

АНАЛИЗ И СИНТЕЗ СИГНАЛОВ И ИЗОБРАЖЕНИЙ

УДК 004.93

ТЕОРЕТИКО-ИНФОРМАЦИОННЫЕ ГРАНИЦЫ ТОЧНОСТИ КЛАССИФИКАЦИИ ОБЪЕКТОВ В ПРОСТРАНСТВАХ ПРЕДСТАВЛЕНИЙ С ЗАДАННЫМИ РАССТОЯНИЯМИ

© М. М. Ланге, С. В. Парамонов

*Федеральный исследовательский центр «Информатика и управление» РАН,
119333, Москва, ул. Вавилова, 42
E-mail: lange_mm@mail.ru, psvpobox@gmail.com*

Исследуются нижние границы вероятности ошибки классификации объектов при фиксированных значениях количества обрабатываемой информации в пространствах представлений объектов с заданными расстояниями. Границы определяются монотонно убывающими функциями наименьшей средней взаимной информации между предъявляемыми объектами и оценками их классов от вероятности ошибки. Обращения таких функций дают нижние границы вероятности ошибки классификации при фиксированных значениях количества обрабатываемой информации. Для древовидных и векторных представлений объектов получены численные реализации указанных границ. Демонстрируется более низкая граница вероятности ошибки в пространстве векторных представлений объектов по сравнению с аналогичной границей в пространстве древовидных представлений. Отмечается возможность понижения границы вероятности ошибки путём комплексирования представлений с различными метриками.

Ключевые слова: классификация, представление объектов, расстояние, вероятность ошибки, взаимная информация, комплексирование представлений.

DOI: 10.15372/AUT20240501

EDN: VVIQVJ

Введение. В задаче классификации объектов в заданном пространстве их описаний (представлений) критерием качества решающих алгоритмов является вероятность ошибки, которая должна уменьшаться с ростом количества информации, используемой для принятия решения. Для вероятностной модели классификации известна нижняя граница вероятности ошибки как функция средней взаимной информации между множеством объектов и возможных решений об их классах [1, 2]. Указанная граница построена на основе модели кодирования сообщений, переданных по каналу с шумом [3], и является модификацией соотношения, известного в теории информации как функция «скорость—погрешность» (rate distortion function) для кодирования дискретных сообщений с допустимой погрешностью по мере Хэмминга [4]. Существенно, что в соответствии с полученной границей наименьшая вероятность ошибки уменьшается с ростом средней взаимной информации между множеством объектов и классов. Поэтому построенная граница вероятности ошибки является теоретическим обоснованием максимизации указанной средней взаимной информации при отборе признаков [5, 6].

Возможность применения методов теории кодирования источников с заданной погрешностью (rate distortion theory) к задачам классификации и анализа данных исследовалась в работах [7–9]. В [7] для модели классификации предлагалось использовать теоретико-информационное соотношение вероятности ошибки и средней взаимной информации. В качестве такого соотношения применялась нижняя граница Шеннона для функ-

ции «скорость—погрешность» в схеме кодирования дискретных сообщений с погрешностью по мере Хэмминга [4]. Позднее указанная граница Шеннона применялась для оценивания наименьшей вероятности ошибки при заданном количестве обрабатываемой информации в модели сегментации изображений [8]. В [9] исследовался теоретико-информационный метод оптимизации алгоритмов сегментации изображений на основе критерия минимизации функционалов информационной избыточности и вариации информации.

Необходимо отметить, что использованная в работах [7, 8] граница Шеннона следует из границы, полученной в [1, 2], когда средняя взаимная информация между множеством данных и множеством классов достигает наибольшего значения, определяемого энтропией множества классов. Поэтому в схемах классификации данных и сегментации изображений граница Шеннона даёт заниженное значение наименьшей вероятности ошибки, к которому можно приблизиться с увеличением средней взаимной информации между множествами данных и классов.

