Cross Entropy-based Automatic Thresholds Setting-Up Method for Sleep Staging System

Chen Chen, Adrien Ugon, Xun Zhang, Amara Amara, Patrick Garda, Jean-Gabriel Ganascia, Amina Kotti, Carole Philippe, Andrea Pinna

Cross Entropy-based Automatic Thresholds Setting-Up Method for Sleep Staging System

Chen CHEN*†, Adrien UGON*, Xun ZHANG†, Amara AMARA†, Patrick GARDA*, Jean-Gabriel GANASCIA*, Amina KOTTI *, Carole PHILIPPE ‡ and Andrea PINNA*
*Sorbonne Université, UPMC Univ Paris 06, LIP6, Paris, France
{first name.last name}@lip6.fr
†Institut Supérieur d'Electronique de Paris, MINARC, Paris, France
{first name.last name}@isep.fr
‡Hôpital Universitaire Pitié-Salpêtrière, Unité des Pathologie du Sommeil, Paris, France
carole.philippe@aphp.fr

***Abstract*—Sleep staging is a fundamental step in diagnosis and treatment of sleep disorders. In current sleep staging systems, normally a set of thresholds should be set up to determine the boundaries in differentiating different linguistic or symbolic features. However, as far as we know, there are no fully satisfying automatic method to do this task. Thresholds are mostly set up manually. In this paper, an automatic thresholds setting-up method based on Cross Entropy is proposed. Person-dependent thresholds can be provided automatically by using Cross Entropy and used in personalized sleep staging analysis while considering individual variability. The feasibility of Cross Entropy has also been evaluated, computational results exhibit that the Cross Entropy-based method is an efficient, convenient and applicable stochastic method for automatically setting-up thresholds in sleep staging system. Compared with manual method, average F-Measures are improved more than 10% for all the stages and up-to 74% for stage N3 by using proposed method.**

***Index Terms*—Cross Entropy, Thresholds, Personalization, Automatic Sleep Staging Systems**

I. Introduction

Sleep disorders are affecting more and more people which deteriorate quality of life and become a significant cause of morbidity and mortality [1]. In diagnosis and treatment of sleep disorders, sleep staging is considered to be the fundamental step. Clinical sleep staging is based on the visual analysis of an overnight polysomnography (PSG) recordings including electroencephalogram (EEG), electrooculogram (EOG) and electromyogram (EMG) by a physician under the guidance of American Academy of Sleep Medicine (AASM) manual [2]. It is a time-consuming and labor intensive task. Inter-rater variability also exists [3].

To release the burden of physicians, different sleep staging systems have been proposed in the past decades [4], [5], [6], [7], [8]. Among existing sleep staging systems, thresholds are widely applied in transforming digital parameters into linguistic or symbolic features to model inference process under the guidance of medical knowledge [5], [6], [7], [8].

In clinical practice, boundaries of linguistic or symbolic features are very flexible. Physicians may adjust the boundaries for each linguistic or symbolic feature according to their experience and patient information. While, in sleep staging systems, thresholds are used to determine the boundaries in differentiating different linguistic or symbolic features. However, as far as we know, there are no fully satisfying automatic setting-up thresholds methods in existing sleep staging systems. Manually predefined values of thresholds have been widely applied in [5], [6], [7], [8] due to the following reasons: 1) To build a mathematical model or a threshold function to differentiate boundary requires a set of data with sufficient quantity and adequate quality; 2) There is lack of uniformity between subjects and thresholds variability exists [9].

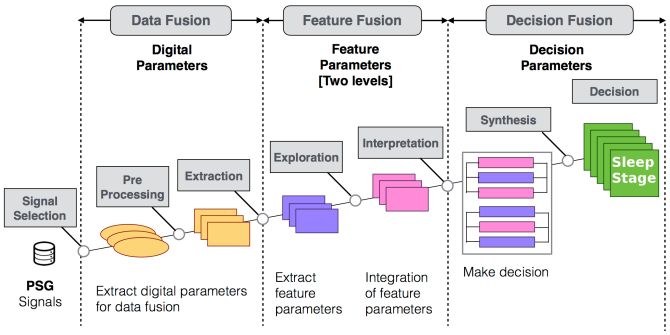
In this paper, we propose an automatic thresholds setting-up method by using Cross Entropy [10] to overcome the limitations of existing thresholds setting-up problems. With the stable and rapid searching ability in selecting the optimal thresholds combination, it can also be adapted to different person to provide personalized thresholds which is suitable to high inter-subject variability sleep staging analysis.

