Monitoring Depth of Hypnosis under Propofol General Anaesthesia
Granger Causality and Hidden Markov Models

Nicoletta Nicolaou[1,2,3] and Julius Georgiou[1,2,3]
[1]*Holistic Electronics Research Lab, University of Cyprus, Nicosia, Cyprus*
[2]*Dept. of Electrical and Computer Engineering, University of Cyprus, Nicosia, Cyprus*
[3]*KIOS Research Centre, University of Cyprus, Nicosia, Cyprus*

Keywords: Brain-Computer Interface, EEG, Anaesthesia, Awareness.

Abstract: Intra-operative awareness is experienced when a patient regains consciousness during surgery. This work presents a Brain-Computer Interface system that can be used as part of routine surgery for monitoring the patient state of hypnosis in order to prevent intra-operative awareness. The underlying state of hypnosis is estimated using causality-based features extracted from the spontaneous electrical brain activity (EEG) of the patient and a probabilistic classification framework (Hidden Markov Models). The proposed method is applied to EEG activity from 20 patients under propofol anaesthesia. The mean discrimination performance obtained was 98% and 85% for wakefulness and anaesthesia respectively, with an overall performance accuracy of 92%. The use of a probabilistic framework increases the anaesthetist's confidence on the estimated state of hypnosis based on the marginal probabilities of the underlying state.

1 INTRODUCTION

Intra-operative awareness occurs in approximately 0.1-0.8 % of surgical patients (Bruhn et al., 2006). The real incidence of awareness, however, is likely to be much higher due to the amnesic effect of certain anaesthetics resulting in some patients having no recollection of regaining awareness. The importance of monitoring depth of anaesthesia in order to prevent intra-operative awareness is apparent considering that, given the estimated 234.2 million major surgical procedures undertaken annually worldwide (Weiser et al., 2008), at least 1,873,600 people are likely to have experienced intraoperative awareness.

The use of the electrical brain activity (EEG) is currently the preferred method for monitoring anaesthetic depth. Devices that monitor the patient state of hypnosis during anaesthesia are a form of a Brain-Computer Interface (BCI) system. Even though the idea of using the spontaneous patient EEG for monitoring purposes is not new (McEwen et al., 1975), a relatively small number of devices are currently commercially available for this purpose. In these devices, the patient's spontaneous EEG activity is continuously monitored for changes that signal regaining of awareness. In incidences of intra-operative awareness the patient cannot communicate this to the anaesthetist due to the co-administration of neuromuscular blockers with the anaesthetic agents. Thus, the patient is essentially put in a chemically-induced 'locked-in' state whereby communication via conventional means is impaired. Despite the existence of commercial EEG-based devices for monitoring patient state during anaesthesia, they are not considered as part of routine anaesthetic practice in the majority of hospitals worldwide. This is mainly attributed to issues relating to robustness and inter-subject variability (Voss and Sleigh, 2007). Inter-subject variability in particular is an issue that cannot be addressed with current systems, as the systems have a universal state indicator scale (from 0-100; 100: fully conscious, 40-60: surgical anaesthesia, 0: no activity) that cannot be calibrated for each individual patient and, thus, does not take into account patient specifics.

In this study the anaesthetic-induced EEG changes are modelled using causality-based features and the underlying state (wakefulness / anaesthesia) is estimated using a probabilistic framework (Hidden Markov Models). The use of causality features increases the system robustness, as it has been shown that such features capture general mechanisms of anaesthetic administration regardless of the particular anaesthetic protocol (Nicolaou et al., 2012; Barrett et al., 2012), and with a high discriminative ability

Nicolaou N. and Georgiou J..
Monitoring Depth of Hypnosis under Propofol General Anaesthesia - Granger Causality and Hidden Markov Models.
DOI: 10.5220/0004679402560261
In *Proceedings of the International Congress on Neurotechnology, Electronics and Informatics* (BrainRehab-2013), pages 256-261
ISBN: 978-989-8565-80-8

(Nicolaou and Georgiou, 2013). The use of a probabilistic framework in such a system is advantageous, as the anaesthetist can also assess the reliability of a state estimation through the marginal state probabilities. Learning an HMM model for each patient also ensures that the system is calibrated for each patient and the estimated states are not affected by intersubject variability.

