COMPUTERS AND BIOMEDICAL RESEARCH **31,** 305–322 (1998)
ARTICLE NO. CO981483

Detection of Transient ST Segment Episodes During Ambulatory ECG Monitoring

Franc Jager,* George B. Moody,† and Roger G. Mark†

**Faculty of Computer and Information Science, University of Ljubljana, 1001 Ljubljana, Slovenia; and †Division of Health Sciences and Technology, Harvard–Massachusetts Institute of Technology, Cambridge, Massachusetts 02139*

Received April 28, 1998

Using the European Society of Cardiology ST-T Database, we have developed a Karhunen–Loève transform-based algorithm for robust automated detection of transient ST segment episodes during ambulatory ECG monitoring. We review current approaches and systems to detect transient ST segment changes and describe the architecture of our algorithm and its development. The algorithm incorporates a single-scan trajectory-recognition technique in feature space using the Mahalanobis distance function between the feature vectors. The main characteristics of the algorithm are detection of noisy beats, correction of the reference ST segment level to correct for slow ST level drift, detection of sudden significant shifts of ST deviation due to shifts of the mean electrical axis of the heart, detection of transient ST episodes, and, by tracking the QRS complex morphology, differentiation between ischemic and nonischemic ST episodes as a result of axis shifts. We compared the algorithm's performance to other recently developed algorithms and estimated its real-world performance.

INTRODUCTION

Investigations have shown that monitoring ST segment changes during ambulatory electrocardiographic (AECG) monitoring is as valuable as stress testing in the detection of coronary artery disease (*1*). No statistically significant differences in mortality were observed between patients with silent ischemia and patients with angina pectoris (*2*). Accurate assessment of the extent of myocardial ischemia requires long-term monitoring while the patient engages in normal daily activities. Reliable automated methods for detecting transient ST segment episodes compatible with ischemia (ischemic ST episodes) are required.

ST changes may also be observed in a variety of contexts other than ischemia, including ventricular hypertrophy, conduction abnormalities, hyperventilation, electrolyte abnormalities, response to medication, mitral valve prolapse, and response to temperature changes. Since such changes may cause significant shifts ($>100\ \mu$V) in the ST segment deviation level, which may mimic real ischemic

0010-4809/98 $25.00

episodes, it is important to identify these effects. The most troublesome of these nonischemic ST episodes are position-related (postural) changes in the electrical axis of the heart (axis shifts) and very slow but significant drift of the ST level provoked by the effects of medication on repolarization, slow (nonpostural) changes in the cardiac electrical axis, or the effects of heart-rate changes on repolarization. Such changes complicate both automated and manual detection of true ischemic episodes. Reliable ST analyzers should be sophisticated enough to distinguish these nonischemic ST changes from clinically significant ischemic ST changes.

Only a few previous computer systems have dealt with nonischemic ST changes. Early work on computer-analysis methodology for retrospectively identifying the beginning and end of ischemic ST episodes was based on plotting electrocardiographic variables in a high-resolution trend format (*3*). Postural changes were identified as those showing an abrupt onset of change in ST level together with an abrupt onset of change in *R*- and *S*-wave amplitude. One important automated detection system for detecting ST change episodes (*4*) included a multipass recognition algorithm that analyzes time-domain ST segment deviation functions and discriminates between ''stable'' and ''unstable'' ST segment baseline time periods. Unstable periods were analyzed in order to identify significant ischemic ST segment deviations, while the reference ST segment level was corrected for nonischemic shifts between stable segments.

Until recently, there was no standard test material to be used as a basis for quantitative, replicable tests of ST analyzers. The arrival of the European Society of Cardiology ST-T Database (ESC DB) (*5*) offered an important new resource. It was created to provide a standard set of material for assessing the quality of algorithms to detect ischemic ST episodes and to quantify ST deviations. In recent years, five major techniques to detect transient ST segment changes have been investigated: traditional time-domain approaches (*6*, *7*) Karhunen–Loève transform (KLT) approaches (*8–12*), neural networks (*13–16*), a combination of the neural network and KLT techniques (*17*), and fuzzy-logic approaches (*18*, *19*). Time-domain systems (*6*, *7*) use the ST segment deviation function calculated as the magnitude of the ST segment vector determined from two ECG leads. The detection procedure in (*7*) applies a considerable preprocessing phase and smoothing of the time series by nonlinear filters. In the detection of ST change episodes, both time-domain systems apply timing and amplitude criteria with two thresholds as used by the human annotators of the ESC DB. Karhunen–Loève (KL) coefficients were first used in (*20*, *21*) in order to represent and visually identify acute ischemic episodes. Our approach, representing ST segment waveshape and robustly detecting transient ST segment episodes in two ECG leads using the KL coefficients, was first presented in (*8*, *9*). Consecutive ST segment feature vectors are classified as normal or deviating in KL feature space. Later, an approach representing the overall ST-T interval using KL coefficients was proposed (*10*) and used to detect ischemia by incorporating filtered and differentiated KL-coefficient time series (*11*). In order to improve the signal-to-noise ratio of the estimation of the KL coefficients in this approach, an adaptive

estimation was also proposed (*12*). Neural network-based systems operate on a single ECG lead. Those classifying ST segments as normal or ischemic incorporate a back-propagation algorithm (*13*), a back-propagation algorithm with adaptive learning rate (*14*), or a counterpropagation algorithm (*15*). A sequence of ischemic ST segments (typically 30 s in duration) forms an ischemic ST episode. Another back-propagation neural network-based system, implementing the three-layer feedforward paradigm, processes the input of absolute ST amplitude and ST slope or ST amplitude only (*16*). An ischemic ST episode is present if the output of the network stays higher than a threshold for a time longer than 20 s. Combination of a neural network and principal components was described in (*17*). This system uses a back-propagation neural network (*16*) which was trained using the KL coefficient time series obtained by the basis functions derived in (*10*). An ST change episode is detected if the duration of deviating samples is longer than 15 s. Only a few systems automatically deal with slow drift of the reference ST segment level and detect shifts of the mean electrical axis responsible for nonischemic ST episodes (*4*, *6–9*). A fuzzy-logic approach to detecting ischemic ST episodes (*18*, *19*) has been introduced in a different way by mimicking the best possible way in which the human expert operates during the evaluation of the episodes, resulting in a classification of detected ischemic ST episodes. The degree of certainty for each time instant considered within an ischemic ST episode is determined. Each detected episode is compared to all previous ones, with the result that two episodes can be similar, related, or different.

This paper describes a two-channel ST detection algorithm based on an orthonormal function transform approach using principal components to detect transient ST segment changes during AECG or during inpatient monitoring. The algorithm incorporates a single-scan trajectory-recognition technique in KL feature space. In detecting ST episodes, it follows the scheme of timing and amplitude criteria defined by the developers of the ESC DB and differentiates between ischemic and nonischemic ST episodes. For the evaluation of performance and comparison of the algorithm to other systems, we used specific performance measures described in (*8*, *9*, *22*, *23*).