This paper is organized as follows. In the next section, a Symbolic Fusion-based sleep staging system is introduced with discussion on the limitations of thresholds setting-up problem. Then, an automatic thresholds setting-up method is proposed based on Cross Entropy. Followed by results and comparison with manual approaches in section four. Lastly, conclusion and further work are provided.

II. Symbolic Fusion-based Sleep Staging System

In [7], [8], a sleep staging system is proposed using symbolic fusion which consists three levels: data fusion, feature fusion and decision fusion as shown in Figure 1. In data fusion, digital parameters are extracted from PSG recordings to maximize the useful information and to minimize noise and artifacts. In feature fusion, digital parameters are transformed into feature parameters. It simplifies the interpretation of digital parameters, and also performs normalization, reduction and matching of digital parameters. In decision fusion, inference method is used to fulfill sleep staging on the basis of feature parameters.

Figure 1: Symbolic Fusion-based Sleep Staging System



In [7], [8], thresholds are used for translating nine digital parameters into 24 features parameters as shown in Table I. However, to define these thresholds, manual interpretation was performed. It is a time-consuming process; individual variability also exists.

Digital Parameters	Features Parameters	Threshold
EEG LowWaveEnergy	High - Middle - Low	2
EEG SleepSpindles	Confident - Not Confident	1
EEG KComplex	High - Middle - Low	2
EEG Delta	High - Low	1
EEG Theta	High - Low	1
EEG Stability	Stable - Not Confident - Unstable	2
EOG EyeMovement	High - Middle - Low - Lowest	3
EOG Correlation	Conjugate - Disconjugate	1
EMG Activity	High - Middle - Low	2

Table I: Thresholds used in Symbolic Fusion-based Sleep Staging System

III. Automatic Thresholds Setting-Up Method

To overcome the limitations in [7], [8], Cross Entropy-based Automatic Thresholds Setting-Up (CEATS) is proposed as shown in Figure 2. It is dedicated to (1) provide optimal thresholds for sleep staging system (2) release the burden of manual interpretation and (3) allow individual flexibility. Cross Entropy is introduced firstly, followed by details of CEATS.

A. Introduction on Cross Entropy

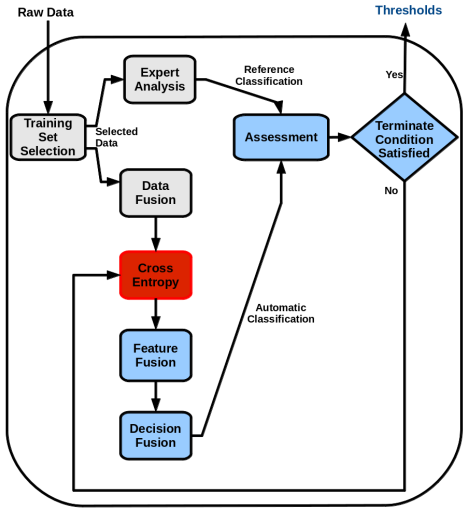
Cross Entropy (CE) was initially proposed to estimate probabilities of rare events for complex stochastic networks by Rubinstein [10] in 1997. It has been extended to solve combinatorial optimization problems in 1999 [11], which turned out to be an effective method.

CE is briefly introduced as follows. In solving the combinatorial optimization problem, a maximization problem can be described as shown in Equation 1.

$$\gamma^* = \underset{x \in \chi}{max} \left[S\left(x\right) \right] \tag{1}$$

γ^* represents the maximal value of S on the domain space χ. To proceed with CE, $f(\cdot; v)$ is defined as a family of Probability Density Functions (PDFs) on χ, with respect to some base measure v. Then γ^* can be estimated by $\ell(\gamma)$ defined in Equation 2.

Figure 2: Cross Entropy-based Automatic Thresholds Setting-Up Method



$$\ell\left(\gamma\right) = \mathbb{P}_u\left(S\left(X\right) \geq \gamma\right) = \mathbb{E}_u I_{\{S(X) \geq \gamma\}} \tag{2}$$

where X is a random vector generated by PDFs with parameter v in $f(x, v)$. $\mathbb{P}_u$ is the probability of the state $\{S(X) \geq \gamma\}$, $\mathbb{E}_u$ is the corresponding expectation operator and $I(\cdot)$ is the indicator function, i.e., $I_{\{S(X) \geq \gamma\}} = 1$ only if $S(X) \geq \gamma$, otherwise, it equals to zero.

