УДК 004.932.4
doi: 10.21685/2227-8486-2024-1-10

ФИЛЬТРАЦИЯ ОДИНОЧНЫХ ЗАТУМАНЕННЫХ/ ЗАДЫМЛЕННЫХ ИЗОБРАЖЕНИЙ С ИСПОЛЬЗОВАНИЕМ ФИЗИКО-ИНФОРМИРОВАННОГО СВЕРТОЧНОГО АВТОЭНКОДЕРА

А. В. Кожевникова[1], М. А. Митрохин[2]

[1, 2] Пензенский государственный университет, Пенза, Россия
[1] alina.kozhevnikova28@mail.ru, [2] mmax83@mail.ru

Аннотация. *Актуальность и цели.* Туман и дым можно считать одними из наиболее фундаментальных явлений, вызывающих ухудшение видимости на изображениях, полученных в естественных условиях. Существует множество подходов к удалению тумана и дыма, и в большинстве из них был достигнут значительный прогресс. Поднимается проблема фильтрации затуманенных/задымленных изображений с использованием технологий нейронных сетей. *Материалы и методы.* Предложенный метод основан на физико-информированном сверточном автоэнкодере. Для достижения наилучших результатов в фильтрации изображений в функцию потерь был интегрирован метод черного канала. Модель обучалась на синтетически и естественно полученных парах чистых и затуманенных/задымленных изображений, сделанных в помещении и на улице. Проводилось сравнение эффективности работы разработанного метода с подходами AOD-net, CAP и Dehaze-Net. *Результаты.* Установлено, что разработанный в рамках исследования алгоритм способен фильтровать туман/дым даже самой высокой плотности. В среднем по качеству фильтрации метод не уступает крупномасштабным нейронным сетям. *Выводы.* Предлагаемый метод может использоваться как часть автоматической системы обработки изображений, работающей в естественных условиях под воздействием атмосферных явлений. По результатам исследования принято решение о доработке параметров модели в направлении устранения проблемы осветления участков сцены изображений с изображением неба.

Ключевые слова: искусственный интеллект, сверточные нейронные сети, автоэнкодер, физико-информированные нейронные сети, фильтрация затуманенных/задымленных изображений, теория черного канала

Для цитирования: Кожевникова А. В., Митрохин М. А. Фильтрация одиночных затуманенных/задымленных изображений с использованием физико-информированного сверточного автоэнкодера // Модели, системы, сети в экономике, технике, природе и обществе. 2024. № 1. С. 139–148. doi: 10.21685/2227-8486-2024-1-10

SINGLE IMAGE DEHAZING USING PHYSICS-INFORMED CONVOLUTIONAL AUTOENCODER

A.V. Kozhevnikova[1], M.A. Mitrokhin[2]

[1, 2] Penza State University, Penza, Russia
[1] alina.kozhevnikova28@mail.ru, [2] mmax83@mail.ru

© Кожевникова А. В., Митрохин М. А., 2024. Контент доступен по лицензии Creative Commons Attribution 4.0 License / This work is licensed under a Creative Commons Attribution 4.0 License.

Abstract. *Background.* Generally, haze can be considered to be one of the most fundamental phenomena causing image visibility degradation. Numerous haze removal approaches have been proposed and most of them have achieved significant progress. In this paper the problem of dehazing images using deep neural network technologies is raised. *Materials and methods.* Our method is based on a physics-informed convolutional autoencoder. To achieve the best results in image dehazing, the black channel prior was integrated into the loss function. The model was trained on synthetically obtained pairs of clean and hazy indoor and outdoor images. The efficiency of the developed method was compared with AOD-net, CAP and Dehaze-Net. *Results*. It is established that the algorithm developed as part of the study is able to dehaze dense images. On the average, the method copes well with its task and is not inferior to large-scale neural networks. *Conclusions.* It was decided to improve the parameters of the model, as well as make small changes to the learning process to eliminate the problem of image lightening having large celestial areas.