В предлагаемой работе исследуются нижние границы вероятности ошибки классификации на множествах представлений объектов с заданными расстояниями. Исследование базируется на модели, в которой условные по классам вероятности объектов в рассматриваемых пространствах представлений определяются экспоненциально убывающими функциями от квадратов расстояний между предъявляемыми объектами и «центральными» представителями классов. Такая модель обеспечивает свойство компактности объектов по классам на множествах исследуемых представлений. Для объектов, заданных полуточновыми изображениями, рассматриваются древовидные представления наборами эллиптических примитивов [1] и представления векторами расстояний [10] от предъявляемых объектов до представителей классов. При этом расстояния, которые образуют компоненты векторов, задаются в пространстве древовидных представлений.

Цель данной работы состоит в том, чтобы продемонстрировать различие граничных значений вероятности ошибки классификации объектов при фиксированном количестве используемой информации в выбранных пространствах представлений и указать на возможность снижения вероятности ошибки путём объединения представлений с разными метриками. Эффект повышения точности классификации за счёт объединения различных представлений для одних и тех же объектов эквивалентен аналогичному эффекту, который достигается путём комплексирования данных различной модальности, получаемых от разных источников [11].

Модель классификации и задача исследования. Рассматривается вероятностная модель, заданная преобразованиями

$$\Omega \xrightarrow{P_{\mathbf{X}^N|\Omega}} \mathbf{X}^N \xrightarrow{Q_{\hat{\Omega}|\mathbf{X}^N}} \hat{\Omega}, \quad (1)$$

где $\Omega = \{\omega_i\}_{i=1}^c$ и $\hat{\Omega} = \{\hat{\omega}_j\}_{j=1}^c$ — множества классов и их оценок, $c \geq 2$; \mathbf{X}^N — множество блоков из $N \geq 1$ объектов. Последовательность $\Omega, \mathbf{X}^N, \hat{\Omega}$ удовлетворяет марковскому свойству, согласно которому элементы каждого множества зависят только от элементов предыдущего множества, и $P_{\mathbf{X}^N|\Omega}, Q_{\hat{\Omega}|\mathbf{X}^N}$ — множества условных по классам распределений блоков объектов и условных по блокам распределений оценок классов соответственно. На множестве Ω заданы априорные вероятности классов.

Используя среднюю взаимную информацию $I_{Q_{\hat{\Omega}|\mathbf{X}^N}}(\mathbf{X}^N; \hat{\Omega})$ и среднюю вероятность ошибки $E_{Q_{\hat{\Omega}|\mathbf{X}^N}}(\mathbf{X}^N; \hat{\Omega})$ в форме средней погрешности по мере Хэмминга как функционалы, зависящие от $Q_{\hat{\Omega}|\mathbf{X}^N}$, в работах [1, 2] для значений $\varepsilon > 0$ введена функция «взаимная

информация—вероятность ошибки»

$$R(\varepsilon) = \min_N \min_{Q_{\hat{\Omega}}|_{\mathbf{X}^N}: E_{Q_{\hat{\Omega}}|_{\mathbf{X}^N}}(\mathbf{X}^N; \hat{\Omega}) \leq \varepsilon} I_{Q_{\hat{\Omega}}|_{\mathbf{X}^N}}(\mathbf{X}^N; \Omega). \quad (2)$$

Здесь внутренний минимум берётся по всевозможным распределениям из множества $Q_{\hat{\Omega}}|_{\mathbf{X}^N}$ при выбранном ε -ограничении вероятности ошибки, а внешний — по длине блоков N . Для функции (2) найдена нижняя граница $\underline{R}(\varepsilon) \leq R(\varepsilon)$ в форме

$$\underline{R}(\varepsilon) = I(\Omega; \mathbf{X}) - h(\varepsilon - \varepsilon_{\min}) - (\varepsilon - \varepsilon_{\min}) \ln(c - 1), \quad (3)$$

