УДК 004.932.2

РАСПОЗНАВАНИЕ ПОКАЗАНИЙ ИЗ ФОТОГРАФИЙ БЫТОВЫХ СЧЕТЧИКОВ ВОДЫ И ЭЛЕКТРОЭНЕРГИИ

Якунин А.Г., Наздрыухин А.С., Дунаев А.С.

ФГБОУ ВО «Алтайский государственный технический университет имени И.И. Ползунова,
Барнаул, e-mail: a.nazdryukhin@gmail.com

Разработан и описан алгоритм для распознавания показаний бытовых счетчиков на изображениях, полученных с фотокамеры смартфона или мобильного телефона. Основное внимание уделялось детектированию на изображении области, соответствующей циферблату счетчика, и последующей обработке изображения циферблата. Детектирование проводилось с помощью поиска и выделения прямоугольного участка, обладающего определенными геометрическими свойствами. Для обработки изображения был разработан специальный метод повышения качества, состоящий из последовательно применяющихся методов шумоподавления, выравнивания освещения и баланса белого, гамма-коррекции, повышения резкости и адаптивной локальной бинаризации. Разработанный алгоритм включает в себя программный модуль, выполняющий оптическое распознавание цифровой информации с помощью сверточной нейронной сети. Разработанный алгоритм устойчив к поворотам, изменению перспективы, умеренным отклонениям в яркости, контрасте, освещенности и другим дефектам изображения, таким как хроматические аберрации, шумы и так далее. Для работы алгоритма не требуются изображения в высоком разрешении, что делает его использование доступным при сохранении достаточно высокой скорости работы. Также были предложены способы улучшения разработанного алгоритма для повышения точности детектирования циферблата и распознавания цифровой информации.

Ключевые слова: показания бытовых счетчиков, обработка изображений, анализ изображений, сверточные нейронные сети, распознавание образов

RECOGNITION OF METERS READINGS FOR MEASURING WATER AND ELECTRICITY CONSUMPTION FROM PHOTOS

Yakunin A.G., Nazdryukhin A.S., Dunaev A.S.

Altai State Technical University, Barnaul, e-mail: a.nazdryukhin@gmail.com

In this paper was developed and described the algorithm that can recognize a indication of household meters from images obtained from the smartphone camera. The main attention was directed to the detection on the image area corresponding to the counter dial and subsequent processing of the image of the dial. The process of detection was carried out by using the search and selection of a rectangular area with certain geometric properties. For image, processing was developed a special method of improving the quality, consisting of successively applied methods of noise reduction, illumination equalization and white balance correction, gamma correction, sharpening, and adaptive local binarization. The developed algorithm includes a software module that performs optical recognition of digital information using a convolutional neural network. The developed algorithm is resistant to rotations, changes in perspective, moderate deviations in brightness, contrast, illumination and other image defects, such as chromatic aberration, noise, etc. High-resolution images are not required for algorithm's work, which makes its use available while maintaining a fairly high speed. Moreover, were proposed some ways to improve the developed algorithm to improve the accuracy of the detection of the dial and the recognition of digital information.

Keywords: domestic meters, image processing, image analysis, convolutional neural network, pattern recognition

Непрерывное развитие современной техники подразумевает под собой разработку и внедрение технологий, позволяющих упростить осуществление тех или иных процессов, возникающих в повседневной жизни.

На сегодняшний день практически в каждом жилом помещении присутствуют счетчики контроля расхода электроэнергии, горячей и холодной воды, которые позволяют осуществлять контроль за процессом распределения соответствующих величин. Наличие подобных систем контроля подразумевает под собой взаимодействие потребителей с производителем услуг посредством отправки текущих показаний счетчиков в соответствующие коммунальные службы с некоторой периодичностью.

Коммунальные службы, в свою очередь, с целью упрощения взаимодействия потребителей с производителем, используют специализированные сайты. Также компаниям требуются сотрудники, осуществляющие периодические проверки целостности счетчиков и корректности предоставляемых показаний.

Для упрощения передачи данных от потребителя к производителю услуг предлагается использовать специальное программное обеспечение для мобильных телефонов или смартфонов, что стало возможным благодаря динамическому развитию цифровых технологий. В частности, современные смартфоны обладают качеством фотографии, достаточным для использования

получаемых изображений с целью распознавания определенных элементов. Более того, в настоящее время смартфон не является предметом роскоши и обладает высоким уровнем доступности и, как следствие, может значительно упростить процесс контроля показателей счетчиков.

