Eyes-Free Barcode Localization and Decoding for Visually Impaired Mobile Phone Users

Vladimir Kulyukin **Aliasgar Kutiyanawala***
Computer Science Assistive Technology Laboratory
Department of Computer Science
Utah State University
Logan, UT, 84322
vladimir.kulyukin@aggiemail.usu.edu, aliasgar.k@aggiemail.usu.edu

Abstract

An eyes-free barcode localization and decoding method is presented that enables visually impaired (VI) mobile phone users to decode MSI (Modified Plessy) barcodes on shelves and UPC barcodes on individual boxes, cans, and bottles. Simple and efficient barcode localization and decoding techniques are augmented with an interactive haptic feedback loop that allows the VI user to align the phone's camera with a fixed surface in the pitch and yaw planes. The method is implemented on a Google Nexus One smart phone running Android 2.1. A laboratory study is presented in which the method was evaluated by one VI and four blindfolded sighted participants.

Keywords: Accessible Shopping, Eyes-Free Barcode Localization & Decoding, Assistive Technology [1]

1 Introduction

Supermarkets are one of the most functionally challenging environments for VI individuals. A modern supermarket has a median area of 4,529 square meters and stocks an average of 45,000 products [1]. VI people typically rely on family members or friends who take them grocery shopping. When these individuals are unavailable, VI shoppers reschedule their shopping trips or rely on store staffers assigned to them when they come to the store. Some staffers are unfamiliar with the store layout, others become irritated with long searches, and still others do not have adequate English skills to read the products' ingredients [2]. These difficulties cause VI shoppers to abandon searching for desirable products or settle for distant substitutes. PeaPod or similar home delivery services provide grocery shopping alternatives. However, such services are not universally available and, when available, require shoppers to schedule and wait for deliveries, thereby reducing personal independence and making spontaneous shopping impossible. To overcome these access barriers, accessible shopping systems are needed that increase personal independence, enable spontaneous shopping, and do not require that supermarkets undergo extensive technological adjustments. Such systems will make VI individuals independent of external assistance and fundamentally improve their quality of life.

In 2006, we began our work on ShopTalk [3], a wearable system for independent blind supermarket shopping. ShopTalk consists of a small computer device (OQO model 01), a Belkin numeric keypad, a Hand Held Products IT4600 SR wireless barcode scanner and its base station, and a USB hub that connects all components. To help carry the equipment, the user wears a small CamelBak backpack. The numeric keypad is attached by velcro to one of the backpack's shoulder straps. To ensure adequate air circulation, the OQO is placed in a wire frame attached to the outside of the backpack. The remaining components - the barcode scanner's base station, the USB hub, and all connecting cables - are placed inside the backpack's pouch. Since the system gives speech instructions to the user, the user has a small headphone. The system was our first attempt (and the first attempt reported in the accessible shopping literature) to use *shelf* barcodes (MSI, not UPC) as topological points for locating products through verbal directions.

In 2007 - 2008, after two single subject studies at Lee's Market Place, an independently owned supermarket in Logan,

---

*Contact Author

[1] Submitted to IPCV 2010

UT, and one single subject study at Sweet Peas, a small independent natural foods store in Logan, UT, ten VI participants were recruited for a longitudinal formal study. The product database was extended to include 4,297 products. The experiment, performed during regular business hours, had each participant shop for the same set of three randomly chosen products five times. The product retrieval rate was 100%. All ten participants found all three products in every run (see [2] for detailed ANOVA analysis and experiment design). A key finding of ShopTalk is that all participants in the sample could execute verbal route directions in supermarkets and product search instructions triggered by barcode scans with 100% accuracy.

