

# МЕТОДИКА КОМПЬЮТЕРНОГО ВЫЯВЛЕНИЯ ТЕКСТОВЫХ ОБЛАСТЕЙ НА ИЗОБРАЖЕНИИ БИРОК, ВЫПОЛНЕННЫХ ЛАЗЕРНОЙ ГРАВИРОВКОЙ НА МЕТАЛЛЕ

**В.Я. Столяров, В.А. Тихомиров**

*Комсомольский-на-Амуре государственный университет, г. Комсомольск-на-Амуре*

В наши дни обработка изображений является важным направлением применения современной вычислительной техники. Известны следующие задачи обработки изображений такие как: фильтрация и восстановление изображений, сегментация изображений, как средства сжатия информации. Проблемы распознавания изображений кроме классической задачи распознавания фигур заданной формы на изображении ставят новые задачи распознавания линий и углов на изображении, распознавания края изображения. В данной исследовательской работе авторы решали проблему локализации текста на изображении.

Для распознавания текста на изображении, необходимо на начальном этапе локализовать область изображения, содержащую текстовые символы. Данная задача актуальна в различных областях: при анализе телевизионных трансляций, поиске информации в медиабазы, административном обслуживании в банках, страховых компаниях и т.д. В статье описывается, оптимизированная авторами, методика обнаружения текстовых областей и выделения символов на изображениях снятых с металлических поверхностей, на которых лазерной гравировкой нанесен некоторый текст. Разработан специализированный алгоритм обработки (очистки) изображения, получаемого с USB-камеры. Алгоритм основан на методах математической морфологии. Цель работы заключалась в создании (усовершенствовании) методики и разработке модуля программы, динамически определяющих текстовую область надписи, нанесенной на металлическую поверхность методом лазерной гравировки. В результате исследования получен оптимизированный алгоритм предварительной обработки текстовых изображений, снятых с металлических поверхностей с помощью USB – устройства. Практическая значимость данного алгоритма обуславливается в использовании в программном обеспечении, разрабатываемом для промышленного предприятия Комсомольский-на-Амуре авиастроительный завод, в качестве программного модуля предварительной обработки изображений перед распознаванием.

*Ключевые слова: обработка изображений, математическая морфология, лазерная гравировка*

## ВВЕДЕНИЕ

На предприятиях возрастает потребность в прикладных программах, способных считывать и распознавать идентификаторы на стандартных текстовых бирках, надписях, маркировках и шильдиках деталей, инструмента и блоков используемого при производстве оборудования. Так как эти объекты подвергаются активной эксплуатации, качество изображения надписей на них, снимаемого для распознавания в компьютерных системах, имеет не самое лучшее качество. В реальном производстве большая доля надписей и маркировок выполняется лазерной гравировкой на металле. Цвет, тон, шероховатость металлической поверхности затрудняют выделение области текстовых маркировок при их компьютерном считывании. Возникает потребность в разработке (усовершенствовании) методик предварительной обработки изображения перед передачей его в стандартный OCR (модуль распознавания текста) в автоматизированных системах машинного зрения. Одним из важных шагов предварительной обработки

изображения является локализация области изображения, содержащей текст.

В связи с этим с каждым днем все больше и больше возрастает интерес к методам распознавания печатных, рукописных и выполненных лазерной гравировкой текстам [11-18], существующие методы и алгоритмы пересматриваются и улучшаются с точки зрения решения задач поиска, распознавания и выявления смысла разнообразной текстовой информации на изображениях и видеозаписях [1 – 6, 8 – 11]. Для решения задачи определения местоположения текста используется большое число различных алгоритмов, основанных на использовании локальных особенностей изображений. Одним из таких алгоритмов является алгоритмы SWT [17] и MSER [18]. Предложенная авторами методика, отличается от подходов, используемых в SWT и MSER алгоритмах. Общая схема предложенной методики представлена на рисунке 1.

В качестве входных параметров метода рассматривается цифровое дискретное изображение размером  $M$  на  $N$  точек, а элементы такого изображения будем называть пикселями. Важно заметить, что алгоритм работает с полутоновыми изображениями, поэтому необходимо на первом шаге

преобразовать исходное цветное изображение в полутоновое.

