

УДК 004.932.72

Д. В. ЗАЕРКО, В. А. ЛИПНИЦКИЙ

ВЫЧИСЛЕНИЕ СРЕДНЕГО РАССТОЯНИЯ МЕЖДУ ПИКАМИ ГОРИЗОНТАЛЬНОЙ ПРОЕКЦИИ

Белорусский государственный университет информатики и радиоэлектроники,
Минск, Республика Беларусь

В ходе анализа и построения алгоритма шумофильтрации на этапе сегментации символьных строк, выяснилась необходимость в описании особой базы или словаря для хранения модельной и скелетной проекции символов алфавита. Однако, при форматировании записей алфавитной проекции в словаре возникли вопросы: сколько значений проекции необходимо для описания единственного символа, а также, как может изменяться данная величина в зависимости от числа проанализированных строк и их символов. Объекты исследования в статье – это вертикальная проекция символов, а так же их ширина сегментов проекции. Предмет изучения – изменение средней ширины символьного сегмента проекции для определенного типа номерных знаков (символьных строк). Основная цель – вычисление и обоснование средней величины сегмента. Информация о средней ширине сегмента проекции, описывающей символ, позволяет определить число значений проекции (координат по оси ox) для хранения самих значений проекций этого символа из некоторого алфавита. Рассматриваются несколько языковых алфавитов, где каждому из них сопоставлен уникальный тип. В частности рассмотрены номерные знаки Республики Беларусь и Словацкой Республики. Используя элементарный статистический аппарат, проведены вычисления и анализ выборок, учитывая случай высокой их вариации. В результате получены средние значения ширины сегментов и значения дисперсии ширины сегментов на примере двух типов алфавитов. Попутно, в ходе исследования, представлен алгоритм использования полученных величин при формировании записей модельной проекции. Алгоритм учитывает «особые» случаи выхода за границы сегментов. Описанные шаги и вычисления возможно применить для большего числа алфавитов, что указывает на возможность описания словарей модельной проекции для них, с последующим широким использованием в шумофильтрации символьных строк.

Ключевые слова: вертикальная проекция, символьная строка, шумофильтрация, средняя ширина сегмента, экстремумы проекций, модельная и скелетная проекции, алфавитная база проекций.

Введение

Задача распознавания символьных строк не является новой [1, 2, 3], и ее практическая реализация в современной инфраструктуре и повседневной деятельности человека широко используется, например, для мониторинга движения автотранспорта, для считывания штрих-кодов на товарах и т. д. Однако, алгоритмы не всегда работают безошибочно, что связано как с техническими, так и с человеческими факторами. Процесс распознавания – это многофакторный и многоступенчатый процесс, скрывающий излишние подробности для пользователей.

Процесс распознавания символьных строк на изображении разбит на множество этапов, в котором каждый последующий напрямую зависит от эффективности предыдущего. Как по-

казывает практика, подготовительные этапы оказывают существенное влияние на конечный результат распознавания. Одним из таких этапов является посимвольная сегментация строки. Зачастую задача символьной сегментации решается путем верного анализа вертикальной или горизонтальной проекции на наличие экстремумов, которые сигнализируют о возможном разрыве между символами.

При разработке шумофильтрирующей модификации алгоритма сегментации символьных строк, на базе анализа горизонтальной проекции [4], возник вопрос верного определения средней ширины символа. Верное определение расстояния между границами экстремумов дает возможность использования этой информации при формировании координатной сетки, где каждый локальный экстремум ука-

зывает на разрыв между символами строки. Ранее, в статье [5], внимание акцентировалось на том, что размер символов строки, подлежащей распознаванию одинаков (например, размер символа автомобильного номера), однако, речь шла лишь о человеческом восприятии. С позиции машинного зрения, которое использует пиксельную матрицу – карту монохромного изображения и ее проекцию на одну из осей, размер символов может разниться по различным причинам. Они связаны как с техническими средствами, так и с неправильной работой алгоритмов обработки и анализа пиксельных изображений. Причинами могут быть:

- искажения пиксельного скелета символа при нормализации символьной строки;
- ошибочный выбор точек сегментации символов (пиков горизонтальной проекции);
- сильной зашумленностью первоначального источника изображения.

В данной работе представлены этапы анализа и вычисления средней ширины символа, необходимой для дальнейших исследований в области очистки символьных строк от шумов различного характера. Результаты, полученные здесь, будут играть важную роль для формирования координатной сетки.

Вычисление средней ширины

Для удобства описания анализируемых данных представим их в матричной форме. Здесь каждый столбец представляет собой одну проанализированную и разбитую на $n \in N$ сегментов символьную строку, а число столбцов в матрице равно числу проведенных измерений символьных строк $m \in M$. Строчный элемент столбца $l \in L$ это ширина сегмента, одного единственного символа. Здесь N, M – множество всех сегментов для каждой символьной строки и множество всевозможных символьных строк, а L – множество возможных отрезков (ширин).