2 METHODS

2.1 Dataset

The data is a subset of EEG data collected from patients during surgery at the Nicosia General Hospital, Cyprus. In this particular study data from 20 male patients (mean age 41.8 ± 20.6) were analysed. The study has been approved by the Cyprus National Bioethics Committee and patients involved gave written informed consent for their participation. The experimental protocol details are described elsewhere (e.g. see (Nicolaou et al., 2012)). In summary, anaesthesia was induced with a propofol bolus and maintained with constant intravenous propofol administration. In most patients this was titrated with an intravenous administration of remifentanil hydrochloride. Lungs were ventilated with an air-oxygen or air-oxygen-N_2O mixture. During surgery boluses of neuromuscular blocking agents and other drugs, such as antibiotics, were administered as required. EEG data were obtained from 19 electrodes based on the international 10/20 system, with a sampling rate of 256 Hz. Data recording was performed throughout the entire surgical duration (awake pre-induction, induction, surgical anaesthesia and recovery of consciousness). Since the exact point at which loss of consciousness occurs after patient induction is not known, the point at which the anaesthetic bolus was administered was considered instead. Recovery of consciousness was defined as the point at which the patient responded (either via voluntary muscular movement or a verbal response) to verbal commands or tactile stimuli by the anaesthetist.

2.2 Granger Causality

Granger Causality (GC) is defined by Wiener as follows: ‘for two simultaneously measured signals, if one can predict the first signal better by incorporating the past information from the second signal than using only information from the first one, then the second signal can be called causal to the first one’ (Wiener, 1956). Mathematically, causality was defined by Granger through the use of regression: for two time series, X_1, and X_2, if X_1 is influenced by X_2, then the addition of past values of X_2 in the regression of X_1 will improve its prediction (Granger, 1980). The performance of the prediction can be assessed through the variances of the fitted regression models. Thus, GC is defined as:

$$GC_{X_2 \to X_1} = ln \frac{\sigma^2_{X_1/X_{1\_}}}{\sigma^2_{X_1/X_1X_2}} \tag{1}$$

where $\sigma^2_{X_1/X_{1\_}}$ and $\sigma^2_{X_1/X_1X_2}$ are the variances of the regression errors, e_{x_j} and $e_{x_1x_2}$ respectively, obtained from the following (auto)regression models ($j = 1, 2$):

$$x_j(t) = \sum_{i=1}^{P} a_{ix_j} x_j(t-i) + e_{x_j}(t) \tag{2}$$

$$x_1(t) = \sum_{i=1}^{P} a_{ix_1x_2} x_1(t-i) + \sum_{i=1}^{P} b_{ix_1x_2} x_2(t-i) + e_{x_1x_2}(t) \tag{3}$$

The variables a_{ix_j}, $a_{ix_1x_2}$ and $b_{ix_1x_2}$ are the regression coefficients for models of order P. The main idea is that if the prediction of X_1 is improved by using past values of X_2 in its prediction, then $\sigma^2_{X_1/X_1X_2} < \sigma^2_{X_1/X_{1\_}}$ and, therefore $GC_{X_2 \to X_1}$ increases. If, however, the past of X_2 does not improve the prediction of X_1, then $GC_{X_2 \to X_1}$ will be close to zero. Therefore, by definition, $GC_{X_2 \to X_1} = 0$ when there is no causality between the signals, and $GC_{X_2 \to X_1} > 0$ otherwise. Similarly, we can also define GC in the opposite direction. A number of factors must be taken into account when considering GC-based analysis, such as the choice of the regression model order, data stationarity and data filtering (more details can be found in (Bressler and Seth, 2010)).

2.3 Hidden Markov Models

Hidden Markov Models (HMMs) belong to the family of Bayesian networks (Rabiner, 1986). They are used to model systems that are assumed to be a Markov process with states that are not directly visible to the observer. The observer can observe only the output of the system. However, the sequence of outputs gives some information about the invisible sequence of discrete states, as each state has a probability distribution associated with it. Transitions between states are also probabilistically described through the transition probability matrix. The HMM model makes two assumptions: (i) the Markov assumption, which states that the current state is dependent only on the previous state; and (ii) the independence assumption, which states that the output observed at a particular time is independent of past outputs and states and depends only on the current state. An HMM can also perform learning, whereby the model parameters that best describe a process can be estimated from a set of examples from the particular process. Training can be both supervised or unsupervised. In supervised learning, the model outputs are equated to the inputs, and the outputs to the corresponding states. The model parameters can then be estimated using maximum likelihood estimation. In our particular case, the output observations correspond to the GC-based features, while the unobserved underlying states are 'wakefulness' and 'anaesthesia'. The HMM is then used to answer the question: "What is the most likely sequence of wakefulness/anaesthesia states that could have generated the observed GC values?".

