

УДК 004.67

**АНАЛИЗ ДАННЫХ
ИССЛЕДОВАТЕЛЬСКИХ ИСПЫТАНИЙ
ПЬЕЗОДАТЧИКОВ ПРИ ВОЗДЕЙСТВИИ
АГРЕССИВНЫМИ СРЕДАМИ**

Новиков С. П., Акулин П. П.

Донской государственный технический
университет, Ростов-на-Дону, Российская
федерация

n_serg7@mail.rupavel.akulin@ya.ru

Рассмотрены проблемы обработки и анализа данных. Выявлены этапы проведения измерений с помощью пьезодатчиков. Описаны и реализованы алгоритмы обработки выходных сигналов сенсоров: определение времени начала рабочего режима, сглаживание измеряемых данных при помощи экспоненциального сглаживания, сглаживание Савицкого-Голя и дискретного преобразования Фурье. Приведено сравнение данных алгоритмов и выявлены их характерные особенности.

Ключевые слова: сенсоры, обработка данных, мониторинг, газы, пьезодатчики, фильтрация сигналов, сглаживание данных.

Введение. При разработке новых типов датчиков и проведении испытаний для определения их качественной и количественной оценки большую роль играет предварительная подготовка получаемых с них данных, чтобы абстрагироваться от шумов или лишних значений, которые снижают точность вычислений. «Проблема установления химического состава различных биообъектов решается различными методами, но в последнее время наблюдается стремительный рост так называемых полевых аналитических технологий. Они предполагают все более возрастающее использование в аналитических процедурах химических сенсоров, биосенсоров и тест-методов» [1]. При этом разработанные устройства генерируют множество данных. Качественная обработка также помогает понять устойчивость работы датчиков, выявить характерные особенности, а также установить пределы применимости при воздействии различными агрессивными средами.

Датчики имеют тенденцию быть чувствительными к более, чем одной физической величине, из-за чего показывают ложный сигнал, когда они подвергаются вибрации. Поэтому необходимо определить технологические этапы и автоматически определять начало рабочего режима. После этого можно проводить анализ данных. Однако в детектируемых сигналах всегда будут присутствовать шумы и ложные сигналы, от которых следует избавиться путем сглаживания измеряемых данных.

Таким образом, необходимо создать систему обработки результирующего сигнала, выполнив следующие задачи:

- выделить этапы проведения измерений;

UDC 004.67

**DATA ANALYSIS OF INVESTIGATION
TEST OF PIEZOSENSORS ON EXPOSURE
TO AGRESSIVE ENVIRONMENT**

Novikov S. P., Akulin P. P.

Don State Technical University, Rostov-on-Don,
Russian Federation

n_serg7@mail.rupavel.akulin@ya.ru

This article discusses the problems of data processing and analysis. The paper identifies measurement stages using the piezoelectric transducer. The authors have described and implemented algorithms for processing the output sensors signals: determining a start time and operation mode, smoothing of the measured data using exponential smoothing, Savitzky-Golay smoothing and discrete Fourier transformation. The comparison of these algorithms is given and their characteristic features are revealed.

Keywords: sensors, data processing, monitoring, gases, piezosensors, signal filtering, data smoothing.

- автоматически определить начало рабочего режима;
- сглаживать измеряемые данные с помощью различных алгоритмов;
- получить количественные и качественные характеристики сигнала;
- провести анализ алгоритмов сглаживания и выбрать предпочтительный.

1. Этапы проведения измерений с помощью пьезодатчиков.

Для проведения количественного анализа необходимо выделить следующие технологические этапы:

- А — выход в рабочий режим — заполнение камеры газом.
- Б — рабочий режим — срабатывание датчика, измерение (получение данных);

Временные интервалы этапов представлены на рис. 1.

