

Интервальный подход к решению задачи распознавания числовых матриц

А.В. ПРОЛУБНИКОВ

Омский государственный университет

e-mail: a.v.prolubnikov@mail.ru

Предлагается подход к решению задачи распознавания числовых матриц, использующий интервальные методы. Рассматриваются результаты вычислительного эксперимента. В качестве приложения разработанного алгоритма рассматривается задача распознавания растровых изображений.

1. Введение

Проблема автоматического анализа формы и состояния объектов, информация о которых представлена в матричном виде, является актуальной во многих отраслях человеческой деятельности. Очевидным примером приложения методов такого анализа является распознавание растровых изображений, которые могут быть представлены матрицами, числовые значения элементов которых соответствуют яркости, цвету или каким-либо другим характеристикам пикселей, совокупность которых представляет изображение. Автоматизация процесса анализа и интерпретации изображений достаточно трудна как в связи с тем, что обработка большого объёма информации, содержащейся в изображениях, требует значительных вычислительных мощностей, так и по той причине, что сами процессы регистрации человеком изображений или каких-либо других объектов, представляемых числовыми матрицами, могут быть достаточно сложны, плохо формализуемы или вообще неизвестны.

Кроме того, при регистрации изображений техническими средствами часто неизбежны погрешности измерения, возможно наличие шума при регистрации изображений, как, например, зашумление отдельных фрагментов или всего 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 произведено распознавание.

Предлагаемый в этой работе подход к распознаванию числовых матриц, применённый, в частности, к распознаванию растровых изображений, использует в качестве эвристики, при помощи которой производится распознавание, минимизацию лебеговой меры объединённого множества решений интервальной системы линейных алгебраических уравнений (ИСЛАУ). ИСЛАУ ставится в соответствие каждой эталонной матрице с учётом распознаваемой зашумлённой матрицы. Характерным отличием представляемого подхода к распознаванию от традиционных является отсутствие стадии обучения алгоритма распознавания.

1.1. Постановка задачи

Матричная постановка задачи следующая. Имеется набор из N эталонных квадратных $n \times n$ -матриц $\{\mathcal{A}^{(k)}\}_{k=1}^N$ с элементами $a_{ij}^{(k)} \in \mathbb{R}$. В ходе зашумления одной из матриц — матрицы $\mathcal{A}^{(p)}$ — получена некоторая матрица \mathcal{A} . Известно, что значение элементов матриц могло быть изменено в пределах интервалов $[a_{ij}^{(k)} - \Delta, a_{ij}^{(k)} + \Delta]$, $\Delta > 0$. Необходимо определить p .

Матрицы $\mathcal{A}^{(1)}, \dots, \mathcal{A}^{(N)}, \mathcal{A}$ без ограничения общности можно считать квадратными. В противном случае, если матрицы из набора эталонных и распознаваемая матрица — $m \times n$ -матрицы и $m < n$, то к каждой матрице добавляются $n - m$ нулевых строк, если $n < m$, добавляется $m - n$ нулевых столбцов.

2. Использование внешнего оценивания множества решений ИСЛАУ для распознавания числовых матриц

2.1. Минимизация лебеговой меры внешней оценки объединённого множества решений ИСЛАУ как эвристика, на основе которой производится распознавание

Определение. Для ИСЛАУ $\mathbf{A}x = \mathbf{b}$, где \mathbf{A} — интервальная $n \times n$ матрица, \mathbf{b} — некоторый интервальный n -мерный вектор, объединённое множество решений — это множество $\Xi(\mathbf{A}, \mathbf{b})$:

$$\Xi(\mathbf{A}, \mathbf{b}) = \{x \in \mathbb{R}^n \mid \exists \mathbf{A} \in \mathbf{A}, \exists \mathbf{b} \in \mathbf{b}: \mathbf{A}x = \mathbf{b}\}. \quad (1)$$

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

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

где a_{ij} — элементы распознаваемой числовой матрицы \mathcal{A} . Элементы матрицы $\mathbf{A}^{(k)}$ — интервалы, характеризующие изменения элементов эталонной матрицы $\mathcal{A}^{(k)}$, необходимые для получения из неё распознаваемой матрицы \mathcal{A} .

