Ultrasound in Med. & Biol., Vol. 25, No. 2, pp. 189–200, 1999
Copyright © 1999 World Federation for Ultrasound in Medicine & Biology
Printed in the USA. All rights reserved
0301-5629/99/$–see front matter

PII S0301-5629(98)00150-1

● *Original Contribution*

COMPUTERIZED QUANTIFICATION OF STRUCTURES WITHIN OVARIAN CYSTS USING ULTRASOUND IMAGES

YAIR ZIMMER,* RON TEPPER,† and SOLANGE AKSELROD*

*Medical Physics Department, Tel Aviv University, Tel Aviv; and †Ultrasound Unit, Department of Obstetrics and Gynecology, Sapir Medical Center, Kfar Saba, Israel

(*Received* 19 *December* 1997; *in final form* 7 *September* 1998)

Abstract—Ovarian cysts are a common type of ovarian mass. The morphology of cysts, as it appears in ultrasound images, is currently used for classification of ovarian pathologies. However, this classification process is based on human interpretation of the sonographic image. In this article, a semiautomatic algorithm for the quantification of ovarian cysts is presented. This algorithm categorizes the structures within a cyst and extracts their quantitative geometric properties (width, characteristic diameter). To assess the validity of the technique, its performance was compared to human classification and manual measurements made by an expert. The results show a good match between automatic evaluations made by a computer and those of an experienced observer, indicating a potential for clinical use. © 1999 World Federation for Ultrasound in Medicine & Biology.

Key Words: **Ovarian cyst, Papillation, Septation, Ultrasound, Image processing, Mathematical morphology, Convex hull, Convex deficiency.**

INTRODUCTION

Ovarian cysts are the most common type of ovarian mass observed in the clinical setting. Being filled with fluid within the ovary, they usually appear in B-scan ultrasound images as relatively dark (hypoechogenic) regions surrounded by bright tissue. An ovarian cyst can be characterized by its size, the nature and appearance of the fluid, the thickness and regularity of its boundaries, and the existence and properties of various structures within it. The link between the morphology of the cyst and its potential malignancy has led to the use of these properties, as well as others, in the differential diagnosis of ovarian cysts. A quantitative analysis of these morphological properties is performed, to assess the malignancy of the ovarian pathology, using scoring systems (DePriest et al. 1993; Finkler et al. 1988; Kurjak and Predanic 1992; Lerner et al. 1994; Sassone et al. 1991). The idea is to score several properties of the ovarian mass, according to a predetermined table, and to use the resulting value (the sum of the individual scores) for classification. However, the scoring process is based on human visual evaluations by an experienced observer, which do not exploit the advantages of computerized analysis.

In this article, a semiautomatic algorithm for the quantification of ovarian cysts is presented. The method refers to the structures within the cyst (papillations, septations), categorizing them and extracting quantitative data about their properties, and does not deal with other morphological parameters (*e.g.*, the turbidity of the fluid, the regularity of the boundaries) that are addressed by scoring systems. Although it currently relies on prior manual identification of the cysts, the algorithm can be applied on automatically segmented images, thus fitting into a streamlined automatic procedure.

To assess the validity of the technique, the performance of the semiautomatic algorithm was compared to evaluations made by an expert gynecologist. A statistical analysis of the results determines the performance of the algorithm. To the best of our knowledge, there were no previous attempts to automatically extract morphological features from ultrasound images of ovarian cysts. The few studies that applied automatic analysis to the ovaries in ultrasound images dealt with ovarian follicles.

THE MORPHOLOGY OF CYSTS IN ULTRASOUND IMAGES

An ovarian cyst appears in an ultrasound image as a relatively dark region, whereas the surrounding solid

Address correspondence to: Solange Akselrod, Ph.D., Medical Physics Department, School of Physics and Astronomy, Tel Aviv University, Tel Aviv 69978, Israel. E-mail: solangea@novpost.tau.ac.il

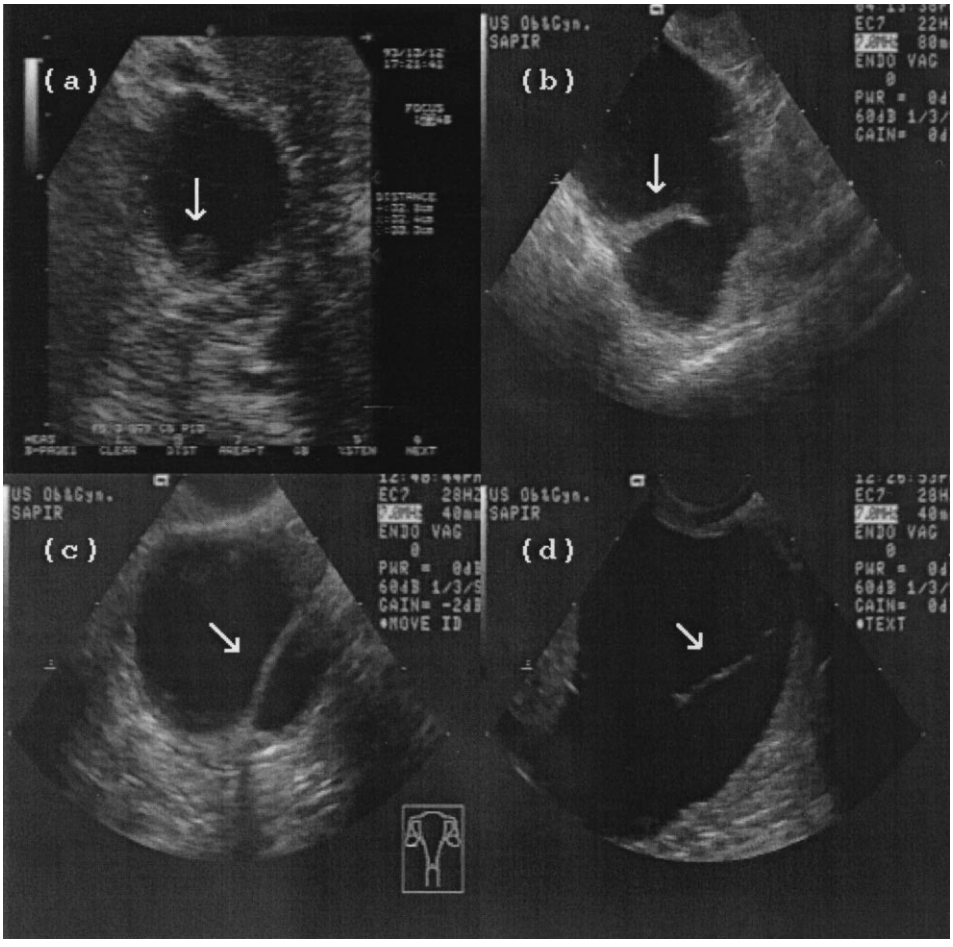

Fig. 1. Examples of various structures within a cyst. (a) A papillation, (b) An incomplete septation, (c) A complete septation, (d) An "internal" sepatation. The structures are indicated by arrows.

tissue appears brighter (see Fig. 1). The cyst may be either unilocular (*i.e.*, composed of a single lumen) or multilocular (*i.e.*, separated into several parts). In the latter case, each part can be treated as a separate cyst.