In 2008 - 2009, ShopTalk was ported onto the Nokia E70 smartphone running Symbian OS 9.1. The smartphone is connected to a small Bluetooth barcode scanner (Baracoda Pencil2) [4]. The port, called *ShopMobile*, was tested in a laboratory study with six sighted blindfolded participants. The experiment had eighty shelf barcodes obtained from the USU bookstore placed on shelves assembled into 3 aisles where aisles 1 and 3 had only one side while aisle 2 had two sides. Each participant successfully retrieved a set of four products.

We are currently developing ShopMobile II, the next version of ShopMobile. Like its direct predecessors, ShopTalk [5] and ShopMobile [4], ShopMobile II uses two data structures to represent the shopping environment. The first data structure is a topological map, which is a directed graph whose nodes are decision points: the store entrance, aisle entrances, and cashier lane entrances. The edges are labeled with directions. Due to the regularity of modern supermarkets, we found it sufficient to have three directional labels: *left*, *right*, and *forward*. The topological map is built at installation time by walking through the store, noting decision points of interest, and then representing them in the map. The second data structure, called the *Barcode Connectivity Matrix* (BCM), is designed to take advantage of the inventory control systems used by many grocery stores. These inventory systems place barcodes on the shelves immediately beneath every product type. Shelf barcodes assist the store personnel in managing the product inventory. Once the locations of all shelf barcodes in the store are known, this information is used to guide the shopper through to the target shelf section, because products are located directly above their corresponding shelf barcodes. More information on how the BCM can be computed from the inventory control databases is available in [2].

Unlike its predecessors, ShopMobile II no longer requires a barcode scanner, because it relies exclusively on computer vision techniques to recognize MSI barcodes on shelves and UPC barcodes on individual products. Vision-based barcode localization and decoding is a well-known research problem. Algorithms have been proposed that look for line sets with mono-oriented gradients to locate the barcode region [6, 7]. Other algorithms extract connected regions with mono-oriented texture to perform local barcode searches in cluttered 3-D scenes [8, 9, 10]. Hough transformation-based methods extract barcode lines by assuming robustness against orientation and size [11]. Various morphological filters and self-learning networks have also been tried [12]. One common disadvantage of these algorithms is that they are very time-consuming and require external servers for image processing. Recently several barcode decoding solutions for camera phones have been proposed [13, 14]. However, these solutions target sighted users in that they require that the phone camera is carefully aligned and centered on a barcode for the decoding to work. An accessible barcode localization and decoding method is proposed in [15]. However, unlike the method presented in this paper, it has not been implemented on a mobile phone yet. We wholeheartedly endorse and commend these R&D efforts. Our approach is based on the hypothesis that *sophisticated vision techniques may not be needed to make barcode localization and decoding accessible to VI mobile phone users, because simple and efficient vision techniques can be augmented with interactive user interfaces that ensure that images have certain properties.*

In this paper, we present an interactive haptic feedback loop that allows the VI mobile phone user to align the camera with a fixed surface in the pitch and yaw planes, which guarantees that the captured image is taken from the mobile phone parallel to the surface. The remainder of our paper is organized as follows. In Section 2, we describe our barcode localization and decoding method. In Section 3, we present an interactive haptic camera alignment loop that enables the VI mobile phone user to align the phone's camera with fixed surfaces in the pitch and yaw planes. In Section 4, we give brief descriptions of the target user and the target use case toward which we are working. We also present a laboratory study with one VI individual and four sighted blindfolded participants. In Section 5, we offer our conclusions.

2 Eyes-Free Barcode Localization and Decoding

A barcode can be viewed as a homogeneous region consisting of alternate black and white lines condensed in a small region. In ShopMobile II, barcodes are characterized as image regions that have the following two properties. We call the first property *alternating frequency* and define it as the number of black to white and white to black transitions along the x-axis. We call the second property *vertical continuity* and define it as the continuity of black and white lines along the y-axis. As an example, consider two one-pixel wide lines A and B placed on the barcode as shown in the left part of Figure 1. These lines can be encoded as bit strings where black pixels map to zeros and white pixels map to ones. The alternating frequency is the number of 0-1 and 1-0 transitions in the bit string. Vertical continuity is measured as the longest common subsequence of the two bit strings. Other string matching techniques can certainly be used provided that there is no impact on the overall efficiency.