Предлагаемый подход основан на предположении, что чем меньше разброс в решениях точечных СЛАУ, составляющих ИСЛАУ $\mathbf{A}^{(k)}x = b$ с некоторой точечной правой частью $b \in \mathbb{R}^n$, тем вероятнее, что именно зашумлением эталонной матрицы $\mathcal{A}^{(k)}$ получена матрица \mathcal{A} .

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

2.2. Выбор вектора правой части ИСЛАУ $\mathbf{A}^{(k)}x = b$

Наряду с матрицей $\mathbf{A}^{(k)}$ вектор правой части b определяет множество $\Xi(\mathbf{A}^{(k)}, b)$, поэтому его выбор важен для эффективности распознавания. Вектор правой части в рассматриваемых ИСЛАУ выбирается точечным для того, чтобы было возможно более точное

внешнее оценивание объединённого множества решений. В случае точечной правой части интервальная оболочка объединённого множества решений имеет покомпонентно меньший радиус, нежели в случае интервальной правой части, что приводит к уменьшению расстояния между объединённым множеством решений и его интервальной оболочкой [4], что, в свою очередь, позволяет производить внешнее оценивание с большей точностью.

В случае точечной правой части b , множество решений $\Xi(\mathbf{A}^{(k)}, b)$ принимает следующий вид:

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

В результате описанных далее преобразований матрицы $\mathbf{A}^{(k)}$ ($k = \overline{1, N}$) становятся неособенными, то есть любая точечная матрица $A \in \mathbf{A}^{(k)}$ обратима. В этом случае

$$\Xi(\mathbf{A}^{(k)}, b) = \{A^{-1}b \mid A \in \mathbf{A}^{(k)}\}.$$

Будем рассматривать множество $\Xi(\mathbf{A}, b)$ как образ множества матриц \mathbf{A} при отображении L_b :

$$L_b : \mathbf{A} \rightarrow \mathbb{R}^n,$$

где для $A \in \mathbf{A}$

$$L_b(A) = A^{-1}b.$$

Пусть α_{ij} — элементы матрицы A^{-1} обратной к матрице $A \in \mathbf{A}$. Для i -ой компоненты вектора $x \in \Xi(\mathbf{A}, b)$ будем иметь:

$$x_i = \sum_{j=1}^n \alpha_{ij} b_j = \frac{1}{\det A} \sum_{j=1}^n A_{ji} b_j,$$

где $A \in \mathbf{A}$ такая, что $x = A^{-1}b$. То есть компоненты вектора b — это веса столбцов, с которыми при получении компонент вектора $x \in \Xi(\mathbf{A}, b)$, $x = L_b(A)$ производится суммирование элементов строк матрицы A^{-1} . A_{ji} — алгебраические дополнения элементов i -го столбца матрицы A .

Если в качестве правой части использовать вектор $e = (1, \dots, 1)^\top$, то компоненты вектора $x = A^{-1}e$ для $A \in \mathbf{A}$ будут в равной мере зависеть от алгебраических дополнений элементов любого столбца, а значит, будут в среднем примерно одинаково зависеть от всех элементов матрицы. Чего не будет при использовании в этом качестве вектора с не равными компонентами. Так, например, если использовать какой-либо из ортов e_i стандартного базиса в \mathbb{R}^n , то получим для $x \in \Xi(\mathbf{A}, e_i)$, $x = L_{e_i}(A)$:

