

Banknotes Recognition in Real Time Using ANN

Yueqiu Ren

A thesis submitted to Auckland University of Technology in
partial fulfilment of the requirements for the degree of Master
of Computer and Information Sciences (MCIS)

2017

School of Engineering, Computer and Mathematical Sciences

Abstract

Financial institutions have adopted various automated banking systems using currency recognition as their main activity, which makes automated currency recognition of significant interest. However, after the review of the literature related to banknote recognition, it turns out that there has not been found any methods implemented or proposed for the recognition of the newly released banknotes. This thesis investigates various methods for achieving banknote real-time recognition using digital image processing. The new Series 7 New Zealand banknotes are considered as an example for intelligent banknote recognition in real time.

Several experiments have been conducted in this study and two groups of training datasets are generated for comparison. One group is composed of banknote images produced by using scanners, and the other group is made up of banknote images captured by webcam. Various combinations of extracted features and classifiers have been analysed. The corresponding results are compared and the performance of each combined method is evaluated. Eventually, the PCA-based composite feature together with the BPNN is the combined method proposed in this thesis. The proposed method has demonstrated excellent performance and comparatively less time-consumption that makes it suitable for real-time applications. To the best of our knowledge, the composite feature containing both colour and texture elements, presented in this thesis has appeared in the field of banknote recognition for the first time. Our contribution is that this research project fills the vacancy of the real-time recognition of the newly released banknotes; and the proposed method paves the way for the future development of multi-currency real-time recognition.

Keywords: Series 7 New Zealand paper currency, real-time banknote recognition, HSV colour quantisation, uniform LBP, the minimum distance classifier, back-propagation neural network, F-measure

Table of Contents

Abstract	I
List of Figures	IV
List of Tables	V
Attestation of Authorship	VI
Acknowledgements	VII
Chapter 1 Introduction	1
1.1 Background.....	1
1.2 Fundamentals.....	2
1.3 Objective of the thesis	6
1.4 Structure of the thesis	7
Chapter 2 Literature review	9
2.1 Overview of DIP-based banknote recognition.....	9
2.2 Digital image processing.....	16
2.2.1 Feature engineering	16
2.2.2 Feature extraction methods.....	19
2.2.3 Dimensionality reduction methods.....	23
2.3 Classification Algorithms.....	29
2.3.1 Minimum distance classifier.....	29
2.3.2 Back-propagation neural network classifier	30
Chapter 3 Methodology	33
3.1 Related work	33
3.2 Research questions and hypotheses	37
3.3 Research Design.....	40
3.3.1 Experimental platform.....	40
3.3.2 Experimental samples.....	41
3.3.3 Combined methods	41
3.3.4 Flowcharts	42
3.3.5 KPI.....	45
Chapter 4 Implementation	47
4.1 Training procedure.....	47
4.1.1 Training set.....	47
4.1.2 Pre-processing	48
4.1.3 Feature extraction	48
4.1.4 Dimensionality reduction	63
4.1.5 Classification model	65
4.2 Testing procedure.....	67
Chapter 5 Findings and discussion	69
5.1 Experimental results.....	69

5.2	Hypothesis validation.....	72
5.3	Discussion	77
Chapter 6	Conclusion	84
6.1	Summary	84
6.2	Novelty of the research	85
6.3	Significance of the research	86
6.4	Limitation and future work	86
References	89
Appendices	98

List of Figures

Figure 1.1 The security features of a \$10 New Zealand banknote – front (Reserve Bank of New Zealand).....	7
Figure 1.2 The security features of a \$10 New Zealand banknote – back (Reserve Bank of New Zealand).....	7
Figure 2.1 An example of LBP operator.....	21
Figure 2.2 Thirty-six rotation-invariant binary patterns	22
Figure 2.3 Uniform LBP based on a neighbour set of eight members on a circle with a radius of one pixel	23
Figure 2.4 A comparison of mapping results between PCA and LDA.....	28
Figure 2.5 A schematic diagram of BPNN	32
Figure 3.1 Series 7 New Zealand banknotes of all different denominations	36
Figure 3.2 Series 6 New Zealand banknotes of all different denominations	36
Figure 3.3 Flowcharts of the camera-based banknote real-time recognition	44
Figure 4.1 The comparison between the banknote images produced by scanner and the images produced by webcam	48
Figure 4.2 Colour histograms with 64 bins for the scanned banknote images	49
Figure 4.3 Colour histograms with 16 bins for the scanned banknote images	50
Figure 4.4 A comparison of before and after use of histogram equalisation for the scanned banknote images	53
Figure 4.5 Thirty-two uniform blocks for the scanned banknote images	54
Figure 4.6 LBP features for the scanned banknote images.....	56
Figure 4.7 Block uniform LBP histograms for the scanned banknote images.....	59
Figure 4.8 LBP histogram part in the colour-LBP histogram for the scanned banknote images.....	62
Figure 4.9 Testing window in MATLAB	68
Figure 5.1 Comparison of the average time cost of recognition using each combined method in the two groups of training sets	71
Figure 5.2 The average time cost of recognition using each combined method.....	71

List of Tables

Table 3.1 Different combinations of extracted features and classifiers	42
Table 4.1 The dimensionality of feature vectors after dimensionality reduction	64
Table 4.2 Parameters of the optimal networks when using the scanned banknote images for training	67
Table 4.3 Parameters of the optimal networks when using the banknote images captured by webcam for training.....	67
Table 5.1 The F-measure of each combined method for recognising each denomination when using scanned banknote images for training.....	70
Table 5.2 The F-measure of each combined method for recognising each denomination when using the banknote images captured by webcam for training.....	70
Table 5.3 The comparison of the average F-measure between the training sets.....	72
Table 5.4 The comparison of the average F-measure between the extracted features....	73
Table 5.5 The comparison of the average F-measure between the classifiers	75

Attestation of Authorship

I hereby declare that this submission is my own work and that, to the best of my knowledge and belief, it contains no material previously published or written by another person (except where explicitly defined in the acknowledgments), nor material which to a substantial extent has been submitted for the award of any other degree or diploma of a university or other institution of higher learning.

Signature: *Yueqin Ren*

Date: 18/07/2017

Acknowledgements

I would like to express my deep gratitude to my parents for their encouragement and support throughout my academic life. I am also extremely grateful to have Dr Wei Qi Yan as my supervisor, who introduced the area of pattern recognition and image processing to me, which opened the door to the exploration of New Zealand banknote recognition. Many thanks to Ms Ksenya Nefiodova for her patience with my enquiries about the Master of Computer and Information Sciences (MCIS) programme as a postgraduate administrator of the School of Engineering, Computer and Mathematical Sciences at AUT. I would also like to say thanks to my fellow members who are also enrolled in MCIS programme of the same term as mine. The valuable academic communication, as well as the friendly little chats with them, makes this hard, but interesting challenge – postgraduate study – much more enjoyable. Last, but not the least, thanks must go to Dr Nicola Saywell and Colin who are always interested in proofreading my work and have made my stay in New Zealand truly memorable.

Chapter 1 Introduction

1.1 Background

Currency started serving as a medium for exchanging goods and services thousands of years ago to replace the ancient barter system where any objects could be swapped if two traders agreed (Maestro, 1993). Even nowadays, currency, as a measurement unit in pricing a transaction, still plays an indispensable role in modern society. For example, it is used as a medium of payment in tackling debt, and as a store of value for savings (Bender, 2006). The monetary form has been extended to cash including coins and banknotes, cashless money like bank cheques, and even electronic data representing currency in bank accounts.

Banknotes can be traced back to the year 1023 when they officially appeared in China for the first time in history, called “jiaozi”, and were later introduced by American colonists for systematic use in the western world (Bender, 2006; Maestro, 1993). In spite of its long history, the worldwide market for banknote printing is still fairly confidential, which is typically justified as the intention of protecting the secure surroundings for the production of this unique product. The printing apparatus and the security inks, as well as the completely automated machinery for excellent accuracy of banknote examination or the highly secure shredding equipment for used notes are not familiar to the general public.

With little revelation of the techniques for the production of banknotes, nevertheless, there is an enormous amount of research starting to reveal the inside story of the banknote, especially in the field of banknote recognition. Nowadays, numerous paper currency recognition systems have been developed through secure analysis, and have had a wide range of applications, such as automated teller machines (ATMs), banknote sorting machines, self-service payment kiosks, and portable devices assisting the visually impaired with recognising banknotes.

Banknote recognition mainly concerns the process of identifying the denomination of a banknote, particularly when a single currency is to be studied (Vishnu & Omman, 2014). In essence, it is the process of classifying the banknote to one of the classes it belongs to (Sargano, Sarfraz, & Haq, 2014). Unlike coins minted by heavy metals, almost all the banknotes worldwide are produced to be as thin as ordinary paper (Tarnoff, 2011). The newly released Series 7 New Zealand banknotes are a typical example. They are printed on paper polymer, which is a type of polypropylene plastic, enabling them to be lightweight (Langwasser, 2014). If weighing coins could be a straightforward way to distinguish different denominations of coins, then apparently, weighing banknotes is not a feasible way to distinguish different denominations of banknotes, due to the intrinsic lightweight property of the raw material. Even so, many other distinctive characteristics of banknotes allow them to be validated and classified by various methods. Those features could be grouped together by their accessibility levels. Level 1 contains the features that can be detected by human sense, such as substrate fidelity, print fidelity, colour fidelity, acoustic fidelity, serial number, holograms, watermark, security thread, security fibre, tactile fidelity, colour-shifting ink, clear window or latent image; Level 2 contains the features that can be detected by minor manipulation, such as micro-text or invisible glowing ink; Level 3 contains the features that could be detected by security analysis, such as magnetic ink, screen traps, manufacture anomalies, materials interaction, intricate patterns, intricate design, or fluorescence eminence (Chambers, 2012).

1.2 Fundamentals

Traditionally, machine learning has been studied either in a supervised paradigm like classification and regression, or in an unsupervised paradigm such as clustering and outlier detection. Supervised learning presumes that the training set has been provided, composed of a set of examples that have been appropriately labelled with the correct output (Preston & Carvalko, 1972; Samarasinghe, 2006). Based on the training set, a supervised learning method generates a model seeking to meet the two targets which are performing as well as possible on the set of training data and generalising as well as

possible to new data (Marsland, 2015; Preston & Carvalko, 1972). It is also called learning from exemplars. On the contrary, in unsupervised learning, the correct output of training data are not provided, or there are no training data at all to speak of (Preston & Carvalko, 1972). Instead, the algorithm attempts to identify the similarities between the inputs, so that the inputs which have something in common are categorised together (Marsland, 2015). Semi-supervised learning initially emerged in the 1970s, when self-training, transduction, and Gaussian mixtures with the expectation-maximisation (EM) algorithm were developed (Zhu, 2011). An explosion of interest in semi-supervised learning occurred in the 1990s, with the development of new theoretical analyses, new algorithms like transductive Support Vector Machines (SVM) and co-training, and new applications in natural language processing and computer vision (Chapelle, Zien, & Schölkopf, 2006). It is a learning paradigm regarding the study of how computers learn in the presence of both labelled and unlabelled data. The purpose of semi-supervised learning is to explore how combining labelled and unlabelled data may change the learning behaviour, and to design algorithms which such a combination can benefit (Zhu & Goldberg, 2009).

Banknote recognition is a typical case of pattern recognition. The definition of the term pattern recognition has provoked wide-ranging discussion in the literature. In general, pattern recognition is interpreted as a branch of machine learning focusing on the recognition of patterns and regularities in data, though it is sometimes described as being synonymous with machine learning (Bishop, 2006). Specifically, it refers to the act of taking in raw data and then taking action based on the category of the pattern (Duda, Hart, & Stork, 2012). It is worth noting that Bezdek (2013) defined pattern recognition in a quite simple, literal way – it is a search for the structure of data. Three key elements can be extracted from the definition. The first element is “the data”, which drawn from realistic scenes and to be processed with the techniques of pattern recognition may be quantitative, qualitative, or both. They can also be depicted as being numerical, linguistic, pictorial, and textual, or any combination thereof. Examples are banknotes, fingerprints, gestures, demographic features, medical records, chemical

constituents, market trends, and so forth. The second element is “the search”, which is concerned with the techniques used to process the data. Until now, statistical pattern recognition has been a powerful technique in the search. Notwithstanding, the type of search to be adopted relies on both the data and the structure supposed to be found. Data is assumed to carry information about the process of generating them. “The structure”, as the third element in the definition, then stands for the approach in which the information can be organised, and hence the relationships between the variables in the process can be recognised by the structure. The relationships can be either invertible association or noninvertible association. Basically, regarding the information, the data contains it, the search recognises it, and the structure represents it (Bezdek, 2013).

Similar to machine learning, pattern recognition can also be categorised depending on the kind of learning method used to produce the output value. Accordingly, there are three types of pattern recognition, which are also commonly perceived as pattern classification, namely, supervised classification, unsupervised classification and semi-supervised classification (Guillaumin, Verbeek, & Schmid, 2010). The main difference among these three types lies in whether the categories of all the experimental samples are known beforehand. Supervised classification often needs a considerable number of samples to be provided whose categories are known in advance (Bow, 2002; Sethi & Jain, 2014). However, it seems difficult to achieve such an essential prerequisite in many practical situations, for example, the cases when the labelled data are expensive or scarce. Under the circumstances, semi-supervised classification could be useful by manipulating readily available unlabelled data or combining supervised classification with strategies known from unsupervised classification, to improve supervised classification tasks (Zhu & Goldberg, 2009).

Generally speaking, syntactic pattern recognition (structural pattern recognition) approach and statistical pattern recognition are the two basic approaches to pattern recognition, though neural pattern recognition remains controversial. Neural pattern recognition (or neural network pattern classification) is occasionally claimed to be a

separate discipline, due to its rather different intellectual pedigree (De Sa, 2012; Schalkoff, 1992), while many researchers perceive it as a close descendant of statistical pattern recognition (Cherkassky, Friedman, & Wechsler, 2012; Jain, Duin, & Mao, 2000). Statistical pattern recognition is a supervised approach to pattern recognition, as the classification of patterns in the training set is known beforehand (De Sa, 2012). It derives classifying functions primarily based on the probabilistic models of feature vector distributions in the classes (De Sa, 2012; Fukunaga, 2013). By analysing the set of training samples, the given statistical models or known discriminant function, based on a certain criterion, n -dimensional vector space is divided into several sections pertaining to the category. In this way, as long as the pattern recognition system figures out which section the object for testing is in, the category of the testing object is supposed to be confirmed.

The variation in patterns within a category is partly caused by environmental noises and the sensors, such as the effects of stain, and the quality of paper or ink on the characters of the writing. Even if the person continuously writes the same character a few times in a row, the written characters would look similar but not be identical, which manifests as the random nature of the pattern itself. Accordingly, when using feature vectors to represent those written characters containing minor shape discrepancies, the points in the feature space corresponding to the feature vectors are distributed in a certain section, rather than converging. Then that certain section can represent the realisation of the set of the random vectors (Li, Liu, & Wang, 2014). For a predetermined distance in the feature space, theoretically, the closer the two points, the more likely that the corresponding patterns are similar. Ideally, the distance between the two patterns of different classes is greater than that of the same class. In addition, for the line segment connecting any two points of the same class, the patterns of all the points on the line segment should belong to the same class. Moreover, the point matching a pattern with a small distortion should be contiguous with that with no distortion. Under such ideal conditions, the feature space can be accurately divided into a range of parts about different classes. On the other hand, when these conditions are not met, its probability

of the class that each feature vector belongs to is estimated, thereby rating the class with the highest probability as the class that point belongs to. It is worth mentioning that a number of techniques can be used to approximate the ideal conditions. The distortion caused by environmental noises and the sensors can be partially eliminated through pre-processing. The distortion resulting from the random nature of the pattern itself, can be effectively controlled by feature extraction and selection. Therefore, in order to achieve a satisfactory distribution of patterns in the feature space, a statistical pattern recognition should at least include data collection, pre-processing, feature extraction and selection, classifiers.

1.3 Objective of the thesis

Automatic machines that are capable of recognising banknotes are widely used in automatic dispensers of a range of products, from beverages to tickets, as well as in many automatic banking operations. Hence, it is essential to have automated currency recognition to carry out successful financial transactions. In this thesis, banknotes are taken into consideration. After reviewing the literature about banknote recognition, it appears that there is little research on New Zealand banknote recognition, and no system found or method implemented or proposed for the newly released Series 7 New Zealand banknotes. Thus, the main purpose of this thesis is to seek out a solution for recognising the new Series 7 New Zealand banknotes, or more specifically, banknote denomination recognition.

In 2014, the Reserve Bank of New Zealand selected the designs for Series 7 New Zealand banknotes and published them in November (Langwasser, 2014). The Canadian Banknote Company Ltd (CBN) won the contract for the design and printing of the new banknote series in an open tender process (Langwasser, 2014). Series 7 New Zealand banknotes were designed and printed to meet a variety of functional, cultural and aesthetic requirements. Compared with Series 6 banknotes, Series 7 banknotes contain features that are much easier for people to identify. Several features of Series 6 notes were enhanced and then embedded into Series 7 notes, and other characteristics of

Series 7 banknotes are new. An overview of the main security features is illustrated on the official website of the Reserve Bank of New Zealand, as shown in Figure 1.1 and Figure 1.2.