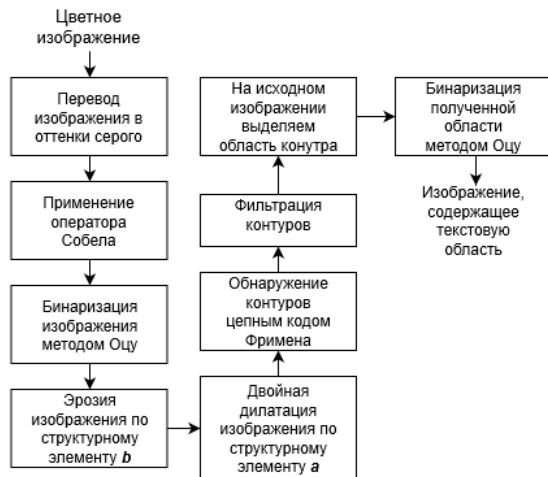


Рис. 1 – Схема, предложенной методики

Разработанный алгоритм обнаружения текстовых областей на изображении базируется на том, что во многих случаях части изображений, содержащие текст, имеют ярко выраженный горизонтальный перепад яркости между буквами и фоновым рисунком, который можно подчеркнуть при помощи вторых горизонтальных пространственных разностей между пикселями изображения [7]. Обнаружение этих областей осуществляется с помощью корреляционной обработки и методов математической морфологии [1].

После преобразования изображения в полутоновое, необходимо обнаружить перепады яркостей на изображении, т.к. данные перепады будут сигнализировать о наличии контура. Для этой цели предлагается применить дискретный дифференциальный оператор, вычисляющий приближение градиента яркости изображения – оператор Собеля [2].

Оператор Собеля основан на свёртке изображения небольшими целочисленными фильтрами в вертикальном и горизонтальном направлениях, поэтому его относительно легко вычислять. Так как требуется обнаруживать исключительно текст, то важны перепады только в горизонтальном направлении [8]. Текстовая область будет содержать большое количество скачков на ограниченном пространственном промежутке, что позволит определить их границы. Для решения данной задачи оператор Собеля принимает вид: размерность ядра свертки = 5x5, порядок частной производной по направлению  $x = 1$ , порядок частной производной по направлению  $y = 0$ . Таким образом формула градиента яркости изображения примет вид:

$$F_p(x, y) = \frac{\partial F_i(x, y)}{\partial x}, \quad (1)$$

где  $F_p(x, y)$  – значение яркости пикселя с координатами  $x$  и  $y$  преобразованного изображения;

$F_i(x, y)$  – значение яркости пикселя с координатами  $x, y$  исходного изображения.

Бинаризация изображения осуществляется методом Оцу [3]. Суть метода Оцу заключается в том, чтобы выставить порог между классом пикселей изображения и классом пикселей фона таким образом, чтобы каждый из них был как можно более «плотным», то есть – к минимизации внутриклассовой дисперсии, которая определяется как взвешенная сумма дисперсий двух классов [2]f1:

$$\sigma_w^2 = w_1\sigma_1^2 + w_2\sigma_2^2, \quad (2)$$

где  $w_1$  и  $w_2$  – вероятности первого и второго классов соответственно. В своей работе Оцу показывает, что минимизация внутриклассовой дисперсии эквивалентна максимизации межклассовой дисперсии, которая равна [9]:

$$\sigma_b^2 = w_1w_2(a_1 - a_2)^2 \quad (3)$$

где  $a_1$  и  $a_2$  – средние арифметические значения для каждого из классов. Особенность этой формулы заключается в том, что  $w_1(t+1), w_2(t+1), a_1(t+1), a_2(t+1)$ , легко выражаются через предыдущие значения что  $w_1(t), w_2(t), a_1(t), a_2(t)$ , где  $t$  - текущий порог.

Данная особенность позволила разработать быстрый алгоритм [4]:

1. Вычисляем гистограмму (один проход через массив пикселей). Далее нужна только гистограмма; проходов по всему изображению больше не требуется, что уменьшает вычислительную сложность алгоритма, которая равна  $O(n)$ .

