

ОБ ОДНОЙ ЭВРИСТИКЕ ДЛЯ РАСПОЗНАВАНИЯ РАСТРОВЫХ ИЗОБРАЖЕНИЙ

А. В. Пролубников

Омский государственный университет им. Ф. М. Достоевского,
644077, Омск, Россия

УДК 519.6

Предлагается эвристика для решения задачи распознавания растровых изображений, состоящая в минимизации меры объединенного множества решений интервальной системы линейных уравнений, которая ставится в соответствие эталонному изображению с учетом распознаваемого изображения. Проведен анализ результатов вычислительного эксперимента.

Ключевые слова: распознавание образов, интервальный анализ.

A new approach to recognition of raster images is presented. For solution of the problem, we construct the systems of interval linear equations for given numerical pattern matrices and use variation of solutions of the systems as a heuristics for recognition. The problem of raster image recognition, where raster images are modified at process of noising, may be considered as a case of the more general problem of numerical matrices recognition. Results of numerical experiments are presented.

Key words: pattern recognition, interval analysis.

1. Постановка задачи и преобразование входных данных. Автоматизация процесса анализа и интерпретации изображений достаточно трудоемка не только потому, что обработка объема информации, содержащейся в изображениях, может потребовать значительных вычислительных мощностей, но и потому, что сами процессы регистрации человеком изображений могут быть достаточно сложны, плохо формализуемы или вообще неизвестны. Кроме того, при регистрации изображений техническими средствами неизбежны погрешности измерения, возможно наличие шума, например зашумление отдельных фрагментов или изображения в целом.

При применении для распознавания растровых изображений алгоритмов распознавания с помощью нейронных сетей [1], параметрических алгоритмов [2], алгоритмов на основе теории морфологического анализа [3] используется предварительная стадия обучения, в ходе которой алгоритм необходимо “обучать изображениям объекта”, полученным при различных условиях его регистрации. Целью процесса обучения является фиксирование некоторых характеристик изображения, по которым может быть проведено распознавание.

Предлагаемый в данной работе подход к распознаванию растровых изображений использует в качестве эвристики, с помощью которой проводится распознавание, минимизацию лебеговой меры объединенного множества решений интервальной системы линейных алгебраических уравнений (ИСЛАУ). ИСЛАУ ставится в соответствие каждому эталонному изображению с учетом распознаваемого изображения. Характерным отличием предлагаемого подхода к распознаванию от традиционных подходов является отсутствие стадии обучения алгоритма распознавания.

1.1. *Постановка задачи.* Растровые изображения представляются в виде матриц. Значения элементов этих матриц соответствуют яркости, цвету или каким-либо другим характеристикам пикселей, совокупность которых представляет изображение. Соответственно задача распознавания растровых изображений может быть рассмотрена в следующей матричной постановке. Имеется N эталонных квадратных $(n \times n)$ -матриц $\{\mathcal{A}^{(k)}\}_{k=1}^N$, соответствующих эталонным изображениям, с элементами $a_{ij}^{(k)} \in \mathbb{Z}_+$. В ходе зашумления одного из изображений с матрицей $\mathcal{A}^{(p)}$ получено некоторое изображение с матрицей \mathcal{A} . Известно, что значения элементов матриц, соответствующих пикселям, могли быть изменены в пределах интервалов $[a_{ij}^{(k)} - \delta, a_{ij}^{(k)} + \delta]$, $\delta > 0$. Необходимо определить p .

Матрицы, соответствующие изображениям $\mathcal{A}^{(1)}, \dots, \mathcal{A}^{(N)}, \mathcal{A}$, без ограничения общности можно считать квадратными. В противном случае, если матрицы, соответствующие эталонным и распознаваемому изображениям, являются $(m \times n)$ -матрицами и $m < n$, к каждой матрице добавляются $n - m$ нулевых строк, если $n < m$, добавляется $m - n$ нулевых столбцов.

Следует отметить, что распознавание изображения с матрицей \mathcal{A} проводится при отсутствии поворотов эталонных изображений, растяжений этих изображений в целом или их отдельных частей и других искажений. Изменение пикселей некоторого эталонного изображения выполняется только в результате зашумления указанного вида — изменения элементов матрицы изображения в заданных интервалах.

1.2. *Предварительная модификация матриц изображений.* Выполним следующую модификацию матриц, поставленных в соответствие эталонным и зашумленному изображениям, подаваемым на вход алгоритма. Каждый элемент как эталонных матриц, так и зашумленной матрицы увеличим на одно и то же значение $v > 0$:

$$a_{ij} := a_{ij} + v, \quad a_{ij}^{(k)} := a_{ij}^{(k)} + v. \quad (1)$$

Преобразования (1) сохраняют расстояние, понимаемое как модуль разности, и между элементами в разных позициях одной матрицы, и между элементами в одинаковых позициях для любой пары матриц из набора входных. Поэтому входная информация, по которой можно различать элементы матриц и проводить распознавание, после проведенных преобразований сохраняется. Эти преобразования можно понимать как перекодирование пикселей растровых изображений, поданных на вход алгоритма распознавания. Под кодированием пикселей растрового изображения будем понимать установление соответствия между значением пикселя изображения и значением представляющего его элемента матрицы изображения.

