Fuzzy logic in computer-aided breast cancer diagnosis: analysis of lobulation[1]

Boris Kovalerchuk [a,b], Evangelos Triantaphyllou [a,b,*],
James F. Ruiz [c], Jane Clayton [d]

[a] *Department of Industrial Engineering, Louisiana State University, 3128 CEBA Building, Baton Rouge, LA 70803-6409, USA*
[b] *Stanley S. Scott Cancer Center, Louisiana State University Medical Center, 2025 Gravier Street, Suite 622, New Orleans, LA 70112, USA*
[c] *Department of Radiology, Woman's Hospital, 9050 Airline Highway, Baton Rouge, LA 70895, USA*
[d] *Department of Radiology, Louisiana State University Medical Center, 1542 Tulane Ave., New Orleans, LA 70112, USA*

Abstract

This paper illustrates how a fuzzy logic approach can be used to formalize terms in the American College of Radiology (ACR) Breast Imaging Lexicon. In current practice, radiologists make a relatively subjective determination for many terms from the lexicon related to breast cancer diagnosis. Lobulation and microlobulation of nodules are two important features in the ACR lexicon. We offer an approach for formalizing the distinction of these features and also formalize the description of intermediate cases between lobulated and microlobulated masses. In this paper it is shown that fuzzy logic can be an effective tool in dealing with this kind of problem. The proposed formalization creates a basis for the next three steps: (i) extended verification with blinded comparison studies, (ii) the automatic extraction of the related primitives from the image, and (iii) the detection of lobulated and microlobulated masses based on these primitives. © 1997 Elsevier Science B.V.

---

* Corresponding author. E-mail: ietrian@lsuvm.sncc.lsu.edu
[1] An early version of this paper was presented in the IEEE-FUZZY '96 Conference in New Orleans, LA, USA, September 9–11, 1996.

1. Introduction

Current diagnostic methods in digital mammography [1,9] are based primarily on neural networks without incorporating fuzzy logic. Nevertheless, it should be mentioned that these methods use degrees of irregularity and circularity which are similar to key concepts in fuzzy logic. These degrees are used as inputs to neural networks [10]. In this paper we apply a fuzzy logic approach for classifying a mass found in a mammogram as lobulated or microlobulated. The lobulated and microlobulated features of a mass are important in breast cancer diagnosis [8].

The proposed analysis is based on the medical definitions of the previous two terms, as given by the American College of Radiology (ACR) Breast Imaging Lexicon. According to this lexicon, a mass has 'lobular' shape if 'it has contours with undulations'. Note that the lexicon defines the notion 'lobular' without any indication of the size or number of undulations and without defining the concept of 'undulation'. The descriptive words in each category describe a continuum from benign to malignant. A lobular mass is most often benign, although a few malignancies will be lobular. Lobular malignancies are usually well differentiated pathologically. Furthermore, a mass with microlobulated margins has a lower chance for malignancy than one with indistinct or spiculated margins. A microlobulated mass would fit into the low-intermediate suspicion category #4 of BI-RADS, and would have a 10–20% chance of malignancy.

In this paper the concept of undulation is defined as the contour between the minima of adjacent concavities. The depth of such concavities may vary from small to very large (as explained in greater detail in Section 2). Therefore, for a formal computer algorithmic analysis, means that if a mass has any one of 'small/medium/large undulation', then the algorithm should classify it as lobular. But this is not necessarily what occurs in a real life situation because a radiologist may take into account the size, the number of undulations, and how deep they are. However, the ACR lexicon does not mention these attributes in the formal definition of lobulation. Therefore, it is likely that different radiologists may have different perceptions about the size and number of undulations sufficient to classify the shape of a mass as lobular.

The term 'microlobulated margins' means (according to the ACR lexicon) that 'the margins undulate with short cycles producing small undulations'. Again, different radiologists may have different perceptions of what 'short cycles' and 'small undulation' mean. The ACR lexicon does not provide a unified framework for defining these terms in a consistent and objective manner and again radiologists are left making subjective and individual decisions regarding these characteristics.

The following two hypothetical examples highlight the need for a unified framework for defining terms related to the shape of masses in mammograms:

1. Suppose that a radiologist has found one 'big' and two 'small' undulations in a given mass. Does this mean that the mass is lobular or microlobular or do both features coexist? Also suppose that for the same mass a second radiologist has decided that there are two 'big' and one 'small' undulations. Again, we have the same question: 'Is this mass lobular or microlobular or do both features coexist?'.

