

DOI: 10.21122/2220-9506-2025-16-4-406-416

Статистический подход к обнаружению и удалению шумовых выбросов в визуализированных данных зондовой электрометрии

О.В. Юхновская, М.А. Гундина, Е.Д. Ромашко, К.В. Пантелеев

Белорусский национальный технический университет,
пр-т Независимости, 65, г. Минск 220013, Беларусь

Поступила 22.04.2025

Принята к печати 24.09.2025

Рост объёмов измерительной информации делает ручную обработку непрактичной, а существующие алгоритмы демонстрируют ограниченную эффективность (до 93 %). Целью работы являлась разработка комбинированного алгоритма на основе синтеза существующих методов для повышения точности и надёжности автоматизированного выявления аномальных значений в данных промышленных измерительных систем, в частности, на обработку изображений, получаемых техническими приборами. Для повышения точности и надёжности обнаружения аномалий предложен комбинированный алгоритм, интегрирующий статистические методы обнаружения выбросов на основе многомерного нормального распределения для фильтрации шумовых аномалий, методы бинаризации и морфологической обработки изображений для выделения геометрических контуров, а также алгоритм анализа связных компонент для локализации области интереса. Особое внимание уделено обработке визуализированных данных (изображений), где ключевой задачей является отделение полезных аномалий от шума. Разработанный метод позволяет автоматически выделять области интереса и сжимать исходные данные более чем вдвое от их объёма без потери информативности. Также разработан подход для локализации нескольких разрозненных областей с аномальной интенсивностью на основе анализа связных компонент. Алгоритм позволяет автоматически присваивать уникальные метки каждому связному объекту и извлекать их характеристики (площадь, периметр, центроид) для последующего анализа. Реализация и тестирование проведены в среде Wolfram Mathematica на примере карт распределения поверхностного электростатического потенциала композитных материалов. Предложенные подходы упрощают последующий анализ измерительных данных и могут быть использованы совместно с алгоритмами машинного обучения.

Ключевые слова: автоматизированная обработка данных, изображения, аномальные значения, локализация области интереса, сегментация изображений

Адрес для переписки:

Пантелеев К.В.
Белорусский национальный технический университет,
пр-т Независимости, 65, г. Минск 220013, Беларусь
e-mail: k.pantsialeyeu@bntu.by

Для цитирования:

О.В. Юхновская, М.А. Гундина, Е.Д. Ромашко, К.В. Пантелеев.
Статистический подход к обнаружению и удалению шумовых выбросов в визуализированных данных зондовой электрометрии.
Приборы и методы измерений.
2025. Т. 16. № 4. С. 406–416.
DOI: 10.21122/2220-9506-2025-16-4-406-416

Address for correspondence:

Pantsialeyeu K.U.
Belarusian National Technical University,
Nezavisimosty Ave., 65, Minsk 220013, Belarus
e-mail: k.pantsialeyeu@bntu.by

For citation:

Yuhnovskaya OV, Hundzina MA, Romashko ED, Pantsialeyeu KU.
Statistical Detection and Removal of Noise Outliers in Visualized Probe Electrometry Data.
Devices and Methods of Measurements.
2025;16(4):406–416. (In Russ.).
DOI: 10.21122/2220-9506-2025-16-4-406-416

DOI: 10.21122/2220-9506-2025-16-4-406-416

Statistical Detection and Removal of Noise Outliers in Visualized Probe Electrometry Data

O.V. Yuhnovskaya, M.A. Hundzina, E.D. Romashko, K.U. Pantsialeyeu

Belarusian National Technical University,
Nezavisimosty Ave., 65, Minsk 220013, Belarus

Received 22.04.2025

Accepted for publication 24.09.2025

Abstract

The increasing volume of measurement information makes manual processing impractical, while existing algorithms demonstrate limited effectiveness (up to 93 %). The aim of this work was to develop a combined algorithm based on a synthesis of existing methods to improve the accuracy and reliability of automated anomaly detection in data from industrial measurement systems, with a particular focus on processing images obtained by technical instruments. To enhance the accuracy and reliability of anomaly detection, a combined algorithm is proposed. It integrates statistical outlier detection methods based on a multivariate normal distribution for filtering noise anomalies, image binarization and morphological processing techniques for highlighting geometric contours, as well as a connected component analysis algorithm for localizing the region of interest. Special attention is paid to the processing of visualized data (images), where the key challenge is separating useful anomalies from noise. The developed method allows for the automatic extraction of regions of interest and compresses the source data by more than half without loss of informativeness. An approach for localizing several disparate regions with anomalous intensity based on connected component analysis has also been developed. The algorithm automatically assigns unique labels to each connected object and extracts their characteristics (area, perimeter, centroid) for subsequent analysis. Implementation and testing were carried out in the Wolfram Mathematica environment using the example of surface electrostatic potential distribution maps of composite materials. The proposed approaches simplify the subsequent analysis of measurement data and can be used in conjunction with machine learning algorithms.

