

## МЕТОД НЕПАРАМЕТРИЧЕСКОЙ КЛАССИФИКАЦИИ В РАСПОЗНАВАНИИ ОБРАЗОВ

*Рассмотрены методы непараметрической классификации в задачах распознавания образов. Кратко описаны два известных метода —  $k$ -NN ( $k$  ближайших соседа) и окно Парзена. Оба метода расширены для приложений с нечетким обучающим набором. Предложен новый гибридный метод, сочетающий преимущества метода  $k$ -NN и окна Парзена.*

Целью исследований в области распознавания образов (pattern recognition) [1–4] является создание систем, обеспечивающих автоматическую классификацию наблюдаемых объектов по заранее предопределенным или даже неизвестным классам (типам, группам, ситуациям). Образ поступает в систему в виде сигналов, например, звука, изображения или показаний измерительных приборов. Результатом распознавания, в общем случае, является вектор значений, который определяет вероятность или степень принадлежности образа к предварительно определенным классам.

Системы распознавания образов успешно используются в приложениях различного уровня сложности: управление роботами, контроль качества в производстве, системы безопасности, распознавание печатных текстов, медицинская диагностика. Распознавание образов тесно связано с такими понятиями как искусственный интеллект, система поддержки принятия решений и экспертная система.

В работах [4, 5] авторы выделяют следующие подходы к построению систем распознавания: статистическая классификация; сравнение с шаблоном; синтаксическое или структурное сравнение; нейронные сети. Эти подходы не являются независимыми; одинаковые методы распознавания могут иметь разную интерпретацию (см., например, [6]). Очевидно, что пока не существует универсального метода и выбор оптимального метода зависит от условий задачи: типа входной информации, сложности образов, требуемой степени обобщения и ожидаемого результата распознавания.

Настоящая статья посвящена методу на основе первого из упомянутых подходов, так называемой *статистической классификации* [7–10]. Особенностью статистического подхода является то, что распознаваемый объект задается точкой в  $D$ -мерном пространстве признаков. Признак представляет собой число, характеризующее образ, например,

для изображения могут использоваться такие признаки, как площадь объекта, параметры гистограммы интенсивностей, фрактальная размерность. Разработка системы распознавания включает такие задачи, как определение классов, выбор признаков, создание классификатора, обучение системы и оценка качества распознавания. На этапе выбора признаков главной целью является нахождение такого пространства признаков, в котором объекты разных классов представлены непересекающимися подмножествами. Эффективность статистического классификатора будет зависеть от того, насколько различимы подмножества классов.

**Задача распознавания образов.** Обычно под задачей распознавания понимается классификация образов после предварительного обучения учителем<sup>1</sup> (supervised learning). В процессе обучения, знания экспертов формализуются и записываются в память системы распознавания. Знание может быть представлено набором обучающих образов, снабженных экспертными оценками. После обучения, система должна уметь обобщать накопленное знание и классифицировать неизвестные образы автоматически.

**Обучение.** Рассмотрим базовую схему непараметрической системы распознавания с учителем (рисунок). Эксперты формируют обучающую выборку из  $N$  образов и определяют множество из  $K$  классов (блок “Выделение классов”).

