Иерархический алгоритм поиска приближенного ближайшего соседа в пространстве пирамидальных представлений изображений

М.М. Ланге[a], А.М. Ланге[a]

[a] *Федеральный исследовательский центр "Информатика и управление" РАН, 119333, ул. Вавилова, 40, Москва, Россия*

Аннотация

Предлагается алгоритм поиска приближенного ближайшего соседа к предъявляемому изображению в пространстве пирамидальных представлений с нарастающей размерностью. Алгоритм обеспечивает логарифмический порядок вычислительной сложности по размерности исходных изображений, но сохраняет линейный порядок сложности по мощности набора данных. Приводятся эмпирические распределения погрешностей поиска относительно ближайшего соседа и численные оценки вычислительной сложности алгоритма для поиска изображений рукописных цифр из базы данных MNIST и для координатной привязки зашумленных изображений к аэрокосмической карте местности из сетевого сервиса Google Maps. Возможно понижение линейного порядка сложности поиска по мощности набора данных за счет предварительного отбора изображений на решающем дереве в пространстве представлений малой размерности.

Ключевые слова: пирамидальное представление; иерархический поиск; приближенный ближайший сосед; погрешность поиска; вычислительная сложность

1. Введение

Существует широкий класс прикладных задач, связанных с поиском в большом наборе данных объекта, близкого по заданной мере к предъявляемому объекту. К таким задачам относится поиск в наборах изображений образцов, сходных с предъявляемыми изображениями, координатная привязка наблюдаемого изображения к цифровой карте местности и другие. Перечисленные примеры относятся к проблеме извлечения изображений (Image Retrieval) [1]. При больших объемах данных решающий алгоритм должен удовлетворять заданным требованиям к допустимой погрешности (точности) поиска и вычислительной сложности (быстродействию). Как правило, эти характеристики находятся в обратной зависимости: с увеличением быстродействия уменьшается точность (растет погрешность) и наоборот. Поэтому необходимо обеспечить баланс этих требований путем варьирования параметрами решающего алгоритма.

В качестве решающих алгоритмов могут быть использованы алгоритмы поиска приближенного ближайшего соседа [2-6] или их модификации. При фиксированной размерности $d \geq 1$ векторного пространства известные алгоритмы с гарантированной точностью, задаваемой допустимой погрешностью $\varepsilon > 0$, реализуют поиск в наборе из n векторов представителя на расстоянии $D \leq (1+\varepsilon)D_{\min}$ от предъявляемого вектора, где $D_{\min} > 0$ – расстояние до ближайшего соседа. Такие алгоритмы используют древовидные структуры данных, которые при фиксированных d и ε позволяют уменьшить порядок роста вычислительной сложности по n по сравнению со сложностью переборного алгоритма. В частности, BBD-алгоритм, использующий решающее Balance Box Decision дерево [5], имеет сложность $O\left(d\lceil 1+6d/\varepsilon\rceil^{d}\log n\right)$, а сложность LSH-алгоритма на основе локального хеширования Locality Sensitive Hashing [6] составляет $O\left(dn^{1/(1+\varepsilon)}\right)$. Для сравнения переборный алгоритм поиска ближайшего соседа ($\varepsilon = 0$) имеет сложность $\Theta(dn)$. Характер зависимости вычислительной сложности указанных алгоритмов от размерности d и допустимой погрешности ε ограничивает их применение для поиска изображений размера $N * N$ из-за чрезмерно высокой размерности $d = N^2$ при $N \geq 100$.

В настоящей работе рассматривается альтернативный алгоритм для быстрого поиска в заданном наборе изображений мощности n приближенного ближайшего соседа к предъявляемому изображению. Предлагаемый алгоритм является модификацией алгоритма, рассмотренного в [7], которая использует пирамидальное представление изображений с многоуровневым разрешением [8-10]. Алгоритм базируется на параметрической стратегии экспоненциального сужения зоны поиска, аналогичной использованной в работе [11]. Такая стратегия поиска обеспечивает вычислительную сложность $O\left(n\log N\right)$, но не гарантирует точности приближенного решения.

Эффективность предложенного алгоритма исследована в терминах эмпирического распределения погрешностей поиска и вычислительного выигрыша относительно алгоритма полного перебора. При различных значениях параметра алгоритма, указанные характеристики получены для поиска изображений рукописных цифр из базы данных MNIST [12] и для координатной привязки зашумленных изображений к цифровой карте участка земной поверхности, взятой на интернет-сервисе Google Maps [13].

2. Формализация задачи

Рассматривается множество изображений $\mathbf{X}$ размера $N*N$, элементы которых принадлежат алфавиту $A=\{0,1,...,q-1\}$ ($q\geq 2$). Предполагается, что каждое изображение $\mathbf{x}\in\mathbf{X}$ имеет, по крайней мере, один ненулевой элемент, а размер изображения является целочисленной степенью числа 2, так что $N=2^L$, где $L>>1$. В общем случае множество $\mathbf{X}$ содержит набор изображений $\hat{\mathbf{X}}$, в котором необходимо найти изображение $\hat{\mathbf{x}}\in\hat{\mathbf{X}}$, ближайшее или достаточно близкое по заданной мере к заданному изображению $\mathbf{x}\in\mathbf{X}$.