Keywords: artificial intelligence, convolutional neural network, autoencoder, physics-informed neural network, image dehazing, dark channel prior

For citation: Kozhevnikova A.V., Mitrokhin M.A. Single image dehazing using physics-informed convolutional autoencoder. *Modeli, sistemy, seti v ekonomike, tekhnike, prirode i obshchestve = Models, systems, networks in economics, technology, nature and society*. 2024;(1):139–148. (In Russ.). doi: 10.21685/2227-8486-2024-1-10

Введение

Получение качественных изображений является важным аспектом машинного зрения, от этого зависит эффективность распознавания и классификации объектов и явлений. Атмосферные и антропогенные явления, такие как туман и дым, заметно снижают качество решения задач. Удаление тумана с изображения является сложной задачей, потому что могут меняться цветовые параметры объектов на изображении. Для того чтобы решить данную проблему, требуются эффективные технологии фильтрации.

В настоящее время существует множество методов фильтрации туманов и задымлений, которые достигли значительных успехов в решении этой задачи. Их можно разделить на две группы: традиционные и созданные с использованием технологий глубоких нейронных сетей (DNN). Традиционные методы фильтрации чаще всего основаны на эмпирических посылках и опираются на методы преобразования яркости изображений или применение пространственной и частотной фильтрации. Основанные на DNN методы используют чаще всего сверточные нейронные сети или их комбинации с другими подходами.

Авторы статьи [1] предложили нелокальный алгоритм на основе методов машинного обучения для оценки характеристик «чистого» изображения, основанный на предположении, что цвета в изображениях без тумана хорошо аппроксимируются несколькими различными цветами, которые образуют плотные кластеры в пространстве RGB.

Статья [2] посвящена методу ослабления света, который был основан на следующем наблюдении: яркость и насыщенность пикселей на размытом изображении резко меняются вместе с изменением насыщенности задымленности или плотности тумана. Было выявлено, что на фотографиях, снятых на улице, большую роль играет атмосфера, более того, чем плотнее задымленности, тем сильнее влияют характеристики общего освещения. Таким образом, для фильтрации тумана с изображений предлагается метод предварительного

ослабления цвета путем создания линейной модели для моделирования глубины сцены размытого изображения.

Такие подходы обычно дают приемлемый результат только для некоторой ограниченной части начальных условий, таких как плотность тумана, общая освещенность и т.п.

В статье [3] авторы предложили сквозную модель фильтрации тумана, основанную на сверточной нейронной сети DehazeNet, она принимает размытое изображение в качестве входных данных и выводит его карту индексов цветопередачи, которая впоследствии используется для восстановления изображения без тумана с помощью модели рассеивания света в атмосфере.

Нейронная сеть AOD-Net [4] представляет собой модель, состоящую из двух частей: k-estimation module, включающего пять сверточных слоев для оценки функции реконструкции изображения, а также описания глубины и относительного уровня фоновой помехи; clean image module – для генерации улучшенного изображения, созданного из слоя поэлементного умножения изображения на функцию реконструкции изображения. Такой новый сквозной интерфейс позволяет легко интегрировать AOD-Net в другие модели DNN.

Статья [5] рассматривает возможности нейронных сетей типа «трансформер» для решения задачи удаления дымки с изображения. В результате эффективной интеграции функций трансформера и сверточной нейронной сети авторы улучшают производительность операций удаления дымки.

Тем не менее все нейросетевые подходы обладают высокой вычислительной емкостью и требуют наличия обширной обучающей базы изображений, обладают зависимостью качества итогового решения от качества обучающего набора данных. Для задачи фильтрации тумана и дыма требуется датасет, содержащий пары абсолютно идентичных изображений, отличающихся только наличием тумана. Создать подобный обучающий набор большого объема для использования в распространенных нейросетевых архитектурах, особенно глубоких, является практически сложной задачей.

Сократить требуемый объем обучающих данных и повысить качество итогового решения возможно при добавлении в процесс обучения нейронной сети априорной информации о физических процессах распространения света. Такой метод основывается на Physics-Informed Neural Network (PINN) [6,7]. Подобный тип нейронных сетей получил распространение при моделировании физических процессов в механике, электродинамике и других областях. Комбинирование сверточных нейронных технологий с физико-информированным подходом для решения задачи улучшения качества изображения является актуальной задачей.

Целью настоящей работы является разработка метода фильтрации тумана с изображений с помощью физико-информированного автоэнкодера с вычислением характеристик черного канала в функции потерь.