Keywords: automated data processing, images, anomalous values, region of interest localization, image segmentation

Адрес для переписки:

Пантелеев К.В.
Белорусский национальный технический университет,
пр-т Независимости, 65, г. Минск 220013, Беларусь
e-mail: k.pantsialeyeu@bntu.by

Address for correspondence:

Pantsialeyeu K.U.
Belarusian National Technical University,
Nezavisimosty Ave., 65, Minsk 220013, Belarus
e-mail: k.pantsialeyeu@bntu.by

Для цитирования:

О.В. Юхновская, М.А. Гундина, Е.Д. Ромашко, К.В. Пантелеев.
Статистический подход к обнаружению и удалению шумовых выбросов в визуализированных данных зондовой электрометрии.
Приборы и методы измерений.
2025. Т. 16. № 4. С. 406–416.
DOI: 10.21122/2220-9506-2025-16-4-406-416

For citation:

Yuhnovskaya OV, Hundzina MA, Romashko ED, Pantsialeyeu KU.
Statistical Detection and Removal of Noise Outliers in Visualized Probe Electrometry Data.
Devices and Methods of Measurements.
2025;16(4):406–416. (In Russ.).
DOI: 10.21122/2220-9506-2025-16-4-406-416

Введение

Актуальной задачей современного приборостроения и промышленной аналитики является разработка эффективных методов автоматизированного выявления аномальных значений (грубых выбросов) в данных, получаемых от технических приборов [1–5]. Сложность современных научных исследований, связанных с обработкой таких данных, приводит к резкому увеличению объёмов измерительной информации, ручная обработка которой становится практически невозможной [1, 2]. Особую значимость эта проблема приобретает при работе с изображениями, получаемыми промышленным оборудованием, где аномалии могут выступать как в роли шума (например, краевые эффекты при сканировании), так и в роли полезного сигнала (дефекты, включения, области с аномальной физической характеристикой) [6].

При этом вопрос построения универсального математического аппарата для надёжного обнаружения и разделения этих типов аномалий остаётся открытым [3–5]. Существующие алгоритмы, основанные на анализе проекций, скользящих окнах, вероятностных нейронных сетей или анализе краёв [7, 8], демонстрируют ограниченную эффективность (до 93 %) и часто требуют априорных знаний о природе наблюдаемых отклонений.

Целью работы являлась разработка комбинированного алгоритма на основе синтеза существующих методов для повышения точности и надёжности автоматизированного выявления аномальных значений в данных промышленных измерительных систем, с особым акцентом на обработку изображений, получаемых техническими приборами.

Подходы и методы определения аномальных значений на изображениях

Кластерный анализ, как совокупность алгоритмов группировки объектов по сходству признаков, предлагает мощный инструмент для решения задачи сжатия массивов информации и выявления аномальных значений. Особую ценность представляет его способность выявлять аномалии без априорных знаний о природе отклонений. В частности, основанный на принципе пространственной локализации

метод предполагает, что нормальные данные концентрируются в плотных кластерах, тогда как аномалии либо значительно удалены от центроидов, либо образуют редкие малочисленные группы. Эта особенность, в сочетании с возможностью работы в режиме реального времени без ресурсоёмкого обучения, делает кластерные подходы критически важными для обеспечения достоверности последующего анализа измерительных данных.

Язык Wolfram Language имеет широкую поддержку иерархических и неиерархических методов кластерного анализа. В иерархических методах каждый экземпляр данных образует отдельный кластер. Затем процесс объединения кластеров происходит последовательно на основании матрицы расстояний.

Методы иерархического кластерного анализа отличаются используемыми мерами близости и алгоритмами классификации. Для функции ClusteringComponent системы можно определить опцию, которая определяет формулу нахождения расстояний.

$\text{ManhattanDistance}[u, v]$ определяется следующим образом:

$$\|u - v\|_1 = \sum_{i=1}^n |u_i - v_i|.$$

$\text{ChessboardDistance}[u, v]$ определяется следующим образом:

$$\|u - v\|_2 = \max_{1 \leq i \leq n} |u_i - v_i|.$$

$\text{EuclideanDistance}[u, v]$ определяется как:

$$\|u - v\|_2 = \sqrt{\sum_{i=1}^n |u_i - v_i|^2}.$$

Одним из фундаментальных подходов к выявлению аномальных значений является использование параметрических статистических методов, опирающихся на свойства распределений вероятностей. Классическим примером служит применение нормального распределения Гаусса для вещественных случайных величин: зная параметры распределения (математическое ожидание и дисперсию), можно идентифицировать аномалии как значения, выходящие за пределы заданного доверительного интервала.

