

Чтение дорожных знаков в изображениях, полученных при неконтролируемых условиях регистрации

И. Н. Приходько*

*Московский государственный университет имени М. В. Ломоносова,
физический факультет, кафедра математического моделирования и информатики
Россия, 119991, Москва, Ленинские горы, д. 1, стр. 2
(Статья поступила 14.05.2018; Подписана в печать 18.05.2018)*

В работе представлен подход к решению задачи чтения дорожных знаков на изображениях, полученных при неконтролируемых условиях регистрации. Условия регистрации учитываются в морфологическом анализе, примененном в данной работе. Эта задача делится на локализацию дорожных знаков определенных цвета и формы и чтение кандидата в дорожные знаки.

PACS: 02.70.-c

УДК: 519.85

Ключевые слова: морфологические методы, форма изображения, анализ изображений, распознавание дорожных знаков, компьютерное зрение, контурный анализ.

ВВЕДЕНИЕ

Характерной чертой для нашего времени является востребованность в работах беспилотников, применяемых в различных сферах. Но как было замечено в книге [1], технологии, развивающиеся невероятным темпом, сдерживаются, в том числе, и проблемами «искусственного интеллекта» в понимании сцен и изображений. Поэтому в настоящее время очень актуальными являются задачи компьютерного зрения для роботов различного характера.

К беспилотникам, которым невероятно важно анализировать и понимать ситуацию вокруг, относятся не только автомобили, но и корабли, самолеты, дроны. Анализу ситуации помогает чтение дорожных или иных указательных знаков. Самый простой путь к автоматизации этого процесса состоит из получения изображения с помощью камеры, обработки, поиска кандидатов в дорожные знаки и последующей классификации этих кандидатов. Последнее нередко, например в исследованиях [2–4], было реализовано с помощью искусственных нейронных сетей. Однако такой подход, хоть и показывает весьма неплохие результаты, очень сильно зависит от обучающей выборки и не способен гарантировать точного результата при любых условиях регистрации. Поэтому наиболее предпочтительным можно считать использование более явных математически обоснованных подходов.

Дорожные знаки (ДЗ) имеют следующие свойства: обладают простыми четкими и контрастными формами, легко разделяются на классы и подклассы в зависимости от формы и цвета. Это упрощает задачу классификации, дополняя ее априорной информацией о принадлежности кандидата в ДЗ к определенному подклассу шаблонов.

Перечисленные свойства предполагают, что уместными подходами к решению задачи чтения ДЗ на изобра-

жении могут считаться методы, основанные на морфологическом анализе, работающем с непосредственно понятием формы, а также контурном анализе, использующем информацию о границах объекта.

1. ЗАДАЧИ ЧТЕНИЯ ДЗ

Предполагается, что на беспилотнике есть регистрирующая аппаратура или любой другой источник изображений, отображающих ту или иную дорожную ситуацию. Под ситуацией понимается определенный набор ДЗ. Для того чтобы осознать эту ситуацию следует правильно прочитать этот набор ДЗ. Для этого процесс чтения ДЗ на изображении разделим на следующие этапы, описанные в статьях [4, 5]:

- получение цветного изображения от регистрирующей аппаратуры;
- предобработка изображения в соответствии со статьей [6];
- локализация кандидатов в ДЗ различными методами, представленными в работах [7–9];
- фильтрация кандидатов в ДЗ согласно статьям [10, 11];
- классификация кандидатов в ДЗ;
- обобщение информации о прочитанных ДЗ и составление дорожной ситуации.

В настоящей работе рассмотрено решение задачи классификации кандидатов в ДЗ с учетом априорной информации получаемой от упомянутых методов предобработки и локализации. При этом под априорной информацией понимается преобладающий цвет знака (синий, красный, желтый) и его форма (круг, квадрат, треугольник и т.д.)

*E-mail: prikhodko_in@mail.ru

2. МОДЕЛЬ ИЗОБРАЖЕНИЯ КАНДИДАТА В ДЗ

Алгоритмы локализации, описанные в работах [7, 8], предоставляют набор k кандидатов в ДЗ в виде цветных изображений $\{f_1, f_2, \dots, f_k\}$. С математической точки зрения изображение представляет собой функцию $f(\cdot)$, заданную на ограниченном подмножестве X плоскости \mathcal{R}^2 и ставящую в соответствие каждому элементу x_i сетки узлов заданной на X значение $f(x_i)$. Причем X называется полем зрения, а $i = \overline{1, N}$, где N число разбиений сетки. В случае аддитивной цветовой схемы RGB каждое значение $f(x_i)$ есть упорядоченный набор из трех чисел, т.е. $f(x_i) = \{R, G, B\}$, где $x_i \in X \subset \mathcal{R}^2$, причем $R, G, B \in [0, 255]$, и каждое из чисел R, G, B отображает яркость в x_i , соответствующую своему цветовому каналу.