2. Начиная с порога  $t = 1$ , проходим через всю гистограмму, на каждом шаге пересчитывая дисперсию  $\sigma_b(t)$ . Если на каком-то из шагов дисперсия оказалась больше максимума, то обновляем дисперсию и  $T = t$ .

3. Искомый порог равен  $T$ .

В рамках решаемой задачи авторы оптимизировали работу алгоритма. Проход через гистограмму делается не от 1 до 254, а от минимальной до максимальной яркости минус единица.

В результате всех вышеописанных действий мы получаем изображение, показанное на рисунке 2.



исходное обработанное

Рис. 2 – Результат применения дифференциального оператора, с последующей бинаризацией.

На этом этапе ставится задача, как выделить именно текстовую область, используя методы математической морфологии [5].

### ОСНОВНАЯ ЧАСТЬ

В контексте машинного зрения морфология относится к описанию свойств формы областей на изображении. Операции математической морфологии изначально были определены как операции над множествами, но скоро выяснилось, что они также полезны в задачах обработки множества точек в двумерном пространстве.

Множествами в математической морфологии представляются объекты на изображении. Легко заметить, то множество всех фоновых пикселей бинарного изображения является одним из вариантов его полного описания [6].

Входными данными для аппарата математической морфологии являются два изображения: обрабатываемое и специальное, зависящее от вида операции и решаемой задачи. Такое специальное изображение принято называть примитивом или структурным элементом. В нашем случае мы задаем два структурных элемента:

1. Структурный элемент  $b$  для эрозии. В результате применения операции эрозии все объекты, меньшие чем структурный элемент, стираются, объекты, соединённые тонкими линиями становятся разъединёнными и размеры всех объектов уменьшаются. В нашем случае эрозия осуществляется в вертикальном по не высокому шаблону, что позволяет очистить изображение от импульсного шума, и более точно локализовать область по высоте символов идентификатора.

2. Структурный элемент  $a$  для дилатации. Дилатация увеличивает область изображения, расширяя его пиксели и тем самым способствуя объединению областей изображения, которые были разделены шумом и др. В нашем случае область текста размывается, это пригодится для выделения контура текстовой области.

Разработанная морфологическая операция является модификацией операции размывания, которая в общем случае описывается следующим образом [6]:

$$(A \ominus b) \oplus b, \quad (4)$$

где  $A$  – исходное изображение,  $b$  – структурный элемент,  $\ominus$  – операция эрозия,  $\oplus$  – операция дилатация.

Предложенная авторами формула имеет следующий вид:

$$((A \ominus b) \oplus a) \oplus a, \quad (5)$$

где  $A$  – исходное изображение,  $b$  – структурный элемент, применяемый при операции «эрозия»,  $a$  – структурный элемент применяемый для операции «дилатация».

То есть – авторы предлагают применять двойную дилатацию для улучшения качества размывания, в результате чего получается область, которая является фоном для текста (рисунок 3). На рисунке 3 слева показано изображение, полученное в результате применения дифференциального оператора Собеля, и последующей его бинаризации, справа – результат применения выведенной авторами формулы к изображению слева.



исходное обработанное

Рис. 4 – Результат применения формулы (5) к предварительно обработанному изображению оператором Собеля и методом Оцу.

Далее над изображением, полученным в результате применения формулы (5), необходимо провести контурный анализ. Контурный анализ производим цепным кодом Фримена [4] с аппроксимацией Teh-Chin [4]. Обнаруженные контуры соединяются, накладываются на исходное изображение, и выполняется бинаризация полученной тестовой области. Результат работы программы представлен на рисунке 5.



исходное обработанное

Рис. 5 – Финальный результат работы программы

На рисунке 5 справа представлено исходное изображение, слева – то же изображение, пропущенное через разработанный авторами фильтр, основанный на морфологической обработке изображений.

### РЕЗУЛЬТАТЫ

Работа, предложенной авторами методики по обнаружению текстовых областей проверялась на реальных изображениях бирки с лазерной гравировкой (подобные рис. 1 слева). Авторы предлагают сравнить предложенную методику, основанную на использовании методов математической морфологии с известными на сегодняшний день алгоритмами SWT и MSER, основная идея работы которых заключается

в том, что буквы и цифры на изображении, как правило, имеют постоянную толщину штриха [10].