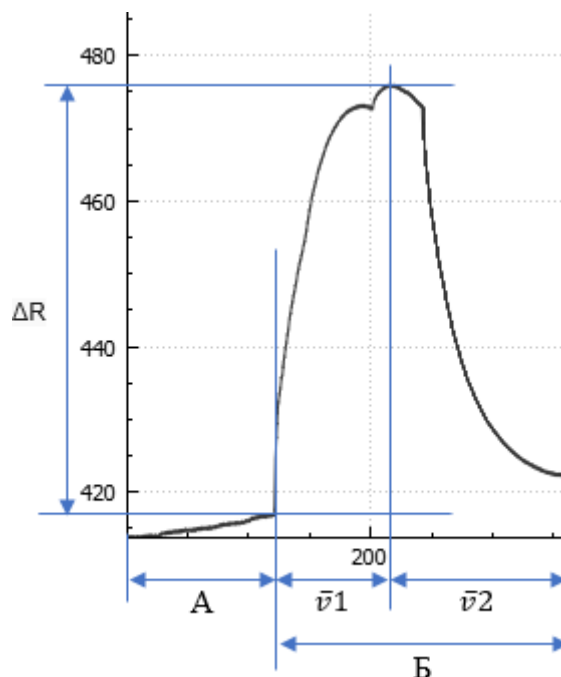


Рис. 1. Временные интервалы технологических этапов

На отрезке А часто наблюдаются ложные вибрации и шумы, после чего начинается резкий рост амплитуды (этап Б) в связи с повышением концентрации газа. После опустошения камеры происходит постепенное спадание амплитуды. Продолжительность рабочего режима, показанного на рис. 1, составляет 350–360 с.

Пьезоэлектрический эффект обнаруживается датчиком. Из этого измерения могут быть сделаны выводы об изменении сопротивления (ΔR), скорости реакции (\bar{v}_1), скорости восстановления (\bar{v}_2). «Давление влияет на изменение скорости реакции именно в газовой среде. Любой газ занимает весь предоставленный объем — молекулы разлетаются, и число их столкновений уменьшается. Если повысить давление, число активных соударений возрастает, соответственно, скорость реакции увеличивается, т.е. изменение скорости реакции в газовой фазе прямо пропорционально изменению давления» [2].

Таким образом, после выделения технологических этапов и автоматического определения начала рабочего режима, можно проводить количественный анализ по изменению сопротивления, времени реакции и восстановления, но перед этим необходимо избавиться от шумов при помощи различных алгоритмов.

2. Алгоритмы обработки выходных сигналов сенсоров

2.1 Определение времени начала рабочего режима

Информационные параметры, которые свидетельствуют о типе и концентрации анализируемого газа могут быть определены только на этапе рабочего режима (рис. 1). Первой задачей программного обеспечения является определение времени начала рабочего режима. Эта задача решается двумя способами: вручную и автоматически. В ручном режиме указанное время определяет оператор с помощью специального маркера. Однако это замедляет обработку данных, поэтому система должна определять начало рабочего режима автоматически, при помощи следующего алгоритма действий:

1) В скользящем окне определяются параметры прямой линии, в которой аппроксимируются измеряемые данные. Такая аппроксимация проводится методом наименьших квадратов (МНК), с помощью которого определяется коэффициент детерминации этой прямой:

$$R = r_{xy}^2,$$

где r_{xy} — коэффициент корреляции:

$$r_{xy} = \frac{\overline{xy} - \bar{x}\bar{y}}{\sigma(x)\sigma(y)}, \quad (1)$$

\bar{x} , \bar{y} , \overline{xy} — средние значения:

$$\bar{x} = \frac{\sum x_i}{n}, \quad (2)$$

$$\bar{y} = \frac{\sum y_i}{n},$$

$$\overline{xy} = \frac{\sum x_i y_i}{n},$$

n — количество данных в скользящем окне, σ — среднеквадратическое отклонение:

$$\sigma(x) = \sqrt{D(x)},$$

$$\sigma(y) = \sqrt{D(y)},$$

где D — дисперсия:

$$D(x) = \frac{\sum x_i^2}{n} - \bar{x}^2,$$

$$D(y) = \frac{\sum y_i^2}{n} - \bar{y}^2. \quad (3)$$

Коэффициент детерминации в формуле (1) оценивает долю дисперсии (изменчивость) значений на оси ординат из формулы (3), которая объясняется с помощью значений на оси абсцисс в простой линейной регрессионной модели. Однако коэффициент детерминации получается путем возведения в квадрат коэффициента корреляции (показывает статистическую взаимосвязь двух или более случайных величин) из формулы (2). Коэффициент детерминации изменяется от 0 до 1, т.е. если значение равно 0,92, то это означает, что в 92% случаев изменения значений оси абсцисс приводят к изменению значений оси ординат [3].

