Recognizing Bangladeshi Currency for Visually Impaired

Mohammad M. Rahman[1], Bruce Poon[1,2(✉)], M. Ashraful Amin[1,3], and Hong Yan[2,3]

[1] Computer Vision and Cybernetics Research, Computer Science and Engineering, Independent University Bangladesh (IUB), Dhaka, Bangladesh
{motiur.rahman0,aminmdashraful}@gmail.com,
bruce.poon@ieee.org
[2] School of Electrical and Information Engineering, University of Sydney, NSW 2006, Australia
h.yan@cityu.edu.hk
[3] Department of Electronic Engineering, City University of Hong Kong, Hong Kong, China

Abstract. Visually impaired people often have to face difficulty when they try to identify denominations of bank notes. Currently in Bangladesh, there is no system that can easily detect the monetary value of the note. Pattern recognition systems developed over the years are now fast enough to do image matching in real time. This enables us to develop a system able to analyze an input frame and generate the value of the paper-based currency in order to aid the visually impaired in their day-to-day life. The proposed system can recognize Bangladeshi paper currency notes with 89.4% accuracy on white paper background and with 78.4% accuracy tested on a complex background.

Keywords: SIFT · SURF · ORB · Visual Assistance

1 Introduction

There are many visually impaired people around the world especially in the developing world. According to a statistics by World Health Organization [1], the total number of visually impaired people in the world is 285 million. 39 million of these people are blind and 246 million of them are affected by vision related problem. About 90% of the total visually impaired population lives in the developing world and most importantly 82% of them are ages 50 years or more. Detecting the value of bank currency is an important aspect if they are to carry out financial activity like any other people.

There have already been a couple of research works on this subject [2][3]. However they were either too broad in their approaches, or the work had been done with some different bank notes. In this project, we concentrate on Bangladesh Bank notes only, and we use algorithms significantly faster than those used in these two research works.

Scale Invariant Feature Transform (SIFT) [4] and Speeded up Robust Features (SURF) [5] are the two methods that are usually used to perform feature matching

X. Wang et al. (Eds.): ICMLC 2014, CCIS 481, pp. 129–135, 2014.
DOI: 10.1007/978-3-662-45652-1_14

between two unknown images. In situation where we need to produce output quickly we need a better solution. Hence, we go for Oriented FAST and Rotated BRIEF (ORB), a method developed in the lab of Open Source Computer Vision (OPENCV) [6].

ORB is much faster than both SIFT and SURF. SURF has a better run-time than SIFT. Each of these three algorithms has been implemented by engineers at OPENCV. In this project, we have tested each of them in desktop environment running Windows 8.1. The mode of vision is via a web-cam that receives input frame from the outside environment, analyzes the frames and then provides information to the user. If there is a paper currency in the frame, it will provide a value of that currency as well. The module runs in real-time, and has the ability to give textual as well as auditory output. The textual output serves as a confirmation for normal users who can see, whereas the auditory output enables visually impaired user to hear.

2 Implementation

As shown in Figure 1, our system first accepts a photograph of a bank note, applies Oriented FAST and Rotated BRIEF (ORB) algorithm on that image and finally creates a database which contains ORB description of all notes.

Fig. 1. Bangladesh banknote pre-processing steps

2.1 Pre-processing and Feature Extraction

In order to initiate the testing procedure, appropriate databases of relevant descriptors are collected. For that, 'good' samples of Bangladesh bank notes are required. The samples collected were mainly drawn from various sources over the Internet. The main reason was that it was difficult to collect bank notes from the Bangladesh Bank (Central Bangladesh Bank) due to various administrative procedures and the bank was located far from the place of our operation. The second reason was to save the extra hassle of scanning the notes via a scanner.

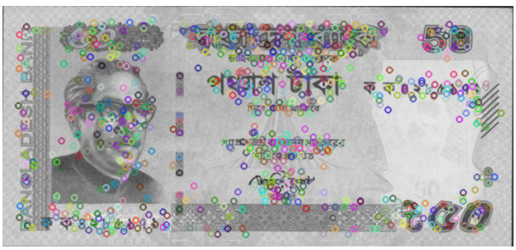

Fig. 2. 50 taka bill with key points

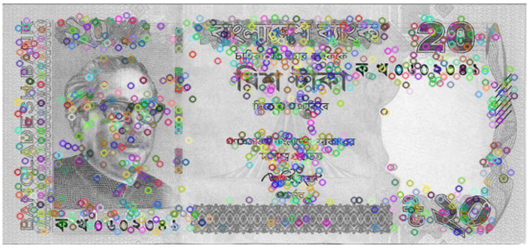

Fig. 3. 20 taka bill with key points

In the above two figures, there are an uncanny amount of similarity between the two bank notes. For a human having normal eye sight, a slight dark environment would confuse the user between the denomination values. It becomes a monumental task for a visually impaired person to differentiate between these two notes.