Figure 1.1 The security features of a \$10 New Zealand banknote – front (Reserve Bank of New Zealand)



Figure 1.2 The security features of a \$10 New Zealand banknote – back (Reserve Bank of New Zealand)

1.4 Structure of the thesis

The remainder of this thesis is organised as follows. Chapter 2 is a comprehensive literature review of paper currency recognition. Banknote recognition models are reviewed from the viewpoint of digital image processing, including single-currency recognition models and multi-currency recognition models. Digital image processing is then introduced, unfolding the elaboration on feature engineering, feature extraction

methods and feature vector dimensionality reduction methods. Among them, colour feature extraction and texture feature extraction are highlighted, and principal component analysis (PCA) and linear discriminant analysis (LDA) are examined as effective vector dimensionality reduction methods. The most frequently used classifiers in the reviewed models are also interpreted, including the minimum distance classifier (MDC) and back-propagation neural network (BPNN) classifier. Chapter 3 outlines the methodology used in this research, including describing the related study, bringing up the research questions along with the hypotheses, and elaborating on research design. Chapter 4 presents the implementation of the experiments, covering the procedures of generation of the dataset, pre-processing, feature extraction, dimensionality reduction, and classification. The outcome of the experiments is analysed in Chapter 5. Finally, Chapter 6 concludes this research, including its novelty, significance, limitations and future work.

Chapter 2 Literature review

Paper currency recognition is a certain kind of application of pattern recognition. This chapter is organised to give an insight into the areas relevant to banknote recognition, primarily from the viewpoint of digital image processing and pattern recognition. In Section 2.1, an overview of digital image processing-based banknote recognition will be presented. In Section 2.2, feature engineering that has a dominant influence on the field of digital image processing will be clarified, followed by the interpretation of feature extraction on colour feature, texture feature and the composite feature in Section 2.2.2 and the explication of dimensionality reduction methods - PCA and LDA in Section 2.2.3. In Section 2.3, the MDC and BPNN classifiers that are the most widely used classifiers for banknote recognition will be elucidated in Section 2.3.1 and Section 2.3.2, respectively.

2.1 Overview of DIP-based banknote recognition

Since each kind of paper currency has unique features, to some extent, the method used for recognition is individual to each currency. Accordingly, some of the existing models for banknote recognition only target currency of a particular country. These are known as single-currency recognition models, such as the recognition model for Australian banknotes (Hinwood, Preston, Suaning, & Lovell, 2006), Bangladeshi banknotes (Jahangir & Chowdhury, 2007), Chinese banknotes (Zhang, Jiang, Duan, & Bian, 2003), Egyptian banknotes (Semary, Fadl, Essa, & Gad, 2015), Ethiopian banknotes (Zeggeye & Assabie, 2016), Euro banknotes (Lee, Jeon, & Kim, 2004), Indian banknotes (Kamal, Chawla, Goel, & Raman, 2015; Sawant & More, 2016; Vishnu & Omman, 2014), Mexican banknotes (García-Lamont, Cervantes, & López, 2012), New Zealand banknotes (Chambers, 2012; Yan & Chambers, 2013; Yan, Chambers, & Garhwal, 2015), Pakistani banknotes (Ali & Manzoor, 2013; Sargano et al., 2014), Persian banknotes (Ahangaryan, Mohammadpour, & Kianisarkaleh, 2012), Saudi Arabian banknotes (Sarfranz, 2015), Sri Lankan banknotes (Gunaratna, Kodikara, & Premaratne, 2008), and United States banknotes (Grijalva, Rodriguez, Larco, & Orozco, 2010).

There are also banknote recognition models aimed at the currency of more than one country, which are known as multi-currency recognition models (Khashman & Sekeroglu, 2005; Pham et al., 2016; Takeda, Nishikage, & Omatu, 1999; Takeda & Omatu, 1995; Youn, Choi, Baek, & Lee, 2015).

These reviewed banknote recognition models can also be grouped by the involved classifiers. The majority of them have utilised the MDC, while others have adopted cross-correlation-based template matching classifier or correlation coefficient-based template matching classifier.

- MDC

A reliable prototype with the aid of a smartphone camera was developed to help visually impaired people recognise United States paper currency in circulation in Ecuador (Grijalva et al., 2010). Each frame was processed to output the audio message communicating the face value of the note in front of the camera. Eigenfaces based on the PCA combined with Mahalanobis distance-based MDC were employed in the system. The prototype demonstrated a recognition rate of 99.838% in ideal conditions, 99.156% in indoor conditions and 95.223% in outdoor conditions, with a processing speed of no less than seven frames per second. Vishnu and Omman (2014) put forward a PCA-based framework for Indian banknote recognition in their paper published in 2014. Five security features including centre number, shape, Reserve Bank of India (RBI) seal, latent image, and micro letter were used for their study. Principal components of the banknote features were extracted and the weight vector similarities were then computed using the Mahalanobis distance. The framework classified Indian currency notes with 96% accuracy in the experiments. Sawant and More (2016) carried out research into Indian paper currency recognition using Euclidean distance-based MDC. Four security features including dominant colour, aspect ratio, ID mark and latent image numbers were taken into consideration for banknote classification as per denomination. The Euclidean distance between the testing sample and the mean value for each class was computed and then the testing sample was allocated to the class with

the minimum distance. The proposed system demonstrated approximately 90% accuracy on Indian banknote recognition.

- Other classifiers belonging to template matching

A simple currency recognition system was developed for Egyptian banknotes (Semary et al., 2015). Digital image processing was applied to the system, including image foreground segmentation, histogram enhancement and Region of Interest (ROI) extraction. Meanwhile, cross-correlation-based template matching was used for classification according to the similarity between images in the database and the ROI part of the testing sample. The recognition accuracy of the proposed method reached 89% under the MATLAB system, with an average recognition time of 10s per banknote. A fast and efficient algorithm was proposed making use of banknote size information and multi-template correlation matching for multi-currency recognition (Youn et al., 2015). Multi-template correlation matching determined the discriminant areas of each banknote that are highly correlated among banknotes of the same type and poorly correlated among banknotes of different types. The correlation coefficient was chosen as the similarity metric. The proposed algorithm was tested on 55 banknotes with different denominations of USD, EUR, KRW, CNY and RUB. It turned out to have 100% classification accuracy for unspoiled note recognition and 99.8% classification accuracy for soiled notes, with the average processing time of 4.83ms per banknote. Zeggeye and Assabie (2016) described their design of automatic recognition of Ethiopian banknotes, where hardware and software solutions worked on taking images of an Ethiopian banknote from a scanner and camera as an input. Four characteristic features, i.e. the dominant colour, the distribution of the dominant colour, the *hue* value, and speeded up robust features (SURF) were extracted as the discriminative features of banknotes. Those features in combination with local feature descriptors were involved in a four-level classification process, as a classification task executed every time one of the four features was extracted. The correlation coefficient-based template matching was implemented for classification. Test results showed that the proposed design had an average recognition rate of 90.42% for genuine Ethiopian currency, with an average

processing time of 1.68 seconds per banknote.

Apart from the group of template-matching classifiers, the group of artificial neural network (ANN) classifiers has also been widely adopted for paper currency recognition, such as BPNN, learning vector quantisation (LVQ) and radial basis function network (RBFN).

- BPNN classifier

The method of improving the recognition rate and the transaction speed of classifying Japanese as well as US banknotes was explored (Takeda & Omatu, 1995). The proposed BPNN consisted of three layers, with each layer containing 16, 16 and 12 nodes. They also introduced the random masks technique to reduce the scale of the network, and proved its effectiveness for time series data and Fourier power spectra directly serving as the inputs to the network. The recognition rate reached more than 92% and the recognition speed was no less than 3 notes per second, by using the proposed method. A few years later, another BPNN-based banking machine was developed using neural weights and mask sets optimised by the genetic algorithm, which is applicable to US dollars, British pounds, French francs, Spanish pesetas, Italian lira, Australian dollars, Korean won, Belgian francs, German marks and Japanese yen (Takeda et al., 1999). Fifty to 100 notes of each kind served as the training set, and over 20,000 notes worked as the testing set. The machine was verified to be able to recognise more than 97% of testing samples, and the recognition speed was enhanced by nearly 20% compared with the method proposed by Takeda and Omatu (1995). A study was conducted where a neural network-based approach for recognising Chinese paper currency RMB was analysed (Zhang et al., 2003). For each RMB image, linear transform and edge detection were used to obtain the image particularly representing the characteristic edge information, which was then divided into different areas as the input vectors to a three-layer BPNN for classification. The results showed that the recognition ratio to the new edition of 100 RMB, new edition of 50 RMB, old edition of 50 RMB, 20 RMB and new edition of 10 RMB were 95%, 99%, 99%, 92% and 98%. It was suggested that

distinctive point extraction algorithms for Euro banknotes should be applied to high-speed banknote-counting machines (Lee et al. 2004). BPNN was employed as the classifier to minimise the recognition error rate. Accordingly, five neural networks were trained, with one for inserting direction and the others for denomination recognition. The experimental results demonstrated a 100% recognition rate on 5, 10, 20, 50, 500 euro and a 95% recognition rate on 100 and 200 euro. A three-layer neural network was trained using a back-propagation learning algorithm for the recognition of the two main currencies used in Cyprus, namely Turkish Lira and Cyprus Pound (Khashman & Sekeroglu, 2005). The BPNN was composed of 100 neurones in the input layer, 30 neurones in the hidden layer and nine neurones in the combined output layer. Validation experiments confirmed the effectiveness of the algorithm by which an overall currency recognition ratio of 95% was obtained, with the average recognition time of 0.05 seconds. In 2006, a novel device named MoneyTalker was invented for visually impaired people to recognise Australian banknotes (Hinwood et al., 2006). The device utilised the broad range of colours and patterns on Australian paper currency to distinguish different denominations. Different algorithms were applied to the classification stage, including BPNN. The device was proved to be capable of reaching a recognition rate of more than 99%, and the time cost of recognition was around 3.01 seconds per note. In 2007, a neural network-based recognition scheme for Bangladeshi banknotes was proposed for the first time, when currency recognition machines were not yet used in Bangladesh (Jahangir & Chowdhury, 2007). In the proposed scheme, Axis Symmetric Masks were implemented in the pre-processing step to decrease the network size and to ensure correct recognition in case of flipped banknotes. The final network was constructed by 20 input nodes, 15 hidden nodes and eight output nodes, and trained via a back-propagation approach. This scheme recognised eight denomination classes of Bangladeshi banknotes, including 1, 2, 5, 10, 20, 50, 100 and 500 Taka, with an average accuracy of 98.57% and an average recognition time of 6.87ms per note. In 2008, a three-layer BPNN-based currency recognition system was developed for Sri Lankan banknotes (Gunaratna et al., 2008). In the process of feature extraction, the Canny algorithm was used for edge detection, and linear transform was

used to remove noise interference. 100, 500, 1000 and 2000 rupee notes were involved in the testing phase of the experiments. As a consequence, the system achieved 100% accurate classifications for all denominations with an average recognition time of around 200ms per note, and had the proven ability to separate classes properly in various image conditions. In 2012, a new Persian banknote recognition system that employed a discrete wavelet transform and multi-layer perceptron (MLP) neural network was presented (Ahangaryan et al., 2012). The three-layer neural network was trained by 138 images of Persian banknotes. After a trial-and-error process, the well-trained BPNN contained 57 input neurones, three hidden neurones and four output neurones. Three hundred and twenty samples of Persian paper currency consisting of 50, 100, 200, 500, 1000, 2000, 5000 and 10000 toman notes served as the testing set and were classified with a recognition rate of over 99%. An empirical approach was presented for Series 6 New Zealand paper currency security analysis, and a working prototype was developed in the MATLAB environment (Chambers, 2012). The research was intended to utilise digital image processing and classification to recognise the denomination of Series 6 New Zealand banknotes. A two-part feature vector composed of colour feature and texture feature was the input to the classifier, while the denomination of the testing banknote was the output. One hundred and sixty-seven banknote images were tested on the prototype. Finally, the feed-forward neural network classifier trained using Bayesian back-propagation regulation learning delivered the optimal performance at the accuracy of 98.6% (Yan & Chambers, 2013; Yan et al., 2015). An intelligent recognition system for Pakistani banknotes was developed in 2014, using a three-layer BPNN for classification (Sargano et al., 2014). The extracted feature in the study included aspect ratio, a set of useful colour features, a binary pattern of lettering block for the banknote, a binary pattern of see-through block, and a binary pattern of identification marks block for the banknote. The system was tested on 175 Pakistani banknotes and the results demonstrated a 100% recognition rate on properly captured images.

- Other ANN classifiers

In 2012, an effective approach to recognise Mexican banknotes was proposed where the colour and texture features were concatenated as a composite feature to investigate, and LVQ networks were adopted for classification (García-Lamont et al., 2012). The five denominations frequently used in daily transactions in Mexico were considered in the research, including 20, 50, 100, 200 and 500 Pesos. A 98.95% recognition rate was obtained with an average processing time of 0.9641 seconds, when applying the proposed approach to Mexican banknotes with the usual damage. An automatic Saudi Arabian paper currency recognition system was introduced by Sarfraz (2015). A few interesting features of the banknote, including the height, width, the area without a mask, the area with the first mask, the area with the second mask, the Euler number, and the correlation with the template image, served as the input vectors of the RBFN classifier adopted in his research. An overall recognition rate of 91.51% was reached on a testing set composed of 110 banknote images, with an accuracy of 95.37%, 91.65%, 87.5% for the non-tilted images, noisy non-tilted images, and tilted images, respectively. The average time taken for each banknote to be recognised was about 3s.

There are other kinds of classifiers occasionally appearing in banknote recognition-related literature, for example, SVM, K-nearest neighbours (KNN) and K-means. A modular approach was presented for the recognition of Indian paper currency (Kamal et al., 2015). Four distinct and unique features, i.e. central numeral, Ashoka emblem, colour band and identification mark, were extracted separately, and SVM was employed as the classifier. The proposed approach was evaluated on over 300 Indian banknotes of different denominations with various physical conditions. It yielded the accuracy of 97.02% for central numeral detection, a 95.11% true positive rate and a 0.09765% false positive rate for emblem recognition, and the accuracy of 100% for the recognition of colour matching in CIE LAB colour space and identification marks. Ali and Manzoor (2013) designed an application for Pakistani paper currency recognition under the programming tool of MATLAB. The extracted banknote features included Euler number, area, height, width and aspect ratio. A KNN classifier was adopted to

contribute to an average accuracy of 98.57%. A feasible solution was put forward for the selection of distinct regions for banknote recognition (Pham et al., 2016). With images captured by a one-dimensional visible light line sensor, similarity mapping was responsible for obtaining the distinct areas. PCA was applied to discriminant features for dimensionality reduction and the classifier based on the K-means algorithm made a final decision on the type of the paper currency. The prototype was equipped with an average recognition speed of 568 images/s, and a 100% recognition rate on the banknotes in the Angolan Kwanza and South African rand databases, a 99.994% recognition rate on the banknotes in United States dollar database, and a 99.675% recognition rate on the banknotes in the Malawian kwacha database.

2.2 Digital image processing

Pattern recognition has been applied to many scientific and technical disciplines, including computer vision. There is a wealth of literature about the applications using a combination of pattern recognition and computer vision. The combined use correspondingly prompts an increase in the popularity of one of the most active research areas, named digital image processing, which grew from electrical engineering as an extension of the signal processing branch (Umbaugh, 2005). Digital image processing deals with arrays of numbers obtained by spatially sampling points of a physical image, mainly concerning the extraction of data, measurements or information from an image by means of semiautomatic or automatic methods (Chakraborty, Nalawade, Manjre, Sarawgi, & Chaudhari, 2016; Pratt, 2013). It diverges from conventional pattern recognition. Normally, image processing is the procedure prior to recognition, providing a sophisticated representation of the key information lying in the image, so as to smooth the way for classifying the image to a certain number of categories in the recognition stage.

2.2.1 Feature engineering

Despite today's advanced computer technologies, discovering knowledge from data is

still not a simple task because of the complicated characteristics of the data generated by a computer. In an attempt to find the best performance of the predictive model, not only does the best algorithm need to be selected, but also the information should be derived from the original dataset as much as possible. The question that arises is how to derive useful information from the dataset. The answer has been found in feature engineering whose objective is to optimise data for facilitating the next phase of analysis.

Features can be any extractable measurement used, and they can be numerical, symbolic, or both (Schalkoff, 1992). For instance, the length, width and height (measured in metric units) of an object are examples of numerical features, whereas the shape of an object could be an example of the symbolic feature. They may be represented by continuous, discrete, or discrete-binary variables signifying the presence or absence of a particular feature. Feature engineering is a process of turning raw data into a set of useful features that describe the data accurately. In feature engineering, feature construction, feature extraction and feature selection are the commonly adopted techniques for data processing.