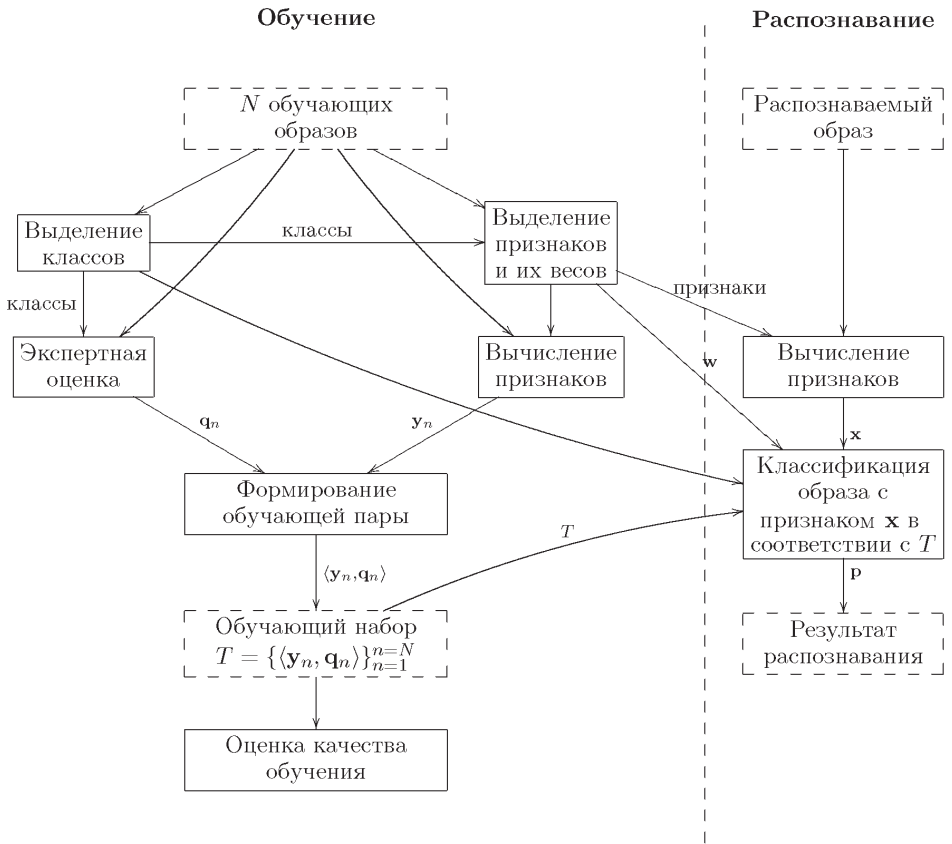
Затем эксперты определяют признаки (блок “Выделение признаков”), на основе которых будет происходить классификация. Другими словами, задается преобразование образа (сигнала) в вектор признаков  $\mathbf{y} \in R^D$ . Признаки могут иметь разный масштаб и вес, для учета которых вводится нормирующий вектор  $\mathbf{w} \in R^D$ . Методы выделения признаков широко освещены в литературе, например [10–12]. В блоке “Вычисление признаков” каждый  $n$ -й обучающий образ сопоставляется с вектором признаков  $\mathbf{y}_n \in R^D$ .

Каждый  $n$ -й образ выборки получает экспертную оценку в виде вектора

$$\mathbf{q}_n = (q_{1,n}, q_{2,n}, \dots, q_{K,n}), \quad 0 \leq q_{j,n} \leq 1, \quad \sum_{j=1}^K q_{j,n} = 1,$$

где  $q_{j,n}$  — степень принадлежности  $n$ -го образа к  $j$ -му классу, а сумма есть условие нормировки. Значение  $q_{j,n}$  можно интерпретировать как *оценку вероятности* — при отсутствии достоверной информации эксперты оценивают вероятность  $j$ -го класса, т.е. условную вероятность

<sup>1</sup>Распознавание без учителя не затрагивается в настоящей работе.



**Общая схема обучения и распознавания образов**

$q_{j,n} = P(j\text{-й класс} | y_n)$  и как *степень принадлежности*, если образ по своей сути может принадлежать к нескольким классам одновременно и классификация задается нечеткими (мягкими) множествами, т.е.  $q_{j,n} = \chi_j(y_n)$ , где  $\chi_j$  — функция принадлежности  $j$ -го класса (membership function).

В обоих случаях методы распознавания оказываются одинаковыми.