Based on the important sampling: take N random samples $X = (X_1, X_2, ..., X_N)$ from an important sampling density g on χ, the unbiased estimator $\widehat{\ell}(\gamma)$ of $\ell(\gamma)$ can be defined as shown in Equation 3.

$$\widehat{\ell}(\gamma) = \frac{1}{N} \sum_{i=1}^{N} I_{\{S(X_i) \geq \gamma\}} \frac{f(X_i; v)}{g(X_i)} = \ell(\gamma) \tag{3}$$

From Equation 3, only one sample suffices to estimate $\ell(\gamma)$ since it is true for all i. While it is difficult to directly compute g because this g depends on the unknown parameter ℓ. Moreover, it is convenient to choose a g in the family of densities $f(\cdot; v)$. The CE method solves this efficiently by finding the minimal Kullback-Leibler distance as shown in Equation 4 which defines the distance between g and $f(\cdot; v)$.

$$\mathcal{D}(g, f) = \int g(x) ln g(x) dx - \int g(x) ln f(x; v) dx \tag{4}$$

In solving combinatorial optimization problems, CE creates a sequence of $f(\cdot; v_1), f(\cdot; v_2), ...$ of PDFs that are driven in the direction of the theoretically optimal density $f(\cdot; v^*)$. $f(\cdot; v^*)$ corresponds to the degenerate density of the optimal solution. In each iteration, it generates a set of samples and the elite samples (in terms of solution quality) would be selected to update the parameters of the PDF $f(x; v)$ parameterized by v. Since the elite samples are selected in each iteration, γ would be improved and can converge quickly to the optimal solution γ^*.

B. Cross Entropy-based Automatic Thresholds Setting-Up Method

We propose CE-based method to set up the personalized thresholds for a Symbolic Fusion-based sleep staging system [7], [8]. As shown in the Figure 2, a training set is initially selected. Clinical sleep staging is performed by an expert to classify the training set into different sleep stages ("*Expert Analysis*"). In parallel, automatic analysis is processed ("*Data Fusion*", "*Feature Fusion*" and "*Decision Fusion*"). In order to assess impact of different thresholds, "*Assessment*" of sleep staging is adopted by comparing results of automatic and expert analysis. With assessment results, CE can generate new thresholds combinations for the next loop. This process is repeated until optimal thresholds combination is found, i.e. until the "*Terminate Condition*" is satisfied.

To realize automatic thresholds selection for this sleep staging system, details of CEATS procedure is described below.

- **Step 1. Initialization**: Define a specified mechanism to generate PDFs. Defined the sample size N, elite sampling rate ρ and the max iteration number G.
  To search for the optimal thresholds combinations for the sleep staging system, normal distribution mechanism is performed to generate PDFs for each threshold which can be presented as $\mathcal{N}(\mu, \sigma^2)$. μ is the mean of the distribution and σ is the standard deviation. Initial value of μ and σ is set according to the distribution of each digital parameters as shown in Table I.
  Different sample sizes and elite sampling rates are analyzed in the next section. The max iteration is 50 which takes the consideration of computational time.
- **Step 2. Sampling**: According to normal distribution density function $\mathcal{N}(\mu, \sigma^2)$, a set of samples are generated. Each sample is composed of 15 different thresholds.
- **Step 3. Fitness function evaluation**: F-Measure is used as the fitness function to evaluate the impact of different thresholds on the sleep staging system. It balances both precision and recall as shown in Equation 5. Precision equals to (Tp/(Tp + Fp)) and Recall equals to (Tp/(Tp + Fn)). The parameters TP, FN, and FP are respectively *True Positives*, *False Negatives*, and *False Positives*, which are used to quantify the quality of a classification.

$$F - Measure = 2 * \frac{Precision * Recall}{Precision + Recall} \quad (5)$$

- **Step 4. Selection**: Rank the values of fitness function and select elite samples in term of F-Measure. In this step, a number of ρN samples with higher F-Measure are selected as elite samples.
- **Step 5. Updating**: Update the PDFs parameters μ and σ. According to the elite samples, new $\widehat{\mu}$ and $\widehat{\sigma}$ are calculated.
- **Step 6. Check terminate conditions**: If one of terminate conditions like, F-Measure reaches to a desired value or iteration reaches to the pre-defined value or the standard deviation σ is close to zero, is satisfied, then the procedure stops.
- **Step 7. Repeat**: If the terminate condition is not satisfied, repeat from Step 2 to Step 6 until one of the terminate conditions is satisfied.

IV. Results

In this section, we present the simulation results obtained by using CEATS and compare them with results obtained with Manually Thresholds Setting-Up (MTS).

