77CH 1220 - 3CS

In Two Volumes

Volume 1 of 2

and

Volume 2 of 2

PROCEEDINGS OF THE 1977 JOINT AUTOMATIC CONTROL CONFERENCE

SPONSORING ORGANIZATIONS

American Institute of Chemical Engineers
American Society of Mechanical Engineers
Institute of Electrical and Electronics Engineers
Instrument Society of America
Society of Manufacturing Engineers

PARTICIPATING ORGANIZATIONS

American Institute of Aeronautics and Astronautics
Society for Industrial and Applied Mathematics
Technical Association of the Pulp and Paper Industry

WP25 - 4:10

A PROTOTYPE SYSTEM FOR AUTOMATED INTERPRETATION OF VECTORCARDIOGRAMS

Donald E. Gustafson
Scientific Systems, Inc.
Cambridge, Massachusetts

A. S. Willsky S. K. Mitter, A. Akant,
J-Y Wang, P. Doerschuk
Massachusetts Institute of Technology
Cambridge, Massachusetts

M. E. Womble, M. C. Lancaster,
J. H. Triebwasser
USAF School of Aerospace Medicine
San Antonio, Texas

ABSTRACT

A new approach to computer interpretation of vectorcardiograms (VCG's) has been developed using tools from the theory of statistical signal analysis. The system consists of a number of program modules, including a preprocessor, waveform detector, rhythm analyzer, waveform morphology feature extractor and pattern recognizer. A prototype system has been developed, implemented and tested and is presently operational at the USAF School of Aerospace Medicine.

I. INTRODUCTION

Computerized interpretation of electrocardiograms (ECG's) and vectorcardiograms (VCG's) is a problem which has been studied for almost two decades. Development of computer programs began in the early 1960's and has progressed to the stage where nine major programs are available for routine clinical use in the United States. Volume has grown to the point where currently at least four million ECG's are processed annually at about 2400 sites [1]. However, even in light of the extensive effort expended and the results obtained in this area, it is fair to say that significant problems remain to be solved. One of these problems is the improvement of diagnostic accuracy, both in morphology and rhythm analysis. Present programs must be overread, so are designed for low miss probabilities, but in turn have false alarm probabilities which are generally excessive.

Another problem area is sensitivity to nondiagnostic morphological changes. A revealing study was recently performed by Bailey, et al [2], in which several of the most popular programs were subjected to almost identical replicas of a series of ECG records. The only difference was that one set was sampled at t = 0, 2, 4, 6,... msec, while the other was sampled at 1, 3, 5, 7,. ...msec. Although there is no discernable visual difference between the two sets, total program diagnostic reproducibility was no more than 90% for the programs tested. These results suggest that noise handling is a problem with these programs and that they are too highly tuned to parameters which may contain significant noise components.

Another difficulty with many programs in present use stems from the use of a hierarchical decision-tree structure consisting of many deterministic decisions. The diagnostic program described here has resulted from an effort to view the problem from another perspective -- one in which the underlying statistical nature of the problem is of paramount importance. An experienced cardiologist depends not only on a few threshold-type decisions but on the weighing of many relevant and interdependent factors which he or she may find difficult to prescribe precisely. Rather than attempt to follow the hierarchical decision structure, our approach has been based on the idea of matching the computer diagnoses to those of cardiologists by varying a set of free parameters of assumed statistical models of the ECG or VCG.

II. PHILOSOPHY OF APPROACH

There are two principal classes of cardiac abnormalities which may be diagnosed using ECG or VCG data: (1) chronic, which produce changes in waveform morphology; (2) rhythm, which produce abnormal temporal patterns of P and R waves. Based on this dichotomy, diagnostic system contains two prinicipal diagnostic subsystems: (1) morphology, (2) rhythm.

Morphology interpretation is accomplished using statistical pattern recognition. Feature extraction is performed using a Karhunen-Loève (K-L) expansion which has desirable noise suppression and data compression properties [3,4]. In addition, the K-L expansion is a conservative approach since signal reconstruction accuracy is the expansion criterion used. Careful visual examination by cardiologists has been utilized to determine the appropriate expansion order. Pattern recognition is performed using a supervised learning technique to partition the feature space into regions which correspond to cardiologist-determined disease classes.

