

УДК: 517.958

Идентификация неоднородного вещества методами импульсной мультиэнергетической томографии

В. Г. Назаров^а, И. В. Прохоров^б, И. П. Яровенко^с

Институт прикладной математики ДВО РАН,
Россия, 690041, г. Владивосток, ул. Радио, д. 7

E-mail: ^а naz@iam.dvo.ru, ^б prokhorov@iam.dvo.ru, ^с yarovenko@iam.dvo.ru

Получено 27.08.2024, после доработки — 10.06.2025.

Принято к публикации 07.07.2025.

В статье рассматриваются математические аспекты проблемы идентификации многокомпонентной рассеивающей среды по данным импульсного мультиэнергетического рентгеновского облучения. Задачи рентгеновской диагностики представляют значительный интерес как с теоретической, так и с практической точки зрения, а радиографические методы незаменимы при неразрушающем контроле изделий.

В рамках математической модели на основе нестационарного интегро-дифференциального уравнения переноса излучения сформулированы обратная задача нахождения коэффициента ослабления по излучению, известному на границе области, и задача идентификации вещества по найденным значениям коэффициента ослабления на дискретном наборе энергий облучения среды. Проведена предварительная обработка широкого списка веществ, представляющих интерес в компьютерной томографии, на предмет возможности их идентификации по приближенно заданному коэффициенту ослабления излучения, характеризующему среду. При анализе степени близости веществ в некоторой норме установлено, что множество всех возможных веществ, потенциально содержащихся в среде, распадается на конечное число непересекающихся кластеров. При достаточно малой длительности зондирующего сигнала рассеивающая составляющая выходящего из среды излучения асимптотически мала. Это обстоятельство позволяет свести обратную задачу для уравнения переноса излучения к задаче обращения преобразования Радона от коэффициента ослабления. Методами численного моделирования на специально разработанном цифровом фантоме анализируется возможность однозначной или частичной идентификации вещества при варьировании длительности зондирующего импульса и числа энергетических уровней облучения среды.

Ключевые слова: импульсная томография, нестационарное уравнение переноса излучения, обратные задачи, коэффициент ослабления, идентификация вещества, мультиэнергетическое зондирование

Работа выполнена в рамках государственного задания ИПМ ДВО РАН (соглашение № 075-00459-25-00).

UDC: 517.958

Identification of inhomogeneous matter by pulsed multienergy tomography methods

V. G. Nazarov^a, I. V. Prokhorov^b, I. P. Yarovenko^c

Institute of Applied Mathematics FEB RAS,
7 Radio st., Vladivostok, 690041, Russia

E-mail: ^a naz@iam.dvo.ru, ^b prokhorov@iam.dvo.ru, ^c yarovenko@iam.dvo.ru

Received 27.08.2024, after completion — 10.06.2025.

Accepted for publication 07.07.2025.

The article considers the mathematical aspects of the problem of identifying a multicomponent scattering medium based on pulsed multienergy X-ray irradiation data. X-ray diagnostics problems are of considerable interest from both theoretical and practical points of view, and radiographic methods are indispensable in non-destructive testing of products.

Within the framework of a mathematical model based on a non-stationary integro-differential equation of radiation transfer, the inverse problem of finding the attenuation coefficient for radiation known at the boundary of the region and the problem of identifying a substance based on the found values of the attenuation coefficient on a discrete set of irradiation energies of the medium are formulated.

A preliminary processing of a wide list of substances of interest in computed tomography was carried out to determine the possibility of their identification by an approximately specified radiation attenuation coefficient characterizing the medium. When analyzing the degree of proximity of substances in a certain norm, it was found that the set of all possible substances potentially contained in the medium is divided into a finite number of non-intersecting clusters. For a sufficiently short duration of the probing signal, the scattering component of the radiation leaving the medium is asymptotically small. This circumstance allows us to reduce the inverse problem for the radiation transfer equation to the problem of inverting the Radon transform from the attenuation coefficient. The possibility of unambiguous or partial identification of a substance by varying the duration of the probing pulse and the number of energy levels of irradiation of the medium is analyzed using numerical modeling methods on a specially developed digital phantom.

Keywords: pulse tomography, non-stationary radiation transfer equation, inverse problems, attenuation coefficient, substance identification, multi-energy probing

Citation: *Computer Research and Modeling*, 2025, vol. 17, no. 4, pp. 621–639 (Russian).

The work was carried out within the framework of the state assignment of the IAM FEB RAS (agreement No. 075-00459-25-00).

Введение

Задачи рентгеновского исследования среды, в частности идентификации неизвестного вещества, представляют значительный интерес как с теоретической, так и с практической точки зрения. Радиографические методы являются незаменимыми в случаях, когда необходим неразрушающий контроль изделия или когда непосредственный доступ к объекту исследования затруднен или нежелателен. Несмотря на обилие научных публикаций по данной тематике, исследователи продолжают искать новые подходы к решению задачи идентификации веществ, поскольку каждая конкретная ситуация может требовать специфических методов и алгоритмов. Среди наиболее известных и широко используемых методов выделяется метод дуальных энергий [Alvarez, Macovski, 1976; Brooks, 1977; Ying et al., 2006; Rebuffel, Dinten, 2007; Клименов и др., 2013; Osipov et al., 2015; Осипов и др., 2017; Ahmed et al., 2022; Jumanazarov et al., 2020].

Значительная часть упомянутых работ посвящена вопросу идентификации веществ таможенного контроля. Не менее важным применением этого метода является дифференцировка мягких тканей, которая широко используется в биологической и медицинской визуализации [Yokhana et al., 2022]. Также метод дуальных энергий применяется при исследовании строения минералов сложного химического состава [Komarskiy et al., 2023]. С общетеоретической точки зрения задачи идентификации рассматривались в работе [Bal et al., 2022], в которой получены локальные условия обратимости и устойчивости преобразования мультиэнергетической компьютерной томографии. Другой подход к решению задачи идентификации предложен в работах [Nazarov, 2021; Назаров, 2021], где рассматривается проблема частичной идентификации химического состава неизвестной среды путем многократного облучения рентгеновскими лучами разных энергий. В [Nazarov, 2021; Назаров, 2021] построена специальная индикаторная функция, которая позволяет установить достаточные условия различимости веществ в рамках предлагаемого автором подхода.

В отличие от упомянутых работ в рамках данного исследования предполагается, что заранее известен список веществ, потенциально содержащихся в исследуемой среде. Каждое вещество характеризуется своими значениями коэффициента ослабления для различных значений энергии излучения. Исходные данные в задаче — набор проекционных данных томографического просвечивания среды на дискретном наборе энергий. Суть метода заключается в выборе из списка веществ тех, чьи коэффициенты ослабления на всем наборе энергий имеют минимальное среднеквадратичное отклонение от значений, восстановленных из проекционных данных. В такой постановке критически важным требованием является высокая точность восстановления коэффициентов ослабления исследуемой среды.

Как правило, в рентгеновской томографии используется коллимация источников и детекторов излучения, что позволяет пренебрегать рассеянием в среде и определять коэффициент ослабления из экспоненциального закона ослабления. В данной работе мы рассматриваем задачу идентификации неизвестной среды без предположения о коллимации источника рентгеновского излучения, при этом для подавления рассеянной компоненты измеряемого сигнала используются источники излучения импульсного типа. В рамках модели, основанной на интегро-дифференциальном уравнении переноса излучения, применение предельно малой длительности зондирующего импульса позволяет выделить баллистическую компоненту решения [Прохоров, Яровенко, 2021; Прохоров, Яровенко, 2022], в результате чего задача определения коэффициента ослабления сводится к хорошо изученной проблеме — обращению преобразования Радона [Natterer, 1986; Herman, Natterer, 2013]. В [Nazarov et al., 2023; Yarovenko, Prokhorov, 2024] доказаны оценки для скорости убывания компоненты выходящего излучения, отвечающей однократному и двукратному рассеянию, при уменьшении длительности зондирующего импульса и путем численного

моделирования показано, что использование ультракоротких импульсов в схемах томографического сканирования позволяет существенно улучшить качество восстановления структуры неизвестной среды.

Настоящая статья является развитием и дополнением результатов работы [Nazarov et al., 2023] и структурно состоит из двух основных частей. Первая часть посвящена предварительному анализу поведения коэффициентов ослабления для вещества (их количество — 31), представляющих интерес в рентгенодиагностике и входящих в справочные таблицы Хаббла – Зельтцера [Hubbell, Seltzer, 1995; Berger et al., 2005]. Значения коэффициентов ослабления излучения для каждого вещества на энергиях 50, 60, 80, 100 и 150 кэВ рассматриваются как векторы в пятимерном евклидовом пространстве. Проанализирована близость векторов с точки зрения расстояния между ними. Выявлено, что множество потенциально содержащихся в среде веществ распадается на конечное число непересекающихся подмножеств (кластеров) веществ с близкими коэффициентами ослабления. Во второй части демонстрируется работа предложенного метода идентификации на специально созданном цифровом фантоме. Определяются значения длительности импульса рентгеновского излучения, при которых задача идентификации вещества может быть полностью или частично решена в рамках разработанного подхода.

