

УДК 004.93'11

ЛОКАЛЬНЫЕ БИНАРНЫЕ ШАБЛОНЫ МЕДИАННОГО ПИКСЕЛЯ – ЭФФЕКТИВНЫЕ ИНФОРМАТИВНЫЕ ПРИЗНАКИ ТЕХНОЛОГИИ РАСПОЗНАВАНИЯ ОБРАЗОВ

*Ефимов И.Н., аспирант Самарский государственный университет путей сообщения,
e-mail: Mr.Efimov.IN@gmail.com*

Ключевые слова: компьютерное зрение, локальные бинарные шаблоны, распознавание образов, информативные технологии.

Введение

Технологии компьютерного зрения находят применения в самых разнообразных областях жизни человека, будь то распознавание дорожных знаков, распознавание номеров транспортных средств, распознавание черт лица человека, эмоций или жестов. Создание вышеперечисленных технологий остаётся сложной технической и теоретической задачей.

Для успешного функционирования технологии распознавания образов необходимо обеспечить: высокую вероятность правильного распознавания, возможность работы в реальном времени и успешное определение подмены распознаваемого объекта. Основной задачей при создании технологии распознавания образов является корректный выбор наиболее эффективного информативного признака.

Информативный признак, представленный в данной статье, может эффективно использоваться как для распознавания образов, так и для определения подмены распознаваемого объекта [1]. Скорость вычисления информативного признака позволяет использовать его в реальном времени.

Вычисление кодовых комбинаций ЛБШ

По данным ряда исследований [3, 4], распознавание образов с помощью локальных бинарных шаблонов (далее ЛБШ) обеспечивает требуемые результаты с точки зрения скорости работы и качества распознавания. ЛБШ доказали свою состоятельность в особенности из-за частичной инвариантности к условиям освещения и вычислительной простоте формирования кодов, что позволяет использовать ЛБШ в реальном времени. ЛБШ являются простыми, но очень эффективными текстурными признаками.

В 2002 году группа под руководством Тимо Оджала [7] предложила метод вычисления кодов ЛБШ. Для представления ЛБШ используется обозначение $LBP_{P,R}$, где P – количество рассматриваемых точек окрестности, R – радиус окрестности. Коды ЛБШ вычисляются путем суммирования пороговых различий интенсивности с центром окрестности:

$$LBP_{P,R}(x,y) = \sum_{i=0}^{P-1} s(g_i - g_c) 2^i, \quad (1)$$

Приводятся описание, достоинства и недостатки оригинального информативного признака, основой которого являются локальные бинарные шаблоны. Автором делается попытка изменения формы и алгоритма построения кодовых комбинаций, что является новизной публикации, а также предлагается сравнительный анализ результатов работы оригинального информативного признака. Показаны преимущества признака применительно к распознаванию образов по сравнению с аналогами.

где x, y – координаты центрального пикселя, i – номер рассматриваемого пикселя, g_c – значение яркости центрального пикселя, g_i – яркость рассматриваемого пикселя, а $s(z)$ равна

$$s(z) = \begin{cases} 1, & z \geq 0; \\ 0, & z < 0. \end{cases}$$

В 2008 году были предложены два новых усовершенствованных ЛБШ [4]: Three-PatchЛБШ и Four-PatchЛБШ. Основное отличие предложенных ЛБШ от существующих в том, что сравниваются фрагменты изображения, а не пиксели, по какой-либо метрике (например, Хи-квадрат). Как следует из названия, шаблон Three-PatchЛБШ (TP) вычисляется путем сравнения значений трех фрагментов (рис. 1 а).

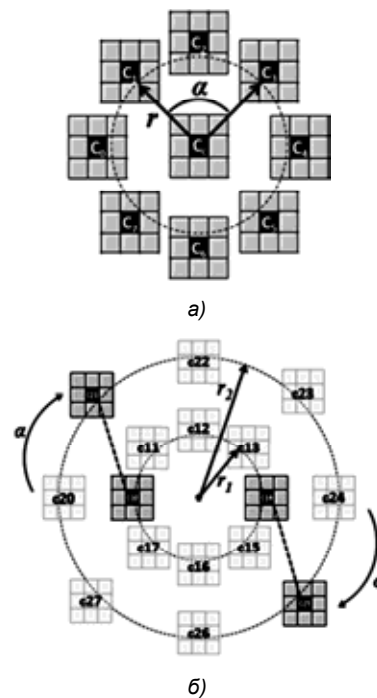


Рис. 1. а) $TP_{8,1,2}$ для 3-х фрагментов для $P=8, r=1$ и $\infty=2$. б) $FP_{8,1,2,1}$ для $P=8, r_1=1, r_2=2$ и $\infty=1$

Алгоритм Three-PatchЛБШ:

Шаг 1. Фрагменты C_i и $C_{i+\alpha}$ сравниваются по некой метрике d со значениями центрального фрагмента C_c , где α – число, определяющее второй фрагмент для сравнения.