Feature construction refers to the process of discovering missing information on the relationship between features as well as augmenting the space of features by creating or inferring additional features (Guyon & Elisseeff, 2006; Wnek & Michalski, as cited in Liu & Motoda, 1998). It usually expands the feature space. In comparison, feature extraction intends to derive informative features from an initial set of measured data by means of functional mapping (Wyse, Dubes, & Jain, as cited in Liu & Motoda, 1998). It often reduces the dimensionality of feature space. For instance, supposing the original dataset contains the features A_1, A_2, \dots, A_n . After applying feature construction to the dataset, m new features $A_{n+1}, A_{n+2}, \dots, A_{n+m}$ are created and added to the original feature set as a whole for further analysis, whereas after feature extraction, features B_1, B_2, \dots, B_m ($m < n$) are obtained via $B_i = F_i(A_1, A_2, \dots, A_n)$ where F_i denotes the mapping function.

On one hand, feature selection differs from feature construction and extraction in that it will not give rise to new features. Feature selection is defined as the process of selecting a subset of features for use from the original ones when the raw data contains many features that are either irrelevant or redundant (Cios, Swiniarski, Pedrycz, & Kurgan, 2007). Even though the primary motivation of feature selection is to help select relevant features, it is also beneficial to other aspects, such as an increase in processing speed and decrease in storage requirements by general data reduction, performance improvement in predictive accuracy by feature set reduction, and data understanding from gaining knowledge about the process of generating or visualising the data (Guyon & Elisseeff, 2006; Liu & Motoda, 2007).

On the other hand, feature construction, feature extraction and feature selection are not entirely independent issues. Perceiving features as the language used for representation, when the language is not enough to express the question, feature construction can enrich the language by creating compound features, while feature selection assists in simplifying the language in the situation where the language comprises more features than needed (Liu & Motoda, 1998). Plus, when some of the constructed features are of little use, feature selection can then remove those features. It is also very normal to combine feature selection and feature extraction in pattern recognition. Feature extraction usually demands a great deal of computational effort, and the extracted features resulting from applying a feature operator or algorithm to the input data may contain noises or errors (Bezdek, 2013). Consequently, several requirements are supposed to be taken into consideration at the very beginning of a pattern system design. For example, selecting and extracting features that are computationally feasible and are likely to contribute to the accuracy rating of a pattern recognition system, and reducing the problematic data to a reasonable amount of information in the procedures of feature selection and extraction, without the loss of valuable information (Schalkoff, 1992). Ultimately, the combined use of feature construction, feature construction and feature selection depends on the target – for a simpler description targeted at maintaining the topological structure of the data or for better classification aimed at boosting the

predictive ability (Liu & Motoda, 1998).

2.2.2 Feature extraction methods

Colour and texture are regarded as the prominent features of paper currency, and they often appear in the literature on banknote recognition. This section introduces the approaches to colour and texture feature extraction.

- **Colour feature extraction**

Colour feature can be used as part of a composite feature or even as a standalone feature for banknote recognition, where the banknotes apply a dominant colour for distinction (García-Lamont et al., 2012). The colour information of an image includes several specific aspects, such as the proportion of each colour and the spatial location of the colours. Among them, each of the aspects can be considered as a standalone colour feature. A colour histogram is a feasible feature extraction method since it can extract the statistics of each of the colours in an image, regardless of the spatial locations of the colours (Bharkad, 2013).

A colour histogram is a graphical representation of the distribution of the composition of colours in an image (Lamsal & Shakya, 2015). It can be built for a variety of colour spaces, though most frequently it is used for three-dimensional colour spaces like RGB or HSV (Lamsal & Shakya, 2015). Taken RGB colour space as an example, *red*, *green* and *blue* are the three channels in the space, so *red* histogram, *green* histogram and *blue* histogram can be constructed separately based on each channel. However, since the number of each channel is finite, a single variable histogram integrating the three histograms is therefore normally used for convenience purposes (Smith & Chang, 1996). Accordingly, in the single-variable histogram, the values of the horizontal axis stand for different colours in an image and the corresponding value of the vertical axis represents the number of pixels of the particular colour in the image. The number of bars in the single variable histogram is consistent with the number of bins determined by how many small intervals the RGB colour space is divided into.

Among various colour spaces, HSV is frequently found in research involving digital image processing, as it is closer to the human conceptual understanding of colours and is able to separate chromatic and achromatic components (Bharkad, 2013). HSV stands for *hue*, *saturation* and *value*, and an HSV colour model utilises these three attributes to differentiate colours (Pathrabe & Karmore, 2011). The *hue*, also called the name of the colour, is the core attribute of a colour, such as red, green, yellow, and so forth, and its value ranges from 0 to 360° corresponding to different basic colours. The *saturation* refers to the purity of a colour with the value from zero to one; the higher the *saturation* value is, the purer the colour is. The *value* means the brightness of a colour, ranging from the value of zero indicating the black to the value of one indicating the white (Zeggeye & Assabie, 2016).

- **Texture feature extraction**

Local binary pattern (LBP) is a type of visual descriptor used for classification in computer vision, and in particular it is a classic tool for texture description. Ojala, Pietikainen and Harwood (1994) formally introduced the original LBP operator in 1994, borrowing from a model of texture analysis in which a texture image could be characterised by its texture spectrum. The foundation of LBP is that an image can be perceived to comprise micro-patterns. From that point, LBP is the first-order circular derivative of patterns produced through concatenating the binary gradient directions, while a histogram of the micro-patterns displays information on the distribution of edges and some other local features in an image (Nanni, Lumini, & Brahnam, 2012).

Specifically, the original LBP operator uses the intensity value of a centre pixel as a threshold to convert each of the eight neighbouring pixels to a binary code ‘0’ or ‘1’. The eight codes will then form an ordered pattern based on the positions relative to the centre pixel. Mathematically, the LBP code for the pixel P can be defined as Equation 2.1 and Equation 2.2.

$$LBP(P) = \sum_{i=0}^7 2^i s(g_i - g_p) \quad (2.1)$$

$$s(t) = \begin{cases} 1, & t \geq 0 \\ 0, & t < 0 \end{cases} \quad (2.2)$$

where g_p is the intensity value of the centre pixel P , g_i is the intensity value of the i -th pixel in clockwise order at the eight neighbours of the pixel P , and $s(\cdot)$ represents the threshold function. Figure 2.1 illustrates an instance of this process in detail.

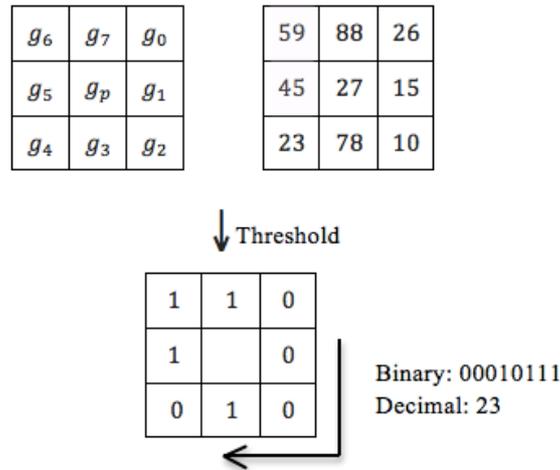


Figure 2.1 An example of LBP operator

One predominant feature of the LBP algorithm is that it is a grey-scale invariant texture primitive statistic because the operator calculates the relative difference, which can be inferred from Equation 2.1. In other words, it is robust to grey-scale changes as the operator itself is invariant against monotonic transformations of the intensity (García-Lamont et al., 2012). In addition, since the LBP operator can be carried out with several operations in a small neighbourhood and a lookup table, it is also famous for its computational simplicity so as to be utilised for real-time analysis (García-Lamont et al., 2012; Nanni et al., 2012).

However, the original LBP operator only covers the texture in every small area with a fixed radius, which is apparently not able to compute the textures on various sizes and frequencies. With regard to the textures at different scales, Pietikäinen, Ojala and Xu (2000) modified the original LBP operator, expanding a 3×3 neighbourhood to a neighbourhood of any size, known as a circular local binary pattern. It applies a circular

neighbourhood and bilinear interpolation at non-integer pixel coordinates to allow any radius and number of pixels in the neighbourhood. Later on, they defined the rotation invariant version of LBP, because the basic circular LBP operator is not rotation invariant, which is undesirable (Pietikäinen et al., 2000). Figure 2.2 displays the 36 rotation-invariant binary patterns appearing in the eight pixels circularly symmetric neighbour set, with the first row containing the nine uniform patterns, where black and white circles correspond to bit values of 0 and 1 in the 8-bit LBP code, respectively.

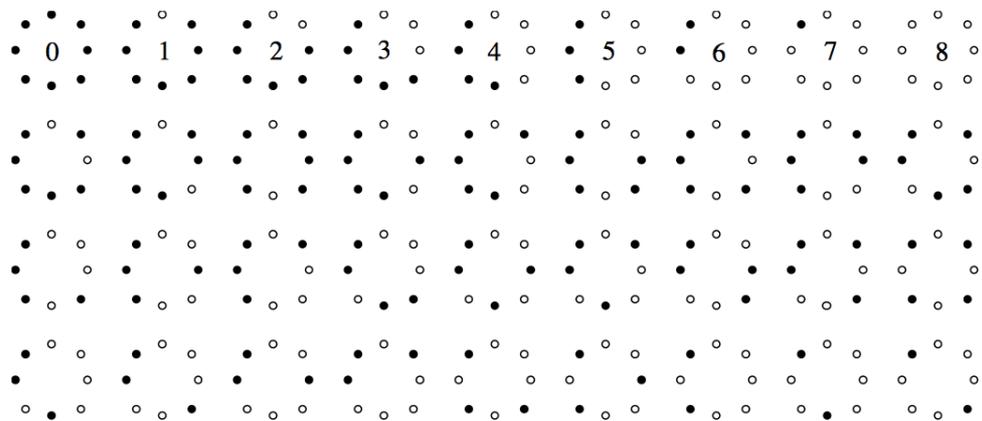


Figure 2.2 Thirty-six rotation-invariant binary patterns

Another functional extension to the original LBP operator is the so-called uniform patterns, which are mainly used to reduce the dimensionality of a feature vector and can also be used along with rotation-invariant LBP (Ojala, Pietikainen, & Maenpaa, 2002). This extension was inspired by the fact that certain binary patterns appear more frequently in textural images than others. The term “uniform” denotes the uniform appearance of the local binary pattern. In other words, there are a limited number of discontinuities or transitions in the circular presentation of the pattern. A local binary pattern is perceived as uniform if it contains no more than two bitwise transitions from zero to one or one to zero. For instance, the patterns 11111111 (0 transition), 00000001 (2 transitions) and 00000110 (2 transitions) are uniform, whereas the patterns 10000110 (4 transitions) and 10101010 (8 transitions) are not uniform. All the uniform patterns based on a neighbour set of eight members on a circle with a radius of one pixel are listed in Figure 2.3. Ojala et al. (2002) assumed that uniform LBP is of great account in

the texture of an image, which was proved by experimental results that uniform patterns account for nearly 90% of all patterns based on a neighbour set of eight members on a circle with a radius of one pixel, and for approximately 70% based on a neighbour set of 16 members on a circle with a radius of two pixels.

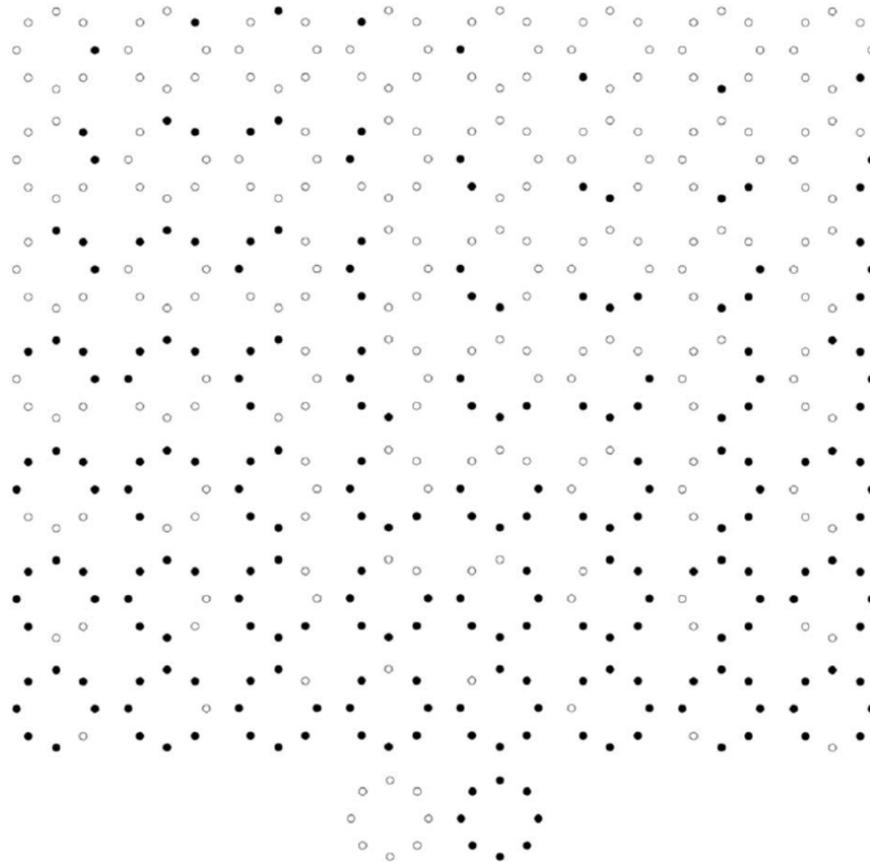


Figure 2.3 Uniform LBP based on a neighbour set of eight members on a circle with a radius of one pixel

2.2.3 Dimensionality reduction methods

This section focuses on demonstrating how PCA and LDA, the two methods for reducing the dimensionality of a feature vector, work by deriving their foundations from the mathematics. The advantages and disadvantages of each method will be discussed afterwards. By addressing both aspects, a better understanding of different methods will be gained.

- **PCA**

PCA is a useful statistical method in modern data analysis and is widely used in scientific disciplines. It is also known as the Hotelling transform or the Karhunen-Loève transform (Cios et al., 2007; Sonka, Hlavac, & Boyle, 2014). As observations are often described by intercorrelated variables with a little noise, PCA is often used to identify the most meaningful basis to re-express a given dataset so that hidden structure in the dataset is discovered and the noise in the dataset is filtered (Kurita, 2014). Specifically, PCA computes principal components which are a set of new orthogonal variables obtained through linear combinations of the original variables. The values of these new variables for the observations are referred to as factor scores, which can be illuminated as the projections of the observations onto the principal components (Kurita, 2014). The first principal component takes up most of the variability of the data, and each subsequent component takes up the remaining variability as much as possible (Sonka et al., 2014).

Digital images are typically represented in a manner of the matrix in computer vision and are expressed by vectors when digital image processing is involved. For example, a square image with N by N will be expressed in a $N \times N$ matrix. The n -dimensional vector can then be transformed to a one-dimensional vector by placing the rows of pixels in the image one after the other in sequence, as shown in Equation 2.3:

$$imageVec = \{x_1, x_2, \dots, x_{N^2}\} \quad (2.3)$$

where the first N elements starting from x_1 to x_N refer to the first row of the image, followed by the next n elements referring to the second row of the image, and so on. So if there are 20 sample images and each of them is represented by an image vector, the large image matrix for expressing those images will be like Equation 2.4:

$$imagesMatrix = \begin{pmatrix} imageVec1 \\ imageVec2 \\ \dots \\ \dots \\ imageVec20 \end{pmatrix} \quad (2.4)$$

In order to avoid the difficulty of computing high dimensional vectors, image space of a high dimensionality needs to be reduced to a subspace of a smaller dimensionality while retaining as much appropriate information from the original images as possible (Ye, Shi, & Shi, 2009). PCA is one of the right approaches to address the issue. In statistics, PCA is a practical method of simplifying a multi-dimensional dataset to be of a lower dimensionality for analysis or visualisation. Through optimal linear transformation, an observed space is divided into orthogonal subspaces with the largest variance (Cios et al., 2007).

Several equations regarding the use of PCA for image classification have been introduced by researchers, such as Equation 2.5, Equation 2.6, and Equation 2.7. Suppose a scenario where a set of N images are distributed in an n -dimensional image space as $\{x_1, x_2, \dots, x_N\}$, and also each image belongs to one of c classes $\{X_1, X_2, \dots, X_c\}$. The n -dimensional image space is then mapped onto an m -dimensional feature space, where normally $m \ll n$. The new feature vector $y_i \in R^m$ will be defined by the linear transformation as shown in Equation 2.5:

$$y_i = W^T x_i \quad i = 1, 2, \dots, N \quad (2.5)$$

where $W \in R^{n \times m}$ is a matrix with orthonormal columns.

If the total scatter matrix S_T is defined as Equation 2.6:

$$S_T = \sum_{i=1}^N (x_i - \mu)(x_i - \mu)^T \quad (2.6)$$

where N is the total number of training images and x_i is the i -th feature vector of samples, and $\mu \in R^n$ represents the mean feature vector of all samples in the training set. Then after the linear transformation W^T , the scatter of the transformed feature vectors $\{y_1, y_2, \dots, y_N\}$ is $W^T S_T W$.

