Image-Based Barcode Detection and Recognition to Assist Visually Impaired Persons

Wendy P. Fernandez[1*], Yang Xian[2*], *Student Member, IEEE,* and Yingli Tian[1,2], *Senior Member, IEEE*
[1] Department of Electrical Engineering, the City College, [2] Department of Computer Science, the Graduate Center,
The City University of New York, New York, USA
wfernan001@citymail.cuny.edu, yxian@gradcenter.cuny.edu, ytian@ccny.cuny.edu

Abstract- **Millions of people worldwide suffer from visual impairment or blindness. One of the major limitations caused by the decrease in visual abilities is the difficulties encountered during shopping. In this paper, we propose a barcode detection and recognition method for a shopping assistant. The proposed method is capable of extracting the essential product information (e.g., ingredients, source, calories, etc.) from the detected barcode region. Suggestions of human-system interaction are provided to enhance the user experience. The shopping assistant facilitates the process of grocery shopping and increases the independency of visually impaired persons. To achieve a real-time detection and recognition, we build the detection framework upon a state-of-the-art parallel segment detector that takes advantages of the property that barcode consists of a group of parallel lines. Barcode recognition is accomplished utilizing a robust phone application that provides the detailed information of the product in real-time. As demonstrated by the experimental results, the proposed method is effective, robust, and shows great potential to be expanded into an assistive system for visually impaired persons.**

I. INTRODUCTION

Visual impairment is the reduction of vision not fixable with standard glasses, contact lens, medicine or surgery [1]. 285 million people are estimated to be visually impaired worldwide and 39 million are blind [2]. Visual impairment, especially blindness, decreases the ability of people to perform many daily activities. One of the most difficult activities is grocery shopping since it commonly relies on human visual system to extract the useful product information. Barcodes are standardized product identifiers utilized to increase information management efficiency. They are the main product tagging method at grocery stores [3]. Product recognition by visually impaired users is often performed based on scanning devices, magnifiers to read the labels or sighted shopping assistants. The resources are expensive and sometimes unavailable especially when it requires the presence of another person. Several phone applications offer barcode readers. Nevertheless, these applications rely on the users to accurately localize the barcode regions, and therefore, are still not suitable and convenient for blind users.

Recently, several image-based approaches have been proposed to achieve 1D and 2D barcode detection. Katona and Nyul [4] presented a framework based on bottom-hat filtering and several morphological operations. Bodnar and Nyul [5] utilized the barcode symmetry for detection. The image was partitioned uniformly and scanned in a circular pattern. Later in [6], the scan-line based approaches were replaced with distance transformation to cluster regions with similar properties. Zamberletti *et al.* [7] proposed a barcode detection system utilizing supervised machine learning techniques combined with Hough transform. In [8], Creusot and Munawar introduced Maximally Stable Extremal Regions (MSER) to cluster the candidate segments in Hough space. Although these methods achieve satisfying detection performance under controlled environments, they are not ideal for the real-world shopping assistive system for visually impaired users with the limitation in computation power and the existence of motion blur when capturing images.

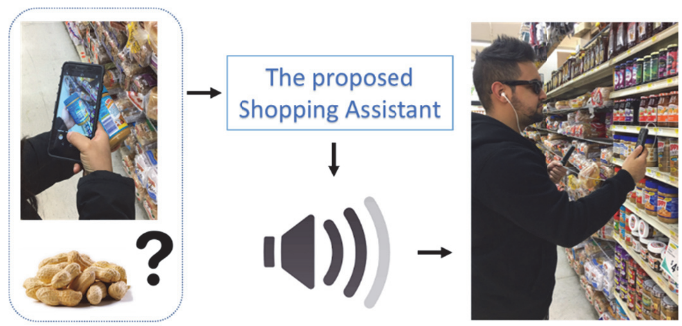

Figure 1. Illustration of the proposed method in assisting grocery shopping of blind users with suggested human-system interaction. The detailed product information is recovered based on the barcode region detected from the image captured by the user. Audio tools could be utilized to return information important to this specific user.

