

Обнаружение, локализация и трансформация двумерного графического кода

А.А. Трубицын¹, М.В. Шадрин²

¹ Рязанский государственный радиотехнический университет имени В.Ф. Уткина, ул. Гагарина 59/1, г. Рязань, 390005, Россия

² ООО «Квантрон Групп», ул. Энгельса, д. 28а, помещ. Н2, Рязань, 390010, Россия

Аннотация

Область притязаний статьи относится к компьютерному зрению. Представляются математические основы и алгоритмы подготовки двумерных графических кодов к декодированию данных. Этапы подготовки включают в себя обнаружение, локализацию и трансформацию кода к удобному для раскодирования виду. Предложенная в работе функция осцилляций яркости служит надежному обнаружению кодов на изображении, преобразование Хафа лежит в основе решения задачи локализации, а формулы обратного преобразования квадрата в произвольный четырехугольник позволяют трансформировать локализованные четырехугольные области в канонические квадраты. Проведено исследование работы алгоритмов на множестве тестовых и реальных изображений, показан несущественный уровень искажений при трансформации обнаруженных и локализованных двумерных кодов в исходный квадратный вид.

Ключевые слова

Компьютерное зрение, графические коды, бинарное изображение, медианный фильтр, метод связанных компонент, преобразование Хафа.

Detection, localization and Transformation of two-Dimensional Graphic Code

A.A. Trubitsyn¹, M.V. Shadrin²

¹ Ryazan State Radio Engineering University named after V.F. Utkin, 59/1 Gagarina str., Ryazan, 390005, Russia

² Kvantron Group LLC, 28a, room H2, Engels str, Ryazan, 390010, Russia

Abstract

The ROI of the article relates to computer vision. Mathematical foundations and algorithms for preparing two-dimensional graphic codes for data decoding are presented. The preparation stages include the detection, localization and transformation of the code to a decoding-friendly form. The brightness oscillation function proposed in the paper serves to reliably detect codes in the image, the Hough transform is the basis for solving the localization problem, and the formulas for the inverse transformation of a square into an arbitrary quadrilateral allow you to transform localized quadrangular areas into squares with a horizontally oriented base. The work of algorithms on a set of test and real images is studied, the practically permissible level of distortion during the transformation of detected and localized two-dimensional codes into the original square form is shown.

Keywords

Computer vision, graphic codes, binary image, median filter, connected component method, Hough transform.

ГрафиКон 2023: 33-я Международная конференция по компьютерной графике и машинному зрению, 19-21 сентября 2023 г., Институт проблем управления им. В.А. Трапезникова Российской академии наук, г. Москва, Россия

EMAIL: assur@bk.ru (А.А. Трубицын); m.shadrin@kvantron.com (М.В. Шадрин)

ORCID: 0000-0002-9337-8947 (А.А. Трубицын); 0009-0009-0472-8553 (М.В. Шадрин)



© 2023 Copyright for this paper by its authors.

Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

1. Введение

При распознавании образов без применения процедур обучения и самообучения решаются следующие задачи:

- обнаружение объекта по набору признаков,
- точная локализация объекта,
- расшифровка информации (например, установление личности человека по изображению лица).

Решения первых двух задач, как ни странно, характеризуются наименьшей степенью разработанности.

В настоящее время на повестке дня стоят задачи обнаружения и локализации фрагментов изображений, в качестве отличительных особенностей которых задана высокая частота колебаний яркости [1]. Такая особенность характеризует многие искусственные объекты человеческой деятельности: предупреждающие разметки на автотрассах, привлекающие внимание баннеры и билборды, наконец, одномерные и двумерные графические код (штрих-коды, Data_Matrix, QR-коды), маркирующие изделия различного назначения. Последние являются примером не столько средств привлечения внимания, сколько удобного способа записи и передачи информации. Роль такого способа доставки информации до заинтересованных потребителей будет только возрастать.