Для оценки средней ширины символа воспользуемся вероятностными характеристиками [6]. В качестве случайной величины l выступает ширина символа (расстояние между экстремумами) из однородной совокупности (предположение) Ω . Оценка средней ширины будет равна среднему арифметическому всех длин $\bar{l} = \frac{1}{N+M} \sum_{j=1}^M \sum_{i=1}^N l_{ij}$, дисперсия

$$\sigma^2 = \frac{1}{N+M} \sum_{j=1}^M \sum_{i=1}^N (l_{ij} - \bar{l})^2, \text{ среднееквадратическое}$$

$$\text{отклонение } \sigma = \sqrt{\frac{1}{N+M} \sum_{j=1}^M \sum_{i=1}^N (l_{ij} - \bar{l})^2} \text{ и коэффициент вариации } V = \frac{\sigma}{\bar{l}}.$$

При разработке программной модификации проекта ANPR были проанализированы $M = 100$ номерных знаков для каждого из двух типов. Для их (т. е. символьной строки на ней) были составлены вертикальные проекции по соответствующим пиксельным матрицам и проведена сегментация на интервалы по локальным экстремумам.

Важным пунктом исследования был анализ нескольких типов строк, совпадающих по числу интервалов разбития. Первый тип представлен номерными знаками Словацкой Республики, где каждая строка разбита, как правило, на $N = 12$ сегментов. Второй тип – номера Республики Беларусь (образца 2004 г. – 12 сегментов. Сегменты получены после анализа экстремумов горизонтальной проекции для каждой символьной строки. Подробный алгоритм построения горизонтальной проекции и его применение, для распознавания автомобильных номеров, можно найти в рамках проекта JavaANPR [1].

В результате вычислений получены следующие значения для словацкого номерного знака и для белорусского номерного знака:

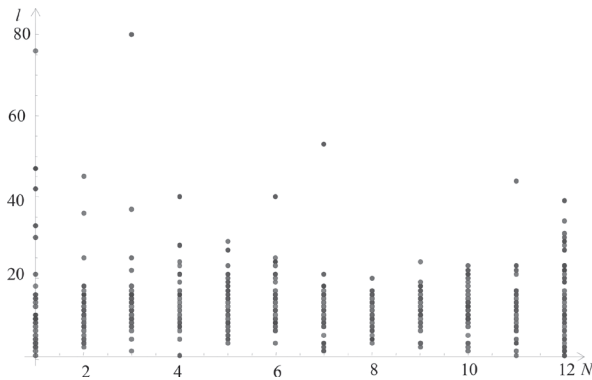
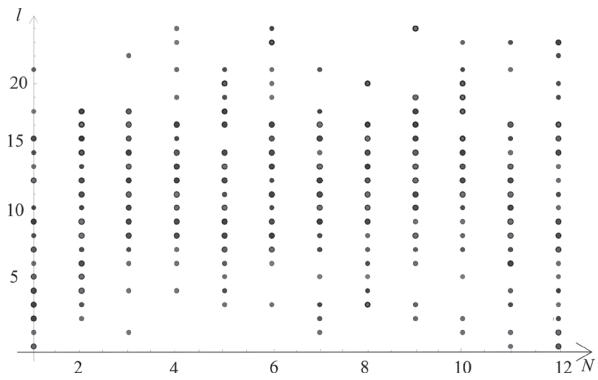
$$\bar{l}_{sk} = 11,87 \approx 12; \sigma_{sk}^2 = 38,49; \sigma_{sk} = 6,20;$$

$$V_{sk} = 0,52; \min l_{sk} = 1, \max l_{sk} = 80;$$

$$\bar{l}_{rb} = 11,80 \approx 12; \sigma_{rb}^2 = 41,43; \sigma_{rb} = 6,43;$$

$$V_{rb} = 0,54; \min l_{rb} = 1, \max l_{rb} = 74.$$

Как известно из статистики, коэффициент вариации позволяет делать вывод об отклонении измеренных значений от среднего арифметического. Чем большее значение V , тем больше разброс вариационного ряда, и, следовательно, меньшая однородность. В нашем случае, оба ряда кажутся неоднородными, т. к. $V > 0,33$, и требуют исключения отдельных, сильно отличающихся значений. Для наглядности представим распределение частоты длин, например для первого типа номеров. Ось Oy описывает ширину l сегментов символа. Ось Ox описывает число N сегментов, на которые раз-

Рис. 1. Распределение частот длин l для первого типаРис. 2. Распределение частот длин l для первого типа, после удаления скачков

бита символьная строка, согласно первоначальному анализу экстремумов.