Barcode readers in phone applications are popular nowadays. Their functionalities range from price comparison [9], inventory control [10], events check-in [11], proof of delivery [12], calories counters [13] to food selection [14-17]. Among all the aforementioned applications, barcode readers emphasizing on food selection provide essential product information in assisting blind users during grocery shopping. Fooducate [15] targets at helping people make healthy choices by grading products based on their nutritiousness and calories. ScanHalal [16] provides basic information such as the ingredients, the source of the product, and others – but at relatively slower speed. Ipiit [17] is sensitive to the quality of the captured input image. In the proposed system, we employ ShopWell [14], which 1) provides the general ingredients, nutrition facts, gradings based on the products' nutritiousness, 2) allows users' inputs to avoid certain ingredients, record certain health conditions, and 3) is robust, efficient, and suitable in assisting the grocery shopping of blind users.

* equal contribution

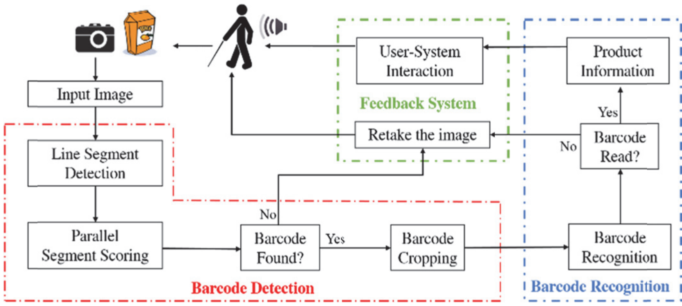

Figure 2. Systematic flowchart of the shopping assistant. It consists of the proposed barcode detection and recognition method, along with the suggested user feedback system.

In this paper, we propose a framework which combines a modified version of Parallel Segment Detector (PSD) [18] and ShopWell [14], a recognition application available for multiple platforms, to achieve the 1D barcode detection and recognition. Suggestions are provided to further enhance the user experience. The rest of the paper is organized as follows. Section II introduces in detail the shopping assistant for blind users. Section III presents the experimental results and analysis for the proposed barcode detection and recognition method. Conclusions are drawn in Section IV.

II. Methodology

The proposed method takes a product image captured by the user as the input. It aims to detect and recognize the barcode to extract the product information. The input image is first fed to the line segment detection algorithm and afterward, each line segment is evaluated to measure its possibility belonging to the barcode region. A barcode candidate region is generated after the scoring function. This candidate region is then cropped, saved, and served as the input to a barcode recognition application (i.e., ShopWell [14]) to extract the necessary product information for the user. In cases where the barcode is not found during the detection process, or the recognition application fails to extract the product information, simple signaling (e.g., vibrations or beep tone) could be utilized as feedback to instruct the user to retake the image. Other suggestions are provided for the feedback system as shown in Sec. II.C.

As illustrated in Fig. 2, the assistive system consists of three components: barcode detection, barcode recognition, and user feedback system. The detailed steps are presented in the following subsections.

A. Barcode Detection

Barcodes are designed under normalized standards by the GS1, i.e., The Global Language of Business [3]. 1D barcodes are represented as parallel lines compacted close to each other. The exploited barcode detector is built upon PSD [18] which has superior performance in detecting parallel lines. Moreover, PSD is effective in reducing the false negatives caused by motion blur during image capture, which fits perfectly in our application.

We first extract all the line segments from the input image utilizing line segment detector (LSD) [19]. Afterward, a bounding box for each line segment is generated based on the coordinates of the corner pixels. Each bounding box is augmented in the perpendicular direction by twice the length of the segment, as suggested in [18]. For the evaluated segment s_i, the segments inside the bounding box form its neighborhood N_i. We then calculate the score value v_i which measures the likelihood of segment s_i belonging to a barcode region, i.e., the likelihood of segment s_i being a bar. To be more specific, this score value is determined same as in [18] by comparing the length, angle, and center of each segment within the neighborhood as presented in Eq. (1):

$$v_i = \sum_{s_k \in N_i} (\varphi_{ik}^{\alpha} \varphi_{ik}^{\rho} \varphi_{ik}^{\chi}), \quad (1)$$

in which φ_{ik}^{α}, φ_{ik}^{ρ} and φ_{ik}^{χ} represent the angle, length, and center distance functions between the evaluated segment s_i and the segment s_k in the corresponding neighborhood N_i. Each of the listed three conditions is determined by a distance function and a pre-set threshold (more details can be found in [18]).