Finally, the PCA method tends to find a projection matrix W_{opt} to maximise the determinant of the total scatter matrix of the projected samples, based on Equation 2.7:

$$W_{opt} = \arg \max_W |W^T S_T W| = [w_1, w_2, \dots, w_m] \quad (2.7)$$

where the project matrix $\{w_i | i = 1, 2, \dots, m\}$ is comprised of the set of n -dimensional eigenvectors of S_T corresponding to m largest eigenvalues.

However, when PCA is used for image compression, the transformation analysis would be slightly different from that of the previous scenario (Smith, 2002). For instance, there are 20 images as an original dataset, and each has N^2 pixels. N^2 20-dimensional vectors will be formed, with each vector comprising all the intensity values from the same pixel from each image. So if 20 eigenvectors are transformed to 16 eigenvectors through PCA, then the final dataset will have 16 dimensions, which has saved 20% of the space. In a word, for image compression, a vector represents each pixel, with each item in the vector being from a different image, whereas for pattern recognition, a vector represents an image, with each item in the vector being a different pixel from that image.

Despite the distinct advantage of dimensionality reduction, applying PCA to images has a few disadvantages. Relationships of a given pixel to pixels in adjacent rows are not taken into consideration as a result of rearranging pixels column by column to a one-dimensional vector (Sonka et al., 2014). Besides, the data type that PCA can process must be vectors rather than image matrix (Liu & Motoda, 2007). Another drawback is in the global nature of the representation; little change or error in the input images affects the whole eigen-representation (Sonka et al., 2014).

- **LDA**

LDA is another linear projection-based method to compress the information residing in a training dataset into a smaller dimensional space. It explores for those vectors in the underlying space that best discriminate among classes instead of those that best describe the data (Ye et al., 2009). With the mathematical explanation, given some independent features describing the dataset, LDA makes the linear transformation on them, which produces the largest mean differences between the desired classes (Rahman, Banik, & Naha, 2014). The formulation relevant to LDA will be further clarified in the manner of equations later in this section, including Equation 2.8, Equation 2.9, Equation 2.10, Equation 2.11 and Equation 2.12.

Since LDA adjusts similarity by increasing inter-class dissimilarity as well as intra-class similarity, now defining the between-class scatter matrix as S_B , the within-class scatter matrix as S_W , and the total scatter matrix as S_T :

$$S_B = \sum_{i=1}^C N_i (\mu_i - \mu) (\mu_i - \mu)^T \quad (2.8)$$

$$S_W = \sum_{i=1}^C \sum_{x_k \in x_i} (x_k - \mu_i) (x_k - \mu_i)^T \quad (2.9)$$

$$S_T = S_B + S_W \quad (2.10)$$

where μ_i is the mean vector of class X_i , and N_i is the number of samples in class X_i . If S_W is non-singular, the optimal projection W_{opt} is chosen as the matrix with orthonormal columns that maximise the ratio of the determinant of the between-class scatter matrix of the projected samples to the determinant of the within-class scatter matrix of the projected samples.

$$W_{opt} = \arg \max_W \frac{|W^T S_B W|}{|W^T S_W W|} = [w_1, w_2, w_3, \dots, w_m] \quad (2.11)$$

where the projection matrix $\{w_i \mid i = 1, 2, 3, \dots, m\}$ is the set of generalised eigenvectors of S_B and S_W corresponding to the m largest generalised eigenvalues $\{\lambda_i \mid i = 1, 2, 3, \dots, m\}$, as shown in Equation 2.12.

$$S_B w_i = \lambda_i S_W w_i \quad i=1, 2, 3, \dots, m \quad (2.12)$$

It is worth noting that there are no more than $c-1$ nonzero generalised eigenvalues. Therefore, the maximum value of m is $c-1$, where c represents the number of classes in the dataset.

Compared with PCA, LDA is less sensitive to large variations in lighting that include not only intensity but also the number and direction of light sources (Rahman et al., 2014). The goal of PCA is to yield project directions that maximise the total scatter across all classes. To realise the target, PCA retains unwanted variations on account of illumination, as the variations between the images owing to lighting and viewing direction are most likely to be much greater than image variations due to changes in content identity (Duda et al., 2012). Thus, although PCA achieves larger total scatter, LDA can be optimal from the discrimination point of view by achieving greater between-class scatter. Figure 2.4 illustrates the kernel idea of PCA and LDA regarding dimensionality reduction, taking a two-dimensional dataset as an example. There are samples of two categories in the dataset, namely, the blue ones and the orange ones, and the blue line is the newly created coordinate axis.

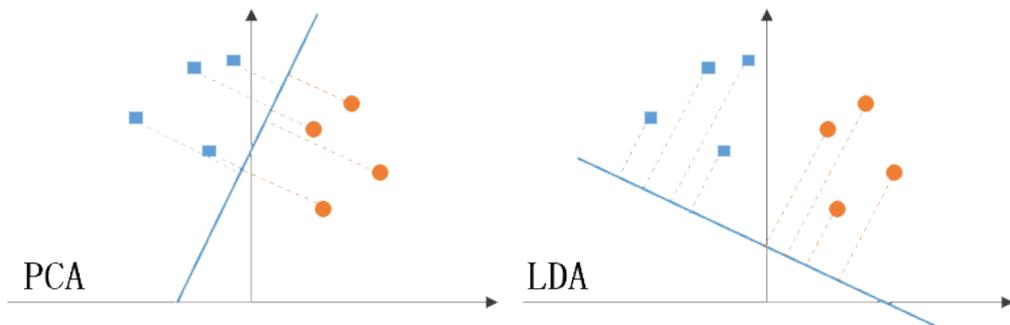


Figure 2.4 A comparison of mapping results between PCA and LDA

Apparently, as shown in Figure 2.4, PCA maps the entire set of data to the coordinate axis that most conveniently represents the set of data. The mapping does not make use

of any known category information within the dataset. Accordingly, even though PCA facilitates the representation of data by reducing dimensionality as well as retaining the original information as much as possible, it might cause difficulty in classification. By contrast, LDA makes the best of the category information that is known beforehand in the dataset. It maps the dataset to another axis to make the data more easily distinguishable. However, LDA still has several limitations (Luo, Ding, & Huang, 2011; Wang & Tang, 2004). Firstly, as a linear analysis algorithm, image matrices must be transformed to vectors before LDA can deal with them. Secondly, LDA may over fit to training data thereby generalising poorly to new testing data. Thirdly, LDA is no longer a good option for data classification when the categories of the dataset heavily rely on variance instead of mean value. Fourthly, LDA can generate $c-1$ -dimensional subspace at most, while c stands for the number of the categories of the dataset. Fifthly, LDA is not suitable for dimensionality reduction on the dataset with non-Gaussian distribution. Sixthly, small sample size problem results in the singularity of the within-class scattering matrix in LDA; hence, failing to obtain the optimal projection direction.

2.3 Classification Algorithms

Classifier design has always been of great interest to the pattern recognition community and there has been significant development in the past decades (Li et al., 2014). Since the primary task of classification is, in essence, to find the model that generated the patterns, classification techniques are unique to the type of candidate models (Duda et al., 2012). In digital image analysis, the classification procedure is to analyse the numerical properties of various image features and then categorise the data into one of the classes based on homogeneous characteristics.

2.3.1 Minimum distance classifier

The minimum distance-based classification algorithm is used to classify unknown image data to one of the target classes which has the greatest similarity, measuring the distances between the input feature vector and all the mean vectors of the target classes

(Althafiri, Sarfraz, & Alfarras, 2012). The distance is defined as an index of similarity, so the smaller the distance, the larger the similarity between the two classes of patterns.

The MDC is easy to implement and computationally simple, and it is also capable of yielding accuracy comparable to other more computationally intensive algorithms like maximum likelihood (Althafiri et al., 2012). In practice, the MDC has a high performance on the condition that the distance between means is large in comparison with the randomness of each class with regard to its mean, which could be the limitation of this classifier (Choi, Lee, & Yoon, 2006; Duda et al., 2012). Three kinds of distance measures are usually adopted in the procedure of minimum distance classification, namely, Euclidean distance, Manhattan distance and Chebyshev distance.

2.3.2 Back-propagation neural network classifier

The first computational model for ANN was created by McCulloch and Pitts in 1943, based on mathematics and algorithms, and named threshold logic. It paved the way for an ANN study splitting into two approaches, with one approach focusing on biological processes in human brain and the other concentrating on the application of ANN to artificial intelligence. In ANN models, neural networks are built from many neurones that are grouped in layers, and neurones can be connected in different manners depending on the specific algorithm used (Jankowski & Grabczewski, 2006). Each neurone can produce a series of real-valued activations, whereby input neurones are activated via sensors recognising the environment, and other neurones are activated via weighted links from formerly active neurones (Schmidhuber, 2015). ANN is an adequate approach for pattern recognition by virtue of self-organisation, parallel processing and generalisation (Cristea, 2009; Sethi & Jain, 2014). However, processing overhead could be the main disadvantage of adopting ANN that needs to be taken into consideration (Vishnu & Omman, 2014). Moreover, a significant number of training samples are required for neural network methods to find the proper classification function, and the necessary characteristic quantity is excessive. All these factors indicate that complex calculation appears to be inevitable, and overfitting is likely to

come about (He, Zhang, Liang, Jin, & Li, 2015).

It is worth mentioning that a key breakthrough in ANN was the back-propagation algorithm that initially appeared in 1974 (Werbos, as cited in Hinton, 2002). Nevertheless, its significance did not receive much attention until it was proved to be a noticeably faster learning algorithm than other earlier algorithms for working with multilayer networks (Rumelhart, Hinton, & Williams, 1985). Learning occurs in the perceptron by modifying connection weights after each piece of data is processed, based on the amount of error in the output compared with the expected result. Mean squared error (MSE) is the common algorithm to measure the error, which is defined as follows:

$$E = \frac{1}{N} \sum_{i=1}^N (h(x_i) - y_i)^2 \quad (2.13)$$

where y_i is the expected output given x_i as an input and $h(x_i)$ is the actual output of the neural network.

The essential processes brought by the BPNN are the involvement of a differentiable transfer function at each node of the network and the adoption of error back-propagation to modify the internal network weights after each training session. The way the back-propagation method calculates the gradient of an objective function regarding the weights of a multilayer stack of modules is an application of the chain rule for derivatives (LeCun, Bengio, & Hinton, 2015). After the MSE is calculated, the gradient of the objective concerning the input of a module is computed by working backwards from the gradient regarding the input of the following module. By means of this mode of propagation throughout all modules, starting from the output to the input, step by step, once the gradients are calculated, it is simple to work out the gradients for the weights of each module.

Figure 2.5 shows a standard BPNN composed of three layers of nodes. Each node in the input and hidden layer is connected to each node in the next layer, and there is no

connection between the nodes in the same layer. Each connection between the nodes has a weighting factor on it. The information flows only one way as per the direction arrow in Figure 2.5. In the beginning, all weighting factors are assigned randomly. At the end of each training session, the total errors of the outputs are computed, and this information is then transmitted back to the network using a back-propagation approach so as to update weighting factors. By means of repeating this process, an effective neural network can be obtained. The input layer consists of two nodes, excluding the bias node. There is no calculation in the input layer; x_1 and x_2 will be passed onto the hidden layer. The hidden layer has two nodes, excluding the bias node. The outputs of the nodes in the hidden layer depend on the value of the input nodes and the weighting factors attached to the connections. For example, the value of the highlighted node V is determined by x_1 , x_2 , w_0 , w_1 and w_2 . The output layer is composed of two nodes and has no bias node. The number of output nodes should equal the number of the categories of experimental samples, corresponding to the output results.

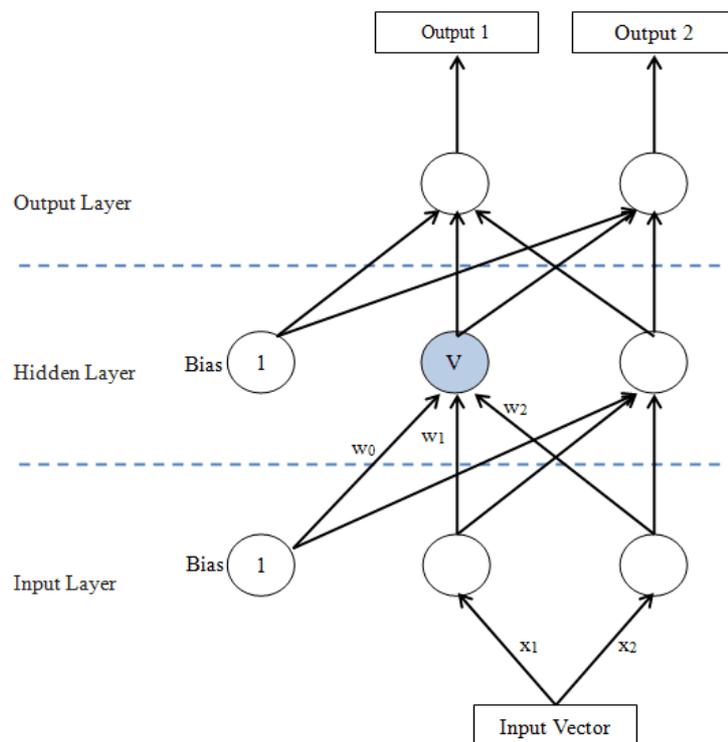


Figure 2.5 A schematic diagram of BPNN

Chapter 3 Methodology

This chapter will present the methods involved in this thesis. In Section 3.1, the related work will be introduced with an emphasis on the study regarding Series 6 New Zealand banknote recognition that has provided direct inspiration for this thesis. In the next section, research questions will be raised with the hypotheses. Subsequently, Section 3.3 is about research design, including five subsections. In Section 3.3.1, the experimental platform will be described, followed by Section 3.3.2 giving an account of the experimental samples. In Section 3.3.3, various combined methods for classification applied in this experimentation will be listed. In Section 3.3.4, the framework of the system will be illustrated by flowcharts indicating the workflow of the training procedure and testing procedure. Finally, the KPI for evaluating the performance of the combined methods will be explained in Section 3.3.5.

3.1 Related work

Visual features such as colour, shape and texture can be extracted by deploying digital image-processing techniques to characterise images. Each of the features is represented by one or more feature descriptors for classification. Colour and texture are the distinct features of a banknote that humans rely on to perceive it; such features are supposed to be applied in machine learning for developing an efficient banknote recognition system. Many studies on banknote recognition have been conducted using digital image processing to extract the colour or texture feature of banknotes.

RGB is the colour space that is commonly selected to present the colour feature of banknotes for banknote recognition, since *red*, *green* and *blue* have unique distributions according to denomination of the banknote. Chae, Kim and Pan (2009) calculated the mean value of each of the three channels, and then the difference between *red* and *green*, *red* and *blue*, *green* and *blue* together served as the colour feature. Solymár, Stubendek, Radványi and Karacs (2011) employed the statistical properties of the colour of banknote in RGB and YcBcR colour space, including the mean, standard deviation,

median values and ratios of standard deviations of RGB values. García-Lamont et al. (2012) used the RGB model to express the colour feature of the banknote, which is based on a Cartesian coordinate system where the RGB colour space is represented by a cube. Correspondingly, the colours were described by the points located inside the cube, defined by vectors that extend from the origin. HSV colour space is also found in banknote recognition research to extract the colour feature of banknotes. An HSV colour model draws on the three attributes - *hue*, *saturation*, and *value*, to differentiate colours (Pathrabe & Karmore, 2011). Aasma and Asma (2016) regarded the mean value of the *hue*, the *saturation*, and the *value* as the predominant colour feature of the banknote. Murthy, Kurumathur and Reddy (2016), and Zeggeye and Assabie (2016) extracted the mean value of the *hue* as the colour feature of the paper currency. Plus, there are still other studies that only make use of the grey-level information of banknotes for colour feature extraction, instead of any colour space. In the research undertaken by Hassanpour, Yaseri, and Ardeshiri (2007), a grey-scale histogram with 52 bins was computed to find the pletitude of different colours in the banknote image. In Chambers' study (2012), a grey-scale histogram with 256 bins were used to characterise the colour feature of the banknote.

A number of methods have been employed to analyse the texture of banknotes for banknote recognition, and grey-level co-occurrence matrix (GLCM) is one of them. It computes the statistical features based on the intensity of a grey-level image, indicating how often the intensity value i occurs horizontally adjacent to the intensity value j (Devi et al., 2016; Hlaing & Gopalakrishnan, 2016). Chambers (2012) selected the *contrast*, *correlation*, *energy* and *homogeneity* from the GLCM features to be part of the texture feature of banknotes, while Hlaing and Gopalakrishnan (2016) extracted the GLCM features including *energy*, *homogeneity*, *correlation*, *contrast* and *entropy* as the texture feature to differentiate various banknotes. Apart from GLCM, the LBP operator also examines the spatial relationship of pixels for texture analysis where the neighbouring pixels are converted to binary codes by using the grey-level value of the centre pixel as threshold (Devi et al., 2016). Usually the LBP feature works in conjunction with a

histogram to describe the texture of banknotes. Accordingly, the LBP histogram was adopted as the feature descriptor for the texture feature of banknotes in the research undertaken by Guo, Zhao and Cai (2010) and the study conducted by García-Lamont et al. (2012). Moreover, Hidden Markov Model (HMM) has been put forward for texture-based feature extraction for banknote recognition, where the texture of banknotes was modelled as a random process. Specifically, the transition matrix was quantised into 10 grey levels and finally the main diagonal values of the matrix were exploited to distinguish different denominations (Hassanpour et al., 2007; Hassanpour & Farahabadi, 2009).

