

МИНОБРНАУКИ РОССИИ
Федеральное государственное бюджетное образовательное учреждение
высшего образования
**«САРАТОВСКИЙ НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ ИМЕНИ Н.Г.
ЧЕРНЫШЕВСКОГО»**

Кафедра физики открытых систем

**Распознавание зашумленных сигналов на основе
искусственных нейронных сетей и других методов
классификации**

АВТОРЕФЕРАТ МАГИСТЕРСКОЙ РАБОТЫ
студента 2 курса 2241 группы
направления 09.04.02 «Информационные системы и технологии»
института физики
Алимова Никиты Владимировича

Научный руководитель

профессор, д.ф.-м.н., профессор _____ А.Н. Павлов

подпись, дата

Зав. кафедрой физики открытых систем

профессор, д.ф.-м.н., профессор _____ А.А. Короновский

подпись, дата

Саратов 2021 г.

ВВЕДЕНИЕ

В проводимых исследованиях задача распознавания зашумленных сигналов будет решаться на примере экспериментальных данных межклеточного электрического потенциала малого нейронного ансамбля. В типичных примерах таких сигналов можно идентифицировать характерные сигналы 2-3 нейронов, каждый из которых генерирует потенциалы действия, немного отличающиеся по форме, и фоновый шум, к которому можно отнести слабые сигналы нейронов, находящихся в отдалении от регистрирующего микро-электрода.

Большинство нейронов мозга взаимодействуют друг с другом, посылая и получая короткие электрические импульсы, так называемые потенциалы действия или короткие сигналы – спайки (spike). При анализе поведения нейронного ансамбля или изучении нейронного кода спайки считаются стереотипными (одинаковыми) событиями. Следовательно, для передачи информации важна не форма каждого сигнала спайка, а точное время его генерации. Тогда мы можем говорить о импульсах, генерируемых нейронами, как о многомерном бинарном процессе.

Обычно предполагается, что каждый нейрон генерирует импульсы одинаковой формы и амплитуды, в то время как сигналы от разных клеток имеют некоторые индивидуальные особенности (хотя их отличия могут быть небольшими). Хотя это предположение является дискуссионным (например, в «берсте» каждый последующий импульс обычно меньше предыдущего), оно достаточно типично для многих практических случаев, поэтому далее работа построена исходя из принятия данного предположения.

В настоящее время существует ряд численных инструментов для сортировки спайков. В данной работе рассматриваются некоторые методы, в том числе искусственные нейронные сети, вейвлет-преобразование, и сравниваются со стандартными методами, наиболее широко используемыми в экспериментальных лабораториях. Несмотря на то, что существующие

методы в целом позволяют решать данную задачу, оптимальная процедура для извлечения признаков импульсов все еще остается сложной задачей.

Целью данной выпускной квалификационной работы является проведение сравнительного анализа методов распознавания зашумленных сигналов на примере экспериментальных данных межклеточных электрических потенциалов малого нейронного ансамбля с использованием математического аппарата искусственных нейронных сетей, вейвлет-анализа и анализа главных компонент.

Материалы исследования. Исследования проводились, используя сигналы внеклеточного электрического потенциала и различные методы их анализа.

Выпускная квалификационная работа содержит введение, две главы (1. Краткие теоретические сведения; 2. Результаты исследований), заключение и список использованных источников. Общий объем работы 60 стр.

ОСНОВНОЕ СОДЕРЖАНИЕ РАБОТЫ

Краткие теоретические сведения. Хотя детали различных методов сортировки спайков могут значительно отличаться, подавляющее большинство известных методов включает ряд общих операций. Эти операции могут быть независимыми, или некоторые из них могут быть включены в единую процедуру улучшения качества сортировки импульсов. Среди широко применяемых подходов можно выделить следующие группы методов извлечения признаков классифицируемых сигналов:

- Пороговый.
- Анализ главных компонент (Principal component analysis - PCA).
- Вейвлет-преобразование (Wavelet Transform - WT).
- Искусственные нейронные сети и другие алгоритмы распознавания.

