Forensic Authentication of Bank Checks

Rajesh Kumar, Gaurav Gupta

Chapter 16

FORENSIC AUTHENTICATION OF BANK CHECKS

Rajesh Kumar and Gaurav Gupta

Abstract This chapter describes an automated methodology for the forensic authentication of bank checks. The problem of check authentication is modeled as a two-class pattern recognition problem. Color and texture features are extracted from images of genuine and counterfeit checks. A support vector machine is utilized to determine check authenticity. Classification experiments involving a dataset of 50 bank checks yielded a detection accuracy of 99.0%. The automated methodology can be used by non-specialist personnel to detect check counterfeiting in a banking environment where large numbers of checks are handled on a daily basis.

Keywords: Bank checks, authenticity, pattern recognition, color, texture

1. Introduction

The counterfeiting of currency notes, bank checks and certificates (e.g., birth, death and degree certificates) is a major problem in developing countries such as India. Even developed countries are encountering serious threats from counterfeiting. Technological advances, such as advanced scanning, color printing and color copying, have opened new avenues for criminals to run their counterfeiting businesses. The counterfeit materials look real to the naked eye. In fact, the quality of counterfeits is so good that even experts are often unable to distinguish them from the originals.

Counterfeit bank checks and currency notes directly affect a national economy. In recent years, the counterfeiting of checks, bank drafts and money orders has increased dramatically – this is evident from the number of alerts issued by the U.S. Federal Deposit Insurance Corporation. The number of alerts was 50 in 2003, 75 in 2004, 168 in 2005, 318 in

2006 and 300 in 2007, a 500% increase in just four years. In 2007 alone, the United States, Canada and other countries jointly intercepted more than 590,000 counterfeit checks with a total face value of approximately $2.3 billion [12]. Clearly, the counterfeiting of checks and other financial documents is a serious concern.

The examination of counterfeit documents, in general, and counterfeit checks, in particular, relies on manual observations of certain built-in security features. In a forensic laboratory environment, counterfeit checks are examined microscopically and under different lighting conditions to identify discrepancies. Unfortunately, such examinations are cumbersome and infeasible in a banking environment where large numbers of checks are processed daily. Automated approaches that can be performed rapidly by non-specialist personnel are required to address the check counterfeiting problem.

Research in the area of automated authentication of security documents is relatively new. Several researchers have studied the problem of determining the authenticity of documents using pattern recognition [3, 9, 11]. Moreover, some researchers [2, 4, 6, 7] have examined printers, scanners and other devices that could indirectly help identify fraudulent documents.

This chapter describes an automated methodology for authenticating bank checks using pattern recognition. The objective was to implement a reliable system for determining the authenticity of large numbers of bank checks in real time. The problem was framed as a two-class classification problem involving pairs of checks: (i) Class I, when the reference and questioned checks are both genuine; and (ii) Class II, when one of the two checks (i.e., the questioned check) is fake. Since bank checks and other important documents incorporate security features based on the printing technology and printed designs, suitable features based on color and texture attributes were extracted and used for classification. A support vector machine (SVM) was trained and subsequently used for classification. The input vector to the support machine was a (dis)similarity index obtained by taking the absolute difference of corresponding elements in the feature vectors of a pair of genuine and fake checks. Classification experiments involving a dataset of 50 bank checks yielded a detection accuracy of 99.0%.

2. Security Features in Bank Checks

Before designing an automated system for authenticating bank checks, it is important to understand the characteristics of checks and their security features. Knowledge of these characteristics and features is

central to modeling bank check authentication as a pattern recognition problem.

Most important documents, including bank checks, have certain features that are considered to be difficult to copy and are, therefore, used for authentication purposes. These features are referred to in the forensic literature as security features. The more valuable or sensitive a document, the more complex are its security features. For example, passports, visas and currency notes have several complex security features; bank checks, official stamp paper and certificates have simpler (and relatively easy to copy) features due to their variations and lesser importance.

