Adaptive Estimation of Karhunen-Loève Series Applied to the Study of Ischemic ECG Records

J García, S Olmos, GB Moody[1], RG Mark[1], P Laguna

Centro Politécnico Superior, Universidad de Zaragoza, Spain
[1]Division of Health Sciences and Technology, Harvard-MIT, Cambridge, USA

Abstract

The Karhunen-Loève transform (KLT) has been used as a tool to analyze the repolarization period in the study of ischemic episodes. The dynamic variations in the ST-T complex shape are shown as variations of the kl series associated with the beat series. In this work we propose an adaptive system to estimate the kl coefficients of the ST-T complex in order to improve the signal-to-noise ratio (SNR) of the estimation (it is obtained around 10dB of improvement). A transversal adaptive linear combiner filter using as reference inputs the KL basis functions and as primary input the concatenation of noisy ST-T complexes from consecutive beats is used. It is shown how the weights of the filter become, after convergence, estimates of the kl series. The Least Mean Squares (LMS) and Recursive Least Squares (RLS) algorithms are studied and compared in the kl time series estimation. It is presented a specific initialization for the LMS which leads to the same performance than the RLS.

1. Introduction

The ST-T complex of the ECG reflects the repolarization phase of the cardiac electrical cycle, and ischemia is usually reflected as changes in the ST-T shape. The Karhunen-Loève Transform (KLT) has been used to model the ST-T complex[1] and has permitted to study its dynamic behavior. We present an adaptive filter to estimate the *kl* series of the ST-T complex that reduces the noise incorrelated with the signal. The SNR improvements of the adaptively estimated *kl* series with respect to the direct estimation (with the inner product) are derived. It is shown that an improvement of 10dB can be achieved, with adequated convergence time to follow ischemic episodes, using the adaptive estimation. This improvement is very useful when using the *kl* series to monitor or detect ST-T transient changes that can be associated with ischemic episodes. The LMS and RLS algorithms are detaily studied in order to determine which one presents the best performance in the *kl* series estimation.

2. The Karhunen-Loève Transform

The KLT is an orthogonal signal-dependent linear transform which is optimal in the sense that it concentrates the maximum signal information in the minimum number of parameters and defines the domain where the signal and noise are more separated. It is needed to build a training set which contains the statistical properties of the signals to be analyzed in order to obtain the basis functions of the KL transform field. The KL basis functions for the ST-T complex were derived from a training set of 100.000 preprocessed beats[1]. Once the basis functions have been derived, each ST-T complex of the ECG signal to be analyzed is represented in the KL transform field by a feature vector; the first components (2-4) of this vector represent almost all (70%-90%) the signal energy[1]. In this way there will be as many kl time series, $kl_n(i)$ (where n represents the KL order and i is the beat number), as KL coefficients are needed to represent the ST-T complex. The dynamic variations in the ST-T complex shape are shown as variations of the $kl_n(i)$ series. The direct way to obtain the $kl_n(i)$ series is from the inner product of the KL basis with the ST-T complexes to be analyzed. This leads to a noisy time series, and the adaptive estimation is used to reduce the noise being a more suitable estimation.

3. The Adaptive Filter

The adaptive estimation allows reduction of noise incorrelated with the signal. We propose the adaptive linear filter to estimate the $kl_n(i)$ time series. The KL basis KL_{ik} are the reference inputs to the transversal filter whose primary input d_k is the concatenation of

ST-T complexes from consecutive beats s_k plus noise n_k (Figure 1). The weights after filtering each ST-T complex will result in the estimated kl coefficients.

