

Министерство науки и высшего образования Российской Федерации  
Санкт-Петербургский политехнический университет Петра Великого  
Физико-механический институт

Высшая школа фундаментальных физических исследований

Работа допущена к защите  
Директор ВШФИ  
\_\_\_\_\_ В.В. Дубов  
«\_\_\_» \_\_\_\_\_ 2023 г.

**ВЫПУСКНАЯ КВАЛИФИКАЦИОННАЯ РАБОТА  
РАБОТА БАКАЛАВРА**

**АЛГОРИТМ АВТОМАТИЗИРОВАННОЙ ОБРАБОТКИ  
РЕНТГЕНОСКОПИЧЕСКИХ ИЗОБРАЖЕНИЙ ДЛЯ АНАЛИЗА ГРУЗОВ  
В ИНСПЕКЦИОННО-ДОСМОТРОВОМ КОМПЛЕКСЕ НА БАЗЕ  
БЕТАТРОНА**

по направлению подготовки 03.03.02 – Физика

Профиль 03.03.02\_01 – Физика атомного ядра и элементарных частиц

Выполнил

студент гр. 5030302/90101 \_\_\_\_\_

А.В. Иванов

Руководители

профессор ВШФИ, д.ф.м.н., \_\_\_\_\_

Я.А. Бердников

в.н.с. научного подразделения \_\_\_\_\_

И.И. Врубель

ООО «Индиком», к.ф.м.н.

Консультант

по нормоконтролю \_\_\_\_\_

И.Г. Голиков

Санкт-Петербург – 2023

**САНКТ-ПЕТЕРБУРГСКИЙ ПОЛИТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ  
ПЕТРА ВЕЛИКОГО**

**Высшая школа фундаментальных физических исследований**

**УТВЕРЖДАЮ**

**Директор ВШФФИ**

**\_\_\_\_\_ В.В. Дубов**

**«\_\_\_\_» \_\_\_\_\_ 2023г.**

**ЗАДАНИЕ**

**по выполнению выпускной квалификационной работы**

**студенту Иванову Александру Викторовичу гр.5030302/90101**

**фамилия, имя, отчество (при наличии), номер группы**

**1. Тема работы: алгоритм автоматизированной обработки рентгеноскопических изображений для анализа грузов в инспекционно-досмотровом комплексе на базе бетатрона**

**2. Срок сдачи студентом законченной работы: 07 июня 2023**

**3. Исходные данные по работе: характеристики инспекционно-досмотрового комплекса, результаты измерений.**

**4. Содержание работы (перечень подлежащих разработке вопросов):**

При подготовке ВКР необходимо решить следующие задачи: 1. Провести анализ литературы в сфере применения инспекционно-досмотровых комплексов на базе бетатрона. 2. Исследовать алгоритмы для кластеризации рентгеноскопического изображения. 3. Испытать выбранные алгоритмы на практике алгоритмов кластеризации.

**5. Перечень графического материала (с указанием обязательных чертежей): Рисунки, графики, чертежи, схемы и таблицы в соответствии с содержанием ВКР**

**6. Дата выдачи задания 02 февраля 2023**

**Руководитель ВКР \_\_\_\_\_ Я.А. Бердников \_\_\_\_\_**

**(подпись)**

**инициалы, фамилия**

Задание принял к исполнению «03» февраля 2023 г.

(дата)

Студент

\_\_\_\_\_

А.В. Иванов

(подпись)

инициалы, фамилия

## **РЕФЕРАТ**

На 40 с., 25 рисунков, 2 таблицы

**КЛЮЧЕВЫЕ СЛОВА:** Кластеризация, инспекционно-досмотровый комплекс, бетатрон, рентгеновское излучение, рентгеноскопическое изображение, обработка изображения, портал

Тема выпускной квалификационной работы: «алгоритм автоматизированной обработки рентгеноскопических изображений для анализа грузов в инспекционно-досмотровом комплексе на базе бетатрона».

Данная работа посвящена обработке и изучению экспериментальных данных, полученных на действующем инспекционно досмотровом комплексе, разработанным компанией “Индиком”. Основным направлением работы является разработка алгоритма кластеризации рентгеноскопического изображения. Проведён литературный обзор по темам: инспекционно-досмотровые комплексы, получение рентгеноскопического изображение, обработка рентгеноскопического изображения. Проведен обзор и описание установки компании “Индиком”. Исследованы различные методы для получения кластеризованного рентгеноскопического изображения.

## **Abstract**

This paper contains 40 pages, 25 pictures, 2 tables.

**KEYWORDS:** Clustering, inspection system, betatron, X-ray radiation, X-ray imaging, image processing

The subject of the graduate qualification work is "Algorithm for Automatic Processing of X-ray Images for Cargo Analysis in an Inspection System Based on a Betatron".

This work is devoted to the processing and study of experimental data obtained at the inspection system developed by LLC "Indicom". The main focus of the work is the application of an algorithm for clustering X-ray images. A literature review was conducted on the topics of inspection-screening complexes, X-ray image acquisition, and processing of X-ray images. An overview and description of the system developed by "Indicom" were provided. Various methods were applied for obtaining clustered X-ray images.

<b>Оглавление</b>	
<b>Введение .....</b>	7
<b>1. Обзор литературы.....</b>	9
1.1. Современный уровень техники .....	9
1.2. Физические принципы работы инспекционно-досмотровых комплексов .....	10
1.3. Известные методы кластеризации.....	11
1.4 Алгоритмы различных методов кластеризации.....	13
1.4.1 Алгоритм метода k- средних .....	13
1.4.2 Алгоритм кластеризации с-средних.....	14
1.4.3 Алгоритм “лидера” .....	15
1.4.4 Алгоритм mean-shift. .....	18
<b>2. Самостоятельная работа.....</b>	20
2.1. Источник рентгеноскопического излучения .....	20
2.2. Особенности досмотровой установки.....	21
2.3. Получение изображений.....	22
2.4. Применение различных методов кластеризации к полученным изображениям .....	24
2.5. Применение методов кластеризации с целью определения разных веществ. ....	32
<b>Заключение.....</b>	37
<b>Список литературы .....</b>	39

## **Введение**

Для досмотра грузовых транспортных средств используется специальное оборудование неразрушающего рентгеновского контроля. Этот метод позволяет обнаруживать незаконные предметы, которые могут быть скрыты в грузовых отсеках, и предотвращать их незаконное перемещение через границы.

Рентгеновский досмотр грузовых автомобилей - это высокоэффективный метод, который широко применяется на границах и в морских портах для проверки транспорта. Рентгеновские лучи проникают через кузов автомобиля, создавая изображение его внутренней структуры на специальной пленке или экране. Этот метод позволяет обнаруживать скрытые предметы, запрещенные к провозу, а также несоответствие декларируемого и перевозимого груза, что помогает снизить уровень преступности и поддерживает безопасность на границах. Многие системы рентгеновского досмотра работают с использованием метода дуальных энергий [16]. При помощи этого метода возможно идентифицировать эффективный атомный номер ( $Z_{eff}$ ) содержимого объекта контроля в крупногрупповом приближении. Для визуализации принадлежности материала к одной из групп

- органика ( $Z_{eff} < 1$ );
- неорганические материалы ( $10 < Z_{eff} < 20$ );
- тяжелые металлы ( $Z_{eff} > 20$ );

где обычно используется цветное представление, позволяющее получить цветное изображение объекта досмотра [2]

Большинство современных досмотровых систем проникающую способность излучения около 300мм по стали, что позволяет обнаруживать предметы, которые могут быть скрыты массивными грузами.

Кроме того, современные досмотровые комплексы оснащены программным обеспечением, которое использует методы кластеризации на основе

искусственного интеллекта для автоматического детектирования нежелательных объектов, что значительно ускоряет процесс проверки.

Существуют как стационарные, так и мобильные порталы, которые могут быть использованы в различных условиях, в зависимости от потребностей. Кроме того, различные источники рентгеновского излучения, такие как рентгеновские трубы или бетатроны, могут использоваться в различных досмотровых системах в зависимости от назначения.

Однако, следует помнить, что каждая досмотровая система имеет свои особенности, и выбор конкретной системы должен основываться на многих факторах, таких как размер и форма предметов, наличие препятствий, преобладающий тип материала и многое другое. Поэтому, перед выбором досмотровой системы, следует провести тщательный анализ и учитывать все эти факторы, чтобы достичь наилучших результатов.

ниже приведены технические характеристики досмотровой радиометрической системы «ШтокТрРН» [1]

Энергия тормозного излучения	6 и 9 МэВ
Проникающая способность по стали	270 мм
Обнаружение стальной проволочки на воздухе	1,0 мм
Обнаружение стальной проволочки за 100 мм стали	3,0 мм
Обнаружение стальной проволочки за 250 мм стали	8,0 мм
Производительность досмотра	25 транспортных средств в час
Габаритные размеры проема портала для проведения досмотра (Ш×В)	4,2×4,7 м
Макс. габариты досматриваемого транспортного средства (Д ×Ш×В)	20×3×4,5 м
Макс. масса транспортного средства	55 тонн

Таблица 1.1 технические характеристики досмотровой радиометрической системы «ШтокТрРН»[1]

# **1. Обзор литературы**

## **1.1. Современный уровень техники**

Современные досмотровые системы имеют множество преимуществ, которые базируются на современных технологиях и инновационных решениях. Программное обеспечение, используемое в современных досмотровых комплексах, позволяет осуществлять досмотр грузов и пассажиров более эффективно и быстро. Для этого используются различные методы, например, такие как кластеризация на основе искусственного интеллекта и другие подходы, которые позволяют детектировать нежелательные объекты автоматически.

Технологии, применяемые в современных досмотровых системах, также включают в себя детектирующие системы, позволяющие с высокой точностью определять интенсивность проникающего гамма-излучения, проходящего сквозь объекты. Это особенно важно для обеспечения безопасности на территории страны, поскольку такие системы могут обнаруживать потенциально опасные предметы и материалы

Кроме того, важную роль в проведении досмотровых процедур играют генерирующие источники ионизирующего излучения. Например, в Институте неразрушающего контроля Томского политехнического университета (ИНК ТПУ), совместно с ООО «Фотон», были созданы следующие источники ионизирующего излучения [3]: ускорители заряженных частиц – бетатроны МИБ – 7,5Д и РХВ – 9Д (Шток – Т), используемые для досмотра грузовых автомобилей. Благодаря данным технологиям досмотровый процесс становится более качественным, безопасным и точным, что уменьшает риски террористических актов и других преступлений.

В той же работе [3] отмечается, что в досмотровых системах для регистрации прошедшего через объект контроля излучения чаще всего применяются три основных типа детекторов:

- люминесцентные экраны совместно с ПЗС-камерой;

- люминесцентные экраны совместно с фотодиодной матрицей;
- сцинтилляционные кристаллы в комплексе с фотодиодной линейкой.

