Math-Net.Ru

Общероссийский математический портал

С. В. Коваценко, О восстановлении сегментации динамических образов, *Дискрет. матем.*, 1996, том 8, выпуск 4, 143–148

DOI: <https://doi.org/10.4213/dm552>

Использование Общероссийского математического портала Math-Net.Ru подразумевает, что вы прочитали и согласны с пользовательским соглашением
<http://www.mathnet.ru/rus/agreement>

Параметры загрузки:

IP: 52.87.193.239

26 апреля 2023 г., 18:23:53

О восстановлении сегментации динамических образов

© 1996 г. С. В. Коваценок

Предложено решение задачи восстановления сегментации искаженной последовательности дискретных образов за время, полиномиально зависящее от длины сигнала.

1. Введение

В распознавании последовательных образов, например, таких как речь, рукописный текст, печатный текст, главную сложность представляет разделение длинного сигнала на элементарные единицы, такие, как, например, фонемы или буквы. Переборное решение таких задач приводит к алгоритмам, время работы которых полиномиально зависит от длины обрабатываемой последовательности. В настоящей работе предложен полиномиально зависящий от длины последовательности образов алгоритм решения подобной задачи, при этом степень полинома определяется разнообразием элементов последовательности.

Выводы, полученные из решаемой математической задачи, были использованы в прикладной задаче анализа зашумленного графического образа текста, когда требуется разделить графический образ слова на части, соответствующие буквам. На основе данного подхода получен алгоритм решения, в котором время полиномиально зависит от длины слова, а степень полинома равна числу различных значений для ширины буквы и промежутка между буквами.

2. Основные понятия и результаты

Зафиксируем положительные числа $T_1, \dots, T_k, d, L, \delta, T_j \neq T_j$ при $i \neq j$. Рассмотрим последовательность чисел $\tau: t_0 < t_1 < \dots < t_{2m+1}$, $m \in \mathbb{N}$, такую, что $t_0 = 0$, $t_1 - t_0 = T_{n_1}$, $t_{2j+1} - t_{2j} = T_{n_j}$, для некоторых $n_1, n_j \leq k$, $t_{2j} - t_{2j-1} = d$, $j = 1, \dots, m$, $0 \leq t_{2m+1} - L < \delta$, $0 < \delta < \min_{1 \leq i \leq k} T_i$. Множество всех таких последовательностей обозначим через Θ . Назовем сигналом длительности L интегрируемое на $[0, L]$ отображение $\sigma: \mathbb{R} \rightarrow \mathbb{R}^+$, такое, что $\sigma(t) = 0$ при $t \notin [0, L]$, $\sigma(t) \neq 0$ для $t \in [0, L]$.

Предположим, что про сигнал известно, что до добавления к нему шума, интегрируемой на отрезке $[0, L + \delta]$ функции, на отрезках $[t_{2j-1}, t_{2j}]$, $j = 1, \dots, m$, где t_{2j}, t_{2j-1}, m взяты из некоторого набора $\tau_0 \in \Theta$, он тождественно равнялся 0, однако сам этот набор неизвестен. Требуется его восстановить.

Введем функционал $F: \Theta \rightarrow \mathbf{R}^+$

$$F(\tau) = \sum_{i=1}^m \int_{[t_{2i-1}, t_{2i}]} \sigma(x) dx.$$

Будем решать задачу A минимизации функционала $F(\tau)$ по $\tau \in \Theta$. Требуется предъявить алгоритм и оценку его сложности. Далее под сложностью алгоритма будем понимать количество элементарных операций (сложение, вычитание, умножение, деление, сравнение), необходимых для задания алгоритма. Имеет место следующее утверждение.

Теорема 1. *Существует алгоритм решения задачи A сложности, не превосходящей*

$$\left(\frac{L+d}{T_1+d} + 1\right) \dots \left(\frac{L+d}{T_k+d} + 1\right) (C + 4k),$$

где C — константа.

Пусть

$$l = \max_{i=1 \dots k} \left\lceil \frac{L+d}{T_i+d} \right\rceil + 1,$$

содержательно, $l-1$ означает максимально возможное число графических образов букв в графическом образе слова. Теорема 1 означает, что сложность не превосходит $l^k(C + 4k)$, т.е. полиномиально зависит от длины сигнала. Таким образом, алгоритм эффективно работает на длинных сигналах при небольшом разнообразии величин T_i , $i = 1, \dots, k$.