Одним из используемых подходов является мультипликативная модель оценки аномальных значений, основанная на двумерном нормальном распределении с весовыми коэффициентами

значимости [9]. Данная модель позволяет эффективно адаптировать математический аппарат к задачам обнаружения выбросов при сохранении интерпретируемости результатов.

Пусть нормально распределённая случайная величина с математическим ожиданием μ и дисперсией σ^2 :

$$X \sim N(\mu, \sigma^2).$$

Плотность вероятности распределения находится по следующей формуле:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}.$$

Если случайные величины X, Y независимы, плотность распределения может быть найдена по формуле:

$$f(x, y) = \frac{1}{2\pi\sigma_x\sigma_y} e^{-\frac{1}{2}\left(\frac{x-\mu_x}{\sigma_x}\right)^2} e^{-\frac{1}{2}\left(\frac{y-\mu_y}{\sigma_y}\right)^2}.$$

Также этот случай можно и далее распространить на m -мерное пространство признаков.

Предположим, что в рассматриваемом наборе данных был только один признак, и этот признак имел нормальное распределение, тогда можно смоделировать алгоритм обнаружения аномалий, используя $f(x)$. Также можно установить некоторое пороговое значение ε , которое будет определять, является ли случай аномальным или нет. ε следует задавать эвристически, и его значение будет зависеть от варианта использования и предпочтительной чувствительности к аномалиям.

Пусть обучающий набор с m элементами $(x^1, x^2, \dots, x^m)$, где x^i является n -мерным вектором в общем случае, верхний индекс обозначает номер рассматриваемого признака.

Для каждого признака x^i вычислим параметры μ и σ^2 .

$$(\mu^1, \mu^2, \dots, \mu^m), (\sigma^{21}, \sigma^{22}, \dots, \sigma^{2m}), \text{ где:}$$

$$\mu^i = \frac{1}{n} \sum_{k=1}^n x_k^i; \quad \sigma^{2i} = \frac{1}{n} \sum_{k=1}^n (x_k^i - \mu^i)^2;$$

i – номер элемента выборки; j – номер признака.

Поскольку параметры являются независимыми, то умножая f^i можно построить функцию правдоподобия:

$$f(x^1, \dots, x^m) = f(\mu^1, \sigma^{21}) \dots f(\mu^m, \sigma^{2m}).$$

По этой формуле определяется вероятность того, что этот новый элемент является аномальным.

Теперь, когда получено значение функции f , проверяем меньше ли оно значения ε .

Условия определения аномального значения:

$$\text{sign}(x) = \begin{cases} 1, & f < \varepsilon \\ 0, & f \geq \varepsilon \end{cases}.$$

Важным является замечание, что, если значение является аномальным по некоторому единичному признаку x^i , то это не означает, что данное значение будет аномальным при учёте всех признаков.

Рассмотрение особенностей удаления значений, не входящих в область интереса, после выявления аномальных значений на снимке, полученном промышленным оборудованием

Для применения разработанного математического аппарата к обработке изображений каждый пиксель рассматривался как независимая выборка, а его координаты (x, y) – как двумерный вектор признаков. Параметры математического ожидания (μ_x, μ_y) и ковариационной матрицы Σ оценивались по всем пикселям исходного изображения (рисунок 1а). Пороговое значение ε для функции правдоподобия $f(x)$ было установлено эмпирически на уровне 0,02. Данное значение было выбрано в результате анализа гистограммы распределения $f(x)$ по всему изображению и позволило эффективно отделить плотное ядро данных, соответствующее полезному сигналу (области образца), от разреженных шумовых выбросов, вызванных краевыми эффектами.

Пиксели, для которых значение $f(x)$ оказывалось ниже порога ε , классифицировались как аномальные (шумовые) и исключались из дальнейшего анализа. Таким образом, массив данных был отфильтрован, и последующие шаги (медианная фильтрация, бинаризация, выделение связных компонент) применялись уже к очищенным данным, что существенно повысило их качество и упростило выделение области интереса.

Отработку методик и алгоритмов автоматической обработки статических изображений, содержащих аномальные значения,

выполняли на картах распределения поверхностного электростатического потенциала композитных полимерных материалов на основе Фторопласта Ф-4, наполненных углеродным наноматериалом. Изображения получены при сканировании поверхности образца бесконтактным зарядочувствительным зондом (аналог сканирующего зонда Кельвина) [10] с шагом сканирования 2 мм по меандру и последующей бикубической интерполяцией измерительных данных. Выбор изображения для отработки алгоритмов локализации области интереса обусловлен наличием шумов, которые можно исключить автоматически для более корректного последующего статистического анализа интересных измерительных данных.