Rhythm analysis is performed by using two subsystems, one for persistent rhythms and the other for transient rhythms [5,6]. Persistent rhythms are identified by using a multiple-model Kalman Filter to track the rhythms. The most likely persistent rhythm is found by examination of the innovations from each filter. Transient rhythms are identified by using a generalized-likelihood ratio (GLR) technique.

The functional flow of the complete system is outlined in Figure 1. Note that there is no interaction between the morphology and rhythm subsystems. However, both morphology and wave timing will be utilized in a postprocessor, yet to be designed, to provide a quite complete rhythm interpretation capability. It is to be emphasized here that we are seeking to perform data compression and determine a set of statistical discriminants which contain all of the diagnostic information required for VCG interpretation.

The use of the 3-lead VCG rather than the 12-lead ECG for prototype design was chosen for several reasons: (1) the methods to be tested are more efficiently mechanized since there are fewer leads (2) the USAFSAM cardiologists were experienced in reading VCG's as well as ECG's (3) VCG data was available for morphology interpretation. Since rhythm diagnosis was concerned only with R-R interval patterns, both VCG and ECG data were used to achieve a suitable rhythm data base for design. The VCG data was recorded in the Frank (X,Y,Z) lead system and digitized at USAFSAM. In the remainder of this paper, we discuss in more detail the philosophy of approach and the development of the integrated system for VCG interpretation.

III. DATA PREPROCESSING

Data preprocessing is divided into two principal functions: (1) baseline removal (2) R wave detection. Starting with a sampled (250 Hz) and digitized signal representation, data preprocessing consists of three steps:

Step 1: Crude Baseline Removal

The baseline is estimated using a symmetric moving window (non-causal filter). Eleven equally spaced points in a 200 msec window are employed. The amplitude frequency response has a slope of about -1/Hz up to the first cutoff point at 1.25 Hz. This step removes most of the low-frequency baseline but does leave artifacts due principally to QRS waves. These are removed in step 3.

Step 2: R Wave Detection

The R wave fiducial points are found using a combination of amplitude (above the baseline) and slope information for the X lead. Let I be the initial 2.4 sec data interval and let $x, \dot{x}$ be, respectively, the value and slope of the x component. Then the tests are, in order:

(i) $|\dot{x}| > 0.15 \max_{t\in I} |\dot{x}|$

(ii) change in sign of $\dot{x}$ detected for two consecutive steps (R wave peak)

(iii) $|x| > 0.20 \max_{t\in I} |x|$, $\dot{x}x > 0.13 \max_{t\in I} (\dot{x}x)$ } at peak

(iv) $|x| < 0.60$ R wave height

$|\dot{x}| < 0.10 \max_{t\in I} |\dot{x}|$

Step 3: Fine Baseline Removal

The QRST complexes are removed from the raw data and then the baselines computed from this rectified data which contains essentially baseline and P waves. This gives a very accurate baseline estimate. At present, one of two methods may be used: (1) the method of step 1, or (2) a method based on spline approximations to the sample means found in successive PR intervals.

The results of these three steps is the production of cleaned up VCG signals, essentially free from baseline, and the location of each QRS complex.

IV. MORPHOLOGY DIAGNOSIS

4.1 Feature Extraction

In this work we modeled the waveforms recorded at the three VCG leads as second order non-stationary random processes. The statistical properties of a second order random process are essentially characterized by its covariance function (with no loss of generality we may assume the processes to be zero mean). Let $(x_t | 0 \le t \le T)$ denote a second order process. It is known that the process can be represented uniquely as

$$x_t(w) = \sum_{n=1}^{\infty} Z_n(w)\phi_n(t) \qquad (4.1)$$

where $Z_n(w)$ are mutually uncorrelated random variables given by

$$Z_n(w) = \int_o^T x_t(w)\phi_n(t)dt \qquad (4.2)$$

and the $\phi_n(t)$ are obtained by solving the eigenvalue-eigenfunction problem.