Шаг 2. Если окрестность C_i ближе по значениям к C_c , чем $C_{i+\alpha}$ к C_c , то $s(z) = 1$, иначе $s(z) = 0$.

Шаг 3. Расчет производится для каждого пикселя по следующему выражению:

$$TP_{P,R,\alpha}(x,y) = \sum_{i=0}^{P-1} s(d(C_i, C_c) - d(C_{i+\alpha}, C_c)) 2^i, \quad (2)$$

где

$$s(z) = \begin{cases} 1, & z \geq \tau; \\ 0, & z < \tau. \end{cases}$$

На практике используется значение большее нуля (например, $\tau = 0,01$), для обеспечения стабильности в расчетах.

Шаблон Four-PatchЛБШ (FP) вычисляется путем сравнения значений четырех фрагментов (рис. 1 б).

Алгоритм Four-PatchЛБШ:

Шаг 1. Определяются 2 окружности фрагментов с радиусами R_1 и R_2 .

Шаг 2. По некой метрике d попарно сравниваются фрагменты $C_{1,i} - C_{2,i+\alpha}$ и $C_{1,i+p/2} - C_{2,i+p/2+\alpha}$, где α – число, определяющее следующий фрагмент для сравнения.

Шаг 3. Если первая пара фрагментов ближе по значениям друг к другу, чем вторая, то $s(z) = 1$, иначе $s(z) = 0$.

Шаг 4. Расчет производится для каждого пикселя по следующему выражению:

$$FP_{P,r_1,r_2,\alpha}(x,y) = \sum_{i=0}^{P-1} s(d(C_{1,i}, C_{2,i+\alpha}) - d(C_{1,i+p/2}, C_{2,i+p/2+\alpha})) 2^i, \quad (3)$$

где

$$s(z) = \begin{cases} 1, & z \geq \tau; \\ 0, & z < \tau. \end{cases}$$

На практике используется значение большее нуля (например, $\tau = 0,01$), для обеспечения стабильности в расчетах. На рис.6представлены иллюстрации к Three-PatchЛБШ и Four-PatchЛБШ.

Вычисление гистограмм кодовых комбинаций ЛБШМП

На данный момент, помимо Three-PatchЛБШ и Four-PatchЛБШ, существуют альтернативные ЛБШ [6, 9]. Одним общим недостатком является вычислительная сложность алгоритмов. Автором разработан новый вид ЛБШ. Предложенный ЛБШ медианного пикселя (далее ЛБШМП) (рис. 2), существенно ускоряет процесс вычисления шаблона, при сохранении свойств, присущих Three-PatchЛБШ и Four-PatchЛБШ.

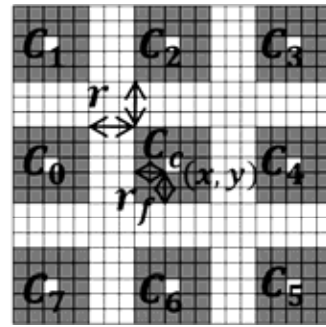


Рис. 2. Схема построения ЛБШМП для $P = 8$, $r = 2$, $P_f = 25$ и $r_f = 2$

Математическая модель ЛБШМП имеет вид:

$$M_{P,r,P_f,r_f}(x,y) = \sum_{i=0}^{P-1} 2^i \cdot S(z_i), \quad (4)$$

где P – количество рассматриваемых фрагментов, r – радиус окрестности, P_f – количество пикселей во фрагменте, r_f – радиус фрагмента, x, y – координаты центрального пикселя, i – номер сравниваемого фрагмента, $i = \overline{0, P-1}$, а выражение s :