A cyst can generally contain two major types of morphological structures: papillations and septations. A papillation is a small region of ovarian tissue penetrating from the boundary of the cyst into its lumen. A septation is a narrow strip of ovarian tissue that penetrates the lumen and divides the volume into several segments. An important difference between papillations and septations is that septations are generally much more elongated than papillations. Examples of various structures found within an ovarian cyst are shown in Fig. 1.

A septation can be classified either as complete or incomplete. When a septation is complete, several separated cystic lumens are generated (hence forming a multilocular cyst). For an incomplete septation, the cystic region remains unilocular. A complete septation also may look incomplete in the image, due to poor contrast. Furthermore, poor imaging parameters or improper viewing angle may diminish the edges of a septation, causing it to appear as a strip located inside the cyst. A similar effect can be observed for papillary structures. Finally, when multiseptated (multilocular) cysts are analyzed, a value describing the maximal thickness possible for a (complete) septation should be predetermined. Cystic lumens separated by a wider strip of tissue can be considered uncorrelated. The existence, number, and dimensions of these structures are important indicators for the potential malignancy of the cyst. It should be emphasized that the partition into complete and incomplete structures is not a common definition in medical sonographic diagnosis; we adopt it only clearly to describe the algorithm.

THE PROPOSED ALGORITHM—AN OVERVIEW

The proposed method identifies the various structures within a cyst, categorizes them, and quantitatively

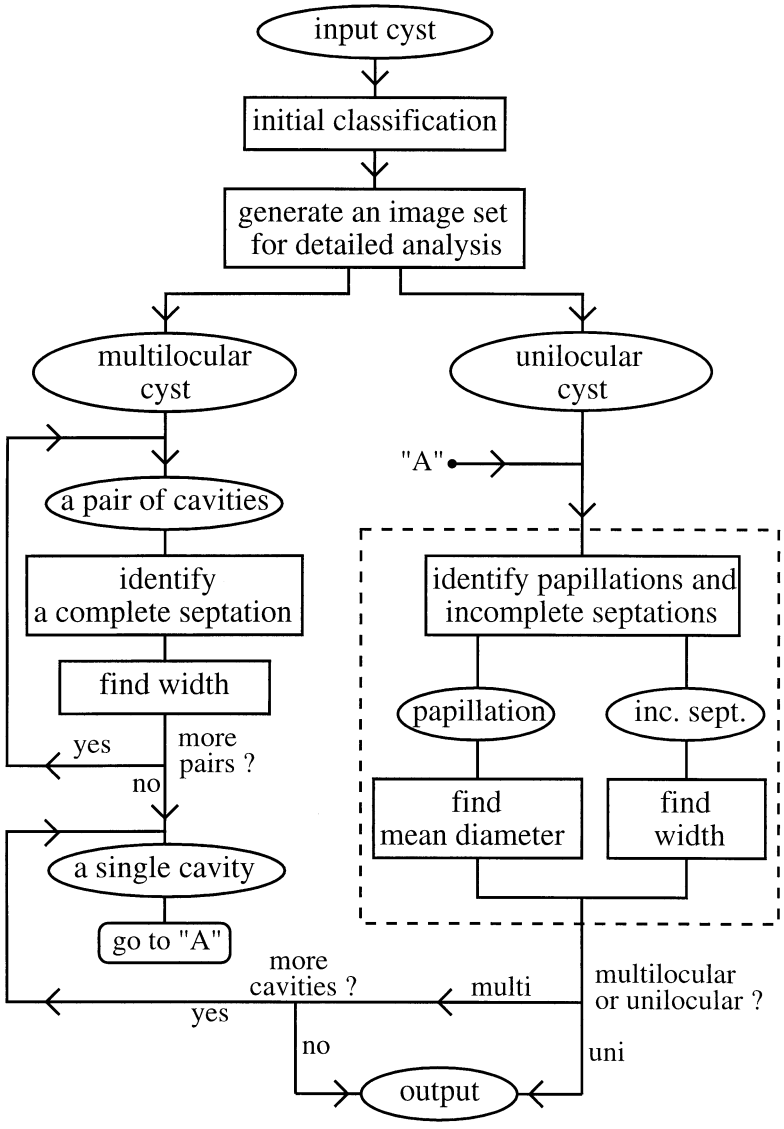

Fig. 2. The main steps of the algorithm. In this flow chart, an ellipse represents data and a rectangle represents an operation.

characterizes their geometry. From this analysis, the number, location, and dimensions of the various structures within each cyst are extracted. A flow chart describing the main steps of the algorithm is provided in Fig. 2. In Fig. 2, data (usually an image) are represented by an ellipse and an operation (a task) is represented by a rectangle.

As illustrated in Fig. 2, the method first separates multilocular cysts from unilocular ones. In case of multiseptated cysts, the algorithm first analyzes pairs of cystic cavities to evaluate the widths of the separating complete septations. Only then, is each cystic region addressed separately and the structures within it (papillations, incomplete septations) identified and analyzed. If the original cyst is unilocular, only the latter part of the analysis is performed, as shown in Fig. 2.

The analysis of papillations and incomplete septations within a single cystic cavity (whether an original unilocular cyst or part of a multiseptated one) is basically divided into classification (which includes identifying both structure types and ruling out false structures) and quantitative computations. Classification (to papillations, incomplete septations, etc.) is performed using several geometric properties of the ovarian structures and the minimum error rate Bayes classifier (Duda and Hart 1973). Then, the characteristic dimensions of each structure (the mean diameter of each papillation, the width of each septation) are assessed using mathematical mor-

phology (see Appendix 2). Both parts (identification, measurements) are described in Fig. 2.