Security features are typically embedded in a document during paper manufacture and/or during printing. The features incorporated at the time of paper manufacture include paper type, thickness, surface roughness and watermarks. The features embedded during printing are the artistic design, printing patterns, micro features and the printing technology itself. Bank checks, like other security documents, have security features embedded in them during both phases.

2.1 Features Embedded During Manufacture

Paper plays an important role in the embedding of security features. Special types of paper are used for security documents because these types of paper are only available to official entities. The type of paper – made from cotton, grass or bamboo – is also a security feature. Different types of paper have unique physical and chemical properties that facilitate authentication. The thickness of the paper used for bank checks is also a distinguishing feature.

Other security features embedded during paper manufacture include watermarks and fluorescent optical fibers. Watermarks are designed into security documents to enhance identification and security. The watermark is actually a thinner area than the rest of the paper. The "dandy roll" used in the paper manufacturing process incorporates a metal representation of the watermark, which pushes paper fibers aside and leaves an imprint on the paper [8]. The imprint or watermark produced in this manner is more transparent than the rest of the paper and is readily visible.

Fluorescent optical fibers are also embedded as a security feature. The fibers are generally visible under ultra-violet light.

Watermarks and optical fibers are difficult to duplicate using scanners and copiers. However, as a result of cost considerations (especially with regard to embedding optical fibers), the vast majority of bank checks

do not have these security features. Therefore, these features are not considered in the proposed methodology for detecting counterfeit bank checks.

2.2 Features Incorporated During Printing

Security features incorporated during the printing process include the type of printing ink, specially-designed fonts and artwork. The printing process varies from conventional offset printing to modern laser printing. Each printing process produces documents of a different quality; thus, the printing process itself incorporates security features in a document. Specialized technology such as intaglio printing is also used to print bank checks. This type of printing produces raised surfaces that can be felt by touching certain areas of a bank check. The security of the bank check is enhanced because the raised surfaces cannot be duplicated using a document reproduction device.

Using inks of different colors helps individualize bank checks. The inks range from conventional dye- or oil-based inks to special magnetic inks. Some inks are thermochromic – they change color when exposed to heat and return to their original color upon cooling; other inks are resistant to solvents. Thermochromic inks are resistant to color-copying and scanning while inks that are resistant to solvents are difficult to erase. Each of these inks can make a bank check distinctive. In addition to printing inks, a magnetic ink or toner is used in bank checks to print the magnetic ink character recognition (MICR) characters used to automate check processing and clearing [10].

Aside from printing processes and printing inks, special artwork is often printed on bank checks to provide additional security features. These may include micro-printing, crisscross lines, MICR characters and the bank logo, among others. Micro-printings are periodically-repeated characters, words or patterns that are distributed throughout a check. Scanning or copying these features may result in the deformation of their shapes [3]. Crisscross lines are intricate lines that are an essential part of security documents; these lines are also resistant to scanning and copying. The MICR characters and bank logo are typically printed in a distinctive manner to enhance the individuality of a bank check.

This work focuses on micro-printings on the backgrounds of bank checks to detect counterfeiting via the application of image processing and pattern recognition methodologies. Counterfeit detection primarily relies on color and texture features extracted from regions of interest that are confined to the backgrounds of bank checks.

