

О. В. ЖВАЛЕВСКИЙ

Санкт-Петербургский институт информатики и автоматизации Российской академии наук, Санкт-Петербург

## АНАЛИЗ СТРУКТУРНОГО СХОДСТВА ПАРНЫХ ТЕНЗОТРЕМОРОГРАММ С ПРИМЕНЕНИЕМ ФУНКЦИЙ РАССТОЯНИЯ МЕЖДУ ВРЕМЕННЫМИ РЯДАМИ

*В докладе рассматриваются парные тензотреморограммы – временные ряды, которые возникают при регистрации усилия, удерживаемого кончиками пальцев двух рук испытуемого при опоре на чувствительную платформу. Основная гипотеза заключается в том, что процесс удержания заданного уровня усилия протекает по-разному у различных категорий испытуемых – как у здоровых испытуемых, так и у людей, имеющих ту или иную патологию. В докладе исследуются различные подходы к анализу структурного сходства парных тензотреморограмм, основанные на применении функций расстояния между временными рядами.*

**Введение.** Будем рассматривать тензотреморограммы (ТТГ) – временные ряды, которые возникают при регистрации усилия, удерживаемого испытуемым кончиками пальцев при опоре на чувствительную платформу [1, 2]. Эти временные ряды используются для диагностики болезни Паркинсона и синдрома паркинсонизма [3, 4]. Удержание испытуемым усилия на определённом уровне заключается в совмещении друг с другом двух меток на экране компьютера, каждая из которых управляется отдельной рукой испытуемого. Парные ТТГ – это временные ряды, которые регистрируются для левой руки и для правой руки испытуемого одновременно в ходе одного и того же измерительного эксперимента. Парные ТТГ объективным образом отражают процесс постоянной коррекции взаимного положения двух меток. Основная гипотеза заключается в том, что процесс удержания усилия протекает различным образом у различных категорий испытуемых – как у здоровых испытуемых, так и у людей, имеющих ту или иную патологию. Предлагаемый доклад посвящен исследованию различных подходов к анализу структурного сходства, имеющегося между парными ТТГ, основанных на применении функций расстояния между временными рядами.

**Постановка задачи.** Структура экспериментальных данных полностью определяется структурой измерительного эксперимента, который состоит из двух различных частей: первая часть заключается в удержании определённого уровня усилия кончиками пальцев; вторая часть заключается в удержании определённого уровня усилия при помощи вытянутых рук. Измерительный эксперимент каждого типа проводится дважды: первый раз удерживается минимально допустимый уровень усилия, во второй раз удерживается максимально допустимый уровень усилия. В соответствии с этим, в сравнительном анализе участвуют четыре различных типа временных рядов, каждый из которых представлен двумя парными ТТГ. Представляется целесообразным использовать здесь подходы к анализу, основанные на применении *функций расстояния* между временными рядами.

Первый подход к анализу временных рядов заключается в *кластеризации* имеющихся временных рядов. В этом случае, временные ряды группируются в *кластеры* или группы, однородные в смысле используемой функции расстояния. Вторым подходом к анализу временных рядов является анализ *внутренней* структуры временных рядов, порождаемой применением биологической обратной связи. В этом случае, в имеющихся временных рядах выделяются характерные *структурные элементы*. При анализе парных ТТГ, такими структурными элементами будут, очевидно, фрагменты с похожим поведением, имеющиеся в различных ТТГ, составляющих общую пару, и отражающие процесс взаимной коррекции положения меток друг относительно друга. В этом смысле, проверка основной гипотезы сводится к выбору такой функции расстояния, которая позволит наиболее точно восстанавливать структуру взаимных коррекций. Таким образом, оптимальный выбор функции расстояния полностью определяется внутренней структурой анализируемых временных рядов.