Теория черного канала

Теория черного канала, или dark channel prior (DCP) [8], представляет собой вывод из наблюдения о том, что в «чистых» изображениях хотя бы один цветовой канал близок к нулю. «Черный канал» такого изображения представлен следующей формулой:

$$J^{dark}(x) = \min_{c\in\{r,g,b\}} \left(\min_{y\in\Omega(x)} \left(J^c(y) \right) \right), \quad (1)$$

где J^c – цветовой канал изображения $\Omega(x)$ и локальный участок в точке x. Черный канал выводится из двух минимумов: $\min_{c\in\{r,g,b\}}$ – минимум каждого пикселя (рис. 1,*б*) и $\min_{y\in\Omega(x)}$ – фильтр минимума (рис. 1,*в*).

а) *б*) *в*)

Рис. 1. Вычисление черного канала «чистого» изображения

Свет на пути к наблюдателю частично поглощается молекулами воздуха и взвешенными в нем жидкими и твердыми частицами (каплями воды, кристаллами льда, пылинками и т.п.) и рассеивается на них. В нашей задаче рассматривается рассеивание света туманом или дымкой, поэтому атмосферу представим в виде системы взвешенных прозрачных частиц (капли воды и пыль). Благодаря рассеянию света некоторый слой атмосферы приобретает в любом направлении от источника определенную яркость, которая меньше, чем при отсутствии тумана или дымки. Визуально черный канал тем интенсивнее, чем выше плотность тумана [8] (рис. 2). Это обуславливается следующим: чем дальше сцена от наблюдателя, тем ниже интенсивность света, а значит, повышается количество темных объектов.

Рис. 2. Черный канал изображений с туманом

Предположив, что окружающий свет, выраженный в виде буквы «A», известен и передача небольшого участка изображения, обозначаемая $t(x)$, – константа, можно получить минимум цветового канала каждого пикселя, находящегося в туманной или задымленной области и предсказать коэффициент пропускания

$$\overline{t}(x) = 1 - \omega * \min_{c}\left(\min_{y \in \Omega(x)}\left(\frac{I^{c}(y)}{A^{c}}\right)\right), \quad (2)$$

где $\omega = 0{,}95$, сохраняет небольшую дымку для более естественного результата.

Архитектура и обучение нейронной сети

Физико-информированные нейронные сети – это класс моделей в машинном обучении, используемый для решения задач методами глубокого обучения с принципами физики. PINN разработаны для аппроксимации физических уравнений путем минимизации функции потерь. Это позволяет предсказывать и моделировать поведение с учетом физических принципов, что делает PINN наиболее востребованными там, где требуется учет физических законов [9].

Для решения проблемы фильтрации туманов и дыма предлагается использовать физико-информированный сверточный автоэнкодер, так как это позволяет избавиться от недостатков генеративных сверточных сетей, заключающихся в ненулевой вероятности появления новых, отсутствующих в исходной сцене, деталей изображения. Модель состоит из энкодера и декодера, каждый имеет четыре полносвязных сверточных слоя с функцией активации ReLU. Для ускорения обучения и получения более четкого изображения на выходе с минимальным количеством эпох была использована пакетная нормализация как для энкодера, так и для декодера (рис. 3).

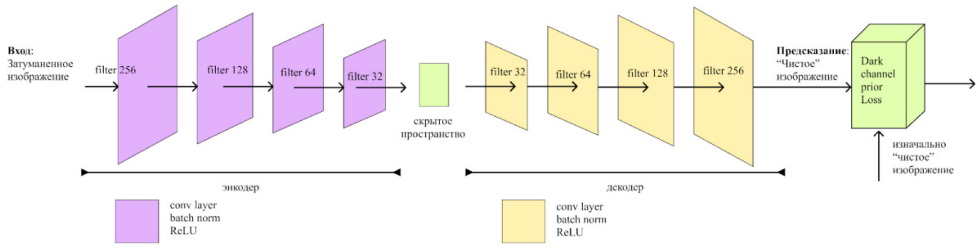

Рис. 3. Архитектура сверточного автоэнокдера

Поскольку довольно сложно получить совершенно одинаковую пару изображений, различающихся только содержанием тумана, на вход нейронной сети подавались синтетически полученные данные.

Обучение проводилось на 130 парах (затуманенные и «чистые») изображений, полученных в помещении (I-HAZE [11], содержащий 35 пар) и на улице (Dense-Haze CVPR 2019 [10], содержащий 55 пар и O-HAZE [12], содержащий 45 пар). Каждая фотография содержит цветовую диаграмму Макбета (Macbeth ColorChecker) [13] для проверки соответствия итоговых цветов сцены с

входными. Поскольку размеры изображений различались, было решено привести их к одной ширине и высоте – 512×512 пикселей. Далее на вход нейронной сети подавалось затуманенное изображение с тремя цветовыми каналами.