The bank notes are being pre-processed so that all of them have approximately of the same size, about 500×250 pixels. The descriptor and detector of the notes are found and stored in a table. Both the detectors and the descriptors are found using Oriented BRIEF and Rotated FAST (ORB) [7]. The number of considered features was 500, the number of pyramid level that was used was 8 and the edge threshold was 31[8]. All of these values were used as default when creating the database of descriptors.

2.2 Testing in Real Time

The upper dashed box represents the training of the system and the lower dashed box represents the testing of the system.

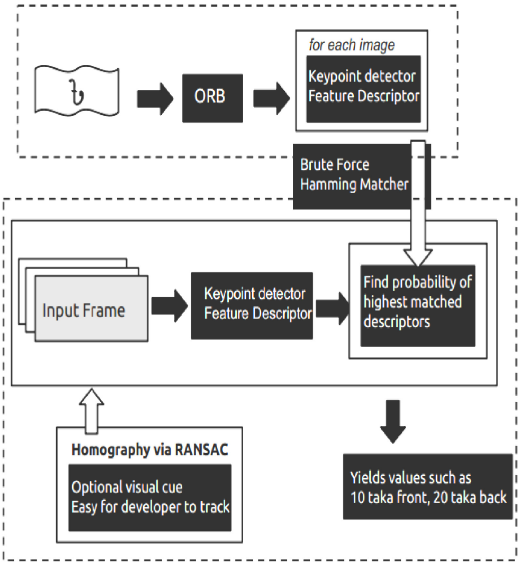

Fig. 4. Overview of the system

Once the descriptors have been computed, they are then stored into a table with the corresponding label of the bank notes. For testing purposes, user input is taken via an HD web-cam. The camera sends about 30 frames per second. On receiving an image, the module computes the detectors found in the image using FAST [9]. Using the detectors, the corresponding descriptors are calculated using BRIEF [10]. However, as BRIEF has little notion of direction embedded in it, it has to be found out using centroid intensity. The next part involves matching of the descriptors of the input frame with that of the training samples.

2.3 Thresholding

The content of the input image often determines the quality of the descriptor that is chosen to represent a good match. In settings where there is no background, the distance between the descriptors will be different from when there is background.

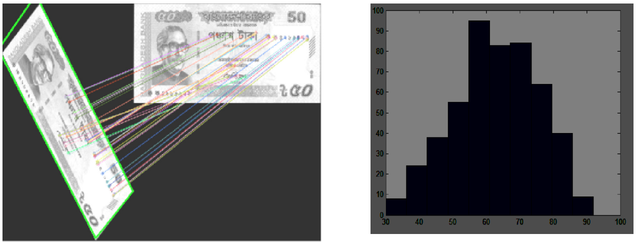

Fig. 5. Money with no background (left) and the distance of the descriptors from training and testing sample (right)

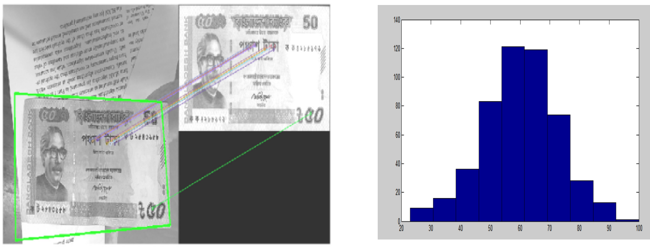

Fig. 6. Money with background (left) and the distance of the descriptors from training and testing sample (right)

In normal situation where there is the presence of background, only descriptors whose distance is less than twice the distance of the minimum distance between the descriptors [11] are used.

$$match\_desc\_dist<2*min(match\_desc\_dist)$$

The frequency of the matched descriptors is analyzed. The highest frequency indicates the presence of a given bank note.

2.4 Homography Calculation

A training image of Taka 50 has to undergo rotation along with translation to match with the image in real time. The phenomenon is known as homography and is calculated using:

$$x' = \frac{a_1 x + a_2 y + a_3}{a_7 x + a_8 y + 1} \tag{1}$$

$$y' = \frac{a_4 x + a_5 y + a_6}{a_7 x + a_8 y + 1} \tag{2}$$

x' and y' are coordinates in the test scene, whereas x and y are coordinates in the training scene. They are bounded to one another in value via the constants a_1 to a_8. The above two equations are examples of over constrained system which can be solved using either Least Squared Error or via Random Sampling Consensus (RANSAC).

