Recognition of Banknotes in Multiple Perspectives Using Selective Feature Matching and Shape Analysis

3 authors:

Carlos Miguel Correia da Costa
Institute for Systems and Computer Engineering of Porto (INESC Porto)
35 PUBLICATIONS   **12** CITATIONS

Germano Veiga
Institute for Systems and Computer Engineering of Porto (INESC Porto)
65 PUBLICATIONS   **174** CITATIONS

Armando Jorge Sousa
University of Porto
89 PUBLICATIONS   **208** CITATIONS

Some of the authors of this publication are also working on these related projects:

Robotic Technologies for Non-Standard Design and Construction in Architecture (RobTech) View project

STAMINA View project

Recognition of Banknotes in Multiple Perspectives Using Selective Feature Matching and Shape Analysis

Carlos M. Costa
INESC TEC
and Faculty of Engineering,
University of Porto
Email: carlos.m.costa@inesctec.pt

Germano Veiga
INESC TEC
Email: germano.veiga@inesctec.pt

Armando Sousa
INESC TEC
and Faculty of Engineering,
University of Porto
Email: asousa@fe.up.pt

Abstract—**Reliable banknote recognition is critical for detecting counterfeit banknotes in ATMs and help visual impaired people. To solve this problem, it was implemented a computer vision system that can recognize multiple banknotes in different perspective views and scales, even when they are within cluttered environments in which the lighting conditions may vary considerably. The system is also able to recognize banknotes that are partially visible, folded, wrinkled or even worn by usage. To accomplish this task, the system relies on computer vision algorithms, such as image preprocessing, feature detection, description and matching. To improve the confidence of the banknote recognition the feature matching results are used to compute the contour of the banknotes using an homography that later on is validated using shape analysis algorithms. The system successfully recognized all Euro banknotes in 80 test images even when there were several overlapping banknotes in the same test image.**

Index Terms—**banknote recognition, feature matching, computer vision, shape analysis**

I. INTRODUCTION

Banknotes play a critical role in our trading society, and although digital currency is becoming popular, physical banknotes still account for a large amount of the local transactions. As such, systems that are able to recognize banknotes can be applied to aid in the manipulation of this type of currency. These kind of systems are critical for visually impaired people [1], since they allow them to be more independent while avoiding the help of untrusted people. Also, they can increase the security and reliability of ATMs [2], by making sure the maintenance operations are performed correctly and the valid banknotes are not replaced with counterfeits. Other less critical applications are related with automatic sorting and counting of banknotes to speedup transactions and money transfers. For these types of systems to be effective and useful, they must be able to recognize folded, wrinkled and worn banknotes in several perspective views, scale dimensions, and should also tolerate cluttered environments with different lighting conditions.

With these goals in mind, it was implemented a computer vision system capable of reliably recognize banknotes using a camera. It starts by preprocessing the input images in order to remove environment noise and improve contrast and brightness. Then relevant keypoints and their associated descriptors are computed to later be used to find the best matching in a database of valid banknotes. The correct matching of keypoint descriptors is critical to ensure the proper recognition of the banknotes. There are several techniques to perform descriptor filtering, such as the ratio test [3] and the homography outlier removal [4]. Although these techniques can yield very good results, a post-processing shape analysis is applied to make sure the results obtained are really banknotes. This is related to the fact that the matching of several parts of wrinkle banknotes may result in the recognition of multiple instances of the same banknote. In addition, images similar to banknotes or from other countries currencies may yield incorrect partial matches. As such, this post-processing phase is critical to ensure the correct recognition of the banknotes. This phase starts by computing the banknote contour using the retrieved homography, and then removes any banknote recognition result that has a convex contour, or has its area, circularity and aspect ratio outside the acceptable ranges for banknotes. To detect multiple banknotes in the same image, the inliers of the last recognized banknote are removed and the process presented earlier is repeated until there are no more valid matches.

In the following section it will be presented an overview of several approaches that can be used to perform banknote recognition. In Section III a detailed description of the implementation will be provided. In Section IV the representative results of the recognition system will be given and in Section V it will be discussed the robustness and reliability of the system. Finally, the conclusions will be given in Section VI.

