was controlled by a small computer OKITAC-3400. This main scan accompanied a sub-scanning controlled by an electronic circuit. In the sub-scanning, the raster scanned circularly a small disk area by changing the scanning diameter and yielded a set of digitized samples. This scanner was used to extract gastric contours using histograms of the gray level occurrence in local regions. They also measured some metrics after compensating noise such as due to the photomultiplier and to the film granularity by taking a local average of the gray level. A problem of the dedicated hardware approaches was the difficulty in incorporating the overall information in the control in order to prevent the contour tracking operator from tracking irrelevant misleading edges.

On the other hand, Soma and Fukushima[6] and Fukushima *et al.*[10] extracted the shadow of a stomach using a dynamic threshold method and segmentation techniques before tracking its contour. In their method, overall property of the gray level distribution was used as well as local property, thus stabilizing the shadow extraction and the contour

tracking.

Application of other algorithms for the contour extraction can be found in a feasibility study by Sugimoto and Uehara[1]. They presented results of equal level tracing, gradient tracing, thresholding, and dynamic thresholding which they call a region-divide method. Nakamura *et al.*[12] used contour-tracking operators different in various portions of the gastric contour and detected the contour by local thresholding.

Although a contour detected by these methods may not be strictly precise, the precision of the contour extraction is required only in certain portions in diagnosing the stomach. A requirement in pattern recognition is that the contour must be represented in a form appropriate for machine manipulation, for example, a representation form from which machine can determine the location of each anatomical position. One representation may be to depict a contour as one depicted in schematic sketches in medical references, particularly, textbooks. However, the sketches do not always correspond to the photographic density of a radiogram; they are drawn by also invoking the knowledge of physicians. Other difficulties in detection of the contour arise from diverse shapes of stomachs, the variation of the background gray level, noise, non-uniform distribution of contrast material, and the overlap of shadows of other organs, such as the duodenum and the intestine. Since useful information is not clearly known for computationally diagnosing a stomach, the required precision and form in the contour extraction is not clear, either. Therefore, the contour extraction study has a close tie with the diagnosing study.

The computational diagnosis, the second concern of the research, includes not only determining the best classification rule but also

finding or identifying numerically tractable features which are useful in classification. Generally, three benefits can be expected from computerizing the diagnosis:

1) The improvement of the diagnosing accuracy by measurement and quantification.
2) The automation of the interpretation process for releasing persons from routine works.
3) The stable and clear-cut diagnosis excluding the effect of the psychological fluctuation of human mood.

As for screening gastric radiograms, the second and third benefits are the most important today.

Some approaches has been taken to the computational diagnosis problem. They can be categorized either descriptive or quantitative. The earliest approach is attributed to Soma *et al.*[5]. They used the curvature of the gastric contour according to the diagnosing process of physicians. Later, this method was refined by incorporating a branching logic based on thresholding[6]. A similar approach was taken by Akatsuka *et al.*[8] and by Sugimoto and Uehara[11] to describe the appearance of a contour.

In contrast to the descriptive methods, Mori *et al.*[9] used three measurements and designed a classification rule for fishhook shaped stomachs in order to discriminate diseased from normal ones. The abnormality included the shrinking of the lesser curvature, the deformity of the apex, and the profile niche on the lesser curvature. One of their interesting approach was "multicontouring" to extract profile niche on the contour. Sugimoto and Uehara[11] reported discrimination of high-tension stomachs from low-tension ones. Hatori

et al.[13] studied on pattern classification of the contour in the apex region using coefficients obtained by the least square approximation of the contour by polynomials. They also used width patterns and curvature patterns of the apex. Recently, they used the integrated curvature function[14], which is similar to the cumulative angular function defined by Zahn and Roskies[15], finding an interesting property of the function when used for description of the apex contour. They extracted several features from this function. Fukushima *et al.*[6] examined five geomerical measurements and twenty eight Fourier descriptors as candidate features for abnormality detection of the gastric contour. They also examined coefficients of two-dimensional orthogonal transforms of the gray level image of the apex region as candidate features obtained without prior judgement for the contour extraction.

Although most of the reports described above claimed promising results, it is nearly impossible to make comparative discussion of these methods because the examined data and their points of interest are different. It is much better, for the present, to examine a variety of methods rather than to select the best, for more approaches and examinations are needed in this study.

1.3 Synopsis of the Dissertation

The author has been engaged in the study on the computerized diagnosis of SPAP images since early seventies. The study has been including image processing for extracting the gastric contour; feature extraction from the gastric contour; and computational diagnosis of stomachs by descriptive and statistical analyses. This dissertation describes approaches and results of the study in somewhat chronological

order.

Chapter 2 briefly introduces the gastric radiography and explains the reason why SPAP images were selected as an object class of gastric radiograms. Important observations in SPAP images are also introduced here. Then, a concepturized scheme for digital analysis of an SPAP image are introduced to give an overall view of the study. This scheme was given to meet both the diagnosing process by physicians and pattern recognition techniques of the days the study began.

Chapter 3 shows some basic properties of an SPAP image and presents a technique of extracting the gastric contour from an image. This technique is based on a dynamic threshold method as well as other fundamental methods for segmentation such as labeling, connected component extraction, thining, distance transformation, inverse distance transformation, etc. This technique has two advantages: (1) existing or established fundamental methods can be used for image processing, and (2) misleading edges can be eliminated before tracking the contour by using global properties such as area and structural information. In this chapter, the concept of the gastric axis is also introduced, although its utilization is discussed later. The final section of this chapter is devoted to constructing a system incorporating the methods described above. It is emphasized that the work is not and will not be completed at anytime because new knowledge and methods will be found by researchers as the study grows. Therefore, it is more important to provide an understandable system than to fix dedicated hardware which might be even time-and-memory-saving. When a large computer is available, this requirement can be easily satisfied because the time and memory problem seldom occurs. For this environment, a system structure

is proposed. This system was also implemented on a small computer, taking advantage of a tool system for digital image processing, which was developed in the course of this study. This is briefly remarked.

In chapter 4, another important subject of the study, the feature extraction and the pattern classification, is discussed. Two kinds of information are assumed as inputs to classification programs: the gastric contour and the gray level image of the apex portion of a stomach. For gastric contours, which were extracted manually, two approaches were studied: a descriptive approach and a quantitative one. In the descriptive approach, the curvature of a contour at its every portion was used as a fundamental feature, which was evaluated by a branching logic using the gastric axis. In the quantitative approach, five geometrical measurements and twenty-eight Fourier descriptors were studied as candidate features. They were examined by the discriminant analysis based on the leaving-one-out method. By the feature selection, sets of only a few features were selected from among the candidate features. It was proved by the leaving-one-out method that contribution of the selected sets in the classification is compatible with that of physicians. For gray level images of the apex, two-dimensional orthogonal transform coefficients were analyzed as candidate features by the same discriminant analysis. Selected feature sets in this case also resulted in discrimination accuracy compatible with physicians. A validation study on the results of the discriminant analyses was made by the clustering of the samples and the features.

In Chapter 5, concluding remarks are given on the body of the research as well as orientation and problems of the research to be settled in the future.

CHAPTER 2 GASTRIC RADIOGRAPHY

2.1 Methodology of Radiographic Observation

The radiographic examination of a stomach consists of three methods: the fluoroscopy, the direct radiography, and the indirect radiography. The essential part of the examination, the fluoroscopy, is conducted to survey the motion and to find the deformity of a stomach. Sometimes, the cine fluorography is conducted to record complete information. The fluorography is also applied as a guide to make radiograms in proper view angles and to manipulate the human body. Direct radiograms are of considerable aid in bringing out deformities in relatively inaccessible locations, in showing small ulcer craters, and in giving a permanent record of details observed during the fluoroscopy[17]. They are taken at precise examination for hospital visitors and for those who marked suspicious findings at mass screening. Although the image quality, sharpness and details, is reduced in indirect radiograms, these are most widely used in mass screening because of their convenience. In the strategy of mass screening in Japan, indirect radiograms are the most important tool since a considerable number of subjects must be examined.

Gastric radiograms include images for observing the fold of the gastric wall after a small amount of contrast material, usually the barium-water mixture, is swallowed; images of a stomach filled with contrast material and air to expand the gastric wall; double contrast images, taken after putting aside the bulk of the contrast material, to observe the texture of the gastric wall whose shadow is enhanced by the remaining thin layer of the contrast material; and images of a filled

stomach which is pressed from outside of the subject's body to observe regions of interest. The position of a subject is also changed on a movable bed. Images taken in such a variety of ways depict different aspects and locations of the gastric wall. Gastric diagnosis is made based on the interpretation of all these images.

In 1974, the Indirect Radiography Standardization Committee of the Japanese Society of Gastric Mass Survey submitted a recommendation with respect to the routine examination in mass screening[4]. The recommendation prescribed eight kinds of radiograms taken of a subject in various positions: the prone position-gastric wall image; the standing position-anteroposterior image; the prone position image; the standing position-oblique view image; the supine position-double contrast image taken from frontal or oblique view. The committee recommended ten methods for taking six images selected from among those described above in proper order (permitting same kinds of images to be taken more than once.) The density and the amount of contrast material were also prescribed by the committee as well as portions to be depicted by, and purposes of, taking respective images. Such standardization is important not only for its primary purpose, the improvement of the screening efficiency, but also for computerizing the interpretation.

The shape of a stomach and the running pattern of the gastric wall are quite different from person to person and from time to time, due to many factors such as the peristalsis of the stomach, the physique, the age, the sex, and the psycological state of the subject. Therefore, it is impossible to strictly define the state of being "normal" on the basis of radiograms. If no abnormal findings can be detected by complete search through the radiograms, the stomach may be

normal. The abnormal state is implied by a lack of the harmony, in other words, the inconsistency of curves[18].

2.2 The Task Domain

To begin the research, the statistics on indirect radiograms were studied for understanding the role of the images and accessible approaches[5]. The most important image classes are of the SPAP image and of the supine position-double contrast image. The detection rate by the SPAP image is about thirty percent which is the highest rate. This image is also fundamental in the radiography, because it shows the position, the shape, the tonus, and the various deformities of a stomach. Therefore, two of the ten recommendations by the Indirect Radiography Standardization Committee[4] state that two SPAP images should be taken earlier and later instants in the radiography. Other images can depict diseases directly inaccessible by an SPAP image. However, since they depict only specific local regions, SPAP images must be used indispensibly.

The shape of a stomach can be determined from the SPAP image: the form of the fishhook, of the steer-horn, or of the intermediate. The fishhook shape shares more than eighty percent of the stomachs of Japanese people.

Frequent diseases of the stomach include ulcer, cancer, polyp, and gastritis. Soma *et al.*[5] investigated the frequency and the location of diseases detected by indirect radiograms. Gastric diseases were located more in the apex, the pyloric antrum, and the lower gastric body. Thus, most diseases were located in and around the apex region. Even when the apex was not diseased, its shape often indicated diseases

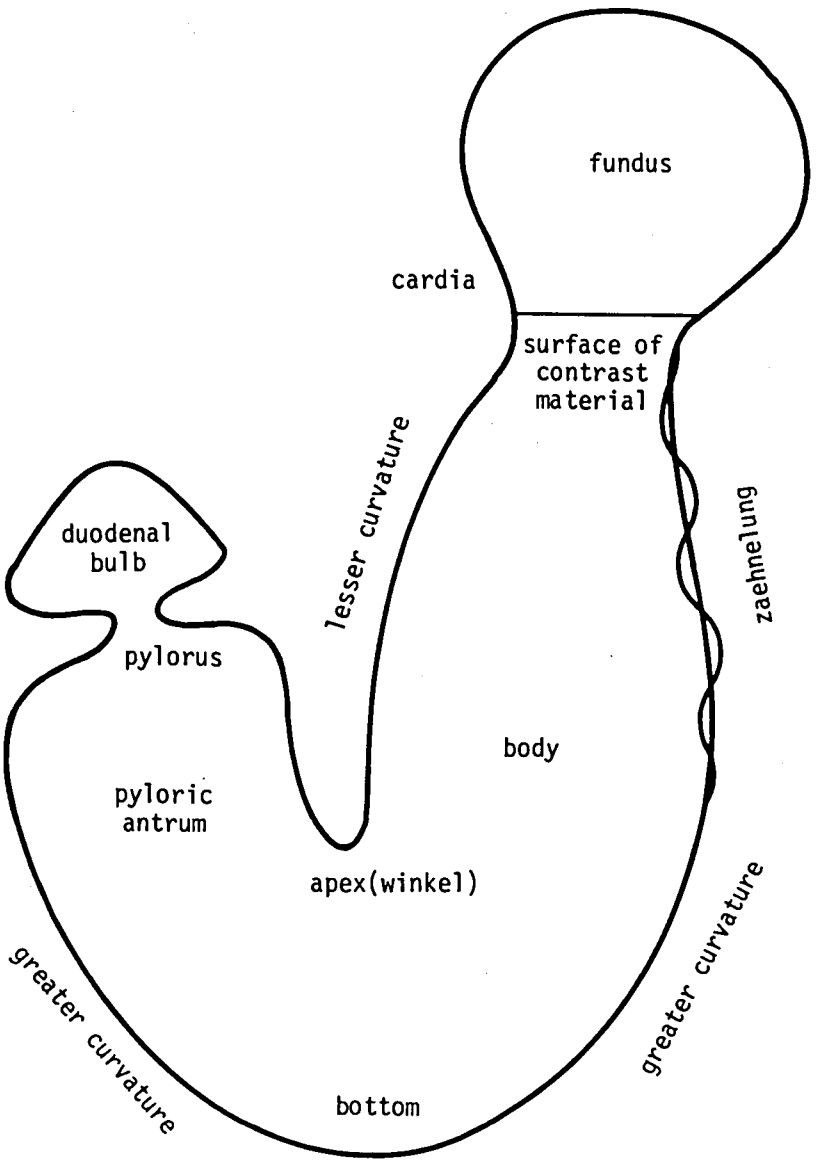


Fig. 2.1 The anatomical terminology of a stomach.

located in other portions.

They also studied abnormalities observed in indirect radiograms and their frequency. As a trend, advanced cancer and gastric polyp were suspected from filling defect; early cancer and ulcer from irregularity, rigidity, and niche on the gastric contour, deformity of the apex, and irregular fold.

2.3 Clues for Interpreting SPAP Images

Since SPAP images retain primary information, interpretation of images of this kind was studied. Because it is very important to know medical and radiological knowledge on the images, clues for interpreting the images are described here. Such knowledge may be incorporated in the design of the computer interpretation system, if possible.

Fig. 2.1 illustrates a pattern of a stomach projected on an SPAP image and the anatomical terminology.

The following is a summary of Chapter C of Part 3 of Ichikawa *et al.*[18]. Most clues may be categorized as either the deformity of the stomach or the local change of its contour. The pliability of the gastric wall is also reflected on such findings.

The deformity of the stomach includes the following:

1. Change of the proportion of the entire shape: in many cases, it can be an initial clue to find various abnormalities. This is because most diseases cause shrinking of the gastric wall and change the proportion.
2. Shrinking of the lesser curvature: shrinking of the contour from the apex to the pylorus, from the apex to the cardia, or of the whole of the lesser curvature.

3. Deformity of the gastric body.
4. Deformity of the apex.
5. Deformity of the pyloric antrum. Recognition of this deformity is often difficult because of insufficient filling of contrast material, the peristalsis, and the fold of the gastric wall.

The local change of the gastric contour includes the following:

1. Niche: the shadow of the ulcer-like hollow on the wall. It is the most easily identified finding. A kind of niche called "Shattenplus im Shattenminus" discriminates cancer from ulcer.
2. Filling defect, caused by fungating or infiltrating of cancer.
3. Rigidity.
4. Linearization.
5. Irregularity.

The rigidity, the linearization, and the irregularity are caused by the projection of a portion which has become rigid due to diseases.

They recommended a procedure of examining an SPAP image:

1. The proportion of the entire shape.
2. The shape of the apex.
3. The contour of the lesser curvature of the body.
4. The shape and the contour of the pyloric antrum.
5. The cardia and the pylorus.
6. The greater curvature and the fundus.

2.4 Computerized Interpretation of SPAP Images

From the discussion leading to this point, a system was conceptualized for interpreting SPAP images by computer[5]. Fig. 2.2 illustrates major procedures and linkage of the system. The system is

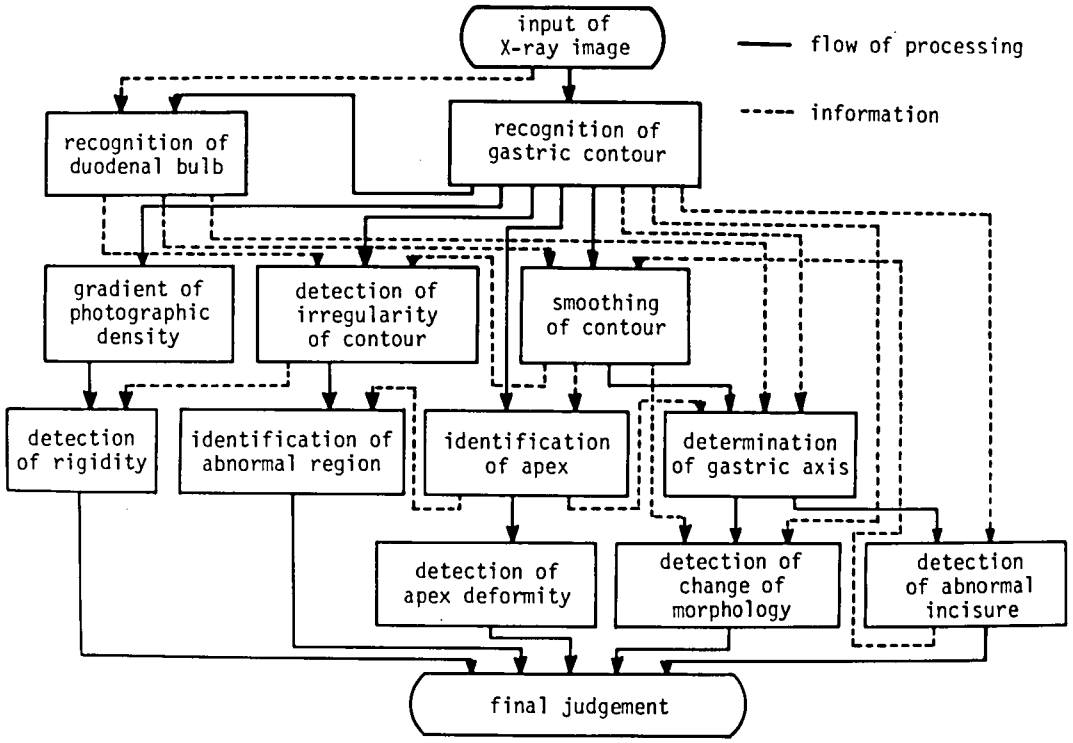


Fig. 2.2 A conceptual configuration for SPAP image interpretation.

composed of two major tasks: the extraction of the gastric contour from an image; and the interpretation of the contour. Although the contour extraction may be called preprocessing for diagnosing an image, it is not a trivial problem: even in those portions where contrast material is filled, an efficient method must be established; portions such as the fundus form only dim shadows due to absence of the material, making this problem more difficult one. Thus, the contour extraction problem makes an essential part of the system design for finding abnormal changes of the contour.

Studies on the interpretation of the contour involves identification of diagnostic features. Three tasks were introduced to computerize the interpretation: obtaining a smooth pattern of a stomach,

identifying the apex, and determining the gastric axis. The smooth pattern was introduced to represent the degree of the inconsistency by the deviation of a contour from its smooth version.

Importance of identifying the apex is well understood. It plays an important role in identifying various portions of a stomach because it is typically observable in SPAP images, particularly pronounced when the stomach is of the fishhook shape. The concept of the gastric axis was introduced for the first time in order to understand the correspondence between points on the lesser curvature and ones on the greater curvature. (It is a problem of the skewed symmetry.) By introducing the gastric axis, we hoped that abnormal incisures, or curvatures, could be discriminated from waves caused by the peristalsis: if two curvatures exist on both sides of the gastric axis in a symmetrical manner, they may be caused by the peristalsis.

Fig. 2.2 has been a guideline of the study as well as the recommendation stated lastly in the previous section.

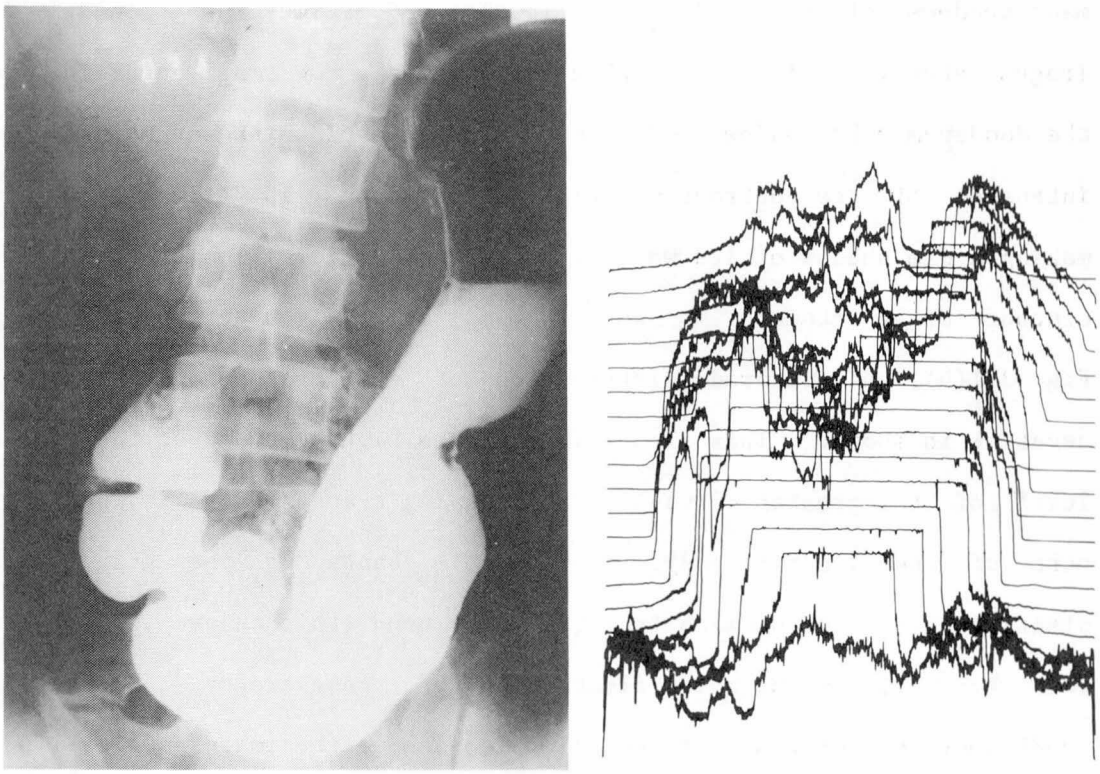
2.5 Summary of the Chapter

This chapter introduced the methodology of the radiographic examination of the stomach and the strategy taken in mass screening. An important role of the SPAP image was described. Frequent diseases of the stomach were also introduced as well as their radiographic findings. The apex was noted as an important portion because of the frequent detection of diseases. Finally, a configuration for computerized diagnosis was conceptualized as a guideline of the study. In this configuration, three tasks were postulated for computer interpretation in addition to those reflecting purely radiological tasks.

CHAPTER 3 DIGITAL IMAGE PROCESSING OF SPAP IMAGES

3.1 Properties of SPAP Images

Mainly due to the X-ray scatter and the radiographic mottle, the maximum resolution of the direct radiogram is about 0.2 mm on the image plane; usually, it is in a range between 0.3 mm and 0.6 mm. Direct radiograms involves dim stripes of the shadow of the lead grid which is used to suppress the X-ray scatter. The width between two adjacent stripes is about the same size. Human eyes also recognize the gastric



(a) (b)

Fig. 3.1 An example of SPAP images.
(a) The original photograph.
(b) Transparency distribution.

contour to the same precision. Therefore, the desirable sampling pitch for a direct radiogram is less than 0.2 mm in the image plane. The resolution of the indirect radiogram is five times lower in the object plane than that of the direct radiogram.

The study began with examination on direct radiograms because their image quality is more controlled. Fig. 3.1(a) is a typical X-ray negative of the direct SPAP image. Basically, the image consists of bright areas, which are made by the shadow of contrast material in the stomach, and the dark background areas. The backbone and the pelvis make shadows of intermediate gray levels. Sometimes, particularly in images taken later in radiographic examination, the small intestine and the duodenum make shadows of the same gray level as the stomach due to intrusion of the contrast material. The fundus is filled with air, making a dim shadow of its wall. Therefore, recognition of the entire contour of a stomach is not as easy as it might be expected. In Fig. 3.1(b), a measured distribution of the transparency of this negative is shown. This figure illustrates high contrast at the middle level of the greater curvature, decreasing transparency in the upper part of the gastric body, considerable change of the gray level discriminating the stomach from the background (this change is called the shading,) and the noisy fluctuation of transparency. Developing conditions and scratches of the film also change the image quality.