$$s(z_i) = \begin{cases} 1, & z_i \geq \tau; \\ 0, & z_i < \tau, \end{cases}$$

$$z_i = \text{med}(C_i) - \text{med}(C_c),$$

$$\text{med}(C_i) = \text{med}(g(0), \dots, g(l), \dots, g(2l)) = g(l),$$

где $\text{med}(C_i)$ – медианное значение яркостей пикселей фрагмента C_i , элементы которого расположены в порядке возрастания значений яркости $g(j)$, т.е. выполняется условие: $g(0) \leq \dots \leq g(l) \leq \dots \leq g(2l)$, j – номер рассматриваемого пикселя фрагмента C_i , $j = \overline{0, 2l}$, $l = 2r_f(r_f + 1) - 1$ – индекс медианного элемента, $\text{med}(C_c)$ – медианное значение центрального фрагмента C_c .

Пороговое значение τ берется большим, чем ноль (например, $\tau = 0,01$), чтобы обеспечить стабильность в расчетах. Выражение для яркости пикселя имеет вид:

$$g(j) = 0,3 \cdot R_j + 0,59 \cdot G_j + 0,11 \cdot B_j, \quad (5)$$

где R_j , G_j и B_j цветовые составляющие красного, зеленого и голубого цвета пикселя. Ниже (рис. 3) представлена схема для нахождения $\text{med}(C_i)$, где каждая ячейка – это значение интенсивности пикселя $g(j)$.

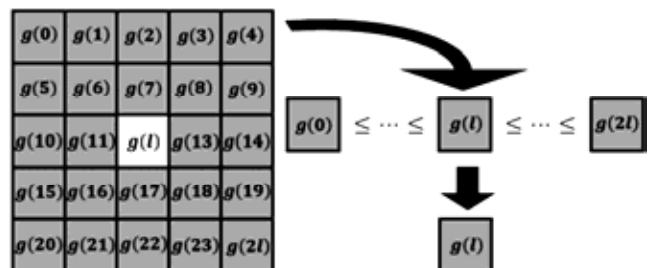


Рис. 3. Поиск медианного значения фрагмента C_i

Дальнейшим развитием ЛБШМП стало изменение формы построения ЛБШ и увеличение числа сравнива-

емых фрагментов (рис. 4). На рис. 4 $b_1 = 8r_f + 4$ и $b_2 = r + 2r_f + 1$ – размеры окрестности фрагментов.

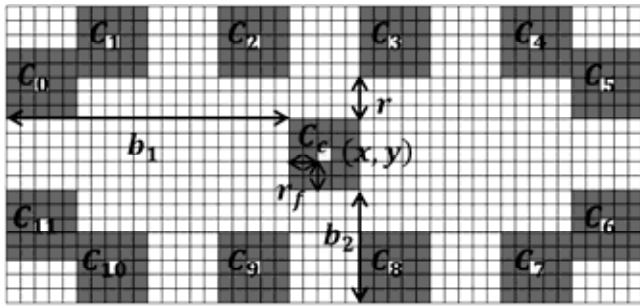


Рис. 4. Схема построения ЛБШМП, с увеличенным числом фрагментов

Для увеличения вероятности правильного распознавания добавлена внешняя структура сравнения. Подобно Four-Patch ЛБШ изменился алгоритм сравнения – фрагменты сравниваются не с центральным блоком, а с фрагментами внешней структуры. Результатом преобразований стала схема, представленная на рис. 5, где $[r/2]$ – оператор выделения целой части от вещественного выражения. Количество и схема размещения фрагментов данной конструкции могут варьироваться в зависимости от требуемого результата.

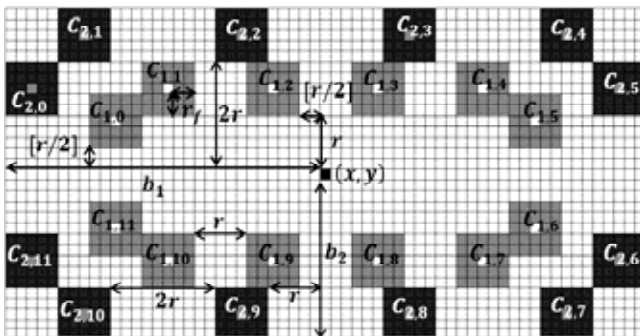


Рис. 5. Схема построения ЛБШМП для $P=12$, $r=5$, $P_f=25$ и $r_f=2$

Для того что бы наглядно увидеть реализацию информативных признаков необходимо: рассчитать для каждого пикселя значение ЛБШ, привести значения ЛБШ к общему множеству (0..256), для этого необходимо нормализовать значения ЛБШ, установить цветовые составляющие красного, зеленого и голубого цвета пикселя равными значению ЛБШ этого пикселя. На рис.6 представлена иллюстрация к ЛБШМП.