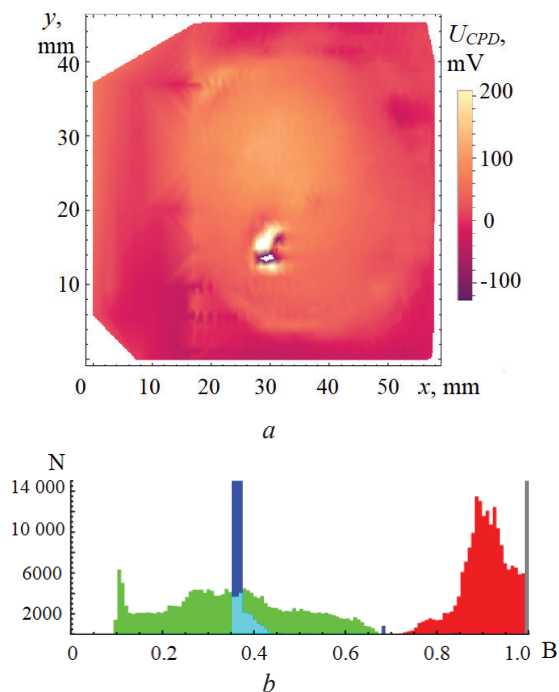


Рисунок 1 – Исходное изображение (а) – карта распределения поверхностного электростатического потенциала полимерного композиционного материала на основе Фторопласта Ф-4, и гистограмма распределения пикселей яркости светового пространства для исходного изображения (b)

Figure 1 – Original image (a): map of the surface electrostatic potential distribution of a polymer composite material based on Fluoroplast F-4, and a histogram of pixel brightness distribution in the light space for the original image (b)

В результате картирования поверхностного электростатического потенциала наблюдается существенное ухудшение пространственного

разрешения в краевых зонах (рисунок 1а). Это проявляется в снижении контрастности цветового кодирования на периферии, отклонении формы карты потенциалов от геометрии образца, в повышенной погрешности измерений потенциала по всей поверхности с максимальными отклонениями в приграничных зонах. Таким образом, областью интереса являлась непосредственно область на изображении, соответствующая области поверхности образца, представляющая собой круг диаметром 50 мм.

Одной из ключевых причин наблюдаемых искажений является эффект краевого поля. В местах резкого изменения геометрии образца линии электрического поля искривляются, создавая аномалии в измерениях. Датчик прибора фиксирует не только потенциал поверхности, но и паразитные поля от боковых граней и подложки (в данном случае – предметный столик), что приводит к размытости показаний. Свой вклад также вносят ограничения разрешающей способности измерительного зонда. В случае методов на основе зонда Кельвина ограничения зонда по углу наклона и расстоянию до поверхности особенно заметно проявляются при сканировании крутых краёв образца.

На рисунке 1b представлена гистограмма распределения яркости пикселей B светового пространства в случае исходного изображения. По оси абсцисс расположены значения уровня яркости заданного цвета, а высоты столбцов гистограммы пропорциональны частоте N зафиксированных значений яркости. В условиях обычного и ровного освещения при правильно подобранной выдержке гистограмма обычно показывает максимум в центральной части, спадая в области теней и яркостей. Гистограмма может описать степень контраста в изображении. Контраст является измерением разницы яркостей между светлыми и тёмными частями изображения. Широкие гистограммы отражают сцены со значительным контрастом, тогда как узкие гистограммы означают, что контраст снижен, и изображение может оказаться плоским или малоинформативным.

Для наглядной визуализации распределения поверхностного электростатического потенциала было применено цветовое кодирование в Hue (тон) модели HSV. Данное представление было выбрано благодаря его способности линейно

и интуитивно отображать непрерывный диапазон скалярных величин на цветовой круг, что облегчает визуальную интерпретацию данных. Все этапы обработки (фильтрация, бинаризация, анализ компонент) выполнялись алгоритмически над исходным массивом измерительных данных (матрицей значений потенциала) в среде Wolfram Mathematica. Визуализации на рисунках 2а и 3 являются результатом этих вычислений.

Преобразование динамического диапазона – «выравнивание гистограммы» (рисунок 2а) применялось к данным для повышения контрастности визуализации. Как показал анализ, прямое пороговое преобразование без данного этапа также приводит к удовлетворительным результатам для выделения области интереса круглой формы. Таким образом, этап выравнивания не является строго обязательным в рамках предложенного алгоритма для данного типа задач и может быть опущен для сокращения вычислительных затрат.