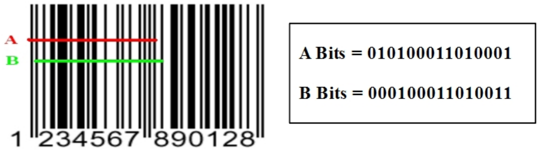

Figure 1: One Pixel Wide Lines on the Barcode.

-1	2	-1
-2	4	-2
-3	6	-3
-4	8	-4
-5	10	-5
-6	12	-6
-7	14	-7
-6	12	-6
-5	10	-5
-4	8	-4
-3	6	-3
-2	4	-2
-1	2	-1

Figure 2: Line Detection Filter Kernel.

For efficiency, the 1024 by 768 image is reduced to 320 by 240 pixels. Each pixel, a packed integer containing the pixel's alpha and RGB components, is mapped to a grayscale value Y between 0 and 255, where $Y = 0.3 \times R + 0.59 \times G + 0.11 \times B$. Since the Google Nexus One's camera has a provision to take a greyscale image, the above step reduced to setting $Y = G$ for efficiency. In our previous work [16], we used a Gabor filter to localize barcodes in images. In the current implementation, the Gabor filter is replaced with a much more efficient line detection filter that convolutes the kernel shown in Figure 2 through the entire image. This kernel is inspired by the line vertical detection kernel based on a 3×3 matrix discussed in [17]. We experimentally extended the kernel to a 13×3 matrix with decreasing coefficients along the y-axis to make it less sensitive to the noise in the image. Given an $m \times n$ matrix, our filter can be generated as follows:

$$f[i][j] = \begin{cases} (\frac{m+1}{2} - |i|) \times \frac{n^2-1}{4} & \text{if } j = 0 \\ -(\frac{m+1}{2} - |i|) \times (\frac{n+1}{2} - |j|) & \text{if } j \neq 0, \end{cases}$$

where $-m/2 \leq i \leq m/2$ and $-n/2 \leq j \leq n/2$.

Since barcode lines pass through this filter along with some other vertical lines generated by graphics and text that may be present in the image, the filtered image is searched in a rasterized pattern with two one pixel wide lines in order to isolate areas with high alternating frequency and high vertical continuity. Fast histogram analysis is performed on all pairs of lines with sufficiently high coefficients for the alternating frequency vertical continuity. After a small set of candidate regions is obtained, each candidate region is mapped to the original image, because, while the reduced resolution is sufficient for fast and adequate barcode localization, it turned out not to be adequate for ZXing [13], which we used for barcode decoding. ZXing is an open source library for decoding several 1D and 2D barcode types, including UPC, but does not support the MSI barcode type. We extended ZXing to decode MSI barcodes and extended the API to accept a set of barcode regions for decoding. Thus, the localized candidate regions are mapped back to the original image scale (1024 by 768 pixels) and cropped from it so that ZXing can process them. Future references to ZXing in the paper will refer to this extended version of ZXing.

We designed an experiment to measure the contribution of our barcode localization algorithm to ZXing. A total of 187 images of real products were captured with a HTC Touch Pro smartphone. In the first run, ZXing alone decoded 91 images. In the second run, the images were first processed with our barcode localization algorithm and then ZXing ran on the cropped candidate regions. This time barcodes were correctly decoded in 133 images, which amounts to an increase of 42 images or 46.15% over the number of images decoded by ZXing alone. The software for this experiment was implemented using Java SE and Netbeans and ran on a Dell Optiplex 960 Core2Duo machine running at 3GHz.

