Analysis of EEG to Quantify Depth of Anesthesia Using Hidden Markov Model

Junbeom Kim, Huh Hyub, Seung Zhoo Yoon, Ho-Jin Choi, Kwang Moo Kim, Sang-Hyun Park*

***Abstract*— Real-time quantification of the patient's consciousness level during anesthesia is an important issue to avoid intraoperative awareness and post-operative side effects. A depth-of-anesthesia (DoA) monitoring method called Bispectral Index (BIS) is generally used for this purpose. However, BIS is known to be inaccurate at the transitory state, and also shows a critical time delay in quantifying the patient's consciousness level.**

This paper introduces a novel method to reduce the response time in the quantification process. This thesis develops a new index called HDoA by analyzing EEG using Hidden Markov Model. The proposed approach is composed by two steps, training and testing. In the training step, two HMM, awakened and anesthetized model are learned based on each training set. In the testing step, by evaluating the probability of producing the testing EEG from two models respectively, the index HDoA is derived. Since the evaluation of DoA using HMM is training based method, it have better performance with more training process.

Experiments show that HDoA has a high correlation with BIS at a steady state, and outperforms BIS in two ways: (1) shorter delay time in transition state, and (2) higher Fisher Score. The validity of HDoA has been tested by 8 real clinical data.

I. Introduction

An accurate index which can predict the depth of anesthesia (DoA) by extracting meaningful in-formation from various biological signals would help anesthesiologists precisely recognize the patient`s state. In particular, many basic research results have shown that the pattern transition of an EEG reading during surgery is highly associated with the DoA [1, 2]. Because the level of consciousness has to reflect the state of the brain activity theoretically, the DoA can be expected to present the degree of brain activity [3].

BIS is regarded as an accurate method which provides a steady state. Despite this, BIS has five deficiencies [4]. First, it is slow in detecting the exact moment of loss of consciousness (LoC). Second, BIS values become unstable when the signal quality is poor. Third, BIS is not restored back to the baseline value after awakening (the baseline value refers to the DoA index before anesthesia). Fourth, BIS values are occasionally inaccurate when a general inhalation anesthesia is not given. Finally, BIS values vary depending on the anesthetic drug.

Among the shortcomings of BIS, the most remarkable drawback is the slow response during the transition state. BIS cannot provide appropriate information about the patient`s state when the state of the patient changes from an awakening state to an anesthetic state [5]. For this reason, BIS provides incorrect index values to the anesthesiologist. This can have deleterious effects on the patient`s health, as the anesthetic agent could be overdosed.

In this study, a new algorithm will be introduced and tested with 8 instances of clinical data to verify it in the following aspects: the performance of transition states and the maintaining of a steady state while under anesthesia, distinctions between awareness and unconsciousness using the Fisher score, and stability given a poor signal quality. We report that the algorithm shows good performance superior to BIS even when considering the advantages of BIS. Thus, we hope that DoA diagnosis technology with EEG signals can be developed with this research.

II. Methods

A. Hidden Markov Model

Hidden Markov models are a general statistical modeling technique for linear problems such as sequences or time series. They have been widely used in speech recognition applications for many years. The key idea is that an HMM is a finite model that describes a probability distribution over an infinite number of possible sequences. The HMM is composed of a number of states. Each state emits symbols according to symbol emission probabilities, and the states are interconnected by state-transition probabilities. Starting from some initial state, a sequence of states is generated by moving from state to state according to the state-transition probabilities until an end state is reached. Each state then emits symbols according to that state's emission probability distribution, creating an observable sequence of symbols. The sequence of states is a Markov chain, because the choice of the next state to occupy is dependent on the identity of the current state. However, this state sequence is not observed and is hidden. Only the symbol sequence generated by these hidden states is observed. The most likely state sequence must be inferred from an alignment of the HMM to the observed sequence.

In general, when using HMMs we are interested in solving one of three problems. First, given an existing HMM and an observed sequence, we want to know the probability that the HMM could generate the sequence (the scoring problem). Second, we want to know the optimal state sequence that the

This research was supported by Basic Science Research Program through the National Research Foundation of Korea(NRF) funded by the Ministry of Education(2013R1A1A2013008).

Junbeom Kim and Sang-Hyun Park are with the KAIST Institute for IT Convergence, KAIST, Daejeon, Korea (phone: +82-42-350-7111; fax: +82-42-350-4290; email: silenteye@kaist.ac.kr, lovephysics@itc.kaist.ac.kr).