In order to understand properties of an image, it may be a shortcut to apply algorithms for digital image processing. In this study, copied 35 mm films of direct radiograms were scanned with a mechanical image scanner, which digitized the gray level into 256 levels of the photographic density, or the logarithm of the opacity.[1] Each of

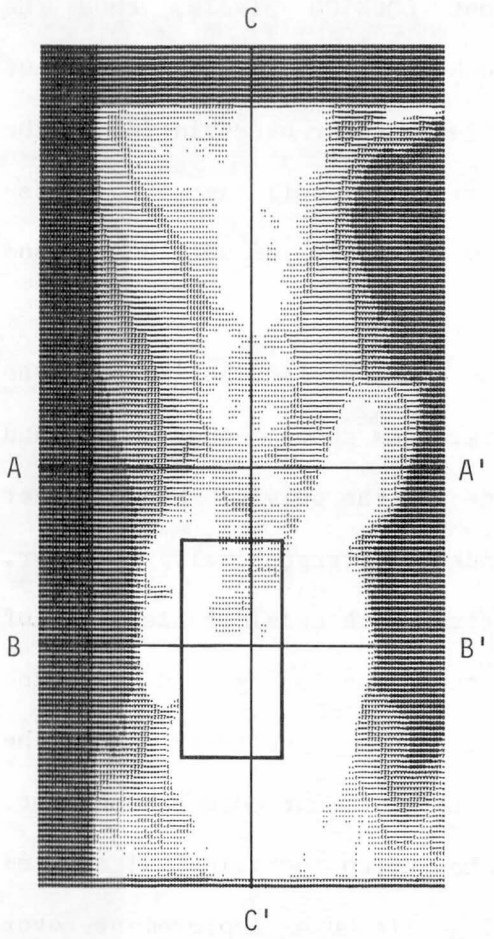
the digitized images consists of about 600X500 pixels, thus the corresponding sampling pitch is about 0.54 mm in the original scale of the direct radiogram. Fig. 3.2 shows a line printer overprinting of the density distribution of the previous image as well as its cross-sectional profiles. Properties described above may be well understood from these profiles.

Fig. 3.3 shows a histogram of the density distribution over the image. It consists of four major clusters. As shown, they correspond to the stomach, bones such as the backbone and the pelvis, the brighter background area, and the darker background area, respectively. However, this correspondence does not occur exactly, with considerable area of mixture. Therefore, binarization of the image using any constant threshold never brings about the gastric contour precisely. On the other hand, differentiation, a primary operation for edge enhancement, does not work, either. A result of the Laplacian operation[19] applied to this image is shown in Fig. 3.4. Although improvement over thresholding is observed in the greater curvature and in the fundus, the lesser curvature is collapsed due to noise.

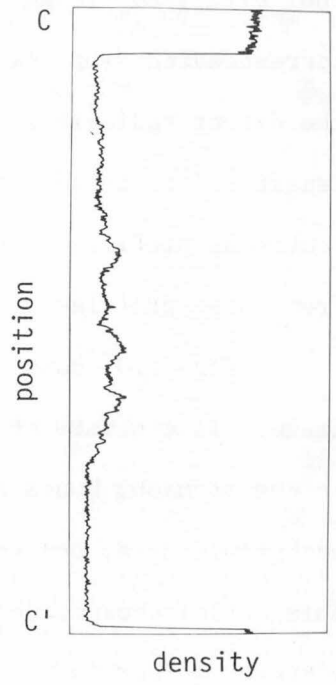
In a small area, thresholding works more preferably than edge enhancement. It was suggested that thresholding is applicable to transparency in the apex region, particularly after the transparency distribution is smoothed[5]. Also in a later experiment about a wider area over the apex, thresholding resulted better than differentiation. This area is indicated by a rectangle in Fig. 3.2(a). Fig. 3.5(a) is a result of a kind of differential filtering defined by a weight pattern

---

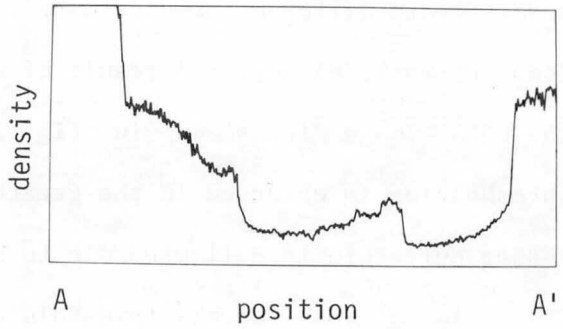
[1]Effects of the photo-copy process and the photo-electric conversion are involved in these digitized data.



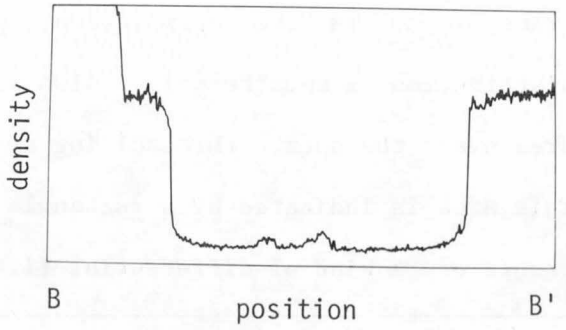
(a)



(b)



(c)

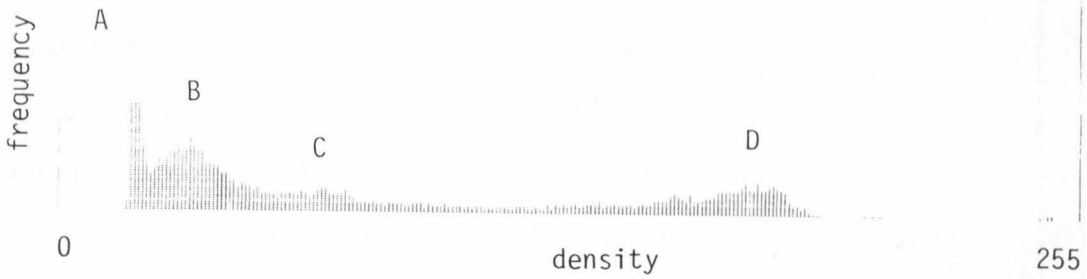


(d)

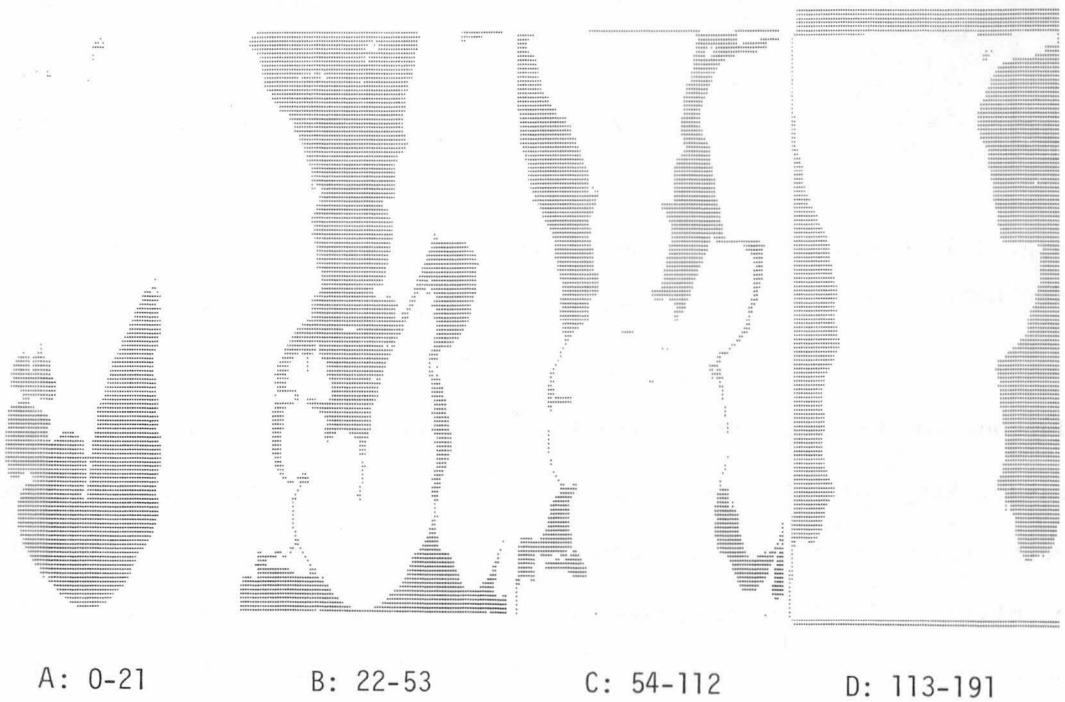
Fig. 3.2 Density distribution of an SPAP image.

(a) An overprinting by a line printer.

(b), (c), and (d) Profiles on the lines $\overline{AA'}$, $\overline{BB'}$, and $\overline{CC'}$ in the overprinting.



(a)



(b)

Fig. 3.3 Density distribution of an SPAP image.

(a) The density histogram.

(b) Regions corresponding to respective density clusters.

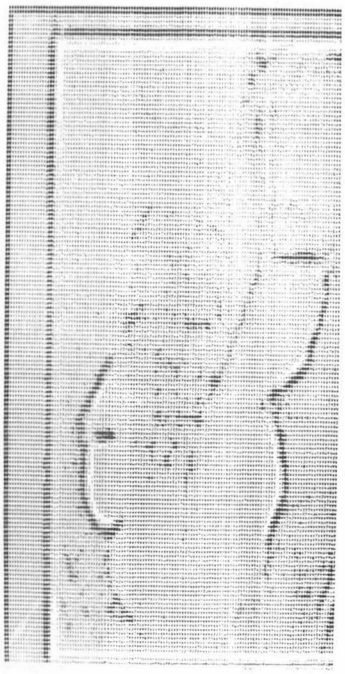


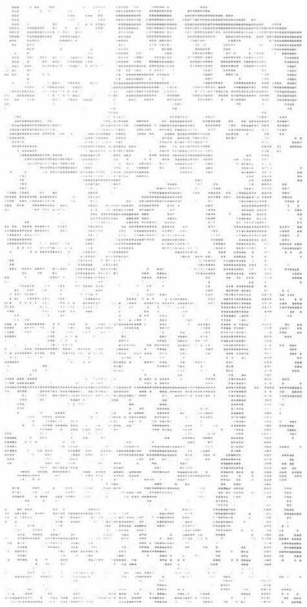
Fig. 3.4 A result of the Laplacian operation.

shown in the same figure; Fig. 3.5(b) is a result of thresholding to extract the region corresponding to the cluster of the brightest density in the histogram. Advantage of thresholding over differentiation is clearly recognized in this example.

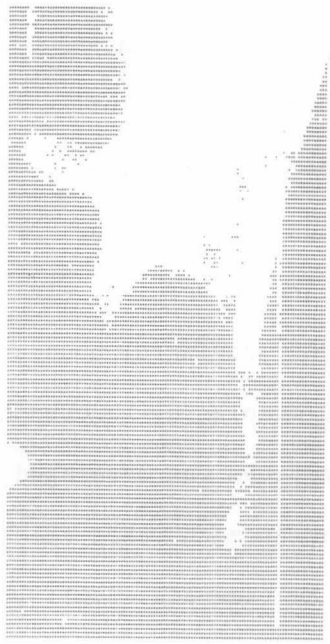
Thus, the space-variant property of the gray level distribution stands as a major problem in extracting the gastric contour.

3.2 Extraction of the Shadow of a Stomach

In order to overcome the problem of the shading and noise, the author studied some methods based on the dynamic threshold concept. This concept was originally introduced by Chow and Kaneko[20]. In their method, they assumed the bi-modality of the histogram of the density distribution in a small partition of an image. To each histogram, they applied curve fitting by a mixture function composed of two Gaussian distributions. They determined local threshold on the basis of the best fit obtained by a nonlinear optimization technique. We devised some methods because the bi-modality of histograms does not always hold; and because much time is consumed by the nonlinear optimization. Three methods are described in chronological order of development in this section, where the image in Fig. 3.2 is refered to as an illustrative example. Results of application to another example are also shown later. As a result, the third method was selected and used in the subsequent experiments.



(a)



(b)

-1	-1	-1	-1	-1	-1	-1	-1	-1
-1	-1	-1	-1	-1	-1	-1	-1	-1
-1	-1	-1	-1	-1	-1	-1	-1	-1
-1	-1	-1	8	8	8	-1	-1	-1
-1	-1	-1	8	8	8	-1	-1	-1
-1	-1	-1	8	8	8	-1	-1	-1
-1	-1	-1	-1	-1	-1	-1	-1	-1
-1	-1	-1	-1	-1	-1	-1	-1	-1
-1	-1	-1	-1	-1	-1	-1	-1	-1

(c)

(d)

Fig. 3.5 Processing of the rectangular region in Fig. 3.2(a).
(a) Differential filtering with the weight pattern in (c).
(b) Thresholding by the threshold which is shown in (d) along with the density histogram.

3.2.1 Partitioning of the Frame

A primitive means to cope with the varying discriminating level is to divide the entire image plane into small subregions to determine thresholds, subregion by subregion. In such a means, subregions including the boundary between the object and the background must be identified; otherwise, false boundaries are extracted in subregions where no meaningful boundary exists. To avoid this malfunction and to perform meaningful thresholding in subregions where true boundary exist, an algorithm was developed. This algorithm is based on the knowledge that the stomach makes the brightest shadow in the radiogram.

Algorithm (see Fig. 3.6):

Step 1: Divide the entire image plane into subregions composed of 100X100 pixels. Initialize all entries of a table which indicates the state of each subregion. The state is "unprocessed,"

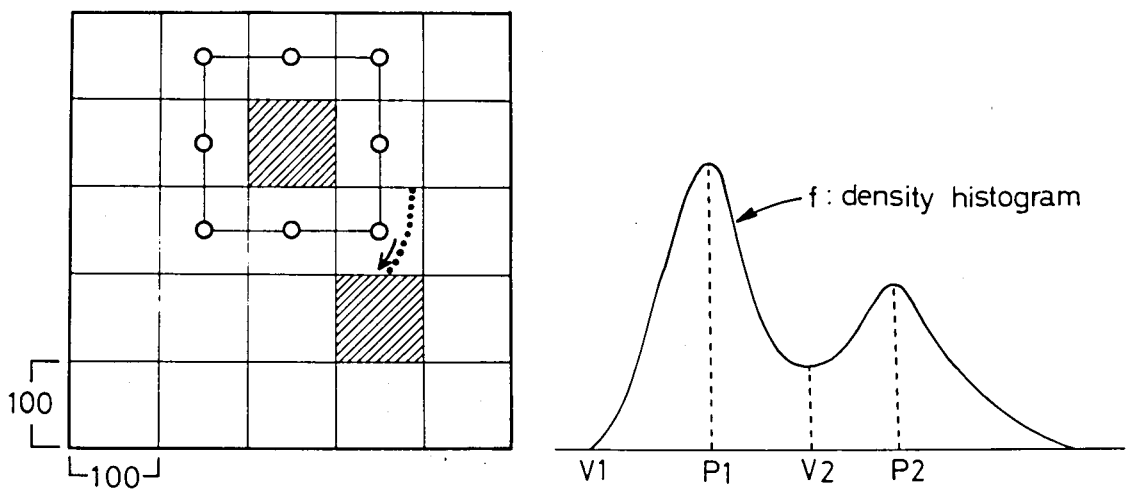


Fig. 3.6 Partitioning of a frame.
(a) Partitioning.
(b) Determination of a threshold.

"processed," or "processing required." The initial state is "unprocessed."

Step 2: Calculate an average density in each subregion and search for the subregion of the minimum average. Occupancy of the stomach in this subregion is greater than that in any other subregion. Since this subregion is probably the nearest to the central portion of the stomach, it is called the central subregion.

Step 3: Find the mode of the density distribution from its histogram in each of the central subregion and its eight neighbor subregions. Set the state of the central subregion to "processing required." Set the state of each of the neighbor subregions also to "processing reqired," if its mode density is nearly equal to the mode density of the central subregion. (This means that its density distribution is similar to that of the central subregion.)

Step 4: Determine a threshold for each of the subregions which are in the state of "processing required." Binarize this subregion and extract the boundary. If this boundary is connected to any "unprocessed" subregion, change its state to "processing required." Change the states of the processed subregions to "processed."

Step 5: Search the table for a subregion in the state of "processing required." If any one exists, then go to the next step; if no one exists, then terminate the processing.

Step 6: Determine a threshold, binarize the subregion, and extract the boundary. If this boundary is connected to any "unprocessed" subregion, change its state to "processing required." Go to Step 5.

A threshold is selected from among three candidate values $V1$, $V2$, and $(P1+P2)/2$ in Fig. 3.6. The values $V1$, $V2$, $P1$, and $P2$ are determined

as densities of valleys and peaks of the density histogram of each subregion. The valleys and the peaks are detected by finding the change of the sign of

$$h_{i+1+l}-h_{i-l}. \tag{3.1}$$

This quantity is in proportion to the difference of local averages of the density histogram

$$(2l+1)\left[\frac{1}{2l+1}\sum_{j=i+1-l}^{j=i+1+l} h_j - \frac{1}{2l+1}\sum_{j=i-l}^{j=i+l} h_j\right],$$

where h_j is the frequency of the density j. In Step 4, a threshold is determined as $V2$, the density of the valley next to the hill of the brightest part. In Step 6, a threshold is selected from the three values as one which is nearest to the threshold of the central subregion.

This method processes only subregions where detected boundaries are connected. Therefore, it prevents detection of false boundaries as well as saves processing time and memory storage. The number of processed subregions is reduced in this method: for example, only ten subregions out of thirty were processed for the example of Fig. 3.2. This result is shown in Fig. 3.7.

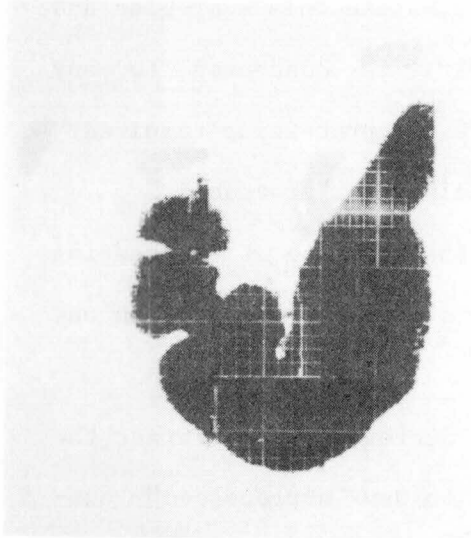


Fig. 3.7 A result of the frame partitioning method.

As a disadvantage of this method, discontinuity might appear at boundaries between subregions if the shading is great. This effect may be reduced if the size of a subregion is reduced. However, the reduction is limited by the number of pixels contained in each subregion since a

meaningful histogram must be made from them.

3.2.2 Uni-Directional Partition and Interpolation

As observed in Fig. 3.2, the shading effect is more involved in the horizontal direction than in the vertical direction. Therefore, an examination was conducted based on the assumption of the uni-directional shading. In this examination, an entire image is divided into strip-like subregions which overlap by 50 percent with the adjacent subregion each other. The subregions have their long sides in the same orientation as the backbone as shown in Fig. 3.8.

Algorithm:

Step 1: Binarize the entire image using one threshold determined by the mode method[19] applied to the histogram of the entire density distribution. This threshold is determined as the density where the first alteration of the sign of the expression (3.1) occurs when the histogram is searched in a direction from the brighter densities to the darker ones.

Step 2: A subregion is called a boundary subregion if 20-80 percent of its pixels have their densities greater than the threshold. For each of the boundary subregions, make the density histogram and find its valley in the brightest part in the same manner as Step 1.

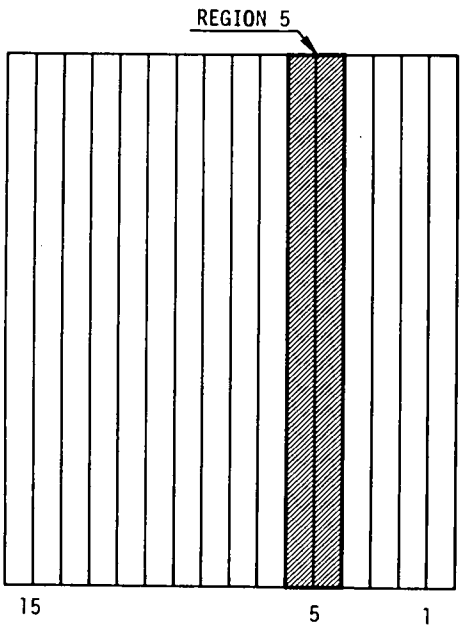
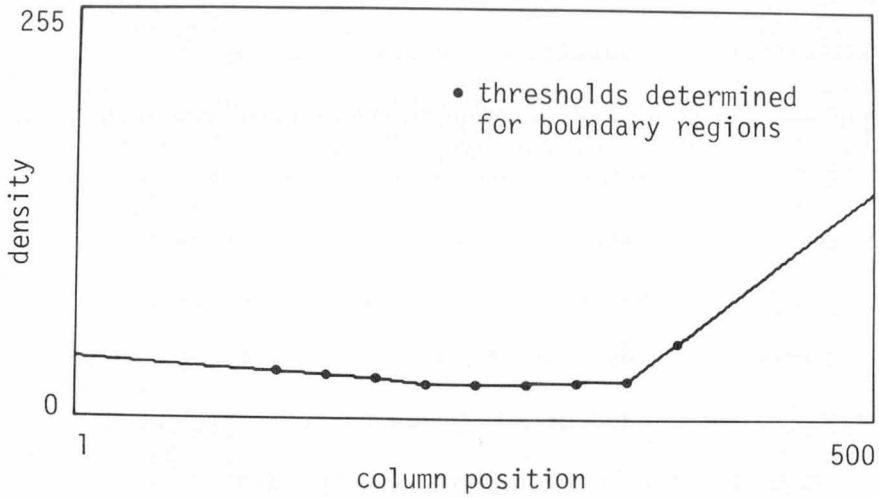
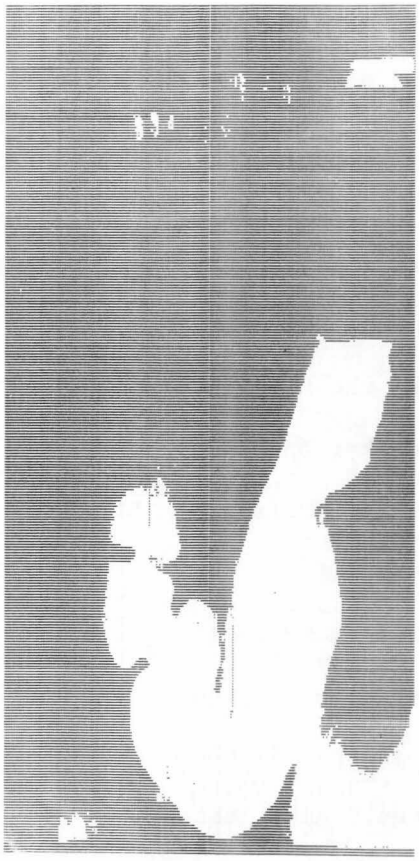


Fig. 3.8 Uni-directional partition of a frame.



(a)



(b)

Fig. 3.9 A result of the uni-directional partitioning method.
(a) Distribution of the threshold. (b) The binarized result.

Determine the threshold for its central column as the density of this valley.

Step 3: Linearly interpolate or extrapolate thresholds for all of the remaining columns.

Step 4: Binarize the image using the thresholds determined column by column.

Fig. 3.9 shows the threshold *versus* the column number as well as the result of this thresholding applied to the example in Fig. 3.2. The threshold varies considerably in columns passing through the upper part of the gastric body. Thus, the precision of the extracted contour is exellent in this region. Note that the backbone is eliminated because its shadow makes a kind of horizontal shading with an abrupt change of the density. A disadvantage observed here is that a part of the shadow of the pelvis is also extracted. This is caused by the vertically unvarying property of the threshold. What is worse, it is connected to the stomach. Disconnected components do not bring about any problems because they can be eliminated by logical operation.

3.2.3 Bi-Directional Partition and Interpolation

The third method is similar to the original version of the method of Chow and Kaneko[20]. Their method is summarized as follows.

Algorithm:

Step 1: Partition the entire image into 7X7 subregions which overlap 25 percent or 50 percent with the adjacent subregion each other.

Step 2: Make the density histogram in each of the subregions.

Step 3: Select subregions with large variance of the density distribution. For each selected subregion, estimate the underlying density distribution. This estimation is made by curve fitting by a linear

mixture of two Gaussian functions to the density histogram.

Step 4: Test the bi-modality of the mixture function of the estimated distribution. For each histogram with appreciable bi-modality, calculate the threshold from the estimated distribution function by the method of maximum likelihood.

Step 5: Perform the regionwise interpolation of the threshold.