В результате проведения преобразований (1) уменьшается отношение радиуса интервала изменения элемента эталонной матрицы δ к модулю этого элемента. Так, если до указанных преобразований для элементов эталонных матриц $a_{ij}^{(k)}$ отношение радиуса интервала, в котором происходит изменение элемента, к величине модуля элемента матрицы составляет $\delta/|a_{ij}^{(k)}|$, то после их проведения это отношение уменьшается с увеличением v и составляет $\delta/|a_{ij}^{(k)} + v|$. Ограничение на величину v рассматривается ниже.

2. Минимизация лебеговой меры внешней оценки объединенного множества решений ИСЛАУ как эвристика, на основе которой проводится распознавание. Входным растровым изображениям поставим в соответствие интервальные матрицы и производим распознавание, оценивая множества решений систем уравнений с этими матрицами. Ниже проведено построение этих матриц.

2.1. *Построение интервальных матриц, соответствующих эталонным изображениям.* В качестве решения интервальной системы линейных алгебраических уравнений будем рас-

смагивать объединенное множество решений этой системы. Компонентами и элементами интервальных векторов и матриц являются замкнутые интервалы. Везде далее интервальность векторов и матриц обозначается квадратными скобками. Интервальный вектор $[\mathbf{x}]$ может пониматься как интервальная оболочка некоторого точечного вектора \mathbf{x} , при этом величины интервалов — возможных погрешностей в измерениях компонент вектора \mathbf{x} — указываются при необходимости.

Определение 1. Для ИСЛАУ $[A][\mathbf{x}] = [\mathbf{b}]$, где $[A]$ — интервальная $(n \times n)$ -матрица; $[\mathbf{b}]$ — некоторый интервальный n -мерный вектор, объединенное множество решений есть множество $\Xi([A], [\mathbf{b}])$:

$$\Xi([A], [\mathbf{b}]) = \{x \in \mathbb{R}^n \mid \exists A \in [A], \exists \mathbf{b} \in [\mathbf{b}]: A\mathbf{x} = \mathbf{b}\}.$$

Матрицам $\{\mathcal{A}^{(k)}\}_{k=1}^N$ с элементами $a_{ij}^{(k)} \in \mathbb{R}$ поставим в соответствие интервальные матрицы $\{[A^{(k)}]\}_{k=1}^N$ с интервальными элементами $[a_{ij}^{(k)}]$, формируемыми следующим образом:

$$[a_{ij}^{(k)}] = [\min\{a_{ij}^{(k)}, a_{ij}\}, \max\{a_{ij}^{(k)}, a_{ij}\}] \quad (2)$$

(a_{ij} — элементы матрицы \mathcal{A} , соответствующей распознаваемому изображению). Элементы матрицы $[A^{(k)}]$ — интервалы, характеризующие изменения элементов эталонной матрицы, необходимые для получения из нее распознаваемой матрицы.

Предлагаемый подход основан на предположении, согласно которому чем меньше разброс в решениях точечных СЛАУ, составляющих ИСЛАУ $[A^{(k)}][\mathbf{x}] = [\mathbf{b}]$ с некоторой точечной правой частью \mathbf{b} из \mathbb{R}^n , тем выше вероятность того, что именно зашумлением эталонного изображения $\mathcal{A}^{(k)}$ получена матрица \mathcal{A} .

Разброс в решениях точечных СЛАУ, составляющих (2), можно оценить через лебегову меру μ множества решений $\Xi([A^{(k)}], [\mathbf{b}])$. Поскольку мера $\mu(\Xi([A^{(k)}], [\mathbf{b}]))$ зависит от взаимного расположения элементов $[A^{(k)}]$ и непрерывно зависит от их изменений, то при изменении достаточно ограниченного числа элементов матрицы эталонного изображения $\mathcal{A}^{(p)}$ в не очень большом диапазоне предложенная эвристика может быть использована для решения поставленной задачи распознавания.

2.2. Выбор вектора правой части ИСЛАУ. Наряду с матрицей $[A^{(k)}]$ вектор правой части $[\mathbf{b}]$ определяет множество $\Xi([A^{(k)}], [\mathbf{b}])$, поэтому его выбор важен для эффективности распознавания. Вектор в правой части рассматриваемых ИСЛАУ выбирается точечным, что позволяет выполнить более точную внешнюю оценку объединенного множества решений. В случае точечной правой части интервальная оболочка объединенного множества решений имеет покомпонентно меньший радиус, чем в случае интервальной правой части, что приводит к уменьшению расстояния между объединенным множеством решений и его интервальной оболочкой [4]. Это, в свою очередь, позволяет выполнять внешнее оценивание с большей точностью.