В проводимых исследованиях сравнение методов проводилось на примере данных, состоящих из более 10000 записей импульсов нейронов, в каждом импульсе по 64 точки. Для упрощения процесса обучения нейронной

сети выделялись признаки, на основании которых возможно охарактеризовать каждый сигнал, например, амплитуда сигнала и коэффициенты дискретного преобразования Фурье.

Рассмотрены некоторые методы классификации сигналов – искусственные нейронные сети (персептрон Розенблатта), метод k ближайших соседей, метод k-средних, анализ главных компонент, методы на основе вейвлет-преобразования.

Результаты исследований. Вначале были рассмотрены результаты применения нейросетевых методов распознавания формы сигналов. Классификация при помощи нейронной сети на основе многослойного персептрона осуществлялась на основе 60% от всей выборки, оставшиеся 40% были использованы для проверки обучения нейронной сети. Обучалась нейронная сеть на основе многослойного персептрона со следующей конфигурацией:

- функция активации ReLu
- 3 нейрона на входном слое
- 3 скрытых слоя
- 3 нейрона на выходном слое

Из рисунка 1 видно, что нейронная сеть хорошо распознала сигналы одного нейрона, но кластеры точек, соответствующих сигналам двух других, пересекаются друг с другом. Чтобы определить, насколько точно нейронная сеть классифицирует два этих типа сигналов, была обучена новая нейронная сеть, конфигурация нейронной сети была следующая:

- функция активации ReLu
- 2 нейрона на входном слое (поскольку на вход подаётся 2 сигнала)
- 2 скрытых слоя
- 2 нейрона на выходном слое

Результаты обучения нейронной сети представлены на рисунке 2.