II. RELATED WORK

There are several approaches that can be used to successfully recognize banknotes [5], and they range from simple but less robust techniques to more advanced and accurate systems [1]. Template matching is one of the simplest recognition methods, which tries to find the banknotes in an image by simple bitmap comparison. But this approach has the limitation

that both the reference banknotes and its targets in the image must have the same size and perspective view. To mitigate this restriction, the comparison could be done in several scales and orientations, but it would not be very efficient. The dynamic template matching proposed in [6] could be used instead, but it still is not the best way to handle the recognition. Another way to tackle this problem would be to perform color and shape segmentation of specific parts of the banknotes, like the method proposed in [7]. But this leads to a very specific implementation that would require fine tunning for each banknote, and it would need a lot of effort to successfully recognize all banknotes from both sides. A more broad implementation could use the size, color and texture of each banknote to perform the recognition [8], but like the previous case, it would require extensive parameterization to recognize each banknote, and would have to take in consideration the accuracy of the distance measurements, because several types of banknotes may have similar sizes. A more robust implementation could use Principal Component Analysis or even adapt the eigenfaces algorithm to try to recognize the banknotes [9], but this can have some problems when the perspective of the reference banknotes is very different from the ones in the image.

For detection of counterfeit banknotes, ultra violet or infra-red light could be used to highlight specific parts of the banknotes that are hard to duplicate and easier to recognize [10]. Other similar technique takes advantage of the fact that specific parts of the banknotes are highlighted when they are illuminated with LEDs with different colors and intensities [11]. Another approach takes in consideration the electromagnetic fields present in sections of the banknotes to perform the recognition [12]. But all these techniques require special hardware that is too expensive. Moreover, they are not meant to be used by visually impaired people.

Some of the most common techniques used to perform banknote recognition rely on machine learning algorithms, such as Support Vector Machines [13], Artificial Neural Networks [14], and Hidden Markov Models [15]. These techniques usually apply some sort of clustering of features before training the classifier, such as the Bag of Keypoints model [16], or try to extract relevant features from the reference images. After the training, the classifiers can be used to recognize the banknotes. Although this is a good approach for general recognition, it may not be very precise in calculating the exact location and contour of the banknotes.

III. Implementation

The following sections present the main processing stages of the implemented banknote recognition system (overview shown in Figure 1). The C++ source code along with the complete results are available at [1]. To speed up development, the Open Source Computer Vision (OpenCV) library was used.

A. Reference image database setup

In order to be able to recognize the target banknotes, a database of valid instances must be computed. This database

[1]https://github.com/carlosmccosta/Currency-Recognition
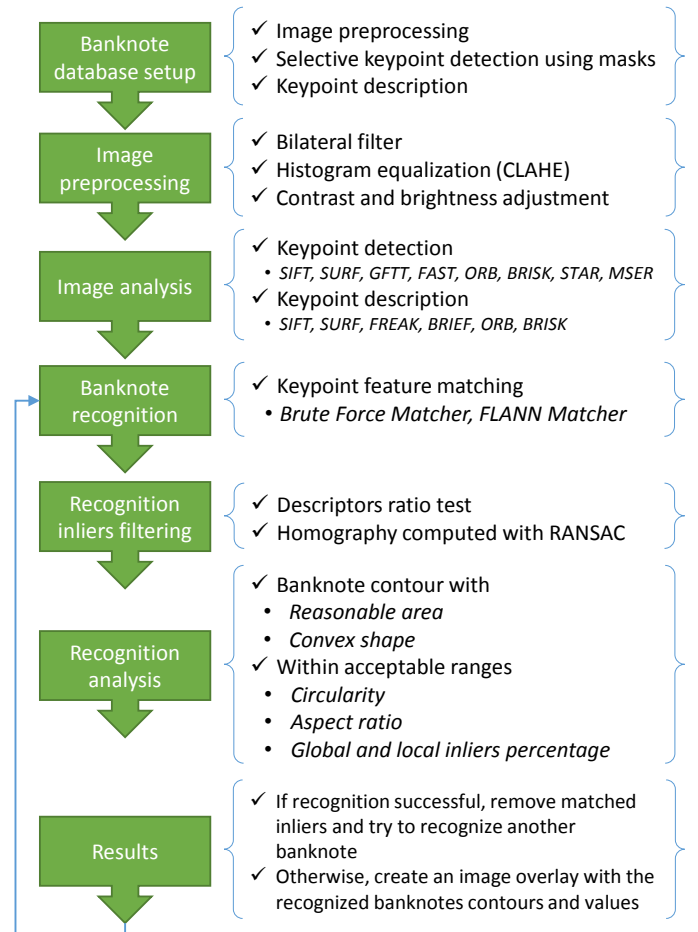