$$x = \frac{1}{\det A} (A_{1i}, \dots, A_{ni}).$$

То есть компоненты векторов, составляющих $\Xi(\mathbf{A}^{(k)}, e_i)$ и являющихся образами матриц $A \in \mathbf{A}$ при отображении L_{e_i} , будут в меньшей степени зависеть от элементов i -ых столбцов матриц $A \in \mathbf{A}^{(k)}$, чем от элементов других столбцов. Элементы из позиций в i -ом столбце матрицы A не будут влиять на значения A_{1i}, \dots, A_{ni} алгебраических дополнений, а значит не будут в той же мере, что и другие элементы матрицы, влиять на компоненты векторов, составляющих $\Xi(\mathbf{A}^{(k)}, e_i)$. Мы исходим из того, что все элементы эталонных матриц и распознаваемой матрицы должны в равной мере учитываться при распознавании.

Таким образом, рассматриваемые далее ИСЛАУ имеют вид:

$$\mathbf{A}^{(k)} x = e. \tag{3}$$

2.3. Вычислительная сложность распознавания

Множество $\Xi^{(k)} \stackrel{def}{=} \Xi(\mathbf{A}^{(k)}, e)$ 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 этого множества с помощью гиперплоскостей, ограничивающих его — это задача с экспоненциально растущей относительно размерности вычислительной сложностью. Поэтому мы ограничимся оцениванием сверху лебеговой меры множества $\Xi^{(k)}$ лебеговой мерой некоторого приближения интервальной оболочки $\square\Xi^{(k)}$ этого множества. Интервальная оболочка $\square\Xi^{(k)}$ — это минимальный по включению интервальный вектор, содержащий $\Xi^{(k)}$, то есть такой брус со сторонами параллельными координатным осям, что $\Xi^{(k)} \subset \square\Xi^{(k)}$, и для любого бруса \mathbf{c} такого, что $\Xi^{(k)} \subset \mathbf{c}$, имеем $\square\Xi^{(k)} \subseteq \mathbf{c}$.

Обозначим как $\mathbf{X}^{(k)}$ получаемое с помощью какого-либо численного метода интервального анализа приближение интервальной оболочки $\square\Xi^{(k)}$. $\mathbf{X}^{(k)}$ — это некоторый брус

$$\mathbf{X}^{(k)} = ([x_1^k, \bar{x}_1^k], \dots, [x_n^k, \bar{x}_n^k])^\top,$$

$\underline{x}_i^k \leq \bar{x}_i^k$, такой что $\square\Xi^{(k)} \subseteq \mathbf{X}^{(k)}$. Лебегова мера μ бруса $\mathbf{X}^{(k)}$ рассчитывается как

$$\mu(\mathbf{X}^{(k)}) = (\bar{x}_1^k - \underline{x}_1^k) \cdot \dots \cdot (\bar{x}_n^k - \underline{x}_n^k). \quad (4)$$

Вычисление внешних оценок $\mathbf{X}^{(k)}$ может быть произведено с помощью методов интервального анализа, разработанных для внешнего оценивания объединённых множеств решений ИСЛАУ. При вычислительной сложности выбранного алгоритма внешнего оценивания $Encl$ равной $C_{Encl}(n)$ общая трудоёмкость алгоритма, производящего распознавание, исходя из сравнения N значений $\mu(\mathbf{X}^{(k)})$, составит

$$C(N, n, Encl) = O(N \cdot C_{Encl}(n)). \quad (5)$$

В случае $C_{Encl}(n) = O(n^2)$ мы получим 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

3.1. Модификация матриц, подаваемых на вход алгоритма

Для улучшения качества распознавания по предложенному алгоритму производится следующая модификация матриц, подаваемых на его вход. Каждый элемент как эталонных матриц, так и зашумлённой матрицы увеличивается на одно и то же значение v ($v > 0$):

$$a_{ij} := a_{ij} + v, \quad (6)$$

$$a_{ij}^{(k)} := a_{ij}^{(k)} + v. \quad (7)$$

Преобразования (6) и (7) 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 производить распознавание, после проведённых преобразований сохраняется.