METHODS

Reference Material

The goal of the European Society of Cardiology ST-T Database was to achieve a comprehensive standard reference for assessing the quality of AECG analyzers. The database consists of 90 two-channel ambulatory ECG excerpts sampled at 250 samples/s, each 2 h in length and with known or suspected ischemia. The database contains 379 episodes of ST change (368 ischemic and 11 nonischemic), annotated in each lead separately. An ST change episode is present if an absolute ST segment deviation of 100 μV from the reference value lasts at least 30 s. The beginning and end of each ST segment change episode are defined at points of absolute deviation of 50 μV from the reference value. Successive episodes are

separated only if there is a baseline interval lasting at least 30 s. ST levels were measured 80 ms after the J point (or 60 ms after the J point if heart rate exceeded 120 bpm). Evaluation of the two-channel ST detector requires the combination of the reference ST annotations from the two leads into one reference annotation stream using the logical OR function. This combining yields a total of 250 ischemic and 9 nonischemic ST episodes.

Overview of Recognition Algorithm

Our development of the recognition algorithm was restricted to a real-time fully automatic single-scan strategy with minimum decision latencies. The algorithm was developed in terms of efficiency, accuracy, and robustness using all 90 records of the ESC DB. A general flowchart is shown in Fig. 1. The algorithm was devised as a postprocessor to the Aristotle arrhythmia detector (*24*) and operates simultaneously on two ECG leads. In the preprocessing step (module 1) performs baseline correction using cubic spline approximation and subtraction technique, low-pass filtering using a six-pole Butterworth filter with a cutoff frequency of 55 Hz, and exclusion of abnormal beats (as determined by Aristotle) and their neighbors. In module 2, the ST segment and QRS complex KL feature vectors are extracted from the ST and QRS pattern vectors of each single beat, and noisy beats are detected in the feature space. During a short learning period, an average beat is composed from the first 50 normal beats, following Aristotle's learning phase and using time-domain techniques (*6*). The purpose of this average beat is to reliably represent the morphology of normal beats to derive the first ST segment and the first QRS complex KL feature vector. The filtered sequences of feature vectors are resampled to be equidistant and are further smoothed (module 3). Correction of the reference ST segment level (module 4), detection of significant axis shifts (module 5), and detection of ST change episodes (module 6) are fully implemented in KL feature space incorporating an ST segment and QRS complex-feature trajectory-recognition technique. The algorithm identifies a nonischemic ST episode (module 7) if the detected ST change episode is accompanied by a significant axis shift at its beginning or end; otherwise the episode is judged to be ischemic.

Feature Selection

The KLT (*25*) is an operation through which a nonperiodic random process can be expanded over a series of orthonormal functions with uncorrelated features. The KL expansion does not require detailed knowledge of the probability structure of the problem and is based on the mutually orthogonal eigenvectors belonging to eigenvalues of the covariance matrix associated with the pattern vectors in descending order. A pattern vector may be represented as a linear combination of KL basis functions; the components of such a feature vector are KL coefficients. The KLT is optimal in the sense that the least mean square error obtained by approximating a given pattern of a random process using any given number of the first few coefficients is minimal in comparison to other

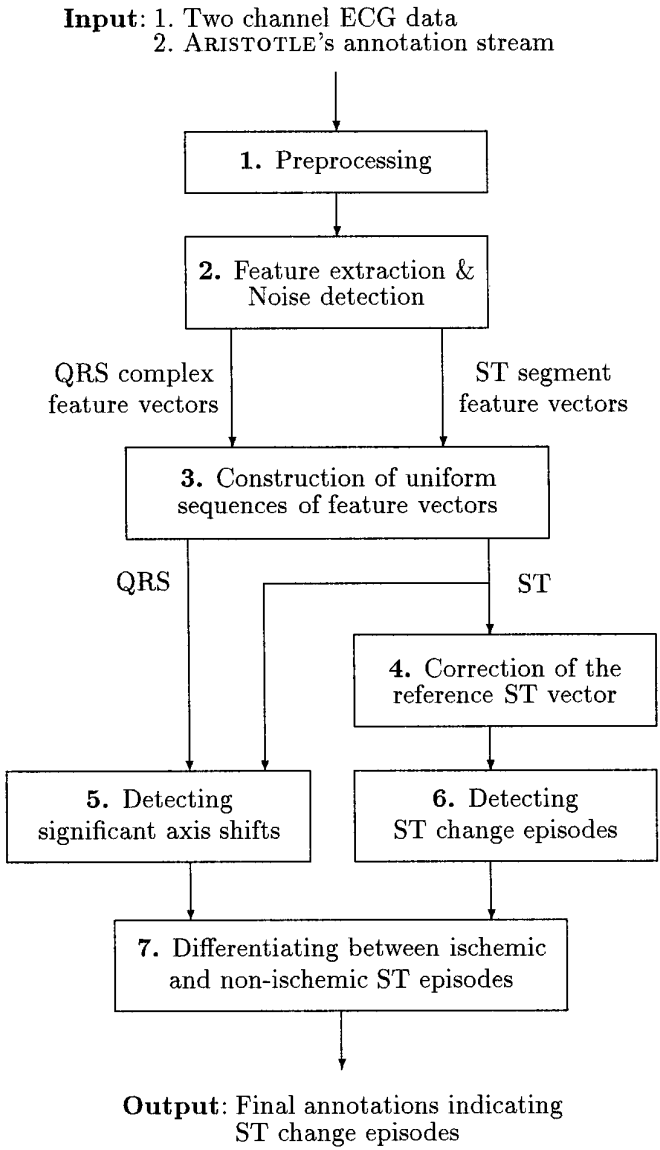

FIG. 1. General flowchart of the KLT-based ST change detection algorithm.

suboptimal transforms. The residual error of a truncated KLT is an optimal estimate of noise content.

To derive the ST segment and QRS complex basis functions, we used all 90 records of the ESC DB. We derived a robust measure of the distribution of pattern vectors $\{\mathbf{x}\}$: robust covariance matrices with outliers excluded. Pattern

vectors which belong to the modeled distribution are expected to be well represented, while the expected value of the residual error of outliers will be high. Abnormal beats (classified by the Aristotle arrhythmia detector (*24*)), their neighbors, and noisy beats (detected by time-domain procedures (*6*)) were discarded. The remaining 744,800 heartbeats were corrected to the isoelectric level using a time-domain procedure to locate the position of the PQ segment (*6*) and were used for computation of both covariance matrices,

$$\mathbf{R} = \sum_{i=1}^{L} p(\omega_i) E\{(\bar{\mathbf{x}}_i - \boldsymbol{\mu}_0))(\bar{\mathbf{x}}_i - \boldsymbol{\mu}_0)^T\}. \quad [1]$$

We considered all intervals of the records containing ischemic or nonischemic ST episodes, and intervals with no significant ST deviation, as separate classes of pattern vectors and calculated their means, $\bar{\mathbf{x}}_i$. All patterns were centralized by subtracting the mean vector, $\boldsymbol{\mu}_0$, obtained over all classes, thus utilizing spatial distribution of pattern classes. The number of pattern classes, L, is 605 and 120 for ST segment and QRS complex covariance matrix, and $p(\omega_i)$ is the a priori probability of the occurrence of the ith class mean. Each set of KL basis functions was derived from 16 discrete samples, taken at 8-ms intervals over each of the two ECG leads. ST segment basis functions start at the point 40 ms after the Aristotle fiducial point (FP) and end at the point FP + 160 ms. The QRS complex basis functions set covers the interval from FP − 96 ms to FP + 24 ms.