2. Suppose that in some study, five out of ten radiologists concluded that a particular mass is lobular, but the other five came to the opposite conclusion. How should we train a computer system to detect a lobular mass by utilizing this contradictory experience? Should we exclude these cases from the training set? However, similar cases may appear again in a real life situation. If we exclude these cases, any trained detection system will diagnose them arbitrarily, although, most properly it should not identify lobular features.

Example 2 illustrates a typical source of intra- and extra-observer variability in mammography and some of its consequences. How can one minimize these problems? This paper proposes a lobular/microlobular mass identification approach which addresses this methodological and practical problem. This approach can also become the basis for analyzing and formalizing other ACR lexicon terms. The proposed approach is designed in a manner which copies the way human experts make decisions. Therefore, this paper will concentrate only on the development of an approach for formalizing lobularity and microlobularity in the masses found in a mammogram. This paper has the following structure. First it discusses the development of the features, which characterize lobularity and microlobularity and the formalization of these features in terms of a fuzzy logic approach (Section 2). Section 3 develops the notions of degrees of lobularity and microlobularity based on formalized features. Finally, the paper ends with some concluding remarks.

2. Formalization with fuzzy logic

In this section we change the previous two definitions slightly. We define a mass to be lobular if it has a contour with some big and deep undulations. The margins of a mass are microlobulated if they have several small concavities (cycles) producing several small and shallow undulations. At a first glance it may appear that we did not improve the precision of the definitions. However, these reformulations are of critical importance. They allow us to apply fuzzy logic and express the original two principal ACR definitions as functions of secondary and easily fuzzifiable terms.

The above considerations involve two important fuzzy terms, namely the terms 'some' and 'several'. These terms have a rather clear meaning when they are used in context with other terms of natural language [2,3]. We can then define a fuzzy set with the fuzzy terms {*few*, *some*, *several*, *many*} for the number of undulations. Note that the number of undulations can be equal to 0, 1, 2, 3,...etc.

For instance, for the fuzzy term 'few' the number of undulations can be set equal to 0. That is, the corresponding family of the four fuzzy membership functions are: $\mu_{\text{few}}(x)$, $\mu_{\text{some}}(x)$, $\mu_{\text{several}}(x)$, and $\mu_{\text{many}}(x)$ (Fig. 1). Some possible sampled values of these membership functions could be: $\mu_{\text{few}}(2) = 1/3$, $\mu_{\text{some}}(2) = 2/3$, $\mu_{\text{some}}(3) = 1$, $\mu_{\text{many}}(2) = 0$, etc. Interviewed radiologists felt comfortable with this formalization. Although one may argue with the numerical values of the above membership functions, the main issue here is that it is possible to effectively quantify fuzzy concepts which are critical for consistent and objective classification of masses as lobular or microlobular.

Next we define the meaning of the terms of the fuzzy set *{small, big}*. This set is crucial in defining the size of undulations. First we need an adequate scale to measure the length of a given undulation. We consider the length of an undulation in relative terms since different masses may have different sizes. For instance, an undulation of 3 mm in length could be considered as 'microlobular' in a large mass while a small mass with the same undulation could be considered as 'lobular'.

Therefore, we first need to compute L; the maximum length of a mass. This approach allows us to estimate the undulation length as a fraction of L. Fig. 2a presents a mass with undulations. Specifically, the curve between points A and B is an undulation. We can formalize the fuzzy terms 'small' and 'big' by characterizing undulations on a scale determined by the relative undulation length (Fig. 2b). According to the membership functions in Fig. 2b, a relative length of more than $L/4$ can be defined as a big undulation, while an undulation of relative length of less that $L/12$ could be considered as a small undulation. Undulations of intermediate length can be assigned intermediate membership values.

Since masses may have varying degree of depth of lobularity one can also define the fuzzy membership functions regarding the ('shallow' or 'deep') depth of undulations. Thus, we introduce a relative measure of the depth of lobularity, which is defined as a fraction of the maximum length (denoted as L) of the mass. This step is similar to those described in the previous fuzzy sets.