ДЗ представляет собой цветное мозаичное изображение, состоящее из сегментов белого, черного, а также одного из основных цветов: синего, красного или желтого. Так как из априорной информации уже известен дополнительный цвет ДЗ, то можно без потери информации преобразовать цветное изображение $f(x_i) = \{R, G, B\}$ в изображение в оттенках серого, которое имеет значение $f(x_i) = Y$, где $Y \in [0, 255]$, и вычисляется в согласии с алгоритмами, изложенными в книге [12].

Таким образом, можно представить ДЗ в виде серого мозаичного кусочно-непрерывного изображения:

$$f(x) = \sum_{i=1}^N c_i \chi_i(x), \quad x \in X, \quad i = \overline{1, N}, \quad (1)$$

где $c_i \in [0, 255]$ — яркость изображения $f(\cdot)$ в точке x_i , $\chi_i(x)$ — индикаторная функция принадлежности x разбиению A_i поля зрения X : $\bigcup_{i=1}^N A_i = X$, N — количество кусочно-непрерывных областей, а индикаторная функция записывается как:

$$\chi_i = \begin{cases} 1, & x \in A_i, \\ 0, & x \notin A_i. \end{cases}$$

3. ЭЛЕМЕНТЫ МОРФОЛОГИЧЕСКОГО АНАЛИЗА

Центральным понятием морфологического анализа, изложенного в книге [13] и статьях [14, 15], является форма изображения, понимаемая как инвариант условной регистрации.

Итак, ДЗ рассматривается как изображение в градациях серого, т.е. под изображением понимается числовая функция $f(\cdot)$, заданная на ограниченном подмножестве X плоскости \mathcal{R}^2 . Значение $f(x)$ функции $f(\cdot)$ в точке $x \in X$ называется яркостью в точке x поля зрения X .

Согласно книге [13] пространство изображений является евклидовым пространством $\mathcal{L}_\mu^2(X)$ с задан-

ной считающей мерой μ . Линейные операции сложения изображений и умножения изображения на число определены следующим образом:

$$\begin{aligned} (f + g)(x) &= f(x) + g(x), \\ (a \cdot f)(x) &= a \cdot f(x), \quad x \in X, \end{aligned}$$

а скалярное произведение и норма заданы как:

$$\begin{aligned} (g, f) &= \sum_{i=1}^n f(x_i)g(x_i), \\ \|f\| &= \left(\sum_{i=1}^n f^2(x_i) \right)^{\frac{1}{2}}. \end{aligned}$$

Изображение $f(\cdot)$ принадлежит классу кусочно-непрерывных изображений $C(X)$, заданному на поле зрения X тогда и только тогда, когда существует конечное измеримое разбиение A_1, \dots, A_N поля зрения X , такое, что на каждом из $A_i, i = 1, \dots, N$, функция $f(\cdot)$ является непрерывной, причем,

$$\varphi(f) = \sup_{x \in X} |f(x)| < \infty.$$

Функционал $\varphi(f)$, определенный на $C(X)$, является нормой элемента $f \in C(X)$ и обозначается как $\|f\|_C$. В этом случае расстояние между изображениями f и g определяется нормой разности $f - g$.

Причем линейное нормированное пространство изображений $C(X)$ не является евклидовым: в нем нельзя определить скалярное произведение, согласованное с нормой $\|\cdot\|_C$.