3 Interactive Camera Alignment

The shopper always takes a picture of a fixed surface *after* the phone's camera is aligned with the surface in the pitch and yaw planes. The alignment is carried out through an iterative haptic control loop. If the product is a box, the user is instructed to find the top of the box, i.e. the side with an opening tab, by touch and then align the camera with the bottom, i.e. the opposite side. This is because on most boxes sold in the U.S. the UPC barcode is on the bottom. Once the bottom is found, the shopper aligns the camera with an edge of the side and clicks a button on the touch screen to start the alignment loop. The current readings of the phone's orientation sensor are taken to define the absolute pitch and yaw planes. The shopper then slowly moves the camera away from the surface. If the camera is misaligned in the pitch or yaw planes, specific haptic signals (vibrations) are issued to re-align it. The signals persist until the camera is aligned. The user is instructed to stop moving the camera away from the surface when she thinks that the camera is approximately 10cm from the surface. Then the picture is taken by a click of a button on the touch screen. If the product is a can, the shopper aligns the camera with the edge of the top or bottom circle. If it is a bottle, the camera is aligned with the edge of the bottom circle.

When the shopper reaches an aisle, she aligns the phone camera with a shelf and captures an image. If a barcode is decoded successfully and happens to be a shelf barcode, the system checks if it is the target barcode. If it is, the system verbally informs the shopper about it. If it is not, the system generates verbal directions to the target product (e.g. "Two shelves up, three barcodes left."). If the barcode cannot be decoded in the image, the system divides the image into sub-images that are likely to contain barcodes. Each sub-image is processed with a barcode decoder. If the barcode is decoded, the system goes into the verbal instruction mode. If the barcode is not decoded, the system assumes that the barcode is rotated by 90 degrees. This assumption is enforced by the interactive haptic alignment loop. The image is rotated by 90 degrees, and the entire processes is repeated again. If part of a barcode is detected, the system gives the user haptic feedback to move the camera slightly left, right, up, or down, depending on where the partial barcode is detected. When the system is unable to find the barcode in the image, it notifies the user through audio feedback (currently a ping) to take another image. The overall flow of control of the barcode localization and decoding method is given in Figure 3. If necessary, user feedback messages can be delivered in a variety of formats: haptic, audio, speech, or a combination thereof.

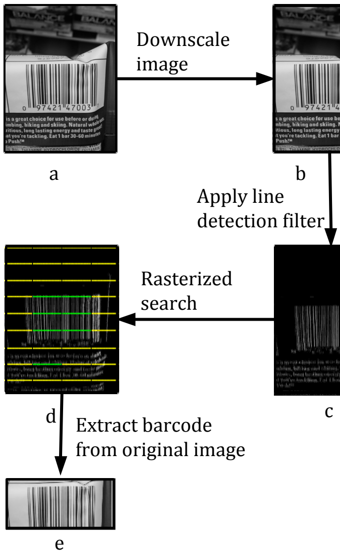

Figure 4: Accessible Barcode Localization and Decoding Method.

4 A Laboratory Study

Before we describe our laboratory studies, we will briefly describe the target user and the target use case toward which we are working. We will call our target user Alice (name changed to protect identity) who was a participant in one of our longitudinal accessible shopping experiments in Lee's Market Place. She is in her mid-twenties, a mother of two children, and a seasoned mobile phone user. Both she and her husband are completely blind and cannot help each other with grocery shopping. Alice has excellent orientation and mobility (O&M) skills, takes independent walks around her neighborhood, uses the Cache Valley transit system, and has a full-time job. Alice is a resident of Logan, Utah, and goes to Lee's Market Place several times a week. ShopMobile II is a mobile solution for individuals like Alice.