Максимальные и минимальные размеры сегментов свидетельствуют о большом расхождении размеров, а так же указывают на возможные ошибки измерения размеров сегментов. На ошибки указывают нечастые и резко отличающиеся в большую сторону значения. Как видно, таких выступающих значений не много (порядка 10) и все они лежат выше $l > 30$. С практической точки зрения они легко объяснимы ошибками сегментации строки по экстремумам (несколько символов склеены между собой) из-за сильного зашумления изображения номерного знака, или из-за неправильного анализа экстремумов по вертикальной проекции. Исключим из выборок явно выделяющиеся ошибочные значения при сегментации (скачки), для обоих типов.

После удаления заметна визуальная полоса, в которой лежат все значения. Характеристики примут следующие значения:

$$\bar{l}_{sk} = 11,11 \approx 11; \sigma_{sk}^2 = 16,82; \sigma_{sk} = 4,10;$$

$$V = 0,36; \min l_{sk} = 1, \max l_{sk} = 25;$$

$$\bar{l}_{rb} = 11,28 \approx 11; \sigma_{rb}^2 = 21,21; \sigma_{rb} = 6,60;$$

$$V = 0,40; \min l_{rb} = 1, \max l_{rb} = 25.$$

Коэффициенты вариации стали близки $V \approx 0,33$, следовательно, можно сделать вывод об относительной однородности выборок. Касательно средних значений, видны значительные изменения дисперсии, а также среднее значение ширины принимает значение, близкое $\bar{l} \approx 11$. Учитывая среднее значение, можно провести правильное формирование координатной сетки для описания модельной и скелетной проекции символов, в которой каждый символ

может быть описан сегментом, состоящим из $i \in [0, \bar{l}]$ значений проекции на ось Ox .

Использование средней ширины

Используя среднюю ширину сегмента символа, и определив цену одного деления сегмента, можно сформировать модельный шаблон символьной проекции. Модельная проекция символа будет представлять собой множество пар вида: $\langle i, p(x_i) \rangle$, где i – порядковый номер позиции в шаблоне символа по оси Ox , ограниченный числом \bar{l} . $p(x_i)$ – значение проекции в каждой точке x_i , $i \in [0, \bar{l}]$.

В дальнейшем полученные величины будем использовать при составлении шаблонов алфавитных символов следующим образом:

1) если ширина анализируемого символа совпадает со средним размером символа (с размером заданной оси), то поточечные значения проекции этого символа будут сохранены при формировании одного из модельных шаблонов символа;

2) если ширина анализируемого символа меньше, чем средний размер символа, то поточечные значения проекции будут также сохранены в качестве шаблона, однако разница будет дополнена 0-ми значениями проекции слева и справа относительно реальной ширины;

3) если ширина анализируемого символа больше, чем средний размер символа, то поточечные значения проекции представляют собой несколько символов, соединенных в один фрагмент.

В последнем случае, дополнение или, наоборот, деление на равные части (на величину средней ширины) проекции будет не корректной, т. к. могут быть утеряны значимые части проекции. Следовательно, она не будет доста-

точно информативной для сравнения. Возможно, для решения этой проблемы подойдет более тонкий анализ экстремумов фрагмента проекции, однако, это лишь предположение, требующее дополнительных исследований в этом направлении.

Заключение

Вычисленная средняя ширина символа, для двух типов номерных знаков, позволяет описать алфавитную базу модельных проекций для их, где каждый элемент алфавита – это символ представленный набором значений проекций, а размер этого набора определяется,

исходя из средней ширины символа для определенного типа номеров. Дисперсия может быть использована для оценки допустимой погрешности, также учитываемая при создании алфавитной базы. Формирование базы, для модельной и скелетной проекции символов алфавита, является одним из важнейших частей составления алгоритма шумофильтрации проекции. Шаги, описанные в данной статье, могут быть проведены для любого алфавита, а это значит, что средние характеристики могут быть вычислены для формирования соответствующих модельных и скелетных шаблонов проекций.

ЛИТЕРАТУРА

1. **Ondrej Martinsky.** Algorithmic and Mathematical Principles of Automatic Number Plate Recognition Systems. // B. Sc. Thesis, Brno University of Technology, Faculty of Information Technology, Department of Intelligent Systems, 2007. – 76 с.
2. **Визильтер Ю. В., Желтов С. Ю., Князь В. А., Ходарев А. Н., Моржин А. В.** Обработка и анализ цифровых изображений с примерами на LabVIEW IMAQ Vision. – М.: ДМК Прес, 2007. – 464 с.
3. **Гонсалес Р., Вудс Р.** Цифровая обработка изображений // М.: Техносфера. 2005. 1072 с.
4. **Заерко, Д. В.** Применение модифицированных алгоритмов JavaANPR для автоматического распознавания номеров автомобилей / Д. В. Заерко, В. А. Липницкий // Информационные технологии и системы 2018 (ИТС – 2018): материалы международной научной конференции, БГУИР, Минск, Беларусь, 25 октября 2018 г. – С. 286–287.
5. **Заерко Д. В.** Фильтрация сегментной проекции символьной строки на этапе сегментации. Д. В. Заерко, В. А. Липницкий. // XVII Международная конференция «Технические средства защиты информации»: материалы международной научной конференции, БГУИР, Минск, Беларусь. 11 июня 2019. Минск. – С. 29.
6. **Гмурман. В. Е.** Теория вероятностей и математическая статистика: учебное пособие / В. Е. Гмурман. – Издание 9-е, дополненное. – М.: Высшая школа, 2003. – 479 с.