Если в качестве правой части использовать вектор $\mathbf{e} = (1, \dots, 1)^T$, то $\mu(\Xi([A^{(k)}], \mathbf{e}))$ будет в равной мере зависеть от всех элементов матрицы, в отличие от случая использования вектора с неравными компонентами вектора правой части. Поскольку все элементы эталонных и распознаваемой матриц должны в равной мере учитываться при распознавании, рассматриваемые ниже ИСЛАУ имеют вид

$$[A^{(k)}][\mathbf{x}] = \mathbf{e}.$$

2.3. *Вычислительная сложность распознавания.* Множество $\Xi^{(k)} \stackrel{def}{=} \Xi([A^{(k)}], \mathbf{e})$ представляет собой объединение не более чем 2^n полиэдров [5], являющихся пересечениями $\Xi^{(k)}$ с ортантами \mathbb{R}^n . Задача нахождения лебеговой меры $\Xi^{(k)}$ имеет большую вычислительную сложность, поскольку описание самого множества с помощью гиперплоскостей, ограничивающих его, — задача с экспоненциально растущей относительно размерности вычислительной сложностью. Поэтому ограничимся выполнением оценивания сверху лебеговой меры множества $\Xi^{(k)}$ лебеговой мерой некоторого приближения интервальной оболочки $\Delta\Xi^{(k)}$ этого множества. Интервальная оболочка $\Delta\Xi^{(k)}$ — минимальный по включению интервальный вектор, содержащий $\Xi^{(k)}$, т. е. брус со сторонами параллельными координатным осям такой, что $\Xi^{(k)} \subset \Delta\Xi^{(k)}$. Для любого бруса $[\mathbf{c}]$, такого что $\Xi^{(k)} \subset [\mathbf{c}]$, имеем $\Delta\Xi^{(k)} \subseteq [\mathbf{c}]$.

Обозначим через $[X^{(k)}]$ получаемое с помощью какого-либо численного метода интервального анализа приближение интервальной оболочки $\Delta\Xi^{(k)}$. $[X^{(k)}]$ — некоторый брус:

$$[X^{(k)}] = ([\underline{x}_1^k, \bar{x}_1^k], \dots, [\underline{x}_n^k, \bar{x}_n^k])^T, \quad \underline{x}_i^k \leq \bar{x}_i^k,$$

такой что $\Delta\Xi^{(k)} \subseteq [X^{(k)}]$. Лебегова мера μ бруса $[X^{(k)}]$ рассчитывается по формуле

$$\mu([X^{(k)}]) = (\bar{x}_1^k - \underline{x}_1^k) \cdots (\bar{x}_n^k - \underline{x}_n^k).$$

Внешние оценки $[X^{(k)}]$ могут быть получены с помощью методов интервального анализа, разработанных для выполнения внешнего оценивания объединенных множеств решений ИСЛАУ. При вычислительной сложности выбранного алгоритма внешнего оценивания $Encl$, равной $C_{Encl}(n)$, общая трудоемкость алгоритма, проводящего распознавание на основе сравнения N значений $\mu([X^{(k)}])$, равна

$$C(N, n, Encl) = O(NC_{Encl}(n)).$$

В случае $C_{Encl}(n) = O(n^2)$ получаем алгоритм с минимальным возможным порядком вычислительной сложности для алгоритмов решения поставленной задачи, поскольку вычислительная сложность любой процедуры обработки элементов $(n \times n)$ -матрицы составляет не менее $O(n^2)$.

3. Алгоритм распознавания. Модификация входных матриц — неотъемлемая часть предлагаемого алгоритма распознавания. Проводимая модификация матриц позволяет повысить качество распознавания и использовать для оценивания множеств решений построенных систем уравнений алгоритм с квадратичной вычислительной сложностью — итерационный интервальный метод Гаусса — Зейделя.

3.1. *Модификации матриц изображений, подаваемых на вход алгоритма.* Матрицы растровых изображений, поданных на вход алгоритма распознавания, подвергаются следующим преобразованиям. Сначала осуществляются указанные выше преобразования (1), после чего эталонным изображениям с учетом зашумленного изображения ставятся в соответствие интервальные матрицы вида (2) и проводится преобразование их диагоналей, необходимое для эффективного применения численных методов интервального анализа. Рассмотрим эти преобразования более подробно.

Уменьшение отношения радиуса изменения элемента к его величине после ряда преобразований (1) позволяет сгладить изменения элементов $\mathcal{A}^{(p)}$, в результате которых получена

матрица \mathcal{A} . До выполнения указанных преобразований существенное изменение относительно малого числа элементов матрицы, соответствующей эталонному изображению, может привести к тому, что распознавание по предлагаемому в работе алгоритму будет невозможно, несмотря на то что большая часть элементов была изменена незначительно. Однако при выборе достаточно большого значения ν изменения значений всех элементов матрицы, соответствующей эталонному изображению, будут малы по сравнению со значениями элементов и во многих случаях распознавание будет возможно.

