Volume 4, Issue 11, November 2014 ISSN: 2277 128X

International Journal of Advanced Research in Computer Science and Software Engineering

Research Paper

Available online at: www.ijarcsse.com

A Study of Computer Vision Techniques for Currency Recognition on Mobile Phone for the Visually Impaired

Dhanashri Mulmule, Amit Dravid
Deptt. of Computer Engg. & SPPU
Savitribai Phule Pune University, Pune, India

Abstract— Currency recognition is a challenging task in computer vision. Various Image processing methods and artificial intelligence techniques have been employed to handle this. We have studied various currency recognition applications employing different keypoint detection, keypoint extraction and keypoint matching approaches. In this paper we also present a study of some recent computer vision techniques like BRIEF, ORB, FAST, AGAST, BRISK, FREAK and earlier techniques like SURF, SIFT. The mobile based currency recognition applications discussed here are meant to support visually impaired and blind users.

Mobile devices are ubiquitous and come with a built in camera having a fair resolution. Currency note images captured by blind users can have several failings: images of the notes may not be ideally aligned/oriented; there can be scale changes due to variation of distance from camera; illumination changes can also lead to differences in images. There is also the possibility of a cluttered background in the image, or the note being partially occluded, folded, worn and/or wrinkled, etc. This paper illustrates computer vision techniques employed till now and studies new improved techniques which can be used in their place. These new methods offer several advantages for efficient currency note recognition in mobile applications. We have also studied blind users' needs and expectations from such applications and how researchers have developed variety of touch screen utilities like Braille keyboard, food product identification, and indoor navigation applications. Also, our in person visit to an NGO for blind helped us to understand their special needs from touch screen mobile devices. This visit has really motivated and has enriched our work in correct way.

Keywords— Currency note recognition, Computer Vision, Visually impaired blind users, Assistive mobile phone applications-FREAK, BRISK, SURF, SIFT

I. INTRODUCTION

Most of us are fortunate as we can see the world with our eyes. But many of us in the world are not that privileged. Visually impaired people have a reduced vision or no vision at all and face difficulties in carrying out day to day activities. Globally 285 million people are estimated to be visually impaired and 39 million are blind and 246 have low vision [7]. India shares the largest burden of global blind population of nearly 15 million people [9]. Our basic intent is to study assistive mobile applications for Indian currency note identification for blind, visually impaired and elderly people. One basic difficulty with these people is identifying banknotes denomination. When exchanging cash sometimes blind people have to rely on assistance from others or keep the banknote folded in certain way in their pockets. In India though we have debit/credit cards, exchanging cash is still an inevitable part of our daily routine. Indian currency have Rs.10,Rs.20,Rs.50,Rs.100,Rs.500 and Rs.1000 banknotes which differ in size, color and have different identification marks but still create difficulty for a blind user in distinguishing the denomination as the currency size variation is negligible and identification marks [8] and colour fades away due to circulation over time or due to dirt.

Now-a-day's touch screen mobile phones, tablets and phablets have built in camera and are widely used by all. These devices are used by visually impaired population with utilities like Talks (Symbian OS) and Google's Talkback. Also now-a-days mobile devices come with good hardware configuration and high speed processor support in economical costs. Such devices can be used to support an automatic banknote recognition application.

We have studied the phases of a mobile based currency recognition system and it has in general the following phases: A currency note image is captured by digital camera built in mobile device. The captured image is in digital format.

Image Pre-processing: Currency image captured appears different if taken under varying lighting conditions. To reduce this effect of light source on image, it is converted into gray scale. In pre-processing, image can be scaled to a particular pixel height and width for further processing. Also some applications does background subtraction to process further only area of interest.

Key-point detection: It mainly identifies points in the digital image at which the brightness changes sharply which can be edges, corners or blobs. The detector should be repetitive in detection of key points. Here each image is searched for locations that are likely to match well in other images [30].