В операции распознавания образов используются гистограммы кодовых комбинаций ЛБШ. Гистограмма ЛБШ, построенная одновременно для всего изображения, содержит только текстурные особенности без информации о расположении кодов на изображении. Для учета информации о расположении особенностей, изображение предлагается разбивать по одной из следующих схем (7 а и 7 б) (в первоначальном способе Т.Анопен предлагается разделять изображение на 7x7 блоков), в каждом регионе вычисляется своя гистограмма кодовых комбинаций. Гистограммы для каждой области объединяются в одну общую (рис. 8), при этом общая гистограмма содержит

текстурные особенности и информацию о расположении кодовых комбинаций ЛБШ.

Ниже приведены примеры комбинаций деления изображения лица на регионы и примеры гистограмм.

Таким образом, математическая модель информативного признака гистограммы кодов ЛБШМП со схемы 5 будет иметь вид:

$$H_{P,r,P_f,r_f}(h,k) = V_k \sum_{x=b_1}^{W_k-b_1-1} \sum_{y=b_2}^{H_k-b_2-1} f(M_{P,r,P_f,r_f}(x,y),h), \quad (6)$$

где P – количество рассматриваемых фрагментов, r – ближайшее расстояние от центра до 1-й окрестности фрагментов (рис. 5), P_f – количество рассматриваемых пикселей во фрагменте, r_f – радиус рассматриваемых фрагментов, h – номер столбца гистограммы, $h = \overline{0, 2^p - 1}$, k – номер региона гистограммы, $k = \overline{1, K}$, K – количество регионов разделяющих изображение, V_k – весовой коэффициент k -го региона, x, y – координаты центрального пикселя, $b_1 = 3r + 6r_f + 3$ и $b_2 = 2r + 2r_f + 1$ – размеры 2-й окрестности, а функция f имеет вид:

$$f(M_{P,r,P_f,r_f}(x,y),h) = \begin{cases} 1, & M_{P,r,P_f,r_f}(x,y) = h; \\ 0, & M_{P,r,P_f,r_f}(x,y) \neq h, \end{cases}$$

где $M_{P,r,P_f,r_f}(x,y)$ – функция расчета кодовой комбинации ЛБШМП, основной операцией которой является сравнение фрагментов окрестностей 1 и 2:

$$M_{P,r,P_f,r_f}(x,y) = \sum_{i=0}^{P-1} s(z_i) 2^i, \quad (7)$$

$$s(z_i) = \begin{cases} 1, & z_i \geq \tau; \\ 0, & z_i < \tau, \end{cases}$$

$$z_i = \text{med}(C_{1,i}) - \text{med}(C_{2,i}),$$

$$\text{med}(C_{q,i}) = \text{med}(g(0), \dots, g(l), \dots, g(2l)) = g(l),$$

где z_i – разность медианных значений фрагментов, g – порядковый номер окрестности, $q = \overline{1, 2}$, i – номер шага сравнения фрагментов, $i = \overline{0, P-1}$, τ – пороговое значение берется большим, чем ноль (например, $\tau = 0,01$), $g(j)$ – значение яркости пикселя выражение для нахождения яркости представлено в 5, $\text{med}(C_{q,i})$ – медианное значение фрагмента $C_{q,i}$, элементы которого, расположены в порядке возрастания значений яркости $g(j)$, т.е. выполняется условие: $g(0) \leq \dots \leq g(l) \leq \dots \leq g(2l)$. На рис. 3 представлен пример вычисления $\text{med}(C_{q,i})$.

Алгоритм вычисления гистограмм ЛБШМП:

Шаг 1. Изображение делится на регионы, для каждого региона, в зависимости от его ценности, устанавливаются значения весов V_k (например, для глаз $V_k = 1, 2$, для носа $V_k = 1, 1$, для рта $V_k = 1$).

Шаг 2. Вычисляется медианное значение для каждого фрагмента окрестности 1 и окрестности 2.