Each segment returned after LSD is evaluated within its neighborhood and the one with the largest scoring value indicates it has the highest probability being a bar belonging to a barcode. Then, this is utilized to further detect the barcode region for the recognition step later. A bisector line is first generated orthogonal to the segment with the highest scoring value in the center. Afterward, this bisector line is explored to estimate the beginning and ending of the barcode region. This is accomplished in a manner different from PSD [18] as shown in Eq. (2):

$$\sigma_{p \in \wp} = \sum_{i=p}^{p-\alpha} |B(i) - B(i-1)| - \sum_{i=p}^{p+\alpha} |B(i) - B(i+1)|, \quad (2)$$

where $\wp$ indicates a set of the sampled pixels along the bisector line. p stands for the evaluated pixel. $B(\cdot)$ represents the intensity value along the bisector line at a given position. Parameter α defines the distance to check on the left side and on the right side. Since the size of the barcode region varies for each input instance, we set α same as the length of the segment with highest scoring value (i.e., the length of a bar). This is based on the observation that in general, barcode region forms a rectangular with four sides in the same magnitude of length. Compared with a fixed setting α for all images, this tailored manner performs normalization process implicitly.

This step aims to capture the difference in the oscillation patterns between the left side and the right side for each pixel along the bisector line. It is based on the observation that pixels within the barcode region tend to have the similar oscillation patterns on both sides and will cancel each other as approaching to the center of the barcode region based on Eq. (2). On the other hand, the aforementioned phenomenon is not observed in general for pixels outside the barcode region along the bisector line. Thus, after the calculation of the σ value for each sampled pixel along the bisector line, we observe a pattern in σ that the changes outside the barcode region are steeper (as shown in Fig. 3(b1),(b2)). And the starting and ending points of the "slower-changed" region provide the hint for estimating the boundaries of the potential barcode region.

Figure 3. Illustrations of the proposed barcode detection and recognition framework on two sample images. For each instance, (a) Input image. (b) Corresponding $\boldsymbol{\sigma}$ values for the sampled pixels along the bisector line. (c) Results generated during the barcode detection step including the segment with the highest score (in red), the bisector line (in green), and the bounding box indicating the barcode region (in blue). (d) Barcode region after cropping. (e) Returned recognition results from ShopWell [14] based on the cropped barcode region. Figure is best viewed in color.

In practice, in order to find the boundaries of the barcode region, we start with the intersection of the segment with the highest score and the bisector line (denote as p_r). Afterward, the search process starts from p_r and expands along the bisector line on both directions. For each direction, the absolute difference in σ value between each pair of the neighboring pixels is monitored. Once the difference exceeds a pre-set threshold t, i.e., the boundary of the "flat" region is hit, the search process terminates. The termination positions of both directions define the starting and ending points of the barcode region. To further ensure the performance, once a candidate region is generated, the ratio r between the width and the height is calculated. If r is larger than a threshold r_u, we will reduce t and repeat the above process; on the contrary, if r is smaller than a threshold r_l, t is increased and the above procedure is repeated. This adjustment reduces the number of failure cases when the barcode region is surrounded by cluttered background. Finally, to fully utilize all the useful information (i.e., the digits below the bars appearing in the barcode), we expand the candidate area evenly along each side. This expanded region is then cropped and utilized as the input to the barcode recognition step. Details of the parameter settings can be found in Sec. III.B.

B. Barcode Recognition

ShopWell [14] is a cellphone application available for multiple platforms. As mentioned, this application provides essential information including the list of ingredients, the nutrition facts, and the grading based on the nutritiousness. More importantly, it provides personalized interaction that allows the users to avoid certain ingredient, record preference diet, and choose for personalized lifestyle. Moreover, this application responses quickly with the detailed product information and is robust in real application.

Fig. 3(e1), (e2) provide two examples of the returned product information based on the barcode regions obtained in the previous detection step. As for the daily use of visually impaired persons, it could save the personal choices as to avoid a list of ingredients (e.g., peanuts, soy, gluten, etc.). Additionally, suggestions are provided based on specific health conditions (e.g., diabetes, low fodmap, high cholesterol, etc.). With the detected barcode region, information retrieval process is more effective and convenient for the users.

C. User-System Interaction

To further enhance the user experience of the proposed shopping assistant, several suggestions are listed in this subsection to facilitate the process. As illustrated in Fig. 2, feedback is necessary in the following scenarios to guide the image capture process: 1) when a candidate barcode region cannot be captured in the detection phase; 2) when the cropped region is a partial barcode and cannot be recognized by the application. Although with the fixed standards, variations still exist in the appearances and positions of each barcode for different products. For example, for products with boxes, in general the barcodes tend to appear in the bottom corner of one side – however, it is still challenging for visually impaired users to find the correct side with the full barcode regions exposed to the camera. It is even more difficult with products in bottles or cans. Therefore, simple signaling embedded in the cellphone (e.g., vibrations or beep tones) could be utilized to instruct the user to retake the image. Moreover, if partial barcode is detected, audio instructions could be provided as to guide the user to rotate the product towards a certain direction.