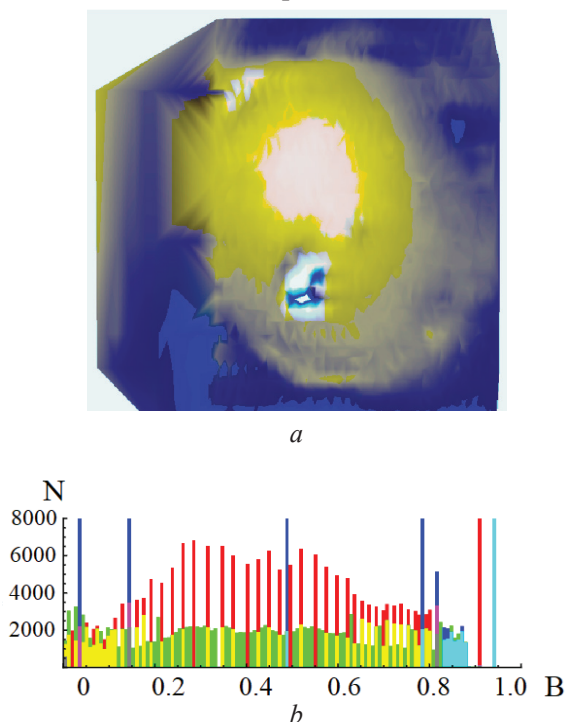


Рисунок 2 – Изображение с выровненной гистограммой распределения яркости (а) и гистограмма распределения яркости (b)

Figure 2 – Image with equalized brightness histogram (a) and brightness distribution histogram (b)

Широкое распространение в системах технического зрения получили усредняющие и

медианные фильтры. К данному изображению применяется медианный фильтр. Затем осуществляется преобразование изображения в полутоновое и удаление незначительных по величине компонент (рисунок 3).



Рисунок 3 – Бинарное изображение после обработки

Figure 3 – Binary image after processing

Одномерный медианный фильтр представляет собой скользящее окно, охватывающее нечётное число элементов изображения. Центральный элемент заменяется медианой всех элементов изображения в окне. Окно перемещается вдоль фильтруемого сигнала, и вычисления повторяются.

Существуют рекомендации, согласно которым начинать следует с медианного фильтра, окно которого охватывает три элемента изображения [12]. Если ослабление сигнала оказывается незначительным, окно фильтра расширяют до пяти элементов.

Для качественного распознавания контуров изображения границы должны быть яркими (что характерно для бинарного изображения), тонкими и без разрывов. Для получения чёткого контура после фильтрации необходимо повысить контрастность изображения (усилить перепады яркости).

Для перевода изображения из полутонового в бинарное применяется пороговое преобразование. Критически важным является выбор величины порога: при слишком высоком пороге могут возникать разрывы контура, при слишком низком могут появляться неинформативные пиксели.

Выделение границ основано на алгоритмах, которые обнаруживают точки цифрового изображения с резкими изменениями яркости или другими видами неоднородностей.

Основной целью обнаружения резких изменений яркости изображения является фиксация

важных перепадов и изменений. Они могут отражать различные предположения о модели формирования изображения, а изменения в яркости изображения могут указывать на изменения глубины, ориентации поверхностей, изменения в свойствах материала и др.

В идеальном случае результатом выделения границ является набор связанных кривых. Эти кривые обозначают границы объектов, грани и отгибки потенциального рельефа, а также отображают изменения положения поверхностей.

Далее осуществляется определение координат прямоугольной области интереса и строятся соответствующие границы этой области (рисунок 4).

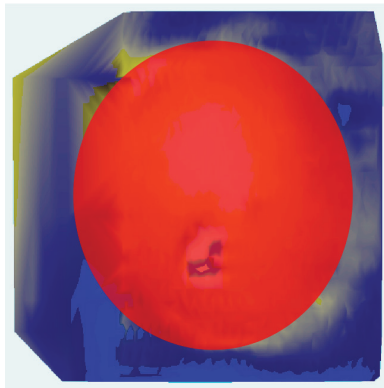


Рисунок 4 – Выделение области интереса

Figure 4 – Region of interest selection

Исходный массив данных проверяется на наличие аномальных значений. На основе этого критерия строится новый массив данных, содержащий только значения из области интереса (ROI). Значения, не входящие в область интереса, отбрасываются. В результате обработки исходной выборки, содержащей 2131 элемент, была получена выборка из 1149 элементов. Это позволило сжать исходные данные до 53,9 % от их первоначального объема (рисунок 5).

Анализ гистограмм распределения (рисунок 6b и c) показывает, что применение алгоритмов локализации области интереса эффективно устраняет ложные экстремумы (шумы), сохраняя при этом информацию о реальных аномалиях (дефекты, включения, адсорбция загрязнений и др.), представляющих исследовательский интерес.

