SPATIAL IMAGE VARIABILITY ANALYSIS

M. S. Pattichis[1], T. Cacoullos[2], and P. Soliz[3]

[1]Department of EECE, The University of New Mexico, Albuquerque, NM 87131-1356
[2]Department of Mathematics, The University of Athens, Athens, Greece
[3]Kestrel Corporation, Albuquerque, New Mexico

ABSTRACT

In our analysis, each image is partitioned into a number of non-overlapping, spatial regions. Each spatial region is viewed as a Region of Interest (ROI). A grade is assigned to each ROI from a set of independent raters. The model is applied in lung image analysis, where a grade of 0/1 is assigned to each ROI. Here, 1 represents success, while 0 represents failure to detect a hypothesized pattern in the region. Methods for establishing spatial symmetry and spatial growth analysis are presented. In addition, a Bayes and a novel summation classifier are used for classifying the entire lung based on their regional grades. When compared against the readers, the two classifiers have a total misclassification error of the order of inter-rater variability error.

1. INTRODUCTION

In image analysis, there is always a need to establish "ground truth" over specific regions of interest (ROIs). The problem is complicated when there is no absolute agreement on what constitutes ground truth. For example, suppose that a region of interest (ROI) is graded by multiple readers, each giving his or her own grade. In this case, it is often found, especially in medical imaging, that there is strong variability between the grades. We use the term "spatial image variability analysis" to refer to image processing methods focused on analyzing the spatial variability resulting from multiple grades over regions of interest (ROIs).

Given an image, let I denote the set of pixels that are in the spatial support of the image. Define a partition of I by

$$I = R_1 \cup R_2 \cup \cdots \cup R_M ,$$

where the R_i denote non-overlapping regions of interest, and M denotes the number of regions. A grade over the ROIs is defined as a random variable X_i of each ROI into a set of finite grades

$$X_i : R_i \to \{0,1,\ldots,N-1\},$$

where N denotes the number of possible grades.

Using the collection of all the grades found for all ROIs $Y = (X_1, X_2, \ldots, X_M)$, a classification decision is made for the entire image. We use the symbol C_i to denote the *i-th* classifier (which could represent a human reader or a computer classification system), and note that

$$C_i : \{0,1,\ldots,N-1\}^M \to \{0,1,\ldots,P-1\},$$

where $\{0,1,\ldots,N-1\}^M$ denotes the M-fold tensor product of the integer set $\{0,1,\ldots,N-1\}$, and P denotes the maximum number of possible classifications.

The rest of the paper is organized into five Sections. In Section 2, we develop a model for lung image analysis. The model is aimed at classifying regions of interest where lung opacities are detected. In Section 3, we develop spatial pattern analysis methods. In this Section, we first look at the problems of : (i) establishing symmetry, and (ii) developing models for modeling opacity growth. In Section 4, we discuss how to develop computer classifiers based on reader-statistics. In the concluding Section, we summarize our findings and discuss future directions.

2. LUNG IMAGE ANALYSIS

In developing an ROI-based system for the lung, we follow the International Labor Organization (ILO) guidelines [1]. In our example, we use chest radiographs of miners.

The lung is divided into six regions of interest $(M = 6)$. Each region of interest is marked with an 'X' if lung opacities are found in the region. There are two possibilities $(N = 2)$ for each random variable X_i, and $i = 1, 2, \ldots, 6$.

Right Lung	Left Lung
R_1 (UR)	R_4 (UL)
R_2 (MR)	R_5 (ML)
R_3 (LR)	R_6 (LL)

The entire chest radiograph is then classified according to opacity shape and size. We restrict our attention to the case where the opacity profusion is of shape type *q*. The size of the opacities is further sub-divided into 12 categories arranged in order of increasing size: *q0/-, q0/0, q0/1, q1/0, q1/1, q1/2, q2/1, q2/2, q2/3 q3/2, q3/3, q3/+*. In order to be able to apply a statistical analysis on sufficiently large populations, we only consider the main profusion categories *q0, q1*, and *q2*. In this new categorization, we assign *qi/j* grades as belonging to the *qi* class. This gives us three main categories $(P = 3)$. Each chest radiograph is graded by two readers, giving two classifiers C_1 and C_2 for establishing ground truth.