С развитием искусственного интеллекта алгоритмы для обработки получаемого изображения выходят на более высокий уровень, что позволяет с большей скоростью проверять объект на соответствие тем или иным требованиям. Развитие технологий в области алгоритмов и машинного обучения позволяет решать сложные задачи и повышать производительность в различных сферах, включая досмотровые комплексы. Исходя из напряженной мировой обстановки таможенные органы становятся основными заказчиками на такие комплексы. Именно поэтому данная работа уделяет особое внимание изучению, выбору и проверки новых алгоритмов автоматической обработки и кластеризации изображений, которые помогут упростить процедуру проверки грузовых транспортных средств.

## **1.2. Физические принципы работы инспекционно-досмотровых комплексов.**

Одним из основных требований к комплексам неразрушающего рентгеновского контроля является минимизация дозы, получаемой водителем и операторами комплекса. В установке, используемой в данной работе, в качестве источника рентгеновского излучения применяется бетатрон МИБ-9. Мощность дозы излучения данного бетатрона на расстоянии 1м от мишени составляет 20 Р/мин. Из-за низкой дозовой нагрузки, получаемой при сканировании в инспекционно-досмотровом комплексе, рентгеноскопическое изображение имеет заметный уровень шумов. Этот шум может повлиять на четкость и точность анализа изображения, что затрудняет извлечение значимой информации из

рентгеноскопических данных. Для того чтобы уменьшить шумовые помехи изображения используются различные алгоритмы кластеризации.

Так как в результате работы комплекса получается цифровое изображение, к данному изображению применимы алгоритмы кластеризации цифровых изображений. Тут следует учитывать, что при проведении рентгеноскопического контроля выходными данными является информация об интенсивности излучения прошедшего через объект контроля. В дальнейшем эти данные будут представлены в виде изображения. Ослабление рентгеновского излучения подчиняется следующему закону:  $I = I_0 e^{-\mu x}$ . Где  $I$  интенсивность излучения после прохождения вещества;  $I_0$ - интенсивность без преграды;  $\mu$ - линейный коэффициент ослабления;  $x$  - толщина вещества.

### 1.3. Известные методы кластеризации

Для кластеризации рентгеноскопического изображения используются алгоритмы из статистических методов обработки данных один из которых называется метод k-средних, рассмотрим данный метод подробнее. Метод кластеризации k-средних один из наиболее популярных методов кластеризации в области обработки рентгеноскопического изображения. Данный алгоритм выбирается из-за его простоты и быстродействия. Алгоритм стремится найти оптимальные центры кластеров и присвоить каждую точку данных к ближайшему кластеру на основе их признаков и схожести. Происходит это за счет минимизации суммарного среднеквадратичного отклонения точек кластера от центров этих кластеров  $V = \sum_{i=1}^k \sum_{x \in S_i} (x - \mu_i)^2$  [ 6 ]. Данный алгоритм заключается в разбиении множества данных на заранее определенное число кластеров k. Основная идея заключается в том, что на каждой итерации

перевычисляется центр масс для каждого кластера, полученного на предыдущем шаге, затем векторы разбиваются на кластеры вновь в соответствии с тем, какой из новых центров оказался ближе по выбранной метрике.

Алгоритм завершается, когда на очередной итерации не происходит изменения расстояния между точками кластера и центром данного кластера. Это происходит за конечное число итераций, так как количество возможных разбиений конечного множества конечно, а на каждом шаге суммарное квадратичное отклонение  $V$  уменьшается, поэтому зацикливание невозможно. Данный алгоритм один из наиболее простых и распространенных, но очень чувствителен к шуму. Также для работы данного алгоритма необходимы входные данные в виде точного количества кластеров, что является недостатком данного алгоритма.

Еще один популярный алгоритм кластеризации - Сдвиг среднего (Mean Shift) [7] – алгоритм, стремящийся найти области повышенной плотности в имеющемся множестве данных. На каждой итерации алгоритма для каждой области находится ее центр масс (центроид), являющийся средним значений принадлежащих ей точек. Затем эти центры фильтруются на этапе постобработки, устранивая слишком близкие точки, и формируется итоговый набор центроид.

Алгоритм используемый в [11] называемый Leader является наиболее подходящим алгоритмом для решения необходимых задач в сфере инспекционно досмотровых комплексов. Данный алгоритм дает достоверные результаты, основным недостатком данного алгоритма является пересегментация изображения.

## **1.4 Алгоритмы различных методов кластеризации.**

### **1.4.1 Алгоритм метода k- средних**

В данном алгоритме разбиение объектов на кластеры сводится к определению центроидов этих кластеров. Число кластеров  $k$  задается исследователем заранее.

Рассмотрим первоначальный набор  $k$  средних (центроидов)  $\mu_1, \dots, \mu_k$  в кластерах  $s_1, s_2, \dots, s_k$ . На первом этапе центроиды кластеров выбираются случайно.

Относим каждый пиксель к тем кластерам, чье среднее (центроид) к ним ближе всего. Каждое пиксель принадлежит только к одному кластеру, даже если его можно отнести к двум и более кластерам.

Затем центроид каждого  $i$ -го кластера рассчитывается по следующему правилу:

$$\mu_i = \frac{1}{s_i} \sum_{x^{(j)} \in s_i} x^{(j)}$$

Таким образом, алгоритм  $k$ -средних заключается в перевычислении на каждом шаге центроида для каждого кластера, полученного на предыдущем шаге.

Алгоритм останавливается, когда значения  $\mu_i$  не меняются:  $\mu_i^{\text{шаг } t} = \mu_i^{\text{шаг } t+1}$

Важно, что неправильный выбор первоначального числа кластеров  $k$  может привести к некорректным результатам. Именно поэтому при использовании метода  $k$ -средних важно сначала провести проверку подходящего числа

кластеров для данного набора данных.

K-means clustering on the digits dataset (PCA-reduced data)  
Centroids are marked with white cross

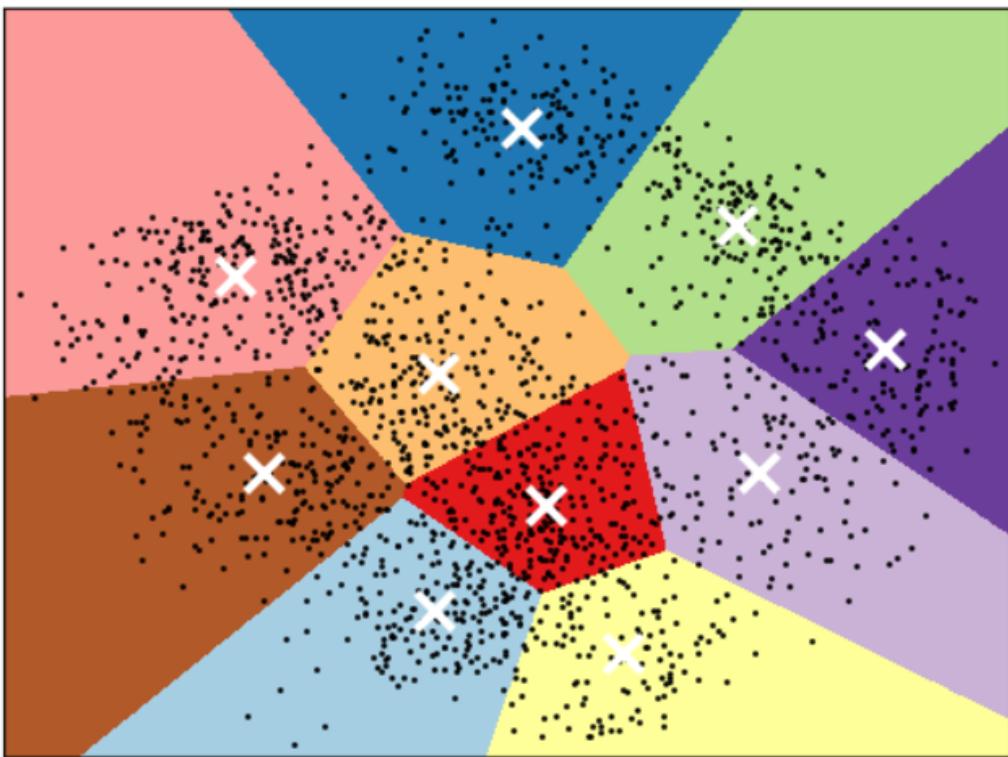


Рис 1.1. Пример кластеризации методом k - средних [8]

#### 1.4.2 Алгоритм кластеризации с-средних

Алгоритм кластеризации с-средних (c-mean) является вариантом алгоритма k-средних, в котором для каждого элемента из рассматриваемого множества рассчитывается степень его принадлежности каждому из кластеров. Основные шаги алгоритма с-средних:

1. Инициализация: Задается начальное количество кластеров  $k$  и случайно выбираются  $k$  точек в пространстве данных в качестве начальных центров кластеров.
2. Вычисление весовых коэффициентов: Для каждой точки данных вычисляются весовые коэффициенты, которые определяют степень принадлежности точки к каждому кластеру. Эти коэффициенты могут

быть вычислены, например, с использованием функции принадлежности, такой как функция Гаусса или функция тяжести.

3. Пересчет центров кластеров: Для каждого кластера вычисляется новый центр, учитывая весовые коэффициенты точек, принадлежащих к данному кластеру. Центр кластера можно вычислить как взвешенное среднее координат точек.
4. Пересчет весовых коэффициентов: Для каждой точки данных пересчитываются весовые коэффициенты на основе расстояний до новых центров кластеров. Часто используется обновление на основе расстояний, где точки, ближайшие к центру кластера, получают больший вес.
5. Проверка критерия сходимости: Проверяется условие сходимости, например, изменение центров кластеров меньше заданного порогового значения или достижение максимального числа итераций. Если условие не выполнено, переходим к шагу 2.
6. Определение кластеров: После завершения алгоритма каждая точка данных присваивается к кластеру с наибольшим весовым коэффициентом.

Алгоритм повторяет шаги 2-6 до тех пор, пока критерий сходимости не будет выполнен. Результатом работы алгоритма является набор кластеров, включающих центры и соответствующие точки данных.

#### **1.4.3 Алгоритм “лидера”**

Для надежного разбиения необходимо принять, что каждая ступень испытуемого материала рассматривается как отдельный сегмент изображения. Кластеры, соответствующие 2,5, 5,0, 7,5, 10,0 и 15,0 см – Шаги толщины материала. Показаны на точечной диаграмме с осями, заданными прозрачностью обозначенной (T) при высокой и низкой энергии (рис. 1.2.).