Любое изображение $\mathbf{x}\in\mathbf{X}$ с указанными размерами допускает пирамидальное представление

$$\mathbf{x}_L=(x_0,...,x_l,\dots,x_L) \tag{1}$$

порядка L, которое содержит последовательность описаний изображения $\mathbf{x}$ с уровнями разрешения $l=0,...,L$ [8].

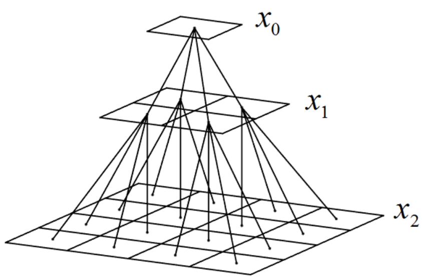


Рис.1. Пирамидальное представление $\mathbf{x}_2=(x_0,x_1,x_2)$ порядка $L=2$.

Пример пирамидального представления порядка $L=2$ дан на рис.1. Описание l-го уровня x_l в (1) является изображением размера 2^l*2^l, получаемым из описания x_{l+1} путем усреднений по непересекающимся группам из четырех смежных элементов. Ненулевое значение элемента в вершине пирамиды x_0 позволяет сформировать нормализованное представление

$$\mathbf{y}_L=(y_1,...,y_l,\dots,y_L) \tag{2}$$

путем деления всех элементов пирамиды (1) на значение элемента вершины x_0. Нормализация элементов уменьшает зависимость представления (2) от средней яркости изображения по сравнению с представлением (1).

Пусть $z(k_{l1},k_{l2})>0$ – значение элемента с индексами $(k_{l1},k_{l2})=1,...,2^l$ в нормализованном описании $y_l\in\mathbf{y}_L$. Для любой пары изображений $\mathbf{x}\in\mathbf{X},\ \hat{\mathbf{x}}\in\mathbf{X}$, имеющих нормализованные описания $\mathbf{y}_L,\ \hat{\mathbf{y}}_L$, вводится мера их различия l-го порядка

$$D_l(\mathbf{x},\hat{\mathbf{x}})=\frac{1}{(2^l\times 2^l)}\sum_{k_{l1}=1}^{2^l}\sum_{k_{l2}=1}^{2^l}|\,z(k_{l1},k_{l2})-\hat{z}(k_{l1},k_{l2})\,|,\ l=1,...,L, \tag{3}$$

где $z(k_{l1},k_{l2})$, $\hat{z}(k_{l1},k_{l2})$ – элементы описаний $y_l\in\mathbf{y}_L$, $\hat{y}_l\in\hat{\mathbf{y}}_L$ соответственно для изображений $\mathbf{x}$ и $\hat{\mathbf{x}}$. Суммирование мер $D_t(\mathbf{x},\hat{\mathbf{x}})$ вида (3) порядка $t=1,...,l\ (l\leq L)$ с весами $w_t=\frac{1}{2}\log_2(2^t*2^t)=t$ порождает взвешенную меру различия порядка l:

$$\tilde{D}_l(\mathbf{x},\hat{\mathbf{x}})=\sum_{t=1}^{l}w_tD_t(\mathbf{x},\hat{\mathbf{x}}),\ 1\leq l\leq L. \tag{4}$$

Точный или приближенный поиск ближайшего соседа для предъявляемого изображения $\mathbf{x}\in\mathbf{X}$ выполняется на подмножестве $\hat{\mathbf{X}}\subset\mathbf{X}$, содержащем n изображений. Алгоритм поиска принимает решение $\hat{\mathbf{x}}^*\in\hat{\mathbf{X}}$ по мере (4) наибольшего порядка L на наборе изображений $\hat{\mathbf{X}}^*\subseteq\hat{\mathbf{X}}$ мощности $n^*\leq n$ в соответствии с решающим правилом

$$\hat{\mathbf{x}}^* = \arg \min_{\hat{\mathbf{x}} \in \hat{\mathbf{X}}^*} \tilde{D}_L(\mathbf{x}, \hat{\mathbf{x}}). \tag{5}$$

Набор $\hat{\mathbf{X}}^*$ в (5) определяется используемой стратегией направленного (иерархического) поиска, которая в случае $\hat{\mathbf{X}}^* = \hat{\mathbf{X}}$ обеспечивает точное решение $\hat{\mathbf{x}}^*_{\mathrm{NN}}$ (Nearest Neighbor), совпадающее с ближайшим соседом, а в случае $\hat{\mathbf{X}}^* \subset \hat{\mathbf{X}}$ – приближенное решение $\hat{\mathbf{x}}^*_{\mathrm{AN}}$ (Approximate Nearest), которое может отличаться или совпадать с ближайшим соседом.

