33rd Annual International Conference of the IEEE EMBS
Boston, Massachusetts USA, August 30 - September 3, 2011

Assessing the impact of signal normalization: preliminary results on epileptic seizure detection

Lojini Logesparan, Alexander J. Casson and Esther Rodriguez-Villegas

***Abstract*— Signal normalization is an essential part of patient independent algorithms, for example to correct for variations in signal amplitude from different parts of the body, prior to applying a fixed threshold for event detection. Multiple methods for providing the required normalization are available. This paper presents a systematic investigation into the effects of five different methods using epileptic seizure detection from the EEG as an illustration case. It is found that, whilst normalization is essential, four of the considered methods actually decrease the ability to detect seizures, counteracting the algorithm aim. For optimal detection performance the effects of the signal normalization illustrated here should be incorporated into future algorithm designs.**

I. INTRODUCTION

Signal normalization is an essential part of patient independent algorithms used for the analysis of physiological signals and the automatic detection of features and salient points. Taking the scalp electroencephalogram (EEG) as an example, the absolute value of the EEG signal can vary widely [1]: with age; between different people; between different parts of the head; and between different subjects states, such as being asleep or awake and during epileptic seizures. Moreover it is possible for the absolute EEG values to vary over time due to changes in the electrical activity of the brain and also due to the varying quality of the electrode connection to the scalp.

To correct for these changes, automated analysis algorithms must utilize normalized, or relative, amplitude values. Here the raw data is corrected by some measure of the average background so that a fixed threshold can be applied during signal classification. There are of course multiple different methods by which the required normalization can be provided. Different techniques can vary in terms of:

- The mathematical function (such as the mean or median) used to calculate the normalization.
- The amount of memory present, that is, the amount of background data used to calculate the normalization.
- Where in the signal processing chain the normalization is applied.

This last option is illustrated in Fig. 1 which shows a generalized seizure detection algorithm. In the top, not-normalized route, input data y is passed to a feature extraction stage which *emphasizes* the features of interest: signal processing is applied such that the *interesting* sections of the input signal are amplified relative to the *non-interesting* sections. The generated signal F is then passed to a classifier such that thresholds can be applied to separate the *interesting* and *non-interesting* sections. To generate a normalized signal N there is then a choice over whether the normalization factor (z) is calculated using the input signal (y) or the feature (F), and whether this factor is used to normalize y or F, illustrated as Routes A, B, and C in Fig. 1.