Из вектора признаков и вектора принадлежностей, заданных экспертами, формируется обучающая пара  $\langle y, q \rangle$ . Результатом обучения является множество обучающих пар

$$T = \{ \langle y_1, q_1 \rangle, \langle y_2, q_2 \rangle, \dots, \langle y_N, q_N \rangle \}.$$

**Распознавание.** На вход системы поступает неизвестный образ. В блоке “Вычисление признаков” рассчитывается вектор  $x \in R^D$  по алгоритму, определенному в процессе обучения. Классифицировать (или распознать) образ означает найти вектор

$$\mathbf{p} = (p_1, p_2, \dots, p_K), \quad 0 \leq p_j \leq 1, \quad \sum_j^K p_{j=1} = 1,$$

где  $(p_j)$  показывает вероятность  $j$ -го класса при условии  $\mathbf{x}$  или, в другой интерпретации, степень принадлежности точки  $\mathbf{x}$  к  $j$ -му классу. Тогда задача построения классификатора (блок “Классификация”) сводится к нахождению функции

$$p_j = P(j\text{-й класс}|\mathbf{x})$$

на основе обучающей информации  $T$ . Для нескольких классов, достаточно использовать векторную запись вероятностей  $\mathbf{p} = P(\mathbf{x})$ .

**Параметрические и непараметрические классификаторы.** Статистические классификаторы можно разделить на две группы с точки зрения способа оценки вероятностей класса:

1. В параметрических методах необходима предварительная информация о форме функции плотности вероятности  $(p_j)$ , и в процессе обучения вычисляется оценка параметров этой функции. Например, часто используется функция нормального распределения

$$p_j = p_j(x, \mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}},$$

где  $\mu$  и  $\sigma$  — параметры, получаемые из обучающего набора  $T$ . Этот подход не эффективен, когда нет предварительной информации о форме функции плотности распределения или она имеет сложный вид (например, несколько локальных максимумов).

2. Непараметрические методы позволяют получить оценку вероятности напрямую из обучающего набора  $T$ :

$$p_j = p_j(x, T).$$

Непараметрические методы могут работать со сложными многомерными функциями распределения. Очевидно, такой подход более универсальный, но ресурсоемкий.

Далее речь пойдет о непараметрических классификаторах. В работе [1] выделяют два эффективных метода непараметрической классификации — окно Парзена (Parzen window) и  $k$ -NN ( $k$  Nearest Neighbours —  $k$  ближайших соседей). Ряд альтернативных подходов [4] обеспечивает высокое качество распознавания, но два упомянутых метода привлекательны своей простотой и универсальностью. В их основе лежит мера сходства/различия между образами.

**Мера сходства/различия.** Пусть необходимо сравнить два образа с признаками  $\mathbf{x}$  и  $\mathbf{y}$ . Распространенной мерой является евклидово расстояние

$$d(\mathbf{x}, \mathbf{y}) = \sqrt{\sum_{d=1}^D (x_d - y_d)^2},$$

заданное в  $D$ -мерном пространстве признаков. При расчете меры важно учитывать масштаб признаков, так как признаки с большим диапазоном значений подавляют остальные. Для этого признаки могут быть нормализованы:

$$d(\mathbf{x}, \mathbf{y}) = \sqrt{\sum_{d=1}^D w_d^2 (x_d - y_d)^2},$$

где вектор нормализующих коэффициентов  $\mathbf{w} = (w_1, \dots, w_D)$  определяется в процессе обучения. В векторе  $\mathbf{w}$  может быть так же учтен вес признака.

Евклидова мера работает с пространствами небольшой размерности ( $D < 30$ ). При больших  $D$  расстояние между двумя случайными точками стремится к одной величине, а все точки находятся на поверхности гиперсферы. В литературе известны и альтернативные меры, такие как  $l_n$ -норма, супремум-норма и расстояние Махаланобиса, однако в задачах распознавания они используются реже.