Погрешность поиска по правилу (5) приближенного ближайшего изображения в заданном наборе определяется различием значений $\tilde{D}_L(\mathbf{x}, \hat{\mathbf{x}}^*_{\mathrm{AN}})$ и $\tilde{D}_L(\mathbf{x}, \hat{\mathbf{x}}^*_{\mathrm{NN}})$, и в случае $\tilde{D}_L(\mathbf{x}, \hat{\mathbf{x}}^*_{\mathrm{NN}}) > 0$ равна

$$\varepsilon^{\mathrm{s}}_{n^*}(\hat{\mathbf{x}}^*_{\mathrm{AN}}, \hat{\mathbf{x}}^*_{\mathrm{NN}}) = \frac{\tilde{D}_L(\mathbf{x}, \hat{\mathbf{x}}^*_{\mathrm{AN}}) - \tilde{D}_L(\mathbf{x}, \hat{\mathbf{x}}^*_{\mathrm{NN}})}{\tilde{D}_L(\mathbf{x}, \hat{\mathbf{x}}^*_{\mathrm{NN}})}. \tag{6}$$

В задаче координатной привязки погрешность определяется отклонением координат (a^*, b^*) найденного по правилу (5) изображения $\hat{\mathbf{x}}^*_{\mathrm{AN}}$ от координат (a, b) предъявляемого изображения $\mathbf{x}$ и в случае изображений размера $N * N$ равна

$$\varepsilon^{\mathrm{g}}_{n^*}(\hat{\mathbf{x}}^*_{\mathrm{AN}}, \mathbf{x}) = \frac{1}{2}\left(\frac{|a - a^*|}{N} + \frac{|b - b^*|}{N}\right). \tag{7}$$

Качество рассматриваемого алгоритма исследуется в терминах распределений

$$\Pr\left\{\varepsilon^{\mathrm{s}}_{n^*}(\hat{\mathbf{x}}^*_{\mathrm{AN}}, \hat{\mathbf{x}}^*_{\mathrm{NN}}) \le \varepsilon\right\} \tag{8}$$

$$\Pr\left\{\varepsilon^{\mathrm{g}}_{n^*}(\hat{\mathbf{x}}^*_{\mathrm{AN}}, \mathbf{x}) \le \varepsilon\right\} \tag{9}$$

погрешностей (6) и (7) с параметром $n^* : 1 \le n^* \le n$, где $\varepsilon \ge 0$ – допустимая погрешность, принимающая значения с некоторым шагом. При фиксированном значении ε вероятности (8) и (9) соответствуют надежности выполнения точности поиска и привязки.

В разделе 3 дается описание алгоритма поиска и приводятся оценки его вычислительной сложности при больших значениях n и N, связанных соотношением $N^2 \ge \log_q n$. В разделе 4 приводится параметрическое семейство эмпирических распределений (8) с параметром n / n^* и численные оценки сложности алгоритма как функции от n / n^*, полученные на множестве изображений рукописных цифр. Аналогичные параметрические семейства эмпирических распределений (9) и оценки вычислительной сложности, полученные для координатной привязки зашумленных изображений к карте участка земной поверхности, приводятся в разделе 5.

3. Алгоритм поиска

Предполагается, что изображения из набора $\hat{\mathbf{X}}$, на котором производится поиск, заданы нормализованными пирамидальными представлениями вида (2) и образуют многоуровневую сеть

$$\hat{\mathbf{Y}}_1, ..., \hat{\mathbf{Y}}_l, ..., \hat{\mathbf{Y}}_L \tag{10}$$

в которой $\hat{\mathbf{Y}}_l$ – подмножество представлений всех изображений из $\hat{\mathbf{X}}$, заданных l уровнями нормализованных пирамид. Алгоритм поиска решения (5) использует стратегию последовательного сужения зоны поиска на уровнях $l = 1, ..., L$ сети (10). Согласно этой стратегии, число анализируемых изображений на l-м уровне определяется экспоненциальной функцией

$$n_l = \left\lfloor n 4^{-\alpha(l-1)} \right\rfloor, \; l = 1, ..., L \tag{11}$$

с коэффициентом $\alpha = (L-1)^{-1} \log_4 (n / n^*)$, где $n^* = 1, 2, ..., n$ - мощность набора $\hat{\mathbf{X}}^* \subseteq \hat{\mathbf{X}}$, на котором принимается решение (5) по представлениям $\hat{\mathbf{Y}}_L$ последнего уровня сети (10). Стратегия отбора изображений, анализируемых на последовательных уровнях сети (10) схематически показана на рис. 2.

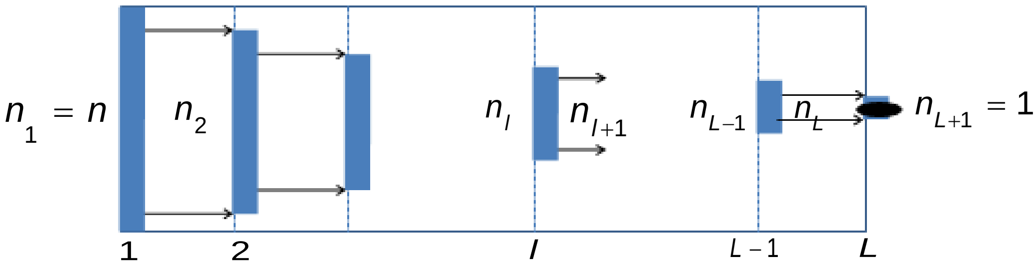


Рис.2. Схема отбора изображений в сети пирамидальных представлений.