We are working toward the following use case. When Alice enters the supermarket, she picks up a shopping basket. Alice also has adequate cane skills to *pull* a shopping cart if she has to. Alice takes out her mobile phone and, if her shopping list is prepared, selects a product by menu browsing. The system tells Alice where the product is located, and if requested, generates a template-based route description. When Alice reaches the aisle, she aligns her phone camera with a shelf and takes a picture. If the camera is not vertically aligned to the shelf, the system uses the phone's orientation sensor to give Alice haptic (vibratory) feedback to align the camera. Alice then takes an image. If part of a barcode is detected, another vibratory signal requests Alice

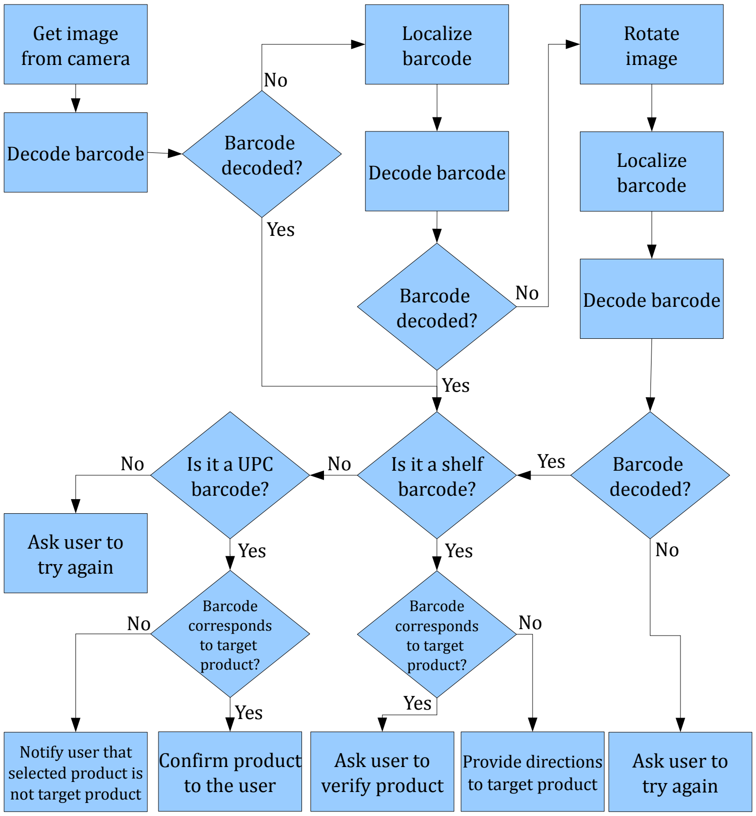

Figure 3: Barcode Localization and Decoding.

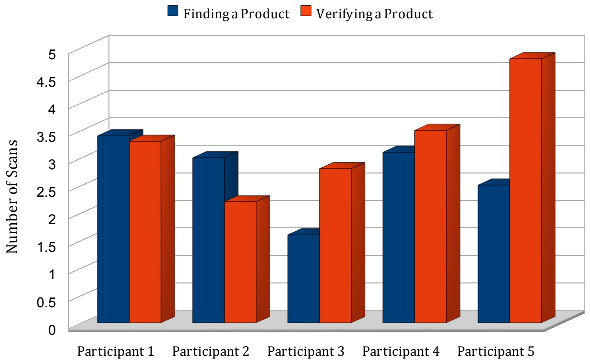

Figure 5: Mean Values for Number of Scans to Retrieve and Verify Products for Each Participant.

to move the phone to the left (right). When the shelf barcode is recognized, the system uses its database to give Alice verbal directions to the target product, e.g. *two shelves up, five barcodes right* or *this is aisle 5, target product is in aisle 2*. When the target shelf barcode is recognized, Alice reaches above the barcode and takes a product from the shelf. If the product's identity cannot be determined manually, e.g. by touch, smell, shake, etc., they system allows Alice to interactively recognize its UPC barcode to verify the product's identity.

We approximated the target use case with a laboratory study. We ported all software to a Google Nexus One smart phone equipped with a five megapixel camera. The phone runs on Android 2.1 on a 1 GHz processor with 512 MB of RAM. A grocery store environment was simulated by assembling a shelf section in our laboratory that consisted of five shelves. Twelve empty boxes of real products were placed on three shelves.