The concept of a lobular mass can now be formulated as follows: A mass is lobular if it has at least 3 undulations with length and depth of not less than $L/4$. We can also formulate the concept of microlobulated mass margins. The mass margins are microlobulated if there are at least 6 undulations with length and depth of not more than $L/12$. These definitions are based on the interdependence of the concepts of size, depth, and number of undulations and can be used to quantify the
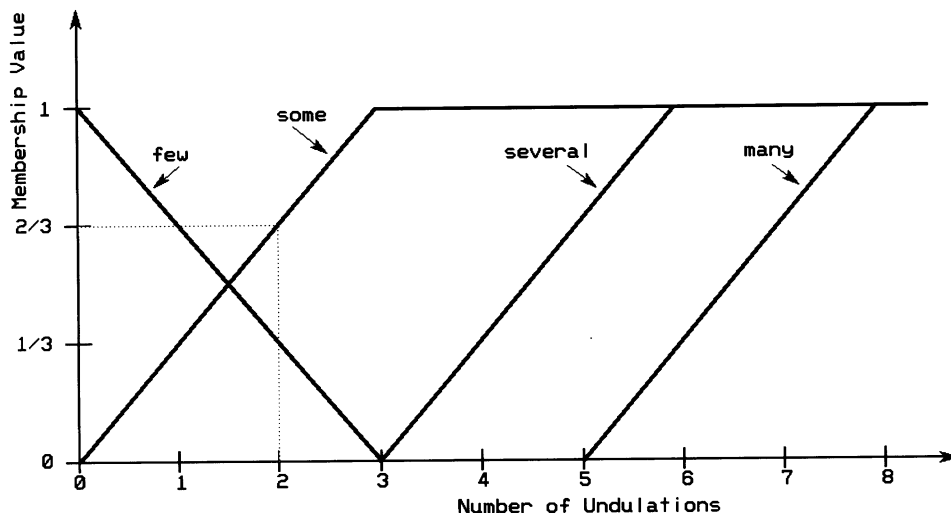

Fig. 1. Membership functions related to the number of undulations.
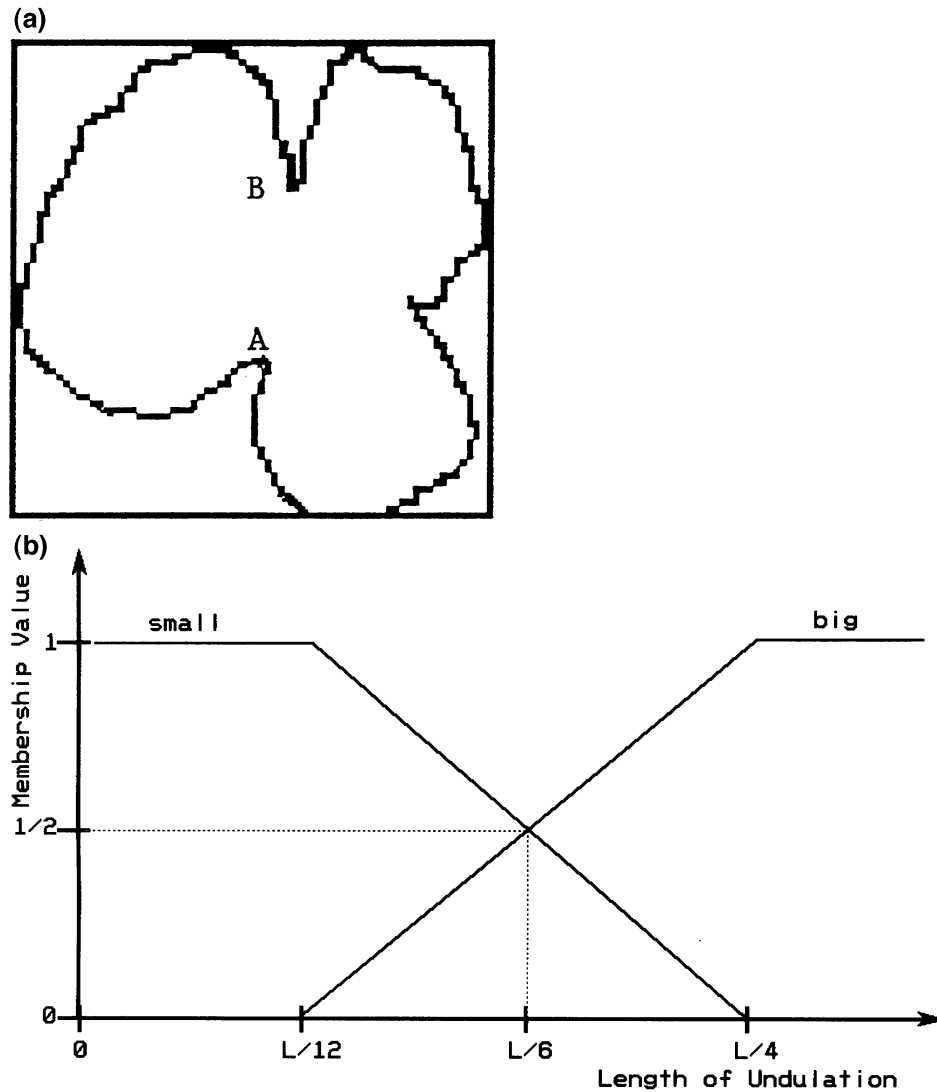
(a)