Step 6: Assign the thresholds to the central pixels of the respective subregions. Perform the pointwise interpolation by a bi-linear method to determine thresholds for all of the remaining pixels.

Step 7: Binarize the image.

In applying the method to our problem, the following had to be taken into account:

1. The number of subregions in partitioning must be more in order to cope with the considerable shading.
2. The curve fitting by nonlinear optimization should be avoided, if possible, to save processing time as well as to let various distributions be acceptable without incorporating any strong assumptions about the form of the distribution function.
3. Thresholds should be determined in as many subregions as possible without the regionwise interpolation in order to minimize the smoothing effect of the interpolation on the "dynamic" threshold.

The following is the revised method.

Algorithm:

Step 1: Partition the entire image into typically 23X19 subregions as shown in Fig. 3.10.

Step 2: Calculate the average and the coefficient of variance of the density in each subregion.

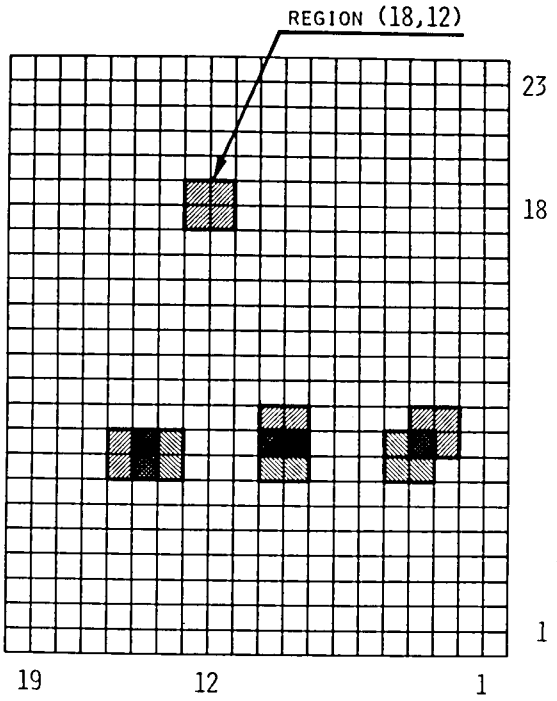


Fig. 3.10 Bi-directional partition of a frame.

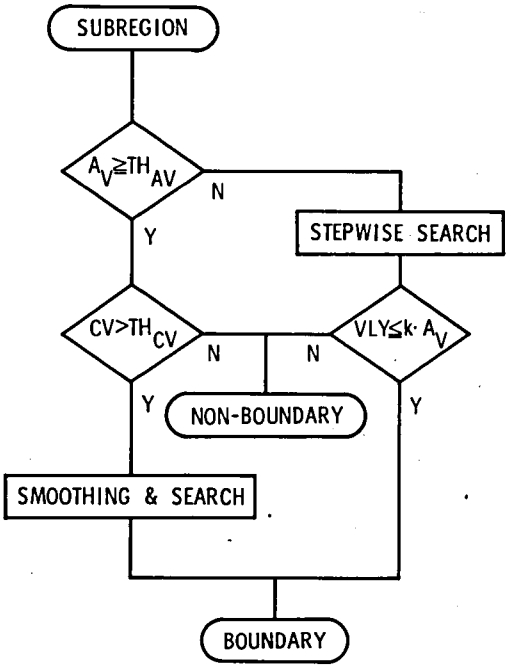
Step 3: For subregions of small average densities or of large coefficients of variance, determine thresholds by the mode method. Disregard those subregions where too large values are obtained. The subregions are called boundary subregions if their thresholds are accepted.

Step 4: Determine the thresholds for the subregions which are not boundary subregions, by interpolation by taking averages of thresholds of neighbor boundary subregions.

Step 5: Assign the thresholds determined in Step 3 and in Step 4 to the central pixels of respective subregions. Determine thresholds for all of the remaining pixels by the bi-linear interpolation.

Step 6: Binarize the image.

In Step 3, a heuristic method was devised. Described below is



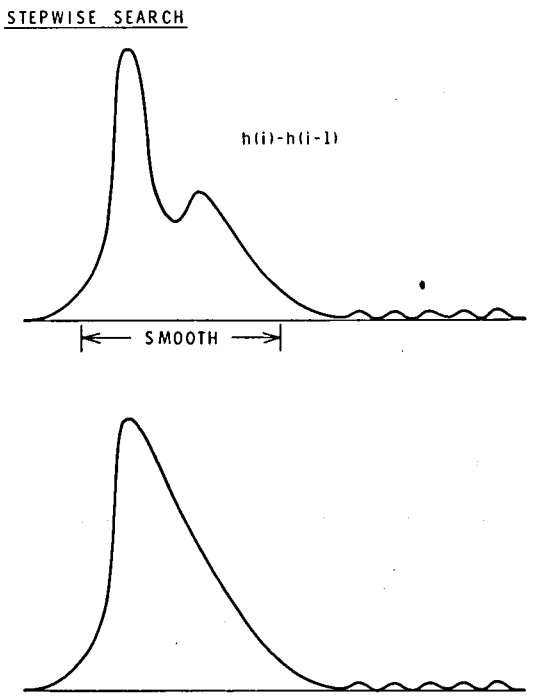
(a)

Fig. 3.11 Determination of boundary subregions and their thresholds.

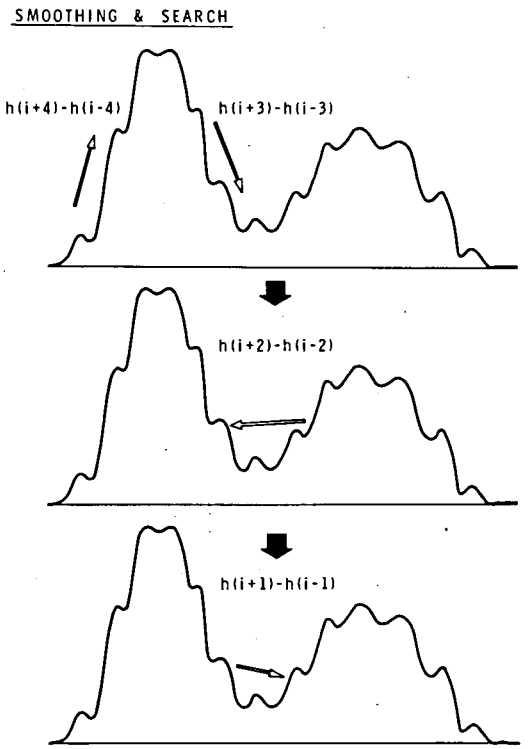
(a) The procedure.

(b) Valley search for concentrated histograms.

(c) Valley search for spread histograms.



(b)



(c)

its meaning. In this method, the valley of a histogram is searched for by evaluating the sign of the expression (3.1). This search is made in a direction from the brighter densities to the darker ones in order to find the valley in the brightest side. In order to determine thresholds which separate the stomach from the background, it is necessary to recognize whether each subregion includes the boundary of the stomach or not. This process is carried out as shown in Fig. 3.11(a):

1. If the average density of a subregion is small, this subregion is almost occupied by the stomach, and the background shares only a small part. Therefore, the small valley must be detected in the histogram, between the big hill corresponding to the stomach and the small hill corresponding to the background. This search is made step by step by setting $l=1$ in the expression (3.1) so as not to lose this valley. This search works because, due to the histogram concentration on a low density (see Fig. 3.11(b),) sample fluctuation of the histogram is not outstanding. If the value thus obtained is not too large, it is accepted as the threshold, while the subregion is recognized as a boundary subregion. Otherwise, it is very probable that the fluctuating skirt of a uni-modal histogram has been detected; therefore, the value is rejected and the subregion is denied to be a boundary subregion.
2. If the coefficient of variance is large, a subregion is shared by the stomach and the background evenly. Therefore, the subregion is recognized as a boundary subregion. Although great fluctuation appears in the histogram due to less concentration of the density, the valley itself is large. Thus the smoothing effect in the expression (3.1) can be set large in this case: the search is

started by using rather large l. Finding any alteration of the sign of this expression, l is set smaller and the search is repeated in the reverse direction, and so forth. It is just like a rocking chair going to standstill. (See Fig. 3.11(c).)

The coefficient of variance is used as a criterion of the boundary subregion because the hills of the darker densities in the histogram tend to spread widely.

Fig. 3.12 shows recognized boundary subregions and their thresholds for the example in Fig. 3.2 as well as its binarized result.

			41	36	36	49	49	47	34	36	35	36	35					23
	58								26		37	28	41					22
	65								34	29	31							21
	80								24	27								20
																		19
									30									18
																		17
																		16
														34				15
																45		14
														27		51		13
												34	29	26	36	41		12
			23	28								15	29	51	37			11
			45	21	22					23		39	24					10
		40	40	23	22				22	23		25	40					9
		40	40	22	22			20	20	22		21	34					8
		42	20		23	23	20	20	22	22	22	24	32					7
		35	36	23	21	20	20	21	21			25	32					6
			31	23	21		20	20	20			34	34					5
			30	27			20	20	20		23	29						4
				30	22	22	20	20	20	24	24							3
	34				19	19			20	20	23							2
	37	30	41	44	34	20	19	19	20	20	21	22	65	49	32	47	37	1
19	18	17	16	15	14	13	12	11	10	9	8	7	6	5	4	3	2	1	

(a)

(b)

Fig. 3.12 A result of the bi-directional partitioning method.
(a) Thresholds for boundary subregions. (b) The binarized result.

The original image size had been reduced to 300X250 by taking a four-pixel average; each subregion is composed of 25X25 pixels. Note that the boundary subregions form the shape of the stomach approximately on the whole. Mainly due to false recognition of some boundary subregions, portions of the backbone and the pelvis were also extracted. Scratches on the film were involved as well as holes caused by noises and incomplete interpolation of the threshold. These false effects can be eliminated logically if they are not connected to the stomach.

Processed examples for another image are shown in Fig. 3.13.

3.3 Extraction of the Gastric Contour

For the purpose of extracting the contour of only one figure, *i.e.*, the shadow of a stomach, irrelevant components should be eliminated from the binarized image. These components are such as extra shadows in the background and noises and holes in the stomach. They can be eliminated without smoothing the binarized image but logically.

Algorithm:

Step 1: Perform labeling[19] to the connected components.

Step 2: Calculate the area of each component by counting the number of pixels where identical labels have been assigned.

Step 3: Extract the component of the largest area.

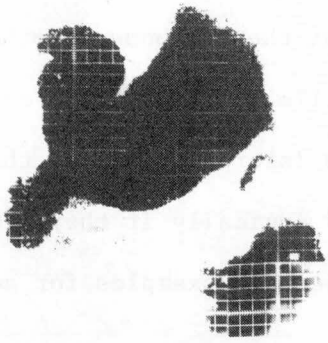
(At this point, the extra shadows are eliminated.)

Step 4: Exchange the label of the largest component with that of the background.

Step 5: Perform Step 1, Step 2, and Step 3.

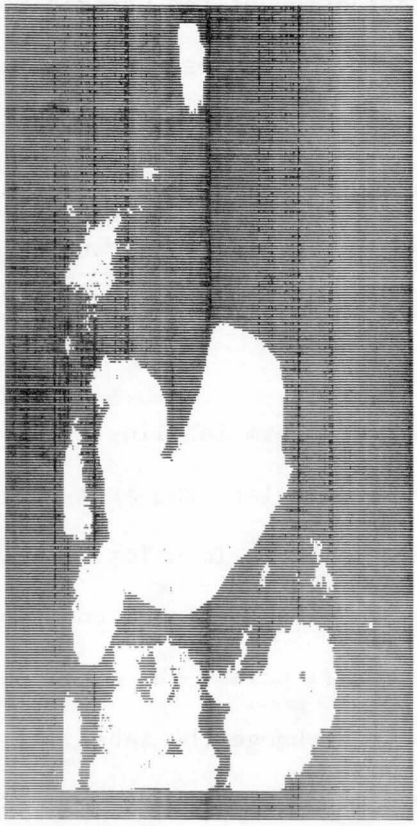
The resultant image contains only one and simply connected figure. An intermediate result after Step 3 is shown in Fig. 3.14. Meaning of this

(a)



(b)

(c)



(d)

Fig. 3.13 Processed examples for another image.

(a) The original photograph. (b) The result of the frame partitioning method. (c) The result of the uni-directional partitioning method. (d) The result of the bi-directional partitioning method.

method may be well understood if we think of the fact that we take pictures of interested objects large. This method is simple but very useful in extracting a figure from a noisy image without blurring edges. For extracting multiple fugures, or for noise filtering, delete components, whose area is smaller than a prescribed value, in Step 3[21].

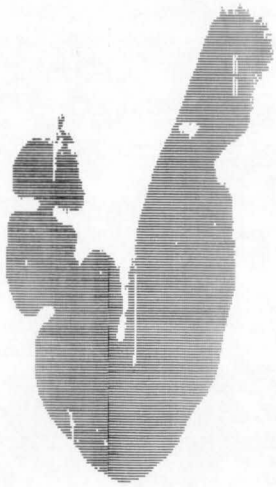


Fig. 3.14 The component of the largest area.

Since the connectivity holds for results of the thresholding, the gastric contour can be extracted by boundary tracking. The tracking operator never lose its way because only the shadow of the stomach exists in the binary image. A typical boundary tracking operator[19] first finds one boundary point by a TV-like raster scan. Then, the operator searches for an adjacent boundary point by circularly scanning the neighbor of this point. This operation is continued successively moving the center of the circle to the new point until the first point is met again. The boundary can be represented as an ordered sequence of its points. In implementing the algorithm, the starting point of each circular search must be changed according to the previous motion of the operator; otherwise, the operator may be "trapped" by a spinal portion of the boundary. The author devised an algorithm for determining the starting point. Let us define the eight neighbor elements of a point A as depicted in Fig. 3.15(a).

Algorithm:

Suppose that the new boundary point A_i has been detected by searching the neighbor of the point A counterclockwise. The succeding counter-

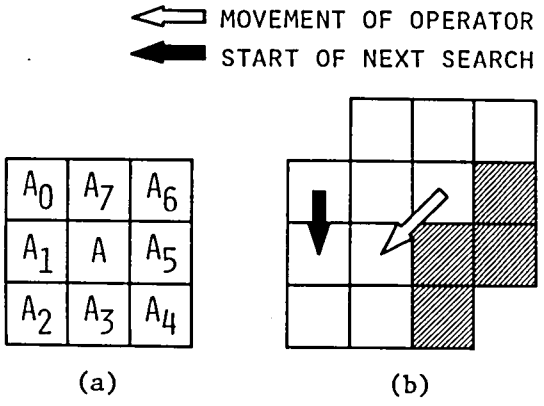


Fig. 3.15 Boundary tracking.
(a) Definition of neighbors.
(b) Beginning of circular search.

clockwise search begins at the point $A_{i-2(modulo\ 8)}$ in relation to the point A_i, which is the new point A, as shown in Fig. 3.15(b). Because the boundary determined by thresholding tends to be biased into the inside of the stomach (see Appendix A,) the points in the outside of the extracted shadow are taken as boundary points. This method ignores scratches cutting into the shadow if their width is of one pixel.

3.4 Structural Information and Its Applications

Soma *et al.*[5] introduced the concept of the gastric axis. It may be useful in understanding the peristalsis of a stomach. It may be also used as the information on the overall flow pattern of the stomach, thus enabling recognition and elimination of the shadow of the intestine, which is often observed when an image is taken after a period of time since the subject swallows contrast material. In such application, thinning, distance transformation, and skeletoning are invoked in order to extract and manipulate structural information of the

largest shadow in a binary image.

3.4.1 Thinning, Distance Transformation, and Skeletoning

There have been two concepts of thinning. One is a method to obtain the medial axis of a figure by shrinking it so that it may be used by computer to recognize the topological structure of the figure.[1] For this purpose, the connectivity of the medial axis must be guaranteed for any figure. Hilditch[22] developed a method of obtaining a medial axis, which is well-defined by the algorithm. To manipulate the medial axis as a linear graph, a method was also given by Hilditch[23] for simultaneous labeling to its nodes and branches. In this method, a node is a terminal or crossing point; a branch is a point sequence connecting two nodes. An abstract representation of the linear graph was given by the author[24]: the linear graph is represented as a table which is a variant of the node-branch incidence matrix. Its entries are locations, labels, and degrees of nodes as well as labels of branches connected to each node. The linear graph becomes a tree in our application because the given figure is simply connected: as a matter of truth, the purpose of Step 4 and Step 5 of the algorithm in Section 3.3 is to obtain a tree-structured medial axis; for the purpose of only extracting the gastric contour, this algorithm can be terminated at Step 3.

The other concept of thinning is usually called skeletoning[25]. A skeleton is a set of minimal number of points neccessary to preserve the complete information of the shape of a figure. Therefore, it is a technique for data compression of a binary image. However, since the connectivity is not guaranteed for a skeleton, it is not suitable for

---

[1]The terms medial axis and gastric axis are used discriminatingly in this article. The latter is an axis of a "tube," without branches.

recognition of the structure of a figure by computer.

Distance transformation, usually performed as preprocessing for skeletoning, is a method of calculating distance of any point in a figure from its boundary. The boundary can be completely restored from the set of the skeleton points and their distance by inverse distance transformation.

Montanari[25] and Levi and Montanari[26] mathematically formalized distance transformation, skeletoning, and inverse distance transformation for any kinds of distance. Suppose T_{ij} is distance of a point (i,j) from the boundary; t_{ijkl} is the distance between the point (i,j) and a point (k,l). Distance T_{ij} is given as a solution of a functional equation of dynamic programming,

$$T_{ij}=min(T_{kl}+t_{ijkl})$$

under constraint that $T_{ij}=0$ for any point outside of the figure. A skeleton point (i,j) is a point where relation

$$T_{ij}\neq T_{kl}-t_{ijkl}$$

holds for any $(k,l)\neq(i,j)$. Inverse distance transformation is given as a process of solving an equation,

$$T_{ij}=max(0,T_{kl}-t_{ijkl}) .$$

They also developed sequential algorithms for efficient implementation of the distance and inverse distance transformations. The author noted that, although it is essentially a parallel operation, the skeletoning is a process of extracting only the skeleton points, thus requiring no additional image storage[24]. All of the algorithms are based on neighborhood operations. The type of distance is defined by permissible paths between pixels in a neighborhood of a pixel. Three primitive distances were exemplified by Montanari[25], called method 0, method 1,

and method 2, respectively. Practically, Method 2 may be appropriate in approximating the Euclidean distance.

3.4.2 Determination of the Gastric Axis

The gastric axis may be established by eliminating extra branches from the medial axis. To start the process, a human observer specifies two terminal points in the medial axis which are the start and end points of the gastric axis. Then, computer manipulates the table of its linear-graph representation to eliminate, from the table, entries related to extra branches by successively deleting the unspecified terminal nodes, or nodes of degree one, and the branches connected to these nodes. This operation is repeated until no unspecified terminal nodes remain. Finally, the nodes and the branches which have no labels in the table are eliminated from the medial axis in the image plane. Then, the residual spines are deleted.

3.4.3 Elimination of the Shadow of the Intestine

If contrast material is included in the intestine, it makes a shadow of the approximately same density as the stomach. If this shadow overlaps that of the stomach, obviously, anatomical knowledge on the structure of the image is necessary for discriminating the stomach from the intestine. It may be permitted for the human operator to estimate and draw the stomach boundary. However, this scheme was not adopted, since one purpose of the study has been to find approaches to machine manipulation of the structure.

The author devised a discriminating procedure on one premise: in the medial axis, terminal nodes of the branches corresponding to extra shadows have been recognized by computer or specified by the human operator.

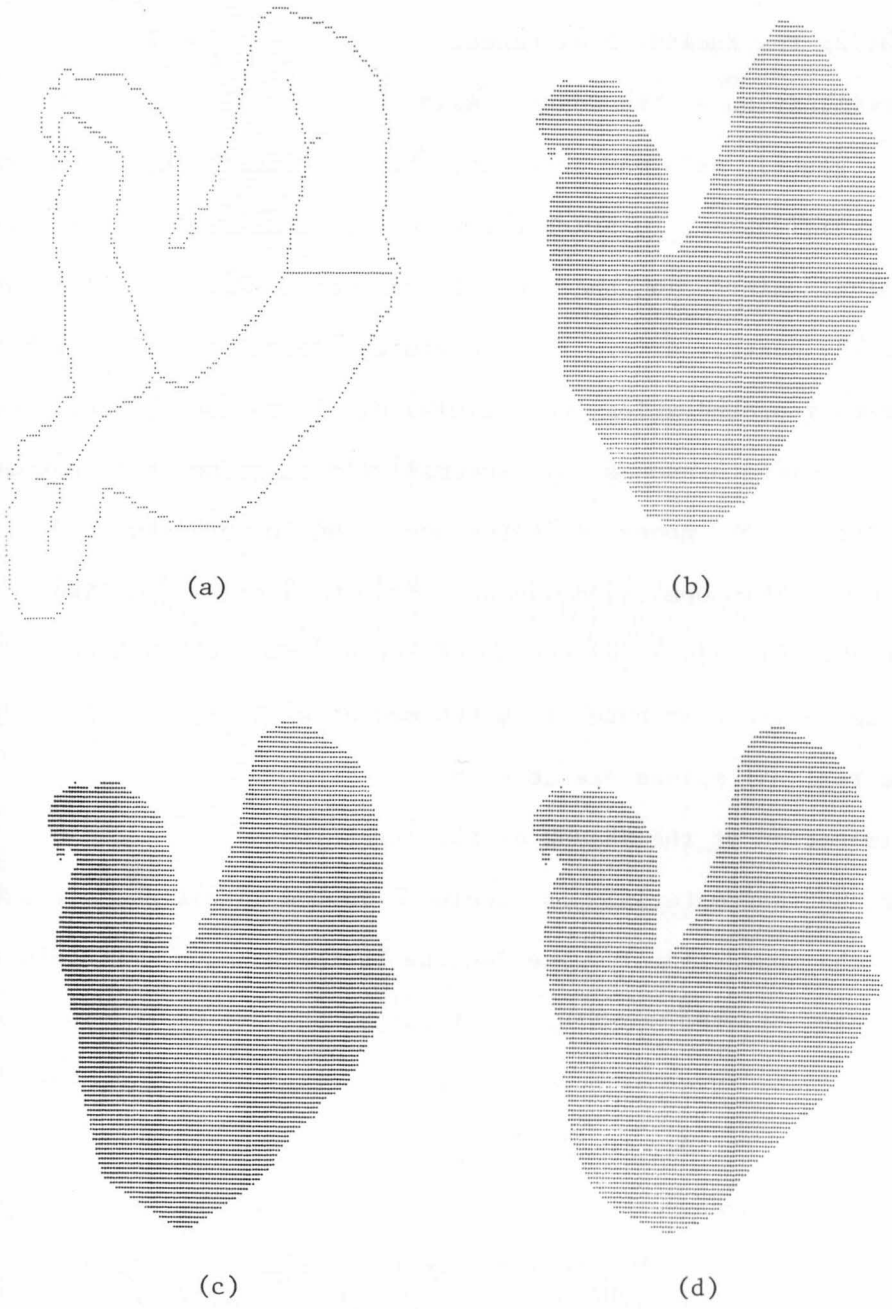


Fig. 3.16 Examples of elimination of the shadow of the intestine.
(a) A figure with the shadow of the intestine at its lower left part and its medial axis.
(b), (c), and (d) The figures reconstructed using the distance defined by method 0, method 1, and method 2, respectively.

Procedure:

Step 1: Delete these terminal nodes and branches from the table of the graph.

Step 2: Delete them from the image plane using the table.

Step 3: Apply the inverse distance transformation to the tailored medial axis after associating its points with their distance obtained by the distance transformation of the original figure.

Examples of the procedure are shown in Fig. 3.16, where three kinds of distance are used for comparison. Because the medial axis does not include all of the skeleton points, reconstructed figures involve distortion. This effect is reduced as the distance gets closer to the Euclidean distance.

3.4.4 Understanding the Structure of a Figure

Akatsuka *et al.*[8] suggested a method of understanding the approximate shape of a stomach using a functional relation of the distance *versus* the relative location of each point on the gastric axis. However, they have not discussed what kind of information is represented by this relation. The eye-observation of various results of the structural methods may be useful in understanding the structural information of a figure. One shortcut to do this is to reconstruct corresponding figures using the distance of the points on the (possibly tailored) axis. The inverse distance transformation can be used for this purpose. The information of any subset of the medial axis may be examined in this manner of "analysis by synthesis." Another shortcut is to compare the axis with a skeleton, a minimal set preserving the boundary information. Fig. 3.17 illustrates examples of such approaches.