Figure 1. Overview of the main processing stages of the banknote recognition system

contains the descriptors associated with the keypoints for each preprocessed banknote (from both sides). To improve the detection of the relevant parts of the banknotes and to avoid the usage of sections that are similar across several banknotes, the keypoint detection algorithm is only applied inside the manually created masks associated with each banknote. Only relevant parts such as the banknote number and unique textures or patterns are included in the banknotes masks (only the white areas shown in Figure 2 will be used to compute the keypoints of a 500 € banknote).
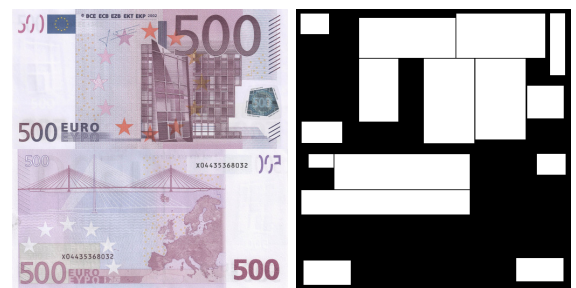

Figure 2. Front and back of a 500 € banknote with associated feature detection masks

B. Preprocessing

To improve the detection of good features and ensure that the system has robust recognition even when the images have considerable noise, a preprocessing step is applied. In a first phase, most of the noise is removed using a bilateral filter. This filter was chosen because it preserves the edges of image blobs, which are very valuable structures in the detection of feature points. After the noise is reduced, a Contrast Limited Adaptive Histogram Equalization (CLAHE) is applied to increase the contrast. This can improve the recognition of the system when the images are taken in low light environments. This technique has better results over the simple histogram equalization because it can be applied to images that have areas with high and low contrast, and also limits the spread of the noise. Finally, it is also possible to manually tune the contrast and brightness. Figure 3 shows the impact of the preprocessing stage in an image which had camera sensor noise and high brightness shifts due to wrinkled plastic.

Figure 3. Impact of preprocessing stage (original image on the left, preprocessed image on the right)

C. Recognition

The recognition is the most critical phase in the system, in which the provided image is analyzed to extract the banknotes monetary value and their contour. The current implementation supports recognition of multiple banknotes in the same image, even if they are partially occluded. The system is able to recognize any type of banknotes. The Euro currency was selected for the computation of the results, but any other currency can be used. To setup the system for other currencies it is only necessary to replace the database images and masks with the intended currency images. To improve the robustness of the system, 3 levels of detail for each banknote are provided (with images having pixel width of 256, 512 and 1024 respectively). For an ideal banknote recognition result, the image resolution of both the reference and the target images should be the same. But converting a high resolution banknote database to the image banknotes resolution has a considerable processing overhead, and as such, to allow the system to be more efficient and able to run in real time, a compromise between precision and computation time was achieved by precomputing the reference images in 3 different scales. At run time, the appropriate level of detail is selected according to the resolution of the target image. The reason for several levels of detail is related to the fact that the geometry of the banknotes changes drastically from a low resolution to a high resolution banknote image. As a result, the computed

keypoints and their associated descriptors will be considerably different and the recognition of the banknotes will likely fail. This can be clearly seen in Figure 4. This approach mitigates this problem by selecting the most similar database image resolution in relation to the banknotes in the image being analyzed.

Figure 4. Impact of image resolution when computing SIFT keypoints from a very low resolution image (left) to a high resolution image (right) of the hologram of a 500 € banknote

1) Feature detection: Feature detection is the initial recognition step in which interesting keypoints for matching are identified in the image. These keypoints are normally selected by analyzing the edges, corners, blobs or even ridges. Also, the keypoints provide a condensed representation of the image, which significantly speeds up matching (compared to bitmap or blob matching). To allow fine tuning of the system, the implementation supports the usage of one of several feature detection algorithm, namely, Scale Invariant Feature Transform (SIFT), Speeded Up Robust Features (SURF), Good Features to Track (GFTT), Features from Accelerated Segment Test (FAST), Oriented FAST and Rotated BRIEF (ORB), Binary Robust Invariant Scalable Keypoints (BRISK), STARand Maximally Stable Extremal Regions (MSER).

