DOI: 10.14529/cmse240402

PRECISE LOCALIZATION OF PDF417 CODE BASED ON FAST HOUGH TRANSFORM

© 2024 D.G. Mitrofanov[1,2], P.K. Zlobin[1], J.A. Shemiakina[1], P.V. Bezmaternykh[1,3]

[1]*Smart Engines Service LLC (pr. 60-letiya Oktyabrya 9, Moscow, 117312 Russia),*

[2]*Lomonosov Moscow State University (GSP-1, Leninskie Gory 1, Moscow, 119991 Russia),*

[3]*FRC "Computer Science and Control" RAS (st. Vavilova 44, bld. 2, Moscow, 119333 Russia)*

E-mail: mitrofanov.dmitryi@gmail.com, zlobin.pavel.k@gmail.com, jshemiakina@smartengines.com, bezmaternyh@isa.ru

Received: 30.09.2024

The PDF417 is a popular barcode symbology which is widely used in a huge variety of business processes. In this paper, we propose an original method for precise PDF417 code localization. It can successfully process projectively distorted images captured via the mobile device cameras. The core of this method is the analysis of the Fast Hough Transform image. This analysis is aimed to: (a) determine the line, corresponding to the vanish point of vertical symbol sides, using the RANSAC algorithm; (b) select the best pair of Hough-points corresponding to the horizontal symbol sides. We also propose the evaluation methodology for assessing the accuracy of precise PDF417 localization and a new dataset SE-PDF417-SYN-400, which consists of 400 synthesized PDF417 images and is publicly available. The accuracy of the proposed method on SE-PDF417-SYN-400 is equal to 0.948, and its error rate is about four times less than the one obtained by the popular ZXing detector. The average running times on iPhone 8 and iPhone 14 Pro Max mobile devices are equal to 77 and 34 ms per image correspondingly.

Keywords: barcode reading, PDF417, Fast Hough Transform, vanish point, RANSAC.

FOR CITATION

Mitrofanov D.G., Zlobin P.K., Shemiakina J.A., Bezmaternykh P.V. Precise Localization of PDF417 Code Based on Fast Hough Transform. Bulletin of the South Ural State University. Series: Computational Mathematics and Software Engineering. 2024. Vol. 13, no. 4. P. 19–34. DOI: 10.14529/cmse240402.

Introduction

Fast and reliable data entry is essential for a huge variety of business processes including identity document (ID) verification and boarding passes validation. The manual approach is very tedious and error-prone, so this is precisely the task where machines could replace humans. To simplify this task, special machine-readable graphical symbols named barcodes were introduced in the middle of the previous century.

Generally, they can be divided into three groups: one-dimensional (1D), stacked, and two-dimensional (2D) codes. The 1D codes are represented as a set of parallel dark thick and thin bars on light background. Typical examples of these types are UPC/EAN, Code 128, Interleaved 2 of 5. Stacked codes may be considered as a set of 1D codes placed on top of each other with some extra meta information bars. The PDF417 symbology [1] is the most representative example of such codes and is the subject of this paper. It is widely used in ID documents issued by American Association of Motor Vehicle Administrators (AAMVA) member jurisdictions and in paper boarding passes supported by the International Air Transport Association (IATA). Samples of these documents are presented in Fig. 1. A general structure of PDF417 code is shown in Fig. 2.

a) The AAMVA sample

b) The IATA sample

Fig 1. The expected PDF417 localization results (in red) for samples taken from the open sources

Nowadays, these codes are commonly captured via mobile device cameras in uncontrolled conditions. In such circumstances, captured barcode images are regularly distorted by perspective transformation [2].

Fig 2. The PDF417 barcode structure. Image is taken from the open source

First goal in the barcode recognition field is to roughly determine the region of interest (ROI) containing the image of a symbol. There are plenty of such detectors, most of them nowadays are based on some supervised learning approaches [3]. The next problem is the precise localization of the code symbol. That usually means finding a quadrangle of symbol corner points further denoted as $Q = \langle a, b, c, d \rangle$ (see red quadrangles in Fig. 1). Let us denote this problem as $\mathcal{P}$. In this paper, a new method for such barcode corners detection problem is presented. By design, it may tackle with perspective distortion of symbols images and is suitable for the usage on mobile devices. Like some other methods, it is based on the usage of the discrete integral image transformation. A typical choice in this case is the famous **H**ough **T**ransform (HT). To overcome its computational "bloating", we propose using its fast discrete approximation, known as **F**ast **H**ough **T**ransform (FHT) [4, 5]. The usage of such approximation is not trivial, so we do provide in-depth description of the resulting FHT image analysis specific for the problem under consideration. We also use RANSAC scheme to estimate bar lines directions that allows to deal with extra lines when the PDF417 symbol is placed on the complicated background.

The contributions of the paper can be formulated as follows. First, we propose the new method for precise PDF417 codes localization. It is based on FHT image analysis and RANSAC scheme. Second, we introduce a way to evaluate the accuracy of such precise localization. Finally, we present a new dataset SE-PDF417-SYN-400 suitable for this problem.

The rest of this paper is organized as follows. In Section 1, we review the related works. Then, in Section 2 we discuss the details of the proposed method. The experiment methodology and results are presented and analyzed in Section 3. The Conclusion concludes this work.

1. Related Works

The barcode localization problem $\mathcal{P}$ implies that we need to detect the position of symbol corners Q within an input image I. Currently, there is a wealth of approaches to deal with this problem. Some of these approaches are not barcode-type specific [6], and some of them hugely rely on the internal barcode structure and so-called "finder patterns". At the same time, these approaches may impose different constraints on the input image. Some of them require the presence of only one target symbol per image, others request the estimation of symbol size in advance, and major part expect a binary image as an input instead of a grayscale one.

A common PDF417 barcode processing pipeline $\mathbb{P}$ contains the following stages [7, 8]: (a) ROI extraction; (b) image preprocessing; (c) symbol localization; (d) symbol segmentation; (e) message decoding. The subject of interest in this paper is the combination of stages (b) and (c).

