

ИДЕНТИФИКАЦИЯ ЧЕЛОВЕКА ПО ЭКГ

И.К. Рысцов^{*}, Л.А. Кизуб^{**}, В.В. Вишнеvский^{**}

^{*}Национальный технический университет Украины «КПИ»,
^{**}Институт проблем математических машин и систем НАН Украины
e-mail: moztart@voliacable.com

Введение

При разработке транспортных ВЕБ-сервисов проекта «Медгpид» [1] мы столкнулись с трудностями проведения парсинга (связывания) деперсонализированных электрокардиограмм с записью конкретного пациента в метабазе данных проекта. В свою очередь, трудоемкая процедура парсинга файлов кардиограмм не позволяет проводить обоснованные процедуры отбора данных для репрезентативных выборок в популяционных исследованиях.

Подходов для выхода из названной проблемы два. Первый связан с организацией взаимодействия гpид-системы с другой информационной системой национального уровня – электронным реестром пациентов. Этот подход описан в докладе «Медицинский гpид и электронный реестр пациентов. Перспективы взаимодействия».

В данном докладе мы попытались провести обзор решений для второго подхода к решению проблемы, а именно, использованию самого телеметрического сигнала для идентификации пациента. В этом случае электрокардиосигнал рассматpивается с точки зрения биометрии.

В течение многих лет ЭКГ изучались и использовались в клинической диагностике. Но только сравнительно недавно было предложено использовать ЭКГ в качестве биометрической информации для идентификации людей [2] - [6]. Одна из причин этого кроется на наш взгляд в противоположности целей, которые ставятся при медицинском и биометрическом подходах. В первом случае нам нужно найти в ЭКГ-сигнале признаки, которые отличают больного человека от здорового, а во втором случае нам нужно найти сходство в ЭКГ-сигналах, принадлежащих одному и тому же человеку, и при этом распознавать отличия в ЭКГ-сигналах, принадлежащих разным людям.

В целом к биометрическому фактору предъявляются следующие основные требования:

- *универсальность* — каждый человек должен обладать этим признаком;
- *измеримость* — этот признак можно быстро измерить;
- *уникальность* — каждый человек обладает уникальным значением этого признака;
- *постоянство* — значение признака не изменяется со временем у одного человека;
- *достоверность* — степень надежности технологии идентификации;
- *изменяемость* — значение этого признака должно быть сложно подделать.

По поводу первого и последнего требований вопросов в отношении ЭКГ почти не возникает, чего нельзя сказать по поводу остальных требований. С измеримостью дело обстоит неплохо. Традиционные ЭКГ на 12 ответвлений выглядят непрактичными с современной точки зрения. В США уже давно ведется, так называемый, мониторинг Холтера [8], в котором используется двухканальная ЭКГ для постоянного мониторинга работы сердца с помощью портативных устройств. Были предложены также жакеты с ЭКГ-датчиками, которые автоматизируют процесс снятия ЭКГ-сигнала, и в недалеком будущем можно ожидать автоматического определения субъекта, одевшего жакет [9].

Основной целью данной работы является исследование точности идентификации субъектов с помощью трехканального ЭКГ-сигнала, которая может быть достигнута при различных технологиях распознавания.

1. Предварительные сведения

Основные технологические этапы работы биометрической системы состоят в следующем:

1. Сбор исходных данных (кардиограмм) осуществляется сенсором;
2. Предварительная обработка входного сигнала осуществляется препроцессором;
3. На третьем и четвертом этапах из сигнала извлекаются признаки и создается шаблон для сравнения. При первом входе в систему происходит регистрация субъекта и его шаблон сохраняется в базе данных.

4. При последующих входах субъекта в систему его шаблон сравнивается с шаблонами, хранящимися в базе данных. Если сравнение было успешным, то субъекту разрешается определенное действие (например, вход в охраняемую зону).

Электрокардиограмма (ЭКГ) это количественная мера электрической активности сердца человека в течение некоторого времени. На рисунке 1 показана типичная форма ЭКГ - сигнала в одном сердечном цикле для 1-го отведения.

