32nd Annual International Conference of the IEEE EMBS
Buenos Aires, Argentina, August 31 - September 4, 2010

Automatic K-complexes Detection in Sleep EEG Recordings using Likelihood Thresholds

S. Devuyst, T. Dutoit, P. Stenuit, and M. Kerkhofs

Abstract— **In this paper, we present an automatic method for K-complexes detection based on features extraction and the use of fuzzy thresholds. The validity of our process was examined on the basis of two visual K-complexes scorings performed on 5 excerpts of 30 minutes. Results were investigated through all different sleep stages. The algorithm provides global true positive rates of 61.72% and 60.94%, respectively with scorer 1 and scorer 2. The false positive proportions (compared to the total number of visually scored K-complexes) are of 19.62% and 181.25%, while the false positive rates estimated on a one 1 second resolution are only of 0.53% and 1.53%. These results suggest that our approach is completely suitable since its performances are similar to those of the human scorers.**

I. Introduction

THE K-complex is an EEG transient event that occurs during sleep, spontaneously or in response to stimuli [1], [2]. It was described for the first time by Loomis *et al.* in 1938, like an association of a fast wave (8-16Hz) and a delta wave [3]. Since this, repeated efforts were made to delineate a standard K-complexes definition but none of them gained broad acceptance because of the high variability of its morphology (Fig. 1).

According to the AASM rules [4], the K-complex is a "*well-delineated negative sharp wave immediately followed by a positive component standing out from the background EEG, with total duration ≥ 0.5 sec, usually maximal in amplitude when recorded using frontal derivations*". To this definition, Rechtschafen and Kales (R&K) criteria [2] add that "*spindle activity (12 to 14 Hz) wave may or may not constitute a part of the complex*".

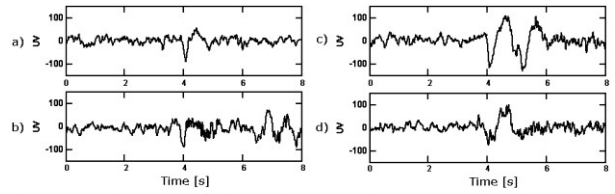

Fig. 1. Various morphology of K-complexes (every fourth second): a) isolated K-complex without spindle, b) isolated K-complex with spindle, c) pair of K-complexes, d) K-complex with blunted negative component.

Manuscript received April 23, 2010. This work was supported in part by the Belgian Network DYSCO (Dynamical Systems, Control, and Optimization), funded by the Interuniversity Attraction Poles Programme, initiated by the Belgian State, Science Policy Office.

S. Devuyst and T. Dutoit are with the TCTS Lab, Université de Mons - UMONS, B-7000 Mons, Belgium (stephanie.devuyst@umons.ac.be; phone: +32 (65) 37.47.20; fax: +32 (65) 37.47.29).

P. Stenuit and M. Kerkhofs are with the SleepLaboratory CHU Vésale, Montigny-le-Tilleul, Belgium.

None of these two standards in sleep analysis mention an absolute or relative amplitude criterion. However, many authors have suggested to use a minimal amplitude of 75uV [5], [6] or a minimum peak to peak amplitude criterion of 100uV [6]. Other works have preferred to require that K-complexes be at least twice as high as the EEG background within the 1 or 5 preceding seconds [7].

Some works have also imposed a maximum duration, generally comprised between 1 to 3 s [6]-[8].

Reliable detection of K-complexes is essential for sleep stage scoring, since they constitute (with the spindles) one of the principal markers of transition from sleep stage 1 or REM to sleep stage 2 (or 3 or 4). Unfortunately, their visual identification is very time-consuming (there are typically 1 to 3 K-complexes per minute in stage 2 of young adults [1]) and rather subjective since it cannot be performed on regular basis. Hence, poor inter-scorers agreements are reported in literature: 50 & 57% for Bremer *et al.* [6] and only 32 & 52% in Sherriff *et al.* [9]. That's why automatic identification of K-complexes is of great interest.

Bremer *et al.* [6] proposed an electronic system based on filtering, pulsers, threshold detector and both analog and digital logic techniques. He obtained a sensitivity of 63% by not distinguishing the various sleep stages, and a sensitivity of 68% by only considering sleep stage 2.

Sherif *et al.* [9] used a mathematical description of the K complex morphology to apply matched filter. The resultant true positive (TP) and false positive (FP) rates were respectively 67.18% and 154,68% on a first visual scoring and 85,44% and 52,43% on a second visual scoring.