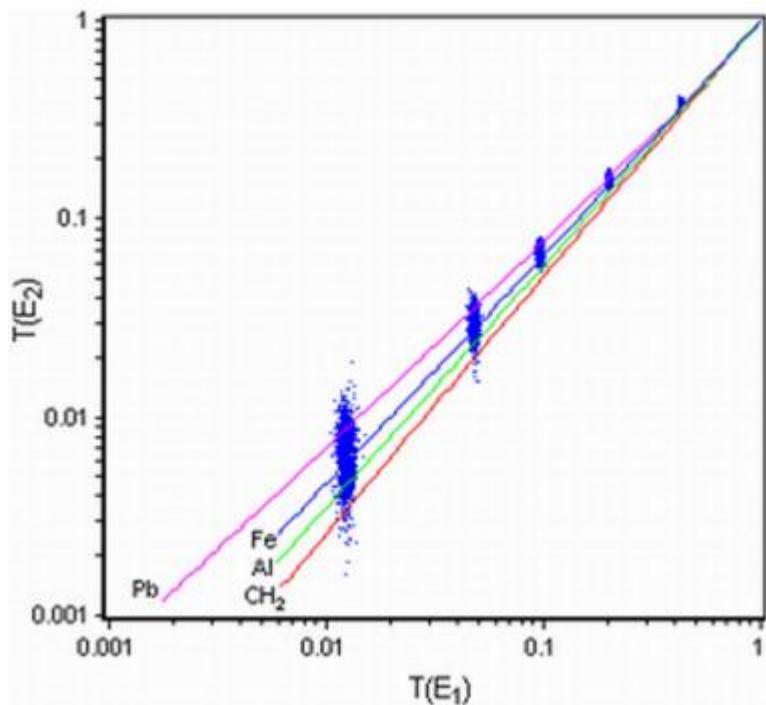


Рис 1.2. Диаграмма ослабления двухэнергетических радиоскопических данных: ось абсцисс: (7,5 МэВ), ось ординат: (4,5 МэВ). Приведены данные для материалов толщиной 2,5, 5,0, 7,5, 10,0 и 15,0 см.

Они представляют собой эллипсы, размеры которых являются функцией распространения измеренной прозрачности. Ребро эллипса образует границу решения для кластера. Если точка данных (пиксель) попадает в границу решения, пикселью присваивается метка соответствующего кластера.

Естественно предположить, что область, заключенная внутри эллипса, должна включать не менее 3 стандартных отклонений по обеим осям. В таком случае при распределении яркости пикселей по нормальному закону не менее 99% пикселей, принадлежащих шагу по толщине, падают в единый кластер.

Это приводит к перекрытию кластеров с одинаковой прозрачностью и даже их слиянию вместе. Этот эффект виден на рис. 1.2. Применение методов

статистической классификации не позволяет выделить кластеры из разных материалов с одинаковой массой. [11]

Алгоритм состоит из следующих шагов:

Шаг 1: Назначьте первый элемент данных,  $P_1$ , кластеру  $C_1$ . Этот набор данных станет лидером кластера  $C_1$ .

Шаг 2: Перейдите к следующему элементу данных, например,  $P_2$ , и вычислите его расстояние от лидера  $P_1$ . Если расстояние между  $P_2$  и лидером  $P_1$  меньше определенного пользователем порогового значения ( $t$ ), то точка данных  $P_2$  присваивается этому кластеру (кластер  $C_1$ ). Если расстояние между лидером  $P_1$  и элементом данных  $P_2$  больше, чем установленное пользователем пороговое значение  $t$ , то создается новый кластер  $C_2$  и  $P_2$  назначается лидером этого нового кластера.

Шаг 3: Для всех остальных элементов данных вычисляется расстояние между точкой данных и лидером кластеров. Если расстояние между точкой данных и лидером любого кластера меньше, чем заданный пользователем пороговый уровень, точка данных присваивается этому кластеру. Однако, если расстояние между точкой данных и лидером любого кластера больше, чем установленное пользователем пороговое значение, создается новый кластер, и этот конкретный элемент данных присваивается этому кластеру и считается его лидером.

Шаг 4: Повторяйте шаг 3, пока все элементы данных не будут назначены кластерам.

Очевидно, что результат разбиения изображения зависит от порядка обработки пикселей, что является недостатком алгоритма. Еще одним недостатком алгоритма кластеризации Leader является пересегментация. Большие сегменты, соответствующие грузовым объектам, могут быть покрыты множеством маленьких пятен с неправильными метками. Этот эффект проявляется на

выходном изображении, где метки кластеров заменяются соответствующими цветами. Ложные цвета маленьких сегментов приводят к шуму и сильно искажают конечное изображение, так как количество маленьких кластеров превышает количество крупных. В результате получается изображение с высоким уровнем шума.

#### 1.4.4 Алгоритм mean-shift.

1. Инициализация: Задается радиус окна сглаживания (bandwidth), который определяет расстояние, на которое будет происходить сдвиг. Данный радиус, в первую очередь, определяется количеством выборок. Каждая точка данных инициализируется как центроид.
2. Оценка плотности: Для каждой точки данных вычисляется плотность на основе окна сглаживания и функции ядра, которая определяет вклад каждой точки в окне сглаживания. Наиболее распространенным ядром является гауссовское ядро.
3. Сдвиг центроидов: Для каждой точки данных вычисляется вектор сдвига, который определяется взвешенной суммой векторов от каждой точки данных до других точек данных, где вес каждого вектора определяется функцией ядра и расстоянием до центроида. Таким образом, центроиды смещаются в сторону более плотных областей данных.
4. Обновление центроидов: Центроиды обновляются путем сдвига на вычисленные векторы сдвига.
5. Повторение шагов 2-4: Шаги оценки плотности, сдвига центроидов и их обновления выполняются итеративно до тех пор, пока центроиды не стабилизируются.

6. Определение кластеров: Кластеры определяются путем присвоения каждой точке данных ближайшего центроида после достижения признака сходимости.

Алгоритм Mean-Shift повторяет шаги 2-5 до достижения сходимости, то есть пока центроиды не перестанут смещаться или смещение станет меньше некоторого порогового значения.

Основной преимуществом алгоритма Mean-Shift является его способность выявлять кластеры произвольной формы и адаптироваться к различной плотности данных. В случае рентгеновских изображений плотность данных связана с интенсивностью пикселей.

## 2. Самостоятельная работа

### 2.1. Источник рентгеноскопического излучения

Многие современные инспекционно досмотровые комплексы используют бетатрон в качестве источника рентгеновского излучения[17]. Бетатрон представляет из себя циклический ускоритель, который ускоряет электроны на круговой орбите при помощи возрастающего магнитного поля и удерживает их на этой орбите этим возрастающим магнитным полем. Получаемый в бетатроне пучок электронов при подаче импульса сброса направляется на мишень, при взаимодействии электронов с которой генерируется жесткое рентгеновское излучение.

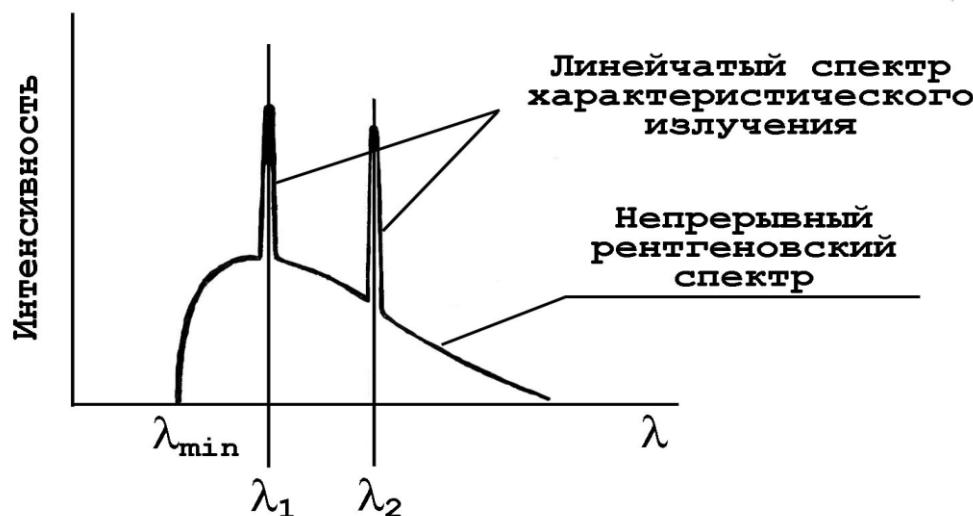


Рис 2.1. Зависимость интенсивности рентгеновского излучения бетатрона от длины волны испускаемого излучения

Основные особенности энергетического спектра излучения мишени бетатрона: Тормозное рентгеновское излучение имеет непрерывный спектр. Это означает, что энергия фотонов может принимать любое значение в определенном диапазоне, а не только дискретные значения, как в случае линейчатого спектра.

Однако у непрерывного спектра есть отчетливая левая граница  $\lambda_{\min}$ . Это означает, что фотоны с длиной волны меньше, чем  $\lambda_{\min}$ , не могут быть созданы в результате генерации тормозного излучения.

Для мишени бетатрона вместе с непрерывным энергетическим спектром тормозного излучения также наблюдается излучение с дискретным набором линий.

## **2.2. Особенности досмотровой установки**

В инспекционно-досмотровом комплексе, разработанном в компании Индиком, используется бетатрон МИБ-9, произведенный Томским политехническим университетом [18]. Этот комплекс в том числе позволяет проводить сканирование объектов в двухэнергетическом режиме 6 и 9 МэВ для реализации метода дуальных энергий. Максимальная энергия тормозного излучения бетатрона составляет 9 МэВ.

Данный комплекс предназначен для инспекционного контроля на предприятиях, аэропортах и других объектах. Мощность дозы бетатрона составляет 20 Р/мин на расстоянии 1 м от мишени. Водитель остается внутри транспортного средства во время контроля, что ускоряет процесс и упрощает конструкцию комплекса. Это позволяет сократить время проведения контроля и повысить эффективность использования ресурсов.

Однако, это конструкционное решение не лишено недостатков. Главный из них - заметный статистический шум сигнала, обусловленный низкой дозовой нагрузкой при обследовании. Это значит, что точность определения атомного номера вещества может быть низкой, что затрудняет работу операторов.

Однако данная проблема может быть решена с помощью различных методов кластеризации, которые позволяют повысить точность измерений.

## 2.3. Получение изображений

Для получения изображения необходимо произвести как минимум две нормировки массива исходных данных об интенсивности излучения, прошедшего через объект контроля. Так как импульсы бетатрона имеют разброс по интенсивности, исходные данные искажены по горизонтальной оси. Кроме того, так как излучение бетатрона характеризуется узкой диаграммой направленности[19], интенсивность излучения, регистрируемого в крайних детекторах детектирующей системы значительно меньше, чем интенсивность излучения, регистрируемого в центральной части. Для исправления первого типа искажения используется коррекция на опорный детектор, который не перекрываются проезжающим объектом. С его помощью устраняется шум сигнала, обусловленный непостоянством дозы излучения от импульса к импульсу. Устранение второго искажения осуществляется методологически: диаграмма направленности излучения бетатрона оценивается при каждом сканировании, после того как объект покидает зону досмотра. Сигналы детекторов, измеренные в моменты, когда объект находится в зоне контроля, корректируются с учетом геометрической формы диаграммы направленности излучения бетатрона.