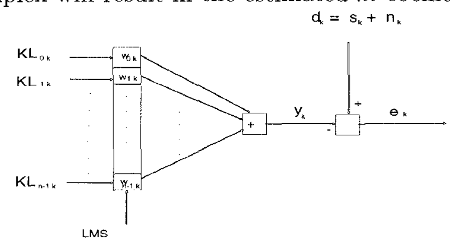

Figure 1: *Adaptive Filter.*

The N samples that compose each ST-T complex are assumed to be the sum of the signal of interest (a deterministic signal component, $s_k = STT_k$) and an incorrelated noise component n_k. If the deterministic component is strictly periodic with a period of N samples, then it satisfies $s_k = s_{k+N}$ for all k. The reference inputs KL_{jk} ($j = 0, ..., n-1$) ($n \leq N$, with n the number of taps in the filter and N the number of samples of each ST-T complex) are formed by concatenating copies of the jth KL basis function; thus $KL_{jk} = KL_{jk+N}$. In the KL vectorial space, $d_k = s_k + n_k$ may be expressed as the sum of all the KL components and the incorrelated noise:

$$d_k = \sum_{j=0}^{N-1} kl_j \, KL_{jk} + n_k \qquad (1)$$

The output of the adaptive filter, y_k, is the signal that we want to be an estimation of s_k, and e_k is the error signal $e_k = s_k + n_k - y_k$ with:

$$y_k = \sum_{j=0}^{n-1} w_{j\ k} \, KL_{jk} \qquad (2)$$

If $\mathbf{KL}_k$ denotes the reference input vector and $\mathbf{W}_k$ the weight vector: $\mathbf{KL}_k = [KL_{0k},, ..., KL_{n-1k}]^T$ and $\mathbf{W}_k = [w_{0k}, ..., w_{n-1k}]^T$, then $y_k = \mathbf{KL}_k^T \, \mathbf{W}_k$. Minimizing the mean squared error $\xi = E[e_k^2]$ using any adaptive algorithm[2], the weight vector converges to the optimal solution $\mathbf{W}^* = \mathbf{R}^{-1}\mathbf{P}$, where $\mathbf{R} = E[\mathbf{KL}_k \, \mathbf{KL}_k^T]$ and $\mathbf{P} = E[d_k \, \mathbf{KL}_k]$. In this case, given the orthonormality conditions of the KL basis and the lack of correlation between the noise n_k and the KL basis $KL_{n\ k}$, $\mathbf{R}$ and $\mathbf{P}$ reduce to $\mathbf{R} = \frac{1}{N}\mathbf{I}$ and $\mathbf{P} = \frac{1}{N}[kl_0, kl_1, ..., kl_{n-1}]^T$, respectively. The optimal weight vector, $\mathbf{W}^*$, that minimizes the mean squared error is given by:

$$\mathbf{W}^* = [kl_0, kl_1, ..., kl_{n-1}]^T \qquad (3)$$

This result means that each weight w_i is an estimate of the ith KL coefficient for s_k. Thus the weight vector is a characterization of the deterministic signal component, and the output signal y_k, in the optimum case, takes the value:

$$y_k = \sum_{j=0}^{n-1} w_i^* \, KL_{j\ k} = \sum_{j=0}^{n-1} kl_j \, KL_{j\ k} \qquad (4)$$

i.e., the projection of s_k onto the subspace spanned by $KL_{j\ k}$ ($i = 0, ..., n-1$) with $n \leq N$. Thus y_k is the nth-order KLT representation of s_k, and $y_k = s_k$ if $n = N$ (i.e., if all of the KL components are included).

The minimum mean squared error, ξ_{min}, will be $\xi_{min} = E[d_k^2] - \mathbf{P}^T\mathbf{W}^*$. Given that the weight vector oscillates around this optimal value, y_k is an unbiased estimate of s_k. The remaining noise due to the misadjustment (M) depends upon the adaptive algorithm used to adjust the weight vector[2]. The elements of the weight vector, evaluated at the end of each ST-T complex, are the adaptive estimates of the KL coefficients of that complex. The quality of the y_k estimation is thus directly related to the quality of the KL estimation. In figure 2, the direct and adaptive (with 10dB of SNR improvement) estimations of $kl_0(i)$, from the lead V3 of the record e0106 from the European ST-T Database, are shown.

