DIAGNOSIS AND DECISION-MAKING FOR AWARENESS DURING GENERAL ANAESTHESIA

J K Backory, M F Abbod and D A Linkens

The University of Sheffield, Sheffield

We describe the design process of a diagnostic system for monitoring the anaesthetic state of patients during surgical interventions under general anaesthesia. This subject has received much attention from the research community especially since the advent of balanced anaesthesia, which has made the detection of awareness much more difficult. Mid-Latency Auditory Evoked Potentials (MLAEPs) obtained during general anaesthesia are used to design a neuro-fuzzy system for the determination of the level of unconsciousness after feature extraction using multiresolution wavelet analysis (MRWA).

The neuro-fuzzy system proves to be a useful tool in eliciting knowledge for the fuzzy system: the anaesthetist's expertise is indirectly coded in the knowledge rule-base through the learning process with the training data. The anaesthetic depth of the patient, as deduced by the anaesthetist from the clinical signs and other haemodynamic variables, noted down during surgery, is subsequently used to label the MLAEP data accordingly. This anaesthetist-labelled data, used to train the neuro-fuzzy system, is able to produce a classifier that successfully interprets unseen data recorded from other patients. This system is not limited, however, to the combination of drugs used here. Indeed, the similar effects of inhalational and analgesic anaesthetic drugs on the MLAEPs demonstrate that the system could potentially be used for any anaesthetic and analgesic drug combination.

We also suggest the use of a closed-loop architecture that would automatically provide the drug profile necessary to maintain the patient at a safe level of sedation.

Keywords: neuro-fuzzy, anaesthesia, wavelets.

INTRODUCTION

Ever since the introduction of muscle relaxants into clinical anaesthesia, there has existed the possibility of not recognising an inadequately anaesthetised patient. Many of the classic signs of light anaesthesia are made unreliable or are ablated by muscle relaxants. Coupled with this, the general tendency to use balanced anaesthetic techniques using several drugs to control each of anaesthesia, analgesia and paralysis to maintain the patient at a lighter plane (for safety) of anaesthesia has increased the risk of awareness. There have been reports of incomplete general anaesthesia by patients who were pharmacologically paralysed while under general anaesthesia [1]. Anaesthetists now use autonomic responses (changes in blood pressure and heart rate, sweating and lacrimation) to determine the depth of anaesthesia (DOA). Unfortunately, these responses are also affected by other drugs such as opioids and anticholinergics, making the responses unreliable. Matters were furthermore made more complicated by the introduction of intravenous anaesthetics. Unlike inhalational anaesthetics, the relation between dose rate and blood level concentration of intravenous agents varies widely between subjects [2], making it inappropriate for diagnosing DOA.

The MLAEP, on the other hand, has been shown to produce graded changes with increasing concentration of anaesthetic drugs and they have also been shown to give the balance between the depression of the nervous system caused by the anaesthetic drugs and arousal caused by surgical stimulation [3]. There has since then been an extensive investigation on the effects of various inhalational and intravenous anaesthetic agents on the various components of the AEP. The outcome points to the fact that the MLAEP may effectively be used as an indicator of anaesthetic depth during surgery.

These encouraging results obtained with the use of MLAEP have prompted this study in which the MLAEP are used to produce a reliable indicator of DOA that may be used in the operating theatre. In various studies, the latencies of the characteristic peaks of the MLAEP are obtained after visual inspection, and these are interpreted by an expert to get the DOA. In an automated system, these features have to be automatically extracted; such a feature extraction method using a neural network as a time-series approximation was described and used in [4]. Previous studies [5,6] have also required that the baseline observations be obtained as these are used with the intra-operative observations to calculate the depth of anaesthesia. Baseline observations are often difficult to obtain and they are also highly corrupted by noise and the large EEG signal. It is now known that anaesthetic drugs at the same potency produce similar graded changes in humans; Schwender et al [7] was able to create MLAEP from individual responses from several patients under the same anaesthetic depth. It was deemed to be feasible to produce a DOA monitor that would not require baseline values. This is investigated

here through the use of MRWA to extract significant features from the MLAEP (8).