$$\int_o^T R(t,s)\phi(s)ds = \lambda\phi(t) \qquad (4.3)$$

where R(t,s) is the covariance kernel of the process $(x_t | 0 \le t \le T)$. This representation is optimal in the sense that

$$\int_o^T \left| x_t(w) - \sum_{n=1}^{N} Z_n(w)\phi_n(t) \right|^2 dt = \mathrm{Min}\left[\int_o^T \left| x_t(w) - \sum_{n=1}^{N} \alpha_n(w)\psi_n(t) \right|^2 dt \;\middle|\; \begin{array}{l} (\alpha_n(w))\, n=1,\dots N \\ \int_o^T \psi_i(t)\psi_j(t)\,dt = \delta_{ij} \end{array} \right] \qquad (4.4)$$

where the eigenvalues of (4.3) form an ordered set $\lambda_o \geq \lambda_1 \ldots \geq \lambda_m \geq \ldots$ and either the sequence terminates in a finite number of steps or $\lambda_n \to 0$ as $n \to \infty$. This expansion is known as the Karhunen-Loeve (K-L) expansion. It should be emphasized that no assumption of stationarity of the process needs to be made for the expansion (4.1) to be valid. In fact, if the process were stationary then (4.1) reduces to the Fourier Series (sine-cosine) expansion.

In order to apply these ideas to the modelling of VCG waveforms, we compute the sample covariance function of the waveforms obtained at each of the three VCG leads, where the sample functions are elements of a representative collection of samples corresponding to both normal and vaious types of abnormal heartbeats. Thus our sample covariance functions are computed from a finite number of second order random variables and the eigenvalue-eigenfunction problem reduces to an eigenvalue-eigenvector problem.

Since $R(t,s)$ is a covariance kernel the eigenvalue-eigenvector problem we have to solve corresponds to a non-negative definite symmetric covariance. The coefficients of the Karhunen-Loeve expansion form a vector in feature space and we are interested in keeping the dimensions of the feature space as low as possible consistent with waveform reconstruction with no loss of essential information. This means that we do not have to solve the eigenvalue-eigenvector problem corresponding to (4.3) completely. Thus if the sample covariance matrix in $n \times n$ and the dimension of the feature space is $m<n$, then we need only the first m principal eigenvalues and eigenvectors. For our problem n is typically 300 and m is 60. This eigenvalue-eigenvector problem can be efficiently solved using a method due to Wilkinson and Reinsch.

We remark that the K-L expansion apparently has not found more widespread use because it was felt that the eigenvalue-eigenvector problem was too large to be solved efficiently.

We have previously remarked that we wish to represent the heart-beat waveforms using the K-L expansion without any essential loss of information. This means that if we compute

$$e_m(w) = \int_o^T |x_t(w) - \hat{x}_t(w:m)|^2 dt$$

where

$$\hat{x}_t(w;m) = \sum_{n=1}^{m} z_n(w)\phi_n(w)$$

then, for example, we want $E(e_m(w)) < \varepsilon$ where ε is some pre-assigned small number. This, however, is a global error criterion and in our problem we are required to make sure that the representation is adequate in some local regions (for example, the P-wave, the S-T segment). In order to ensure this it may be more efficient to use a weighted K-L expansion, i.e., the $\phi_n(t)$ are now orthonormal in the following sense:

$$\int_o^T w(t)\phi_i(t)\phi_j(t)dt=\delta_{ij} \quad \text{where } w(t)>0,\ 0 \leq t \leq T,$$

is a properly weighted function. We have experimented with this approach and have obtained reasonably satisfactory results [4].

Another possibility is to use a segmented K-L expansion. This approach is being extensively investigated in the M.I.T. doctoral dissertation of A. Akant

Finally, we remark that the K-L expansion has desirable noise suppression properties so that 60 Hz filtering and complex artifact rejection logic is not necessary. Only gross artifacts, which are easily detected and eliminated, need be considered prior to feature extraction.

4.2 Clustering

Once the features are found (the coefficients of the K-L expansion), morphology interpretation is performed using a supervised pattern recognition algorithm in feature space. The algorithm used is based on representing each morphological class by a hyperellipse, defined by the class mean and covariance matrix.

Partitioning Algorithm

Let $\alpha_i(j)$ be the feature vector for the jth record in the ith class. These are determined apriori by USAF cardiologists based on careful examination-enhanced VCG records. The algorithm then proceeds as follows: first, the class means $\bar{\alpha}_i$; $i = 1,2,\ldots,C$ are computed. Then the class covariance matrices P_i and their inverses, are found. However, the inverses were not found directly for two reasons:

(i) P_i may be ill-conditioned numerically, due to insufficient data;

(ii) possible ill-conditioning of numerical inversion for large dimensions (up to 60 are required).