Для минимизации ошибок между результатом нейронной сети и «чистым» изображением была создана функция потерь на основе теории черного канала и среднеквадратичной ошибки (mean squared error):

$$\mathrm{MSE} = \mathrm{MSE}_{\mathrm{AE}} + \mathrm{MSE}_{\mathrm{DCP}}, \tag{3}$$

где

$$\mathrm{MSE}_{\mathrm{AE}} = \frac{1}{n}\sum_{i=1}^{n}\left(y_i - \overset{=}{y}_i\right)^2, \tag{4}$$

$$\mathrm{MSE}_{\mathrm{DCP}} = \frac{1}{n}\sum_{i=1}^{n}\left(\overset{=}{y}_i - J^{dark}(x)_i\right)^2. \tag{5}$$

В качестве алгоритма численной оптимизации использовался Adam (adaptive moment estimation) с начальной скоростью сходимости 0.004. Размер батча составлял 4.

Результаты эксперимента

На рис. 4,*а*,*е* изображены входные и выходные тренировочные пары изображений, которые были взяты из датасетов Dense-Haze CVPR 2019 [10], I-HAZE [11] и O-HAZE [12] и проведены через три различных метода фильтрации (DehazeNet [3], CAP [2], AOD-Net [4], наш метод), находящихся в свободном доступе.

а) *б*) *в*) *г*) *д*) *е*)

Рис. 4. Результаты фильтрации задымленных изображений различных методов: *а* – входное изображение; *б* – DehazeNet; *в* – CAP; *г* – AOD-Net; *д* – наш метод; *е* – чистое изображение

Можно заметить, что изображения, полученные в результате работы DehazeNet (рис. 4,*б*) и CAP (рис. 4,*в*), слишком темные, на CAP (рис. 4,*в*) также появляются зеленые артефакты. AOD-Net (рис. 4,*г*) хорошо справился с фильтрацией тумана низкой плотности, а более высокую ему не удалось отфильтровать так, чтобы были четко видны все объекты. Наш метод наиболее близок к оригинальному изображению, однако на фото с участками неба данные области слишком светлые, за счет чего фотографии высветляются.

Обычно для задач восстановления изображений используются метрики пикового отношения сигнала к шуму (PSNR) [15] и индекса структурного сходства (SSIM) [14].

PSNR определяет уровень искажений при сжатии и включает подсчет среднеквадратичной ошибки (MSE) [16]. Диапазон принимаемых значений – от 0 до 100, измеряется в логарифмической шкале в децибелах и представляет собой числовое значение, которое количественно определяет сходство между входными и выходными изображениями.

Дано два изображения f и g, оба имеют разрешение M×N, PSNR между f и g вычисляется по следующей формуле:

$$\mathrm{PSNR}(f,g)=10\log_{10}\left(\frac{255^2}{\mathrm{MSE}(f,g)}\right), \tag{6}$$

где

$$\mathrm{MSE}(f,g)=\frac{1}{\mathrm{MN}}\sum_{i=1}^{M}\sum_{j=1}^{N}\left(f_{ij}-g_{ij}\right)^2. \tag{7}$$

Чем выше значение PSNR, тем выше качество итогового изображения.

SSIM коррелирует с качеством и восприятием зрительной системы человека (цветовая модель HVS). Вместо использования традиционных методов суммирования ошибок SSIM моделирует искажение изображения как комбинацию трех факторов: потери корреляции, искажения яркости и искажения контрастности [17]. Данная метрика вычисляется по формуле

$$\mathrm{SSIM}(f,g)=l(f,g)c(f,g)s(f,g), \tag{8}$$

где

$$\begin{cases} l(f,g)=\dfrac{2\mu_f\mu_g+C_1}{\mu_f^2+\mu_g^2+C_1} \\ c(f,g)=\dfrac{2\sigma_f\sigma_g+C_2}{\sigma_f^2+\sigma_g^2+C_2}, \\ s(f,g)=\dfrac{\sigma_{fg}+C_3}{\sigma_f\sigma_g+C_3} \end{cases} \tag{9}$$

Первая часть формулы (9) – это функция, которая сравнивает яркость двух изображений. Она максимальна и равна 1, только когда $\mu_f=\mu_g$. Вторая часть сравнивает контрастность f и g, измеряется отклонением σ_f и σ_g и

принимает максимальное значение, когда $\sigma_f = \sigma_g$. Третья функция измеряет коэффициент корреляции между двумя изображениями.