One of the most popular steps in stage (b) is an application of some binarization procedure to the input image [9]. It simplifies the task of further image analysis and increases the overall runtime performance. By design, barcode symbol elements are black and white, and the usage of binarization step seems to be inherent. Nevertheless, in uncontrolled environment the complexity of barcode binarization task drastically increases due to uneven lighting conditions, presence of shadows or reflections. Moreover, binarization methods may be sensitive to noise and can severely damage some barcode patterns making their further decoding impossible [10]. Today, there is no generally accepted method for barcode binarization. Some special, "symbology oriented" binarization methods, including ones for PDF417, have already been published [11, 12]. But their usage generally does not make much difference.

Another conventional image preprocessing technique for PDF417 detection is mathematical morphology (MM). The application of properly selected MM filters can enhance the barcode image and simplify the task of edge detection or the stage of symbol segmentation [8, 13, 14]. The outcome of MM filter depends on the shape and size of the structured element (SE) being used. If this size is too small, over-segmentation occurs, and vice versa. This SE size must be in accordance with the barcode module size, which can vary a lot in presence of perspective transformation. Therefore, the size of SE can not be set in advance in general case. In paper [15], authors propose to iterate through the set of various structural elements and return the best result. These iterative strategies suffer from significantly increased method running times, which makes them undesirable for the mobile devices.

There are localization methods which detect the “start” and the “stop” patterns of the symbol [7] (see Fig. 2). The detection of such patterns commonly includes some template matching strategy. This step is especially dependent on binarization accuracy. The stripes in these patterns can be very thin and, thus, can be easily corrupted by binarization procedure. Having detected precise positions of these patterns we can estimate the Q of the whole barcode. Another approach is to extract the contours from the image and analyze them [7, 16]. The contours of “start” and “stop” patterns can drastically simplify this process.

The Hough transform is a commonly used tool in the barcode localization problem. The example of HT application for 1D codes detection is presented in papers [8, 17]. In the paper [18], a promising method for localizing general matrix codes was proposed. It is based on the special modification of HT and its analysis, but unlike our method, it requires high-resolution images. As for PDF417 codes, the paper [19] introduces the method of their localization and segmentation based on the analysis of HT image. Unfortunately, in that work the problem of extra elements placed near the PDF417 code symbol is not addressed.

The goal of analysis is to determine lines with strong edges corresponding to code’s columns and estimate their vanish point. These lines are expected to be corresponded to the maxima of HT image and the vanish point corresponds to a line on HT image passing through these maxima. The naive HT calculation is notoriously slow. In this paper, a new method of code localization based on FHT analysis instead of HT is presented. This method does not require preliminary binarization step.

Thus, in the literature we found three groups of approaches to deal with $\mathcal{P}$ problem, which are based on: (a) preliminary image processing with mathematical morphology; (b) special patterns detection; (c) HT image analysis.

Unfortunately, at this point there some problems which preclude objective comparison of methods proposed in listed papers. Some of these works, like [19], does not provide any information about resulting performance or accuracy. Another do assess their methods over private datasets, which are commonly very small. Moreover, the evaluation methodology for the $\mathcal{P}$ problem is not established. The commercial barcode readers that promote themselves as effective tools for PDF417 recognition focus only on the end-to-end metric, i.e. the number of correctly read codes. They do not provide any details about their internal subsystems and their performance. Therefore, these solutions can not be used to evaluate the $\mathcal{P}$ problem. In order to overcome all these issues we propose: (a) the $\mathcal{P}$ problem oriented dataset; (b) the evaluation methodology; (c) the baseline.

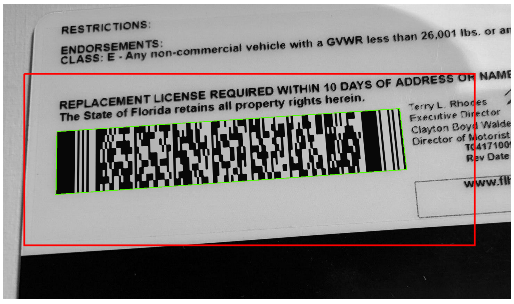


Fig 3. The image shows two examples: a red quadrangle indicating a ROI and a green quadrangle demonstrating the result of localization

2. Proposed Algorithm

Let $I_{W\times H}$ be the region of interest containing the whole PDF417 symbol $S_{R\times C}$, where R and C denotes the number of rows and columns respectively. It is assumed that $I_{W\times H}$ is a grayscale image with certain constraints: (a) there is a gap between the symbol and the input image bounds (at least 5 pixels); (b) the center of image $I_{W\times H}$ lies within the symbol S; (c) the area of S occupies at least quarter of the image area. The location of the symbol is determined as an ordered quadrangle of code corners $Q = \langle a, b, c, d\rangle$. Thus, our goal is to construct an algorithm $\mathcal{A}_{full}$, which is able to detect such a quadrangle Q within the image $I_{W\times H}$ (see Fig. 1).

The target quadrangle Q may be defined as intersection of four lines which lay on symbol sides. At this point we assume that the code is oriented "mainly horizontally": the total length of parallel projection of sides (a, b) and (c, d) on axis X is bigger than on axis Y (Fig. 1). Let us denote the pairs of horizontal and vertical sides as $\mathtt{S}_h = \{(a, b), (c, d)\}$ and $\mathtt{S}_v = \{(b, c), (a, d)\}$ respectively. In order to find these pairs, we propose to analyze the FHT image of $I_{W\times H}$. Let denote this image as FHT_{full}.

At first stage, the "strongest" points are detected in FHT_{full}. Having found these points, the corresponding sides are determined using two completely different approaches, namely $\mathcal{A}_H$ and $\mathcal{A}_V$. Now, let us describe these algorithms in depth.

Stage 1: *"strongest" points extraction.* To reduce the running time and simplify the process of parameter fine-tuning, we scale input image $I_{W\times H}$ in J. Both its sides are scaled equally, and the scale factor f is calculated by Equation (1):

$$f = \frac{D}{\max(W, H)}. \tag{1}$$

In this work, the value of D is set to 800.