MID-LATENCY AUDITORY EVOKED POTENTIAL

The Auditory Evoked Potential (AEP) are brain responses to auditory stimuli, usually loud clicks through earphones. The AEP which lasts for about 1 second is made up of three parts: the brainstem response which is the first 10 ms of the response, the MLAEP which lasts from 10 to about 50 ms, and the Late Cortical Response which lasts from 50 to 1000 ms. The MLAEP contains three characteristic peaks (Na, Pa and Nb) which have been studied by several investigators. The changes in latencies and amplitudes of these peaks with several anaesthetic drugs (intravenous and inhalational) and with surgical stimuli have been shown to correlate well with observed anaesthetic depth as well as with signs of awareness.

The amplitudes of the AEPs are typically in the order of a few microvolts and are embedded in the ongoing EEG waveform whose amplitude is typically a few tens of microvolts. Ensemble averaging has been an effective tool for the enhancement of the AEP waveform. It is generally assumed that the noise (background EEG) will be uncorrelated with each presentation and will tend to cancel itself due to the averaging process. In this research work, averages of 200 individual responses were created. This corresponds to a recording time of about 30 s, which is an acceptable sample time for prediction of DOA. The system developed at the Northwick Park Hospital was used to record the AEPs. The stimulus rate for the clicks was set to 6.1224 Hz and after each click the first 120 ms of the response, corresponding to 121 points at a sampling rate of 1 KHz, was stored.

PATIENTS AND METHODS

Fourteen patients, ASA I or II, were studied after obtaining their informed consent. They were all premedicated with 10-20 mg temazepam. All the patients were also given fentanyl as part of the balanced anaesthetic technique. After loss of consciousness in the anaesthetic room, they were intubated. In some cases, if apnoea occurred, the patients were manually ventilated. The patient was then brought to the operating theatre. In the set of experiments a Target Control Infusion (TCI) system was used for ten of the patients, three were anaesthetised using a manually controlled infusion pump, and the last patient was anaesthetised using enflurane.

During the surgical procedure, the DOA as determined by the anaesthetist was noted at relevant stages and these were used to label the data collected. Four DOA levels, Awake/Light (AWAKE), OK Light (OKL), OK and OK Deep (OKD), were used as they were though to be the most clinically significant.

FEATURE EXTRACTION AND VALIDATION

The advantages of using a MRWA of event-related potentials (of which the AEP is one) was demonstrated by Samar (8). An analysis in the time domain is usually carried out by selecting the few peaks of interest, measuring their amplitudes and/or latencies, and discarding the rest of the waveform information. Also, this time-series analysis can easily be corrupted by noise, even when the signal has been band-pass filtered. This method loses all the information coded within the peaks, intermittent peaks, time-relations between peaks, slopes and other higher derivatives. The wavelet transform is able to retain this information.

The MRWA using Daubechies Wavelets (9) was used to decompose the signal into approximations at different scales of resolution. The Daubechies 6 Wavelet with 12 wavelet coefficients, found to give very satisfactory results, was used to carry out a decomposition of the original MLAEP signal into its detail components and the residual component. Before the AEP is analysed using the MRA, it is padded with zeroes to make the sequence length 128 samples instead of the 121; the MRA requires the length be a modulo 2 number. Since each decomposition level is also accompanied by decimation by 2, the complete decomposition process produces six detail sequences and a residual sequence, giving a total of 128 wavelet coefficients (due to the orthogonality of the transform). The six detail components are called D1, D2, D3, D4, D5 and D6 and the number of coefficients in each sequence are 64, 32, 16, 8, 4, and 2 respectively. Thus each wavelet coefficient of D1 spans a time of 2 ms, that of D2 spans a time of 4 ms, and similarly for the others. Also, D1 contains the highest frequency components and D6 the low-frequency components of the AEP signal.