Key-point description: It is the most important part of feature vector generation. Each region around a detected key point location is converted into a more compact and stable (invariant) descriptor that can be matched against other descriptor. The aim of descriptor formation is to construct a descriptor with unique and distinguishable key points. It is the most challenging task to generate feature vector out of detected key points [30].

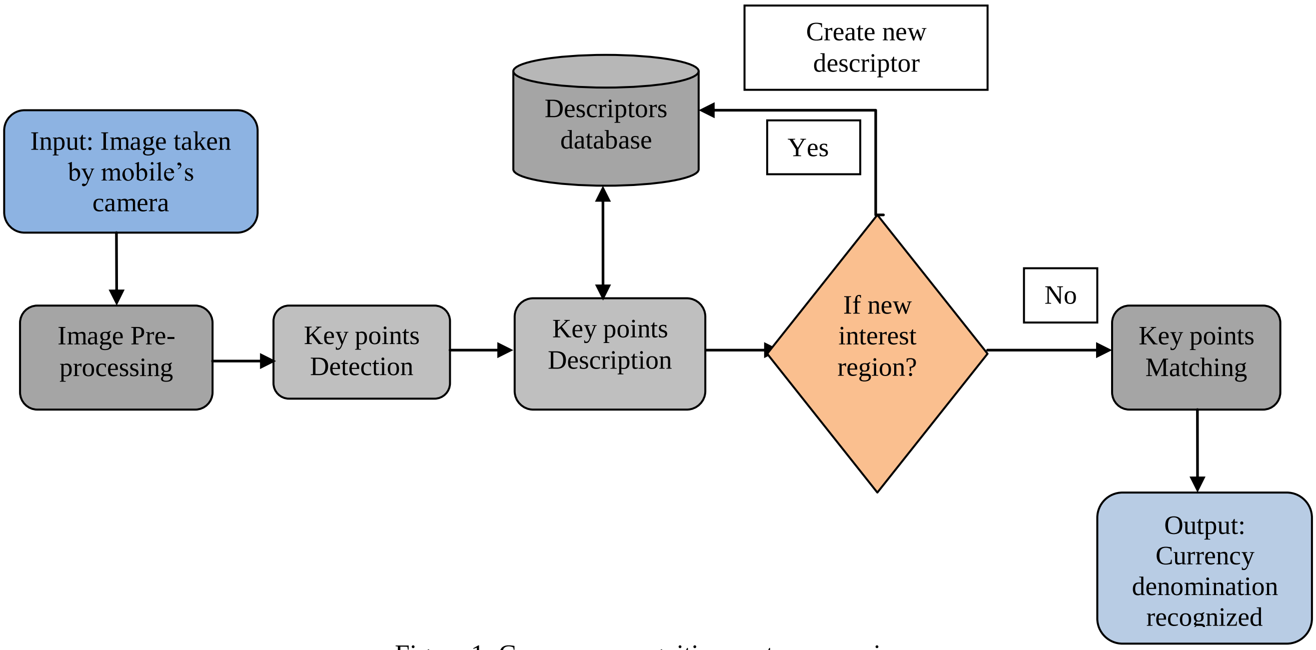


Figure 1: Currency recognition system overview

Key-point matching: In this last phase, extracted features are compared with already known features from database. Best matches are found between two feature vectors. Different classification approaches like nearest neighbor and nearest to second nearest neighbor ratio, k-mean clustering etc. are being used for matching monetary features between training and testing images.

We have studied the published literature on camera based currency recognition applications. While capturing currency note images by a visually impaired user, it is not assured that the entire currency bill is captured and the image may not always contain the place where the currency amount is written.There are real world challenges in capturing scenarios like image can be captured in poor illumination conditions, currency notes may be folded, dirtied, partially occluded, different orientation viewpoints, scale changes, worn and wrinkled bills, etc. To cope with such challenges, more robust, scale and rotation invariant computer vision techniques are to be chosen [18, 19, 20, 21].