Математическая модель. Постановка задачи идентификации

Рассмотрим интегро-дифференциальное уравнение, описывающее нестационарный процесс взаимодействия излучения с веществом, вида [Прохоров, Яровенко, 2021; Прохоров, Яровенко, 2022]

$$\left(\frac{1}{c} \frac{\partial}{\partial t} + \omega \cdot \nabla_r + \mu(r, E) \right) I(r, \omega, E, t) = \sigma(r, E) \int_{\Omega} p(r, \omega \cdot \omega', E) I(r, \omega', E, t) d\omega'. \quad (1)$$

Функция $I(r, \omega, E, t)$ интерпретируется как плотность потока частиц в момент времени $t \in [0, T]$, в точке $r \in \mathbb{R}^3$, движущихся со скоростью c в направлении единичного вектора $\omega \in \Omega = \{ \omega \in \mathbb{R}^3 : |\omega| = 1 \}$ с энергией $E \in [\underline{E}, \bar{E}]$. Функции μ и σ имеют смысл коэффициентов ослабления и рассеяния, а через p обозначена индикатриса рассеяния.

Неоднородный объект, подлежащий идентификации, содержится в цилиндре G с центром в начале координат, диаметром основания d и высотой l , $G = \left\{ r = (r_1, r_2, r_3) \in \mathbb{R}^3 : \sqrt{r_1^2 + r_2^2} < \frac{d}{2}, -\frac{l}{2} < r_3 < \frac{l}{2} \right\}$. Обозначим через G_0 плотное подмножество области G , допускающее разбиение на конечное число попарно непересекающихся подобластей $G_1, G_2, \dots, G_p$:

$$G_0 = G_1 \cup \dots \cup G_p, \quad \bar{G}_0 = \bar{G}, \quad G_i \cap G_j = \emptyset \text{ при } i \neq j,$$

и относительно коэффициентов уравнения (1) предположим, что $\mu(r, E) = \mu_i(E)$, $\sigma(r, E) = \sigma_i(E)$, $p(r, \omega \cdot \omega', E) = p_i(\omega \cdot \omega', E)$ при $r \in G_i$, где $\mu_i(E)$, $\sigma_i(E) \in L^\infty[\underline{E}, \bar{E}]$, $p_i(\omega \cdot \omega', E) \in L^\infty([-1, 1] \times [\underline{E}, \bar{E}])$. Все функции неотрицательные, справедливо $\sigma \leq \mu$, и индикатриса рассеяния p удовлетворяет условию нормировки

$$\int_{\Omega} p(r, \omega \cdot \omega', E) d\omega' = 1$$

для почти всех $(r, E) \in G \times [\underline{E}, \bar{E}]$. Вне области G функции μ, σ продолжены нулем, то есть вне G излучение со средой не взаимодействует.

Области G_i заполнены химическими соединениями, входящими в состав неоднородной среды G и характеризуются коэффициентами ослабления и рассеяния $\mu_i(E)$, $\sigma_i(E)$. При томографии среды коэффициенты можно определить лишь приближенно, причем для различных энергий погрешность определения может быть разная. Поэтому задача однозначной идентификации веществ, входящих в заранее составленный список, нетривиальна и требует формализации, разработки алгоритмов ее решения и верификации на широком наборе веществ, представляющих интерес в медицинской или промышленной томографии.

Обозначим через Π_ω плоскость, касательную к границе области G и перпендикулярную направлению ω , $\Pi_\omega = \{r \in \mathbb{R}^3 : r \cdot \omega = \frac{d}{2}\}$. Очевидно, что плоскости $\Pi_{-\omega}$ и Π_ω параллельны и находятся на расстоянии d друг от друга. Сканирование объекта осуществляется путем синхронного поворота плоскостей с источниками излучения $\Pi_{-\omega^*}$ и детекторами $\Pi_{+\omega^*}$ в горизонтальных сечениях $\omega^* \in \Omega^* = \{\omega = (\omega_1, \omega_2, \omega_3) \in \Omega : \omega_3 = 0\}$.

Как правило, в задачах томографии требуется определить хотя бы один из коэффициентов уравнения (1), если для каждого $\omega^* \in \Omega^*$ его решение I_{ω^*} известно на множестве $\Pi_{-\omega^*} \times \Omega^* \times [\underline{E}, \bar{E}] \times [0, T]$ и $\Pi_{+\omega^*} \times \Omega^* \times [\underline{E}, \bar{E}] \times [0, T]$.

Рассматривая прямую задачу для уравнения (1), мы будем опускать параметрическую зависимость решения уравнения $I_{\omega^*}(r, \omega, E, t)$ от направления ω^* , которое характеризует положение внешнего источника облучения. Для краткости изложения введем ряд обозначений $X = G \times \Omega \times [\underline{E}, \bar{E}] \times [0, T]$, $X_0 = G \times \Omega \times [\underline{E}, \bar{E}] \times \{t = 0\}$, $\Omega_{-\omega^*} = \{\omega \in \Omega : -\omega^* \cdot \omega > 0\}$, $Y^- = \Pi_{-\omega^*} \times \Omega_{-\omega^*} \times [\underline{E}, \bar{E}] \times [0, T]$, $X^- = Y^- \cup X_0$ и присоединим к уравнению (1) начальные и граничные условия

$$I|_{X_0} = h_0(r, \omega, E), \quad (2)$$

$$I|_{Y^-} = h_{ext}(z, \omega, E, t), \quad (3)$$

где h_0 характеризует состояние процесса в начальный момент времени, а функция h_{ext} описывает плотность потока излучения падающего на объект. На множестве $X_0 \times Y^-$ построим функцию

$$h(z, \omega, E, t) = \begin{cases} h_0(z, \omega, E), & \text{если } (z, \omega, E, t) \in X_0, \\ h_{ext}(z, \omega, E, t), & \text{если } (z, \omega, E, t) \in Y^-, \end{cases}$$

и объединим условия (2) и (3) в одно начально-краевое условие следующего вида:

$$I|_{X^-} = h(r, \omega, E, t). \quad (4)$$

Относительно функции h в условии (4) предполагается, что $h \in L^\infty(X^-)$.

В прямой задаче требуется определить функцию I из уравнения (1) и условия (4) при заданных μ , σ , p , c , h .

Зачастую решение начально-краевой задачи (1), (4) понимается в некотором обобщенном смысле, при этом дифференциальное выражение в левой части уравнения (1) имеет смысл производной в точке (r, ω, t) в направлении вектора $(\omega, \frac{1}{c})$. Такой подход позволяет избежать довольно значительных трудностей при определении решения задачи.

Для простоты изложения мы ограничимся случаем, когда серийное облучение среды, зависящее от направления ω^* , осуществляется прямоугольными импульсами длительностью ε вида

$$h(\xi, \omega, E, t) = \begin{cases} \frac{1}{\varepsilon}, & (\xi, \omega, t) \in \Pi_{-\omega^*} \times \Omega_{-\omega^*} \times (0, \varepsilon), \\ 0, & (\xi, \omega, t) \notin \Pi_{-\omega^*} \times \Omega_{-\omega^*} \times (0, \varepsilon). \end{cases} \quad (5)$$

С физической точки зрения такое определение функции h накладывает ограничение только на длительность импульсов, не предполагая ни коллимацию зондирующего потока по направлению, ни пространственную локализацию источников излучения, которые не редко используется в томографии для подавления паразитного влияния рассеяния в среде при нахождении коэффициента ослабления излучения.

Сформулируем обратную задачу для уравнения (1), которую также будем называть задачей томографии [Прохоров, Яровенко, 2021; Прохоров, Яровенко, 2022; Nazarov et al., 2023; Yarovenko, Prokhorov, 2024].

Определение 1. Под обратной задачей будем понимать задачу определения функции μ из соотношений (1), (4) и дополнительного условия

$$\int_{d/c}^{d/c+\varepsilon} I(\eta, \omega^*, E, t) dt = H(\eta, \omega^*, E), \quad (\eta, \omega^*, E) \in \Pi_{\omega^*} \times \Omega^* \times [\underline{E}, \bar{E}], \quad (6)$$

в которых величины c, d, ε и функции $h(\xi, \omega^*, E, t), H(\eta, \omega^*, E)$ при $(\xi, \omega^*, E, t) \in \Pi_{\omega^*} \times \Omega^* \times [0, T] \times [\underline{E}, \bar{E}]$, $(\eta, \omega^*, E) \in \Pi_{\omega^*} \times \Omega^* \times [\underline{E}, \bar{E}]$ заданы.

Перейдем к постановке задачи идентификации неизвестного вещества [Nazarov et al., 2023]. Как правило, даже визуальный анализ реконструированной функции $\mu'(r, E_i)$ позволяет достаточно просто определить границы области G_j . В особо трудных случаях, например для слабоконтрастных сильно рассеивающих включений, можно воспользоваться алгоритмами восстановления границ областей G_j , основанными на применении индикатора неоднородности [Anikonov et al., 2002]. Поскольку в пределах каждой подобласти коэффициент ослабления считается зависящим только от энергии, при идентификации веществ нет необходимости использовать значения $\mu'(r, E_i)$ в каждой точке $r \in G$. Для количественной оценки коэффициента ослабления $\mu'(r, E_i)$ в подобластях G_j можно ограничиться конечным набором функционалов

$$M_j(E_i) = \frac{1}{\text{mes}\{G'_j\}} \int_{G'_j} \mu'(r, E_i) dr, \quad j = 1, \dots, p, \quad (7)$$

где G'_j — некоторое подмножество области G_j , $G'_j \subset G_j$.