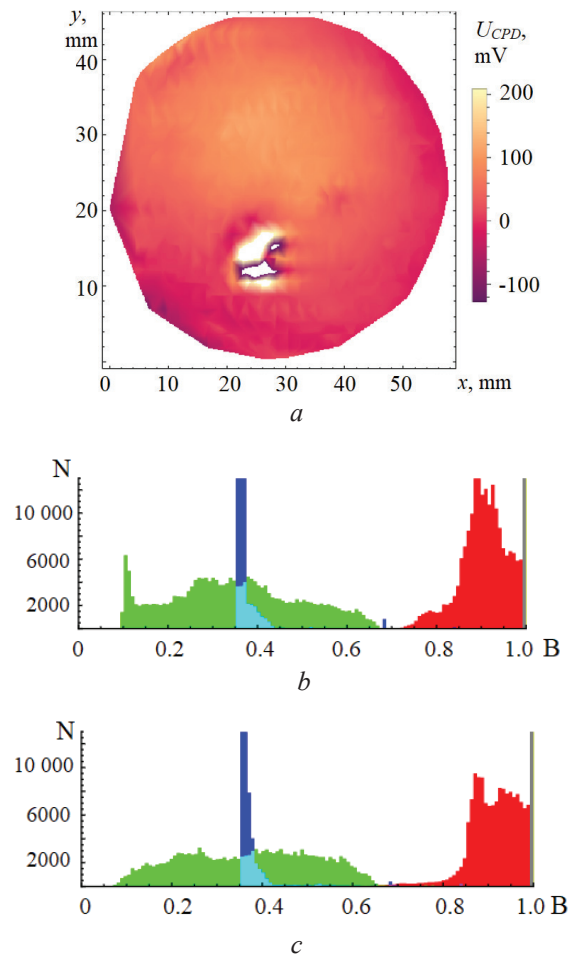


Рисунок 5 – Автоматически выделенная область интереса в исходной системе координат (a). Гистограммы распределения поверхностного электростатического потенциала до (b) и после (c) выделения области интереса из карт распределения поверхностного электростатического потенциала

Figure 5 – Automatically selected region of interest in the original coordinate system (a). Histograms of surface electrostatic potential distribution before (b) and after (c) region of interest extraction from surface electrostatic potential distribution maps

Таким образом, предложенный алгоритм позволил сжать исходные данные в 1,85 раза (или до 53,9 % от исходного объема), что существенно упрощает последующую обработку.

Анализ изображения, содержащего несколько областей с аномальной интенсивностью

При обработке снимков, полученных промышленным оборудованием, часто возникает задача анализа изображений, содержащих

несколько разрозненных областей с аномальной интенсивностью (как повышенной, так и пониженной). В таких условиях традиционные подходы, основанные на разрастании регионов, выделении границ или кластеризации, оказываются неэффективными. Это связано с тем, что указанные области могут сливаться в процессе обработки, что приводит к потере информативности исходного изображения и невозможности корректно выделить отдельные аномалии. Для выделения нескольких областей с аномальной интенсивностью наиболее эффективно использовать алгоритм, основанный на анализе связных компонентов.

Понятие связности между элементами изображения является одним из фундаментальных понятий, которые упрощают определение большого числа других понятий, связанных с цифровыми изображениями, например, такими как область и граница. Концепция смежности пикселей интуитивна понятна, но для корректного описания формализуем её.

Рассмотрим пиксел изображения с координатами (r, c) . Четвёркой его соседей по вертикали и горизонтали будет следующее множество:

$$\{(r+1, c), (r-1, c), (r, c+1), (r, c-1)\}.$$

Таким же подходом определяется и множество диагональных элементов:

$$\{(r+1, c+1), (r-1, c+1), (r+1, c-1), (r-1, c-1)\}.$$

Объединение двух вышеописанных множеств называется восьмёркой соседей.

Таким образом, для двумерного изображения существуют две стандартные связности:

– 4-связность – пиксели соединяются, если их края соприкасаются. Два соседних пикселя являются частью одного и того же объекта, если они оба включены и соединяются по горизонтали или вертикали;

– 8-связность – пиксели соединяются, если их края или углы соприкасаются. Два смежных пикселя являются частью одного и того же объекта, если они оба соединяются по горизонтали, вертикали или диагонали.

Два подмножества A и B называются смежными, если существует хотя бы один пиксел из A , смежный с пикселем из B .

Дискретным путём длины n между двумя пикселями p и q называется неповторяющаяся

последовательность пикселей с координатами, где каждый соседний являются смежными.

Пусть S – некоторое подмножество элементов изображения. Два его элемента p и q называются связными в S , если между ними существует путь, целиком состоящий из элементов подмножества S . Для любого пикселя p из множества S совокупность всех пикселей, связанных с ним в S , называется компонентой связности.

Рассмотрим результат действия алгоритма выделения областей интереса на основе связности, на примере анализа изображения, приведённого на рисунке 6. Связным объектом будем считать множество точек, в котором каждая пара точек удовлетворяет некоторому критерию связности.