Designing a user interface for the visually impaired users is also a challenging task. The touch-based mobile devices have opened new ways of user interaction for every kind of user. Nektarios Paisios [11] presented comparative evaluation of four different touch screen text entry methods which are: 1) QWERTY text-entry method is identical to Apple's VoiceOver and consists in the traditional computer keyboard layout with a screen reading software; 2) MultiTap which is the touch screen counterpart of the original keypad text-entry method; 3) With NavTouch users can perform gestures anywhere on the screen, and 4) BrailleType provides Braille approach and the touch screen serves as a representation of the Braille cell. Schneider et al. [15] have presented an algorithm which identifies facial features on android platform on mobile phone and even authenticates the user. Zhao et al. [16] presents an android application which authenticates mobile users by touch gestures like flick up/down, zoom in/out, and pinch in/out.

Blind users face enormous challenges while doing any day to day tasks which are very simple for a normal person to do. Like recognizing food products, indoor navigation at unknown place, outdoor navigation are always a challenging activities to name a few. There are many mobile applications supported by different mobile operating systems like iOS and android that aid various assistive applications for users with special need. Matusiak et al. [10] designed a mobile application whose one module recognizes scanned objects to a database of objects, e.g. food or medicine containers. The two other modules were meant to detect major colors and locate direction of the maximum brightness regions in the captured scenes. Oliveira et al. [13] have designed four assistive applications for blind and visually impaired users: 1) a simple to operate mobile navigational guide that can help its visually impaired users repeat paths that were already walked once in any indoor environment; 2) Mobile Brailler, with several prototype methods of text entry on a modern touch screen mobile phone that are based on the Braille alphabet and thus are convenient for visually impaired users; 3) Enables visually impaired users to leverage the camera of a mobile device to accurately recognize currency bills; 4) enables visually impaired users to determine whether a pair of clothes, in this case of a tie and a shirt, can be worn together or not, based on the current social norms of color-matching.

We carried out our own informal survey by visiting Niwant Andha Mukta Vikasalay (NAMV), Vidya Nagar, Pune, (Maharashtra state, India). This is a non government organisation which works for blind students who have passed 12th grade and are willing to stand on their own feet. They provide academic support for further studies, computer labs and facilitate earn and learn environment. The organization have Braille printing machines which is run by these students for printing and supplying educational Braille books to other parts of the Maharashtra state. They educate these youngsters for various computer programming languages and they are proficient in languages like C, C++, Java etc. to name a few.

We interacted with a few bright students from this organisation who were well aware of latest mobile applications. The blind individuals were using all latest mobile applications along with assistive applications. Applications they were using are banking, email readers, outloud readers, google docs, google map, facebook, twitter, an outdoor navigation applications and clothes colour detecting applications. They are mostly using android touch screen phones with Google's Talkback utility. Earlier they were using Symbian based Talks utility as screen reader on older mobile phones. We received a feedback that they feel the need for a mobile for recognizing Indian currency. This feedback has fuelled our interest to do research in currency recognition applications. Thanks to these interactions we came to know more about a real blind user's perspective and his/her expectations from such assistive applications.

An efficient banknote recognition requires sufficient and distinctive monetary features extraction, which should be invariant to scale, viewpoint changes and must be reliable. We studied earlier and recent keypoint detectors and descriptors such as Scale Invariant Feature Transform (SIFT) [17] and Speeded Up Robust Feature (SURF) [18], Binary Robust Independent Elementary Features (BRIEF)[19], Oriented Fast and Rotated BRIEF (ORB) [29], The Binary Robust Invariant Scalable Keypoints (BRISK)[20], coined Fast Retina Keypoint (FREAK)[21]. We have also studied recent keypoint detectors like Fast Accelerated Segment Test (FAST) [27], and Adaptive and Generic corner detection based on Accelerated Segment Test (AGAST) [28].