В результате проведения преобразований (6) и (7) уменьшается отношение радиуса интервала изменения элемента эталонной матрицы к модулю этого элемента. Так, если до преобразований (6) и (7) 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 ростом v и составляет $\Delta/|a_{ij}^{(k)} + v|$.

Такое уменьшение отношения радиуса Δ к величине модуля элемента матрицы необходимо по следующей причине. Если значение Δ близко к модулям значений элементов эталонных матриц, то интервальные матрицы $\mathbf{A}^{(k)}$, построенные в соответствии с (2), будут содержать такие точечные матрицы, что векторы $x = L_e(A) = A^{-1}e$ для $A \in \mathbf{A}^{(k)}$ могут столь сильно отличаться друг от друга, что производить распознавание, исходя из близости или удалённости этих векторов для всех матриц из $\mathbf{A}^{(k)}$, минимизируя по k меру множества $\Xi^{(k)}$, нельзя. Если же радиус Δ мал относительно модулей элементов зашумляемой эталонной матрицы, что мы будем иметь после проведения преобразований (6) и (7), то такое распознавание во многих случаях возможно.

Так, при распознавании монохромных изображений, если пикселям белого цвета соответствуют элементы матриц равные 1, пикселям чёрного цвета — 0, а изменению цвета пикселей при зашумлении соответствует изменение значения элемента матрицы с 0 на 1 или наоборот, то величина Δ возможного изменения элемента матрицы будет или больше модуля элемента, или равна ему. Для таких матриц распознавание по предложенному подходу с использованием какого-либо метода внешнего оценивания объединённого множества решений ИСЛАУ будет невозможно, тогда как оно эффективно производится после модификации (6) и (7) входных матриц с $v = 10$, когда $\Delta/|a_{ij}^{(k)} + v| \leq 0.1$.

Проведение преобразований (6) и (7) входных точечных матриц эквивалентно следующему преобразованию интервальных матриц вида (2):

$$\mathbf{A}^{(k)} := \mathbf{A}^{(k)} + v\mathbf{E}, \quad (8)$$

где \mathbf{E} — матрица с элементами $\mathbf{E}_{ij} = [1, 1]$, $i, j = \overline{1, n}$.

После проведения преобразования (8) матриц $\mathbf{A}^{(k)}$ модифицируем их диагональные элементы:

$$\mathbf{A}^{(k)} := \mathbf{A}^{(k)} + \mathbf{D}, \quad (9)$$

где \mathbf{D} — диагональная интервальная матрица с элементами $\mathbf{D}_{ii} = [D, D]$. D задаётся следующим образом:

$$D = \max_{1 \leq k \leq N} D^{(k)}, \quad (10)$$

где

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

Матрицы $\mathbf{A}^{(k)}$ после выполнения преобразований (9) — это H -матрицы, что позволяет в ходе распознавания выполнять с ними необходимые вычисления по нахождению внешних оценок $\mathbf{X}^{(k)}$.

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

$$ij\text{-ый элемент } \langle A \rangle = \begin{cases} |a_{ij}|, & \text{если } i = j, \\ -|a_{ij}|, & \text{если } i \neq j. \end{cases}$$

Определение. 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 и $s > \rho(P)$, $\rho(P)$ — спектральный радиус P . Матрица $\mathbf{A} \in \mathbb{IR}^{n \times n}$ называется интервальной M -матрицей, если каждая вещественная матрица $A \in \mathbf{A}$ является M -матрицей.

Определение. Матрица $A \in \mathbb{R}^{n \times n}$ называется H -матрицей, если её компарант является M -матрицей. Матрица $\mathbf{A} \in \mathbb{IR}^{n \times n}$ называется интервальной H -матрицей, если каждая вещественная матрица $A \in \mathbf{A}$ является H -матрицей.

