Managing the Noisy Glaucomatous Test Data by Self Organising Maps

Xiaohui Liu, Gongxian Cheng, and John Wu

Abstract — **One of the main difficulties in obtaining reliable data from patients in glaucomatous tests is the measurement noise caused by the learning effect, inattention, failure of fixation, fatigue etc. Using Kohonen's self-organising feature maps, we have developed a computational method to distinguish between the noise and true measurement. This method has been shown to provide a satisfactory way of locating and rejecting noise in the test data, an improvement over conventional statistical methods.**

I. INTRODUCTION

DIAGNOSIS of visual function losses in glaucomatous patients depends on the analysis of the data collected from corresponding psychophysical tests [4]. One of the main difficulties in analysing such data is the measurement noise caused by patient's learning effect, inattention, failure of fixation, fatigue etc. This implies that the data collected cannot be guaranteed to accurately reflect the visual function of an individual, therefore making the clinician's diagnostic process a challenging one.

One of the traditional approaches to dealing with the measurement noise is to quantify the data variance, e.g. short-term fluctuation [8]. This kind of statistical information is expected to make the clinician aware of the potential noise in the data and therefore s/he can take this information into account when interpreting the data.

The clinician's interpretation using the statistical information is, however, ultimately a subjective one, depending on his/her experience and knowledge. Although the statistical information gives a global measurement about how reliable an individual test is, it don't tell *when* and *where* during the test the noise actually occurred, which is extremely useful for eliminating noise and monitoring patient's behaviour during the span of the test.

This has led us to seek a novel way of dealing with the noise in the test data. In this paper, we introduce a computational method that shows promise in overcoming the difficulties associated with conventional techniques. This method utilises the transition trajectories in Kohonen's self-organising maps (SOM) [5] to demonstrate patients' behaviour during the glaucomatous psychophysical test and these trajectories can be used to show when and where the noise occurred in the test data. The real noise can then be filtered out, and therefore, the clinician can make more accurate decisions based on quality test data.

We have used 263 clinical visual function test records (2630 input vectors to the SOM) from glaucoma patients and suspects at the Moorfields Eye Hospital in London to experiment with the proposed method. In particular, 91 pairs of repeated tests are used to see how successful this method is in identifying and deleting the noise. The findings are very encouraging.

X Liu and G Cheng are with the Department of Computer Science at Birkbeck College, University of London, London WC1E 7HX, U.K. J Wu is with the Department of Preventive Ophthalmology, Institute of Ophthalmology, University of London, London EC1V 9EL, U.K. and is associated with the Moorfields Eye Hospital, London.

II. THE MOTION SENSITIVITY PERIMETRY

The Motion Sensitivity Perimetry (MSP) was proposed in [3] and has been shown to be useful and reliable measurements of glaucoma progressing for early glaucoma patients [10]. The test examines 6 locations (L1,L2,L3,L4,L5 and L6) within the visual field by four different stimuli (S1, S2, S3, S4), which consist of both motion and flicker.

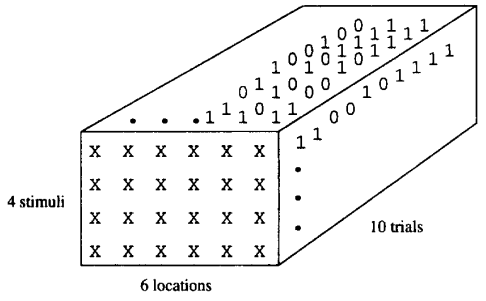

Figure 1. Data structure for each MSP (1=yes, 0=no)

In order to investigate the test behaviour as a function of time, the test strategy used a Single Amplitude Trial (SAT) for ten trials. The principle for SAT was to obtain the patient's behaviour to the same test condition during a given number of trials. A reliable patient's behaviour pattern should be different from an unreliable pattern detected by this strategy. To keep the same test conditions, therefore, the sequence of presentations in MSP used a look-up table rather than by random order. In the table, the sequence of test locations had two patterns (P1:L1+L2+L3+L4+L5+L6, P2:L2+L4+L6+L1+L3+L5). The whole sequence for MSP

