Proceedings of the 2005 IEEE
Engineering in Medicine and Biology 27th Annual Conference
Shanghai, China, September 1-4, 2005

Information Flow and Data Reduction in the ECG Interpretation Process

Ryszard Tadeusiewicz, *Senior Member, IEEE*
Piotr Augustyniak, *Member, IEEE*
AGH University of Science and Technology
30, Mickiewicza Ave.
30-059 Krakow Poland

Abstract— **In result of ECG interpretation process the diagnostic outcome summarizes all the principal information included in the raw recording. This process is investigated in this paper in data flow and data reduction aspects being of particular impotrance for implementation of distributed ECG interpretation in a wireless network. The detailed analysis of existing software and cardiologists societies guidelines reduces the whole interpretation process to the network of modules interconnected by data busses. Each module is then described by the computational reliability and by expected data streams at its inputs and outputs, from which the average data reduction ratio is computed. These parameters are consequently used for a design of interpretation architecture improving the diagnostic reliability and oriented to appropriate task sharing between the remote wearable recorder and the interpretive center.**

I. Introduction

The implementations of automatic ECG interpretation are currently widespread in many devices ranging from large servers to small wearable cardiomonitors. The appearance of autonomous recorders for real-time monitoring and interpretation, communicating over a wireless digital link [1-4] implied new prerequisites for the interpretive software architecture. The first of considered aspects is the diagnostic reliability directly affecting the interpretation autonomy and the need for external supervision. The second aspect is the local data reduction ratio which is the main condition for maintaining the communication service expenses within the margins of acceptance.

Despite our studies bases on algorithms from various manufacturers [5-9], all investigated applications follow a very similar architecture originating from their version histories or the upgradeable modules concept. This functional-growth architecture is very convenient for manufacturers tailoring the software for diverse users from the same bricks [10]. Nevertheless that approach neither optimizes the diagnostic reliability, neither aims at the effective reduction of data stream.

Scientific work financed from the State Committee for Scientific Research resources in years 2004-2007 as a research project No. 3 T11E 00127

In this paper we investigate the existing interpretive software for electrocardiographs and generalize rules concerning the optimal architecture for satisfying both previously mentioned criteria: high data reduction ratio and high immunity to the errors at the subsequent stages of processing chain.

II. Material and Methods

The aim of our research is to explore the opportunity for improvement of wearable ECG recorders performance opened by the structural rearrangement of the interpretive software. Therefore, we consider diagnostic subroutines as black-boxes and never attempt to modify the interpretive algorithms or their mathematical foundations. In the experimental part of the research we used standard ECG databases: MIT-BIH [11] and CSE [12] recommended for testing of the software performance and a standard interpretive software designed to be embedded in a stand-alone ECG machine. In order to estimate the probability of each function call, we assume that particular heart diseases are proportionally represented in the database.

A. Pursuit of the Incorrect Interpretation Probability

Each heuristic subroutine in the ECG interpretation chain shows a non-zero probability of inappropriate processing and incorrect outcome. Despite of applying very thorough testing procedures, no software engineer is able to foresee all possible signal recording conditions combined with all possible heart diseases. In a typical processing chain, subsequent procedures use results of the previous ones, thus misinterpretations and inaccuracies may compensate or cumulate [13]. Unfortunately, the error-cumulative scenario is much more probable, because statistically writing, the correct and accurate result is a singularity in a cloud of all possible outcomes. Three approaches aiming at reduction of overall error probability may be applied separately or in combination:

- reducing of the processing chain length,
- using the most accurate procedures at the front of processing chain,
- applying the auto-assessment functions.

0-7803-8740-6/05/$20.00 ©2005 IEEE.

The first two methods were subject of our studies, since the auto-assessment functions are often already implemented as a part of diagnostic subroutines. Moreover, the investigated methods do not require the additional computation power.

B. Investigations of the Data Reduction Efficiency

Each interpretive subroutine transforms the data fed to its inputs into results at its outputs. Since the whole interpretation process leads from a collection of signal samples (sometimes of diverse nature) to the final diagnostic outcome, from a statistical viewpoint it is a sort of data reduction tasks. For each specialized subroutine the information stream reduction ratio can be quantitatively measured by comparing the expected throughput of the outputs and the inputs. The effective data reduction at the beginning of the process, postulated by the wearable recorders implementation can be achieved with use of the following methods:

- putting the most reduction-effective procedures in front of the processing chain,
- putting the most frequently used procedures in front of the processing chain.