2) Сравниваем коэффициент детерминации с пороговым значением. Если его величина соответствует заданному порогу, то принимается решение о начале рабочего режима. Когда график начинает резко возрастать, то значение коэффициента детерминации начинает резко

убывать и стремиться к нулю. В ходе тестирования выяснилось, что точку старта можно определить при пороговом значении коэффициента детерминации выше 0,8. Поэтому можно остановиться на среднем значении равном 0,92.

Как было сказано ранее, опыт длился 350–360 с, поэтому в программном обеспечении интервал скользящего окна был задан 355 с. Этот параметр можно изменить в настройках программы, если длительность опыта будет изменена.

Результат определения начала рабочего режима представлен на рис. 2.

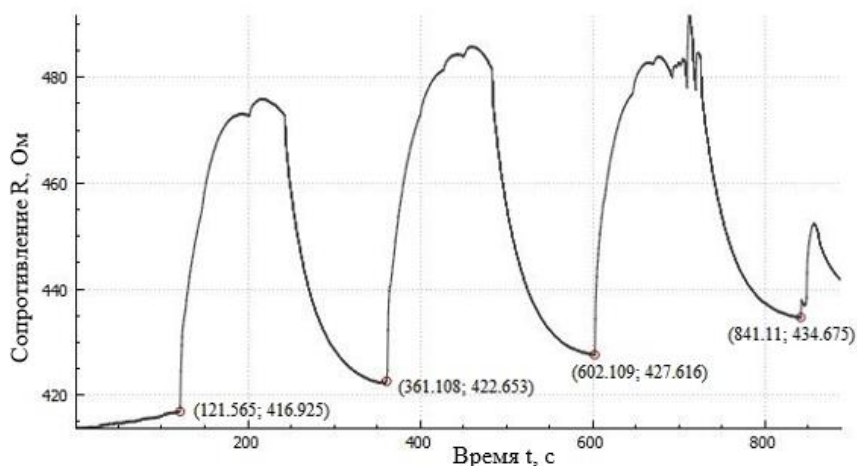


Рис. 2. Результат определения начала рабочего режима

На рис. 2 отмечены значения точек начала рабочего режима, что соответствует поставленной задаче. Еще стоит отметить, что после резкого увеличения давления затухания могут быть более длительными, поэтому необходимо найти точки начала для каждого возбуждения датчиков.

Таким образом, можно автоматически определять начало рабочего режима, чтобы избежать шумы и ложные сигналы при возникновении вибрации мембраны в пьезодатчике.

2.2 Сглаживание измеряемых данных

Для повышения воспроизводимости результатов анализа производится предварительное сглаживание исходных данных. Сглаживанию подвергаются все измеряемые данные, в том числе и те, которые не относятся к рабочему режиму. Рассмотрим три известных алгоритма:

- 1) экспоненциальное сглаживание;
- 2) сглаживание методом Савицкого-Голея;
- 3) сглаживание после преобразования Фурье.

Данные алгоритмы отличаются друг от друга своими методами обработки и сглаживания, но какой из них наиболее применим, рассмотрим подробнее.

2.2.1 Алгоритм экспоненциального сглаживания

Данный алгоритм применяется для сглаживания данных временных рядов с использованием функции экспоненциального окна. В то время как в простом скользящем среднем значения прошлых наблюдений взвешиваются одинаково, экспоненциальные функции используются для назначения экспоненциально убывающих весов во времени.[4]

Сглаженное значение $S_i(t_i)$ сигнала $S(t)$ момент времени t_i равно:

$$S_t = \begin{cases} C_1 & : t = 1 \\ S_{t-1} + p(C_t - S_{t-1}) & : t > 1 \end{cases}$$

где: s_t — сглаженный ряд, c_t — исходный ряд, p — коэффициент сглаживания, который выбирается экспериментально ($0 < p < 1$).

