Personalized sleep staging system using evolutionary algorithm and symbolic fusion

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

To cite this version:

Chen Chen, Adrien Ugon, Xun Zhang, Amara Amara, Patrick Garda, et al.. Personalized sleep staging system using evolutionary algorithm and symbolic fusion. 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC'16), Aug 2016, Orlando, FL, United States. pp.2266 - 2269, 10.1109/EMBC.2016.7591181 . hal-01396595

HAL Id: hal-01396595
https://hal.sorbonne-universite.fr/hal-01396595

Submitted on 14 Nov 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Personalized Sleep Staging System using Evolutionary Algorithm and Symbolic Fusion

Chen Chen[1,2], Adrien Ugon[1], Xun Zhang[2], Amara Amara[2], Patrick Garda[1], Jean-Gabriel Ganascia[1], Carole Philippe[3] and Andrea Pinna[1]

Abstract— **This paper presents a novel system for automatic sleep staging based on evolutionary technique and symbolic intelligence. Proposed system mimics decision making process of clinical sleep staging using Symbolic Fusion and considers personal singularity with an adaptive thresholds setting up system using Evolutionary Algorithm. It proved to be an effective and promising system in personalizing sleep staging. This system can also be integrated with other medical systems to realize remote sleep monitoring or home-care.**

I. INTRODUCTION

Sleep is an essential part that contributes to refreshment and self-repairs in our daily life. However, the prevalence of sleep disorders is approximate to 10% of the population [1], including sleep apnea, insomnia and so on. Sleep disorders deteriorate quality of life [2] and become a significant cause of morbidity and mortality [3], [4]. Therefore, it is crucial to diagnose and treat sleep disorders in time. That's why a clinical sleep analysis is needed and sleep staging is its fundamental step.

Polysomnography (PSG) is the gold standard diagnostic test for sleep disorders. It consists in recording different physiological signals during a sleep period. Resulting curves are visual analysed by sleep experts who assess sleep cycles using a subset of the PSG signals (brain activity (3 derivations EEG), eyes movement (2 EOG) and chin muscle tone (1 EMG)). Guidelines for the visual scoring of sleep and associated events have been defined by the American Academy of Sleep Medicine (AASM) [5]. Five sleep stages are defined : W (Wakefulness), N1, N2, N3 and REM (Rapid Eye Movements). The whole recording is divided in 30-seconds epochs. Each epoch should be assigned to a sleep stage. Each sleep stage is defined by rules based on a list of criteria, combining the observation of specific visual patterns on the PSG signals. AASM guidelines provide definitions of sleep events and rules to score sleep stages. Visual patterns to identify sleep stages are defined by frequency ranges, amplitude, or other parameters that can be measured on the signals using standard signal processing methods. Although guidelines have been published to explain how to score sleep in general situation, it has been shown that a variability between subjects in the shape of recording waves exist [6].

Visual scoring of sleep stages is a tedious and time-consuming task. The inter and intra scorers variability is a real issue [7], although partially solved by the publication of international guidelines. However, it is still acknowledged that scoring practices need to be improved [8]. To support it, automatic sleep staging methods have gained a wide spread attention in researches. However, the automated systems that on the market are not accurate enough to make the sleep expert sufficiently confident in their use. How to win the trust of the sleep expert ? The symbolic fusion is a decision method to abstract scattered raw data to build a rich knowledge. Symbolic fusion can be conducted through the 3-levels framework designed by B. Dasarathy [9] : data level, features level and decision level. As a decision support method, symbolic fusion is faithful to the decision-making process of AASM manual. For each sleep stage, visual patterns should be extracted from the data, that correspond to one or more of the criteria of the AASM. Visual patterns are then combined, according to AASM rules, to take a decision about the sleep stage to assign to an epoch. An automatic sleep staging approach based on symbolic fusion has been proposed in [10]. The solution proposed by the authors still have some limitations. Some AASM criteria need to be added and fusion strategies need to be adapted to the new patterns within the symbolic fusion structure. This is due to the fact that the default symbolic fusion is still not enough « smart » to take into account the uniqueness of the person. In order to do that, an additional module should be developed.

Differential Evolution (DE) is a popular and efficient method of evolutionary algorithm which has been successfully applied to solve multi-parameter problems in diverse domains like mechanical engineering design or chemical engineering [11]. It owns many advantages : 1) DE can mimic natural biological evolution and provide a fast and stable convergence ; 2) It is less sensitive to initial population ; 3) It is a parallel search method ; and 4) It can improve fitness function value iteratively.