Определение 2. Пусть задан $S = \{S_1, \dots, S_m\}$ — список всевозможных веществ с коэффициентами $\mu_1(E), \dots, \mu_m(E)$, которые могут содержаться в среде G , и известны $M_j(E_i)$ для всех $i = 1, \dots, q, j = 1, \dots, p$. Задачей идентификации будем называть задачу определения списка $S^* = \{S_1^*, \dots, S_p^*\}$, в котором коэффициент $\mu_j^*(E)$ соответствующий веществу S_j^* , доставляет минимум функционалу

$$J_j(f) = \sqrt{\sum_{i=1}^q (M_j(E_i) - f(E_i))^2} \quad (8)$$

для всех $j = 1, \dots, p$.

На данную задачу будем также ссылаться как на задачу идентификации неизвестной среды.

Асимптотическое решение обратной задачи для уравнения переноса излучения

Обозначим через $L_{r,\omega}$ луч, исходящий из точки $r \in \mathbb{R}^3$ в направлении ω , $L_{r,\omega} = \{r + \omega\tau : \tau > 0\}$, а через $d(r, -\omega)$ — расстояние от точки $r \in G$ до плоскости $\Pi_{-\omega^*}$ в направлении $-\omega$, и пусть $d(r, -\omega, t) = \min\{d(r, -\omega), ct\}$. Если луч $L_{r,\omega}$ не имеет общей точки с плоскостью $\Pi_{-\omega^*}$, то очевидно, что $d(r, -\omega, t) = ct$. Обозначим через

$$W = \left\{ I \in L^\infty(X) : \left(\frac{1}{c} \frac{\partial}{\partial t} + \omega \cdot \nabla_r \right) I \in L^\infty(X) \right\}$$

и введем в рассмотрение операторы $\mathcal{A} : L^\infty(X) \rightarrow W(X)$ и $S : L^\infty(X) \rightarrow L^\infty(X)$, которые определяются выражениями

$$\mathcal{A}f(r, \omega, t) = \int_0^{d(r, -\omega, t)} \exp \left(- \int_0^\tau \mu(\eta - \tau' \omega) d\tau' \right) f \left(r - \omega\tau, \omega, t - \frac{\tau}{c} \right) d\tau, \quad (9)$$

$$Sf(r, \omega, E, t) = \sigma(r) \int_\Omega p(r, \omega \cdot \omega', E) f(r, \omega', E, t) d\omega'. \quad (10)$$

Решение прямой задачи для уравнения переноса эквивалентно решению уравнения интегрального типа [Прохоров, Яровенко, 2021]:

$$I = I_0 + \mathcal{A}SI, \quad (11)$$

для которого справедливо представление в виде равномерно сходящегося ряда Неймана:

$$I = \sum_{n=0}^{\infty} (\mathcal{A}S)^n I_0, \quad (12)$$

где функция

$$I_0(r, \omega, E, t) = h \left(r - d(r, -\omega, t)\omega, \omega, E, t - \frac{d(r, -\omega, t)}{c} \right) \exp \left(- \int_0^{d(r, -\omega, t)} \mu(r - \tau\omega, E) d\tau \right) \quad (13)$$

имеет смысл интенсивности нерассеянного поля, а функция $I_n = (\mathcal{A}S)^n I_0$ при $n = 1, 2, \dots$ описывает n -кратно рассеянное поле. Для простоты предполагается $p = \frac{1}{4\pi}$, что характерно для изотропного типа рассеяния в среде.

В [Nazarov et al., 2023; Yarovenko, Prokhorov, 2024] для функции I_1 , описывающей однократно рассеянное поле, доказана следующая оценка:

$$\Psi_1(\eta, \omega^*, \varepsilon) = \int_{d/c}^{d/c+\varepsilon} I_1(\eta, \omega^*, t) dt \leq \frac{\bar{\sigma}d}{2} \Phi(\varepsilon), \quad \eta \in \Pi_{\omega^*}, \quad (14)$$

где $\bar{\sigma} = \sup_{r \in G} \sigma$ и

$$\Phi(\varepsilon) = \frac{1}{2} \left(\frac{c\varepsilon}{d} \ln \left| 1 + \frac{d}{c\varepsilon} \right| + 1 - \frac{d}{c\varepsilon} \ln \left| 1 + \frac{c\varepsilon}{d} \right| \right). \quad (15)$$

Нетрудно видеть, что функция $\Phi(\varepsilon) \rightarrow 0$ при $\varepsilon \rightarrow 0$ и $\Phi(\varepsilon) \rightarrow 1$ при $\varepsilon \rightarrow \infty$, поэтому интеграл в левой части (14) стремится к нулю при $\varepsilon \rightarrow 0$, а при $\varepsilon \rightarrow \infty$ ограничен сверху константой $\frac{\bar{\sigma}d}{2}$.

При проведении широкомасштабной серии вычислительных экспериментов на цифровых фантомах биологических объектов с помощью метода Монте-Карло мы убедились, что для фиксированного ε монотонно возрастающая последовательность

$$\Psi_m(\eta, \omega^*, \varepsilon) = \int_{d/c}^{d/c+\varepsilon} \sum_{n=1}^m I_n(\eta, \omega^*, t) dt$$

при увеличении m достаточно быстро стабилизируется. Причем, так же как и функция $\Psi_1(\varepsilon)$ в неравенстве (14), выражения $\Psi_m(\varepsilon)$ при $\varepsilon \rightarrow 0$ сходятся к нулю для произвольного $m \geq 1$.

Учитывая вид функции h , описывающей прямоугольный импульс длительности ε , получаем

$$\begin{aligned} \Psi(\eta, \omega^*, \varepsilon) &\equiv \int_{d/c}^{d/c+\varepsilon} I(\eta, \omega^*, t) dt = \int_{d/c}^{d/c+\varepsilon} I_0(\eta, \omega^*, t) dt + \Psi_\infty(\eta, \omega^*, \varepsilon) = \\ &= \exp\left(-\int_0^d \mu(\eta - \omega^* \tau) d\tau\right) + \Psi_\infty(\eta, \omega^*, \varepsilon). \end{aligned} \quad (16)$$

Таким образом, при малых ε имеет место асимптотическое равенство

$$\int_0^d \mu(\eta - \tau \omega^*) d\tau = -\ln |\Psi(\eta, \omega^*, \varepsilon)|, \quad \eta \in \Pi_{\omega^*}. \quad (17)$$

Левая часть соотношения (17) представляет собой преобразование Радона функции μ в некоторой горизонтальной плоскости $\eta_3 = \text{const}$. Для обращения преобразования Радона разработано множество алгоритмов [Natterer, 1986; Herman, Natterer, 2013]. При проведении расчетов, необходимых для решения задачи идентификации, мы использовали алгоритм свертки и обратной проекции [Natterer, 1986]. Так как равенство (17) носит асимптотический характер, то, вне зависимости от используемого алгоритма обращения, рост ошибки восстановления функции μ при увеличении параметра ε неизбежен, что является серьезным препятствием для успешного решения задачи идентификации.

Компьютерная обработка и кластеризация списка веществ, потенциально содержащихся в неоднородной среде

Перед проведением вычислительных экспериментов по решению задачи идентификации для конкретных веществ и энергий представляется весьма полезным изучить вопрос о том, как расположены в соответствующем пространстве вещества, выбранные в качестве гипотетически возможных. Это позволит заранее оценить максимально допустимую абсолютную ошибку определения коэффициентов ослабления излучения $\mu(s, E)$, при которой возможна однозначная идентификация вещества в каждой подобласти. Рассмотрим данный вопрос более детально.

Как говорилось выше, мы рассматривали 31 вещество и пять значений энергий $E_1 = 50$, $E_2 = 60$, $E_3 = 80$, $E_4 = 100$, $E_5 = 150$ кэВ, для которых нам известны значения коэффициентов ослабления $\mu(s, E)$ из таблиц Хаббла–Зельцера [Hubbell, Seltzer, 1995; Berger et al., 2005]. Пронумерованный перечень этих веществ приводится в табл. 1.