2.4 Methodology

The ability to separate consciousness and anaesthesia using the spontaneous changes in the patient's EEG activity was investigated. The particular methodology consists of the following steps:

1. For each patient, the 19 electrodes were split into the following five grids: left frontal (LF: electrodes Fp1, F7, F3, T3, C3), right frontal (RF: Fp2, F8, F4, C4, T4), left posterior (LP: T5, P3, O1), right posterior (RP: T6, P4, O2), and midline (Z: Fz, Cz, Pz). The average EEG activity over each of the five grids was then estimated. The particular groupings were chosen such that broad areas corresponding to frontal and posterior activity were obtained, as fronto-posterior interactions appear to play an important role in (un)consciousness.
2. For each patient, the continuous averaged EEG data obtained over each of the five grids were windowed into 2-s non-overlapping segments.
3. Using the manual markers in the EEG record, the windows corresponding to wakefulness (class A) and anaesthesia (class B) were identified.
4. For each 2-s segment, pairwise fronto-posterior GC features were estimated, resulting into the following 4-dimensional feature vector:
   $F_C^i = [GC^i_{LF\rightarrow LP}, GC^i_{RF\rightarrow LP}, GC^i_{LF\rightarrow RP}, GC^i_{RF\rightarrow RP}]$ where $C \in \{A,B\}$ corresponds to one of the two classes, and $i = 1,...,N_C$ denotes the i^{th} 2-s segment from all the available segments of each class (N_C). The order of the fitted regression models was set to 6.
5. The 'Bayes Net Toolbox' for Matlab is used for HMM modelling (Murphy, 2001). An HMM model is trained using 40% of the available data for each class. The HMM outputs are modelled as continuous Gaussians. The estimated prediction probabilities are then obtained for all available data.

The specific parameters, such as duration of the windows and the order of the fitted regression models were based on previous investigations (Nicolaou et al., 2012). The performance of the HMM for each patient is then estimated via the average specificity (4), sensitivity (5) and accuracy (6) estimated over $B = 50$ bootstrap repetitions. $T_{ru}P(T_{ru}N)$ is the number of true positives (negatives), and $T_{ot}P(T_{ot}N)$ is the total number of 'ground truth' positive (negative) examples of each class.

$$SP = \frac{T_{ru}P}{T_{ot}P} \tag{4}$$

$$SE = \frac{T_{ru}N}{T_{ot}N} \tag{5}$$

$$AC = \frac{1}{2}\left(\frac{1}{B}\sum_{b=1}^{B} SP_b + \frac{1}{B}\sum_{b=1}^{B} SE_b\right) \tag{6}$$

3 RESULTS AND DISCUSSION

Figures 1 and 2 show examples of the marginal probabilities for wakefulness and anaesthesia obtained from one of the 50 bootstrap repetitions and for two randomly chosen patients (S20 and S2 respectively). For visualisation purposes a moving average filter ($n = 10$) was applied to the marginal probabilities. The probabilistic framework allows the anaesthetist to associate a likelihood to a particular decision. It can be seen from the figures that the marginal probabilities track the transitions between wakefulness and anaesthesia well: as expected, the marginal probability for wakefulness is close to 1 prior to anaesthetic induction and after recovery of consciousness, while it drops below 0.5 during surgical anaesthesia (and *vice versa* for the marginal probability for anaesthesia). This implies that the Granger Causality-based features provide a good representation of the two states and that the HMM is able to track the successions of the two states successfully. The presence of artefacts, such as the use of diathermy, may cause brief distortion in state estimations. This can be seen in figure 1, where artefacts caused by diathermy introduce spurious brief changes in the estimated marginal probabilities. In such cases, it is easy for the anaesthetist