0-7803-7523-8/02/$17.00 ©2002 IEEE

Table 1. Left-Right agreement based on pooled data. The weighted kappa statistic between right and left lungs was 0.72 with a 95% confidence interval of 0.65-0.78. The simple kappa statistic was 0.65 with a 95% confidence interval of 0.58-0.72.

		Right			
Left	**0**	**1**	**2**	**3**	**Total**
0	6	8	0	0	14
1	9	116	13	1	139
2	0	18	78	5	101
3	0	2	9	29	40
Total	15	144	100	35	294

3. SPATIAL ANALYSIS

3.1 Spatial Symmetry

Spatial symmetry can be addressed by comparing sums of variables. For example, to investigate right-left symmetry, we form the right lung sum

$$S_{123} = X_1 + X_2 + X_3,$$

and compare it against the left lung sum

$$S_{456} = X_4 + X_5 + X_6.$$

The results are shown in Table 1. In perfect left-right symmetry, all the off-diagonal entries would be zero. The agreement between the two lung grades is evaluated using both simple and weighted kappa statistics, as shown in the caption of table 1 [2]. For this example, the overall inter-rater agreement was found to be very good [2].

Similarly, if examining other types of symmetries, such as up-down, or rotational symmetries, we simply form an appropriate sum of variables that we require to remain fixed.

3.2 Spatial Growth Analysis

Spatial growth analysis can only be effectively addressed if we have a time sequence of images, as generated in longitudinal studies. Often, longitudinal studies are limited in the number of cases that can be analyzed through time. Yet, assuming that the disease progresses through time, becoming more and more severe, it is possible to study the progression by investigating the pattern distributions as sorted by severity.

This idea is demonstrated in Table 2. The spatial-patterns are sorted by severity, from the normal case q0 to q1, and then q2. By examining the most-likely patterns in each case, it is clear that there is growth from the top to the middle, and middle to bottom.

Table 2. Pattern-distribution by class: q0, q1, and q2. Each pattern represents $X_1X_2X_3 - X_4X_5X_6$. The table shows the classifications of a single reader, selected over a population where there is strong agreement (the percentage of chest radiographs where there was disagreement, was only 6%).

q0 patterns

Pattern	Occurrence	Percentage
000-000	96/126	76%
000-100	1/126	1%
001-101	1/126	1%
100-000	6/126	5%
100-100	18/126	14%
101-001	1/126	1%
110-100	1/126	1%
110-110	1/126	1%
111-111	1/126	1%

Most-likely pattern is 000-000 occurring 76%.

q1 patterns

Pattern	Occurrence	Percentage
000-100	1/26	4%
100-100	13/26	50%
100-110	1/26	4%
110-100	1/26	4%
110-110	6/26	23%
111-111	4/26	15%

Most-likely pattern is 100-100 occurring 50%.

q2 patterns

Pattern	Occurrence	
110-110	4/6	67%
111-110	1/6	17%
111-111	1/6	17%

Most-likely pattern is 110-110 occurring 67%.
(The sample size for the q2-population was small, only 4 cases)

4. SPATIAL PATTERN CLASSIFICATION

We consider two computer-based classifiers: (i) a Bayes classifier based on the two readers, and (ii) a summation classifier based on the spatial symmetry and spatial growth analysis.

The implementation of the Bayes classifier was based on the average of both readers. We note that the original reader grades are given in the range of *q0/-* to *q3/+* as described in Section 2. Recall that these categories are arranged in 12 categories, in increasing degree of severity. Hence, their average represents the average severity found by the two readers. We note that Bayes classification is always possible in our formulation. This is primarily due to the fact that there are always finitely many

possibilities that are easily implemented in multidimensional integer arrays.