Для матриц, построенных в соответствии с (2), (8), (9), компарант $\langle \mathbf{A}^{(k)} \rangle$ — M -матрица. То есть каждая построенная матрица $\mathbf{A}^{(k)}$ — H -матрица, 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 $\mathbf{X}^{(k)}$, исходя из лебеговой меры которых производится распознавание. Как начальное приближение может быть выбран брус $([-B, B], \dots, [-B, B])^\top \in \mathbb{IR}^n$, заведомо содержащий объединённые множества решений $\Xi^{(k)}$ для некоторого $B > 0$.

3.2. Выбор начального приближения для множества $\Xi^{(k)}$ и ограничение на величину v

Без ограничения общности можно считать, что все элементы a_{ij} матриц $A \in \mathbf{A}^{(k)}$ положительны. В противном случае можно провести предварительные преобразования (6) и (7) входных матриц с некоторым таким $v_0 > 0$, что значение v_0 больше модулей всех отрицательных элементов эталонных и распознаваемой матриц.

После преобразований (8) и (9) матриц $\mathbf{A}^{(k)}$, построенных в соответствии с (2), любая матрица $A \in \mathbf{A}^{(k)}$ — матрица с диагональным преобладанием. Пусть для матриц $A \in \mathbf{A}^{(k)}$

$$R_i(A) \stackrel{def}{=} a_{ii} - \sum_{j \neq i} a_{ij},$$

$$R_* \stackrel{def}{=} \min_{1 \leq i \leq n} R_i(A).$$

Для матрицы A справедлива [7] оценка:

$$\|A^{-1}\|_\infty \leq \frac{1}{R_*(A)}.$$

Из (10) следует, что $R_*(A) \geq v(n-1)$, поэтому

$$\|A^{-1}\|_\infty \leq \frac{1}{v(n-1)}. \quad (11)$$

Если $x \in \Xi^{(k)}$, то для $A \in \mathbf{A}^{(k)}$ такой, что $x = A^{-1}e$, будем иметь:

$$\|x\|_\infty \leq \|A^{-1}\|_\infty \|e\|_\infty,$$

то есть

$$\|x\|_\infty \leq \frac{1}{v(n-1)}. \quad (12)$$

Значение в правой части (12) позволяет задать начальное приближение для интервального метода Гаусса-Зейделя — брус $([-B, B], \dots, [-B, B])^\top$, где

$$B = \frac{1}{v(n-1)}. \quad (13)$$

Из (12) также следует, что рост v может привести к такому уменьшению нормы векторов, составляющих множества $\Xi^{(k)}$, что лебеговы меры приближений интервальных оболочек этих множеств, получаемых с помощью какого-либо алгоритма внешнего оценивания, не будут отражать специфику эталонных матриц. Это будет происходить как в силу того, что с ростом 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 $\mu(\mathbf{X}^{(p)})$ от прочих значений $\mu(\mathbf{X}^{(k)})$, $k \neq p$. При проведении численных экспериментов полагалось $v = 10a$, где a — максимальное значение модуля элемента в эталонных матрицах.

3.3. Алгоритм распознавания и его вычислительная сложность

Предлагается следующий алгоритм распознавания числовых матриц.

Алгоритм распознавания числовых матриц

Вход: Эталонные матрицы $\{\mathcal{A}^{(k)}\}_{k=1}^N$ и зашумлённая матрица \mathcal{A} .

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

1. По матрицам $\{\mathcal{A}^{(k)}\}_{k=1}^N$ и матрице \mathcal{A} в соответствии с (2), (8), (9) сформировать матрицы $\{\mathbf{A}^{(k)}\}_{k=1}^N$.
2. Рассчитать с помощью метода *Encl* значения $\mathbf{X}^{(k)}$, $k = \overline{1, N}$.
3. Найти p такое, что $\mu(\mathbf{X}^{(p)}) = \min_{1 \leq k \leq N} \mu(\mathbf{X}^{(k)})$, и выдать p в качестве результата распознавания.