0-7803-1901-X/94 $4.00 ©1994 IEEE

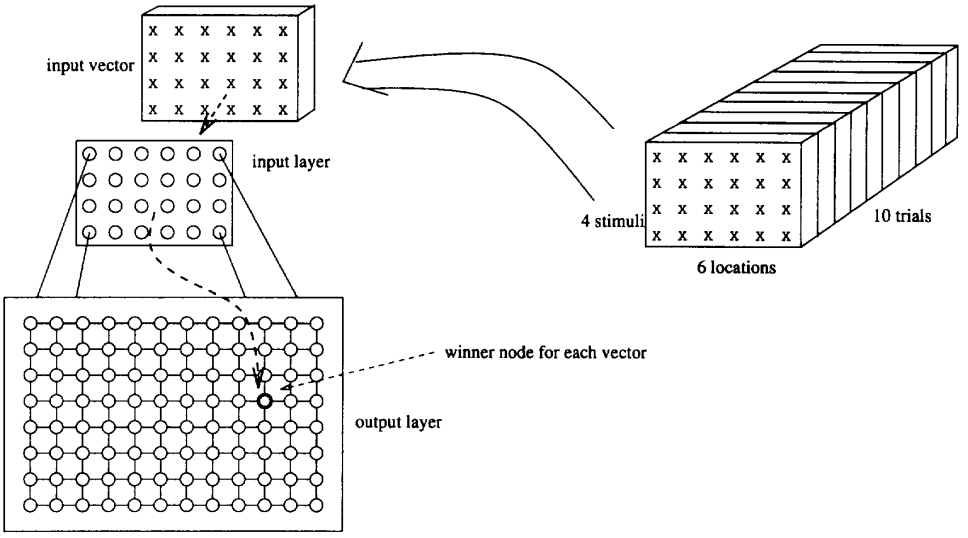

Figure 2. Apply the SOM to the MSP

was (S4P1+S2P2+S1P1+S3P2) × (number of trials). The test data basically constitutes a three dimensional array, as shown in Figure 1.

There are ten trials for each test and each trial leads to 24 data elements: 6 locations are tested using 4 different stimuli. Each element is either 1 which shows that the patient does see the motion or flicker, or 0 which shows the patient does not. The sensitivity of a patient is then defined by how often s/he can see the stimuli.

III. APPLICATION OF THE SOM TO THE MSP

As discussed above, each patient is subjected to ten repeated trials during a single test and each trial has a fixed pattern involving 6 locations and 4 stimuli. Therefore it would be natural to view the testing process as a patient going through ten identical test cycles. It would also be natural to assume that the response pattern should be similar from one test cycle to another for a patient having a *reliable* test. This assumption, however, is often violated by factors such as the learning effects, fatigue etc., which lead to significant differences between patient's response patterns. Therefore, *noise* is typically involved in the response patterns of a psychophysical test.

The proposed method for filtering out unreliable data consists of three steps. Firstly, Kohonen's learning technique is used to train a network capable of generating maps which reflect the patient's test behaviour, as illustrated in equation (1).

$$\frac{dw_i}{dt} = \alpha(t)\gamma(i,t)(x_j - w_i) \qquad (1)$$

The α is a monotonical decay function with time and γ is a neighbourhood function, two of the most popular being the "bubble" adaptation[5] or a Gaussian-type function [9]. The vector of connection weights between input and output neurons is represented by $w_i(i=1,2,...,M)$ where M is the number of output neurons.

Each response pattern for each test cycle (6 × 4 matrix) is used as an input vector to the self-organising map and each winner node is produced on the output map [Figure 2]. In all, 2630 trial data vectors corresponding to 263 tests are used to train the network and the whole data set is reiteratively submitted 100 times in random orders.

Secondly, an effort is made to find a network which shows better *neighbourhood preservations*, i.e. similar input patterns are mapped onto identical or closely neighbouring neurons on the output map. This step is important as we want to map similar response patterns from patients onto similar neurons. We have used the *topographical product* (TP)[1] as a measurement for this purpose where

$$TP = \frac{1}{2M(M-1)} \sum_{j=1}^{M} \sum_{k=1}^{M-1} \frac{1}{k} \log\left(\prod_{l=1}^{k} R_1(j,l) R_2(j,l)\right)$$

While R_1 and R_2 can be calculated as follows:

$$R_1(j,k) = \frac{d^V(w_j, w_{n_k^A(j)})}{d^V(w_j, w_{n_k^V(j)})}$$

$$R_2(j,k) = \frac{d^A(j, n_k^A(j))}{d^A(j, n_k^V(j))}$$

Where $d^V(w_j, w_{n_k^A(j)})$ denotes the distances measured in the input space between w_j and $w_{n_k^A(j)}$ and $d^A(j, n_k^A(j))$ denotes the distances measured in the output space between j and kth

nearest neighbour of j, $n_k^A(j)$. The TP indicates the magnitude of neighbourhood violation. Therefore, the smaller the value of TP is, the better the neighbourhood preservation would become.