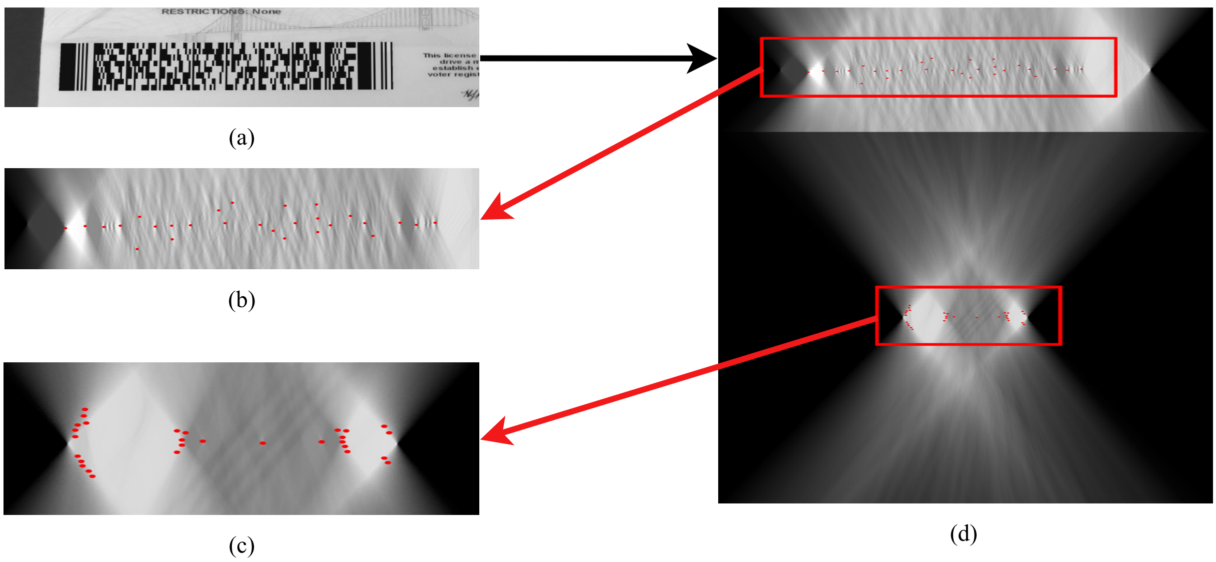


Fig 4. The "strongest" points $\langle P_h, P_v\rangle$ for the source image. (a) Source image I; (b) detected P_h points; (c) detected P_v points; (d) Sets $\langle P_h, P_v\rangle$ on the FHT image

Then, the image $FHT_{full}(J) = FHT_h(J) \cup FHT_v(J)$ is calculated. Here, FHT_h and FHT_v parts correspond to horizontal and vertical lines, respectively (Fig. 4), and the sign '$\cup$' denotes the procedure of their concatenation. This procedure is described in details in [20]. After that, the local maxima M of FHT_{full} image are calculated. Then, they are sorted by intensity and are examined from the "highest" to the "lowest" ones. For every such element e, all other maxima in window of size W_{size} centered in e are suppressed, thus filtering out some maxima. The size W_{size} is set to 17 pixels in this work. After that, the sets of brightest points P_h and P_v from Hough spaces FHT_h and FHT_v respectively are selected (Fig. 4) and sorted by their intensity. Further, all but the first N_h is set to 50 and N_v is set to 30 points are removed from consideration and the "survived" ones form the target pair of sets $\langle P_h, P_v \rangle$. The whole procedure of the proposed method $\mathcal{A}_{STP}$ of $\langle P_h, P_v \rangle$ extraction is summarized in Algorithm 1 and its application for a sample image is demonstrated in Fig. 4.

Algorithm 1: "strongest" points $\langle P_h, P_v \rangle$ extraction

Data: grayscale image $I_{W \times H}, D > 0, W_{size} > 0, N_h > 0, N_v > 0$
Result: $\langle P_h, P_v \rangle$

$J \leftarrow ScaleImage(I_{W \times H}, D)$;
$FHT_{full} \leftarrow FastHoughTransform(J)$;
$M \leftarrow ExtractLocalMaxima(FHT_{full})$;
$S \leftarrow NonMaximaSuppression(M, W_{size})$;
$\langle P_h, P_v \rangle \leftarrow TakeBestMaxima(S, N_h, N_v)$;
$return\ \langle P_h, P_v \rangle$;

Stage 2: *horizontal lines detection.* The goal of this algorithm $\mathcal{A}_H$ is to choose the best pair of "strongest" points $\mathtt{R} = \langle R_l, R_u \rangle$ and, therefore, determine a pair of corresponding symbol code sides $\mathtt{S}_h = \langle H_l, H_u \rangle$.

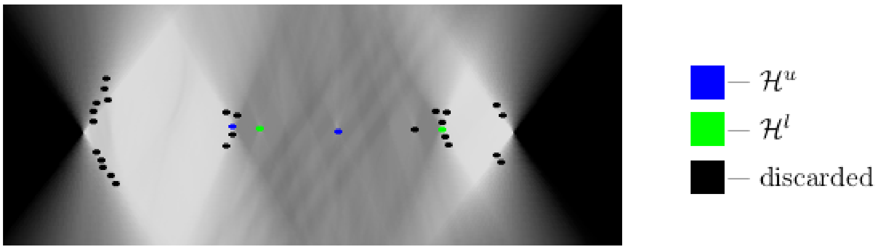
Let us consider the surrounding area of the upper side H_u of the symbol on source image $I_{W \times H}$. There should be predominantly "white" pixels above it and predominantly "black" pixels below it. As for the lower side H_l, the intensity of pixels are arranged in the opposite way (see Fig. 2). Relying on this observation, we divide the points P_h (obtained on the previous stage) into three sets: (a) $\mathcal{H}^u$ — candidate points for H_u; (b) $\mathcal{H}^l$ — candidate points for H_l; (c) the other ones. The third set of points is discarded from further consideration. Let us generate the set of all possible pairs $C_h = \mathcal{H}^u \times \mathcal{H}^l$. On FHT_h image the area corresponding to the barcode symbol should not contain major intensity changes. So, the pairs of points forming a non-monotonous segment are filtered out. To do this, the third central statistical moment M_3 of the segment is calculated. Then, the value $|M_3|$ is compared with the threshold T_M. If it is greater, than the pair is filtered out.

Resulting set is denoted as C_f. Among its elements, the pair $\langle R_l, R_u \rangle = \max\limits_{\langle i,j \rangle \in C_f} (intensity(i) + intensity(j))$, $i \in \mathcal{H}^l, j \in \mathcal{H}^u$ is chosen as the best one. Thus, the pair $\mathtt{S_h} = \langle H_l, H_u \rangle$ can be restored. The algorithm of obtaining a straight line from a point on FHT image is shown in [20].