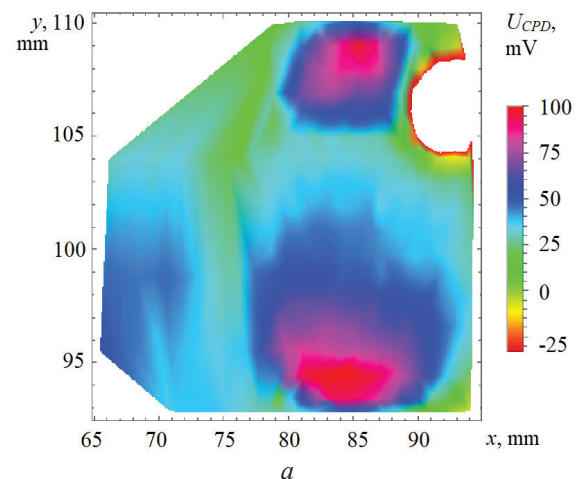


Рисунок 6 – Распределение плотности исходных данных (a) и выделение контрастных областей на изображении (b)

Figure 6 – Density distribution of the original data (a) and extraction of contrast regions in the image (b)

Под выделением связных компонент понимают присвоение уникальной метки

каждому объекту изображения. При последующем анализе данных метки служат в качестве идентификаторов при обращении к объектам. Так, например, при использовании функции ComponentMeasurements можно с помощью данных идентификаторов обращаться к каждой компоненте связности на изображении. Такой подход позволяет автоматизировать последовательность действий и провести ряд операций в цикле с каждой компонентой связности.

Такая группировка смежных пикселей позволяет получить необходимые для последующего анализа свойства объектов, такие как, например, периметр и площадь области интереса. Поэтому задача выделения связанных компонентов является фундаментальной задачей обработки изображений.

На рисунке 7 приведены результаты локализации нескольких областей интереса на изображении.

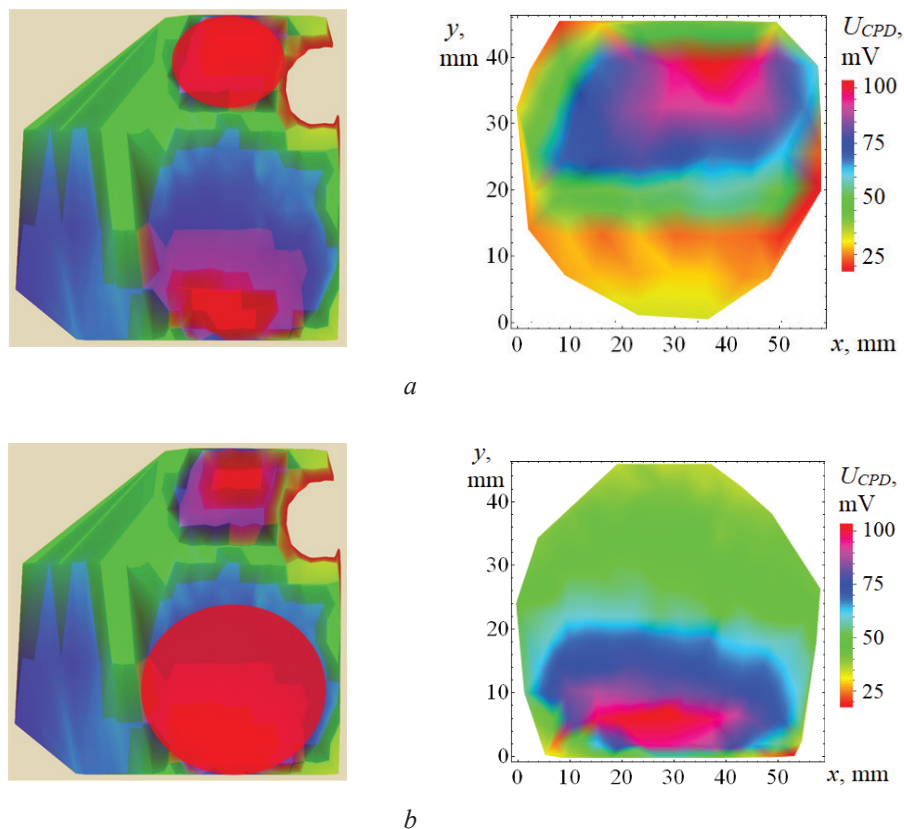


Рисунок 7 – Локализация «верхней» (а) и «нижней» (б) областей интереса на изображении

Figure 7 – Localization of the "upper" (a) and "lower" (b) regions of interest in the image

Следует отметить, что при построении контуров могут быть учтены дополнительные параметры для наилучшего соответствия, в частности, измерение формы, свойства граничного контура, свойства ограничивающего прямоугольника, особенности области, учёт топологических свойств, свойства гистограммы для компонента, свойства интенсивности изображения, статистические измерения и др.

Важным преимуществом предложенного алгоритма является его гибкость и адаптивность.

Ключевые параметры, такие как порог обнаружения аномалий (ϵ), размер ядра медианного фильтра и тип связности (4- или 8-связность) для анализа компонент, не являются жёстко фиксированными. Они могут быть настроены в зависимости от специфики исследуемого объекта (его геометрии, материала) и качества исходных данных (уровня шума, контрастности). Это не препятствует автоматизации, а, напротив, указывает на универсальность метода. Для работы с различными типами объектов

может быть разработан модуль предварительного анализа, который на основе оценки характеристик входного изображения будет автоматически подбирать оптимальный набор параметров для последующей обработки, что является темой наших дальнейших исследований.