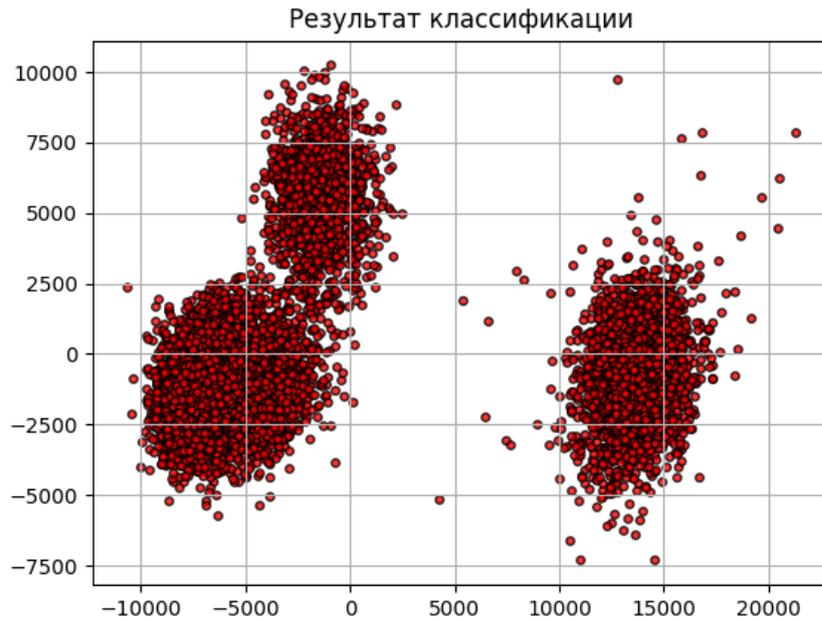


Рисунок 1. Результат классификации при помощи нейронной сети

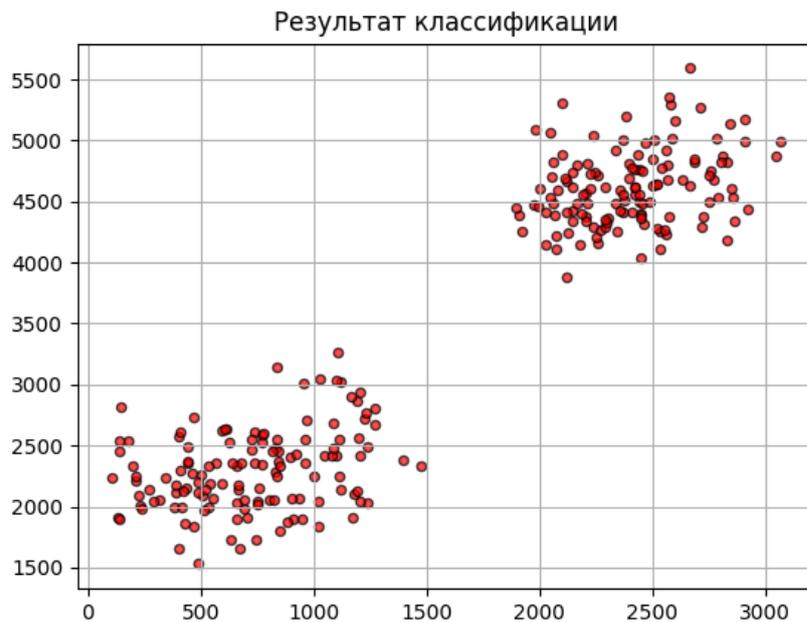


Рисунок 2. Результат классификации для двух перекрывающихся областей.

Точность распознавания сигнала составила 1, что означает корректную классификацию всех сигналов, поданных на вход нейронной сети. Поскольку часть данных использовалась в обучении нейронной сети, точность распознавания сигналов очень высока. Если данных будет больше, то,

возможно, натренированная нейронная сеть не сможет распознать все импульсы, поданные на вход, корректно.

Применение метода ближайших соседей также позволяет провести классификацию и достаточно точно разделить сигнал разных типов. Один класс отчетливо отделим от двух других. Далее определим, насколько точна оценка, применив алгоритм классификации только для двух данных классов. Для этого использовался заранее обученный алгоритм для более точной классификации. Сгенерировано два перекрывающихся друг друга сигнала, результат работы классификатора можно наблюдать на рисунке 3.

