

# РАЗРАБОТКА ПРОГРАММНЫХ АЛГОРИТМОВ ОБРАБОТКИ БИНАРНЫХ ИЗОБРАЖЕНИЙ АВТОМАТИЗИРОВАННОЙ СИСТЕМЫ РАСПОЗНАВАНИЯ СИМВОЛЬНЫХ ДАННЫХ, ПРЕДСТАВЛЕННЫХ В РАСТРОВОМ ФОРМАТЕ

Мостовщиков Д.Н.<sup>1</sup>, Могилатов Р.К.<sup>2</sup>, Усов А.Е.<sup>3</sup>

<sup>1</sup>Мостовщиков Дмитрий Николаевич – старший системный архитектор,  
Li9, Inc., г. Феникс;

<sup>2</sup>Могилатов Роман Константинович - технический руководитель,  
Python SoftServe Inc., г. Роли;

<sup>3</sup>Усов Алексей Евгеньевич – ведущий системный архитектор,  
Li9, Inc., г. Феникс,  
Соединенные Штаты Америки

**Аннотация:** проведен анализ современных программных алгоритмов обработки бинарных изображений с целью распознавания символьных данных, представленных в растровом формате, которые характеризуются низким разрешением и высоким уровнем шумов. Показано, что на основе морфологических методов обработки бинарных изображений можно существенно уменьшить время машинного анализа входных данных и нагрузку на вычислительный ресурс соответствующей аппаратно-программной платформы. В результате проведенного анализа разработана математическая модель, которая позволяет обобщить методику построения морфологических алгоритмов восстановления бинарных изображений с целью дальнейшего распознавания символьных данных, представленных в растровом формате. Эффективность применения данной методики оценивалась на основе целевых функций точности восстановления и времени обработки входных данных, а также уровня нагрузки на вычислительный ресурс системы машинного анализа.

**Ключевые слова:** растровое представление графической информации, бинарные изображения, предварительная обработка изображения, морфологические методы, дилатация, эрозия, целевые функции.

## DEVELOPMENT OF SOFTWARE ALGORITHMS FOR PROCESSING BINARY IMAGES OF AUTOMATED SYSTEM FOR RECOGNIZING CHARACTER DATA PRESENTED IN RASTER FORMAT

Mostovshchikov D.N.<sup>1</sup>, Mogylatov R.K.<sup>2</sup>, Usov A.Ye.<sup>3</sup>

<sup>1</sup>Mostovshchikov Dmitriy Nikolaevich – Senior Systems Architect,  
LI9, INC., PHOENIX;

<sup>2</sup>Mogylatov Roman Kostantinovich – Technical Leader,  
PYTHON SOFTSERVE INC., RALEIGH;

<sup>3</sup>Usov Aleksey Yevgenyevich – Lead Systems Architect,  
LI9, INC., PHOENIX,  
UNITED STATES OF AMERICA

**Abstract:** an analysis of modern software algorithms for processing binary low resolution and noisy images was conducted in order to recognize symbolic data presented in raster format. It is shown that morphological methods of processing binary images allow to significantly reduce the time of machine analysis of input and the load of the computing resource of the corresponding hardware and software platform. As a result of the analysis, a mathematical model has been developed. It allows generalizing the methodology for building of morphological algorithms for recovering binary images in order to further recognize symbolic data presented in raster format. The effectiveness of the application of this technique was estimated on the basis of the objective functions of the reconstruction accuracy and the processing time of the input data, as well as the level of load on the computational resource of the raster image analysis system.

**Keywords:** raster representation of graphic information, binary images, image preprocessing, morphological methods, dilatation, erosion, target functions.

УДК 004.93'12

### Введение

В связи с глобальной цифровизацией в области современных информационных технологий в значительной степени расширился инструментарий и выросла эффективность средств хранения, передачи и обработки данных. В свою очередь, это привело к появлению концепций глубинного анализа и формирования на основе больших объемов данных баз знаний. Основным заданием при этом является построение системы машинного анализа цифровых данных, в частности, задачи выделения и классификации визуальных объектов при обработке массивов графических данных. Стандартные подходы, которые применяются с целью увеличения точности анализа графических данных, включают в себя следующие группы:

- увеличение качества фото-регистрации и сканирования аналоговых источников графических данных [1, 2];
- использование при оцифровке входных данных методик с адаптивным порогом [3, 4] и экспертной оценки качества цифровых данных;

- хранение цифровых данных с высокими показателями разрешающей способности и динамического диапазона [5, 6], а также использование адекватной цветовой схемы;
- использование при восстановлении входных данных моделей распределения статистического шума и нейросетевых алгоритмов предварительной обработки [7, 8];
- применение программных и нейросетевых алгоритмов сегментирования матриц изображений [9, 10];
- применение программных и нейросетевых алгоритмов при выделении и классификации визуальных объектов [11, 12].