In this paper we proposed a method to personalize the sleep staging using the symbolic fusion and differential evolution. We modified the Dasarathy symbolic structure by introducing an additional personalization module based on evolutionary algorithm between the Data and Feature level. This paper is organized as follows. Next section describes the limitations of symbolic fusion and presents a method to personalize sleep staging. Followed by results

[1]Chen Chen, Adrien Ugon, Patrick Garda, Jean-Gabriel Ganascia and Andrea Pinna are with Laboratoire d'Informatique de Paris 6, Pierre and Marie Curie University, Paris, France. {first name.last name}@lip6.fr

[2]Xun Zhang and Amara Amara are with Institut Supérieur d'Electronique de Paris, MINARC, Paris, France. {first name.last name}@isep.fr

[3]Carole Philippe is with Hôpital Universitaire Pitié-Salpêtrière, Unité des Pathologie du Sommeil, Paris, France. carole.philippe@aphp.fr

and comparison with the symbolic fusion proposed in [10]. Lastly, discussion and conclusion are provided.

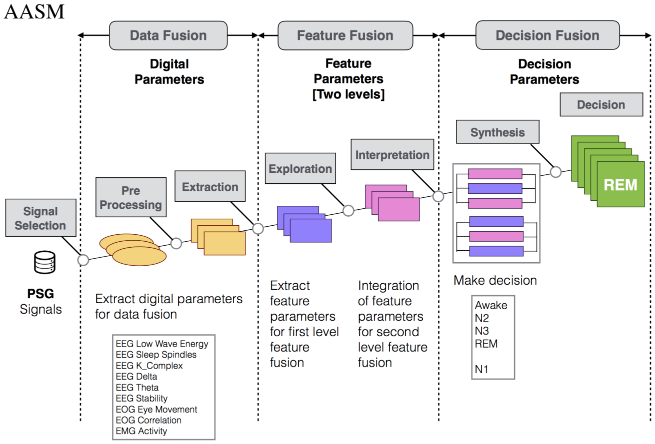
II. METHODS

Before describing the evolutionary algorithm and explaining its implementation in the symbolic fusion, a quick overview on the limitations of symbolic fusion to personalize the sleep staging is presented.

A. Sleep Staging by Symbolic Fusion

An existing system based on symbolic fusion has been designed to realize classification of sleep stages [10]. This method uses the three-level architecture defined by B. Dasarathy to abstract data. Figure 1 presents the architecture of symbolic fusion : data fusion, feature fusion and decision fusion.

Fig. 1. Sleep staging workflow using symbolic fusion in accordance with AASM



In data fusion, time-domain analysis, frequency-domain analysis, and non-linear analysis have been applied to maximize the useful information and to minimize noise and artifacts. Nine digital parameters were extracted as shown in data fusion section in Figure 1. In feature fusion, digital parameters were transformed into feature parameters. In decision fusion, inference method is used to fulfill sleep staging on the basis of feature parameters. The link between digital and feature parameters is explained via thresholds in Table I.

However, visually interpretation is needed to perform thresholds setting-up before the feature fusion. These thresholds are used to differentiate the boundary of different feature parameters. The boundary for low, middle or high feature is determined by a visual setting of threshold for digital parameter. Total 15 thresholds are defined for translating nine digital parameters into 24 feature parameters. In this paper, an Adaptive Thresholds Setting Up (ATSU) system is proposed to automatically provide optimal values of thresholds.

B. Personalized Sleep Staging Workflow

In order to personalize the automatic sleep staging, we propose an ATSU system to set up the thresholds automatically using Evolutionary Algorithm. This algorithm is

Digital Parameters	Features Parameters	Threshold
EEG Low Wave Energy	High - Middle - Low	2
EEG Sleep Spindles	Confident - Not Confident	1
EEG K_Complex	High - Middle - Low	2
EEG Delta	High - Low	1
EEG Theta	High - Low	1
EEG Stability	Stable - Not Confident - Unstable	2
EOG Eye Movement	High - Middle - Low - Lowest	3
EOG Correlation	Conjugate - Disconjugate	1
EMG Movement Activity	High - Middle - Low	2

TABLE I
CORRELATION BETWEEN DIGITAL AND FEATURE PARAMETERS VIA THRESHOLD SETTING UP

adopted because of its robustness and ability in searching and discovering combinatorial thresholds [12]. With the feasibility of evolutionary technique and the elasticity of symbolic intelligence, steps of the new workflow of sleep stages detection is shown in Figure 2; details of ATSU system architecture are shown in Figure 3.