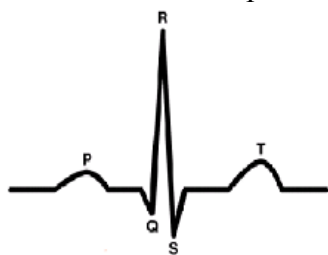


Рис. 1. Типичная форма ЭКГ- волны для одного сердцебиения

Цикл состоит из Р-волны, QRS-волны и Т-волны. QRS-волна, обычно называется QRS-комплексом и содержит наиболее важную информацию о сердцебиении. С другой стороны, есть надежда, что QRS-комплекс характеризует и самого человека. Наличие R-пика указывает на появление сердцебиения. Интервал между R-пиками двух сердцебиений называется RR-интервалом, и характеризует сердечный ритм.

Существуют различные способы получения кардиограмм, однако наибольшее распространение

получила стандартная ЭКГ в 12 ответвлениях, где есть три главных направления ориентации проводников (рис. 2). Можно говорить о многомерности ЭКГ сигнала, если он получен в 2-х ответвлениях, то двумерный, в трех - трехмерный и т.д. Можно говорить также о векторных (контурных) кардиограммах, объединяя одномерные сигналы с разных ответвлений в один сложный (комплексный) сигнал, как это показано на следующем рисунке.

С математической точки зрения это соответствует переходу от двух временных сигналов $f(t)$, $g(t)$, полученных с разных ответвлений, к плоской кривой C , заданной в параметрическом виде $x = f(t)$, $y = g(t)$, $0 \leq t \leq T$.

2. Сбор исходных данных

На форму ЭКГ- сигнала могут оказать влияние следующие факторы:

1. Способ получения сигнала (частота, число ответвлений, тип прибора).
2. Состояние субъекта: положение (лежа, сидя, стоя), эмоции (спокоен, возбужден), здоровый, больной, наличие сердечных патологий.

В литературе встречаются разные сценарии сбора кардиограмм. Во многих работах авторы проверяют свои алгоритмы идентификации на публичных базах кардиограмм, которые есть как в Европе, так и в США [9]. В работе [10] кардиограммы собирались от 22 здоровых людей (20 мужчин, 2 женщины). Сигнал снимался с двух ортогональных

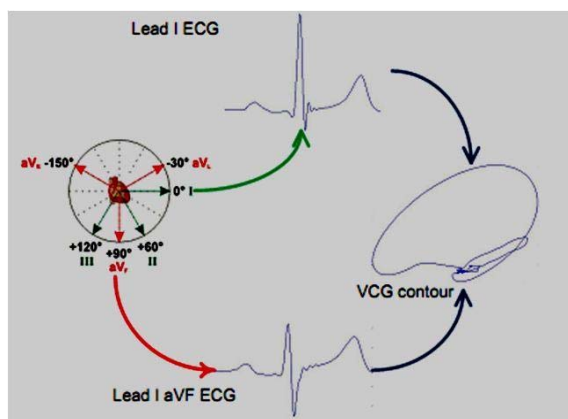


Рис. 2. Векторная кардиограмма (контур)

ответвлений (рис. 2) в сидячем положении. Замер начинался сразу, когда приходил пациент, чтобы сохранить ритм, связанный с предыдущей активностью. На рисунке 3 показана вариация сердечного ритма, полученного в процессе сбора данных.

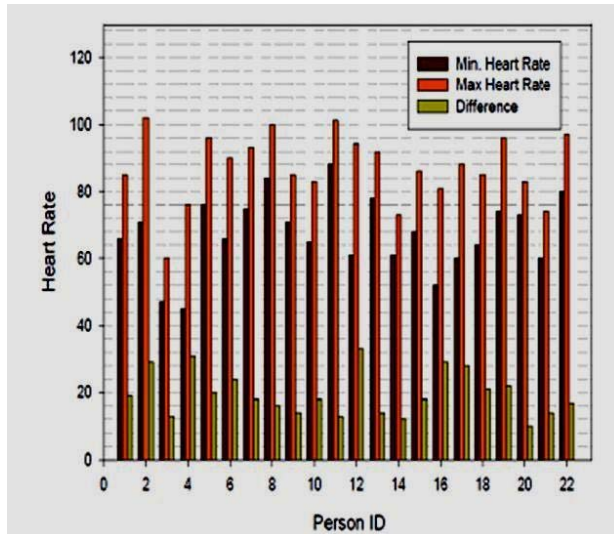


Рис. 3. Вариация сердечного ритма

вызванное дыханием, потоотделением, движением пациента, а также слабыми контактами на электродах. В спектре этот шум расположен на частоте меньшей 1 герца.

2. Влияние частоты питания (50 герц), вызванное плохим заземлением ЭКГ – оборудования.