**Окно Парзена.** Пусть дана выборка  $\{\mathbf{y}_n\}$  из  $N$  точек. Тогда оценка вероятности в точке  $\mathbf{x}$  будет

$$\hat{p}(\mathbf{x}) = \frac{1}{N\sigma} \sum_{n=1}^N g\left(\frac{d(\mathbf{x}, \mathbf{y}_n)}{\sigma}\right), \quad (1)$$

где  $g(\cdot)$  — “окно” или ядро — гладкая сглаживающая функция;  $\sigma$  — относительный размер окна. В пределе  $N \rightarrow \infty$  и  $\sigma \rightarrow 0$  функция  $\hat{p}(\mathbf{x})$  приближается к действительной плотности распределения  $p(\mathbf{x})$  [13]. Окном часто является гауссовская функция, тогда

$$\hat{p}(\mathbf{x}) = \frac{1}{N\sigma\sqrt{2\pi}} \sum_{n=1}^N \exp\left(-\frac{d^2(\mathbf{x}, \mathbf{y}_n)}{2\sigma^2}\right). \quad (2)$$

Основной проблемой данного подхода является выбор размера окна  $\sigma$ . В случае, когда точки  $\mathbf{y}_n$  неравномерно распределены в пространстве признаков,  $\sigma$  должна адаптивно выбираться исходя из их локальной плотности. В работе [14] предложен метод адаптивной подстройки

размера окна на основе оценки дисперсии в точке  $\mathbf{x}$ . Так же следует отметить, что выражение (1) описывает идеальный случай и не учитывает возможные ошибки признаков  $\mathbf{x}$  и  $y_n$ .

Вычисление суммы (1) ресурсоемко в случае большой выборки. Ряд методов [15, 16] был предложен для решения этой проблемы, в частности на базе оптимального сокращенного набора из  $\{y_n\}$  [17].

***k* ближайших соседей (*k*-NN).** Пусть дана выборка  $\{y_n\}$  из  $N$  точек. Выберем число соседних точек  $k < N$ , например  $k = \sqrt{N}$  [1]. Тогда оценка ПР в точке  $\mathbf{x}$  будет

$$\hat{p}(\mathbf{x}) = \frac{k}{NV_k(\mathbf{x})},$$

где  $V_k(\mathbf{x})$  — объем наименьшей гиперсферы, которая содержит  $k$  точек, ближайших к  $\mathbf{x}$ . Другими словами, для того чтобы отнести точку  $\mathbf{x}$  к тому или иному классу, необходимо зафиксировать  $k$  и сравнить размеры (радиусы) окрестностей  $V_k(\mathbf{x})$ , покрывающие  $k$  точек из  $\{y_n\}$ , ближайших к  $\mathbf{x}$ . Тогда класс с наименьшим  $V_k(\mathbf{x})$  будет наиболее вероятным.

Недостатком метода  $k$ -NN является то, что в нем не учитываются расстояния между рассматриваемой точкой  $\mathbf{x}$  и точками выборки  $\{y_n\}$  внутри окрестности  $V_k(\mathbf{x})$ . Очевидно, ближайшие точки должны иметь больший вес, чем остальные. Другой проблемой метода является выбор оптимального  $k$  в случае неоднородности выборок по каждому классу.

**Новый метод: нечеткий обучающий набор.** В рассмотренных подходах Парзена и  $k$ -NN предполагается, что все элементы обучающего набора имеют одинаковый вес и принадлежат только к одному из классов, т.е.

$$q_{j,n} \in \{0, 1\}, \quad \sum_{j=1}^K q_{j,n} = 1,$$

где  $q_{j,n}$  — экспертная оценка,  $j$  — номер класса и  $n$  — номер обучающего элемента. Расширим оба рассмотренных метода для классификации с нечетким обучающим набором, для которого

$$0 \leq q_{j,n} \leq 1, \quad \sum_{j=1}^K q_{j,n} = 1.$$

**Окно Парзена.** Для реализации взвешенного суммирования добавим множитель  $q_{j,n}$  в выражение (1):

$$\hat{p}_j(\mathbf{x}) = \frac{1}{N\sigma\sqrt{2\pi}} \sum_{n=1}^N q_{j,n} \exp\left(-\frac{d^2(\mathbf{x}, y_n)}{2\sigma^2}\right). \quad (3)$$

