Automated Performance Evaluation of Real-Time QRS-Detection Devices

SETH SUPPAPPOLA, PhD, YING SUN, PhD

This paper presents a QRS-detector error logger and the implementation of a fully automated system for evaluation of QRS-detector performance. The error logger requires a known reference, such as an electrocardiogram (ECG) database, that provides ECG data and reference pulses wherever QRS complexes occur. These reference pulses are used to validate detection pulses generated by a QRS detector in real time. Using an eight-bit microcontroller chip, the error logger was constructed at a very low materials cost. Applications include automated analysis of QRS-detector performance and optimization of parameters in QRS-detection algorithms. (BIOMEDICAL INSTRUMENTATION & TECHNOLOGY 1995;29:41–49)

Because of the extreme variability of electrocardiographic (ECG) signal morphology, QRS detection has always been a challenging engineering task, and, as evidenced in the literature, it still sparks ongoing research interest.[1-12] Most often, when a real-time QRS detector is designed, especially if it is a digitally based detector, much of the evaluation is done by computer simulation using a database. In the past, most research on QRS detectors has been done using custom sets of ECG test data.[1-5] And although much current research uses this approach,[6,7] due to the increased accessibility of digital ECG databases, many researchers are moving toward using common databases for their evaluations.[8-11] In the United States, there are two such databases: the American Heart Association (AHA) Ventricular Arrhythmia Database and the Massachusetts Institute of Technology–Beth Israel Hospital (MIT–BIH) ECG database. Both of these databases were originally distributed in analog format, but they are now most common in digital format, which has further standardized the ECG data, since distortions intrinsic to features of analog recorders such as tape-head alignment and motor speed are no longer a factor.

There are many advantages to using a digital QRS detector over an analog detector.[9] Consequently, there is much interest in digital detectors. A designer may develop an algorithm using one of the aforementioned databases, but although he or she may test the algorithm on hundreds of thousands of beats, it is still only a simulation. And although some very thorough analog evaluations have been done on digital detectors using common databases,[11] it is nevertheless our objective to provide an *automated* system that will analyze, in real time, QRS detectors in environments that are approximately clinical in the sense that only an actual patient is missing.

Compared with a computer simulation, the proposed evaluation system has several advantages. First, since the error logger is fully independent of the detection scheme used, a designer may evaluate either digital or analog QRS detectors directly. Second, the error logger is fully independent of the ECG signal source. Therefore, the error logger may be used with any database (analog or digital format), so long as the signal source can provide a reference pulse whenever it produces a QRS complex. Third, it is easy to assess what effects various changes in the ECG signal will have on QRS-detector performance. For example, the dynamic range of a QRS detector can be established by varying the gain of the ECG before it is fed to the detector. Or, the effect of noise can be studied — one simply mixes the ECG with the desired contaminant (white noise, 60-Hz noise, etc.) before it is fed to the QRS detector. Or, the effects of signal processing can be investigated by operating on the ECG signal before it is fed to the QRS detector (filtering, overdrive, etc.). Thus, we evaluate the QRS detector in the way in which it will be used, as opposed to through hypothetical computer simulation.

In this paper we present the design for a piece of electronic equipment for the evaluation of both analog and digital QRS detectors. Then, by way of an example application, we demonstrate its use while providing a guideline for using the device to test QRS detectors.

Received from the University of Rhode Island, Department of Electrical Engineering, Kingston, Rhode Island 02881-0805.

Address correspondence and reprint requests to Dr. Sun.