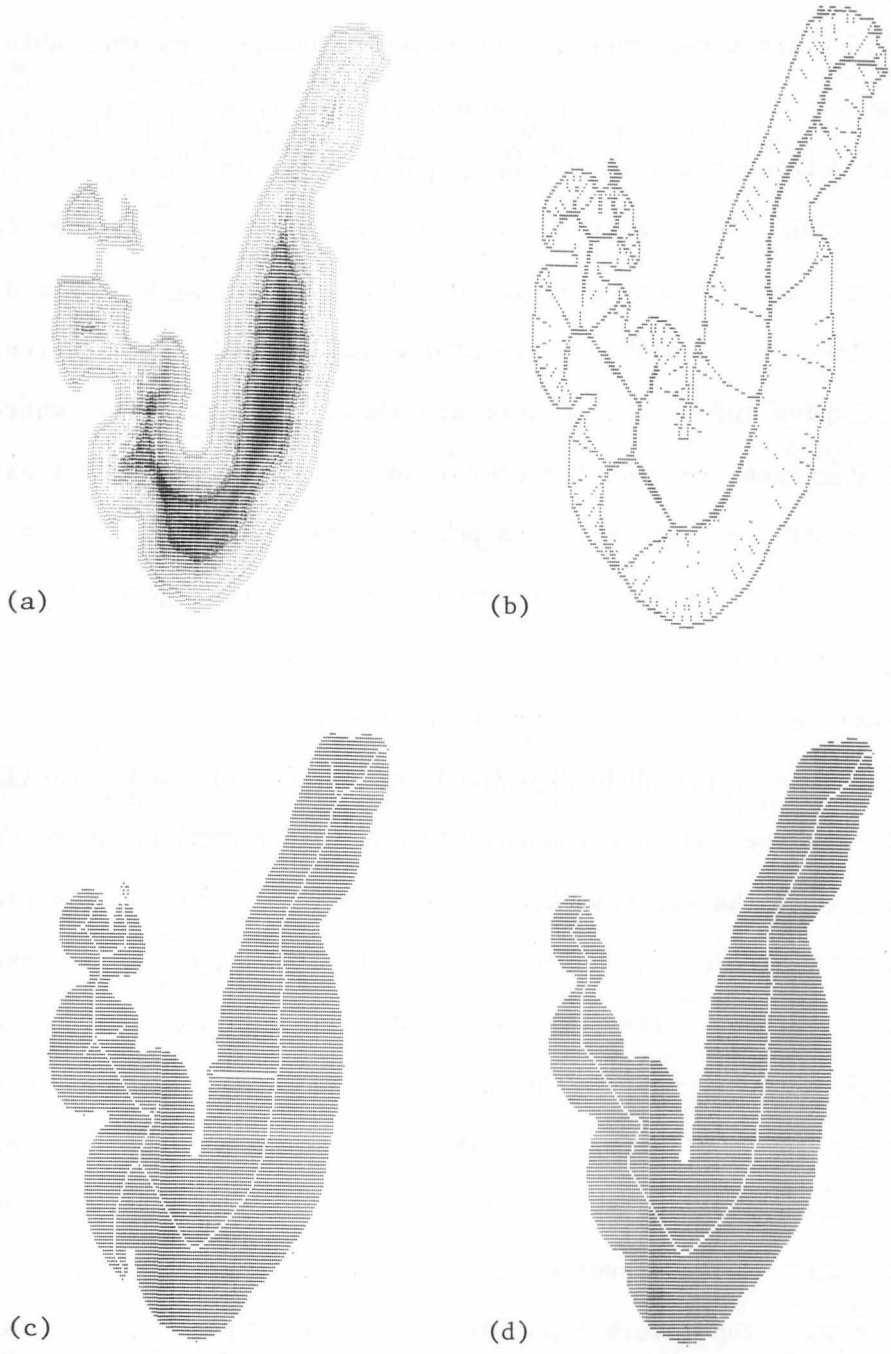


Fig. 3.17 Examples of structural information.
(a) Distance distribution. (b) A skeleton.
(c) and (d) Figures reconstructed from the medial axis and the gastric axis, respectively.

3.5 A System Configuration for Digital Image Processing

3.5.1 The Basic Concept

The study on the computerized diagnosis of SPAP images of stomachs involves not only designing the system but also analyzing the images as in cases of any other biomedical images and, more generally, natural images. Often, the analysis forms a main part of such a study. The knowledge and the technique for its implementation are usually unknown at the beginning; rather, they are acquired with advancement of the study. Thus, the subject is inevitably open-ended. For such subjects, it is useful to design a prototype system which can be revised and improved easily by researchers with a variety of background of radiology, medicine, and engineering. The following may be required of such a prototype:

1. Interactive system. It will facilitates flexible installation of the knowledge and flexible control of the processing flow as well as quick evaluation of results of such actions.
2. Programming language. A popular and standardized high level language should be usable so that researchers can commit easily in revision of the system.
3. Changeability to an autonomous system. Procedures determined in the course of the study should be fixed in the system as well as their control parameters so that the system requires as little human intervention as possible.

From this point of view, a software system was composed for the research and development. The strategy taken here are,

1. Separating the system into a set of unit routines executed individually,

2. Structuring each program into basic modules,
3. Using FORTRAN, a high level programming language, and
4. Programming attaching more importance to readability than to processing time and memory size.

Nowadays, the computing cost has been reduced due to the development of fast and large computers, while software is becoming larger and larger. Therefore, a high level language should be used even for practical systems because of its portability, flexibility, readability, and writability. Thus, the user may obtain intermediate results of each unit routine, input them into his own programs, or feed results of his programs to another unit routine. On the other hand, he may modify any routine or write a new one using the provided module programs. If he wants to apply a number of routines in series, he may make a macro procedure. If such prototypes and their modified versions are circulated among different research groups, the researchers may be able to "talk" with each other using the realized systems as tools just like mathematical expressions.

As a first step, a system was created for implementation on a general purpose large computer and was tested.

3.5.2 The System Configuration

The system consists of eleven unit routines R1, R2, ..., R11 as depicted in Fig. 3.18. Each routine is a main program which calls the subroutine modules listed in Table 3.1 successively in a hierarchical structure. The user may activate individual routines; or he may make a new program calling the subprogram modules and activate it. Functions of the respective routines are described below:

R1. Window specification. It extracts a specified window region with

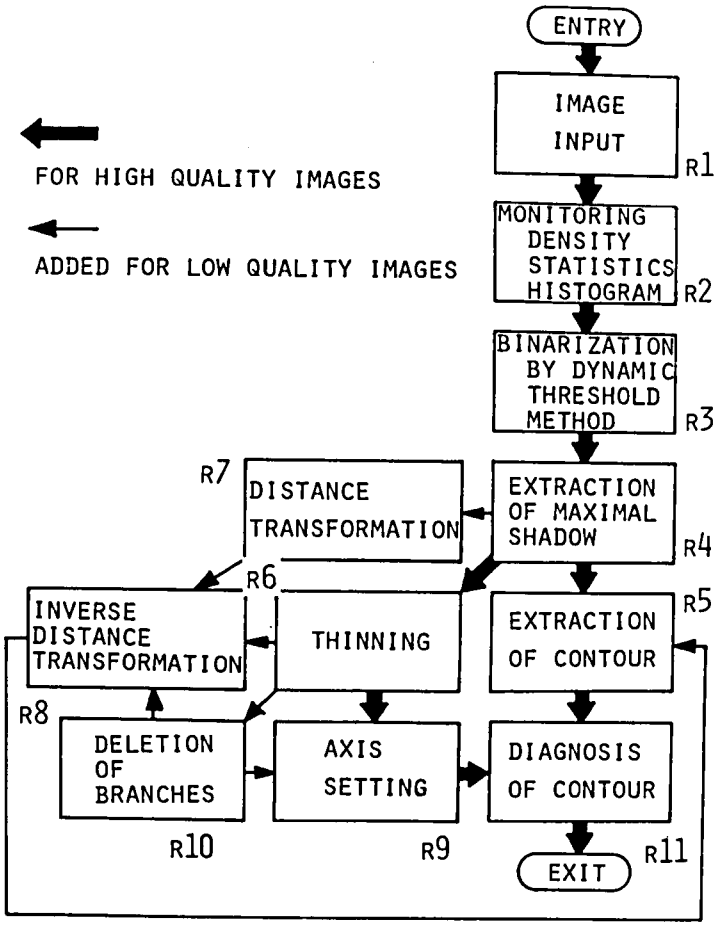


Fig. 3.18 A system configuration for SPAP image processing.

pixels of specified apperture from a large image stored on magnetic tape.[1] The resolution is reduced by a factor of NXN by assigning the average density of NXN neighbor pixels to the corresponding pixel of the window.

R2. Image property observation. It displays the density distribution and its histogram on the line printer. It also calculates and

[1]For the purpose of developing an algorithm or a software system, it is recommended that test data are stored in a digital form for reproducibility. Otherwise, temporary change of the quality of image films and photo-electric conversion noise disable performance evaluation.

Table 3.1 Subprogram modules

Name	Function
REGNMT	reads an image from MT.
GRYIMG	prints a gray-level image.
SBLPRT	prints a line of symbols for one image line.
HSTGRM	prints a density histogram.
STTSTC	calculates statistics.
DYNTHR	performs dynamic thresholding.
HSTGM1	calculates a regionwise histogram.
STRVLY	searches a histogram valley strictly.
FSTVLY	searches a histogram valley with smoothing.
REGINT	performs regionwise interpolation.
PNTINT	performs pointwise interpolation.
BNRYMG	prints a binary image.
FRAME1	sets frame edge points to "0."
LBLING	performs labeling.
MXCLST	extracts the largest connected component.
SMOCON	smoothes a contour in an image area.
INVERT	inverts the binary level.
CONTOR	extracts a contour.
IMGCLR	clears an image area.
SETDTA	sets data in an image area.
THINNG	performs thinning.
THNLBL	performs labeling of thinned lines.
GRAPH1	represents as a linear graph.
FRAME2	sets frame edge and adjacent points to "0."
DSTRF0	distance transformation (method 0.)
DSTRF1	distance transformation (method 1.)
DSTRF2	distance transformation (method 2.)
SKLTN0	skeletoning (method 0.)
SKLTN1	skeletoning (method 1.)
SKLTN2	skeletoning (method 2.)
INVDS0	inverse distance transformation (method 0.)
INVDS1	inverse distance transformation (method 1.)
INVDS2	inverse distance transformation (method 2.)
SELDTA	selects data of specified points.
PMNTND	specifies permanent nodes.
DELBR1	deletes branches from a linear graph table.
DELBR2	deletes branches from an image area.
DELBR3	deletes residual spines.
BRSRCH	searches branching.
CHKBR	Searches branching from a branch.
DELSPN	deletes spines.
ORGCNT	coordinate-system conversion.
RANGE	determines the range of a stomach.
WINKEL	determines the apex.
PYLORS	determines the pylorus.
SMOOTH	smoothes a contour.
DEVIAT	calculates a deviation curve.
SEIKI	normalizes the pixel value.
WGRAPH	prints a deviation curve.

reports statistics such as the average, the standard deviation, and the coefficient of variance of the density.

R3. Binarization. It binarizes an image by the third version of the dynamic threshold method in Section 3.2.

R4. Extraction of the largest shadow. It logically eliminates background shadows as well as noises and holes in the stomach by the method described in Section 3.3.

R5. Extraction of the contour. It extracts the gastric contour by tracking the stomach boundary counterclockwise as described in Section 3.3.

R6. Thinning. It performs thinning of a stomach and represents its medial axis in a tabular form by taking the axis as a linear graph as described in Section 3.4.

R7. Distance transformation. It performs distance transformation. It also produces the skeleton of a stomach for comparison between the (possibly tailored) medial axis and the skeleton to understand the loss of information. Montanari's three kinds of distance are available.

R8. Inverse distance transformation. It performs inverse distance transformation using one of the three kinds of distance.

R9. Determination of the gastric axis. It determines the gastric axis as described in Section 3.4. Then, the axis is represented as a sequence of its points arranged in order from the fundus to the pyloric antrum. This representation is made by a line tracking; residual spines are deleted at the same time.

R10. Deletion of extra branches. It deletes all specified branches from the medial axis as described in Section 3.4.

R11. Interpretation of the contour. It extracts features from a gastric contour and classifies the stomach. The final version of this routine is described in detail in Section 4.3.

In Fig. 3.18, only the routines specified by the thick arrows are usually invoked. However, if the intestine or bones make shadows as bright as the stomach, the routines specified by the thin arrows should be also invoked. The flow of processing may be determined as below after the routine R4. If no background shadows overlap the stomach, the gastric contour will be passed to the routine R11 after it is extracted by the routine R5. On the contrary, if some shadow overlaps the stomach, the routine R10 deletes branches corresponding to the extra shadow from the medial axis after this is obtained by the routine R6; the routine R7 calculates distance of each point; and then, the routine R8 reconstructs the figure of the stomach using the distance of the points on the tailored axis. This figure will be processed by the routine R5, which passes the gastric contour to the routine R11. On the other hand, the gastric axis is obtained by the routine R9 from the medial axis and is also passed to the routine R11 in the both cases. The flow control itself involves a difficult problem and the human operator must do it at present. Some control parameters are to be specified at runtime for convenience of changing their values.

3.5.3 CPU Time and Memory Size

The system was implemented on a general purpose large computer, FACOM-M190. The approximate CPU time for processing an image of 300X250 pixels was thirty seconds for the routine R1; three seconds for each of the routines R2, R5, R8, R9, and R10; five seconds for the routine R3; and ten seconds for each of the routines R4, R6, R7, and R11.[1] The

Table 3.2 Examples of CPU time

Routine	Example 1	Example 2
R1	30.73 sec.	31.00 sec.
R2	3.51	3.55
R3	5.09	5.29
R4	9.14	9.72
R5	1.92	2.19
R6	9.65	11.89
R7	10.99	10.83
R8	3.64	3.60
R9	2.37	2.40
R10	2.28	2.29
R11	10.13	9.51
Total	89.45	92.27
Average	8.13	8.39

total time was about ninety seconds for one image. Examples are shown in Table 3.2. The program size of the system is about 3500 FORTRAN source lines. The working memory was prepared for respective routines as required: 75 kilo-words for the image; 3 kilo-words for each of the row and column numbers of the gastric contour[2], the medial axis, or the skeleton; and 1200 words for the table of a linear graph.

3.6 Remarks on Implementation on a Small Computer

In conducting this research, the author impressively felt the

[1]Calculation of Fourier descriptors was not included in the routine R11 at this time. See also Section 4.2 and Section 4.3.
[2]Actually, a contour needed less than 2 kilo-words. The Freeman's coding is not used.

need of designing a flexible system convenient as a tool for image processing study. Thus, a tool system was actually made and implemented on a small computer NOVA01 equipped with standard peripherals under its operating system RDOS. The hardware also includes an ITV image reader and a CRT image display but is not flourished with expensive devices such as image processing hardware. The main strategy of this work was to develop a root system which can be readily implementable in any computer system without expensive dedicated hardware, for only such systems can be acceptable at many laboratories due to financial problems. This system can be used at two levels: the command level and the programming level. A beginner can use this system by simply issuing a command from the console. If a stream of procedures are to be performed successively, he can also stack the procedures to create a macro command in advance. While values of many parameters are given in dialogue with the system, most of them can be also fixed in order to minimize repetitive and cumbersome operation of specifying identical values. A more advanced user can modify and append programs appropriate for his problem. For ease of programming, the system is written in FORTRAN and a set of virtual array subroutine subprograms is provided for manipulating large images. The design principle of the system, the system management, supporting utilities, and processing programs were published elsewhere[27]. The programs for the gastric contour extraction were implemented on this tool system. Table 3.3 shows examples of the processing time for the gastric contour extraction. For comparison, FFT calculation time is also included. The processing time grows rapidly as the image size grows. This is mainly due to the frequent access to images resident on magnetic disk, through the buffer

Table 3.3 Examples of process time

Routine	Image size 64×64	128×128	256×256
R1	0.4 min.	0.9 min.	2.3 min.
R2	0.6	1.9	6.8
R3	1.7	4.3	40.6
R4	3.2	10.6	51.2
R5	0.2	0.5	1.3
R6	1.5	13.9	134
R7	0.2	1.0	3.6
R8	0.6	2.7	11.1
R9	0.8	2.7	7.3*
R10	0.8	2.7	5.1*
Total	10.0	41.2	263
FFT	1.6	8.5	43.5
FFT on NOVA3/D	0.8	4.1	19.0

* includes about 1.7 minutes for human intervention.

memory of the small size. Therefore, the time will be reduced if recent computer with fast and large main memory and quick access disk are used. The reduction is estimated to be by a factor of two or three from an implementation experience of the tool system on a small computer NOVA3D.

3.7 Summary of the Chapter

This chapter has been devoted to extraction of the gastric contour from an SPAP image by digital image processing. Athough it is called preprocessing from a conventional point of view in pattern recognition, it forms an important part of the computerized

interpretation: the precision of the extracted contour is crucial in analyzing the image. In this chapter, after preliminary examination on the property of the images, binarization techniques have been developed based on the dynamic threshold concept. As postprocessing of the binarization, logical filtering by labeling is used to eliminate shadows separate from the stomach; while methods of handling the structural information are used to eliminate shadows connected to the stomach. These techniques were arranged to make a prototype system which was implemented on a large computer. This system requires at most two kinds of human intervention: the flow control according to existence of extra shadows and specification of some nodes. The gastric contour extracted by this system may be useful in diagnosing the deformity of a stomach and in extracting the gastric contour more precisely. In order to detect fine changes on the gastric contour, more sophisticated methods must be developed to attain precision. This system was also implemented on a small computer in its general purpose image processing software which was made to realize some requirements as described in Section 3.5.

CHAPTER 4 FEATURE EXTRACTION AND PATTERN CLASSIFICATION

4.1 General Approaches

Approaches to classifying samples into pattern classes, or pattern recognition, may be categorized to be either descriptive or quantitative. Descriptive approaches use syntactic methods for inference, for example, based on the formal language theory. A famous example of descriptive approaches is found in Ledly[28] in the chromosome analysis. He described contours of chromosomes using primitives selected as different types of curve segments; then, he applied the formal language theory for classification. On the other hand, quantitative approaches use statistical methods known as the multivariate analysis, particularly, the discriminant analysis and the clustering. These statistical methods are also called supervised pattern recognition and unsupervised pattern recognition, respectively. Quantitative approaches have been used for a long period in biology and in medicine. In biology, a main concern has been to classify many variants of species in natural ways. This is called natural classification. In medicine, a main concern has been to classify findings and symptoms from a standpoint of human interest to know whether they are of a normal state or of a diseased state.

In the both approaches, the first step is to extract features from samples. The features are, however, of a somewhat different property. In descriptive approaches, features are primitives, or element patterns. They are identified by template matching. On the other hand, in quantitative approaches, features are measurements and transforms of measurements. Sometimes, the primitives are identified by

quantitative evaluation; a number of primitives thus obtained for different observations are summarized by a descriptive approach in order to give a final decision. Therefore, the difference between the two approaches comes from the level they are used: if a main concern of a research is to identify primitives, it appears as a quantitative approach; if it is to analyze relations of primitives, it appears as a descriptive approach. Selection of the approaches depends on the way a researcher recognizes the world of objectives: whether an objective should be recognized as a whole or by analyzing the context of its primitives. Therefore, primitives may be called local features, while measurements may be called global features. The remaining part of this chapter is devoted to description and quantification of a gastric contour; and quantification of an image of the apex region.

4.2 Description of a Gastric Contour

In order to describe the appearance of a contour, the deviation curve and the gastric axis are introduced.

4.2.1 The Deviation Curve

As mentioned in Chapter 1, abnormalities are recognized on the basis of a lack of harmony of the shape of a stomach; in other words, inconsistency or discontinuity of the local property of curves of its contour. Such concepts are ambiguous but lead us to an idea of examining the curvature. In fact, when physicians inspect gastric radiograms, they draw an ideal contour for each stomach in their minds. The deviation of a real shape from its ideal one give them hints about symptoms. This deviation is related to the curvature of the contour.

The ideal shape differs person by person as does the shape of a

stomach. The author has used a smoothed contour of each stomach as its ideal one, or its standard pattern. It is obtained by the moving average method. This method is applied to the original contour which is represented as a sequence of points in the Cartesian coordinate system. Then, deviation from the smoothed contour is calculated for any point on the original contour. The sign of the deviation is defined to be positive if the original point exists outside of the smoothed contour; and negative if inside. Actually, the sign is determined by a local property of the contour: the side of the smoothed point with regard to the tangent line of the original contour at the point corresponding to the smoothed point. This method is valid if the original contour is not extremely irregular. In this way, a deviation curve is formed, which is assumed as the original contour stretched along its smoothed one. Thus, a two-dimensional figure is reduced to a one-dimensional wave form.

4.2.2 Interpretation of a Deviation curve

Interpretation of the deviation at any location of the contour involves identification of feature points, establishment of the gastric axis, and judgement on the deviation.

4.2.2.1 Identification of feature points

It is difficult to define anatomical feature points of a stomach because a stomach is just a deformable tube. It has no fixed points nor angles of fixed morphology; it has no bulbs of particular shapes. However, in radiograms, three portions may be identifiable: the cardia, the apex, and the pylorus. These portions are more frequently observed in SPAP images than in other kinds of images. Unfortunately, the cardia is not clearly observable; and the pylorus is often occluded by the shadow of the pyloric antrum even in SPAP images. On the contrary, the

apex is often observable unless the peristalsis is prominent. In addition, its pathological importance in diagnosis has been recognized as described in Chapter 2. The experiment in this section referred to the apex and the uppermost points of both the fundus side and the pyloric antrum side. Details about the identification of various portions are described in Section 4.3.

4.2.2.2 Establishment of the gastric axis

Physicians always imagine the three-dimensional structure of a stomach on the basis of their anatomical knowledge and can draw the axial path from the cardia to the pylorus on the radiogram. The concept of the gastric axis was introduced to let computer understand the two-dimensional shadow of a stomach as a projection of a three-dimensional tubular container. The peristalsis may be distinguished from abnormal incisures by the symmetry about the gastric axis.

A method of establishing a gastric axis has been introduced in Section 3.4, where the axis was determined from the binarized shadow of a stomach by thinning.

The author also devised another method.[1] It establishes the axis from a delineated figure, the gastric contour. For a fishhook-shaped stomach, middle points of horizontal slices of the stomach are successively connected in the portions above the level of the apex; below the apex level, middle points of radial slices centered on the apex are successively connected. For a steer-horn shaped stomach, middle points of vertical slices are successively connected in the portion in the lefthand side of the apex. In the both cases, if a slice

---

[1]Historically, this method was examined first[5].

gives more than one line segment due to winding of the contour, the point nearest to the center is considered. After the entire axis is obtained, it is smoothed in its excessively discontinuous portions.

4.2.2.3 Local feature description of a contour

A variety of features may be calculated from the deviation curve. The author has used the depth and the width of each curve segment whose both ends are two adjacent zero-crossing points. This means that the deviation curve is manipulated as a boxcar wave form. In a boxcar, the signed hight and the length correspond to the extremum value and the width of the curve segment, respectively. If a segment includes more than one extremum, one of the greatest magnitude is adopted. Symptomatic changes on the contour may be put into certain descriptions of boxcars.

1. Linearization: the boxcar is long and of small magnitude.
2. Niche: the boxcar is short and of large magnitude.
3. Abnormal incisure: the boxcar is long and of large magnitude; and no similar incisures exist in a symmetrical position on the curvature in the opposite side about the gastric axis in the contour geometry.
4. Irregularity: many extreme points appear in an interval of the deviation curve; in other words, the average interval between the extreme points is short.

If none of these conditions holds, the corresponding curve segment of the contour may be regarded as normal.

This logic is depicted in Fig. 4.1. Actually, some noise thresholds were incorporated. For example, thresholds θ_S and θ_L ($0<\theta_S<<\theta_L$) were used instead of the zero level for the boxcar interpretation. Each point of the deviation curve corresponds uniquely

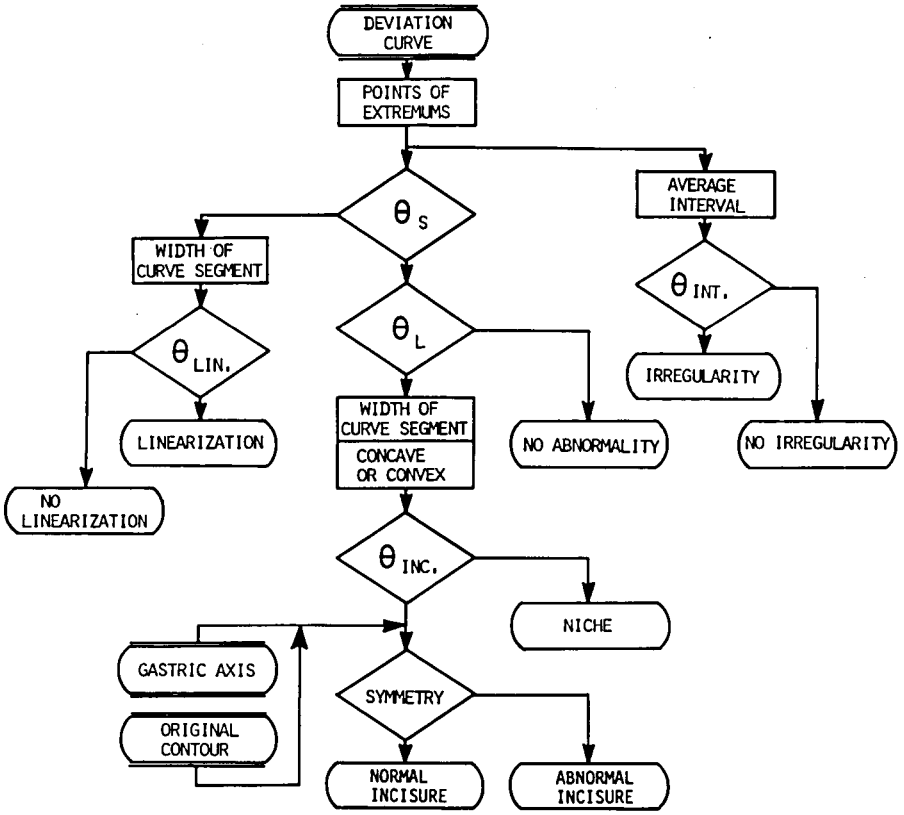


Fig. 4.1 A logic for description of a gastric contour.

to some one of the original contour. Therefore, interpretation of the deviation curve can be represented in terms of the original contour.