Despite there already being much research on banknote recognition, there is only one study regarding New Zealand banknotes, which is specifically about Series 6 New Zealand banknote recognition (Chambers, 2012). The Series 6 New Zealand banknotes are the most similar edition of notes to Series 7 New Zealand banknotes among all the paper currency worldwide, even though they have their unique characteristics, as shown in Figure 3.1 and Figure 3.2. Because of this, Chambers' (2012) study is the most significant reference for this research.



(a) 5 NZD



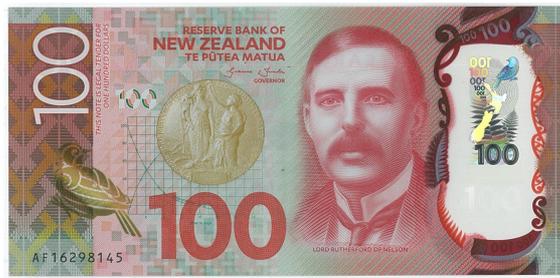
(b) 10 NZD



(c) 20 NZD



(d) 50 NZD



(e) 100 NZD

Figure 3.1 Series 7 New Zealand banknotes of all different denominations



(a) 5 NZD



(b) 10 NZD



(c) 20 NZD



(d) 50 NZD



(e) 100 NZD

Figure 3.2 Series 6 New Zealand banknotes of all different denominations

According to the literature, banknote recognition systems based on digital image processing could be divided into two groups. One group focuses on analysing the

banknote as a whole image, and the other group targets at the analysis of specific areas of interest, known as ROI, such as an investigation exclusively into the serial number zone of notes. In Chamber's (2012) study, the banknote image was analysed as a whole, using digital image processing and classification to recognise the denomination. Two dominant categories of the main features were used, namely, colour feature and texture feature. For extracting the colour feature, a grey-level histogram was generated using 256 bins to describe the frequency of dark to light colour. From the histogram, six shape descriptor metrics were obtained, including *kurtosis*, *central moment*, *mean*, *variance*, *standard deviation*, and *skew*. Five texture features were extracted with the *entropy* level and the four features from the GLCM, namely, *correlation*, *contrast*, *energy* and *homogeneity*. The extracted colour feature and texture feature were concatenated as a composite feature, which was directly input as a scalar to a feed-forward neural network classifier trained by Bayesian regulation back-propagation. Finally, the banknote in question was classified into its respective denomination, along with a measurement of the similarity between the existing samples and the suspect banknote.

3.2 Research questions and hypotheses

This thesis is primarily concerned with an investigation into how to utilise digital image-processing techniques to facilitate banknote real-time recognition tasks. Feature extraction plays a decisive role in classification. The features derived from data can directly affect the performance of the classification model (Kwak & Choi, 2002). Accordingly, the assumption on the subtle relationship between feature extraction and classification was mentioned by Duda et al. (2012): if a perfect feature extractor produced a representation that would make the classifier of little account, and conversely, if an omnipotent classifier did not require a complicated feature extractor to help out. By reviewing the literature, it turns out that the ideal combination of feature extractor and classifier relies more on experimental study than theoretical speculation. Particularly, feature extraction is much more problematic and domain specific than classification, requiring knowledge of the domain, just as a great feature extractor for face recognition is probably applicable neither to the recognition of vehicle licence

plates nor to the recognition of hand gestures. Nonetheless, there is no doubt that the more representative the features are, the better the performance of the model would be, since the ultimate goal of feature engineering is to improve the performance of the model.

With great interest stimulated by Chambers' (2012) study on Series 6 New Zealand banknote recognition as well as the assumption of the subtle relationship between feature extraction and classification from Duda et al. (2012), this research attempts to answer the core question of **how to achieve real-time recognition of the new Series 7 New Zealand banknotes using digital image-processing techniques.**

Although the time taken for each note to be recognised is a factor to consider for real-time recognition, the recognition rate is still given the priority on the evaluation of recognition performance. Accordingly, three sub-questions about the recognition rate are derived from the core research question, which are regarding training set, features and classifiers, correspondingly. The hypotheses corresponding to the sub-questions are also formulated.

1) Scanner and camera are the most commonly used hardware to generate images. So in which way should the training images be produced to yield better recognition results?

Hypothesis 1.1: The banknote images captured by webcam are likely to have a lower resolution due to the low configuration of ordinary webcams in the current market. By contrast, the images produced by scanner are of high resolution, retaining much more original and detailed information on banknotes, which are therefore hypothesised to result in better recognition results.

Hypothesis 1.2: The training set made up of the banknote images captured by webcam is hypothesised to produce better recognition results, as those training images are generated in the way that simulates the testing environment to the maximum extent.

2) Colour and texture are the most discriminative features of Series 7 New Zealand banknotes. So which feature is able to yield better recognition results for Series 7 New Zealand banknote real-time recognition - colour feature, texture feature or the composite feature containing both colour and texture features?

Hypothesis 2.1: The colour feature, as the predominant feature of the banknote, is hypothesised to yield better recognition results, because the colour difference between the different denominations is dramatic.

Hypothesis 2.2: The banknote of each denomination has a unique pattern, which is easy to distinguish, such as portraits of different key characters, different background designs, graphics of the various colour-changing birds, and so forth. So the texture feature is hypothesised to be the optimum feature that is able to produce better recognition results.

Hypothesis 2.3: The combined use of the colour elements and texture elements is hypothesised to result in better recognition results than solely adopting either colour or texture feature, as each feature part in the composite feature is supposed to contribute to banknote real-time recognition.

3) The MDC and the BPNN are often employed as effective classifiers to deal with banknote recognition. So which one of them is able to yield better results for Series 7 New Zealand banknote real-time recognition?

Hypothesis 3.1: BPNN classifiers require a large number of samples to train a superior network, and have the defect of being liable to stick to a local optimum. Unlike BPNN, the MDC has the capability to work well with a small sample size. Additionally, the MDC has been widely employed in a single-currency recognition model and has the proven ability to produce an excellent recognition rate. So the MDC is hypothesised to result in better recognition results.

Hypothesis 3.2: Among the classifiers occurring in banknote recognition literature, BPNN is a popular classifier frequently employed in both single-currency recognition models and multi-currency recognition models. Thus, the BPNN classifier is hypothesised to have excellent ability in banknote classification as per denomination, and is therefore capable of generating better recognition results.

3.3 Research Design

Paper currency recognition is a significant branch of pattern recognition. Being a typical case of pattern recognition, digital image processing-based banknote recognition is made up of a few steps; they are image acquisition, pre-processing, feature extraction, dimensionality reduction, and classification decision. Image acquisition is the action of obtaining an image with the help of hardware-based sources like a digital camera or scanning equipment. Then the simple pre-processing procedure is implemented on the obtained images. Once the image is pre-processed, a succession of complex image-processing methods, including feature extraction and dimensionality reduction, are applied to the images for further analysis. Finally, the classifier analyses the numerical properties of various image features and organises data into categories.

3.3.1 Experimental platform

MATLAB is chosen to be the platform to explore the automatic real-time banknote recognition for Series 7 New Zealand banknotes, taking advantage of its matrix-based language to naturally express computational mathematics, and its built-in graphics that make it convenient to visualise and gain insights from data. The research will be undertaken on MATLAB student version R2010a, operated on the Microsoft Windows 7 64-bit operating system of a Lenovo Y430P laptop. The laptop embeds Intel(R) Core(TM) i7-4710MQ CPU @ 2.50GHZ processor, with 8GB RAM. It will use a Lenovo easy camera - the front-facing webcam of the laptop to demonstrate the real-time recognition.

3.3.2 Experimental samples

The experimental samples are composed of all denominations of New Zealand banknotes including 5 NZD, 10 NZD, 20 NZD, 50 NZD, and 100 NZD, which are collected randomly in the circulation market. Three hundred banknote samples were collected in total, with 60 samples for each denomination. Two hundred and fifty of them served as the training samples, with 50 training samples for each denomination, while the rest 50 samples worked as testing samples, with 10 testing samples for each denomination.

3.3.3 Combined methods

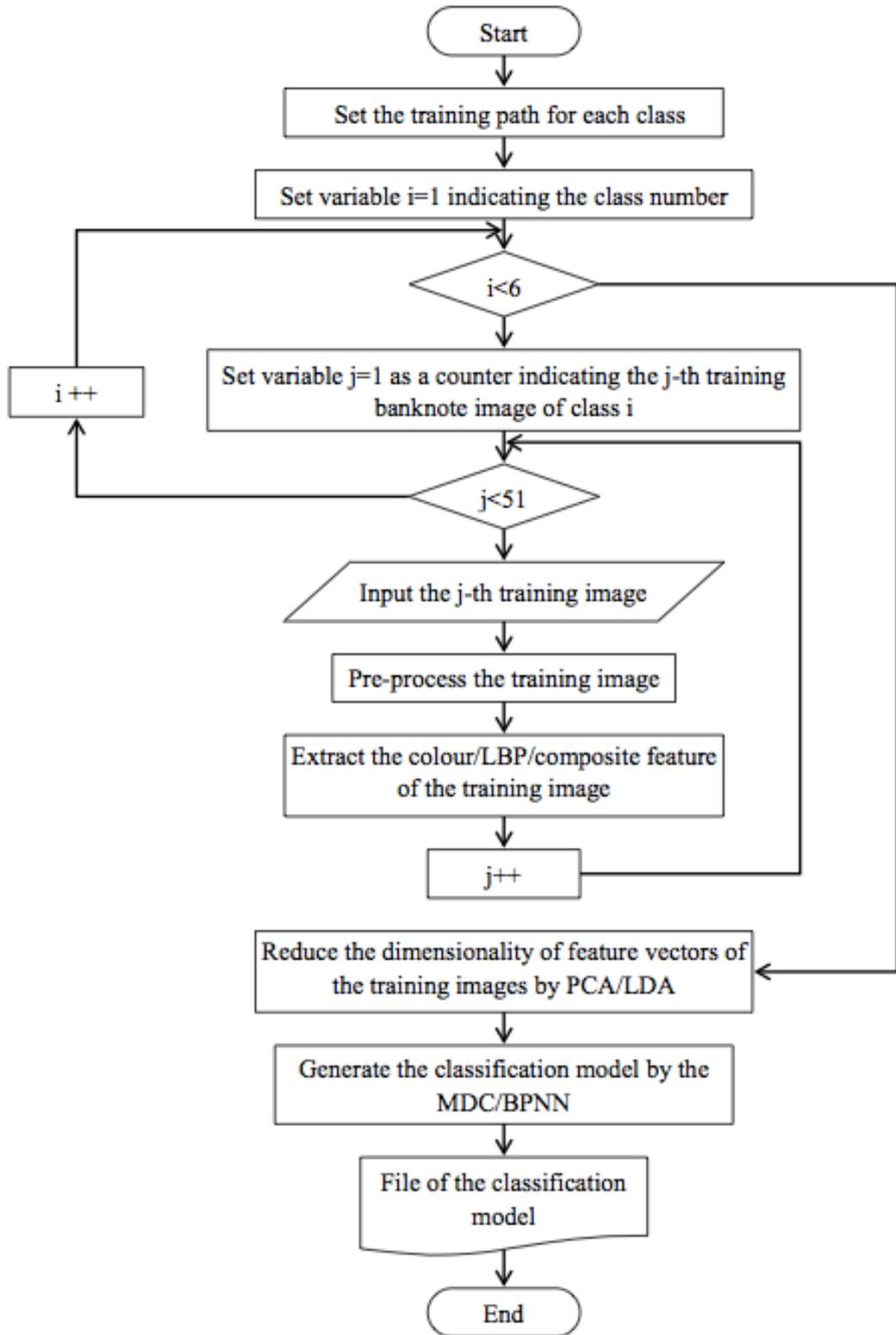
In this thesis, the colour histogram works as the feature for the colour of banknotes, and the LBP histogram is the feature for the texture of banknotes. Since this thesis focuses on real-time recognition, effective dimensionality reduction algorithms - PCA and LDA, are utilised in experimentation to build up computing speed, thereby adapting to real-time recognition. Euclidean distance-based MDC and BPNN classifiers are employed respectively to classify the testing banknote into the correct category. The combinations of extracted features and classifiers investigated in this project are listed in Table 3.1.

Table 3.1 Different combinations of extracted features and classifiers

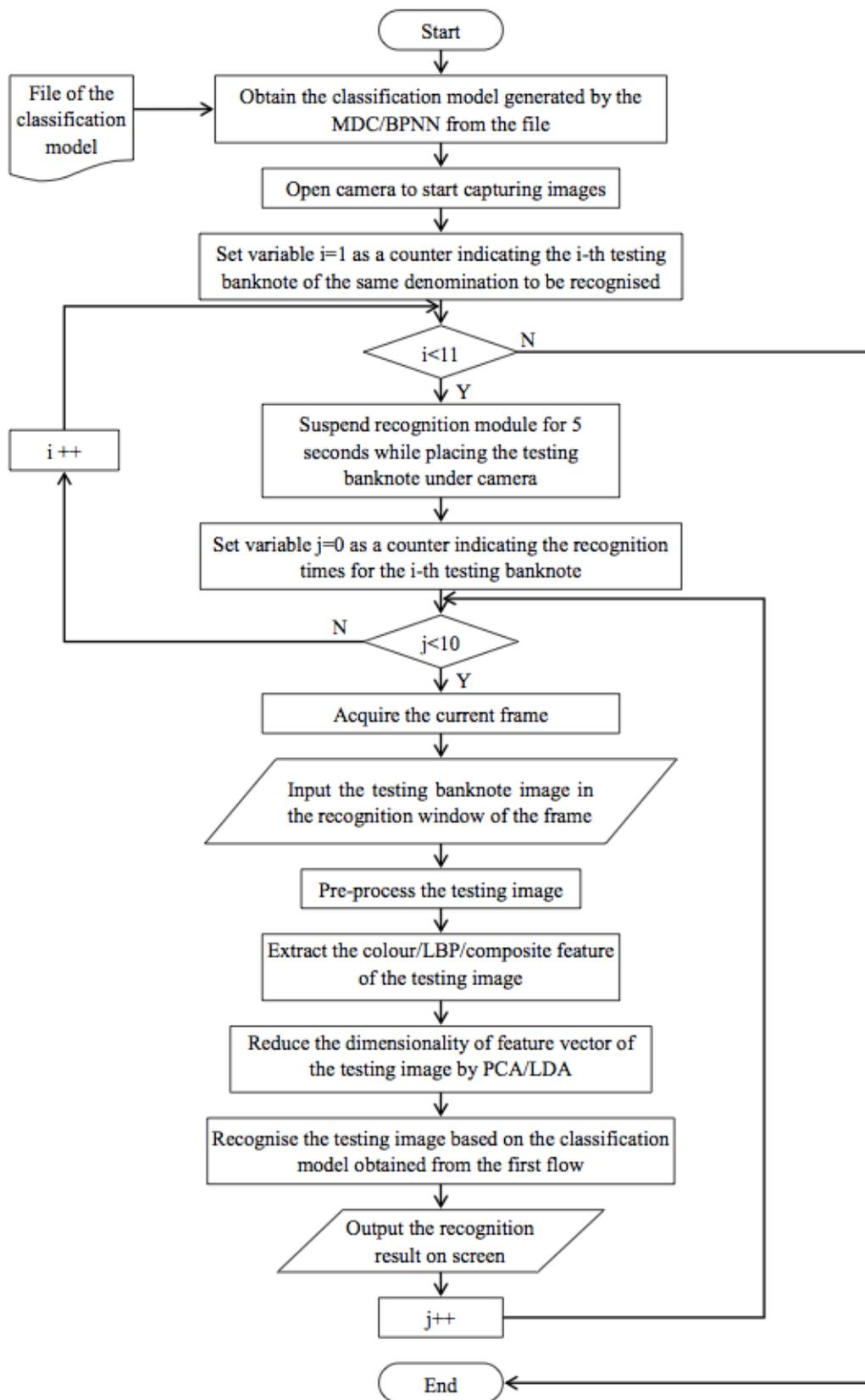
Combination No.	Features	Dimensionality reduction methods	Classifiers
1	Colour	PCA	MDC
2	LBP	PCA	MDC
3	Colour+LBP	PCA	MDC
4	Colour	LDA	MDC
5	LBP	LDA	MDC
6	Colour+LBP	LDA	MDC
7	Colour	PCA	BPNN
8	LBP	PCA	BPNN
9	Colour+LBP	PCA	BPNN
10	Colour	LDA	BPNN
11	LBP	LDA	BPNN
12	Colour+LBP	LDA	BPNN

3.3.4 Flowcharts

The flowcharts of the camera-based banknote real-time recognition are shown in Figure 3.3, with the training procedure shown in Figure 3.3 (a) and the testing procedure shown in Figure 3.3 (b).