В представляемой работе ставится задача разработки алгоритмов надежного обнаружения и точной локализации фрагментов изображений с высокочастотными пространственными колебаниями интенсивности (яркости), ограниченными на плоскости изображения рамками четырехугольников. Ярким примером таких объектов как раз являются графические коды. Таким образом, при обнаружении и локализации объектов будут использованы два признака – высокая частота пространственных колебаний их яркости и простая четырехугольная геометрическая форма. Поскольку многие алгоритмы декодирования информации требуют определенной стандартной ориентации графических кодов, то после локализации произвольного четырехугольника в случае распознавания двумерного кода возникает необходимость трансформации его в канонический квадрат.

Основная суть предлагаемых методов решения поставленной задачи заключается в использовании строгого математического аппарата без применения эвристических алгоритмов, страдающих сильной неустойчивостью по отношению к вариациям входных данных и особенно к изменению класса этих данных.

2. Обнаружение двумерного графического кода

С целью выделения областей с осцилляцией интенсивности чаще всего вычисляют градиент изображения или некоторые другие со схожими свойствами дифференциальные операторы.

К часто используемым операторам градиента относятся операторы Робертса, Превиттса, Собеля, Кирша, операторы Лапласа и LOG [2]. С учетом времени вычислений и шумоподавления в [3, 4] рекомендуется использовать улучшенный метод ядра Собеля для вычисления градиента и определения направления. К каждому пикселю (x, y) применяются четыре оператора ядра Собеля, представляющие направления $0^\circ, 45^\circ, 90^\circ$ и 135° , и максимальное из четырех значений используется в качестве градиента пикселя (x, y) . Затем градиентное изображение необходимо преобразовать в бинарное.

Однако непосредственное применение методики [3, 4] к исходному изображению не дает нужного эффекта, поскольку не приводит к отчетливому выделению нужной области. Как отмечается в [3] изображение нужно предварительно уменьшить до четверти исходного с помощью билинейной интерполяции, затем преобразовать в серое изображение и, наконец, усилить контрастность изображения. Такая последовательность операций увеличивает временные затраты, не является строго обоснованной и поэтому не может считаться надежной в общем случае. Это связано с тем, что градиент функции по своему физическому смыслу указывает направление максимального возрастания/убывания функции и количественно

характеризует крутизну функции в каждой точке (на разных участках). Поэтому градиент не совсем подходит для обнаружения сильно осциллирующих функций.

В настоящей работе в качестве индикатора осцилляций интенсивности на различных участках бинарного изображения предлагается использовать новую точечную характеристику $v(x, y)$, представляющую собой суммарное количество p_x и p_y скачков интенсивности “black-white” и “white-black” по x и y осям в окрестности каждого пиксела (x, y) бинарного изображения. Назовем двумерную функцию $v(x, y)$ «функцией осцилляций». За окрестность пиксела принимается квадратное окно размером $k \times k$ с центром в текущем пикселе (x, y) (рисунок 1а). Причем допустима оценка $v(x, y)$ в двух нормах: $v(x, y) = \max(p_x, p_y)$ или $v(x, y) = (p_x + p_y)/2$, которые на практике, однако, дают почти одинаковые результаты. На первый взгляд вычисление двух функций $p_x(x, y)$ и $p_y(x, y)$ с помощью скользящего окна по всему исходному бинарному изображению может показаться затратной операцией, однако это не так, поскольку здесь могут быть использованы интегралы изображений, а точнее интегралы предварительно сформированных изображений $i_x(x, y)$ и $i_y(x, y)$. Каждый пиксел этих изображений занимает 1 бит информации, т.е. может принимать значение 0 или 1 в зависимости от того, наблюдается ли относительный скачок интенсивности у соседей справа или слева, тогда $i_x(x, y)=1$, иначе $i_x(x, y)=0$ (рисунок 1б); либо относительный скачок интенсивности у соседей сверху или снизу, тогда $i_y(x, y)=1$, иначе $i_y(x, y)=0$ (рисунок 1в). Еще раз подчеркнем, что исходным для анализа и формирования изображений $i_x(x, y)$ и $i_y(x, y)$ служит бинарное изображение. В примере, представленном на рисунке 1 количество скачков интенсивности $p_x(x, y)=46$ и $p_y(x, y)=2$, поэтому величина функции осцилляций в пикселе (x, y) равна $v(x, y) = \max(p_x, p_y)=46$.