Estimation of Dimensionality of Feature Vectors

The first five coefficients of the ST and QRS basis function sets contribute approximately 97 and 93%, respectively, to the total expected power of approximation. To estimate the optimal number of ST segment and QRS complex KL coefficients for distinguishing between noisy and nonnoisy patterns, we used a method originally introduced in (*26*). As an estimate of separability, we used histograms of the normalized residual errors, $r_{\mathbf{x}}(N) = \|\tilde{\mathbf{x}}_N - \mathbf{x}\|/\|\mathbf{x}\|$, of a truncated KLT for noisy and nonnoisy patterns, where $\tilde{\mathbf{x}}_N$ is an approximated pattern. The residual errors for both groups of patterns decrease as the dimensionality of approximation, N, increases. Average residual errors, $\bar{r}(N)$, for nonnoisy patterns are lower than those for noisy ones. Table 1 summarizes the differences between the average residual errors for noisy and nonnoisy ST segment and QRS complex patterns of the ESC DB. The differences increase for small values of N, but after that decrease beyond $N = 5$. We conclude that 5 KL coefficients for ST segment and 5 for QRS complex are sufficient and necessary for separating between noisy and nonnoisy patterns and for accurately representing the features. The dimensionality of 5 for QRS complex is the same as previously estimated (*26*). The trajectory-recognition procedures of the algorithm use the Mahalanobis distance measure, d, between feature vectors. All components of feature vectors are scaled by the corresponding eigenvalues so that their standard deviation ζ is 1. Subtle features of pattern vectors corresponding to basis functions with lower standard deviations are emphasized, and waveforms with different mor-

TABLE 1

Differences between Average Residual Errors for Noisy and Nonnoisy ST Segment and QRS Complex Patterns

N	1	2	3	4	5	6	7	8	
$\Delta\bar{r}_{H}^{(st)}(N)$	0	1	2	7	8	7	7	6	[%]
$\Delta\bar{r}_{A}^{(st)}(N)$	3	7	10	13	14	14	14	12	[%]
$\Delta\bar{r}_{A}^{(qrs)}(N)$	1	1	2	2	3	2	2	2	[%]

Note. N—dimensionality of approximation; $\Delta\bar{r}_{H}^{(st)}(N)$—noisy patterns were annotated by human annotators of the ESC BD; $\Delta\bar{r}_{A}^{(st)}(N)$, $\Delta\bar{r}_{A}^{(qrs)}(N)$—annotator was time-domain algorithm for noise detection (*6*).

phologies are expected to be well discriminated. We assume that the features are independent and normally distributed, so that d^2 is distributed as χ^2_N, i.e., as a χ^2 random variable with $N(5)$ degrees of freedom. In the present case, the expected mean value, m, of d^2 is $N\zeta^2 = 5$, and its expected standard deviation, σ, is $\sqrt{2N}\ \zeta^2 \approx 3.16$. Some architecture parameters and thresholds of the algorithm are expressed in terms of this expected mean and standard deviation.

Noise Detection

In module 2, the algorithm derives a five-dimensional KL feature vector for the ST segment, $\mathbf{s}(i)$, and another five-dimensional feature vector for the QRS complex, $\mathbf{q}(i)$, for each single isoelectric-corrected (*6*) heartbeat. The KL representations of noisy beats are expected to have relatively large residual errors and to differ markedly from those of neighboring beats. Based on these observations, the algorithm considers beat i noisy if the normalized residual error for the ST segment, $r_s(i)$, or for QRS complex, $r_q(i)$, exceeds a certain percentage,

$$(r_s(i) > T_r) \vee (r_q(i) > T_r), \qquad [2]$$

where T_r is the residual threshold, $T_r = 25\%$, or if the ST or QRS feature vector differs sufficiently from those of the past few beats,

$$d^2(\mathbf{s}(i), \tilde{\mathbf{s}}(i)) > \Theta \vee d^2(\mathbf{q}(i), \tilde{\mathbf{q}}(i)) > \Theta, \qquad [3]$$

where $\tilde{\mathbf{s}}(i)$ and $\tilde{\mathbf{q}}(i)$ are the mean feature vectors representing the recursively updated dominant beat morphology for the last 15 normal beats, and Θ is a distance threshold (8.16, i.e., $m + 1\sigma$). We empirically set both thresholds in such a way as to achieve high specificity according to human annotations about noisy beats in the ESC DB, while removing as few heart beats as possible.

Characterization of Feature-Vector Time Series

In module 3, sequences of KL feature vectors are smoothed (15-point moving average), resampled at a constant rate of 0.5 samples/s, and further smoothed (9-point moving average). Distance functions of the first order for ST segment,

$F_s(k)$, and QRS complex, $F_q(k)$, form the basis for the feature trajectory recognition technique

$$F_s(k) = d(\mathbf{s}(k), \mathbf{s}(1))$$
$$F_q(k) = d(\mathbf{q}(k), \mathbf{q}(1)), \quad [4]$$

where $\mathbf{s}(k)$ and $\mathbf{q}(k)$ are ST segment and QRS complex feature vectors, k is the sample number, and $\mathbf{s}(1)$ and $\mathbf{q}(1)$ are the first ST and QRS feature vectors. We characterized trend plots of several time-domain metrics and of KL coefficient representation of ST segment and QRS complex data for the records of the ESC DB. Visualizing features in trend plot form permits us to display information globally and is better at revealing significant events than visual examination of raw signals. Figure 2 shows trend plots for record e0601. During true ischemic ST episodes (see Fig. 2, lead 1) the time series of parameters describing the QRS complex shape do not change significantly, and the ST segment deviation function varies smoothly to form a triangular shape. The characteristics of an axis shift (see Fig. 2, lead 0) are sudden step changes in some or all QRS-related parameters, which are stable before and after the shift. Axis shifts typically last from 30 s to more than 2 min. When an axis shift occurs, the mean electrical axis changes its direction. A shift in the mean electrical axis may or may not cause a shift in ST segment level. We defined axis shifts as significant when they were associated with significant concurrent changes of ST segment deviation level with an absolute value of $>100\ \mu$V. An axis shift interval which starts and/or ends in a significant axis shift defines a nonischemic ST change episode. During nonischemic ST episodes, the ST segment level remains stable and forms a rectangular shape with a maximum ST deviation of $<300\ \mu$V. Significant ($>100\ \mu$V) slow drift of ST segment level (see Fig. 2, lead 1, prior to and during second ischemic episode) is similar to a real ischemic episode but is generally long in duration and of maximum ST deviation of $<200\ \mu$V.