Соответствующие подходы оптимизации машинного анализа зачастую можно отнести к методам экстенсивного развития — в ряде случаев они приводят к существенному усложнению и увеличению себестоимости аппаратной платформы системы машинного анализа. Помимо необходимости модернизации подсистемы регистрации данных, это связано с увеличением нагрузки на вычислительный ресурс, информационное хранилище и пропускающую способность информационных каналов [13-16]. Оценка роста нагрузки может быть проведена на уровне обработки, передачи и хранения данных, объем которых экспоненциально растет с ростом качества и объема информации, которая подлежит анализу.

В связи с этим на сегодняшний день можно отметить тенденцию снижения качества графических данных сканирования и предварительной обработки текста и схем, что приводит к большому количеству ошибок на этапе распознавания данных методами машинного анализа. В значительной степени это связано с тем, что в данном случае подразумевается работа с большими объемами графических данных, базовые элементы которых (символьные элементы текста, элементы схем и т.п.) характеризуются низкой вариативностью, что значительно упрощает их распознавание даже при низком качестве фото-регистрации. Кроме того, данный класс графических данных представляет собой набор монохромных изображений, что дает возможность их отображения в виде бинарных изображений и, таким образом, существенно уменьшает объем данных на этапе оцифровки.

Рассмотрим пример обработки аналогового образца текстовых данных, который переводится в цифровую форму и классифицируется на уровне выделения символьного ряда системой машинного анализа (рис. 1).