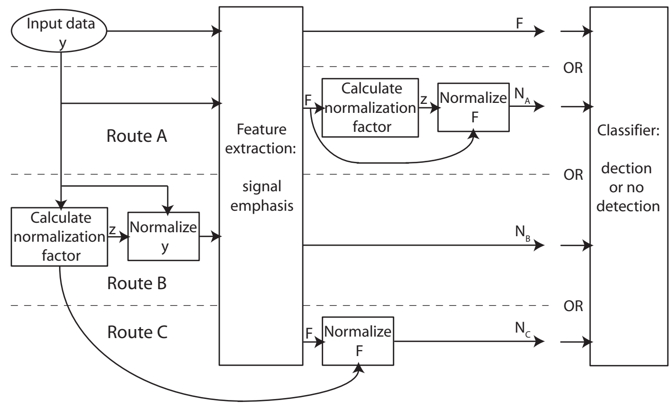


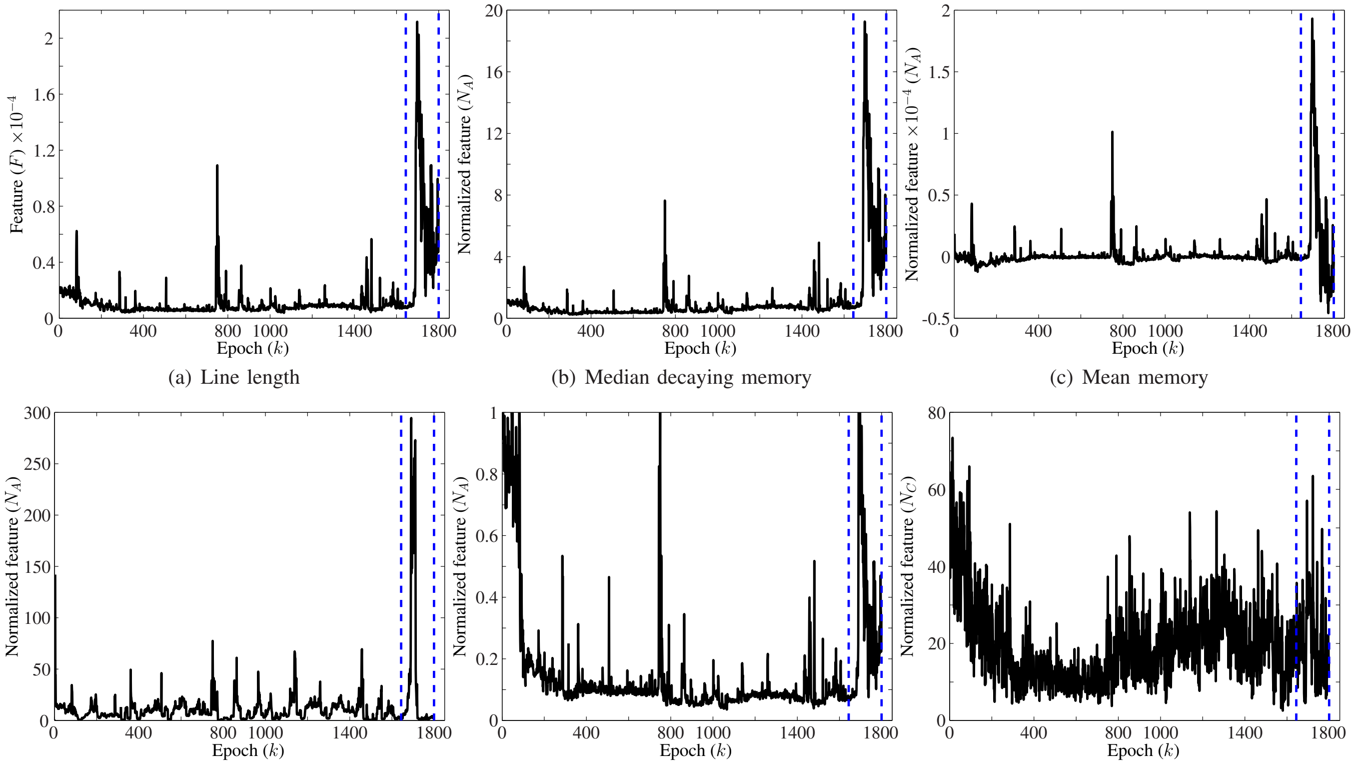
Fig. 1. Signal normalization can be provided with topologies A, B, or C: there is a choice over whether the normalization factor (z) is calculated using the input signal (y) or the feature (F), and whether this is used to normalize y or F.

Regardless of the precise technique used, the key requirement for the normalization is that the raw data is modified to correct for broad level amplitude changes, and that doing this has a minimal effect on the overall algorithm performance. As an example, consider the case of an epileptic seizure detection algorithm [2]. Here F should be large when a seizure is present and small when no seizure is present. However, seizures are often associated with larger raw EEG amplitudes [1]. If the normalization factor z also increases during the seizure the effect of calculating a normalization F/z is that z reduces the effective value of F. Thus, rather than aiding, normalization makes the seizure detection more difficult. [2] suggested using median based normalization, rather than the standard deviation, to overcome this.

To the authors' knowledge, however, no systematic investigation into the impact of multiple different techniques has been considered previously in the literature. This study presents such an investigation for five different normalization techniques used with a simple EEG seizure detection algorithm. The investigation presented is inevitably preliminary only: there are numerous different normalization methods which could be investigated, and the assessment methodology required isn't obvious *a priori*. We present preliminary

The authors are with the Department of Electrical and Electronic Engineering, Imperial College London, UK. Email: {lojini.logesparan04, acasson, e.rodriguez}@imperial.ac.uk.