Correction of the Reference ST Segment Level

Feature samples of the ST time series may be divided in general into those forming a class of ''normal''ST segments and those far from this class. Correction of the reference ST segment level (module 4) to account for slow drift is performed by updating the mean ST feature vector of normal ST segments in the ST segment feature trajectory, $\mathbf{s}(k)$, using the Mahalanobis distance of second order, d^2, as a measure between feature vectors. The center of the class of normal ST segments, $\bar{\mathbf{s}}(k)$, is updated if a newly observed ST feature vector, $\mathbf{s}(k)$, is ''close enough.'' After each newly observed current ST feature vector, $\mathbf{s}(k+1)$, tentative new values, $\mathbf{S}(k+1)$, for the new mean ST feature vector, $\bar{\mathbf{s}}(k+1)$, are determined. This tentative new mean ST feature vector is determined as an exponentially weighted sum of the current mean ST feature vector, $\bar{\mathbf{s}}(k)$, and the current ST feature vector, $\mathbf{s}(k+1)$,

$$\mathbf{S}(k+1) = \frac{1}{\alpha}((\alpha - 1)\bar{\mathbf{s}}(k) + \mathbf{s}(k+1)), \quad [5]$$

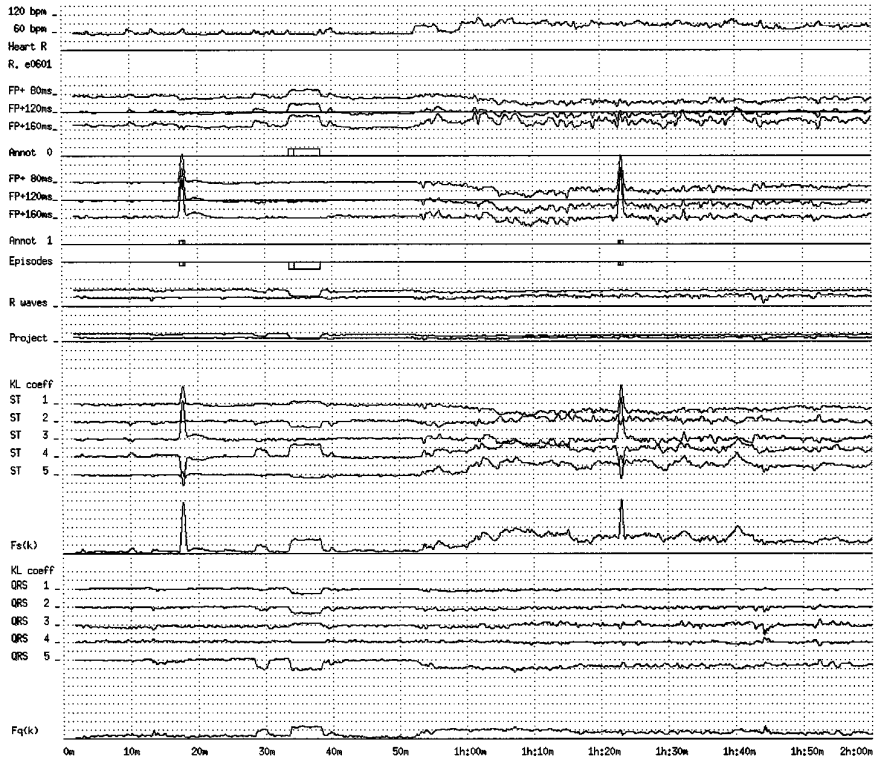

FIG. 2. Trend plots of record e0601 of the ESC DB. From top to bottom: heart rate, ST segment deviations for both leads at 80, 120, and 160 ms after the Aristotle fiducial point (resolution, 100 μV) with the corresponding reference ST episode annotation streams (record contains one nonischemic ST episode in lead 0 and two ischemic ST episodes in lead 1); combined reference ST annotation stream in the sense of logical OR function; *R*-wave amplitudes and projections of the mean electrical axis to both lead axes (resolution, 1 mV); and time series of the first 5 KL coefficient representation of ST segments and QRS complexes (resolution, 1 ζ; eigenvalue, λ_i) with corresponding distance measures, $F_s(k)$ and $F_q(k)$, when using the first 5 KL coefficients (resolution, 1 ζ).

due to the nonstationariness of the ST feature trajectory, where α is the time constant. The updating procedure fits the specific nature of the ST segment feature trajectory. The updated estimate of the mean ST feature vector, $\mathbf{S}(k+1)$ is taken as the new mean ST feature vector, $\bar{\mathbf{s}}(k+1)$, provided two conditions are satisfied. Due to the nature of the slow drift which does not cause extreme excursions of ST segment feature vectors over the feature space, the new mean ST feature vector has to be sufficiently close to the initial ST feature vector,

$$d^2(\mathbf{S}(k+1), \mathbf{s}(1)) < \Psi. \qquad [6]$$

A newly observed current ST feature vector, $\mathbf{s}(k+1)$, has to be close to the

current mean ST feature vector, $\bar{\mathbf{s}}(k)$, or closer to the initial feature vector, $\mathbf{s}(1)$, than the current mean,

$$d^2(\mathbf{s}(k+1), \bar{\mathbf{s}}(k)) \leq \Phi(k+1) \vee d^2(\mathbf{s}(k+1), \mathbf{s}(1)) < d^2(\bar{\mathbf{s}}(k), \mathbf{s}(1)). \quad [7]$$

The left rule ensures that the residual, $\mathbf{s}(k+1) - \bar{\mathbf{s}}(k)$ (drawn from Eq. [5]), of the exponentially weighted mean, which is proportional to the value of the current feature vector, cannot be high or have a large influence on the next position of the mean feature vector. The right rule ensures that current ST feature vectors which are closer to the initial ST feature vector (they are likely to be there) than the current mean should be used for updating, even if the mean ST feature vector is far from the initial ST feature vector. Since there is no other fixed and established class in the ST feature space, only excursions of ST feature vectors over the feature space, updating is dependent on predefined feature-space boundaries. Time constants and feature-space boundaries were set empirically. The time constant α was set to 450 (15 min), while the thresholds Ψ and $\Phi(k+1) = \Phi_1$ were set to 11.32 $(m + 2\sigma)$ and 1.02 $((m + \sigma)/8)$, respectively. After an extreme ST episode or after an ST episode ended in slow drift of the reference ST segment level, the current ST feature vector may be far from the mean ST feature vector. In such periods, the ST feature vector used for updating should be taken from the wider neighborhood of the mean ST feature vector. Whenever the distance function, $d^2(\mathbf{s}(k+1), \mathbf{s}(1))$, exceeds $m + 3\sigma$ or $m + 9\sigma$, $\Phi(k+1)$ is increased for the following $\beta = 900$ samples (30 min) to $\Phi_2 = 3.62$ $((m + 3\sigma)/4)$ or to $\Phi_3 = 8.36$ $((m + 9\sigma)/4)$, respectively. Correcting the reference ST segment level for record e0704 of the ESC DB containing severe slow drift is shown in the center part of Fig. 3. Switching of $\Phi(k)$ may be observed.