The Matrix Inversion Lemma was used, which necessitated using an initial estimate of P_i and P_i^{-1}. The assumed form was $P_i = f_i I$ with f_i a scalar size parameter which was varied to give best results.

After the clusters have been formed, it is necessary to evaluate the clustering for diagnostic purposes. Given a point α in feature space, the distance of α from the ith cluster is computed as

$$d_p(i,\alpha) = -\frac{1}{2}d(i,\alpha) - \frac{1}{2}\sum_j \ln(P_i)_{jj} \qquad (4.7)$$

where $d(i,\alpha)$ is the Mahalanobis distance between α and the ith cluster

$$d(i,\alpha) = (\alpha - \bar{\alpha}_i)^T P_i^{-1} (\alpha - \bar{\alpha}_i) \qquad (4.8)$$

Under a gaussian assumption $d_p(i,\alpha)$ is proportional to the log probability of α for the ith class.

V. RHYTHM ANALYSIS

The basis for our approach to rhythm analysis is the determination of a set of dynamic models that accurately describe the sequential behavior of the elapsed times between consecutive heartbeats ("R-R intervals"). Specifically, a number of arrhythmias are characterized by persistent patterns of R-R intervals, while others involve abrupt changes in the interval pattern. For each of these we develop a simple model that generates an interval pattern with the corresponding statistical properties. Using these models, we then apply two statistical techniques for the identification of persistent patterns and for the detection of abrupt changes.

It is this aspect that is novel in approach to arrhythmia analysis. We have developed very simple phenomenological models that accurately describe the statistical behavior of R-R interval patterns. Using these models we can apply powerful statistical techniques to develop an ECG analysis system that is simple and robust and whose performance can be accurately determined as a function of a very small number of design parameters. Our treatment is necessarily brief. For a complete development, we refer the reader to [5,6].

5.1 Modelling of R-R Interval Patterns

Let $y(k)$ denote the actually observed kth R-R interval. We think of y as being the output of an R-R pattern generator, which is characterized by the state vector $x(k)$. The output $Hx(k)$ represents the ideal kth R-R interval, which differs from $y(k)$ by the noise $v(k)$, which arises from two sources:

i) the unavoidable errors in computing R-R intervals, caused by inaccuracies in locating the fiducial points.

ii) variations due to the fact that actual rhythms are never "textbook perfect", rather there are small, apparently random variations about the ideal underlying pattern.

Models are first described for several persistent rhythms.

Small Variation: This class exhibits small but random deviations from the mean value of the R-R intervals (e.g., normal sinus rhythm).

$$x(k) = x(k-1), \quad y(k) = x(k) + v(k) \qquad (5.1)$$

where v is zero mean Gaussian and white, with variance R_s. The initial mean $m(0)$ and variance $P(0)$ of $x(0)$, as well as R_s are design parameters, reasonable values for which can be determined with the aid of the statistical techniques described in [5].

Large Variation: This class is characterized by large but random variations in the R-R intervals (e.g., sinus arrhythmia). The model for this class is also given by (5.1), the only difference being that the variance of v is taken to be $R_1 > R_s$.

Period Two Oscillator: This class is characterized by R-R intervals which are alternately long and short (e.g., bigeminy).

$$x(k) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} x(k-1) \qquad (5.2)$$

$$y(k) = (1,0)x(k) + v(k) \qquad (5.3)$$

where the initial mean $m(0)$ and covariance $P(0)$ of $x(0)$ and the variance R_2 of v are free parameters to be determined by some statistical means.

Period Three Oscillator: This class exhibits an R-R interval sequence with a period of 3 (e.g., trigeminy).

$$x(k) = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} x(k-1) \qquad (5.4)$$

$$y(k) = (1 \quad 0 \quad 0)x(k) + v(k) \qquad (5.5)$$

Again $m(0)$, $P(0)$, and R_3 (the variance of v) are free parameters.

We now turn our attention to the models of transient events. All such events are modeled as sudden, unpredictable changes on an otherwise normal record. Thus the basic model is (5.1) with covariance of $v = R_s$, and the various transient events are modeled as changes in this pattern.

