An SCP Compatible 12-Lead Electrocardiogram Database
for Signal Transmission, Storage, and Analysis

CC Chiang[1,2], YC Yang[1], WC Tzeng[1], WD Tseng[3], JC Hsieh[1]

[1]Department of Bioinformatics, Chung Hua University, Hsinchu, Taiwan
[2]Wei Gong Memorial Hospital, Miao-Li, Taiwan
[3]Department of Mathematics, University of British Columbia, BC, Canada

Abstract

12-lead Electrocardiogram (ECG) is one of the most important non-invasive and low-cost diagnostic examinations in clinical practice. However, different manufactory proprietary forms obstacle electronic ECG data transmission and data exchange. The primary objective of this study was to develop an inter-hospital web-based 12-lead ECG database server for ECG record exchange and analysis in the local area of Miao-Li County in Taiwan. The server was developed using software such as PHP, MySQL, and Matlab for SCP compatible format ECG record file transmission, format conversion, record storage, and signal analysis. In the past year, our efforts resulted in the collection of more than three thousands complete 12-lead SCP-ECG files. The results indicated that (1) SCP format compatible ECG record files can be exchanged and can be converted into ASCII and XML formats through the established server; (2) authorized end users can browse the database and analyze ECG signals by Matlab signal processing related toolboxes. In conclusion, the established ECG database server can provide effective medical informatics services such as online ECG pattern recognition and diagnosis for clinical physicians and ECG signal processing for researchers. In the future, our work may include other medical signals such as holter ECG and exercise ECG. More efforts will also extend to establish a country-wide inter-hospital medical signal database..

1. Introduction

Electrocardiogram has been the most commonly used biomedical test for the diagnosis of various heart diseases. However, due to the use of proprietary data format in almost all the commercial electrocardiograph, ECG's so obtained are not compatible among different brands of electrocardiograph manufactures. This not only makes the exchange of patient's ECG records difficult and inefficient, but hinders the research work in the analysis of relationship between ECG data and various heart diseases. To facilitate the acquiring of raw ECG data for research work and to promote the use of open ECG file format, we design a web-based ECG system in this study. The ECG database system, which includes a web-based interface for controlling the system, managing ECG data, and a Matlab-based toolbox for analysis and displaying ECG data [1]. At present, the database is collecting ECG data from patients with acute myocardial infarction, hyperkalemia, and hypokalemia. The collected data then can be analyzed using the programs in the Matlab-based toolbox to search for parameters which are ECG related characteristics for these two heart diseases. In addition, a utility to output the ECG data in XML format will be developed to facilitate the exchange of ECG data. To facilitate the compatibility of ECG data among electrocardiographs from different manufactures, OpenECG project has been promoting SCP-ECG (Standard Communications Protocol for Computer-Assisted Electrocardiography) standard format for over a decade [2]. Currently, although most of electrocardiograph manufactures claim that they have been supporting SCP-ECG format, the files output from commercial electrocardiographs still can only be interpreted by their proprietary viewers. Since there is a vast amount of these so called SCP-ECG data, they are valuable not only as patient's records, but also as raw data for researchers. Hence, in this study, we try to develop programs to decode the SCP-ECG files generated by the commercial electrocardiographs and store the results in the designed database.

2. Methods

2.1. SCP format analysis

A recorded SCP file is composed of two parts: (1) header information ; (2) compressed ECG waveform data.

2.1.1. Header information

The information stored in the header part are global parameters, annotations, and comments. The storage format in the header part is based on ASCII standard, and the basic storage unit is byte. There are a total of seventeen set of data rows to record those information in the header part. However, the start location and the full length of each data row are dependent on the location carrier tag represented by binary ODOA in two bytes. Basically, the extracted ECG parameters such as heart rate, p wave duration, QT interval, etc. are shown in 2nd data row and the half parts of the 6th to the 10th rows, and the ECG axis degrees are shown from the 13th to the 15th rows. The machine type and filter setting are shown on the 17th row, and the annotations are stored on portions of the 6th to the 15th rows. As shown in Figure 1, the extracted parameters of header part and their corresponding locations are described there.

```
0= 0D 0A 00 02 7D    *****              28-Sep-2003  10:29:19    Unknown
1= 0D 0A 00 03 03
2= 0D 0A 00 04 03
3= 0D 0A 00 05 09    *****
4= 0D 0A 00 06 5B    Rate   87   . AGE NOT ENTERED, ASSUMED TO BE 50 YEARS 1
5= 0D 0A 00 07 63    PR    166   . NORMAL SINUS RHYTHM, RATE  87...........
6= 0D 0A 00 08 63    QRSD   78   . BORDERLINE LOW VOLTAGE IN FRONTAL LEADS.
7= 0D 0A 00 09 63    QT    358   . CONSIDER ANTERIOR INFARCT..............
8= 0D 0A 00 0A 0D    QTc    43
9= 0D 0A 00 0B 03
10= 0D 0A 00 0C 0C   --Axis-
11= 0D 0A 00 0D 0D   P      32
12= 0D 0A 00 0E 0D   QRS    11 |
13= 0D 0A 00 0F 72   T       8   - BORDERLINE ECG -             Un
14= 0D 0A 00 10 70
15= 0D 0A 00 11 7D                     *****-***  25mm/s  10mm/mV  F~0.5Hz-40Hz \
16= 0D 0A 00 12 70
```

Figure 1: Header information in an SCP-compatible file.