In Least Squared Error, the error between the estimated and the actual output is minimized by a certain iterative procedure. In RANSAC, two points are randomly chosen to fit a line. The error between the estimated and the actual output is found out. If that error is above a threshold then two random points are again chosen. If the error is lower than the threshold, the loop stops.

3 Experimental Results

The system was developed using OpenCV 2.4.6 [6] and mexopencv [12]. As an input medium, Logitech HD web-cam was used. The web-cam was portable to scan two types of currency transformation (full and half folded). The experiment was also conducted using a contrastive background.

Fig. 7. Full and half folded 50 taka note with background (above), full and half folded 20 taka note without background (bottom)

In this experiment, 8 different bank currencies were used. Two sides, the front side as well as the backside of the currencies were matched against each other, forming a 16×16 confusion matrix. In Table 1, the confusion matrix is provided. Here note that the most distinct notes are 10 and 1000 taka and the most misclassified notes are those of 20 and 100 denominations.

Table 1. Confusion matrix for all the paper currency notes, tested with plain background as shown in Figure 7

	2 front	2 back	5 front	5 back	10 front	10 back	20 front	20 back	50 front	50 back	100 front	100 back	500 front	500 back	1000 front	1000 back	Sum
2 front	91.9	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	98.1
2 back	0.4	93.3	0.4	0.4	0.4	0.4	0.8	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	99.9
5 front	0.4	0.4	93.7	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	99.9
5 back	0.4	0.4	0.4	93.9	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	100.1
10 front	0.8	0.4	0.4	0.4	93.7	0.4	22.2	0.4	3.3	0.4	0.4	0.4	0.4	0.4	0.4	0.4	125.1
10 back	0.4	0.4	0.4	0.4	0.5	94.0	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	100.3
20 front	0.4	0.4	0.4	0.4	0.4	0.4	56.4	0.4	0.4	0.4	0.8	0.4	0.4	0.4	0.4	0.4	63.0
20 back	0.4	0.4	0.4	0.4	0.4	0.4	0.4	94.0	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	100.2
50 front	2.0	0.4	0.4	0.4	0.4	0.4	13.6	0.4	90.0	0.4	0.4	0.4	0.4	0.4	0.4	0.4	111.0
50 back	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	93.9	0.4	0.4	0.4	0.4	0.4	0.4	100.2
100 front	0.4	0.4	0.4	0.4	0.4	0.4	1.2	0.4	0.8	0.4	67.5	0.4	0.4	0.4	0.4	0.4	74.9
100 back	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	93.7	0.4	0.4	0.4	0.4	99.9
500 front	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	93.6	0.4	0.4	0.4	99.8
500 back	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	93.3	0.4	0.4	99.5
1000 front	0.4	0.4	0.4	0.4	0.4	0.4	1.6	0.4	0.8	0.4	26.3	0.4	0.4	0.4	93.8	0.4	127.6
1000 back	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	94.0	100.3
Sum	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	1599.8

In Table 2, note the recognition accuracy is not as good as those shown in Table 1. The system is most confused if a flipped 100 taka note is presented. From the confusion matrix, it can be seen that 50 taka note is the most distinct note while 50 taka flipped note is the most misclassified note.

Table 2. Confusion matrix for all the paper currency notes, tested with complex background as shown in Figure 7