Следующие теоремы дают оценки сложности решения задачи A при фиксированном l и растущем k , и при фиксированном k и растущем l .

Теорема 2. *Пусть $l > 1$, $k > 1$. Тогда существует алгоритм решения задачи A сложности, не превосходящей*

$$\frac{e}{\sqrt{2\pi}} (C + 5k)(k + 1) e^{C_2 \sqrt{k} \ln k},$$

где

$$C_2 = 5\sqrt{l} \left(\frac{1}{2} + \frac{1}{\ln l} \right).$$

Теорема 3. *Пусть $k > 2$, $l > 2k(k + 1)$. Тогда существует алгоритм решения задачи A сложности, не превосходящей*

$$\frac{e}{\sqrt{2\pi}} (C + 5k)(k + 1) \frac{l^k}{k!}.$$

3. Доказательства теорем

Доказательство теоремы 1. Рассмотрим отображение V , сопоставляющее каждому узлу $\vec{n} = (n_1, \dots, n_k)$, $n_i \in \mathbf{Z}^+$, $i = 1, \dots, k$, сетки k -мерного параллелепипеда

$$P = \left[0, \left[\frac{L+d}{T_1+d}\right]\right] \times \dots \times \left[0, \left[\frac{L+d}{T_k+d}\right]\right]$$

число, равное

$$V(\vec{n}) = \int_{[t, t+d]} \sigma(x) dx, \quad (1)$$

где

$$t = \sum_{i=1}^k n_i(T_i + d) - d.$$

Совокупность узлов с приписанными числами $V(\vec{n})$ назовем весовой сеткой. Рассмотрим отображение $Z(\vec{n})$ с такой же областью определения (соответствующую сетку назовем ценовой), значения которого удовлетворяют условиям

$$\begin{aligned} Z(0, \dots, 0) &= 0, \\ Z(n_1, \dots, n_k) &= \min(Z(n_1 - 1, n_2, \dots, n_k) + V(n_1, \dots, n_k), \dots, \\ &\quad Z(n_1, \dots, n_{k-1}, n_k - 1) + V(n_1, \dots, n_k)), \quad (n_1, \dots, n_k) \neq (0, \dots, 0), \end{aligned} \quad (2)$$

причем если в наборе $(n_1, \dots, n_k)$ компонента $n_i = 0$, то среди аргументов функции $\min$ из (2) нет i -го аргумента.

Заметим, что движение по ребрам сетки, начинающееся в $(0, \dots, 0)$, у которого каждая координата в процессе движения не уменьшается, соответствует подбору некоторого набора $\tau \in \Theta$, если мы приходим в узлы $(n_1, \dots, n_k)$ такие, что

$$L \leq T_1 n_1 + \dots + T_k n_k + (n_1 + \dots + n_k - 1)d < L + \delta. \quad (3)$$

Заметим также, что для всех $\vec{n} \in P$ число $Z(\vec{n})$ равняется минимальной цене пути из тех, по которым можно из $(0, \dots, 0)$ прийти в узел $\vec{n}$. Таким образом, среди целочисленных точек гиперплоскости, определяемых решением (3), есть узел, которому соответствует оптимальный набор τ_0 .

Итак, во-первых, надо вычислить все числа из области значений отображения V , здесь для каждого узла необходимо вычислить интеграл типа (1), определенный на интервале, левый конец которого определяется по ходу заполнения весовой сетки, т.е. с каждым узлом связаны 2 операции сложения: $+T_i$, $+d$; к узлам весовой сетки приписываются и эти числа (обозначим их через $D(n_1, \dots, n_k)$), чтобы не перевычислять заново суммы вида

$$T_1 n_1 + \dots + T_k n_k + (n_1 + \dots + n_k - 1)d.$$

Сложность вычисления каждого интеграла можно оценить сверху не зависящей от k и L константой C_I .