2.2.2 Алгоритм Савицкого-Голея

Данный алгоритм также позволяет фильтровать шумы и получать гладкую линию зависимости. Если ставится задача прогнозирования или получения аналитической формулы, то имеет смысл подбирать модель с помощью МНК. Однако, если нужно только избавиться от шумов, то будем использовать формулу линейного фильтра:

$$y_i = \bar{h}\bar{W} = \sum_{k=1}^n h_k W_k,$$

где вектор \bar{h} — шаблон или маска фильтра, y_i — сглаженное значение, а другие значения, попавшие в модель «скользящего окна», изображенную на рис. 3, обозначим через:

$$\bar{W} = [Y_{i-r}, \dots, Y_i, \dots, Y_{i+r}]'.$$

Пусть даны значения Y_1, Y_2, \dots, Y_N в точках X_1, X_2, \dots, X_N , причем точки заданы с равным шагом, $\Delta X = X_{i+1} - X_i = \text{const}$. Рассмотрим точку i . Зададим количество точек n , по которым будет строиться модель — ширину окна. Естественно, считать, что точки, которые лежат слева от оцениваемой, ничем не хуже, чем точки, которые лежат справа. Следовательно, ширина окна должна быть нечетным числом: r точек остаются слева, r — справа, и сама точка i . В сумме $n = 2r + 1$.

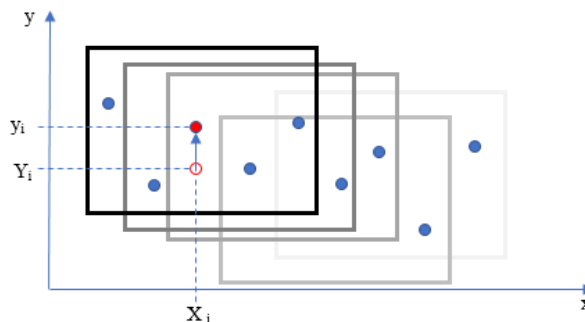


Рис. 3. Модель «скользящее окно»

Шаблон бывает полиномиальный и простой. При использовании полинома необходимо вычислять коэффициент при неизвестном значении методом МНК, что весьма затратно для ресурсов процессора. Поэтому воспользуемся простым шаблоном, который бывает двух видов: модифицированный прямоугольный и треугольный.[5]

Для прямоугольного шаблона вес центральной точки необходимо увеличить, по сравнению с остальными, например так: $\bar{h} = \frac{1}{6} [1 \ 1 \ 2 \ 1 \ 1]$.

Треугольный шаблон строится таким образом, чтобы веса равномерно возрастали от краев к центру, например: $\bar{h} = \frac{1}{16} [1 \ 2 \ 3 \ 4 \ 3 \ 2 \ 1]$.

2.2.3 Дискретное преобразование Фурье

Дискретное преобразование Фурье позволяет разложить спектр сигнала на составляющие, в котором будут отображаться полезный сигнал и шум. Формула дискретного преобразования Фурье выглядит следующим образом:

$$F[n] = \sum_{k=0}^{N-1} f[k] e^{-i \frac{2\pi}{N} nk},$$

где $F[n]$ — результирующее значение преобразования, N — общее количество значений, $f[k]$ — исходное значение сигнала, n и k — индексы массива значений, i — мнимая единица, т.к. преобразование осуществляется из N комплексных чисел.[6]

Обратное преобразование Фурье отличается лишь знаком мнимой части и результирующее значение делится на общее количество значений:

$$f[n] = \frac{1}{N} \sum_{k=0}^{N-1} F[k] e^{i\frac{2\pi}{N}nk}.$$

2.3 Сравнение алгоритмов сглаживания

При реализации трех алгоритмов были составлены их сравнительные характеристики. На рис. 4 изображен отрезок графика исходных значений, который сильно зашумлен.