**Методика структурного анализа.** Для того, чтобы провести сравнительный анализ двух временных рядов, необходимо рассматривать фрагменты одного временного ряда и исследовать вхождение этих фрагментов в другой временной ряд. Если взять произвольный фрагмент *первого* временного ряда и, затем, перемещать окно соответствующей ширины вдоль *второго* временного ряда, то для каждого допустимого положения окна можно вычислить расстояние

между соответствующими фрагментами. Результатом построения будет *дистанционный профиль* данного (второго) временного ряда по отношению к заданному фрагменту первого временного ряда. Каждому фрагменту первого временного ряда соответствует свой дистанционный профиль (и, наоборот). Если, теперь, рассмотреть различные фрагменты первого временного ряда (например, перемещая вдоль *первого* временного ряда окно соответствующей ширины), то можно получить *дистанционную матрицу*. Каждой заданной заранее длине фрагментов и каждой процедуре перебора различных фрагментов соответствует своя дистанционная матрица. Фиксируя некоторое пороговое значение (меру сходства), в анализируемом (втором) временном ряде выделяются фрагменты, которые можно считать (в смысле заданной функции расстояния и в смысле заданной меры сходства) *вхождениями* заданных фрагментов первого временного ряда во второй временной ряд. Построение и анализ дистанционных профилей, которые строятся для *парных* временных рядов, имеет особое значение. Действительно, если взять некоторый *начальный* фрагмент одного из двух временных рядов, составляющих пару, и построить соответствующий ему дистанционный профиль для второго временного ряда, составляющего ту же пару, то, обнаруживая ближайший минимум, можно выделить во втором временном ряде фрагмент, похожий на исходный фрагмент первого временного ряда. Если теперь, взять выделенный фрагмент второго временного ряда и построить дистанционный профиль первого временного ряда по отношению уже к этому фрагменту, то, обнаруживая ближайший минимум уже в новом дистанционном профиле, можно выделить в первом временном ряде фрагмент, похожий на ранее выделенный фрагмент второго временного ряда. Последовательно выделяя таким способом фрагменты в обоих временных рядах, можно постепенно восстановить структуру взаимных коррекций. При этом, следует учитывать, что выделяемые фрагменты могут перекрываться, а сами «коррекции» могут многократно повторяться в виде «эха». По своей сути, речь идёт об организации процедуры последовательного сдвига одного временного ряда относительно другого. При этом, каждый такой сдвиг должен обеспечивать минимальное расстояние между локальными фрагментами обоих временных рядов, вычисляемое в пределах некоторого окна.

**Предварительные результаты.** Будем использовать для оценки корректности основной гипотезы стандартное евклидово расстояние и его аналоги. Это позволит получить первичное представление об анализируемых временных рядах и построить надёжное основание для применения, в дальнейшем, других функций расстояния между временными рядами.

**Преобразования временных рядов.** Как показывают вычисления, непосредственное применение евклидова расстояния (или его аналогов) не позволяет каким-либо образом различать ТТГ, относящиеся к различным категориям. При этом, само значение евклидова расстояния (или его аналога) существенным образом зависит от амплитуды. Именно поэтому, оказывается целесообразным использовать то или иное преобразование (вроде нормализации и/или стандартизации), позволяющее привести анализируемые временные ряды к единой шкале. Как показывают вычисления, наиболее удачный способ нормировки заключается в вычислении максимального отклонения для каждой отдельной пары ТТГ и нормировки данным значением исходного евклидова расстояния, вычисляемого для двух парных временных рядов. Ещё один успешный вариант преобразования связан с удалением из анализируемого временного ряда локального тренда, которое также позволяет нивелировать различие в амплитуде, имеющееся между различными временными рядами. Аналогичные результаты получаются и при дифференцировании исходных временных рядов с последующими нормализацией и/или стандартизацией.

**Построение дистанционного профиля.** Как показывают вычисления, дистанционный профиль, построенный для парных ТТГ, имеет, периодическую структуру, что является ожидаемым, в силу того, как осуществляется процесс регистрации физиологических сигналов. В то же время, если рассматривать фрагменты, извлекаемые из ТТГ одного класса (то есть – ТТГ, относящимся к определённой категории испытуемых), и строить дистанционные профили для ТТГ другого класса, то уже не будет такой явной периодической структуры, при этом, сами расстояния будут большими. Как показывают вычисления, ширина окна и сама структура взаимных сдвигов временных рядов, составляющих общую пару, друг относительно друга различается для различных ТТГ – как для ТТГ, относящихся к различным классам, так и для ТТГ одного и того же класса, но относящихся к различным типам измерительных экспериментов, что подтверждает целесообразность организации *составного* измерительного эксперимента, содержа-

шего внутри себя несколько различных тестов, соответствующих различным условиям проведения измерительного эксперимента. Также было обнаружено, что учёт локальных деформаций парных временных рядов, позволяет точнее определять сходство соответствующих подпоследовательностей, а значит, и точнее определять величину самого сдвига, причём эта величина оказывается информативным показателем.