Под суммой (3) теперь стоит произведение. Первый сомножитель представляет собой вес, соответствующий вероятности  $j$ -го класса для  $n$ -го элемента обучающего набора. Второй сомножитель, гауссовская функция, задает вес в зависимости от расстояния между  $y_n$  и  $x$ . Чем больше расстояние, тем меньше вклад в сумму от данного элемента.

**$k$  ближайших соседей ( $k$ -NN).** Для выражения

$$\hat{p}(\mathbf{x}) = \frac{k}{NV_k(\mathbf{x})}$$

заменяем **определение**  $V_{j,k}(\mathbf{x})$  — объем наименьшей гиперсферы, которая содержит  $k$  точек  $j$ -го класса из обучающего набора, ближайших к  $x$  — на **определение**  $V_{j,Q(k)}(\mathbf{x})$  — объем наименьшей гиперсферы, такой что

$$\sum_{\substack{1 \leq l \leq N \\ y_l \in V_{j,Q(k)}(\mathbf{x})}} q_{j,l} \geq Q_j(k),$$

где

$$Q_j(k) = \frac{k \sum_{n=1}^N q_{j,n}}{N} \quad (4)$$

определяет суммарный вес точек в объеме  $V_{j,Q(k)}(\mathbf{x})$  для  $j$ -го класса и  $k > 0$  имеет такой же смысл, как в  $k$ -NN, но может принимать дробное значение.

Будем называть такой метод  $Q$ -NN. Его основное отличие от  $k$ -NN заключается в том, что число точек обучающего набора, используемое для классификации в данной окрестности, не является фиксированным, а выбирается адаптивно в зависимости от их веса. Чем больше вес точек  $y_n$  в окрестности  $x$ , тем меньше точек рассматривается. Формула (4) — один из простых способов выбора суммарного веса окрестности. На практике проще оказывается выбрать одно значение экспериментально сразу для всех классов:

$$Q = Q_1 = \dots = Q_K.$$

**Новый метод: гибрид окна Парзена и  $k$ -NN.** Для решения главной проблемы рассмотренных подходов — выбора значения  $\sigma$  и  $k$  — рассмотрим следующий адаптивный метод. Запишем оценку вероятности Парзена (3) как функцию, зависящую от  $x$  и  $\sigma$ :

$$\hat{p}_j(\mathbf{x}, \sigma) = \frac{1}{N\sigma\sqrt{2\pi}} \sum_{n=1}^N q_{j,n} \exp\left(-\frac{d^2(\mathbf{x}, y_n)}{2\sigma^2}\right). \quad (5)$$

На практике обучающий набор обычно не нормирован ввиду того, что экспертная оценка  $q_{j,n}$  является неравномерной по пространству признаков и несбалансированной между классами, т.е.

$$\sum_{n=1}^N q_{i,n} \neq \sum_{n=1}^N q_{j,n}, \quad i \neq j, \quad i, j \in 1, 2, \dots, K.$$

Поэтому целесообразно рассматривать нормированную оценку вероятности

$$\tilde{p}_j(\mathbf{x}, \sigma) = \hat{p}_j(\mathbf{x}, \sigma) / \sum_{j=1}^K \hat{p}_j(\mathbf{x}, \sigma). \quad (6)$$

Нормировочный коэффициент  $1/(N\sigma\sqrt{2\pi})$  выражения (6) сокращается в отношении (5), поэтому далее при расчетах его будем опускать.