The research leading to these results has received funding from the European Research Council under the European Community's 7th Framework Programme (FP7/2007-2013) / ERC grant agreement no. 239749.



(a) Line length
(b) Median decaying memory
(c) Mean memory
(d) Standard deviation memory
(e) Peak detector
(f) Signal range

Fig. 2. 1800 epochs of analyzed scalp EEG data. (a) shows the calculated line length feature (1). (b)–(f) show the resulting signals after normalizing the line length by five different methods. The two dashed vertical lines show the start and end of an expert marked seizure.

work for establishing a suitable analysis framework and generate initial results and directions. Also, although it is likely that similar effects are present in a number of fields, to keep the analysis here tractable we only consider the EEG seizure detection case.

II. METHODS

A. Analysis methods

In this study, the feature F to be normalized has been selected as the line length, a low computational complexity, amplitude-dependent, linear feature which is commonly used for seizure detection [3], [4]. The line length of a sampled signal $y(n)$ is calculated as the sum of the instantaneous gradient of the signal [3]:

$$F(k) = \sum_{n} |y(n-1) - y(n)| \tag{1}$$

where n is the sample number within a short epoch of data, and k is the epoch being analyzed. Here, each epoch is generated as a non-overlapping 2 s section of EEG data.

Fig. 2(a) shows the calculated F from 1800 epochs of a single channel of scalp EEG data. An expert marked seizure is present between epochs 1644 and 1800 and as expected the line length increases in these epochs compared to non-seizure (interictal) epochs. The feature has thus successfully *emphasized* the seizure to be detected.

The aim now is to qualitatively and quantitatively assess the impact of different normalization techniques on the line length seizure emphasis shown in Fig. 2(a). Qualitative results are generated by plotting the signal $N(k)$ resulting from normalization of the line length signal given in Fig. 2(a). A quantitative comparison is then provided by calculating $N(k)$ for a set of four EEG recordings each containing one expert marked seizure. All recordings are approximately one hour long with 23 EEG channels. $N(k)$ is calculated separately for each channel and the resulting distribution of all $N(k)$ values in seizure and non-seizure epochs then found.

B. Normalization techniques

In this preliminary study five different normalization techniques, based upon methods previously reported in the literature, are investigated. These five methods are defined as follows. All methods require up to 120 epochs to be present to be calculated, and no seizures are present in the first 120 epochs of the analyzed data.

1) Median decaying memory [5]–[7]: calculated here as $N_A = F/z$ where

$$\begin{aligned} z(k) =& (1-\lambda)\text{median}\{F(k-1)\cdots F(k-120)\} \\ &+ \lambda z(k-1) \end{aligned} \tag{2}$$

and $\lambda = 0.99923$ with initial conditions of $z(1) = F(1)$.

2) Mean memory [3]: calculated as $N_A = F - z$ where

$$z(k) = \text{mean}\{F(k-1)\cdots F(k-120)\} \tag{3}$$

3) Standard deviation memory [5], [6]: calculated here as $N_A = F/z$ where z is the standard deviation of 30 s (15 epochs) of F ending one minute (30 epochs) before the current epoch:

$$z(k) = \text{std}\{F(k-31)\cdots F(k-46)\} \tag{4}$$

4) Peak detector [8]: calculated here as $N_A = F/z$ where

$$z(k) = \begin{cases} F(k) & \text{if } F(k) > z(k-1), \\ z(k-1) & \text{if } F(k) \leq z(k-1). \end{cases} \tag{5}$$

with initial conditions of $z(1) = F(1)$.