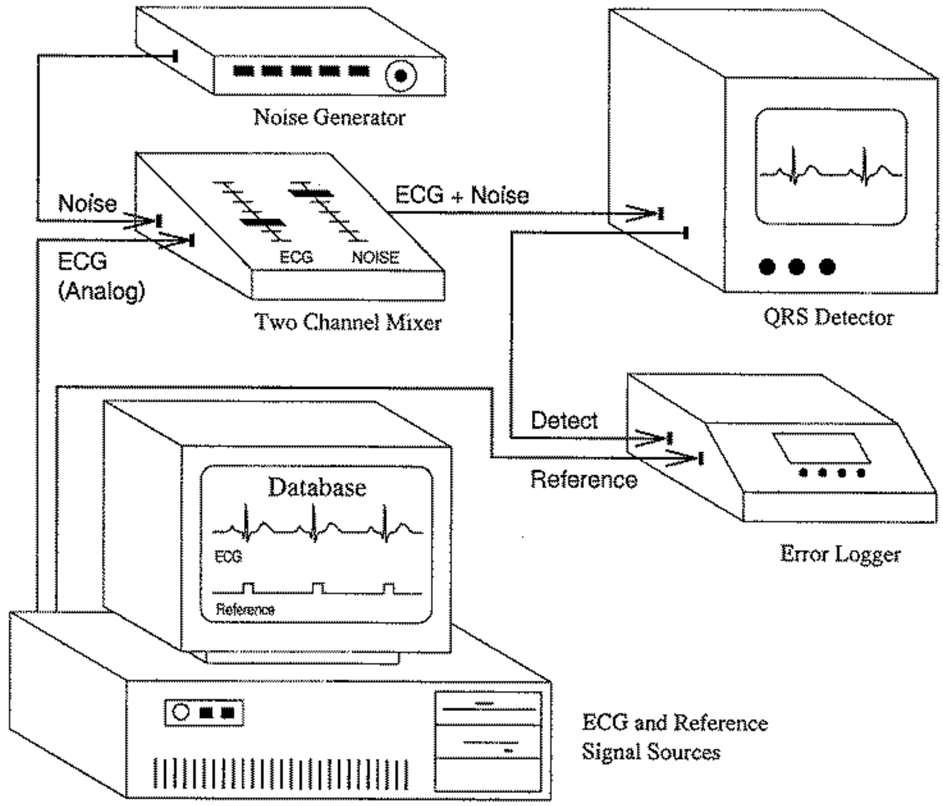

Figure 1. Typical setup of the complete QRS-detector evaluation system using the error logger.

METHOD

An example setup of the QRS-detector evaluation system is shown in Figure 1. It consists of four major components: 1) a known ECG reference source, such as a database, that can provide an analog ECG signal coincidental with a digital logic signal that indicates the occurrence of a QRS complex; 2) an ECG mixer and noise source for adding noise contaminants to the ECG; 3) a hardware QRS detector, which can be any physical device that accepts ECG as input and locates the QRS complexes in the signal in real time; and 4) the QRS-detector error logger, which is presented in this paper. Of course, the mixer is an optional component, since it is necessary only if one desires to study the effects of noise.

The mixer we chose to use is a straightforward summing circuit which offers no filtering or offset capability so that we minimally distort the ECG which is generated from the database. For more sophisticated testing, AAMI has published a standard[13] that includes a more complex circuit for adding noise, filtering, attenuation, and testing pacemaker pulse transmission.

We have often been asked why, since our system uses a digital database, we did not use the computer to simply add the desired noise signal to the ECG data before they are presented to the system under scrutiny. The answer is twofold: first, we wanted to keep the system highly modular so that any ECG signal reference could be used, and second, we wanted to add white noise, for example, to the digital ECG data before conversion would bandlimit the resulting noise to the Nyquist frequency of the ECG data (this would limit our ability to study the effect of high-frequency noise on the QRS detector under scrutiny).

Hardware

As can be seen from the schematic in Figure 2, the electronic hardware for the error logger comprises a very simple system, thereby simplifying construction and keeping cost extremely low. The heart of the error logger is based on the Motorola MC68HC811E2 microcontroller unit (MCU). This MCU integrates a 6800-based eight-bit microprocessor with ¼ kbyte of RAM and 2 kbytes of EEPROM. A quartz oscillator provides the system clock, and a Maxim 690 microprocessor supervisory circuit provides a power-on reset. The microcontroller is interfaced to an LCD display for data display and has four softkeys for user control. The electronics are packaged in an aluminum enclosure, and the four softkeys are mounted just below the LCD on the enclosure's front panel, as indicated in Figure 1. The softkey switches are debounced under software control. The bottom row of the LCD is used by the software to display the softkeys' functions.

The Reference and Detect inputs are optoisolated to prevent ground loops and to eliminate the need for a common signal ground. The inputs draw only 180 μA with TTL-level input signals.* The resulting logic signals are fed into Schmitt trigger inverters and drive LED indicators that light when the respective inputs are high.

* Although TTL levels are probably the most common, due to the flexibility of the optoisolated inputs, any differential input voltage from 3 V to over 1,000 V can be used to indicate a high level.

Software

The software for the error logger was written entirely in assembly language. It was composed on an IBM-compatible PC, assembled, and then downloaded to an MC68HC11 development system with in-circuit emulation capabilities. When the software was fully debugged, the program was stored in the EEPROM of an MC68HC11 microcontroller, which was then installed in the error logger.†

† This prototype device was developed at the University of Rhode Island for research purposes; therefore, complete units are not available, nor are programmed microcontrollers. However, the assembly source (fully commented) will gladly be provided to anyone who sends an Internet e-mail request for such to sun@ele.uri.edu. Moreover, the electrical schematic completely describes all the electronics except the 5-VDC source. The packaging is not described in detail, but it is a trivial part of the design.