Таблица 1. Пронумерованный перечень веществ, потенциально содержащихся в среде

1. Пластик А-150	17. Винилтолуин
2. Жировая ткань	18. Полиэтилен
3. Сухой воздух	19. Пластик Мулар
4. Аланин	20. Полиметилметакрилат
5. В-100 пластик	21. Полистирол
6. Бакелит	22. Тефлон
7. Кровь в целом	23. Полихлорвинил
8. Костная ткань	24. Ткань гонад
9. Ткань мозга	25. Мягкая ткань
10. Ткань груди	26. Мягкая ткань 4к (4-компонентная)
11. Пластик С-552	27. Тк.-эквив. газ
12. Р-р сульфата аммония	28. Тк.-эквив. газ 2
13. Глазная линза	29. Вода
14. Ткань легких	30. Этанол
15. Ткань мускульная	31. Метанол
16. Ткань яичника	

Таким образом, каждому веществу (с номером s) в 5-мерном евклидовом пространстве $\mathbb{R}^5$ ставится в соответствие точка

$$\mu(s, E) = (\mu(s, E_1), \mu(s, E_2), \dots, \mu(s, E_5)).$$

Для упрощения обозначений там, где это не вызывает путаницы, каждое вещество обозначается целочисленным номером s . На первом этапе вычислялась норма каждого вещества:

$$\|\mu(s, E)\| = \sqrt{\sum_{k=1}^5 [\mu(s, E_k)]^2},$$

а также расстояние между всевозможными парами веществ: $d(s, p) = \|\mu(s, E) - \mu(p, E)\|$. В итоге была получена симметрическая матрица размером 31×31 для 465 различных пар веществ. Несложно увидеть, что при идентификации любого вещества s наибольший интерес представляет такое вещество $p = p(s) \neq s$, для которого расстояние $d(s, p(s)) > 0$ является минимальным среди всевозможных значений расстояний $d(s, q)$. Это вещество $p(s)$ создает наибольшую трудность для правильной идентификации вещества s . Далее это обстоятельство будет пояснено на примере конкретных расчетов. Теоретически таких веществ $p(s)$ может быть несколько, однако на практике подобная ситуация является исключительной.

В таблице 2 для каждого вещества s приведены ближайшее вещество $p(s)$ и расстояние между ними. Для более компактного представления результатов таблица поделена на две секции таким образом, чтобы числа $d(s, p(s))$ в ней шли в невозрастающем порядке. В столбцах таблицы 2 под символом s идут (без повторов) номера и названия всех рассмотренных веществ. В столбцах под символом $p(s)$ также идут номера и названия веществ $p(s)$. Однако здесь имеется семь повторов: это вещества с номерами 3, 4, 5, 10, 14, 16, 20, каждое из которых присутствует здесь ровно два раза. В то же время здесь отсутствуют вещества с номерами 8, 11, 12, 13, 15, 26, 28. В столбцах таблицы под символом $d(s, p(s))$ указаны расстояния между соответствующими точками в $\mathbb{R}^5$.

Отсутствие некоторых веществ в столбцах под символом $p(s)$ указывает на то, что они не являются ближайшими ни к одному другому веществу. Повторы некоторых веществ в столбцах под символом $p(s)$ означает, что они являются ближайшими в $\mathbb{R}^5$ одновременно к нескольким

другим веществам. Значения $d(s, p(s))$ для некоторых пар веществ одинаковы, и такие повторы образуют 11 пар.

Помимо этого, были найдены нормы в $\mathbb{R}^5$ для веществ из полного списка. Было установлено, что эти нормы изменяются от значения $4,342 \cdot 10^{-4}$ для $s = 27$ до $1,191$ для $s = 8$, то есть более чем в 2000 раз. При этом основной вклад в норму всегда вносит первая компонента $\mu(s, E_1)$ вектора $\mu(s, E)$ (напомним, что $E_1 = 50$ кэВ).

Таблица 2. Расстояния $d(s, p(s))$ от вещества s до ближайшего к нему вещества $p(s)$. Значения меры близости приведены в порядке убывания

s	$p(s)$	$d(s, p(s))$	s	$p(s)$	$d(s, p(s))$
8. Костная ткань	23. Полихлорвинил	$2,715 \cdot 10^{-1}$	29. Вода	10. Ткань груди	$4,769 \cdot 10^{-3}$
23. Полихлорвинил	5. Пластик В-100	$1,520 \cdot 10^{-1}$	10. Ткань груди	29. Вода	$4,769 \cdot 10^{-3}$
22. Тефлон	5. Пластик В-100	$1,351 \cdot 10^{-1}$	31. Метанол	30. Этанол	$3,390 \cdot 10^{-3}$
5. Пластик В-100	22. Тефлон	$1,351 \cdot 10^{-1}$	30. Этанол	31. Метанол	$3,390 \cdot 10^{-3}$
11. Пластик С-552	4. Аланин	$1,075 \cdot 10^{-1}$	13. Глазная линза	16. Ткань яичника	$2,944 \cdot 10^{-3}$
19. Пластик Мулар	4. Аланин	$3,744 \cdot 10^{-2}$	25. Мягкая ткань	7. Кровь в целом	$1,826 \cdot 10^{-3}$
4. Аланин	19. Пластик Мулар	$3,744 \cdot 10^{-2}$	7. Кровь в целом	25. Мягкая ткань	$1,826 \cdot 10^{-3}$
12. Р-р сульфата аммония	1. Пластик А-150	$2,792 \cdot 10^{-2}$	24. Ткань гонад	9. Ткань мозга	$1,139 \cdot 10^{-3}$
1. Пластик А-150	20. Полиметил-метакрилат	$1,023 \cdot 10^{-2}$	9. Ткань мозга	24. Ткань гонад	$1,139 \cdot 10^{-3}$
20. Полиметил-метакрилат	6. Бакелит	$1,004 \cdot 10^{-2}$	15. Ткань мускульная	14. Ткань легких	$1,065 \cdot 10^{-3}$
6. Бакелит	20. Полиметил-метакрилат	$1,004 \cdot 10^{-2}$	16. Ткань яичника	14. Ткань легких	$5,143 \cdot 10^{-4}$
18. Полиэтилен	2. Жировая ткань	$9,113 \cdot 10^{-3}$	14. Ткань легких	16. Ткань яичника	$5,143 \cdot 10^{-4}$
2. Жировая ткань	18. Полиэтилен	$9,113 \cdot 10^{-3}$	28. Тк.-эквив. газ 2	3. Сухой воздух	$2,768 \cdot 10^{-4}$
21. Полистирол	17. Винилтолуин	$8,416 \cdot 10^{-4}$	27. Тк.-эквив. газ	3. Сухой воздух	$3,171 \cdot 10^{-5}$
17. Винилтолуин	21. Полистирол	$8,416 \cdot 10^{-4}$	3. Сухой воздух	27. Тк.-эквив. газ	$3,171 \cdot 10^{-5}$
26. Мягкая ткань 4к	10. Ткань груди	$6,067 \cdot 10^{-3}$			

На основе полученной информации была построена таблица 3. В ней указаны все рассмотренные вещества (31 шт). Запись $n \rightarrow m$ или $m \leftarrow n$ означает, что вещество m является ближайшим (в $\mathbb{R}^5$) к веществу n , а запись $n \leftrightarrow m$ означает, что вещество m является ближайшим к веществу n и, одновременно с этим, вещество n является ближайшим к веществу m . Вещественное число, стоящее в скобках справа от номера вещества s , есть норма $\|\mu(s, E)\|$.

Например, в строке $N = 1$ символьная запись означает следующее. Для вещества $s = 27$ норма $\|\mu(27, E)\| = 4,342 \cdot 10^{-4}$, а для вещества $s = 3$ норма $\|\mu(3, E)\| = 4,639 \cdot 10^{-4}$, и точки $\mu(27, E)$ и $\mu(3, E)$ являются единственными ближайшими (в $\mathbb{R}^5$) друг к другу. В свою очередь, вещество $s = 3$ является единственным ближайшим к веществу $s = 28$, и норма $\|\mu(28, E)\| = 7,401 \cdot 10^{-4}$. Кроме этого, вещество $s = 28$ не является ближайшим ни к какому другому веществу $q \in S$. Все это легко проверяется по данным из табл. 2. Таким образом, вещества с номерами 27, 3 и 28 образуют подмножество в S , которое мы будем обозначать символом K_1 и называть кластером в $S = \{S_1, \dots, S_{31}\}$.

Аналогично устроены все остальные кластеры K_i , $2 \leq i \leq 11$, в табл. 3, причем кластеры попарно не пересекаются и их объединение есть множество всех рассмотренных веществ S .

$$S = K_1 \cup \dots \cup K_{11}, \quad K_i \cap K_j = \emptyset \quad \text{при } i \neq j. \quad (18)$$

С формальной точки зрения можно дать следующее определение кластера.