5) Signal range [5], [6]: calculated here as $N_C = F/z$ where

$$z(k) = \max_k\{y(n)\} - \min_k\{y(n)\}. \tag{6}$$

III. RESULTS AND DISCUSSION

A. Qualitative results

Fig. 2 (b)-(f) show how the normalized line length ($N(k)$) varies for the five normalization methods used. The median decaying memory (Fig. 2(b)) is seen to preserve the *emphasis* of the seizure with essentially just a change in the amplitude value being provided. In contrast the mean memory (Fig. 2(c)) preserves the emphasis at the start of the seizure, but towards epoch 1800 the values are decreased, making it more difficult to detect the end of the seizure. The standard deviation memory (Fig. 2(d)) acts similarly although it also modifies the artifacts present: the artifact seen in the line length at epoch 750 is reduced in significance, but other smaller artifacts are highlighted. Both the peak detector (Fig. 2(e)) and signal range (Fig. 2(f)) perform relatively poorly, removing the emphasis of the seizure provided by the line length. In these cases it would not be possible to threshold the normalized feature to uniquely detect the epileptic seizure.

B. Quantitative results

Fig. 3 gives box plots demonstrating the distribution of $N(k)$ between seizure epochs (shaded) and non-seizure epochs (non-shaded) for four EEG recordings from different subjects. The boxes represent the 25^{th} percentile (bottom line), median (middle line) and 75^{th} percentile (top line) of the distribution with the maximum and minimum values also shown. The general trend for the feature (or normalized feature) to be increased in seizure epochs is clearly seen.

Ideally, there should be no overlap between the $N(k)$ values in seizure and non-seizure epochs, allowing them to be completely separated for 100% classification accuracy. This is not possible for any of the plots, and instead a trade-off between the sensitivity (fraction of correct classifications) and specificity (fraction of incorrect classifications) must be accepted. Two possible positions for a fixed detection threshold, determining this trade-off, are shown in Fig. 3. Firstly, the green dotted *detect* line is drawn at the lowest percentile of the seizure epochs across patients, and indicates a threshold that would ensure that at least 75% sensitivity is achieved for all subjects. The red dashed *reject* line is drawn at the highest percentile for the non-seizure epochs across patients, and indicates a threshold that would ensure that at least 75% specificity is achieved for all subjects. Ideally the reject line would be below the detect line showing that at least 75% can be achieved for both sensitivity and specificity.

This is only achieved in the median decaying memory case (Fig. 3(b)). This method thus provides both normalization and enhances the detection performance compared to just thresholding the raw feature. The other four methods provide normalization, but the distance between these lines is increased compared to the raw data, indicating that a worse trade-off between sensitivity and specificity will be obtained when using any fixed threshold.