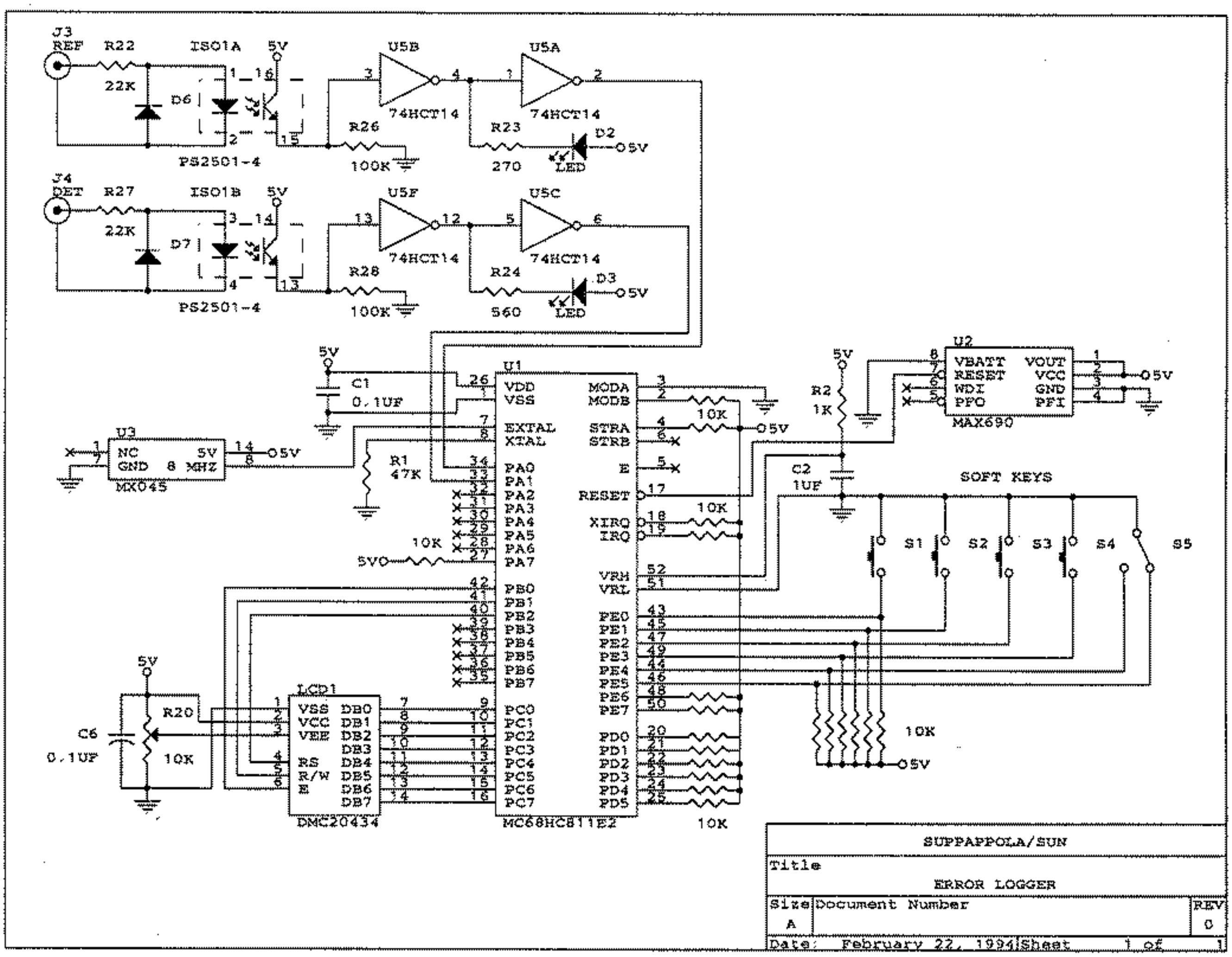

Figure 2. Electrical schematic of the microcontroller-based error logger.

The error logger accumulates and displays several statistical parameters in real time: false positive (FP), false negative (FN), and true positive (TP). It also tallies the number of beats analyzed. To determine whether or not a QRS complex was correctly detected (i.e., TP) the concept of a valid detection interval (VDI) is employed.[8] The VDI is an interval in which if a detector makes a detection, it is considered to be a TP, but any additional detections in the VDI are treated as FPs.

The VDI has three parameters: Lead, Lag, and Delay, as shown in Figure 3. Let t_R be the rising edge of a reference pulse, and let t_{VDI} be a point shifted from t_R by an amount Delay such that $t_{VDI} = t_R + \text{Delay}$. Then, Lead is the duration of the VDI before t_{VDI}, and Lag is the duration of the VDI after t_{VDI}. Hence, the VDI is the interval

$$t_R + \text{Delay} - \text{Lead} \leq \text{VDI} \leq t_R + \text{Delay} + \text{Lag} \quad (1)$$

There are basically four rules for determining detection statistics based on the VDI principle:

1. If a detector makes a detection in the VDI it is considered to be a TP.

2. Only one detection in the VDI is considered to be a TP. Any additional detection is scored as a FP. This allows only one TP per beat.