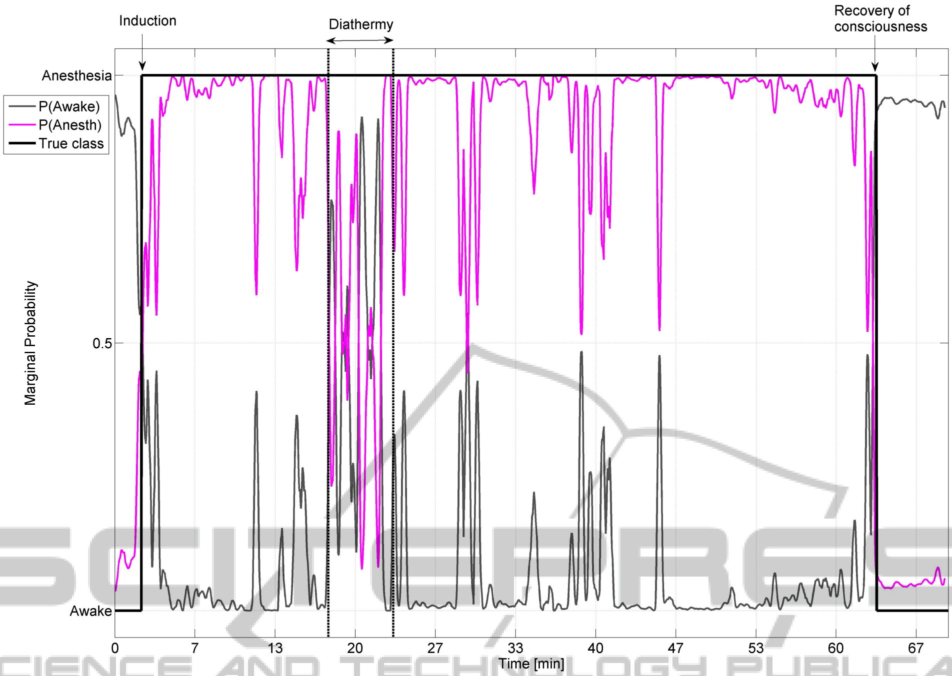

Figure 1: Estimated state probabilities for wakefulness (black) and anaesthesia (maroon) for patient S20. The state with a probability greater than '0.5' is the classified state. Large peaks in the marginal probabilities during anaesthesia are caused by diathermy artefacts, while the peaks after induction are caused by tracheal intubation.

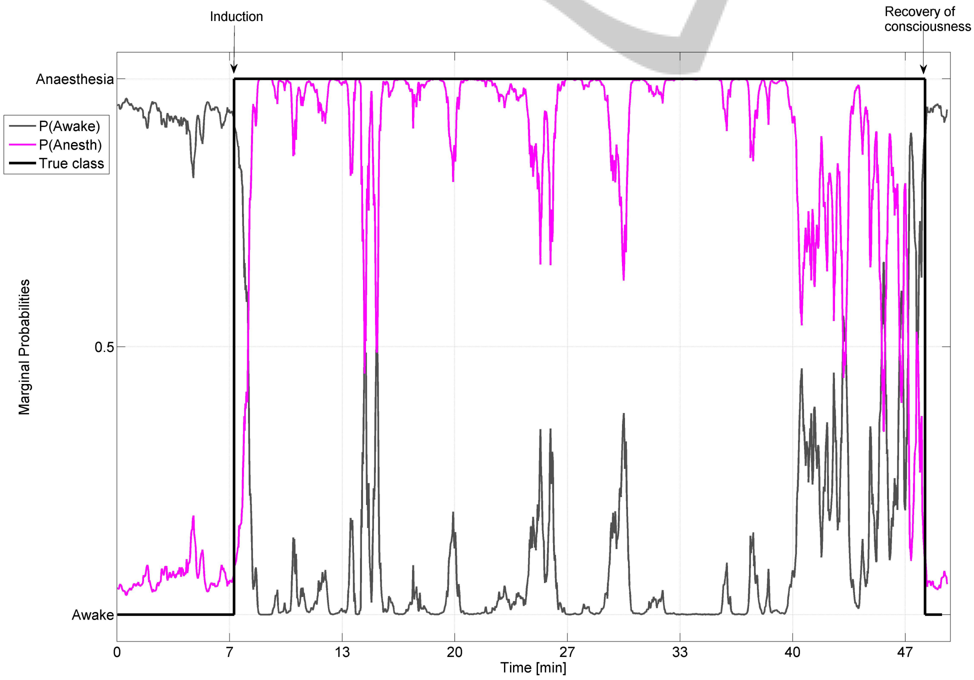

Figure 2: Estimated state probabilities for wakefulness (black) and anaesthesia (maroon) for patient S2. The state with a probability greater than '0.5' is the classified state. Large peaks in the marginal probabilities during anaesthesia are caused by diathermy artefacts.