Для сравнения работы алгоритмов для локализации текста используются следующие метрики [10]:

1. Точность локализации. Эта метрика отражает вероятность того, что выделенные текстовые области являются корректными;

2. Полнота локализации. Эта метрика отражает вероятность того, что корректные текстовые области были выделены;

3. Время работы алгоритмов.

Результаты сравнения работы алгоритмов представлены в таблице 1.

**Табл. 1. Результаты сравнения работы алгоритмов**

Метод	Точность локализации, %	Полнота локализации, %	Время работы с одним изображением, с.
SWT	91	83	0.23
MSER	93	88	0.21
<b>Методика, предложенная авторами</b>	<b>96</b>	<b>85</b>	<b>0.17</b>

Как видно из табл. 1, предложенная авторами методика показывает высокую скорость в сравнении с другими известными алгоритмами. Не самый лучший показатель полноты локализации, обусловлен тем, что предложенная авторами методика не изотропна, т.к. применение дифференциального оператора к изображению осуществляется только в горизонтальном направлении.

**ЗАКЛЮЧЕНИЕ**

Основная цель работы была достигнута. На практике был реализован, предложенный авторами, оптимизированный алгоритм предварительной обработки текстовых изображений, снятых с металлических поверхностей с помощью USB-устройств. Алгоритм показал высокое качество выделения текстовых групп из изображений, что позволило, в дальнейшем, при передаче этих изображений в стандартные OCR системы получить 100% распознавание снятых текстовых надписей.

**СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ**

1. Гонсалес, Р Цифровая обработка изображений [Текст]: учеб. пособие / Р. Гонсалес, Р. Вудс; перевод с англ.: Л.И. Рубанова, П.А. Чочиа; под ред. П.А. Чочиа. – М.: Техносфера, 2012. – 1104 с.  
 2. Otsu, N., «A Threshold Selection Method from Gray-Level Histograms,» IEEE Transactions on Systems, Man, and Cybernetics, Vol. 9, No. 1, 1979, pp. 62-66.  
 3. Визильтер, Ю. А. Обработка и анализ цифровых изображений с примерами на LabView и Imaq Vision [Текст]: учеб.

пособие для вузов / Ю.А. Визильтер, В.А. Князь, С.Ю. Желтов; под ред. В.А. Мовчан.. – М.: ДМК-Пресс, 2016. – 464с.

4. Фурман, Я. А. Введение в контурный анализ; приложение к обработке изображений и сигналов [Текст]: учеб. пособие для вузов / Я.А. Фурман, А.К. Передреев, А.В. Кревецкий; под ред. Я.А. Фурман.. – М.: Физматлит, 2003. – 592с.

5. Красильников, Н.А. Цифровая обработка 2D- и 3D-изображений [Текст]: учеб. пособие для вузов / Н.А.Красильников.– СПб.: BHV, 2011. – 608 с.

6. Ghosh P., Deguchi K. Mathematics of shape description: A morphological approach to image processing and computer graphics [Wiley, 2008]

7. Goutsias J. Mathematical morphology and its applications to image and signal processing / Goutsias J., Vincent L., Bloomberg D., Kluwer, 2002. – 456p.

8. K. Jung, K. Kim, A. K. Jain, “Text information extraction in images and video: a survey”, Pattern Recognition, p. 977 – 997, Vol 5. 2004.

9. Шапиро, Л Компьютерное зрение [Эл. издание]: отдельное издание / Л. Шапиро, Дж. Стокман; перевод с англ.: – М.: Бином, 2006. – 752с.

10. Андрианов А.И. Локализация текста на изображениях сложных графических сцен / Современные проблемы науки и образования, 2013. № 3

11. Кузьмицкий, Н.Н. Обнаружение фрагментов текста на изображениях реальных сцен на базе сверточной нейросетевой модели / Информатика. –2015. – № 2(46). – С. 12-21.