	2 front	2 back	5 front	5 back	10 front	10 back	20 front	20 back	50 front	50 back	100 front	100 back	500 front	500 back	1000 front	1000 back	sum
2 front	93.9	3.1	0.4	0.4	1.2	0.7	0.4	0.4	0.4	1.7	0.4	13.9	1.6	0.4	0.4	0.4	119.9
2 back	0.4	66.8	0.4	2.8	5.3	0.7	1.2	0.4	0.4	19.6	1.7	13.9	0.4	0.4	0.4	0.4	115.3
5 front	0.4	0.4	93.8	0.4	0.4	0.7	0.8	0.4	0.4	0.9	4.2	1.0	0.4	0.4	0.4	0.4	105.6
5 back	0.4	0.4	0.4	87.2	0.4	0.7	0.4	0.4	0.4	5.2	0.4	1.0	0.4	0.4	0.4	0.4	99.1
10 front	0.4	0.4	0.4	0.4	82.0	2.2	11.8	0.4	0.4	15.7	8.9	8.9	0.4	0.4	0.9	0.4	134.0
10 back	0.4	0.4	0.4	0.4	0.4	84.6	0.4	0.4	0.4	0.4	0.4	1.0	0.4	0.4	0.4	0.4	91.4
20 front	0.4	0.4	0.4	0.4	0.8	0.7	53.2	0.4	0.8	0.9	1.3	2.0	0.4	0.4	2.2	0.4	65.2
20 back	0.4	0.4	0.4	0.4	0.4	0.7	0.4	93.9	0.4	0.4	0.4	1.0	0.4	0.4	0.4	0.4	101.0
50 front	0.4	22.6	0.4	4.4	5.3	0.7	28.0	0.4	93.3	17.8	14.8	5.9	1.2	0.4	2.2	0.4	198.4
50 back	0.4	0.4	0.4	0.4	0.4	0.7	0.4	0.4	0.4	33.5	0.4	1.0	0.4	0.4	0.4	0.4	40.6
100 front	0.4	0.4	0.4	0.4	0.8	1.5	0.4	0.4	0.4	0.4	62.8	1.0	0.4	0.4	0.4	0.4	71.1
100 back	0.4	0.4	0.4	0.4	0.4	0.7	0.4	0.4	0.4	0.4	0.4	41.6	0.4	0.4	0.4	0.4	48.2
500 front	0.4	0.9	0.4	0.4	0.4	0.7	0.4	0.4	0.4	0.4	2.5	1.0	90.5	0.4	0.4	0.4	100.2
500 back	0.4	0.4	0.4	0.4	0.4	0.7	0.4	0.4	0.4	0.4	0.4	1.0	0.4	93.8	0.4	0.4	101.0
1000 front	0.4	1.8	0.4	0.8	0.8	2.9	0.8	0.4	0.4	0.9	0.4	5.0	1.6	0.4	89.6	0.4	107.2
1000 back	0.4	0.4	0.4	0.4	0.4	0.7	0.4	0.4	0.4	1.3	0.4	1.0	0.4	0.4	0.4	93.7	101.7
sum	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	1600.0

4 Conclusions

This experiment shows that it is possible to detect and enumerate values of an unknown bank note, even if they are very close to another in composition. One of the other goals of this experiment is to extend it to the mobile world as well. Even if it is obvious that it is impossible for every visually impaired people in a developing country to own a smart-phone, it is not that hard for people living in the developed world. Implementing this algorithm to recognize bank currencies for such demography will definitely improve their economic lifestyle. However, on Windows 8.1 platform, the proposed system can only recognize Bangladeshi paper currency notes with 89.4% accuracy on white paper background and with 78.4% accuracy tested on a complex background.

Acknowledgement. This project is supported by a grant from Independent University Bangladesh (IUB). We would like to thank Mr. Forhan Noor for helping us with the experiments.

References

1. WHO Media Centre fact sheet no. 282 (October 2013). http://www.who.int/mediacentre/factsheets/fs282/en/
2. Digman, M., Elder, C.: Mobile Banknote Recognition and Conversion. https://stacks.stanford.edu/file/druid:yt916dh6570/Elder_Digman_Foreign_Bill_Recognition.pdf
3. Bruna, A., Farinella, G.M., Guarnera, G.C., Battiato, S.: Forgery Detection and Value Identification of Euro Banknotes. Sensor (Basel) **13**(2), 2515–2529 (2013). doi: 10.3390/s1302022515. http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3649387/ (February 18, 2013)
4. Lowe, D.: Distinctive Image Features from Scale-Invariant Keypoints. International Journal of Computer Vision (2004). http://www.cs.ubc.ca/~lowe/papers/ijcv04.pdf
5. Bay, H., Ess, A., Tuytelaars, T., Gool, L.V.: Speeded-Up Robust Features (SURF). ftp://vision.ee.ethz.ch/publications/articles/eth_biwi_00517.pdf
6. Open Source Computer Vision. http://opencv.org/
7. Perception, M., Group, R.: Chubu University, Japan. http://www.vision.cs.chubu.ac.jp/CVR/pdf/Rublee_iccv2011.pdf
8. Open CV, Feature Detection and Description. http://docs.opencv.org/modules/features2d/doc/feature_detection_and_description.html#orb-orb
9. Rosten, E., Drummond, T.: Machine learning for high-speed corner detection. European Conference on Computer Vision (May 2006). http://www.edwardrosten.com/work/rosten_2006_machine.pdf
10. Calonder, M., Lepetit, V., Strecha, C., Fua, P.: BRIEF: Binary Robust Independent Elementary Features. In: Daniilidis, K., Maragos, P., Paragios, N. (eds.) ECCV 2010, Part IV. LNCS, vol. 6314, pp. 778–792. Springer, Heidelberg (2010)
11. Issues with imgldx in Descriptor Matcher mexopencv, Stack Overflow. http://stackoverflow.com/questions/20717025/issues-with-imgidx-in-descriptormatcher-mexopencv
12. Yamaguchi, K.: GitHub.com/mexopencv. https://github.com/kyamagu/mexopencv