Таблица 3. Кластеры и нормы $\|\mu(s, E)\|$ в $\mathbb{R}^5$

N	Номера веществ s в кластере N (в скобках $\ \mu(s, E)\ $)
1	27 $(4,342 \cdot 10^{-4}) \leftrightarrow 3 (4,639 \cdot 10^{-4}) \leftarrow 28 (7,401 \cdot 10^{-4})$
2	30 $(3,279 \cdot 10^{-1}) \leftrightarrow 31 (3,305 \cdot 10^{-1})$
3	18 $(3,818 \cdot 10^{-1}) \leftrightarrow 2 (3,887 \cdot 10^{-1})$
4	17 $(4,042 \cdot 10^{-1}) \leftrightarrow 21 (4,126 \cdot 10^{-1})$
5	26 $(4,171 \cdot 10^{-1}) \rightarrow 10 (4,224 \cdot 10^{-1}) \leftrightarrow 29 (4,235 \cdot 10^{-1})$
6	24 $(4,390 \cdot 10^{-1}) \leftrightarrow 9 (4,400 \cdot 10^{-1})$
7	15 $(4,421 \cdot 10^{-1}) \rightarrow 14 (4,430 \cdot 10^{-1}) \leftrightarrow 16 (4,434 \cdot 10^{-1}) \leftarrow 13 (4,437 \cdot 10^{-1})$
8	25 $(4,464 \cdot 10^{-1}) \leftrightarrow 7 (4,479 \cdot 10^{-1})$
9	12 $(4,737 \cdot 10^{-1}) \rightarrow 1 (4,666 \cdot 10^{-1}) \rightarrow 20 (4,742 \cdot 10^{-1}) \leftrightarrow 6 (4,840 \cdot 10^{-1})$
10	19 $(5,333 \cdot 10^{-1}) \leftrightarrow 4 (5,707 \cdot 10^{-1}) \leftarrow 11 (6,774 \cdot 10^{-1})$
11	22 $(8,629 \cdot 10^{-1}) \leftrightarrow 5 (7,877 \cdot 10^{-1}) \leftarrow 23 (9,208 \cdot 10^{-1}) \leftarrow 8 (1,191)$

Определение 3. Пусть $X = \{x_1, x_2, \dots, x_m\}$ — метрическое пространство, состоящее из m точек. Непустое подмножество $K = \{y_1, y_2, \dots, y_n\} \subset X$ будем называть кластером в X , если K удовлетворяет следующим условиям:

- (i) если $y \in K$ и $x = x(y) \in X$ — ближайшая к y точка из X , отличная от y , то $x \in K$;
- (ii) если $z \in X$ — точка, отличная от $y \in K$, для которой y является ближайшей, то $z \in K$;
- (iii) никакое собственное подмножество $K_1 \subset K$ не удовлетворяет одновременно условиям (i) и (ii), в которых K заменено на K_1 .

Покажем, что кластеры всегда существуют. Выберем произвольную точку в X , например x_1 , и построим множества (первого уровня) U_1 и D_1 следующим образом. Во множество U_1 включим все точки из $X \setminus x_1$, для которых x_1 является ближайшей, а во множество D_1 включим все точки из $X \setminus x_1$, которые являются ближайшими для x_1 . Множество U_1 может оказаться пустым, но множество D_1 всегда не пусто и может содержать несколько точек.

Для каждой точки d_{1k} из D_1 аналогичным образом построим множества (второго уровня) $D_2 d_{1k}$ и $d_{1k} U_2$, а затем $(D_2 D_1) = \bigcup_k (D_2 d_{1k})$ и $(D_1 U_2) = \bigcup_k (d_{1k} U_2)$. Если U_1 не пусто, то для каждой точки $u_{1s} \in U_1$ аналогичным образом построим множества (второго уровня) $u_{1s} U_2$ и $D_2 u_{1s}$, а затем $(U_1 U_2) = \bigcup_s (u_{1s} U_2)$ и $(D_2 U_1) = \bigcup_s (D_2 u_{1s})$. Можно показать, что цепочка вложенных множеств

$$x_1 \subset [D_1 \cup x_1 \cup U_1] \subset [(D_2 D_1) \cup (D_1 U_2) \cup D_1 \cup x_1 \cup U_1 \cup (D_2 U_1) \cup (U_1 U_2)] \subset \dots \subset S$$

стабилизируется за конечное число шагов, удовлетворяет условиям (i), (ii), (iii) и, таким образом, представляет собой кластер в X , порожденный точкой x_1 .

Несложно убедиться, что нормы веществ в таблице 3 монотонно возрастают (слева направо и сверху вниз), за исключением двух случаев, в кластерах $N = 9$ и $N = 11$.

Это обстоятельство наводит на мысль, что, по-видимому, точки $\mu(s, E_1, E_2, \dots, E_5)$, $s = 1, 2, \dots, 31$, в пространстве $\mathbb{R}^5$ расположены достаточно близко к некоторой прямой. Ответить

на этот вопрос в пространстве $\mathbb{R}^5$ сразу затруднительно, поскольку термин «достаточно близко» несколько расплывчат. Однако при рассмотрении двумерных подмножеств, вопрос решается легко (и положительно). Остановимся на этом подробнее.

В рамках исследования был проведен анализ расположения точек, соответствующих коэффициентам ослабления веществ, входящих в список потенциального интереса, в двумерных подпространствах, определяемых различными парами энергетических уровней. Рисунок 1 иллюстрирует характерное поведение графиков, полученное для подпространства, образованного энергиями E_3 и E_5 , при этом по оси абсцисс отложены значения $\mu(s, E_3)$, а по оси ординат — $\mu(s, E_5)$. Полученные данные указывают на наличие четко выраженной линейной корреляции между значениями $\mu(s, E_3)$ и $\mu(s, E_5)$, с линией тренда, проходящей через начало координат. Дополнительные расчеты выявили, что все точки лежат внутри конуса с раствором в $3,17^\circ$, если за ось конуса принять прямую, проходящую через начало координат и точку, соответствующую веществу $s = 8$. Аналогичная картина наблюдалась при различном выборе пар энергий для $\mathbb{R}^2$. Проведение подобных экспериментов в исходном пятимерном пространстве $\mathbb{R}^5$, при использовании пяти ранее указанных энергий, показало, что все точки располагались внутри конуса с раствором в $15,4^\circ$. Подобная тенденция сохраняется и при дальнейшем увеличении размерности пространства за счет добавления дополнительных энергий. В частности, при использовании всего набора энергетических уровней, присутствующих в таблицах Хаббла – Зельцера (36 значений, соответствующих энергиям от 0,001 до 20 МэВ), раствор конуса составляет $3,86^\circ$.

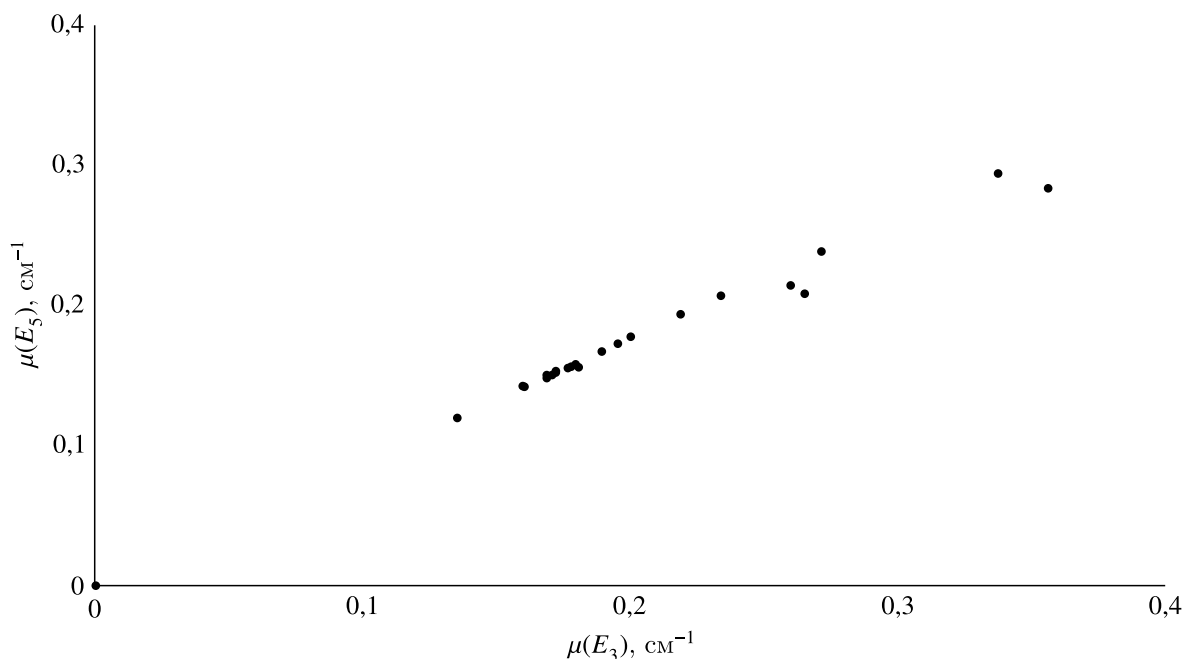


Рис. 1. Геометрическое расположение точек, соответствующих коэффициентам ослабления исследуемых веществ в двумерном подпространстве, образованном энергетическими уровнями E_3 и E_5 . По горизонтальной оси отложены значения коэффициентов ослабления $\mu(s, E_3)$, по вертикальной оси — значения $\mu(s, E_5)$, где s — номер вещества. График демонстрирует четко выраженную линейную корреляцию между коэффициентами ослабления на различных энергетических уровнях с линией тренда, проходящей через начало координат

Важно отметить, что при детальном рассмотрении многие точки на рис. 1 фактически представляют собой кластеры, состоящие из нескольких веществ. Например, точка, ближайшая к началу системы координат, является кластером из веществ с номерами 3, 27 и 28, что в точности соответствует кластеру $N = 1$ из табл. 3. Следующая по удаленности от начала координат

точка представляет собой кластер из веществ 30 и 31 и т. д. Всего в двумерных случаях было выявлено 12 кластеров, которые по составу входящих в них веществ схожи с кластерами, представленными в табл. 3. Этот факт указывает на устойчивость кластерной структуры при изменении размерности пространства.

