

Одним из главных недостатков вычислений на GPU является то, что время на загрузку и выгрузку данных в/из памяти графического процессора может превысить время, сэкономленное за счёт параллельной обработки. Поэтому необходимо определить класс задач, для которых использование GPU имеет смысл.

Для тестирования данного подхода были проведены сравнительные эксперименты на разных типах задач. Сравнялось время выполнения вычислений с использованием классического Spark и SparkCL. Результаты экспериментов показали ожидаемый прирост производительности в задачах, которые требуют большого объёма вычислений для каждого элемента данных, и в которых каждый элемент данных имеет большую размерность. К таким задачам можно отнести, например, обработку изображений.

Для расширения класса задач, для которых оправдано использование связки Spark и OpenCL необходим механизм, позволяющий сохранять данные в памяти GPU по завершении работы каждого ядра. Это позволит уменьшить издержки, связанные с обменом данных между центральным и графическим процессорами.

Исследование поддержано проектом CERES. Centers of Excellence for young REsearchers (Reg.no. 544137-TEMPUS-1-2013-SK-JPHES),



Список использованных источников:

1. Jian, L. Parallel data mining techniques on Graphics Processing Unit with Compute Unified Device Architecture (CUDA) / L. Jian, C. Wang, Y. Liu, S. Liang, W. Yi, Y. Shi // The Journal of Supercomputing. - 2013. - Vol. 64, iss.3 — p. 942-967.
2. Keckler, S. GPUs and the Future of Parallel Computing. / S. W. Keckler, W. J. Dally, B. Khailany, M. Garland, D. Glasco // IEEE Micro. - 2011. - Vol. 31, No 5 - p. 7-17.
3. Segal, O. SparkCL: A Unified Programming Framework for Accelerators on Heterogeneous Clusters [Электронный ресурс] / O. Segal, P. Colangelo, N. Nasiri, Z. Qian, M. Margala. - Режим доступа: <https://arxiv.org/ftp/arxiv/papers/1505/1505.01120.pdf> - Дата доступа: 28.03.2017.

ПРИМЕНЕНИЕ МЕТОДОВ МАШИННОГО ОБУЧЕНИЯ ПРИ ДИАГНОСТИКЕ ЗАБОЛЕВАНИЙ НА ОСНОВЕ ДАННЫХ ЭКГ

*Белорусский государственный университет информатики и радиоэлектроники
г. Минск, Республика Беларусь*

Голиков А.В. – аспирант кафедры ЭВМ

Фролов И.И. – к.т.н. доцент

Всем известно, что диагнозы сердечно-сосудистых заболеваний можно ставить по данным электрокардиограмм. Однако, профессором, доктором медицинских наук В.М.Успенским предложен новый метод диагностики, который позволяет диагностировать широкий спектр заболеваний внутренних органов по ЭКГ. Выявлено, что различные расстройства вносят уникальные характеристики в ЭКГ-сигнал, а значит знаки приращений интервалов и амплитуд последовательных кардиоциклов можно использовать для диагностики информации о состоянии здоровья человека и наличии расстройств. Многие заболевания сказываются на работе сердца задолго до проявления клинических симптомов, что позволяет использовать ЭКГ для ранней диагностики.

Суть метода заключается в использовании свойства variability сердечного ритма (интервалов кардиоциклов), которое позволяет определить текущее состояние здоровья человека. Для диагностики важны знаки приращений интервалов и амплитуд последовательных кардиоциклов. Для обеспечения высокого качества диагностики требуется использовать электрокардиографию с разрешением более 500 Гц.

Технологию информационного анализа данных по ЭКГ можно разбить на 2 этапа: этап предварительной обработки ЭКГ-сигнала, и этап машинного обучения.

На этапе предварительной обработки осуществляется:

- 1) демодуляция сигнала — вычисление амплитуд, интервалов и углов по кардиограмме длиной 600 кардиоциклов;
- 2) дискретизация — преобразование данных ЭКГ-сигнала в 599-символьную строку в 6-буквенном алфавите (кодограмму);
- 3) векторизация — преобразование в вектор $63=216$ частот триграмм.

На этапе машинного обучения происходит:

- 1) формирование последовательностей триграмм, которые характеризуют определенное состояние здоровья или недугов; проведение статистического анализ информативности признаков (триграмм);
- 2) обучение модели классификации – на базе существующих данных ЭКГ здоровых и больных пациентов;

3) оценка качества диагностики.