а) изображение; б) представление – $LBP_{8,1}$; в) представление – $TP_{8,1,2}$.

г) Представление – $FP_{8,1,2,1}$; д) Представление – $M_{12,5,8,1}$.

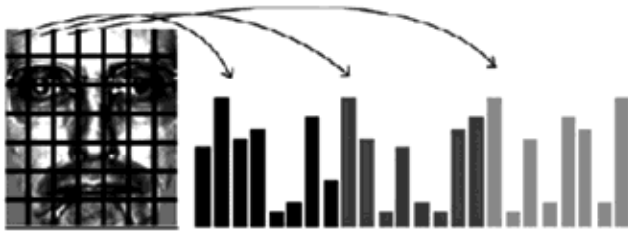
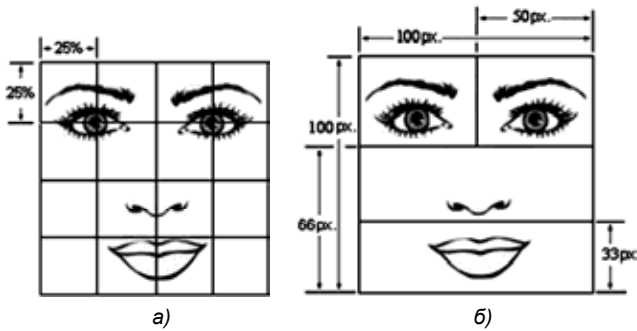
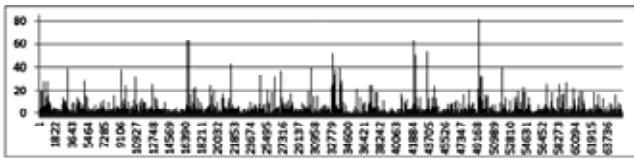


Рис. 7. Пример построения общей гистограммы

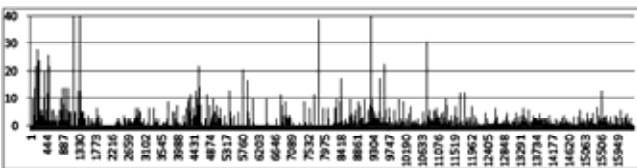


а)

б)



е)



е)

Рис. 8. а) схема из 16 регионов – 4×4 ;

б) схема из 4 регионов – 3×2 , с областями глаз, носа и рта;

в) гистограмма из 16 регионов – 4×4 ;

г) гистограмма из 4 регионов – 3×2 ,

с областями глаз, носа и рта

Шаг 3. Для каждой связки фрагментов, вычисляется z_i . Если z_i меньше порогового τ , то результатом будет – 0, иначе 1.

Шаг 4. Результаты сравнения переводятся в десятичную систему счисления и объединяются в одну общую кодовую комбинацию.

Шаг 5. Полученное число на шаге 4 является номе-

ром столбца h гистограммы H_{P_r, P_f, r_f} , к значению которого добавляется 1.

Шаг 6. Собранные гистограммы регионов объединяются в одну общую гистограмму с учетом веса региона.

Результаты работы

Информативные признаки тестировались на базе данных (далее БД) (табл. 1), полностью включающей изображения баз Калифорнийского технологического института [10], TheIndianFaceDatabase[5] и выборочно FERET[8]. Изображения подбирались таким образом, чтобы избежать перекрытий регионов лица, больших поворотов головы ($>20^\circ$) и сильных эмоций. В базе содержатся изображения лиц людей различных рас, возрастов и мимики, изображения подготавливались с различными условиями съемки – освещением, масштабом и ракурсом. Алгоритм предобработки и приведения изображений к общей форме приведен в [2].

Таблица 1.

	Тестовая выборка	Обучающая выборка
Зарегистрированные	1000 изображений, 100 человек	200 изображений, 100 человек
Отсутствующие в БД	200 изображений различных лиц	

Применение в роли информативного признака гистограмм ЛБШМП, позволило добиться ожидаемых результатов распознавания. Для сопоставления информативных признаков использовалась функция расчета расстояния – ядро пересечения гистограмм. При построении информативных признаков использовалась схема на рис. 7 б. Процент случаев выдачи доступа зарегистрированному пользователю составил 98,9, а процент выдачи доступа человеку, не имеющему допуск – 3,5. Использование информативного признака гистограмм кодов ЛБШМП (рис. 5) увеличивает вероятность правильного распознавания на 13 % и 7 % по сравнению с Three-PatchЛБШ и Four-PatchЛБШ на подобранной БД.