Цель исследования: разработка, реализация и тестирование алгоритма для распознавания показаний бытовых счетчиков на изображениях, полученных с фотокамеры смартфона или мобильного телефона с помощью технологий компьютерного зрения.

Материалы и методы исследования

Алгоритм состоит из четырех основных этапов (рис. 1).

Детектирование циферблата счетчика. Данный этап осуществлялся с помощью определения контура циферблата, полученного детектором границ Кэнни [1]. Алгоритм для выделения контуров Кэнни основан на модели ступенчатых перепадов яркости. Задача нахождения оптимального детектора не может быть решена аналитически, только с помощью численных методов. Однако Кэнни показал, что хорошим приближением к оптимальному детектору ступенчатых перепадов является первая производная гауссиана. Для нахождения градиента изображения, сглаженного с помощью фильтра Гаусса, можно использовать оператор Собеля или – для более точных вычислений – оператор Собеля –Шарра [2], который определяется по формулам

$$G_x = \begin{pmatrix} 3 & 0 & -3 \\ 100 & -10 \\ 3 & 0 & -3 \end{pmatrix}; G_y = \begin{pmatrix} 3 & 10 & 3 \\ 0 & 0 & 0 \\ -3 & -10 & -3 \end{pmatrix}.$$

Тогда результат свертки изображения и матриц G_x и G_y будет являться соответствующими координатами вектора градиента. Далее по формулам можно получить G – модуль вектора градиента (1) и θ – угол направления градиента (2) для каждого пикселя.

$$G = \sqrt{G_x^2 + G_y^2}, \quad (1)$$

$$\theta = \arctan\left(\frac{G_y}{G_x}\right). \quad (2)$$

При вычислении θ должна округляться до ближайшего значения к $0^\circ, 45^\circ, 90^\circ, 135^\circ$.

Чтобы детектор имел высокую точность выделения границ, используются методы подавления точек, соответствующих немаксимальным значениям векторов градиента, и порогового преобразования с гистерезисом [1].

После выделения контуров мы следуем простой эвристике, которую можно обозначить следующими пунктами:

1. Циферблат должен быть расположен близко к центру изображения.

2. Циферблат должен иметь прямоугольную форму, то есть контуры циферблата, выделенные алгоритмом Кэнни, подходят на четырехугольник при аппроксимации (в данной работе использовался алгоритм Рамера – Дугласа – Пекера [3]).

3. Ширина циферблата гораздо больше, чем его высота.

Используя данные замечания, можно отфильтровать все контуры так, чтобы остался только контур циферблата. Результат показан на рис. 2.

Поскольку снимок может иметь некоторые искажения перспективы (на рис. 2, в – циферблат имеет трапециевидный контур), следует произвести пространственную коррекцию. Для этой цели использовался алгоритм преобразования по четырем опорным точкам на изображении (рис. 2, в): (x'_i, y'_i) , $i = 1, \dots, 4$. Далее, если обозначить матрицу преобразования как M , то коррекция перспективы выполняется по следующим формулам:

$$(x_k, y_k) \rightarrow \frac{1}{w} (x'_k, y'_k) (x'_k, y'_k, w)^T = \\ = M (x_k, y_k, 1)^T \quad w = \{w', w' \neq 0\infty, \text{ иначе},$$

где (x'_i, y'_i) и (x_i, y_i) – координаты i -го пикселя исходного и преобразованного изображения соответственно. То есть преобразование выполняется таким образом, чтобы опорные точки стали углами прямоугольника.

Рис. 1. Блок-схема работы алгоритма

Рис. 2. а) исходная фотография счетчика, б) выделенные с помощью алгоритма Кэнни границы, в) изображение с отмеченным расположением циферблата

Рис. 3. Блок-схема алгоритма обработки изображения циферблата

Коррекция изображения циферблата. В данной работе предполагалось, что изображения получены с фотокамеры смартфона или мобильного телефона. Такие фотографии имеют множество дефектов, например цифровой шум, нарушения экспозиции, неравное освещение, геометрические абберации или различные хроматизмы.

Для компенсации данных дефектов использовался метод повышения качества изображения для более четкого выделения цифр на циферблате. Предварительная обработка изображения выполнялась с помощью специально подобранных алгоритмов нормализации изображения, применяемых последовательно (рис. 3). Более подробно данный подход описан в [4].