The matching filtering was also tested by Woerst *et al*. [10]. In this work, the EEG was filtered with 3 asymmetrical patterns scaled to durations of 0.5s, 1s and 1.5s. The filtered signals were then summed up and compared to a threshold. The sensitivity was calculated on a 1 second resolution and showed to be very good (93%). Unfortunately, the corresponding specificity still remained high (32%).

Another templates-based method using a dynamic time warping measurement has been suggested by Kerkeni *et al.* [11]. However, comparable results were obtained with a FP rate of almost 34% for a TP rate of 80%.

The Rosa and Paiva algorithm [7] is based on a stochastic model for the generation of K-complexes. This model consists of feedback loops of rhythms driven by white noise and pulses. The subsequent detector corresponded to the model working in the inverse way. The TP rates were

comprised between 83% and 92% while the false positive rate varied between 32% and 157%.

Finally, trained artificial neuronal networks were used to classify and recognize K-complexes. In this class, Jansen [8] tried to do a direct automatic K-complexes recognition (where the network inputs were the samples of the bandpass filtered EEG). However he achieved poor performances with detection rate ranging from 42% to 67%. Then, Bankman *et al*. [12] proposed to use features-based neural networks detection, where 14 features were taken on significant points of the possible K-complex. Much better performance where obtained with a sensitivity of 90% for about 8% of FP.

All these results are not easily comparable since they depend on the evaluation method. Algorithms for classification problem [7], [8], [11] and [12] use EEG segment of fixed length containing K-complexes or not. The corresponding false positive rate is calculated as follow:

$$FPrate = 1 - specificity = \frac{nbr\ of\ False\ Positive}{(nbr\ of\ False\ Positive\ + nbr\ of\ True\ Negative)} \quad (1)$$

Algorithms for detection problem use a whole night recording. In this case the false positive rate can either be calculated on a 1 second resolution like above [10], or approximated by looking at the proportion of false positives compared to the number of real K-complexes [6], [9]:

$$"FPrate" \approx FPproportion = \frac{nbr\ of\ False\ Positive}{(nbr\ of\ True\ positive + nbr\ of\ False\ Negative)} \quad (2)$$

In this study we computed these two parameters so that our results can be compared with those of the other existing methods. In addition, we preferred to examine the sleep stage influence on the visual/automatic agreement rate rather than restricting the K-complexes identification to stage 2. To limit the number of false detections, we used a multi-level approach whose certain procedures are inspired from the above literature. Nevertheless, the originality of this work is the using of likelihood thresholds to combine their outcomes.

To check the validity of our process, we compared the automatic detection with two visual K-complexes scorings performed independently by two experts. The corresponding results are reported in section III.

II. Method

A. Recording

Data used in this study were recorded at the Sleep Laboratory of the André Vésale hospital (Montigny-le-Tilleul, Belgium). The training data set consists of 2 whole-night recordings and 1 excerpt of 2 hours coming from 3 healthy subjects. The testing data set consists of 5 excerpts of 30 minutes extracted from whole-night recordings of 5 other healthy subjects. The sampling rate (*fsampling*) was 200Hz. These records were given independently to two experts for K-complexes scoring. The corresponding whole night recording were also previously scored in sleep stages according to the Rechtschaffen and Kales criteria [2], but no note of these stages quotations was given to the scorers.

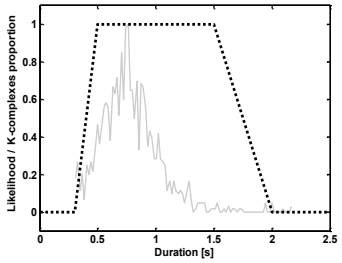

Fig. 2. In solid line: normalized distribution of K-complexes according to their duration feature. In dotted line: the corresponding thresholding curves providing the likelihood to be a real K-complex

B. Automatic analysis

The present algorithm is based on features extraction and the use of likelihood thresholds. The strategy is to limit the number of false detections thanks to a multi-level approach. At each stage, a new characteristic of the pseudo K-Complex is compared to a threshold to confirm (or not) its legitimacy. Nevertheless, to not reduce the TP rate in the same way, we chose to use fuzzy thresholds.

Hence, each pseudo K-complex is no more simply rejected or accepted on some level, but the algorithm gives to it a likelihood of corresponding to a real K-complex. This likelihood is computed according to its feature value and according to some "thresholding curve" defined beforehand.

For example, by considering the thresholding curve providing a likelihood to be a real K-complex according to its duration (illustrated by dotted line in Fig. 2), one can see that a pseudo K-complexes of 0.4s duration will still be accepted but with a likelihood of only 0.5.