где $I(\Omega; \mathbf{X})$ — средняя взаимная информация между Ω и \mathbf{X} , $h(z) = -z \ln z - (1 - z) \ln(1 - z)$. Функция $\underline{R}(\varepsilon)$ монотонно убывает на отрезке $\varepsilon_{\min} \leq \varepsilon \leq \varepsilon_{\max}$, $\underline{R}(\varepsilon_{\min}) = I(\Omega; \mathbf{X})$ и $\underline{R}(\varepsilon_{\max}) = 0$. Наименьшая вероятность ошибки ε_{\min} уменьшается с ростом $I(\Omega; \mathbf{X})$, а $\varepsilon_{\max} = (c - 1) \min_{i=1}^c P(\omega_i)$, где $P(\omega_i)$ — априорная вероятность класса $\omega_i \in \Omega$. В случае $I(\Omega; \mathbf{X}) = H(\Omega)$ имеем $\varepsilon_{\min} = 0$ и $\underline{R}(\varepsilon)$ совпадает с границей Шеннона [4]. В силу монотонности границы (3) обратная функция $\underline{R}^{-1}(I)$ даёт нижнюю границу вероятности ошибки классификации объектов на множестве \mathbf{X} при фиксированных значениях количества информации $I \leq I(\Omega; \mathbf{X})$, а величина $1 - \underline{R}^{-1}(I)$ соответствует верхней границе точности классификации при заданных значениях количества обрабатываемой информации.

Вычисление характеристик $I(\Omega; \mathbf{X})$ и ε_{\min} в (3) требует знания совместного распределения на множествах Ω и \mathbf{X} . Априорное распределение классов на множестве Ω считается равномерным с вероятностями $1/c$, а условные по классам вероятности объектов определяются с использованием расстояний $d(\mathbf{x}^*, \mathbf{x}_i^*) \geq 0$, $i = 1, \dots, c$, между объектом $\mathbf{x} \in \mathbf{X}$ и представителем i -го класса $\mathbf{x}_i \in \mathbf{X}$ в пространстве представлений \mathbf{X}^* объектов множества \mathbf{X} .

С учётом представлений $\mathbf{x}^* \in \mathbf{X}^*$, $\mathbf{x}_i^* \in \mathbf{X}^*$ условные по классам вероятности объектов имеют вид

$$P_{\mathbf{X}^*|\Omega}(\mathbf{x}^*|\omega_i) = \frac{\exp(-w_i^* d^2(\mathbf{x}^*, \mathbf{x}_i^*))}{\sum_{\mathbf{x}^* \in \mathbf{X}^*} \exp(-w_i^* d^2(\mathbf{x}^*, \mathbf{x}_i^*))}, \quad (4)$$

где $w_i^* > 0$, $i = 1, \dots, c$ — свободные параметры. В качестве представителей классов выбираются центральные объекты с представлениями

$$\mathbf{x}_i^* = \arg \min_{\hat{\mathbf{x}}^* \in \mathbf{X}_i^*} \sum_{\mathbf{x}^* \in \mathbf{X}_i^*} d^2(\mathbf{x}^*, \hat{\mathbf{x}}^*) \quad (5)$$

в подмножествах $\mathbf{X}_i^* \subset \mathbf{X}^*$, $i = 1, \dots, c$.

Условные распределения (4) относительно представителей классов (5) и равномерное априорное распределение классов позволяют вычислить значения $I(\Omega; \mathbf{X}) = I(\Omega; \mathbf{X}^*)$ и $\varepsilon_{\min} = \varepsilon_{\min}^*$ в границе (3). Реализации границы строятся на множествах изображений лиц \mathbf{X}_f и подписей \mathbf{X}_s от $c = 25$ персон, по 40 объектов от каждой персоны. Для указанных данных рассматриваются множества древовидных представлений лиц \mathbf{X}_f^t и подписей \mathbf{X}_s^t наборами эллиптических примитивов [1] и множества представлений \mathbf{X}_f^v и \mathbf{X}_s^v векторами расстояний [10] объектов до центральных представителей классов в пространстве древовидных представлений. Используемые древовидные представления строятся путём дихотомического разбиения объектов на непересекающиеся сегменты и аппроксимации сегмен-