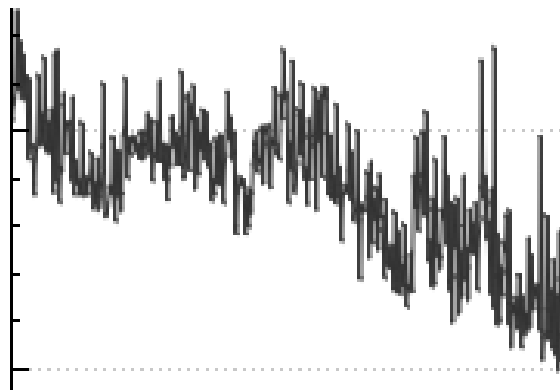


Рис. 4. Отрезок графика исходных значений

1) Экспоненциальное сглаживание;

Результаты сглаживания представлены на рис. 5 при разных коэффициентах.

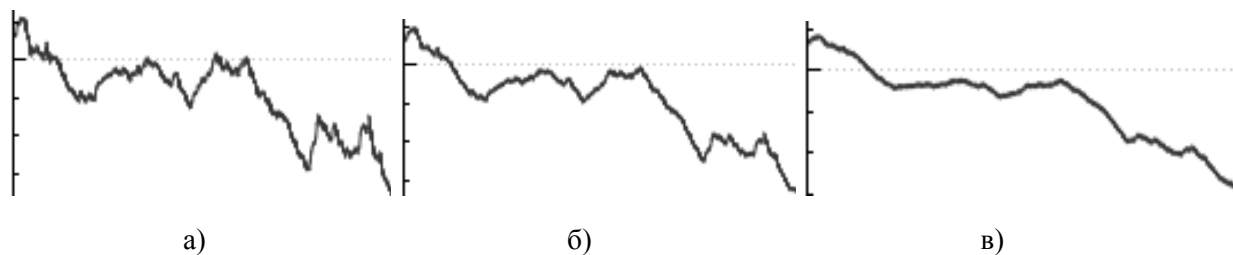


Рис. 5. Результаты сглаживания при коэффициенте

$p = 0,1$ (а), $p = 0,05$ (б) и $p = 0,2$ (в)

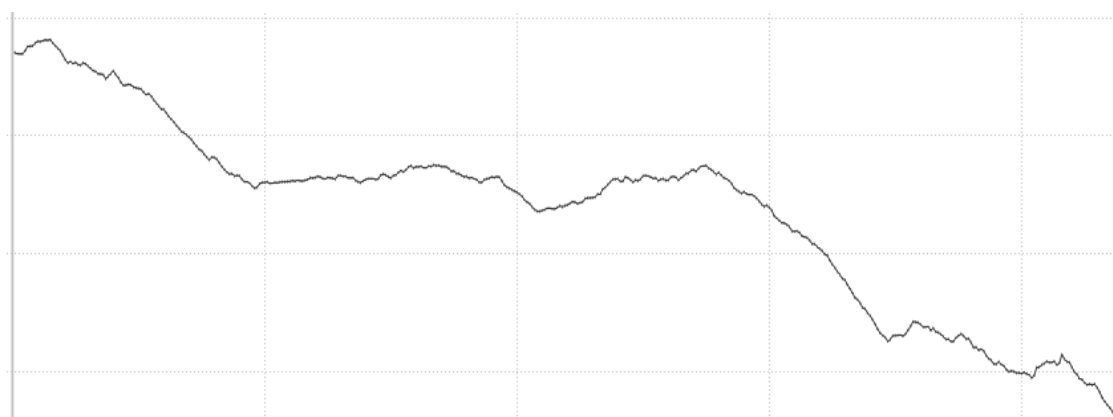


Рис. 6. Увеличенный вид графика при коэффициенте $p = 0,02$

2) Сглаживание Савицкого – Голея;