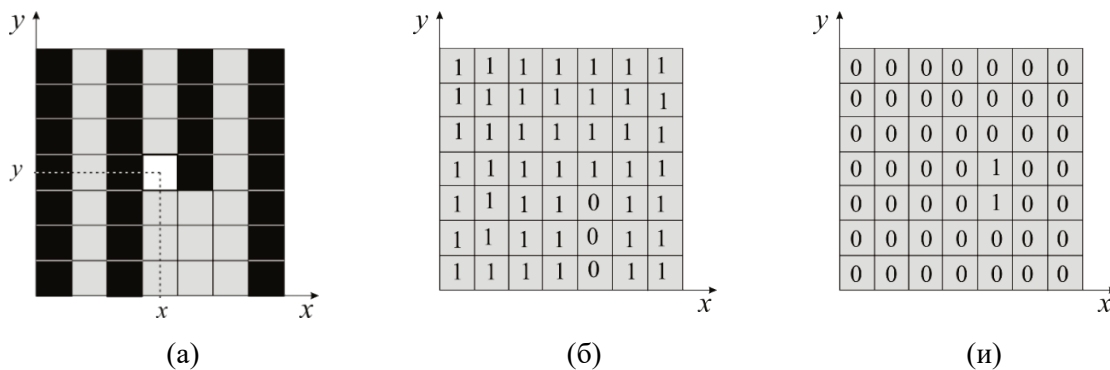


Рисунок 1 – Пример вычисления функции осцилляций: (а) модель фрагмента бинарного изображения в окне размером 7×7 с центральным пикселем (x, y) , (б) функция $i_x(x, y)$ в том же окне, (в) функция $i_y(x, y)$ в том же окне

Как работает функция осцилляций для бинарных изображений демонстрируется рисунком 2. Фрагменты изображения с высокой частотой колебаний интенсивности превращаются в сплошные фигуры при обработке изображения на рисунке 2а функцией осцилляций, а от сплошных одноцветных фигур остаются лишь контуры (рисунок 2б). Бинаризация функции осцилляций лишь подчеркивает указанный факт (рисунок 2в).

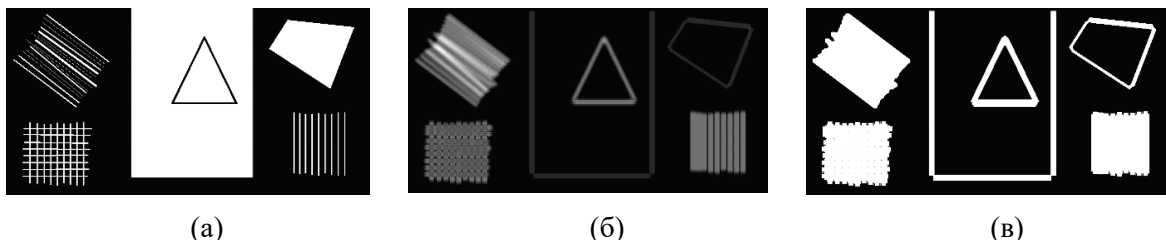


Рисунок 2 – Вычисление функции осцилляций при обработке бинарных изображений: (а) исходное бинарное изображение, (б) двумерная функция осцилляций исходного изображения в оттенках серого, (в) бинаризованная по методу Оцу функция осцилляций

Что касается бинаризации исходного изображения и функции осцилляций, то здесь можно отметить следующее. Поскольку в нашей задаче не акцентируется внимание на мелких деталях

изображения, то выбор падает на пороговые методы бинаризации. В пороговых методах пиксели изображения подразделяются на фоновые и объектные по отношению к некоторому порогу интенсивности. Порог может быть глобальным, вычисляемым по различным характеристикам изображений (средняя интенсивность, гистограмма интенсивностей и т.д.) (см., например, метод Оцу [5]), или локальным, в каждом пикселе величина которого определяется по интенсивности пикселей окружения [6, 7, 8]. Отдельного места в ряду этих методов заслуживают появившиеся сравнительно недавно улучшенные локальные методы, где порог в каждом пикселе определяется на основе анализа его яркости, скользящих среднего и дисперсии [9].