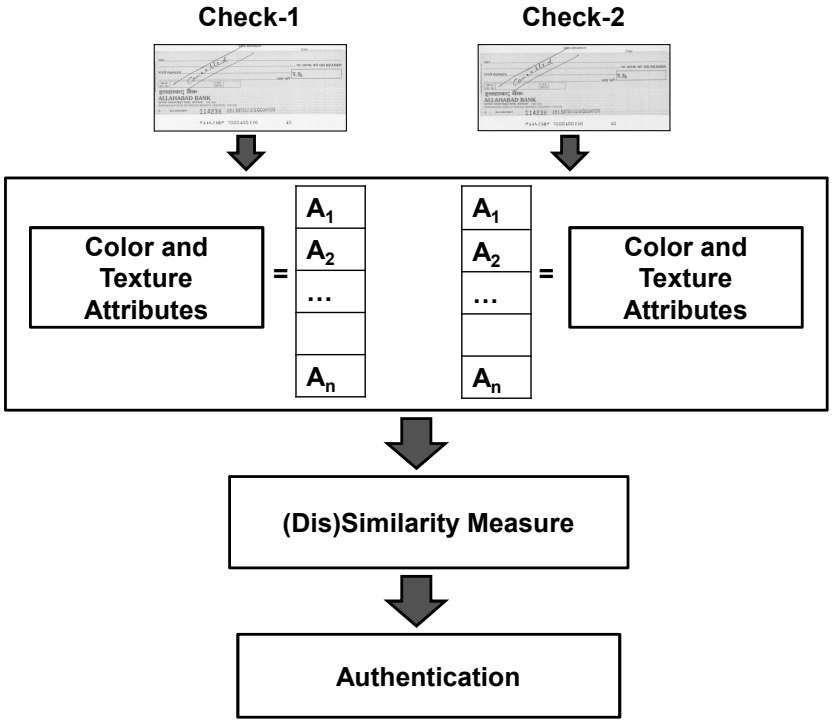

Figure 1. Bank check authentication methodology.

3. Bank Check Authentication Methodology

The authentication of a bank check can be framed as two-class classification problem. The two classes are: (i) Class I, when the questioned check and the reference check are both genuine; and (ii) Class II, when the questioned check is fake and the reference check is genuine.

Figure 1 presents the proposed bank check authentication methodology. To model the problem in a pattern recognition framework, images of the reference and questioned checks are first captured using a flatbed scanner. Some of the areas in the scanned checks with background designs are selected as the regions of interest. From the regions of interest on both types of checks, 2D histograms of the hue and saturation as color features and the gray level co-occurrence matrix of the intensity component as texture features are extracted. The feature vectors of the two checks are then combined to create a single feature vector.

After pairing the feature vectors, the resulting vector is submitted to a support vector machine classifier for authentication. Prior to the testing phase, the support vector machine is trained using a sufficient

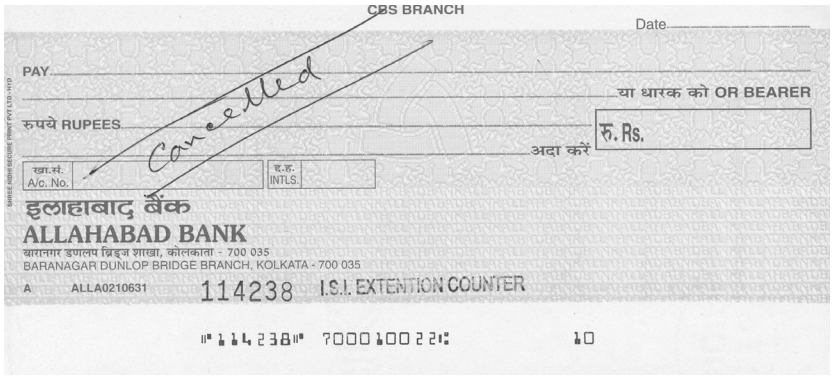

Figure 2. Original check.

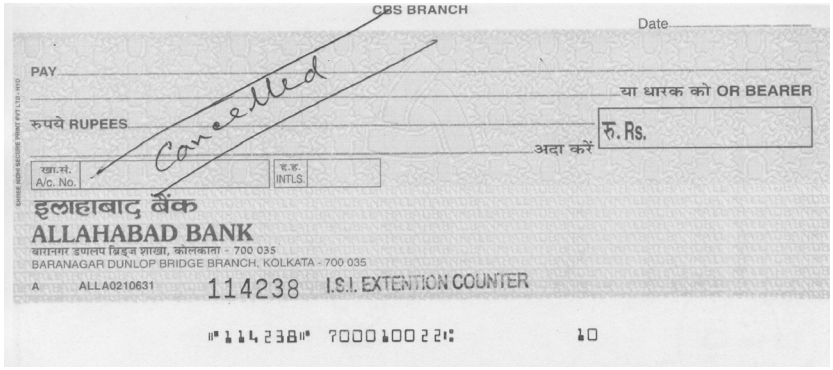

Figure 3. Counterfeit check created from the original check.

number of genuine-genuine and genuine-fake check pairs from a bank check image dataset.