Detection of Significant Axis Shifts

The underlying architecture of the axis-shift detection procedure (module 5) is a bandpass differentiation of the ST segment and QRS complex distance functions of the first order, $F_s(k)$ and $F_q(k)$ (Eq. [4]). Searching for significant axis shifts in the first-order distance functions is more straightforward. The operator used to compute the derivative of the distance functions behaves like a bandpass filter. High frequencies are attenuated (rejecting spikes and noises), while the characteristic of the low-frequency region of the filter's transfer function is approximately linear (detecting step changes). An additional requirement to detect significant axis shifts is that there be "flat" intervals of distance function before and after the step change. A step change in either function is characterized by an interval during which the function is approximately constant or flat (its mean absolute deviation from its mean value must be less than $\Sigma_1 = 0.33$ over an interval of $N_F = 108$ samples, i.e., 216 s), followed by an interval during which it changes significantly (the N_M sample moving average value, $N_M = 37$ samples, must change by at least $\Sigma_2 = 1.11$ within the next $N_S = 37$ samples, i.e., 74 s), followed by an interval during which the function is again approximately constant

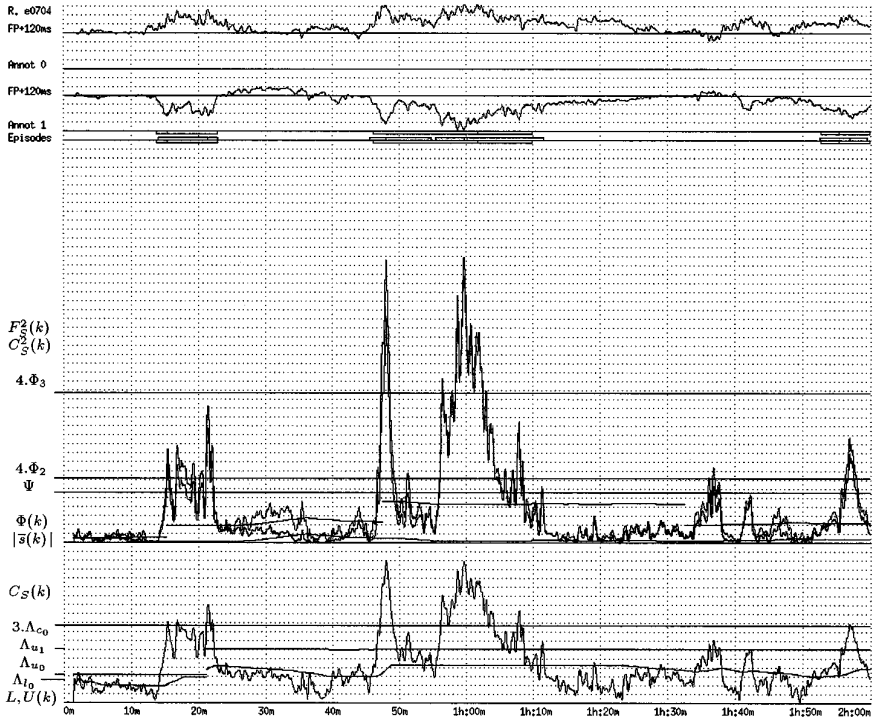

FIG. 3. Correction of reference ST segment level and detection of ST change episodes for record e0704 of the ESC DB. Top: ST segment deviation levels for both leads at 120 ms after the Aristotle fiducial point (resolution, 100 μV) with the corresponding reference ST episode annotation streams. Combined reference ST annotation stream is along and below the line "Episodes." Above this line are algorithm-annotated ST episodes. Middle: ST segment distance function of the second order, $F_S^2(k)$ (top), corrected distance function, $C_S^2(k)$ (bottom), piecewise continuous threshold $\Phi(k)$, and projected mean ST feature vector (magnitude), $|\bar{\mathbf{s}}(k)|$, (resolution, 2ζ). Bottom: The corrected ST segment distance function of the first order, $C_S(k)$, and piecewise continuous lower, $L(k)$, and upper, $U(k)$, decision thresholds (resolution, 0.5 ζ).

or flat (defined as for the first flat interval). The thresholds were estimated empirically from trend plots of the functions. Significant axis shift is considered in progress whenever a step change is present simultaneously in the ST segment and QRS distance function.

Detection of ST Change Episodes

ST change episodes are detected (module 6) by sequentially classifying samples of the first-order ST segment distance function after correction of the reference ST segment level,

$$C_s(k) = d(\mathbf{s}(k), \bar{\mathbf{s}}(k)), \qquad [8]$$

as normal and deviating. The distance function of the first order is convenient for applying time-domain human-defined criteria for identifying ST change episodes. Consecutive samples of $C_s(k)$ which are "far" from the class of normal ST segments form ST episodes. There is only one known class, i.e., class of normal ST segments, so the classification depends on predefined decision thresholds. The mean and variance of feature vectors $C_s(k)$ differ from record to record. The projected feature space ST trajectory, $C_s(k)$, shows significant inter- and intrarecord variability, and its distribution is not known a priori. Sequences of normal ST segments represent stationary intervals, which in fact are only piecewise stationary due to noisy periods which cause bumps and ripples in the $C_s(k)$, while sequences of deviating ST segments show continuous and slowly varied amplitude, and thus represent time-dependent or nonstationary intervals (ST change episodes) which differ considerably in their duration and in the magnitude. These observations led us to devise a procedure which incorporates adaptive decision thresholds. We followed timing and amplitude criteria for identification of ST change episodes in the raw ECG signals (*5*) as defined by the developers of the ESC DB (see Reference Material). We strictly retained timing criteria with two thresholds sequentially classifying ST feature vectors and identifying the exact beginning, peak extrema, and exact end of ST episodes, while the thresholds, lower $L(k)$ (equivalent to 50 μV) and upper $U(k)$ (equivalent to 100 μV), are adaptive. Higher adaptive decision thresholds along bumps or ripples of $C_s(k)$ result in the avoidance of false episode detections. Since the samples of $C_s(k)$ which are "close" to the class of normal ST segments can be classified definitely as normal, and the samples which are far definitely as deviating, the decision thresholds are adaptive only within a predefined "guard zone." The adaptive decision thresholds are driven by a heuristic descriptor of trend, i.e., exponentially weighted mean of the corrected ST distance function,

$$\overline{C}_s(k) = \frac{1}{\gamma}((\gamma - 1)\overline{C}_s(k-1) + C_s(k)), \qquad [9]$$

estimating nonstationary behavior of $C_s(k)$. Both adaptive decision thresholds, $L(k)$ and $U(k)$, are defined as

$$\begin{aligned} L(k) &= \overline{C}_s(k) \\ U(k) &= 2 \cdot \overline{C}_s(k). \end{aligned} \qquad [10]$$