Так как (1) принадлежит пространству $C(X)$, то форма ДЗ может быть записана как:

$$\begin{aligned} \mathcal{V}_f &= \left\{ f(x) = \sum_{i=1}^m c_i \chi_i(x), \right. \\ &\left. x \in X, \quad c_i \in (-\infty, \infty), \quad i = \overline{1, m} \right\}. \quad (2) \end{aligned}$$

Пусть множество всех изображений есть евклидово пространство $\mathcal{L}_\mu^2(X)$, а множество \mathcal{V}_f является выпуклым замкнутым множеством в $\mathcal{L}_\mu^2(X)$. С формой (2), как с подпространством $\mathcal{L}_\mu^2(X)$, взаимно однозначно связан оператор $\Pi_{\mathcal{V}_f}$ ортогонального проецирования на \mathcal{V}_f . Этот оператор каждому элементу $g \in \mathcal{L}_\mu^2(X)$ ставит в соответствие его единственную ортогональную проекцию $\Pi_{\mathcal{V}_f} g \in \mathcal{V}_f$, определяемую как ближайшее к $g \in \mathcal{L}_\mu^2(X)$ изображение $\Pi_{\mathcal{V}_f} g$ из \mathcal{V}_f .

Ортогональный проектор $\Pi_{\mathcal{V}_f}$ на форму \mathcal{V}_f может быть записан как:

$$\Pi_{\mathcal{V}_f} \equiv \Pi_f = \sum_{i=1}^m \frac{(g, \chi_i)}{\|\chi_i\|^2} \chi_i.$$

Таким образом, задача классификации ДЗ представляется в виде решения задачи на минимум:

$$\|g - \Pi_{f_k} g\|^2 \sim \arg \min_k, \quad (3)$$

где f_k — изображение-образец, принадлежащее форме \mathcal{V}_{f_k} , g — изображение, которое проверяется на принадлежность к форме \mathcal{V}_{f_k} , Π_{f_k} — ортогональный проектор на форму \mathcal{V}_{f_k} , $k = \overline{1, n}$, n — количество изображений-образцов, причем $\|g - \Pi_{f_k}g\|^2$ считается мерой близости изображения g к форме \mathcal{V}_{f_k} .

4. КЛАССИФИКАЦИЯ ДЗ

Алгоритм классификации выполняется следующим образом. Предположим, с предыдущих алгоритмов локализации ДЗ поступило изображение $h \in \mathcal{L}_\mu^2(X)$, о котором известно, что это круглый синий ДЗ. Теперь следует определить к какому из возможных ДЗ с подходящими характеристиками он относится. В первую очередь следует избавиться от ненужной информации о фоне. Предлагается сделать это следующим образом:

$$g = h\chi_{\text{интерес}},$$

где $\chi_{\text{интерес}}$ — индикаторная функция интересующей области, не включающей фон (рис. 1). Т.к. эти преобразования производятся и с шаблонами ДЗ, то фон совершенно никак не будет влиять на результат.

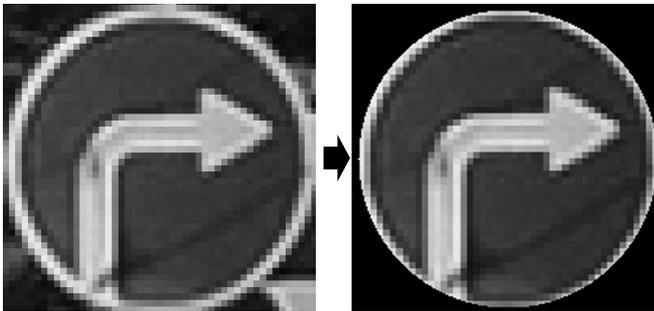


Рис. 1: Удаление фона с классифицируемого изображения

Теперь следует приступить к решению задачи на минимум (3). Программа, реализующая морфологический классификатор, была написана на языке Python 3, с использованием `numpy` и `OpenCV` для базовых операций с матрицами и изображениями.

Для наглядного примера работы морфологического классификатора возьмем $n = 2$ шаблонов, один из которых соответствует классифицируемому знаку, причем только в рамках данного примера изображения увеличиваются до размеров 200×200 пикселей.

Следующий этап представляет собой вычисление $\|g - \Pi_{f_k}g\|^2$ и поиск минимума по k . Стоит отметить, что $g - \Pi_{f_k}g$ принадлежит пространству изображений $\mathcal{L}_\mu^2(X)$ и называется морфологической разностью, т.к. наглядно отображает различия между изображениями f_k и g (рис. 2). Согласно ПДД, каждому ДЗ присвоен свой номер и название, поэтому в дальнейшем для идентификации ДЗ может быть использован лишь номер.