Если значение δ близко к модулям значений элементов эталонных матриц, то интервальные матрицы $[A^{(k)}]$, построенные в соответствии с (2), будут содержать точечные матрицы, такие что векторы вида $\mathbf{x} = A^{-1}\mathbf{e}$ при $A \in [A^{(k)}]$ будут существенно различаться. В этом случае проводить распознавание, основываясь на близости или удаленности данных векторов для всех матриц в $[A^{(k)}]$, минимизируя по k меру множества $\Xi^{(k)}$, нельзя. Если же радиус δ мал по сравнению с модулями элементов матрицы, сопоставляемой с зашумляемым эталонным изображением, что имеет место в результате проведения преобразований (1), то такое распознавание во многих случаях возможно. Например, при распознавании монохромных изображений, в случае если пикселям белого цвета соответствуют элементы матриц, равные 1, пикселям черного цвета — 0, а изменению цвета пикселей при зашумлении соответствует изменение значения элемента матрицы с 0 на 1 или наоборот, величина δ возможного изменения элемента матрицы не будет превышать модуля элемента. Для таких матриц распознавание по предложенному подходу с использованием какого-либо метода внешнего оценивания объединенного множества решений ИСЛАУ будет невозможно, тогда как оно эффективно выполняется после модификации (см. (1)) входных матриц с $\nu = 10$, когда $\delta/|a_{ij}^{(k)} + \nu| \leq 0,1$.

Вместе с тем с ростом ν уменьшается норма векторов, составляющих множества $\Xi^{(k)}$. Это может привести к тому, что лебеговы меры приближений интервальных оболочек указанных множеств, получаемых с помощью какого-либо алгоритма внешнего оценивания, не будут отражать специфику эталонных матриц. Это будет происходить как в силу того, что с увеличением модулей элементов в матрицах и ростом диагонального преобладания будут нивелированы различия между самими оцениваемыми множествами $\Xi^{(k)}$, так и в силу наличия погрешностей используемых численных методов. Для эффективного распознавания необходимо достаточное отклонение одного из значений $\mu([X^{(p)}])$ от прочих значений $\mu([X^{(k)}])$, $k \neq p$. При проведении вычислительных экспериментов, рассматриваемых ниже, величина ν полагалась равной 10 a (a — максимальный по модулю элемент во всех матрицах, поданных на вход алгоритма).

Выполнив преобразования (1) и построив в соответствии с (2) матрицы $[A^{(k)}]$, модифицируем их диагональные элементы:

$$[A^{(k)}] := [A^{(k)}] + [D] \quad (3)$$

$[D]$ — диагональная интервальная матрица с элементами $[D_{ii}] = [d, d]$. Матрица $[D]$ задается следующим образом:

$$d = \max_{1 \leq k \leq N} d^{(k)}. \quad (4)$$

Здесь

$$d^{(k)} = 2 \max_{1 \leq i \leq n} \sum_{j \neq i} |[A^{(k)}]_{ij}|.$$

Поскольку значение d выбирается в соответствии с (4), матрицы $[A^{(k)}]$ после преобразований (3) представляют собой H -матрицы, что позволяет в ходе распознавания выполнять с их помощью необходимые вычисления по нахождению внешних оценок $[X^{(k)}]$.

Определение 2. Компарантом матрицы $A = (a_{ij}) \in \mathbb{R}^{n \times n}$ называется матрица того же размера, обозначаемая как $\langle A \rangle$, такая что

$$\langle A \rangle_{ij} = \begin{cases} |a_{ij}|, & i = j, \\ -|a_{ij}|, & i \neq j. \end{cases}$$

Определение 3. Матрица $A = (a_{ij}) \in \mathbb{R}^{n \times n}$ называется M -матрицей, если она представима в виде $A = sI - P$, где P — неотрицательная матрица; $s > \rho(P)$; $\rho(P)$ — спектральный радиус P . Интервальная матрица $[A]$ называется интервальной M -матрицей, если каждая вещественная матрица $A \in [A]$ является M -матрицей.

Определение 4. Матрица $A \in \mathbb{R}^{n \times n}$ называется H -матрицей, если ее компарант является M -матрицей. Если каждая вещественная матрица $A \in [A]$ является H -матрицей, то $[A]$ называется интервальной H -матрицей.

В соответствии с (2), (3), компаранты матриц из $[A^{(k)}]$ — M -матрицы. Иными словами, каждая построенная матрица $[A^{(k)}]$ является H -матрицей, что гарантирует улучшение любого достаточно широкого начального интервального вектора-приближения интервальным методом Гаусса — Зейделя [6] и позволяет получать для множеств $\Xi^{(k)}$ их внешние оценки $[X^{(k)}]$, исходя из лебеговой меры которых проводится распознавание. В качестве начального приближения может быть выбран брус $([-B, B], \dots, [-B, B])^T$, заведомо содержащий объединенные множества решений $\Xi^{(k)}$ для некоторого $B > 0$.