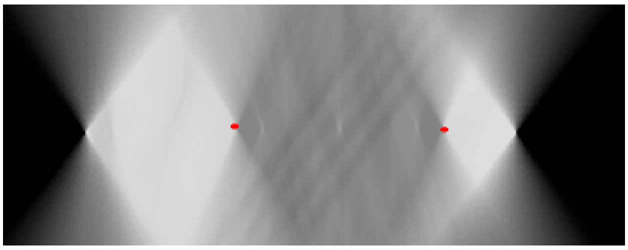
Algorithm 2: horizontal sides $\mathtt{S}_h$ calculation

Data: grayscale image $FHT_h, P_h, T_M > 0$
Result: $\langle H_l, H_u \rangle$

$\mathcal{H}^l, \mathcal{H}^u \leftarrow SplitPoints(FHT_h, P_h)$;
$C_h \leftarrow CombinePairs(\mathcal{H}^l, \mathcal{H}^u)$;
$C_f \leftarrow FilterPairs(C_h)$;
$\mathtt{R} \leftarrow SelectBestPair(C_f)$;
$\langle H_l, H_u \rangle \leftarrow RestoreSides(\mathtt{R})$;
$return\ \langle H_l, H_u \rangle$;



a) Classification of the “strongest” Hough points



b) The best pair of points

Fig 5. Example of horizontal sides selection

Stage 3: *vertical lines detection.* Now we need to find $\mathtt{S}_v = \langle V_l, V_r \rangle$ — the external vertical sides of the symbol. But this time, we expect to have the similar elements inside the code which separate the columns of PDF417 symbol. Such elements can be observed in Fig. 2. Thus, we need to choose another approach for their detection. The lines corresponding to the edges of guard patterns and internal edges share the same vanish point vp. Let denote the image of vp in FHT_v space as l. This image l is represented as a straight line. We expect that the major part of points from P_v lies on this line. The other part of the points is considered as “outliers”. To robustly determine the parameters of line l, we use the well-known RANSAC method [21]. The intensity profile P_l corresponding to detected line l is extracted from image FHT_v. Such profile is presented in Fig. 6. The next step is to find the PDF417 guard patterns in this profile. This can be easily done by matching the known in advance pattern [1] with the given profile. After that, the external edges of such patterns are taken as the symbol sides $\mathtt{S}_v$. The full algorithm $\mathcal{A}_V$ is presented in Algorithm 3.

Algorithm 3: vertical sides L_v calculation

Data: grayscale image FHT_v, P_v

Result: $\langle V_l, V_r \rangle$

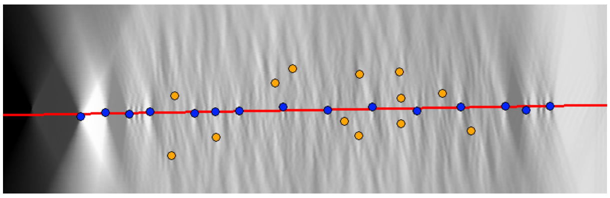
$l \leftarrow RANSAC(FHT_v, P_v)$;

$P_l \leftarrow CalculateProfile(l)$;

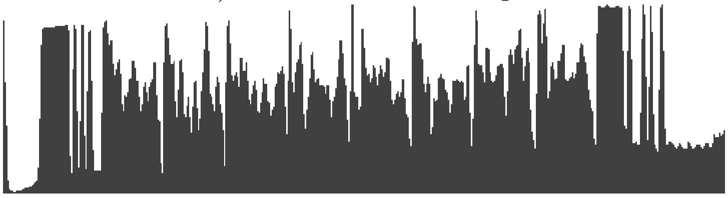
$\langle G_l, G_r \rangle \leftarrow FindGuardPatterns(P_L)$;

$\langle V_l, V_r \rangle \leftarrow GetExternalEdges(G_l, G_r)$;

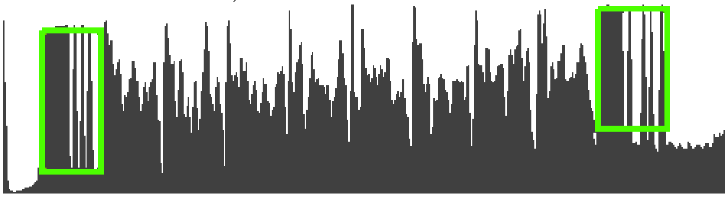
return $\langle V_l, V_r \rangle$;



a) Line l on FHT_v image



b) Profile for the line l



c) Found guard patterns

Fig 6. Example of vertical sides calculation

Stage 4: *symbol quadrangle restoration*. The final step is to restore the required quadrangle $Q = \langle a, b, c, d \rangle$. It can be calculated as cross product of detected sides: $a = V_l \times H_u$, $b = V_r \times H_u$, $c = V_r \times H_l$, $d = V_l \times H_l$, where the sign '$\times$' denotes the intersection of lines.

The whole algorithm $\mathcal{A}_{full}$ for Q detection is outlined in Algorithm 4.

Algorithm 4: restoration of the symbol quadrangle Q

Data: grayscale image $I_{W\times H}, D > 0, W_{size} > 0, N_h > 0, N_v > 0$

Result: Q

$(P_v, P_h) \leftarrow \mathcal{A}_{STP}(I_{W\times H}, D, W_{size}, N_h, N_v)$;

$\mathtt{S}_h \leftarrow \mathcal{A}_H(FHT_h, P_h)$;

$\mathtt{S}_v \leftarrow \mathcal{A}_V(FHT_v, P_v)$;

$return\ IntersectionOfLines(\mathtt{S}_h, \mathtt{S}_v)$;

Fig 7. The localization result Q

3. Experiments

3.1. Evaluation Methodology

Let us consider a quadrangle Q found by algorithm $\mathcal{A}_{full}$. We need to define whether this quadrangle is correctly found or not. In order to do it, we have to know in advance the expected "ground truth" result $G_I = \langle Q_I, R_I, C_I \rangle$. Here, Q_I stands for a quadrangle of symbol location within the image $I_{W \times H}$; R_I and C_I denote the expected numbers of symbol rows and columns respectively.