12. Yin, X.-C. Multi-orientation scene text detection with adaptive clustering / X.-C. Yin, W.-Y. Pei, J. Zhang, H.-W. Hao // IEEE Transactions on Pattern Analysis and Machine Intelligence. – 2015. – Vol. 37, Issue 9. – P. 1930-1937. – DOI: 10.1109/TPAMI.2014.2388210

13. Zuo, Z.-Y. Multi-strategy tracking based text detection in scene videos / Z.-Y. Zuo, S. Tian, X.-C. Yin // 13th International Conference on Document Analysis and Recognition (ICDAR). – 2015. – P. 66-70. – DOI: 10.1109/ICDAR.2015.7333727

14. Koo, H.I. Scene text detection via connected component clustering and nontext filtering / H.I. Koo, D.H. Kim / IEEE Transactions on Image Processing. – 2013.– Vol. 22, Issue 6. – P. 2296-2305. – DOI: 10.1109/TIP.2013.2249082

15. Болотова, Ю.А. Распознавание автомобильных номеров на основе метода связанных компонент и иерархической временной сети / Ю.А. Болотова, В.Г. Спицын, М.Н. Ру домёткина // Компьютерная оптика. – 2015. – Т. 39, № 2. –С. 275-280. – DOI: 10.18287/0134-2452-2015-39-2-275-280

16. Столяров В.Я .Оценка эффективности применения контурной обработки к изображениям с лазерной гравировкой / В.Я.Столяров, В.А.Тихомиров, сборник Международных научно-практических конференций "Актуальные научные исследования в современном мире", София, Болгария, 2018 (World of Sciens), С. 84-88.

17. Epshtein B., Ofek E., Wexler Y. Detecting text in natural scenes with stroke width transform // Proceedings of International Conference on Computer Vision and Pattern Recognition, 2012. P. 2963-2970

18. Neumann L., Matas J. A method for text localization and recognition in real-world images //10th Asian conference on Computer vision, 2012. P. 770-783.

*Столяров Виталий Яварович – аспирант, преподаватель, ФГБОУ ВО КнАГУ, тел. (984)1758172, e-mail: stolyarov.vitalij@list.ru*

*Тихомиров Владимир Александрович – заведующий кафедры математического обеспечения и применения ЭВМ, профессор, кандидат технических наук ФГБОУ ВО КнАГУ,, e-mail: kmorevm@knastu.ru*

# METHOD FOR COMPUTER IDENTIFICATION OF TEXT AREAS ON THE LABEL IMAGES IMPLEMENTED BY LASER ENGRAVING

V.J. Stolyarov, V.A. Tikhomirov

*Komsomolsk-on-Amure State University, Komsomolsk – on – Amure*

Abstract – Nowadays, image processing is an important area of application of modern computer technology. The following image processing tasks are known such as filtering and restoring images, image segmentation, as information compression tools. The problems of image recognition, in addition to the classical task of recognizing figures of a given shape in an image, pose new problems of recognizing lines and angles in an image, recognizing the edges of an image. In this research work, the authors solved the problem of localizing text in an image.

To recognize text in the image, it is necessary at the initial stage to localize the image area containing text characters. This task is relevant in various fields: in the analysis of television broadcasts, the search for information in media outlets, administrative services in banks, insurance companies, etc. The article describes a technique optimized by the authors for detecting text areas and highlighting characters in images taken from metal surfaces on which some text is laser-engraved. A specialized algorithm for processing (cleaning) an image received from a USB camera has been developed. The algorithm is based on methods of mathematical morphology.

The purpose of the work was to create (improve) the methodology and develop a program module that dynamically determines the text area of the inscription deposited on a metal surface by laser engraving. Because of the study, an optimized algorithm for the preliminary processing of text images taken from metal surfaces using a USB device was obtained. The practical significance of this algorithm is determined by the use in the software developed for the industrial enterprise Komsomolsk-on-Amur aircraft factory as a software module for image pre-processing before recognition.

