

Vision-based Precise Cash Counting in ATM Machines

Hossein Farid Ghassem Nia and Huosheng Hu

*School of Computer Science and Electronic Engineering
University of Essex, Colchester CO4 3SQ, United Kingdom*

Email: hfarid@essex.ac.uk; hhu@essex.ac.uk

Abstract — The demand to use ATM machines has increased tremendously worldwide during the last decade. Traditional cash counting in ATM machines is based on the mechanical mechanism that is in touch with cash and may cause some damages from time to time. This paper proposes a novel visual based counting technology for cash counting in ATM machines. To achieve precise counting, intelligent decision making algorithms are deployed to evaluate the counting performance so that the machine will not produce incorrect counts. The experimental results showed that the proposed method can eliminate false positive counts effectively.

Keywords: Vision-based cash counting, Decision making, ATM machines

I. INTRODUCTION

Automated Teller Machines (ATM) have revolutionized the banking operation worldwide, in which accurate cash-counting technique is essential. In line with increased demand for more complex and faster operations, it is necessary to use more intelligent methods that can count cash faster and also reliably. Computer vision is one of the techniques that can be used for cash counting, which has the advantage of contactless for a wider range of materials and preventing fraud. For example, the ATMs would be able to count and dispense vouchers, documents, credit cards and etc. Due to sensitivity of counting task and complexity of operations, it is vital that systems make safe decision autonomously in order to achieve the accurate count. Meanwhile, systems should avoid decisions that may be risky or unsafe.

Generally, a key challenge in cash counting is the assurance of precise counting. In other words, it is vital to guarantee that there is no false positive count. To improve reliability, several vision sensors may be employed to carry out counting tasks concurrently. The possible outputs from multiple vision sensors include (i) true positive – correct count and confirm the count; (ii) true negative – incorrect count and reject count; (iii) false positive – incorrect count but confirm count. A decision making engine should be deployed to deal with uncertainty and fuse the results safely.

This paper proposes a novel decision making technique to be employed in ATM machines. This system is to replace mechanical counting mechanism with contactless computer-vision enabled counting. It captures an image from stack of notes and uses some image analysis methods to count the notes. The proposed decision making method aims to eliminate false positive errors since these errors are very costly and should be prevented from cash counting machines.

The rest of this paper is organized as follows. Section 2 presents some background information and provides a clear

explanation of the problem. The relevant literatures for decision making are also discussed in this section. Section 3 describes the proposed visual counting method. Experimental results are presented in Section 4 to show the feasibility and performance of the proposed method. Finally, a brief conclusion and summary are given in Section 5.

II. BACKGROUND

A. problem description

The question to be addressed in this paper is how to eliminate false positive counts in a cash counting machine that employs vision systems. Traditionally, a single camera may be used to capture an image and count the notes using images analysis techniques. However, due to the nature of imaging sensor used and other unpredictable parameters, sensors may not be able to count the notes accurately. Since vision sensors are not working in an ideal situation, the count may not be accurate. Some of the parameters that may affect operation are lighting condition, position, presentation of notes and etc.

To overcome to this limitations, it is necessary to use several cameras with the multi view vision technique. Meanwhile, each image would be analysed by independent image processing techniques. Figure 1 shows the traditional approach for vision-based note counting in a) and also multiview approach for note counting in b).

Using several sensors and analysis techniques will lead to different counting results that may aggregate to a single point or may not. An intelligent decision making system should be used to decide which sensors have the most reliable count and whether there is an agreement on count number within all analysis methods. Experiments showed that using conventional methods such as aggregating and majority voting [11] is not sufficient. The reason is that there is no ground truth for the table of results.

An example may clarify the problem with majority voting. If 45% of the agents that are representing true class but 55% are representing false value, majority voting simply chose false values because they are in majority. A simple solution is to put a weight on each agent. In our system, there is no gold standard to obtain weights in a dynamic platform. In addition, on every execution, images are different so there is no particular pattern among the set of results. On the other hand, due to the nature of application and the variety of materials, it is not practical to obtain prior information regarding the input data. Therefore, in the lack of dataset, it is not possible to use supervised learning.