An example of description by this procedure is shown in Fig. 4.2 along with its deviation curve. It should be noted that, in spite of the rather simple logic, some symptomatic changes were properly described. For example, profile niche near the apex is represented by the context "↑C↑" in Fig. 4.2. However, interpretation of the local features in the context of their topological and geometrical relations is difficult in general.

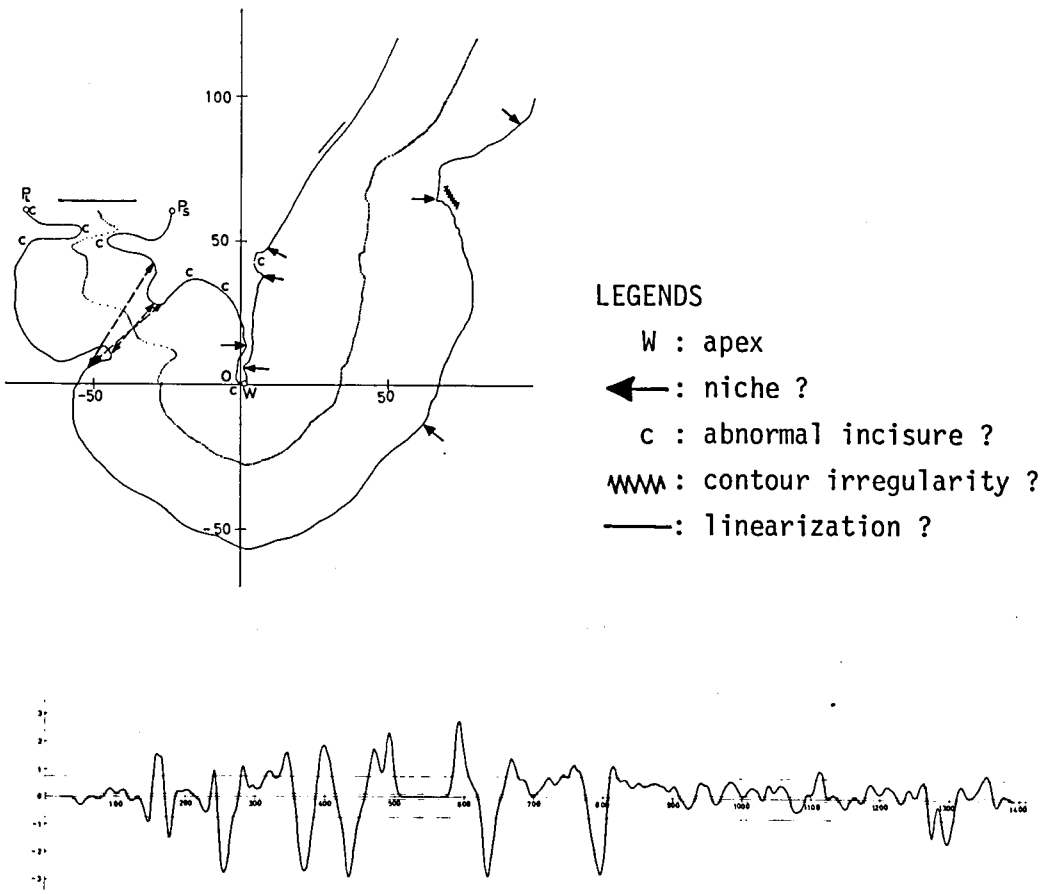


Fig. 4.2 An example of the description of gastric contours and its deviation curve.

4.3 Quantitative Analysis of Gastric Contours and Apex Images

This section describes two approaches to the feature extraction from SPAP images and their evaluation. A particular interest is placed on the deformity of the apex, because symptomatic changes tend to appear in this region. One approach is the extraction of features from a gastric contour. As proposed in a study on computational analysis of heart diseases[29], some geometrical measurements and Fourier descriptors were examined as candidate features. The other is the extraction by a two-dimensional orthogonal transform of a gray level image of the apex region. In the both approaches, the validity of individual features and sets of features was evaluated using the same methods. Therefore, in this section, the evaluation methods are described first, and then the feature extraction and the results of the respective approaches are shown. A two-class problem and a three-class problem are postulated. These are referred to as N/A (normal/abnormal) classification; and N/U/C (normal/ulcer/cancer) classification.

The sample set used in the experiment is composed of eleven samples of gastric cancer, twenty-four samples of ulcer, and twenty normal samples. Diseases were limited to those located in the apex region. Diagnoses of the diseased samples had been verified through pathological examination; the normal samples through direct radiography and endoscopy. The experiment analyzed their indirect radiograms taken in mass screening.

4.3.1 Methods for Evaluation of Discriminating Features

As a premise in pattern classification, patterns are expected to be separable to some extent in the feature space even if not completely.[1] To prove the separability, the following may be examined:

1. To get insight of the distribution of sample patterns through inspection of scattergrams by eyes.
2. To make a discriminant function and evaluate its discriminating performance by the discriminant analysis.
3. To group the sample patterns into clusters by the clustering and inspect those contained in the respective clusters.

Inspection of scattergrams is a most convenient means, particularly when the distributing property is unknown. However, the dimension of feature spaces is limited to at most two. The discriminant analysis is a direct method to evaluate the discriminating performance of a classifier. The third method is applied without assuming *a priori* classes. It is useful in inspecting the distributing property in feature spaces whose dimensionality is high, particularly, higher than two.

In addition to these examinations, the effectiveness of features and discrimination methods is determined in comparison with the discriminating performance of human experts.

4.3.1.1 The discriminant function

For the discriminant analysis, many kinds of discriminant functions have been developed such as linear classifiers, piecewise linear classifiers, and nonlinear classifiers[30,31]. The classifiers may be parametric or nonparametric. For simplicity, a linear classifier obtained by a parametric method was used in this study. In this section, a brief description of this classifier is given.

For a two-class problem, a Bayes' classifier is given by

$$h(x)=\frac{1}{2}(x-m_1)^T\Sigma_1^{-1}(x-m_1)-\frac{1}{2}(x-m_2)^T\Sigma_2^{-1}(x-m_2)+\frac{1}{2}\ln\frac{|\Sigma_1|}{|\Sigma_2|}\lessgtr\ln\frac{p_1}{p_2}\rightarrow x\varepsilon\begin{cases}class\ 1\\ class\ 2\end{cases}$$

[1]The terms classification and discrimination are synonymously used in order to match the terminology of both the pattern recognition and the discriminant analysis.

if the normal distribution is assumed for each class. In this expression, x is a pattern vector; m_1 and m_2 are the mean vectors, $\sum_1$ and $\sum_2$ are the variance-covariance matrices, and p_1 and p_2 are *a priori* probabilities of the class *1* and the class *2*, respectively. The function $h(x)$ denotes the Bayes' discriminant function.

When $\sum_1=\sum_2=\sum$, it becomes a linear classifier

$$h(x)=(m_2-m_1)^T\sum^{-1}x+\frac{1}{2}(m_1^T\sum^{-1}m_1-m_2^T\sum^{-1}m_2)\lessgtr \ln\frac{p_1}{p_2} \rightarrow x\varepsilon\begin{cases}class\ 1\\ class\ 2\end{cases} \qquad (4.1)$$

Rewriting this rule, an expression

$$(x-m_1)^T\sum^{-1}(x-m_1)-(x-m_2)^T\sum^{-1}(x-m_2)\lessgtr \ln\frac{p_1}{p_2} \rightarrow x\varepsilon\begin{cases}class\ 1\\ class\ 2\end{cases} \qquad (4.2)$$

is obtained. The expression $(x-m_i)^T\sum^{-1}(x-m_i)$ is squared Euclidean distance between the pattern vector x and the prototype pattern vector m_i after whitening of the distribution. This distance is called the Mahalanobis' distance. That is, a Bayes' linear classifier is a Mahalanobis' distance classifier. When $p_1=p_2=\frac{1}{2}$,

$$(x-m_1)^T\sum^{-1}(x-m_1)\lessgtr(x-m_2)^T\sum^{-1}(x-m_2) \rightarrow x\varepsilon\begin{cases}class\ 1\\ class\ 2\end{cases}$$

This gives a decision boundary which is the perpendicular bisector of the line joining the prototype pattens m_1 and m_2 in the whitened space. This is called the minimum Mahalanobis' distance classifier. The Bayes' classifier is the optimum classifier under the Fisher's criterion

$$f=[E\{h(x)|1\}-E\{h(x)|2\}]^2/[Var\{h(x)|1\}+Var\{h(x)|2\}]$$

where $E\{h(x)|i\}$ and $Var\{h(x)|i\}$ are the conditional expectation and the conditional variance of $h(x)$ in the class i. This criterion is a measure of the inter-class distance compared with the intra-class distance. When $\sum_1=\sum_2$, the solution maximizing this criterion is given by the rule (4.2). Therefore, the discriminant function of this rule is also called the Fisher's discriminant function.

When the classification risk,

c_{ij}=cost associated with deciding $x \varepsilon$ class j when $x \varepsilon$ class i is considered, a minimum-risk Bayes' classifier is given by exchanging p_1 with $(c_{12}-c_{11})p_1$ and p_2 with $(c_{21}-c_{22})p_2$ in (4.1) and in (4.2).

In the discriminant analysis which follows, the minimum Mahalanobis' distance classification rule is used. This means that the assumption is

1. the class distribution is normal with an identical variance-covariance matrix and
2. the minimum risk Bayes' classifier is used with

$$(c_{12}-c_{11})p_1=(c_{21}-c_{22})p_2 \ .$$

The mean vectors and the variance-covariance matrix are estimated on the basis of the design sample set:

$$\hat{m}_i = \frac{1}{n_i} \sum_{j=1}^{n_i} x_{ij}$$

$$\hat{\sum} = \frac{1}{\sum_{i=1}^{g}(n_i-1)} \sum_{i=1}^{g} \sum_{j=1}^{n_i} (x_{ij}-\hat{m}_i)(x_{ij}-\hat{m}_i)^T$$

where n_i is the number of samples in the class i; g is the number of classes; x_{ij} is the j-th sample vector of the class i; $\hat{m}_i$ is an estimated mean vector; and $\hat{\sum}$ is an estimated variance-covariance matrix. These estimates can be shown to be unbiased.

4.3.1.2 Methods of the discriminant analysis

The discriminant analysis includes the estimation of the error classification rate and the feature selection.

Error estimation. In designing a classification rule, its goal is to minimize the loss due to false classification of samples which are unknown to the classifier. Since the underlying distribution is partially or completely unknown in general, the loss must be estimated on the basis of a limited number of samples. It is known that the

samples used in designing a classifier are well discriminated by this classifier. This discrimination test is called the replacement method. Thus, the estimate by the replacement method is biased optimistically. This bias is a monotonically increasing function of the ratio of the number of features to the number of samples. To reduce this bias, the number of the samples must be much more than the number of the features[32]. To obtain an unbiased estimate, the test sample set must be independent of the design sample set. If a large number of samples are available, the independence can be achieved by partitioning the set of all samples into the design set and the test set. Unfortunately, however, in many cases including studies on medical diagnosis, it is difficult or impossible to obtain a large number of samples whose true classes are known. The jack-knife test[29] and the leaving-one-out test[31] are reasonable selection for problems in such situations. Usually, once a useful set of features and a form of the discriminant function are found according to such methods, the discriminant function is finally designed on the basis of all available samples in order to use as much information as possible. In the jack-knife test, a fraction of all the samples are excluded from the design set and tested by the designed classifier. Changing the excluded samples, this design-test cycle is repeated. The error rate is estimated by totaling the results obtained in all repetitions. The leaving-one-out test is a variant of the jack-knife test. In this test, a discriminant function is designed on the basis of the sample set from which one sample is excluded at each repetition. The unbiased property of the leaving-one-out estimate has been proved experimentally[33].

Feature selection. Typically, the discriminant analysis also includes

the selection of the best set of features. The purpose of the feature selection is twofold. One purpose comes from the "principle of parcimony": to save cost for computation. The other is much more important[32]. In designing a linear discriminant function, the coefficients for respective features must be estimated on the basis of a set of available samples, *i.e.*, the design set. Since such samples are limited in number, the estimation involves statistical error. Such error is accumulated linearly in calculation of the discriminant for any test sample. The more the features are used in the discriminant function, the more the erroneous terms become. Therefore, the discrimination accuracy becomes worse even if meaningful features are included. If meaningless features are included, the discrimination is violated more. If the discriminant function is piecewise linear or nonlinear, equivalently it involves more features. Therefore, a harder situation occurs. Thus, the feature selection is essential in the design of a discriminant function.

The strict means for selecting a best feature set is to examine the discriminating performance for all combinations of candidate features. However, this exhaustive search is not feasible in reasonable computation time if there are many candidate features. For such a case, sequential search procedures are known to give suboptimal solutions[33]. Let us consider selection from n candidate features. In the backward sequential search, the cadidates are discarded one by one. At the beginning of the k-th step, $n-k+1$ features are selected, while other $k-1$ features are discarded. At this step, one feature is excluded from the selected set; and the performance of discrimination by other $n-k$ features is evaluated. This evaluation is made for each feature of the

selected set. Then, a feature is discarded if it results the best performance when excluded. At the end of this step, $n-k$ features are selected. On the contrary, in the forward sequential search, the candidates are selected one by one. At the beginning of the k-th step, $k-1$ features are selected, while other $n-k+1$ features are not. At this step, one of the unselected features is appended to the selected set; and the performance of discrimination by this augmented set is evaluated. This evaluation is made for each of the unselected features. Then, a feature is selected if it results the best performance when appended. At the end of this step, k features are selected. In the both search procedures, the best feature set is selected by comparing the discriminating performance at each step. The computation time is less in the forward search, while a better feature set is obtained by the backward search.

In this study, the error rate was estimated by the replacement method and the leaving-one-out method. For the feature selection, the forward sequential search was applied. It is known that, in the course of the forward search, the error rate decreases first and then increases again as the number of features increases[33]. Therefore, in most cases examined, minimal feature sets giving a local minimum of the error rate were adopted in order to save time for calculation.

4.3.1.3 The clustering

The clustering is a method of grouping given samples into assumable number of clusters in a "natural" way. It is useful in inspecting the distribution of samples, particularly when the feature space is of high dimensionality. The clustering is also applied to the feature selection. When a set of features are clustered into groups of

similar features, representative features are selected one by one from respective feature clusters.

The principle of the clustering is to gather samples which are located near in the feature space or which have a high similarity in some sense[34]. The nearness is measured by the inter-sample distance such as the Mahalanobis' distance,

$$(x_r - x_s)^T \sum^{-1} (x_r - x_s)$$

where x_r and x_s are sample patterns, and $\sum$ denotes the variance-covariance matrix. (This variance-covariance matrix is not of respective classes but of the entire distribution, because true classes are assumed unknown.) The Euclidean distance is generally not applicable because measuring units of respective features are different. The measure of similarity is given by the correlation coefficient between samples,

$$\sum_{k=1}^{p} \frac{(x_{rk} - \bar{x}_r)(x_{sk} - \bar{x}_s)}{\sqrt{\sum_{j=1}^{p} (x_{rj} - \bar{x}_r)^2} \sqrt{\sum_{j=1}^{p} (x_{sj} - \bar{x}_s)^2}}, \qquad \bar{x}_r = \frac{1}{n} \sum_{j=1}^{n} x_{rj}$$

where p is the number of the features and x_{rj} is the j-th feature of the r-th sample.

Methods of the clustering are either hierarchical or nonhierarchical. The hierarchical method begins by assuming that each sample forms one cluster. Similar clusters are merged step by step to give a chart like a tournament diagram, which is called the dendrogram. Depending on the definition of the measure of the inter-cluster distance, various methods have been developed such as described below. The non-hierarchical method begins by assuming a number of seed samples. Then, splitting and merging of clusters are iterated until a prescribed convergence criterion is satisfied. Typical non-hierarchical methods

are the K-mean method and the ISODATA method.

In this study, hierarchical methods were applied. Because of the unknown distributing property assumed in the clustering, all sorts of existing hierachical methods were used. These methods differ in the definition of the distance (or the similarity) between samples and between clusters. As to the inter-sample distance, the normalized Euclidean distance, the Mahalanobis' distance, and the inter-sample correlation coefficient were examined; as to the inter-cluster distance, the nearest neighbor method, the farthest neighbor method, the centroid method, the mode method, the median method, the group average method, the flexible method, the Ward's method, and the WPG method were examined[35].[1] In this section, only the results supporting those of the discriminant analysis are given. The cluster property of the samples is shown not only in a form of the dendrogram but also in forms of the confusion matrix and the error classification rate. For the latter forms of representation, a name of a cluster was determined as the name of the majority of the samples classified into that cluster.

The hierarchical methods were also used for inspecting the property of the features with regard to the feature selection. This inspection was made after exchanging the samples with the features. That is, a set of m-dimensional n samples was manipulated as a set of n-dimensional m samples after transposing the data matrix[33].

4.3.2 Feature Extraction from a Gastric Contour and Its Evaluation

Physicians examine the proportion of the entire shape of a stomach at an earlier stage in inspecting its still radiogram.

[1]Only the first four methods were applied when the inter-sample correlation coefficient was used as a measure of the inter-sample distance.

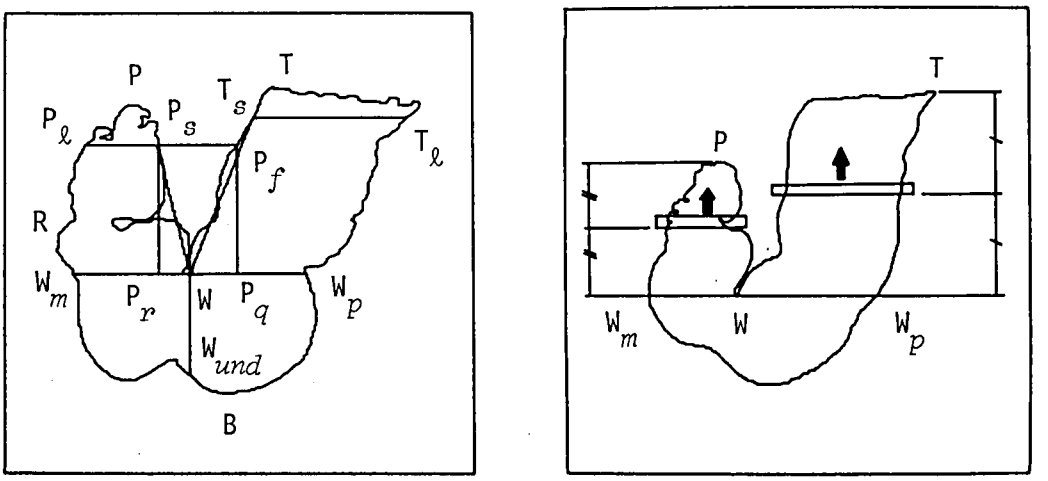
Therefore, some features of this kind may exist, although difficulty may arise in quantification of the shape due to its change through the peristalsis. Such features may represent the first impression when a physician sees a radiogram.

This section describes a method of extracting such features from a gastric contour. The features are calculated automatically without human intervention when a contour is given to computer. Experimental results are shown on the separability of the samples in some feature spaces.

4.3.2.1 Feature extraction

A gastric contour is assumed to be given as a simply-connected closed curve. Some geometrical measurements and Fourier descriptors were examined as candidate features.

Geometrical measurements. The following quantities were calculated on



(a) (b)

Fig. 4.3 Points detected for the measurement.
(a) Points detected.
(b) Determination of the limits.

the basis of the feature points shown in Fig. 4.3.

$$GM1 = \widehat{T_s WP_s} / \widehat{P_l W_{und} T_l}$$

$$GM2 = \widehat{WP_s} / \overline{WW_{und}}$$

$$GM3 = \angle T_s WP_s$$

$$GM4 = \widehat{T_s WP_s} / \overline{W_m W_p}$$

$$GM5 = A_1 / A_2$$

where A_1 is area of the region surrounded by the straight line $\overline{P_s P_f}$ and the curvature $\widehat{P_f WP_s}$; A_2 is area of the rectangle $\square P_f P_q P_r P_s$. The quantities *GM4* and *GM5* are calculated only for fishhook-shaped stomachs. The measurements were devised to meet the expertise knowledge of physicians and radiologists: the quantities *GM1*, *GM2*, and *GM4* are related to the shrinking of the lesser curvature; the quantities *GM3* and *GM5* are related to the deformity of the apex. Note that *GM1*, *GM2*, *GM4*, and *GM5* are normalized quantities which are independent of the size of the stomach. The feature points are defined by an algorithm as follows.

a) The apex: The apex W is determined by a counterclockwise search of the gastric contour. This search starts from the uppermost point T and finds a point where the gradient to the vertical axis

$$\{x(j+1)-x(j-1)\} / \{y(j+1)-y(j-1)\}$$

exceeds a threshold θ_{tw} for the first time. The search is iterated from roughly smoothed contours to moderate ones, and then finally to the original, in order to avoid false detection of any point on irregular portions of the contour. At each iteration, the search begins with a neighbor of the point detected by the previous search. Validity of the obtained point is further checked as to its relative position in the circumscribing rectangle of the gastric contour. Suppose that the position of this point is

(x_w, y_w) and that $\cdot_{min}$ and $\cdot_{max}$ denote the minimum and the maximum of the coordinate values of the contour points. Then if

$$(y_w - y_{min}) / (y_{max} - y_{min}) > \theta_{wy}$$

that point is rejected and the search is further continued; if

$$(x_w - x_{min}) / (x_{max} - x_{min}) < \theta_{wx}$$

that point is rejected and, instead, a point straight above the lowermost point B is selected. The former check is to avoid detecting points in irregular portions of the surface of contrast material; the latter allows for steer-horn shaped stomachs. The apex determined by this algorithm agrees well with one identified by physicians.

b) The points W_m, W_p, and W_{und}: Each of these points is specified on the greater curvature, having its position (x,y) satisfying either $x=x_w$ or $y=y_w$.

c) The cardia limit $\overline{T_l T_s}$: This line is established at the level where abrupt change more than θ_{cl} occurs in the number of the contour points contained in a horizontal slit (see Fig. 4.3,) as the slit moves upwards at intervals of five points, starting from the middle level between the apex W and the uppermost point T. The length of the narrow side of the slit is 25 points; the length of its long side is virtually adjusted to cover only the gastric body at each level of the slit. (Actually, this is equivalently done by counting the contour points of the gastric body.)

d) The antrum limit $\overline{P_l P_s}$: If the stomach is of the fishhook shape, this line is established at the level where abrupt change more than θ_{al} occurs in the number of the antrum contour points included in a horizontal slit, as it moves upwards starting from the middle level

between the apex W and the point P which is located uppermost on the curvature between the apex and the point R located on the extreme left. If the stomach is of the steer-horn shape, a vertical slit is used: it moves similarly but to the left, locating the points P_l and P_s on the both sides of the point R.

e) The point P_f: the intersection between the elongation of the line $\overline{P_l P_s}$ and the lesser curvature.

Although the cardia and the pylorus may be used as feature points in order to represent the shape of a stomach numerically, the cardia is not clearly observable and the pylorus is often occluded by the shadow of the pyloric antrum in radiograms. Therefore, in this case, positions of these points must be estimated or some subsidiary points must be used. One approach is to measure only a portion which is filled with contrast material. The purpose of establishing the two limits described above is to avoid, in the measurement, the influence of the irregular surface of contrast material in this approach.

In order to determine the shape of a stomach, the level of the point P is examined: if

$$(y_p - y_w) / (y_w - y_{min}) > \theta_{sh} ,$$

then the shape is of the fishhook; otherwise, it is of the steer-horn.