3.2. Алгоритм распознавания и его вычислительная сложность. Предлагается следующий алгоритм распознавания растровых изображений, использующий предложенную эвристику.

Вход: матрицы $\{\mathcal{A}^{(k)}\}_{k=1}^N$, соответствующие эталонным изображениям, и матрица \mathcal{A} , соответствующая зашумленному изображению.

Выход: индекс p матрицы $\mathcal{A}^{(p)} \in \{\mathcal{A}^{(k)}\}_{k=1}^N$ — результат распознавания.

1. С использованием матриц $\{\mathcal{A}^{(k)}\}_{k=1}^N$ и матрицы \mathcal{A} в соответствии с (1)–(3) сформировать матрицы $\{[A^{(k)}]\}_{k=1}^N$.

2. С помощью метода *Encl* вычислить значения $[X^{(k)}]$, $k = \overline{1, N}$.

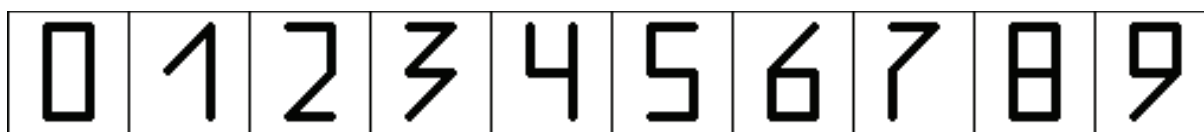
3. Найти p , такое что $\mu([X^{(p)}]) = \min_{1 \leq k \leq N} \mu([X^{(k)}])$, и выдать p в качестве результата распознавания.

Вычислительная сложность алгоритма распознавания, использующего в качестве *Encl* интервальный метод Гаусса — Зейделя ($Encl = GS$), равна

$$C(N, n, GS) = O(NN_{GS}n^2).$$

Здесь n — размерность квадратных матриц изображений; N — число эталонных матриц; N_{GS} — количество итераций интервального метода Гаусса — Зейделя, которое в ходе экспериментов полагалось равным 25.

4. Вычислительный эксперимент. Для исследования эффективности предложенного алгоритма в качестве данных для тестирования использовались изображения цифр: черно-белые (бинарные) изображения и изображения в градациях серого цвета с разрешением 20×20 , 35×35 и 50×50 пикселей (см. рисунок). Кроме того, проведены эксперименты с



Эталонные изображения цифр

цифрами шрифтов Times New Roman, Arial, Courier New. Наиболее сложным для распознавания является шрифт, представленный на рисунке. Элементы матриц тестовых изображений принимают два значения:

$$a_{ij}^{(k)} = \begin{cases} c_1, & \text{если пиксел в позиции } ij \text{ белого цвета,} \\ c_2, & \text{если пиксел в позиции } ij \text{ черного цвета.} \end{cases}$$

В случае если изображения черно-белые, $c_1 = 1$ и $c_2 = 0$, в случае если изображения представлены в градациях серого цвета, c_1 и c_2 могут принимать целые значения в диапазоне от 0 до 255.

Будем называть уровнем шума (в процентах) целое число Q , лежащее в интервале $[0, 100]$. Изображения зашумлялись следующим образом. Генератор случайных чисел давал целые числа в интервале $[0, 100]$. Просматривался каждый пиксел изображения и анализировалось число, сгенерированное генератором случайных чисел на очередном шаге. Если число попало в интервал $[0, Q]$, то пиксел зашумлялся. В ходе зашумления бинарного изображения, когда зашумление осуществляется инвертированием пикселей, т. е. заменой пиксела белого цвета на пиксел черного цвета или наоборот, при нулевом уровне шума ($Q = 0$) получается чистое изображение, при 50 %-м уровне шума ($Q = 50$) инвертируется в среднем 50 % пикселей, при 100 %-м уровне шума изображение инвертируется полностью.

Для каждой упорядоченной пары изображений (A, B) проводилось зашумление изображения A , после чего запускался алгоритм распознавания. Для каждой такой пары выполнялось по 100 испытаний. Доля правильных решений (процент распознавания) P вычислялась следующим образом:

$$P = \frac{\text{Количество правильных ответов}}{\text{Количество испытаний}} \cdot 100 \%$$

4.1. *Сравнение результатов распознавания с помощью предложенной эвристики с результатами распознавания с помощью известных алгоритмов.* В ходе эксперимента по распознаванию монохромных изображений значение $\nu = 10$ для преобразований (1) было выбрано таким образом, чтобы инвертирование пиксела — изменение его значения с 0 на 1 или наоборот — приводило к изменению элементов $a_{ij}^{(k)}$ модифицированных матриц эталонных изображений в интервале $[a_{ij}^{(k)} - \delta, a_{ij}^{(k)} + \delta]$ с радиусом $\delta \leq a_{ij}^{(k)}/10 \leq 0,1$.