PRELIMINARY STEPS

Preliminary classification

The initial step in analyzing the cyst is sorting it into one of two categories: unilocular or multilocular. We use images in which the cysts ("objects") are represented by a specific gray level whereas the rest of the image ("background") is represented by another gray level. The required initial classification then can be performed easily by counting the number of unconnected cystic regions in the image.

The following detailed analysis is performed on images containing either a single cyst (for most structures) or two cystic parts (for analyzing a complete septation). We therefore construct this set of images prior to structure analysis (see second rectangle from top in Fig. 2). As part of this preliminary step, the obtained images are processed further so that each cystic structure is indicated separately. The mathematical technique used is described in detail in the following sub-section.

Isolating the various cystic structures

Structure separation begins by computing the convex hull of the cyst, which is the smallest polygon circumscribing the cyst (see Fig. 3). Appendix 1 provides a mathematical background about the convex hull and presents a simple method for computing it. Subtracting the original object from its convex hull provides the convex deficiency of the cyst. The convex deficiency of the cyst is composed of several separated regions, each associated with a single cystic structure (*e.g.*, the gray regions in Fig. 3a). This description is also valid for images containing two cystic regions (Fig. 3b).

The regions composing the convex deficiency may often overestimate the corresponding cystic structures (*e.g.*, the outer parts of the septations in Figs. 3a and 3b). On other occasions, two small cystic structures may connect. Consequently, we use the convex deficiency of the cyst as an upper limit for the area representing cystic structures (*i.e.*, structures within a cyst or between two cystic regions).

Then, a new image is formed by applying morphological closing with a binary disk (a roundly shaped structuring element) on the original object (the cyst itself). A detailed discussion on mathematical morphology in general and morphological closing with a disk in particular is provided in Appendix 2. The obtained image exhibits partial filling of the concavities in the cyst (which correspond to cystic structures), leaving a crater in the outer part of the concavity (Fig. 4b). Because closing with a disk of radius r fills only cavities wider than r, a smaller disk results in less filling and a deeper crater. The radius of the applied disk for each cystic structure was determined as the minimal radius (from the series r = 20, 30, and 40 pixels), which fills more than 50% of the corresponding part of the convex deficiency. When all three disks filled a smaller portion of the region (probably leading to underestimation of the cystic structure), the convex deficiency and not the result of morphological closing was used for describing that cystic structure.

As can be seen from Fig. 4b, the craters formed by morphological closing cause a small underestimation of the corresponded structures. To eliminate this artifact, we take each region that was added to the cyst by morphological closing (if a single region in the original convex deficiency was broken into several separate subregions, we address each one of them separately) and replace it by its convex hull (computed as explained in Appendix 1). The results obtained are displayed in Fig. 4c.

The images used in the rest of the algorithm are those obtained after the steps noted previously (e.g., Fig. 4c). In these images, the area added to the cyst can be viewed as a corrected convex deficiency. We henceforth refer to this area as the convex deficiency (rather than using the term for the region associated with the actual convex hull).

ANALYZING COMPLETE SEPTATIONS

The analysis of a complete septation is performed on an image containing two "objects" (cysts) surrounded by "background," after computing (as described previously) the corrected convex hull and convex deficiency (Fig. 4c is an example). We identify and isolate the only part of the convex deficiency that is connected to both cysts, because it contains the analyzed septation. We furthermore identify its boundary segment, which is connected to one of the objects (this segment is henceforth called "edge 1"), and the segment connected to the other object (henceforth called "edge 2").

The septation between the two cysts can be visually described by a straight line passing from a point on edge 1 to a point on edge 2. Because the distance between the two opposite edges varies significantly, our problem is to find a pair of edge points, yielding a "correct" septation (*i.e.*, whose width agrees with manual evaluation). It should be noted that we seek the length of the line (width of septation) rather than its exact location.

To save computation time, we first select the shorter segment among edge 1 and edge 2. Then, for every point on the selected edge, we find the shortest line connecting this point and a point on the other edge. The set of

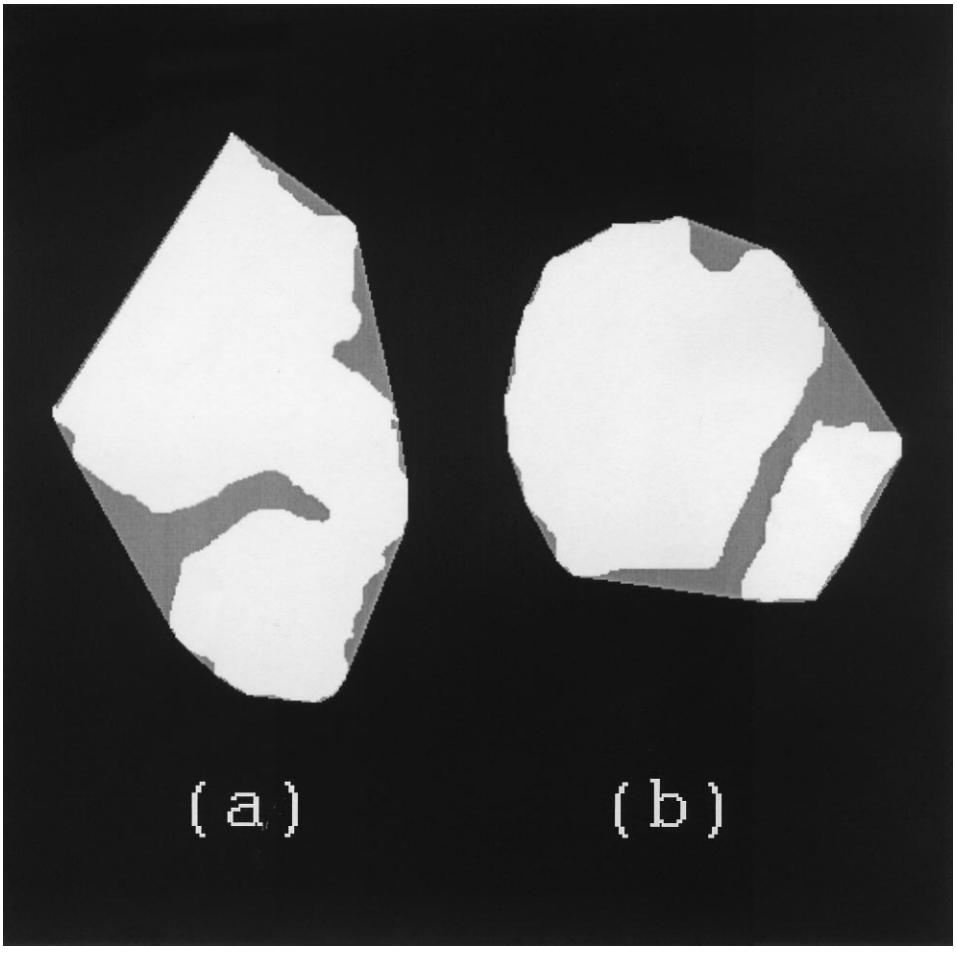

Fig. 3. Examples of the computed convex hull of a cyst. (a) Corresponds to the image in Fig. 1b, and (b) corresponds to the image in Fig. 1c. The images are magnified for display. In this figure, the original cyst is white and the convex deficiency is gray.