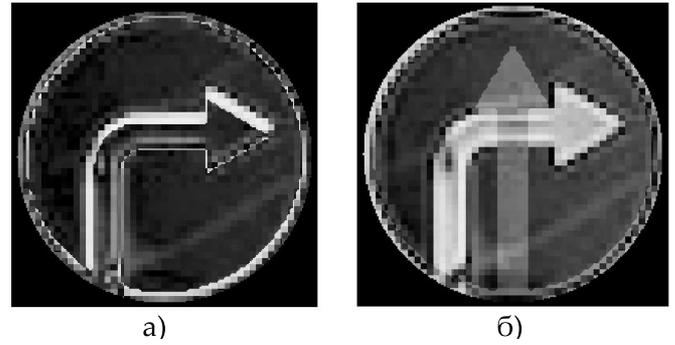


Рис. 2: Морфологические разности между классифицируемым изображением g и шаблонами (а) f_1 (№ 4.1.2), соответствующими классифицируемому ДЗ, и (б) f_2 (№ 4.1.1)

Квадрат нормы морфологической разности для шаблона f_1 равен $\|g - \Pi_{f_1}g\|^2 \simeq 7033^2 \simeq 50 \cdot 10^6$ (рис. 2,а). Для f_2 мера близости изображений $\|g - \Pi_{f_2}g\|^2 \simeq 10017^2 \simeq 100 \cdot 10^6$ (рис. 2,б). Для удобства чтения, значение v_m меры близости приведем к виду:

$$v_m = 100 \frac{\|g_i - \Pi_{f_m}g_i\|^2}{\min_k \|g_i - \Pi_{f_k}g_i\|^2} - 100, \quad (4)$$

где $m = \overline{1, n}$. Таким образом, получим $v_1 = 0$, $v_2 \simeq 102$. Этот вид позволяет сразу оценить на сколько процентов от минимума отличаются отброшенные кандидаты и судить о корректности результатов. Так как минимуму соответствует наибольшая близость между изображениями, то изображение g считается относящимся к ДЗ изображенному на f_1 .

Итак, в рамках данного примера был получен следующий результат: на анализируемом изображении h , прочитан ДЗ № 4.1.2 «Движение направо».

Рассмотрим работу морфологического классификатора при $n = 13$ и размерах изображений 25×25 пикселей, что более соответствует реальным условиям регистрации. В целях сокращения объема вычислительных операций при непосредственно классификации, все проекторы на форму изображений-шаблонов Π_{f_k} были рассчитаны заранее. В третьем столбце h_2 табл. I отображены приведенные значения v_m мер близости изображений f_k к $h_2 \equiv h$ в соответствии с (4). Минимумы, равные нулю, указывают на принадлежность к форме \mathcal{V}_{f_i} .

Кроме уже знакомого изображения h_2 , в классификации также участвовали изображения h_1, h_3, h_4, h_5 (рис. 3), которые тоже были верно классифицированы.

Небольшие отклонения от минимума (по сравнению с остальными столбцами) в столбце h_4 (табл. I) указывают на небольшие сложности при классификации этого изображения. Эти трудности вызваны тем, что ДЗ был зарегистрирован под значительным углом, и знак в результате преобразований оказался повернут на небольшой угол, поэтому его форма плохо совпа-

Таблица I: Классификация изображений h_i

№	ДЗ f_i	h_1	h_2	h_3	h_4	h_5
4.1.1	Движ. прямо	0	101	110	23	130
4.1.2	Движ. направо	217	0	146	25	106
4.1.3	Движ. налево	271	78	143	17	134
4.1.4	Движ. прямо или направо	171	99	153	23	130
4.1.5	Движ. прямо или налево	163	94	0	19	125
4.1.6	Движ. направо или налево	168	73	112	19	130
4.2.1	Объезд препят. справа	198	99	145	0	135
4.2.2	Объезд препят. слева	201	77	140	25	0
4.2.3	Объезд препят. спр. или сл.	219	66	141	24	131
4.3	Круговое движение	226	99	143	25	137
4.4	Велосипедная дорожка	226	99	143	25	137
4.5	Пешеходная дорожка	153	91	135	18	120
4.6	Ограничение мин. скорости	220	81	139	19	108

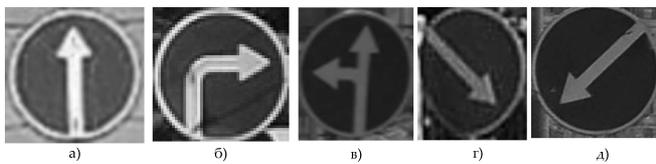


Рис. 3: Классифицируемые изображения: (а) h_1 , (б) h_2 , (в) h_3 , (г) h_4 , (д) h_5