An important disadvantage of the Bayes classifier is that it is only optimal when the underlying statistical distributions are known exactly. In our example, the distributions are estimated from the data. Given the exponential growth in the number of patterns (here, $N^M = 64$), there is a need for large sample populations in order to ensure accurate density estimation. This is demonstrated in the lack of statistics for many of the possible spatial patterns presented in Table 2. It is also shown in the number of '*' that are present in the Bayes-classifier Karnaugh maps of Figure 1 (see description below).

The summation classifier is based on the spatial analysis Section. From Section 3.2, we know that as the severity of the disease increases from *q0* to *q2*, the number of marked regions also increases. This leads us to consider the total sum

$$S = X_1 + X_2 + X_3 + X_4 + X_5 + X_6,$$

as a single measure for classification. Naturally, any classifier under consideration must also satisfy the symmetries established on the data. In our example, it is clear that left-right symmetry is satisfied. Another important feature of the summation classifier is that it is symmetric with respect of spatial opacity growth. In other words, the summation classifier would work equally well whether increased severity occurs as a result of opacity region growth from the lowest regions to the top regions, or it is a result of region growth from the top to the bottom regions. This spatial growth pattern is known to hold true for some opacity types that are not of type *q*.

For minimum misclassification error, the summation classifier is implemented by:

$$\text{Classify } y \text{ in } q0, \text{ if } s \le 1,$$
$$\text{Classify } y \text{ in } q1, \text{ if } 2 \le s \le 4,$$
$$\text{Classify } y \text{ in } q2, \text{ if } s = 5, 6,$$

where $y = (x_1, x_2, x_3, x_4, x_5, x_6)$. When compared against each reader, both the Bayes and the summation classifiers produced a total error of approximately 35%, which is the same as the total error of comparing each reader against the other.

To visualize each binary classifier, we use a modification of Karnaough maps [3]. The classifiers are illustrated in Figure 1. In Figure 1, both the Bayes and the summation classifiers are shown. We use four 4-variable K-maps for visualizing each 6-variable binary classifier. The top two rows are for describing the Bayes classifier, while the lower two rows are for describing the summation classifier. We use the symbol '-' to represent the absence of a mark, while using the symbol '*x*' to represent the presence of a mark. In each cell of each map, we plot the q-classification for the pattern: (i) 0 for $q0$, (ii) 1 for $q1$, (iii) 2 for $q2$, and (iv) '*' when the pattern has not been observed (Bayes classifier only).

For plotting the maps, we use the upper right and upper left variables X_1X_4 to select between the four maps. For example, for the summation classifier, when both regions are marked, we have $X_1X_4 = xx$ which select the lower-right 4-variable map in Figure 1.

The right-left symmetry is demonstrated by requiring that: (i) the $X_1X_4 = -x$ and $X_1X_4 = x-$ tables be identical, and (ii) that the $X_2X_5 = x-$ column be identical to the $X_2X_5 = -x$ column, in each 4-variable K-map. Note that the symmetry is perfect for the summation classifier, while it is nearly perfect for the estimated Bayes classifier. In fact, in the Bayes classifier, the symmetry would be satisfied if some unobserved patterns in the top rows of the second and third K-maps were correctly classified (described by $X_1X_4 = -x$ and $X_3X_6 = --$, and $X_1X_4 = x-$, $X_3X_6 = --$ respectively).

5. CONCLUDING REMARKS

A general method for analyzing images in terms of regional grades over regions of interest has been developed. Examples of establishing spatial symmetry and spatial growth are presented in terms of the regional grades of lung images. Furthermore, a Bayes and a summation classifier are developed for classifying the entire lung from its regional grades. The function of each classifier is summarized in terms of a collection of Karnaugh map plots.

In the future, our methods will be extended to the case of multiple raters, and a larger number of regions of interest. A clear challenge with many ROIs would be to develop efficient methods for manipulating the resulting classifiers.

6. ACKNOWLEDGMENTS

Funding for the study was through a grant from the National Institute for Occupational Safety and Health (NIOSH), Grant 2R44 OHRRGM 03595.