3. Any detection outside the VDI is considered to be a FP.

4. If no detection is made in the VDI (i.e., the beat is missed), a FN is scored.

These statistics are accumulated in real time, and the software algorithm to do so is given in Figure 4. The instantaneous detection delay is also computed. It is defined as the time from the onset of the reference pulse to the onset of the detection pulse.

The error logger software is based on a 1-ms interrupt cycle. Thus, the instantaneous-detection-delay measurement has an accuracy of 1 ms. A clock accumulating the elapsed time is also maintained. LCD updates are time-multiplexed so that when several parameters need to be adjusted simultaneously, the update does not exceed the 1-ms period. The input signals (Reference and Detect) are sampled once every millisecond; however, the software acknowledges only low to high (L–H) logic transitions on these inputs. Thus, to ensure detection of the L–H transition, input signals should remain low for at least 2 ms before a L–H transition and should remain high for at least 2 ms thereafter. Notice, though, that since only the rising edges of the Reference and Detect signals are considered, it is not necessary that the Reference and Detect pulses overlap when using the VDI principle to analyze a QRS detector.

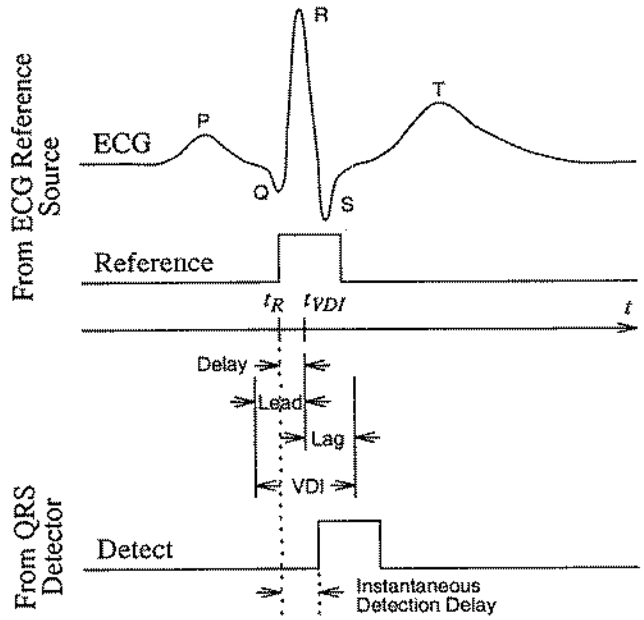

Figure 3. Explanation of the valid detection interval (VDI).

All setup information is entered by the user via softkeys. This includes all the parameters of the VDI (i.e., Lead, Lag, and Delay). Values for these parameters from 0 to 999 ms may be entered. Softkey functions are also provided to initiate and pause the accumulation of detection statistics.

Receiver Operating Characteristic Curves

To aid in the interpretation of results from various tests, we use receiver operating characteristic (ROC) curves.[14,15] Essentially, these are plots of percentage TP versus percentage FP. Typically, it is desired to have the detector operate in the upper left corner of the plot, as this point maximizes the percentage TP while minimizing the percentage FP. We do not delve into the intricacies of interpreting ROC curves in general here, since on a given ROC curve there is no single operating point that would be optimal for *every* application.

To generate the data for these plots we use the following definitions:

$$\%\,\text{TP} = \frac{\text{Number TP}}{\text{Number beats}} \times 100\% \quad (2)$$

$$\%\,\text{FP} = \frac{\text{Number FP}}{\text{Number beats}} \times 100\% \quad (3)$$