to assess whether such changes are true or whether they are indeed spurious, e.g. if diathermy is being utilised at the particular moment. However, the presence of artefacts does not necessarily induce spurious state changes, as the marginal probabilities may not go over/under the 0.5 threshold; this can also be seen in figures 1 and 2. Moreover, changes in the marginal probabilities that are persistent rather than transient is another cause of alert for the anaesthetist, as such changes could be another indication that the patient may be regaining consciousness.

Table 1 shows the mean performance of the proposed method for each patient and overall. Patient-wise, a mean sensitivity, specificity and accuracy of 0.98, 0.85 and 0.92 is obtained. This is comparable to results from other studies, which range from 64-93% (for more details see (Nicolaou et al., 2012) and references within). For 12 patients the overall accuracy is more than 90%. The lower specificity (performance for anaesthesia) compared to sensitivity (performance for wakefulness) is expected if we take into account that no artefact removal has been performed, thus some misclassification due to artefacts during surgery is expected. For 3 patients the mean specificity is between 0.63-0.66. This could be mainly attributed to the small number of training features resulting from the small number of available features for wakefulness for the particular patients. Thus, this has a negative effect on the generalisation ability of the HMM classifier for the anaesthesia state, given that the small number of training features cannot be expected to capture all feature attributes associated with anaesthesia.

Table 1: Mean performance of GC-based HMM classification for 20 patients (S1,...S20). SE: Sensitivity, SP: Specificity, AC: Accuracy. The maximum (best) classification is indicated by '1' (corresponding to 100%).

PATIENT	SE	SP	AC
S1	1.00	0.98	0.99
S2	1.00	0.94	0.97
S3	0.89	0.93	0.91
S4	1.00	0.93	0.97
S5	1.00	0.93	0.97
S6	1.00	0.66	0.83
S7	1.00	0.96	0.98
S8	1.00	0.87	0.94
S9	1.00	0.81	0.91
S10	0.95	0.88	0.92
S11	0.95	0.79	0.87
S12	0.94	0.65	0.80
S13	0.98	0.94	0.96
S14	0.95	0.77	0.86
S15	1.00	0.63	0.82
S16	0.96	0.78	0.87
S17	1.00	0.90	0.95
S18	1.00	0.78	0.89
S19	0.91	0.87	0.89
S20	0.99	0.92	0.96
TOTAL	0.98	0.85	0.92

An important advantage of the proposed methodology is its clinical applicability. Individual patient variability is taken into account through calibration of the BCI system for each patient, as opposed to commercially available systems that employ a universal 0-100 scale without system calibration to patient specifics. This calibration is also possible due to the ability of HMMs to learn incrementally, thus facilitating a more higher-level learning through the addition of new information. The probabilistic framework adds credibility by associating a probability likelihood to each decision. This strengthens the anaesthetist's decisions regarding the assessment of the underlying patient state of hypnosis by allowing a given degree of certainty to their actions.

The proposed methodology has some limitations. These are mainly related to the particular features utilised: Granger Causality (GC) has been the recipient of some criticism regarding the effects of stationarity, volume conduction, filtering and regression model order to the estimated GC values (Bressler and Seth, 2010; Florin et al., 2010). However, the nature of causality itself implies that causality-based measures likely capture more general mechanisms of anaesthetic action. If these specific issues are taken into account causality can, thus, constitute a robust and reliable feature of anaesthetic-induced changes in spontaneous brain activity. A more extensive discussion on GC, general limitations posed by GC and potential solutions can be found in (Nicolaou et al., 2012).