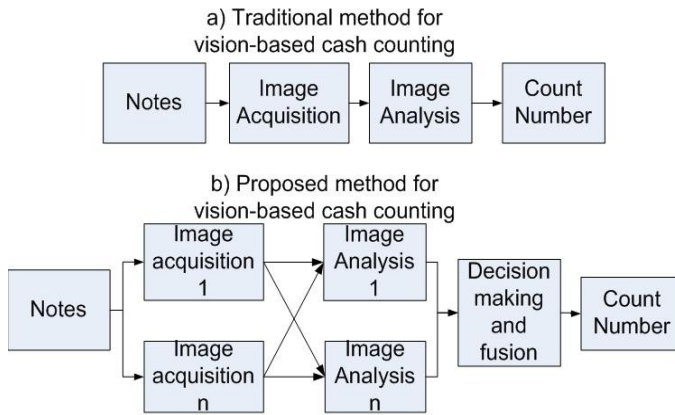


Figure 1- system architecture for vision-based cash counting

Table 1- Table of count results

	Image analysis method 1	...	Image analysis method n
Frame view 1	Count _{1,1}	...	Count _{n,1}
...
Frame view m	Count _{1,m}	...	Count _{n,m}

In cash counting, it is essential that systems produces the true positive or true negative results, only. In this project, one of the challenge is that all the images, patterns and counts will change in every single count. So, the decision making engine should be able to judge about reliability of sensors and which sensors and methods are representing the true postive results. The system should also make these decisions autonomously and safly.

B. Classifier fusion and Decision making theory

Decision making can be defined as selecting one expert agent among several agents. However, the selected option may not be the safest, most reliable or most accurate option. Then, decision making can be redefined as aggregating classifiers. Classifier fusion is technique to bring together the sum output of set of classifiers to make a final decision. Multiple Classifier Fusion (MCF) can be used to draw a conclusion from set of results e.g. Table 1[13, 14]. For cash counting, it is necessary to draw a conclusive count based on the bank of results, e.g. table 1. For this purpose, there are two possible solutions, one result can be selected as reliable count or alternatively, all counts fused together as it is shown in [15]. But in the lack of ground truth and training data, a single count is not reliable.

Majority voting [11] cannot provide reliable results and adaptive majority function [9] needs a training set to obtain weights. Jacobs et al introduces a new supervised classifier fusion using an adaptive combination of local experts [16]. In their method, they treated each classifier as an expert. Hence they could split the training set into subsets where local experts handle the small subsets. Similarly, Alpaydin et al. approximated the global class discrimination using a framework in [17]. Their framework is based on linear per-

ceptron. However, the performance of these methods depends on the quality of training sets.

Niu et al. combined majority voting and Bayesian theory for motor fault diagnosis [12]. They used several classifiers where each classifier produces independent classes. However, in the lack of dataset, it is not possible to train their method. In addition, there is no discussion regarding execution time hence their method, may not be suitable for real time applications.

To satisfy strict requirements of cash counting, Multi Criteria Decision Making may be employed. The use of probabilistic methods in MCDM has been explored in [1, 2]. Peng proposed decision making framework based on Bayesian networks influence diagram. They could achieve a practical solution in dealing with MCDM problems. Their method also benefits from modelling of noisy interactions among agents [1]. The general use of fuzzy set theory in MCDM is explored in [3, 4, 5, 6]. In addition, in many areas, the decision making is done in linguistic environment [7].

Goumas et al. introduced a new multiple classifier fusion technique to enhance positioning accuracy of objects in an optical inspection system [8]. It was confirmed that their proposed method, for certain objects, could not reach a good performance in compare to naive Bayes and Dempster–Shafer combiner [10].

Another important point is that these methods will not work effectively in the lack of complete dataset. In addition, it is not discussed that how these method can provide guarantee for true positive/negative outcomes. In addition, if the data is noisy, they may produce incorrect result. Due to nature of the cash counting task, we do not have prior knowledge about the input data and level of noise. Therefore, the machine should be able to deal with all possible inputs. Because of these limitations, it is mandatory to create a novel method that can satisfy the requirements.

III. THE PROPOSED METHOD

Based on the literature review and also strict system requirements, the proposed decision making framework should be able to eliminate false positive results. Some restrictions in design and development include the lack of training data set, limited computation power and hard deadline for real time count. The proposed method should be flexible so it can adapt itself to new materials on the fly.