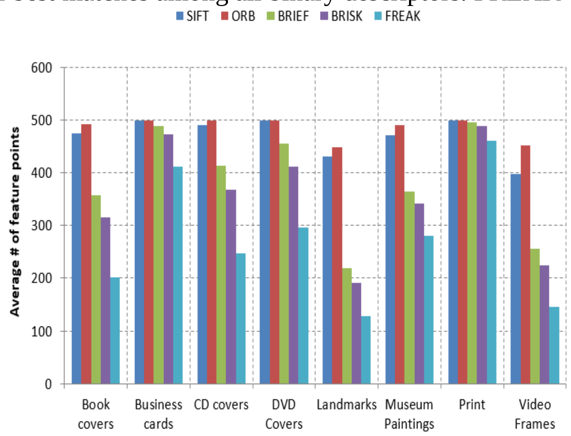
II. RELATED WORK

Computer vision techniques involve methods for acquiring, processing, analysing, and understanding images. Since past decade several methods of keypoint detection, description and matching are developed for fast computation and to provide robustness. These techniques allow comparison of "keypoints" in images even when the two images have changes in orientation, scale, illumination. To an extent, these techniques are able to compare images despite artifacts introduced by image compression techniques (like jpeg). Earlier techniques to name like Scale Invariant Feature Transform (SIFT) by D. Lowe [17], a 128-dimensional vector is obtained from a grid of histograms of oriented gradient. Its high descriptive power and robustness to illumination change have ranked it as the reference keypoint descriptor for the past decade. It is computationally expensive. Speeded Up Robust Features (SURF) proposed by Bay et al. [18] creates a grid around the keypoint and divides each grid cell into sub-grids. At each sub-grid cell, the gradient is calculated and is binned by angle into a histogram whose counts are increased by the magnitude of the gradient, all weighted by a Gaussian. These grid histograms of gradients are concatenated into a 64-dimensional vector. SURF is fast because it uses integral images, which drastically reduces the number of operations for simple box convolutions, independent of the chosen scale. It outperforms previous techniques in terms of repeatability, robustness and distinctiveness. The Binary Robust Invariant Scalable Keypoints (BRISK) proposed by Leutenegger et al. [20] is a detector and descriptor. BRISK detector is scale-invariant and it searches keypoints in the original image plane. It is faster than SURF. The Binary Robust Independent Elementary Feature (BRIEF) [19] proposed by Calonder et al. It uses a sampling pattern consisting of 128, 256, or 512 comparisons equating to 128, 256, or 512 bits, with sample points selected randomly from an isotropic Gaussian distribution centered at the feature location. It has low computational and storage requirements. The Oriented Fast and Rotated BRIEF (ORB) by Rublee et al. [29], is similar to BRISK, is a refined, scale-invariant version of FAST. It has fixed point operations and have low memory requirement compared to previous techniques. The Fast Accelerated Segment Test (FAST) [27] detector is high-speed corner detector and it is based on the Accelerated Segment Test. The authors have presented a machine learning approach to create decision trees that allows FAST to classify a candidate point with only a few pixel tests, thus speeding-up the detection process. Mair et al. [28] proposed adaptive and generic corner detector based on accelerated segment test (AGAST). It outperforms previous detectors in repeatability and has low memory requirements. Another recent descriptor, coined as Fast Retina Keypoint (FREAK) by Alahi et al. [21], computes a cascade of binary strings by comparing image intensities over a retinal sampling pattern. It has overlapping receptive fields and kernel size changes exponentially. It uses 43 receptive fields leading to approximately one thousand pairs. To estimate the rotation of keypoints detected, it sums the estimated local gradients over selected 45 pairs. FREAK implementation takes advantage of Single Instruction and Multiple Data (SIMD) instructions on Intel processor since the operations are performed in parallel. Experiments in [21] shows that FREAKs are in general faster to compute with lower memory load and also more robust than SIFT, SURF or BRISK.