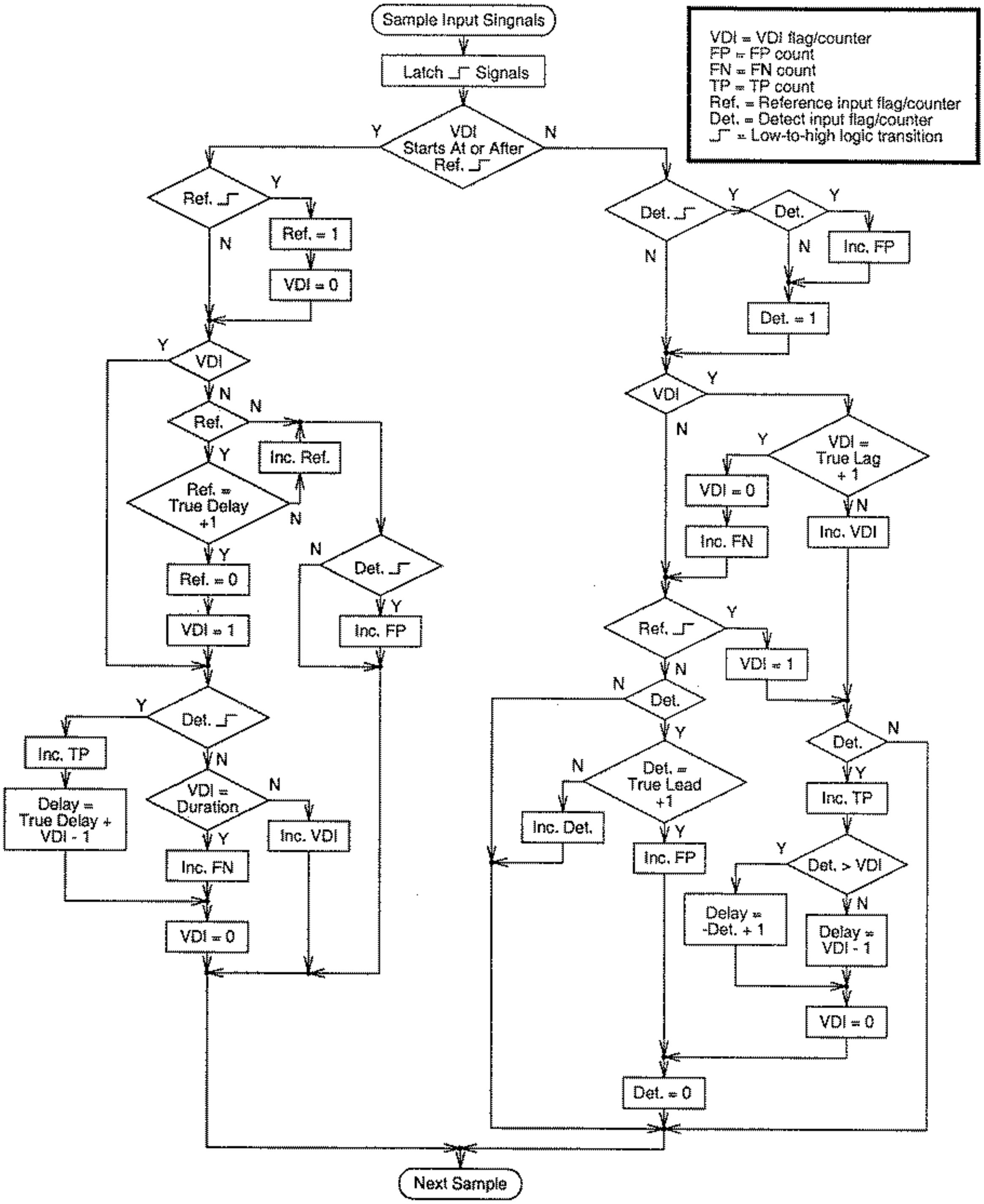

Figure 4. Flowchart for real-time determination of false-positive (FP), false-negative (FN), and true-positive (TP) statistics, and instantaneous detection delay.

Notice that since the number of TPs can not exceed the number of beats, the percentage TP is upper-bound by 100%. However, the number of FPs can exceed the number of beats, therefore, under these definitions, the percentage FP *can* exceed 100%.

To generate the complete ROC curves, it is necessary to run through the entire test data set for *each* point on the curve. The procedure for a noise test is:

1. Select a signal-to-noise ratio (SNR) and compute appropriate signal levels.

2. Use a test signal and an oscilloscope to set the appropriate level for the signal and noise using a signal mixer.

3. Initiate the error logger.

4. Begin reproducing the ECG data test set.

5. After the entire ECG test set has been presented to the QRS detector, use equations 2 and 3 to compute the point on the ROC curve.

6. Repeat the previous steps until sufficient data have been produced to generate the ROC curve.

A dynamic range test would be performed similarly except no noise would be added, and only the gain of the ECG would be adjusted.

FULLY AUTOMATED ANALYSIS SYSTEM RESULTS

To demonstrate the use of the error logger, we use it to evaluate and compare two QRS detectors: one is an analog design, an actual circuit that was extracted from a piece of production medical equipment, an intra-aortic balloon pump (IABP), and the second is a digital detector, which is under development and is intended to replace the analog design. We use two tests to compare and evaluate the approaches: 1) a dynamic range test, and 2) a 60-Hz power-line-interference test.