3.1 Bank Check Image Dataset

The dataset used in the experiments comprised images of 25 genuine checks (five checks from five different banks) and 25 counterfeit checks. The counterfeit checks were created by scanning the 25 genuine checks using a flatbed scanner. The 25 scanned images were then reproduced using a high quality color printer on 100 GSM paper so that the printed checks looked like their original counterparts. After the counterfeited checks were printed, color images of all 50 checks (25 genuine and 25 fake) were captured at 300 dpi using a flatbed scanner. Figure 2 shows an original check and Figure 3 shows the counterfeit check created from the original check.

In a real scenario, a used check is sent for forensic analysis when its authenticity is suspected. A used check usually contains handwritten or printed information and a signature. Extracting features from an entire check introduces some bias. Keeping these facts in mind, regions containing background patterns (i.e., micro-printing) were selected for feature extraction purposes. Since a reasonable number of security features were embedded in the check backgrounds, the goal was to detect the discrepancies in the security features that resulted from the scanning and printing processes used to create the counterfeit checks.

3.2 Color and Texture Feature Extraction

The color and texture of the bank checks were assumed to be the principal features for the pattern recognition problem underlying check authentication. The reason is that, although the genuine and fake checks look similar, the changes in color and texture due to scanning and/or printing can be used to distinguish between the two types of checks. Specifically, color scanning and color printing are different tasks that involve different proprietary technologies.

To capture color information, the RGB color components of the check images were converted to the HSI (hue, saturation and intensity) color space as follows:

$$H = \frac{\sqrt{3}(G - B)}{2R - G - B} \tag{1}$$

$$S = 1 - \frac{min(R, G, B)}{255} \tag{2}$$

$$I = \frac{(R + G + B)}{3} \tag{3}$$

The HSI color space has been shown to be close to human perception and has been used in a number of computer vision problems, including face recognition [1]. Since color information in the HSI space is only contained in the hue and saturation components, color attributes were extracted as 2D hue-saturation histograms. The hue provides chromaticity (color spectrum) information while saturation expresses the purity of the color. Thus, a 2D histogram of hue and saturation gives the relative distributions of a particular hue (color) and its purity (saturation). To obtain a hue-saturation histogram, the hue was considered to range from 0° to 359° with a bin size of 30° while saturation ranged from 0 to 1 in a linear scale with a bin size of 0.2. This approach yielded a total of 60 features capturing color information.

To capture texture information, the gray level co-occurrence matrix (GLCM) of the intensity component was computed. The gray level co-

occurrence matrix [5] is a matrix of relative frequencies with which gray values of two pixels, separated by a distance d and at an angle θ with the horizontal axis, occur on an image. To compute the features, for a given distance d, a joint probability matrix was determined by summing the gray level co-occurrence matrices for different values of θ and then normalizing the result.

Let $p_{\alpha\beta}$ be the joint probability of co-occurrence of two pixels with intensities α and β separated by (d, θ) in polar coordinates. Some commonly-used features based on a gray level co-occurrence matrix (of size $S \times S$) are defined as follows [5]:

$$Contrast = \sum_{\alpha,\beta=0}^{S-1} p_{\alpha\beta}(\alpha - \beta)^2 \tag{4}$$

$$Homogeneity = \sum_{\alpha,\beta=0}^{S-1} \frac{p_{\alpha\beta}}{1 + (\alpha - \beta)^2} \tag{5}$$