In the present version of the program, θ_{tw}=1.732 (60 degrees,) θ_{wy}=0.75, θ_{wx}=0.2, θ_{cl}=25 points, θ_{al}=10 points, and θ_{sh}=0.15.

Fourier descriptors. A plane closed curve can be represented by a set of trigonometric functions by the Fourier expansion, a method of functional approximation of periodic functions. Fourier descriptors are defined on the basis of this expansion. Two typical sets of the Fourier descriptors are one for the cumulative angular function[15] and one for

the complex-domain representation[36]. The latter was used here, which is related to the concept of Lissajous' figure.

A plane curve can be represented in a complex plane by $z(t) = x(t)+y(t)\sqrt{-1}$, where t is the path length from an arbitrarily chosen starting point of the curve; $x(t)$ and $y(t)$ denote the position of a point in a real Cartesian coordinate system; and $z(t)$ denote that in a complex Cartesian coordinate system. If the curve is closed, the complex function $z(t)$ is periodic and

$$z(t+T) = z(t) \ .$$

Therefore, it can be expanded into a Fourier series, thus

$$z(t) = \sum_{n=-\infty}^{\infty} c_n \ exp(jn\omega_0 t)$$

where

$$c_n = \frac{1}{T} \int_T z(t) \ exp(-jn\omega_0 t) \ dt$$

and $\omega_0 = \frac{2\pi}{T}$. In a discrete system,

$$z(l) = \sum_{k=0}^{N-1} c(k) \ exp(j\frac{2\pi}{N} kl)$$

$$c(k) = \frac{1}{N} \sum_{l=0}^{N-1} z(l) \ exp(-j\frac{2\pi}{N} kl) \ , \quad k=0,1,2,\ldots,N-1,$$

where $c(k)$ and $z(l)$ correspond to c_n and $z(t_0+l\Delta t)$, with Δt being the sampling length along the path. The coefficients $c(k)$'s represent proportion and phase shift of the sinusoidal waves of respective frequencies. By selecting a limited number of coefficients, the closed curve is approximated by a set of sinusoidal waves. A class of Fourier descriptors is defined on the basis of this expansion[36]. In this experiment, Fourier descriptors were defined by

$$FDk = |c(k+1) / c(1)| \ , \quad k=1,2,\ldots,28,$$

where $c(k)$ is the coefficient of the k-th harmonic. To exclude the

effect of the location and the size of a stomach, the zero-th order harmonic was not used and the other harmonics were normalized by the first order harmonic.

4.3.2.2 Analyzed data

Gastric contours of the sample radiograms were traced by persons without medical training. This means that we assumed the precision of the contour which would be recognized without expertise on the gastric radiography. (It should be noted that the gastric contour can not always be recognized clearly. The precision assumed here is a reasonable goal of the contour extraction, if it is to be performed by computer.) Each of the tracings were represented as a point sequence in a digital coordinate system. This representation is the same as one directly obtained by counterclockwise boundary-tracking in digital image processing. The interval between vertically or horizontally adjacent pixels is about one millimeter in the scale of a direct radiogram (which is of about the same scale as a real body.)

Two types of gastric contours may be traced from a radiogram: the contour of the gastric wall and that of contrast material. These are distinctly recognized, particularly, in portions of the upper body and of the pyloric antrum, because contrast material is not filled completely in these portions. The both types were traced, if possible, in order to see

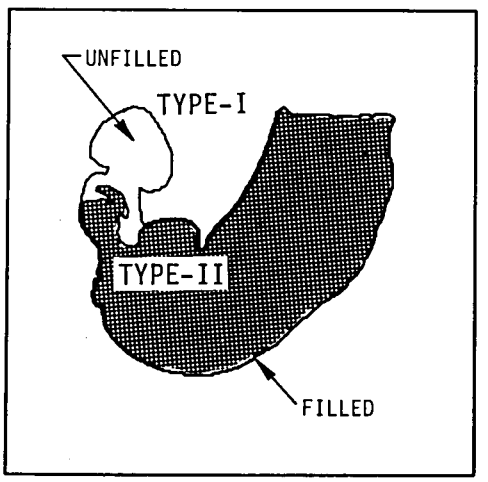
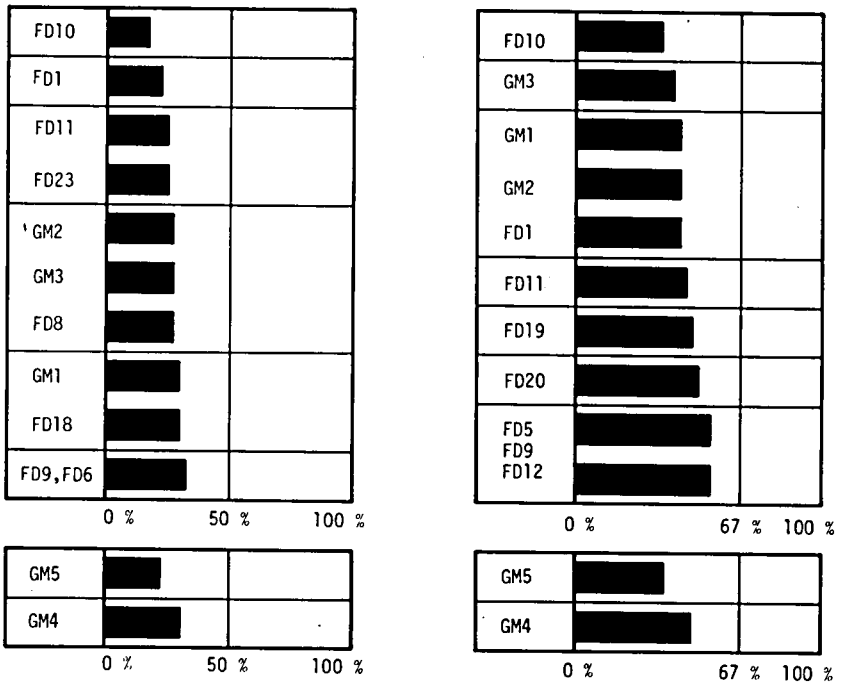


Fig. 4.4 Types of the gastric contour.



(a) (b)

Fig. 4.5 The error rate in classification of the type-I contours using each of the discriminating features. The features GM4 and GM5 were evaluated for only the fishhook shaped stomachs. (a) N/A classification. (b) N/U/C classification.

the sensitivity of the classification error to the precision of the contours. The portion of the fundus was not traced, because its contrast is weak and also because it is truncated by the frame of a radiographic film. Hereafter, a contour of the gastric wall is called the type-I; one of contrast material is called the type-II. (See Fig. 4.4.) It seems very difficult for computer to extract type-I contours.

Forty-six type-I contours and forty-seven type-II contours could be traced as closed curves from indirect radiograms of the fifty-five

Table 4.1 The error rate in classification of the gastric contours using the most discriminating feature set

(a) Both the fishhook shaped and the steer-horn shaped stomachs

Cases				
Contour type	I	I	II	II
Sample size	40	40	44	44
Classification	N/A	N/U/C	N/A	N/U/C
Candidate features				
31V	**7.5 %	***25.0 %	22.7 %	36.4 %
8PC of 31V	12.5	27.5	25.0	34.1
28FD	(12.5)	(25.0)	(25.0)	(40.9)
8PC of 28FD	15.0	30.0	31.8	45.5
6 M.D.s	11.3	29.2		

(b) Only the fishhook shaped stomachs

Cases				
Contour type	I	I	II	II
Sample size	36	36	40	40
Classification	N/A	N/U/C	N/A	N/U/C
Candidate features				
33V	8.3 %	25.0 %	10.0 %	30.0 %
8PC of 33V	11.1	27.8	22.5	37.5

Notes:

1. The first column denotes sets of candidate features for the feature selection: the candidate features are GM1, GM2, GM3, and the Fourier descriptors (31V); the first eight principal components (8PC); the Fourier descriptors (28FD); GM4, GM5, and 31V (33V).

(Continued on the next page.)

Table 4.1 (Continued.)

The result of classification by six physicians is shown in the row denoted by "6 M.D.s."

2. The best cases are indicated by asterisks: the features are FD10, FD1, GM2, and FD19 (**); FD10 and GM2 (***).
3. Parentheses indicate that the order of selecting features was determined by the replacement method in order to save time for calculation. (The number of features composing the best set and the error rate were determined by the leaving-one-out method.)

samples which were available. Forty type-I contours and forty-four type-II contours were analyzed which satisfy an assumption of the feature extraction program: the highest point of the contour is located in the side of the fundus rather than of the pylorus. The two types were analyzed separately.

4.3.2.3 Results

Scattergrams. Distribution of the samples was examined through its projections onto some two-dimensional feature spaces including ones composed of principal components of the features. The projections implied feasibility of separation of the diseased samples from the normal ones by a linear classifier.

Discriminant analysis. Fig. 4.5 shows error rates of more discriminating features obtained in classification of the type-I contours, when the features were used individually. Table 4.1 shows

error rates attained by best sets of features. These features were selected, in the N/A classification, from among those whose error rates were less than forty percent when used individually; or, in the N/U/C classification, less than 56.7 percent. Table 4.1 includes other cases of the feature selection: from the first eight principal components of all features, from only the Fourier descriptors, and from the first eight principal components of the Fourier descriptors. The type-I contours were also inspected by six physicians independently, who were experienced or less experienced. This result is also included. Table 4.2 shows confusion matrices estimated by the leaving-one-out method and by the replacement method in classification by the selected feature set; while Table 4.3 shows actual confusion in classification by the physicians.

Clustering. The samples were clustered in a variety of merging processes depending on the methods. However, separation was recognized between the normal samples and the diseased ones through most dendrograms; separation between the cancer samples and the ulcer samples was not clear. As shown in Table 4.4(a), the error rate in the best N/A classification was 22.5 percent for the set of the type-I contours containing both the fishhook shaped and steer-horn shaped stomachs, in the original 31-feature space; Fig. 4.6(a) shows the dendrogram in this case. For the set of only the fishhook shaped stomachs, the error rate was 13.9 percent in the original 33-feature space. However, smaller values were obtained when the clustering was performed in reduced feature spaces composed of features selected by the discriminant analysis; the error rate was approximately ten percent as shown in Table 4.4(b). In this case, separation of the clusters was more apparent as

shown in the dendrogram of Fig. 4.6(b). The separation of the type-II contours was generally worse. When the steer-horn shaped stomachs were included, the error rate was 27.3 percent in the selected feature space; for only the fishhook shaped stomachs, it was 12.5 percent. The error rates all agreed well with those obtained by the discriminant analysis.

For the type-I contours, the clustering of the features resulted in a dendrogram shown in Fig. 4.7. As pointed out by Toriwaki *et al.*[33], the features selected through the discriminant analysis tended to belong to different clusters. This property was clearly recognized in the N/A classification and supports the validity of the selected features.

4.3.3 Feature Extraction from an Apex Image and Its Evaluation

This section decribes another approach to the feature extraction, which is rather mathematical than pattern-based: no decision is made as preprocessing such as feature point detection. The underlying concept used here is pattern matching. The hypothesis was that pattern matching works in classification of certain local regions even if the shape of a stomach changes as a whole due to the peristalsis. Lower sequency or frequency components of a two-dimensional orthogonal transform of an apex image were used for the purpose of discriminating abnormal changes of the apex.

4.3.3.1 Feature extraction

Assume that an image is represented as a matrix $[F]$ whose elements are gray levels of corresponding pixels. A two-dimensional orthogonal transform, or unitary transform to be precisely, of this image is given by

$$[\alpha] = [U]^T[F][V] \tag{4.3}$$

Table 4.2 Confusion matrices of classification of the type-I contours of both the fishhook shaped and the steer-horn shaped stomachs by the discriminant analysis

		Assigned classes		
		N	A	Error
Verified	N	14	2	12.5 %
diagnoses	A	1	23	4.2
Overall error classification				7.5

Sample size = 40

Feature set = (FD10,FD1,GM2,FD19)

		Assigned classes			
		N	U	C	Error
Verified	N	12	2	2	25.0 %
diagnoses	U	1	9	3	30.8
	C	1	1	9	18.2
Overall error classification					25.0

Sample size = 40

Feature set = (FD10,GM2)

Note: Identical matrices were obtained by both the leaving-one-out method and the replacement method.

Table 4.3 Confusion matrices of classification of the type-I contours of both the fishhook shaped and the steer-horn shaped stomachs by six physicians

		Assigned classes		
		N	A	Error
Verified	N	88	8	8.3 %
diagnoses	A	19	125	13.2
Overall error classification				11.3

Sample size = 40×6

		Assigned classes			
		N	U	C	Error
Verified	N	88	5	3	8.3 %
diagnoses	U	13	56	9	28.2
	C	6	34	26	60.6
Overall error classification					29.2

Sample size = 40×6

Table 4.4 Confusion matrices of classification of the type-I contours of both the fishhook shaped and the steer-horn shaped stomachs by the clustering

(a) Clustering in the 31-feature space

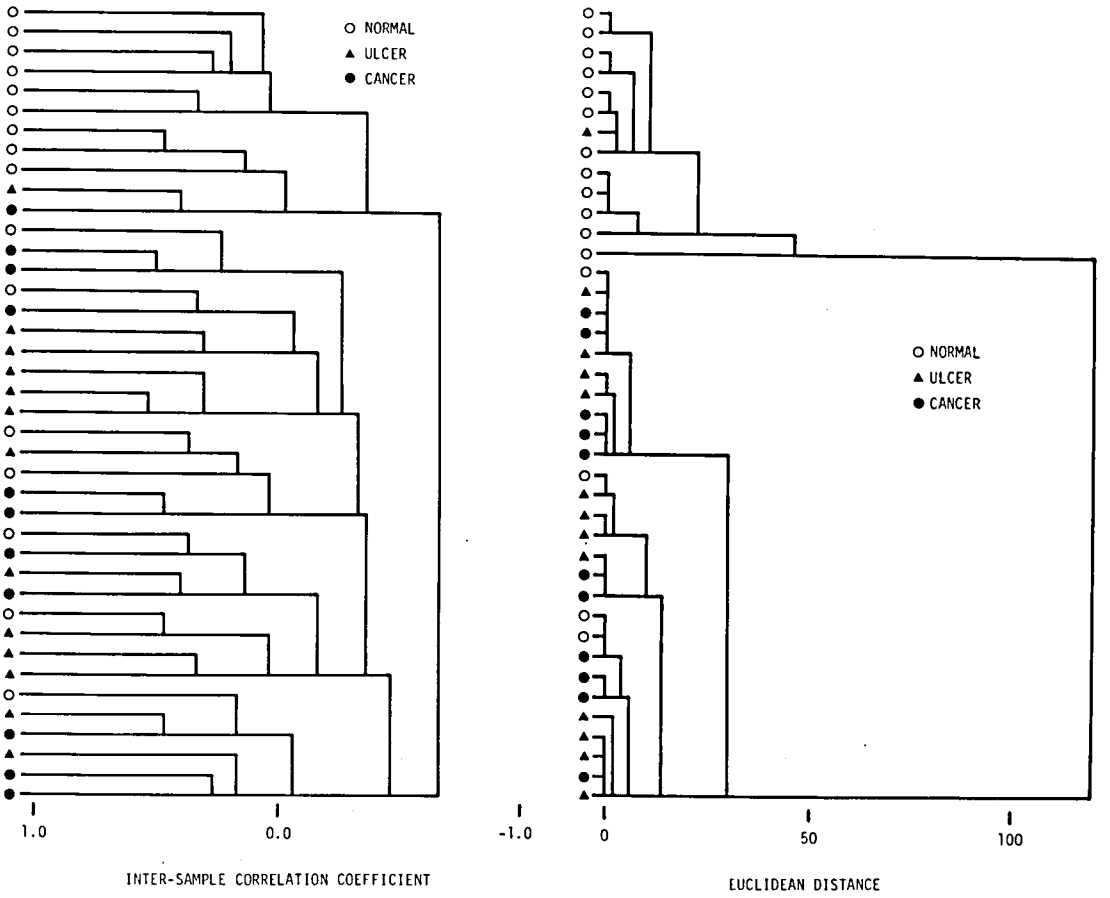
		Assigned classes		
		N	A	Error
Verified	N	9	7	43.8 %
diagnoses	A	2	22	8.3
Overall error classification				22.5

Sample size = 40
Feature set = (GM1,GM2,GM3,FD1,..,FD28)
The farthest neighbor method with the inter-sample correlation coefficient measure.

(b) Clustering in the selected feature space

		Assigned classes		
		N	A	Error
Verified	N	12	4	25.0 %
diagnoses	A	1	23	4.2
Overall error classification				12.5

Sample size = 40
Feature set = (FD10,FD1,GM2,FD19)
The Ward's method with the normalized Euclidean distance measure.



(a) (b)

Fig. 4.6 Dendrograms of the clustering of the type-I contours.
(a) Clustering in the 31-feature space. Feature set = (GM1, GM2, GM3, FD1,...,FD28). Method: the farthest neighbor method with the inter-sample correlation coefficient measure.
(b) Clustering in the selected feature space. Feature set = (FD10, FD1, GM2, FD19). Method: the Ward's method with the normalized Euclidean distance measure.

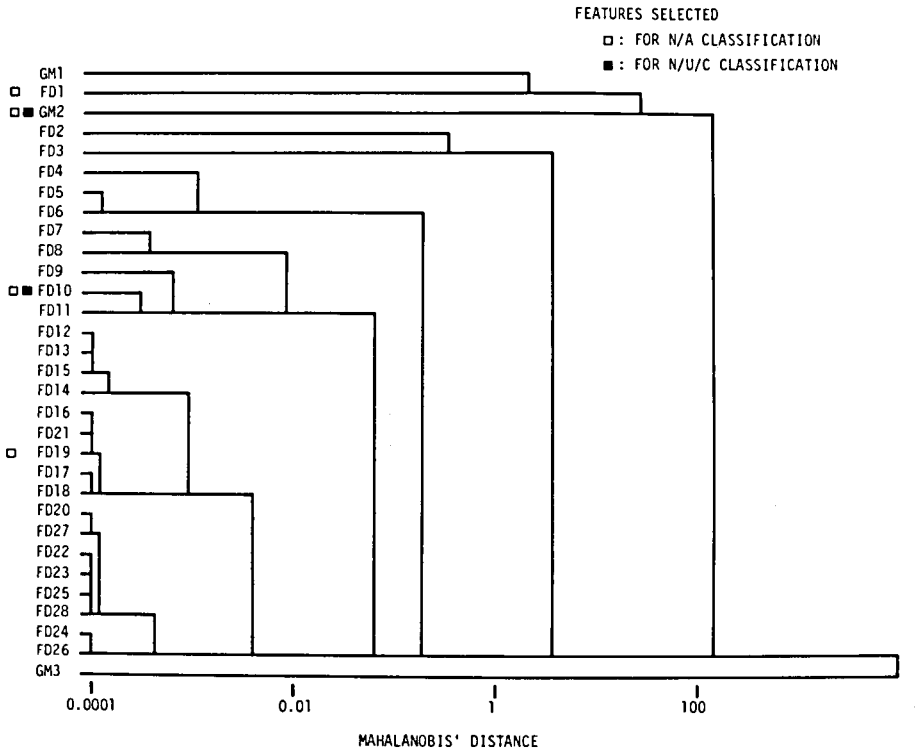


Fig. 4.7 A dendrogram of the clustering of the contour features. The marks indicate the selected best features. Samples: the type-I contours of both the fishhook shaped and the steer-horn shaped stomachs. Method: the flexible method with the Mahalanobis' distance measure.

where the matrices $[U]$ and $[V]$ are unitary matrices deduced from complete sets of orthogonal functions of the transform. Let us define

$$[U]=(u_0,u_1,\ldots,u_{N-1})$$

and

$$[V]=(v_0,v_1,\ldots,v_{N-1}),$$

where the column vectors $u_0,u_1,\ldots,u_{N-1}$ and $v_0,v_1,\ldots,v_{N-1}$ are sampled version of the orthogonal basis functions. Rewriting the expression (4.3), we obtain an expansion of the original image into a set of the basis images:

$$[F]=[\overline{U}][\alpha][\overline{V}]^T$$

or

$$[F]=\sum_{i=0}^{N-1}\sum_{j=0}^{N-1}\alpha_{ij}u_i v_j^T$$

where α_{ij}'s are the components of the matrix $[\alpha]$, and the bar over the matrices indicates the complex conjugate. If the basis functions are arranged in ascending order of the frequency or sequency, the coefficients represent proportions of respective basis images $u_i v_j^T$ in order of roughness, *i.e.*, from rough appearances to modifying details. The meaning of the feature extraction by an orthogonal transform can be better understood by interpreting the extraction of lower frequency or sequency components as approximating the original image by a limited number of basis images of the transform. Its visual meaning is to take a vague or unfocused observation of the image. Therefore, the method may extract features representing a first impression of a physician when he sees the image. The information retained by lower sequency components is proportions of more monotonous basis images, which form a blurred version of the original image.

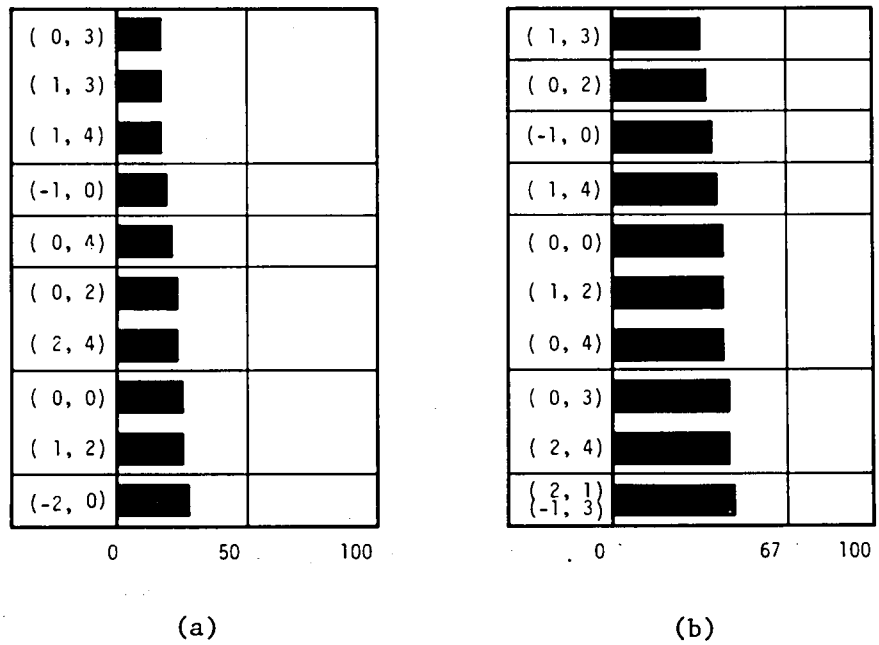
In order to see the sensitivity of the classification error to the type of the transform, two types were examined: the slant transform[37,38] and the Fourier transform[39]. These were chosen because they are commonly used or suboptimal in data compression of gray-level images. The both were programmed according to the fast algorithms in ordered form[40,41]. The Fourier transform was applied to both original images and images preprocessed by the Hanning window, which is often used to reduce the spectrum leakage. The spectrum leakage is a phenominon caused by limiting the space band, or the window, for calculation of the transform. In order to reduce this

effect, various windows with tapered outskirts have been developed[42]. The Hanning window is a weighting function given by

$$w_{ij} = [0.5-0.5\ cos\{2\pi i / (N-1)\}][0.5-0.5\ cos\{2\pi j / (N-1)\}]$$

Its visual meaning is to emphasize the central portion of an image.

Moduli of lower sequency components were examined as candidate features: in the case of the Fourier transform, thirty-five components contained in a semicircle centered on the zero-th component which was shifted to the center of the transformed matrix; in the case of the slant transform, thirty components contained in a quadrant centered on



(a) (b)

Fig. 4.8 The error rate in classification of the apex images using each of the discriminating features obtained by the Hanning window-Fourier transform operation. The numbers in parentheses represent the spatial frequency of a feature as "(the vertical frequency, the horizontal frequency)." (a) N/A classification. (b) N/U/C classification.

Table 4.5 The error rate in classification of the apex images using the most discriminating feature set

Classification	N/A	N/U/C
Candidate features		
35FT	14.9 %	34.0 %
8PC of 35FT	17.0	36.2
35H.FT	**6.4	***21.3
8PC of 35H.FT	10.6	23.4
30ST	12.8	31.9
8PC of 30ST	19.1	34.0
M.D.	2.1 - 21.3 %	19.1 - 36.2 %

Sample size = 47

Notes:

1. The first column denotes sets of candidate features for the feature selection: the candidate features are the Fourier transform features (35FT); the first eight principal components (8PC); the Hanning window-Fourier transform features (35H.FT); the slant transform features (30ST). The result of classification by one physician is shown in the row denoted by "M.D." Two figures in both sides of a hyphen represent a range of his error rate.
2. The best cases are indicated by asterisks: the features are $|\alpha_{1,4}|$, $|\alpha_{-1,0}|$, and $|\alpha_{0,0}|$(**); $|\alpha_{1,3}|$, $|\alpha_{-1,0}|$, $|\alpha_{0,0}|$, $|\alpha_{3,2}|$, $|\alpha_{-1,4}|$, and $|\alpha_{-1,2}|$ (***). $|\alpha_{i,j}|$ denotes a feature whose spatial frequency is (i,j).