A. Subjects, Recordings and Preprocessing

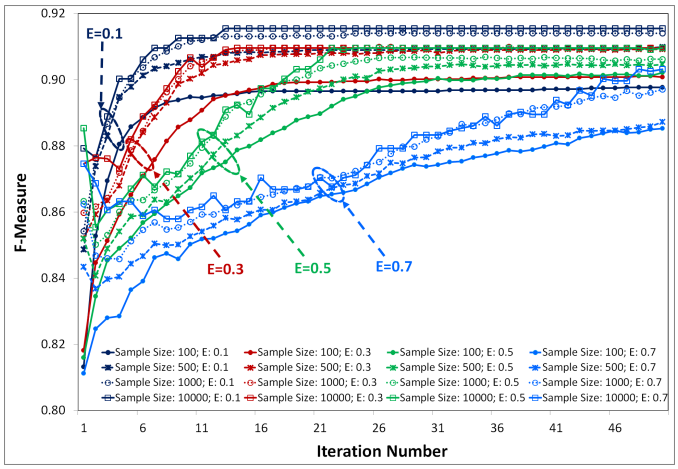
Overnight PSG signals were recorded from 14 subjects (9 males and 5 females) ranging from 22 to 65 years old (μ = 42, σ = 15.8) in La Pitié-Salpêtrière hospital (AP-HP), in Paris. AHI (average number of apneas and hypopneas per hour of sleep) ranges from 0 to 35.3 (μ = 13.7, σ = 19). PSG recordings were segmented into 30s epoch and manually scored by experts into five different stages : W, N1, N2, and N3 and R according to AASM manual. By using Symbolic Fusion, nine digital parameters have been extracted from raw PSG signals as shown in [7], [8].

B. Evaluation of Cross Entropy Parameters

Sample sizes of 100, 500, 1000 and 10000 with elite sampling rates 0.1, 0.3, 0.5 and 0.7 were set to evaluate the performance of CEATS.

Figure 3 shows F-Measure dependence on iteration number for different sample sizes of stage W classification for one patient using CEATS. Due to the stochastic ability of Cross Entropy, F-Measure in Figure 3 are mean values of 20 independent runs. Large sample size potentially increases the sample diversity and may provide fast convergence speed. However, it also increases the computational time. For the same sample size, the increase of elite sampling rate slows down the convergence speed. According to the simulation results, sample size of 500 with elite sampling rate of 0.1 provides optimum between F-Measure and computational complexity.

Figure 3: Cross Entropy with Different Sample Sizes and Elite Sampling Rates

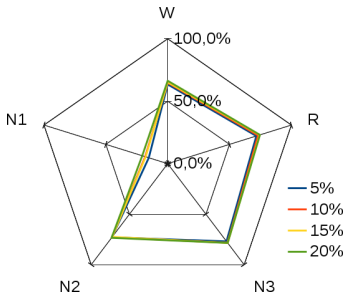


C. Evaluation of Training Set Selection

To evaluate impact of training set on CEATS, four values were used to balance the size of training set *vs* evaluation set : 5% *vs* 95%, 10% *vs* 90%, 15% *vs* 85% and 20% *vs* 80%. CEATS selects the optimal thresholds according to different training sets and pass these thresholds to the corresponding evaluation sets. Training set is randomly selected.

Figure 4 shows F-Measure of evaluation sets that use different training sets in CEATS. It shows that with reduction of training set size used in CEATS, it almost has no impact on the F-Measure for stage W, N2, N3 and R; only very slight difference for stage N1. In order to reduce the size of training sets, 5% is adopted in this paper.

Figure 4: Radar Chart of Different Training Sets



D. Comparison between CEATS and MTS

F-Measure is used to evaluate the impact of different thresholds on the sleep staging system. CEATS can provide personalized thresholds automatically for sleep staging system; while MTS needs manual efforts in setting-up thresholds. Comparison between CEATS and MTS [7], [8] is shown in Table II. It shows the increased F-Meausre value of each patient by applying CEATS in comparison to MTS method.