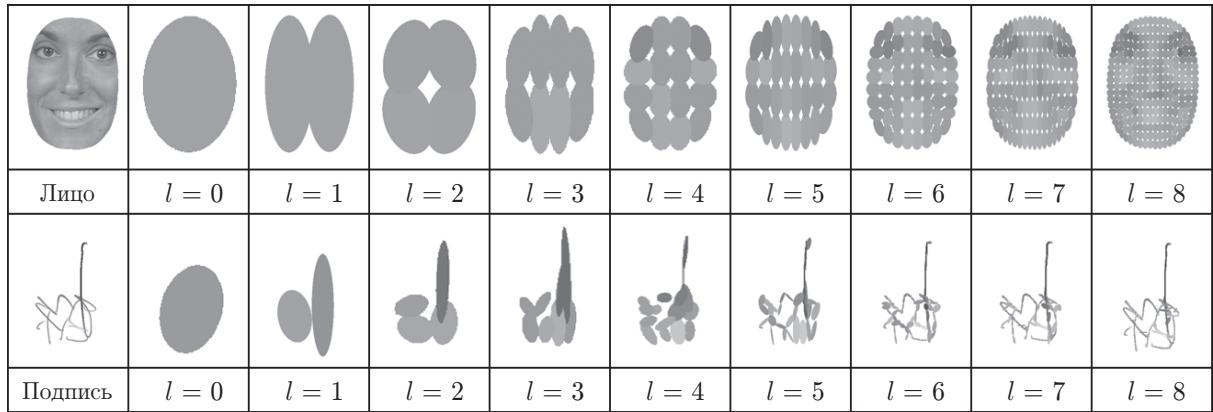


Рис. 1. Примеры древовидных представлений изображений лица и подписи

тов эллиптическими примитивами, размеры и ориентация которых определяются эллипсами рассеяния соответствующих сегментов. Примитивы имеют яркостную окраску, которая определяется средней яркостью пикселей аппроксимируемых сегментов. Для каждого объекта строится набор эллиптических примитивов, которые соответствуют вершинам бинарного дерева глубины L . Уровни дерева с номерами $l = 0, \dots, L$, содержащие 2^l примитивов, образуют многоуровневое представление объекта, в котором параметры примитивов задаются в собственных координатах объекта и нормируются относительно параметров примитива нулевого уровня. Указанные операции обеспечивают инвариантность представлений всех уровней относительно сдвига, поворота, масштаба и уровня яркости представляемого объекта. Примеры представлений лица и подписи эллиптическими примитивами, образующими бинарные деревья глубины $L = 8$, даны на рис. 1. В таких представлениях l -й уровень дерева содержит 2^l примитивов, которые соответствуют вершинам этого уровня.

На множестве древовидных представлений \mathbf{X}^t объектов используется расстояние $d(\mathbf{x}^t, \hat{\mathbf{x}}^t) = d_L(\mathbf{x}, \hat{\mathbf{x}})$, $\mathbf{x}^t \in \mathbf{X}^t$, $\hat{\mathbf{x}}^t \in \mathbf{X}^t$, которое определяется метрикой $d_L(\mathbf{x}, \hat{\mathbf{x}}) \geq 0$ с квадратичным ядром для пары объектов $\mathbf{x} \in \mathbf{X}$, $\hat{\mathbf{x}} \in \mathbf{X}$, представленных деревьями глубины L [1]. Множество векторных представлений \mathbf{X}^v образовано векторами $\mathbf{x}^v = (x_1^v, \dots, x_c^v)$ длины c , компоненты которых $x_i^v = d(\mathbf{x}^t, \mathbf{x}_i^t)$, $i = 1, \dots, c$, заданы расстояниями между древовидным представлением \mathbf{x}^t объекта и древовидными представлениями \mathbf{x}_i^t , $i = 1, \dots, c$, центральных объектов вида (5) в соответствующих классах. Для любой пары объектов с векторными представлениями $\mathbf{x}^v \in \mathbf{X}^v$, $\hat{\mathbf{x}}^v \in \mathbf{X}^v$ введено среднеквадратичное расстояние

$$d(\mathbf{x}^v, \hat{\mathbf{x}}^v) = \left(\frac{1}{c} \sum_{i=1}^c (\mathbf{x}_i^v - \hat{\mathbf{x}}_i^v)^2 \right)^{1/2},$$