In this work, the evaluation is interpreted as a binary value $B(Q, G) \in \{0, 1\}$. Here, one stands for correct answers, zero — otherwise. We propose to calculate this binary value $B(Q, G)$ as comparison of some quadrangle score $E(Q, G) \in \mathbb{R}^+$ with the predefined threshold T: $B(Q, G) = [E(Q, G) < T]$. Here, $[\cdot]$ denotes Iverson notation; the value $T = 0.017$, it is chosen experimentally.

Now, let us define the way of $E(Q, G)$ calculation. At first, we need to introduce the rectified quadrangle TQ with sizes $\mathcal{T}_W$ and $\mathcal{T}_H$. Then, $\mathcal{T}_W = 17 \cdot C_I$, and $\mathcal{T}_H = 3 \cdot R_I$. The left-top of TQ corner has $(0, 0)$ coordinates. Let the height of a module be $3px$, and the width of a module be $1px$. Now let calculate the parameters of projective transform H, which maps rectangle TQ into quadrangle Q: $RQ = H(Q)$. Let us calculate the maximum L_2 distance between corresponding corners of TQ and RQ. This value normalized by the perimeter of TQ is taken as evaluation value $E(Q)$. The whole algorithm for a single quadrangle is presented in Algorithm 5.

Algorithm 5: $E(Q_I, G_I)$ calculation

```
Data: Q_I, G_I = ⟨Q_I, R_I, C_I⟩
Result: E(Q_I) ∈ ℝ+
T_W, T_H = 17 · C_I, 3 · R_I;
TQ = Quad((0,0), (T_W, 0), (T_W, T_H), (0, T_H));
H = ProjectiveTransform(TQ, FQ);
RQ = H(Q);
P = 2 · (T_W + T_H);
for i from 1 to 4 do
    d_i = ||TQ_i − RQ_i||^2 / P;
end
return max_{i∈[1...4]} (d_i);
```

In fact, we have to evaluate this score not for a single image, but for the sequence of ones. The dataset of images with corresponding ground truth values is defined as follows: $\mathbb{D}_{syn} = \{(I_{W \times H}, G) \mid I_{W \times H} \in \mathbb{I}, G \in \mathbb{G}, |\mathbb{I}| = |\mathbb{G}|\}$. Here, $\mathbb{I}$ represents the set of input images with PDF417 symbols, $\mathbb{G}_I$ determines the ground truth for every image $I_{W \times H} \in \mathbb{I}$. N denotes the total number of items in the dataset $\mathbb{D}_{syn}$. Thus, the overall accuracy $\mathcal{S}$ of the proposed method over the dataset $\mathbb{D}_{syn}$ is evaluated as follows:

$$\mathcal{S}(\mathcal{A}_{full}, \mathbb{D}_{syn}) = \frac{1}{N} \cdot \sum_{I \in \mathbb{I}} B(\mathcal{A}_{full}(I), G_I). \qquad (2)$$

The bigger the value of $\mathcal{S}$ is, the better accuracy it shows.

3.2. Data Description

To evaluate the accuracy $\mathcal{S}$ of proposed method, a dataset $\mathbb{D}_{real}$ of real PDF417 codes was collected. For every image in it, the corresponding ground truth values were manually specified. In addition to $\langle Q, R, C \rangle$ they also include original textual message M. The real PDF417 codes often contain some private information, thus, it is not recommended to publish them into public domain due to legal issues. One way to overcome these issues is to synthesize the dataset $\mathbb{D}_{syn}$ from $\mathbb{D}_{real}$ in three steps.

At first, original messages M for the encoding are generated. Since a significant part of real images with PDF417 codes includes ID cards, it is desirable to simulate a similar message structure so as to approximate synthetic PDF417 codes to the real ones. The proposed message structure includes, partially or completely, the following details in random order: (a) first and second name; (b) geographical location; (c) some strings of letters and digits. For example, the message «HIJKLMNO XYZ1234 Lombardy Italy 9876543210 Laura Thomas PQRSTUVW» is encoded in the barcode in Fig. 7.

At the second stage, a rectangular code image S is generated from a given message M according to the PDF417 symbology specification [1]. The final step is to paste an image S into some background taken from $\mathbb{D}_{real}$ knowing the original coordinates of symbols' boundaries Q. There is a number of constraints for such pasting. In order for to fit the requested Q, the image S must be properly scaled. It is necessary to impose restrictions on the min/max scaling of each side. To calculate the min/max valid side dimensions of Q, one has to calculate the min/max width and height of the quadrilateral. Thus, two ranges of lengths are obtained: one for width and one for height. In this case, there are two difficulties: (a) variability in the ratio of width and height; (b) variability in the dimensions of images, which leads to a large range of dimensions of S. In order to avoid excessive compression or stretching of S, adjusting it to the Q, it is necessary to generate a large number of PDF417 of all valid sizes and dimensions from the detected ranges and choose visually the best one from them.

Original colors of barcode modules for S are black and white. But naive pasting of such binary images produces unnatural results. So, the intensities of white and black colors on the real barcode are calculated and replaced with locally averaged ones in order to adopt to real image. The ground truth values for the dataset $\mathbb{D}_{syn}$ are automatically populated from corresponding items in $\mathbb{D}_{real}$ during the generation process. Finally, the $\mathbb{D}_{syn}$ dataset contains 400 images. This dataset is freely available for download at ftp://smartengines.com/se-pdf417-syn-400.

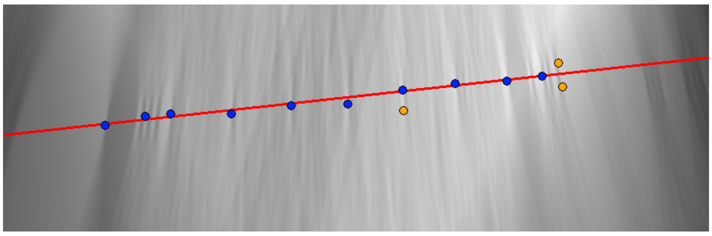
3.3. Results and Discussion

First of all, we have to set the baseline for the $\mathcal{P}$ problem. As was mentioned before, we can not use any commercial barcode reader, because they do not provide the required information. So, for these needs we use the well-known open source ZXing barcode reader (3.5.3 version). It can return required information of Q even in case of further code recognition failure. Let denote $\mathcal{Z}$ the method of PDF417 corner points detection. It uses binarization and template matching strategy for this problem solving.