3. Локализация двумерного графического кода

После бинаризации функции осцилляций областью интереса становится набор сплошных белых областей на черном фоне (рисунок 2в). Для пространственной локализации этих обнаруженных областей предлагается следующая последовательность алгоритмов.

1) Маркирование отдельных областей бинарного образа функции осцилляций методом связанных компонент. Для связывания компонент традиционно используют рекурсивные методы закрасивания областей (компонент), многопроходные [10] и двухпроходные [11] методы, в которых каждому пикселу в итоге присваивается лейбл принадлежности отдельной компоненте изображения.

2) Выделение граничного контура каждой компоненты. Для изображений в градациях серого выделение граничного контура является целой научной проблемой [12, 13]. Однако для изображений с отдельными монохромными областями такие контура могут быть построены последовательным сканированием по строкам и столбцам.

3) Преобразование Хафа массива граничных пикселей каждой компоненты. Задача построения многоугольника, локализирующего обнаруженную область по ее контуру, может быть решена с помощью преобразования Хафа [14], которое позволяет обнаружить отрезки прямых линий на изображениях в координатах ρ и θ . Параметр ρ есть длина нормали к прямой, проведенной из начала координат, а θ – угол между этой нормалью и осью абсцисс. В преобразовании Хафа количество пикселей, лежащих на одну прямую, хранится в ячейке двумерного массива, называемого аккумулятором. Чем больше таких пикселей, тем больше значение накапливается в соответствующей ячейке аккумулятора. Поэтому обнаружение k ярко выраженных прямых, проходящих через сравнительно большее количество пикселей на плоскости, сводится к обнаружению первых k локальных максимумов аккумулятора. В этом заключена суть алгоритма обнаружения прямых на основе преобразования Хафа.

4) Поиск первых k локальных максимумов дискретной двумерной функции. Для поиска первых k -локальных максимумов предлагается следующий алгоритм. Первый максимум находится как глобальный максимум функции. Затем организуется вычислительный цикл, в котором индикаторный уровень опускается на 1 до появления следующего максимума. Определение каждого следующего после 1-го максимума осуществляется с помощью проверки трех условий: превышение индикаторного уровня, превышение по величине ближайших соседей и отличие на некоторый интервал углов по расположению от всех ранее найденных максимумов.

5) Определение вершин локализирующего k -угольника путем перехода из пространства Хафа в декартовую систему координат.

На рисунке 3 показаны этапы обнаружения и локализации трех QR-кодов, размещенных на гранях куба (рисунок 3а): 1-й шаг – пороговая бинаризация исходного изображения методом Оцу (рисунок 3б), 2-й шаг – вычисление двумерной функции осцилляций (рисунок 3в), 3-й шаг – пороговая бинаризация по среднему уровню (рисунок 3г), 4-й шаг – выделение связанных компонент I, II и III (рисунок 3д), 5-й шаг – выделение внешних контуров связанных компонент I, II и III (рисунок 3е), 6-й шаг – преобразование Хафа каждой из связанных компонент (рисунок 3ж, где показана зависимость (синограмма) интенсивности $z_{ij} = z(\rho_i, \theta_j)$ и четыре первых максимума, пронумерованные по возрастанию угла θ для контура связанной компоненты I, рисунок 3з, где представлена трехмерная зависимость интенсивности $z_{ij} = z(\rho_i, \theta_j)$), 7-й шаг – локализация областей I, II и III четырехугольниками (рисунок 3и).

Отметим, что результаты получены после предварительного подавления импульсных шумов на исходном изображении переключающим медианным фильтром, который исключает не только одиночные импульсы, но и многопиксельные [15].