(a) Flowchart of the training procedure



(b) Flowchart of the testing procedure

Figure 3.3 Flowcharts of the camera-based banknote real-time recognition

3.3.5 KPI

The F-measure, originally derived from the field information retrieval, is nowadays routinely exploited as a performance metric for multi-class classification systems (Dittimi, Hmood, & Suen, 2017; Ni'am, Faisal, & Arif, 2014; Sokolova & Lapalme, 2009). In classification, the precision of classification is defined as the number of true positives divided by the sum of true positives and false positives, which is expressed by Equation 3.1. The recall of classification is defined as the number of true positives divided by the total number of elements that belong to the positive classes, which is represented by Equation 3.2. The F-measure (F_1 score) is defined as the harmonic mean of precision and recall, which is shown by Equation 3.3.

$$Precision = \frac{TP}{TP+FP} \quad (3.1)$$

$$Recall = \frac{TP}{TP+FN} \quad (3.2)$$

$$F - measure = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (3.3)$$

where TP stands for true positives or hit, representing the number of items correctly labelled as belonging to the positive class; FP is false positive or false alarm, representing the number of items incorrectly labelled as belonging to the positive class; FN refers to false negatives or missing correct acceptance, representing the number of items that are not labelled as belonging to the positive class but should have been. Precision and recall reach their best value at 100% and worst at 0%, while F-measure reach its best value at one and worst at zero. For example, a 100% precision for classifying \$5 indicates that every item labelled as belonging to the category \$5 does indeed belong to \$5, and a 100% recall indicates that every item from the category \$5 was labelled as belonging to \$5.

A large number of banknote recognition-related studies employ the accuracy to measure the recognition rate of the proposed method, with the accuracy equal to correct

recognition times divided by the total of recognition times. However, the accuracy is not suitable for reflecting the performance of the various combined methods to be investigated in this research project. A typical example is the situation in which all testing banknotes of the different denominations are recognised as one particular denomination, for example, \$5. In that situation, the accuracy of recognising \$5 is 100%, which incorrectly reflects the performance of the combined method for recognising each denomination, whereas the F-measure of recognising \$5 is approximately 0.3333, resulting from the precision being 20% and the recall being 100%. Clearly, F-measure is a more comprehensive KPI for this research. Thus, precision and recall, the two essential components of F-measure, will be calculated based on the experimental results; eventually, F-measure (F_1 score) is adopted as the primary KPI to evaluate the various combined methods in this research.

On the other hand, as this thesis pertains to real-time recognition, the recognition time is also taken into account in the assessment of recognition performance using the combined methods, especially in the situation where the methods are of the same F-measure. Specially, the recognition time is reflected by the time cost of recognition that counts from the camera capturing the testing banknote to outputting the recognition result to the screen. Nevertheless, F-measure is given priority in this research project, which is supposed to be more significant than the recognition time.

Chapter 4 Implementation

This chapter will provide a more detailed description of the experiments. The figures with respect to the scanned banknote images will be inserted as examples in this chapter, including explanations on the generating of the training set, pre-processing procedure, the procedures of feature extraction including the extraction of the colour feature, the texture feature, and the composite feature containing colour and texture elements, the steps to reduce the dimensionality of the feature vector using PCA and LDA respectively, and the process of creating the classification model using the MDC and the BPNN, respectively. Section 4.2 will describe the details of the testing procedure.

4.1 Training procedure

4.1.1 Training set

Training samples need to be converted into digital images for computers to deal with. Two hundred and fifty training samples, with 50 training samples for each denomination, are processed in different ways – via scanner or webcam. Accordingly, the training set is divided into two groups in this research project. One group is composed of the scanned banknote images at 300 dpi, and the other group consists of the images captured by webcam at 96 dpi. The two groups of training sets are separately used in this research project for comparison, as shown in Figure 4.1.



(a) The banknote images of different denominations produced by scanner



(b) The banknote images of different denominations produced by webcam

Figure 4.1 The comparison between the banknote images produced by scanner and the images produced by webcam

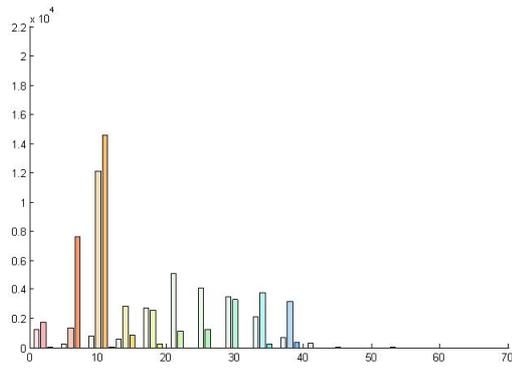
4.1.2 Pre-processing

All the original banknote images of different denominations, either taken by webcam or scanned by scanner, have an aspect ratio of 2:1. In order to reduce computation load as well as maintain the scale, the training images are resized to 400×200 in MATLAB.

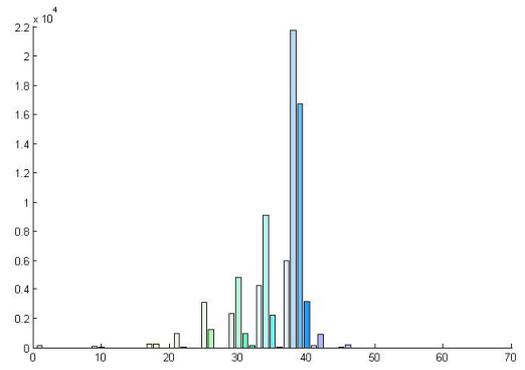
4.1.3 Feature extraction

- **Colour feature extraction**

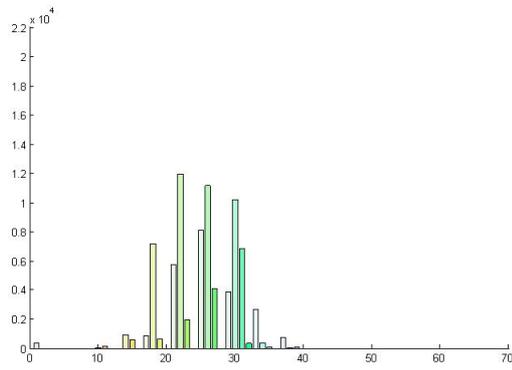
The input image is converted from RGB to HSV mode in MATLAB. Considering that PCA, one of the dimensionality reduction algorithms to be used in the experiments, is sensitive to illumination, the *value* representing the brightness of a colour is discarded to ensure the stable condition of PCA. Then, the remaining components are quantised in accordance with human eyes' sensitivity to different colours, dividing the *hue* into 16 portions, and the *saturation* into four portions (Peng, Zhu, & Lin, 2012; Zhao, Yan, & Zhang, 2007). In this way, the colour histogram with 64 bins for each denomination is formed, as shown in Figure 4.2.



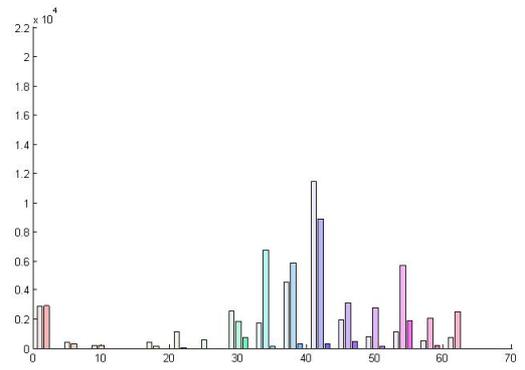
(a) 5 NZD



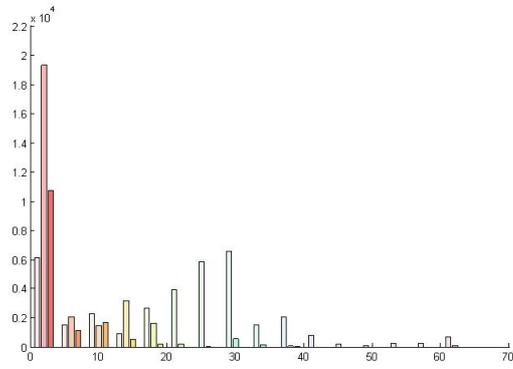
(b) 10 NZD



(c) 20 NZD



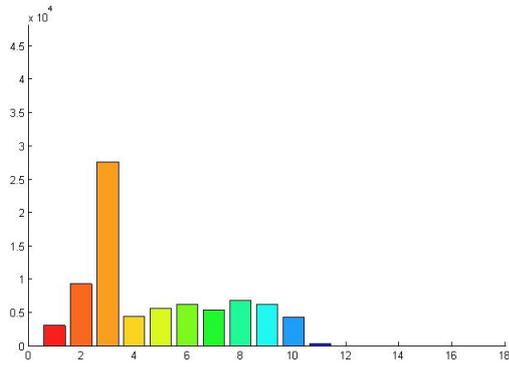
(d) 50 NZD



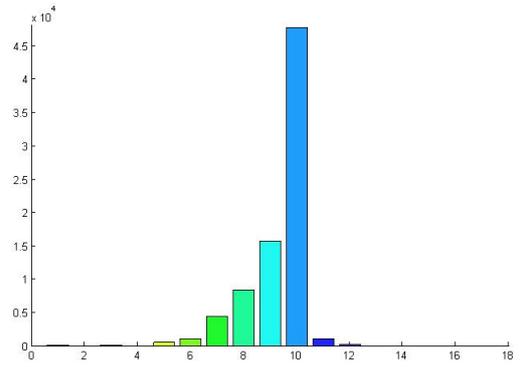
(e) 100 NZD

Figure 4.2 Colour histograms with 64 bins for the scanned banknote images

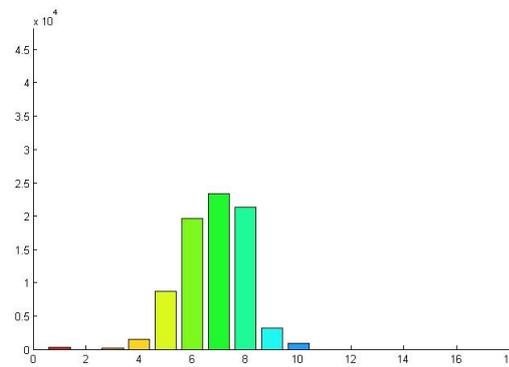
It is observed that the adjacent columns have very similar colours in the colour histogram with 64 bins, where each of the 16 portions of the *hue* is combined with each of the four portions of the *saturation*. Therefore, we discarded the *saturation* as well to highlight the *hue* difference in the colour histograms. Finally, the colour histograms with 64 bins are transformed to the more distinguishable colour histograms with 16 bins, as displayed in Figure 4.3.



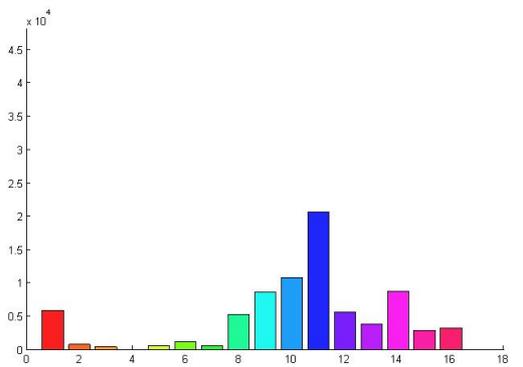
(a) 5 NZD



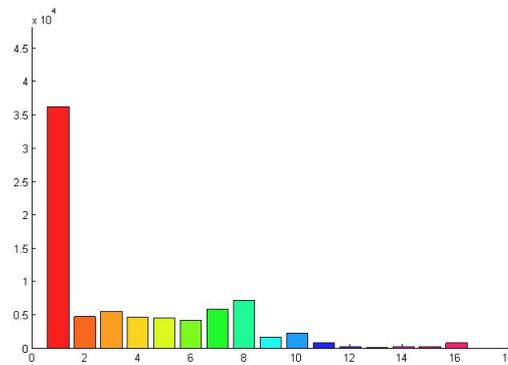
(b) 10 NZD



(c) 20 NZD



(d) 50 NZD



(e) 100 NZD

Figure 4.3 Colour histograms with 16 bins for the scanned banknote images

In training, the *ColourHistogramMatrix* is generated to store the extracted colour features of all training images. As there are 50 training samples for each of the five different denominations, and each training sample is represented by a colour histogram with 16 bins, the dimension of the *ColourHistogramMatrix* is 250×16 . The first 50 vectors are the colour histogram vectors for 5 NZD, followed by 50 colour histogram

vectors for 10 NZD, 50 colour histogram vectors for 20 NZD, 50 colour histogram vectors for 50 NZD, 50 colour histogram vectors for 100 NZD in sequence, as shown in Equation 4.1.

$$\text{ColourHistogramMatrix} = \begin{pmatrix} 5\text{NZDClrHstVec1} \\ \dots \\ 5\text{NZDClrHstVec50} \\ 10\text{NZDClrHstVec1} \\ \dots \\ 10\text{NZDClrHstVec50} \\ 20\text{NZDClrHstVec1} \\ \dots \\ 20\text{NZDClrHstVec50} \\ 50\text{NZDClrHstVec1} \\ \dots \\ 50\text{NZDClrHstVec50} \\ 100\text{NZDClrHstVec1} \\ \dots \\ 100\text{NZDClrHstVec50} \end{pmatrix} \quad (4.1)$$

- **Texture feature extraction**

The input image is converted to a grey-scale image in MATLAB. To enhance the contrast of the grey levels, the grey-scale image is passed through histogram equalisation in MATLAB. The default function transforms an intensity image to an image with 64 discrete grey levels. By histogram equalisation, the histogram of an image is equalised, and frequencies of intensity are uniformly distributed over the whole intensity range. A comparison of banknote images between before the use of histogram equalisation and after the use of histogram equalisation is displayed in Figure 4.4, giving a direct view on the usefulness of histogram equalisation for the extraction of the banknotes' texture feature.



(a) 5 NZD



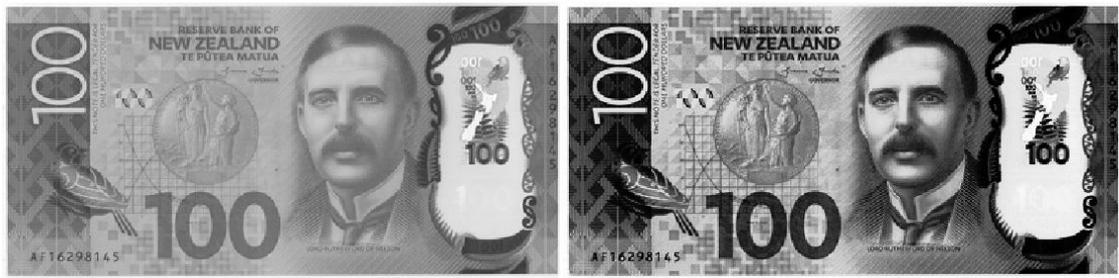
(b) 10 NZD



(c) 20 NZD



(d) 50 NZD



(e) 100 NZD

Figure 4.4 A comparison of before and after use of histogram equalisation for the scanned banknote images

The LBP operator is employed to extract the texture feature of banknotes in the experiments. If the LBP feature of the banknote is directly extracted as a whole, then there is only one LBP histogram that would be obtained. However, if the banknote is divided into several blocks, then the relative position of each block can be somehow recorded, and 32 LBP histograms would be obtained based on 32 blocks. In this way, each of the blocks can be compared with the corresponding block of the other banknotes via LBP histogram, thereby improving the recognition rate. Accordingly, before the implementation of the LBP algorithm, each banknote is partitioned into four rows and eight columns, generating 32 blocks with each size of 50×50 , as shown in Figure 4.5.