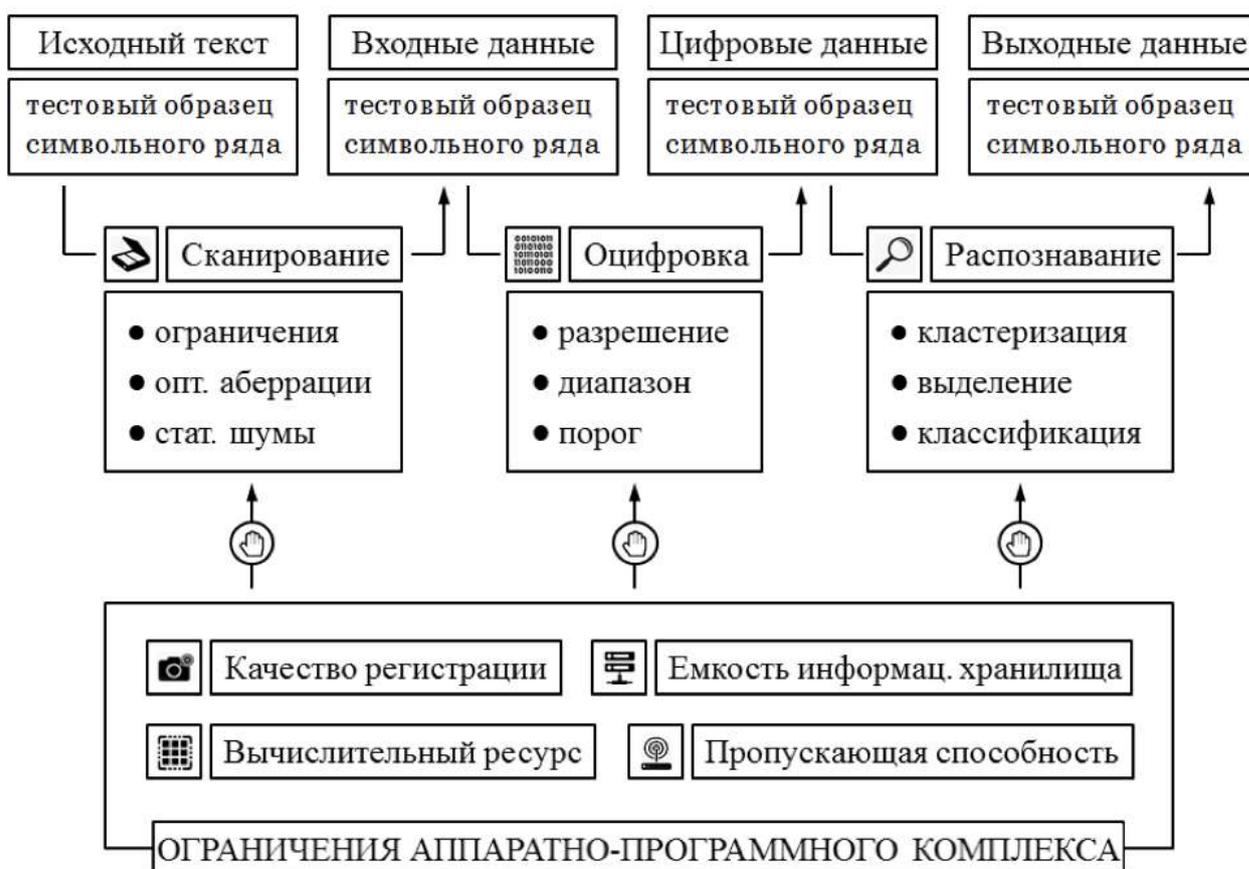


Рис. 1. Диаграмма выявления источников ошибок при распознавании символьных рядов

На этапе сканирования графические данные регистрируются в низком качестве с искажениями, связанными со статистическими шумами и оптическими aberrациями, а также отсутствием унифицированных требований по настройке параметров яркости и контраста. Далее, на этапе оцифровки и перевода в бинарный формат, графические данные дополнительно искажаются в соответствии с минимальным динамическим диапазоном и низким разрешением матрицы изображения, а также использованием стандартных пороговых алгоритмов, без учета анализа особенностей изображения и применения комплексных алгоритмов с адаптивным порогом. При распознавании символьного ряда в восстановленном тексте появляются ошибки, для исключения которых необходимо использовать ресурсоемкие программные и нейросетевые алгоритмы машинного анализа.

Таким образом, комплексность задачи распознавания символьных данных и оцифрованных монохромных схем связана с ограничениями, которые накладываются на вычислительный ресурс и инфраструктуру передачи данных системы машинного анализа, а также максимальное время обработки входных данных. При этом данный класс графических данных составляет значительную часть общего объема данных, который подлежит введению в цифровой документооборот и обработке на уровне машинного анализа с целью формирования баз знаний. В их числе материалы большой исторической значимости, источники которых были утеряны, а также документация, которая представляет высокую экономическую и стратегическую значимость (документация, связанная с производством, энергетикой, оборонным комплексом, хранением токсичных отходов и т.д.), что указывает на **высокую актуальность** данного класса задач.

При анализе **современных исследований и публикаций в профильных изданиях**, посвященных данной проблеме, были рассмотрены методы оптимизации алгоритмов предварительной обработки [5-8], выделения и классификации визуальных объектов [9-12], а также соответствующей аппаратной платформы [13-16]. Был указан приоритет использования алгоритмов, которые базируются на морфологических методах предварительной обработки и восстановления бинарных искажений [17-21]. При этом для построения алгоритмов предварительной обработки с целью эффективного выделения символьного ряда и элементов схем, представленных в растровом формате, необходимо построить целостную математическую модель обработки графических данных морфологическими методами с минимальной нагрузкой на вычислительный ресурс и минимальным временем обработки данных, что выделяется как **нерешенная часть** вопроса в рамках общего исследования. **Целью работы**, таким образом, стала разработка комплексной методики предварительной обработки бинарных изображений морфологическими методами, эффективность которой определяется через целевые функции восстановления исходных данных, времени обработки и нагрузки на вычислительный ресурс аппаратно-программной платформы.

#### Основная часть исследования

Базовая модель применения морфологических методов предварительной обработки и восстановления бинарного изображения включает в себя представление двумерной матрицы исходного изображения  $\mathbf{G}_I: \{g_i(x_i, y_i)\}$  размерности  $X_I \times Y_I$  и двумерной матрицы графического примитива  $\mathbf{G}_P: \{g_p(x_p, y_p)\}$  размерности  $X_P \times Y_P$ , а также набор логических операторов. Исходное изображение и графический примитив, представленные в бинарном формате, могут быть формализованы на основе следующей системы уравнений:

$$\begin{cases} g_i(x_i, y_i) = \begin{cases} 0 & \text{для всех } \begin{cases} x_i \in [1; X_I] \\ y_i \in [1; Y_I] \end{cases} \\ 1 & \text{для всех } \begin{cases} x_p \in [1; X_P] \\ y_p \in [1; Y_P] \end{cases} \end{cases}, \text{ где } \begin{cases} X_I \geq X_P \\ Y_I \geq Y_P \end{cases}. \end{cases} \quad (1)$$