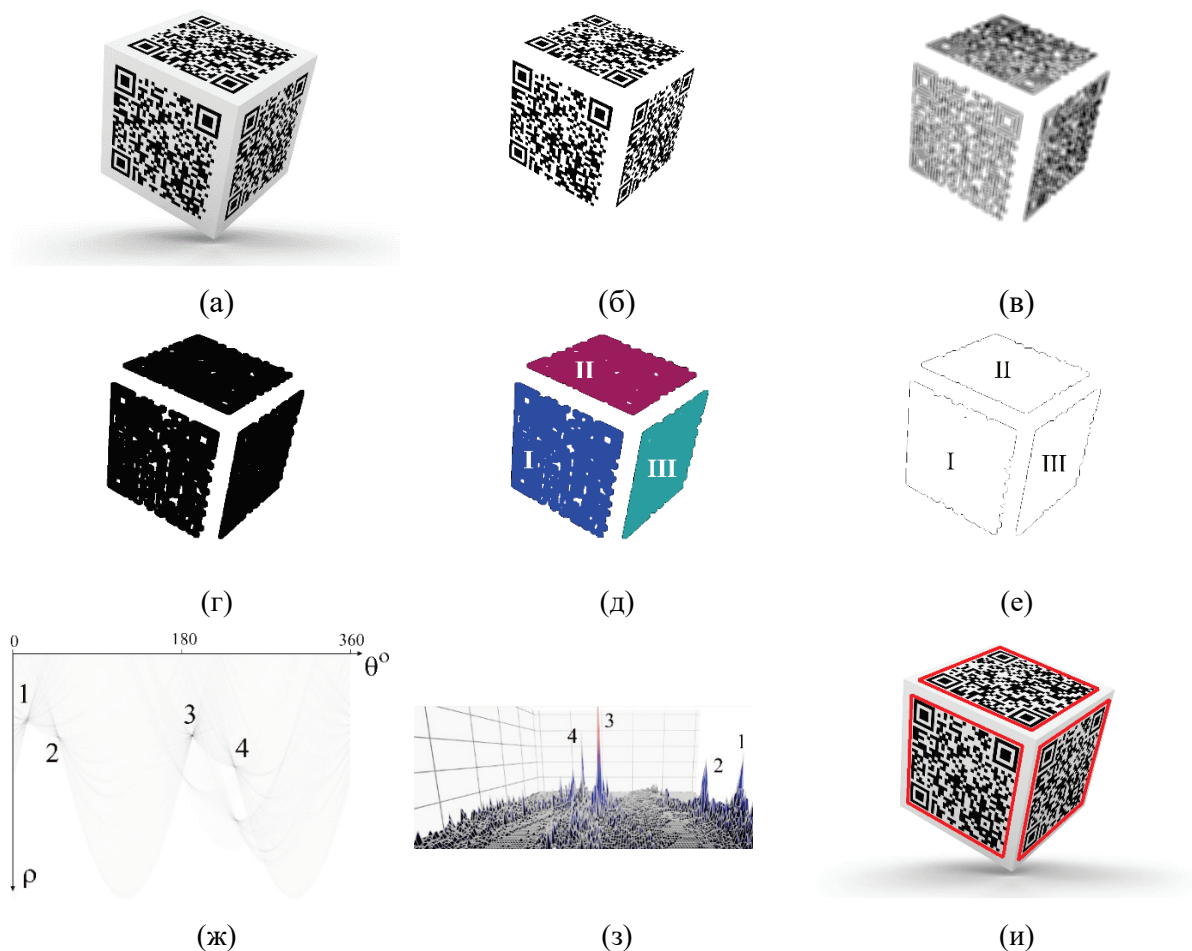


Рисунок 3 – Обнаружение и локализация трех QR-кодов: (а) исходное изображение, (б) бинаризованное методом Оцу изображение, (в) функция осцилляций интенсивности, (г) бинаризованная по среднему уровню функция осцилляций интенсивности, (д) связанные компоненты I, II и III, (е) контуры связанных компонент I, II и III, (ж) синограмма контура I компоненты, (з) преобразование Хафа внешнего контура I компоненты, (и) выделенные красными четырехугольниками QR-коды

4. Трансформация произвольного четырехугольника в канонический квадрат

Для трансформации локализованного четырехугольника ($1'2'3'4'$) в квадрат (1234) использовалось преобразование (1), предложенное в работе [16]. Проведенные исследования показали, что наименьшие искажения наблюдаются в системе координат, связанной с центром четырехугольника. Формулы преобразования и демонстрирующий графический материал (рисунок 4) представляют случай именно такой системы координат. За координаты центра 0 четырехугольника приняты среднеарифметические значения координат его вершин.