Selected Detail components were used to create the features used, and the selection was based on the results obtained from a student t-test ($p<0.05$) on the detail components. The components D6 and the residual were not used since they contain mostly the residual EEG. The early and late components of details D1-D4, not part of the MLAEP, were removed, and the energy contained in the remaining samples of each detail was computed. For D5, the core of the MLAEP signal, each of the four samples (D5_1, D5_2, D5_3 and D5_4) was analysed individually. The first and last samples of D5, as expected, did not produce consistently significant difference when the data was obtained at different DOA levels, and were thus discarded. Thus, the feature vector used consisted of the six values D1-D4, D5_2 and D5_3.

The features from the first nine patients were used to construct a training/validation set, and the data from the remaining five patients were used for testing. The student *t-test* was again used to test for significant difference between the data, for all the patients, between the different DOA levels, as shown in Table 1. There is no significant difference between AWAKE and OKL for D5_2 and D5_3, since the AEPs still have high frequencies at these DOA levels.

A set of 1000 features, with 250 patterns from each of the classes was created for training the classifier.

NEURO-FUZZY CLASSIFIER

The primary bottleneck of knowledge-based systems is the tedious rule-base acquisition. It is the determination of these fuzzy rules from the data that plays an important role in the performance of the fuzzy system. While for some systems, the knowledge obtained from the expert is sufficient to create a fuzzy system for classification, this is not true for all. Still, for many systems where the expert knowledge is available, this is usually dependent on one or very few experts' opinions. In the cases where enough data are available, it may prove to be better to directly implement these fuzzy rules from the data through a learning process. The membership functions determined from the experts are also rarely optimal in terms of carrying out the classification process as required: they differ from person to person as well as from time to time.

Neural networks, where the network weights can be used to represent knowledge, can be substituted for fuzzy systems in the event that the rule base acquisition proves to be difficult. The adaptivity of the neural network's weights means that knowledge may be imparted to the system from input/output data. However, this learning in neural networks would produce a blackbox model which is not transparent, unlike fuzzy systems. Furthermore, *a priori* knowledge cannot be used to improve the neural network's performance. In the last decade, researchers have been showing a large interest in the combination of both techniques and the term neuro-fuzzy system has often been coined in the literature (10).

When input/output data are available, it might prove sufficient or indeed beneficial to implement the fuzzy rules from these data. In our case, the data obtained during surgery was labelled and could thus be used to train the neuro-fuzzy system. The expert knowledge is acquired during data collection through interaction with the anaesthetist. This expert knowledge is then implemented in the fuzzy classifier by using the labelled data set to optimise the rules that were obtained. The neuro-fuzzy system used in this paper is the one reported in Bersini and Bontempi (11). This trainable Fuzzy Inference System (FIS) is based on the Takagi and Sugeno approach (12) and uses gaussian membership functions. The centres of the rules are initialised using a fuzzy clustering algorithm (13). The centres and widths of the membership functions are optimised using the gradient-based Levenberg-Marquardt algorithm (14) and the outputs are found using the pseudo-inverse method. The system uses the 10-fold cross-validation method: train on 9 subsets and test on the remaining subset. The FIS architecture is shown in Figure 1.

An FIS network of 16 rules was found to give the best compromise between network complexity and performance. Although performance could be increased using a more complex network, this was not significantly so, and thus the less complex network of 16 rules was used.

DECISION MAKING

Decision-making during general anaesthesia involves determining the amount of anaesthetic drug required to maintain a stable level of DOA. When the TCI system is used for infusion, the decision-making unit determines the concentration of drug in plasma that the TCI should aim for. We have developed a simulated closed-loop infusion system where the target concentration to be aimed from by the TCI system is determined by the DOA level.

The effects of the anaesthetic and analgesic drugs on the patients are modelled using fuzzy logic patient models. The anaesthetic patient model produces a set of features based on the concentration of anaesthetic in plasma, the level of stimulation by surgery and the patient sensitivity. The level of stimulation itself is derived from the analgesic patient model which characterises the effect of the analgesic drug on the level of pain perceived by the patient. An intelligent controller then determines the correct target concentration based on the concentration level trend and the DOA level trend.