Rhythm Jump: This class is characterized by a sudden change in the heart rate.

$$x(k) = x(k-1) + v\delta_{\theta,k} \qquad (5.6)$$

Here v is the unknown size of the shift in the average R-R interval at the unknown time θ. Also, δ_{ij} is the Kronecker delta ($\delta_{ij} = 0$, $i \neq j$, while $\delta_{ij} = 1$).

Noncompensatory Beat: This class is characterized by the presence of a single lengthened or shortened R-R interval (e.g., SA block, PAC)

$$x(k) = x(k-1) + v[\delta_{\theta,k} - \delta_{\theta,k}] \qquad (5.7)$$

i.e., $x(k) = x(0)$ for $k \neq 0$, and $x(0) = x(0) + v$.

Compensatory Beat: This class is characterized by an isolated premature QRS complex followed by a compensatory pause before the following beat (i.e., a PVC).

$$x(k) = x(k-1) + v[\delta_{\theta,k} - \delta_{\theta,k-1} - \delta_{\theta,k-2}] \qquad (5.8)$$

Double Noncompensatory Beat: This class is characterized by two consecutive shortened or lengthened R-R intervals.

$$x(k) = x(k-1) + v[\delta_{\theta,k} - \delta_{\theta,k-2}] \qquad (5.9)$$

5.2 An Arrhythmia Detection Technique

Examining the models for the persistent rhythm classes, we see that they are all linear systems. Thus, given a sequence of observed R-R intervals, we can use the multiple hypothesis method [5] -- consisting of a bank of Kalman Filters, one for each of the models -- to compute the probabilities for each of the persistent rhythm categories based on analysis of the filter residuals.

In the case of the transient categories, the generalized likelihood ratio (GLR) technique [6] has been implemented. This approach involves the implementation of a Kalman Filter based on the small variation model. The residuals of this filter are then fed into several matched filters that compute most likely times and the likelihood ratios for each type of transient event. Estimates of the jump v are also obtained.

The prototype system consists of both of these subsystems with several additional features designed to:

a) shorten the response time of the multi-filter in the case of a shift from one normal rhythm pattern to another one;
b) improve the distinguishability of small variation from the period two and three oscillators and from large variations;
c) speed up the GLR identification process by looking at a narrow "window" of the most recent data;
d) provide an initialization procedure for all filters to enhance detection of events at the beginning of an ECG record;
e) reset the system subsequent to the detection of a transient event.

VI. RESULTS

Fig. 2. depicts the result of one test on real ECG data using the rhythm subsystem. The format of the figure is as follows: At the top of the figure is the actual ECG waveform being processed. The small vertical lines beneath the waveform are the R-wave detector. The multifilter probabilities (in percentage form) are displayed next, where the R-R intervals (between the present R-wave and the preceding one) are measured in units of 4 ms. The symbol "OUT" is used to indicate a multifilter outlier -- an indication that a previously identified persistent pattern has been interrupted.

The GLR likelihood ratios are plotted below the multifilter results. Again, the horizontal axis is actual time, while the numbers given represent the running estimate of the mean R-R interval, as produced by the small variation GLR filter. We note that in addition to the categories, "noncompensatory" (N), "compensatory" (C), and "double noncompensatory" (D), we have the category "warning" (W), which indicates the preceding R-R interval is fine but that the present one is aberrant. This signal tells us that when we look at the next interval, we should be able to decide among the various transient categories. The actual GLR decisions are located beneath the plot. In addition, W, C, N, and D, two other symbols are used. "JD" denotes the detection of a double noncompensatory and indicates that it might really be a jump -- a question that is resolved upon looking at the next interval. Finally, at the top of the GLR plot, the times at which the GLR filter is adjusted are indicated. "JUMP" indicates that the filter estimate has been adjusted following the detection of a jump. "OUT" indicates that the GLR filter has been reinitialized. This only happens if the multifilter has locked onto period two or three oscillation and an outlier has been detected.