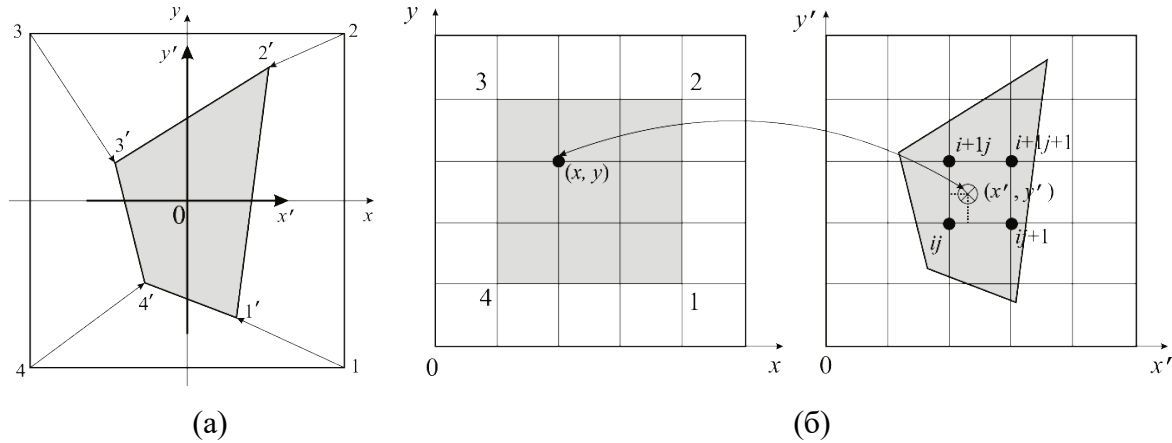


Рисунок 4 – Преобразование четырехугольника (1'2'3'') в квадрат (1234): (а) схема обратного преобразования, (б) схема вычисления интенсивности пикселей квадрата

Формулы преобразования для каждой из четырех вершин имеют вид

$$\begin{aligned}
 x'_k &= c_1 x_k + c_2 y_k + c_3 x_k y_k + c_4, \\
 y'_k &= c_5 x_k + c_6 y_k + c_7 x_k y_k + c_8, \\
 k &= 1, 2, 3, 4.
 \end{aligned}
 \tag{1}$$

Поскольку полное число уравнений для четырех преобразуемых вершин равно восьми, то получаем замкнутую систему уравнений относительно восьми неизвестных коэффициентов $c_i, i=1,2 \dots 8$. Решение системы удастся получить в виде последовательности достаточно простых выражений

$$\begin{aligned}
 c_1 &= \frac{c_{15}^+ + c_{15}^-}{2}, c_5 = \frac{c_{15}^+ - c_{15}^-}{2}, c_2 = \frac{c_{26}^+ + c_{26}^-}{2}, c_6 = \frac{c_{26}^+ - c_{26}^-}{2}, \\
 c_3 &= \frac{c_{37}^+ + c_{37}^-}{2}, c_7 = \frac{c_{37}^+ - c_{37}^-}{2}, c_4 = \frac{c_{48}^+ + c_{48}^-}{2}, c_8 = \frac{c_{48}^+ - c_{48}^-}{2}, \\
 c_{15}^\pm &= \frac{b_1^\pm a_{11} - b_2^\pm a_{01}}{a_{00} a_{11} - a_{10} a_{01}}, c_{26}^\pm = \frac{b_1^\pm - a_{10} c_{15}^\pm}{a_{01}}, \\
 c_{37}^\pm &= \frac{s_1^\pm - s_2^\pm - (x_1 - 2)c_{15}^\pm - (y_1 - y_2)c_{26}^\pm}{g_{01}}, c_{48}^\pm = s_1^\pm - c_{15}^\pm x_1 - c_{26}^\pm y_1 - c_{37}^\pm x_1 y_1, \\
 b_k^\pm &= (s_k^\pm - s_{k+1}^\pm)g_{k+1,k+2} - (s_{k+1}^\pm - s_{k+2}^\pm)g_{k,k+1}, k = 1, 2, \\
 a_{00} &= (x_1 - x_2)g_{23} - (x_2 - x_3)g_{12}, a_{01} = (y_1 - y_2)g_{23} - (y_2 - y_3)g_{12}, \\
 a_{10} &= (x_2 - x_3)g_{34} - (x_3 - x_4)g_{23}, a_{11} = (y_2 - y_3)g_{34} - (y_3 - y_4)g_{23}, \\
 g_{i,i+1} &= x_i y_i - x_{i+1} y_{i+1}, i = 1, 2, 3, \\
 s_k^\pm &= x'_k \pm y'_k, k = 1, 2, 3, 4.
 \end{aligned}$$