Over past several years, much work has been done on various assistive applications and devices for people with special needs. Further there are many new and general mobile image recognition applications like Google Goggles (www.google.com/mobile/goggles) which identifies landmarks and logos, to recognize book and CD covers, SnapTell (www.snaptell.com) like applications are in use. Over past few decades many currency recognition applications were developed but with limitations and having specific environments. Hasanuzzaman et al. [1] have a novel camera-based computer vision technology to automatically recognize American Dollar banknotes for assisting visually impaired people. They have proposed a component-based framework based on Speeded Up Robust Features (SURF). Features selected as reference regions were descriptive enough for each banknote category for matching with query images. A variety of conditions including occlusion, rotation, scaling, cluttered background, illumination change, viewpoint variation, and worn or wrinkled bills are considered and further by blind subjects it have been verified on proposed algorithm. Further they proposed automatic thresholding to learn all the thresholds from training images so as to extend the application for other countries currency recognition. Non currency images are handled and spatial clustering is used to guide the user if there is no currency bill in the camera view. Although this banknote recognition system provides robustness with features like high accuracy and high efficiency they rely on SURF which was developed in 2006. Nektarios et al. [2] proposed US Dollar bill recognition application using camera of a Smartphone to recognize partial and even distorted images of paper bills. An adaptation of Scale Invariant Feature Transform is used on a stream of continuous images. SIFT algorithm is used to generate the key-points of all the training images and the speed of recognition is enhanced by employing the K-means clustering technique to separate and reduce the number of the feature vectors of the training. They achieved around 93% detection accuracy among four categories of US currency bills and found misclassification for currency bills poorly shot images. Paisios et al. [3], have applied SIFT detectors and descriptors with Nearest Neighbor and Nearest to Second Nearest Neighbor Ratio approaches to improve the accuracy of classification of test keypoints. Xu Liu [4] proposed a currency recognition application which processed 10frames/second to capture high quality image. Ada-boost framework is used for training images and efficient background subtraction and perspective correction algorithms are implemented in this US dollar currency recognition application. It supported smartphone platforms like Symbian and Windows Mobile. The recognition result given to user via the phone speaker or can be communicated through vibration. Papastavrou et al. [5] proposed a Euro, polish Zloty and most of the US Dollar denominations recognition application based on computer vision and pattern recognition technique. They converted the image into gray scale by a novel adaptive segmentation algorithm. The application is portable and runs in real time on smartphones symbian, windows mobile and iphone with recognition speed 0.1 seconds. Jahangir et al. [6] have designed an application which takes scanned images of banknotes and then fed into a Multilayer Perceptron, trained by Backpropagation algorithm, for recognition. They used axis Symmetric Masks in preprocessing stage to reduce the network size. The application could recognize currently available 8 Bangladeshi denominations successfully with an average accuracy of 98.57%.

III. DISCUSSION

We have studied comparison and evaluation of the traditional or floating point descriptors against the binary descriptors. Heinly et al. [24] have done a performance evaluation of traditional and novel binary features. For detectors, they found that FAST has the highest entropy, which attributed to its good spread of points. Detectors like SIFT, BRISK, and SURF, have the next highest entropies. Further they noted that binary descriptors showed speedup and algorithmic efficiency which is important in mobile device applications. For descriptors, they found the precision of BRISK is higher than its binary competitors like BRIEF, ORB. It achieved very high quality matches. SIFT gave the best performance, and proved to be the most robust to changes of image viewpoints. Cameron Schaeffer [26] has done a comparison of speed and accuracy of BRISK and FREAK with SURF in the context of a pedestrian detection image database. SURF underperformed compared to FREAK and BRISK in accuracy. FREAK provides a higher rate of matching correspondences. SURF scored higher recall and lower precision than BRISK and FREAK. Bekely et al. [25] evaluated performance of binary descriptors like BRIEF, ORB, BRISK and FREAK. They used Stanford Mobile Visual Search (SMVS) data set because binary descriptors are mainly used in mobile applications. They used Oxford dataset only for validation. Their result showed that BRISK keypoint descriptor gives highest percentage of precision and largest number of best matches among all binary descriptors. FREAK offered comparably good result with respect to BRISK.