$$Energy = \sum_{\alpha,\beta=0}^{S-1} p_{\alpha\beta}^2 \tag{6}$$

$$Correlation = \sum_{\alpha,\beta=0}^{S-1} p_{\alpha\beta}[\frac{(\alpha - \mu_\alpha)(\beta - \mu_\beta)}{\sigma_\alpha \sigma_\beta}] \tag{7}$$

where:

$$\mu_\alpha = \sum_{\alpha,\beta=0}^{S-1} \alpha p_{\alpha\beta} \tag{8}$$

$$\mu_\beta = \sum_{\alpha,\beta=0}^{S-1} \beta p_{\alpha\beta} \tag{9}$$

$${\sigma_\alpha}^2 = \sum_{\alpha,\beta=0}^{S-1} (1 - {\mu_\alpha}^2) p_{\alpha\beta} \tag{10}$$

$${\sigma_\beta}^2 = \sum_{\alpha,\beta=0}^{S-1} (1 - {\mu_\beta}^2) p_{\alpha\beta} \tag{11}$$

The four features (contrast, homogeneity, energy and correlation) were extracted for three distances ($d = 5$, 10 and 15). This yielded a total of 12 ($= 4\times$ 3) texture features extracted from the gray level co-occurrence matrix. Upon combining the color and texture features, a feature vector of dimension 72 was created to represent a bank check.

3.3 Authentication

The authentication of a bank check was formulated as a two-class pattern recognition problem. Thus, for a given pair of bank checks, one of them known to be genuine (reference check), it is necessary to determine whether or not the two checks are similar; in other words, whether or not the second check (questioned check) is genuine.

A support vector machine classifier was used to solve the two-class pattern recognition problem. The input to the support vector machine classifier must be a vector that represents both the checks. Thus, a (dis)similarity index obtained by taking the absolute difference of the corresponding elements in the feature vector of the two checks was used as the input vector.

Let $v_r = (v_{r_1}, v_{r_2}, \cdots, v_{r_n})$ and $v_q = (v_{q_1}, v_{q_2}, \cdots, v_{q_n})$ be the feature vectors of the reference and questioned checks, respectively, where $n = 72$. Then, the combined vector obtained by pairing the two vectors is given by:

$$v = (|v_{r_1} - v_{q_1}|), (|v_{r_2} - v_{q_2}|), \cdots, (|v_{r_n} - v_{q_n}|) \tag{12}$$

Following the usual leave-one-out pattern recognition strategy, the dataset was divided into two parts, one for training and the other for testing. A four-fold cross-validation over the training data was used to select the parameters of a support vector machine with a radial basis function (RBF) kernel. The optimal parameter, which were selected via cross-validation, were utilized for classifier design. The support vector machine was trained using the selected parameters and evaluated using the testing set. The classifier was designed to output 1 or -1 corresponding to both checks being genuine or one of the checks being counterfeit, respectively.

4. Results and Discussion

The experimental evaluation was conducted using a dataset of 50 checks (25 genuine and 25 fake). The 25 genuine checks came from five different banks, five checks per bank. Each genuine check was used to create one counterfeit check, yielding the 50 checks in the dataset. Following the leave-one-out strategy, all the checks from four of the banks

(i.e., 20 genuine and 20 fake checks) were kept for training. The checks from the remaining (fifth) bank (five genuine and five fake) were used for testing.

Consider the checks from a particular bank (five genuine and five fake checks). Although they were taken from different sources, the five genuine checks from the bank are similar (contemporary checks) and have similar security features. A pair of two genuine checks is formed in ${}^5C_2 = 10$ different ways. Similarly, a pair of genuine and fake checks from the ten checks is formed in $5 \times 5 = 25$ ways. Thus, for all the checks corresponding to a bank, there are ten input vectors for Class I (i.e., genuine-genuine) pairs and 25 input vectors for Class II (genuine-fake) pairs. To avoid bias towards any class, ten pairs were randomly selected out of the 25 vectors for Class II.