2) Keypoint descriptor extraction: The feature description step associates to each keypoint a description of its surroundings, in order to allow the matching of keypoints. This normally involves the computation of n-jets or local histograms and the final result is a vector in an n-dimensional space characterizing each keypoint. In order to detect instances with different perspective views, these descriptors must be scale and rotation invariant. Also, they should tolerate different lightning conditions. There are several algorithms that can accomplish this task, and as such, they were included in the implementation and can be selected to fine tuning the system. It was included the SIFT, SURF, Fast Retina Keypoint (FREAK), Binary Robust Independent Elementary Features (BRIEF), ORBand BRISKfeature descriptors.

3) Descriptors matching: In order to detect multiple banknotes in the same image, a correct matching between the image descriptors and the reference banknotes descriptors must be established. This initial matching can be performed using either a brute force or a heuristic approach. In the brute force approach, each descriptor in the image is compared with all descriptors in the reference image to find the best correspondence. In the heuristic approach using the Fast Library for Approximate Nearest Neighbors (FLANN), several optimizations are employed to reduce the computation time. The main speedups are due to appropriate selection of which keypoints to match and also the usage of efficient data

structures to speed up n-dimensional search (such as k-d trees).

4) Inliers filtering: After the initial matching, an inliers filtering phase is applied. It starts by applying a ratio test [3] and then refines the results with the computation of a homography. In the ratio test, each image keypoint descriptor is associated with the two best reference image descriptors. This allows to decide if the matching is correct or not, by computing the ratio between the distances of these keypoint descriptors. The rationale behind it is that if the ratio is close to one then there are two points with equivalent match probability, and as such, it is very likely that this is an incorrect match and should be discarded.

The refinement of the inliers is performed with the computation of a homography that allows the mapping of the positions in the database reference image to the positions in the target image, in which the banknotes to be recognize reside. This is achieved by using the Random Sample Consensus (RANSAC) method to find the homography that best fits the detected keypoints. Since this is a RANSAC method, it iteratively tries to find better results by randomly selecting the supporting keypoints until there is a high confidence on the results (which occurs when the found homography transformation applied to the reference keypoints yields a high percentage of keypoints positions close enough to the target keypoints) or the maximum number of 5000 iterations is reached. The rationale behind using such method is that the geometry of the banknotes is planar, and if it is assumed that in most cases the banknotes that are going to be recognize have also planar geometry, then a homography can be used to analyze if a match is correct or not. This classification is performed by applying the transformation computed by the homography to each reference keypoint and check if a high percentage of the resulting points in the target image are close to the position of the matched keypoints.

Having the inliers, two approaches can be used to decide if a valid banknote was found or not. One technique relies in the computation of the global inliers ratio in relation to the number of keypoints, and considers that there was a correct match if this ratio is above a given threshold. This is the most appropriate method for most of the banknote recognition cases. Another method that may yield better results when the banknotes are partially occluded, is to consider a correct match when one of the components of the masks (shown in Figure 2) have the local inliers ratio above a given threshold. The idea behind this approach is to consider each patch of the image represented in the mask as a unique identifier of a banknote. As such, if this local patch is correctly detected, then there is a high confidence that the recognition of that banknote was successful.

5) Shape analysis: When a banknote recognition is considered valid, it undergoes a post-processing step in which its contour shape is analyzed. This step is crucial to avoid the multiple detection of the same banknote when it is wrinkled or folded. Also, it removes any recognition that have a contour shape that can not be associated with a banknote.

The initial filtering is performed by removing any result
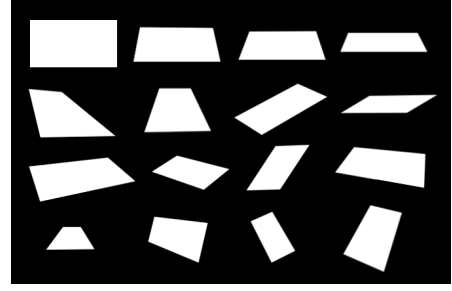

Figure 5. Database of valid instances of banknote shapes

which have a contour with a very low area in relation to the whole image. In a second step, any result without a convex contour is also removed. This is applied because a banknote has a convex quadrilateral shape. It will never have a concave shape, even when it is folded. The reason for this is that the contour is not computed from the banknote borders. Instead a homography is used to map the 4 corners of the reference image to the positions in the image in which the banknotes reside. A final step removes any result that has a contour with a circularity or aspect ratio outside the acceptable bounds. These bounds are retrieved from a database of valid shapes (shown in Figure 5).