Our experiments have shown that using a Gaussian-type neighbourhood function gives much better neighbourhood preservations than a neighbourhood iteration set (i.e., the "bubble" adaptation). This has been shown to be true in some other cases [2, 7].

Having obtained a well-performed network, a 12×8 map in our case, the final step is to generate the behaviour maps for individual patients and and analyse these maps to identify and reject the measurement noise. As far as each patient is concerned, there would be ten winner nodes and nine transitions on the output map. These transitions constitute a transition trajectory, which graphically illustrates how patient's behaviour changed from one trial to the other [Figure 3].

As one of the key SOM features is that similar input vectors would lead to similar winner nodes, here we have the general principle for identifying the noise from the data: if most of the winner nodes are centered around one particular region, then few of the remaining winner nodes indicate that they constitute the unreliable parts of the test, or "measurement noise". For example, the nodes 8 are 9 in Figure 3 are the noise and therefore can be discarded.

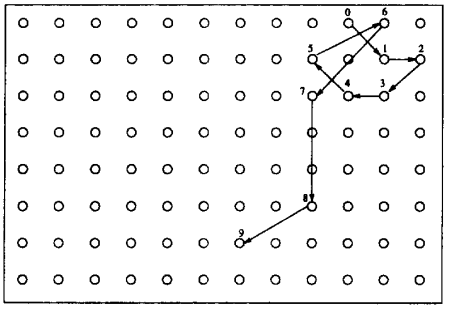

Figure 3. A transition trajectory in the output map

It should be noted that each neuron on the output map is likely to have a number of input vectors associated with it, and these input vectors in turn determines the physical meaning of the neuron such as average sensitivity, response time, etc. Using the meaning of the neurons and the transition trajectory on the map, one can explain patient's test behaviour in depth, for example, whether the noise was due to fatigue, learning effect, or other reasons. Details of this study can be found in [6].

IV. THE RESULTS

In this section, we present the experimental results in applying the proposed method to a set of clinical test data (2630 data vectors) from the Moorfields Eye Hospital, London. To find out how successful this method is in achieving its objective, we use the idea of *reproducibility* of the test results.

As glaucoma is a long term progressing disease, the visual function should remain more or less the same during a short period of time. Therefore results from such two repeated tests within this time period should be very close. However, this is not always true under real clinical situations as noise is involved in each test, perhaps for different reasons. Thus it is not surprising to note that there are a large number of repeated tests, which were conducted within an average time span of one month, whose results showed disagreements to various degrees.

As one of the main reasons for the disagreement is the noise, i.e. the unreliable portion of the test, with the two tests, it is natural to assume that the sensitivity results of the two tests should agree (to various degrees) after the noise is discarded. This then constitutes a strategy for evaluating our proposed method for eliminating noise from data.

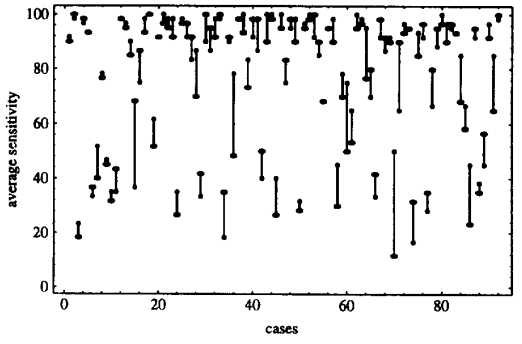

Figure 4(a). Before deletion

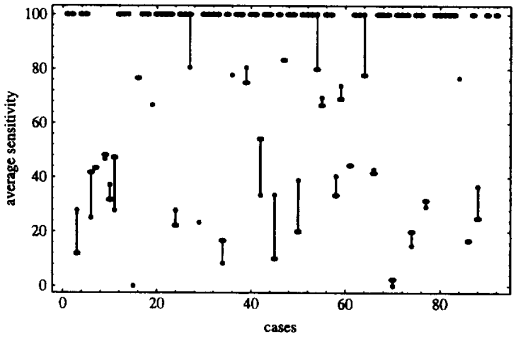

Figure 4(b). After deletion

We have chosen 91 pairs of records for this purpose. The average sensitivity values of these tests are contrasted in Figure 4(a) where the dot is used to indicate the result of the first test, the oval is used for the result of the second test, and the difference between the two results for each case is

illustrated by the line in between them. The same results after the rejection of noise by the proposed method are given in Figure 4(b).