Huh Hyub and Seung Zhoo Yoon are with the Department of Anesthesiology and Pain Medicine, College of Medicine, Korea University, Seoul, Korea (email: clumania@hanmail.net, yoonsz70@gmail.com).

Ho-Jin Choi is with the Department of Computer Science, KAIST, Daejeon, Korea (email: hojinc@kaist.ac.kr).

Kwang Moo Kim is with the Charm Engineering Co.,Ltd. Gyeong-gi, Korea (email: kmkim@charmeng.com).

978-1-4244-7929-0/14/$26.00 ©2014 IEEE

HMM would use to generate the sequence (the alignment problem). Third, given a large amount of data, we want to find the structure and parameters of the HMM which best account for the data (the training problem).

B. *Overall Strategy*

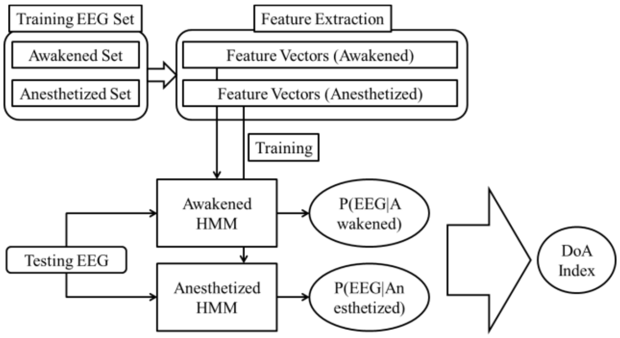


Figure 1. Overall strategy of the proposed approach

An outline of the EEG analysis algorithm with the HMM is shown in Fig. 1. The overall process consists of the two separate steps of training and testing. EEG data from a deep anesthesia state and a relatively awake state are used for the training of an anesthetized state and an awakened state, respectively. After the training process, the anesthetized state and the awakened state have their own separate models and parameters. For testing, feature vectors are extracted from EEG signals using the same algorithm used with the training process. The feature vectors are then quantized using the same centroid vectors that were used for the training process. Finally, the quantized sequence is divided into short sequences with 64 observations, and each short sequence is evaluated by means of likelihood evaluation.

The most important aspect when using a HMM for DoA is the extraction of the feature vector. The raw EEG itself cannot be used as a feature vector directly, as before any other signal processing steps, the information is not exploited in the EEG, which is a noise-like signal. Due to the observation of the dynamics of the EEG signal from an awakened state to an anesthetized state showing a decline in the spectrum in the gamma band, a process to change the domain from the time to the frequency domain is needed. The details of the feature vector extraction process will be discussed in section C.

C. *Feature Vector Extraction*

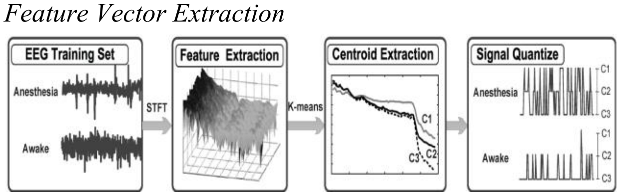


Figure 2. Feature vector extraction and quantized by extracted centroid

For training, the feature vectors of the EEG data are extracted for every epoch. With the k-means algorithm, the feature vectors are quantized to the nearest centroid vector. From the quantized sequence, the model parameters which best match the sequence can be estimated.

In the feature extraction process, preprocessed EEG signal is divided into epochs with a window size of 128, and every epoch overlaps 64 samples with the previous and next epoch. Short-time Fourier trans-form (STFT) is performed on each epoch, and the result of STFT is considered as the feature vectors with 128 elements. The extracted feature vectors are quantized and serve as a sequence for the hidden Markov model parameter extraction process. The quantized sequences have information about changes in the EEG signal in the frequency domain to quantify the depth of anesthesia. The parameter extraction process using the sequences to define two HMMs for awakened and anesthetized states will be discussed in the next section.