Results from different frames and image analysis are presented in table 1 where rows represent frame number and columns represent image analysis method. Each cell contains the count results for the n^{th} analysis method and m^{th} frame view. During experiments, it was noted that the statistical information in the result table are key indicator to show the reliability of results. After further investigations, it was concluded that the best practical method is to extract statistical information from the count table. Using this method, information is generated in real time. Hence a new architecture is needed to put all elements of system together. The proposed algorithm has four primary steps. The first step is to create a hypothesis. The second step is to update probabil-

ity density function. The third step is to extract statistical information. The final and the fourth step is rule based decision making. We explain these steps below.

1) *First step: hypothesis creation*

The first step is to create a hypothesis. We need to select a candidate (guess accurate count from table of counts) and then would be able to decide about confidence of candidate. The ultimate aim is that our decision making engine can decide whether the hypothesis is true positive/negative or false positive. Obviously, this part is vital in correct operation of machine. The closer the candidate number is to true positive count then it is more likely that system can make correct decision with higher confidence. We call this process hypothesis creation. In fact the role of decision engine is to assess the reliability and confidence of hypothesis. To create hypothesis, the mode of count table is calculated. The equation 1 is showing the hypothesis where C_{ij} represents the count in respective count table. The mode can also be interpreted as the most frequent value in table.

$$\text{Hypothesis} = \text{Mode}(C_{ij}) \quad (1)$$

2) *Second step: probability density function*

The second step is to adjust probability density function distribution over the table of counts. Experiments showed that position of cells in table is a key factor. This is simply because of the fact that some frames have better view to notes and also some methods may work better for that particular frame. Using PDF, all the single count results will have a weight that would be used for weighted average. The PDF is modified on every execution based on the statistical data of count table

$$f(x, y) = A \exp \left(- \left(\frac{(x - x_0)^2}{2\sigma_x^2} + \frac{(y - y_0)^2}{2\sigma_y^2} \right) \right). \quad (2)$$

The PDF is in form of 2D Gaussian shape. The x_0 and y_0 are center of the PDF which have the biggest weight. The location of x_0 and y_0 is cell in the count table which is the most reliable count. Position of cell is determined in the first execution. The sigma illustrates the width of PDF which is proportional to the size of count table. Having the PDF, a weighted average of count table can be calculated.

3) *Third step: statistical information extraction*

The third step is to extract statistical information from table of counts. In order to draw a framework to assess the hypothesis, some statistical information need to be extracted from the count table. Having a count table, the mean can be calculated as below:

$$M = \frac{1}{n * m} \sum_i^n \sum_j^m C_{ij} \quad (3)$$

The above equation calculates the mean over the entire count table. C_{ij} represents the count in respective count table. The size of count table is also shown by n and m . Standard

deviation is also another important factor that reflects count deviation in the count table. The standard deviation can be calculated using the following equation.

$$\sigma = \sqrt{\frac{1}{n-1} \sum_{i,j}^n (C_{ij} - M)^2} \quad (4)$$

where σ shows the standard deviation.

To have better estimation of the deviation in the system, it is necessary to calculate the true standard deviation or error bar.

$$S_e = \frac{\sigma}{\sqrt{n}} \quad (5)$$

where S_e is representing the true standard deviation.

There should be also some other parameters that indicate how many times the analysis methods produced different count, let us call it H . An example can clarify the importance of H . For example, if one count has big error in large number of counts, it may be acceptable but its standard deviation is high. But if 60% of counts have small variation, the standard deviation would be small but the candidate value is unreliable.

Hence we need another parameter which can illustrate the number of counts that are different from mean of table. The H is the number of counts in the count table that their value is different from the mean of all counts in the count table. This parameter can provide an indication of the variation in the count table. Having the true standard deviation error and also propagation of variations, then we can use the following equation to give a score to abnormality in the count table.

$$S_{ABN} = \frac{H * S_e}{n} \quad (6)$$

where S_{ABN} indicates the abnormality score in the count table. The higher the score is the more abnormal the count table is. The ideal score is zero which indicates no variation in the system.