Выходные данные каждого из вышеописанного метода были вручную проведены через каждую из метрик, и средние значения записаны в табл. 1.

Таблица 1

Количественные оценки с различных методов
на основе синтетических наборов данных

Метод	PSNR	SSIM
DehazeNet	27.87	0.336
CAP	28.10	0.504
AOD-Net	28.07	0.394
Наш метод	**28.75**	**0.413**

Данные табл. 1 отражают показатели PSNR и SSIM нашего метода и трех других.

Вывод

В работе предложен метод фильтрации задымленных или затуманенных изображений, построенный на основе физико-информированных сверточных нейронных сетей. Метод использует теорию черного канала (DCP) для оценки физических характеристик распространения света, учитываемых в функции потерь, что позволяет снизить требования к размеру обучающей выборки и ускорить процесс обучения.

Установлено, что модель лучше фильтрует изображения, сделанные в помещении, так как черный канал наиболее чувствителен к небесным участкам. Установлено, что разработанная модель хорошо справляется с фильтрацией туманов и задымлений различных плотностей, однако немного изменяет яркость сцены. Дальнейшим развитием предлагаемого метода является поиск способов устранения засветки ярких областей изображения.

Список литературы

1. Berman D., Avidan S., Treibitz T. Non-local image dehazing // Proceedings of the IEEE conference on computer vision and pattern recognition. 2016. P. 1674–1682.
2. Zhu Q., Mai J., Shao L. A fast single image haze removal algorithm using color attenuation prior // IEEE Transactions on Image Processing. 2015. Vol. 24. P. 3522–3533.
3. Cai B., Xu X., Jia K. [et al.]. Dehazenet: An end-to-end system for single image haze removal // IEEE Transactions on Image Processing. 2016. Vol. 25 (11). P. 5187–5198.
4. Li B., Peng X., Wang Z. [et al.]. AOD-Net: All-in-one dehazing network // Proceedings of the IEEE International Conference on Computer Vision. 2017. P. 4770–4778.
5. Guo C., Yan Q., Anwar S. [et al.]. Image Dehazing Transformer with Transmission-Aware 3D Position Embedding // IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). 2022. P. 5802–5810. doi: 10.1109/CVPR52688.2022.00572
6. Thakur S., Raissi M., Mitra H., Ardekani A. Temporal Consistency Loss for Physics-Informed Neural Networks. 2023.
7. McGowan E., Gawade V., Guo W. A Physics-Informed Convolutional Neural Network with Custom Loss Functions for Porosity Prediction in Laser Metal Deposition. 2022.
8. He K., Sun J., Tang X. Single Image Haze Removal Using Dark Channel Prior Fellow // IEEE Transactions on Pattern Analysis and Machine Intelligence.

9. Cuomo S., di Cola V. S., Giampaolo F. [et al.]. Scientific Machine Learning through Physics-Informed Neural Networks: Where we are and What's next // Journal of Scientific Computing. 2022. Vol. 92 (88).
10. Dense-Haze CVPR 2019. URL: https://www.kaggle.com/datasets/rajat95gupta/hazing-images-dataset-cvpr-2019 (дата обращения: 20.12.2023).
11. I-HAZE: a dehazing benchmark with real hazy and haze-free indoor images. URL: https://data.vision.ee.ethz.ch/cvl/ntire18//i-haze/ (дата обращения: 20.12.2023).
12. O-HAZE: a dehazing benchmark with real hazy and haze-free outdoor images. URL: https://data.vision.ee.ethz.ch/cvl/ntire18//o-haze/ (дата обращения: 20.12.2023).
13. ColorChecker (Macbeth) Chart. URL: https://poynton.ca/notes/color/GretagMacbeth-ColorChecker.html (дата обращения: 20.12.2023).
14. Nilsson J., Akenine-Möller T. Understanding SSIM : preprint. 2020.
15. Faragallah O. S., El-Hoseny H., El-Shafai W. [et al.]. A Comprehensive Survey Analysis for Present Solutions of Medical Image Fusion and Future Directions // IEEE Access. Vol. 9. P. 11358–11371.
16. Краткий обзор методов измерения сходства между изображениями. URL: https://skine.ru/articles/258694/ (дата обращения: 20.12.2023).
17. Horé A., Ziou D. Image quality metrics: PSNR vs. SSIM // IEEE Computer Society. 2010. P. 2366–2369.