Вычислительная сложность алгоритма распознавания, использующего в качестве *Encl* интервальный метод Гаусса-Зейделя: $Encl = GS$, составит

$$C(N, n, GS) = O(N \cdot N_{GS} \cdot n^2), \quad (14)$$

где n — размерность квадратных матриц изображений, N — число эталонных матриц, N_{GS} — количество производимых итераций интервального метода Гаусса-Зейделя, которое в ходе проведённых экспериментов полагалось равным 20.

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 , 50×50

и 100×100 пикселей (рис. 1). Кроме того, были проведены эксперименты с цифрами шрифтов Times New Roman, Arial, Courier New. Наиболее сложным для распознавания шрифтом является представленный на рис. 1. Элементы матриц тестовых изображений принимают два значения:

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

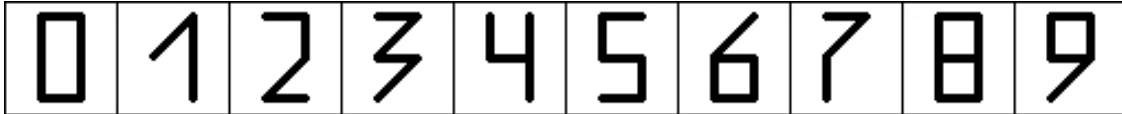


Рис. 1. Эталонные изображения цифр

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

Назовём уровнем шума в $Q\%$ 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 попадало в интервал $[0, Q]$, то пиксел зашумлялся. В ходе зашумления бинарного изображения, когда зашумление происходит инвертированием пикселей, то есть заменой пиксела белого цвета на пиксел чёрного цвета или наоборот, при 0% -ом уровне шума ($Q = 0$) получается чистое изображение, при 50% -ом уровне шума ($Q = 50$) инвертируется в среднем 50% пикселей, при уровне шума в 100% изображение инвертируется полностью.

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

$$P = \frac{\text{количество правильных ответов}}{\text{количество испытаний}} \times 100.$$

В ходе проведённого эксперимента по распознаванию монохромных изображений значение $v = 10$ для преобразований (6) и (7) было выбрано так, чтобы инвертирование пиксела — изменение его значения с 0 на 1 или наоборот — давало изменение 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 отражены в таблицах 1 и 2 и проиллюстрированы графиком на рис. 2. В таблице 1 показаны результаты распознавания при уровне шума от 31% до 50% для матриц, представляющих тестовые изображения с разрешением 20×20 , 35×35 , 50×50 и 100×100 пикселей. В таблице 2 представлены результаты расчётов для значений $Q = 45\%$ и $n = 50$. При уровне шума до 30% процент распознавания составляет не менее 99.9% . Графики на рис. 2 иллюстрируют рост процента распознавания при росте размерности распознаваемых матриц.

% шума	31	32	33	34	35	36	37	38	39	40
$n = 20$	99.57	99.6	99.37	99.19	99	98.56	98	97.27	95.97	94.96
$n = 35$	99.9	99.97	99.97	99.84	99.83	99.71	99.47	99.36	99.02	98.31
$n = 50$	100	100	99.99	99.98	99.97	99.99	99.93	99.89	99.82	99.74
$n = 100$	100	100	100	100	100	100	100	100	100	100
% шума	41	42	43	44	45	46	47	48	49	50
$n = 20$	93.32	90.67	88.07	84.1	79.87	75.28	69.16	63.06	57.02	49.72
$n = 35$	97.52	96.55	94.46	91	87.22	82.18	76.16	67.86	60.13	49.41
$n = 50$	99.51	99.27	98.36	97.29	95.23	90.6	83.99	74.61	62.69	49.86
$n = 100$	100	99.98	99.97	99.86	99.68	99.08	97.23	99.08	75.09	50.68

Т а б л и ц а 1. Процент распознавания тестовых изображений при шуме от 31% до 50%

45%	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	—

Т а б л и ц а 2. Результаты распознавания тестовых изображений при уровне шума 45%, $n = 50$

$S, \%$	0	5.4	7.4	16.2	23.5
$P, \%$	100	99.93	99.79	99.72	99.81
Δ	10	25	50	75	100

Т а б л и ц а 3. Процент задач S при уровне шума 44%, $c_1 = 110$, $c_2 = 120$.

