

Д. В. Исакевич, В. В. Исакевич

Программная реализация устройства хранения и анализа электрокардиограмм

Аннотация. Рассмотрена программная реализация подсистемы сжатия данных устройства хранения и анализа электрокардиограмм (заявка на полезную модель РФ №2015130397/14(04627) от 23.07.2015 [1]) на базе вычислительного комплекса «Монокуб-РС» производства ЗАО «МЦСТ». Описаны конструкция и предварительные характеристики подсистемы (производительность и точность восстановления данных).

Ключевые слова и фразы: собственный вектор, сжатие, электрокардиограмма, биомедицинские данные, главная компонента, айгеноскоп.

Введение

Сжатие электрокардиограмм (ЭКГ) является актуальной технической задачей в связи с ростом объёмов хранимых и обрабатываемых записей — что неизбежно при переходе к индивидуальной медицине.

Такое сжатие возможно как известными стандартными методами сжатия (дельта-кодирование, кодирование Хаффмана и по длинам серий и др. [2]), так и специальными методами, учитывающими специфические особенности электрокардиографических сигналов. Далее ограничим рассмотрение специальными методами, памятуя о возможности последующего применения к сжатой ЭКГ известных стандартных методов.

Для сжатия электрокардиограмм (ЭКГ) широко используются представления в базисах вейвлет-функций [3–7]. Вместе с тем, общеизвестны методы (дискретное разложение Карунена–Лоэва, анализ главных компонент, анализ сингулярного спектра), использующие оптимальное представление многомерных данных в адаптивном

базисе собственных векторов ковариационной матрицы. Такое представление обеспечивает наименьшее (для конкретных данных) число значимых коэффициентов разложения. Эти методы широко применяются для обработки и анализа многомерных данных, а также снижения их размерности и сжатия [8, 9]. При работе с одномерными временными рядами используется приём развёртки ряда в траекторную матрицу с последующим сингулярным разложением [10].

Очевидно, что развёртка временного ряда в траекторную матрицу искусственно повышает размерность пространства, в котором ведётся анализ; таким образом, траекторная матрица избыточна. Эта избыточность приводит к тому, что спектр собственных значений оказывается весьма широким [11] (линия с квадратами на рисунке 2), а сжатие ЭКГ — неэффективно.

Ранее для анализа ЭКГ авторами была предложена конструкция «Кардиоайгеноскоп» (патент на полезную модель №128470 РФ [12]), использующая синхронизацию с R-пиками ЭКГ для развёртки временного ряда в матрицу (матрицу ансамбля). Авторами было показано, что такая синхронизация позволяет получить ансамбль, имеющий наименьшую размерность и наиболее экономное представление [11]. Такой ансамбль предложено называть синхронным.

Представление ЭКГ в базисе собственных векторов ковариационной матрицы синхронного ансамбля используется в предложенной конструкции устройства хранения и анализа электрокардиограмм [1]. Данная конструкция состоит из трёх блоков — кардиоайгеноскопа [12], блока авторизации и доступа, а также блока задания режимов работы, контроля сжатия, хранения и выборки. Ключевым блоком устройства хранения и анализа ЭКГ является кардиоайгеноскоп — именно он осуществляет сжатие ЭКГ и восстановление сигнала из сжатой формы. Рассмотрим реализацию этого блока.

Описание программной реализации и результатов

Прототип подсистемы сжатия ЭКГ реализован на базе вычислительного комплекса «Монокуб-РС» производства ЗАО «МЦСТ». Прототип функционирует на базе ОС «Эльбрус» и представляет собой программу, используемую в консольном режиме для обработки исходных записей ЭКГ. Каждая из исходных записей содержит последовательность значений отсчётов ЭКГ в текстовом формате представления чисел с плавающей точкой. Сжатая запись представляет

собой файл, содержащий (в текстовом формате представления чисел):

- (1) частоту дискретизации;
- (2) оценку времени, затраченного на вычисления;
- (3) число выявленных R-пиков и их положение на временной оси;
- (4) число собственных векторов, использованных для сжатия, и сами собственные векторы;
- (5) коэффициенты разложения синхронного ансамбля по собственным векторам.

Работа программы сжатия организована в соответствии с алгоритмами, описанными авторами ранее [11]. Из исходной записи ЭКГ удаляется медленно изменяющаяся компонента. В оставшейся после удаления части ЭКГ путём порогового детектирования определяется положение R-пиков; за R-пик принимается середина интервала, на котором порог превышен. Положение R-пиков может быть определено и другими способами, например, корреляционным детектированием [13, 14].