Index terms: image processing, mathematical morphology, laser engraving

## REFERENCES

1. Gonsalez, R Digital image processing [text]/R. Gonsalez, R. Woods; translate from English.: L.I. Rubanova, P.A. Chochia; under the editorship. P.A. Chochia. – M.: Technosfera, 2012. – 1104 p.
2. Otsu, N., «A Threshold Selection Method from Gray-Level Histograms,» IEEE Transactions on Systems, Man, and Cybernetics, Vol. 9, No. 1, 1979, pp. 62-66.
3. Vizilter, Yu. A. Digital image processing and analysis with example LabView and Imaq Vision [text]: / Yu.A. Vizilter, V.A. Knyaz, S.Yu. Zheltov; under the editorship. V.A. Movchan.. – M.: DMK-Press, 2016. – 464p.
4. Phurman, Ya. A. Introduction to contour analysis [text]: / Ya. A. Phurman, A.K. Peredreev, A.V. Krevetskiy; under the editorship. Ya.A. Phurman.. – M.: Phymathlit, 2003. – 592p.
5. Krasilnikov, N.A. Digital 2-D and 3-D image processing [text] / N.A. Krasilnikov.– SPb.: BHV, 2011. – 608 c.
6. Ghosh P., Deguchi K. Mathematics of shape description: A morphological approach to image processing and computer graphics [Wiley, 2008]
7. Goutsias J. Mathematical morphology and its applications to image and signal processing / Goutsias J., Vincent L., Bloomberg D., Kluwer, 2002. – 456p.
8. K. Jung, K. Kim, A. K. Jain, "Text information extraction in images and video: a survey", Pattern Recognition, p. 977 – 997, Vol 5. 2004.
9. Shapiro, L Computer vision / L. Shapiro, G. Stockman; translate from English.: – M.: Binom, 2006. – 752p.
10. Andrianov A.I Localization of text on images of complex graphic scenes / Modern problems of science and education, 2013. № 3
11. Kuzmitskiy NN. Detection of text objects in images of realscenes based on convolutional neural network model [InRussian]. Informatics 2015; 2(46): 12-21.
12. Yin, X.-C. Multi-orientation scene text detection with adaptive clustering / X.-C. Yin, W.-Y. Pei, J. Zhang, H.-W. Hao // IEEE Transactions on Pattern Analysis and Machine Intelligence. – 2015. – Vol. 37, Issue 9. – P. 1930-1937. – DOI: 10.1109/TPAMI.2014.2388210
13. Zuo, Z.-Y. Multi-strategy tracking based text detection in scene videos / Z.-Y. Zuo, S. Tian, X.-C. Yin // 13th International Conference on Document Analysis and Recognition (ICDAR). – 2015. – P. 66-70. – DOI: 10.1109/ICDAR.2015.7333727
14. Koo, H.I. Scene text detection via connected component clustering and nontext filtering / H.I. Koo, D.H. Kim / IEEE Transactions on Image Processing. – 2013.– Vol. 22, Issue 6. – P. 2296-2305. – DOI: 10.1109/TIP.2013.2249082
15. Bolotova YuA, Spitsyn VG, Rudometkina MN. Licenseplate recognition algorithm on the basis of a connected components method and a hierarchical temporal memory model. Computer Optics 2015; 39(2): 275-280. DOI:10.18287/0134-2452-2015-39-2-275-280
16. Stolyarov V. Ya . Evaluation of the effectiveness of applying contouring to images with laser engraving / V. Ya. Stolyarov, V.A.Tikhomirov, collection of international scientific and practical conferences "Actual scientific research in the modern world ", Sofia, Bulgaria, 2018 (World of Sciens), C. 84-88.
17. Epshtein B., Ofek E., Wexler Y. Detecting text in natural scenes with stroke width transform // Proceedings of International Conference on Computer Vision and Pattern Recognition, 2012. P. 2963-2970
18. Neumann L., Matas J. A method for text localization and recognition in real-world images //10th Asian conference on Computer vision, 2012. P. 770-783.

*Stolyarov Vitalij Yavarovich – postgraduate, lecturer, Komsomolsk-on-Amure State University, (984)1758172, e-mail: stolyarov.vitalij@list.ru*

*Tikhomirov Vladimir Alexandrovich – head of the department of software and applications, docent, candidate of Engineering Sciences, e-mail: kmopevm@knastu.ru.*