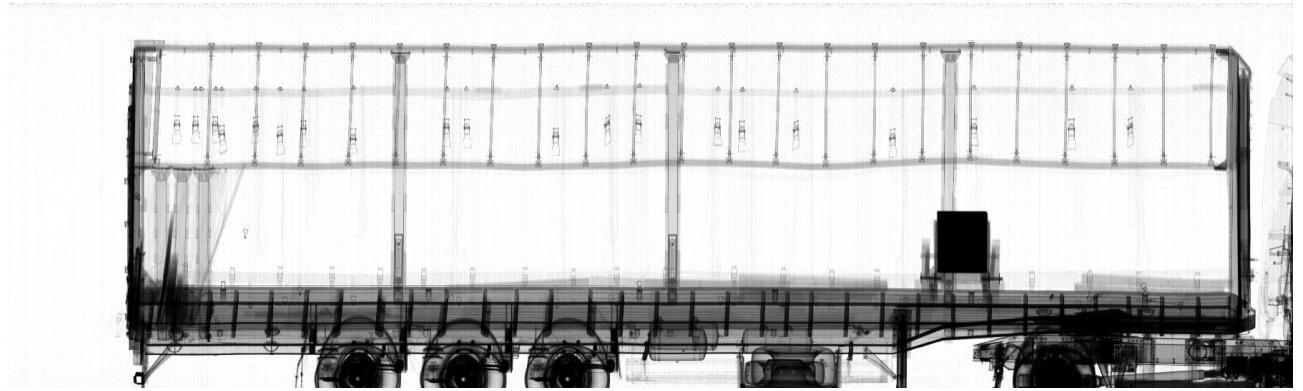


Рис 2.2. Изображение полученное в инспекционно досмотровом комплексе, с линейным представлением интенсивности

На рисунке 2.2 показано рентгеноскопическое изображение, полученное в моноэнергетическом режиме, где видны конструкция прицепа транспортного средства и тестовый поглотитель предельной толщины, выполненный из стали. Изображение, полученное в моноэнергетическом режиме, не позволяет определить эффективный атомный номер вещества, но позволяет оценить массовую толщину фрагментов объекта контроля. В основе интерпретации рентгеновского изображения лежит принцип ослабления рентгеновского излучения. Каждый пиксель на изображении определяется отношением интенсивности рентгеновского излучения за объектом контроля к интенсивности излучения без объекта контроля. Мы можем выразить это отношение с помощью уравнения:  $\frac{I}{I_0} = e^{(-\mu x)}$ , где  $I$  - интенсивность после прохождения через вещество,  $I_0$  - начальная интенсивность,  $\mu$  - коэффициент ослабления, а  $x$  - толщина вещества.

Преобразуя получаем:

$$\ln(I / I_0) = -\mu x.$$

Таким образом, значение ( $\ln(I / I_0)$ ) в каждом пикселе изображения будет пропорционально толщине материала, через который прошло рентгеновское излучение в этой точке.

Наибольшее распространение получил массовый коэффициент ослабления равный  $\frac{\mu}{\rho}$ , где  $\mu$  - линейный коэффициент ослабления,  $\rho$  - плотность вещества.

С учетом массового коэффициента ослабления закон ослабления для гамма-излучения принимает вид:

$$\frac{I}{I_0} = e^{-(\mu/\rho)\rho t}$$

отсюда:

$$\ln(I / I_0) = -(\mu/\rho)\rho t$$

где  $t$  - толщина вещества, через которое прошло рентгеновское излучение.

Множитель  $\rho t$  называется массовая толщина поглотителя. Данный параметр выражается в граммах на квадратный сантиметр ( $\text{г}/\text{см}^2$ ). Массовая толщина может быть использована для оценки плотности или концентрации вещества в образце или объекте на рентгеновском изображении.

Преобразование интенсивности излучения к массовой толщине, позволяет получить дополнительное, более физичное представление объекта досмотра.

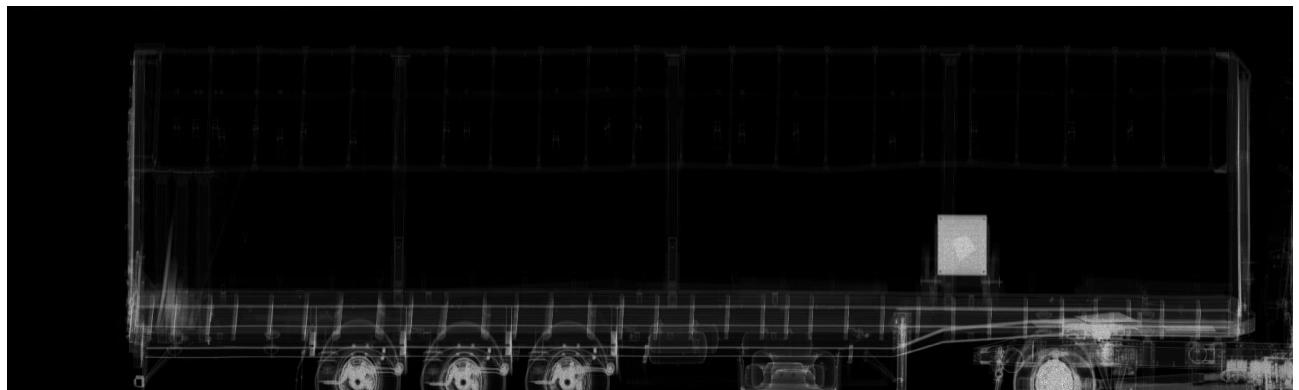


Рис. 2.3. Изображение полученное в представления массовой толщины объектов  
На изображении 2.3. яркость каждого пикселя приведена к диапазону массовых толщин от 0 до  $255 \text{ г}/\text{см}^2$ . Также на изображении отчетливо виден тест-объект, находящийся за стальным препятствием и предназначенный для проведения теста на проникающую способность.

## **2.4. Применение различных методов кластеризации к полученным изображениям**

Кластеризация методом к-средних

В ходе сравнения изображений, представленных на рисунках 2.4-2.12, можно заметить, что наилучшие результаты кластеризации достигаются при разделении на 10 кластеров. При этом стоит отметить, что данные результаты основаны на рассмотрении как логарифмического, так и не логарифмического изображений.

В представлении массовой толщины можно наблюдать объект, который находится за стальной преградой. Это означает, что алгоритм кластеризации смог правильно выделить этот объект в отдельный кластер, даже на пределе проникающей способности. В случае линейного представления, можно заметить, что оптимальный уровень шума достигается при разделении изображения на 10 кластеров, и различные предметы, находящиеся на пути прохождения излучения, обозначаются разными цветами.

Таким образом, в данном исследовании было выявлено, что оптимальное количество кластеров для данного набора данных составляет 10. Кластеризация на основе алгоритма k-mean позволила выделить различные объекты и структуры на изображениях, а использование логарифмического представления улучшило видимость объектов с разной плотностью материалов.

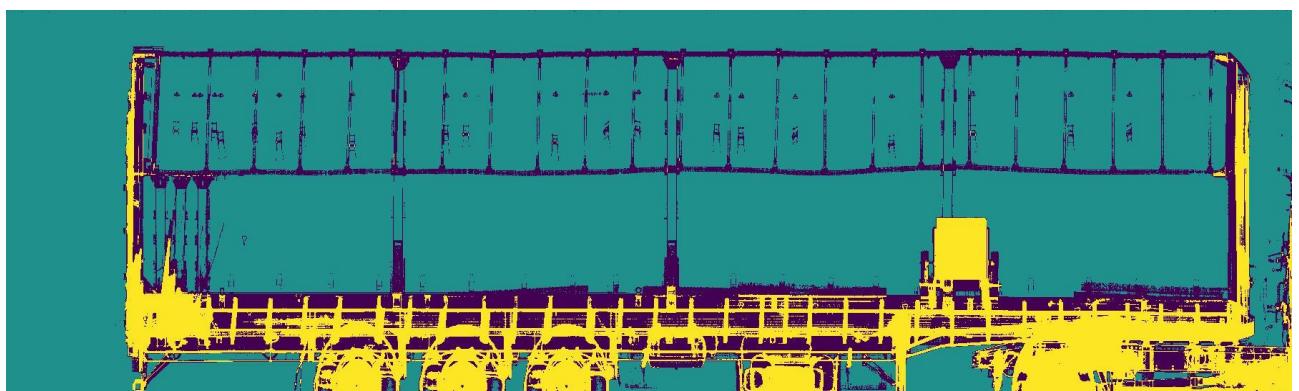


Рис 2.4. кластеризация методом к-средних линейного представления объекта при разбиении изображения на 3 кластера

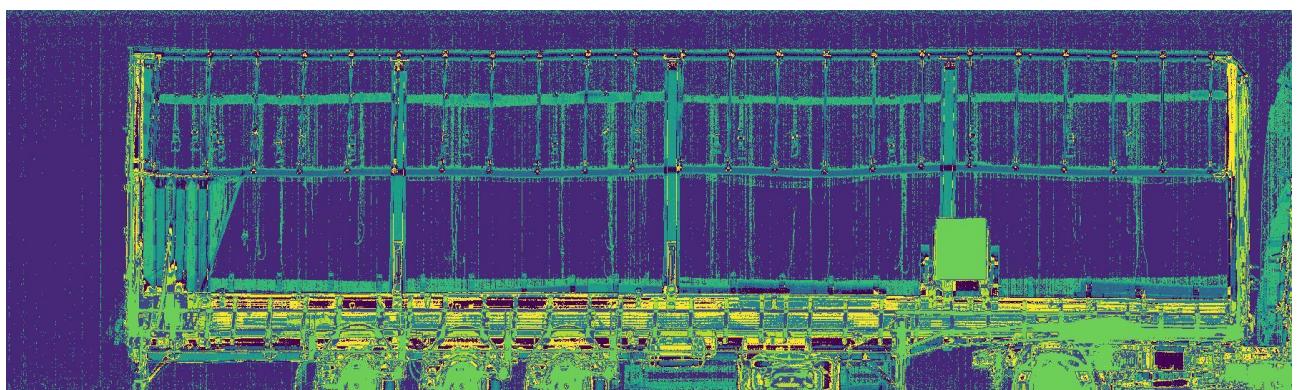


Рис 2.5. кластеризация методом к-средних линейного представления объекта при разбиении изображения на 10 кластеров

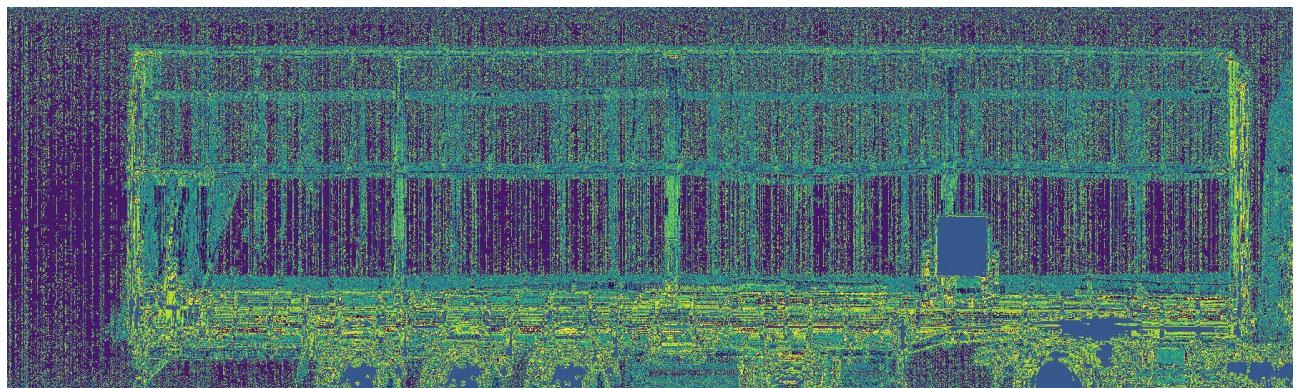


Рис. 2.6. кластеризация методом к-средних линейного представления объекта при разбиении изображения на 30 кластеров