Table 4.6 Confusion matrices of classification of the apex images by the discriminant analysis

		Assigned classes		
		N	A	Error
Verified diagnoses	N	18	2	10.0 %
	A	1	26	3.7
Overall error classification				6.4

Sample size = 47

Feature set = $(|\alpha_{1,4}|, |\alpha_{-1,0}|, |\alpha_{0,0}|)$

Hanning window-Fourier transform.

Leaving-one-out/replacement method.

		Assigned classes			
		N	U	C	Error
Verified diagnoses	N	19	1	0	5.0 %
	U	2	10	4	37.5
	C	1	2	8	27.3
Overall error classification					21.3

Sample size = 47

Feature set = $(|\alpha_{1,3}|, |\alpha_{-1,0}|, |\alpha_{0,0}|, |\alpha_{3,2}|, |\alpha_{-1,4}|, |\alpha_{-1,2}|)$

Hanning window-Fourier transform.

Leaving-one-out method.

(Continued on the next page.)

Table 4.6 (Continued.)

		Assigned classes			
		N	U	C	Error
Verified	N	19	1	0	5.0 %
diagnoses	U	2	11	3	31.3
	C	0	2	9	18.2
Overall error classification					17.0

Sample size = 47

Feature set = ($|\alpha_{1,3}|, |\alpha_{-1,0}|, |\alpha_{0,0}|, |\alpha_{3,2}|, |\alpha_{-1,4}|, |\alpha_{-1,2}|$)

Hanning window-Fourier transform.

Replacement method.

the zero-th component.

4.3.3.2 Analyzed data

The sample radiograms were enlarged, and their printed positives were read into the computer through a vidicon TV camera. Only the region of the apex was digitized. Each of the digitized images consists of 64X64 pixels; the gray level was quantized into about one hundred actual levels. The positives were placed in a fixed position in order to avoid the shading effect of the TV camera. Ambiguous images had been discarded, and therefore, the analyzed images totaled forty-seven. Gray levels of the input images were linearly normalized so that the minimum and the maximum values would become constants.

4.3.3.3 Results

Scattergrams. Distribution of the samples was examined through its

projections onto some two-dimensional spaces composed of principal components of the features. The projections implied feasibility of separation of the diseased samples from the normal ones by a linear classifier.

Discriminant analysis. The samples were best classified when the Fourier transform was applied to the Hanning-windowed images. Fig. 4.8 shows error rates of more discriminating features in this case, when the features were used individually. Table 4.5 compares error rates attained by best sets of features which were selected in the same way as the best sets of the contour features. The printed positives of the samples were inspected also by an experienced physician. His judgement was given with a certainty index as in the common manner of medical diagnosis. Table 4.5 includes also the error rate of the judgement, which is represented by a range obtained by optimistic and pessimistic assessment of the judgement. In the pessimistic assessment, a judgement was classified to be correct only if he is certain of it; in the optimistic assessment, a judgement was classified to be correct even if he is uncertain. Table 4.6 shows confusion matrices estimated by the leaving-one-out method and by the replacement method in classification by the selected feature set; while Table 4.7 shows actual confusion in classification by this physician.

Clustering. The trend of the clustering of both the samples and the features was similar with the trend in the case of the gastric contours. Although the samples were clustered in a variety of merging processes, separation was recognized between the normal samples and the diseased ones through most dendrograms; separation between the cancer samples and the ulcer samples was not clear. The error rate in the N/A

classification was 10.6 percent in the original 35-feature space in the Hanning window-Fourier transform case (see Table 4.8(a)); Fig. 4.9(a) shows a dendrogram in this case. In the space composed of the features selected by the discriminant analysis, the error rate was obtained as shown in Table 4.8(b). These values agree well with the result of the discriminant analysis. As shown in Fig. 4.9(b), separation was recognized more clearly in the dendrogram obtained by the clustering in the selected feature space than in the dendrogram of Fig. 4.6 for the gastric contours.

The clustering of the Hanning window-Fourier transform features resulted in a dendrogram shown in Fig. 4.10. Here again, the selected features tended to belong to different clusters. This trend was also recognized in the clustering of the slant transform features.

4.3.4 Discussion

Some similarities were noted between the results of the experiments on the gastric contours and on the apex images:

a) The error rates in classification by the best feature sets were comparable to those by the physicians.

b) The samples were separable both in the discriminant analysis and in the clustering, similarly: the normal samples and the diseased ones were clearly separable, while the cancer samples and the ulcer samples were not.

Since the experiments were conducted to prove the feasibility of classification, no sophistication was made about its method. We may expect better accuracy of classification by elaborating the form of the discriminant function and the method of the feature selection. Moreover, some papers also reported successful results in classification

Table 4.7 Confusion matrices of classification of the apex images by a physician

		Assigned classes		
		N	A	Error
Verified diagnoses	N	20	0	0.0 %
	A	1	26	3.7
Overall error classification				2.1

Sample size = 47

Optimistic assessment.

		Assigned classes		
		N	A	Error
Verified diagnoses	N	18	2	10.0 %
	A	8	19	29.6
Overall error classification				21.3

Sample size = 47

Pessimistic assessment.

(Continued on the next page.)

Table 4.7 (Continued.)

		Assigned classes			
		N	U	C	Error
Verified	N	20	0	0	0.0 %
diagnoses	U	0	13	3	18.8
	C	1	5	5	54.5
Overall error classification					19.1

Sample size = 47

Optimistic assessment.

		Assigned classes			
		N	U	C	Error
Verified	N	18	2	0	10.0 %
diagnoses	U	2	11	4	31.3
	C	6	5	1	90.9
Overall error classification					36.2

Sample size = 47

Pessimistic assessment.

Note: It was permitted for the physician to assign more than two classes to one sample when it was impossible to make definite judgement. Thus, the error rate here is defined by 1-(correct classification rate).

Table 4.8 Confusion matrices of classification of the apex images by the clustering using the Hanning window-Fourier transform features

(a) Clustering in the 35-feature space

		Assigned classes		
		N	A	Error
Verified	N	19	1	5.0 %
diagnoses	A	4	23	14.8
Overall error classification				10.6

Sample size = 47

Feature set = (35 features)

The centroid method with the inter-sample correlation coefficient measure.

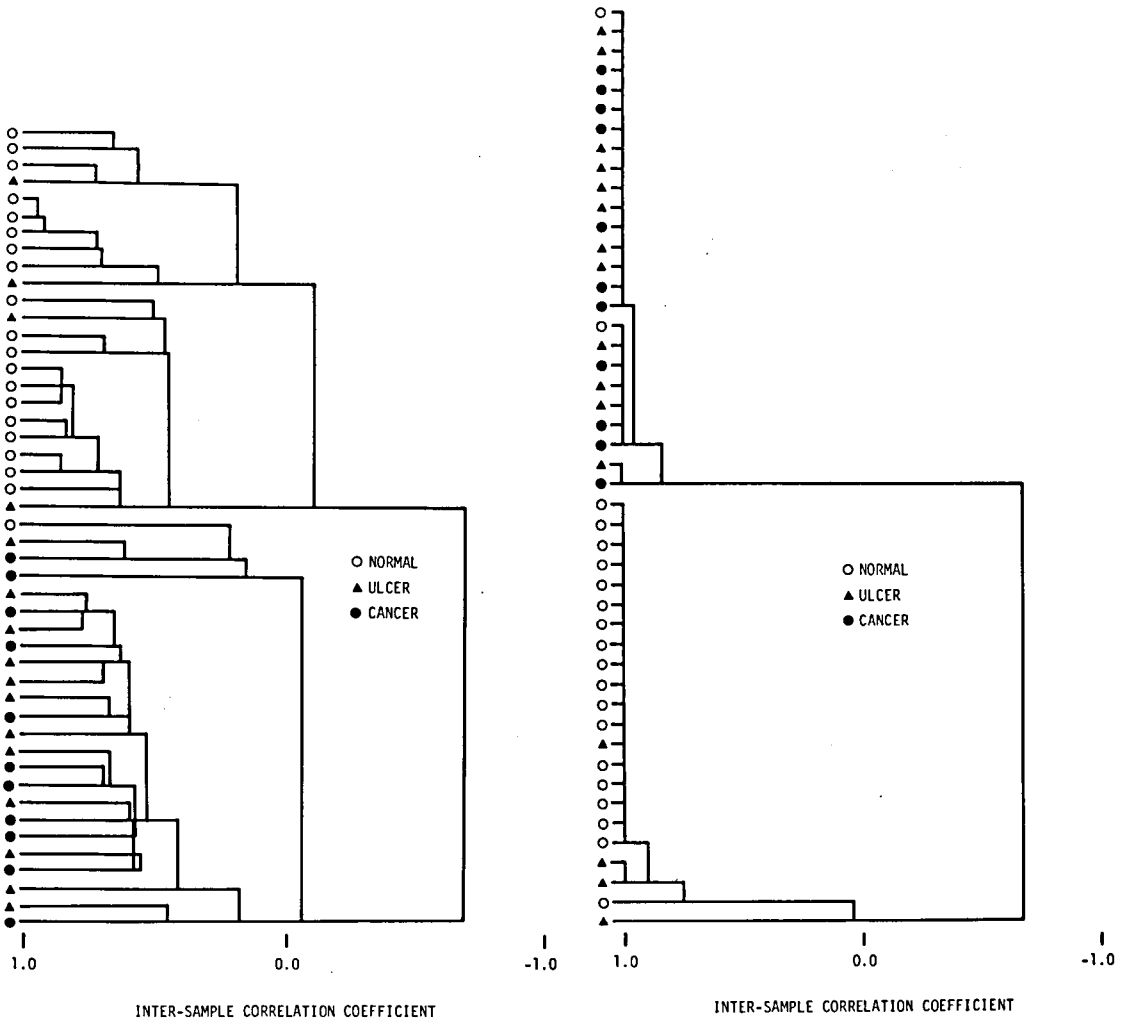
(b) Clustering in the selected feature space

		Assigned classes		
		N	A	Error
Verified	N	18	2	10.0 %
diagnoses	A	4	23	14.8
Overall error classification				12.8

Sample size = 47

Feature set = $(|\alpha_{1,4}|, |\alpha_{-1,0}|, |\alpha_{0,0}|)$

The centroid method with the inter-sample correlation coefficient measure.



(a) (b)

Fig. 4.9 Dendrograms of the clustering of the apex images using the Hanning window-Fourier transform features.

(a) Clustering in the 35-feature space. Method: the centroid method with the inter-sample correlation coefficient measure.

(b) Clustering in the selected feature space. Feature set = $(|\alpha_{1,4}|, |\alpha_{-1,0}|, |\alpha_{0,0}|)$. Method: the centroid method with the inter-sample correlation coefficient measure.

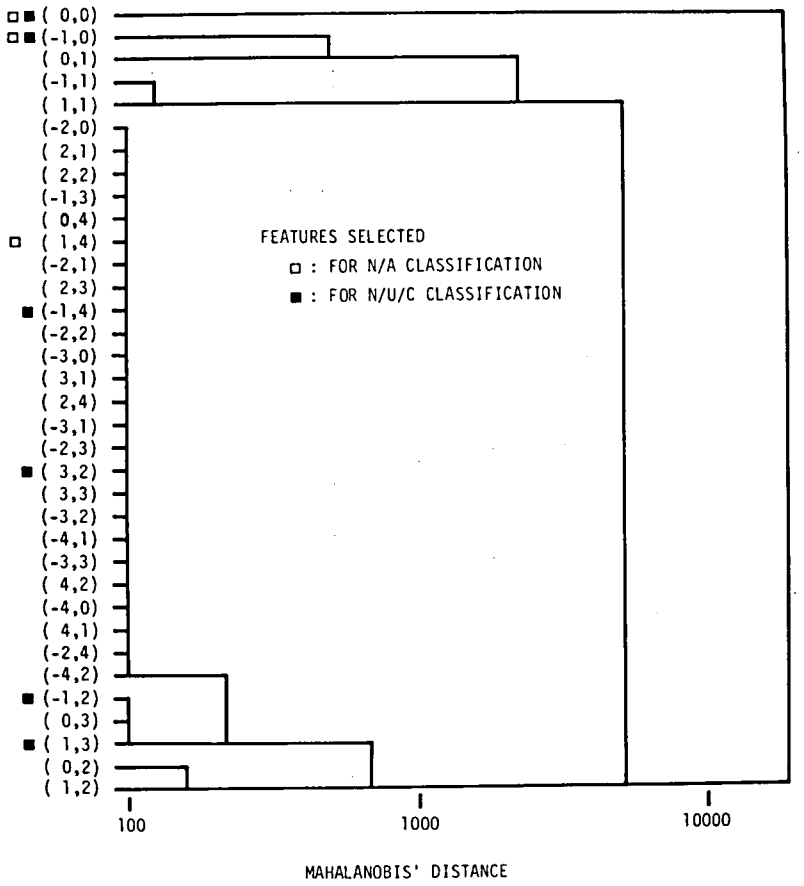


Fig. 4.10 A dendrogram of the clustering of the Hanning window-Fourier transform features. The marks indicate the selected best features. Method: the median method with the Mahalanobis' distance measure.

of the apex pattern[9,14]. Therefore, the N/A classification of the apex will be one of the most feasible tasks in the computerized interpretation of gastric radiograms. (Remember that the apex is easily detected by computer.) However, when applying a discriminant method to mass screening, the *a priori* probability and the loss function must be incorporated into the discriminant function. Thus, the method should be based on, for example, the Bayes' classification rule.

The apex images were discriminated a little better than the gastric contours. This may be because of limiting the abnormalities to those of the apex. A few more properties were noted in the respective cases. As to the gastric contours,

a) the type-I contours were classified better than the type-II contours, and

b) each of the features (the five geometrical measurements and some of the Fourier descriptors) was discriminating.

As to the apex images, the following were noted:

a) The error rate did not depend much on the type of the orthogonal transform. However, the Hanning window was very useful.

b) The same discriminant analysis was applied to another class of gastric radiograms of the same samples: the prone position images of the stomachs filled with contrast material. The classification accuracy was worse than that of the SPAP images. The error rate was 27.5 percent in the N/A classification; 43.1 percent in the N/U/C classification. The same physician who examined the SPAP images resulted the error rate of 7.8-29.4 percent in the N/A classification; 21.6-41.2 percent in the N/U/C classification. Therefore, this class of gastric radiograms contains subsidiary

information. This may be because the natural shape is deformed in taking the radiograms. Details of this examination are described in Appendix C.

As to the feature selection, the experiments resulted in a trend similar to one reported recently[33], particularly, in the N/A classification which was successful:

a) The feature selection from the original features was compared with that from their first eight principal components. The result favored the former. In addition, in the latter case, the order of selecting the features was not in order of the principal component sequence. Therefore, the principal component analysis is not appropriate in the feature selection, or in the dimensionality reduction of the feature space.

b) The selected features tended to belong to the different feature clusters. This means that these features were not similar in a sense of the clustering; and that they were selected in a complementary manner regardless of the discriminating capability of the respective features.

In the clustering, two specific properties were found. These properties are noted here for documentation.

a) Among the hierarchical methods, the nearest neighbor method and the mode method resulted much confusion. This may be due to the chaining effect and the dispersion of the samples, respectively.

b) Better separation was realized in the clustering of the samples, when the inter-sample correlation coefficient was used as a measure of similarity. Probably, this means that the samples are dispersing on both sides of a hyperplane in the feature spaces

rather than forming spherical clusters.

Although the estimated error rates agreed well with those of physicians, a problem must be noted lastly. That is, more samples must be examined in order to evaluate the generality of the results, particularly, the bias of the sample distribution from the natural distribution. (The bias in the error rate estimation is expected to be minimum, because the leaving-one-out method was used.)

4.4 Summary of the Chapter

This chapter has been devoted to the feature extraction and the classification of a stomach. In the descriptive approach, local properties of a gastric contour were examined to describe its appearance. The gastric axis was established to discriminate abnormal incisures from the peristalsis. Contextual interpretation of the description is still an open problem. In the quantitative approaches, the statistical analyses were applied to a set of confirmed samples. By the discriminant analysis based on the minimum Mahalanobis' distance classification rule, the geometrical measurements and some of the Fourier descriptors have been proved of their usefulness in classification of gastric contours. The error rate was estimated by the leaving-one-out method, and then the best sets of discriminating features were selected by the forward sequential search. In the best N/A classification, the estimated error rate was 7.5 percent; in the best N/U/C classification, it was 25 percent. By the same method of the discriminant analysis, usefulness was also proved for features obtained by two-dimensional orthogonal transforms of the apex images. Particularly, the Hanning window-Fourier transform application was noted of its

discriminating contribution. The selected best feature set for N/A classification attained the error rate of 6.4 percent; the best feature set for N/U/C classification attained that of 21.3 percent. The figures agreed well with the results of the clustering and inspection by physicians. The selected features tended to belong to different clusters of the candidate features. This may be a support of the feature selection. Computational feasibility was thus confirmed as to classification of the deformity of the apex, a portion of medical importance.

A computer program has been designed for analyzing a gastric contour. It was called the routine R11 in Section 3.5. Its final version has been implemented on a small computer. Given a contour, it determines the feature points, makes the deviation curve, calculates quantitative features, and displays intermediate results on CRT. It also calculates specified coefficients of the forward Fourier transform of a contour as well as calculates the inverse Fourier transform for specified coefficients.[1] A synthesized figure is displayed on CRT, giving a useful means in analyzing the information relevant to the specified harmonics by visual inspection of the figure.

---

[1] Care must be and was taken in calculation of the Fourier descriptors by the fast Fourier transform (FFT) algorithm. See Appendix A.2.

CHAPTER 5 CONCLUSION

This dissertation discussed an overall view on approaches to computerizing the interpretation of SPAP images, an important class of gastric radiograms. The goal of this study has been mainly to construct a system for screening gastric diseases. A configuration of the diagnosing system was conceptualized for the first time through this study, where two major tasks were postulated: the extraction of the gastric contour and its interpretation to make decision on abnormalities. Also, the concept of the gastric axis was introduced in order to manipulate structural information of a stomach. Conclusion is summarized below.

Contour extraction:

1. A method was developed for the contour extraction. This method is based on a dynamic threshold concept and also uses techniques for binary image processing. A gastric contour extracted by this method may be precise enough for extracting morphological features from it for diagnosing the deformity of the stomach.
2. A dedicated software system was developed for implementation on a large computer system. In this system, the flexibility of processing and of modification was emphasized. This is because of the need of such systems in implementing new knowledge easily for finding efficient methods and for improving them. Since this need occurs in most studies on pattern recognition of natural images, a generalized version of the software system was designed and implemented on a small computer system.

Feature extraction and pattern classification. One descriptive approach and two quantitative approaches were examined for the interpretation:

1. In the descriptive approach, the deviation curve was introduced. The appearance of a gastric contour was described on the basis of a branching logic using this deviation curve and the gastric axis.
2. In one of the quantitative approaches, gastric contours were analyzed by statistical methods of pattern recognition using five geometrical measurements and twenty-eight Fourier descriptors as candidates of discriminating features. The separability of the samples was proved by methods of the eye-inspection of scattergrams, the discriminant analysis based on the minimum Mahalanobis' distance classification rule, and the clustering of the samples and of the features. By the feature selection in the discriminant analysis, the best feature set was obtained with 7.5 percent error in the N/A classification; and with 25.0 percent error in the N/U/C classification. These figures were estimated by the leaving-one-out method; and agreed well with the result of inspection of the gastric contours by six physicians. In the N/U/C classification, confusion occurred between classes of ulcer and cancer. (This is not crucial in mass screening application.)
3. In the other quantitative approach, the same statistical methods were applied to apex images. The candidate features of this case are moduli of lower frequency/sequency components obtained by two-dimensional orthogonal transforms of the images. The trend was similar to that of the gastric contour analysis. The minimum error rate was 6.4 percent in the N/A classification; and 21.3 percent in the N/U/C classification. These figures also agreed well with the result of inspection of the images by a physician.

Thus, the experiments resulted a promising perspective of computerized

interpretation of SPAP images. Since the usefulness was proved of computational manipulation of the gastric contour, (a) the contour extraction study was assured to be meaningful and (b) the possibility was recognized in introducing human assistance in the contour extraction task as a shortcut. In addition, verifying instances were provided for statistical methods of pattern recognition: the similarity between the feature selection and the clustering of features; and the difference between the feature selection and the principal component analysis in discrimination problems.

The study is fundamental but not closed. Thus, the subjects must be noted which require further investigation:

1. The contour extraction methods must be improved for precise extraction, which is important for examining rigidity, linearization, and irregularity of a contour partly. The feature extraction and the pattern classification must be studied on contours extracted by digital image processing. The digital image processing must allow for a variety of gray-level distribution of images and various shapes of stomachs.
2. The statistical analyses must be conducted on a large number of verified samples which represent the natural distribution. Therefore, such samples must be collected. Methods must be also developed for detecting abnormalities located in portions other than the apex.
3. Methods must be developed for inspecting all of the six or more radiograms taken of one subject in order to make diagnosis. In fact, in the experiments in the previous chapter, inaccuracy of the physicians arose, because only the SPAP images were shown to them.

REFERENCES

[1] *Digital Image Processing of Biomedical Images*, K. Preston, Jr. and M. Onoe, eds. New York: Plenum Press, 1976.

[2] *Digital Picture Analysis*, A. Rosenfeld, ed. Berlin: Springer-Verlag, 1976.

[3] Health and Welfare Statistics Association of Japan, Special issue on trend of health care in Japan, *Indeces of Health and Welfare*, vol. 26, no. 9, p. 67, 1979.

[4] Indirect Radiography Standardization Committee of the Japanese Society of Gastric Mass Survey, *Gastric Cancer and Mass Survey*, vol. 28, pp. 114-116, 1974.

[5] T. Soma, S. Fukushima, K. Kawai, and K. Ida, "Computer diagnosis of roentgenography of the stomach—preliminary study of the upright barium-filled picture—," *Japanese Journal of Medical Electronics and Biological Engineering*, vol. 11, no. 2, pp. 100-107, 1973.

[6] T. Soma and S. Fukushima, "Feature extraction and quantitative diagnosis of gastric roentgenograms," in *Digital Image Processing of Biomedical Images*, K. Preston, Jr. and M. Onoe, eds. New York: Plenum Press, 1976, pp. 323-334.

[7] K. Miyawaki, "Pattern recognition," *Japanese Journal of Medical Electronics and Biological Engineering*, vol. 8, no. 5, pp. 379-383, 1970.

[8] T. Akatsuka, T. Isobe, and O. Takatani, "Feature extraction of stomach radiograph," in *Proc. 2nd Int'l Jnt. Conf. on Pattern Recognition*, Copenhagen, 1974, pp. 324-328.

[9] H. Mori, H. Niki, E. Nakagomi, and M. Ozeki, "An experiment in computer screening of gastric barium-filled X-ray films in erect anteroposterior position," *Japanese Journal of Medical Electronics and Biological Engineering*, vol. 15, no. 7, pp. 457-464, 1977.

[10] S. Fukushima, S. Utsunomiya, and T. Soma, "Extraction of contour of stomach from radiograms," *ibid.*, vol. 15, no. 6, pp. 383-388, 1977.

[11] N. Sugimoto and M. Uehara, "Situation of studies on pictorial pattern recognition," *FUJITSU*, vol. 24, no. 7, pp. 191-208, 1973.

[12] S. Nakamura, E. Nishimura, and C. Yoshimoto, "Measurement and processing of gastric images using a gastric-radiographic image data base," in *Proc. 9th Japanese Jnt. Conf. on Image Technology*, Tokyo, 1978, pp. 91-94.

[13] M. Hatori, T. Iwazuka, Y. Taki, and S. Nishida, "Classification of contrast radiographs, especially gastric angle shape, by computer," *Japanese Journal of Radiographic Image Research*, vol. 7, no. 1, pp. 33-40, 1977.

[14] M. Hatori, S. Hasegawa, and Y. Taki, "Automatic gastric angle shape classification on contrast radiographs," in *Proc. Int'l Symp. on Medical Info. Syst.*, Osaka, 1978, pp. 411-414.

[15] C. T. Zahn and R. Z. Roskies, "Fourier descriptors for plane closed curves," *IEEE Trans. Comput.*, vol. C-21, no. 3, pp. 269-281, 1972.

[16] S. Fukushima, H. Tsujinaka, K. Sakurai, and T. Soma, "Experiments on abnormality discrimination of gastric radiograms," *Tech. Rep. of IECE Japan*, PRL80-76, Jan. 1981.

[17] L. W. Paul and J. H. Juhl, *The Essentials of Roentgen Interpretation*, New York: Hoeber Medical Division, 1965.