Another component in the feedback system is the tailored requests input by the user. The available speech-to-text and text-to-speech applications could be utilized to translate the users' requests and return related information in audio.

Figure 4. Sucessful detection results from sample images of the self-collected barcode dataset. Figure is best viewed in color.

III. Experimental Results

A. Dataset

To evaluate the proposed barcode detection and recognition method, a new dataset is proposed collected from grocery stores and supermarkets containing a variety of products that are commonly used in the daily life. The images are captured utilizing a Huawei Honor 5 cellphone camera with a resolution of 8 megapixels. It consists of 206 images with different rotations and resolutions of the barcode region to best mimic the real-world situation. Several examples from the dataset are illustrated in Figs. 4 and 5.

B. Parameter Settings

To generate the σ values along the bisector line, as mentioned previously, α in Eq. (2) is set to the same value as the length of the segment with the highest score. $\wp$ is generated by sampling every other pixel along the bisector line. Later, during the search of the barcode region boundaries, the pre-set threshold t used to measure the difference between neighboring pixels is set to 5,000. In practice, the comparison is not performed at every pixel along the bisector line, but at an evenly sampled way for efficiency concern. The two thresholds r_u and r_l used for controlling the ratio are set to 3 and 1, respectively.

C. Experimental Results

The proposed method is evaluated utilizing the self-collected barcode dataset. Instances with barcode regions appearing in different orientations and zoomings are tested and reported.

As illustrated in Fig. 3, first the segment with the highest possibility (marked in red) is accurately generated. Afterward, the bisector line (marked in green) is created to assist the search for the boundaries of the potential barcode region starting from the intersection of the above listed two lines. The final bounding box (marked in blue) is generated based on the

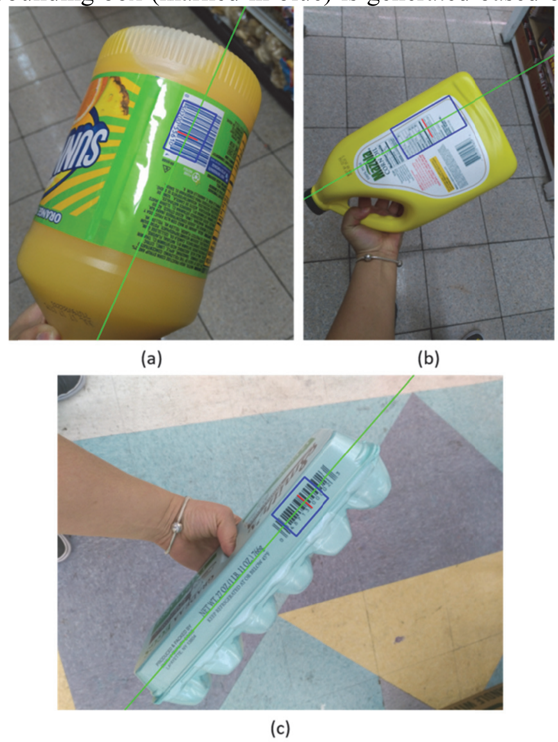

Figure 5. Unsucessful detection results from sample images of the self-collected barcode dataset corresponding to three different categories. Details in text. Figure is best viewed in color.

results returned from the searching followed with an expansion along all edges. The expansion introduces other useful information (i.e., digits beneath the bars) to further facilitate the recognition step later. The region defined by the bounding box is then cropped, saved, and served as the input to the recognition step. Two examples are provided in Fig. 3 with recognition results returned by ShopWell containing essential information of the product such as nutrition and ingredients. Suggestions and recommendations are also provided based on the user's profile. As suggested, text-to-speech applications could be utilized to return certain information to the user by audio.

Fig. 4 provides more examples indicating the successful detection results for the proposed framework. The recognition results are not included here due to the limit of space. In general, among the 206 images within the dataset, the proposed detection method successfully generates a bounding box that covers the whole barcode region in 145 of them (i.e., the accuracy is 70.39%). The 61 failure cases can be roughly classified into three categories: 1) The segment with the highest score overlaps a bar, but not fully; 2) The segment with the highest score does not overlap a bar; 3) The ratio of the ground-truth barcode region is larger than the pre-set threshold r_u (3 in our implementation). Corresponding examples can be found in Fig. 5. Moreover, among the failure cases, 55.74% (i.e., 34 images) of them belong to the first category.