Полученные положения R-пиков используются для формирования синхронного ансамбля. Синхронный ансамбль представляет собой матрицу, в качестве столбцов которой берутся отрезки временного ряда исходной ЭКГ, включающие заданное число отсчётов слева и справа от R-пиков. Таким образом, положение R-пиков во всех столбцах матрицы синхронного ансамбля оказывается одним и тем же — что и обуславливает низкую размерность представления. Из каждого столбца ансамбля удаляется компонента, представляющая линейный и/или квадратичный тренд, таким образом, чтобы первый и последний отсчёт столбца были равны нулю.

Для полученного синхронного ансамбля \mathbf{Y} размера $M \times L$ вычисляются собственные векторы-столбцы \mathbf{v}_k (рисунок 1) ковариационной матрицы

$$(1) \quad \mathbf{K} = \frac{\mathbf{Y}\mathbf{Y}'}{L}$$

и собственные значения λ_k этой матрицы, нормированные к её следу (рисунок 2, линия с кружками). Собственные векторы и собственные значения вычисляются с использованием метода исчерпывания [15]; поэтому прежде всего находятся те собственные векторы, которые имеют наибольшие собственные значения. Расчёт прекращается тогда, когда сумма первых m найденных нормированных собственных

значений λ_k приблизится к единице с заданной точностью ε

$$(2) \quad 1 - \frac{\sum_{k=1}^m \lambda_k}{\text{Tr} \mathbf{K}} < \varepsilon$$

Выражение (2) означает, что первые m собственных векторов описывают не менее $(1 - \varepsilon)$ от средней энергии ЭКГ на конечном интервале анализа, синхронизированном с R-пиками [11]; при этом рассматривается ЭКГ, из которой удалён линейный тренд. Иначе говоря, представление в ортонормированном базисе $m < M$ собственных векторов позволяет описать ЭКГ с относительной среднеквадратичной погрешностью не более ε .

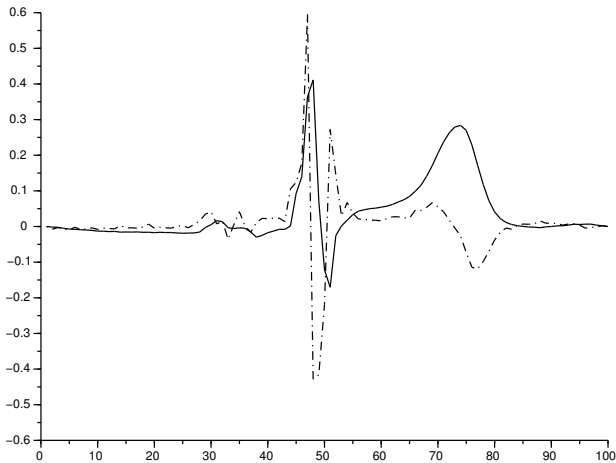


Рис. 1. Собственные векторы №1 (сплошная линия) и №2 (штрих-пунктирная линия) ковариационной матрицы синхронизированного с R-пиками ансамбля для временного ряда ЭКГ. Частота дискретизации 100 Гц. Длина интервала анализа 100 отсчётов. Объём ансамбля — 50 элементов.

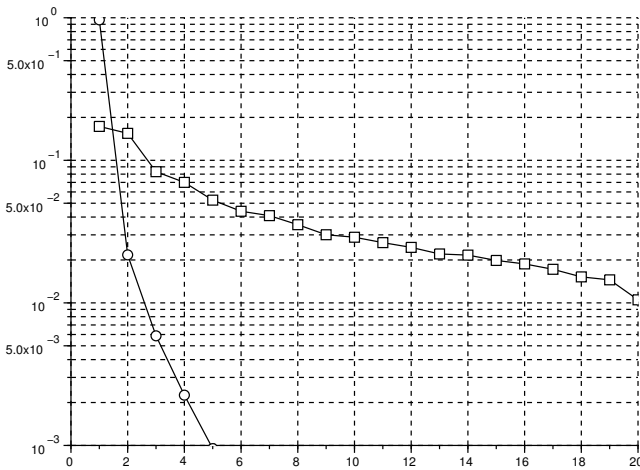


Рис. 2. Нормированные к следу ковариационной матрицы собственные значения, упорядоченные по убыванию. Линия с квадратами — собственные значения, полученные для траекторной матрицы после удаления из её столбцов постоянной составляющей. Линия с кружками — собственные значения для синхронизированного с R-пиками ансамбля. Исходная ЭКГ (диагноз «Здоров») [16, 17]. Частота дискретизации 100 Гц. Число строк траекторной матрицы — 100 отсчётов. Число столбцов траекторной матрицы — 50 элементов.