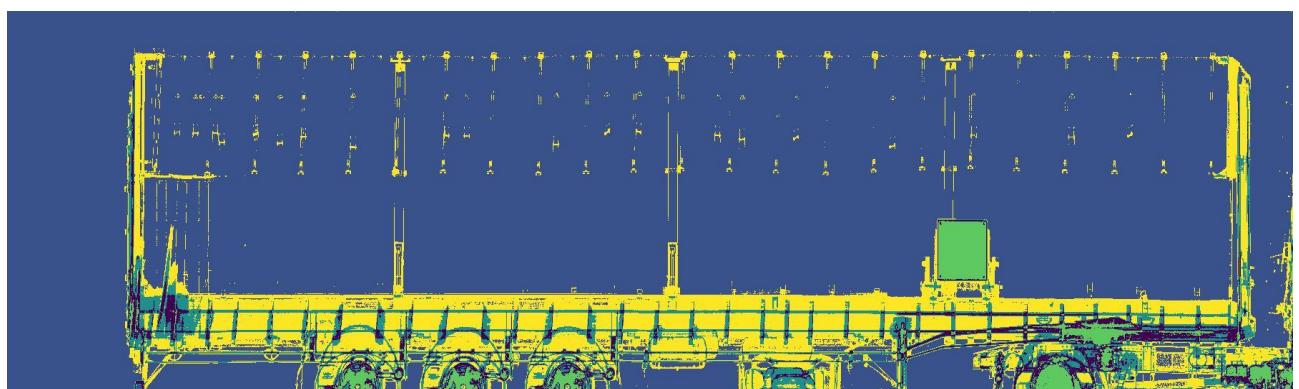


Рис 2.7 кластеризация методом к-средних представления массовых толщин при разбиении изображения на 5 кластеров

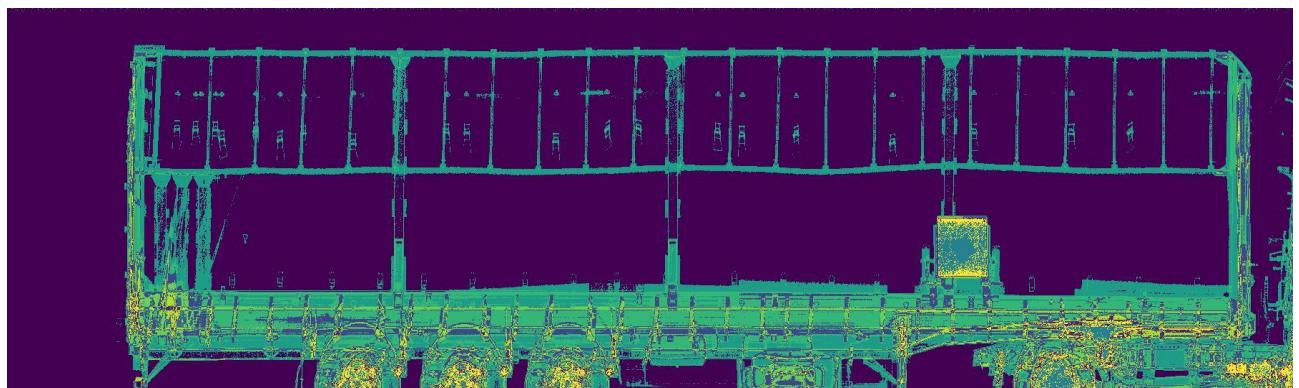


Рис 2.8. кластеризация методом к-средних представления массовых толщин при разбиении изображения на 10 кластеров

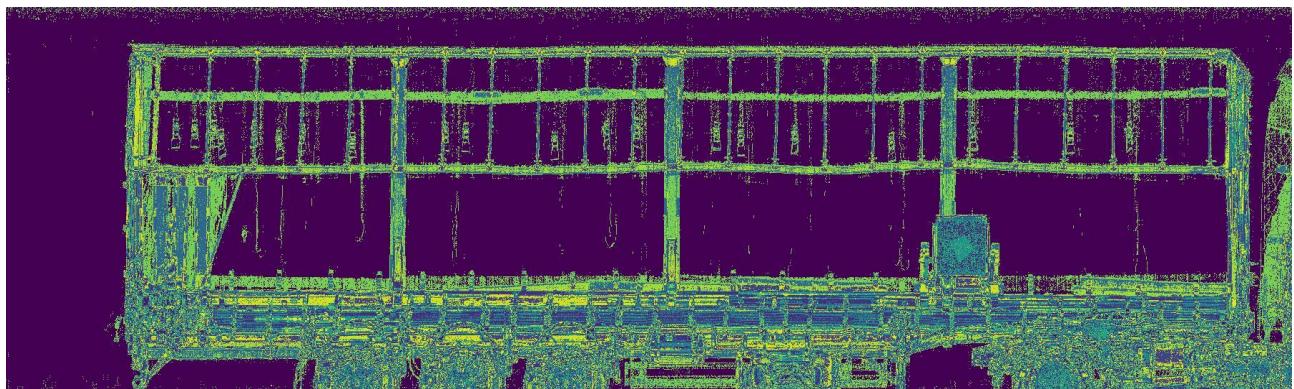


Рис 2.9. кластеризация методом к-средних представления массовых толщин при разбиении изображения на 30 кластеров

### Кластеризация методом с-средних

Применение метода с-средних для кластеризации рентгеноскопического изображения показанное на рисунках 2.13-2.19. Из данных изображений видно, что наилучших результатов кластеризации требуется разделить изображения на разное количество кластеров. В частности, для линейного представления ОПТИМАЛЬНОЕ количество кластеров составляет 10, в то время как для представления массовой толщины ОПТИМАЛЬНОЕ количество кластеров составляет 30. Это различие объясняется особенностями алгоритма кластеризации с-среднего.

На представлении массовых толщин можно заметить объект, который находится за стальной преградой. Это указывает на то, что алгоритм кластеризации смог правильно выделить этот объект в отдельный кластер, несмотря на наличие преграды. Представление массовых толщин может обеспечить лучшую видимость объектов с разной плотностью материалов.

В случае линейного представления можно заметить, что оптимальный уровень шума достигается, и различные предметы, находящиеся на пути прохождения излучения, обозначаются разными цветами. Это говорит о том, что алгоритм кластеризации смог правильно разделить объекты на изображении в соответствии с их характеристиками и отнести их к соответствующим кластерам.

Таким образом, в данном исследовании было выявлено, что оптимальное количество кластеров для не логарифмированного изображения составляет 10, а для логарифмированного изображения - 30. Кластеризация с использованием алгоритма с-среднего позволила успешно выделить различные объекты и структуры на изображениях.

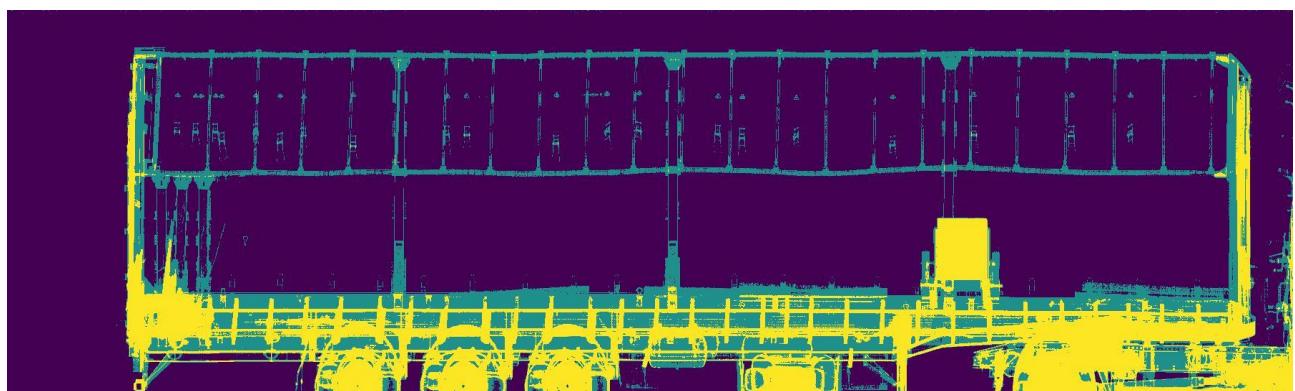


Рис 2.10. кластеризация методом с-средних линейного представления при разбиении изображения на 3 кластера

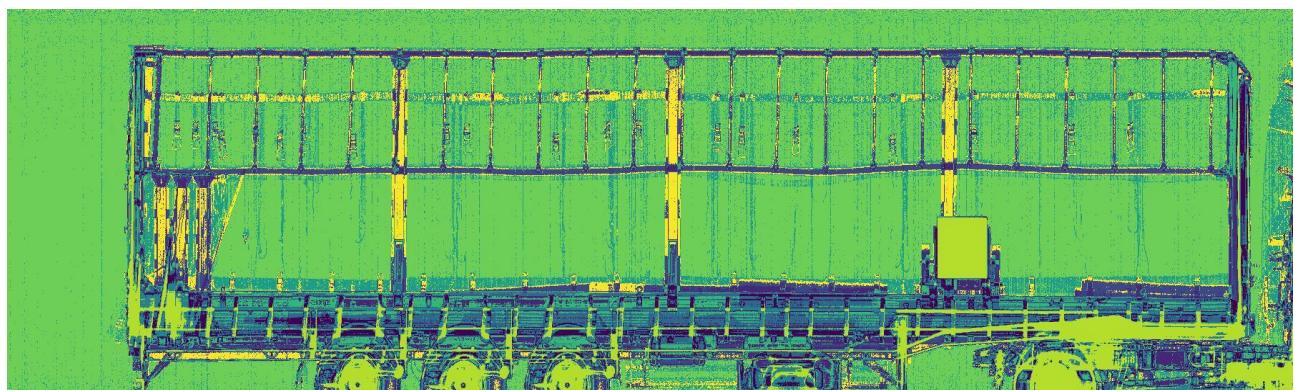


Рис 2.11 кластеризация методом с-средних линейного представления при разбиении изображения на 10 кластеров