Объединим методы Парзена и  $Q$ -NN в гибридный метод, суть которого заключается в адаптивном выборе  $\sigma$  в соответствии с принципом  $Q$ -NN:

$$\sigma = \min \{ \sigma' | \exists \hat{p}_j(\mathbf{x}, \sigma') \geq Q, \quad 1 \leq j \leq K \}, \quad (7)$$

где

$$\hat{p}_j(\mathbf{x}, \sigma') = \sum_{n=1}^N q_{j,n} \exp \left( \frac{d^2(\mathbf{x}, \mathbf{y}_n)}{2\sigma'^2} \right). \quad (8)$$

Значение  $Q_j(k)$  выбирается для каждого класса в зависимости от суммарного веса экспертных оценок “Ошибка!” и плотности точек  $\mathbf{y}_n$  в пространстве признаков. Один из способов выбора оценки “Ошибка!” — формула (4).

Альтернативным подходом к адаптивной подстройке  $\sigma$  является не минимальное значение в условии (7), а максимальное различие нормированных вероятностей:

$$\sigma = \min (\max \arg_{\sigma} \tilde{p}_j(\mathbf{x}, \sigma)) : \quad \exists \tilde{p}_j(\mathbf{x}, \sigma) \geq Q_j, \quad 1 \leq j \leq K,$$

Гибридный метод может быть реализован при помощи алгоритмов [15–17, 19], оптимизирующих поиск ближайших соседей и вычисление оценки  $\hat{p}_j$ .

**Выводы.** 1. Рассмотрены методы непараметрической классификации  $k$ -NN и окно Парзена. Эти методы расширены для приложений с нечетким обучающим набором. Введение нечеткости при помощи вектора  $\mathbf{q}_n$  в обучающий набор позволяет задавать классификацию в терминах теории вероятности или мягких множеств, а так же учитывать такие факторы, как точность измерений признаков.



2. Получен новый гибридный метод, объединяющий преимущества подходов  $k$ -NN и Парзена. Гибридный метод повышает точность распознавания образа с признаками  $\mathbf{x}$  за счет адаптивной подстройки оконного параметра  $\sigma$  в зависимости от плотности и веса обучающих пар  $\langle \mathbf{y}_n, \mathbf{q}_n \rangle$  в окрестности  $\mathbf{x}$ . Описанный гибридный метод апробирован в экспертной системе распознавания бумажных документов.

## СПИСОК ЛИТЕРАТУРЫ

1. D u d a, R. O. Pattern Classification and Scene Analysis. second ed. / R.O.Duda, P.E. Hart, D.G. Stork. – NY: John Wiley & Sons, 2000.
2. T h e o d o r i d i s, S. Pattern Recognition / S. Theodoridis, K. Koutroumbas. — New York: Academic Press, 1999.
3. D a v i e s, E. R. Machine Vision: Theory, Algorithms, Practicalities / E.R. Davies. – Oxford: Academic Press, 1997.
4. J a i n, A. K. Statistical pattern recognition: A review / A.K.Jain, R.P.W.Duin, J. Mao // IEEE Transactions on Pattern Analysis and Machine Intelligence. – 2000. – Vol. 22, no. 1. – P. 4–37. [citeseer.ist.psu.edu/jain99statistical.html](http://citeseer.ist.psu.edu/jain99statistical.html).
5. S c h a l k o f f, R. Pattern recognition: Statistical, structural and neural approaches / R. Schalkoff. – Wiley, 1992.
6. S c h u r m a n n, J. Pattern classification, a unified view of statistical and neural approaches / J. Schurmann. – John Wiley & Sons, 1996.
7. W e b b, A. Statistical Pattern Recognition / A. Webb. – London: Arnold, 1999.
8. F u k u n a g a, K. Introduction to statistical pattern recognition, second edition / K. Fukunaga. – Academic Press, 1990.
9. M c L a c h l a n, G. Discriminant Analysis and Statistical Pattern Recognition / G. McLachlan. – John Wiley & Sons, 1992.
10. D e v r o y e, L. A probabilistic theory of pattern recognition / L. Devroye, L. Gyorf, G. Lugosi. – Springer, 1996.
11. H a s t i e, T. The Elements of Statistical Learning / T. Hastie, R. Tibishirani, J. Friedman. – Berlin: Springer, 2001.
12. V a p n i k, V. Statistical Learning Theory / V. Vapnik. – New York: John Wiley & Sons, 1998.
13. P a r z e n, E. On estimation of a probability density function and mode / E. Parzen // Annals of Math. Statistics. – 1962. – Vol. 33. – P. 1065–1076.
14. K a t k o v n i k, V. Nonparametric density estimation with adaptive varying window size. – Signal Processing Laboratory, Tampere University of Technology. – 2000. <http://www2.mdanderson.org/app/ilya/Publications/eurotoparzen.pdf>
15. I z e n m a n, A. J. Recent developments in nonparametric density estimation / A. J. Izenman // J. Am. Statistical Assoc. – 1991. – Vol. 86. – P. 205–224.
16. J e o n, B. Fast parzen density estimation using clustering-based branch and bound / B. Jeon, D. Landgrebe // IEEE Trans. Pattern Analysis and Machine Intelligence. – 1994. – Vol. 16, no. 9. – P. 950–954.
17. G i r o l a m i, M. Probability density estimation from optimally condensed data samples / M. Girolami, C. He // IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE. – 2003. – Vol. 25, no. 10. – P. 1253–1264.
18. V a h o. A Data structures and algorithms / A. V.Aho, J.E. Hopcroft, J.D. Ullman. – Addison-Wesley, 1983.