(b)

Fig. 2. (a) A mass with undulations and (b) Membership functions related to the length of undulations.

concepts of lobular and microlobular masses objectively and consistently. We present the fuzzy logic structures for the lobular and microlobular concepts in Figs. 3 and 4, respectively.

Fig. 3 shows the structure of a mass with three undulations. Each undulation is presented with its length and depth. All these undulations are big and deep. Hence, all membership functions are equal to 1.00 and according to our formalization such a mass is lobular. Similarly, Fig. 4 shows a microlobulated mass with 6 undulations and all of them are small and shallow.

-->	undulation 1	length	= big	μ_{big}(undulation1)	= 1.00
		depth	= deep	μ_{deep}(undulation1)	= 1.00
MASS -->	undulation 2	length	= big	μ_{big}(undulation2)	= 1.00
		depth	= deep	μ_{deep}(undulation2)	= 1.00
--->	undulation 3	length	= big	μ_{big}(undulation3)	= 1.00
		depth	= deep	μ_{deep}(undulation3)	= 1.00

Fig. 3. Fuzzy logic structures for a lobular mass.

The previous definitions allow some masses to be classified as both lobular and microlobulated without any contradiction if the mass has at least nine undulations (of which three are lobular and six are microlobular). That is, one just needs to join the structures given in Figs. 3 and 4. Cases of an intermediate nature can also be formalized. Examples of such cases are depicted in Fig. 5.

We take the three biggest and deepest undulations and compute the minimum of their membership function values for the terms 'big' and 'deep'. We define this value as the degree of lobularity (or DL). For instance, for the mass described in Fig. 5 the minimum for the first three undulations is 0.70, that is, for this case DL = 0.70. Similarly, it can be easily verified that the degree of microlobularity (or DM) computed with the remaining 6 undulations is 0.60. Such estimates can be used as inputs for a breast cancer computer-aided diagnostic (CAD) system.

If the number of undulations is less than 3, we combine the membership functions for the length and depth with a membership function for the number of undulations (as defined in Fig. 1). In this combination, we compute the minimum of these three values in accordance with standard fuzzy logic practice. We analyze the arguments for the minimum in the next section. Now let us consider, for instance, the mass with two undulations presented in Fig. 6.

Fig. 6 allows to compute 0.60 as the corresponding degree of lobularity (DL), while Fig. 1 shows that $\mu_{some}(2) = 0.66$ for a case with 2 undulations. Thus, their minimum of 0.60 characterizes the lobularity of this mass. It is important to state here that the proposed fuzzy membership functions are only indicative. Their exact

--->	undulation 1	length	= small	μ_{small}(undulation1)	= 1.00
		depth	= shallow	$\mu_{shallow}$(undulation1)	= 1.00
--->	undulation 2	length	= small	μ_{small}(undulation2)	= 1.00
		depth	= shallow	$\mu_{shallow}$(undulation2)	= 1.00
MASS --->	undulation 3	length	= small	μ_{small}(undulation3)	= 1.00
		depth	= shallow	μ_{deep}(undulation3)	= 1.00
--->	undulation 4	length	= small	μ_{small}(undulation4)	= 1.00
		depth	= shallow	$\mu_{shallow}$(undulation4)	= 1.00
--->	undulation 5	length	= small	μ_{small}(undulation5)	= 1.00
		depth	= shallow	$\mu_{shallow}$(undulation5)	= 1.00
--->	undulation 6	length	= small	μ_{small}(undulation6)	= 1.00
		depth	= shallow	$\mu_{shallow}$(undulation6)	= 1.00

Fig. 4. Fuzzy logic structures for a microlobulated mass.