Database	W	N1	N2	N3	R
Patient 1	+0.540	+0.147	+0.193	+0.782	+0.271
Patient 2	+0.507	+0.087	+0.152	+0.796	+0.832
Patient 3	+0.118	+0.105	+0.308	+0.716	+0.111
Patient 4	+0.235	+0.055	+0.495	+0.530	+0.218
Patient 5	+0.144	+0.504	+0.430	+0.742	+0.517
Patient 6	+0.412	+0.185	+0.339	+0.620	+0.321
Patient 7	+0.481	+0.057	+0.283	+0.724	+0.687
Patient 8	+0.509	+0.067	+0.418	+0.840	+0.759
Patient 9	+0.235	+0.155	+0.181	+0.834	+0.714
Patient 10	+0.552	+0.129	+0.464	+0.824	+0.777
Patient 11	+0.580	+0.073	+0.225	+0.757	+0.728
Patient 12	+0.477	+0.080	+0.451	+0.808	+0.327
Patient 13	+0.436	+0.169	+0.359	+0.638	+0.249
Patient 14	+0.464	+0.026	+0.223	+0.770	+0.667

Table II: F-Measure Improvements by applying CEATS

In Table II, F-Measure for all the stages is increased by using CE-based method. For these 14 patients, average 40.6%, 13.1%, 32.3%, 74.1% and 51.3% F-Measure are improved for stage W, N1, N3 and R, respectively.

CEATS proved to be an efficient method which can automatically set up personalized thresholds for each subject. It can be applied to realize personalized sleep staging. To our knowledge, this is the first study to explore CE in thresholds setting-up problems for the sleep staging applications and it can also be extended to other domains.

V. Conclusion

In this paper, a CE-based Thresholds Setting-Up method is proposed for automatic sleep staging system. It automatically provides optimal thresholds for releasing the burden of manual interpretation of thresholds and considering individual variability. Instead of building a mathematical model or a thresholds function, CE is a way to search for the optimal thresholds combination among the possible thresholds combinations space. With robustness and rapid searching ability, CE demonstrates to be an effective method in solving combinatorial thresholds setting-up problems. For the further work, proposed method can be extended to realize complete personalized sleep disorders analysis and plan to be integrated in an embedded system to realize remote sleep monitoring.

VI. Acknowledgment

This work was financially supported by China Scholarship Council (CSC) and under collaboration of Laboratoire d'Informatique de Paris 6 and Institut Supérieur d'Electronique de Paris. We would also like to thank Dr C. Philippe and Dr A. Kotti for the collaboration and for providing the database and help in the field of medicine.

References

[1] J. L. Hossain and C. M. Shapiro, "The prevalence, cost implications, and management of sleep disorders: an overview," *Sleep and Breathing*, vol. 6, no. 2, pp. 85–102, 2002.

[2] A. A. of Sleep Medicine, C. Iber *et al.*, *The AASM manual for the scoring of sleep and associated events: rules, terminology and technical specifications*. American Academy of Sleep Medicine, 2007.

[3] H. Danker-hopfe, P. Anderer, J. Zeitlhofer, M. Boeck, H. Dorn, G. Gruber, E. Heller, E. Loretz, D. Moser, S. Parapatics *et al.*, "Interrater reliability for sleep scoring according to the rechtschaffen & kales and the new aasm standard," *Journal of sleep research*, vol. 18, no. 1, pp. 74–84, 2009.

[4] N. Schaltenbrand, R. Lengelle, and J.-P. Macher, "Neural network model: application to automatic analysis of human sleep," *Computers and Biomedical Research*, vol. 26, no. 2, pp. 157–171, 1993.

[5] S.-F. Liang, Y.-H. Chen, C.-E. Kuo, J.-Y. Chen, and S.-C. Hsu, "A fuzzy inference system for sleep staging," in *Fuzzy Systems (FUZZ), 2011 IEEE International Conference on*. IEEE, 2011, pp. 2104–2107.

[6] S.-F. Liang, C.-E. Kuo, Y.-H. Hu, and Y.-S. Cheng, "A rule-based automatic sleep staging method," *Journal of neuroscience methods*, vol. 205, no. 1, pp. 169–176, 2012.

[7] A. Ugon, "Fusion symbolique et données polysomnographiques," Ph.D. dissertation, Université Pierre et Marie Curie (UPMC, Paris 6), 2013.

[8] C. Chen, X. Liu, A. Ugon, X. Zhang, A. Amara, P. Garda, J.-G. Ganascia, C. Philippe, and A. Pinna, "Symbolic fusion: A novel decision support algorithm for sleep staging application," in *Proceedings of the 5th EAI International Conference on Wireless Mobile Communication and Healthcare*. ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering), 2015, pp. 19–22.

[9] J. Y. Maggard, "Automation of sleep staging," 2010.

[10] R. Y. Rubinstein, "Optimization of computer simulation models with rare events," *European Journal of Operational Research*, vol. 99, no. 1, pp. 89–112, 1997.

[11] R. Rubinstein, "The cross-entropy method for combinatorial and continuous optimization," *Methodology and computing in applied probability*, vol. 1, no. 2, pp. 127–190, 1999.