$S, \%$	22.8	37.5	47.3	46.4	46.4
$P, \%$	99.71	99.6	99.8	99.72	99.82
Δ	10	25	50	75	100

Т а б л и ц а 4. Процент задач S при уровне шума 44%, $c_1 = 119$, $c_2 = 120$.

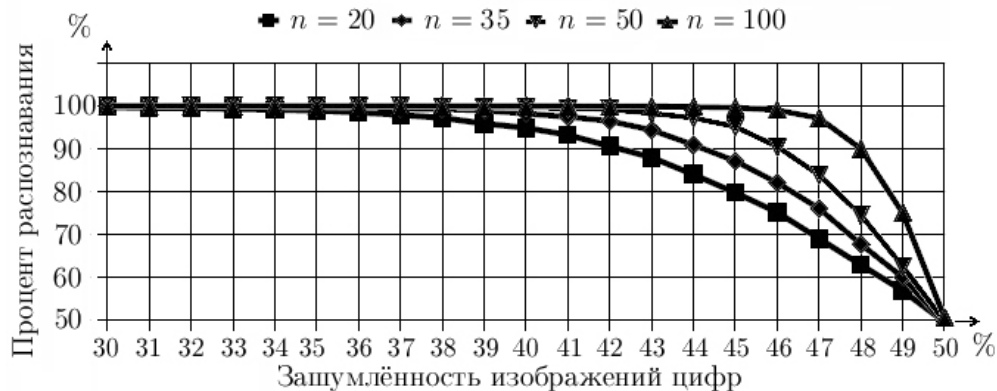


Рис. 2. Повышение процента распознавания тестовых изображений при росте разрешения

Как показывает вычислительный эксперимент, предложенная эвристика и алгоритм, основанный на её расчёте, 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 $Q = 45\%$ [8]. Для алгоритмов распознавания с помощью нейронных сетей процент распознавания не превышает 90% для уровня шума Q от 46% и выше [1]. Как видно из приведённых результатов, при тех же значениях уровня шума предложенный подход даёт тот же или больший процент распознавания при надлежащем разрешении изображений.

Также было проведено сравнение эффективности распознавания с помощью предложенной эвристики с эффективностью распознавания при помощи другой эвристики, также используемой при распознавании растровых изображений, представляющей собой минимизацию евклидова расстояния $\rho(\mathcal{A}, \mathcal{A}^{(k)})$ между матрицей \mathcal{A} , сопоставляемой распознаваемому изображению, и матрицами $\mathcal{A}^{(k)}$, сопоставляемым эталонным изображениям:

$$\rho(\mathcal{A}, \mathcal{A}^{(k)}) = \sqrt{\sum_{i=1}^n \sum_{j=1}^n (a_{ij} - a_{ij}^{(k)})^2}. \quad (15)$$

Результатом распознавания является номер эталонного изображения p , дающий минимум (15). Минимизация расстояния ρ между эталонной матрицей $\mathcal{A}^{(k)}$ и распознаваемой \mathcal{A} часто позволяет производить успешное распознавание при решении задачи в рассматриваемой нами постановке. Вместе с тем, возможны ситуации, когда для пары матриц $\mathcal{A}^{(1)}$ и $\mathcal{A}^{(2)}$ и распознаваемой матрицы \mathcal{A} будем иметь

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

тогда как для большего количества позиций ij в этих матрицах будем иметь

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

и \mathcal{A} в действительности получена в результате зашумления матрицы $\mathcal{A}^{(2)}$.

Был проведён эксперимент, который показал, что с ростом радиуса Δ интервалов, в которых происходят изменения элементов эталонных матриц, если $|c_1 - c_2| < \Delta$, растёт доля задач, в которых процент распознавания с помощью предложенной эвристики оказывается большим, чем при распознавании с помощью евклидова расстояния.