Table 2 - Rule based decision making

if weighted average=Hypothesis and $S_{ABN} < \text{threshold}$
Hypothesis is correct
if weighted average \neq Hypothesis and $S_{ABN} < \text{threshold}$
if H is small
Hypothesis is correct
if weighted average \neq Hypothesis and $S_{ABN} < \text{threshold}$
if (Hypothesis - (M \pm σ) > 0)
Hypothesis is incorrect
if weighted average \neq Hypothesis and $S_{ABN} > \text{threshold}$
Hypothesis is incorrect
if weighted average=Hypothesis and $S_{ABN} > \text{threshold}$
Hypothesis is incorrect

4) Rule based decision making

So far, a hypothesis was created and then the PDF function was updated. Later, S_{ABN} was calculated to determine the reliability of the count table. In the next stage, it is necessary to assess the trueness of the hypothesis. For this purpose, the rule based decision making is used to make the safest decision. Table 2 shows a simplified rule based decision making engine. The rules of engine can be extended or modified according to requirements. If hypothesis is correct then the hypothesis would be shown to user as true positive count. If the hypothesis is incorrect then a rejection will be shown to user. Figure 2 shows the architecture of the proposed decision making system.

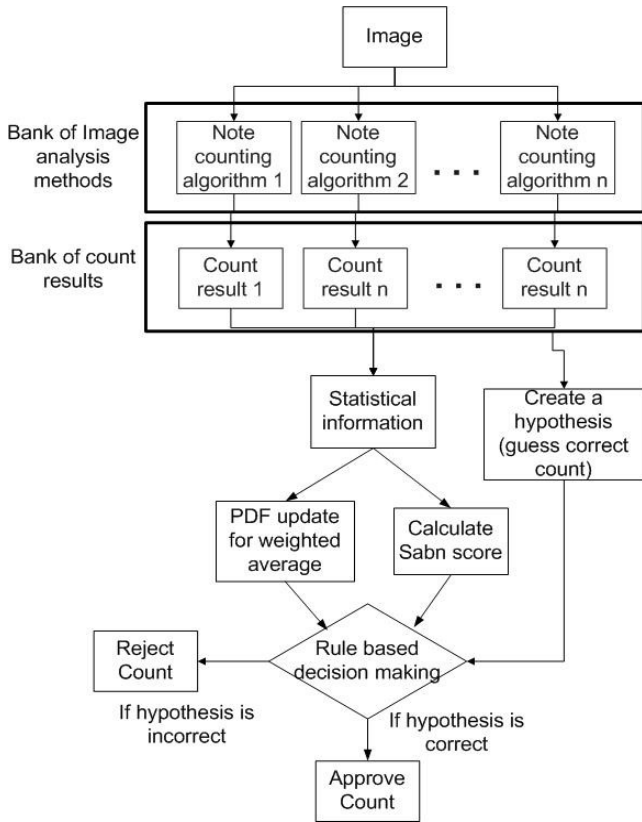


Figure 2 - Proposed method

IV. EXPERIMENTAL RESULTS

This section explains the experimental results regarding the proposed decision making system. Since the decision making engine is the final part of the system that analyze results and produces the output to the user, assessing the decision accuracy can be interpreted as evaluating the overall performance of the vision based cash counting system. On the other hand, if the decision making engine fails, the user may get a false positive count.

Although the desired output is true positive or true negative, the rate between the true positive and true negative is also important factor, i.e. a 60% rejection in count (true negative) will make the operation a frustrating and long process.

Hence, user experience with the machine is based on the decision making engine performance. In addition, feedback from user regarding the accuracy and reliability of the machine also can be used to evaluate the overall performance of the system.

There have been some issues regarding assessing the performance of machine and decision making engine. Due to the nature of application, there are limited numbers of samples that can be used for field tests. Beside all of these difficulties, a huge portion of project time was spent to test the performance of machine, using different approaches.

There are two test levels that are used to test the algorithms, namely unit test and system test. In the unit test, each component was separately tested to make sure the output is valid. Some of the units that were tested include the hypothesis creation and statistical rules extraction. In most tests, the defect test approach was chosen. In the system test, the system operation from user point of view was selected. The final output was investigated and compared against the true values (hand counted values). Defect testing approach was selected to find any potential problems. System test can provide us useful information regarding overall performance of the developed algorithm.

One of the challenges in conducting experiments was data generation. Meanwhile, the decision making algorithm or classifier fusion system should be fully tested to provide a high confidence in operation. To achieve this goal, the classifier fusion algorithm was tested on three types of data: offline data, simulated data and live data. Each of these data types are discussed in details below.