Базовый набор операторов при этом включает бинарные операторы конъюнкции, дизъюнкции, импликации и эквивалентности. Математическое представление применения данных операторов при работе с матрицами бинарных изображений, может быть формализовано следующим образом:

$$\begin{cases} g_i(x_i, y_i) \vee g_p(x_p, y_p) = \begin{cases} 0 & \text{при } \begin{cases} g_i(x_i, y_i) = 0 \\ g_p(x_p, y_p) = 0 \end{cases} \\ 1 & \text{при } \begin{cases} g_i(x_i, y_i) = 1 \\ g_p(x_p, y_p) = 1 \end{cases} \end{cases} \\ g_i(x_i, y_i) \wedge g_p(x_p, y_p) = \begin{cases} 0 & \text{при } \begin{cases} g_i(x_i, y_i) = 0 \\ g_p(x_p, y_p) = 0 \end{cases} \\ 1 & \text{при } \begin{cases} g_i(x_i, y_i) = 1 \\ g_p(x_p, y_p) = 1 \end{cases} \end{cases} \\ g_i(x_i, y_i) \leftrightarrow g_p(x_p, y_p) = \begin{cases} 0 & \text{при } g_i(x_i, y_i) \neq g_p(x_p, y_p) \\ 1 & \text{при } g_i(x_i, y_i) = g_p(x_p, y_p) \end{cases} \\ g_i(x_i, y_i) \Rightarrow g_p(x_p, y_p) = \begin{cases} 0 & \text{при } g_i(x_i, y_i) < g_p(x_p, y_p) \\ 1 & \text{при } g_i(x_i, y_i) \geq g_p(x_p, y_p) \end{cases} \end{cases}, \quad (2)$$

причем каждая логическая операция рассчитывается для всех позиций элементов матрицы  $\mathbf{G}_P$ , которая накладывается на матрицу  $\mathbf{G}_I$  с постепенным сдвигом в рамках сканирования матрицей графического примитива матрицы изображения, что на математическом уровне может быть выражено как:

$$\begin{cases} \begin{cases} x_i \in [x; x + X_P] \\ y_i \in [y; y + Y_P] \\ x_p \in [1; X_P] \\ y_p \in [1; Y_P] \end{cases}, \text{ где } \begin{cases} x \in [0; X_I - X_P] \\ y \in [0; Y_I - Y_P] \end{cases}. \end{cases} \quad (3)$$

Помимо набора бинарных операторов, математическая модель включает в себя унарный оператор отрицания, который представляется в следующем виде:

$$\neg g_I(x_I, y_I) = \begin{cases} 0 & \text{при } g_I(x_I, y_I) = 1 \\ 1 & \text{при } g_I(x_I, y_I) = 0 \end{cases} \text{ для } \begin{cases} x_I \in [1; X_I] \\ y_I \in [1; Y_I] \end{cases}, \quad (4)$$

на основе чего может быть получен бинарный оператор антиэквивалентности, как отрицание оператора эквивалентности, а также оператор разницы, который определяет элементы множества  $G_I$ , не входящие в множество  $G_P$ , как отрицание оператора импликации:

$$g_I(x_I, y_I) \oplus g_P(x_P, y_P) = \neg(g_I(x_I, y_I) \leftrightarrow g_P(x_P, y_P)). \quad (5)$$

$$g_I(x_I, y_I) / g_P(x_P, y_P) = \neg(g_I(x_I, y_I) \Rightarrow g_P(x_P, y_P)). \quad (6)$$

На основе дополнительных логических операторов появляется возможность варьировать математическое представление основного набора, что в дальнейшем позволит увеличить скорость обработки данных и уменьшить нагрузку на вычислительный ресурс аппаратно-программной платформы системы машинного анализа.