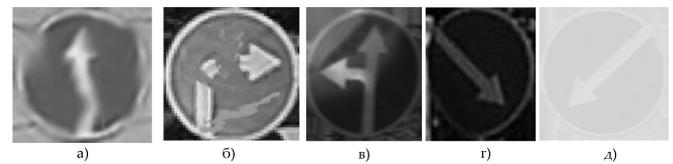


Рис. 4: Искаженные классифицируемые изображения: (а) h_1 , (б) h_2 , (в) h_3 , (г) h_4 , (д) h_5

ла с шаблоном но, несмотря на это, минимум хорошо выражен.

5. ОСОБЕННОСТИ ПОДХОДА

Как известно из книги [13], методы морфологического анализа хорошо помехозащищены, практически не зависят от условий регистрации, имеют низкие требования к предобработке изображений, а также при анализе малых изображений достаточно быстры.

Ради демонстрации помехозащищенности метода, были искажены ранее рассматриваемые классифицируемые изображения h_i , $i = \overline{1,5}$ (рис. 4). Искажения вносились на основе гипотетических проблем при реальных условиях регистрации таких как: размазывание ДЗ из-за высокой скорости автомобиля, замусоренность или визуальное заграждение части ДЗ, отблески, а также засвеченность или затемненность ДЗ.

Результаты классификации искаженных знаков представлены в виде таблицы (табл. II). Как можно заметить, нули по-прежнему соответствуют верным результатам, хотя разница между минимумом и отброшенными шаблонами стала заметно меньше. В столбце, соответствующем h_5 , пришлось округлять значения до сотых (а не до целых), чтобы оценить разницу между значениями.

Что касается неинвариантности метода относительно угла поворота, то предполагается, что ее можно устранить дополнительными методами. Например, с помощью методов контурного анализа, реализованных в библиотеке OpenCV [16], можно на основе расчета моментов контуров оценить точку центра масс контура и затем сопоставить разницу в местоположении с точкой геометрического центра контура. Таким образом для некоторых видов знаков можно получить предполагаемый угол наклона ДЗ. В рамках данной работы этот простой подход позволил определить направление стрелки на ДЗ: № 4.1.1, № 4.2.1 и № 4.2.2, что позволило повернуть нечитаемый знак h (рис. 5,а) на угол, соответствующий предполагаемому шаблону f , и провести его точное чтение (рис. 5).

ЗАКЛЮЧЕНИЕ

В ходе исследования было установлено, что классификатор, построенный на методах морфологического анализа, безошибочно считывает ДЗ (табл. I и II) независимо от условий регистрации, при маленьком размере исследуемого изображения и при существенных шумах и искажениях. Низкие требования к предобработке и грамотное построение алгоритмов позволяют быстро решать эту задачу, удовлетворяя условиям ре-

Таблица II: Классификация искаженных изображений h_i

№	ДЗ f_i	h_1	h_2	h_3	h_4	h_5
4.1.1	Движ. прямо	0	36	24	23	4.01
4.1.2	Движ. направо	90	0	26	24	3.16
4.1.3	Движ. налево	89	29	24	16	4.22
4.1.4	Движ. прямо или направо	70	35	27	23	3.79
4.1.5	Движ. прямо или налево	50	33	0	19	3.85
4.1.6	Движ. направо или налево	73	25	20	19	3.79
4.2.1	Объезд препят. справа	65	37	24	0	4.22
4.2.2	Объезд препят. слева	85	29	25	25	0
4.2.3	Объезд препят. спр. или сл.	91	25	24	24	4.01
4.3	Круговое движение	98	29	26	24	3.79
4.4	Велосипедная дорожка	98	29	26	24	3.79
4.5	Пешеходная дорожка	40	34	25	18	3.58
4.6	Ограничение мин. скорости	90	27	24	19	3.16

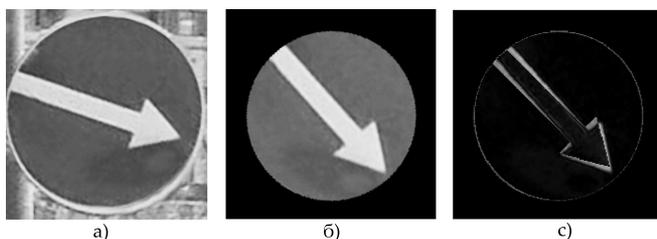


Рис. 5: Поворот знака с помощью методов контурного анализа: (а) нечитаемое изображение h , (б) h' повернутое изображение h , (в) морфологическая разность $h' - \Pi_f h'$, указывающая на почти полное совпадение h' с шаблоном f

ального времени, что в совокупности позволяет применять морфологический классификатор в реальных условиях регистрации.