По найденным m собственным векторам \mathbf{v}_k для матрицы ансамбля \mathbf{Y} вычисляются коэффициенты разложения (главные компоненты [10]) \mathbf{H}

$$(3) \quad \mathbf{H} = \mathbf{Y}' \mathbf{v}_k$$

Отметим, что общее число получаемых коэффициентов разложения $m \times L$ для синхронного ансамбля оказывается существенно меньше, чем $M \times L$. На этом и основан эффект сжатия.

Для испытания программы сжатия использованы ЭКГ с частотой дискретизации 1000 Гц (длительность записи — до 100 сек), взятые

из открытой базы ЭКГ [16, 17]. Каждая запись подвергалась предварительной децимации с коэффициентом $K = 10$; таким образом, обрабатывались временные ряды с частотой дискретизации 100 Гц. Интервал анализа ЭКГ при построении синхронного ансамбля содержал 50 отсчётов временного ряда слева от R-пика и 49 отсчётов справа от R-пика; общая длина интервала анализа — 100 отсчётов (одна секунда). Наибольшая относительная погрешность представления ЭКГ $\varepsilon = 10^{-2}$.

Как правило, представление ЭКГ с заданной точностью требует от двух до четырёх собственных векторов; случай, представленный на рисунке 2, является типичным. Больше число собственных векторов требуется в случаях, когда ЭКГ носит нерегулярный характер (например, при аритмическом синдроме) и/или содержит существенные помехи. В первом случае пациент требует особого внимания, во втором — собственные векторы, содержащие помехи, можно исключить из рассмотрения.

Достигнутый с использованием программы коэффициент сжатия равен 12.5. Полученный результат близок к теоретической оценке коэффициента сжатия ЭКГ [1], рассчитанной для обработанных записей

$$(4) \quad K_{\text{сж.}}^{\text{теор.}} = \frac{L \cdot M}{Z \cdot M + Z + Z \cdot L + L + 1} \approx 13.2$$

где L — объём синхронного ансамбля (число R-пиков);

M — длина интервала анализа ЭКГ;

Z — число собственных векторов, используемых для представления ЭКГ на интервале анализа.

Результаты сравнения коэффициентов сжатия, полученных с использованием разработанной программы сжатия и стандартных архиваторов, приведены в таблице 1.

Оценка времени, затраченного на вычисления, составила в среднем 10.5 мсек для одной записи по одному отведению; с учётом временных затрат на ввод-вывод — в среднем 47 мсек для одной записи по одному отведению. Все вычисления производились с одинарной точностью. Оптимизация при компиляции программы не производилась. Для создания прототипа использованы библиотека математических вычислений EML и компилятор LCC, поставляемые с вычислительным комплексом «Монокуб-РС».

Таблица 1. Эффективность разных способов сжатия ЭКГ

Способ сжатия	Коэффициент сжатия
Программа сжатия ЭКГ	12.5
Архиватор zip	4.6
Архиватор tar + Архиватор gzip	4.6
Программа сжатия ЭКГ + Архиватор zip	28.5
Программа сжатия ЭКГ + Архиватор tar + Архиватор gzip	29.7

Заключение

Представленная программная реализация сжатия электрокардиограмм на базе вычислительного комплекса «Монокуб-РС» позволяет осуществлять многократное сжатие электрокардиограмм. Один вычислительный комплекс способен обслуживать полумиллионный город при ежемесячном обследовании каждого жителя. При таком обслуживании суммарный объём записей, накапливаемых ежемесячно, не превышает 7 Гбайт — что позволяет сохранить все ЭКГ города за 20 лет на одном носителе.

Список литературы

- [1] Исакевич В.В., Исакевич Д.В., Сушкова Л.Т., Аль-Барати Б.С.. Устройство хранения и анализа ЭКГ: Заявка на патент РФ на полезную модель, 23.07.2015, № 2015130397/14(04627). ↑ 1, 2, 6.
- [2] Смит С.. *Форматы сжатия данных* // Электронные компоненты, 2009, № 8, с. 83–87. ↑ 1.
- [3] Djohan A., Nguyen T.Q., Tompkins W.J. *ECG compression using discrete symmetric wavelet transform* // Engineering in Medicine and Biology Society, 1995., IEEE 17th Annual Conference IEEE, 1995. Vol. 1, p. 167–168. ↑ 1.
- [4] Istepanian R., Petrosian A., et al. *Optimal zonal wavelet-based ECG data compression for a mobile telecardiology system* // Information Technology in Biomedicine, IEEE Transactions on, 2000. Vol. 4, no. 3, p. 200–211. ↑ 1.
- [5] Chen J., Itoh S. *A wavelet transform-based ECG compression method guaranteeing desired signal quality* // IEEE Transactions on Biomedical Engineering, 1998. Vol. 45, no. 12, p. 1414–1419. ↑ 1.
- [6] Miaou S.-G., Yen H.-L., Lin C.-L. *Wavelet-based ECG compression using dynamic vector quantization with tree codevectors in single codebook* // Biomedical Engineering, IEEE Transactions on, 2002. Vol. 49, no. 7, p. 671–680. ↑ 1.
- [7] Kim B., Yoo S., Lee M. *Wavelet-based low-delay ECG compression algorithm for continuous ECG transmission* // Information Technology in Biomedicine, IEEE Transactions on, 2006. Vol. 10, no. 1, p. 77–83. ↑ 1.