--->	undulation 1	length	= big	μ_{big}(undulation1)	= 0.80
		depth	= deep	μ_{deep}(undulation1)	= 0.70
--->	undulation 2	length	= big	μ_{big}(undulation2)	= 0.73
		depth	= deep	μ_{deep}(undulation2)	= 0.71
--->	undulation 3	length	= big	μ_{big}(undulation3)	= 0.90
		depth	= deep	μ_{deep}(undulation3)	= 0.80
MASS --->	undulation 4	length	= small	μ_{small}(undulation4)	= 0.90
		depth	= shallow	μ_{shallow}(undulation4)	= 0.80
--->	undulation 5	length	= small	μ_{small}(undulation5)	= 0.90
		depth	= shallow	μ_{shallow}(undulation5)	= 0.70
--->	undulation 6	length	= small	μ_{small}(undulation6)	= 0.60
		depth	= shallow	μ_{deep}(undulation6)	= 0.70
--->	undulation 7	length	= small	μ_{small}(undulation7)	= 0.67
		depth	= shallow	μ_{shallow}(undulation7)	= 0.97
--->	undulation 8	length	= small	μ_{small}(undulation8)	= 0.80
		depth	= shallow	μ_{shallow}(undulation8)	= 1.00
--->	undulation 9	length	= small	μ_{small}(undulation9)	= 0.84
		depth	= shallow	μ_{shallow}(undulation9)	= 0.79

Fig. 5. Structural descriptions for a fuzzy lobular and microlobulated mass.

form can be determined from a consensus approach among radiologists and/or by using historic data. In the next section we present the DL and DM ideas formally.

3. Degrees of lobularity and microlobularity

Radiologists use an informal approach in determining the lobularity and microlobularity of a mass. To maintain consistency in these evaluations and increase objectivity, we need to formalize these concepts. Let us first consider the two masses depicted in Fig. 7. Intuitively, the first mass has deep undulations, while the second mass has shallow undulations. Different measures can be created to formalize this distinction. Fig. 7a shows two distances d_1 and d_2, defined between the points A and C and between the points B and E, respectively, for undulation 1 (i.e., U_1). If each of them is no less than $L/4$, then the undulation is deep (see also Fig. 7b). If these distances are no more than $L/12$, then undulation 1 is shallow (see also Fig. 7a). This situation indicates that formally the depth D of the undulation closely depends on the pair d_1, and d_2 (this concept is not to be confused with the one of depth of lobularity which was defined in Section 2). The method used to compute these values was considered in [6].

	--->	undulation 1	length	= big	μ_{big}(undulation1) = 0.80
			depth	= deep	μ_{deep}(undulation1) = 0.70
MASS					
	--->	undulation 2	length	= big	μ_{big}(undulation2) = 0.60
			depth	= deep	μ_{deep}(undulation2) = 0.60

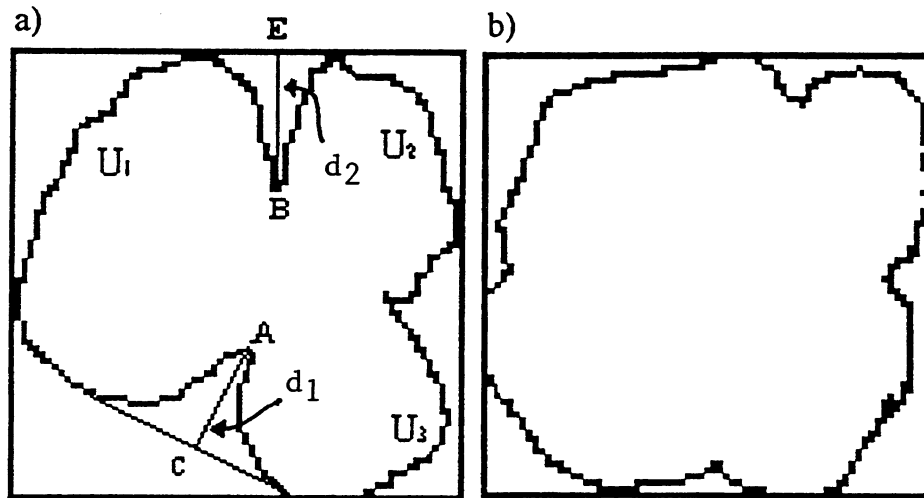Fig. 6. Fuzzy logic structures for a mass with less than three undulations.