Результаты вычислительного эксперимента приведены в табл. 1, 2. В табл. 1 показаны результаты распознавания при уровне шума Q от 31 до 48 % для матриц, представляющих тестовые изображения с разрешением 20×20 , 35×35 и 50×50 пикселей. Из табл. 1 следует, что с увеличением разрешения изображений, подаваемых на вход алгоритма, процент распознавания P растет, при этом среднеквадратичное отклонение данной величины по всем 90 упорядоченным парам изображений цифр при $Q \leq 45$ уменьшается.

В табл. 2 представлены результаты расчетов при $Q = 45 \%$, $n = 50$. При уровне шума до 30 % процент распознавания составляет не менее 99,9 %.

Таблица 1
Процент распознавания тестовых изображений
при различных уровнях шума Q

$Q, \%$	$P, \%$ ($\sigma, \%$)		
	$n = 20$	$n = 35$	$n = 50$
31	99,78 (1,07)	99,96 (0,42)	100,00 (0)
32	99,70 (1,11)	99,98 (0,20)	100,00 (0)
33	99,53 (1,44)	99,91 (0,59)	99,97 (0,23)
34	99,22 (2,32)	99,88 (0,61)	99,97 (0,23)
35	98,84 (2,47)	99,78 (1,41)	99,97 (0,23)
36	98,61 (2,65)	99,73 (1,01)	99,91 (0,64)
37	97,83 (3,33)	99,51 (1,68)	99,94 (0,38)
38	97,36 (3,47)	99,33 (1,99)	99,90 (0,62)
39	96,20 (4,09)	99,22 (1,74)	99,84 (0,94)
40	95,12 (4,66)	98,99 (1,60)	99,73 (0,99)
41	93,18 (5,14)	97,61 (3,46)	99,56 (1,52)
42	90,33 (5,82)	96,30 (3,46)	99,31 (1,70)
43	87,39 (5,95)	94,60 (4,05)	98,52 (2,99)
44	82,27 (5,54)	91,21 (5,00)	96,92 (3,91)
45	80,01 (6,06)	87,86 (5,35)	94,60 (5,06)
46	74,51 (5,46)	83,08 (5,05)	90,52 (5,37)
47	68,90 (5,27)	75,71 (6,09)	84,49 (6,44)
48	62,87 (5,72)	67,60 (6,29)	75,03 (7,00)

Результаты вычислительного эксперимента показывают, что предложенная эвристика и алгоритм, основанный на ее расчете, по качеству распознавания не уступают известным алгоритмам, применяемым для распознавания растровых изображений. Так, для параметрических алгоритмов “Кора”, “R-метод”, “TEMP” и “CORAL” процент распознавания изображений при тех же условиях, при которых проводился вычислительный эксперимент, составляет 80 % при $Q = 43 \%$, для алгоритма, основанного на монохромной морфологии, — 80 % при $Q = 45 \%$ [7]. Для алгоритмов распознавания с помощью нейронных сетей процент распознавания не превышает 90 % для уровня шума $Q \geq 46 \%$ [1]. Из приведенных результатов следует, что при одних и тех же значениях уровня шума предложенный подход дает тот же или больший процент распознавания при надлежащем разрешении изображений.

4.2. *Распознавание с помощью предложенной эвристики и распознавание с помощью матричных норм разности матриц в качестве эвристики.* Проведено сравнение эффективности распознавания с помощью предложенной эвристики и эффективности распознавания с помощью матричных норм, которые также могут быть использованы при распознавании растровых изображений.

Для эвристики, представляющей собой минимизацию евклидова расстояния между матрицей \mathcal{A} , сопоставляемой с распознаваемым изображением, и матрицами $\mathcal{A}^{(k)}$, сопоставляемыми с эталонными изображениями:

$$\rho(\mathcal{A}, \mathcal{A}^{(k)}) = \sqrt{\sum_{i=1}^n \sum_{j=1}^n (a_{ij} - a_{ij}^{(k)})^2},$$

Таблица 2

Результаты распознавания тестовых изображений при $Q = 45\%$, $n = 50$

$A \setminus B$	0	1	2	3	4	5	6	7	8	9
0	—	97	92	100	91	86	96	94	85	97
1	98	—	98	98	98	98	94	92	98	99
2	94	94	—	94	99	97	100	96	98	89
3	99	96	96	—	97	96	97	91	100	95
4	100	93	99	96	—	92	99	99	93	95
5	94	97	95	97	93	—	94	99	90	96
6	96	93	100	95	97	94	—	95	99	98
7	99	99	99	91	98	96	96	—	99	99
8	76	92	93	96	89	84	92	97	—	94
9	96	97	93	89	94	95	98	97	96	—