С целью минимизации яркостных искажений и исключения появления пустых пикселей на изображении квадрата при трансформации четырехугольника формулы вида (1) представляются в виде обратного преобразования, т.е. квадрата в четырехугольник (рисунок 4а). В рамках такого подхода центр получаемого квадрата совмещается с центром исходного четырехугольника, причем квадрат полностью покрывает этот четырехугольник. В таком случае сторона квадрата может иметь в достаточной мере произвольную длину, лишь бы вершины квадрата оказались в пределах области монитора. После фиксации вершин квадрата определяются коэффициенты $c_1 - c_8$. При построении изображения квадрата целочисленные координаты его пикселей пробегают весь выбранный диапазон по x и y (рисунок 4б), а полученные (нецелочисленные) координаты

x' и y' (1) позволяют вычислить интенсивность в точке (x', y') (в исходной системе координат) по формулам двумерной интерполяции, например, билинейной, с использованием значений интенсивности ближайших четырех пикселей (i, j) , $(i+1, j)$, $(i, j+1)$ и $(i+1, j+1)$ исходного изображения. Вычисленная таким образом интенсивность точки (x', y') транслируется в интенсивность пиксела (x, y) на изображении квадрата (см. рисунок 4б).

Восстановление тестовых квадратных изображений, полученных при съемке под разными углами, показало практическое отсутствие искажений в диапазоне углов съемки от 0° до 40° - 50° по отношению к нормали. Еще большее увеличение угла съемки приводит к искажениям, затрудняющих дальнейший анализ двумерных кодов.

На рисунок 5 представлен пример трансформации локализованных QR-кодов (см. рисунок 3и) в канонические квадраты.

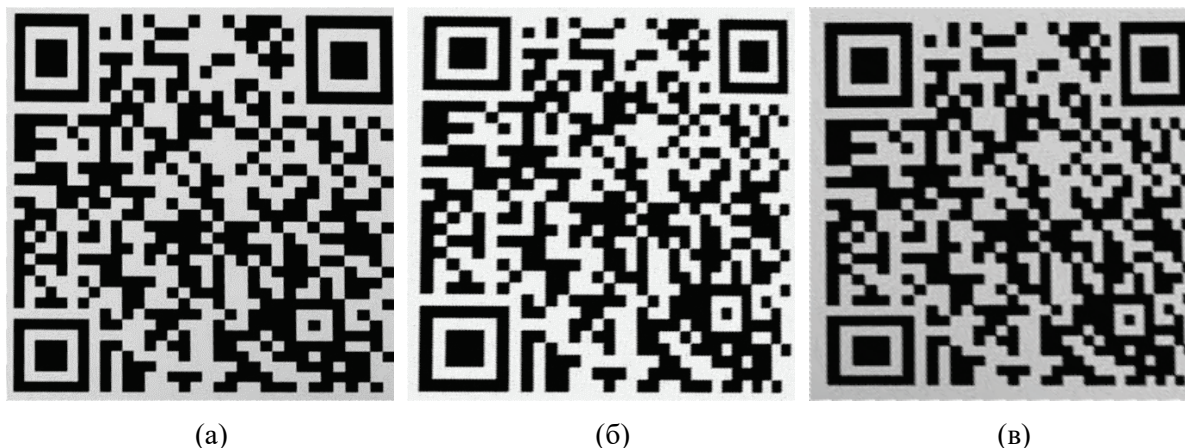


Рисунок 5 – Трансформация QR-кодов, расположенных на гранях I (а), II (б) и III (в) куба на рисунке 3 и, поэтому зарегистрированных при разных углах съемки, в канонические квадраты