Fig. 7. Masses (a) with deep and (b) shallow undulations.

The values of $\mu_{\mathrm{deep}}(d_1)$ and $\mu_{\mathrm{deep}}(d_2)$ are computed by using the corresponding membership function in Fig. 2 so that the previous two measures are transformed into a single degree of lobularity for a given undulation. Recall that we use the same membership functions for the length and depth of undulations. This is done by substituting the terms, 'big' for 'deep' and 'small' for 'shallow'. Next, we compute *min* $\{\mu_{\mathrm{deep}}(d_1), \mu_{\mathrm{deep}}(d_2)\}$, which could be considered as the degree of depth of the undulation. That is:

$$\mu_{\mathrm{deep}}(\mathrm{undulation}) = \min\{\mu_{\mathrm{deep}}(d_1), \mu_{\mathrm{deep}}(d_2)\}.$$

Similarly, we define the degree of shallowness of an undulation as:

$$\mu_{\mathrm{shallow}}(\mathrm{undulation}) = \min\{\mu_{\mathrm{shallow}}(d_1), \mu_{\mathrm{shallow}}(d_2)\}.$$

Observe that the length of undulation1 (i.e., U_1) is measured as the length of the mass margin between points A and B (see also Fig. 7a).

Now one can define the Degree of Lobularity (DL) of a mass as follows:

$$\mathrm{DL(mass)} = \min\{\mu_{\mathrm{some}}(k), \min_{k \geq i \geq 1}\{\mu_{\mathrm{big}}(U_i), \mu_{\mathrm{deep}}(U_i)\}, \tag{1}$$

where U_1, U_2,..., U_k are undulations such that:

$$\min_{k \geq i \geq 1}\{\mu_{\mathrm{big}}(U_i), \mu_{\mathrm{deep}}(U_i)\} \geq 0.50.$$

Similarly, we define the degree of microlobularity (DM) of a mass with k undulations:

$$\mathrm{DM(mass)} = \min\{\mu_{\mathrm{several}}(k), \min_{k \geq i \geq 1}\{\mu_{\mathrm{small}}(U_i), \mu_{\mathrm{shallow}}(U_i)\}\}, \tag{2}$$

where U_1, U_2,..., U_k are undulations such that:

$$\min_{k \geq i \geq 1}\{\mu_{\text{small}}(U_i), \mu_{\text{shallow}}(U_i)\} \geq 0.5.$$

For the extreme case of $k = 0$, we have $\mu_{\text{some}}(k) = 0$ and $\mu_{\text{several}}(k) = 0$ (see Fig. 1). Therefore, both degrees of lobularity and microlobularity are equal to 0, i.e., the outcome corresponds to what is expected with common sense.

There are some theoretical and experimental arguments for the general case (e.g. [2,4,5]) justifying Eqs. (1) and (2). However, we can also use some additional arguments derived from this mammographic problem. A consistent computer-based breast cancer diagnostic system should refuse to diagnose a mammogram with a significant number of doubtful features. We can express how doubtful a given feature is by some degree between 0 and 1, with the highest degree of doubt given at 0.50. The values of DL and DM are examples of such degrees. For these uncertain (doubtful) features, a CAD system can suggest the presence of a particular feature, but only with some degree of reliability. This reliability can be very low. Also, this degree of reliability depends on the particular values of the DL and DM measures. Therefore, the equations used to define DL and DM become even more critical.

This situation can be explained with a modified example from Fig. 3. Assume that the first five membership functions for undulations are equal to 1.00 and the sixth function is equal to 0.60 (i.e., $\mu_{\text{deep}}(\text{undulation3}) = 0.60$). Then Eq. (1) gives us a 'pessimistic' estimate, i.e., low degree of certainty for the presence of lobularity, DL = 0.6. Substituting in Eq. (1) the minimum operation for the maximum will give us an 'optimistic' assessment, i.e. high degree of lobularity, DL = 1.00 for this case. In the last 'optimistic' estimate we ignore and lose the warning information (i.e., the fact that $\mu_{\text{deep}}(\text{undulation1}) = 0.60$). The value 0.60 suggests that one should be cautious and study the case in great detail. However, no warning information is lost if we use the 'pessimistic' min operation in Eqs. (1) and (2). Therefore, for critical questions regarding cancer diagnosis, we see that the 'pessimistic' strategy is the safest. We also consider statements with a low degree of reliability as a preliminary suggestion indicating that we need to switch the set of features to a higher level of detail in order to fully evaluate the complexity of a given case. Some experiments by Doi et al. [1] and Wu et al. [9], have shown that relatively simple cases can be diagnosed within a small feature space. For more complicated cases we need a pathologically confirmed training sample with more features and a specifically designed diagnostic method. A CAD system designed as above, will have switching capabilities based on the described approach.