которое позволяет вычислить на множестве \mathbf{X}^v условные по классам вероятности вида (4) с соответствующими представителями классов вида (5).

Численные реализации границы $R(\varepsilon)$. Методика вычисления характеристик $I(\Omega; \mathbf{X})$ и ε_{\min} границы вида (3) подробно изложена в работе [2]. Вычисление указанных характеристик выполнено с использованием матриц расстояний на множествах древовидных представлений лиц \mathbf{X}_f^t [12] и подписей \mathbf{X}_s^t [13]. Матрицы расстояний получены для древовидных представлений глубины $L = 8$, содержащих информативные уровни $l = 1, \dots, 8$. Численные реализации границы $R(\varepsilon)$ показаны на рис. 2 сплошными кривыми f и s для изображений лиц и подписей. Пунктирная кривая соответствует границе Шеннона, когда $I(\Omega; \mathbf{X}) = H(\Omega) = \ln c = 3,219$ и $\varepsilon_{\min} = 0$. Кривые на рис. 2, *a* вычислены на множествах

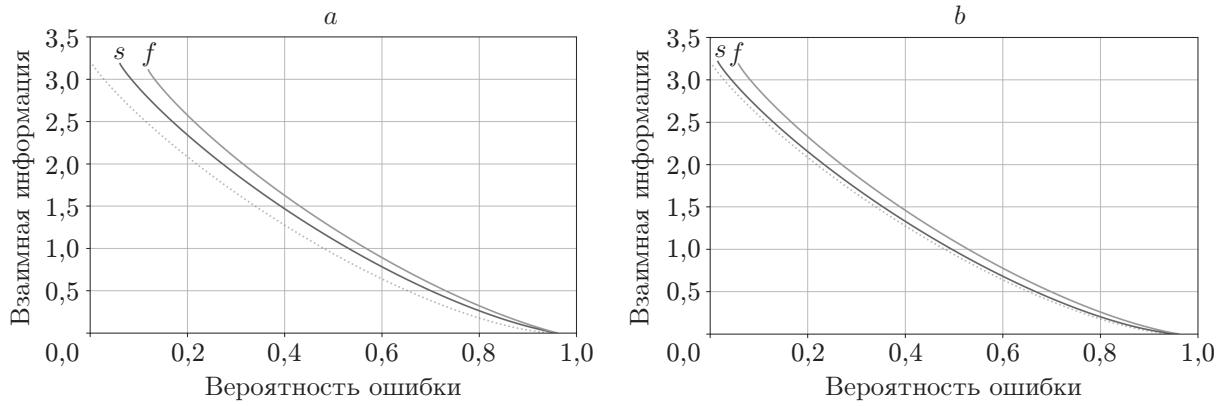


Рис. 2. Реализации границы $\underline{R}(\varepsilon)$ на множествах изображений лиц (f) и подписей (s) для древовидных представлений: $I(\Omega; \mathbf{X}_f^t) = 3,110$, $\varepsilon_{f \min}^t = 0,092$; $I(\Omega; \mathbf{X}_s^t) = 3,184$, $\varepsilon_{s \min}^t = 0,052$ (а) и векторных представлений: $I(\Omega; \mathbf{X}_f^v) = 3,187$, $\varepsilon_{f \min}^v = 0,060$; $I(\Omega; \mathbf{X}_s^v) = 3,215$, $\varepsilon_{s \min}^v = 0,015$ (б)

древовидных представлений \mathbf{X}_f^t и \mathbf{X}_s^t , кривые на рис. 2, а — на множествах векторных представлений \mathbf{X}_f^v и \mathbf{X}_s^v . Поскольку для множеств представлений объектов каждой модальности выполняется соотношение $I(\Omega; \mathbf{X}^v) > I(\Omega; \mathbf{X}^t)$, которое обеспечивает $\varepsilon_{\min}^v < \varepsilon_{\min}^t$, реализации границы $\underline{R}(\varepsilon)$ в векторном пространстве представлений проходят ближе к границе Шеннона по сравнению с реализацией границы в пространстве древовидных представлений.