RESULTS

The FIS network of 16 rules gave a correct classification of 78.9 % on the training/validation data. The number of incorrect classifications in each of the four DOA levels were: AWAKE 44, OKL 69, OK 76 and OKD 22. The number of incorrect classifications of OK and OKL are higher because, clinically, there is little difference between these two levels. Figure 2 below shows the results obtained for patient 1 who underwent varicose veins surgery and who was also paralysed. The period of actual OKD observed was because the surgical interference during that period was not high. The classifier correctly identifies the period

of OKD and OK during the course of this surgical procedure. The short periods of OKD observed 55 to 75 minutes after start of surgery when the TCI value was 5000 ng ml^{-1}, is acceptable because there was only intermittent surgical interference then. At the end of surgery, the DOA was still OK because recording was stopped before the patient was awake.

CONCLUSION

This research work has demonstrated that the use of MLAEP for determining the depth of anaesthesia is possible. Wavelet features derived from the MLAEP signal recorded during surgery can be used to diagnose the DOA level during the surgical procedure. The student *t-test* showed that there is a significant difference between the wavelet features recorded at different levels of DOA. We were also able to show that there was little patient inter-variability as there was a significant difference between the features from 9 patients. The neuro-fuzzy classifier successfully extracted rules from the training data and was successfully able to classy features from new patients. We also suggest that the use of a pseudo closed-loop system, where the anaesthetist maintains overall control, could be used to advise the anaesthetist on the amount of anaesthetic to be used.

REFERENCES

1. Tracy J, 1993, "Awareness in the operating room: a patient's view", Memory and Awareness in Anesthesia, P.S. Sebel, B. B. Bonke, and E. Winograd, eds., Prentice Hall, New Jersey, 349-353

2. Thornton C, Heneghan C, Navaratnarajah M, Bateman P and Jones J, 1985, "Effect of Etomidate on the auditory evoked response in man," BJA, 57, 554-561

3. Thornton C and Newton D E F, 1989, "The auditory evoked response: a measure of depth of anaesthesia", Baillière's Clin. Anaesthesiology, 3, 559-585

4. Linkens D A, Abbod M F and Backory J, "Fuzzy logic control of depth of anaesthesia using auditory evoked responses", IEE Colloquium, Fuzzy Logic Controllers in Practice, London, 4/1-4/6

5. Linkens D A, Elkfafi M and Peacock J E, 1996, "Intelligent processing of evoked potentials for monitoring depth of anaesthesia," 16th International Symposium on Computing in Anaesthesia and Intensive Care, Rotterdam, The Netherlands

6. Linkens D A, Shieh J S, and Peacock J E, 1994, "Machine-learning rule-based fuzzy logic control for depth of anaesthesia," Proc. of the IEE Int. Conf. on Control '94, Coventry, 31-36

7. Schwender D, Golling W, Klasing S, Faber-Züllig E, Pöppel E and Peter K, 1994, "Effects of surgical stimulation on midlatency auditory evoked potentials during general anaesthesia with propofol/fentanyl, isoflurane/fentanyl and flunitrazepam/fentanyl", Anaesthesia, 49, 572-578

8. Samar V J, Swartz K P, and Raghuveer M R, 1995, "Multiresolution analysis of event-related potentials by wavelet decomposition", Brain and Cognition, 27, 398-438

9. Daubechies I, 1988, "Orthonormal bases of compactly supported wavelets", Comm. on Pure and Appl. Math., 41, 909-996

10. Isermann R, 1997, "Special Issue: Application of neuro-fuzzy systems - Preface", Fuzzy Sets and Syst., 89, 275

11. Bersini H and Bontempi G, 1997, "Now comes the time to defuzzify neuro-fuzzy models", Fuzzy Sets and Syst., 90, 161-169

12. Takagi T and Sugeno M, 1985, "Fuzzy identification of systems and its applications to modeling and control", IEEE Transactions on Systems, Man and Cybernetics, 15, 116-132

13. Bezdek J C and Adderson I M, 1985, "An application of the c-varieties clustering algorithms to polygonal curve fitting", IEEE Trans. on Syst. Man and Cybernetics, 15, 637-641.