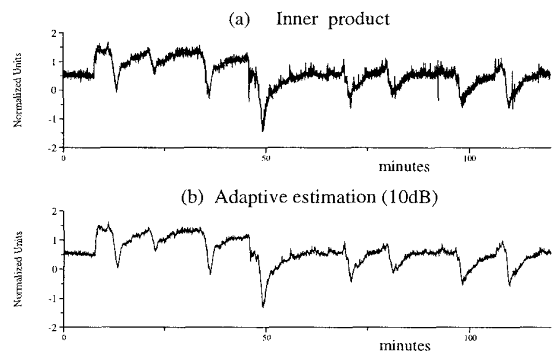

Figure 2: *Estimation of $kl_0(n)$ in record e0106.*

3.1. The LMS Algorithm

First, we have used the LMS algorithm [2]

$$\mathbf{W}_{k+1} = \mathbf{W}_k + 2\mu e_k \, \mathbf{KL}_k \qquad (5)$$

The condition that assures the convergence of the algorithm is [3] : $0 < \mu < \frac{1}{3 \, tr[\mathbf{R}]} = \frac{N}{3n}$. The time constant τ for the convergence of the MSE is:

$$\tau = \frac{1}{4 \, \mu\lambda} = \frac{N}{4\mu} \qquad (6)$$

where $\lambda = \frac{1}{N}$ is the eigenvalue of the matrix $\mathbf{R}$ (all the eigenvalues are identical). τ is expressed in sampling intervals. The gain constant, μ, thus controls the stability and the speed of convergence. Given an appropriate choice of μ, the estimate of the weight vector may be obtained within a single beat ($\tau < N$) if necessary. Thus adaptive filtering may be used in principle even for tracking beat-by-beat ST-T variations. To measure the excess of mean squared error we calculate the misadjustment [2] $M = \frac{Excess\, MSE}{\xi_{min}}$, which for the LMS algorithm can be approximated by:

$$\mathbf{M} \simeq \mu\ tr[\mathbf{R}] = \mu \frac{n}{N} \tag{7}$$

The mean square error $\xi = \xi_{min}(1 + M)$ will be:

$$\xi = \left(\frac{1}{N} \sum_{j=n}^{N-1} kl_j^2 + E[n_k^2] \right) \left(1 + \frac{\mu\, n}{N}\right) \tag{8}$$

The MSE thus depends on the noise power, the power in the ST-T complex not represented by the first n kl_n coefficients, and the gain constant, μ. Note that the dependence on the KLT order n is not evident, since an increase in n value increases the $(1 + \frac{\mu\, n}{N})$ factor and decreases the $\sum_{j=n}^{N-1} kl_j^2$ factor. Thus, the optimum solution minimizes n and maximizes $\sum_{j=0}^{n-1} kl_j^2$; this property is intrinsic to the KLT. Given that at the steady state the estimated signal y_k is orthogonal with the error e_k [2], the $Excess\, MSE$ is the excess of error power introduced in y_k, and the signal-to-noise ratio of this estimation, SNR_y, will be:

$$SNR_y = \frac{\frac{1}{N} \sum_{j=0}^{n-1} kl_j^2}{\left(\frac{\mu\, n}{N}\right) \left(\frac{1}{N} \sum_{i=n}^{N-1} kl_j^2 + E[n_k^2] \right)} \tag{9}$$

If we consider that the ST-T energy is strongly concentrated in the n first coefficients, we can neglect the term $\sum_{j=n}^{N-1} kl_j^2$, obtaining $SNR_y \simeq SNR_d \frac{N}{\mu\, n}$ where SNR_d is the SNR of the original d_k signal. Comparison of this SNR_y with that obtained from the direct estimation of $kl_n(i)$ will give the SNR improvement ($\triangle SNR$) achieved by the adaptive system. Direct $kl_n(i)$ estimation yields a SNR (SNR_y^{direct}) that can be estimated if we assume the noise is white and then its PSD is uniformly distributed in the KL domain $SNR_y^{direct} \simeq SNR_d \frac{N}{n}$. Thus the SNR improvement obtained using the adaptive filter is:

$$\triangle SNR_{LMS} = \frac{SNR_y}{SNR_y^{direct}} = \frac{1}{\mu} \tag{10}$$

Thus we find that, for appropriately chosen values of μ ($\mu < 1$), the adaptive estimate of $kl_n(i)$ is cleaner than the $kl_n(i)$ time series obtained from the inner product. It is possible to get a $\triangle SNR$ of 10dB ($\mu = 0.1$) with $\tau = 2.5$ beats. The choice of μ obviously involves the typical trade-off between SNR improvement and rate of convergence and is limited by the need to track changes occurring within a few beats in typical cases.