Алгоритм поиска. Для предъявляемого изображения $\mathbf{x} \in \mathbf{X}$, на последовательных уровнях $l = 1, ..., L-1$ сети (10) вычисляются значения меры различия $\tilde{D}_l\,(\mathbf{x}, \hat{\mathbf{x}})$ вида (4) для n_l изображений из набора $\hat{\mathbf{X}}$ и среди них отбираются n_{l+1} изображений с наименьшими значениями $\tilde{D}_l\,(\mathbf{x}, \hat{\mathbf{x}})$; на уровне $l = L$ среди $n_L = n^*$ изображений отбирается ближайшее с наименьшим значением $\tilde{D}_L\,(\mathbf{x}, \hat{\mathbf{x}})$, которое дает решение (5).

В случае $1 \le n^* < n$ параметр α>0 в (11) обеспечивает экспоненциальное сужение зоны иерархического поиска, которое приводит к нахождению приближенного ближайшего соседа на наборе изображений $\hat{\mathbf{X}}^* \subset \hat{\mathbf{X}}$ мощности n^*; в случае $n^* = n$ параметр $\alpha = 0$ приводит к переборному поиску ближайшего соседа на наборе $\hat{\mathbf{X}}^* \equiv \hat{\mathbf{X}}$ мощности n. В обоих случаях вычисление меры различия изображений производится с использованием рекурсии

$$\tilde{D}_l\,(\mathbf{x}, \hat{\mathbf{x}}) = \tilde{D}_{l-1}(\mathbf{x}, \hat{\mathbf{x}}) + w_l D_l(\mathbf{x}, \hat{\mathbf{x}}),\ l = 1, ..., L \qquad (12)$$

при начальном условии $\tilde{D}_0\,(\mathbf{x}, \hat{\mathbf{x}}) = 0$.

Вычислительная сложность алгоритма определяется числом элементарных операций, затрачиваемых на вычисление меры на всех уровнях сети (10), и на сортировку значений меры на последовательных уровнях для отбора ближайших изображений согласно (11), включая отбор решения на последнем уровне. В работе [7] получена асимптотическая оценка вычислительной сложности сформулированного иерархического алгоритма поиска при больших значениях мощности n набора изображений $\hat{\mathbf{X}}$, размере изображения N, удовлетворяющем условию $N^2 \ge \log_q n$ (q –размер алфавита), и соотношении $n / n^* \ge N^2 / 4$, обеспечивающем коэффициент сужения зоны поиска $\alpha \ge 1$. Асимптотика сложности иерархического алгоритма с указанными параметрами имеет вид $C_{n^* \le 4n/N^2} = O\left(n \log N\right)$. Для сравнения переборный алгоритм имеет вычислительную сложность $C_{n^* = n} = \Omega\left(nN^2\right)$. Из приведенных оценок следует, что доля сложности иерархического поиска приближенного решения относительно сложности переборного поиска ближайшего соседа убывает с увеличением размера изображения как $O\left(N^{-2} \log N\right)$.

Приведенная оценка вычислительной сложности иерархического алгоритма имеет линейный порядок по мощности набора изображений, в котором производится поиск. Понижение порядка сложности по n может быть достигнуто путем замены параметра n в (11) величиной $n_0 < n$. В частности, используя BBD дерево [5] для набора изображений, заданных верхними уровнями пирамидальных представлений, можно в пространстве малой размерности $d = 4$ отобрать $n_0 = \lfloor \log n \rfloor$ изображений с заданной допустимой погрешностью $\varepsilon \ge 1$. Поскольку вычислительная сложность такого отбора равна $O\left(d^d \log n\right)$, общая вычислительная сложность алгоритма составит $O\left((\log n) \log N\right)$.

Численные оценки сложности алгоритма получены для случая $n_0 = n$ путем вычисления суммарного числа элементарных операций, затрачиваемых на вычисление меры и сортировку вставками значений меры со сложностью $m \log_2 m$ на наборе из m элементов [14]. Поэтому при фиксированных n и $N = 2^L$ вычислительная сложность алгоритма равна

$$C_{n^*} = C_{n^*}^{\mathrm{msr}} + C_{n^*}^{\mathrm{srt}}, \qquad (13)$$

где

$$C_{n^*}^{\mathrm{msr}} = \sum_{l=1}^{L} n_l 4^l \le n \sum_{l=1}^{L} 4^l \left(n / n^*\right)^{-\frac{l-1}{L-1}}, \qquad (14)$$

$$\begin{aligned} C_{n^*}^{\mathrm{srt}} &= (n-1)[n^*=n] + \left((n^*-1) + \sum_{l=1}^{L-1} n_l \log_2 n_l\right)[n^*<n] \\ &\le (n-1)[n^*=n] + n\log_2 n\left(\sum_{l=1}^{L-1}(n/n^*)^{-\frac{l-1}{L-1}} - \sum_{l=1}^{L-1}\frac{l-1}{L-1}(n/n^*)^{-\frac{l-1}{L-1}}\right)[n^*<n] \\ &+ (n^*-1)\left(1 + n\log_2 e\sum_{l=1}^{L-1}\frac{l-1}{L-1}(n/n^*)^{-\frac{l-1}{L-1}}\right)[n^*<n] \end{aligned} \qquad (15)$$

соответственно затраты на вычисление меры и на сортировку, $[f]$ – индикатор f. В случае $n^* = n$ формулы (13), (14) и (15) дают оценки вычислительной сложности переборного поиска точного решения, а в случае $n^* < n$ – оценки сложности иерархического поиска приближенного решения.