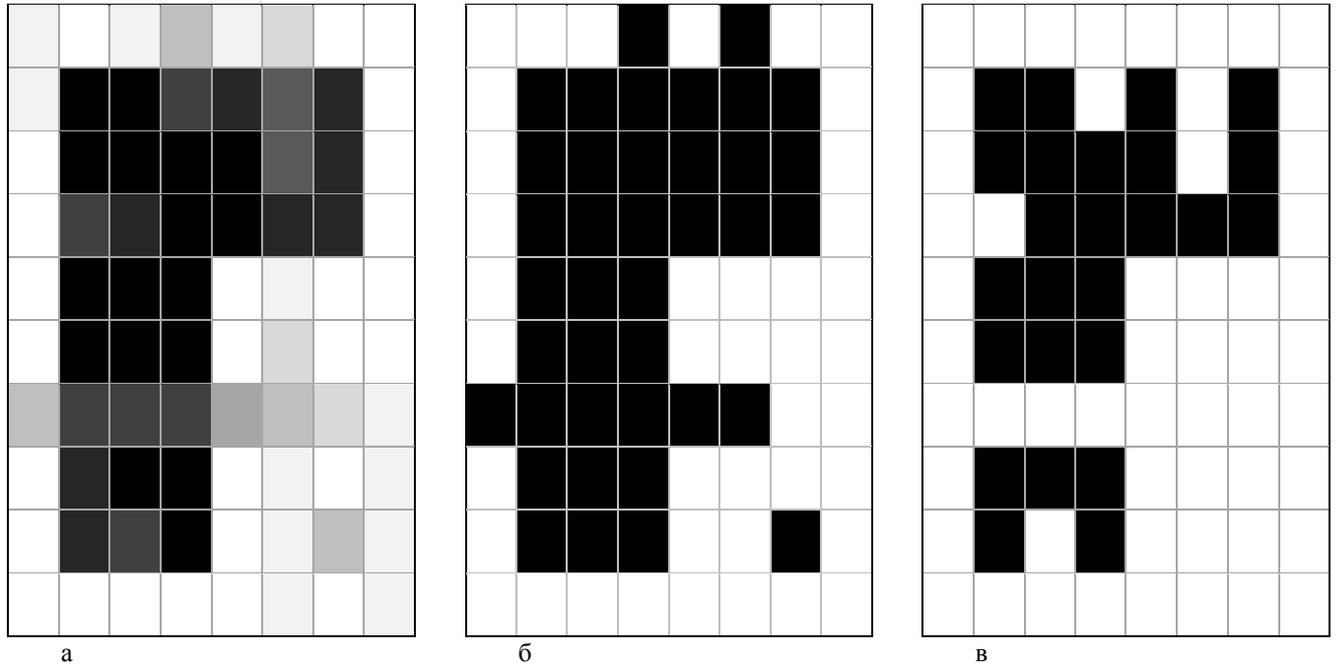
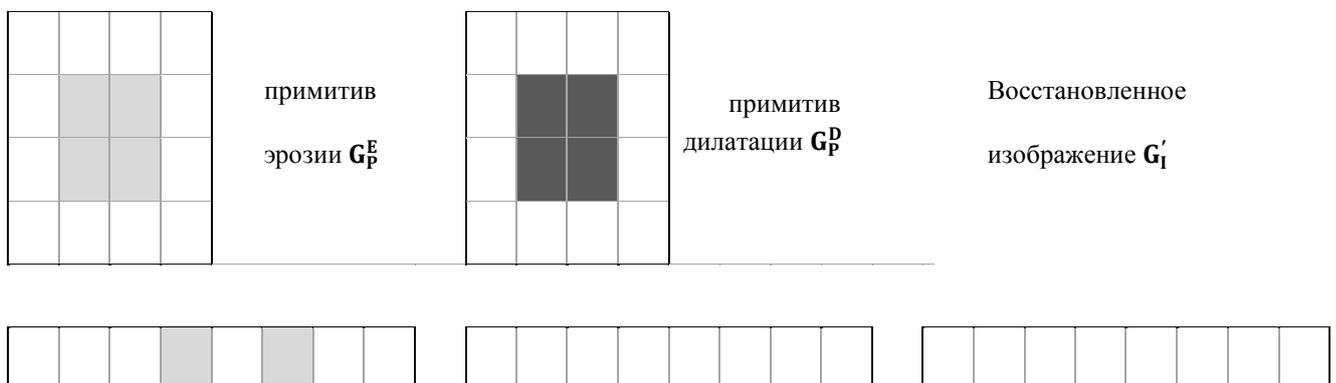


Рис. 2. Исходное изображение (а), искаженное в связи с низким (б) и высоким пороговым (в) значением

Рассмотрим пример морфологической обработки бинарного изображения. Пусть фрагмент текста (рис. 2-а) искажен при оцифровке. В случае, представленном на рис. 2-б, матрица изображения включает дополнительные элементы (пиксели), а в случае, представленном на рис. 2-в, на изображении символа присутствуют разрывы. На основе логических операторов можно построить процедуры дилатации, как множества перемещений примитива  $G_P^D$  по матрице изображения  $G_I$ , при которых примитивы совпадают, по меньшей мере, в одном элементе, и процедуры эрозии, как множества перемещений примитива  $G_P^E$  по матрице изображения  $G_I$ , при которых оба примитива полностью совпадают.

В случае наличия шумов как лишних элементов к матрице изображения последовательно применяется процедура эрозии (рис. 3-а), которая удаляет данные элементы, но также и удаляет граничные элементы символа, а далее процедура дилатации (рис. 3-б), которая восстанавливает толщину структурных элементов (рис. 3-в). Как можно видеть, исходное изображение при этом полностью восстанавливается (рис. 3-в полностью соответствует рис. 2-а).