3.2. The RLS Algorithm

The aim of RLS algorithm is to obtain, at each time k, a multiple linear regression model of the inputs and desired responses of the filter up to this time, in a recursive way. Formally, we want to minimize a cost function $\mathcal{E}(k)$ which exponentially weights the differences between the desired response $d(i)$ and filter output. It it expressed as $\mathcal{E}(k) = \sum_{i=1}^{k} \lambda^{k-i} \left[d(i) - \mathbf{W}^T(k)\mathbf{X}(i)\right]^2$, where λ is a *forgetting factor*, $\mathbf{X}(i)$ represents the KL basis function vector, and $\mathbf{W}(k)$ is a weight vector adapted to minimize $\mathcal{E}(k)$. The RLS algorithm produces after convergence an unbiased estimation of the desired signal [4]. The exponential nature of the estimators, i.e., the finite window effect leads to an *estimation-noise* which results in a misadjustment $\mathcal{M}$ of the output signal from its optimal setting [4]:

$$\mathcal{M} = \frac{1 - \lambda}{1 + \lambda} n \tag{11}$$

where n is the number of taps (basis functions) in the adaptive filter. In nonstationary environments a $\lambda < 1$ value is selected so that the algorithm presents finite memory and tracking capability. This choice results in a convergence time of [4]: $\tau = \frac{1}{1-\lambda}$. Following a similar analysis than with LMS, we get the improvement in SNR using RLS algorithm versus direct estimation:

$$\triangle SNR_{RLS} = \frac{SNR_y}{SNR_y^{direct}} = \frac{1}{N} \frac{1 + \lambda}{1 - \lambda} \tag{12}$$

When λ approaches to unity the improvement $\triangle SNR$ will be better, but the time constant τ will be greater.

4. Comparison Between RLS and LMS

The improvement of the RLS versus the LMS is obtained to compare the efficiency of the two algorithms:

$$\triangle SNR_{RLSvsLMS} = \frac{\triangle SNR_{RLS}}{\triangle SNR_{LMS}} = \frac{\mu}{N} \frac{1 + \lambda}{1 - \lambda} \tag{13}$$

In figure 3 the iso-improvement curves as function of its features parameters (μ (gain constant in LMS) and λ (forgetting factor in RLS)) can be seen: the equal $\triangle SNR$ curves correspond to sampling rates of 1000, 360 and 250 Hz, i.e. to N = 600, 216 and 150 samples in the ST-T complex, respectively. The region above each curve corresponds to a better improvement in SNR produced by the RLS algorithm and the region

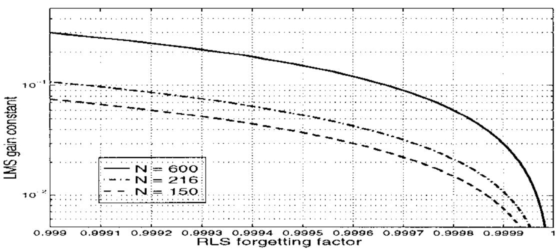

Figure 3: *Equal ΔSNR curves.*

below conversely. Because of the KL eigenvectors are orthonormal and all eigenvalues of the correlation matrix are equal the LMS algorithm will show the same time constant for each tap and the convergence rate will not be limited by the smallest eigenvalue. Thus, the RLS algorithm will not represent a great improvement over the LMS performance. To obtain the same SNR improvement it is necessary to select μ and λ related accordingly to (13) and this leads roughly (λ near 1) to the same rate of convergence for both algorithms:

$$M_{LMS} = \mu \frac{n}{N} = N \frac{1-\lambda}{1+\lambda} \frac{n}{N} = M_{RLS} \qquad (14)$$

$$\tau_{LMS} = \frac{N}{2\mu} = \frac{1}{2} \frac{1+\lambda}{1-\lambda} \approx \tau_{RLS} \qquad (15)$$