7. REFERENCES

[1] *Guidelines for the use of ILO international classification of radiographs of pneumoconiosis*, 1980, International Labour Office: Geneva.

[2] Fleiss, J. *Statistical Methods for Rates and Proportions*, Wiley, 1981.

[3] McCluskey, E. J., *Introduction to the Theory of Switching Circuits*, McGraw-Hill, 1965.

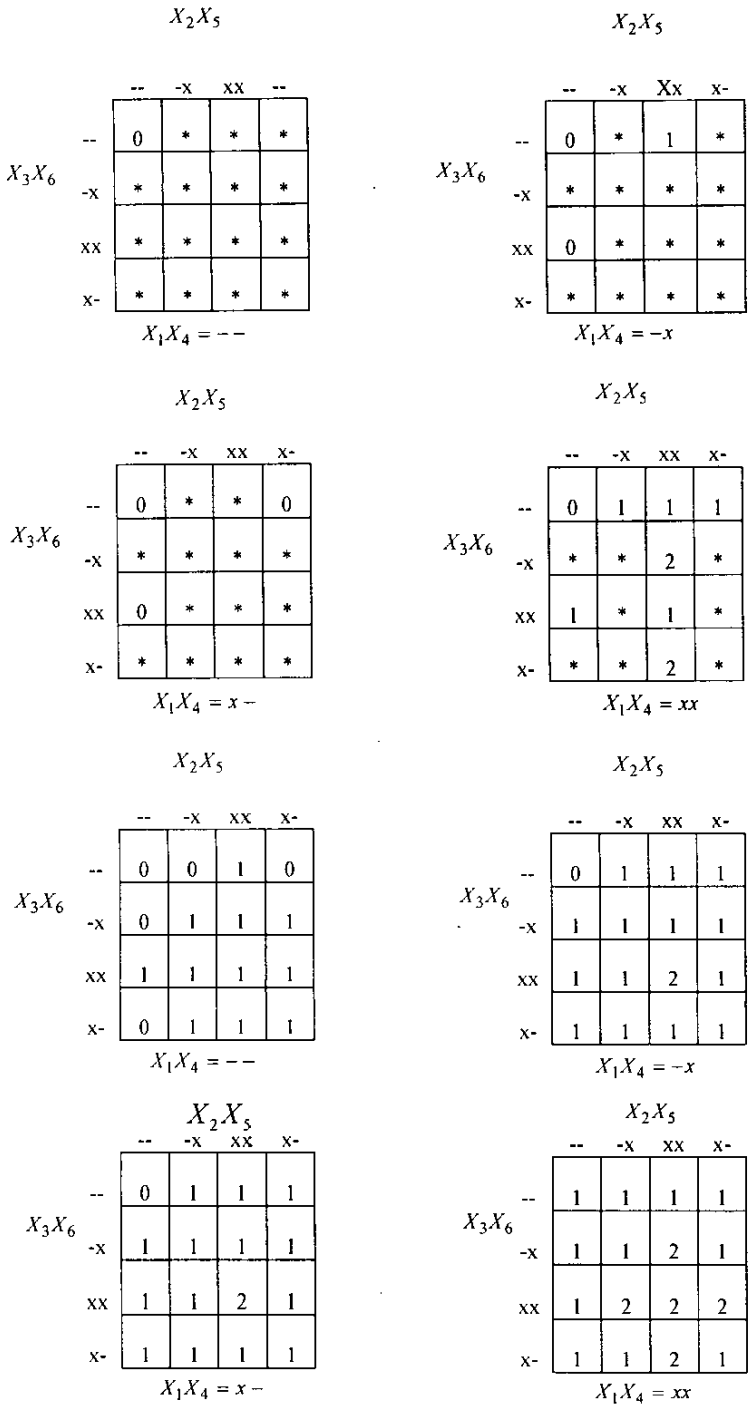

Figure 1. Summary of Bayes and Summation classifier results. A '*' implies that the pattern was never observed.