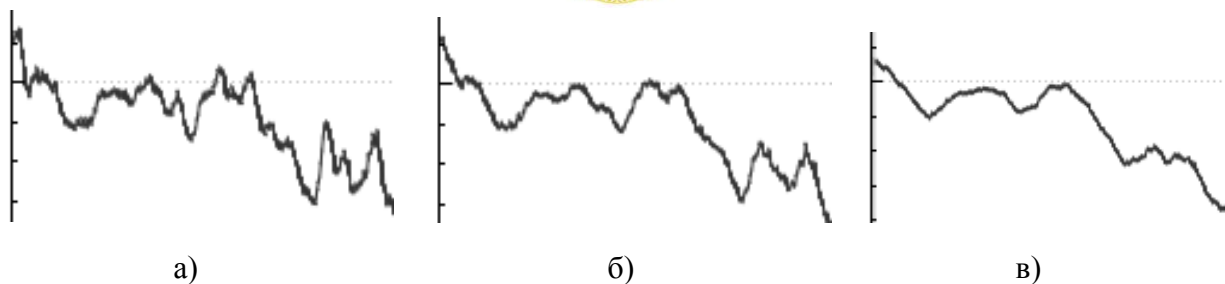


Рис. 7. Результаты сглаживания прямоугольным шаблоном «окна» при радиусе $r = 5$ (а), $r = 10$ (б) и $r = 20$ (в)

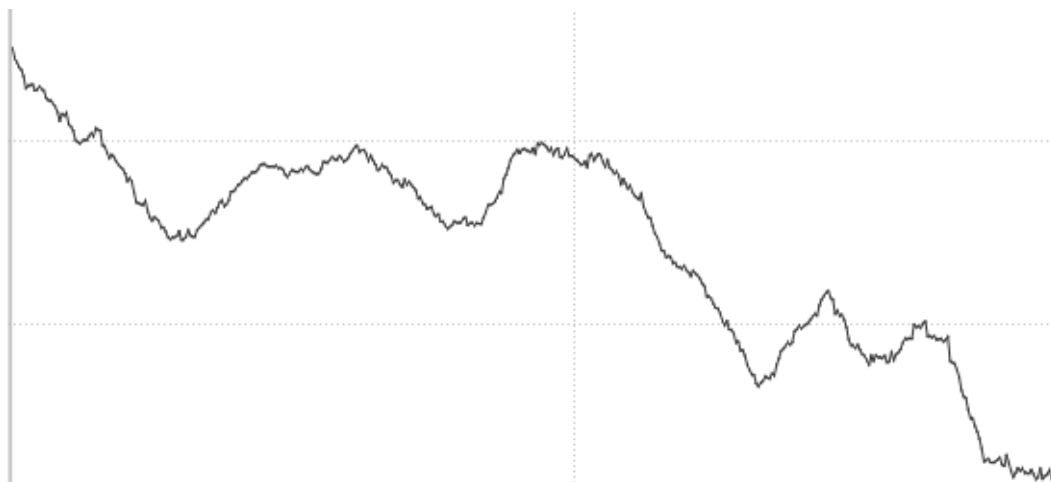


Рис. 8. Увеличенный вид графика с прямоугольным шаблоном «окна» при радиусе $r = 20$

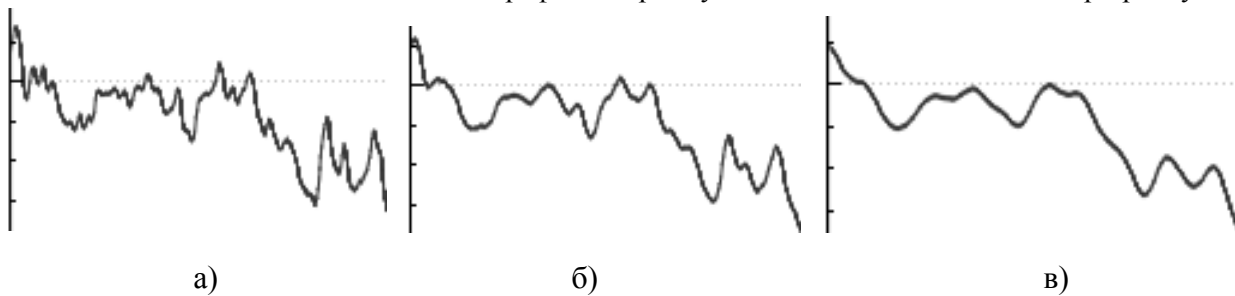


Рис. 9. Результаты сглаживания треугольным шаблоном «окна» при радиусе $r = 5$ (а), $r = 10$ (б) и $r = 20$ (в)