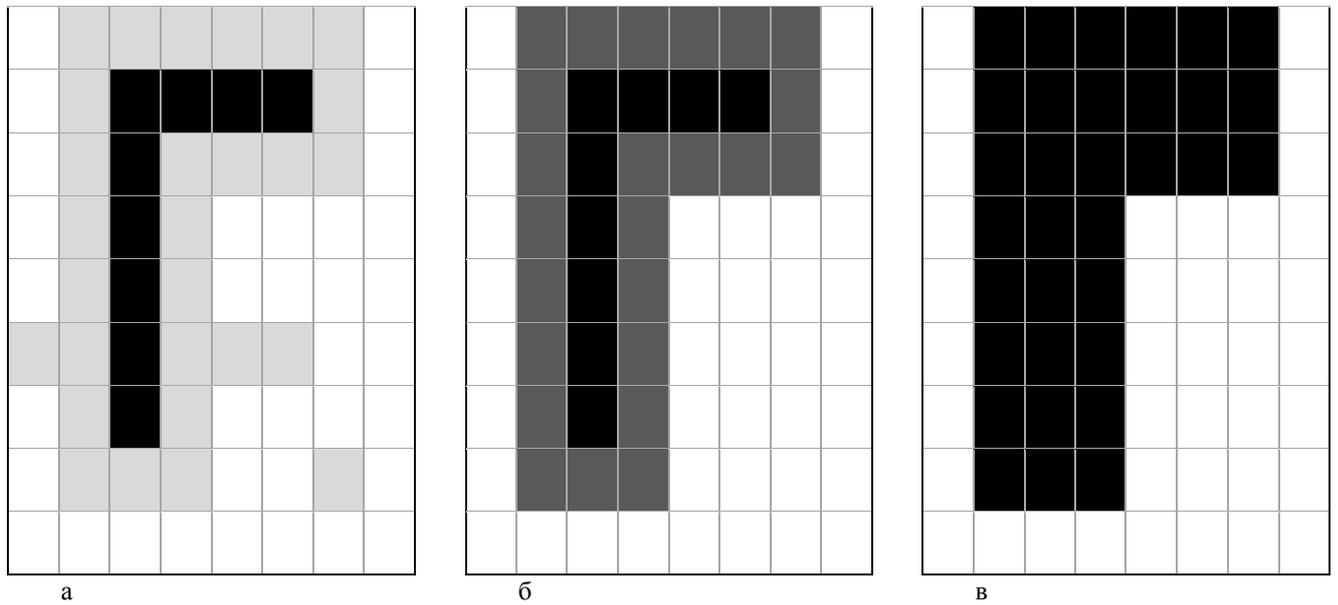


Рис. 3. Алгоритм последовательного применения процедуры эрозии (а) и дилатации (б) с целью удаления из матрицы изображения шумов (в)

Аналогично, при наличии разрывов в структурных элементах изображения (рис. 2-в) следует использовать указанные процедуры в обратном порядке. За счет применения процедуры дилатации устраняются разрывы (рис. 4-а), но также и увеличивается толщина структурных элементов, а далее при помощи процедуры эрозии толщина элементов возвращается к исходной (рис. 4-б). Таким образом, в данном случае также исходное изображение, искаженное полностью, восстанавливается при помощи морфологических методов (рис. 4-в полностью соответствует рис. 2-а).