minimal distances obtained is preserved as candidates for the width of the septation.

Contrary to intuition, the mean (or even the median) distance cannot be used as a reliable assessment of the septation width, because of the large values contributed by lines located where edge 1 is far from edge 2. However, this contribution is spread over a relatively large range of values, whereas many lines are characterized by values close to the correct width (otherwise this width is meaningless). Consequently, we generated a histogram of the computed distances, rounding each value to the nearest 0.5 pixel (using a 0.5-pixel precision is sufficient for our purposes). Then, the most frequent value (*i.e.*, the maximal peak) in the histogram is expected to represent the correct septation width. To avoid various artifacts, this histogram was smoothed before peak detection using a conventional three-cell Gaussian filter. In other words, each cell in the histogram was weighted with its two nearest neighbors using the formula:

$$h(k) = 0.25 \cdot h_0(k-1) + 0.5 \cdot h_0(k) + 0.25 \cdot h_0(k+1) \quad (1)$$

where h_0 is the k^{th} cell in the original histogram, and h(k) is the k^{th} cell in the filtered histogram.

Finally, peak detection is applied on the filtered histogram and the global peak is taken as the evaluated width of the complete septation. In case the width obtained is larger than a predefined value (the structure is too thick to be considered a septation), we reject the result. When several complete septations are found in the same image, the procedure is repeated for every septation (*i.e.*, every pair of separated cysts).

ANALYZING PAPILLATIONS AND INCOMPLETE SEPTATIONS

The initial image

We precede the analysis described in the following by computing the corrected convex hull and convex

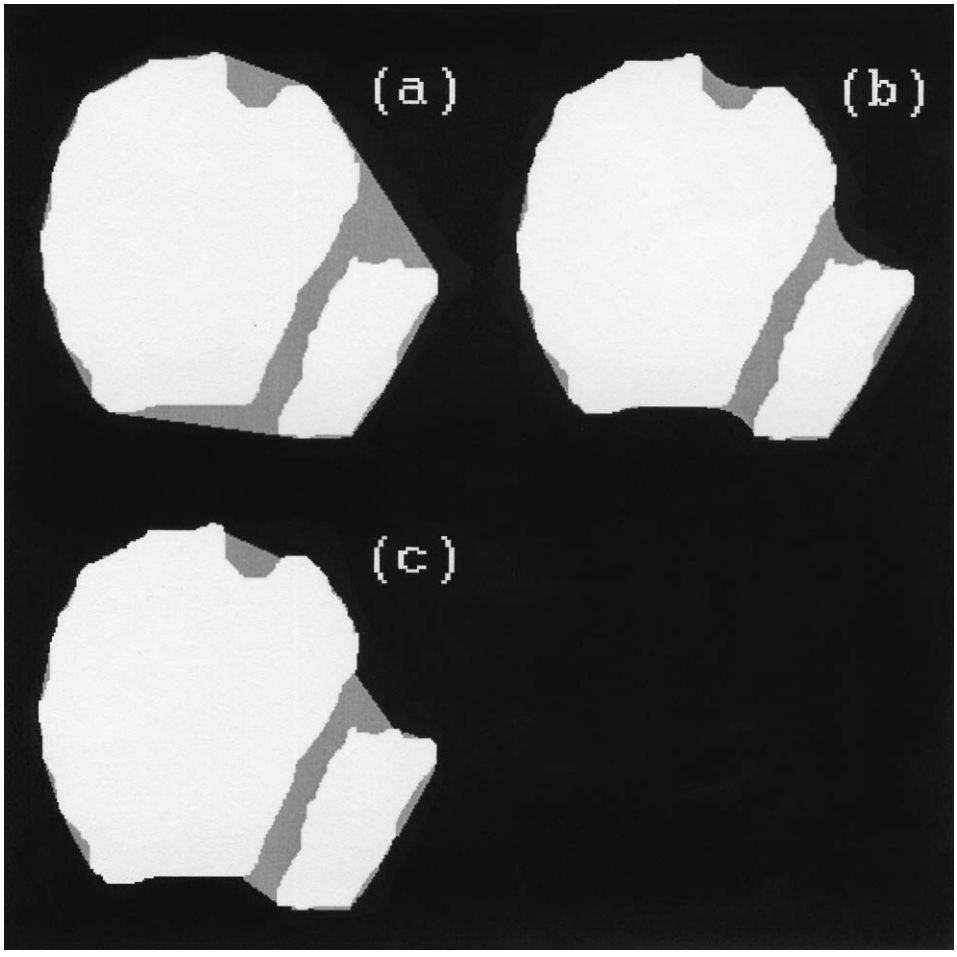

Fig. 4. Correcting the convex deficiency. (a) The initial convex deficiency (same as in Fig. 3b). (b) The result obtained after morphological closing with a disk. (c) The final result. The images are magnified for display.

deficiency of an image containing only the selected cyst ("object") surrounded by "background." Excluding rare cases, each part of the convex deficiency includes either a single structure or none. Generally, a specific part of the convex deficiency can represent a papillation, an incomplete septation, or a region that is irrelevant to the medical diagnosis (henceforth called "side") and appears as a result of local concavities in the boundary of the cyst.

When the original image (the B-scan) is faint or suffers from poor contrast, the region linking the structure to the surrounding tissue may disappear and the structure may appear as located entirely inside the cyst. The scanning angle also can induce a similar effect. Luckily, these regions are included (as separate parts) in the convex deficiency and, thus, can be geometrically analyzed. Again, such a region can contain a papillation (henceforth called "internal papillation"), a septation (henceforth called "internal septation"), or a meaningless structure generated due to small bright noises within the cyst.