4. Conclusions

Radiologists often make relatively subjective determinations for many features related to breast cancer diagnosis. We have formalized some important features from the ACR breast imaging lexicon, i.e., lobulation and microlobulation of nodules. This formalization is the basis the next three steps:
1. extensive radiologic validation;

2. automatic detection of lobulation/microlobulation in a mammographic image; and
3. similar formalizations of the other terms from the breast imaging lexicon.

This study suggests that fuzzy logic can be an effective tool in dealing with this kind of medical problem.

It should also be stated here that the ACR breast imaging lexicon involves many concepts which could be defined in a fuzzy logic approach similar to the proposed lobulation analysis. However, as has been shown in our previous work with breast calcifications [6], the various features presented in the lexicon pose a broad range of problems requiring tailored solutions. Analyzing all the concepts covered in the ACR lexicon (approximately 30), is outside the scope of this paper and will require a sequence of similar papers.

However, such a goal is of great importance in breast cancer diagnosis, since as it has been demonstrated that traditional artificial intelligence and statistical methods of pattern recognition and diagnosis may be dramatically unreliable [7].

The proposed fuzzy logic approach is both feasible and effective because this type of fuzzy logic approach has been successful in many other areas (see also the study by Kovalerchuk and Klir, 1995 [2]). Its application in breast cancer diagnosis is only matter of time. The proposed fuzzy logic approach, when it is considered with the success of the application of fuzzy logic in many other domains, has the potential of opening a new and promising direction for effective and early breast cancer diagnosis.

Acknowledgements

References

[1] K. Doi, M.L. Giger, R.M. Nishikawa, K.R. Hoffmann, H. Macmahon, R.A. Schmidt, and K.-G. Chua, Digital radiography: a useful clinical tool for computer-aided diagnosis by quantitative analysis of radiographic images, *Acta Radiologica* 34(5) (1993) 426–439.

[2] B. Kovalerchuk, and G. Klir, Linguistic context spaces and modal logic for approximate reasoning and fuzzy-probability comparison. In: *Proceedings of the Third International Symposium on Uncertainty Modeling and*, *NAFIPS'95* (IEEE Computer Society Press, College Park, Maryland, 1995) a23–a28.

[3] B. Kovalerchuk, Linguistic context spaces: necessary frames for correct approximate reasoning, *Int. J. Gen. Sys*. 24(4) 1996 (in print).

[4] B. Kovalerchuk and B. Dalabaev, T-norms as scales, in: *The First European Congress of Fuzzy and Intelligent Technologies* (Aachen, Germany, 1993) 1482–1487.

[5] B. Kovalerchuk and V. Taliansky, Comparison of empirical and computed values of fuzzy conjunction, *Fuzzy Sets Sys*. 46(2) (1992) 49–53.

[6] B. Kovalerchuk, E. Triantaphyllou, and J.F. Ruiz, Monotonicity and Logical Analysis of Data: A Mechanism for Evaluation of Mammographic and Clinical Data, *Proceedings of the 13-th Symposium for Computer Applications in Radiology (SCAR)* (Denver, CO, June 6–9, 1996) 191–196.

[7] B. Kovalerchuk, E. Triantaphyllou, and J.F. Ruiz, The reliability issue of computer-aided breast cancer diagnosis, *Appl. Intell.* (1997) in press.

[8] L. Tabar and P.B. Dean, *Teaching Atlas of Mammography* (Verlag, New York, NY, 1986).

[9] Y. Wu, M. Giger, C. Vyborny, R. Schmidt, and C. Metz, Artificial neural networks in mammography: application to decision making in the diagnosis of breast cancer, *Radiology* 187(1) (1993) 81–87.

[10] Y. Wu, K. Doi, M. Giger, C. Metz and W. Zhang, Reduction of false positives in computerized detection of lung nodules in chest radiographs using artificial neural networks, discriminant analysis and a rule-based scheme, *J. Digital Imaging* 17(4) (1994) 196–207.

.