Формулы (12), (13) и (14) использованы для получения численных оценок относительной сложности $C_{n^* \le n} / C_{n^* = n}$ при значениях $n/n^* = 2^k, k = 0,1,...$ и параметрах n и $N = 2^L$, с которыми проведены эксперименты по поиску рукописных цифр (раздел 4) и по координатной привязке изображений к карте местности (раздел 5). Величина, обратная относительной сложности, соответствует вычислительному выигрышу иерархического алгоритма по сравнению с алгоритмом перебора.

4. Оценки качества и сложности поиска рукописных цифр

Экспериментальные характеристики качества алгоритма поиска получены на наборе полутоновых изображений рукописных цифр из базы данных MNIST [12]. Вычислительный эксперимент выполнен с помощью кода, написанного на языке MATLAB [15]. Примеры изображений рукописных цифр даны на рис. 3.

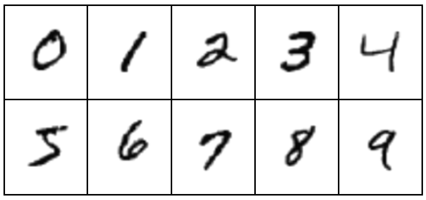


Рис.3. Примеры рукописных цифр из базы данных MNIST.

Цифры на изображениях нормированы по размеру и центрированы в поле изображения. Параметры изображений: $N = 32,\ q = 256$; число уровней представления изображений $L = \log_2 N = 5$; мощность набора данных $\hat{\mathbf{X}}$ равна $n = 60000$. Параметры $N,\ q,\ n$ удовлетворяют необходимому условию $N^2 > \log_q n$. Набору $\hat{\mathbf{X}}$ предъявлялось 10000 изображений, не входящих в набор данных $\hat{\mathbf{X}}$, так что $\tilde{D}_L(\mathbf{x}, \hat{\mathbf{x}}^*_{\mathrm{AN}}) \ge \tilde{D}_L(\mathbf{x}, \hat{\mathbf{x}}^*_{\mathrm{NN}}) > 0$. Для каждого $\mathbf{x}$ и соответствующей пары $\hat{\mathbf{x}}^*_{\mathrm{AN}},\ \hat{\mathbf{x}}^*_{\mathrm{NN}}$ вычислялась погрешность поиска вида (6) и на 10000 предъявляемых изображениях строилось семейство эмпирических распределений вида (8) с параметром $n/n^* = 2^k, k = 0,5,...,10$ при значениях допустимой погрешности $0 \le \varepsilon \le 0,1$. Графики семейства распределений даны на рис. 4. Численные оценки сложности алгоритма поиска представлены значениями $C_{n^* \ge 1} / C_{n^* = n}$ в точках $n/n^* = 2^k, k = 0,1,...10$ ($N = 32, n = 60000$). Полученные оценки сложности алгоритма поиска даны в таблице 1.

Таблица 1. Оценки относительной сложности иерархического алгоритма поиска изображений рукописных цифр с параметрами $N = 32, q = 256, n = 60000$

$\log_2(n/n^*)$	0	1	2	3	4	5	6	7	8	9	10
$C^{\mathrm{msr}}_{n^* \le n} / C_{n^* = n}$	0.9993	0.5325	0.2885	0.1597	0.0908	0.0535	0.0329	0.0214	0.0146	0.0107	0.0082
$C^{\mathrm{srt}}_{n^* \le n} / C_{n^* = n}$	0.0007	0.0356	0.0285	0.0237	0.0205	0.0182	0.0166	0.0154	0.0146	0.0139	0.0134
$C_{n^* \le n} / C_{n^* = n}$	1.0000	0.5681	0.3170	0.1834	0.1113	0.0717	0.0496	0.0368	0.0292	0.0246	0.0216

Из приведенных распределений и оценок сложности следует, что при значениях $8 \le \log_2 n/n^* \le 10$ иерархический алгоритм реализует нулевую погрешность поиска ближайшего соседа ($\varepsilon^{\mathrm{s}}_{n^*} = 0$) с вероятностью 0,980–0,997 и обеспечивает вычислительный выигрыш в 34–46 раз по сравнению с переборным алгоритмом. С ростом допустимой погрешности $\varepsilon > 0$, вероятность нахождения ближайшего соседа с точностью $\varepsilon^{s}_{n^*} \le \varepsilon$ увеличивается при тех же значениях вычислительного выигрыша.