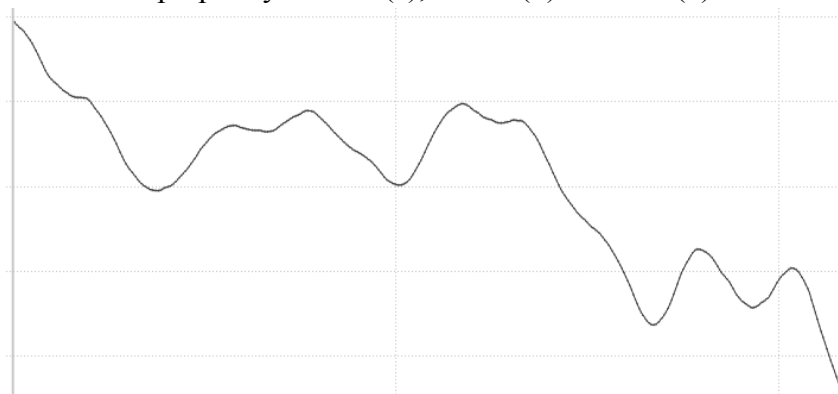


Рис. 10. Увеличенный вид графика треугольным шаблоном «окна» при радиусе $r = 20$

3) Дискретное преобразование Фурье;

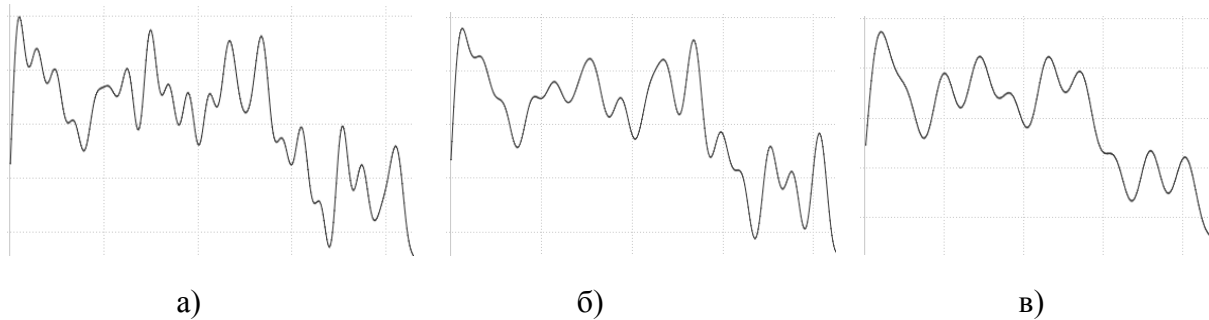


Рис. 11. Результаты сглаживания при уровне шума $L = 100$ (а), $L = 500$ (б) и $L = 1000$ (в)