D. *Hidden Markov Model Parameter Extraction*

HMM parameter extraction from a given sequence is a key part of the training process. The HMM requires three model parameters, π, A and B, for specifications. Here, π implies the initial state distribution, while A represents the state transition probability distribution matrix and the B represents the observation symbol probability distribution in each state. For convenience, the compact notation $\lambda = (\pi, A, B)$ is used to indicate the three parameters. HMM parameter extraction is a process to find λ that maximizes the probability of the observation sequence. Because there is no analytical method to find the parameters, parameter extraction chooses λ such that $P(x|\lambda)$ is locally maximized using the Baum-Welch algorithm as an expectation-maximization (EM) method.

Awakened and anesthetized hidden Markov models are learned in the training state. The notations of the two models are as follows, with the denotations AW (awakened) and AN (anesthetized), respectively.

$$\lambda_{AW} = (\pi_{AW}, A_{AW}, B_{AW}) \tag{1}$$

$$\lambda_{AN} = (\pi_{AN}, A_{AN}, B_{AN}) \tag{2}$$

Each HMM calculate probability of producing the given testing sequences, $P(\vec{x}|\lambda_{AW})$ and $P(\vec{x}|\lambda_{AN})$, which will be used for derive final index to quantify depth of anesthesia. Likelihood evaluation and formula for DoA will be described in the next section.

E. *Likelihood Evaluation*

During the testing process, the testing sequences are evaluated by using method based on trained HMM parameters. The evaluating method computes the probability that the observed sequence was produced by the trained models. Since there is two HMM models from equation (1) and (2), λ_{AN} and λ_{AW}, for deep anesthetic state and awakened state respectively, two probabilities are computed from the models which are $P(\vec{x}|\lambda_{AN})$ and $P(\vec{x}|\lambda_{AW})$. Log of the ratio of two probabilities is used to estimate the depth of anesthesia.

$$\mathrm{HDoA} = \frac{\left(\log\left(\frac{P_{AW}}{P_{AN}}\right) + m\right)}{K} \tag{3}$$

Evaluated depth of anesthesia is numerically adjusted by adding a constant m and dividing a constant K as shown in equation (3), to achieve values from 0 to 100. The scaling constant m and K are fixed to 220 and 3.5 by results of experimentation.

F. *Preprocessing and Post-processing*

The measured EEG signal is vulnerable to various noise and artifacts [2]. For example, an electronic device generates 60Hz noise, thus easily contaminating an EEG signal. A notch filter is applied to re-move this noise. A wavelet-based de-noising technique is also incorporated in the pre-processing procedure to remove ocular artifacts generated from a patient`s eye movements [6]. Two main techniques are used in the post-processing procedure. The first is the moving average technique, which smoothens the calculated DoA index (BIS, which is our comparison target; the same moving average technique is also exploited for post-processing [7]). The second method is used when the normalized EEG signal changes abruptly. In this situation, the EEG signal is likely to be affected by various artifacts, and hence the final DoA index will also be inaccurate. Therefore, this index is replaced by a weighted average value from the previous 15-second index values. High weighting factors are given to recent index values, and low weighting factors to old index values.

The classification of the training and testing set is based on the BIS value. Because the BIS is re-liable at a steady state, a score that exceeds 95 is regarded as training for the awakened state while a score below 20 is regarded as the training set for the anesthetized state. The baseline is determined by the maximum and minimum values of the BIS data overall, which are 100 and 15, respectively. The others are regarded as the testing set.

III. Results

For experimentation, written consent was obtained from all patients after approval of the trial by the Institutional Review Board (No. MD11004-001). We studied 8 ASA physical status I-II patients, aged 25-60 years, who underwent general anesthesia conducted by a senior anesthetist at Korea University Anam Hospital.

A. *Analysis in Steady State*

Many previous clinical studies have shown that BIS can give appropriate index values in a steady state [8]. Thus, verification in a steady state is performed with the Pearson correlation between each method and the BIS. In order to evaluate the accuracy of the HDoA, epochs of steady-state behavior were extracted from each case using a decision algorithm based on the BIS value. There is no clear definition of what a steady state is, but it can be defined as when the index values are within 5 from the average index value in a 3-minute epoch. Therefore, if the index changes abruptly or BIS gives inappropriate values when the signal quality is poor, this state cannot be determined as a steady state. A higher correlation means that the index provides more reliable values in a steady state.