Использование информативного признака гистограмм кодов ЛБШМП (рис. 5) увеличивает скорость работы в 1,7 и 2 раза по сравнению с Three-PatchЛБШ и Four-PatchЛБШ, соответственно. Ниже в табл. 2. выполнен сравнительный анализ скоростных характеристик с аналогами алгоритма, выполняемых на изображении размером 100×100 пикселей.

Таблица 2.

Алгоритм	Время выполнения
Three-Patch ЛБШ	107 мс
Four-Patch ЛБШ	124 мс
ЛБШМП	61 мс

Заключение

В данной работе гистограммы ЛБШМП хорошо зарекомендовали себя в процессе распознавания лиц. Разработанные информативные признаки найдут свое применение в самых различных областях компьютерного зрения, таких как распознавание дорожных знаков, распознавание номеров транспортных средств и т.д.

Основной задачей, которую необходимо решить при использовании гистограмм ЛБШМП, является наиболее корректный подбор параметров информативного признака.

Представленный алгоритм вычисления информативных признаков можно модифицировать: добавлять новые фрагменты или изменять расположение фрагментов, с условием сохранения основной идеи исследования.

Статья подготовлена при поддержке фонда содействия развитию малых форм предприятий в научно-технической сфере в номинации «Информационные технологии» («У.М.Н.И.К.» I полугодие г. Самара 2014).

Литература

1. Ефимов И.Н. Методы повышения достоверности идентификации обучаемых при дистанционном обучении / И.Н. Ефимов // Дистанционное и виртуальное обучение – 2013. – № 3 – 62-73 с.
2. Ефимов И.Н. Биометрическая идентификация в дистанционной подготовке кадров на ж. д. транспорте / И.Н. Ефимов, А.М. Косолапов // Вестник транспорта поволжья – 2013. – № 1 – 62-66 с.

3. Ahonen T. Face Recognition with Local Binary Patterns, 2004. – 469-481 с.
4. Ahonen T. Face description with local binary patterns: Application to face recognition / T. Ahonen, A. Hadid, M. Pietikäinen // IEEE Trans. Pattern Anal. Mach. Intell. – 2006. – Т. 28 – № 12 – 2037–2041с.
5. Jain V. The Indian face database / V. Jain, A. Mukherjee – 2002.
6. O'Connor B. Facial Recognition using Modified Local Binary Pattern and Random Forest / B. O'Connor, K. Roy // Int. J. Artif. Intell. Appl. – 2013. – Т. 4 – № 5.
7. Ojala T. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns / T. Ojala, M. Pietikainen, T. Maenpaa // IEEE Trans. Pattern Anal. Mach. Intell. – 2002. – Т. 24 – № 7.
8. Phillips P.J. The facial recognition technology (FERET) database / P. J. Phillips // IEEE Trans. Pattern Anal. Mach. Intell. – 2004. – Т. 22.
9. Trefný J. Extended set of local binary patterns for rapid object detection / J. Trefný, J. Matas // Proc. Comput. Vis. Winter – 2010.
10. Weber M. Frontal face dataset, california institute of technology // – 1999.

LOCAL BINARY PATTERNS MEDIAN PIXEL'S – EFFECTIVE TECHNOLOGY OF PATTERN RECOGNITION

Efimov I.N,

In this paper author propose the new informative features local binary patterns median pixel's (LBPMP). The author changed the shape and the algorithm for constructing local binary patterns. Author compared the advantages of the new feature in relation to pattern recognition.

УВАЖАЕМЫЕ КОЛЛЕГИ!

Для тех, кто не успел оформить подписку на первое полугодие 2015 года через ОАО «Роспечать», сохраняется возможность приобретения журналов непосредственно в редакции по адресу: г. Москва, ул. Авиамоторная, дом 8, Научный Центр МТУСИ, ком. 612. Российское научно-техническое общество радиотехники, электроники и связи им. А.С. Попова, метро «Авиамоторная», или оформить Заказ в соответствии с требованиями, выставленными на сайте журнала: www.dsra.ru.

Справки по телефонам:

(8-495) 362-42-75; (+7 903) 221-79-79 (Алексеева Любовь Ильинична).

E-mail: rntores@mail.ru