3. Высокочастотный шум, который появляется внутри отдельного сердцебиения.

Для удаления этих шумов может использоваться цифровой фильтр (Батерворда) с полосой пропускания от 1 до 40 герц, но можно использовать для этой цели и вейвлеты.

Затем выделяются отдельные R-пики, т.е. определяются моменты времени, когда появляются R-пики. Проблема выделения R-пиков исследовалась на протяжении 30 лет и для ее решения было предложено множество методов и программ, на которых мы здесь не можем здесь останавливаться [11].

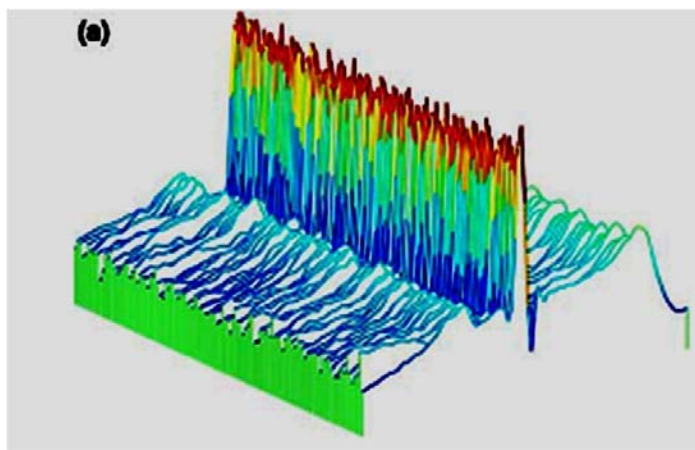


Рис. 4. Матрица одномерного ЭКГ- сигнала

что матрица S содержит большую часть полезной информации о сигнале (за исключением RR-интервалов) и графически ее можно изобразить следующим образом.

В нашем случае база кардиограмм сложилась исторически и была получена на обычной аппаратуре с частотой 500 герц (сигнал из 12 ответвлений, пациент находится в положении лежа, в спокойном состоянии, но может иметь патологии). После предварительного преобразования сигнал в систему поступает в трех ортогональных ответвлениях (рис. 2).

3. Предварительная обработка сигналов

Исходный ЭКГ - сигнал всегда содержит высокочастотные и низкочастотные шумы, которые можно разделить на три вида:

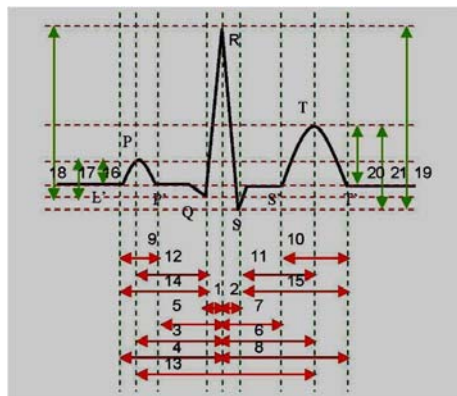
1. Низкочастотный шум: «блуждание» базовой (изолинии) линии,

После этого сердцебиения выравниваются по положениям R-пику и усекаются до окна длительностью в 0,2-0,8 секунды (100-400 замеров при частоте 500 герц) с центром в R-пике. Этот размер окна выбирается эмпирически, поскольку нужно по возможности сохранить Р и Т-пики, и избежать перекрытия R-пиков. В результате предварительной обработки, одномерный сигнал $F(t)$ преобразуется в матрицу S размера $M \times N$, где, а M - число сердцебиений в сигнале, а N – размер окна. Заметим,

4. Выделение признаков и построение шаблонов

Выделение признаков, которые действительно характеризуют субъекта, является самой трудной проблемой при идентификации. В большинстве работ по матрице сигнала сначала получают типовое сердцебиение (паттерн). При этом используют как простое усреднение, так и более тонкие статистические методы. Это сразу на порядок понижает размерность задачи.

На паттерне локализуются опорные точки, по которым вычисляются расстояния, называемые дистанционными признаками. Эти признаки можно разделить на темпоральные (расстояния по горизонтали) и амплитудные (расстояние по вертикали). Этот процесс иллюстрируется следующим рисунком [12].



соседей». Он помечает новый входной вектор, анализируя K его ближайших соседей в обучающих данных. Для распознавания также часто используется метод опорных векторов (support vector machine) или нейронные сети.