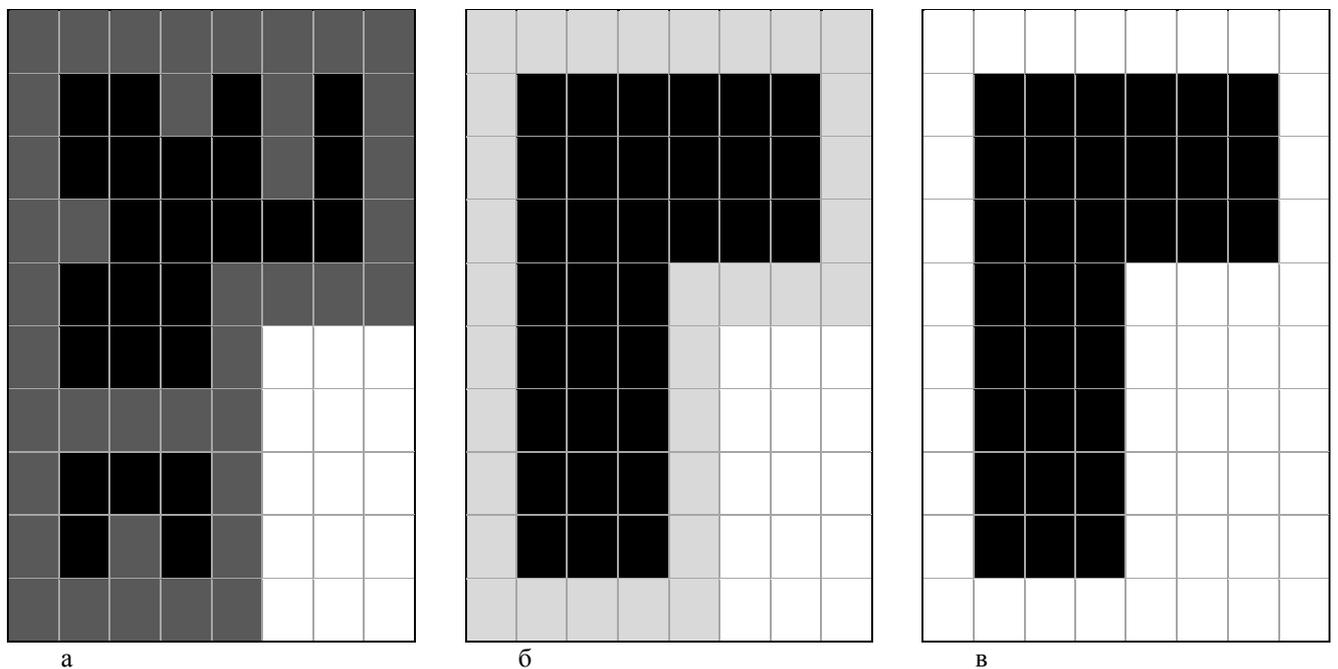
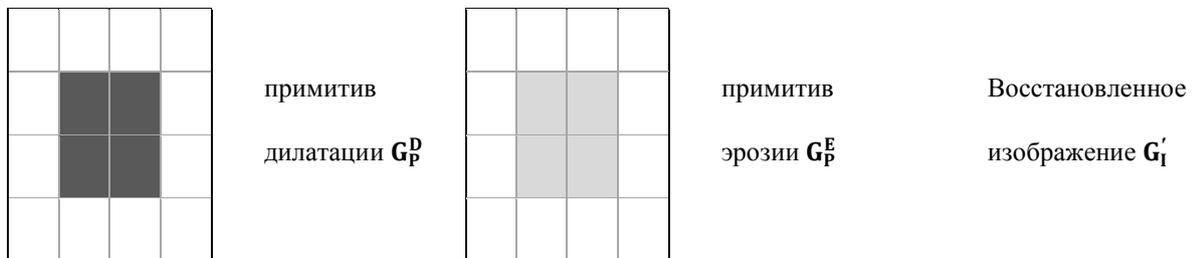


Рис. 4. Алгоритм последовательного применения процедуры дилатации (а) и эрозии (б) с целью устранения на матрице изображения разрывов (в)

Настройка параметров и оптимизация алгоритма восстановления изображения морфологическими алгоритмами в соответствии с поставленной задачей проводится через соотнесение элементов матрицы исходного изображения  $\mathbf{G}_I: \{g_I(x_I, y_I)\}$  и восстановленного изображения  $\mathbf{G}'_I: \{g'_I(x_I, y_I)\}$ . Это позволяет построить целевую функцию восстановления изображения как:

$$\Delta F_A = \sum_{y_I=1}^{Y_I} \sum_{x_I}^{X_I} |g'_I(x_I, y_I) - g_I(x_I, y_I)|. \quad (7)$$

Таким образом, задача оптимизации применения морфологических методов может быть решена через вычисление глобального минимума целевой функции  $\Delta F_A$  при ограничениях на время обработки данных и нагрузку на вычислительный ресурс. Аргументами целевой функции выступают графические примитивы дилатации и эрозии, выражение данных функций через логические операторы и последовательность выполнения соответствующих процедур.

#### **Выводы**

В результате проведенного исследования были проанализированы современные подходы, применяемые при построении алгоритмов обработки монохромных изображений с целью распознавания символьных данных, представленных в растровом формате. Показано, что на основе морфологических методов можно существенно уменьшить время машинного анализа графических данных и нагрузку на вычислительный ресурс при восстановлении матрицы изображения с низким разрешением и высоким уровнем шумов. В результате проведенного анализа разработана математическая модель, которая включает в себя набор бинарных логических операторов, матрицы исходного изображения, матрицы искаженного изображения, матрицы восстановленного изображения и матрицы графических примитивов операций дилатации и эрозии. Оптимизацию морфологических алгоритмов было предложено проводить через поиск локального минимума целевой функции для сопоставления элементов матриц исходного и восстановленного изображения, в условиях ограничений на время обработки данных и нагрузку на вычислительный ресурс системы машинного анализа.