результатом распознавания является номер эталонного изображения p , дающий минимум ρ . Минимизация расстояния ρ между эталонной матрицей $\mathcal{A}^{(k)}$ и распознаваемой матрицей \mathcal{A} часто позволяет проводить успешное распознавание при решении задачи в рассматриваемой постановке. Вместе с тем возможны ситуации, когда для пары матриц $\mathcal{A}^{(1)}$ и $\mathcal{A}^{(2)}$ и распознаваемой матрицы \mathcal{A} имеет место нервенство

$$\rho(\mathcal{A}, \mathcal{A}^{(1)}) > \rho(\mathcal{A}, \mathcal{A}^{(2)}),$$

тогда как для большего количества позиций ij в этих матрицах — неравенство

$$|a_{ij} - a_{ij}^{(1)}| < |a_{ij} - a_{ij}^{(2)}|,$$

и \mathcal{A} действительно будет получена в результате зашумления матрицы $\mathcal{A}^{(1)}$.

Проведен эксперимент, показавший, что с увеличением радиуса δ интервалов, в которых происходят изменения элементов эталонных матриц при $|c_1 - c_2| < \delta$, растет доля задач, в которых процент распознавания с помощью предложенной эвристики оказывается большим, чем при распознавании с помощью евклидова расстояния.

Пусть S — доля задач в общем числе решаемых задач распознавания, в которых предложенная эвристика дает верное распознавание, тогда как минимизация евклидова расстояния не позволяет провести правильное распознавание. Для того же набора эталонных изображений, представленных в градациях серого цвета со значениями $c_1 = 110$ и $c_2 = 120$, в ходе эксперимента с тем же выбором зашумляемых элементов, изменяемых в интервалах радиуса δ , получены значения S , приведенные в табл. 3. Там же приведены значения S при $c_1 = 119$ и $c_2 = 120$. Доля задач, в которых ситуация оказалась обратной при уровне шума Q от 30 до 50%, не превышает 0,01% числа всех решенных задач распознавания для данного значения Q .

Подобная тенденция увеличения доли задач S при росте δ сохраняется и при других значениях c_1, c_2 , если значение $|c_1 - c_2|$ достаточно мало по сравнению со значением δ .

Для эвристики, состоящей в минимизации расстояния, определяемого по формуле

$$\rho'(\mathcal{A}, \mathcal{A}^{(k)}) = \sum_{i=1}^n \sum_{j=1}^n |d_{ij} - a_{ij}^{(k)}|,$$

могут быть построены задачи, демонстрирующие сглаживание при распознавании с помощью предложенной в работе эвристики больших изменений малого числа пикселей зашумляемого эталонного изображения и позволяющие проводить распознавание исходя из изменений большего числа менее зашумленных пикселей изображения.

Таблица 3

Доля задач S при $Q = 44\%$, $c_2 = 120$

δ	$c_1 = 110$		$c_1 = 119$	
	$S, \%$	$P, \%$	$S, \%$	$P, \%$
10	0	100,00	22,8	99,71
25	5,4	99,93	37,5	99,60
50	7,4	99,79	47,3	99,80
75	16,2	99,72	46,4	99,72
100	23,5	99,81	46,4	99,82

Пусть \mathcal{A} — матрица некоторого изображения. С использованием матрицы \mathcal{A} построим две матрицы. Первая матрица — матрица \mathbf{A} , в которой относительно большому зашумлению подвергнуто малое число элементов, вторая — та же матрица, в которой относительно малому зашумлению подвергнуто большое число элементов. Обе матрицы могут быть построены таким образом, что при распознавании минимизацией $p'(\rho')$ в качестве результата будет получена вторая матрица.

Положим $\mathcal{A}^{(1)} := \mathcal{A}$, $\mathcal{A}^{(2)} := \mathcal{A}$. Построим матрицы $\mathcal{A}_M^{(1)}$ и $\mathcal{A}_M^{(2)}$, так чтобы в каждой из них 90% случайно выбранных элементов были равны 1, а 10% — 0. Для каждой из этих матриц процедура заполнения проводится независимо, вероятности выбора 1 и выбора 0 в качестве значения элемента матрицы одинаковы.

Пусть η_1, η_2 — независимые случайные величины, с равными вероятностями принимающие значения -1 и 1 , ξ_1, ξ_2 — независимые случайные величины, имеющие равномерное распределение на отрезке $[0, 1]$, δ_1, δ_2 — некоторые заданные радиусы интервалов, задающие границы области изменения элементов матриц $\mathcal{A}^{(1)}$ и $\mathcal{A}^{(2)}$ соответственно ($\delta_1 > \delta_2$).