Полученные результаты имеют важное практическое значение для решения задачи идентификации веществ. Выявленная кластерная структура расположения веществ в энергетическом пространстве позволяет определить группы веществ, которые могут оказаться трудно различимыми в процессе идентификации.

Кратко остановимся на вопросе о том, насколько правомерно мы используем термин «кластер» в нашей работе в отношении множества $K \subset X$, определенного условиями (i), (ii), (iii). Основная причина, по которой мы производим кластеризацию множества S , состоит в следующем. Мы хотим найти во множестве S такие подмножества веществ, которые представляют наибольшую трудность для однозначного решения задачи идентификации неизвестного вещества. Задача нахождения таких подмножеств решается до того, как мы приступаем к просвечиванию исследуемого тела рентгеновским излучением. Эта работа начинается с построения матрицы расстояний $d(s, p)$, $s, p \in S$, которая играет главную роль в большинстве методов кластерного анализа [Mandel, 1988; Дюран, Оделл, 1977; Миркин, 2011]. Затем, проводя обработку данных матрицы $d(s, p)$, мы получаем табл. 2, содержащую наиболее важную для нас информацию. И наконец, используя условия (i), (ii), (iii), мы решаем одну из основных задач кластерного анализа, которая, согласно [Дюран, Оделл, 1977; Миркин, 2011], состоит в разбиении множества S на попарно не пересекающиеся классы (кластеры). Результаты этой работы представлены в табл. 3. Из сказанного ранее несложно увидеть, что в нашем случае в каждом кластере K_i , $1 \leq i \leq 11$, существует единственная пара веществ, расстояние между которыми является минимальным среди всех расстояний между любыми парами веществ данного кластера. Для K_1 это пара $27 \leftrightarrow 3$, для K_2 — пара $30 \leftrightarrow 31$ и т. д. Именно эти пары веществ представляют наибольшую трудность для однозначного решения задачи идентификации. Исходя из сказанного, мы считаем, что использование слова «кластер» в нашей работе является правомерным.

Отметим еще один важный вопрос. Как влияют ошибки в исходных данных на результаты кластеризации множества S ? Исходными данными для S являются значения коэффициентов ослабления излучения $\mu(s, E_k)$ для всевозможных веществ $s \in S$ и энергий E_k , которые брались из таблиц [Hubbell, Seltzer, 1995]. Согласно [Hubbell, Seltzer, 1995] их точность — не менее четырех значащих цифр, т. е. не хуже 0,01 %. Мы считаем, что этой точности вполне достаточно для того, чтобы утверждать, что разбиение множества S на кластеры происходит корректно. В этом легко убедиться, если посмотреть, как в табл. 3 меняются нормы точек на границах соседних кластеров, т. е. при переходе от кластера K_i к кластеру K_{i+1} , $1 \leq i \leq 10$. Ясно, что коэффициенты $\mu'(s, E_k)$, найденные по результатам просвечивания неизвестной среды и решения уравнения переноса, могут иметь гораздо большую ошибку, чем 0,01 %. Соответствующие им точки $\mu'(s, E)$ при этом могут находиться очень далеко от точек $\mu(s, E)$, но это уже не имеет никакого отношения к процедуре кластеризации.

Вычислительные эксперименты на цифровом фантоме

Для тестирования предложенного алгоритма нами был разработан цифровой фантом, который использовался в дальнейших численных экспериментах. Фантом представляет собой цилиндр высотой и диаметром 10 см, сделанный из полиэтилена. Внутри фантома расположены три цилиндрические полости, заполненные водой, этиловым и метиловым спиртом соответственно. Геометрия томографического сканера и схематичное изображение применяемого в численных экспериментах цифрового фантома приведены на рис. 2, а. Заметим, что данный фантом представляет собой модель слабоконтрастного объекта, так как коэффициенты ослабления фоновой

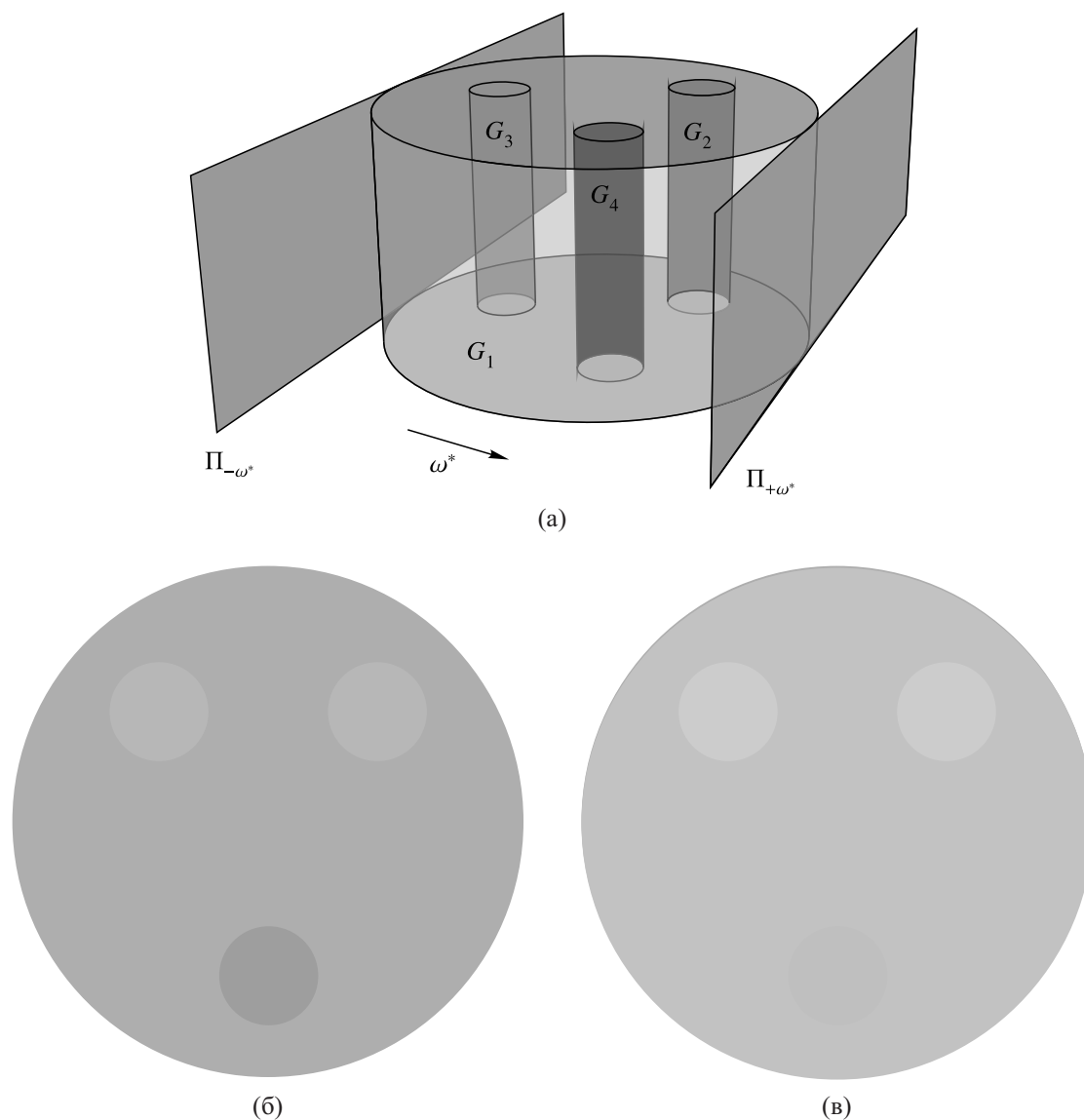


Рис. 2. Цифровой фантом: а) геометрия томографического сканера и схематичное изображение структуры фантома; б) результаты реконструкции по проекционным данным при длительности импульса 3 пс на энергии излучения 50 кэВ; в) результаты реконструкции на энергии излучения 150 кэВ. Коэффициенты ослабления фоновой среды и включений фантома близки друг к другу на всем диапазоне энергий, что затрудняет решение задачи идентификации, причем контрастность томограмм падает с ростом энергии источника излучения

среды и всех включений довольно близки друг к другу на всем диапазоне энергий. Более того, два включения содержат спирты, коэффициенты ослабления которых практически совпадают. Таким образом, для уверенного распознавания веществ требуется очень хорошее качество реконструкции коэффициента ослабления среды.

В численных экспериментах моделировался томограф, содержащий 600 детекторов, дискретизация по углу составляла 800 направлений. Количество детекторов и угловых направлений выбиралось достаточно большим, чтобы снизить влияние погрешности дискретизации на качество реконструкции коэффициента ослабления среды. Для описания серийного облучения среды использовался импульсный источник, определяемый формулой (5). Длительность применяемых

импульсов варьировалась в экспериментах и составляла 3, 10, 30 пикосекунд. Данные о значениях коэффициентов, входящих в состав среды на различных энергиях, рассчитывались при помощи базы данных Xcom [Berger et al., 2005].