REFERENCES

1. **Ondrej Martinsky.** Algorithmic and Mathematical Principles of Automatic Number Plate Recognition Systems. // B. Sc. Thesis, Brno University of Technology, Faculty of Information Technology, Department of Intelligent Systems, 2007.
2. **Vizilter U. V., Zheltov S. U., Knjaz V. A., Hodarev A. N., Morzhin A. V.** Obrabotka i analiz cifrovyyh izobrazhenij s primerami na LabVIEW IMAQ Vision – M.: DMK Pres, 2007. – 464 p.
3. **Gonzalez, R. C. and Woods, R. E.** Digital Image Processing. 3rd Edition, Prentice Hall, Upper Saddle River, 2008. – 1074 p.
4. **Zaerko, D. V.** Primenenie modifitsirovannyh algoritmov JavaANPR dlja avtomaticheskogo raspoznavaniya nomerov avtomobilej / D. V. Zaerko, V. A. Lipnitski // Information technologies and systems 2018 (ITS– 2018): materials of the international scientific conference, BSUIR, Minsk, Rep. Belarus, 25 October 2018 г. – P. 286–287.
5. **Zaerko, D. V.** Fil'tracija segmentnoj proekcii simvol'noj stroke na jetape segmentacii. D. V. Zaerko, V. A. Lipnitski // XVII International conference «Tehnicheskie sredstva zashhity informacii»: materials of scientific conference, BSUR, Minks, Rep. Belarus. 11 June 2019. Минск. – P. 29.
6. **Gmurman V. E.** Teorijaverojatnostej i matematicheskajastatistika: uchebnoeposobie/ V. E. Gmurman. – 9-th Revised Edition, 2003. – 479 p.

Поступила
01.07.2019

После доработки
12.09.2019

Принята к печати
01.10.2019

ZAERKO D. V., LIPNITSKI V. A.

CALCULATING THE AVERAGE DISTANCE BETWEEN HORIZONTAL PROJECTION PEAKS

During the analysis and construction of the noise filtering algorithm at the stage of segmentation of character strings, the need to describe a special database or dictionary for storing model and skeletal projections of alphabet characters became clear. However, when formatting entries for alphabetical projection in the dictionary, the following questions arose: how many projection values are needed to describe a single character, and also how this value may change depending on the number of

strings analyzed and their characters. The objects of research in the article are the vertical projection of the characters, as well as their width of the projection segments. The subject of research is the change of the average width of the projection segment for a certain type of license plates (character string). The main goal is to calculate and justify the average size of the segment. Data about the average width of the projection segment that describes a character allows to determine the number of projection values (coordinates along the ox axis) to store the projection values of this character from an alphabet. Several language alphabets are considered, where each of them is associated with a unique type. In particular, license plates of the Republic of Belarus and the Slovak Republic are considered. Using an elementary statistical apparatus, calculations and analysis of samples were carried out, taking into account the case of their high variation. As a result, the average values of the segment width and the variance of the segment width were obtained using two types of alphabets as an example. In research an algorithm for using the obtained values in the formation of model projection records is presented. The algorithm takes into account «special» cases of going beyond the segment boundaries. The described steps and calculations can be applied to a larger number of alphabets, which indicates the possibility of describing model projection dictionaries for them, with subsequent widespread use of character strings in noise filtering.

Keywords: vertical projection, symbolic string, noise filtering, average segment width, projection extreme points, model and skeleton projection, alphabetic projection dictionary.



Заерко Денис Владимирович, аспирант кафедры информатики БГУИР, область научных исследований – машинные алгоритмы распознавания объектов, машинное обучение, анализ данных. E-mail: zaerko1991@gmial.com.

Zaerko. D. V., postgraduate student, Informatics department of BSUR, E-mail: zaerko1991@gmial.com.



Липницкий Валерий Антонович, профессор, доктор технических наук, область научных интересов – алгебра и ее приложения, защита информации от помех и несанкционированного доступа. E-mail: valipnitski@yandex.by.

Lipnitski. V. A. Doctor of technical sciences, Professor, Informatics department of BSUR. E-mail: valipnitski@yandex.by.