Увеличение средней взаимной информации между множеством предъявляемых объектов и множеством классов и, как следствие, уменьшение наименьшей вероятности ошибки могут быть достигнуты комплексированием древовидных и векторных представлений. Комплексирование предполагает совместное использование метрик, заданных на множествах представлений \mathbf{X}^t и \mathbf{X}^v , которые на ансамбле $\mathbf{X}^t \mathbf{X}^v$ порождают условные по классам вероятности объектов

$$P_{\mathbf{X}^t \mathbf{X}^v | \Omega}(\mathbf{x}^t, \mathbf{x}^v | \omega_i) = \frac{\exp \{-(w_i^t d^2(\mathbf{x}^t, \mathbf{x}_i^t) + w_i^v d^2(\mathbf{x}^v, \mathbf{x}_i^v))\}}{\sum_{(\mathbf{x}^t, \mathbf{x}^v) \in \mathbf{X}^t \mathbf{X}^v} \exp \{-(w_i^t d^2(\mathbf{x}^t, \mathbf{x}_i^t) + w_i^v d^2(\mathbf{x}^v, \mathbf{x}_i^v))\}}, \quad (6)$$

где суммирование выполняется по парам представлений всех объектов. Распределения вида (6) совместно с равномерным распределением классов дают среднюю взаимную информацию $I(\Omega; \mathbf{X}^t \mathbf{X}^v)$ и наименьшую вероятность ошибки $\varepsilon_{\min}(\mathbf{X}^t \mathbf{X}^v)$, значения которых для ансамблей древовидных и векторных представлений лиц и подписей отражены в таблице. Для сравнения в таблице даны аналогичные характеристики, полученные на множествах древовидных и векторных представлений объектов от указанных источников. Очевидно, что для рассматриваемых ансамблей представлений объектов численные реализации границы $\underline{R}(\varepsilon)$ вида (3) с характеристиками, приведёнными в таблице, должны быть ближе к границе Шеннона по сравнению с реализациями нижней границы, показанными на рис. 2, для отдельных представлений.

Из сопоставления приведённых в таблице характеристик, полученных для древовидных и векторных представлений объектов и для ансамбля этих представлений, следует, что поведение функции «взаимная информация — вероятность ошибки» обеспечивает на ансамбле представлений потенциальную возможность повышения точности классификации при фиксированных значениях количества анализируемой информации. Увеличение

Таблица

Средняя взаимная информация и наименьшая вероятность ошибки на множествах древовидных и векторных представлений изображений лиц и подписей, а также на ансамбле этих представлений

Space: \mathbf{X}^*	$\mathbf{X}_f^t \mathbf{X}_f^v$	\mathbf{X}_f^t	\mathbf{X}_f^v	$\mathbf{X}_s^t \mathbf{X}_s^v$	\mathbf{X}_s^t	\mathbf{X}_s^v
$I(\Omega; \mathbf{X}^*)$	3,189	3,110	3,187	3,217	3,184	3,215
$\varepsilon_{\min}(\mathbf{X}^*)$	0,057	0,092	0,060	0,005	0,052	0,015

средней взаимной информации между ансамблем и множеством классов и, соответственно, понижение наименьшей вероятности ошибки могут быть обеспечены увеличением размера ансамбля. В этом случае в условных по классам вероятностях (6) под знаком экспонент используются взвешенные суммы квадратов расстояний по всем представлениям, входящим в ансамбль.