6) Detection of multiple banknotes: In order to recognize several banknotes in the same image, the steps presented in previous sections were performed to every banknote reference image, and the best match was chosen. After the retrieval of the best match, its inliers were removed from the keypoint set, and these two steps were repeated again for every reference image, in order to find another banknote. This is done until no valid match is found.

Only the inliers were removed from the keypoint set in order to be able to successful recognize partially occluded banknotes. A less robust method (not used), that will likely fail, removes the keypoints that are inside the banknote contour. Although it might require less computation time to perform, it will fail to detect banknotes that are on top of each other, because part of their keypoints will be removed when one of the banknotes is detected.

IV. RESULTS

The system was tested with 80 test images (overview of banknotes value distribution shown in Table I) that contained banknotes in the most common conditions, such as different perspective views, cluttered environments, partially occluded banknotes and also multiple banknotes per image (67 images had only 1 banknote, 11 images had 2 banknotes and 2 images had 3 banknotes).

With the proper selection of the keypoint detection and description algorithm (depending on the image contents), the system successfully recognized all the 95 banknotes in the 80 test images. The most successful configurations are shown in Table II, which was built by manually inspecting each test image recognition results and selecting the configuration

which successfully recognized all banknotes in the image and achieved better contour estimation.

In Figures 6 to 9 are shown some representative results of the implemented banknote recognition system, showing detection of banknotes with background clutter, perspective distortion, folding and partial occlusion.

TABLE I
TESTING DATASET OVERVIEW

Banknote value	5 €	10 €	20 €	50 €	100 €	200 €	500 €
Nº of banknotes	15	12	19	19	6	9	15

TABLE II
SELECTION OF THE CONFIGURATIONS WITH THE BEST RECOGNITION RESULTS (1 CONFIGURATION PER TEST IMAGE)

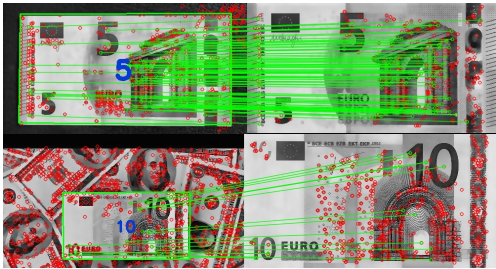| Detector	Descriptor	Images with 1 banknote	Images with 2 banknotes	Images with 3 banknotes
SIFT	SIFT	37	7	2
SURF	SURF	24	3	0
GFTT	SIFT	3	1	0
FAST	SIFT	1	0	0
BRISK	BRISK	1	0	0
ORB	ORB	1	0	0

Figure 6. Detection of a banknote in an ideal perspective view with (bottom) and without (top) background clutter (using SIFT detector, SIFT descriptors and BFMatcher)
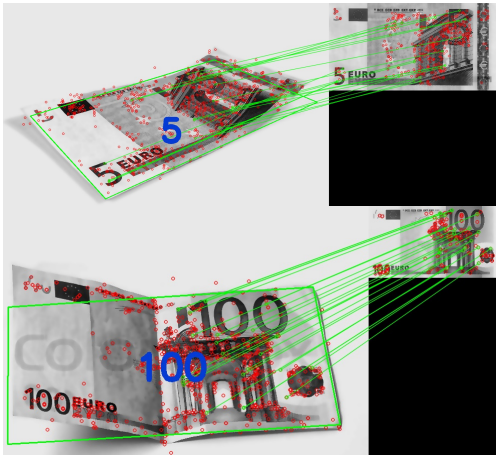

Figure 7. Detection of banknotes with perspective distortion and folding (left image used SURF detector, SURF descriptors and BFMatcher while the right image used SIFT detector, SIFT descriptors and BFMatcher)
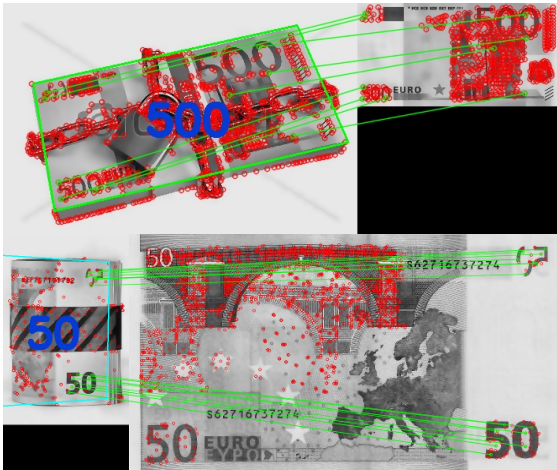