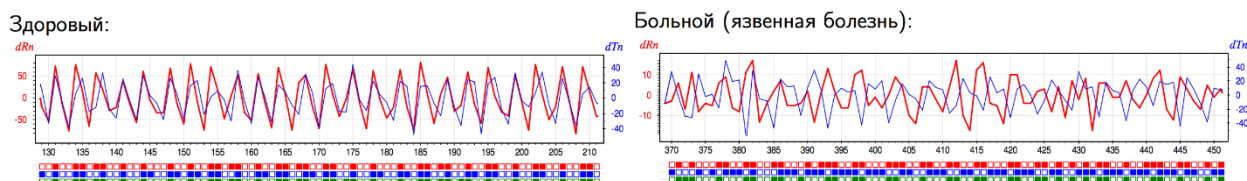


Рис. 1 – Различия в знаках приращения интервалов (dR_n), амплитуд (dT_n) и углов ($d\alpha_n$) кардиоциклов у здоровых и страдающих различными заболеваниями людей.

Технология применяется на протяжении последних 15 лет, за это время накоплены данные о двадцати тысячах больных пациентов, включающие несколько десятков заболеваний. Произведенные эксперименты на базе моделей машинного обучения по одной электрокардиограмме демонстрируют высокую точность определения широкого спектра заболеваний. В настоящее время компания CardioQVARK занимается разработкой мобильного устройства для проведения первичной диагностики, в основе которого лежит описанный метод. Устройство позволяет снимать данные ЭКГ пациентов с высокой частотой квантования сигнала – до 20 кГц, что является существенным техническим преимуществом, задающий новый для компактных приборов уровень функциональности, глубины и детальности диагностики, сравнимый с медицинским оборудованием.

Вданный момент устройство нельзя назвать популярным, им пользуется около 3000 человек. Собранных данных недостаточно много. Но, учитывая потенциал устройства, его функционал можно будет расширить с ростом количества собранных данных - появится возможность для построения более сложных моделей диагностирования заболеваний, расширить список распознаваемых болезней, в конечном счете избавиться от необходимости в частых консультации высококвалифицированного врача.

Список использованных источников:

1. Успенский В.Н. Информационная функция сердца. Теория и практика диагностики заболеваний внутренних органов методом информационного анализа электрокардиосигналов. // Москва: Экономика и информатика, 2008. 116 с.
2. Успенский В.М., Воронцов К.В. Статистическая проверка технологии информационного анализа электрокардиосигналов для диагностики заболеваний внутренних органов // V Международная конференция «Математическая Биология и Биоинформатика», 19 окт.-24 окт. 2014 года, Пущино, Московская обл. Российская Федерация.
3. Бекмачев А., Садовский С., Сунцова О. Российский прорыв: чехол-кардиограф для смартфона CardioQVARK // Электронный-ресурс, 2016. Режим доступа: http://cardioqvark.ru/science/science_9.html

МЕТОДЫ РЕШЕНИЯ ЗАДАЧИ СЛЕПОГО РАЗДЕЛЕНИЯ СИГНАЛОВ

Белорусский государственный университет информатики и радиоэлектроники
г. Минск, Республика Беларусь

Киклевич У.С.

Лавникевич Д.А. – ассистент каф. ЭВМ

Одной из актуальных задач цифровой обработки сигналов является задача разделения и выявления отдельных звуковых волн из сплошного звукового потока. Если слушатель находится в помещении, полном источников звука, то возникает вопрос, как мозгу удастся понять и идентифицировать каждый источник? Данная проблема получила название «эффект коктейльной вечеринки». Таким образом, Cocktail Party Problem – психоакустический феномен, который относится к удивительной способности человека анализировать и распознавать один источник звукового потока в шумной среде [1]. Для решения этой проблемы наш мозг использует 3 последовательных шага:

1. Анализ. Самое важное, что входит в процесс анализа звука, это определение его пространственного расположения. Стоит отметить, что поступающий в мозг поток звуковой информации из одного направления группируется, а из разных – разделяется.
2. Признание. Процесс признания включает в себя распознавание шаблонов речи и нейробиологические механизмы, с помощью которых человек может относительно легко идентифицировать сегрегированный звук из нескольких потоков.
3. Синтез. Процесс синтеза представляет собой реконструкцию отдельных звуковых волн из целого потока. Это позволяет человеку получить информацию от источника и обработать ее [2]. Синтез является самой важной частью в распознавании звука для мозга, и это представляет самую большую проблему для реализации.

Человеческая система восприятия состоит из двух ушей и слуховых путей. Эта сложная система