Заключение

Разработан и апробирован комплексный подход к автоматизированному выявлению аномальных значений в данных промышленных измерительных систем, с особым акцентом на обработку изображений. Новизна исследования заключается в синтезе методов статистического анализа (многомерное нормальное распределение) и алгоритмов обработки изображений (анализ связанных компонент), что позволило создать адаптивный алгоритм, не требующий априорных знаний о природе отклонений.

Важным практическим результатом работы является разработка метода локализации областей интереса на изображениях, полученных промышленным оборудованием. Предложенный алгоритм продемонстрировал высокую эффективность при обработке сложных изображений, содержащих множественные разрозненные аномалии. Реализованный подход позволяет не только идентифицировать, но и количественно характеризовать области интереса через такие параметры как площадь, периметр и положение центра.

Практическая апробация методики подтвердила её эффективность. Разработанный алгоритм позволил сократить объём обрабатываемых данных более чем вдвое при сохранении информативной ценности и устранении артефактов, связанных с краевыми эффектами. Эффективность предложенного комбинированного метода составила 100 % для задачи подавления шумовых аномалий при сохранении полезных сигналов, что превосходит показатели традиционных алгоритмов (до 93 %).

Научно-практическая значимость работы заключается в создании универсального инструментария для обработки измерительных данных различной физической природы (тепловые, оптические, электростатические

поля). Перспективность предложенных подходов заключается в их совместимости с современными технологиями машинного обучения.

Дальнейшее развитие работы предполагает адаптацию методики для обработки потоковых данных в реальном времени и интеграцию с глубокими нейронными сетями для повышения точности классификации аномалий.

Таким образом, представленные в исследовании методы и алгоритмы создают основу для нового поколения интеллектуальных систем анализа измерительной информации, способных эффективно работать с большими объемами разнородных данных в условиях промышленной эксплуатации.

References /

Список использованных источников

1. Bell G, Hey T, Szalay A. Beyond the Data Deluge. Science. 2009;323(5919):1297-1298.
DOI: 10.1126/science.117041
2. Liu Y, Wang H, Zhang Q. Robust anomaly detection for industrial sensor networks with uncertain data quality. IEEE Transactions on Industrial Informatics. 2023;(7):8765-8775.
DOI: 10.1109/TII.2022.3221234
3. Chandola V, Banerjee A, Kumar V. Anomaly detection: A survey. ACM Computing Surveys. 2009;41(3):1-58.
DOI: 10.1145/1541880.154188
4. Qiao Y. [et al.] A multihead attention self-supervised representation model for industrial sensors anomaly detection. IEEE Transactions on Industrial Informatics. 2023;20(2):2190-2199.
DOI: 10.1109/TII.2023.3280337
5. Ruff L. [et al.] A unifying review of deep and shallow anomaly detection. Proceedings of the IEEE. 2021;109(5):756-795.
DOI: 10.1109/JPROC.2021.3052449
6. Weihua W. License Plate Recognition Algorithm Based on Radial Basis Function Neural Networks. Intelligent Ubiquitous Computing and Education, 2009 International Symposium on. 2009;38-41.
7. Anagnostopoulos C, Anagnostopoulos L, Loumos V, Kayafas E. A license plate-recognition algorithm for intelligent transportation system applications. IEEE Trans. Intell. Transp. Syst. 2006;7(3):377-392.
DOI: 10.24425/ijet.2023.144361
8. Jiao J, Ye Q, Huang Q. A configurable method for multi-style license plate recognition. Pattern Recognition. 2009;42(3):358-369.
DOI: 10.1016/j.patcog.2008.08.016

9. Гундина, М.А. Особенности процесса определения аномальных значений при обработке изменительной информации / М.А. Гундина, П.С. Богдан, О.В. Юхновская // Вестник Белорусско-Российского университета. – 2024. – № 2. – С. 96–103.

DOI: 10.24412/2077-8481-2024-2-96-103

9. Hundzina MA, Bohdan PS, Yukhnouskaya VV. Features of the process of determining the number. Be-

larusian-Russian University Bulletin. 2024;83(2):96-103.

DOI: 10.24412/2077-8481-2024-2-96-103

10. Zharin AL, Mikitsevich UA, Svistun AI, Pantsialeu KU. Universal Digital Probe Electrometer for Testing Semiconductor Wafers. *Devices and Methods of Measurements*. 2023;14(3):161-172.

DOI: 10.21122/2220-9506-2023-14-3-161-172

11. Gonzalez RC, Woods RE. Digital image processing. 4th ed. Pearson. 2018;1168 p.