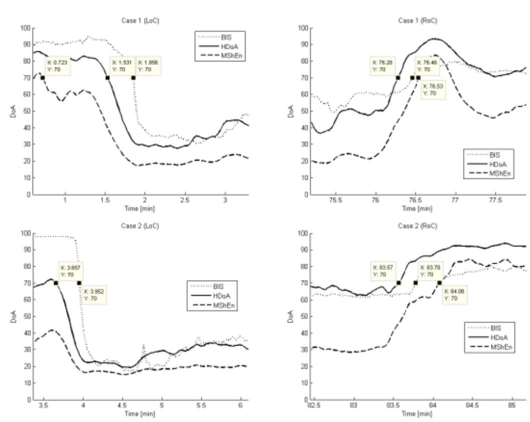
The correlation coefficient in steady state was 0.8831 in average as stated in Table I. A high correlation means that HDoA can provide reliable values in a steady state. The correlation coefficient in transitory state was 0.5210. There are two reasons for this discrepancy between all states and a steady state. First, HDoA has a fast response in the transition state, which means that the index values change abruptly from the awakening state to the anesthetic state. Second, BIS provides instable index values when the signal quality is not appropriate. In the following subsections, detailed explanations of these two considerations are given. These differences stem from improvements to the shortcomings of BIS and hence can be interpreted as positive aspects of HDoA.

B. *Analysis in Transitory State*

Among the shortcomings of BIS, the most remarkable drawback is its slow response in a transition state. BIS cannot give appropriate information about a patient`s state when the state of the patient changes from an awakening state to an anesthetic state [5]. Hence, BIS provides incorrect index values to the anesthesiologist. This can have deleterious effects on the patient`s health because the anesthetic agent could be overdosed. Like the BIS algorithm, although HDoA also applies a 15-second moving average as post-processing to smooth the final index values, the tracking speed of HDoA is faster than that of BIS, which means that the delay time of HDoA is shorter. Although the delay times depend on the patient`s state and the external environment, HDoA shows shorter delay times by 6 to 20 seconds than BIS in all cases as written in Table 1 (on average, shorter delay times by 15.10 seconds than BIS in LOC and by 12.15 seconds in ROC). We assume that an index value of 70 corresponds to the inflection point where the patient starts to be anesthetized. From Fig. 3, we see that HDoA (bold line) reacts to the patient`s anesthetic state faster than BIS (short dashed line).

Table I. Correlation in Steady State and Time Response Comparison between HDoA and BIS

cases	**Time Differences (sec)**		**Correlation in Steady State**
	LoC	***RoC***	
Case 1	19.50	10.80	0.9415
Case 2	17.70	12.60	0.8896
Case 3	10.74	18.00	0.7866
Case 4	10.20	18.00	0.9096
Case 5	10.56	6.00	0.9000
Case 6	18.06	7.20	0.9500
Case 7	17.04	15.60	0.9301
Case 8	16.98	9.00	0.7570
Average	15.10	12.15	0.8831
Standard Deviation	3.89	4.70	0.0673



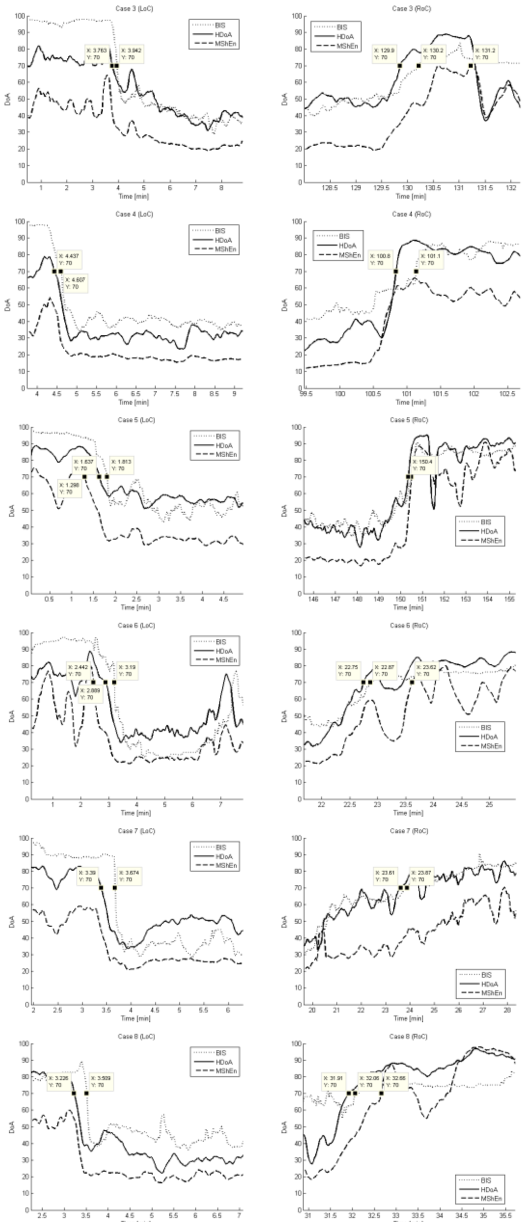


Figure 3. Time response comparison between HDoA and BIS.