Feature selection

A variety of geometric parameters can be used for shape analysis. We chose to use three relatively simple variables as our shape features. Two of them are the area and roundness of the shape. The third variable provides the relative part of its boundary touching the cyst.

The area of the inspected region, easily obtained by pixel counting, is used to eliminate small deficiency parts. Because the convex deficiency of a cyst usually contains a large number of small "sides" (regions formed near local boundary concavities), a preliminary step eliminating these false structures saves much of the computation time and simplifies the rest of the classification algorithm. The same principle applies for cases in which the false structure is located entirely within the cyst. The criterion used for rejecting false structures was an area smaller than 1% of the area of the cyst.

The compactness (also called "roundness") of a shape (Jain 1989) is a measure that assesses the extent to which a shape is elongated. The reason for using this

feature is that septations generally are elongated, whereas papillations are not.

The compactness of a shape is given by:

$$\text{comp} = \frac{(\text{perimeter})^2}{4\pi \cdot \text{area}} \quad (2)$$

The compactness is minimized for a disk, for which it equals 1.

The last variable used we call "pop." This parameter provides the relative part of the shape's perimeter (in percent) connected to the cyst and is formally defined as:

$$\text{pop} = 100 \times \frac{\text{length of portion touching cyst}}{\text{total length of perimeter}} \quad (3)$$

This geometric feature is selected because different types of structures are related to significantly different "pop" values. For an incomplete septation, most of the boundary (typically 75–95%) touches the object (the cyst) and only a relatively small fraction is connected to the background. For a papillation, a smaller part of the boundary (typically 60–80%) is connected to the cyst. Finally, only about half the perimeter (usually 50–60%) touches the cyst for a "side." When "internal" structures are involved, "pop" obviously equals 100, because the entire structure is surrounded by the cyst.

Although each variable provides a means for classification, a single parameter generally cannot be used alone, due to an overlay between the values obtained for different types of structures. Hence, a multivariate analysis is required. The details of the selected technique are discussed in the following.

Classifying the structures

After eliminating small false structures, two basic situations must be addressed: shapes representing structures touching the cyst and shapes related to "internal" structures. The latter case can be identified easily by checking whether the "pop" of the shape equals 100 (or very close to it, to avoid artifacts).

When the analyzed shape is connected to the cyst, we perform a bivariate analysis in a two-dimensional feature space containing the compactness of the shape and the relative part of its boundary touching the cyst. We chose to use the well-known minimum error rate Bayes classifier (Duda and Hart 1973), assuming that each class (papillation, incomplete septation, side) is bivariate normal. The mathematical principles of minimum error rate classification are provided in Appendix 3.

We describe the available data about a specific shape as a feature vector x in a two-dimensional feature space. Then, this shape belongs to class j, which yields a minimal value of the following expression:

$$F_k = (x - \mu_k)^t \Sigma_k^{-1} (x - \mu_k) + \ln(|\Sigma_k|) \quad (4)$$

where: $k = 1, 2, 3$; μ_k = the mean vector for class k; and Σ_k = the covariance matrix for class k.

To apply the technique, the mean vectors and the covariance matrices characterizing the three classes must be known. Hence, we first computed these variables using a training sample set (which included only cases for which the structure type was evident). Then, the classification technique was applied on a test group of unidentified shapes using the obtained values.

When the analyzed shape is fully contained in the cyst ("internal" structures), we perform a univariate analysis based on the compactness of the structure. For this analysis, no training is required. Instead, the means and variances of the Gaussians are taken from the data computed for papillations and incomplete septations. Sides are not considered, because it is assumed that all "internal" false structures previously were eliminated based on their small size. Then, the "internal" structures are sorted using one-dimensional minimum error rate classification.

Quantitative analysis

The method used to extract the quantitative data from the identified structures is based on morphological erosions (see Appendix 2) with various structuring elements. In this technique, the characteristic dimension of the examined object (namely, width or characteristic diameter) is assessed as twice the number of successive erosions (with the same structuring element) required to eliminate the entire object.

For a septation (incomplete or "internal"), the structure was eroded with a 3×3 square (Fig. 5a), and the obtained value was used as the evaluated septation width. Because successive erosions with a square evaluate the short axis (*i.e.*, actual thickness) of the object, using a square is suitable for elongated objects such as septations. Papillations, however, are not elongated and, hence, are better represented by their mean diameter.

For a papillation (including "internal" ones), the structure was eroded with four different three-pixel-long structuring elements, each centered at the examined pixel: a horizontal line (Fig. 5b), a vertical line (Fig. 5c), and two diagonal segments (Figs. 5d and 5e). For each structuring element, the number of erosions required to erase the object was found, hence providing evaluations of its thickness along four different axes. Then, the median of these four values (computed by excluding the two extreme results and finding the average of the two remaining ones) was computed and used as the evaluated mean diameter of the papillation.

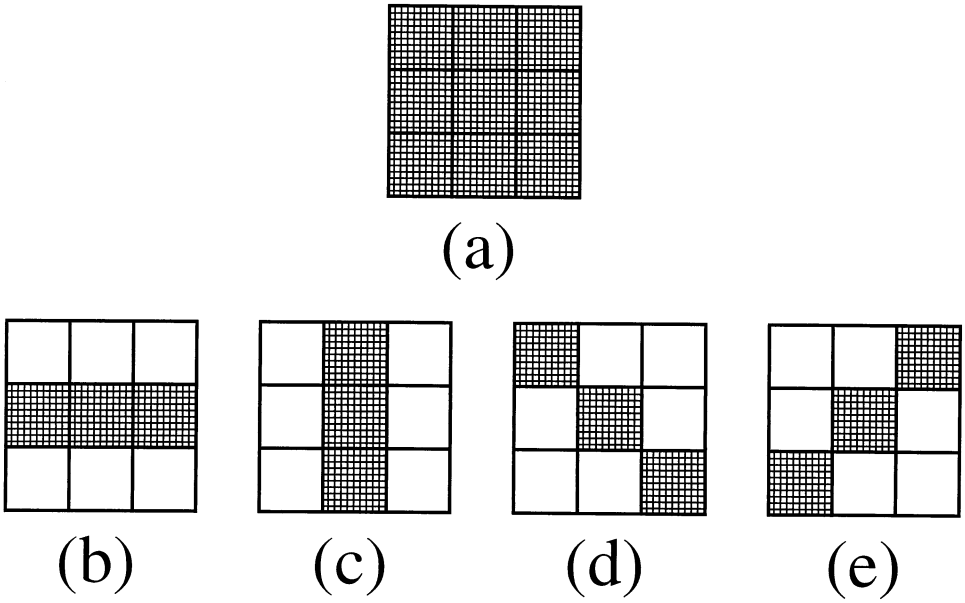

Fig. 5. The structuring elements used for morphological erosions during quantitative analysis. (a) A square, (b) A horizontal line, (c) A vertical line, (d, e) Diagonal segments. The width of a septation was evaluated using the square, whereas the characteristic diameter of a papillation was assessed using the four structuring elements shown in (b)–(e).