The record shown in Fig. 2 contains several bursts of ventricular tachycardia plus one isolated PVC. The bursts tend to drive the multifilter towards large variation, although several inliers (IN) are signaled during the calm periods between the bursts. Each of the bursts consists of two shortened R-R intervals followed by a pause. The GLR system consistently detects the two shortened intervals as a jump or double non-compensatory (JD). In several cases the pause was long enough to be classified as compensatory (C), while in the others it was shorter

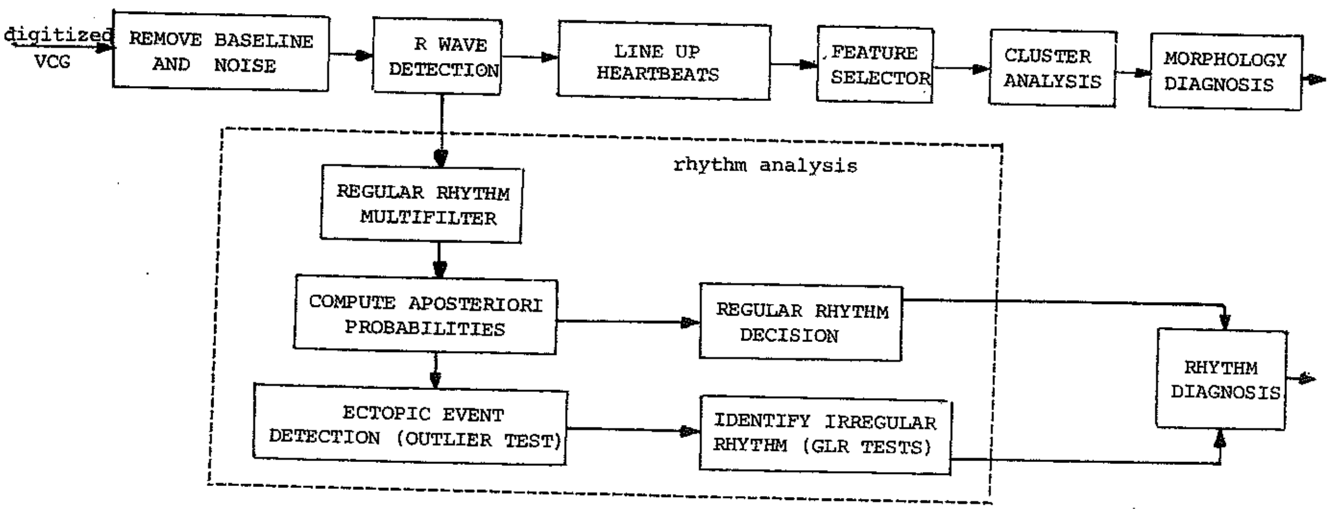

Fig. 1: VCG Interpretation System Functional Flow

lending to a "D" classification. The isolated PVC is classified as a compensatory premature (C).

An ensemble of 936 VCG records were employed to develop and test the K-L expansion technique. Approximately one half of these were normal, with the other half containing examples from 22 cardiologist-determined disease classes. This was done to give maximum variability to the data, and thus provide good reconstructions for both normal and abnormal records. It was estimated that 60 coefficients were sufficient to give reconstructions accurate enough to retain diagnostic information. The worst pointwise reconstruction error was evaluated for each record. An example of the most severe error encountered is shown in Fig. 3. The maximum error is 0.092 mv, for the z axis reconstruction. The integrated absolute errors are shown for each axis. The average pointwise error is obtained by dividing by 100.

VII. CONCLUSIONS

A new approach is being taken to the problem of automated interpretation of electrocardiograms. Rather than relying on deterministic decision rules and a hierarchical decision tree structure, the approach is an attempt to make full use of several powerful and relevant tools of modern filtering and detection theory. The approach utilizes the concepts of data compression, pattern recognition and sufficient statistics to generate a parsimonious set of discriminants which contain all of the useful diagnostic information. With this approach, it is felt that the program will be insensitive to noise and more sensitive to diagnostic discriminants than present programs. This conjecture has been borne out in recent experimental results, which will be the subject of future papers.

References

1. E. Drazen, "Automated Electrocardiography in the United States", Arthur D. Little Report on Contract 230-75-0212, August 1976.
2. J. J. Bailey, et al., "A Method for Evaluating Computer Programs for Electrocardiographic Interpretation", Circulation, Vol. 50, 1975: III Reproducibility Testing and Source of Program Errors, p. 88.
3. M.E. Womble, S.K. Mitter and J.S. Halliday, "Classifying Vectorcardiograms via Signal Pattern Recognition", C.S. Draper Lab.Rep. R-766, Oct. 1973.
4. D.E. Gustafson, A.S. Willsky, S.K. Mitter, J.Y. Wang, A. Akant, W.C. Kessel and P.C. Doerschuk, "Automated VCG Interpretation Studies Using Signal Analysis Techniques", C.S. Draper Lab. Report R-1044, January 1977.
5. D.E. Gustafson, A.S. Willsky and J.Y. Wang, "Cardiac Arrhythmia Detection and Classification Through Signal Analysis", C.S. Draper Lab. Report R-920, July 1975.
6. D.E. Gustafson, A.S. Willsky and J.Y. Wang, "Detection and Identification of Transient Cardiac Arrhythmias Using Signal Analysis", C.S. Draper Lab. Report R-935, December 1975.