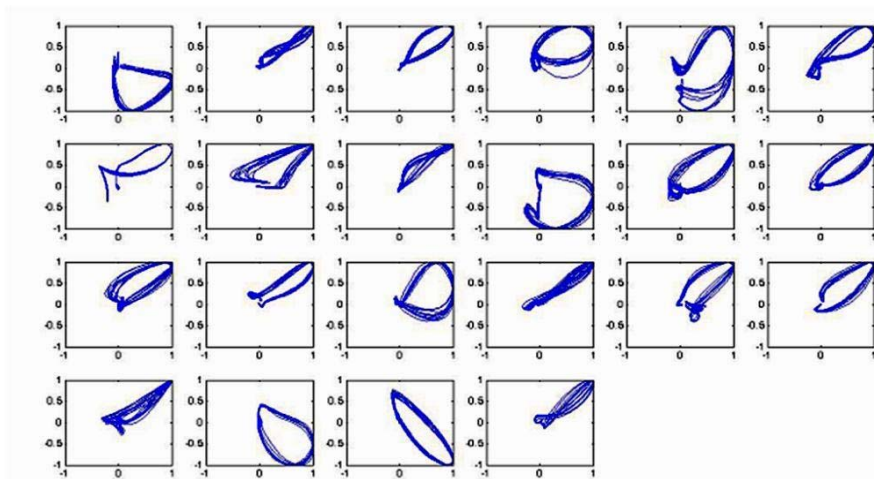


Рис. 6. 10 контуров на каждого субъекта

Исходя из приведенного обзора, использование ЭКГ в качестве биометрического сигнала для идентификации пациента вполне перспективно. В ближайшее время мы планируем провести

сравнительные исследования описанных методов на имеющейся у нас базе данных ЭКГ из более чем 7000 кардиограмм.

Список литературы

1. Вишневский В.В. Грид-система для массового накопления и обработки цифровых электрокардиограмм // Украинський журнал телемедицини та медичної телематики. - 2013. - Т. 11, № 1. - С. 202-208.
2. L. Biel, O. Petersson, L. Philipson and P. Wide, "ECG analysis: a new approach in human identification," IEEE Trans. on Instrumentation and Measurement, vol. 50, no. 3, pp. 808-812, 2001.
3. T.W. Shen, W.J. Tompkins, Y.H. Hu, "One-lead ECG for identity verification," IEEE EMBS, vol. 1, no., pp. 62-63, 2002.
4. S.A. Israel, W.T. Scruggs, W.J. Worek, and J.M. Irvine, "Fusing face and ECG for personal identification", Applied Imagery Pattern Recognition Workshop, vol., no., pp. 226-231, 2003.
5. S. A. Israel, J.M. Irvine, A. Cheng, M.D. Wiederhold, and B.K. Wiederhold, "ECG to identify individuals", Pattern Recognition, vol. 38, no.1, pp. 133-142, 2005.
6. F. Agraftoti and D. Hatzinakos, "Fusion of ECG sources for human identification," IEEE ISCCSP, vol., no., pp.1542-1547, 2008.
7. Holter Monitor – Texas Heart Institute: <http://www.texasheart.org/HIC/Topics/Diag/diholt.cfm>.
8. C. Ye, M.T. Coimbra and B.V.K. Vijaya Kumar, "Arrhythmia Detection and Classification using Morphological and Dynamic Features of ECG Signals," IEEE EMBS, 2010.
9. MIT-BIH Normal Sinus Rhythm Database: <http://physionet.org/physiobank/database/nsrdb/>.
10. M. Abdelraheem, H. Selim, T. Abdelhamid, "Human Identification Using the Main Loop of the Vector cardiogram", Amer. Journal of Signal Processing, vol. 2, no. 2, pp. 23-29, 2012.
11. F. Agraftoti, F. Bui, D. Hatzinakos. Medical Biometrics: The perils of ignoring time dependency, in *Proceedings of 3rd Int. Conf. on Biometrics: Theory, Applications, and Systems*, 2009. Washington, 2009, pp. 1-6.
12. K. Plataniotis, D. Hatzinakos and Y. Wang, "Integrating analytic and appearance attributes for human identification from ECG signals", in *Proc of Biometrics Symposiums*, Baltimore, Maryland, USA, pp. 1-6, 2006.
13. C. Ye, M. Coimbra, B. Kumar, "Investigation of human identification using two-lead electrocardiogram (ECG) signals," in *Proceedings of 4th Int. Conf. on Biometrics: Theory Applications and Systems*, 2010, pp. 1-8.