The particular challenge is the consolidation of all functions having access to the raw signal at their inputs. Unfortunately, the meaning of many signal-derived parameters depends on advanced calculations in long processing chains.

C. Optimization of Software Architecture

As shown in fig. 1, a typical data flow diagram for the basic ECG interpretation process contains many references to the raw signal. Certainly, the software architecture rearrangements are constrained by the logical flow of ECG diagnostic procedure. In some rare cases the data processing chain has to follow a specified order providing first the general information (e.g. a heart beat was detected) and then more precise details (e.g. morphology type or waves length). Within these constraints the reduction-effective and frequently used procedures were identified in the first experiment.

First experiment aimed at estimating a set of statistical parameters for each interpretive procedure:

- δ - outcome relative inaccuracy [%],
- ε - probability of false outcome [%],
- r - data reduction ratio,
- p - probability of use [%] (depending on the frequency of related disease occurrence).

In order to favorite more frequent, more accurate and more reduction-effective functions in access to the data streams, each procedure were attributed a priority level derived from estimated statistical parameters (table 1).

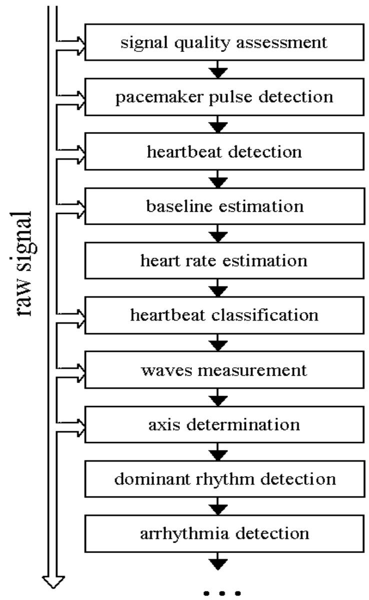

Fig. 1. Typical data flow diagram for the basic ECG interpretation process

TABLE 1.
BASIC ECG INTERPRETATION PROCEDURES, THEIR STATISTICAL PARAMETERS AND ATTRIBUTED PRIORITY LEVELS

Procedure Name	Statistical Parameters				Priority Level
	δ	ε	r	p	
signal quality assessment	10	3,3	20	97	1
pacemaker pulse detection	<1	8,3	70	3	4
heartbeat detection	1,5	2,5	70	100	2
baseline estimation	3	<1	20	97	3
heart rate estimation	<1	<1	1	100	1
heartbeat classification	10	3	50	88	1
waves measurement	3	5	100	85	2
axis determination	3	5	300	85	3
dominant rhythm detection	0	8,5	1,5	100	1
arrhythmia detection	0	10	1,3	80	2
……					

D. Data Busses Concept

A second, complementary approach to the interpretive software optimization bases on a concept of data busses. These inter-procedure information channels are sorted by the value of expected throughput (fig. 2). Each data flow was assigned a throughput level combining statistical parameters of the data:

- average data stream when used,
- frequency of use (depending on data refresh rate)
- probability of use (depending on the frequency of related disease occurrence).

The concept of data busses limits the access to the unprocessed data representation to the subroutines of high peak throughput and high frequency of use. Consequently, several such procedures may be re-designed towards the use of a common and very reliable data interface. In case of waves measurement and axis determination procedures, highly dependent on the previous computation results, two solutions were considered:

- continuous direct connection to the raw signal bus and approximate estimation of calculation starting points,
- occasional connection to a buffer cashing a copy of raw signal dependent on the detection points position.

Since waves measurement and axis determination performance dramatically drops with the use of estimate instead of accurate detection points, the second method was finally applied.

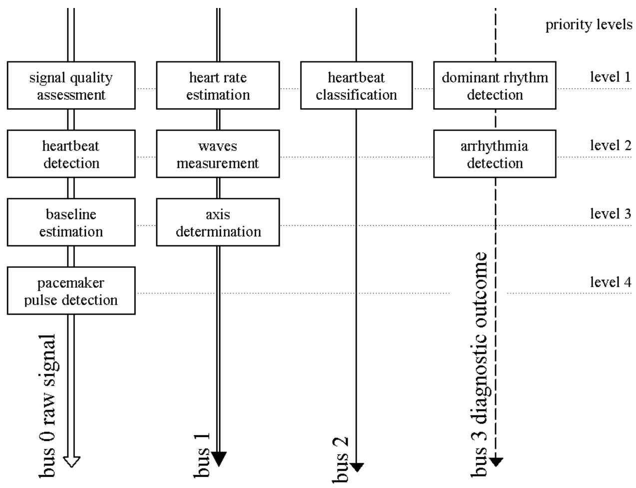

Fig. 2. Data bus concept combined with interpretation procedures priority levels

The final architecture of ECG interpretation software optimized for reliability and early reduction of data stream (fig. 3) contains three raw signal access points:

- common interface for signal quality estimation, baseline estimation and pacemaker detection procedures,
- heartbeat detection filtered input
- waves measurement and axis determination filtered off-line buffer (provided also for ST measurement and P-averaging not considered here).