Каждый эксперимент включал в себя три этапа. На первом этапе мы использовали метод Монте-Карло, чтобы получить функции, описывающие проекционные данные, соответствующие различным длительностям импульсов ε_i и энергиям E_j входящего излучения. При вычислении проекционных данных моделировалось 10^7 траекторий и учитывалось до 10 актов взаимодействия излучения с веществом. Такие параметры обеспечили хорошую точность вычисления проекционных данных, так что относительная ошибка плотности потока выходящего излучения не превосходила 0,5 %.

На втором этапе мы решали задачу обращения преобразования Радона для найденных проекционных данных с помощью алгоритма свертки и обратной проекции [Natterer, 1986]. Примеры такой реконструкции по проекционным данным, соответствующим длительности импульса 3 пс, для энергий излучения 50 и 150 кэВ приведены на рис. 2, б и рис. 2, в. После этого для каждого из включений мы находили среднее значение коэффициента ослабления, соответствующее длительностям импульсов ε_i и энергиям E_j входящего излучения. На последнем этапе мы вычисляли расстояния от восстановленного коэффициента μ до коэффициентов ослабления веществ, содержащихся в табл. 1. После этого полный список веществ ранжировался по увеличению расстояния и вещество, расстояние до которого было наименьшим, объявлялось «наиболее вероятным» и включалось в соответствующий список.

Для выяснения того, как решение задачи идентификации зависит от длительности импульса используемого излучения и коэффициента ослабления излучения вещества, которое подвергается идентификации, была создана компьютерная программа. Ниже приводится таблица, содержащая основные результаты выполненных расчетов. Пусть $\mu_i(s, E)$ — референтное значение из таблиц [Hubbell, Seltzer, 1995; Berger et al., 2005] коэффициента ослабления s -го вещества на энергии E в области G_i , $\mu_i^*(E, \varepsilon)$ — решение задачи идентификации при длительности импульса ε ,

$$R_i = R_i(s, \varepsilon) = \frac{\mu_i(s, E) - \mu_i^*(E, \varepsilon)}{\mu_i(s, E)} \quad (19)$$

— относительная ошибка результата расчетов для энергии E . При фиксированных s, ε, E наибольшее значение R_i обычно принимала при $E = 50$ кэВ, поэтому далее значения μ_i, μ_i^* и R_i приводятся именно для этой энергии.

В первом столбце таблицы 4 указаны истинное вещество, находящееся в соответствующей подобласти G_i , и его номер s в табл. 1. Во втором столбце — референтные (истинные) коэффициенты ослабления $\mu_i(s)$ этих веществ. В третьем столбце — значения коэффициентов μ_i^* и R_i , найденные для импульса длительностью $\varepsilon = 3$ пс. Там же указаны два вещества, ближайшие по метрике (8) к идентифицируемому веществу, коэффициент ослабления которого был найден в результате ранее описанных расчетов при $\varepsilon = 3$ пс. При этом всегда первое вещество оптимальное и самое близкое, а второе — следующее по расстоянию. В четвертом и пятом столбцах — значения, найденные для импульса длительностью $\varepsilon = 10$ пс и $\varepsilon = 30$ пс соответственно.

Прокомментируем полученные результаты. Во всех случаях значения показателя ошибки R монотонно возрастают с ростом длительности зондирующего импульса ε , при этом все численные значения R_i оказались положительными. Таким образом, ввиду отсутствия в определении величин R_i модуля, все значения μ_i^* оказались меньше соответствующих им (точных) значений $\mu_i(s)$. Аналогичная недооцененность коэффициента ослабления характерна при его восстановлении с применением подавления рассеяния при помощи ультракоротких импульсов, так же как и рост ошибки с увеличением значений коэффициента ослабления неизвестного вещества [Прохоров, Яровенко, 2021; Nazarov et al., 2023; Yarovenko, Prokhorov, 2024].

Таблица 4. Решение задачи идентификации для списка веществ S при длительности зондирующих импульсов $\varepsilon = 3, 10, 30$ пс. Приведены значения коэффициентов ослабления μ_i, μ_i^* для $E = 50$ кэВ, названия оптимального вещества и следующего за ним ближайшего к оптимальному вещества из списка S

Вещество Номер в списке Подобласть G_i	$\mu_i(s)$	$\varepsilon = 3$ пс μ_i^*, R_i 2 ближайших вещества	$\varepsilon = 10$ пс μ_i^*, R_i 2 ближайших вещества	$\varepsilon = 30$ пс μ_i^*, R_i 2 ближайших вещества
Полиэтилен $s = 18$ G_1	0,19381	0,18866; 0,0266 Полиэтилен Жировая ткань	0,18463; 0,0474 Полиэтилен Жировая ткань	0,18087, 0,0668 Полиэтилен Метанол
Этанол $s = 30$ G_2	0,17014	0,16935; 0,00464 Этанол Метанол	0,16666; 0,0205 Этанол Метанол	0,16240, 0,0455 Этанол Метанол
Метанол $s = 31$ G_3	0,17307	0,17221; 0,00499 Метанол Этанол	0,16948; 0,0208 Этанол Метанол	0,16514, 0,0458 Этанол Метанол
Вода $s = 29$ G_4	0,22689	0,22393; 0,0130 Мягкая ткань 4к Вода	0,22050; 0,0281 Мягкая ткань 4к Вода	0,21467, 0,0538 Винилтолуин Жировая ткань

Наихудшие результаты относятся к веществу 29 (вода), которое не удается идентифицировать даже при использовании достаточно коротких импульсов. Так, при $\varepsilon = 3$ пс и $\varepsilon = 10$ пс ближайшим веществом к воде оказывается вещество под номером 26 (мягкая ткань четырехкомпонентная), которое входит в тот же кластер $N = 5$ (см. таблицу 3). Увеличение длительности импульса до $\varepsilon = 30$ пс приводит к росту ошибки идентификации настолько, что ближайшим веществом оказывается вещество под номером 17 (винилтолуин), которое лежит в смежном кластере $N = 4$.

Для включений, содержащих спирты, результаты оказались достаточно хорошими. Для вещества 31 (метанол) при $\varepsilon = 3$ пс результат идентификации оказался правильным, а при $\varepsilon = 10$ пс и $\varepsilon = 30$ пс ближайшим к метанолу оказался этанол. Отметим, что оба спирта лежат в одном кластере (кластер $N = 2$), при этом ситуация, что этанол определяется верно во всех экспериментах, в отличие от метанола, связана с уже упоминавшейся выше недооцененностью восстановленных значений коэффициента ослабления и является следствием того, что этанол имеет коэффициент ослабления меньше, чем метанол. При этом уменьшение длительности импульса до $\varepsilon = 3$ пс снижает уровень ошибки в проекционных данных настолько, что позволяет дифференцировать данные вещества. И наконец вещество под номером 18 (полиэтилен) уверенно распознавалось во всех тестах.

Заключение

В данной статье задача идентификации неизвестного вещества формулируется как задача поиска элемента наилучшего приближения из списка потенциально возможных веществ, содержащихся в исследуемой среде. В такой постановке критически важными требованиями являются высокая точность восстановления коэффициентов ослабления исследуемой среды и различимость различных веществ, входящих в потенциальный список.

В рамках исследования был проведен предварительный анализ поведения коэффициентов ослабления для веществ (их количество — 31), представляющих интерес в рентгенодиагностике и входящих в справочные таблицы Хаббла – Зельцера [Hubbell, Seltzer, 1995; Berger et al., 2005],

в результате которого было выявлено существование кластерных структур, которые устойчиво сохранялись при изменении размерности рассматриваемого энергетического пространства. Данный результат имеет важное практическое значение для решения задачи идентификации веществ. В частности, выявленная кластерная структура расположения веществ в энергетическом пространстве позволяет определить группы веществ, которые могут оказаться трудно различимыми в процессе идентификации. Эта информация может быть использована для оптимизации алгоритмов идентификации и оценки их ограничений. Более того, понимание кластерной структуры данных позволяет более точно оценить влияние погрешностей входных данных на результаты идентификации. Это особенно важно при работе с экспериментальными измерениями, где неизбежно присутствуют шумы и систематические ошибки. Кроме того, знание о близко расположенных веществах дает возможность уточнить оценки точности самого алгоритма идентификации, определив его разрешающую способность для различных групп веществ. В данной работе основное внимание было уделено снижению ошибок, вызванных рассеянием, и другие аспекты, влияющие на качество реконструкции коэффициента ослабления в томографии (такие как геометрические искажения, неидеальность спектрального разрешения детекторов и т.д.), остались за рамками данного исследования. Конечно, с практической точки зрения это накладывает определенные ограничения на нашу работу. Тем не менее даже такие скромные достижения в снижении влияния рассеяния представляют интерес и заслуживают внимания при решении сложных задач, таких как идентификация веществ с помощью методов рентгеновской диагностики.