Во-вторых, для вычисления каждого значения отображения Z надо (согласно (2)) k сложений и для выбора минимального из k чисел нужно k сравнений,

кроме того, выполняется проверка: не является ли узел $(n_1, \dots, n_k)$ целочисленной точкой, принадлежащей гиперплоскости, на нее требуется одно сравнение $D(n_1, \dots, n_k)$ с $L + \delta$.

Начиная с узла ценовой сетки, которому соответствует оптимальный набор τ_0 , используя формулу (2), обратным ходом мы восстановим сам набор τ_0 . При этом ни в каком узле мы не побываем дважды, а для каждого проходимого узла мы берем предшествующие ему (т.е. которые только по одной координате на единицу меньше) и, путем не более, чем k сложений и k сравнений, определяем, из которого из них проистекал минимум, даваемый формулой (2).

Осталось заметить, что число узлов в k -мерном параллелепипеде P не превосходит

$$\left(\frac{L+d}{T_1+d} + 1\right) \dots \left(\frac{L+d}{T_k+d} + 1\right).$$

На основе вышесказанного видно, что сложность алгоритма не превосходит

$$\left(\frac{L+d}{T_1+d} + 1\right) \dots \left(\frac{L+d}{T_k+d} + 1\right) ((2 + C_I) + (k + k + 1) + (k + k)),$$

что не превосходит

$$(C + 4k)l^k,$$

здесь $2 + C_I$ соответствует весовой сетке, $k + k + 1$ — ценовой сетке и $k + k$ соответствует обратному ходу.

Доказательство теоремы 2. Для доказательства теоремы 2 модифицируем вышеуказанный алгоритм. Рассмотрим симплекс с ребрами длин

$$a_i = \left\lceil \frac{L+d}{T_i+d} \right\rceil, \quad i = 1, \dots, k,$$

выходящими из начала прямоугольной системы координат параллельно осям. Изменение алгоритма состоит в том, что теперь мы рассматриваем только те части вышеуказанных сеток, узлы которых попали в симплекс. При этом для каждого узла добавляется k операций сравнения $(D(n_1, \dots, n_k) - (L + \delta))/T_j$, $j = 1, \dots, k$, с единицей. Таким образом, вместо коэффициента $(C + 4k)$ мы получаем коэффициент $(C + 5k)$. Осталось оценить количество целочисленных узлов симплекса. Рассмотрим фигуру F , ограниченную гиперплоскостью G_1 , гиперплоскостями $x_1 = -1, \dots, x_k = -1$ и гиперплоскостями $x_1 = a_1, \dots, x_n = a_n$. Эта фигура целиком содержит в себе симплекс, причем для любого целочисленного узла симплекса $(n_1, \dots, n_k)$ кубик со стороной 1, у которого одна из вершин есть $(n_1, \dots, n_k)$, а соседние с ней есть $(n_1 - 1, n_2, \dots, n_k)$, $(n_1, n_2 - 1, n_3, \dots, n_k)$, $\dots$, $(n_1, \dots, n_{k-1}, n_k - 1)$, принадлежит фигуре F . Поэтому количество целочисленных узлов симплекса не превосходит количества единичных кубиков, а значит, и объема фигуры F . Объем этой фигуры $V(F)$ складывается из объема симплекса и объемов слоев: из точки $(0, \dots, 0)$ параллельно координатным осям выходят ребра длин $a_{i_1}, \dots, a_{i_m}$ и $k - m$ ребер единичной длины, $m = 1, \dots, k$,

поэтому

$$\begin{aligned} V(F) &= \frac{a_1 \dots a_k}{k!} + \frac{a_1 \dots a_{k_1} + \dots + a_2 \dots a_k}{(k-1)!} + \dots \\ &\quad + \frac{a_1 a_2 + \dots + a_{k-1} a_k}{2!} + \frac{a_1 + \dots + a_k}{1!} + 1 \\ &\leq \binom{k}{k} \frac{l^k}{k!} + \binom{k}{k-1} \frac{l^{k-1}}{(k-1)!} + \dots + \binom{k}{2} \frac{l^2}{2!} + \binom{k}{1} \frac{l}{1!} + \binom{k}{0}. \end{aligned}$$