C. Analysis using Fisher Score

The metric Fisher score is introduced to compare the performance of different types of depth of anesthesia algorithms [9]. This metric is a general method that serves to evaluate algorithms whose purpose is classifications between two different states. If an algorithm presents a high Fisher score, the algorithm can discriminate between the two states well. Here, the two states are an awakening state and an anesthetic state. Table II indicates the Fisher scores of BIS, MShEn, CAI, ICep and HDoA. From this table, we can conclude that HDoA is a superior algorithm in terms of the Fisher score.

$$\text{Fisher Score} = \frac{\left|MEAN(\overrightarrow{f_{AW}}) - MEAN(\overrightarrow{f_{AN}})\right|^2}{Var(\overrightarrow{f_{AW}}) + Var(\overrightarrow{f_{AN}})} \quad (4)$$

TABLE II. THE FISHER SCORE OF THE DOA METHODS.

Methods	Fisher Score
HDoA	62.64
BIS	47.11
MShEn	58.62
ICep	60.43

IV. CONCLUSION

In this paper, HDoA, an improved algorithm for measuring the depth of anesthesia, is introduced. This algorithm achieves several improvements compared to BIS, which is the most widely used method. BIS was introduced in 1996 and commercialized a few years later. Although the reliability of BIS has been identified through many clinical tests, it still has many shortcomings. HDoA retains the advantages of BIS and avoids its drawbacks. The superiority of MsCAI has been demonstrated via 8 actual clinical tests. HDoA offers the following two key strengths: a higher Fisher score than other algorithms, high correlation with BIS in a steady state, a low delay time in a transition state.

Future work would be more focused on validation with sufficient clinical data and to maintain stability in case of poor signal quality.

ACKNOWLEDGMENT

This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (2013R1A1A2013008).

REFERENCES

[1] I. J. Rampil, "Elements of EEG signal processing." *International journal of clinical monitoring and computing*, vol. 4, issue 2, 1989, pp. 85-98.

[2] S. Sanei, and J. A. Chambers, *EEG signal processing*. Wiley .com, 2008.

[3] L. C. Jameson, and T. B. Sloan, "Using EEG to monitor anesthesia drug effects during surgery," *Journal of clinical monitoring and computing*, vol. 20, issue 6, 2006, pp. 445-472.

[4] M. Jospin, P. Caminal, E. W. Jensen, H. Litvan, M. Vallverdú, M. M. Struys, and D. T. Kaplan, "Detrended fluctuation analysis of EEG as a measure of depth of anesthesia," *Biomedical Engineering, IEEE Transactions on*, vol. 54, issue 5, 2007, pp. 840-846.

[5] T. Zikov, S. Bibian, G. A. Dumont, M. Huzmezan, and C. R. Ries, "Quantifying cortical activity during general anesthesia using wavelet analysis," *Biomedical Engineering, IEEE Transactions on*, vol. 53, issue. 4, 2006, pp. 617-632.

[6] T. Zikov, S. Bibian, G. A. Dumont, M. Huzmezan, and C. R. Ries, "A wavelet based de-noising technique for ocular artifact correction of the electroencephalogram," in *Engineering in Medicine and Biology*, 2002.

[7] N. G. Chamoun, "An introduction to bispectral analysis for the electroencephalogram," *Journal of clinical monitoring,* vol. 10, issue 6, 1994, pp. 392-404.

[8] M. F. O'Connor, S. M. Daves, A. Tung, R. I. Cook, R. Thisted, and J. Apfelbaum, "BIS monitoring to prevent awareness during general anesthesia," *Anesthesiology*, vol. 94, issue. 3, 2001, pp. 520-522

[9] C. M. Bishop, and N. M. Nasrabadi, *Pattern recognition and machine learning,* vol. 1, 2006, New York: springer, p. 740.