Figure 8. Detection of partially occluded banknotes (left image used GFTT detector, SIFT descriptors and BFMatcher while the right image used SIFT detector, SIFT descriptors and BFMatcher)
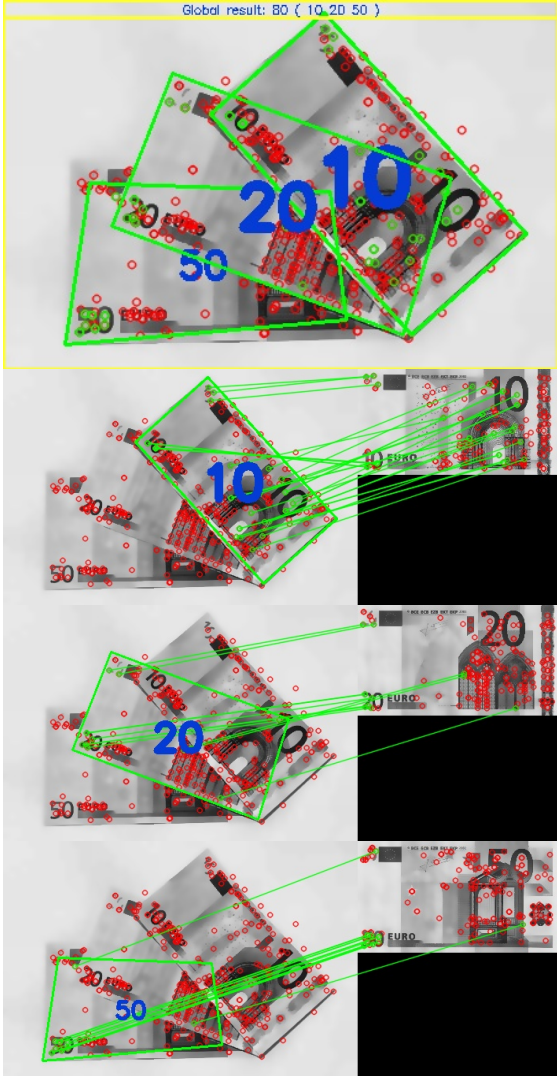

Figure 9. Detection of overlapping banknotes (using SIFT detector, SIFT descriptors and BFMatcher)

V. ANALYSIS OF RESULTS

Analyzing Figures 6 to 9 it can be seen that the proposed recognition system managed to correctly detect the Euro banknotes with perspective distortion and folding (shown in Figure 7) and also with partial occlusion (presented in Figure 8). Moreover, it was able to detect several overlapping banknotes in the same image (seen in Figure 9). Optimal contour estimation was achieved with banknotes without folding and few perspective distortion (example in Figure 6).

Several configurations were tested by combining different feature detectors with feature descriptors, and also using different approaches to perform the descriptor matching and inliers filtering. Analyzing Table II it can be seen that the configuration using SIFT as detector and descriptor in conjunction with a brute force matcher achieved best recognition results. The best performance of SIFT is mainly related to the fact that it can select feature points that can be reliably detected even if the objects are in different perspective views. Moreover, it can compute descriptors that are robust to different lighting conditions. For real-time use, the SURF feature detector and descriptor algorithms are more suitable, since they achieve similar results with significant less computation time. This is achieved mainly due to the simplification of the computations by using integral images. The brute force matcher, although slower than FLANN, achieved better results because unlike the heuristic approach, it matches all descriptors in order to find the best correspondences. However, if the system is to be used in real time, FLANN can be employed with very similar results and lower computation time. In relation to the inliers filtering method, the best results were achieved when the best matched reference image was selected based on the global inliers ratio. However for test images in which most of the banknotes regions were occluded, the local inliers ratio performed better. This occurred because the local matching of patches avoids the removal of recognition results that have low inliers ratio.