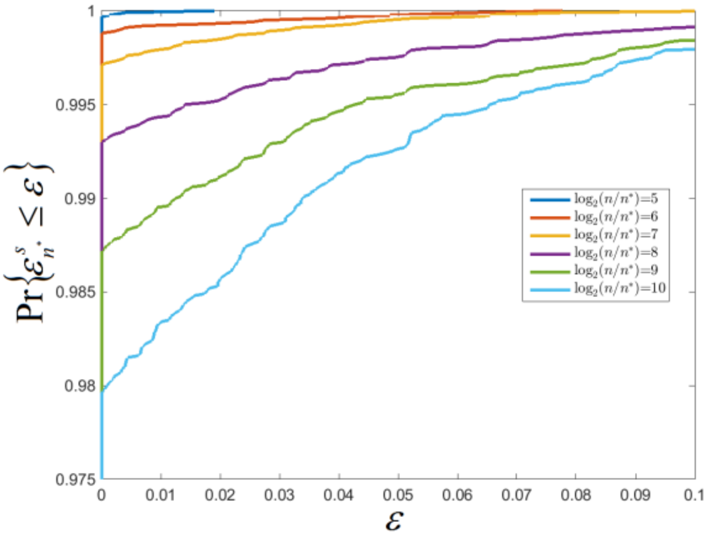
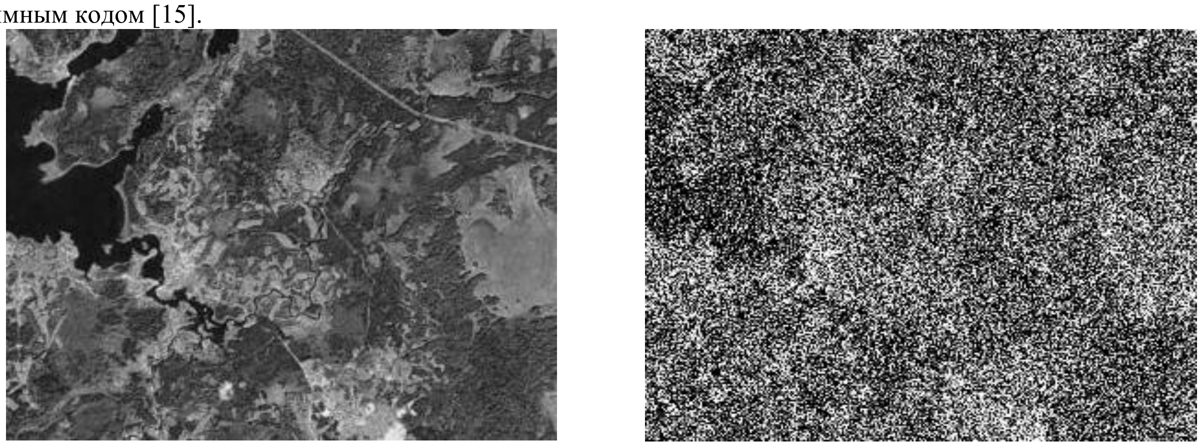


Рис. 4. Эмпирические распределения погрешностей поиска рукописных цифр на наборе изображений с параметрами $N = 32,\ q = 256,\ n = 60000$.

5. Оценки качества и сложности координатной привязки изображений

Экспериментальные оценки эффективности алгоритма поиска для координатной привязки к карте местности получены на цифровом аэрокосмическом снимке участка земной поверхности [13]. Обработка данных выполнена программным кодом [15].



(а) (б)

Рис.5. Аэрокосмический снимок участка земной поверхности (а) без шума; (б) с гауссовым шумом с нулевым средним и дисперсией $\sigma^2 = 0,25$.

Размеры аэрокосмического снимка 300*236 (в пикселях), количество уровней яркости $q = 256$. Использованный снимок и его зашумленная версия даны на рис. 5. Привязка в координатах снимка проводилась по изображению размера 64*64 ($N = 64$). Мощность набора данных $\hat{\mathbf{X}}$ определялась количеством всевозможных изображений указанного размера, взятых на снимке с шагом в два пикселя, и равна $n = 10353$. Параметры N, q, n удовлетворяют необходимому условию $N^2 > \log_q n$. Набору $\hat{\mathbf{X}}$ предъявлялись поочередно все n изображений с аддитивным гауссовым шумом с дисперсией $0 \le \sigma^2 \le 0,25$. Для любого предъявляемого изображения $\mathbf{x} \in \hat{\mathbf{X}}$ с координатами (a,b) вычислялись координаты (a^*,b^*) найденного алгоритмом изображения $\hat{\mathbf{x}}^*_{\mathrm{AN}}$ и погрешность привязки вида (7). При различных значениях σ^2 строились семейства эмпирических распределений (9) с параметром n/n^* в диапазоне значений погрешности $0 \le \varepsilon \le 2,5$. На рис. 6 даны семейства распределений, полученные для четырех значений дисперсии шума при $n/n^* = 2^k, k = 0,3,6,9$. В таблице 2 приведены численные оценки относительной сложности $C_{n^* \le n} / C_{n^* = n}$ как функции от n/n^*. Последний столбец таблицы содержит оценки с параметром $n^* = 1$, которые дают наибольший вычислительный выигрыш (более 260) иерархического алгоритма относительно полного перебора.