IV. Conclusions

In this paper, we have presented a framework that takes an input image (or multiple if no barcode captured at first attempt) of a shopping product captured by the visually impaired users and returns the detailed information of this product based on the detected barcode region. The proposed barcode detection and recognition method is explained and evaluated utilizing a self-collected barcode dataset. We also include suggestions to assist the users in taking the photos in cases where there is no or partial barcode in the captured image and allowing effective information retrieval for specific queries. Our future work will be focusing on building an end-to-end system, accuracy improvement, user interface study, and system evaluation and refinement with the participation of blind users.

Acknowledgements

This work was supported in part by NSF grants EFRI-1137172, IIP-1343402, and IIS-1400802.

References

[1] Center for Disease Control and Prevention (CDC). “Blindness and vision impairment.” URL: www.cdc.gov/healthcommunication/toolstemplates/ entertainmented/tips/blindness.html, 2011 [Accessed on Mar 9, 2017].

[2] World Health Organization (WHO). “Visual impairment and blindness.” URL: www.who.int/mediacentre/factsheets/fs282/en/, 2014 [Accessed on Mar 9, 2017].

[3] GS1US. “Barcoding basics.” URL: https://www.gs1us.org/upcs-barcodes-prefixes/barcoding-basics, [Accessed on Mar 13, 2017].

[4] M. Katona and L.G. Nyul, “A novel method for accurate and efficient barcode detection with morphological operations,” in *Signal Image Technology and Internet Based Systems (SITIS)*, pp. 307–314, 2012.

[5] P. Bodnar and L. G. Nyul, “A novel method for barcode localization in image domain,” in *Image Analysis and Recognition*, vol. 7950 of *Lecture Notes in Computer Science*, pp. 189–196. Springer Berlin Heidelberg, 2013.

[6] P. Bodnar and L. G. Nyul, “Barcode detection with uniform partitioning and distance transformation,” in *IASTED International Conference on Computer Graphics and Imaging*, pp. 48–53, 2013

[7] A Zamberletti, I Gallo, and S. Albertini, “Robust angle invariant 1d barcode detection,” in *IAPR Asian Conference on Pattern Recognition (ACPR),* pp. 160–164, 2013

[8] C. Creusot and A. Munawar, “Real-time barcode detection in the wild,” in *IEEE Winter Conference on Applications of Computer Vision (WACV),* pp. 239–245, 2015

[9] Purch, “Shopsavvy.” URL: https://www.shopsavvy.com, [Accessed on Mar 13, 2017].

[10] RomiSys, “Inventory Droid.” URL: http://www.romisys.com /pages/inventory-droid.php, [Accessed on May 1, 2017].

[11] Evenbrite, “Evenbrite Organizer.” URL: https://www.eventbrite.com/ how-it-works/, [Accessed on May 1, 2017].

[12] Kelgray, “POD, Fully Hosted Barcode Proof of Delivery Solution” URL: http://www.kelgray.co.uk/products/proof-of-delivery.php, [Accessed on Mar 13, 2017].

[13] Under Armour Connected Fitness, “MyFitnessPall” URL: https://www.myfitnesspal.com/, [Accessed on May 1, 2017].

[14] ShopWell, “ShopWell.” URL: http://www.shopwell.com/, [Accessed on Mar 13, 2017].

[15] Fooducate, “Fooducatel.” URL: http://www.fooducate.com/, [Accessed on Mar 13, 2017].

[16] Ummah Labs, “Features.” URL: http://www.scanhalal.com/#features, [Accessed on Mar 12, 2017].

[17] ipiit, “The Food Ambassador.” URL: https://www.ipiit.com/, [Accessed on Mar 12, 2017].

[18] C. Creusot and A. Munawar, “Low-computation egocentric barcode for the blind,” in *IEEE International Conference on Image Processing (ICIP),* pp. 2856-2860, 2016

[19] R.G. von Gioi, J. Jakubowicz, J.-M. Morel, and G. Randall, “Lsd: A fast line segment detector with a false detection control,”*, IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 32, no. 4, pp. 722–732, April 2010.