The palette of ST episodes varies from those showing just slight excursions of the ST distance function up to those showing extreme excursions. Therefore the upper bound of the guard zone, $\Lambda_u(k)$, is also adaptive, permitting adaptive decision thresholds to reach higher values, thus easily avoiding bumps and ripples of the $C_s(k)$ in the intervals containing extreme ST episodes, while the lower bound of the guard zone, Λ_{l_0}, is fixed. The upper bound is driven by a heuristic rule involving the trend, $\overline{C}_s(k)$, and instantaneous distance function, $C_s(k)$, and switches between values Λ_{u_0} and Λ_{u_1},

$$\Lambda_u(k) = \begin{cases} \Lambda_{u_1} = 2\Lambda_{c_0} + \frac{1}{2}\Lambda_{w_0} & \text{if } (\overline{C}_s(k) > \Lambda_{c_0}) \wedge (C_s(k) > 3 \cdot \Lambda_{c_0}) \\ \Lambda_{u_0} & \text{if } 2 \cdot \overline{C}_s(k) \le \Lambda_{c_0} \\ \Lambda_u(k-1) & \text{otherwise,} \end{cases} \quad [11]$$

where Λ_{c_0} is the center of the narrower guard zone and Λ_{w_0} is its width. Whenever $\overline{C}_s(k)$ crosses Λ_{c_0} and the value of the corrected ST distance function, $C_s(k)$, itself crosses $3 \cdot \Lambda_{c_0}$, the guard zone is widened; whenever the quantity $2 \cdot \overline{C}_s(k)$ falls below Λ_{c_0}, the guard zone is reset to its initial width. Outside the guard zone, both decision thresholds are defined by the guard zone boundaries. $L(k)$ is bounded by the center of the guard zone, $\Lambda_c(k) = (\Lambda_{l_0} + \Lambda_u(k))/2$, and the quantity $\frac{2}{3} \cdot \Lambda_{c_0}$ and $U(k)$ by the upper, $\Lambda_u(k)$, and lower, Λ_{l_0}, bounds of the guard zone. The development of this module is described in Appendix. The lower part of Fig. 3 shows an example of detecting extreme ST change episodes in the presence of bumps and ripples of the corrected ST distance function in the record e0704.

Differentiation of Ischemic and Nonischemic ST Episodes

In module 7, a detected ST episode is considered nonischemic if a significant axis shift is in progress anywhere within the interval of $N_A = 75$ samples (2.5 min) surrounding the beginning or end of the ST episode and if the peak ST deviation during the episode does not exceed 300 μV in either lead.

EVALUATION

Since the total number of ST episodes in the ESC DB is small, the records were not divided into development and test sets. To estimate the performance of the algorithm, we used performance matrices assessing its ability to differentiate ischemic and nonischemic ST episodes (*8*, *9*) and standard performance measures of sensitivity and positive predictivity for detection of ischemic ST episodes and ischemic ST durations (*9*, *22*, *23*). To assess real-world performance, we used the bootstrap statistical technique (*27*). Table 2 summarizes the raw performance obtained and expected ''real-world'' performance. The algorithm correctly detected (see sensitivity matrix) 213 ischemic ST episodes and missed 37 (34 + 3) ischemic episodes. Among the algorithm-annotated ischemic ST episodes (see positive predictivity matrix), 206 were actually ischemic episodes and 33 (31 + 2) were falsely detected. Of the 9 nonischemic episodes, 7 were correctly detected and 2 were detected but judged as ischemic. Of the 15 nonischemic episodes annotated by the algorithm, 6 were actually nonischemic, 3 were actually ischemic, and 6 were falsely detected. Gross and average ischemic ST episode sensitivity, IE Se, and positive predictivity, IE +P, were over 85% and over 86%. Ischemic ST duration sensitivity, ID Se, and positive predictivity, ID +P, were over 75% and over 74%. The 5% confidence limits for the minimum expected performance when using 10,000 bootstrap trials were as follows: IE Se and IE +P were over 80%, while ID Se and ID +P were close to or greater than 70%.

TABLE 2

Report on Performance of the KLT-Based ST Change Detection Algorithm Using the ESC DB

	Se matrix			+P matrix		
	KLT algorithm			KLT algorithm		
	Isch	Non-is	No dev	Isch	Non-is	No dev
Reference						
Isch	213	3	34	206	3	—
Non-is	2	7	0	2	6	—
No dev	—	—	—	31	6	—

Note. Se—sensitivity; +P—positive predictivity; TP_S—true positives for sensitivity; FN—false negatives; TP_P—true positives for positive predictivity; FP—false positives; IE Se, IE +P—ischemic ST episode sensitivity and positive predictivity; ID Se, ID +P—ischemic ST duration sensitivity and positive predictivity; [*g*]—gross; [*a*]—average.
Ischemic ST episode detection (raw statistics):

$TP_S = 213$, $FN = 37$, $TP_P = 206$, $FP = 33$
IE Se = 85.2% [*g*], 87.1% [*a*] IE +P = 86.2% [*g*], 87.7% [*a*]
ID Se = 75.8% [*g*], 78.2% [*a*] ID +P = 78.0% [*g*], 74.1% [*a*].

Bootstrap (lowest expected performance—5% confidence limits on 10,000 trials):

IE Se = 80.6% [*g*], 82.2% [*a*] IE +P = 81.1% [*g*], 82.9% [*a*]
ID Se = 69.6% [*g*], 73.2% [*a*] ID +P = 72.5% [*g*], 69.3% [*a*].

Our previously developed time-domain ST detection algorithm (*6*) showed gross IE Se and IE +P of 83 and 84%. The reported average IE Se and IE +P of the other time-domain system (*7*) was 84 and 85%. The KLT-based ischemia detector (*11*) which incorporates filtered and differentiated KLT-coefficient time series yielded sensitivity in detecting ischemia of 66% and specificity[1] of 51%. Single-lead ischemia detection systems using a back-propagation neural network algorithm (*13*) and a counterpropagation algorithm (*15*) were evaluated using only part of the ESC DB. The back-propagation neural network algorithm with adaptive learning rate (*14*) yielded IE Se and IE +P of 88.6 and 78.4%. A three-layer feedforward paradigm using a back-propagation algorithm (*16*) revealed sensitivity and positive predictive accuracy in detecting ischemic ST episodes of about 80 and 90%. A combination of a neural network and principal components (*17*) resulted in about 80% for sensitivity and positive predictive accuracy in detecting ischemic ST episodes. These two systems (*16*, *17*) were evaluated using

[1] The particular performance measures (*8*, *9*, *22*, *23*) are based on the concept of match between reference- and algorithm-annotated ST episodes when assessing the ability to detect ST episodes (events have a time dimension). Therefore, the concept of specificity (the fraction of rejections which are correct) is not applicable, since the number of true negatives is undefined. However, in assessing the ability to detect duration of ischemia, each ST segment or sample of the feature time series is considered a dimensionless event, thus the concept of specificity could be used.

a nonstandard matching test between the reference and the algorithm-annotated ST episodes.

DISCUSSION AND CONCLUSION

We developed a new KLT-based algorithm to detect transient ST segment episodes from the surface ECG. The KLT-based approach was effective in rejecting noise resulting in smooth ST deviation functions, thus retaining useful information necessary for analysis. The novel features of the algorithm incorporating a single-scan trajectory-recognition technique are noise detection in KL feature space, correction of the reference ST segment level, detection of significant axis shifts, and differentiation between ischemic and nonischemic ST change episodes. Strong points of our algorithm are efficient correction of the reference ST segment level resulting in good ischemic ST duration sensitivity and positive predictivity and robust performance which does not vary significantly over a range of input data. The bootstrap statistical test showed robustness, i.e., relatively narrow performance distribution between raw performance statistics and the 5% confidence limits estimating the minimum expected performance in the real world. Robust algorithms are preferred since they are less likely to fail completely, even though it is likely that the robustness is achieved at the expense of performance.