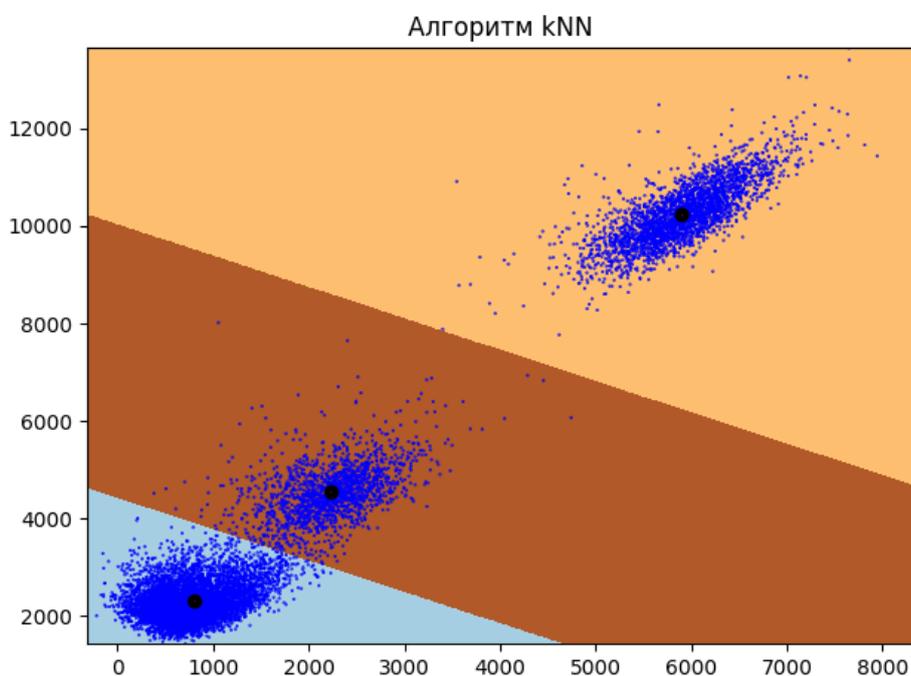


Рисунок 3. Результат классификации методом ближайших соседей

Результаты классификации методом K-средних представлены на рисунке 4. Как видно из этого рисунка, сигналы одного нейрона хорошо отделимы от сигналов двух других. Данные сгруппированы в три кластера, каждый из кластеров отделен цветом, центры каждого из кластеров отмечены на графике черной точкой. Метод выполняется быстро, работает достаточно точно, при условии, что заранее известно примерное количество кластеров, которые необходимо получить.

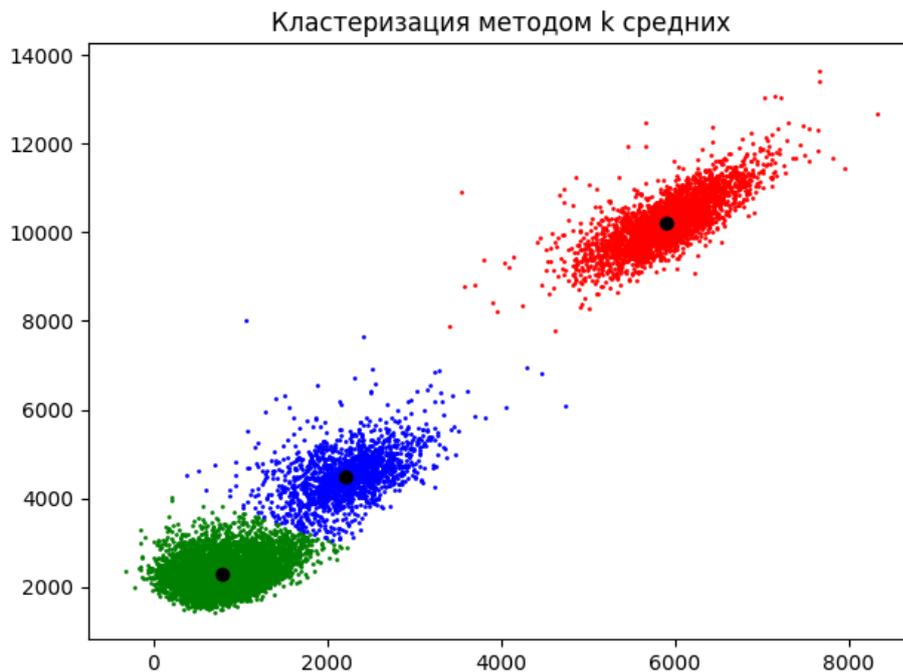


Рисунок 4. Результат классификации методом k-средних

Также в работе была реализована классификация сигналов методом главных компонент. Проведены тестовые исследования на характерных формах спайков с добавлением белого шума (эксперимент с известными заранее формами сигналов). Несколько примеров для разного уровня шума приведены на рисунках 5 и 6. На рисунке 7 показаны результаты для реальных данных. Из приведенных рисунков видно, что эффективность классификации сигнала методом главных компонент различна в зависимости от типа шума.

В случае белого шума облака сигналов хорошо различаются, с увеличением амплитуды шума увеличивается и разброс облака. Затем облака смешиваются и отличить один сигнал от другого не представляется возможным. Результат предсказаний методом PCA составил 0.980.