These thresholding curves were established on the basis of the training database. For each feature, we have examined the distribution of visually identified K-complexes according to their feature value and we have selected the best corresponding "fitting" curves. Such normalized distribution and the associated thresholding curve is illustrated in Fig. 2.

The characteristics were selected so as to reflect the visual criteria as well as possible. For the majority of them, they were extracted from significant points of the possible K-complex. These significant points are similar to those of Bankman *et al*. [12] and are illustrated in Fig. 3:

- *val_min* and *t_min* (X) correspond to the minimal value of the pseudo K-complex.
- *val_max* and *t_max* (*) correspond to the maximal value of the EEG in the interval [*t_min*, *t_min*+1s].
- *t_end* (□) is the first time greater than *t_max* where the EEG is lower than -5uV.
- *t_start* (o) corresponds to the first local maximum which is higher than *val_min*/2, met by scanning the EEG from right to left starting from *t_min*. If this local maximal value is greater than 0uV, then *t_start* is moved to correspond to the first sample higher than 0uV.
- *t_mid1*(>) is to the first value greater than 0uV met by scanning the EEG from left to right starting from *t_min*.
- *t_mid2*(<) is to the first value lower than -5uV met by scanning the EEG from right to left starting from *t_end*.

First, the algorithm detects all local minima. For each one of them, it computes the difference of amplitude with those of the preceding and following local maxima. Then these two values are summed. If the corresponding result is sufficient, the local minimum is considered to be the negative sharp wave of a possible K-complex.

Then, the algorithm checks if this negative sharp wave is immediately followed by a positive component, by computing a feature that reflects the continuity of the growing side of the K-complex:

f2=(t_mid2-t_mid1)/(t_end-t_start)

The duration of the K-complexe is represented by:

f3=(t_end-t_start)

A set of features are then extracted to verify that the K complex amplitude is sufficient. A minimum peak to peak amplitude is first required (although the related threshold is low):

f4=(val_max - val_min)

Then, the automatic process makes sure that pseudo K complex amplitude is sufficient compared to the EEG background. To this end, we thought using the 2 or 3 seconds preceding the K-complex to represent the EEG background. However, in case of K-complexes occurring by pair, the second micro-event was never detected, due to the influence of the first one on the measurement. So we decided to accept K-complexes whose amplitude was either greater than the mean EEG amplitude deduced from the 2 seconds preceding *t_start*, or greater than the mean EEG amplitude deduced from the 2 seconds preceding (*t_start-2*):

$$f5= (val\_max - val\_min)/min(amplEEG_{t.start-2:\ t.start}, amplEEG_{t.start-4:\ t.start-2})$$

Unfortunately, this was not sufficient because of the presence of punctual artifacts. Therefore we also compute the overall EEG background amplitude on the 15 seconds surrounding the pseudo K-complex.

f6=(val_max - val_min)/mean_EEG_amplitude_on_15s

Finally, if this mean background EEG amplitude is too high the potential K-complex is also rejected:

f7=mean_EEG_amplitude_on_15s

To ensure that the amplitude of the negative component is at least 50% of the positive amplitude component (as required in [5]), the algorithm uses the following feature:

f8=abs(val_min)/ val_max

Concerning the sharpness of the first negative wave compared to the second positive wave, a relevant criteria is:

f9=(t_end- t_mid1)/ (t_mid1- t_start)

However, we found that this ratio can sometimes be lower than 1. That's why we associated to it a lower threshold value and considered other characteristics. Hence, the sharpness of the negative wave was represented by:

f10=abs(val_min)/((t_mid1- t_start) fsampling)*

This measurement can produce inadequate estimates of the sharpness when the waveform is complex. That's why we also used the least square acceleration (LSA) filtering [13] which computes the 2nd-derivative of the best approximating parabola to the negative sharp wave.

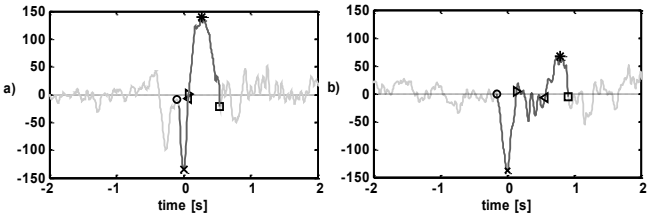

Fig. 3. Indications of t_start (o), t_end (□), t_mid1 (>), t_mid2 (<), val_min (x) and val_max (*) on a real K complex (a) and a non K-complex (b).

f11=LSA filtering of the negative sharp wave

Finally, the following characteristic is used:

*f12=(val_max - val_min) /((t_end- t_start)*fsampling)*

All these features are extracted from significant points of the pseudo K-complex and make it possible to eliminate a great amount of false positives. However, a lot of them were still found in sleep stage 3. Indeed, delta waves occur often by burst in this stage, so that the average EEG background amplitude computed on the 15 seconds surrounding the micro-event is not sufficient.