We recruited one VI individual and four sighted individuals to test the system. The sighted individuals were blindfolded. Each participant was given a ten minute training session in which the system was shown to him/her and then asked to retrieve and verify ten randomly selected products. For each participant, the system logged the image associated with each scan, the result of the scan and the method

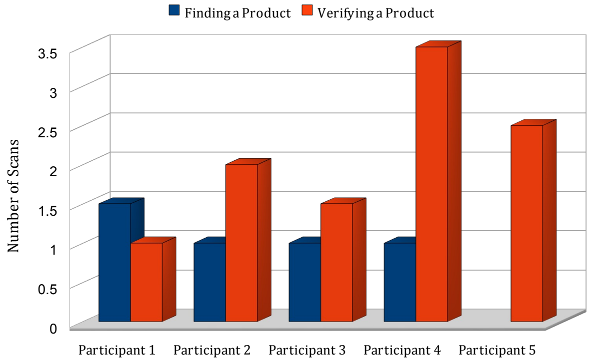

Figure 6: Median Values for Number of Scans to Retrieve and Verify Products for Each Participant.

by which the barcode was decoded.

All participants were able to retrieve all products successfully. Figure 5 shows the average number of scans for product retrieval and verification. It took an average of 2.72 scans for a participant to retrieve a product and an average of 3.32 scans to verify a product. The STD values are 4.4 for product retrieval and 5.6 for product verification. The median values for product retrieval and product verification are shown in Figure 6.

In general, the median values for product retrieval are lower than the median values for product verification. For participant 5 (a sighted individual), the median value for product retrieval is 0. The system is designed so that that the shopper has the option of picking up a product from the shelf without having to scan the MSI shelf barcode below it, because some shoppers can locate the target product by counting shelf barcodes by touch.

Figure 7 shows a distribution of the result of the scans for all participants. ZXing alone decoded about 13% of all barcode scans. When used in conjunction with our barcode localization algorithm it decoded another 16% of all barcode scans. Rotating the image before localizing and decoding the barcode results in about 4% more barcode scans to be decoded. Barcode localization causes the number of barcodes decoded to increase from 38 to 97, an increase of over 155%.

In the informal post-experiment interviews, all participants said that our touch screen user interface was very simple to use but too basic. The interface consisted of a single button encompassing the entire touch screen of the Google Nexus One phone. The barcode decoding procedure was initiated either by finger tapping on any part of the screen or by tapping on the phone's trackball, a small joystick-like hardware component below the touch screen. All participants preferred tapping the screen to tapping the trackball, which resulted in many accidental images with no barcodes present in them, which negatively affected the barcode localization and decoding performance. In our future work, we will design a more sophisticated approach (e.g. finger gestures) to eliminate accidental images.

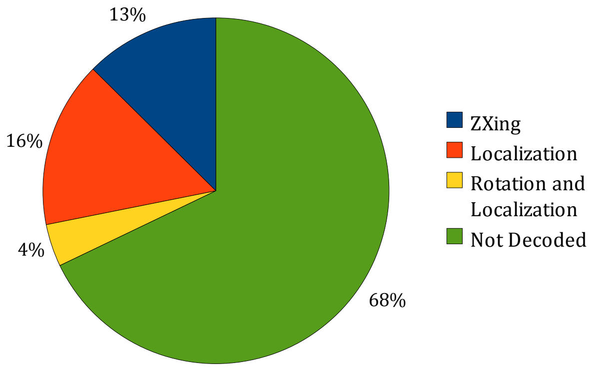

Figure 7: Results of Barcode Scans.