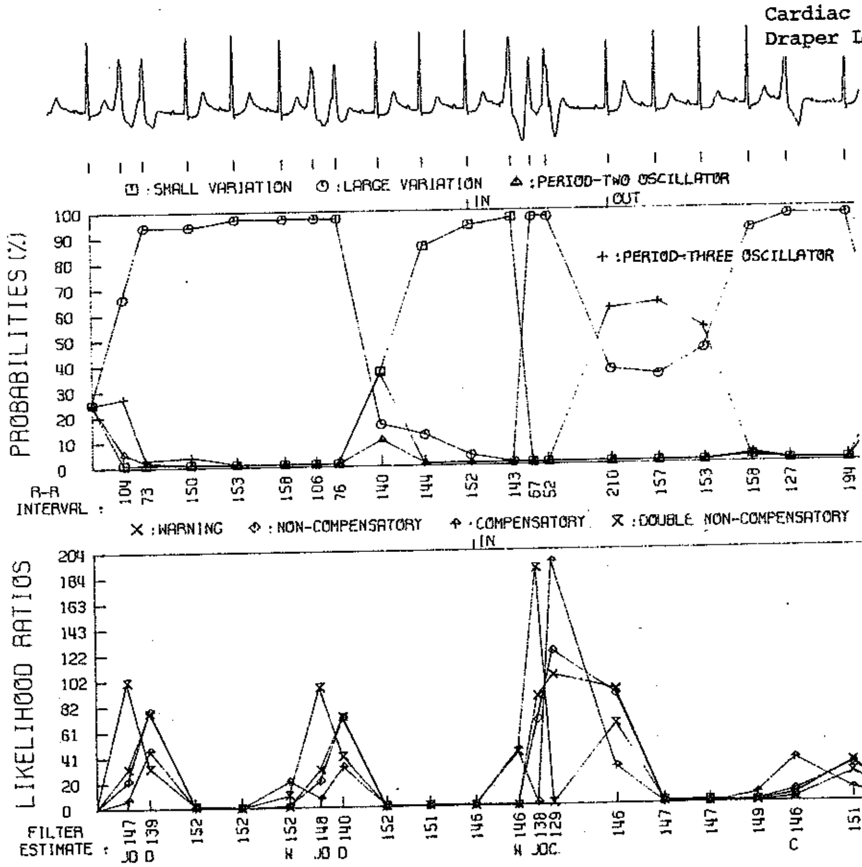

Fig. 2 Rhythm Analysis Example

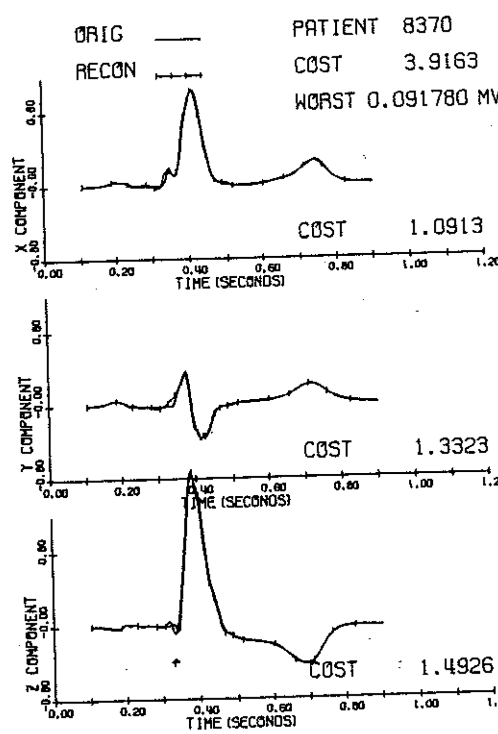

Fig. 3 Example of Worst Reconstruction Error