So their performances become equal looking at the adaptive trade-off between ΔSNR and τ but with higher simplicity for the LMS. However the initial convergence rate is much better in RLS algorithm [4] as it can be seen in the kl_0 estimated on the record e0103 of the *European ST-T Database* (figures 4(b) and 4(c)) for equivalent λ and μ values.

5. Ischemic ECG Analysis

In figure 4 it can be seen the kl_0 series corresponding to the lead V4 of the record e0103 from the European ST-T Database estimated by different methods. The adaptive estimation shows a large SNR improvement: 19dB ($\mu = 0.012, \lambda = 0.99984$). The LMS estimation can lead to errors in the initial detection of ischemic episodes as it is shown comparing figures 4(b) and 4(c) where there are two episodes with the first one underestimated with the LMS because of the convergence time. In the rest of the record the signal tracking is essentially the same with the LMS and RLS. Because of roundoff errors, when $\lambda < 1$ is used, the RLS algorithm becomes unstable and and need to be periodically restarted. It has been proposed an *ad-hoc* initialization for the LMS weights (instead to zero, to the inner product of the first ST-T complex with each basis function) which gives a better initial convergence rate. The result (fig. 4(d)) is that the same tracking properties are obtained with the new initializated LMS than RLS. Thus, this initializated LMS is the best suited for the estimation.

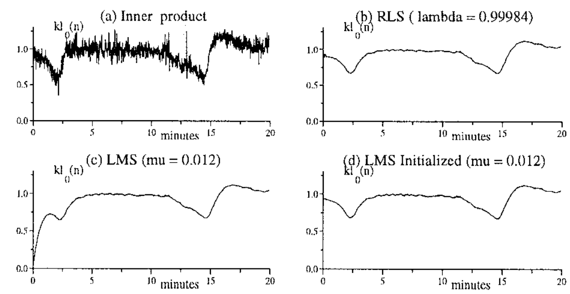

Figure 4: *Estimation of $kl_0(n)$ in record e0103.*

6. Conclusions

The adaptive estimation of $kl_n(i)$ series has been studied in the analysis of the dynamic ST-T complex behavior. $kl_n(i)$ series are better estimated in the adaptive way, with a significant SNR improvement. A transversal linear combiner filter has been used achieving 10dB of SNR improvement with an adequate convergence time ($\tau = 2.5$ beats). The performances of LMS and RLS algorithms have been studied and a specific initialization for the LMS algorithm has been presented. This results in the best suited for the *kl* series estimation.

Acknowledgements

This work was supported by grant TIC94-0608:C02-02, from CICYT, and by PIT 06/93 from DGA.

References

[1] P. Laguna, G. B. Moody, and R. G. Mark, "Analysis of the cardiac repolarization period using the KL transform: Aplications on the ST-T database.", in *Computers in Cardiology*. IEEE Computer Society Press, 1994, pp. 233–236.

[2] B. Widrow and S. D. Stearns, *Adaptive Signal Proccessing*, Prentice-Hall, Englewood Cliffs, New Jersey, 1985.

[3] A. Feuer and E. Weinstein, "Convergence analysis of LMS filters with uncorrelated Gaussian data", *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 33, pp. 222–230, 1985.

[4] Evangelos Eleftheriou and David D. Falconer, "Tracking properties and steady-state performance of RLS adaptive filter algorithms", *IEEE Trans. Acous. Sig. Proc.*, vol. 34, no. 5, pp. 1097–1109, October 1986.

Address for correspondence:
José García
Dep. Ing. Eléctrica, Electrónica y Comunicaciones
Maria de Luna 3. 50015-Zaragoza (SPAIN)
E-mail: josechu@tsc1.cps.unizar.es