Fig. 2. Personalized sleep staging workflow

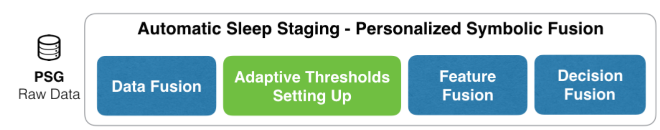
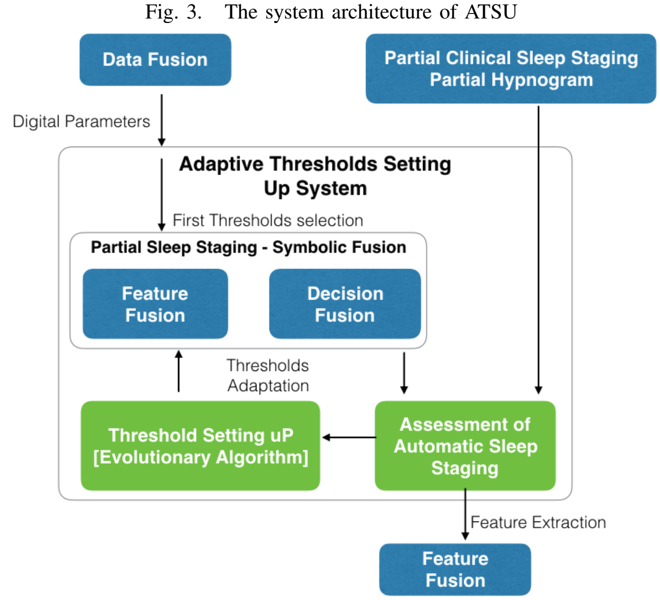


Fig. 3. The system architecture of ATSU



The ATSU has the digital parameters extracted from the PSG signals as input, a first sleep detection by feature and decision fusion is made only on a subset of the epochs with a default value of the thresholds regardless of the person. The same epochs subset is also scored by the sleep expert. Then we compare the result of the first partial symbolic fusion with the sleep expert scoring and an assessment is made in the « *Assessment of Automatic Sleep Staging* » block. If the assessment is positive, the thresholds can be used for full sleep staging analysis by the feature and decision fusion. If the assessment is negative, we need to find new values of the thresholds in order to fit as much as possible to the partial sleep expert scoring. This is performed in the « *Threshold Setting Up* » block by an Evolutionary Algorithm. Once the new values of thresholds are generated, a new threshold adaptation cycle will start. Details are described in the following paragraph.

1) Differential Evolution Algorithm and Assessment of Automatic Sleep Staging Specification: Before using an evolutionary algorithm to optimize a set of values, it is necessary to define the population that will be used. Our population is composed of NP individuals, each described by a D-dimensional vector. In this work we used the classical strategy that is commonly noted DE/rand/1/bin. « *rand* » specifies that the vector to be mutated is selected randomly, « *1* » is the number of pairs of individuals used in construction of mutation vector, « *bin* » is the crossover scheme. Details of DE for ATSU are described as follows.

Initial population generation The initial population $x_{i,G}, i = 1, 2, ..., NP$ is generated randomly and is composed of 15-dimensional individuals which corresponds to 15 thresholds. Three different population sizes (NP=60, 90, 120) are compared with the analysis provided in Section III-B. Generation G is set to 50, which means if termination condition does not reach within 50 loops, it will terminate and provide the best threshold combination of the 50^{th} generation.

Fitness function evaluation F-Measure is used as the fitness function in this study and is evaluated in the Assessment Automatic Sleep Staging block. F-Measure balances both precision and recall as shown in Equation 1. Pr and Re represents Precision and Recall, respectively. 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 classification results.

$$F - Measure = 2 * \frac{Precision * Recall}{Precision + Recall} \quad (1)$$

Check termination conditions The process terminates if the fitness function reaches a desired value(F-Measure=0.92, e.g as shown in Figure 4). Otherwise, the process continues on the mutation, crossover and selection.