Список литературы (References)

- Дюран Б. С., Оделл П. Л. Кластерный анализ. — М.: Статистика, 1977.
Duran B. S., Odell P. L. Cluster analysis. — A survey. — Berlin, Heidelberg, New York: Springer-Verlag, 1974. (Russ. ed.: Dyuran B. S., Odell P. L. Klasternyi analiz. — Moscow: Statistika, 1977.)
- Клименов В. А., Осипов С. П., Темник А. К. Идентификация вещества объекта контроля методом дуальных энергий // Дефектоскопия. — 2013. — № 11. — С. 40–50.
Klimenov V. A., Osipov S. P., Temnik A. K. Identification of the substance of a test object using the dual energy method // Russ. J. Nondestruct. Test. — 2013. — Vol. 49. — P. 642–649. — <https://doi.org/10.1134/S1061830913110053> (Original Russian paper: Klimenov V. A., Osipov S. P., Temnik A. K. Identifikatsiya veshchestva ob'ekta kontrolya metodom dual'nykh energii // Defektoskopiya. — 2013. — No. 11. — P. 40–50.)
- Мандель И. Д. Кластерный анализ. — М.: Финансы и статистика, 1988.
Mandel' I. D. Klasternyi analiz [Cluster analysis]. — Moscow: Finansy i statistika, 1988 (in Russian).
- Миркин Б. Г. Методы кластер-анализа для поддержки принятия решений: обзор. — М.: Высшая школа экономики, 2011.
Mirkin B. G. Metody klaster-analiza dlya podderzhki prinyatiya reshenii: obzor [Cluster analysis methods for decision support: a review]. — Moscow: Vysshaya shkola ekonomiki, 2011 (in Russian).
- Назаров В. Г. Метод гиперплоскостей в задаче идентификации неизвестного вещества // Сиб. журн. индустр. матем. — 2021. — Т. 24, № 3. — С. 39–54.
Nazarov V. G. Method of hyperplanes in the problem of identification of an unknown substance // J. Appl. Ind. Math. — 2021. — Vol. 15. — P. 451–465. — <https://doi.org/10.1134/S199047892103008X> (Original Russian paper: Nazarov V. G. Metod giperploskosti v zadache identifikatsii neizvestnogo veshchestva // Sib. zhurn. industr. matem. — 2021. — Vol. 24, No. 3. — P. 39–54.)
- Осипов С. П., Удод В. А., Ван Я. Распознавание материалов методом дуальных энергий при радиационном контроле объектов // Дефектоскопия. — 2017. — Т. 8. — С. 35–56.
Osipov S. P., Udod V. A., Wang Y. Identification of materials in X-ray inspections of objects by the dual-energy method // Russ. J. Nondestruct. Test. — 2017. — Vol. 53. — P. 568–587. — <https://doi.org/10.1134/S1061830917080058> (Original Russian paper: Osipov S. P., Udod V. A., Wang Y. Raspoznavanie materialov metodom dual'nykh energii pri radiatsionnom kontrole ob'ektov // Defektoskopiya. — 2017. — Vol. 8. — P. 35–56.)
- Прохоров И. В., Яровенко И. П. Задача определения коэффициента ослабления для нестационарного уравнения переноса излучения // Ж. вычисл. матем. и матем. физ. — 2021. — Т. 61, № 12. — С. 2095–2108.

- Prokhorov I. V., Yarovenko I. P.* Determination of the attenuation coefficient for the nonstationary radiative transfer equation // *Comput. Math. Math. Phys.* — 2021. — Vol. 61. — P. 2088–2101. — <https://doi.org/10.1134/S0965542521120101> (Original Russian paper: *Prokhorov I. V., Yarovenko I. P.* Zadacha opredeleniya koeffitsienta oslableniya dlya nestatsionarnogo uravneniya perenosa izlucheniya // *Zh. vychisl. matem. i matem. fiz.* — 2021. — Vol. 61, No. 12. — P. 2095–2108.)
- Прохоров И. В., Яровенко И. П.* Повышение качества томографических изображений при облучении среды импульсами различной длительности // *Доклады Российской академии наук. Математика, информатика, процессы управления.* — 2022. — Т. 505, № 1. — С. 71–78.
- Prokhorov I. V., Yarovenko I. P.* Improving the quality of tomographic images of a medium using irradiation with pulses of different duration // *Doklady Mathematics.* — 2022. — Vol. 106, No. 1. — P. 272–278. — <https://doi.org/10.1134/S1064562422040135> (Original Russian paper: *Prokhorov I. V., Yarovenko I. P.* Povyshenie kachestva tomograficheskikh izobrazhenii pri obluchenii sredy impul'sami razlichnoi dlitel'nosti // *Doklady Rossiiskoi akademii nauk. Matematika, informatika, protsessy upravleniya.* — 2022. — Vol. 505, No. 1. — P. 71–78.)
- Ahmed O. M. H., Song Y., Xie Z.* Material identification approach based on the counting technique and beam hardening correction under industrial X-ray computed tomography: a simulation study // *Braz. J. Phys.* — 2022. — Vol. 52. — P. 26. — <https://doi.org/10.1007/s13538-021-01012-4>
- Alvarez R. E., Macovski A.* Energy-selective reconstructions in X-ray computerized tomography // *Phys. Med. Biol.* — 1976. — Vol. 21. — P. 733–744. — <https://doi.org/10.1088/0031-9155/21/5/002>
- Anikonov D. S., Kovtanyuk A. E., Prokhorov I. V.* Transport equation and tomography. — Utrecht, Bosto: VSP, 2002.
- Bal G., Gong R., Terzioglu F.* An inversion algorithm for P-functions with applications to multi-energy CT // *Inverse Probl.* — 2022. — Vol. 38. — P. 035011. — <https://doi.org/10.1088/1361-6420/ac4d0a>
- Berger M. J., Hubbell J. H., Seltzer S. M., Chang J., Coursey J. S., Sukumar R., Zucker D. S.* XCOM: photon cross section database. — Gaithersburg, MD, USA: National Institute of Standards and Technology, 2005.
- Brooks R. A.* A quantitative theory of the Hounsfield unit and its application to dual energy scanning // *Comput. Assist. Tomog.* — 1977. — Vol. 1. — P. 487–493. — <https://doi.org/10.1097/00004728-197710000-00016>
- Herman G., Natterer F.* Mathematical aspects of computerized tomography. — Oberwolfach: Springer Science & Business Media, 2013.
- Hubbell J. H., Seltzer S. M.* Tables of X-ray mass attenuation coefficients and mass energy absorption coefficients 1 keV to 20 MeV for elements $Z = 1$ to 92 and 48 additional substances of dosimetric interest. — NISTIR-5632. — Gaithersburg, MD, USA: National Institute of Standards and Technology, 1995.
- Jumanazarov D., Koo J., Busi M., Poulsen H. F., Olsen U. L., Iovea M.* System-independent material classification through X-ray attenuation decomposition from spectral X-ray CT // *NDT & E Int.* — 2020. — Vol. 116. — P. 102336. — <https://doi.org/10.1016/j.ndteint.2020.102336>
- Komarskiy A., Korzhenevskiy S., Ponomarev A., Chepusov A.* Dual-energy processing of X-ray images of beryl in muscovite obtained using pulsed X-ray sources // *Sensors.* — 2023. — Vol. 23. — P. 4393. — <https://doi.org/10.3390/s23094393>
- Natterer F.* The mathematics of computerized tomography. — Chichester, UK: Wiley, 1986.
- Nazarov V. G.* The problem of identification an unknown substance by the radiographic method // *J. Phys. Conf. Ser.* — 2021. — Vol. 2099. — P. 012043. — <https://doi.org/10.1088/1742-6596/2099/1/012043>
- Nazarov V. G., Prokhorov I. V., Yarovenko I. P.* Identification of an unknown substance by the methods of multi-energy pulse X-ray tomography // *Mathematics.* — 2023. — Vol. 11, No. 15. — P. 3263. — <https://doi.org/10.3390/math11153263>
- Osipov S. P., Libin E. E., Chakhlov S. V., Osipov O. S., Shtein A. M.* Parameter identification method for dual-energy X-ray imaging // *NDT & E Int.* — 2015. — Vol. 76. — P. 38–42. — <https://doi.org/10.1016/j.ndteint.2015.08.003>

- Rebuffel V., Dinten J.-M.* Dual-energy X-ray imaging: Benefits and limits // *Insight*. — 2007. — Vol. 49. — P. 589–594. — <https://doi.org/10.1784/insi.2007.49.10.589>
- Wang Z., Bovik A., Sheikh H., Simoncelli E.* Image quality assessment: from error visibility to structural similarity // *IEEE Trans. Image Process.* — 2004. — Vol. 13. — P. 600–612. — <https://doi.org/10.1109/TIP.2003.819861>
- Yarovenko I. P., Prokhorov I. V.* An extrapolation method for improving the quality of tomographic images using multiple short-pulse irradiations // *Journal of Inverse and Ill-posed Problems*. — 2024. — Vol. 32, No. 1. — P. 57–74. — <https://doi.org/10.1515/jiip-2023-0022>
- Ying Z., Naidu R., Crawford C.* Dual energy computed tomography for explosive detection // *J. X-ray Sci. Technol.* — 2006. — Vol. 14. — P. 235–256.
- Yokhana V. S. K., Arhatari B. D., Abbey B.* Materials separation via the matrix method employing energy-discriminating X-ray detection // *Appl. Sci.* — 2022. — Vol. 12. — P. 3198. — <https://doi.org/10.3390/app12063198>