One of the major findings is that the results from the two repeated tests have much better agreements after the noise is rejected. This is indicated by the following two measurements. First, more than 80% of the tests after the rejection of noise have reached almost total agreement (less than 1.0% error), while only less than 48% of the tests agreed in the original data set. This is reflected by the fact that there are many more cases on Figure 4(a) where the dot and oval are overlapping or very close than those on Figure 4(b). Second, if one calculates the mean difference between the two tests, 5.6 (95% Confidence Interval: 4.3 - 7.0) is the figure for the original data, while 3.4 (95% Confidence Interval: 1.8 - 4.9) is obtained after the noise is eliminated. This is indicated by the observation that the lines between the two tests are in general shortened in Figure 4(b). These findings have shown that the proposed method does provide an effective way of identifying and discarding the noisy data.

In addition, noise deletion may also be of direct diagnostic assistance to the clinician. One of the difficulties for the clinician is that the result from one test suggests that the patient is *normal* (no glaucoma), while the result from the other test shows that the patient is *abnormal* (having glaucoma of some kind). It has been found that the average sensitivity value of 75% appears to be the golden line that divides the normal and abnormal groups [10]. Since much better agreement is shown between the two repeated tests after the deletion of noise, there would be fewer cases whose test results are split by the golden line. This is indeed the case with our data as shown in Figure 4: there are quite a few conflicting cases in Figure 4(a), while only one or two such cases exist in Figure 4(b).

V. CONCLUDING REMARKS

In this paper we introduce an alternative way of dealing with noisy data. Instead of measuring and providing information on the amount of noise in the data, we try to explicitly identify and then discard the noise so that quality data can be used for decision making. To find out how successful this method is in achieving its objective, we have also suggested an evaluation strategy using the concept of reproducibility of the test results.

The proposed method has been used to model patient's behaviour during the psychophysical test and to show when and where the noise occurred in the test data. The results from this study suggest that the explicit treatment of noise in the data using self-organising maps presents a promising approach to identifying and eliminating measurement noise in these tests.

ACKNOWLEDGEMENTS

This work is supported by the International Glaucoma Society, British Council for Prevention of Blindness, and a research grant from Birkbeck College (No. 92/30). We would also like to thank our colleagues at Birkbeck College and Institute of Ophthalmology for helpful discussions.

REFERENCES

[1]. Bauer, H U and Pawelzik, K R, "Quantifying the Neighbourhood Preservation of Self-Organising Feature Maps," *IEEE Trans. on Neural Networks*, vol. 3, pp. 570-9, 1992.

[2]. Erwin, E, Obermayer, K, and Schulten, K, "Self-Organising Maps: Stationary States, Metastability and Convergence Rate," *Biological Cybernetics*, vol. 67, pp. 35-45, 1992.

[3]. Fitzke, F W, Poinoosawmy, D, Nagasuberamanian, S, and Hitchings, R A, "Peripheral Displacement Threshold in Glaucoma and Ocular Hypertension," *Perimetry Update 1988/89*, pp. 399-405, 1989.

[4]. Hitchings, R, "Psychophysical Testing in Glaucoma," *BJO*, vol. 77, pp. 471-72, 1993.

[5]. Kohonen, T, *Self-Organisation and Associative Memory*, Springer-Verlag, 1989.

[6]. Liu, X, Cheng, G, and Wu, J X, "Identifying the Measurement Noise in Glaucomatous Testing: an Artificial Neural Network Approach," *Technical Report, No BBKCS-9312*, Birkbeck College, University of London, 1993.

[7]. Lo, Z-P and Bavarian, B, "On the Rate of Convergence in Topology Preserving Neural Networks," *Biological Cybernetics*, vol. 65, pp. 55-63, 1991.

[8]. Mills, R P etc, "Estimates of Variance in Visual Field Data," *Perimetry Update 1990/91*, pp. 93-101, 1991.

[9]. Ritter, H and Schulten, K, "On the Stationary State of Kohonen's Self-Organising Sensory Mapping," *Biological Cybernetics*, vol. 54, pp. 99-106, 1986.

[10]. Wu, J X, "Visual Screening for Blinding Diseases in the Community Using Computer Controlled Video Perimetry," *Ph.D Thesis, University of London*, 1993.