С использованием значений элементов $\mathcal{A}_M^{(1)}$ модифицируем 10% элементов матрицы $\mathcal{A}^{(1)}$, определяемых элементами $(\mathcal{A}_M^{(1)})_{ij}$: если $(\mathcal{A}_M^{(1)})_{ij} = 0$, то

$$(\mathcal{A}^{(1)})_{ij} := (\mathcal{A}^{(1)})_{ij} + \lfloor \eta_1 \xi_1 \cdot \delta_1 \rfloor,$$

и модифицируем 90% элементов $\mathcal{A}^{(2)}$ в позициях, определяемых элементами $(\mathcal{A}_M^{(2)})_{ij}$: если $(\mathcal{A}_M^{(2)})_{ij} = 1$, то

$$(\mathcal{A}^{(2)})_{ij} := (\mathcal{A}^{(2)})_{ij} + \lfloor \eta_2 \xi_2 \cdot \delta_2 \rfloor.$$

Здесь как “ $\lfloor \cdot \rfloor$ ” — обозначена операция взятия целой части числа.

Будем считать, что распознаваемое изображение — это изображение с матрицей \mathcal{A} , а эталонные — изображения с матрицами $\mathcal{A}^{(1)}$ и $\mathcal{A}^{(2)}$.

Если элементы матрицы \mathcal{A} случайным образом принимают целые значения в интервале $[110, 120]$ с одной и той же вероятностью для каждого значения, то при $\delta_1 = 60$ и $\delta_2 = 10$ получим матрицы, которые можно понимать как матрицы некоторых растровых изображений. Выбирая из сгенерированных матриц $\mathcal{A}^{(1)}$ и $\mathcal{A}^{(2)}$ такие, что $\rho'(\mathcal{A}, \mathcal{A}^{(1)}) > \rho'(\mathcal{A}, \mathcal{A}^{(2)})$, в более чем 95% случаев запуска алгоритма распознавания с помощью предложенной эвристики в результате получим $\mathcal{A}^{(1)}$, что естественно считать правильным распознаванием, поскольку 90% элементов \mathcal{A} и $\mathcal{A}^{(1)}$ совпадают, тогда как 90% элементов \mathcal{A} и $\mathcal{A}^{(2)}$ различаются. Иными словами, алгоритм проводит распознавание на основе большей части пикселей эталонного изображения, не измененных или слабо измененных в ходе зашумления, уменьшая влияние на распознавание относительно небольшого числа (10% в приведенном примере) сильноизмененных пикселей.

Приведенные примеры задач распознавания и результаты экспериментов с ними показывают, что для эвристик, состоящих в минимизации расстояний ρ и ρ' , представляющих собой нормы матриц разностей эталонных и распознаваемого изображений, характерна чувствительность к значительным изменениям относительно малого количества пикселей эталонных изображений при зашумлении, что может приводить к ошибкам распознавания, в то время как предложенная в работе эвристика позволяет выполнять распознавание, учитывая значения изменений большинства пикселей эталонных изображений.

5. Выводы. Предложен алгоритм распознавания растровых изображений. Используемая алгоритмом эвристика состоит в минимизации лебеговой меры внешней оценки объединенного множества решений интервальной системы линейных алгебраических уравнений, сопоставляемой с эталонной матрицей с учетом распознаваемой. Алгоритм не требует обучения и имеет квадратичную относительно размерности входных матриц вычислительную сложность.

Как показывают проведенные вычислительные эксперименты, предложенная эвристика может быть эффективно применена для решения задачи распознавания растровых изображений. Процент распознавания увеличивается с ростом разрешения изображений.

Список литературы

1. DEMUTH H., BEALE M. Neural network toolbox for use with MATLAB. User's guide. Version 4. [Electron. resource]. <http://cs.mipt.ru/docs/comp/eng/develop/software/matlab/nnet/main.pdf>.
2. КИРНОС Э. А., ПЫТЬЕВ Ю. П. О параметрических алгоритмах распознавания // Вестн. Моск. ун-та. Сер. 3. Физика. Астрономия. 2003. № 1. С. 16–18.
3. KIRNOS E. A., PYT'EV YU. P., DJUKOVA E. V. Training the kora type algorithms // Pattern Recognit. Image Anal. 2002. V. 12, N 1. P. 19–24.
4. ШАРЫЙ С. П. Конечномерный интервальный анализ. [Электрон. ресурс]. <http://www.nsc.ru/interval/Library/InteBooks/SharyBook.pdf>.
5. OETTLI W. On the solution set of a linear system with inaccurate coefficients // SIAM J. Numer. Anal. 1965. V. 2, N 1. P. 115–118.
6. NEUMAIER A. Interval methods for systems of equations. Cambridge: Cambridge Univ. Press, 1990.
7. КИРНОС Э. А. Сравнительный анализ морфологических методов интерпретации изображений: Автореф. дис. ... канд. физ.-мат. наук. М.: Изд-во Моск. гос. ун-та, 2004.

*Пролубников Александр Вячеславович — канд. физ.-мат. наук,
доц. Омского государственного университета
им. Ф. М. Достоевского; e-mail: a.v.prolubnikov@mail.ru*

Дата поступления — 15.08.12