IMAGE ACQUISITION

In this study, we have used B-scan transvaginal sonographic images. These images were partly obtained using the Acuson 128XP/10 scanner (Acuson, Mountain View, CA, USA) with a 5- to 7-MHz transvaginal probe. The rest were generated by the Aloka SSD-680 (Aloka, Tokyo, Japan) with a 5-MHz transvaginal probe. All the images were recorded on VHS video cassettes and later digitized into the computer using a DT-2853 frame grabber (Data Translation, Marlboro, MA, USA), yielding images containing 512 pixels $\times$ 512 pixels $\times$ 256 gray levels. The algorithm was developed on a Silicon Graphics workstation, using version 4.2c of Matlab for multivariate analysis.

RESULTS

The initial data base

The algorithm was tested on 99 ultrasound images displaying ovarian cysts with a variety of morphological appearances. Each cyst contained from none to several morphological structures. The boundaries of each cyst were manually traced by an expert physician (R.T.), who also identified (and classified) each structure and (for most structures) manually evaluated its characteristic parameters (diameter of papillation, width of septation). From these boundaries, images isolating each cystic structure (as a separate part of the corrected convex deficiency) were generated.

A basic assumption in our performance evaluation process was that each structure could be classified definitively by the expert. In fact, the expert was required to classify a small number of structures that he considered borderline cases. Such ambiguous structures exhibit geometric characteristics of two different structure types (usually, a papillation and a septation). When the structure actually has borderline characteristics, its categorization is meaningless. When this ambiguity is caused by a poor viewing angle, the ultrasonographer simply can select a better angle and use the obtained image (rejecting the ambiguous one) for automatic analysis. Therefore, we believe that the performances of the algorithm for cysts with borderline structures should (for the most part) be a minor consideration in the evaluation process.

As a result, we have excluded from the obtained data set the structures that the expert defined as ambiguous (two papillations and four incomplete septations). It should be emphasized that only true borderline cases were excluded, leaving structures with various degrees of similarity to "classic" papillations and septations.

Structure classification

The features (pop, comp) of each structure (excluding complete septations) were computed in order to apply the bivariate analysis. From the set of obtained structures, a group of 47 structures (including 14 papillations, 12 incomplete septations, and 21 sides) was selected as a training set. All members of this set exhibited the char-

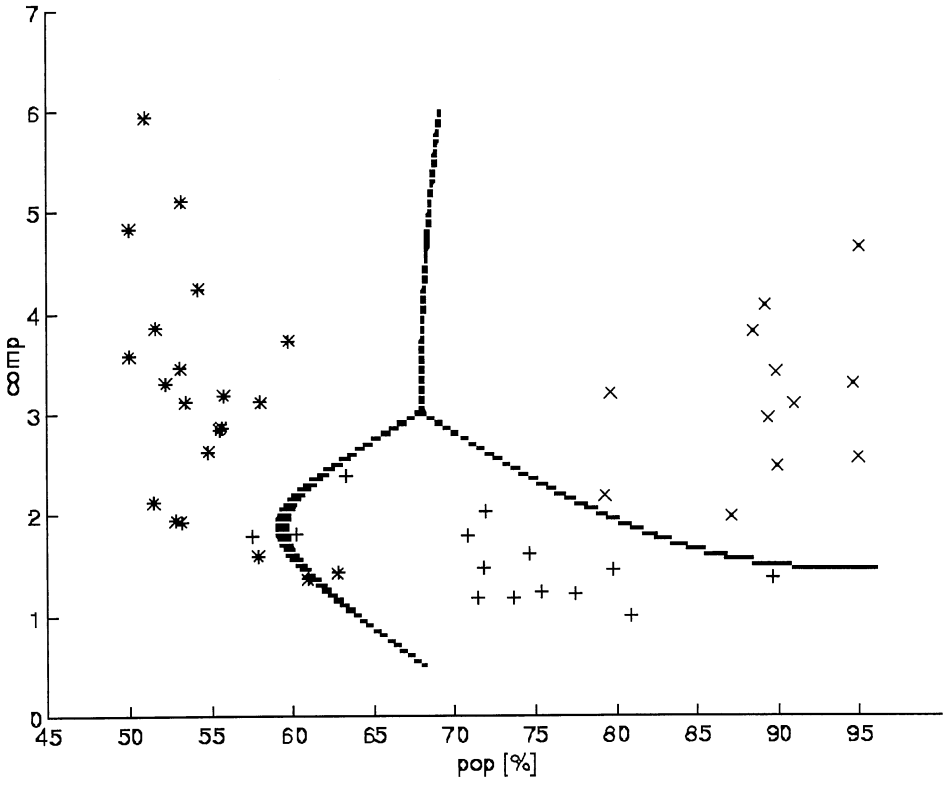

Fig. 6. Partition of feature space obtained from minimum error rate classification using the training group. Each structure in the training sample set is described by a symbol in the figure: a papillation is indicated by "+," a septation is indicated by "×," and a side by "*." When displayed in feature space, the training group forms three clusters. As a result of training, the feature space is divided into three regions (related to papillations, septations, and sides). The boundaries between them are illustrated by the curves.

acteristic (*i.e.*, "classic") appearance of the associated structures. As shown in Fig. 6, the training set is described by three clusters. The curves divide the feature space into three parts (related to papillations, septations, and sides) according to the results obtained from analyzing the training data set.