14. Bishop C M, 1995, "Neural Networks for Pattern Recognition", Oxford University Press, New York.

Table 1 -   Results of carrying out a student *t-test* test of significance between different DOA levels.

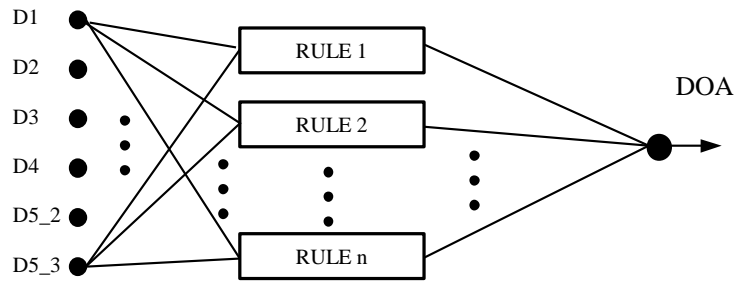|	D5_2	D5_3	D4	D3	D2	D1
AWAKE-OKL	0.248	0.129	0.000	0.000	0.000	0.000
OKL-OK	0.000	0.000	0.000	0.000	0.000	0.000
OK-OKD	0.000	0.000	0.000	0.000	0.002	0.000

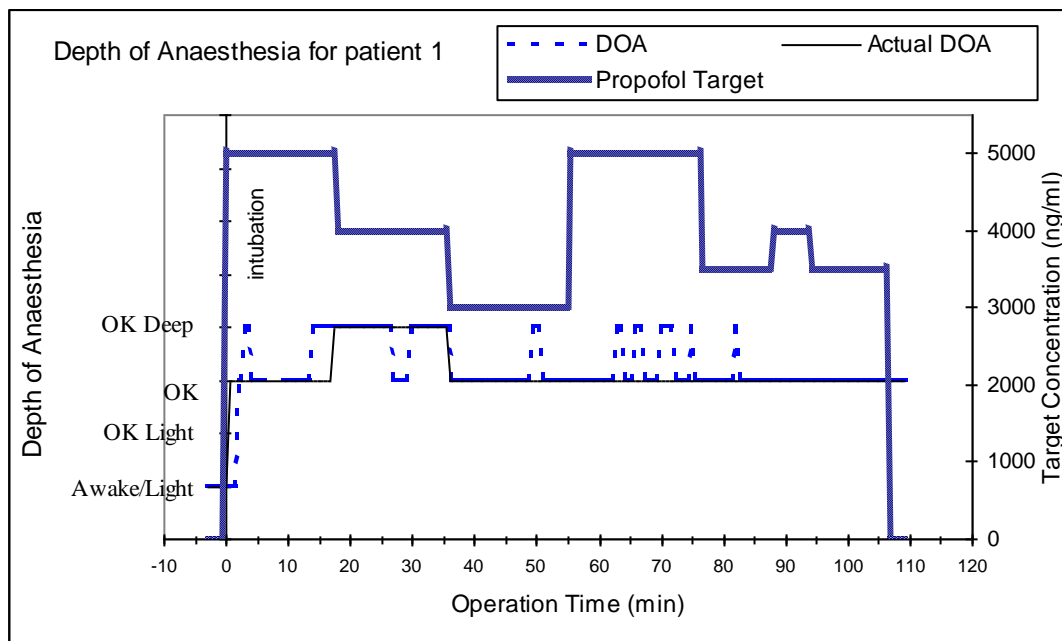Figure 1:   The neuro-fuzzy inference system architecture

Figure 2:   Results of classifying MLAEP features from patient Jay15 using the neuro-fuzzy classifier