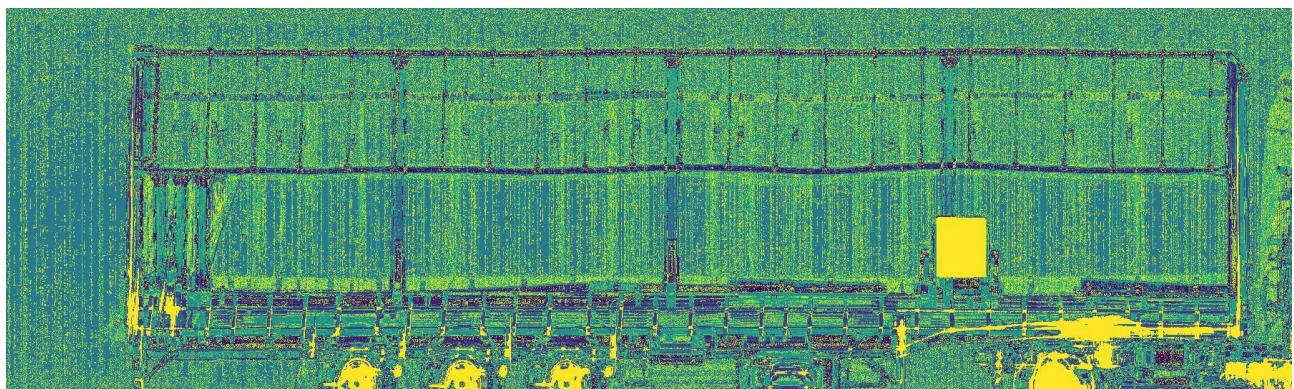
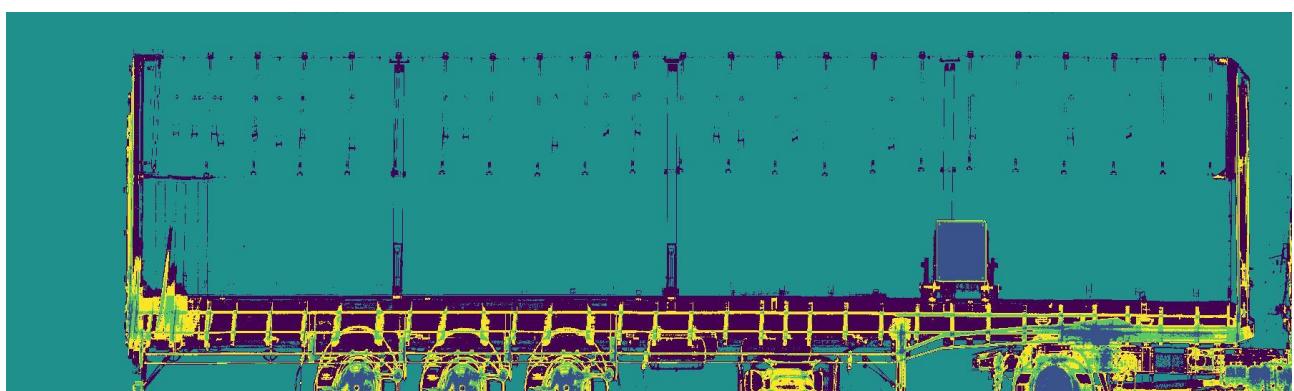
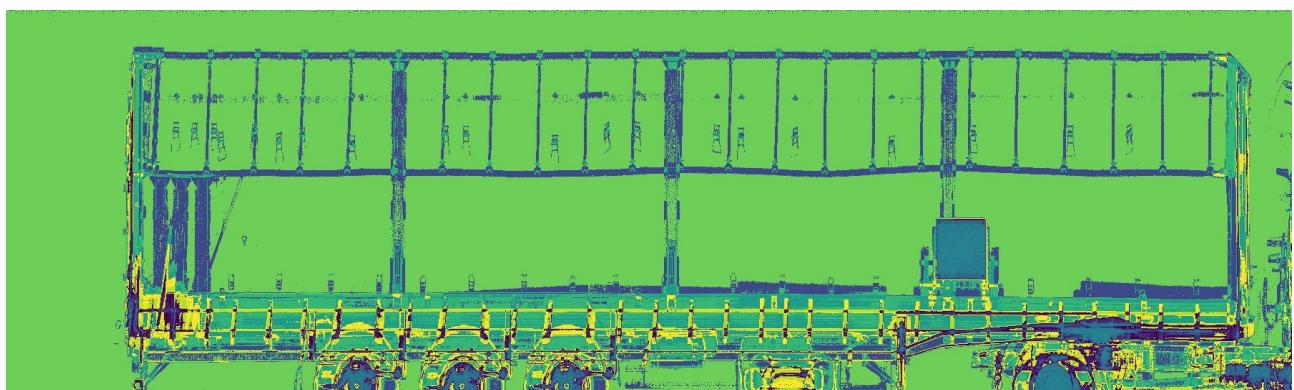


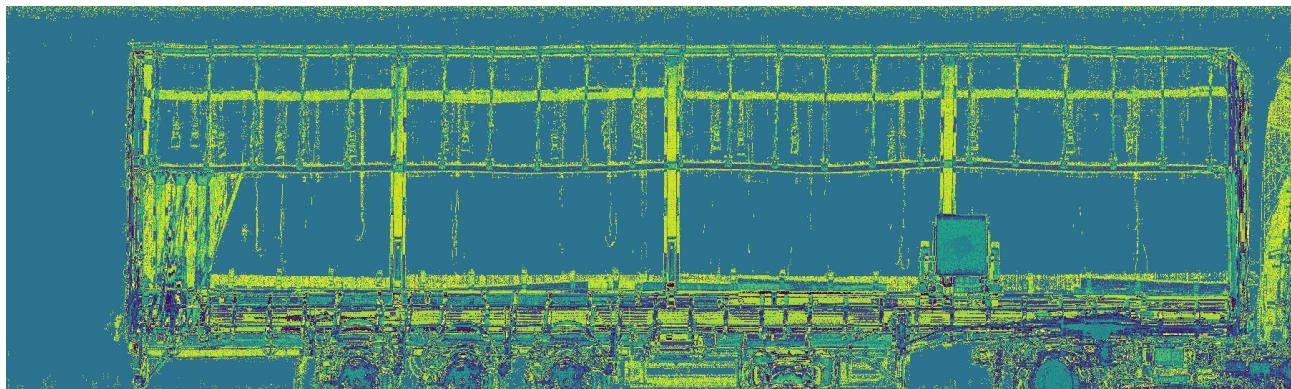
Рис 2.12 кластеризация методом с-средних линейного представления при разбиении изображения на 30 кластеров



2.13. кластеризация методом с-средних представления массовых толщин при разбиении изображения на 5 кластеров



2.14. кластеризация методом с-средних представления массовых толщин при разбиении изображения на 10 кластеров



2.15. кластеризация методом с-средних представления массовых толщин при разбиении изображения на 30 кластеров

### Кластеризация методом сдвига

Анализируя изображения 2.20-2.28, можно заметить, что при кластеризации линейного представления оптимально использовать большие значения выборок, превышающие 1000. Это означает, что для достижения наилучших результатов кластеризации требуется использовать достаточно большое количество точек данных. Большая выборка позволяет учесть большее количество деталей и особенностей изображения, что приводит к более точной кластеризации. В случае представления массовых толщин наблюдается противоположная тенденция. Здесь оптимальным является использование меньшей выборки данных. Меньшая выборка позволяет более отчетливо увидеть объект, который находится за препятствием, так как логарифмирование улучшает видимость объектов с различной плотностью материалов. При использовании меньшей выборки можно более точно выделить этот объект и сосредоточиться на его кластеризации. Таким образом, в данном исследовании было выявлено, что оптимальные значения выборок для кластеризации зависят от типа изображения. Для не логарифмированного изображения требуется

использовать большие значения выборок, превышающие 1000, в то время как для логарифмированного изображения оптимально использовать меньшую выборку данных. Это связано с особенностями визуализации объектов на изображении.

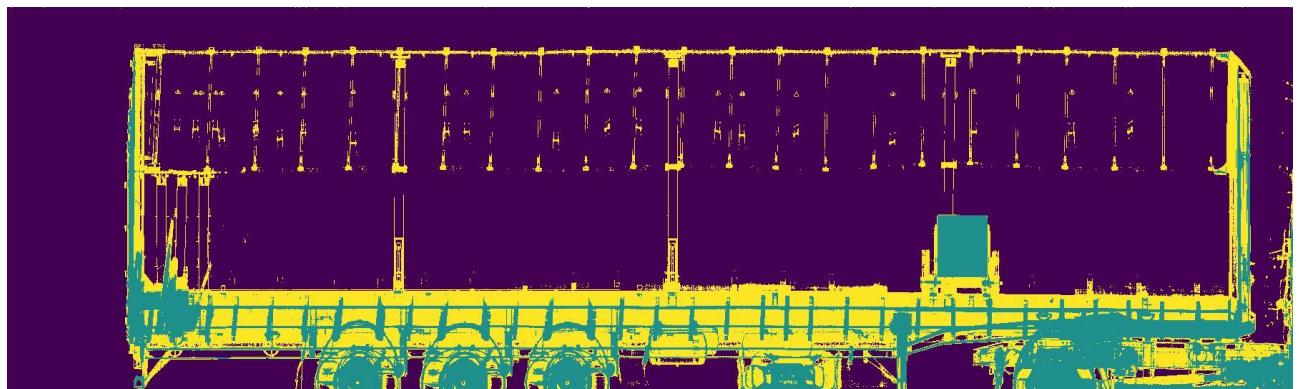


Рис 2.16. кластеризация методом сдвига линейного представления при 1000 выборок

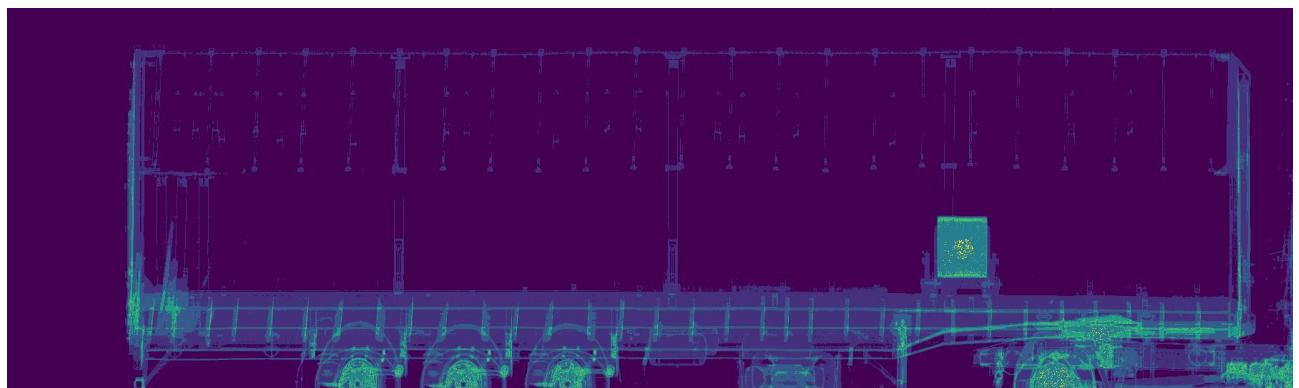


Рис 2.17. кластеризация методом сдвига представления массовых толщин при 100 выборках

### Кластеризация методом лидера

Анализируя полученные результаты (рис 2.18-2.19) что алгоритм кластеризации лидера (Leader clustering) проявляет высокий уровень шума и избыточную сегментацию изображения при различных настройках. Высокий уровень шума указывает на то, что алгоритм кластеризации лидера не справляется с

эффективным удалением шумовых точек и неконтролируемым вариациям в данных. Наличие высокого уровня шума может затруднить последующий анализ и интерпретацию данных. Избыточная сегментация означает, что алгоритм кластеризации лидера разделяет изображение на слишком мелкие и многочисленные сегменты, что может привести к потере смысла и трудностям в интерпретации результата.