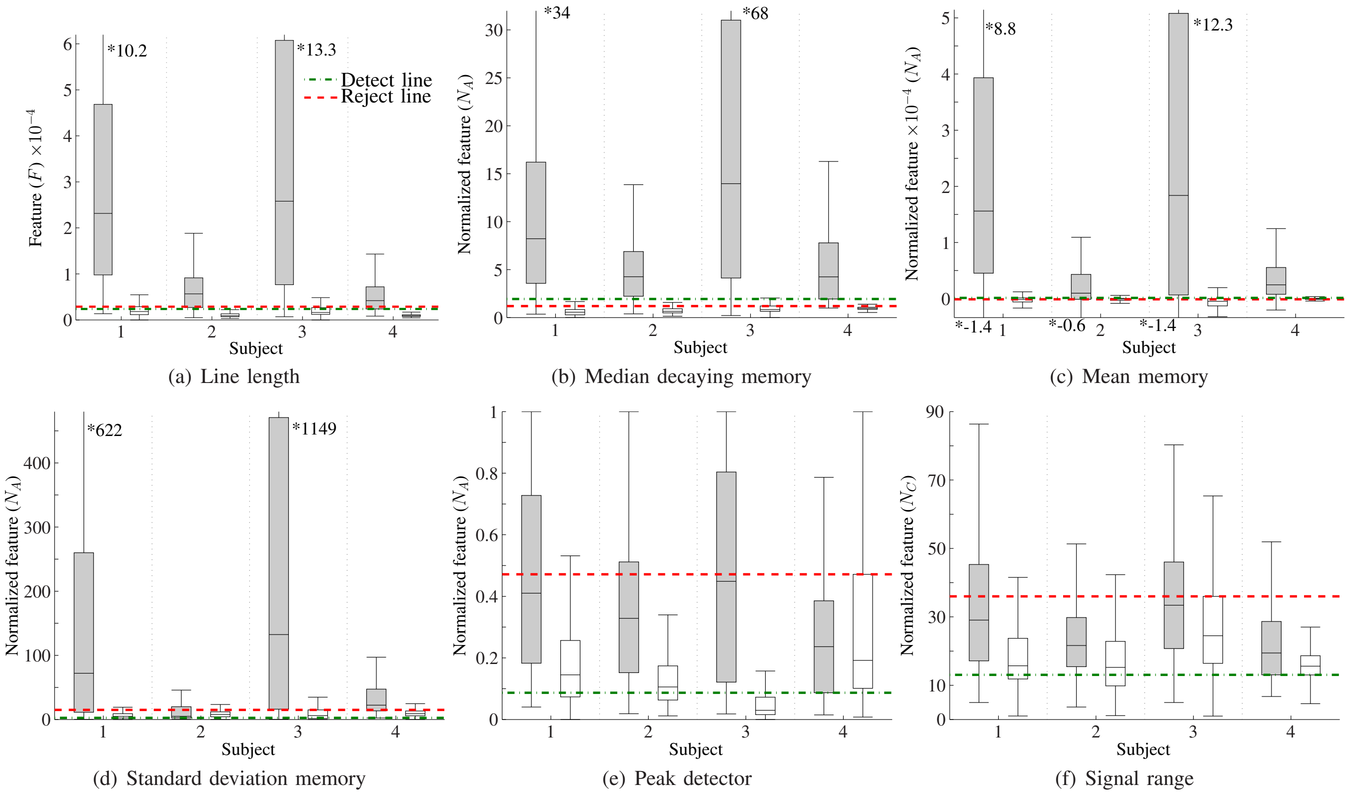
C. Discussion, limitations and future work

The results here clearly indicate that choices are available with regards to the normalization utilized, but normalization methods are not all equal. Most seizure detection algorithms dedicate significant attention to the feature extraction used (the line length here). For example they may consider multiple different transforms, or options within any one method such as the choice of mother wavelet to use in the wavelet transform. The results here demonstrate that for optimal algorithm performance similar attention needs to be given to the normalization used within the algorithm.

The peak detection and signal range methods used here perform comparatively poorly in terms of aiding the emphasis of the seizure over the background data. However both methods guarantee that the normalized feature can only take values within a bounded region—(0,1] for the peak detector. With the other methods considered the normalized values can still in principle take on any value.

Within these other methods, the median decaying memory achieves the best performance as it provides both normalization and aids the separation of seizure and non-seizure epochs. It is noted, however, that although the median decaying memory only needs 120 epochs for it to be calculable, the constant λ from (2) controls how long previous values of z affect the current calculation. $\lambda = 0.99923$ corresponds to a memory *half-life* of approximately 30 minutes and in our experimentation it was found that depending on the choice of initial value for z it can take several hours for the measure to reach a steady-state value. This effect is not reflected in the results here where only one hour EEG long records have been used. It is thus possible that in the analysis of long term continuous recordings the median decaying memory is of a differing utility to that suggested here. The investigation of this is left to future work.

The other calculation methods used here do not include similar long transient effects: they have a maximal 120 epoch transient response. The standard deviation memory method also introduces a one minute delay. By varying these memory factors, for example such that the delay is longer than the duration of any typical seizure, it may be possible to *postpone* changes to the normalization factor z. The changes in the normalization provided would still occur, but not during the seizure itself, potentially making their impact less critical. Such changes in the memory of the normalization method are akin to investigating entirely different normalization techniques and again it is noted that it is not possible to consider all such possibilities in this preliminary work. For example, it would also be of interest to investigate other normalization methods such as the envelope detector [9] and Wilcoxon rank sum test [10], and to assess



(a) Line length (b) Median decaying memory (c) Mean memory

(d) Standard deviation memory (e) Peak detector (f) Signal range

Fig. 3. Box plots showing the distribution of the feature and normalized features in seizure (shaded) and non-seizure (non-shaded) epochs from four subjects. The trend for features to increase during seizure epochs is clearly seen. The green dotted *detect* line indicates a fixed detection threshold that could be used to ensure that at least 75% sensitivity is achieved for all subjects. In contrast the red dashed *reject* line indicates a fixed detection threshold that could be used to ensure that at least 75% specificity is achieved for all subjects. Asterisk (*) numbers indicate the maximum or minimum value of the box plot where it cannot be drawn directly for scaling reasons.

the impact on non-linear signal emphasis features such as those considered in [11].