Mutation A mutated population is generated, that will be used for crossover. Mutation uses the Equation 2. $v_{i,G+1}$ is the i^{th} individual of the mutated population of G+1 generation. The formula explains how to generate mutated individual $v_{i,G+1}$ from 3 randomly selected individual in previous generation. Mutation Factor *F* is used to control the amplification of the differential variation $(x_{r2,G} - x_{r3,G})$. r1, r2, r3 are randomly selected in [1, NP].

$$v_{i,G+1} = x_{r1,G} + F * (x_{r2,G} - x_{r3,G}) \quad (2)$$

Crossover Crossover is used to decide if the j^{th} parameter of i^{th} individual of the crossed-over population should be mutated or not. Two kinds of random values are generated : random value $j_r\,(i)$ is a position generated for i^{th} individual indicating that the $j_r\,(i)^{th}$ parameter of the i^{th} individual will be mutated ; random value $rb\,(i,j)$ is used to decide if the j^{th} parameter of i^{th} individual is mutated or not as shown in Equation 3.

$$u_{j,i,G+1} = \begin{cases} v_{j,i,G+1}, \text{ if } \quad rb\,(i,j) \leq CR \quad \text{or} \quad j = j_r\,(i) \\ x_{j,i,G}, \text{ otherwise} \end{cases} \quad (3)$$

where :

- $u_{j,i,G+1}$ is the j^{th} parameter of i^{th} individual in crossed-over population of G+1 generation ;
- $v_{j,i,G+1}$ is the j^{th} parameter of i^{th} individual in mutated population of G+1 generation ;
- $x_{j,i,G}$ is the j^{th} parameter of i^{th} individual in initial (non-mutated) population of G generation ;
- $rb\,(i,j)$ is a uniform random number ($rb\,(i,j) \in [0,1]$) ;
- CR is the crossover constant ($CR \in [0,1]$) ;
- $j_r\,(i)$ is the randomly generated parameter of i^{th} individual, ($j_r\,(i) \in [1,15]$) ;
- i is the individual ($i \in [1, NP]$) ;
- j is the parameter ($i \in [1, 15]$) ;
- G is the generation.

Selection Greedy criterion is used to decide whether an individual will be a member of next generation. The crossed-over individual is selected if the value of its fitness function is higher in comparison to fitness value of the initial individual, otherwise the initial individual is selected. In our case, the fitness function is the F-Measure *f*. This is formalized in Equation 4.

$$x_{i,G+1} = \begin{cases} u_{i,G+1}, \; if \quad f\,(u_{i,G+1}) > f\,(x_{i,G}) \\ x_{i,G+1}, \text{ otherwise} \end{cases} \quad (4)$$

III. RESULTS

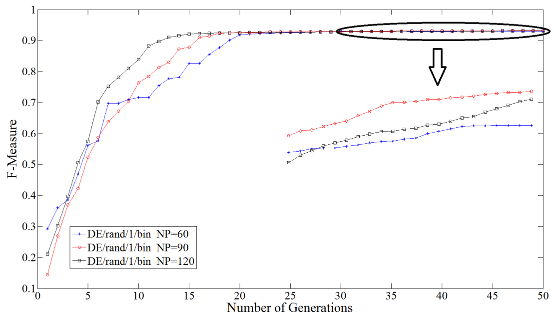
A. Subjects and Recordings

Overnight PSG signals were recorded from 12 subjects (2 males and 10 females) ranging from 26 to 67 years old ($mean$ = 53.42 years, std = 14 years) in Tenon hospital. PSG recordings of EEG, EOG and EMG were segmented into 30s epoch and manually scored by experts into five different stages : W, N1, N2, and N3 and REM according to AASM manual.

B. ATSU Evaluation

For each subject, 20% of the total epochs were randomly selected as a set and uniformly distributed in different stages to be the take into account during the ATSU. With this epochs selection, a population size of 90 for DE is used for adaptive thresholds selection. Figure 4 shows maximum F-Measure of Stage N3 in each generation with population sizes of 60, 90 and 120. Faster convergence is achieved with increase in the population size at the cost of greater computational complexity. Population size of 90 provides optimum between accuracy and computational complexity.

Fig. 4. F-Measure in each generation for ATSU



C. Comparison between Personalized and Non-Personalized Sleep Staging

We compared the F-Measure between the two sleep staging systems using symbolic fusion, one with a manual setting-up of thresholds and another with an automatic method based on the DE which can realize personalized sleep staging. In Table II, F-Measure for all the stages is increased for ATSU based personalized sleep staging in comparison to non-personalized sleep staging. For stage REM, F-Measure increases over 6% ; for stage W, N2 and N3, F-Measure increases 11%, 10% and 15 % respectively.

Sleep Stage	F-Measure Non-Personalized Symbolic Fusion	F-Measure Personalized Symbolic Fusion
W	0.779	0.868
REM	0.712	0.757
N1	Not evaluated	Not evaluated
N2	0.575	0.659
N3	0.834	0.921

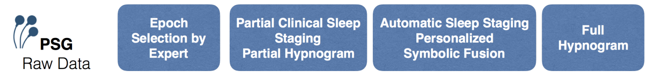
TABLE II
F-MEASURE COMPARISON

These results are encouraging and prove the feasibility of a Personalized Sleep Staging System to help sleep experts in the PSG analysis. In this paper, we don't evaluate the accuracy of the symbolic fusion itself (as already mentioned the symbolic fusion for sleep staging presented in [10] is still incomplete). With these results we prove how to make automatic sleep staging using the symbolic fusion according to AASM and taking into account the personal singularity based on differential evolution.