Оценим сверху

$$\max_{i=1, \dots, k} l^i \binom{k}{i} / i!, \quad k > l.$$

Воспользуемся неравенствами [4]

$$k^k e^{-k} \sqrt{2\pi k} < k! < k^k e^{-k+1/(12k)} \sqrt{2\pi k},$$

и

$$\frac{l^i}{i!} \binom{k}{i} \leq \frac{k^k}{(k-i)^{k-i}} \left(\frac{p}{i^2}\right)^i \left(e^{1/(12k)} \sqrt{\frac{k}{4\pi^2 i^2 (k-i)}}\right).$$

Последнее выражение в скобках не превосходит $e/\sqrt{2\pi}$, так как $i \geq 1$, $k/(i(k-i)) \leq k/(k-1) \leq 2$.

Рассмотрим функцию

$$f(x) = \left(\frac{p}{x^2}\right)^x (k-x)^{x-k}$$

и найдем ее экстремумы на отрезке $[1, k-1]$. Из $f'(x_0) = 0$ приходим к квадратному уравнению $ex_0^2 + px_0 - pk = 0$, отбрасывая отрицательный корень которого, имеем $x_0 = \sqrt{p^2 + 4epk} - p/(2e)$. Легко видеть, что $C_m \sqrt{k} \leq x_0 \leq C_b \sqrt{k}$ для всех $k > l$, где $C_m = \sqrt{l}/2$, $C_b = \sqrt{l}$. Используя эти оценки, приходим к оценке

$$f(x_0) \leq e^{-k \ln k} e^{(C_b - C_m) \sqrt{k} \ln k} e^{-C_b C_m} e^{\sqrt{k} C_q},$$

где

$$C_q = C_b \ln p - c_m \ln C_m^2 - \sqrt{l+1} \ln \left(1 - \frac{C_b}{\sqrt{l+1}}\right).$$

Если взять $C_r = C_q / \ln l$, то $C_q \leq \ln k C_r$ для всех $k > l$. Отсюда следует, что

$$k^k f(x_0) \leq e^{\sqrt{k} \ln k (C_m + C_b + C_r)} \leq e^{C_p \sqrt{k} \ln k},$$

где $C_p = 5\sqrt{l}(1/2 + 1/\ln l)$.

Осталось рассмотреть граничные случаи $i = 0, 1, k-1, k$. Можно показать, что числа $l^i \binom{k}{i} / i!$, соответствующие этим случаям, не превосходят $e^{2\sqrt{kl} \ln k}$ при $k > le$, $l > 1$. $C_p > 2\sqrt{l}$. Значит,

$$V(F) \leq (k+1) e^{C_p \sqrt{k} \ln k}$$

для всех $k > le$, $l > 1$. Отсюда следует доказываемая оценка.

Доказательство теоремы 3. Воспользуемся построенным при доказательстве теоремы 2 алгоритмом. Перепишем x_0 в виде $x_0 = (\sqrt{l^2 + 4lk} - l) / 2$. Если $l > 2k(k-1)$, то $(l + 2k)^2 > l^2 + 4lk > (l + 2k - 1)^2$ и, значит, $k \geq x_0 \geq k - 1$. Таким образом, $\max_{i=0, \dots, k} l^i \binom{k}{i} / i!$ не превосходит максимального из чисел $l^k / k!$, $kl^k / (k-1)!$, lk , 1. Из $l > 2k(k-1)$ следует, что $l \geq k^2$ при $k > 1$ и, следовательно, $l^k / k! \geq kl^k / (k-1)!$; из $l \geq k^2$ легко получить, что $kl^k / (k-1)! \geq lk$ при $k > 2$. Итак, максимальным из этих чисел является $l^k / k!$. Отсюда следует доказываемая оценка.

Автор благодарит Д. Н. Бабина, под руководством которого выполнена эта работа.

Список литературы

1. Болтянский В. Г. *Математические методы оптимального управления*. Наука, Москва, 1960.
2. Беллман Р. *Динамическое программирование*. Мир, Москва. 1966.
3. Шевченко В. П. *Качественные вопросы целочисленного программирования*. Наука, Москва. 1995.
4. Корн Г., Корн Т. *Справочник по математике*. Наука, Москва. 1970.

Статья поступила 23.10.96.