#### **Список литературы / References**

1. *Zou J., Li Z., Guo Z. & Hong D. (2019). Super-Resolution Reconstruction of Images Based on Microarray Camera. Computers, Materials & Continua, 60 (1), 163–177. doi: 10.32604/cmc.2019.05795.*
2. *Rowcliffe J.M., Carbone C., Jansen P.A., Kays R. & Kranstauber B. (2011). Quantifying the sensitivity of camera traps: an adapted distance sampling approach. Methods in Ecology and Evolution, 2 (5). 464–476. doi: 10.1111/j.2041-210x.2011.00094.*
3. *Hazra T.K., Mazumder A., Maity D. A New Perspective of Wavelet Based Image Denoising Using Different Wavelet Thresholding. (2016). International Journal of Science and Research (IJSR). 5 (8). 1504-1509. https://doi.org/10.21275/art20161212*
4. *Akram J. (2018). An application of waveform denoising for microseismic data using polarization-linearity and time-frequency thresholding. Geophysical Prospecting. 66 (5), 872–893. https://doi.org/10.1111/1365-2478.12597.*
5. *Petro A. & Sbert C. (2013). Selective Contrast Adjustment by Poisson Equation. Image Processing On Line, 3, 208-222. doi:10.5201/ipol.2013.41.*
6. *Karim T. & Tasneem T. (2014). Analytical Adjustment of Image Contrast. International Journal of Computer Applications. 98 (20), 44-49. doi:10.5120/17303-7794.*
7. *Zhang C., Liu W. & Xing W. (2018). Color image enhancement based on local spatial homomorphic filtering and gradient domain variance guided image filtering. Journal of Electronic Imaging, 27(06). 1. doi:10.1117/1.jei.27.6.063026.*
8. *Gnouma M., Ladjaillia A., Ejbali R., Zaied M. (2018). Stacked sparse autoencoder and history of binary motion image for human activity recognition. Multimedia Tools and Applications. 78 (2), 2157-2179. doi:10.1007/s11042-018-6273-1.*
9. *Arbelaez P., Maire M., Fowlkes C., Malik J. (May 2011). "Contour Detection and Hierarchical Image Segmentation" (PDF). IEEE Transactions on Pattern Analysis and Machine Intelligence. 33 (5): 898–916.*
10. *Yuan Jiangye, Gleason Shaun S., Cheriyadat Anil M. (2013). "Systematic benchmarking of aerial image segmentation". IEEE Geoscience and Remote Sensing Letters. 10 (6): 1527–1531.*
11. *Adam B., Zaman F., Yassin I., Abidin H. & Rizman Z. (2018). Performance evaluation of faster R-CNN on GPU for object detection. Journal of Fundamental and Applied Sciences, 9 (3S), 909.*
12. *Wei X., Xie C., Wu J. & Shen C. (2018). Mask-CNN: Localizing parts and selecting descriptors for fine-grained bird species categorization. Pattern Recognition, 76, 704-714.*
13. *Dimitrov V. (2016). Cloud Programming Models (MapReduce). Encyclopedia of Cloud Computing, 596-608. doi:10.1002/9781118821930.ch49.*
14. *Gupta H. Load Balancing In Cloud Computing. (2017). International Journal of Recent Trends in Engineering and Research, 3(3), 260-267. doi:10.23883/ijrter.2017.3076.uimcu*
15. *Takabi H., Joshi J.B.D. and Ahn G.-J. "Security and privacy challenges in cloud computing environments," IEEE Security & Privacy, vol. 8, no. 6, pp. 24–31, 2010.*

16. *Petrov V.V., Zichun Le., Kryuchyn A.A., Shanoylo S.M., Fu M., Beliak Ie.V., Manko D.Yu., Lapchuk A.S., Morozov E.M.* Long-term storage of digital information July 2018 DOI: 10.15407/Akademperiodyka. 360.148ISBN: 9789663603605.
17. *Vincent L. & Heijmans H.* (2018). Graph Morphology in Image Analysis. *Mathematical Morphology in Image Processing*, 170-203. doi:10.1201/9781482277234-6.
18. *Sadi A., EL Chakik A. & Elmoataz A.* (2014). Nonlocal PDEs Morphology on Graph: A Generalized Shock Operators on Graph. *Lecture Notes in Computer Science*, 409–416. [https://doi.org/10.1007/978-3-319-07998-1\\_47](https://doi.org/10.1007/978-3-319-07998-1_47).
19. *Roerdink J.B.* (2018). Mathematical Morphology with Noncommutative Symmetry Groups. *Mathematical Morphology in Image Processing*, 205-254. doi:10.1201/9781482277234-7.
20. *Serra J.* Introduction to Mathematical Morphology. (2017). *Image Processing and Mathematical Morphology*, 25–34. <https://doi.org/10.1201/9781420089448-5>.
21. *Schmitt M.* (2018). On Two Inverse Problems in Mathematical Morphology. *Mathematical Morphology in Image Processing*, 151-169. doi:10.1201/9781482277234-5.