For data, we used the AHA Ventricular Arrhythmia Database and software previously developed[16] to perform the conversion back to an analog signal. Furthermore, since we were studying the effects of noise-and-signal dynamic ranges, not the effects that various ECG signal morphologies have on the detectors, we chose database records with which we expected the detectors to perform reasonably well under no-noise conditions. However, to keep the flavor of the ECG somewhat varied, when constructing our set of test data, we selected eight 30-minute records, one from each of the eight series in the AHA database, thereby comprising a test set of four hours of ECG and almost 17,000 beats.

For either experiment, the level of the ECG must be established. Since the QRS complex is nonstationary in the ECG, and of random amplitude, this is no trivial task. However, since most real-time QRS-detection methods either directly or indirectly inspect the slew rate of the R-wave, we defined the signal level to be the peak-to-peak amplitude of the QRS complex. To establish a signal level for our test data, we used an oscilloscope to manually examine the signal level coming off our Data Translation digital-to-analog converter board installed in the PC. We examined each of the eight records to estimate the peak-to-peak voltage levels of the QRS complexes (excluding PVCs), and then averaged the estimates from the eight records to obtain an overall signal level for our set of test data.

To compute the gain of the instrument, we used a sinusoidal test signal at 30 Hz, which is in the upper range of the frequency content of the QRS complex. This proved to be convenient, and ensured that a consistent signal was available for setting gains on the ECG mixer.

The detectors under investigation were designed to detect R-waves while the R-wave occurred. Therefore, we used a VDI with a Delay of 0 ms, a Lead of 10 ms,

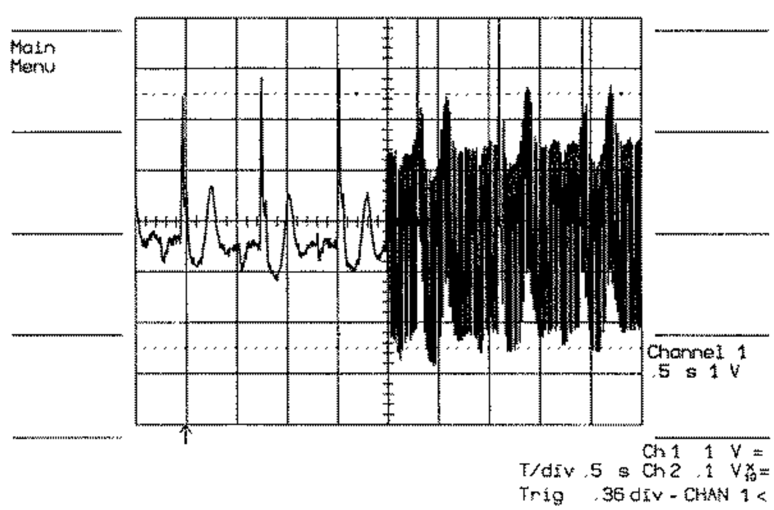

Figure 5. Oscilloscope screendump showing 60-Hz interference at a signal-to-noise ratio of 0 dB, where the noise was turned on at the center of the trace.

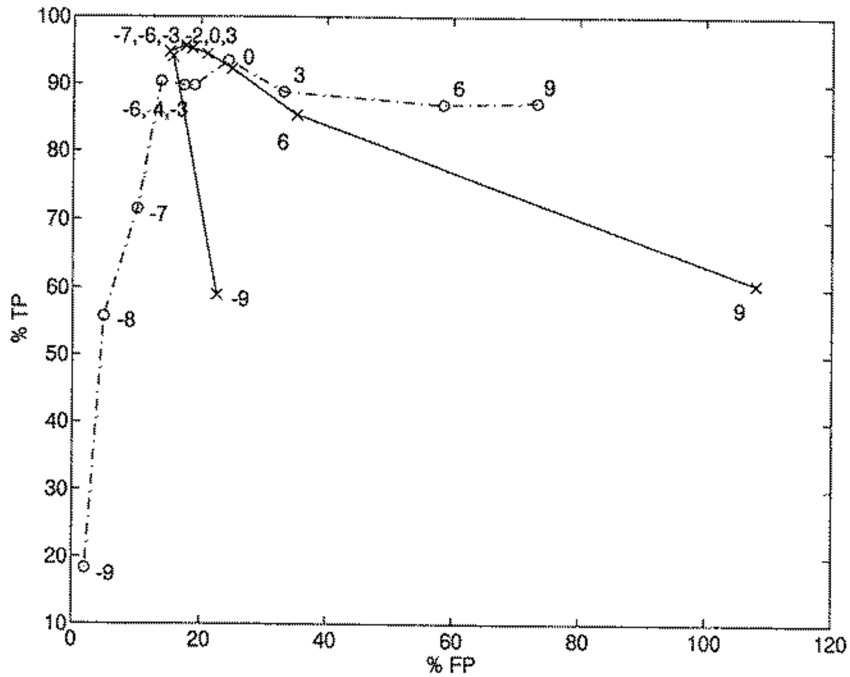

Figure 6. Receiver operating characteristic curves for the dynamic range test as performed on the analog (○) and digital (×) detectors. Numbers are gain in dB; 0 dB was chosen arbitrarily.