Пусть S — процент задач от общего числа решаемых задач распознавания, где предложенная эвристика даёт верное распознавание, тогда как минимизация евклидова расстояния правильного распознавания произвести не позволяет. Для того же набора эталонных изображений, но 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 , представленные в таблице 3. При $c_1 = 119$ и $c_2 = 120$ рост S показан в таблице 4. Процент задач, где ситуация оказывалась обратной при уровне шума Q от 30 до 50 процентов, не превысил 0.01% от числа всех решённых задач распознавания для данного значения Q .

Подобная тенденция к росту доли задач S при росте Δ сохраняется и при других значениях c_1 и c_2 , если значение $|c_1 - c_2|$ достаточно мало относительно значения Δ .

5. Выводы

Предложен алгоритм распознавания числовых матриц. Используемая алгоритмом эвристика состоит в минимизации лебеговой меры внешней оценки объединённого множества решений интервальной системы линейных алгебраических уравнений, сопоставляемой эталонной и распознаваемой матрице. Алгоритм не требует обучения и имеет квадратичную относительно размерности входных матриц вычислительную сложность.

Проведённые вычислительные эксперименты свидетельствуют о высокой эффективности предложенного подхода в приложении к задаче распознавания растровых изображений и о росте процента распознавания с ростом разрешения распознаваемых изображений.

Автор благодарен С.П. Шарому за внимание к работе и высказанные им замечания.

Список литературы

- [1] DEMUTH H., BEALE M. Neural Network Toolbox for Use with MATLAB. User's Guide. Version 4 // [Электронный ресурс] URL: <http://cs.mipt.ru/docs/comp/eng/develop/software/matlab/nnet/main.pdf> (дата обращения 23.09.2011)
- [2] КИРНОС Э.А., ПЫТЬЕВ Ю.П. О параметрических алгоритмах распознавания // Вестник Московского Университета. Серия 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 Recognition and Image Analysis. 2002. Vol. 12. № 1. P. 19–24.
- [4] ШАРЫЙ С.П. Конечномерный интервальный анализ. [Электронный ресурс] URL: <http://www.nsc.ru/interval/Library/InteBooks/SharyBook.pdf> (дата обращения 23.09.2011)
- [5] OETTL W. On the solution set of a linear system with inaccurate coefficients // SIAM Journal on Numerical Analysis. 1965. Vol. 2, № 1. P. 115–118.
- [6] NEUMAIER A. Interval Methods for Systems of Equations. Cambridge: Cambridge University Press, 1990.

- [7] AHLBERG J.H., NILSON E.N. Convergence properties of the spline fit // J. SIAM. 1963. V. 11, № 1. P. 95–104.
- [8] КИРНОС Э.А. Сравнительный анализ морфологических методов интерпретации изображений: Автореф. дис. ... канд. физ.-мат. наук, (05.13.18 — математическое моделирование, численные методы и комплексы программ). Москва, 2004. РГБ ОД, 61:05-1/118.

ИНТЕРВАЛЬНЫЙ ПОДХОД К РЕШЕНИЮ ЗАДАЧИ РАСПОЗНАВАНИЯ ЧИСЛОВЫХ МАТРИЦ

А.В. Пролубников

Предлагается подход к решению задачи распознавания числовых матриц, использующий интервальные методы. Рассматриваются результаты вычислительного эксперимента. В качестве приложения предложенного подхода рассматривается задача распознавания растровых изображений.

Ключевые слова: распознавание образов, интервальный анализ.

AN INTERVAL APPROACH TO THE MATRIX RECOGNITION PROBLEM

A.V. Prolubnikov

We introduce an approach for matrix recognition, which uses methods of interval analysis. Application of the presented approach for recognition of raster image is considered. Test results for raster images of digits are presented.

Key words: pattern recognition, interval analysis.