The bivariate classification technique was tested on 46 papillations, nine incomplete septations, and 41 sides. The algorithm classified correctly 33 papillations (71.7% of the tested ones), seven septations (77.8% of the cases), and 37 sides (90.2% of the examined sides). The univariate classification algorithm was tested on one "internal" papillation and seven "internal" septations. The papillation and five of the septations (71.4%) were identified correctly. Combining these results with those obtained using the bivariate algorithm shows correct classification of 72.3% of the papillations (34 of 47 cases), 75% of the septations (12 of 16 cases), and 90.2% of the sides (37 of 41 cases).

For structures of a specific type, correct classification can be regarded as "positive," whereas misclassification can be regarded as "negative." Then, the success rate is, in fact, the sensitivity of the classification technique for this structure type. Similarly, the specificity, positive predictive value (PPV), and negative predictive value (NPV) of the algorithm can be computed for each structure type. The results obtained are presented in Table 1.

Quantitative analysis

A quantitative analysis was performed on 61 cases (including 40 papillations, 18 incomplete septations, one "internal" papillation, and two "internal" septations) that were correctly classified and for which manual measurements were available. This data set contained structures from both groups (training, test) of the analysis. The values obtained for each structure were translated from pixels to millimeters, using the calibration data provided by the scanner. For the papillations (including the "internal" one), the mean difference between the manual and the automatic values of the mean diameter (obtained by averaging absolute differences) was 1.7 mm. For the septations (incomplete and "internal"), this difference between manual and automatic measures for the septa-

Table 1. Performance of structure classification.

Structure type	No. of cases	Sensitivity† (%)	Specificity (%)	PPV† (%)	NPV† (%)
Papillation*	47	72.3	89.5	85.0	79.5
Incomplete septation*	16	75.0	94.3	70.6	95.4
Side	41	90.2	84.1	78.7	93.0

* Includes "internal" structures.
† Sensitivity equals the success rate.
NPV = negative predictive value; PPV = positive predictive value.

tion width was 0.9 mm. It appears that the error obtained for papillations is much larger than the one obtained for septations. This is not surprising, because the morphological algorithm is more suited to measure the shortest dimension (*i.e.*, width) of an object rather than to measure its mean diameter.

The proposed technique for computing the width of a complete septation was tested on 48 cases, and the results were compared to manual evaluations. The mean difference obtained between manual and automatic evaluations was 0.9 mm. This result, as well as the values obtained for papillations and incomplete septations, are presented in Table 2. Combining the values obtained for the various structure types, we can conclude that the characteristic dimension of a cystic structure obtained by the proposed method matches manual evaluation to within 1–2 mm.

DISCUSSION

Structure classification

Table 1 shows that a success rate of approximately 75% was obtained for structure classification (72.3% for papillations and 75% for incomplete septations). We believe that this result indicates that our classification algorithm performs reasonably (thus having a promising potential for clinical use). Unfortunately, there are no similar automatic algorithms with which to compare.

Although the results obtained are satisfactory, a slightly different classification approach may be considered. This commonly used approach allows rejection of a data point that is unrecognizable (forming, in our case, four structure types, namely, "papillation," "septation," "side," and "rejected"). Borderline cases then can be included in the analysis. In addition, the size of the rejection region in feature space (*i.e.*, the region related to the "rejected" structures) can be determined according to the level of confidence desired for classification.

It should be noted that a large number of small structures (papillations and septations) was missed due to our restriction on the structure size. However, for such small structures (<1% of the cyst), the morphological differences between the three structure types (papillations, septations, and sides) become blurred, and this yields a large number of false-positives (sides misclassified as papillations or septations). Furthermore, such small structures are usually of little medical importance.

Table 2. Performance of quantitative analysis.

Structure type	No. of cases	Mean error† (mm)
Papillation*	41	1.7
Incomplete septation*	20	0.9
Complete septation	48	0.9

* Includes "internal" structures.
† Mean difference between manual and automatic evaluations.

Quantitative analysis

We have found that the characteristic dimension of a cystic structure computed by the algorithm agrees with manual evaluation to within 1–2 mm. This resolution is more than reasonable, considering that commonly used scoring systems (DePriest et al. 1993; Lerner et al. 1994; Sassone et al. 1991) use a single threshold of 3 mm to separate small cystic structures (both papillations and septations) from large ones. Hence, the proposed quantitative technique is satisfactory and presents a good match with manual evaluations.

Automatic segmentation of the image, performed prior to the proposed algorithm, can provide the boundary of the cyst and replace the manual tracing applied in this study. Segmentation of medical images has been discussed widely in the literature. An automatic thresholding technique, especially designed to segment ultrasound images, was described by Zimmer et al. (1996). Combining automatic segmentation with the hereby proposed method for structure analysis can lead to an automatic morphological analysis of ovarian cysts.

SUMMARY

In this article, a semiautomatic algorithm for the quantification of characterizing structures located within ovarian cysts was presented. The technique itself is fully automatic, although generating the initial binary input image requires manual steps. This algorithm categorizes automatically the various structures within a cyst and extracts their quantitative properties. The results obtained were compared to human classification and manual measurements made by an expert. When applied to routinely available ultrasound images, the qualitative analysis (*i.e.*, structure classification) succeeded in about 75% of the cases. The quantitative analysis provided a characteristic dimension that agrees with manual evalu-

ation to within 1–2 mm. It is concluded that the algorithm performs well and enables reliable analysis of the examined structures, thus having a potential for clinical use. Combining the proposed method with a segmentation technique can provide an automatic solution for the analysis of ovarian cysts.

Acknowledgements—This research is supported by a grant from the Ministry of Science, and by the Abramson Center for Medical Physics.

REFERENCES

Andrew AM. Another efficient algorithm for convex hulls in two dimensions. Info Proc Lett 1979;9:216–219.

Bentley JL, Faust MG, Preparata FP. Approximation algorithms for convex hulls. Comm ACM 1982;25:64–68.

Day AM. Planar convex hull algorithms in theory and practice. Comput Graphics Forum 1988;7:177–193.

DePriest PD, Van Nagell JR, Gallion HH, et al. Ovarian cancer screening in asymptomatic postmenopausal women. Gynecol Oncol 1993; 51:205–209.

Duda RO, Hart PE. Pattern classification and scene analysis. New York: John Wiley & Sons, 1973, ch. 2.

Finkler NJ, Benacerraf B, Lavin PT, Wojciechowski C, Knapp RC. Comparison of serum CA 125, clinical impression, and ultrasound in the preoperative evaluation of ovarian masses. Obstet Gynecol 1988;72:659–664.

Jain AK. Fundamentals of digital image processing. Englewood Cliffs, NJ: Prentice Hall, 1989, ch. 9.10.