(a) 5 NZD



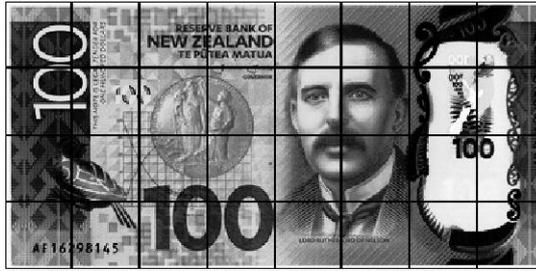
(b) 10 NZD



(c) 20 NZD



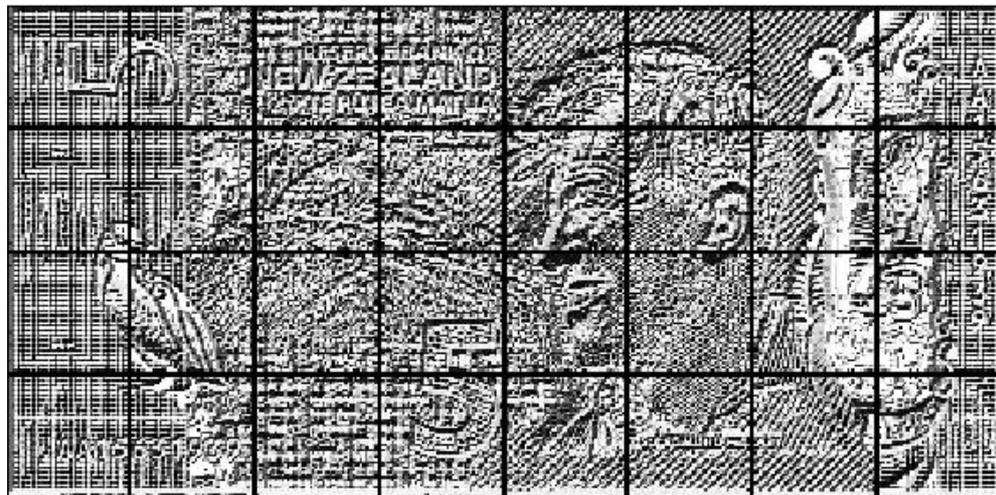
(d) 50 NZD



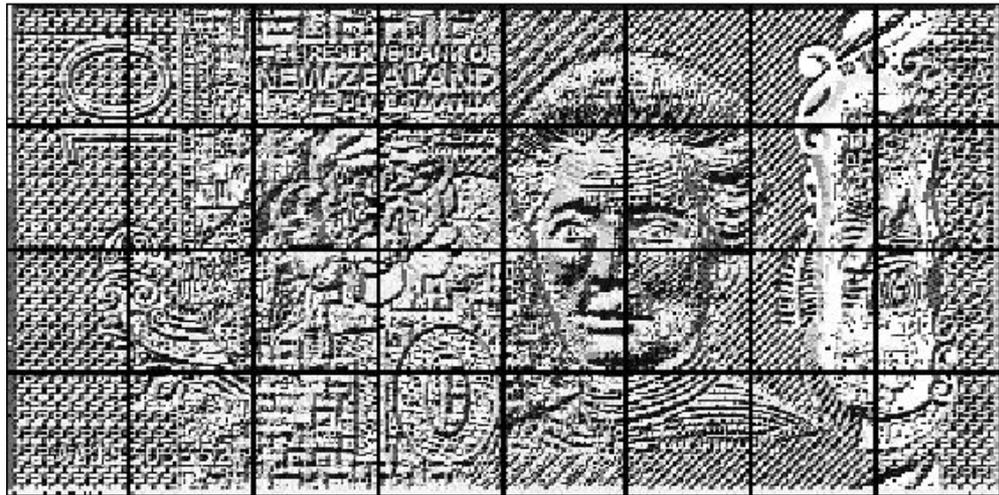
(e) 100 NZD

Figure 4.5 Thirty-two uniform blocks for the scanned banknote images

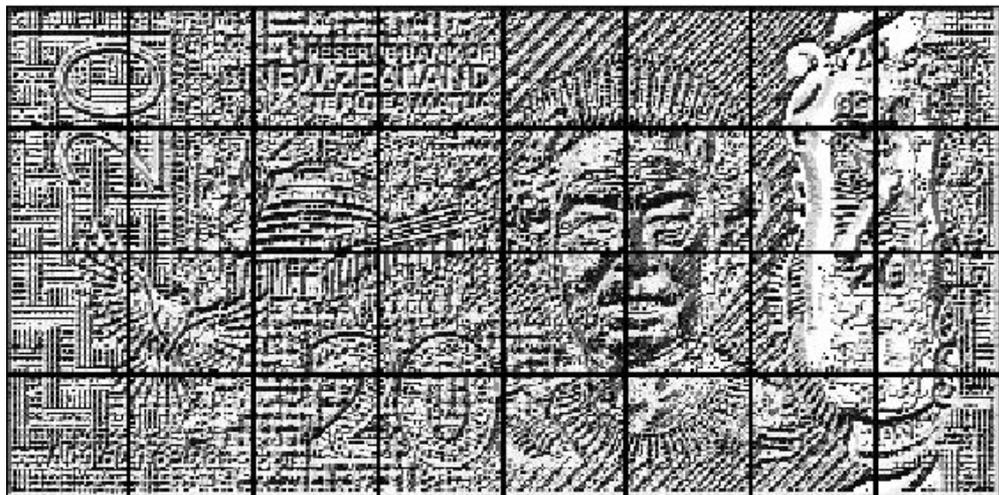
The LBP operator is then implemented on each block in order. For each pixel in a block, its grey-scale value is compared with its eight neighbouring pixels in the order of the neighbouring pixels, i.e. the neighbouring pixel on its left-top, top, right-top, left, right, left-bottom, bottom and right-bottom. When the grey-scale value of the neighbouring pixel is greater than that of the central pixel, denote ‘1’, and otherwise, ‘0’. Following this rule, an 8-digit binary number composed of ‘0’ and (or) ‘1’ represents each pixel, forming the LBP code of each pixel. Different LBP codes correspond to certain micro-features; accordingly, the LBP features for the banknotes of various denominations are displayed in Figure 4.6.



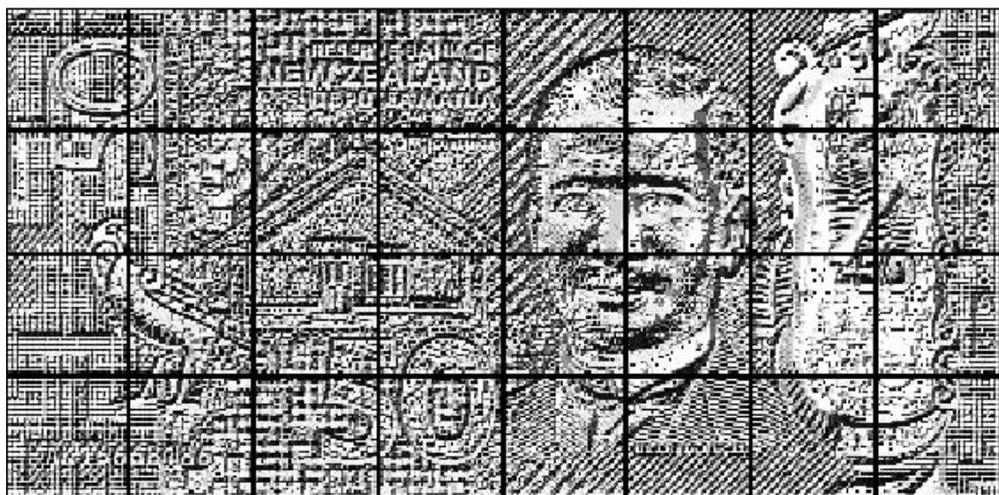
(a) 5 NZD



(b) 10 NZD



(c) 20 NZD



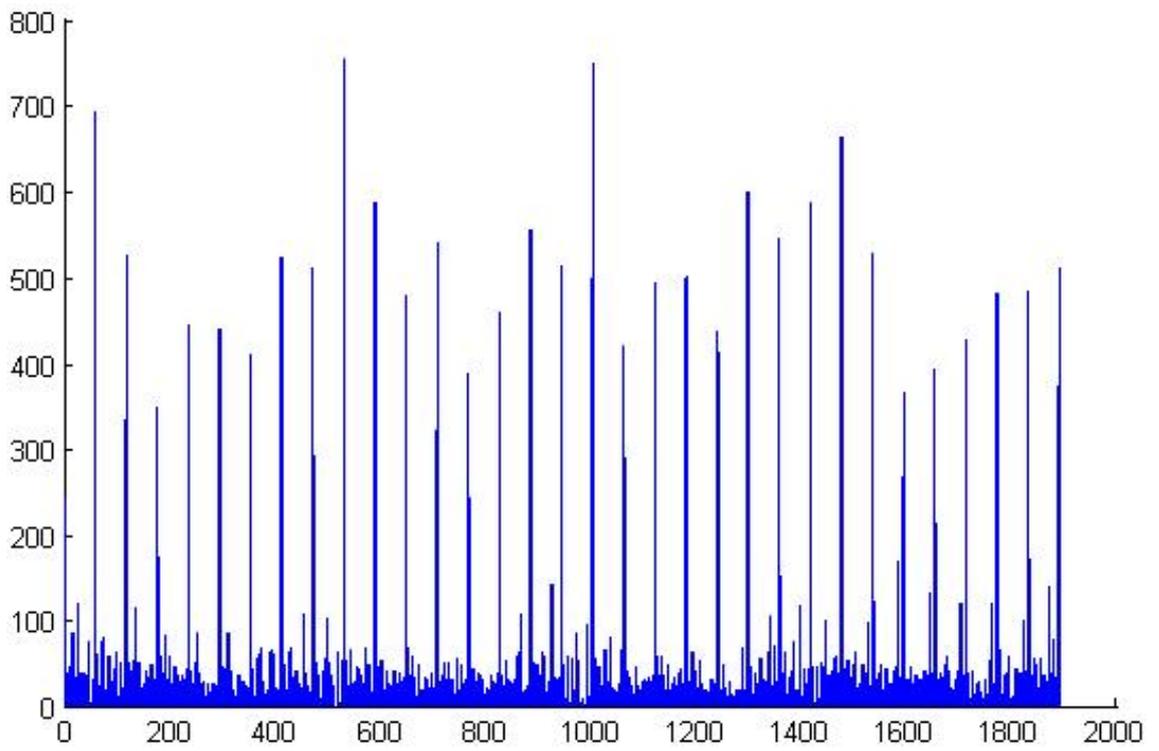
(d) 50 NZD



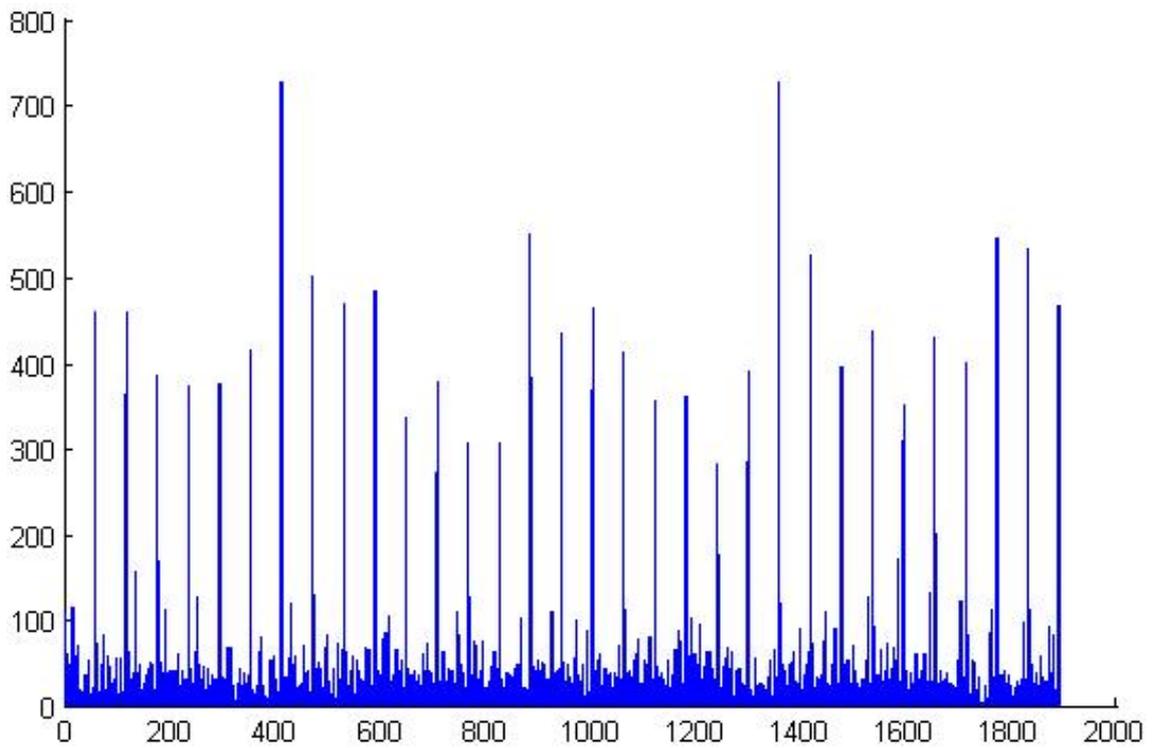
(e) 100 NZD

Figure 4.6 LBP features for the scanned banknote images

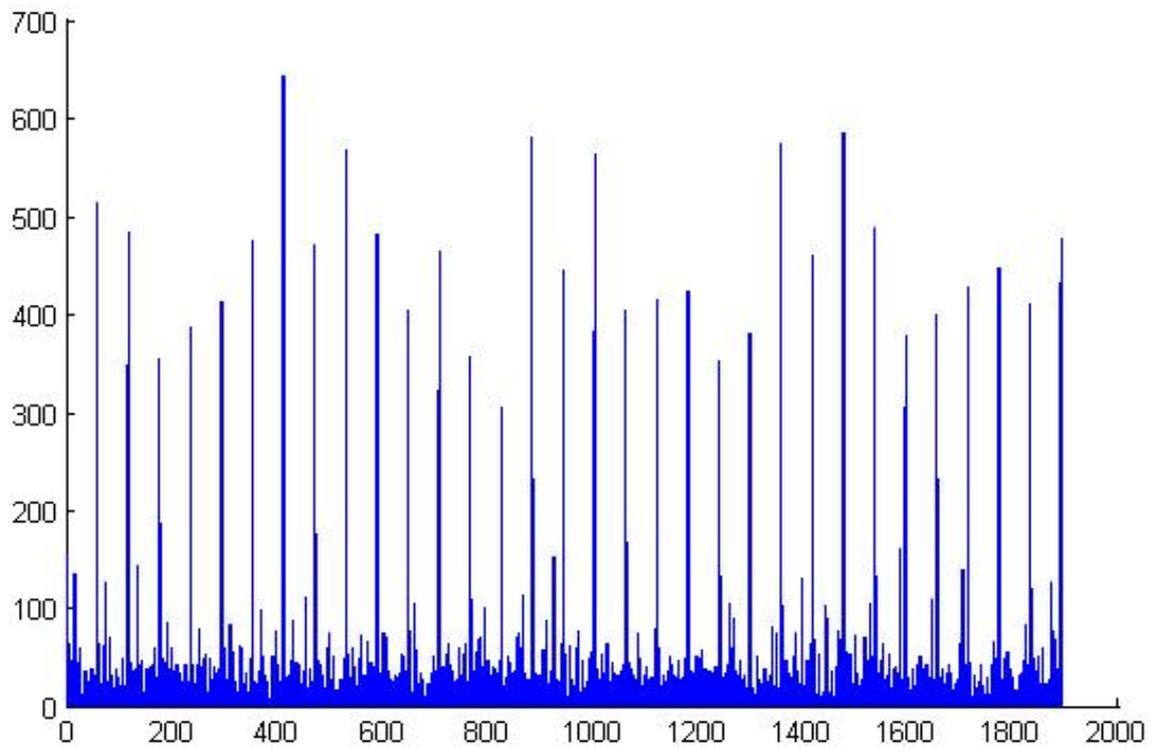
The so-called uniform pattern is then applied to optimise those local binary patterns for creating an efficient uniform LBP histogram. In the computation of the uniform LBP histogram, each uniform pattern is described by a separate bin in the histogram, and all non-uniform patterns are described by a single bin. As a result, the uniform LBP histogram contains 59 bins corresponding to 58 uniform patterns, plus one non-uniform pattern. Since each banknote comprises 32 blocks, and each block generates a uniform LBP histogram with 59 bins, the number of the columns of the block uniform LBP histogram for the whole banknote should be $59 \times 32 = 1888$. Figure 4.7 illustrates the block uniform LBP histograms for the banknotes of different denominations that are generated in such a way.