In addition, as previously mentioned, no artefact removal was performed in this study. Prior to anaesthetic induction the EEG signals may contain artefacts originating both from the patient, such as body movement, eye blinks and speech, and external sources, such as placement of monitoring equipment by the hospital staff. During induction the EEG is usually contaminated by artefacts due to tracheal intubation. During anaesthesia the administered anaesthetics and muscle relaxants ensure that there is no muscle, eye or movement activity from the patient. Artefacts during anaesthesia are mainly caused by the use of diathermy equipment, but some artefacts due to other surgical stimuli could also be present. The latter is not likely as we excluded surgical procedures performed in close proximity to the EEG sensors, such as ear-nose-throat surgeries. Despite having performed no artefact removal, it is not likely that discrimination between wakefulness and anaesthesia is performed solely due to the presence of these artefacts. On the contrary, the performance of the proposed method could be improved by removal of artefacts prior to feature estimation, as the differences between the two states would be more prominent.

4 CONCLUSIONS

The spontaneous EEG activity of anaesthetised patients can be used to assess their underlying state of hypnosis. This type of a BCI system can revolutionise routine surgery and aid towards avoidance of intra-operative awareness. The use of causality-based features in a probabilistic framework adds to the reliability and robustness of such a system. State decisions are supported by related probabilities, thus strengthening the weight of each individual state assessment by both the system and the anaesthetist. Future work will investigate whether causality features obtained from individual electrode pairs can provide similar discriminatory ability in order to eliminate the need for a full set of EEG sensors to capture anaesthetic-induced changes in causality.

ACKNOWLEDGEMENTS

The authors would like to thank Dr. Pandelitsa Alexandrou and Dr. Saverios Hourris (Depr. of Anaesthesia, Nicosia General Hospital, Cyprus), the hospital staff and the anonymous volunteers who participated in this study. This work was co-funded by the Republic of Cyprus and the European Regional Development Fund through the Cyprus Research Promotion Foundation (DESMI 2008). Grants: 'New Infrastructure Project/Strategic/0308/26' and 'DIDAKTOR/DISEK /0308/20'.

REFERENCES

Barrett, A., Murphy, M., Bruno, M.-A., Noirhomme, Q., Boly, M., Laureys, S., and Seth, A. (2012). Granger causality analysis of steady-state electroencephalographic signals during propofol-induced anaesthesia. *PLoS ONE*, 7(1):e29072.

Bressler, S. and Seth, A. (2010). Wienergranger causality: A well established methodology. *NeuroImage*, 58(2):323–329.

Bruhn, J., Myles, P., Sneyd, R., and Struys, M. (2006). Depth of anaesthesia monitoring: what's available, what's validated and what's next? *British Journal of Anaesthesia*, 97:85–94.

Florin, E., Gross, J., Pfeifer, J., Fink, G., and Timmermann, L. (2010). The effect of filtering on granger causality based multivariate causality measures. *NeuroImage*, 50:577588.

Granger, C. (1980). Testing for causality: a personal viewpoint. *J Econ Dyn Control*, 2:329–352.

McEwen, J., Anderson, G., Low, M., and Jenkins, L. (1975). Monitoring the level of anesthesia by automatic analysis of spontaneous eeg activity. *IEEE Trans on Biomedical Engineering*, BME-22(4):299–305.

Murphy, K. (2001). The Bayes Net Toolbox for Matlab. *Computing Science and Statistics: Proceedings of the Interface*, 33.

Nicolaou, N. and Georgiou, J. (2013). Neural networkbased classification of anesthesia/awareness using granger causality features. *Clinical EEG and Neuroscience*, doi:10.1177/1550059413486271.

Nicolaou, N., Hourris, S., Alexandrou, P., and Georgiou, J. (2012). EEG-based automatic classification of 'awake' versus 'anesthetized' state in general anesthesia using granger causality. *PLoS ONE*, 7(3):e33869.

Rabiner, L. (1986). An introduction to hidden Markov models. *IEEE ASSP Magazine*, 3(1):4–16.

Voss, L. and Sleigh, J. (2007). Monitoring consciousness: the current status of EEG-based depth of anaesthesia monitors. *Best Pract Res Clin Anaesthesiol*, 21:313–325.

Weiser, T., Regenbogen, S., Thompson, K., Haynes, A., Lipsitz, S., Berry, W., and Gawande, A. (2008). An estimation of the global volume of surgery: a modelling strategy based on available data. *The Lancet*, 372:139–144.

Wiener, N. (1956). *The theory of prediction*, volume 1. McGraw-Hill, New York.