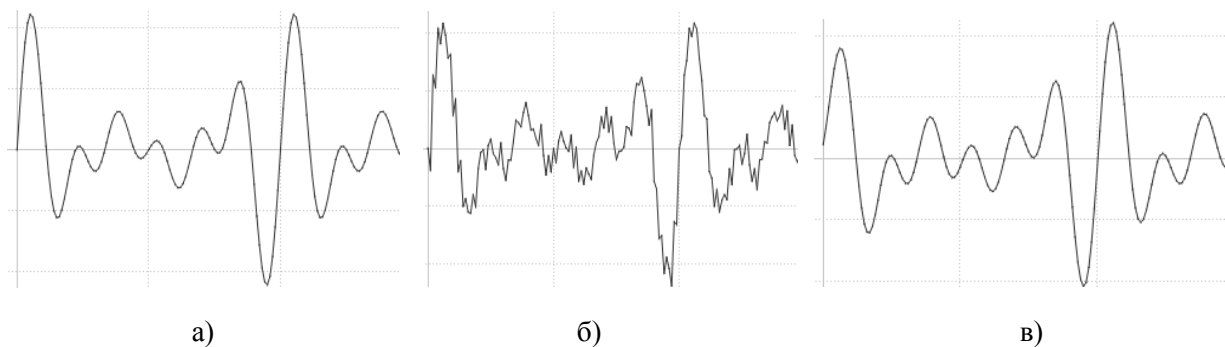


Рис. 12. Результат сглаживания (в) зашумленного гармонического сигнала (б), который идентичен исходному (полезный сигнал) (а)

Таким образом, можно сказать, что все алгоритмы сглаживания справляются со своей задачей в той или иной степени.

Алгоритм экспоненциального сглаживания и Савицкого-Голея показали практически одинаковый результат. Однако следует заметить, что в увеличенном виде заметны микрошумы (рис. 6, 7), кроме треугольного шаблона «окна» (рис. 8).

Преимущество экспоненциального сглаживания заключается в простоте и скорости алгоритма.

Преимущество фильтра Савицкого-Голея состоит в том, что он лучше фильтрует сильные шумы. Однако данный алгоритм требует больше вычислений, в отличие от экспоненциального сглаживания, но т.к. этот алгоритм выполнен через шаблон «окна», без использования МНК, то это не критично.

Дискретное преобразование Фурье сильно искажает исходный сигнал (рис. 10), даже несмотря на то, что шумы убираются. После этого было решено поменять тип сигнала на гармонический, в котором шумы убираются отлично и преобразованный сигнал восстанавливается в исходный в должном виде.

Сглаживание данных. Вывод

Таким образом, были применены на практике наиболее эффективные алгоритмы сглаживания данных, которые отличаются результатами сглаживания для разных типов сигнала и выбора участка. Они отличаются сложностью реализации, и, следовательно, временем обработки. Поэтому каждый вид сглаживания подбирается экспертом под определенный случай.

Заключение. На основании получаемых с датчиков данных выделены основные задачи по обработке. Создана система обработки результирующего сигнала, предусматривающая дальнейшее развитие.

Решены поставленные задачи:

- выделены этапы проведения измерений;
- реализован алгоритм автоматического определения начала рабочего режима;
- построена система фильтрации измеряемых данных с помощью различных алгоритмов;
- получены количественные и качественные характеристики сигнала;
- проведен анализ алгоритмов сглаживания и определены критерии выбора.

Подобная система обработки может использоваться при исследованиях вновь создаваемых и уже существующих датчиков при воздействии на них агрессивными средами.

Библиографический список

1. Никитская, Л. М. Разработка методов и устройств для измерения концентрации фенилаланина в водных растворах : автореф. дис. ... кандидата технич. наук / Л. М. Никитская. — Тамбов, 2012. — 19 с.
2. Лолита Околькова. Изменение скорости реакции. — Режим доступа : <https://distant-lessons.ru/izmenenie-skorosti-reakcii.html> (дата обращения: 12.03.2019).
3. Лоран, П. Ж. Аппроксимация и оптимизация / П. Ж. Лоран. — Москва : Мир, 1975. — 496 с.
4. Грешилов, А. А. Математические методы построения прогнозов / А. А. Грешилов, В. А. Стакун, А. А. Стакун. — Москва : Радио и связь, 1997. — 112 с.
5. Иткин, В. Ю. Интерполяция и сглаживание данных в пакете matlab / В. Ю. Иткин, О. Н. Кочуева. — Москва : РГУ нефти и газа (НИУ) имени И. М. Губкина, 2016. — 131 с.
6. Арфкен, Г. Дискретная ортогональность — дискретное преобразование Фурье / Г. Арфкен. — 3-е изд. — Орландо, Флорида : Academic Press, 1985. — 792 с.