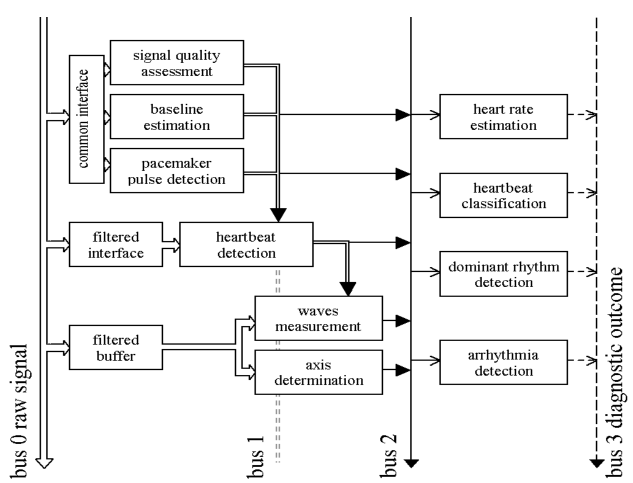

Fig. 3. The final architecture of ECG interpretation software optimized for reliability and early data stream reduction

The group of functions accessing the raw signal issues a complete description of a heartbeat (bus 2), which is not the diagnostic outcome, but contains all the meta-data necessary for further processing and thus the raw signal is no longer necessary. These data appear occasionally, once per heartbeat, but even for as high heart rate as 180 bps, the data stream is 8.2 times lower than for a 12-lead 500 sps raw signal. Should the remaining interpretation procedures be performed by another node of a distributed network, breaking the local processing here is a convenient solution.

III. Experimental Results

The architecture optimization was performed on a standard ECG interpretation software provided for tests by a regional manufacturer. The structured source code was written in C++ programming language. The original target application was the interpretive bedside ECG recorder, and the purpose of the optimization was a manufacturer's interest in migrating to a wearable computer platform. Two factors were assessed separately for the original and the modified architectures:

- average reduction of data rate at the subsequent processing stages,
- average inaccuracy and error probability for selected diagnostic parameters.

In result of architecture optimization, the processing stages could not be set in a similar way for the original and for the modified software. Therefore, with regard to the identical testing conditions, we used the processing time as an estimation of the interpretation progress. The processing stages were set at every 20 percent of the total interpretation time. The reason for this relative approach was twofold:

- particular ECG files varies in processing time,
- although not intended, the software re-design shortened the average processing time.

Table 2. compares the average data reduction ratio on subsequent stages of interpretation process. The right column highlights the difference between the original and optimized architectures, and proves that significant data reduction was achieved by the modified architecture on the early stages of interpretation process.

TABLE 2.
AVERAGE DATA REDUCTION RATIO [%] ON SUBSEQUENT STAGES OF INTERPRETATION PROCESS

Interpretation Progress ([%] of total processing time)	Data Reduction Related to Raw Signal [%]		Data Reduction Gain [%]
	Original Architecture	Optimized Architecture	
0	100	100	0
20	78	47	31
40	54	31	23
60	32	22	10
80	14	12	2
100	8	8	0

The second aspect of architecture optimization, the result accuracy, was tested accordingly to the international standards requirements [14]. The quantitative results for both architectures are summarized in table 3. The comparison of the diagnostic reliability for isolated procedures (table 1) with the corresponding results of the whole processing chain (table 3) leads to the conclusion that in case of the optimized architecture, the overall reliability of each parameter is much less affected by the remaining procedures of ECG interpretation chain.

TABLE 3.
DIAGNOSTIC PARAMETERS QUALITY ACHIEVED BY THE ORIGINAL AND THE OPTIMIZED ARCHITECTURES[1].