**Заключение.** В предлагаемом докладе рассматривается задача выявления структурного сходства между парными ТТГ, возникающего в результате введения в схему организации измерительного эксперимента биологической обратной связи, и исследуются подходы к анализу этого сходства, основанные на применении функций расстояния между временными рядами. Полученные предварительные результаты с применением стандартного эвклидова расстояния и его аналогов подтверждают основную гипотезу. Это обосновывает необходимость в более глубоких и масштабных исследованиях. Первое направление исследований связано с применением различных «гибких» или «эластичных» функций расстояний (вроде DTW [5, 6] и LCSS [7]). Такие функции позволят более точно выделять похожие друг на друга фрагменты, имеющиеся в парных временных рядах, поскольку эти функции допускают локальные искажения как по времени, так и по амплитуде. Второе направление исследований связано с построением различных символьных представлений анализируемых временных рядов (вроде специализированного представления SAX [8]). Удачно выбранное символьное представление позволит обнаруживать связанные друг с другом фрагменты, имеющиеся в парных временных рядах, поскольку такие фрагменты будут описываться одними и теми же (или близкими друг к другу) символьными подпоследовательностями. И, наконец, третье направление исследований связано с разработкой специализированных алгоритмов кластеризации фрагментов временных рядов и построением новых комбинированных функций расстояния между временными рядами, которые наилучшим образом учитывают парную структуру ТТГ. В силу особой структуры экспериментальных данных и особой постановки основной задачи, все результаты (как уже полученные, предварительные результаты, так и предполагаемые в будущем результаты) являются принципиально новыми.

#### ЛИТЕРАТУРА

1. Романов С.П., Якимовский А.Ф., Пчелин М.Г. Метод тензометрии для количественной оценки тремора // Физиологический журнал им. И.М. Сеченова. 1996. Том 82, № 2. С. 118-123.
2. Романов С.П., Пчелин М.Г., Якимовский А.Ф. Характеристики изометрически регистрируемого тремора при поражении экстрапирамидной системы // Российский физиологический журнал им. И.М. Сеченова. 1997. Том 83, № 3. С. 133–139.
3. Дик О.Е., Романов С.П., Ноздрачев А.Д. Энергетические и фрактальные характеристики физиологического и патологического тремора руки человека // Физиология человека, 2010, том 36, № 2, с. 92–100.
4. Дик О.Е., Ноздрачев А.Д. Нелинейная динамика произвольных колебаний руки человека при двигательной патологии // Физиология человека, 2015, том 41, № 2, с. 53–59.
5. Berndt D.J., Clifford J. Using Dynamic Time Warping to find patterns in time series. *Proceedings of the 3rd International Conference on Knowledge Discovery and Data Mining*. 1994. PP. 359-370.
6. Keogh E.J., Ratanamahatana C.A. Exact indexing of dynamic time warping. *Knowledge and Information Systems*. 2005. Vol.7. PP. 358-386.
7. Vlachos M., Gunopulos D., Kollios G. Discovering similar multidimensional trajectories. *Proceedings 18th International Conference on Data Engineering (ICDE)*. 2002. PP. 673-684.
8. Lin J., Keogh E., Lonardi S., Chiu B. A symbolic representation of time series, with implications for streaming algorithms. *Proceedings of 8th ACM SIGMOD Workshop on Research Issues Data Mining and Knowledge Discovery—DMKD'03*. 2003. P. 2.

O.V. Zhvarevsky, (St. Petersburg Institute for Informatics and Automation of the Russian Academy of Sciences, St. Petersburg)

#### **Paired Tenzotremorogramms Structure Similarity Analysis Based on Time Series Distance Functions**

The paper is concern with paired tenzotremorogramms that is a result of fingers force registration (for both human arms) via sensitive platform with tenzosensor. The main hypothesis is: the dynamics of force hold process differs for different patient category (for healthy patients and for patients who have some pathology). In the paper some approaches to paired tenzotremorogramms structure similarity analysis are investigated. These approaches are based on application of time series distance functions.