Previous systems to detect ischemic ST episodes, e.g., (*4*), have been evaluated on a wide variety of databases (or not at all) and have used a variety of performance measures. Thus, their performances cannot be compared. At the present time, ST episode detection systems may be evaluated using the ESC DB and the performance metrics described in (*22*, *23*), thus facilitating the development and comparison of ST detection systems. The performance of our algorithm is comparable to the best of other reported systems which were evaluated using standard methods. Our system is unique, however, in its capability to differentiate between ischemic and nonischemic ST episodes. Furthermore, bootstrap analysis predicts that our system should show robust real-world performance.

We conclude that future generations of ST analyzers must include procedures for correcting the reference ST segment level and for detecting significant axis shifts in order to differentiate ischemic from nonischemic ST episodes. The records of the ESC DB are of 2-h duration and the entire database contains only nine annotated nonischemic ST episodes and another three events which seem to be mixed episodes (a nonischemic episode containing ischemic ones). Development of more advanced ST detection algorithms will require more extensive research databases, ideally consisting of 24-h ambulatory records containing a rich assortment of ischemic and nonischemic ST episodes (*28*). Even with access to richer research databases, however, it may be difficult to achieve major improvements in the performance of single-pass (real-time) algorithms. Tiny ischemic ST episodes are often hidden in global ST events, ST segment drifts, or noise, and their detection may require the use of multipass algorithms capable of utilizing much wider contextual information.

APPENDIX

Development of the procedure to detect ST change episodes (module 6) consisted of three steps (*9*). In the first step, we divided the ESC DB roughly into two sets, one containing records with extreme ST episodes (set 1) and the other containing the remaining records (set 2). By optimizing the decision thresholds for each set independently, we expected better overall performance. To identify extrema ST episodes, we used a heuristic rule detecting extrema excursions of instantaneous distance, $C_s(k)$,

$$(\overline{C}_s(k) > \vartheta_0) \wedge (C_s(k) > 3 \cdot \vartheta_0), \qquad [12]$$

where $\overline{C}_s(k)$ is the trend (Eq. [9]), ϑ_0 is the trend threshold, and k is the sample number. Any change of $C_s(k)$ given a record satisfying this rule was considered an extreme episode. Records of the ESC DB containing at least one such extreme episode were removed. The trend threshold ϑ_0 was set to 1.0 and time constant γ of trend, $\overline{C}_s(k)$, to 450 samples (15 min). The number of removed records (set 1) was 33. After dividing the database, we set constant values for decision thresholds, $U(k) = 2 \cdot L(k) = \vartheta_0$, and employed the detector performance characteristic curve of gross IE Se versus IE +P for set 2 by varying ϑ_0 to obtain the optimal decision threshold, ϑ_1, satisfying the adjusting constraint of approximately equal IE Se and IE +P. Since Eq. [12] serves in the final algorithm for switching the upper bound of the guard zone in the presence/absence of the extreme ST episodes (Eq. [11]), and since it is convenient to keep the rest of the decision boundaries and thresholds equal for both sets, it is necessary to obtain the trend threshold ϑ_0 and the optimal decision threshold ϑ_1 for set 2, which are equal. For this reason, the process of dividing the database was performed in several iterations. If $|\vartheta_1 - \vartheta_0| > \varepsilon$, we repeated the process of dividing the database by setting $\vartheta_0 = \vartheta_1$. We ended the process of dividing after three repetitions with 27 rejected records and with $\vartheta_1 = 1.49$. Gross ID Se and ID +P showed that the lower decision threshold $L(k)$ has a large influence on these performances and has to be bounded. This step showed that $L(k) \approx \frac{2}{3} \cdot \vartheta_1$ is acceptable. For set 1, we set $U(k) = 2 \cdot L(k) = \vartheta_2 > \vartheta_1$ and, following approximately equal gross IE Se and IE +P as adjusting criteria, ended in $\vartheta_2 \approx \frac{3}{2} \cdot \vartheta_1$. The upper decision threshold, $U(k)$, revealed a large influence on these ST episode performance measures.

In the second step of developing the procedure, we employed a guard zone for both sets. The upper center of the guard zone, Λ_{c_1}, was set to ϑ_2 and the lower, Λ_{c_0}, to ϑ_1. The relationship between the centers of the guard zone, $\Lambda_{c_1} = \frac{3}{2} \cdot \Lambda_{c_0}$, the lowest possible value for the lower decision threshold, $\frac{2}{3} \cdot \Lambda_{c_0}$, and the time constant, γ, were retained from the previous step, and the lower bound of the guard, Λ_{l_0}, was assumed to be fixed. Within the guard zone, both decision thresholds were left adaptive, $U(k) = 2 \cdot \overline{C}_s(k)$ and $L(k) = \overline{C}_s(k)$. Employing the detector performance characteristic curve of gross IE Se versus IE +P for each set of records by modifying the width of the guard zone, Λ_w, while retaining the initial center of the guard zone, $\Lambda_{c_0} = \vartheta_1$, showed that the width

of the guard zone (set 2) may be $\Lambda_{w_0} = 0.4$ and the upper bound of the widened guard zone (set 1) $\Lambda_{u_1} = 2 \cdot \Lambda_{c_0} + \frac{1}{2} \cdot \Lambda_{w_0}$. Furthermore, using the detection performance characteristic curve of gross ID Se versus ID +P for both sets, it was shown that the highest possible value for the $L(k)$ may be constrained by Λ_c. Thus, the entire architecture can be described by two parameters, the initial center of the guard zone, Λ_{c_0}, and its initial width, Λ_{w_0}. Average performance improvement in this second step of employing the guard zone was about 3%.

In the third step, we kept the architecture and employed the detector performance characteristic of gross IE Se versus IE +P by modifying Λ_{c_0} and Λ_{w_0}. The adjusting constraint used was approximately equal IE Se and IE +P. We ended in $\Lambda_{c_0} = 1.46$ and $\Lambda_{w_0} = 0.32$. Average performance improvement of this step was about 1%.

ACKNOWLEDGMENTS

This work was supported in part by the Ministry of Science and Technology of the Republic of Slovenia by providing a scholarship to F.J. and by the Massachusetts Institute of Technology.