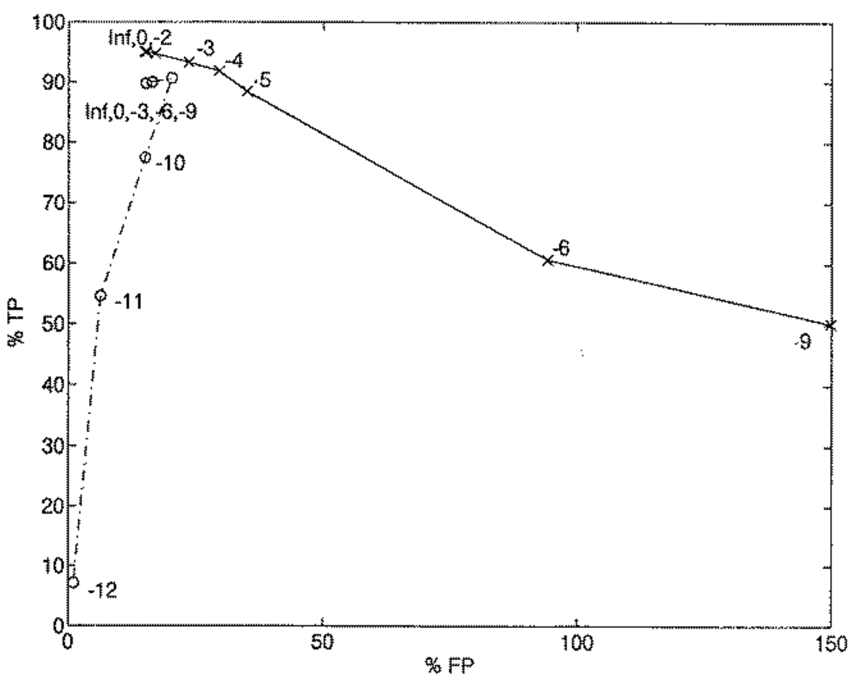

Figure 7. Receiver operating characteristic curves for the 60-Hz interference test as performed on the analog (○) and digital (×) detectors. Numbers are signal-to-noise ratios in dB; "Inf" = ∞ (infinity).

and a Lag of 140 ms. The Lead of 10 ms with 0 ms Delay is required since the AHA ECG Database annotation is not consistently placed exactly at the onset of the R-wave but may occur close to the peak. These VDI parameters were chosen through empirical observations.

The dynamic-range test analyzes the performances of the detectors as a function of input ECG signal level. To vary the level of the ECG signal, one channel of the ECG mixer was used as an amplifier to set various levels of gain, while the other mixer channel was left open. Since the dB scale is based on relative values, not absolute values, an arbitrary gain (from input of mixer to input of detector) was defined as 0 dB. Various other signal levels were then used for the two detection schemes to construct the ROC curves in Figure 6.

Observe that the percentage TP is not a function of percentage FP as the curve doubles back on itself at the left edge. This nonstandard ROC shape is a result of the fact that both very low (approximately no signal) and very high (amplifier overdrive) signal levels result in poor performance; thus, the percentage TP and the percentage FP do not trade off as expected. The 60-Hz power-line-interference test analyzes performance as a function of amplitude of an additive 60-Hz noise level. The aforementioned two-channel signal mixer was used to vary the amounts of 60-Hz interference added to the ECG while the gain of the ECG channel was held constant at a level at which the detector performed well in the dynamic range test. In Figure 5, 60-Hz noise is shown at a signal-to-noise ratio of 0 dB according to our definition. The resulting ROC curves are shown in Figure 7.

For our example application, the IABP with balloon deflation triggered by R-wave detection, the ROC curves generated with the aid of the QRS-detector error logger tell us the following: 1) In terms of a minimum-distance criterion to the minimum error operating point (upper left corner of the ROC where TP = 100% and FP = 0%), the digital detector comes closer to this mark for both the dynamic-range test and the 60-Hz-interference test. 2) For the 60-Hz test, the digital detector performs better as the signal-to-noise ratio decreases because it tends toward FP, while the analog detector tends toward FN (false negative, % FN = 100% − % TP). For the IABP, FPs are more tolerable than FNs, since to deflate the balloon more often than necessary does not increase assistance to the heart but does little harm to the patient; however, not deflating the balloon when the heart contracts provides negative assistance by impeding the flow of blood. 3) As an estimate of dynamic range, by examining the "knee" of the ROC curves we see that the analog detector has about −6 dB to 0 dB → 6 dB tolerance of signal dynamic range, whereas the digital detector has about −7 dB to 3 dB → 10 dB tolerance of signal dynamic range. Thus, the digital detector is a more robust detector in terms of varying ECG signal strength. 4) Based on the previous conclusions, it is clear that the IABP would benefit if the digital detector were to replace the analog detector.