That why we also examine the frequency contents of the 5 seconds surrounding the K-complex, and calculate the relative power in the delta band:

f13=relative power in the delta band

Once the algorithm is applied to an EEG, one then obtained a set of likelihoods of being a K-complex computed for each characteristic. To combine these various values, we chose to select the minimal likelihood and to decrease it proportionally to the number of likelihoods < 0.6:

$$global\_likelihood = minimal\_likelihood * (1 - (0.1 * nbr\ of\ likelihood < 0.6)) \quad (3)$$

Lastly, as we know that it exists a correlation between the EEG channels with respect to the K-complex waveform, we computed the final likelihood of obtaining a real K-complex by averaging the global likelihoods obtained on the central EEG channel and on the frontal EEG channel.

III. Results

Once the final likelihood is calculated for each pseudo K-complex, it is necessary to select a minimal value of it, for which detection will be confirmed. In this way, we chose to use the value of 0.69, which correspond to the optimal point of the ROC curve obtained by varying the minimal likelihood from 0.5 to 1 on the testing data set. Then we computed the detection results through the different sleep stages. The numbers of K-complexes detected by the system and/or the scorers are reported in table I and the corresponding confusion matrix are reported in Table II.

From the Table I, it can be noticed that for a total of 209 K-complexes scored by scorer 1 and 64 K-complexes scored by scorer 2, there was a mutual agreement on only 43. This corresponds to sensitivities of only 20.57% and 67.18% respectively. Moreover, the detection system agreed with 33 of these K-complexes, which corresponds to an agreement rate of 76.74% (when a K-complex is considered as real when both scorers marked it as such).

TABLE I
SUMMARIZATION OF DETECTED RESULTS ON THE 5 EXCERPTS OF THE TESTING DATA BASE. MINIMAL LIKELIHOOD=0.69

	Sleep stages						Total
	W	REM	S1	S2	S3	S4	Total
Nbr. total scored by system	4	0	0	117	51	3	176
Nbr. total scored by scorer #1	1	0	0	143	60	5	209
Nbr. total scored by scorer #2	0	0	0	37	18	9	64
Nbr. scored by only system	3	0	0	27	10	1	41
Nbr. scored by only scorer #1	0	0	0	49	19	2	70
Nbr. scored by only scorer #2	0	0	0	3	4	8	15
Nbr. scored by only system & scorer #1	1	0	0	64	29	2	96
Nbr. scored by only system & scorer #2	0	0	0	4	2	0	6
Nbr. scored by only scorer #1 & scorer #2	0	0	0	7	2	1	10
Nbr. scored by system & scorer #1 & scorer #2	0	0	0	23	10	0	33

TABLE II
CONFUSION MATRIX (WITHOUT DISTINGUISH THE DIFFERENT SLEEP STAGES). MINIMAL LIKELIHOOD=0.69

	Yes #1	No #1		Yes #2	No #2		Yes #1	No #1
Yes syst	129	47	Yes syst	39	137	Yes #2	43	21
No syst	80	~8744	No syst	25	~8799	No #2	166	~8770

On table II, we can observe that the TP rates of the automatic system are of 61.72% and 60.94%, respectively with scorer 1 and scorer 2. Concerning the false detection rate we can see that the proportion of false positives (automatically detected) compared to the number of K-complexes scored by scorer 1 is of 19.62% and the proportion of false positives compared to the number of K-complexes scored by scorer 2 is of 181.25%.

These values can seem high, but there are actually analogous with those of human scorers. Indeed, the proportion of false positives detected by scorer 1 compared to the number of K-complexes scored by scorer 2 is of 22.49% and the proportion of false positives detected by scorer 2 compared to the number of K-complexes scored by scorer 1 is of 214.06 %.

Moreover, if we consider that the mean duration of K-complexes is 1 second and if we remember that the total duration of the testing database is 5*1800 seconds, we can approach the number of true negative by:

$$nbr\ of\ True\ Negative \approx 5*1800 - nbr\ of\ False\ Positive - nbr\ of\ True\ Positive - nbr\ of\ False\ Negative$$

This corresponds to a FP rate of only 0.53% with the visual scoring 1 and of 1.53% with the visual scoring 2.