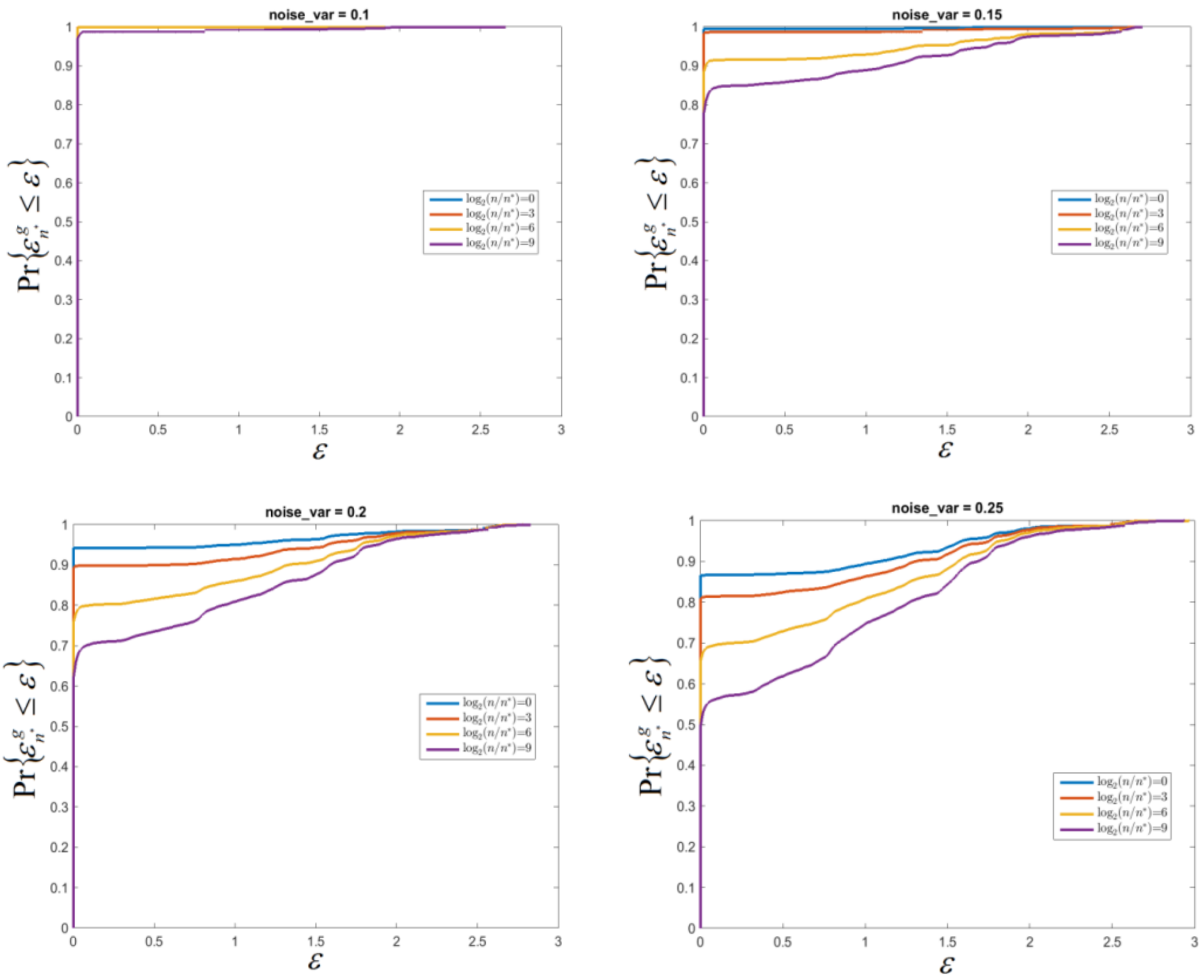


Рис. 6. Эмпирические распределения погрешностей координатной привязки при различных дисперсиях гауссова шума и параметрах $N = 64, q = 256, n = 10353$.

Таблица 2. Оценки относительной сложности иерархического алгоритма привязки изображений к карте местности с параметрами $N = 64,\ q = 256,\ n = 10353$

$\log_2(n/n^*)$	0	3	4	5	6	7	8	9	10	13.3378
$C^{\text{msr}}_{n^* \le n} / C_{n^* = n}$	0.9998	0.1504	0.0824	0.0462	0.0265	0.0158	0.0098	0.0063	0.0028	0.0018
$C^{\text{srt}}_{n^* \le n} / C_{n^* = n}$	0.0002	0.0059	0.0050	0.0044	0.0039	0.0036	0.0034	0.0032	0.0034	0.0027
$C_{n^* \le n} / C_{n^* = n}$	1.0000	0.1563	0.0874	0.0506	0.0304	0.0194	0.0132	0.0095	0.0062	0.0046

Экспериментально установлено, что при отсутствии шума ($\sigma^2 = 0$), иерархический алгоритм с параметром $n^* \ge 1$ с вероятностью единица дает нулевую погрешность привязки $\varepsilon^{g}_{n^*} = 0$ и, следовательно, по качеству эквивалентен алгоритму перебора. В случае шума с дисперсией $\sigma^2 \le 0{,}10$ погрешность $\varepsilon^{g}_{n^*} = 0$ реализуется практически с единичной вероятностью при более, чем 100-кратном вычислительном выигрыше иерархического алгоритма относительно переборного алгоритма. Семейства распределений, полученные при значениях $\sigma^2 = 0{,}15; 0{,}20; 0{,}25$, демонстрируют динамику соотношения показателей качества привязки и быстродействия иерархического алгоритма. При любой фиксированной дисперсии шума σ^2 и заданной допустимой погрешности привязки ε, вероятность реализации точности привязки $\varepsilon^{g}_{n^*} \le \varepsilon$ уменьшается с ростом вычислительного выигрыша алгоритма (увеличением n/n^*). При фиксированных значениях ε и n/n^* вероятность реализации точности $\varepsilon^{g}_{n^*} \le \varepsilon$ уменьшается с увеличением дисперсии шума. Необходимо отметить, что представление цифровой карты местности набором всевозможных изображений, взятых с шагом в один пиксель, должно привести к увеличению вероятности требуемой точности при фиксированных значениях ε, n/n^* и σ^2.