Finally, it should be noted that in terms of the actual numbers, real-world performance should be better, as the data in the AHA database are particularly ectopic, intentionally, to give designers difficult ECGs to work with. In practice, more normal beats should be encountered.

DISCUSSION

We have presented an automated QRS-detector analysis system, the QRS detector error logger, which has applications in developing QRS detectors and in their evaluation. In an example application, we have used the error logger in the design and evaluation of a digital QRS detector to replace the antiquated analog design in a commercial IABP.

The importance of the error logger is that it permits designers to evaluate QRS detectors in approximately clinical environments, in that only the patient is simulated — the ECG database replaces the patient's ECG, and the entire QRS detection system is evaluated as if a patient were present.

The error logger has the potential for assisting medical equipment manufacturers during the design and verification phases of product development by providing a means for obtaining thorough and very realistic results for QRS-detector performance data. However, to use such a device, it is necessary to have access to an annotated ECG database to provide reference data. Nevertheless, when developing a QRS detector, the error logger can be used to obtain ROC curves as a function of various component values in the analog case, or algorithm parameters in the digital case. From these curves a designer can choose the operating point that best suits his or her desired application, and select parameter values based on such. Essentially, the error logger operates in real time and under very realistic conditions and thereby removes any uncertainty that might arise when using a computer simulation to evaluate a QRS detector. ■

REFERENCES

1. *Lindecrantz KG, Lilja H. New software QRS detector algorithm suitable for real time applications with low signal-to-noise ratios. J Biomed Eng. 1988;10:280-4.*
2. *Yu BC, Liu CS, Lee M, Chen CY, Chiang BN. A nonlinear digital filter for cardiac QRS complex detection. J Clin Eng. 1985;10(3):193-201.*
3. *Thakor NV, Webster JG, Tompkins WJ. Optimal QRS detector. Med Biol Eng Comput. 1983;21:343-50.*
4. *Thakor NV, Webster JG. Design and evaluation of QRS and noise detectors for ambulatory ECG monitors. Med Biol Eng Comput. 1982;20:709-14.*
5. *Okada M. A digital filter for QRS complex detection. IEEE Trans Biomed Eng. 1979;26:700-3.*
6. *Jadvar H, Jenkins JM, Stewart RE, Schwaiger M, Arzbaecher RC. Computer analysis of the electrocardiogram during esophageal pacing cardiac stress. IEEE Trans Biomed Eng. 1991;38:1089-99.*
7. *Friesen GM, Jannett TC, Jadallah MA, Yates SL, Quint*

SR, Nagle HT. A comparison of the noise sensitivity of nine QRS detection algorithms. IEEE Trans Biomed Eng. 1990;37:85-98.

8. *Suppappola S, Sun Y. Nonlinear transforms of ECG signals for digital QRS detection: a quantitative analysis. IEEE Trans Biomed Eng. 1994;41:397-400.*
9. *Sun Y, Suppappola S, Wrublewski TA. Microcontroller-based real-time QRS detection. Biomed Instrum Technol. 1992;26:477-84.*
10. *Hamilton PS, Tompkins WJ. Quantitative investigation of QRS detection rules using the MIT/BIH arrhythmia database. IEEE Trans Biomed Eng. 1986;33:1157-65.*
11. *Pan J, Tompkins WJ. A real-time QRS detection algorithm. IEEE Trans Biomed Eng. 1985;32:230-6.*
12. *Börjesson PO, Pahlm O, Sörnmo L, Nygårds M. Adaptive QRS detection based on maximum a posteriori estimation. IEEE Trans Biomed Eng. 1982;29:341-51.*
13. *Association for the Advancement of Medical Instrumentation, American National Standard for Cardiac Monitors, Heart Rate Meters and Alarms, ANSI/AAMI EC13-1983. Arlington, VA: AAMI, 1992.*
14. *Van Trees HL. Detection, Estimation, and Modulation Theory. Vol. I. New York: John Wiley and Sons, 1968:37-47.*
15. *Pahlm O, Sörnmo L. Software QRS detection in ambulatory monitoring — a review. Med Biol Eng Comput. 1984;22:289-97.*
16. *Suppappola S, Sun Y. Software to retrieve, review, and convert the AHA arrhythmia ECG database (abstr). J Am Coll Cardiol. 1992;19(3):7A.*