Kurjak A, Predanic M. New scoring system for prediction of ovarian malignancy based on transvaginal color doppler sonography. J Ultrasound Med 1992;11:631–638.

Lerner JP, Timor-Tritsch IE, Federman A, Abramovich G. Transvaginal ultrasonographic characterization of ovarian masses with an improved, weighted scoring system. Am. J Obstet Gynecol 1994; 170:81–85.

Preparata FP, Shamos MI. Computational geometry: An introduction, 2nd ed. New-York: Springer-Verlag, 1985, ch. 3–4.

Sassone AM, Timor-Tritsch I, Artier A, Westhoff C, Warren WB. Transvaginal sonographic characterization of ovarian disease: evaluation of a new scoring system to predict ovarian malignancy. Obstet Gynecol 1991;78:70–76.

Zimmer Y, Tepper R, Akselrod S. A two-dimensional extension of minimum cross entropy thresholding for the segmentation of ultrasound images. Ultrasound Med Biol 1996;22:1183–1190.

APPENDIX 1

The convex hull

The convex hull of a shape is the smallest convex object that contains the original shape. The area added to the original shape (*i.e.*, the difference between the original object and its convex hull) is called the convex deficiency. The convex hull is extensively discussed in the literature, and numerous algorithms have been proposed in order to compute it. A variety of algorithms for finding the convex hull of a finite set of points in the plan is summarized by Preparata and Shamos (1985) and Day (1988). Although a shape in a binary image can be viewed as a finite set of points (located at the centers of the pixels), relatively few techniques addressed the convex hull in images.

The upper convex hull (U-hull) of a digital planar object is composed of straight segments bounding the object from above. Similarly, its lower convex hull (L-hull) is composed of straight segments bounding it from beneath. The complete convex hull of the shape can be obtained by combining the U-hull with the L-hull. Modification by Andrew (1979) of Graham's algorithm is based on this idea, and a similar notion is used by Bentley et al. (1982). However, these algorithms and others apply complex techniques to move from point to point along the upper (or the lower) convex hull.

We propose to simplify the algorithm for finding the U-hull and the L-hull (and, hence, the convex hull itself). Initially, the rectangle bounding the original object is found. Then, this region is scanned column after column, and the uppermost and lowermost object pixels in each column are separately collected. Next, we connect every pair of upper boundary points with a straight segment and construct a list of all the points on those segments. A similar list is generated from the lower boundary points of the object.

When the first list is sorted according to columns, the uppermost point in each column belongs to the U-hull (because the U-hull is the upper bound of the segments represented in that list). Similarly, the set of lowermost points in the sorted second list belongs to the L-hull. Finally, the convex hull itself is obtained by filling all the pixels between the U-hull and the L-hull in each column.

The proposed method is simpler than similar techniques and, therefore, easier to implement. Furthermore, it uses only part of the boundary points for computation; for an image containing $N \times N$ pixels, the maximal number of employed boundary points is 2N. An example of the result obtained used this method is presented in Fig. 3.

APPENDIX 2

Mathematical morphology

Mathematical morphology is a mathematical technique based on set theory, which provides a quantitative description of geometric objects and structures. Morphological operations are performed on the object using a moving shape called "structuring element." Implementations of mathematical morphology in image processing have been described extensively in the literature.

The two simplest morphological operations are erosion and dilation. In the binary case, the erosion of an object X by a structuring element B is defined as the group of translations of B that are completely included in X, and the dilation of an object X by a structuring element B is defined as the group of translations of B that are at least partly included in X. In practice, a mask (whose shape defines the structuring element) is moved

from pixel to pixel in the image; a pixel belongs to the eroded image if all the pixels covered by the mask (when the mask is centered on the examined pixel) belong to the object (*i.e.*, have a value "1"). Similarly, a pixel belongs to the dilated pixel if at least one pixel under the mask belongs to the object. Morphological opening of an object means erosion followed by dilation, and morphological closing of an object means dilation followed by erosion.

When the image is eroded using a roundly shaped structuring element (a disk), a pixel belongs to the eroded image only when all pixels within a radius r around it (r being the radius of the structuring element) belong to the object. Similarly, a pixel belongs to the dilated image only when at least one object pixel is found within a distance r from it. When erosion with a disk of radius r is applied on the results of dilation with the same structuring element, morphological closing of the image with this disk is obtained.

APPENDIX 3

Minimum error rate classification

Suppose that a region, which is represented in the feature space by a n-dimensional vector x, belongs to one of J classes. According to Bayes rule (Duda and Hart 1973):

$$P(j/x) = \frac{P(x/j) \cdot P(j)}{P(x)} \tag{5}$$

where P(j) = the *a priori* probability, and P(j/x) = the *a posteriori* probability.

Using the well-known minimum error rate classifier (Duda and Hart 1973), the vector x belongs to the class j for which P(j/x) is maximized. Obviously, P(x) remains the same for all classes. Furthermore, P(j) is determined as $\frac{1}{J}$ for all classes, due to lack of *a priori* information. Hence, the vector x is classified by finding the class for which the conditional probability P(j/x) is maximized.

We can easily identify P(j/x) as the probability density of class j, because it provides the probability of having a general vector x in class j. Assuming that P(x/j) is multivariate normal for each class j, we can write:

$$P(x/j) = \frac{1}{(2\pi)^{\frac{n}{2}}|\Sigma_j|^{\frac{1}{2}}} \cdot \exp\left[-\frac{1}{2}(x - \mu_j)^t \Sigma_j^{-1}(x - \mu_j)\right] \tag{6}$$

where μ_k = the mean vector, and Σ_k = the covariance matrix.

A general solution can be obtained by finding where in the feature space each probability density is larger than all others. This partition of the feature space into regions also can be performed using the logarithm of P(x/j), where:

$$1n[P(x/j)] = -\frac{1}{2}(x - \mu_j)^t \Sigma_j^{-1}(x - \mu_j) - \frac{n}{2} 1n(2\pi) - \frac{1}{2} 1n(|\Sigma_j|). \tag{7}$$

Omitting the constant term $\frac{n}{2}1n(2\pi)$, we can now say that a vector x belongs to class j if the following expression is minimal for the class j:

$$F_k = (x - \mu_k)^t \Sigma_k^{-1}(x - \mu_k) + 1n(|\Sigma_k|) \tag{8}$$

where k = 1, 2, . . . , J.