6. Заключение

Исследована эффективность иерархического алгоритма поиска в заданном наборе изображений приближенного ближайшего соседа к заданному изображению в терминах эмпирического распределения значений погрешности и вычислительной сложности алгоритма. Рассматриваемый алгоритм предназначен для реализации поиска с заданным быстродействием и надежностью выполнения требования по точности. Эффективность алгоритма продемонстрирована на двух источниках изображений: на наборе изображений рукописных цифр из базы MNIST и на наборе пересекающихся изображений аэрокосмической карты местности, взятой из интернет-сервиса Google Maps. При фиксированной вероятности реализации требуемой точности показана возможность "размена" точности и вычислительной сложности путем варьирования параметром алгоритма. Полученные экспериментальные результаты показали достаточно высокую надежность требуемой точности поиска в наборе изображений рукописных цифр и требуемой точности координатной привязки зашумленного изображения к цифровой карте местности. Понижение порядка роста вычислительной сложности поиска от мощности набора изображений может быть достигнуто на модификации иерархического алгоритма, которая объединяет представление с многоуровневым разрешением и решающее дерево.

Благодарности

Работа поддержана проектами РФФИ №15-07-07516 и № 15-07-09324.

Литература

[1] Datta, R. Image retrieval: Ideas, influences, and trends of the new age / R. Datta, D. Joshi, J. Li, J. Wang // ACM Computing Surveys. – 2008. – Vol. 40, no. 2. – P. 1–60. DOI: 10.1145/1348246.1348248.
[2] Friedman, J. An algorithm for finding best matches in logarithmic expected time / J. Friedman, J. Bentley, R. Finkel // ACM Trans. Math. Software. – 1977. – Vol. 3, no. 3. – P. 209–226. DOI: 10.1145/355744.355745.
[3] Cleary, J. Analysis of an algorithm for finding nearest neighbors in Euclidean space // ACM Trans. Math. Software. – 1979. – Vol. 5, no. 2. – P. 183–192. DOI: 10.1145/355826.355832.
[4] Soleymani, M. An efficient nearest neighbor search method / M. Soleymani, S. Morgera // IEEE Transactions on Communications. – 1987. – Vol. 35, no. 6. – P. 677–679. DOI: 10.1109/TCOM.1987.1096830.
[5] Arya, S. An optimal algorithm for approximate nearest neighbor searching in fixed dimensions / S. Arya, D. Mount, N. Netanyahu [et al.] // J. ACM. – 1998. – Vol. 45, no. 6. – P. 891–923. DOI: 10.1145/293347.293348.
[6] Andoni, A. Near-optimal hashing algorithms for approximate nearest neighbor in high dimensions /A. Andoni, P. Indyk // Commun. ACM. – 2008. – Vol. 51, no. 1. – P. 117–122. DOI: 10.1145/1327452.1327494.
[7] Lange, M.M. Algorithm of approximate search for the nearest digital array in a hierarchical data set / M.M. Lange, S.N. Ganebnykh // Machine Learning and Data Analysis. – 2016. – Vol. 2, no. 1. – P. 6_16. DOI:10.21469/22233792.2.1.01.
[8] Rosenfeld, A. Quadtrees and pyramids: Hierarchical representation of images / A. Rosenfeld // Pictorial Data Analysis. – Springer, 1983. – P. 29–42. DOI: 10.1007/978-3-642-82017-5.
[9] Jackins, C. Quad-trees, oct-trees, and k-trees: A generalized approach to recursive decomposition of Euclidean space / C. Jackins, S. Tanimoto // IEEE Transactions on PAMI. – 1983. – Vol. 5, no. 5. – P. 533–539. DOI:10.1109/TPAMI.1983.4767433.
[10] Samet, H. The quadtree and related hierarchical data structures //ACM Comput. Surv. – 1984. –Vol. 16, no. 2. – P. 187–260. DOI: 10.1145/356924.356930.
[11] Lange, M.M. Multiresolution data representation for fast image gridding / M.M. Lange, N.A. Novikov // Proceedings of the Scientific-Technical Conference on Computing Vision in Control Systems. – ISR RAS, 2012. –P. 242–249.
[12] Thomas Moeslund's gesture recognition database [Electronic resource]. – URL: http://www-prima.inrialpes.fr/FGnet/data/12-MoeslundGesture/database.html (accessed 14.01. 2017).
[13] Network service Google Maps [Electronic resource]. – URL: http://www.maps.google.com. (accessed 14.01.2017).
[14] Cormen, T. Introduction to Algorithms / T. Cormen, C. Leiserson, R. Rivest, C. Stein. – 3rd edition. – The MIT Press, 2009. – ISBN: 9780262033848.
[15] Algorithm of approximate search for the nearest neighbor [Electronic resource]. – URL: http://sourceforge.net/projects/edivis/files/ (accessed 14.01.2017).