A total of 80 input vectors from four banks were designated for training and 20 input vectors from one bank were designated for testing. Of the 80 training vectors, 75% were randomly chosen for inner-level training and the remaining 25% for validation. This method was repeated five times so that the checks from each of the five banks could be utilized for testing purposes. The overall performance was computed as the mean of the five repetitions.

Experiments involving the trained support vector machine classifier yielded 99% overall accuracy for bank check authentication. The high accuracy may be due to several factors. One is that the counterfeit checks were created by scanning genuine checks and printing counterfeits in a pristine laboratory environment. This does not capture the security features of the genuine checks adequately and reliably. Moreover, in a real scenario, criminals would be likely to use more sophisticated techniques and equipment to create fake checks. The second reason is the small dataset. The third reason is the potential for the extracted features to capture color and texture information in very precise manner; gray level co-occurrence matrix features have been demonstrated to produce good results in a number of applications.

5. Conclusions

The proposed automated methodology for the forensic authentication of bank checks is implemented as a two-class pattern recognition problem involving pairs of checks: (i) Class I, when the reference and questioned checks are both genuine; and (ii) Class II, when one of the two checks (i.e., the questioned check) is fake. Color (2D hue-saturation histograms) and texture (gray level co-occurrence matrix of the intensity component) features were extracted from images of genuine and counterfeit checks. A

trained support vector machine was utilized to determine check authenticity. Classification experiments involving a dataset of 50 bank checks yielded a detection accuracy of 99.0%. The automated methodology can be used by non-specialist personnel to detect check counterfeiting in a banking environment where large numbers of checks are handled on a daily basis.

References

[1] B. Chanda and D. Majumdar *Digital Image Processing and Analysis*, Prentice Hall of India, New Delhi, India, 2005.

[2] C. Chen and C. Chiu, A fuzzy neural approach to design of a Wiener printer model incorporated into model-based digital halftoning, *Applied Soft Computing*, vol. 12(4), pp. 1288–1302, 2012.

[3] U. Garain and B. Halder, On automatic authenticity verification of printed security documents, *Proceedings of the Sixth Indian Conference on Computer Vision, Graphics and Image Processing*, pp. 706–713, 2008.

[4] G. Gupta, S. Saha, S. Chakraborty and C. Mazumdar, Document frauds: Identification and linking fake documents to scanners and printers, *Proceedings of the International Conference on Computing: Theory and Applications*, pp. 497–501, 2007.

[5] R. Haralick, K. Shanmugam and I. Dinstein, Texture features for image classification, *IEEE Transactions on Systems, Man and Cybernetics*, vol. SMC-3(6), pp. 610–621, 1973.

[6] E. Kee and H. Farid, Printer profiling for forensics and ballistics, *Proceedings of the Tenth ACM Workshop on Multimedia and Security*, pp. 3–10, 2008.

[7] N. Khanna, A. Mikkilineni and E. Delp, Scanner identification using feature-based processing and analysis, *IEEE Transactions on Information Forensics and Security*, vol. 4(1), pp. 123–139, 2009.

[8] K. Koppenhaver, *Forensic Document Examination: Principles and Practice*, Humana Press, Totowa, New Jersey, 2007.

[9] C. Lampert, L. Mei and T. Breuel, Printing technique classification for document counterfeit detection, *Proceedings of the International Conference on Computational Intelligence and Security*, vol. 1, pp. 639–644, 2006.

[10] Reserved Bank of India, Mechanized Cheque Processing Using MICR Technology – Procedural Guidelines, Mumbai, India (`www.rbi.org.in/scripts/PublicationsView.aspx?id=4551`), 2014.

[11] A. Roy, B. Halder and U. Garain, Authentication of currency notes through printing technique verification, *Proceedings of the Seventh Indian Conference on Computer Vision, Graphics and Image Processing*, pp. 383–390, 2010.

[12] U.S. Department of Justice and Public Safety Canada, Public Advisory: Special Report on Counterfeit Checks and Money Orders, Washington, DC and Ottawa, Canada (`www.justice.gov/opa/documents/08public-advisory-counterfeit.pdf`), 2008.