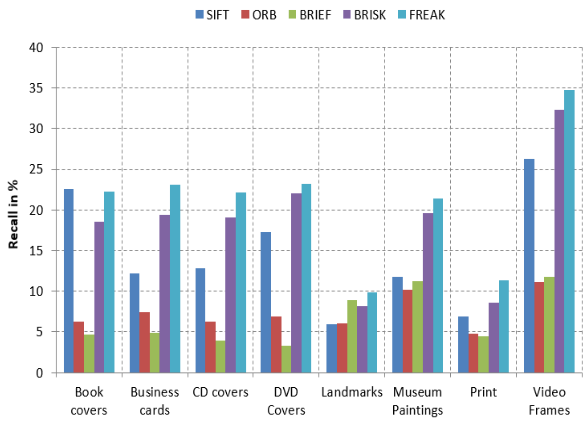


Figure 2: Average number of feature points for binary descriptors. Figure 3: Recall in Percentage

(Figure 2 and Figure 3 Source: Evaluation of binary keypoint descriptors [25])

Table 1compares SIFT, SURF, BRISK and FREAK descriptors [21], description and matching time. 1500 Keypoints are detected per image of size 800x600. The computation times correspond to the description and matching of all keypoints. It shows FREAK is faster than BRISK although the latter is two orders of magnitude faster than SIFT and SURF.

Table 1: Computation time of all detected keypoints

Time per keypoint	SIFT	SURF	BRISK	FREAK
Description in [ms]	2.5	1.4	0.031	0.018
Matching time in [ns]	1014	566	36	25

(Source: FREAK: Fast Retina Keypoint [21])

Further we studied a comprehensive comparison paper [22] by Canclini et.al, which has considered various combinations of pairs of detector/descriptor and has demonstrated their performances in terms of processing time, matching accuracy etc. They presented Average processing time of visual descriptor algorithms for the computation of

500 descriptors. They observed the computation of binary descriptors like BRIEF, ORB, BRISK and FREAK is particularly efficient compared to that of non binary descriptors. For instance, BRISK is 6 times faster than SURF and 20 times faster than SIFT.

IV. CONCLUSIONS

In this paper we have studied different currency note recognition applications using various computer vision techniques. These applications have been designed for blind users as assistive tools with suitable user interfaces. Different currency recognition applications have helped us to understand various computer vision issues related to image capturing, and in choosing keypoint detectors and descriptors. An outcome of this study is the comparative study of traditional and recent detectors and descriptors. This clearly indicates that much improvement in terms of performance is possible with the use of newer methods like FREAK and BRISK in place of SURF and SIFT. Our informal survey with blind people has clearly shown that blind people are quite adept at using smartphones, given a proper user interface, and they stand to benefit much from the development of an assistive Indian currency note recognition application. We hope to use this study as a motivation to actually develop an android based application for the said purpose.

REFERENCES

[1] Hasanuzzaman, Faiz M., Xiaodong Yang, and YingLi Tian. "Robust and effective component-based banknote recognition for the blind." *Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on* 42.6 (2012): 1021-1030.

[2] Paisios, Nektarios, et al. "Recognizing currency bills using a mobile phone: an assistive aid for the visually impaired." *Proceedings of the 24th annual ACM symposium adjunct on User interface software and technology*. ACM, 2011.

[3] Paisios, Nektarios, Alexander Rubinsteyn, and Lakshminarayanan Subramanian. "Exchanging cash with no fear: A fast mobile money reader for the blind." *Workshop on Frontiers in Accessibility for Pervasive Computing. ACM*. 2012.

[4] Liu, Xu. "A camera phone based currency reader for the visually impaired." *Proceedings of the 10th international ACM SIGACCESS conference on Computers and accessibility*. ACM, 2008

[5] Papastavrou, Stavros, Demetris Hadjiachilleos, and Georgios Stylianou. "Blind-folded recognition of bank notes on the mobile phone." *ACM SIGGRAPH 2010 Posters*. ACM, 2010.

[6] Jahangir, Nadim, and Ahsan Raja Chowdhury. "Bangladeshi banknote recognition by neural network with axis symmetrical masks." *Computer and information technology, 2007. iccit 2007. 10th international conference on*. IEEE, 2007.