A. Test using offline data

The first set of experiments to test the software is based on offline data. Offline data is consisted of images which are captured from real samples during operation and saved to hard disk for further examination. To have a complete dataset, over 300 images were collected using various samples. In the offline test, this database was used to generate the count results.

In this method, the count table which was generated by image analysis algorithms was used as input. Based on the output of algorithms, accuracy of proposed decision making framework could be evaluated. Below is the table, which highlights the findings.

Table 3 -test results for offline data

Image batch ID	Number of images in Batch	False positive	True positive	True negative
001	65	0%	95%	5%
002	73	0%	94%	6%
003	98	0%	91%	9%
004	41	0%	97%	3%

Table 3 summarises the outcome of this test. The first column shows the batch ID. A program selected a random number of images in random order to create each batch. As it can be seen from table 3, no false positive results were observed. In addition, true negative rate is also acceptable. The average true positive rate is 94.25% which demonstrates a smooth and reliable operation.

B. Test using simulated data

The second type of test is test using simulated data. We added noise to images to generate simulated data and test our algorithm. Table 4 shows the experimental results for simulated data. In this experiment, the images are affected by noise. Both Speckle and Gaussian noise were added to input images. In addition, the count table was also affected by Poisson noise. These levels of noise in the system created a great deal of uncertainty in the system. The main reason to conduct this test was to see how the proposed method performs in facing uncertainty in the system. This experiment showed that in facing the noisy data and in high uncertainty, system is reliable and avoids false positive results. However, true negative rate has increased due to noisy data.

Table 4 - test results using simulated data

Image batch ID	Number of images in Batch	False positive	True positive	True negative
011	49	0%	86%	14%
012	93	0%	83%	17%
013	35	0%	81%	19%
014	71	0%	80%	20%

C. Test using live data

The third and final types of test is based on live data tests. The third category is carried out while the algorithm is implemented in the software and the test is carried out by two independent volunteer who had no previous knowledge about machine. The aim of this experiment was to find probable cases that system fails. The defect test approach was used. In this method the volunteer were taught to try to fool the system. They placed the sheets with different angles and various gaps in front of camera. They could also move sheets forward and backward to make the objects out of the focus. The volunteers were also allowed to create angles to any number of sheets.

The main reason to design this defect test was due to the fact that the machine may be used in harsh unknown environment. So the machine should be able to cope with the worst case scenario. The volunteers were asked to count 20 batches of materials (notes) randomly. Over 40 different materials were presented to volunteers. All of the material had known quantity and it was confirmed by hand count and machine count before the experiment. The volunteers were

asked to count the material 3 times with normal and standard situation so they can be assured that they know the quantity of material.

Then the volunteers were asked to manipulate and move the materials based on their own preferences and desire. They were asked to test if they can get a miscount on the machine. They were asked; in the event of occurrence of miscount; to record the miscount and also save the image on the hard disk for further investigation. The machine were earased prior to experiment so machine did not contain any information about materials were presented.

In the live test, it is not aimed to not investigate on the count table. The purpose of this test is to assess the output of system to investigate if system produces the false positive results. If it does, under what circumstances the machine is producing the false positive results.

Table 5 shows the result of operation conducted by two volunteers. These results confirmed that the machine did not produce any false positive results under the stress. Similar experiments also conducted by developers, engineers, sales people and also customers. In none of the experiments no false positive results was observed. These experiments proved that the proposed method in this paper is working and it can eliminate false positive results

Table 5 - results of live test

Volunteer No.	Number of materials	Number of counts	False positive rate
1	20	209	0%
2	20	214	0%

V. CONCLUSION AND FUTURE WORK

In this paper, a visual cash counting technique has been developed for the precise cash counting of ATM machines. A novel decision making engine has been deployed to eliminate the false positive counts due to the sensitive nature of security and financial tasks. In order to assess the proposed algorithm, several experiments were conducted with three types of data, namely offline data, simulated data, and real data. Experimental results are presented and the feasibility and performance of the proposed decision making engine are highlighted.

The offline test was firstly executed using saved images on the hard disk. Each image was fed into image analysis programs and the respective count table was created. Simulated data were generated using offline data that were heavily affected by noise. Then, the real data was used to test the system, in which the quantity of materials was known. The outcome was recoded and analysed. It was clear that the proposed method could provide a reliable results. No false positive results were observed.