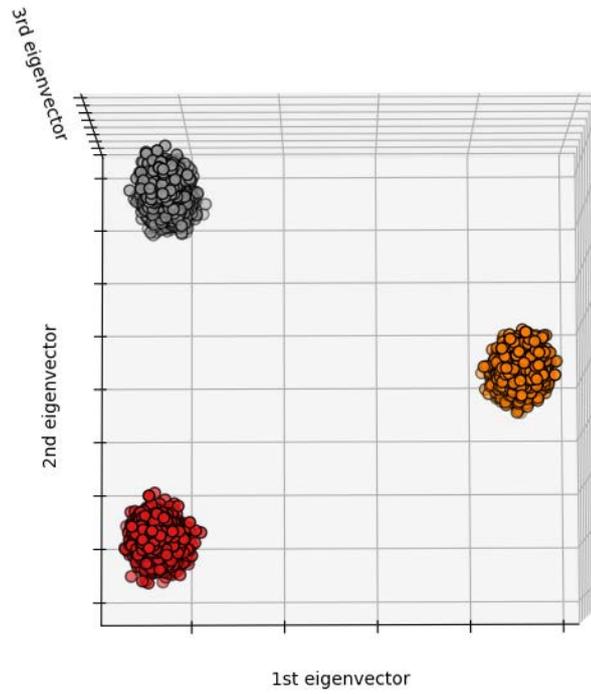


Рисунок 5. Классификация сигналов методом PCA с амплитудой белого шума, равной 100

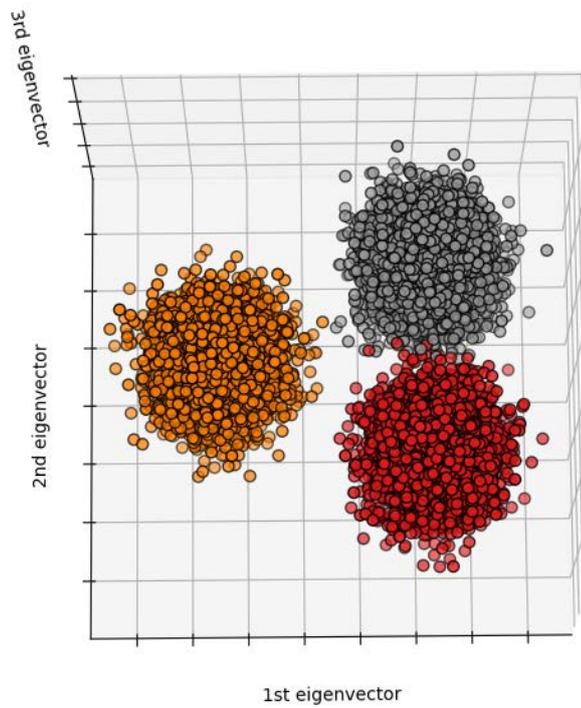


Рисунок 6. Классификация сигналов методом PCA с амплитудой белого шума, равной 400