Проведенный в рамках настоящих исследований анализ данных трансформации показал, что QR-коды, восстановленные до канонического квадрата, вполне пригодны для дальнейшей обработки с целью раскодирования. Визуальный сравнительный анализ рисунков 5а, 5б и 5в подтверждает сделанный вывод.

5. Заключение

Предложенная в работе функция осцилляций позволяет обнаружить фрагменты изображений, характеризующиеся высокочастотными колебаниями интенсивности, примером которых являются двумерные графические коды. Метод связанных компонент и преобразование Хафа являются средством локализации обнаруженных кодов в пределах прямолинейных границ четырехугольных областей. Произвольные четырехугольные области с малым уровнем искажений могут быть трансформированы в канонические квадраты по формулам преобразования (1).

6. СПИСОК ИСТОЧНИКОВ

- [1] Trubitsyn A.A., Shadrin M.V, Serezhin A.A. Localization of image fragments with high frequency intensity oscillation // Journal of Autonomous Intelligence .2023. Vol 6. No 2. P.1-16.
- [2] Burger W, Burge M. Digital Image Processing. London: Springer-Verlag London Ltd., 2016. 811 p.
- [3] Fang L, Xie C. 1-D barcode localization in complex background // International Conference on Computational Intelligence and Software Engineering (Wuhan, China, Dec 10-12, 2010) Proceedings of the IEEE / IEEE, 2010. P. 1-3.

- [4] Santosa F., Goh M. Bar Code Decoding in a Camera-Based Scanner: Analysis and Algorithm // *SIAM Journal on Imaging Sciences*. 2022. Vol 15. No 3. P. 1017-1040.
- [5] Otsu N. A Threshold Selection Method from Gray-Level Histograms // *IEEE Trans. Sys. Man. Cyber.* 1979. Vol 9. No 1. P. 62-66.
- [6] Bradley D, Roth G. Adaptive thresholding using the integral image // *Journal of Graphics Tools*. 2007. Vol 12. No 2. P. 13- 21.
- [7] Niblack W. An introduction to Digital Image Processing. Englewood Cliffs, N.J: Prentice-Hall International, 1986. 215 p.
- [8] Sauvola J, Pietikainen M. Adaptive document image binarization // *Pattern Recognition*. 2000. Vol 33. No 2. P. 225-236.
- [9] Koleda P, Hrková M. Global and Local Thresholding Techniques for Sawdust Analysis // *Acta Facultatis Technicae*. 2018. Vol XXIII. No 1. P. 33-42.
- [10] R.M. Haralick, Some Neighborhood Operations. *In Real Time/Parallel Computing Image Analysis*. NY: Plenum Press, 1981. P. 11-35.
- [11] Rosenfeld A., Pfaltz P. Sequential Operations in Digital Picture Processing // *Journal of the Association for Computing Machinery*. 1966. Vol. 12. P. 471-494.
- [12] Canny J. A Computational Approach to Edge Detection // *IEEE Trans. on Pattern Analysis and Machine Intelligence*. 1986. Vol 8. No 6. P. 679-698.
- [13] Bradski G, Kaehler A. *Learning OpenCV : computer vision with the OpenCV library*. Sebastopol, CA: O'Reilly. 2008. 555 p.
- [14] Hough PVC. Method and means for recognizing complex patterns. U.S. Patent 3,069,654. Dec. 18, 1962.
- [15] Trubitsyn A, Grachev E. Switching median filter for suppressing multi-pixel impulse noise. *Computer Optics*. 2021. Vol 45. No 4: P. 580–588.
- [16] Gonzalez R, Rafael R. *Digital image processing*. NY: Pearson. 2018. 1168 p.