Заключение. Исследованы реализации функции «взаимная информация—вероятность ошибки» на множествах древовидных и векторных представлений классифицируемых данных с известными метриками. Благодаря монотонности такой функции, её обращение даёт нижнюю границу вероятности ошибки классификации в выбранном пространстве представлений при фиксированных значениях количества обрабатываемой информации. Кроме того, отклонение нижней границы вероятности ошибки от единицы соответствует верхней границе точности классификации объектов в выбранном пространстве их описаний. Для множеств изображений лиц и подписей продемонстрирована более высокая эффективность векторных представлений по сравнению с древовидными. Исследованные векторные представления данных могут найти применение для высокоточной классификации изображений с применением решающих алгоритмов на основе опорных векторов и нейронных сетей. Продемонстрирована возможность понижения нижней границы вероятности ошибки при фиксированных значениях количества анализируемой информации путём комплексирования представлений с различными метриками.

СПИСОК ЛИТЕРАТУРЫ

1. Lange A. M., Lange M. M., Paramonov S. V. Tradeoff Relation between Mutual Information and Error Probability in Data Classification Problem // Comput. Math. and Math. Phys. 2021. **61**, Iss. 7. P. 1181–1193.
2. Lange M., Paramonov S. On a Lower Bound to Classification Error Probability in an Ensemble of Data Sources // Proc. of the Int. Conf. on Information Technology and Nanotechnology (ITNT). Samara, Russia, 20–24 Sept., 2021. DOI: 10.1109/ITNT52450.2021.9649088.
3. Dobrushin R., Tsybakov B. Information transmission with additional noise // IRE Trans. Information Theory. 1962. **8**, Iss. 5. P. 293–304. DOI: 10.1109/TIT.1962.1057738.
4. Gallager R. G. Information Theory and Reliable Communication. N.-Y.: Wiley, 1968. 588 p.
5. Brown G., Pocock A., Zhao M.-J., Luján M. Conditional Likelihood Maximisation: A Unifying Framework for Information Theoretic Feature Selection // Journ. Machine Learning Research. 2012. **13**, Iss. 2. P. 27–66.
6. Fleuret F. Fast Binary Feature Selection with Conditional Mutual Information // Journ. Machine Learning Research. 2004. **5**. P. 1531–1555.
7. Sethi I. K., Sarvarayudu G. P. Hierarchical classifier design using mutual information // IEEE Trans. Pattern Analysis and Machine Intell. 1982. **PAMI-4**, Iss. 4. P. 441–445.

8. **Rigau J., Feixas M., Sbert M.** An information theoretic framework for image segmentation // Proc. of the Int. Conf. on Image Processing (ICIP). Singapore, 24–27 Oct., 2004. **2**. P. 1193–1196. DOI: 10.1109/ICIP.2004.1419518.
9. **Мурашов Д. М.** Метод комбинирования карт сегментации изображений на основе минимизации информационной избыточности и вариации информации // Автометрия. 2022. **58**, № 5. С. 34–42. DOI: 10.15372/AUT20220504.
10. **Duin R. P. W., De Ridder D., Tax D. M. J.** Experiments with a featureless approach to pattern recognition // Patt. Recogn. Lett. 1997. **18**, Iss. 11–13. P. 1159–1166.
11. **Белов А. М., Денисова А. Ю.** Классификация изображений ДЗЗ с использованием алгоритма комплексирования данных различных сенсоров // Компьютерная оптика. 2020. **44**, вып. 4. С. 627–635. DOI: 10.18287/2412-6179-CO-735.
12. **Distance-matrices-face.** SourceForge. URL: <https://sourceforge.net/projects/distance-matrices-face/> (дата обращения: 05.07.2024).
13. **Distance-matrices-signature.** SourceForge. URL: <http://sourceforge.net/projects/distance-matrices-signature> (дата обращения: 05.07.2024).

Поступила в редакцию 05.07.2024

После доработки 22.07.2024

Принята к публикации 24.07.2024