- [8] Каханер Д., Молер К., Нэш С.. Численные методы и программное обеспечение. М.: Мир, 2001. — 576 с. ↑ 2.
- [9] Айвазян С.А., Бухштабер В.М., Енюков И.С., Мешалкин Л.Д.. Прикладная статистика: Классификация и снижение размерности. Часть 3. М.: Финансы и статистика, 1989. — 608 с. ↑ 2.
- [10] Данилов Д.А., Жиглявский А.А.. Главные компоненты временных рядов: метод «Гусеница». СПб.: Санкт-Петербургский университет, 1997. ↑ 2, 5.
- [11] Исакевич Д.В., Исакевич В.В.. Кардиоайгеноскопия — новая полезная модель обработки электрокардиограмм: Электронное издание. М.: Издательство «Перо», 2014. — 139 с. ↑ 2, 3, 4.
- [12] Исакевич В.В., Исакевич Д.В., Батин А.С.. Кардиоайгеноскоп: Патент РФ на полезную модель, 27.05.2013, № 128470. ↑ 2.
- [13] Зайченко К.В., Краснова А.И.. *Методика решения проблемы высокоточной синхронизации квазипериодических сигналов* // Электроника и связь. 4, 2008, с. 27–29. ↑ 3.
- [14] Зайченко К.В., Зяблицкий А.В.. *Высокоточная синхронизация электрокардиосигналов сверхвысокого разрешения* // Известия высших учебных заведений России. Радиоэлектроника, 2010, № 1, с. 40–48. ↑ 3.
- [15] Фаддеев Д.К., Фаддеева В.Н.. Вычислительные методы линейной алгебры. СПб.: Издательство «Лань», 2002. — 736 с. ↑ 3.
- [16] Boussejot R., Kreiseler D., Schnabel A. *Nutzung der EKG-Signaldatenbank CARDIODAT der PTB über das Internet* // Biomedizinische Technik. 40, Ergänzungsband 1, 1995, p. 317. ↑ 5, 6.
- [17] Kreiseler D., Boussejot R. *Automatisierte EKG-Auswertung mit Hilfe der EKG-Signaldatenbank CARDIODAT der PTB* // Biomedizinische Technik. 40, Ergänzungsband 1, 1995, p. 319. ↑ 5, 6.

Об авторах:

Даниил Валерьевич Исакевич

Автор подробной типологии собственных векторов временных рядов, быстрой ($O(M^2)$) оценки размерности представления сигнала в базисе собственных векторов, ряда полезных моделей и программ для автоматизации управленческого учёта, анализа и обработки данных.

e-mail:

eigenoscope@yandex.ru

Валерий Викторович Исакевич

Автор понятия «айгеноскопия», ряда изобретений, полезных моделей и программ в областях связи, анализа и обработки сигналов, а также управленческого учёта, в том числе Meditools — первого (ещё под MS DOS) отечественного программного пакета для ведения учёта в медицинских учреждениях и анализа медико-демографических данных.

e-mail:

eigenoscope@yandex.ru

Пример ссылки на эту публикацию:

Д. В. Исакевич, В. В. Исакевич. «Программная реализация устройства хранения и анализа электрокардиограмм», *Программные системы: теория и приложения*, 2015, ??:?, с. ??–??.

URL

<http://psta.psiras.ru/read/>

Daniel Isakevich, Valery Isakevich. *Software implementation of electrocardiogram storage and analysis utility*.

ABSTRACT. A software implementation of data compression subsystem in the electrocardiogram storage and analysis utility is described. The preliminary characteristics of data compression subsystem are presented. (*in Russian*).

Key Words and Phrases: eigenvector, compression, electrocardiogram, biomedical data, principal component, eigenoscope.

Sample citation of this publication

Daniel Isakevich, Valery Isakevich “Software implementation of electrocardiogram storage and analysis utility”, *Program systems: theory and applications*, 2015, ??:?, pp. ??–??. (*In Russian*.)

URL

<http://psta.psiras.ru/read/>