For the dataset $\mathbb{D}_{syn}$ (see section 3.2) the value of $\mathcal{S}(\mathcal{Z}, \mathbb{D}_{syn})$ is equal to 0.803. Thus, the 321 quadrangles of 400 are correctly detected according to metric $B(Q, G) \in \{0, 1\}$ (see section 3.1). As for the proposed method, its accuracy value $\mathcal{S}(\mathcal{A}_{full}, \mathbb{D}_{syn})$ is equal to 0.948. Thus, the 379 quadrangles of 400 are correctly detected. It means that the error rate of the $\mathcal{A}_{full}$ method is about four time less than for the $\mathcal{Z}$ method. The l lines are detected correctly in 391 cases.

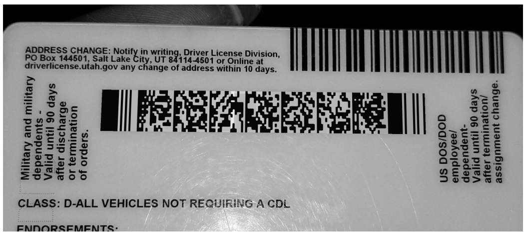
Let us examine the list of images where the proposed method does not produce the expected result. The 21 errors are distributed into three equal classes (i), (ii), and (iii). The class (i) contains errors associated with the incorrect choice of a pair of horizontal sides $\mathtt{S}_h$. The class (ii) consists of cases where the algorithm correctly found the line l, but failed at guard pattern detection Fig. 8. The class (iii) includes the cases when the input image contains extra objects that create "strong" outliers on the FHT image Fig. 8.

a) Example error image from (ii)

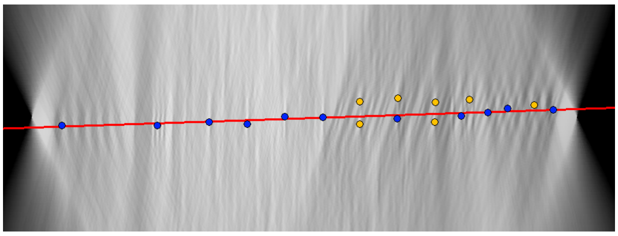


b) Line l on FHT_v image for image (a)

c) Profile for line l for image (b). Green box indicates the guard patterns which were not found



d) Example error image from (iii)



e) Line l on FHT_v image for image (d)

f) Profile for line l for image (e). Blue box corresponds to the guard pattern which is not distinguishable on the profile

Fig 8. The proposed method error samples

The errors of the (i) and (ii) classes can be corrected by adding some steps into the proposed method. For the first class, the set of M first alternatives can be examined. The value of M can be estimated experimentally according to the limitations of the computational resources. For the second class, it is not obligatory to precisely determine the guard pattern positions. Instead of that, only the outer borders of the barcode itself can be detected. After that, the internal "barcodeness" level can be measured by some extra check. Another way to deal with this problem is to increase the initial resolution of the image being analyzed. In such case, the peaks and valleys in guard pattern structure are supposed to be better distinguishable. As for the last class, some errors from it can not be corrected within the proposed method due to its "integral nature". Thus, the better initial localization of barcode region is, the better results it shows. Another approach is to inspect some regions on original image in advance and use the inspection results during the candidate selection process.

Fig 9. Examples when the quadrangle was found correctly on images containing extra objects

To estimate the running time of the proposed method on dataset $\mathbb{D}_{syn}$ we used iPhone 8 and iPhone 14 Pro Max mobile devices. The running time of the proposed method implementation are equal to 29.2 and 12.9 seconds correspondingly, thus, the average running times are equal to 77 and 34 milliseconds per image.

Conclusion

In this paper, we propose the precise method of PDF417 code localization based on Fast Hough Transform usage and RANSAC scheme application. This method does not require commonly used preprocessing steps like binarization, edge detection or morphological filtering. It utilizes only preliminary image scaling which helps to normalize its running time and tune up its parameter values. The average running times measured on iPhone 8 and iPhone 14 Pro Max mobile devices are equal to 77 and 34 ms per image respectively. The computational experiments were conducted on an original dataset SE-PDF417-SYN-400, containing 400 synthesized PDF417 symbols. It is freely available for download at ftp://smartengines.com/se-pdf417-syn-400. It exhibits that the error rate of the proposed method is about four time less than for ZXing detector.

As for future research, the proposed method can be extended to other types of barcodes, such as linear or other stacked codes. In this case, the part related to the "strongest" points extraction from FHT image remains the same, but the part related to symbol edges selection has to be adopted to barcode type.

References

1. Information technology — Automatic identification and data capture techniques — PDF417 bar code symbology specification: Standard / International Organization for Standardization. 2015. DOI: `10.1007/978-981-19-0386-1_49`.
2. Chen C., Kot A.C., Yang H. A two-stage quality measure for mobile phone captured 2D barcode images. Pattern Recognition. 2013. Vol. 46, no. 9. P. 2588–2598. DOI: `10.1016/j.patcog.2013.01.031`.
3. Wudhikarn R., Charoenkwan P., Malang K. Deep Learning in Barcode Recognition: A Systematic Literature Review. IEEE Access. 2022. Vol. 10. P. 8049–8072. DOI: `10.1109/access.2022.3143033`.
4. Ginkel M. van, Hendriks C.L., Vliet L.J. van A short introduction to the Radon and Hough transforms and how they relate to each other. Quantitative Imaging Group, Imaging Science & Technology Department, TU Delft. 2004. P. 1–9.
5. Aliev M., Ershov E., Nikolaev D. On the use of FHT, its modification for practical applications and the structure of Hough image. The Proceedings SPIE. The 11th International

Conference on Machine Vision (ICMV), Munich, Germany, November 1–3, 2018. Vol. 11041. 2019. DOI: 10.1117/12.2522803.