[18] H. Ichikawa, T. Yamada, and I. Doi, *Practicing Gastric Radiography—for Detecting Early Cancer—* Tokyo, Bunko-do, 1965.
[19] A. Rosenfeld, *Picture Processing by Computer*, New York: Academic Press, 1969.
[20] C. K. Chow and T. Kaneko, "Boundary detection of radiographic images by a threshold method," *IBM Report*, RC3203, 1970.
[21] S. Fukushima, S. Utsunomiya, and T. Soma, "On thresholding of noisy images," in *Proc. of Hokkaido Regional Joint Convention of Japanese Societies for Electrical and Electronic Engineering*, Muroran, Hokkaido, Japan, 1976, p. 86.
[22] J. Hilditch, "Linear skeletons from square cupboards," in *Machine Intelligence 4*, B. Meltzer and D. Michie, eds. pp. 403-420. Edinburgh Univ. Press, 1969.
[23] J. Hilditch, "An application of graph theory in pattern recognition," in *Machine Intelligence 3*, D. Michie, ed. pp. 325-347, Edinburgh Univ. Press, 1968.
[24] S. Fukushima and T. Soma, "A system for pattern analysis of radiograms of stomach," *Japanese Journal of Medical Electronics and Biological Engineering*, vol. 16, no. 3, pp. 198-204, 1978.; S. Fukushima and T. Soma, "Processing of radiograms of stomach," in *Proc. 4th Int'l Jnt. Conf. on Pattern Recognition*, Kyoto, 1978, pp. 894-896.
[25] U. Montanari, "A method for obtaining skeletons using a quasi-Euclidean distance," *J. of ACM*, vol. 15, no. 4, pp. 600-624, 1968.
[26] G. Levi and U. Montanari, "A grey-weighted skeleton," *Info. and Control*, vol. 17, pp. 62-91, 1970.
[27] S. Fukushima, Y. Kimura, and T. Soma, "An interactive image processing system implemented on small scale hardware," *Memoirs of Faculty of Industrial Arts, Kyoto Technical University, Science and Technology*, vol. 30, pp. 66-80, 1981.
[28] R. S. Ledley, "High speed automatic analysis of biomedical pictures," *Science*, vol. 146, no. 3641, pp. 216-223, 1964.
[29] C. A. Harlow, S. J. Dwyer III, and G. Lodwick, "On radiographic image analysis," in *Digital Picture Analysis*, A. Rosenfeld, ed. Berlin: Springer Verlag, 1976, pp. 65-150.
[30] J. T. Tou and R. C. Gonzalez, *Pattern Recognition Principles*, Reading, Massachusetts: Addison-Wesley, 1974.
[31] K. Fukunaga, *Introduction to Statistical Pattern Recognition*, New York: Academic Press, 1972.
[32] D. H. Foley, "Considerations of sample and feature size," *IEEE Trans. Info. Theory*, vol. IT-18, pp. 618-626, 1972.
[33] J. Toriwaki, S. Kimura, S. Yokoi, and T. Fukumura, "Experimental evaluation of the performance of sequential feature selection procedure in statistical pattern classification," *Trans. IECE of Japan*, vol. J64-D, pp. 268-275, 1981.
[34] M. Kawaguchi, *Introduction to Multivariate Analysis II*, Tokyo: Morikita-Shuppan, 1978.
[35] *FACOM OS IV CLUSTER Manual*, Tokyo: Fujitsu, 1977.
[36] C. W. Richard, Jr. and H. Hemami, "Identification of three-dimensional objects using Fourier descriptors of the boundary curve," *IEEE Trans. Syst., Man, and Cybern.*, vol. SMC-4, pp. 371-378, 1974.
[37] H. Enomoto and K. Shibata, "Orthogonal transform coding system for

television signals," *IEEE Trans. Electromagn. Compat.*, vol. EMC-13, pp. 11-17, 1971.
[38] W. K. Pratt, W. H. Chen, and L. R. Welch, "Slant transform image coding," *IEEE Trans. Commun.*, vol. COM-22, no. 8, pp. 1075-1093, 1974.
[39] J. W. Cooley and J. W. Tukey, "An algorithm for the machine calculation of complex Fourier series," *Math. Comput.*, vol. 19, pp. 297-301, 1965.
[40] N. Ahmed and K. R. Rao, *Orthogonal Transforms for Digital Signal Processing*, Berlin: Springer-Verlag, 1975.
[41] J. P. Crettez, "The relationship between two fast slant transforms in ordered form," in *Proc. 4th Int'l Jnt. Conf. on Pattern Recognition*, Kyoto, 1978, pp. 445-448.
[42] F. J. Harris, "On the use of windows for harmonic analysis with the discrete Fourier transform," *Proc. IEEE*, vol. 66, pp. 51-83, 1978.

LIST OF PUBLICATIONS

[1] T. Soma, S. Fukushima, N. Murata, K. Ida, and K. Kawai, "Roentgenographic computer diagnosis of the stomach —frontal view of upright barium-filled picture—," *Tech. Rep. of IECE Japan*, MBE71-8, July, 1971.

[2] T. Soma, S. Fukushima, B. Kondo, and K. Ida, "Roentgenographic computer diagnosis of the stomach," *ibid.*, IT72-36, Feb. 1973.

[3] T. Soma, S. Fukushima, K. Kawai, and K. Ida, "Computer diagnosis of roentgenography of the stomach—Preliminary study of the upright barium-filled Picture—," *Japanese Journal of Medical Electronics and Biological Engineering*, vol. 11, no. 2, pp. 100-107, 1973.

[4] S. Fukushima, S. Utsunomiya, T. Soma, and K. Kawaguchi, "Extraction of shadow of stomach by dynamic threshold methods," *Tech. Rep. of IECE Japan*, IE75-77, Nov. 1975.

[5] T. Soma and S. Fukushima, "Feature extraction and quantitative diagnosis of gastric roentgenograms," in *Digital Processing of Biomedical Images*, K. Preston, Jr. and M. Onoe, eds. New York: Plenum Press, 1976, pp. 323-334.

[6] S. Fukushima, S. Utsunomiya, and T. Soma, "Extraction of contour of stomach from radiograms," *Japanese Journal of Medical Electronics and Biological Engineering*, vol. 15, no. 6, pp. 383-388, 1977.

[7] S. Fukushima, T. Oohashi, and T. Soma, "A system for pattern analysis of radiograms," *Tech. Rep. of IECE Japan*, IE77-62, Dec. 1977.

[8] S. Fukushima and T. Soma, "A system for pattern analysis of radiograms of stomach," *Japanese Journal of Medical Electronics and Biological Engineering*, vol. 16, no. 3, pp. 198-204, 1978.

[9] S. Fukushima and T. Soma, "Processing of radiograms of stomach," in *Proc. 4th Int'l Jnt. Conf. on Pattern Recognition*, Kyoto, 1978. pp. 894-896.

[10] Y. Kimura, S. Fukushima, and T. Soma, "An interactive image processing system implemented on a small scale hardware," *Tech. Rep. of IECE Japan*, PRL78-84, March 1979.

[11] H. Tsujinaka, S. Yamaguchi, E. Tsujimoto, S. Fukushima, and T. Soma, "Abnormality discrimination of radiograms of stomach—experiments using contours and two-dimensional orthogonal transforms—," *ibid.*, MBE79-21, July 1979.

[12] S. Fukushima, T. Soma, K. Hayashi, and Y. Akasaka, "Approaches to computerized diagnosis of stomach radiograms," in *Proc. 3rd World Conf. on Medical Informatics*, Tokyo, 1980, pp. 769-773.

[13] S. Fukushima, H. Tsujinaka, K. Sakurai, and T. Soma, "Experiments on abnormality discrimination of gastric radiograms, 2," *Tech. Rep. of IECE Japan*, PRL80-76, Jan. 1981.

[14] S. Fukushima, Y. Kimura, and T. Soma, "An interactive image processing system implemented on small scale hardware," *Memoirs of Faculty of Industrial Arts, Kyoto Technical Univ., Science and Technology*, vol. 30, pp. 66-80, 1981.

[15] S. Fukushima, T. Soma, Y. Akasaka, K. Hayashi, and K. Kawai, "Experiments on discrimination of abnormal stomachs by computer on the basis of standing position-anteroposterior radiograms," *Japanese Journal of Medical Electronics and Biological Engineering*, vol. 20, no. 2, pp. 86-93, 1982.

APPENDIX A SUPPLEMENTARY REMARKS TO THE ALGORITHMS

A.1 A Property of a Threshold and Its Application

A threshold value determined from a gray-level histogram of an image depends on the ratio between the area of the object and that of the background. Assume that an image is composed of two subsets and a transition region between them. Assume its gray-level profile is as shown in Fig. A.1(c), where p_1, p_2, and p_t are proportion of the respective regions; μ_1 and μ_2 are gray levels of respective subsets. In the presence of additive noise with normal distribution $N(0,\sigma^2)$, the histogram is given by

$$f(x) = \int_{-\infty}^{\infty} f_s(y) f_N(x-y)\, dy$$

where

$$f_s(x) = p_1 f_1(x) + p_2 f_2(x) + p_t f_t(x)$$

$$p_1 + p_2 + p_t = 1$$

$$f_i(x) = \delta(x-\mu_i) \;, \quad i=1,2$$

$$f_t(x) = \begin{cases} 1/(\mu_2-\mu_1) \;, & \mu_1 < x < \mu_2 \\ 0 \;, & elsewhere \end{cases}$$

$$f_N(x) = (1/\sqrt{2\pi}\sigma)\, exp(-x^2/2\sigma^2)$$

By the mode method, the threshold is determined as a solution of

$$f'(x_\theta) = 0 \quad and \quad f''(x_\theta) > 0 \tag{A.1}$$

Define the signal-to-noise ratio of an image by

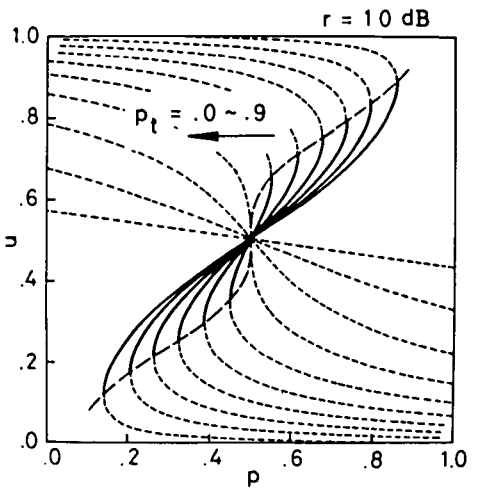
$$r = (\mu_2-\mu_1)/\sigma$$

and put

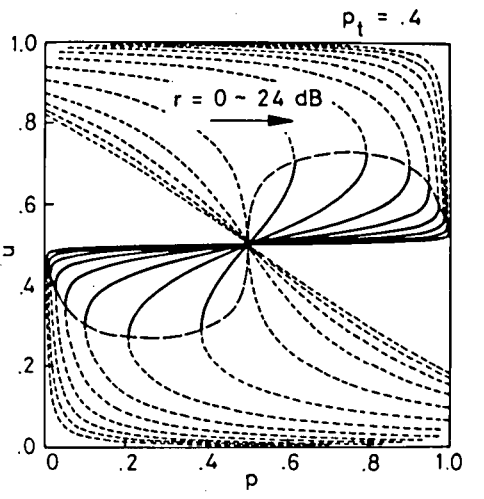
$$u = (x_\theta-\mu_1)/(\mu_2-\mu_1)$$

and

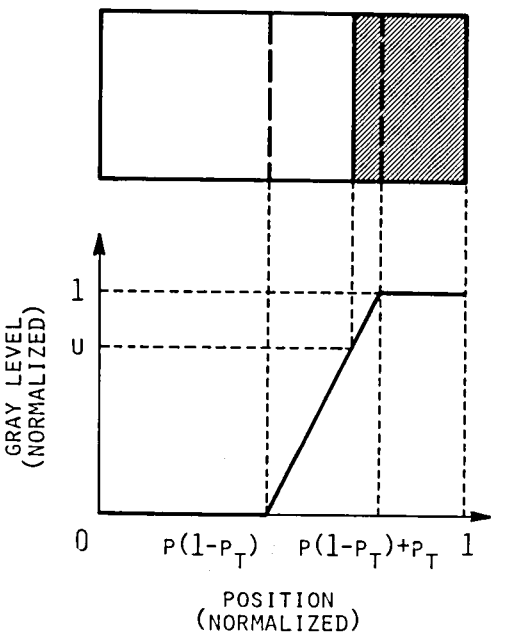
$$p = p_1/(1-p_t)$$



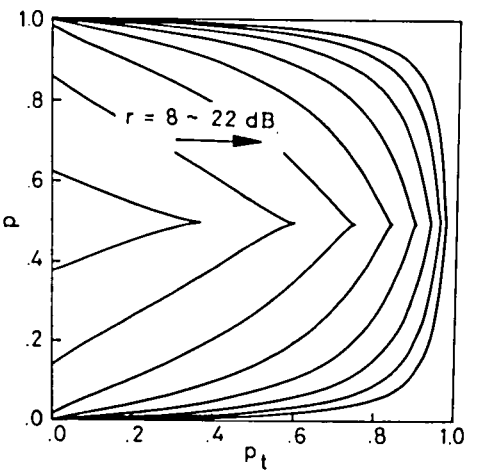
(a)



(b)



(c)

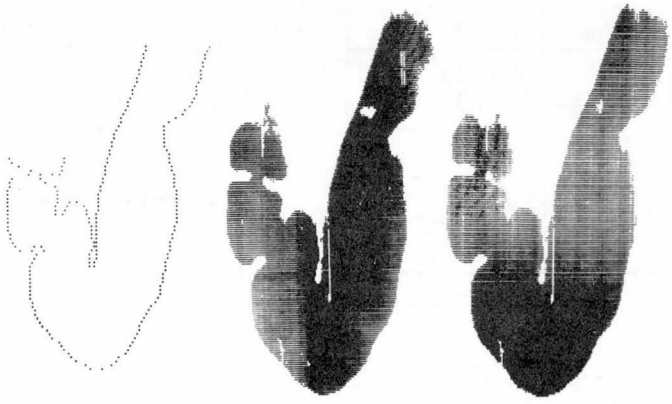


(d)

Fig. A.1 A property of a threshold determined by the mode method. (a) and (b) Trends of the threshold. (c) The model of the density distribution. (d) The condition for a valley to exist in the density histogram, *i.e.*, for the mode method to be applicable.

Then, the threshold depends on p, p_t, and r; and is given, for example, as shown in Fig. A.1(a) and in Fig. A.1(b). The broken lines correspond to the peaks of the histogram. For the mode method to be applicable, a valley must exist in the histogram. This means that the point (p_t, p) must fall in the inner part of the shape of U or V in Fig. A.1(d), which was obtained by eliminating x_θ from the expression (A.1). The threshold is biased towards the gray level of the smaller subset.

This property can be used for modification of the threshold. An example is shown in Fig. A.2. In this example, heavy shading caused erroneous extraction of the gastric contour in the upper part of the gastric body by the third method of the dynamic thresholding. For correcting the result, three quarters of pixels were decimated in the region which has been identified as the background; and the dynamic threshold method was applied again.



(a) (b) (c)

Fig. A.2 An example of the threshold modification. (a) The true contour. (b) The stomach extracted by the dynamic threshold method. (c) The stomach extracted after the threshold modification by pixel decimation.

A.2 Calculation of Fourier Descriptors

Fourier descriptors are calculated by the discrete Fourier transform of a sequence of contour points. Since the number of the contour points is generally not an integer power of a prime number, the fast Fourier transform (FFT) algorithm is not applicable. On the other hand, the spectrum leakage occurs if extra points are discarded or dummy points are appended in order to apply FFT. Let us consider the Fourier descriptors of a circle of radius R. This circle can be represented as

$$z(t) = R\ exp(j\omega_0 t), \quad \omega_0 = \frac{2\pi}{T_0}$$

where the period T_0 corresponds to the path length from the start point to the end point. The function $z(t)$ can be expanded into a Fourier series and the amplitude of its n-th harmonic is

$$c_n = \begin{cases} R, & n=1 \\ 0, & n\neq 1 \end{cases}$$

Assume that dummy points are appended to make the number of points an integer power of two for numerical calculation. By this modification, the fundamental period becomes $T_1=\alpha T_0$ $(1\leq\alpha<2)$. Assuming that analog figures are regarded as an identical pattern, each amplitude must be normalized by the amplitude of the first harmonic to make pattern matching. Define this normalized amplitude as

$$d_n = |c_n / c_1| .$$

Consider three cases.

Case 1:

$$z(t) = R\ exp(j\omega_0 t), \quad 0\leq t\leq T_1 .$$

Then

$$d_n = |sinc\{\pi(\alpha-n)\} / sinc\{\pi(\alpha-1)\}| .$$

Case 2:

$$z(t) = \begin{cases} R\ exp(j\omega_0 t) , & 0 \leq t \leq T_0 \\ R , & T_0 < t \leq T_1 \end{cases}$$

Then

$$d_n = \left| [sinc\{\pi(1-n/\alpha)\}+sinc(\pi n/\alpha)] \ / \ [sinc\{\pi(1-1/\alpha)\}+sinc(\pi/\alpha)] \right| \quad .$$

Case 3:

$$z(t) = \begin{cases} R\ exp(j\omega_0 t) , & 0 \leq t \leq T_0 \\ 0 , & T_0 < t \leq T_1 \end{cases}$$

Then

$$d_n = \left| sinc\{\pi(1-n/\alpha)\} \ / \ sinc\{\pi(1-1/\alpha)\} \right| \quad .$$

In *Case 1* and *Case 3*, when $\alpha \to 2$, d_2 appears very large. *Case 2* is the most moderate; but even in this case, d_n depends on α considerably. It should be noted that in *Case 3* the figure is not a circle but one with a dot at its center. This is remarked here because null dummy points are appended sometimes in practice.

As for features for discriminating patterns, it is sufficient to calculate only required terms. Since relatively small number of terms are required usually, the fast algorithm is in no need even for saving time. When M terms are required, time saved by the fast algorithm is

$$\frac{2N\ log_2 N}{MN} = \frac{2\ log_2 N}{M}$$

If N=512 and M=4, then this ratio is greater than one, which means no saving.

APPENDIX B EXPERIMENTS ON ERROR ESTIMATION AND FEATURE SELECTION

In this appendix, some results are shown to provide evidences of the bias in the error estimation by the replacement method and of the necessity of the feature selection. See Table B.1. Note also that the sequence of the feature selection has no correspondence with the order of the Fourier descriptors. In addition, inappropriate property may be recognized in the feature selection by taking principal components in an ordered sequence determined by the magnitude of eigenvalues. Thus, the principal component analysis has been denied of its role for selecting discriminating features. See Fig. B.1.

Table B.1 A comparison of the error rate estimation by the leaving-one-out method and the replacement method

Classification	N/A		N/U/C	
Method	L-O-O	Repl.	L-O-O	Repl.
Features				
31V	31.7 %	0.0 %	57.5 %	0.0 %
5PC of 31V	20.0	7.5	40.0	30.0
FS from 31V	7.5	7.5	25.0	25.0

Contour type = I, Sample size = 40

Notes:

1. The sample set contains both the fishhook shaped and the steer-horn shaped stomachs.
2. Legends:
   L-O-O: the leaving-one-out method.
   Repl.: the replacement method.
   31V: GM1, GM2, GM3, and the twenty-eight Fourier descriptors.
   5PC of 31V: the first five principal components of 31V.
   FS from 31V: features selected from 31V. (See Table 4.1.)

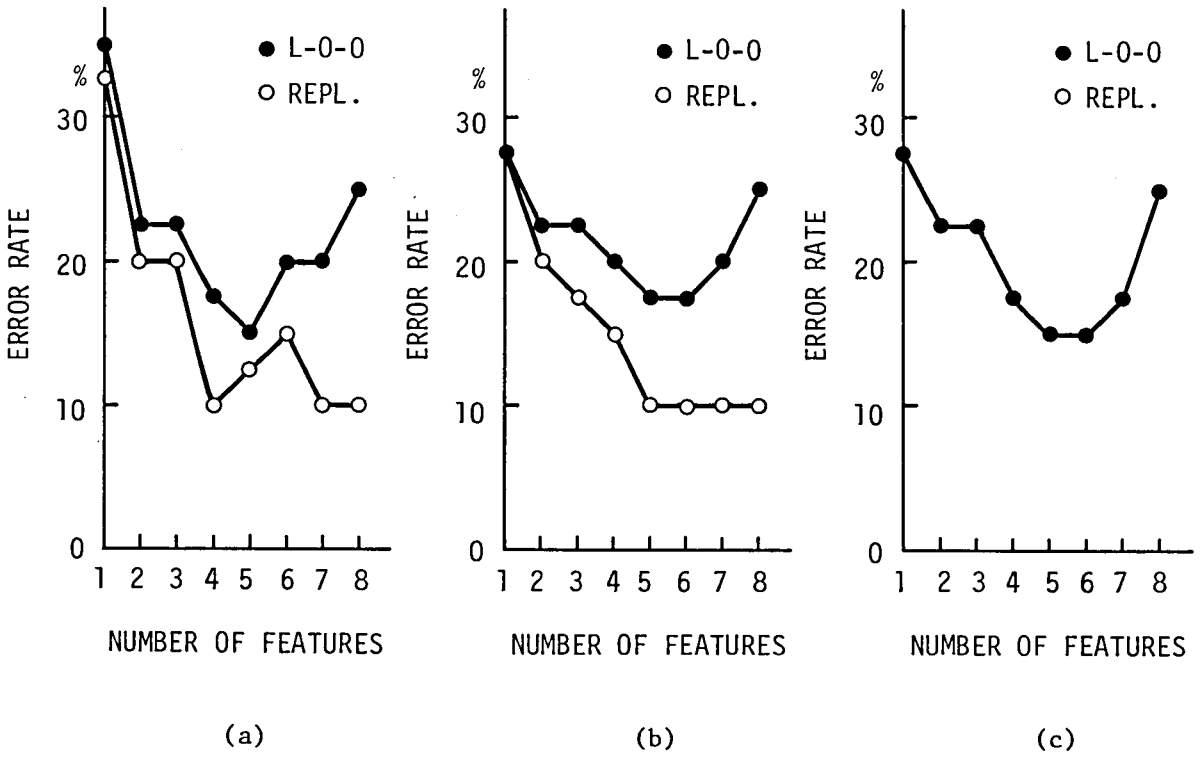


Fig. B.1 A comparison of error rate estimation by the leaving-one-out method (L-O-O) and the replacement method (REPL.) in N/A classification of the type-I contours using the first eight principal components of the Fourier descriptors.
(a) Feature set augmentation in order of the principal components. (b) Augmentation in order of the forward sequential search based on the replacement method.
(c) Augmentation in order of the forward sequential search based on the leaving-one-out method.

APPENDIX C DISCRIMINANT ANALYSIS OF PRONE APEX IMAGES

Prone position-posteroanterior images were analyzed by the discriminant analysis using orthogonal-transform features of the apex. Most of the samples were the same as used in the experiment in Chapter 4. (Because of discarding images of the bad quality, the two sample sets are not identical.) Table C.1 compares the error rate attained by sets of features selected for respective feature classes. Table C.2 shows confusion matrices of the best N/A classification and of the best N/U/C classification. The classification was worse than that of the SPAP images. This may be because the natural shape of a stomach is deformed by various stimuli in taking a prone position image.

Table C.1 The error rate in classification of the prone apex images using the most discriminating feature set

Classification	N/A	N/U/C
Candidate features		
35FT	33.3 %	52.9 %
8PC of 35FT	41.2	51.0
35H.FT	**27.5	45.1
8PC of 35H.FT	29.4	58.8
30ST	35.3	***43.1
8PC of 30ST	39.2	60.8
M.D.	7.8 - 29.4 %	21.6 - 41.2 %

Sample size = 51

Notes:

1. The first column denotes sets of candidate features for the feature selection: the candidate features are the Fourier transform features (35FT); the first eight principal components (8PC); the Hanning window-Fourier transform features (35H.FT); the slant transform features (30ST). The result of classification by one physician is shown in the row denoted by "M.D." Two figures in both sides of a hyphen represent a range of his error rate.
2. The best cases are indicated by asterisks: the feature set is composed of $|\alpha_{0,0}|$ (**); $|\alpha_{0,5}|$ (***).

Table C.2 Confusion matrices of classification of the prone apex images by the discriminant analysis

		Assigned classes		
		N	A	Error
Verified	N	16	4	20.0 %
diagnoses	A	10	21	32.3
Overall error classification				27.5

Sample size = 51

Feature set = ($|\alpha_{0,0}|$)

Hanning window-Fourier transform.

		Assigned classes			
		N	U	C	Error
Verified	N	14	6	0	30.0 %
diagnoses	U	8	13	1	40.9
	C	4	3	2	77.8
Overall error classification					43.1

Sample size = 51

Feature set = ($|\alpha_{0,5}|$); slant transform.

Note: Identical matrices were obtained by both the leaving-one-out method and the replacement method.