5 Conclusion

We presented an eyes-free vision-based barcode localization and decoding method that enables visually impaired (VI) mobile phone users to decode MSI (Modified Plessy) barcodes on shelves and UPC barcodes on individual boxes, cans, and bottles. We showed how simple and efficient barcode localization and decoding techniques were augmented with an interactive haptic feedback loop that allows the VI user to align the phone's camera with a fixed surface in the pitch and yaw planes. Our method is implemented on a Google Nexus One smart phone running Android 2.1. We described a laboratory study in which the method was evaluated by one VI and four blindfolded sighted participants. All participants were able to retrieve and verify all the products successfully. Our touch screen user interface should be improved to eliminate accidental images. Experiments are being planned with VI individuals in a real grocery store.

Acknowledgments

This research has been supported, in part, through NSF grant IIS-0346880.

References

[1] Food Marketing Institute Research. *The Food Retailing Industry Speaks 2006.* Food Marketing Institute 2006.

[2] Nicholson J., Kulyukin V., Coster D. *ShopTalk: Independent Blind Shopping Through Verbal Route Directions and Barcode Scans* The Open Rehabilitation Journal, ISSN: 1874-9437 Volume 2, 2009, pp. 11-23, DOI 10.2174/1874943700902010011.

[3] Nicholson J. and Kulyukin V. *ShopTalk: Independent Blind Shopping = Verbal Route Directions + Barcode Scans.* Proceedings of the 30-th Annual Conference of the Rehabilitation Engineering and Assistive Technology Society of North America (RESNA 2007), June 2007, Phoenix, Arizona, Avail. on-line and on CD-ROM.

[4] Janaswami, K. (2010). *ShopMobile: A Mobile Shopping Aid for Visually Impaired Individuals.* M.S. Report, Department of Computer Science, Utah State University, Logan, UT, to appear.

[5] Kulyukin V., Nicholson J., and Coster D. *ShopTalk: Toward Independent Shopping by People with Visual Impairments.* Proceedings of the 8th ACM Conference on Computers and Accessiblity (ASSETS 2008), Halifax, Canada.

[6] Normand N. and Viard-Gaudin C. *A Two-Dimensional Bar Code Reader.* Pattern Recognition, Vol. 3, pp. 201-203, 1994.

[7] Viard-Gaudin C., Normand N., and Barba D. *Algorithm Using a Two-Dimensional Approach.* In Proceedings of the Second Int. Conf. on Document Analysis and Recognition, No. 20-22, pp. 45-48, October 1993.

[8] Ando S., and Hontanj H., *Automatic Visual Searching and Reading of Barcodes in 3-D Scene, In Proceedings of the IEEE Int. Conf. on Vehicle Electronics.* No. 25-28, pp. 49-54, September 2001.

[9] Ando S. *Image Field Categorization and Edge/Corner Detection from Gradient Covariance.* IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 22, No. 2, pp. 179-190, February 2000.

[10] Ando S., *Consistent Gradient Operators.* IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 22, No. 3, pp. 252-265, March 2000.

[11] Muniz R., Junco L., and Otero A., *A Robust Software Barcode Reader using the Hough Transform.*, In Proceedings of International Conference on Information Intelligence and Systems, No. 31, pp. 313-319, November 1999.

[12] Arnould S., Awcock G. J., and Thomas R., *Remote Bar-code Localization Using Mathematical Morphology.*, Image Processing and its Applications, Vol. 2, No. 465, pp. 642-646, 1999.

[13] *The Zebra Crossing Barcode Decoding Library* http://code.google.com/p/zxing/

[14] Occipital, LLC. *RedLaser* http://redlaser.com/

[15] Tekin, E. and Coughlan, J.M., *An algorithm enabling blind users to find and read barcodes* WACV09, 2009

[16] Kutiyanawala A., Qi X., Tian J. *A Simple and Efficient Approach to Barcode Localization.* 7^{th} International Conference on Information, Communication and Signal Processing, Macau 2009

[17] Fisher R., Perkins S., Walker A. and Wolfart E. *Line Detection.* http://homepages.inf.ed.ac.uk/rbf/HIPR2/linedet.htm