Interpretation Domain	Statistical Parameters			
	Original Architecture		Optimized Architecture	
	δ	ε	δ	ε
pacemaker pulse detection	2.8	9.3	1.5	9.0
heartbeat detection	2.5	3.5	1.7	2.9
baseline estimation	4.3	1.3	4.3	1.3
heart rate estimation	1.0	1.2	1.0	1.2
heartbeat classification	14	7.1	12	4
waves measurement	5.1	7.5	3.3	5.3
axis determination	6.3	7.8	3.7	5.1
dominant rhythm detection	0	10.5	0	8.8
arrhythmia detection	0	13	0	11.8

This improvement is achieved by shortening the processing chains. Consequently, the dependence of subsequent functions input data and the results of previous processing stages is looser and no longer favorites the cumulative error propagation.

IV. DISCUSSION

The work presented in this paper was motivated by recent changes in cardiac monitoring techniques made towards the applications of modern digital communication technology. The classical approach to the ECG interpretation processing chain was revised and important software architecture modifications were proposed to overcome two principal drawbacks:

- necessity of raw signal access on the advanced processing stages,
- cumulative error propagation resulting from data dependencies in the processing chain.

Both aspects were thoroughly studied and applied to the real interpretive software, taking an opportunity of co-operation with the ECG equipment manufacturer. The modular software was modified only at the subroutines interconnection level without changes or adjustment of mathematical methods. Main result is the relative improvement of diagnostic outcome accuracy and data stream reduction, rather than their absolute values. Therefore any manufacturer may check his software for the concordance with the guidelines issued herein.

The aim of our research was fully achieved. We proved that the software architecture optimization is suitable for interpretation improvement in the following areas:

- moves reduction-effective functions to the front of processing chain and consequently reduces the inter-procedures data flow, and thus lowers the communication costs in case of distributed processing,
- reduces the cumulative error propagation by the parallel use of multiple short processing chains instead of one long chain,
- reduces the interpretation processing time and the required computational power, and thus extends the wearable devices autonomy time.

Certainly, the results are expected even more promissing if the particular processing parameters were fully independent. In such case, however, many processing steps have to be repeated on the same signal. Nevertheless, as a future consideration we propose to completely rewrite each of the interpretive function with a goal of minimize the use of previously computed parameters.

[1] The meaning of δ and ε is explained in section II.C

REFERENCES

[1] F. Chiarugi, D. Trypakis, V. Kontogiannis, PJ. Lees, CE. Chronaki, M. Zeaki, N. Giannakoudakis, D. Vourvahakis, M. Tsiknakis, "Continuous ECG Monitoring in the Management of Pre-Hospital Health Emergencies" Computers in Cardiology 2003, Vol. 30, pp. 205-208

[2] GD. Pinna, R. Maestri, E. Gobbi, MT. La Rovere, JL. Scanferlato "Home Telemonitoring of Chronic Heart Failure Patients: Novel System Architecture of the Home or Hospital in Heart Failure Study" Computers in Cardiology 2003, Vol. 30, pp. 105-108,

[3] K. A. Banitsas, P. Georgiadis, S. Tachakra, D. Cavouras Using handheld devices for real-time wireless Teleconsultation". Proc. 26th Annual International Conference of the IEEE EMBS, 2004 pp. 3105-3108

[4] Bar-Or, J. Healey, L. Kontothanassis, JM. Van Thong "BioStream: A system architecture for real-time processing of physiological signals". Proc. 26th Annual International Conference of the IEEE EMBS, 2004 pp. 3101-3104

[5] IBM Electrocardiogram Analysis Program Physician's Guide (5736-H15) 2-nd edition, 1974

[6] HP M1700A Interpretive Cardiograph Physician's Guide ed. 4. Hewlett-Packard, 1994

[7] DRG MediArc Premier IV Operator's Manual version 2.2. 1995

[8] ECAPS-12C User Guide: Interpretation Standard revision A. Nihon Kohden. 2001

[9] CardioSoft Version 6.0 Operator's Manual. GE Medical Systems Information Technologies, Inc. Milwaukee, 2005

[10] M. Paoletti, C. Marchesi "Low computational cost algorithms for portable ECG monitoring units" IFMBE Proc. Medicon 2004, paper 231

[11] G. Moody MIT/BIH Arrhythmia Database Distribution, Massachusetts Institute of Technology, Division of Health Science and Technology, Cambridge, MA, 1993

[12] JL. Willems Common Standards for Quantitative Electrocardiography 10-th CSE Progress Report, 1990. Leuven: ACCO publ.,

[13] E. Straszecka, J. Straszecka "Uncertainty and imprecision representation in medical diagnostic rules" IFMBE Proc, Medicon 2004, paper 172

[14] IEC 60601-2-47 Medical electrical equipment: Particular requirements for the safety, including essential performance, of ambulatory electrocardiographic systems. 2001