VI. CONCLUSIONS

The proposed recognition system successfully recognized the 95 banknotes in the 80 test images, even when they had significant perspective distortion or were partially occluded. It was also robust enough to handle folded and wrinkled banknotes with different kinds of illumination. This was achieved by carefully identifying the regions of the banknotes that had unique features in order to avoid the usage of structures that were similar between banknotes. This technique in conjunction with the usage of reference images with several levels of detail were crucial to improve the correct matching of keypoints descriptors and ensure the correct recognition of the banknotes. The system was configured to recognize Euro banknotes, but can easily be reconfigured to detect other currencies. The achieved results make it a viable option to be used by visually impaired people or to improve automatic banknote counting machines and even increase the security of Automated Teller Machines (ATMs) by detecting counterfeit banknotes.

Future work would include the test of the system using the banknotes under ultra-violet and infra-red light in order to detect with higher confidence counterfeit banknotes and also integrate the system with a speech synthesizer in order to be usable by blind people.

REFERENCES

[1] F. Hasanuzzaman, X. Yang, and Y. Tian, "Robust and effective component-based banknote recognition for the blind," *IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews*, vol. 42, no. 6, pp. 1021–1030, November 2012.

[2] H. Sako, T. Watanabe, H. Nagayoshi, and T. Kagehiro, "Self-defense-technologies for automated teller machines," in *International Machine Vision and Image Processing Conference (IMVIP)*, September 2007, pp. 177–184.

[3] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *International Journal of Computer Vision (IJCV)*, vol. 60, no. 2, pp. 91–110, 2004.

[4] D. L. Baggio, D. M. Escriva, N. Mahmood, R. Shilkrot, S. Emami, K. Levgen, and J. Saragih, *Mastering OpenCV with practical computer vision projects*. Packt Publishing, 2012.

[5] D. Pawade, P. Chaudhari, and H. Sonkambale, "Comparative study of different paper currency and coin currency recognition method," *International Journal of Computer Applications (IJCA)*, vol. 66, no. 23, pp. 26–31, March 2013.

[6] K. Nishimura, "Banknote recognition based on continuous change in strictness of examination," in *ICCAS-SICE*, August 2009, pp. 5347–5350.

[7] Z. Solymár, A. Stubendek, M. Radványi, and K. Karacs, "Banknote recognition for visually impaired," in *20th European Conference on Circuit Theory and Design (ECCTD)*, August 2011, pp. 841–844.

[8] H. Hassanpour, A. Yaseri, and G. Ardeshiri, "Feature extraction for paper currency recognition," in *9th International Symposium on Signal Processing and Its Applications (ISSPA)*, February 2007, pp. 1–4.

[9] F. Grijalva, J. Rodriguez, J. Larco, and L. Orozco, "Smartphone recognition of the u.s. banknotes' denomination, for visually impaired people," in *IEEE ANDESCON*, September 2010, pp. 1–6.

[10] B.-K. Kim, E.-C. Lee, B.-M. Suhng, D.-Y. Ryu, and W.-H. Lee, "Feature extraction using fft for banknotes recognition in a variety of lighting conditions," in *13th International Conference on Control, Automation and Systems (ICCAS)*, October 2013, pp. 698–700.

[11] M. Radványi, Z. Solymár, A. Stubendek, and K. Karacs, "Mobile banknote recognition: Topological models in scene understanding," in *Proceedings of the 4th International Symposium on Applied Sciences in Biomedical and Communication Technologies (ISABEL)*. ACM, 2011, pp. 185:1–185:5.

[12] S. Qian, X. Zuo, Y. He, G. Tian, and H. Zhang, "Detection technology to identify money based on pulsed eddy current technique," in *17th International Conference on Automation and Computing (ICAC)*, September 2011, pp. 230–233.

[13] C.-Y. Yeh, W.-P. Su, and S.-J. Lee, "Employing multiple-kernel support vector machines for counterfeit banknote recognition," *Applied Soft Computing*, vol. 11, no. 1, pp. 1439–1447, January 2011.

[14] S. Gai, G. Yang, and M. Wan, "Employing quaternion wavelet transform for banknote classification," *Neurocomputing*, vol. 118, pp. 171–178, 2013.

[15] H. Hassanpour and P. M. Farahabadi, "Using hidden markov models for paper currency recognition," *Expert Systems with Applications*, vol. 36, no. 6, pp. 10 105–10 111, August 2009.

[16] G. Csurka, C. R. Dance, L. Fan, J. Willamowski, and C. Bray, "Visual categorization with bags of keypoints," in *Workshop on Statistical Learning in Computer Vision (ECCV)*, 2004, pp. 1–22.