References

1. Berman D., Avidan S., Treibitz T. Non-local image dehazing. *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2016:1674–1682.
2. Zhu Q., Mai J., Shao L. A fast single image haze removal algorithm using color attenuation prior. *IEEE Transactions on Image Processing*. 2015;24:3522–3533.
3. Cai B., Xu X., Jia K. et al. Dehazenet: An end-to-end system for single image haze removal. *IEEE Transactions on Image Processing*. 2016;25(11):5187–5198.
4. Li B., Peng X., Wang Z. et al. AOD-Net: All-in-one dehazing network. *Proceedings of the IEEE International Conference on Computer Vision*. 2017:4770–4778.
5. Guo C., Yan Q., Anwar S. et al. Image Dehazing Transformer with Transmission-Aware 3D Position Embedding. *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. 2022:5802–5810. doi: 10.1109/CVPR52688.2022.00572
6. Thakur S., Raissi M., Mitra H., Ardekani A. *Temporal Consistency Loss for Physics-Informed Neural Networks*. 2023.
7. McGowan E., Gawade V., Guo W. *A Physics-Informed Convolutional Neural Network with Custom Loss Functions for Porosity Prediction in Laser Metal Deposition*. 2022.
8. He K., Sun J., Tang X. Single Image Haze Removal Using Dark Channel Prior. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 2011;33(12):2341–2353.
9. Cuomo S., di Cola V.S., Giampaolo F. et al. Scientific Machine Learning through Physics-Informed Neural Networks: Where we are and What's next. *Journal of Scientific Computing*. 2022;92(88).
10. *Dense-Haze CVPR 2019*. Available at: https://www.kaggle.com/datasets/rajat95gupta/hazing-images-dataset-cvpr-2019 (accessed 20.12.2023).
11. *I-HAZE: a dehazing benchmark with real hazy and haze-free indoor images*. Available at: https://data.vision.ee.ethz.ch/cvl/ntire18//i-haze/ (accessed 20.12.2023).
12. *O-HAZE: a dehazing benchmark with real hazy and haze-free outdoor images*. Available at: https://data.vision.ee.ethz.ch/cvl/ntire18//o-haze/ (accessed 20.12.2023).
13. *ColorChecker (Macbeth) Chart*. Available at: https://poynton.ca/notes/color/GretagMacbeth-ColorChecker.html (accessed 20.12.2023).
14. Nilsson J., Akenine-Möller T. *Understanding SSIM: preprint*. 2020.
15. Faragallah O.S., El-Hoseny H., El-Shafai W. et al. A Comprehensive Survey Analysis for Present Solutions of Medical Image Fusion and Future Directions. *IEEE Access*. 2021;9:11358–11371.

16. *Kratkiy obzor metodov izmereniya skhodstva mezhdu izobrazheniyami = A brief overview of methods for measuring similarity between images*. (In Russ.). Available at: https://skine.ru/articles/258694/ (accessed 20.12.2023).
17. Horé A., Ziou D. Image quality metrics: PSNR vs. SSIM. *IEEE Computer Society*. 2010:2366–2369.

Информация об авторах / Information about the authors

Алина Вадимовна Кожевникова
студентка кафедры вычислительной техники,
Пензенский государственный университет
(Россия, г. Пенза, ул. Красная, 40)
E-mail: alina.kozhevnikova28@mail.ru

Alina V. Kozhevnikova
Student of the sub-department of computer engineering,
Penza State University
(40 Krasnaya street, Penza, Russia)

Максим Александрович Митрохин
доктор технических наук, доцент,
заведующий кафедрой вычислительной техники,
Пензенский государственный университет
(Россия, г. Пенза, ул. Красная, 40)
E-mail: mmax83@mail.ru

Maksim A. Mitrokhin
Doctor of technical sciences,
associate professor,
head of the sub-department of computer engineering,
Penza State University
(40 Krasnaya street, Penza, Russia)

Авторы заявляют об отсутствии конфликта интересов /
The authors declare no conflicts of interests.

Поступила в редакцию/Received 26.10.2023
Поступила после рецензирования/Revised 02.12.2023
Принята к публикации/Accepted 04.03.2024