6. Gaur P., Tiwari S. Recognition of 2D Barcode Images Using Edge Detection and Morphological Operation. International Journal of Computer Science and Mobile Computing. 2014. Vol. 3, no. 4. P. 1277–1282.
7. Liu F., Yin J., Li K., Liu Q. An improved recognition method of PDF417 barcode. Chinese Conference on Pattern Recognition (CCPR), Chongqing, China, October 21–23, 2010. IEEE. 2010. P. 1–5. DOI: 10.1109/CCPR.2010.5659332.
8. Kim Y.J., Lee J.Y. Algorithm of a Perspective Transform-Based PDF417 Barcode Recognition. Wireless Personal Communications. 2016. Vol. 89, no. 3. P. 893–911. DOI: 10.1007/s11277-016-3171-6.
9. Xiao Y., Yi J., Qiao G. Automatic Localization of Multi-type Barcodes in High-Resolution Images. International Conference in Communications, Signal Processing, and Systems (CSPS), Changbaishan, China, July 23–24, 2021. Springer. 2021. P. 392–399. DOI: 10.1007/978-981-19-0390-8_103.
10. Bezmaternykh P.V., Vylegzhanin D.V., Gladilin S.A., Nikolaev D.P. 2D barcodes generative recognition. Artificial Intelligence and Decision Making. 2010. No. 4. P. 63–69.
11. Xiao Q., Liu M., Liu Y. A new binarization method for PDF417 bar code by camera phone. IEEE International Conference on Automation and Logistics (ICAL), Qingdao, China, September 1–3, 2008. IEEE. 2008. P. 1904–1908. DOI: 10.1109/ICAL.2008.4636470.
12. Chen R., Yu Y., Xu X., *et al.* Adaptive binarization of QR code images for fast automatic sorting in warehouse systems. Sensors. 2019. Vol. 19, no. 24. P. 5466. DOI: 10.3390/s19245466.
13. Yang X., Gao X., Jia S.Q., Lu Q.Y. A Method for Extracting QR Code from Complex Background Based on Morphology. Applied Mechanics and Materials. 2013. Vol. 239. P. 1466–1471. DOI: 10.4028/www.scientific.net/AMM.239-240.1466.
14. Li D., Zhang L., Jin X. An improvement for PDF417 code authentication on mobile phone terminals based on code feature analysis and watermarking. Multimedia Systems. 2022. Vol. 28, no. 5. P. 1585–1596. DOI: 10.1109/ICIEA.2011.5975630.
15. Li J.-H., Wang W.-H., Rao T.-T., *et al.* Morphological segmentation of 2-D barcode gray scale image. International Conference on Information System and Artificial Intelligence (ISAI), Hong Kong, China, June 24–26 2016. IEEE. 2016. P. 62–68. DOI: 10.1109/ISAI.2016.0022.
16. Trummer M., Denzler J. Reading out 2D Barcode PDF417. 2007. DOI: 10.1007/978-1-84628-945-3_19.
17. Zamberletti A., Gallo I., Albertini S., Noce L. Neural 1D barcode detection using the Hough transform. IPSJ Transactions on Computer Vision and Applications. 2015. Vol. 7. P. 1–9. DOI: 10.2197/ipsjtcva.7.1.
18. Szentandrási I., Herout A., Dubská M. Fast detection and recognition of QR codes in high-resolution images. Proceedings of the 28th spring conference on computer graphics, Budmerice, Slovakia, May 2–4, 2012. 2012. P. 129–136. DOI: 10.1145/2448531.2448548.

19. Hu D., Chen X., Yu D., Li D. Algorithm for detecting the rows boundary of the PDF417 barcode. / ed. by L. Zhang, J. Zhang, M. Liao. Oct. 2005. 60431Z. DOI: 10.1117/12.654952.

20. Aliev M., Kunina I., Nikolaev D., Polevoy D. On the practical aspects of computing the Hough image by the Brady-Yong algorithm. Informatsionnye protsessy. 2023. Vol. 23, no. 2. P. 250–273. DOI: 10.53921/18195822_2023_23_2_250.

21. Fischler M.A., Bolles R.C. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. Communications of the ACM. 1981. June. Vol. 24, no. 6. P. 381–395. DOI: 10.1145/358669.358692.

УДК 519.6 **DOI: 10.14529/cmse240402**

ТОЧНАЯ ЛОКАЛИЗАЦИЯ КОДА PDF417 НА ОСНОВЕ БЫСТРОГО ПРЕОБРАЗОВАНИЯ ХАФА

© 2024 Д.Г. Митрофанов[1,2], П.К. Злобин[1], Ю.А. Шемякина[1], П.В. Безматерных[1,3]

[1] *ООО «Смарт Энджинс Сервис» (117312 Москва, пр. 60-летия Октября, д. 9),*

[2] *Московский государственный университет имени М.В. Ломоносова (119991 Москва, ул. Ленинские горы, д. 1),*

[3] *ФИЦ «Информатика и управление» РАН (119333 Москва, ул. Вавилова, д. 44, кор. 2)*

E-mail: mitrofanov.dmitryi@gmail.com, zlobin.pavel.k@gmail.com, jshemiakina@smartengines.com, bezmaternyh@isa.ru

Поступила в редакцию: 30.09.2024

PDF417 — это популярная символика штрихкода, которая широко используется в удостоверениях личности и транспортных системах. В этой статье мы предлагаем оригинальный метод локализации символа PDF417. Он может успешно обрабатывать проективно искаженные изображения, снятые с помощью камер мобильных устройств. Основой этого метода является анализ изображения быстрого преобразования Хафа. Целью этого анализа является: (a) определение линии, соответствующей точке схода вертикальных сторон символа, с использованием алгоритма RANSAC; (b) выбор лучшей пары точек Хафа, соответствующих горизонтальным сторонам символа. Мы также предлагаем методологию оценки для оценки точности локализации PDF417 и новый набор данных SE-PDF417-SYN-400, который состоит из 400 синтезированных изображений PDF417 и находится в открытом доступе. Точность предлагаемого метода на SE-PDF417-SYN-400 составляет 0.948, частота ошибок предлагаемого метода примерно в четыре раза меньше, чем у детектора ZXing. Среднее время работы на мобильных устройствах iPhone 8 и iPhone 14 Pro Max равно 77 и 34 мс на изображение.

Ключевые слова: распознавание штрихкодов, PDF417, быстрое преобразование Хафа, точка схода, RANSAC.

ОБРАЗЕЦ ЦИТИРОВАНИЯ

Mitrofanov D.G., Zlobin P.K., Shemiakina J.A., Bezmaternykh P.V. Precise Localization of PDF417 Code Based on Fast Hough Transform // Вестник ЮУрГУ. Серия: Вычислительная математика и информатика. 2024. Т. 13, № 4. С. 19–34. DOI: 10.14529/cmse240402.