[7] http://www.who.int/mediacentre/factsheets/fs282/en/

[8] http://www.rbi.org.in/currency/Security%20Features.html

[9] http://timesofindia.indiatimes.com/india/India-has-largest-blind-population/articleshow/2447603.cms

[10] Matusiak, K., P. Skulimowski, and P. Strurnillo. "Object recognition in a mobile phone application for visually impaired users." *Human System Interaction (HSI), 2013 The 6th International Conference on*. IEEE, 2013.

[11] Nektarios Paisios, "Mobile Accessibility Tools for the Visually Impaired" 2012

[12] Mawafo, Josette C. Tagatio, W. A. Clarke, and P. E. Robinson. "Identification of facial features on android platforms." *Industrial Technology (ICIT), 2013 IEEE International Conference on*. IEEE, 2013.

[13] Oliveira, João, et al. "Blind people and mobile touch-based text-entry: acknowledging the need for different flavors." *The proceedings of the 13th international ACM SIGACCESS conference on Computers and accessibility*. ACM, 2011.

[14] Dakopoulos, Dimitrios, and Nikolaos G. Bourbakis. "Wearable obstacle avoidance electronic travel aids for blind: a survey." *Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on* 40.1 (2010): 25-35.

[15] Schneider, Christian, et al. "Feature based face localization and recognition on mobile devices." Control, Automation, Robotics and Vision, 2006. ICARCV'06. 9th International Conference on. IEEE, 2006

[16] Zhao, Xi, et al. "Mobile User Authentication Using Statistical Touch Dynamics Images." (2014).

[17] Lowe, David G. "Distinctive image features from scale-invariant keypoints." *International journal of computer vision* 60.2 (2004): 91-110.

[18] Bay, Herbert, Tinne Tuytelaars, and Luc Van Gool. "Surf: Speeded up robust features." *Computer Vision-ECCV 2006*. Springer Berlin Heidelberg, 2006. 404-417.

[19] Calonder, Michael, et al. "Brief: Binary robust independent elementary features." *Computer Vision–ECCV 2010*. Springer Berlin Heidelberg, 2010. 778-792.

[20] Leutenegger, Stefan, Margarita Chli, and Roland Yves Siegwart. "BRISK: Binary robust invariant scalable keypoints." *Computer Vision (ICCV), 2011 IEEE International Conference on*. IEEE, 2011.

[21] Alahi, Alexandre, Raphael Ortiz, and Pierre Vandergheynst. "Freak: Fast retina keypoint." *Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on*. Ieee, 2012.

[22] Canclini, A., et al. "Evaluation of low-complexity visual feature detectors and descriptors." *Digital Signal Processing (DSP), 2013 18th International Conference on*. IEEE, 2013.

[23] Brady, Erin, et al. "Visual challenges in the everyday lives of blind people." *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 2013.

[24] Heinly, Jared, Enrique Dunn, and Jan-Michael Frahm. "Comparative evaluation of binary features." Computer Vision–ECCV 2012. Springer Berlin Heidelberg, 2012. 759-773.

[25] Bekele, Dagmawi, Michael Teutsch, and Tobias Schuchert. "Evaluation of binary keypoint descriptors." ICIP. 2013.

[26] Schaeffer, Cameron. "A Comparison of Keypoint Descriptors in the Context of Pedestrian Detection: FREAK vs. vs. BRISK." 2012

[27] Rosten, Edward, and Tom Drummond. "Machine learning for high-speed corner detection." *Computer Vision–ECCV 2006*. Springer Berlin Heidelberg, 2006. 430-443.

[28] Mair, Elmar, et al. "Adaptive and generic corner detection based on the accelerated segment test." *Computer Vision–ECCV 2010*. Springer Berlin Heidelberg, 2010. 183-196.

[29] Rublee, Ethan, et al. "ORB: an efficient alternative to SIFT or SURF." *Computer Vision (ICCV), 2011 IEEE International Conference on*. IEEE, 2011.

[30] Richard Szeliski, Computer Vision: Algorithms and Applications, Microsoft Research © 2010