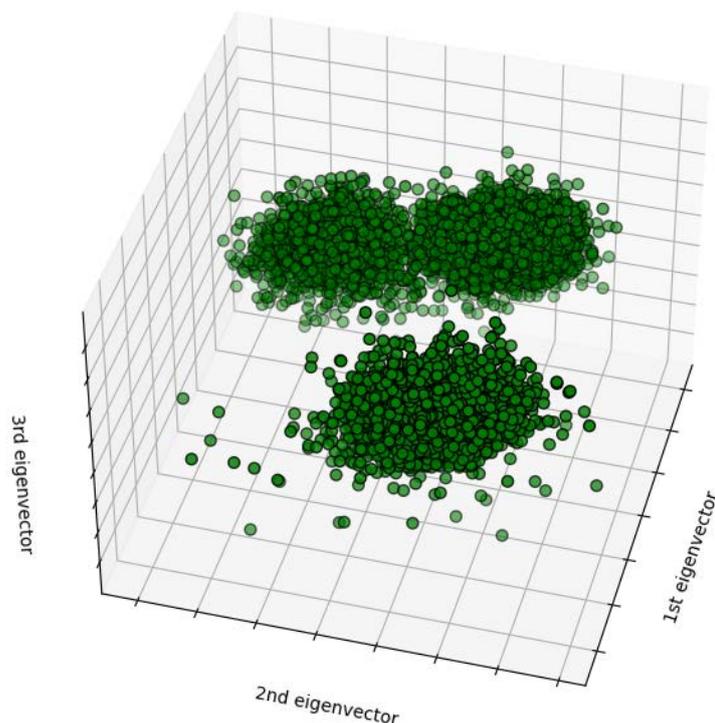


Рисунок 7. Классификация сигналов методом PCA по экспериментальным данным

Также были использованы подходы на основе вейвлет-анализа. При помощи вейвлетов сигналы, соответствующие реальным данным, были разделены на группы (классы) по различным характеристикам. Классификация сигналов с помощью вейвлетов Гаусса первого порядка показала хороший результат, при увеличении шума облако параметров становится больше, и, как можно заметить на рисунках, облака смешиваются, что затрудняет дальнейшую классификацию данным методом. Размер облаков у вейвлетов типа «МНАТ» оказался меньше, но два из трёх сигналов нейрона различимы с трудом. Результат в значительной степени зависит от выбранных параметров, типа шума и вейвлета. Тип вейвлета необходимо выбирать оптимальным для решения задачи, опираясь на имеющиеся данные. Некоторые примеры полученных результатов показаны на рисунках 8-10.

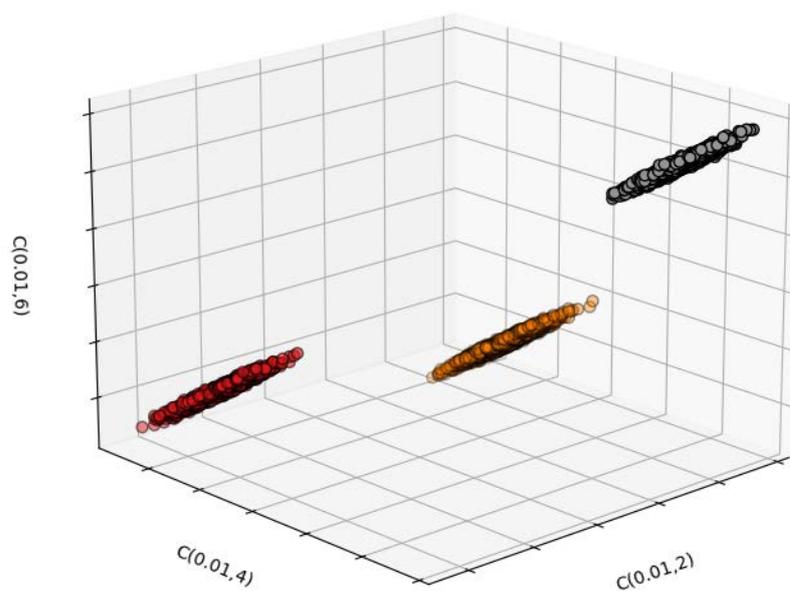


Рисунок 8. Результат применения вейвлет-преобразования к сигналу с белым шумом, амплитуда сигнала равна 50