19. O m a c h i, S. A fast algorithm for a  $k$ -NN classifier based on branch and bound method and computational quantity estimation. <http://citeseer.ist.psu.edu/611715.html>.

Статья поступила в редакцию 5.07.2005

Валерий Михайлович Черненко родился в 1941 г., окончил МВТУ им. Н.Э. Баумана в 1964 г. Д-р техн. наук, профессор, заведующий кафедрой “Системы обработки информации и управления” МГТУ им. Н.Э. Баумана. Академик Международной академии информатизации. Автор 105 научных работ в области моделирования и системного анализа.

V.M. Chornenkiy (b. 1941) graduated from the Bauman Moscow Higher Technical School in 1964. D. Sc. (Eng.). professor, head of “Systems of Data Processing and Control” department of the Bauman Moscow State Technical University. Academician of the International Academy of Information Technology. Author of 105 publications in the field of modeling and system analysis.

Николай Валентинович Птицын родился в 1979 г., окончил МГТУ им. Н.Э. Баумана в 2002 г. Аспирант кафедры “Системы обработки информации и управления” МГТУ им. Н.Э. Баумана. Автор 8 научных работ в области системного анализа.

N.V. Ptitsyn (1979) graduated from the Bauman Moscow State Technical University in 2002. Post-graduate of “Systems of Data Processing and Control” department of the Bauman Moscow State Technical University. Author of 8 publications in the field of system analysis.

УДК 519.6

Н. В. П т и ц ы н

## **РАЗНОСТНЫЙ МЕТОД ИНТЕГРАЛЬНОГО ПРЕОБРАЗОВАНИЯ ИЗОБРАЖЕНИЯ**

*Рассмотрен метод дискретного интегрального преобразования, основанный на вычислении разности между соседними суммами. Исследовано приложение метода к обработке изображения. Представлено сравнение с известными алгоритмами быстрой свертки.*

Дискретные интегральные преобразования [1, 2] играют важную роль в задачах распознавания образов. С их помощью строятся фильтры для подавления шума, выделения признаков, сегментации и других операций, связанных с обработкой сигналов, таких как звук и изображение. Интегральные преобразования являются ресурсоемкими, особенно для многомерных сигналов и больших ядер. Настоящая статья представляет новый метод интегрального преобразования изображения, который в определенных случаях превосходит существующие методы с точки зрения точности, удобства реализации и быстродействия.