Литература

1. Information technology — Automatic identification and data capture techniques — PDF417 bar code symbology specification: Standard / International Organization for Standardization. 2015. DOI: 10.1007/978-981-19-0386-1_49.
2. Chen C., Kot A.C., Yang H. A two-stage quality measure for mobile phone captured 2D barcode images // Pattern Recognition. 2013. Vol. 46, no. 9. P. 2588–2598. DOI: 10.1016/j.patcog.2013.01.031.
3. Wudhikarn R., Charoenkwan P., Malang K. Deep Learning in Barcode Recognition: A Systematic Literature Review // IEEE Access. 2022. Vol. 10. P. 8049–8072. DOI: 10.1109/access.2022.3143033.
4. Ginkel M. van, Hendriks C.L., Vliet L.J. van A short introduction to the Radon and Hough transforms and how they relate to each other // Quantitative Imaging Group, Imaging Science & Technology Department, TU Delft. 2004. P. 1–9.
5. Aliev M., Ershov E., Nikolaev D. On the use of FHT, its modification for practical applications and the structure of Hough image // The Proceedings SPIE. The 11th International Conference on Machine Vision, Munich, Germany, November 1–3, 2018. Vol. 11041. 2019. DOI: 10.1117/12.2522803.
6. Gaur P., Tiwari S. Recognition of 2D Barcode Images Using Edge Detection and Morphological Operation // International Journal of Computer Science and Mobile Computing. 2014. Vol. 3, no. 4. P. 1277–1282.
7. Liu F., Yin J., Li K., Liu Q. An improved recognition method of PDF417 barcode // Chinese Conference on Pattern Recognition (CCPR), Chongqing, China, October 21–23, 2010. IEEE. 2010. P. 1–5. DOI: 10.1109/CCPR.2010.5659332.
8. Kim Y.J., Lee J.Y. Algorithm of a Perspective Transform-Based PDF417 Barcode Recognition // Wireless Personal Communications. 2016. Vol. 89, no. 3. P. 893–911. DOI: 10.1007/s11277-016-3171-6.
9. Xiao Y., Yi J., Qiao G. Automatic Localization of Multi-type Barcodes in High-Resolution Images // International Conference in Communications, Signal Processing, and Systems (CSPS), Changbaishan, China, July 23–24, 2021. Springer. 2021. P. 392–399. DOI: 10.1007/978-981-19-0390-8_103.
10. Безматерных П.В., Вылегжанин Д.В., Гладилин С.А., Николаев Д.П. Генеративное распознавание двумерных штрихкодов // Искусственный интеллект и принятие решений. 2010. No. 4. P. 63–69.
11. Xiao Q., Liu M., Liu Y. A new binarization method for PDF417 bar code by camera phone // IEEE International Conference on Automation and Logistics (ICAL), Qingdao, China, September 1–3, 2008. IEEE. 2008. P. 1904–1908. DOI: 10.1109/ICAL.2008.4636470.
12. Chen R., Yu Y., Xu X., *et al.* Adaptive binarization of QR code images for fast automatic sorting in warehouse systems // Sensors. 2019. Vol. 19, no. 24. P. 5466. DOI: 10.3390/s19245466.

13. Yang X., Gao X., Jia S.Q., Lu Q.Y. A Method for Extracting QR Code from Complex Background Based on Morphology // Applied Mechanics and Materials. 2013. Vol. 239. P. 1466–1471. DOI: 10.4028/www.scientific.net/AMM.239-240.1466.

14. Li D., Zhang L., Jin X. An improvement for PDF417 code authentication on mobile phone terminals based on code feature analysis and watermarking // Multimedia Systems. 2022. Vol. 28, no. 5. P. 1585–1596. DOI: 10.1109/ICIEA.2011.5975630.

15. Li J.-H., Wang W.-H., Rao T.-T., *et al.* Morphological segmentation of 2-D barcode gray scale image // International Conference on Information System and Artificial Intelligence (ISAI), Hong Kong, China, June 24–26 2016. IEEE. 2016. P. 62–68. DOI: 10.1109/ISAI.2016.0022.

16. Trummer M., Denzler J. Reading out 2D Barcode PDF417. 2007. DOI: 10.1007/978-1-84628-945-3_19.

17. Zamberletti A., Gallo I., Albertini S., Noce L. Neural 1D barcode detection using the Hough transform // IPSJ Transactions on Computer Vision and Applications. 2015. Vol. 7. P. 1–9. DOI: 10.2197/ipsjtcva.7.1.

18. Szentandrási I., Herout A., Dubská M. Fast detection and recognition of QR codes in high-resolution images // Proceedings of the 28th spring conference on computer graphics. 2012. P. 129–136. DOI: 10.1145/2448531.2448548.

19. Hu D., Chen X., Yu D., Li D. Algorithm for detecting the rows boundary of the PDF417 barcode // / ed. by L. Zhang, J. Zhang, M. Liao. Oct. 2005. 60431Z. DOI: 10.1117/12.654952.

20. Aliev M., Kunina I., Nikolaev D., Polevoy D. On the practical aspects of computing the Hough image by the Brady-Yong algorithm // Informatsionnye protsessy. 2023. Vol. 23, no. 2. P. 250–273. DOI: 10.53921/18195822_2023_23_2_250.

21. Fischler M.A., Bolles R.C. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography // Communications of the ACM. 1981. June. Vol. 24, no. 6. P. 381–395. DOI: 10.1145/358669.358692.

Митрофанов Дмитрий Георгиевич, студент специалитета механико-математического факультета, Московский государственный университет имени М.В. Ломоносова (Москва, Российская Федерация), лаборант-программист, ООО «Смарт Энджинс Сервис» (Москва, Российская Федерация)

Злобин Павел Константинович, лаборант-программист, ООО «Смарт Энджинс Сервис» (Москва, Российская Федерация)

Шемякина Юлия Александровна, научный сотрудник-программист, ООО «Смарт Энджинс Сервис» (Москва, Российская Федерация)

Безматерных Павел Владимирович, программист первой категории, Федеральный исследовательский центр «Информатика и управление» Российской академии наук (Москва, Российская Федерация), научный сотрудник-программист, ООО «Смарт Энджинс Сервис» (Москва, Российская Федерация)