2.1.2. ECG waveform

In the waveform storage part, a total of 37 sections can be found in a recorded SCP file. The beginning of each section of waveform part can be identified by a tag, represented by a two-byte binary number 1DXX. A two-byte binary number following the tag number is used to represent the half of full section length. 1D is the carrier tag and XX represents the waveform sequence tag. Therefore, the real section tag can be found correctly if the first section tag is identified. The standard 12-lead ECG waveform data recorded in a SCP file are stored in twelve short blocks and three long blocks. Each block is composed of two sections. However, each section can be divided into two subsections, which are called major and minor sections with the size ratio of about 3:1. The compressed ECG waveforms are stored in major sections in each waveform blocks. The data compression is based on the principle of modified Huffman encoding [2]. The values of ECG amplitude are represented by 3-bit, 5-bit,

7-bit, and 11-bit binary numbers respectively. The corresponding values of binary number length are summarized in Table 1.

Table-1: The corresponding values represented by modified Huffman Encoding

Bit length	Negative values	Positive values
3	-2,-1	0,1
5	-4,-3	2,3
7	-12~-5	4~11
11	-140~-13	12~139
16	Below -141	Above 140

2.2. Detection of ECG characteristic points

The extracted ECG waveform data from SCP files were stored as ASCII format and then were loaded into Matlab computing environment to detect ECG characteristic points. In this study, a dyadic wavelet using the first derivative of Gauss function was adopted to perform wavelet transform at scale 2^j, where j is an integer set, [1 2 3 4] representing a dilation scale factor. The dyadic wavelet transform can be expressed by equation (1)

$$Wf(2^j,\tau) = \frac{1}{\sqrt{2^j}} \int_{-\infty}^{\infty} f(t)W(\frac{t-\tau}{2^j})dt \qquad (1)$$

where Wf represents the transform coefficients at scale j, $f(t)$ represents the ECG signals, τ represents a shift factor, and W represents the mother wavelet which is the first derivative of real Gauss function.

2.2.1. QRS wave detection

The algorithms we used on detecting R-wave, P-wave, and T-wave were modified from Li's paper [3] and Sahambi's paper [4]. The transform coefficients, Wf, at scale $j=1$, were used to detect the R-peak and Q-wave and S-wave turn points. The corresponding location of R-wave location of each cardiac cycle can be identified on the minimum point of one-pair selected peaks (local maximum points in $j=1$ coefficients). The selected two peaks corresponding R-wave is based on the distance between two peaks. However, the onset of Q wave and the end of S wave were located by approximated turn points of two selected peaks.

2.2.2. Detection of T wave and P wave

Once the R wave was identified, two peaks of transform coefficients at scale j=3 after 120 ms of R wave were selected to locate T wave. The T wave peak was located on the minimum point between two selected peaks. The onset point was located at the point of 10% less than the left peak, and the end point of T wave was located approximately at the point of 10 % less than the right peak. In the same manner, The P wave can be identified by scale j=3.

3. Results

3.1. Database system

A recorded SCP file was decoded based on the modified Huffman encoding principle shown in Table 1 [2]. The extracted header data were placed in the *.tx0 in the ASCII format, and the waveform data including 12-lead short recorded data, with the recording time of 2.544 sec, and 3 long leads (I , II, III) recorded data, with the recording time of 9.552 sec, were stored in the files *.tx1, and *.tx2 respectively. For further inter-hospital data exchange, the hospital and department code stored in header part, will be identified and placed in the decoded files (tx0, tx1, tx1) under the hospital specific directories. The detail tree-like file structure was shown in Figure 2.
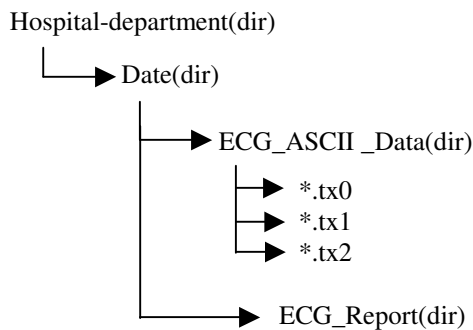
Hospital-department(dir)
    → Date(dir)
        → ECG_ASCII _Data(dir)
            → *.tx0
            → *.tx1
            → *.tx2
        → ECG_Report(dir)

Figure 2: File structure of the ECG database system.