Однако стоит отметить, что алгоритм не инвариан-

тен относительно поворота и, в качестве априорной информации, требует размеры классифицируемого изображения. Но в рамках решаемой задачи эти недостатки могут быть устранены использованием дополнительных методов, описанных в настоящей статье.

Благодарности

Автор выражает благодарность профессору Ю.П. Пытьеву за разработку и развитие морфологического анализа, а также сотрудникам кафедры математического моделирования и информатики за конструктивную критику и советы на научно-исследовательских семинарах.

Цель работы была поставлена кандидатом физико-математических наук Е.Н. Терентьевым.

- [1] Визильтер Ю.В., Желтов С.Ю., Бондаренко А.В. Обработка и анализ изображений в задачах машинного зрения. М.: Физматкнига, 2010.
- [2] Shustanov A., Yakimov P. Proceedings of the 14th International Joint Conference on e-Business and Telecommunications. 2017. **5**, P. 42.
- [3] Xiong B., Izmirli O. International Workshop on Image Processing and Optical Engineering, International Society for Optics and Photonics. 2012. **8335**, P. 83350B.
- [4] Krsak E., Toth S. Acta Electrotechnica et Informatica. 2011. **11**, №4. P. 31.
- [5] Якимов П.Ю. Издательство СГАУ, Материалы Международной конференции и молодежной школы «Информационные технологии и нанотехнологии». 2016. P. 977.
- [6] Якимов П.Ю. Компьютерная оптика. 2013. **37**, №3. P. 60.
- [7] Шемарулин И.А., Карпычев В.Ю. Труды НГТУ им. П.Е. Алексеева. 2016. №2. P. 60.
- [8] Терентьев Е.Н., Shilin-Terentyev N.E., Приходько И.Н., Фаршакова И.И. Научная конференция «Ломоносовские чтения — 2018». Секция физики. P. 162.
- [9] Медведев М.В., Кирпичников А.П., Синичкина Т.А. Вестник технологического университета **19**, №18. P. 143.
- [10] Keser T., Kramar G., Nozica D. In Proceedings of the IEEE International Conference on Smart Systems and Technologies (SST) P. 12.
- [11] Daraghmi Y-A, Hasasneh A.M. International Journal of Signal Processing Systems. 2016. **19**, №18. P. 12.
- [12] Burger Wilhelm, Burge Mark J. Principles of Digital Image Processing: Core Algorithms. Springer Science & Business Media, 2010.
- [13] Пытьев Ю.П., Чуличков А.И. Методы морфологического анализа изображений. М.:Физматлит, 2010.

- [14] *Пытьев Ю. П. Ж.* вычисл. матем. и матем. физ. 2013. **53**:12. P. 2100.
[15] *Pyt'ev Yu. P.* Pattern Recognition and Image Analysis. 1997. **7**, № 12. P. 467.
[16] *Kaehler A., Bradski G.* Learning OpenCV 3: Computer Vision in C++ with the OpenCV Library. O'Reilly Media, Inc, 2017.

Reading traffic signs in images obtained under uncontrolled registration conditions

I. N. Prikhodko

*Department of Computer Modeling and Informatics, Faculty of Physics, Lomonosov Moscow State University
Moscow 119991, Russia
E-mail: prikhodko_in@mail.ru*

The paper presents an method to solving the problem of reading road signs on images obtained under uncontrolled registration conditions. The registration conditions are taken into account in the morphological analysis applied in this work. This task is divided into the localization of traffic signs of certain colors and shapes, and reading the candidate for road signs.

PACS: 02.70.-c.

Keywords: morphological methods, image form, image analysis, road sign recognition, computer vision, contour analysis.

Received 14 May 2018.

Сведения об авторе

Приходько Ирина Николаевна — студентка магистратуры; e-mail: prikhodko_in@mail.ru.