REFERENCES

1. Quuyumi, A., Crake, T., Wright, C., Mockus, L., and Fox, K. The role of ambulatory ST-segment monitoring in the diagnosis of coronary artery disease: Comparison with exercise testing and thallium scintigraphy. *Eur. Heart J.* **8,** 124 (1987).
2. Detry, J. M. R., Robert, A., Luwaert, R. J., Melin, J. A., Brohet, C. R., DeKock, M., D'Hondt, A. M., Dewael, C., and Vanbutsele, R. Prognostic value of angina pectoris during exercise test in patient with myocardial ischemia. *In* "Computers in Cardiology 87," pp. 321–324. IEEE Press, Washington, DC, 1988.
3. Gallino, A., Chierchia, S., Smith, G., Croom, M., Morgan, M., Marchesi, C., and Maseri, A. Computer system for analysis of ST segment changes on 24 hour Holter monitor tapes: Comparison with other available systems. *J. Am. Coll. Cardiol.* **4,** 245 (1984).
4. Shook, T. L., Valvo, V., Hubelbank, M., Feldman, C. L., and Stone, P. H. Validation of a new algorithm for detection and quantification of ischemic ST segment changes during ambulatory electrocardiography. *In* "Computers in Cardiology 87," pp. 57–62. IEEE Press, Washington, DC, 1988.
5. Taddei, A., Distante, G., Emdim, M., Pisani, P., Moody, G. B., Zeelenberg, C., and Marchesi, C. The European ST-T database: Standard for evaluating systems for the analysis of ST-T changes in ambulatory electrocardiography. *Eur. Heart J.* **13,** 1164 (1992).
6. Jager, F., Mark, R. G., and Moody, G. B. Analysis of transient ST segment changes during ambulatory ECG monitoring. *In* "Computers in Cardiology 91," pp. 453–456. IEEE Press, Los Alamitos, CA, 1991.
7. Taddei, A., Costantino, G., Silipo, R., Emdin, M., and Marchesi, C. A system for the detection of ischemic episodes in ambulatory ECG. *In* "Computers in Cardiology 95," pp. 705–708. IEEE Press, Piscataway, NJ, 1995.
8. Jager, F., Mark, R. G., Moody, G. B., and Divjak, S. Analysis of transient ST segment changes during ambulatory ECG monitoring using the Karhunen–Loève transform. *In* "Computers in Cardiology 92," pp. 691–694. IEEE Press, Los Alamitos, CA, 1992.
9. Jager, F. Automated detection of transient ST-segment changes during ambulatory ECG-monitoring, Ph.D. thesis. University of Ljubljana, 1994.
10. Laguna, P., Moody, G. B., and Mark, R. G. Analysis of the cardiac repolarization period using the KL transform: Applications on the ST-T database. *In* "Computers in Cardiology 94," pp. 233–236. IEEE Press, Los Alamitos, CA, 1995.

11. Laguna, P. Model-based estimation of cardiovascular features. *In* ''Analysis of Non-invasive Cardiovascular Signals in Ischemic Heart Disease: Methods, Development and Evaluation,'' pp. 20. CNR Inst. of Clinical Physiology, Pisa, 1996.
12. Garcia, J., Olmos, S., Moody, G. B., Mark, R. G., and Laguna, P. Adaptive estimation of Karhunen–Loève series applied to the study of ischemic ECG records. *In* ''Computers in Cardiology 96,'' pp. 249–252. IEEE Press, Piscataway, NJ, 1996.
13. Stamkopoulos, T., Strintzis, M., Pappas, C., and Maglaveras, N. One-lead ischemia detection using a new backpropagation algorithm and the European ST-T database. *In* ''Computers in Cardiology 92,'' pp. 663–666. IEEE Press, Los Alamitos, CA, 1992.
14. Maglaveras, N. Neural networks in ischemia detection and ECG processing. *In* ''Analysis of Non-invasive Cardiovascular Signals in Ischemic Heart Disease: Methods, Development and Evaluation,'' pp. 22. CNR Inst. of Clinical Physiology, Pisa, 1996.
15. Strintzis, M. G., Stalidis, G., Magnisalis, X., and Maglaveras, N. Use of neural networks for electrocardiogram (ECG) feature extraction recognition and classification. *Neural Network World* **3–4,** 313 (1992).
16. Silipo, R., Taddei, A., and Marchesi, C. Continuous monitoring and detection of ST-T changes in ischemic patients. *In* ''Computers in Cardiology 94,'' pp. 225–228. IEEE Press, Los Alamitos, CA, 1995.
17. Silipo, R., Laguna, P., Marchesi, C., and Mark, R. G. ST-T segment change recognition using artificial neural networks and principal component analysis. *In* ''Computers in Cardiology 95,'' pp. 213–216. IEEE Press, Piscataway, NJ, 1995.
18. Presedo, J., Vila, J., Delgado, M., Barro, S., Palacios, F., and Ruiz, R. A proposal for the fuzzy evaluation of ischaemic episodes. *In* ''Computers in Cardiology 95,'' pp. 709–712. IEEE Press, Piscataway, NJ, 1995.
19. Presedo, J. Fuzzy-logic approach to ischemia analysis. *In* ''Analysis of Noninvasive Cardiovascular Signals in Ischemic Heart Disease: Methods, Development and Evaluation,'' pp. 21. CNR Inst. of Clinical Physiology, Pisa, 1996.
20. Marchesi, C., Varanini, M., and Guidi, M. Reliable identification of acute episodes during ECG and hemodynamic monitoring in ICU. *In* ''Computers in Cardiology 80,'' pp. 315–318. IEEE Press, Los Angeles, CA, 1981.
21. Buemi, M., Macerata, A., Taddei, A., Varanini, M., Emdin, M., and Marchesi, C. Monitoring patient status through principal components analysis. *In* ''Computers in Cardiology 91,'' pp. 385–388. IEEE Press, Los Alamitos, CA, 1991.
22. Jager, F., Moody, G. B., Taddei, A., and Mark, R. G. Performance measures for algorithms to detect transient ischemic ST segment changes. *In* ''Computers in Cardiology 91,'' pp. 369–372. IEEE Press, Los Alamitos, CA, 1991.
23. American National Standard Institute. Ambulatory electrocardiographs. ANSI/AAMI EC38-1994 American National Standard Institute, Inc., Association for the Advancement of Medical Instrumentation, 1994.
24. Moody, G. B., and Mark, R. G. Development and evaluation of a 2-lead ECG analysis program. *In* ''Computers in Cardiology 82,'' pp. 39–44. IEEE Press, Los Angeles, CA, 1982.
25. Tou, J. T., and Gonzales, R. C. ''Pattern Recognition Principles.'' Addison–Wesley, Reading, MA, 1974.
26. Moody, G. B., and Mark, R. G. QRS morphology representation and noise estimation using the Karhunen–Loève transform. *In* ''Computers in Cardiology 89,'' pp. 269–272. IEEE Press, Los Alamitos, CA, 1990.
27. Albrecht, P., Moody, G. B., and Mark, R. G. Use of the ''bootstrap'' to assess the robustness of the performance statistics of the arrhythmia detector. *J. Ambulatory Monitoring* **1**(2), 171 (1988).
28. Jager, F., Moody, G. B., Taddei, A., Antolič, G., Zabukovec, M., Škrjanc, M., Emdin, M., and Mark, R. G. Development of a long term database for assessing the performance of transient ischemia detectors. *In* ''Computers in Cardiology 96,'' pp. 481–484. IEEE Press, Piscataway, NJ, 1996.