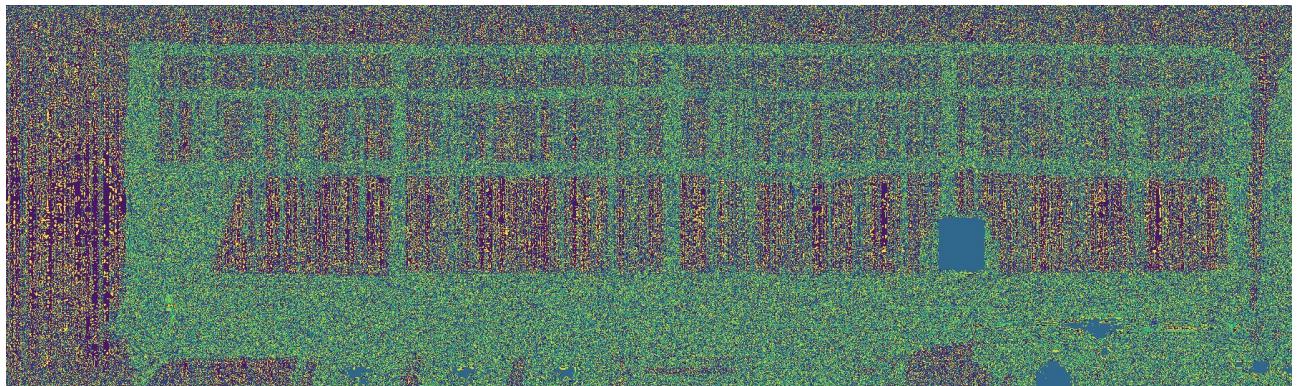


Рис 2.18. кластеризация методом лидера линейного представления при пороге в 10 пикселей

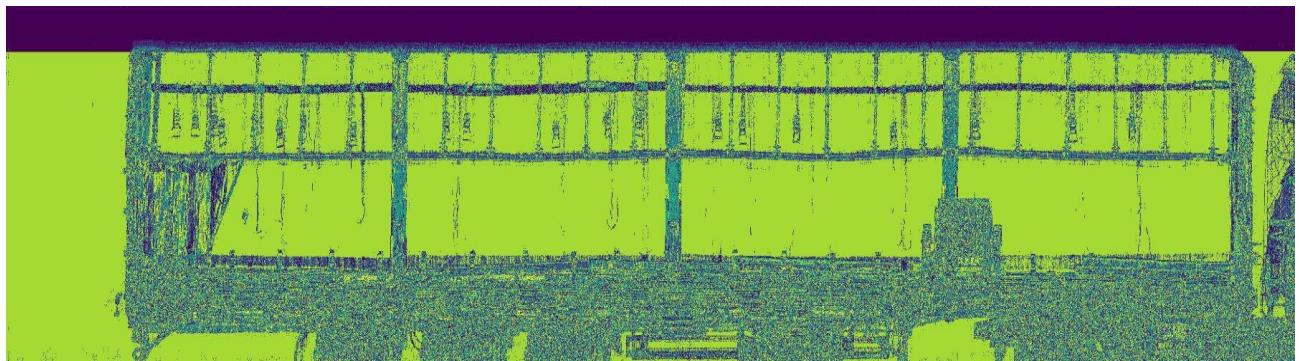


Рис 2.19. кластеризация методом лидера представления массовых толщин при пороге в 10 пикселей

## 2.5. Применение методов кластеризации с целью определения разных веществ.

Были получены результаты работы инспекционно досмотрового комплекса для транспортного средства с грузом в виде фантомов, предназначенных для проведения теста на способность установки идентифицировать группы материалов по эффективному атомному номеру. В каждом измерении на

испытательный стенд были помещены четыре массива наборных пластин, выполненные из одного материала, но образующие различные толщины, приведенные в таблице ниже.

Таблица 2. 1

Данные измерений

Номер измерения	Материал	Толщины материалов(см)
1	листы свинца	1, 2, 4, 6
2	листы полиэтилена	10, 20, 40, 60
3	листы стали	2, 4, 6, 9

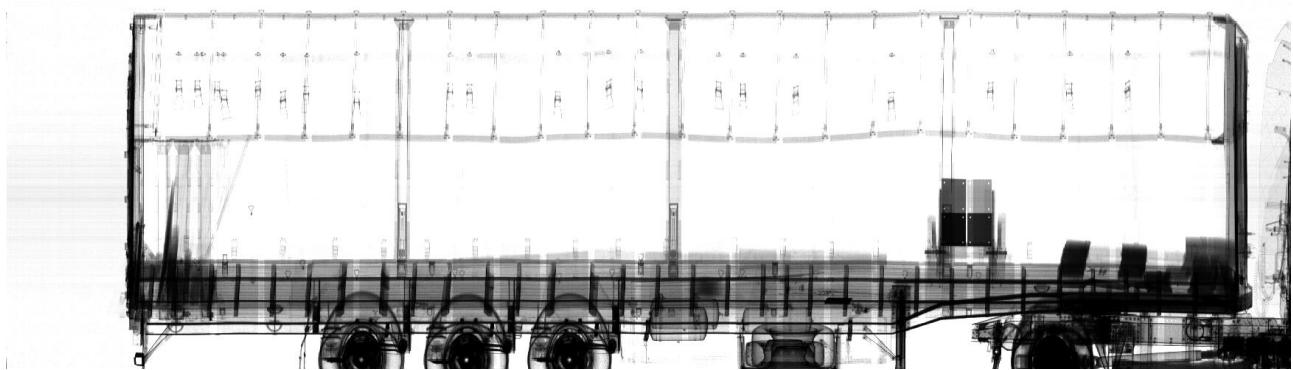


рис 2.20. Линейное представления грузовика с фантомом из свинца

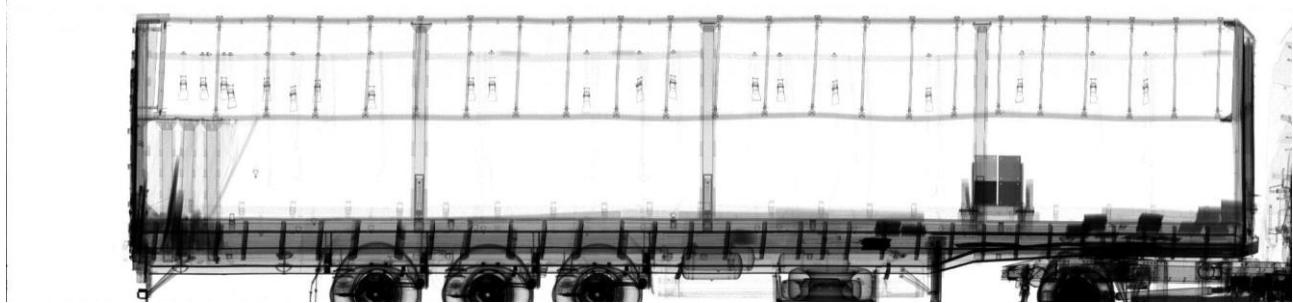


рис 2.21. Линейное представления грузовика с фантомом из полиэтилена

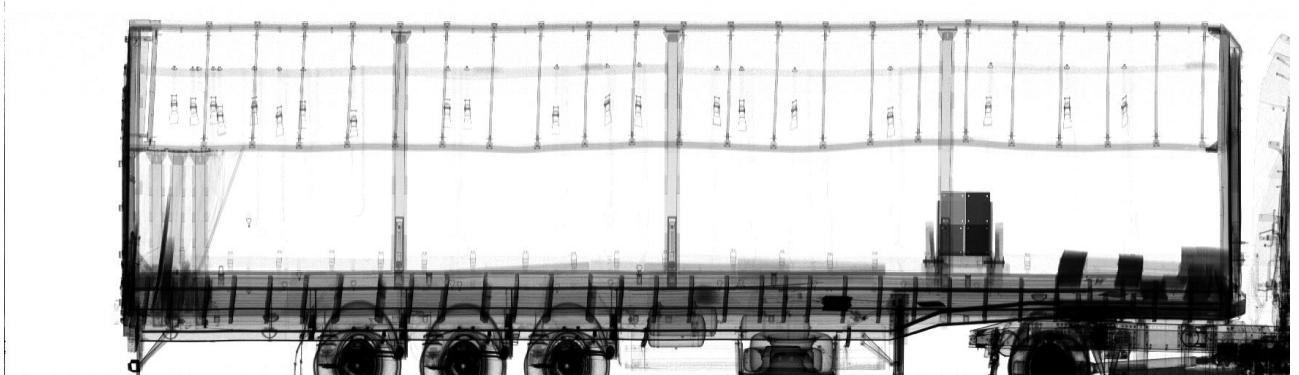


рис 2.22. Линейное представления грузовика с фантомом из стали

В данном случае будет рассматриваться кластеризация только одного фрагмента изображения (рис 2.20-22) в котором данные фантомы находятся. Будут представлена кластеризация методом k и с средних так как исходя из предыдущих результатов именно эти 2 метода показали наименьший уровень шума в выходном изображении, а также лучшее распознавание различных объектов.