III. CONCLUSIONS

It has been shown that the K-complexes are used less than 5 percent of the time in scoring sleep because of the higher number of sleep spindles [6]. However, this does not prevent from developing an accurate K-complex detection system for these 5%.

Here, we presented a method based on features extraction and likelihood thresholds. The detection performances were evaluated on the basis of two human scorings performed independently. TP rates of 61.72% and 60.94% were respectively obtained with scorer 1 and 2. These results are comparable with those of the other existing methods. However, our algorithm presents the advantage of a rather simple implementation. On the other hand, these results are within the range of the agreement rates between scorers (which are of 20.57% and 67.18% on our data base).

The proportion of FP is much higher with the second scorer (181.25%) than with the first scorer (19.62%). This is due to the fact that the second expert marked much less micro-events as K-complex. Nevertheless, by looking at the FP rates estimated on a 1 second resolution, the results are much more encouraging since the system provides a FP rate of only 0.53% with visual scoring 1 and of 1.53% with visual scoring 2.

These results suggests that our approach is completely suitable for an automatic K-complex detection although it is not appropriated for an on line detection.

REFERENCES

[1] M. Kryger, T. Roth and W. Dement, *Principles and practice of sleep medicine*, 4th edition, Elsevier Saunders, ISBN: 1416003207, 2005.

[2] A. Rechtschaffen ,and A. Kales, "A manual of standardised terminology and scoring system for sleep stages in human subjects," U.S. Government Printing Office, Washinghton, DC; 1968.

[3] A. Loomis, E. Harvey, G. Hobart, "Distribution of disturbance patterns in the human electroencephalogram, with special reference to sleep," *J. Neurophysiol*., vol 1, pp 413-430, 1938.

[4] C. Iber, S. Ancoli-Israel, A. Chesson and SF. Quan, "The AASM manual for the scoring of sleep and associated events : rules, terminology and technical specifications, " American Academy of Sleep Medicine, Westchester, Illinois (IL), 2007

[5] A. Rodenbeck, R. Binder, P. Geisler, H. Danker-Hopfe, R. Lund, F. Raschke, HG. Wee and H.Schulz, "A review of sleep EEG patterns. Part I: A compilation of amended rules for their visual recognition according to Rechtschaffen and Kales," *Somnologie*, Vol 10(4), pp 159-175, 2006

[6] G. Bremer, JR. Smith and I. Karacan, "Automatic detection of the K-complex in sleep electroencephalograms," *IEEE Trans. Biomed. Eng.*, vol 17(4), pp 314-23, 1970.

[7] AC. Da Rosa, B. Kemp, T. Paiva Lopes, FH.Da Silva and HA. Kamphuisen, "A model-based detector of vertex waves and K complexes in sleep electroencephalogram ," *Electroencephalogr. Clin. Neurophysiol.*; vol 78(1), pp71-79 1991

[8] Jansen BH., "Artificial neural nets for K-complex detection.," *IEEE Eng Med Biol Mag.*, vol 9(3),pp 50-52, 1990.

[9] O. Sheriff, B. Pagnrek, S. Mamouhd and R. Broughton, "Automatic detection of K-complex in the sleep EEG," *Int. Electrical and Electronic Conf. and Exp.*, vol 81,1977

[10] M. Woertz, T. Miazhynskaia, P. Anderer and G. Dorffner, "Automatic K-complex detection: comparison of two different approaches," *Abstracts of the ESRS*, JSR 13(Suppl.1),1, Prague 2004.

[11] N. Kerkeni, L. Bougrain, M. Hédi Bedoui, F. Alexandre and M. Dogui, "Reconnaissance automatique des grapho-éléments temporels de l'électroencéphalogramme du sommeil," *Traitement et Analyse de l'Information : Méthodes et Applications –TAIMA,* Tunisie, 2007.

[12] N. Bankman, VG. Sigillito, R. A. Wise and P. L. Smith, "Feature-based detection of the K-complex wave in the human electroencephalogram using neural networks," *IEEE transactions on biomedical engineering*, vol. 39, no12, pp. 1305-1310, 1992.

[13] M. G. Frei, R. L. Davidchack and I. Osorio, "Least squares acceleration filtering for the estimation of signal derivatives and sharpness at extrema," *IEEE Transactions on Biomedical Engineering*, vol.46, no.8, 1999.