Основная идея подавления шумов на изображении состоит в том, что яркость целевого пикселя в некотором блоке размером m пикселей берется как соотношение яркостей пикселей во всем блоке (количество пикселей в блоке лучше выбирать как квадрат некоторого натурального числа).

Как правило, методы для удаления шумов отличаются друг от друга только способами выбора блока. В данной работе использовался метод NL-means [5], определяемый формулой вида

$$NL[u(p)] = \frac{1}{C(p)} \int y(d[B(p), B(q)]) u(q) dq,$$

где $d[B(p), B(q)]$ – евклидово расстояние между блоками с центрами p и q , $y(x)$ – функция веса, $C(p)$ – фактор нормализации, $u(q)$ – значение неотфильтрованного пикселя изображения. Интеграл идет по всем точкам в блоке.

Для выравнивания освещения и регулирования баланса белого последовательно применялись алгоритмы SSR [6] и Gray World [7]. Фильтрация SSR вычисляется по формуле

$$f := \log \log f - \log \log (f * g),$$

где $f * g$ – результат свертки исходного изображения и двумерного фильтра Гаусса.

Алгоритм Gray World для трехканального изображения с интенсивностями $R, B, G \in [0, 255]$ по каждому пикселю выглядит следующим образом:

$$scale = \frac{R_{avg} + B_{avg} + G_{avg}}{3}; R := \frac{R * scale}{R_{avg}};$$

$$B := \frac{B * scale}{B_{avg}}; G := \frac{G * scale}{G_{avg}},$$

где $R_{avg}, B_{avg}, G_{avg}$ – средние значения интенсивностей каждого из каналов изображения.

Гамма-коррекция [1, 8] используется для регулирования контраста и яркости изображения. Яркость каждого пикселя для скорректированного изображения определяется как

$$p := 255 \left(\frac{p}{255} \right)^\gamma,$$

где γ – коэффициент гамма-коррекции.

Повышение резкости. Пусть g – гауссиан, e – единичный фильтр, a – коэффициент резкости, тогда, при выполнении перехода к свертке, получим формулу для повышения резкости f .

$$f + a(f - f * g) = f * ((1 + a)e - ag).$$

Тогда изображение с повышенной резкостью определяется по следующей формуле:

$$f := f * ((1 + a)e - ag).$$

Бинаризация изображения [1]. Невозможно гарантировать, даже после выполненных выше преобразований, что яркости нужных нам областей будут схожи настолько, что можно будет подобрать глобальный коэффициент бинаризации. Поэтому в качестве алгоритма бинаризации была выбрана адаптивная бинаризация, которая зависит от яркости изображения в конкретном пикселе и размера окна бинаризации, а не от координат пикселя, как в случае глобальной бинаризации. Иначе говоря, в подходе адаптивной бинаризации рассматриваются пиксели в некотором окне, для которого некоторым способом определяется оптимальный коэффициент бинаризации. Например, по среднему значению интенсивностей пикселей.

Выделение цифровой информации из циферблата. Для выделения цифровой информации из циферблата использовалось бинарное изображение циферблата, полученное в предыдущем разделе. Далее на нем необходимо выделить все контуры и отфильтровать слишком маленькие или слишком большие. Это представляется воз-

можным по той причине, что циферблаты счетчиков имеют одинаковое количество цифр (шесть или семь) и поэтому при масштабировании изображений циферблатов к некоторому заданному разрешению контуры цифр будут соразмерны.

Тогда алгоритм выделения цифр можно описать следующим образом:

1. Приведение бинаризованного изображения циферблата к фиксированному размеру (рис. 4, б).

2. Выделение всех контуров на изображении (рис. 4, в).

3. Для каждого контура находим описывающий его минимальный прямоугольник.

4. Если ширина и высота прямоугольников лежат в некотором фиксированном диапазоне, то данный прямоугольник содержит искомую цифру (рис. 4, г).

Результат выделения показан на рис. 4, д.

Рис. 4. а) исходное изображение циферблата, б) изображение циферблата, полученное с помощью пороговой бинаризации, в) изображение с выделенными контурами, г) изображение с помеченными прямоугольниками с цифрами, д) полученные цифры в бинарном виде

Распознавание найденных цифр. Для распознавания цифр использовалась модель сверточной нейронной сети [9, 10]. Характерное свойство сверточной нейронной сети – устойчивость к искажениям, поворотам и переносам целевого объекта на изображении благодаря свертке с помощью локальных шаблонов. Такие шаблоны встраиваются в модель, формируя пространственные иерархии, как и объекты в реальном мире.