Для кластеризации будет рассматриваться данный фрагмент изображения(рис.2.23) после в представлении массовых толщин

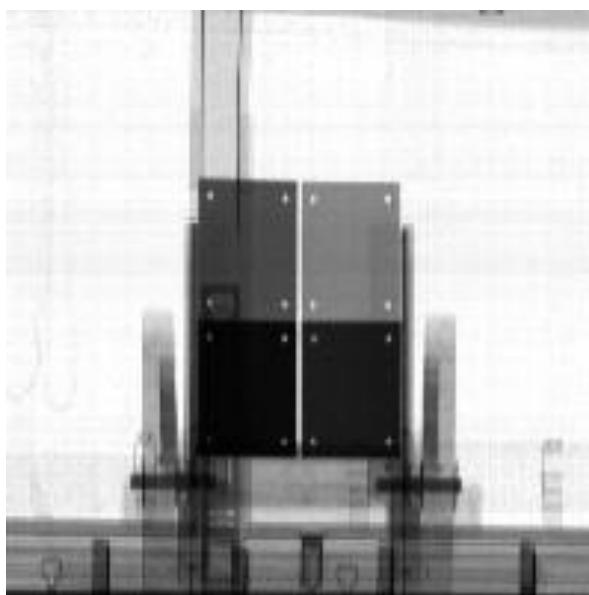


рис 2.23. Фрагмент изображения для кластеризации

Методом k-средних при разбиении на 10 кластеров

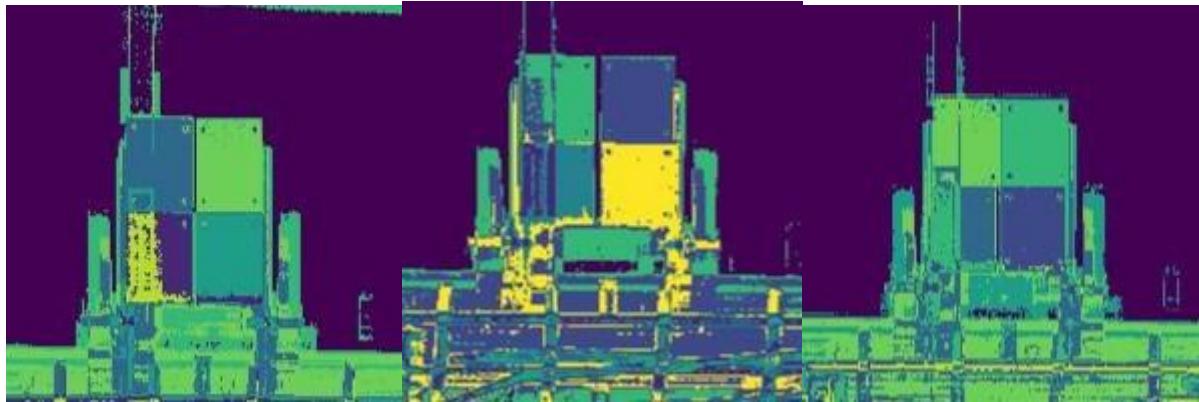


Рис.2.24. Кластеризованное изображение фантомов различных материалов(слева направо(свинец, полиэтилен, сталь) методом к-средних  
Методом с-средних при разбиении на 10 кластеров

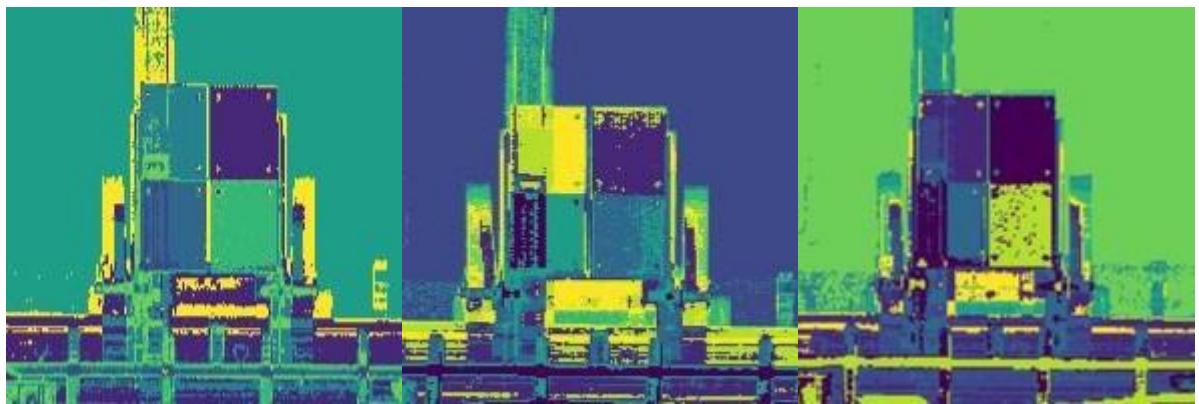


Рис.2.25.. Кластеризованное изображение фантомов различных материалов(слева направо(свинец, полиэтилен, сталь) методом с-средних

Анализируя полученные изображения 2.24-2.25, можно ясно увидеть, что объекты разной толщины находятся в разных кластерах благодаря применению алгоритмов кластеризации. Это означает, что при использовании метода дуальных энергий на таких рентгеновских изображениях можно снизить уровень шума на выходном изображении.

При кластеризации объектов разной толщины в различные кластеры, метод дуальных энергий позволяет более точно определить объекты, состоящие из одного материала. Это может быть особенно полезно для снижения уровня шума на выходном изображении. Целостные объекты, полученные при

кластеризации, будут определяться как объекты с однородными свойствами материала, что позволяет более эффективно удалять шумы и артефакты. Применение метода дуальных энергий в анализе рентгеновских результатов обеспечивает улучшение качества изображения за счет точного определения и классификации объектов на основе их материалов. Это способствует снижению шума и улучшению визуального восприятия результатов.

Таким образом, исходя из полученных изображений, можно сделать вывод, что алгоритм кластеризации позволяет отнести объекты разной толщины к различным кластерам, что в свою очередь приводит к снижению уровня шума на выходном изображении при использовании метода дуальных энергий. Это предоставляет более чистое и четкое изображение, что может быть полезно в рентгеновской диагностике и других областях, где важно точное определение объектов и уменьшение шума.

## **Заключение**

В данной исследовательской работе были рассмотрены устройство и основные принцип работы инспекционно-досмотрового комплекса на базе бетатрона. Были проанализированы основные функциональные возможности подобных систем, а также способы их применения в системах обеспечения безопасности. Основной целью работы было изучение и применение на практике способов цифровой обработки рентгеноскопических изображений, принадлежащих к семейству методов кластеризации. В рамках этого были изучены физические основы получения радиоскопических изображений в инспекционно-досмотровых комплексах излучения. Также было исследовано влияние представления на результат проведения кластеризации.

Все результаты, представленные в работе, получены на действующей установке инспекционно-досмотрового комплекса в рамках проведения испытаний, подтверждающих его основные функциональные параметры.

В ходе исследования также были использованы следующие методы кластеризации: метод k-средних, метод c-средних, mean shift , алгоритм лидера. Алгоритмы k-средних и c-средних дают приемлемые результаты, их основным преимуществом является четкое разбиение изображения на кластеры. Основной недостаток этих двух алгоритмов в необходимости заранее знать количество кластеров на которое будет происходить разбиение, что не всегда удобно.

Алгоритм mean-shift не требует указания изначального количества кластеров, однако он плохо подходит для разбиения находящегося небольших по размеру и близких друг к другу кластеров.

Алгоритм кластеризации лидера, не требует указания количества кластеров, однако, данный алгоритм очень чувствителен к шуму на изображении и разделяет изображение на слишком большое количество мелких кластеров, что дополнительного усиливает шум. Отмечается, что выбор метода кластеризации влияет на результаты анализа, так как каждый метод прикладывается к

различным представлениям рентгенографического изображения. Показано, что применение методов кластеризации к рентгеноскопическому изображению позволяет увеличить качество рентгеноскопического изображения на этапе представления результатов досмотра. В частности выполнение процедуры кластеризации позволяет упростить анализ изображений или использовать ее результаты для проведения идентификации материалов по эффективному атомному номеру с меньшим уровнем цифрового шума.

Итак, данное исследование представляет собой комплексный анализ теоретических и практических аспектов устройства инспекционно-досмотровых комплексов, исследования методов кластеризации рентгеноскопического изображения и их применения в конкретном контексте.

## **Список литературы**

- 1) Усачев Е.Ю. Досмотровые комплексы для контроля грузового автотранспорта, разработанные МИРЭА // Информационные технологии неразрушающего контроля : сборник научных трудов Российской школы конференции с международным участием, Томск, 27-30 октября 2015 г. — Томск : Изд-во ТПУ, 2015. — [С. 73-82].
- 2) Рыжиков В.Д., Ополонин А.Д., Волков В.Г., Лисецкая Е.К., Галкин С.Н., Воронкин 52 Е.Ф. Трехэнергетическая цифровая радиография для разделения веществ с малым эффективным атомным номером // Вісник НТУ «ХПІ» – 2013 – № 34 (1007) – С. 43– 51
- 3) Клименов В.А., Касьянов В.А., Лебедев М.Б., Москалев Ю.А., Темник А.К., Штейн М.М., Чахлов С.В. Современное состояние и перспективы создания конкурентоспособных на мировом рынке систем цифровой радиографии (СЦР)// Контроль. Диагностика – 2011 – Специальный выпуск – С. 25–29
- 4) Санитарные правила и нормативы СанПиН 2.6.1.2523-09 «Нормы радиационной безопасности НРБ-99/2009» (утв. постановлением Главного государственного санитарного врача РФ от 7 июля 2009 г. N 47).
- 5) Клименов В.А., Осипов С.П.. Темник А.К. Идентификация вещества объекта контроля методом дуальных энергий // Дефектоскопия – 2013 – № 11 – С. 40–50
- 6) Котелина Н.О. Кластеризация изображения методом k-средних. // Вестник Сыктывкарского университета- Выпуск 3 (2019)
- 7) Comaniciu D., Meer P. Mean shift: a robust approach toward feature space analysis // IEEE Transactions on Pattern Analysis and Machine Intelligence. — 2002. — May. — Vol. 24, no. 5. — P. 603–619.

- 8) Интернет ресурс: <https://scikit-learn.ru/clustering/>
- 9) Чернявский О. Н. Анализ решений задачи кластеризации изображений лиц в сфере киберкриминалистики // Издательство СПБГУ 2019
- 10) Д. С.Черезов, Н. А. Тюкачев ОБЗОР ОСНОВНЫХ МЕТОДОВ КЛАССИФИКАЦИИ И КЛАСТЕРИЗАЦИИ ДАННЫХ // Воронежский государственный университет 2010
- 11) S. Ogorodnikov and V. Petrunin Processing of interlaced images in 4–10 MeV dual energy customs system for material recognition// Efremov Scientific Research Institute, St.-Petersburg, Russia 2002
- 12) Кудрина Мария Александровна Использование преобразования Хафа для обнаружения прямых линий и окружностей на изображении // Известия Самарского научного центра РАН. 2014. №4-2.
- 13) C. Tomashi and R. Manduchi, in Proceedings of the 1998 IEEE International Conference on Computer Vision (New Delhi, Bombay, India, 1998).
- 14) J. Sijbers et al., Magn. Reson. Imaging 15, 679–688 (1997).
- 15) D. Perion et al., International Patent No. WO 00/43760 (2000).
- 16) Штейн, А .М .Распознавание материалов методом дуальных энергий в досмотровых комплексах с бетатронными источниками излучения,Известия Томского политехнического университета [Известия ТПУ]// 2017
- 17) А.С. Алимов//Практическое применение электронных ускорителей//МОСКОВСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ имени М.В. ЛОМОНОСОВА// 2011
- 18) Чахлов В. Л. Разработка приборов и методов неразрушающего контроля качества в НИИинтроскопии Томского политехнического университета / В. Л. Чахлов, О. А. Сидуленко // Известия Томского политехнического университета [Известия ТПУ]. — 2000. — Т. 303, вып. 1. — [С. 111-132].

19)        В.А.Москалев, Г.И. Сергеев// ИНДУКЦИОННЫЙ УСКОРИТЕЛЬ  
ЭЛЕКТРОНОВ - БЕТАТРОН// Издательство Томского политехнического  
университета 2012