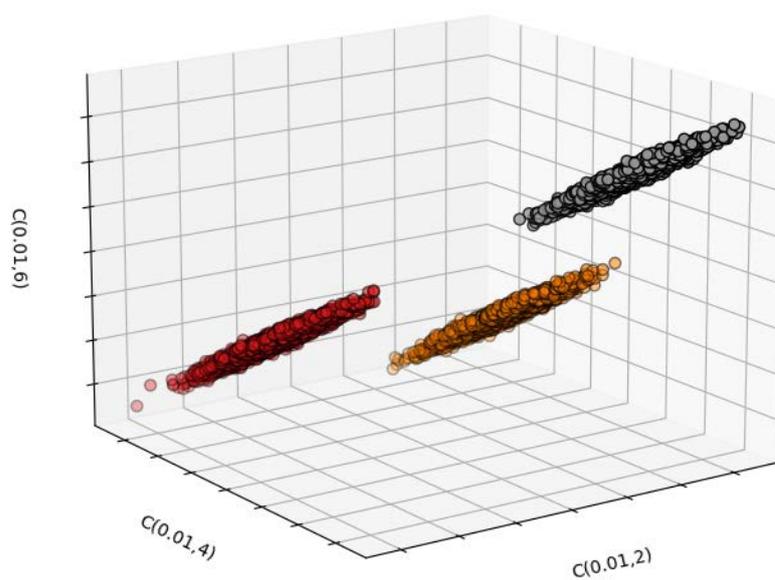


Рисунок 9. Результат применения вейвлет-преобразования к сигналу с белым шумом, амплитуда сигнала равна 100

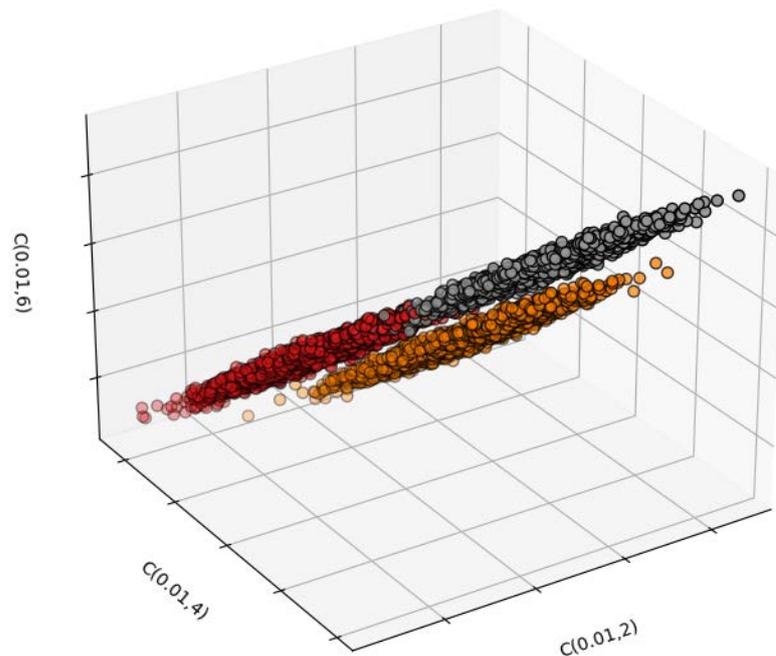


Рисунок 10. Результат применения вейвлет-преобразования к сигналу с белым шумом, амплитуда сигнала равна 250

ЗАКЛЮЧЕНИЕ

В работе были изучены и реализованы следующие алгоритмы классификации:

- на основе многослойного персептрона;
- с применением метода К-ближайших соседей;
- на основе метода К-средних;
- с применением анализа главных компонент;
- с использованием вейвлет-анализа.

Методы классификации были применены к экспериментальным данным, представляющим собой записи сигналов внеклеточного электрического потенциала, содержащим сигналы группы из трёх близко располагающихся нейронов.

Методы К-ближайших соседей и К-средних работали дольше, чем остальные. Время кластеризации составляло от 4 до 7 минут.

Метод К-средних выделил три кластера корректно, но не всегда данные хорошо различимы, и поэтому необходимо знать хотя бы примерное количество кластеров для оценки корректности результата классификации.

Можно отметить, что все методы справились с задачей с точностью более 90%, однако нейронная сеть, метод анализа главных компонент и метод вейвлет-анализа справились с зашумленным сигналом лучше.

Метод анализа главных компонент, вейвлет-анализ и нейронная сеть были сравнимы по скорости. Но вейвлет-анализ и анализ главных компонент, по сравнению с нейронной сетью на основе многослойного персептрона, не требуют предварительной разметки данных. С учетом специфики решаемой задачи каждый из них может применяться для классификации нейронных потенциалов действия.