IV. CONCLUSIONS

Normalization is an essential tool for correcting broad level amplitude differences in recorded signals, for example between different patients, to allow patient independent classification. This paper has systematically investigated five previously reported normalization techniques in terms of their impact on the performance of a simple seizure detection algorithm. All five methods provide signal normalization, but the mean memory, standard deviation memory, peak detector and signal range methods did this at the cost of reducing the detection performance. In contrast, the median decaying memory actually improved the differentiation between seizure and non-seizure epochs.

It is thus clear that in addition to selecting suitable signal processing bases for algorithm development, significant attention must also be given to selecting a suitable signal normalization basis. Preliminary directions for doing this have been provided here, however further work is necessary to determine the effect on non-linear features and using a wider subset of normalization techniques at different stages within the signal processing chain.

REFERENCES

[1] P. E. M. Smith and S. J. Wallace, *Clinicians' guide to epilepsy*. London: Arnold, 2001.

[2] J. J. Halford, "Computerized epileptiform transient detection in the scalp electroencephalogram: Obstacles to progress and the example of computerized ECG interpretation," *Clin. Neurophysiol.*, vol. 120, no. 11, pp. 1909–1915, 2009.

[3] R. Esteller, J. Echauz, T. Tcheng, B. Litt, and B. Pless, "Line length: an efficient feature for seizure onset detection," in *IEEE EMBC*, Istanbul, Oct. 2001, pp. 1707–1710.

[4] K. Patel, C. Chern-Pin, S. Fau, and C. J. Bleakley, "Low power real-time seizure detection for ambulatory EEG," in *Pervasive Health*, London, Apr. 2009, pp. 1–7.

[5] L. Kuhlmann, A. N. Burkitt, M. J. Cook, K. Fuller, D. B. Grayden, L. Seiderer, and I. M. Y. Mareels, "Seizure detection using seizure probability estimation: Comparison of features used to detect seizures," *Ann. Biomed. Eng.*, vol. 37, pp. 2129–2145, 2009.

[6] L. Kuhlmann, M. J. Cook, K. Fuller, D. B. Grayden, A. N. Burkitt, and I. M. Y. Mareels, "Correlation analysis of seizure detection features," in *ISSNIP*, Sydney, Dec. 2008, pp. 309–314.

[7] I. Osorio, M. G. Frei, and S. B. Wilkinson, "Real-time automated detection and quantitative analysis of seizures and short-term prediction of clinical onset," *Epilepsia*, vol. 39, no. 6, pp. 615–627, 1998.

[8] L. Logesparan and E. Rodriguez-Villegas, "Improving phase congruency for EEG data reduction," in *IEEE EMBC*, Buenos Aires, Sep. 2010, pp. 642–645.

[9] A. J. Casson and E. Rodriguez-Villegas, "Toward online data reduction for portable electroencephalography systems in epilepsy," *IEEE Trans. Biomed. Eng.*, vol. 56, no. 12, pp. 2816 –2825, 2009.

[10] R. Meier, H. Dittrich, A. Schulze-Bonhage, and A. Aertsen, "Detecting epileptic seizures in long-term human EEG: A new approach to automatic online and real-time detection and classification of polymorphic seizure patterns," *J. Clin. Neurophysiol.*, vol. 25, no. 3, pp. 119–131, 2008.

[11] C. J. Stam, "Nonlinear dynamical analysis of EEG and MEG: Review of an emerging field," *Clin. Neurophysiol.*, vol. 116, no. 10, pp. 2266–2301, 2005.