3.2. ECG report generation

The extracted header and waveform data stored in *.tx0, *.tx1, *.tx2 files, were reloaded into the Matlab computing environment through a Matlab web-server to generate clinical-like 12-lead paper ECG report. As shown in Figure 3, the clinical ECG report was saved in pdf format. For further application, the generated report will be converted to DICOM and XML format.
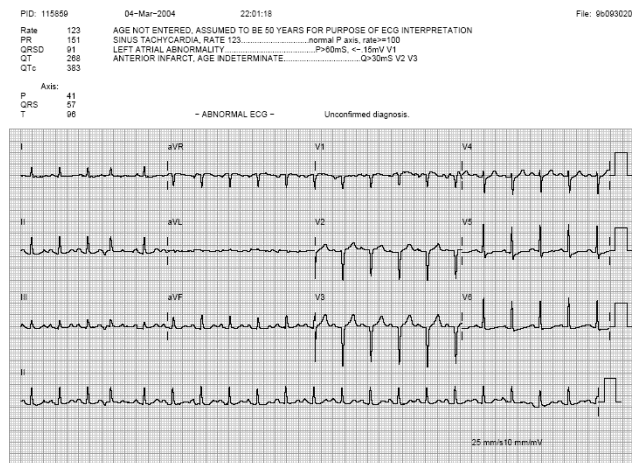

Figure 3: A clinical-like ECG report generated in Matlab environment

3.3. ECG signal processing

As shown in Figure 4, the characteristic points of ECG were identified by the wavelet transforms. The peaks of QRS wave, P wave and T wave were marked by the symbol *, and the onset/end points for waves were marked by the +. In Figure 4, the ECG signals were extracted from lead II.
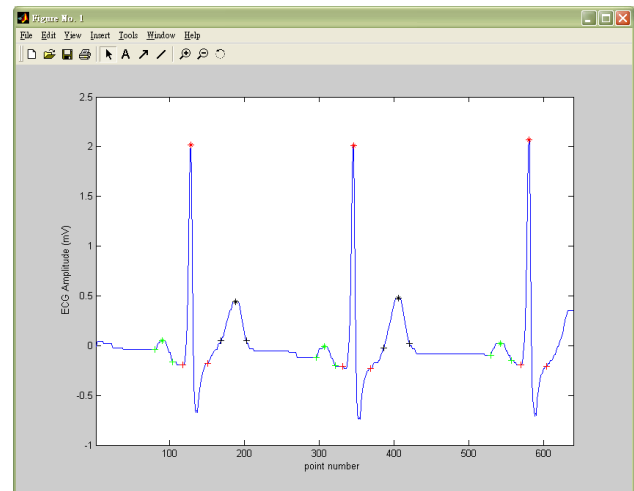

Figure 4: Using wavelet transforms to identify the characteristic points of ECG.

4. Discussion and conclusions

The complete web-based ECG database in this study was constructed as shown in Figure 5. The recorded SCP

compatible files can be uploaded into the database through web to perform format conversions (from SCP to ASCII and XML). Generally, the ECG web server was constructed by Apache and PHP for interactive effects. The extracted data including header information and waveform data from SCP files were stored in MySQL database. For analysis of ECG signal, Matlab scripts were implanted into the system under a Matlab web server to perform ECG-signal processing and result reporting. Until now, three common diseases, acute myocardial infarction, hyperkalemia, and hypokalemia in clinical emergency practice were collected and stored in the database. In order to integrate with the current medical health information system, the printable paper 12-lead ECG report will be needed for directly converting to DICOM format.
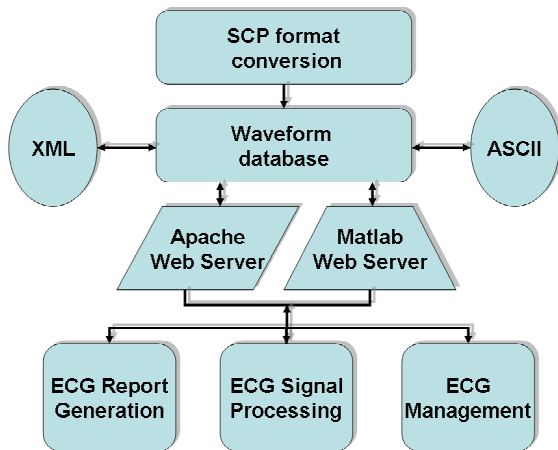
Figure 5: An SCP-ECG database system implanted with Matlab application code.

Acknowledgements

References

[1] Matlab web-server and wavelet toolbox manual. Matlab Inc. http://www.mathworks.com
[2] SCP-ECG manual. Open ECG project. http://www.openecg.net
[3] Li C, Zheng C, and Tai C. Detection of ECG characteristic points using wavelet transform. IEEE transaction on biomedical engineering 1995:42(1):22-28.
[4] Sahambi JS, Tandon SN, and Bhatt RKP. Using wavelet transforms for ECG characterization. IEEE Engineering in medicine and biology 1997:1:77-83.

Address for correspondence

Jui-Chien Hsieh, Ph.D.

Department of Bioinformatics
Chung Hua University
No.707, Sec.2, WuFu Rd.
Hsinchu, Taiwan 300, R.O.C.
E-mail address : jchsieh@chu.edu.tw