Поскольку модель сверточной нейронной сети обладает огромным количеством параметров, для ее обучения требуется массивный набор данных, репрезентатив-

ный относительно классов распознавания. Сформировать подходящий набор данных сложно, поэтому в качестве основы использовалась база рукописных цифр MNIST, дополненная изображениями для каждого класса, полученными с помощью аугментации [9] (рис. 5) изображений цифр с циферблата счетчиков.

Точность распознавания в метрике Ассигасу [10] равна 98,58 % для тестовой выборки.

Рис. 5. Пример аугментации

Результаты исследования и их обсуждение

Детектирование циферблата основано на геометрических соображениях, однако циферблаты некоторых счетчиков могут им не удовлетворять. Например, счетчик может иметь циферблат с круглой формой или высота циферблата может быть больше ширины. В таких случаях имеет смысл разработать специализированную модель машинного обучения для детектирования циферблата на счетчике. Также возможен поиск и распознавание цифр напрямую на изображении счетчика, без детектирования циферблата. Однако последний способ не позволяет корректно обработать изображение.

В данной работе распознавание цифровой информации производится с помощью модели сверточной нейронной сети, обученной на наборе данных MNIST, дополненном изображениями цифр со счетчиков. Однако гораздо лучше будет сформировать новый набор данных, состоящий из цифр, характерных для циферблатов бытовых счетчиков. И хоть такие цифры представлены в машинописных шрифтах, при их извлечении могут образоваться некоторые иска-

жения формы (как, например, у цифр, показанных на рис. 4, д). Поэтому часть набора данных, и достаточно большую, необходимо будет формировать вручную непосредственно из изображений счетчиков. Также необходимо будет учесть то, что для цифр могут использоваться различные шрифты, поэтому составить репрезентативный набор данных будет достаточно сложно.

Заключение

В данной работе был описан алгоритм, который помогает распознать цифровую информацию из изображений бытовых счетчиков, полученных с помощью фотокамеры смартфона или мобильного телефона. Алгоритм достаточно эффективен, в него включены методы, которые позволяют произвести подавление шумов и других искажений изображения, а также выполнить поворот и пространственную коррекцию циферблата бытового счетчика. Данные методы позволяют подготовить изображения для распознавания цифр.

Также были предложены пути дальнейшей разработки, которые, предположительно, позволят улучшить представленный алгоритм.

Список литературы

1. Гонсалес Р., Вудс Р. Цифровая обработка изображений. М.: Техносфера, 2012. 1104 с.
2. Jähne B., Scharf H., Körkel S. Principles of Filter Design. Handbook of Computer Vision and Applications. 1999. Vol. 2. P. 125–151.
3. Douglas D., Peucker Th. Algorithms for the reduction of the number of points required to represent a digitized line or its caricature, The Canadian Cartographer. 1972. Vol. 10. No. 2. P. 112–122. DOI: 10.3138/FM57-6770-U75U-7727.
4. Наздрюхин А.С., Храмов И.Н., Тушев А.Н. Обработка изображений товарных чеков для выделения и распознавания текстовой информации // Вестник Дагестанского государственного технического университета. Технические науки. 2019. № 4. С. 113–122.
5. Chambolle A., Caselles V., Novaga M., Cremers D., Pock T. An introduction to Total Variation for Image Analysis: Lecture Notes. Linz, 2010. 87 p.
6. Ma Z., Wen J. Single-scale Retinex sea fog removal algorithm fused the edge information. Journal of Computer-Aided Design and Computer Graphics. 2015. Vol. 27. P. 217–225.
7. Li B., Xu D., Lee M., Feng S. A Multi-Scale Adaptive Grey World Algorithm. IEICE Transactions. 2007. Vol. E90-D. No. 7. P. 1121–1124.
8. Babakhani R., Zarei P. Automatic gamma correction based on average of brightness. Advances in Computer Science: an International Journal. 2015. Vol. 4. No. 6. P. 156–159.
9. Николенко С.И. Глубокое обучение. СПб.: Питер, 2018. 481 с.
10. Флах П. Машинное обучение. М.: ДМК Пресс, 2015. 400 с.