To show the proposed approach can perform well in the difficult and harsh situations, the defect test was selected to find any potential bug. The system was tested by two

volunteers using various random materials. It was confirmed that the machine did not produce any false positive result.

The future work will be focused on testing the developed decision making method on different ATM machines to further verify its accuracy.

ACKNOWLEDGMENTS

This research is financially supported by a Knowledge Transfer Partnership project funded by UK Technology Strategy Board and Vaccumatic Limited at Colchester, Grant No. KTP007916, as well as a British Council grant "Cooperative Autonomy of Micro Aerial Vehicles", Research Exchanges with Beijing Institute of Technology, China. Authors wish to thank John Gan, David Long and Rob Greenwood for their valuable advices and support on this project.

REFERENCES

[1] W. Watthayu, Y. Peng, "A Bayesian network based framework for multi-criteria decision making," Proceedings of the 17th International Conference on Multiple Criteria Decision Analysis (2004), Whistler, British Columbia CA.

[2] N. Fenton, M. Neil, "Making decisions: using Bayesian nets and MCDA," Journal of Knowledge Based Systems, Vol 14, 2001, pp. 307-325

[3] R.E. Bellman, L.A. Zadeh, "Decision-making in a fuzzy environment," Journal of Management Science, Vol. 17, 1970, pp. 141-164

[4] C.C. Robert, R. Fullér, "Fuzzy multiple criteria decision making: recent developments," Journal of Fuzzy Sets and Systems, Vol. 78, 1996, pp. 139-153

[5] R.A. Ribeiro, "Fuzzy multiple attribute decision making: A review and new preference elicitation techniques," Journal of Fuzzy Sets and Systems, Vol.78, 1996, pp.155-181

[6] R.R. Yager, "On the evaluation of uncertain courses of action," Journal of Fuzzy Optimization and Decision Making, Vol. 1, 2002 pp. 13-41

[7] N. Fenton, W. Wang, "Risk and confidence analysis for fuzzy multicriteria decision making," Journal of Knowledge-based Systems, Vol. 19, 2006, pp. 430-437

[8] S.K. Goumas, I.N. Dimou, M.E. Zervakis, "Combination of multiple classifiers for post-placement quality inspection of components: a comparative study," Journal of Information Fusion, Vol 11, No 2, 2010, pp. 149-162

[9] G. Latif-Shabgahi and S. Bennett, "Adaptive Majority Voter: A Novel Voting Algorithm for Real-Time Fault-Tolerant Control Systems," in Proceedings of the 25th EUROMICRO Conference, 1999, pp.113-120

[10] L. Xu, A. Krzyzak, C.Y. Suen, "Methods of combining multiple classifiers and their applications to handwriting recognition," IEEE Transactions on Systems, Man and Cybernetics, Vol. 22, No. 3, 1992, pp. 418-435

[11] R. Battiti, A. Colla, "Democracy in neural nets: voting schemes for classification," Journal of Neural Networks, Vol. 7, No. 4, 1994, pp. 691-707

[12] Niu G., Han T., Yang B. and Tan A. C. C., "Multi-agent decision fusion for motor fault diagnosis," Journal of Mechanical Systems & Signal Processing, Vol. 21, No. 3, 2007, pp.1285- 1299

[13] D. Ruta, B. Gabrys, "An overview of classifier fusion methods," Computing and Information System, Vol. 7 No. 1, 2000, pp. 1-10

[14] L. Kuncheva, "A theoretical study on six classifier fusion strategies," IEEE Transactions on Pattern Analysis and

Machine Intelligence, Vol. 24, No. 2, February 2002, pp. 281-286

[15] K. Woods, W.P. Kegelmeyer, K. Bowyer, "Combination of multiple classifiers using local accuracy estimates," IEEE Transactions on Pattern Analyses and Machine Intelligence, Vol. 19, 1997, pp. 405-410

[16] R.A. Jacobs, M.I. Jordan, S.J. Nowlan, G.E. Hinton, "Adaptive mixtures of local experts," Journal of Neural Computation, Vol. 3, 1991, pp.79-87.

[17] E. Alpaydin, M.I. Jordan, "Local linear perceptrons for classification," IEEE Transactions on Neural Networks, Vol. 7, No. 3, 1996, pp.788-792.