IV. DISCUSSION AND CONCLUSION

In the literature there are few works that try to validate an automated sleep staging system according to the AASM and the experience of sleep experts. Symbolic fusion and evolutionary algorithm proved to be a promising approach to help the experts in the PSG analysis. A new way to do PSG is proposed, Figure 5 shows the new workflow for personalized sleep staging and its possible use to help experts. Only a few number of epochs need to be analysed by expert instead of overall PSG examen ; and the remaining epochs are ananlysed by the automated sleep staging system.

Fig. 5. The new workflow for personalized sleep staging



Moreover, an evolutionary algorithm is implemented in order to take into account the personal singularity. The combination of evolutionary technique and symbolic intelligence proposed in Personalized Sleep Staging System supplements the shortcomings of symbolic fusion and overcomes personal singularity limitation. This is a first proof of concept, for the further work, we need to define an effective way to choose the good number of generations in order to optimize the computational time for the thresholds setting up (e.g. for G=50 we have 37.81 sec with a Intel Core 5@1.6 GHz and 4 GB of RAM). We also need to understand, with the sleep experts, more efficient way to choose the optimal epochs for the partial clinical sleep staging.

ACKNOWLEDGMENT

This work is financially supported by China Scholarship Council (CSC) and under collaboration of Laboratoire d'Informatique de Paris 6 and Institut Supérieur d'Electronique de Paris.

REFERENCES

[1] S. Ram, H. Seirawan, S. K. Kumar, and G. T. Clark, "Prevalence and impact of sleep disorders and sleep habits in the united states," *Sleep and Breathing*, vol. 14, no. 1, pp. 63–70, 2010.
[2] C. D'Ambrosio, T. Bowman, and V. Mohsenin, "Quality of life in patients with obstructive sleep apnea : Effect of nasal continuous positive airway pressure—a prospective study," *Chest*, vol. 115, no. 1, p. 123–129, 1999.
[3] N. Punjabi, "The epidemiology of adult obstructive sleep apnea," *Proc Am Thorac Soc*, vol. 5, no. 2, pp. 136–143, 2008.
[4] M. Arzt, T. Young, L. Finn, J. Skatrud, and T. Bradley, "Association of sleep-disordered breathing and the occurrence of stroke," *Am J Respir Crit Care Med*, vol. 172, no. 11, p. 1447–1451, 2005.
[5] A. A. of Sleep Medicine *et al.*, *AASM manual for the scoring of sleep and associated events*. American Academy of Sleep Medicine, 2007.
[6] H. G. Jo, J. Y. Park, C. K. Lee, S. K. An, and S. K. Yoo, "Genetic fuzzy classifier for sleep stage identification," *Computers in Biology and Medicine*, vol. 40, no. 7, pp. 629–634, 2010.
[7] S. T. Kuna, R. Benca, C. A. Kushida, J. Walsh, M. Younes, B. Staley, A. Hanlon, A. I. Pack, G. W. Pien, and A. Malhotra, "Agreement in computer-assisted manual scoring of polysomnograms across sleep centers," *Sleep*, vol. 36, no. 4, pp. 583–589, 2013.
[8] T. Penzel, X. Zhang, and I. Fietze, "Inter-scorer reliability between sleep centers can teach us what to improve in the scoring rules," *J Clin Sleep Med*, vol. 9, no. 1, pp. 81–87, 2013.
[9] B. V. Dasarathy, "Sensor fusion potential exploitation-innovative architectures and illustrative applications," *Proceedings of the IEEE*, vol. 85, no. 1, pp. 24–38, 1997.
[10] A. Ugon, J.-G. Ganascia, C. Philippe, H. Amiel, and P. Lévy, "How to use symbolic fusion to support the sleep apnea syndrome diagnosis," vol. 6747, pp. 45–54, 2011.
[11] R. Storn and K. Price, "Differential evolution–a simple and efficient heuristic for global optimization over continuous spaces," *Journal of global optimization*, vol. 11, no. 4, pp. 341–359, 1997.
[12] D. B. Fogel, "The advantages of evolutionary computation." in *BCEC*. Citeseer, 1997, pp. 1–11.