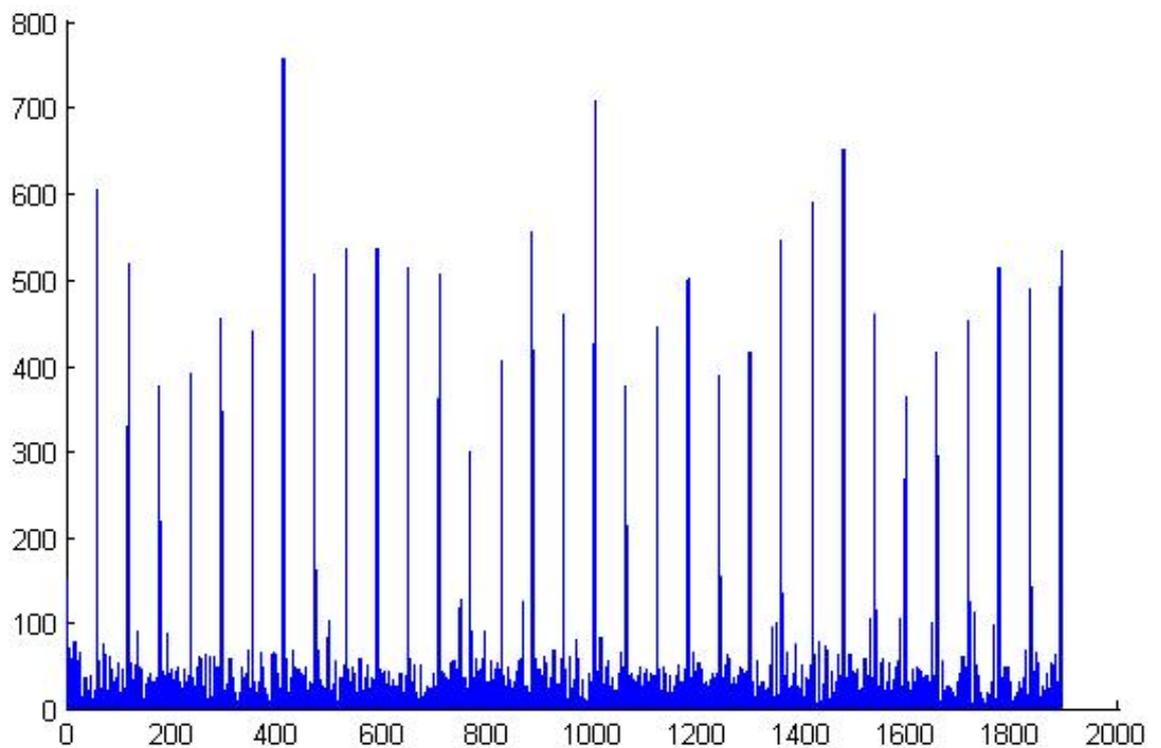
(a) The block uniform LBP histogram for 5 NZD



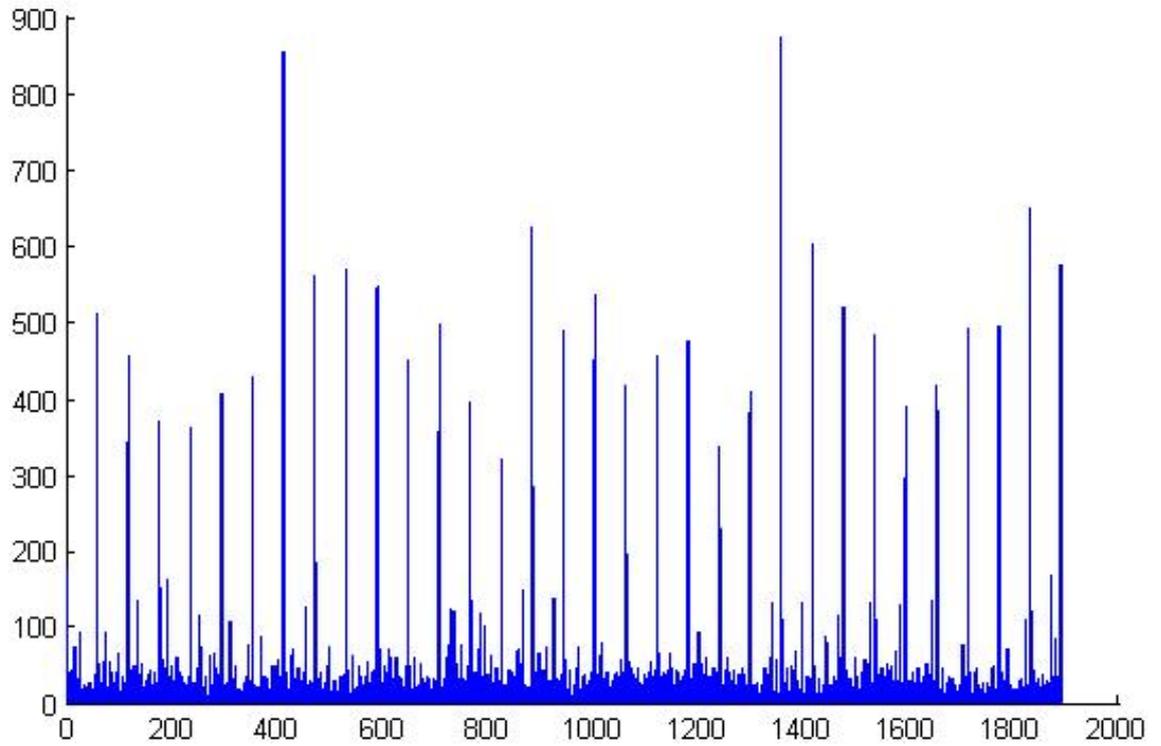
(b) The block uniform LBP histogram for 10 NZD



(c) The block uniform LBP histogram for 20 NZD



(d) The block uniform LBP histogram for 50 NZD



(e) The block uniform LBP histogram for 100 NZD

Figure 4.7 Block uniform LBP histograms for the scanned banknote images

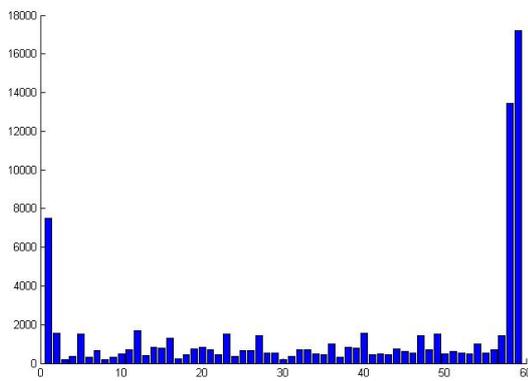
In training, the *LBPHistogramMatrix* is generated to store the extracted LBP features of all training images. As there are 50 training samples for each of the five different denominations, and each training sample is represented by a block uniform LBP histogram with 1888 bins, the dimension of the *LBPHistogramMatrix* is 250×1188 . The first 50 vectors are the block uniform LBP histogram vectors for 5 NZD, followed by 50 block uniform LBP histogram vectors for 10 NZD, 50 block uniform LBP histogram vectors for 20 NZD, 50 block uniform LBP histogram vectors for 50 NZD, 50 block uniform LBP histogram vectors for 100 NZD in sequence, as shown in Equation 4.2.

$$LBPHistogramMatrix = \begin{pmatrix} 5NZDLBPHstVec1 \\ \dots \\ 5NZDLBPHstVec50 \\ 10NZDLBPHstVec1 \\ \dots \\ 10NZDLBPHstVec50 \\ 20NZDLBPHstVec1 \\ \dots \\ 20NZDLBPHstVec50 \\ 50NZDLBPHstVec1 \\ \dots \\ 50NZDLBPHstVec50 \\ 100NZDLBPHstVec1 \\ \dots \\ 100NZDLBPHstVec50 \end{pmatrix} \quad (4.2)$$

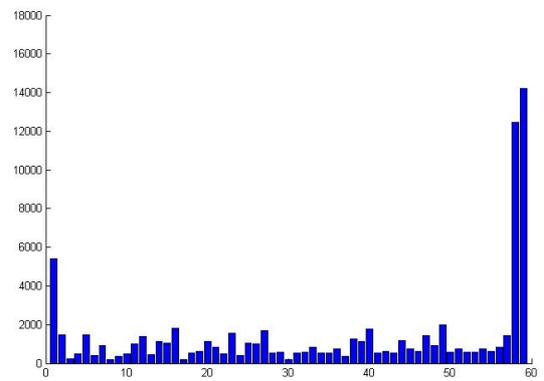
- **Composite feature extraction**

In this thesis, the colour and texture elements are also served together as a composite feature to study. Considering that dimensionality reduction methods will be employed after feature extraction, the appropriate selection for the dimensions of colour part and texture part is crucial to the dimensionality-reduced composite feature. Both the colour part and texture part utilise histograms as feature descriptors and, also, the colour histogram and LBP histogram for a 400×200 banknote are both based on the corresponding values of its 80,000 pixels. Based on this, for a banknote whose bin numbers in the colour histogram and in the LBP histogram differ greatly, the statistics reflected by the vertical axis of the two histograms will accordingly have a difference of magnitude. When dimensionality reduction methods are applied to the composite feature of this banknote, the obtained colour-LBP histogram will rely on the statistics of the histogram that has a larger magnitude, indicating that the composite feature will be more dependent on a certain feature. Thus, to balance the impact of the colour part and texture part on the composite feature, the colour histogram with 64 bins and the LBP histogram with 59 bins are concatenated to form the colour-LBP histogram with 123 bins. Specifically, in the colour-LBP histogram, the colour part is based on the 16

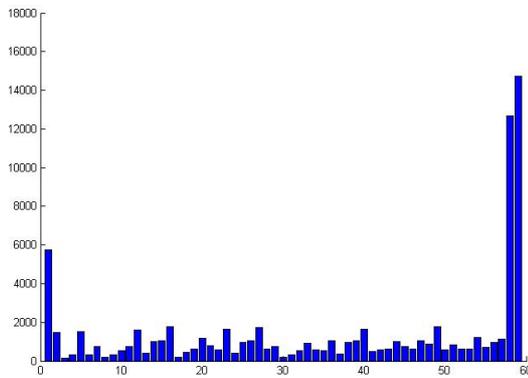
portions of the *hue* and the four portions of the *saturation* of a banknote image in an HSV colour space. The 64 bins in the colour histogram are ordered as follows: H_1S_1 , H_1S_2 , H_1S_3 , H_1S_4 , H_2S_1 , H_2S_2 , H_2S_3 , H_2S_4 , $H_{16}S_1$, $H_{16}S_2$, $H_{16}S_3$, $H_{16}S_4$. Besides, unlike the way of solely extracting the LBP feature where the banknote image is partitioned into 32 blocks, for the LBP feature part of the composite feature, the banknote image is considered as a whole to extract the LBP feature. With the additional help of the uniform patterns, the uniform LBP histogram with 59 bins is generated as the representative of the LBP part in the colour-LBP histogram, as shown in Figure 4.8.



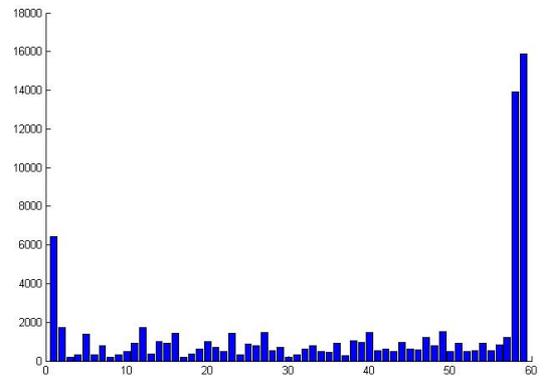
(a) 5 NZD



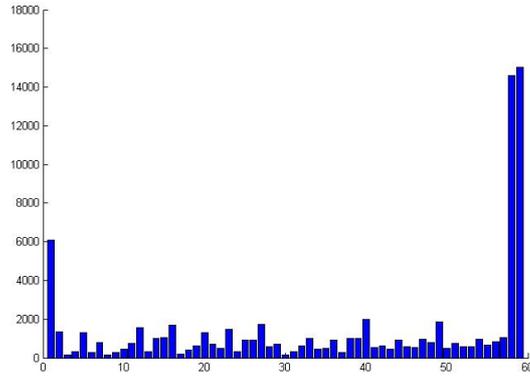
(b) 10NZD



(c) 20 NZD



(d) 50 NZD



(e) 100 NZD

Figure 4.8 LBP histogram part in the colour-LBP histogram for the scanned banknote images

In training, the *Colour-LBPHistogram* is generated to store the extracted composite features of all training images. As there are 50 training samples for each of the five different denominations, and each training sample is represented by a colour-LBP histogram with 123 bins, the dimension of the *Colour-LBPHistogram* is 250×123 . The first 50 vectors are the colour-LBP histogram vectors for 5 NZD, followed by 50 colour-LBP histogram vectors for 10 NZD, 50 colour-LBP histogram vectors for 20 NZD, 50 colour-LBP histogram vectors for 50 NZD, 50 colour-LBP histogram vectors for 100 NZD in sequence, as shown in Equation 4.3.

$$\text{Colour-LBPHistogram} = \begin{pmatrix} 5\text{NZDClrHstVec1} & 5\text{NZDLBPHstVec1} \\ \dots & \dots \\ 5\text{NZDClrHstVec50} & 5\text{NZDLBPHstVec50} \\ 10\text{NZDClrHstVec1} & 10\text{NZDLBPHstVec1} \\ \dots & \dots \\ 10\text{NZDClrHstVec50} & 10\text{NZDLBPHstVec50} \\ 20\text{NZDClrHstVec1} & 20\text{NZDLBPHstVec1} \\ \dots & \dots \\ 20\text{NZDClrHstVec50} & 20\text{NZDLBPHstVec50} \\ 50\text{NZDClrHstVec1} & 50\text{NZDLBPHstVec1} \\ \dots & \dots \\ 50\text{NZDClrHstVec50} & 50\text{NZDLBPHstVec50} \\ 100\text{NZDClrHstVec1} & 100\text{NZDLBPHstVec1} \\ \dots & \dots \\ 100\text{NZDClrHstVec50} & 100\text{NZDLBPHstVec50} \end{pmatrix} \quad (4.3)$$

4.1.4 Dimensionality reduction

In training, PCA and LDA are separately implemented on the *ColourHistogramMatrix*, the *LBPHistogramMatrix* and the *Colour-LBPHistogramMatrix* mentioned above to reduce the dimensionality of the feature vectors. The following two paragraphs explain the process of *ColourHistogramMatrix* dimensionality reduction. As for the *LBPHistogramMatrix* and the *Colour-LBPHistogramMatrix*, the dimensionality reduction to them is done in the same way, so an explanation is omitted from this thesis.

This paragraph is the explanation of how to utilise PCA to reduce the dimensionality of the colour feature vector. After the *ColourHistogramMatrix* is generated, the PCA algorithm is implemented on the *ColourHistogramMatrix*. More specifically, the covariance matrix of the *ColourHistogramMatrix* is first generated. Then, the eigenvectors and eigenvalues of the covariance matrix are calculated. By abandoning the insignificant eigenvectors whose eigenvalues are relatively small, a new eigenvector matrix with fewer rows is produced, and later used for projecting the original *ColourHistogramMatrix*. *ColourHistogramMatrix* multiplies by the new eigenvectors matrix, obtaining the projection vectors of the colour histograms in PCA space, namely, the PCA colour feature vectors of the training images.

This paragraph explains how to utilise LDA to reduce the dimensionality of the colour feature vector. Firstly, calculating the mean of colour feature vector for each class defined as μ_i . Then the mean vector for the whole *ColourHistogramMatrix* is calculated and defined as μ . So, the between-class scatter is $S_B = \sum_{i=1}^C N_i (\mu_i - \mu) (\mu_i - \mu)^T$, where $N_i = 50$, representing the number of samples in each class, and $C = 5$, representing the number of classes. The next step is to calculate the within-class scatter S_w . For each training sample, its colour feature vector x_k minus the mean colour feature vector for each class μ_i , obtains the vector $(x_k - \mu_i)$. By multiplying the vector $(x_k - \mu_i)$ with its transposed vector $(x_k - \mu_i)^T$, a 16×16 square matrix is produced, corresponding to the particular training sample. The within-class scatter S_w

is then obtained by summing up the 250 square matrices that correspond to the 250 training samples. To express it in mathematical way, the within-class scatter $S_w = \sum_{i=1}^c \sum_{x_k \in x_i} (x_k - \mu_i)(x_k - \mu_i)^T$, where $C = 5$, representing the number of classes. The next step is to maximise the value of S_B/S_w , attempting to increase the value of S_B as well as decreasing the value of S_w . By applying the Lagrange Multiplier, this attempt can be translated to calculate the eigenvalues and eigenvectors of the matrix produced by S_B/S_w operation. The eigenvectors with high eigenvalues will be considered to be able to effectively classify different categories in LDA space. As a result, after applying dimensionality reduction to the colour feature vector using LDA, several eigenvectors corresponding to the highest eigenvalues are selected to form a new eigenvector matrix that is later used for projecting the original *ColourHistogramMatrix* to LDA space. *ColourHistogramMatrix*, multiplied by the new eigenvectors matrix, obtains the projection vectors of the colour histograms in LDA space, i.e. the LDA colour feature vectors of the training images.

In terms of the criterion for deciding the extent of reduction on the dimensionality of the feature vector, in the experiments, the eigenvectors whose eigenvalues account for 95% of the total eigenvalues are retained after the procedure of dimensionality reduction. The corresponding dimensionality of features vectors after the implementation of PCA or LDA is shown in Table 4.1.

Table 4.1 The dimensionality of feature vectors after dimensionality reduction

Characterisation	Dimensionality of the feature vector	
	Scan	Webcam
Colour_PCA	4	4
LBP_PCA	99	96
Colour+LBP_PCA	4	4
Colour_LDA	4	4
LBP_LDA	4	4
Colour+LBP_LDA	4	4

4.1.5 Classification model

The MDC and BPNN classifiers are separately employed for creating the classification model. In terms of the MDC, the procedure it participates in at the training stage is calculating the mean vector for each class. Taken the colour feature as an example, in the training stage, after PCA is applied to the *ColourHistogramMatrix*, the final PCA matrix contains 250 PCA feature vectors, with 50 feature vectors for each denomination. Thus, the average of the 50 feature vectors for the same denomination is considered as the PCA colour feature for that particular denomination. Similarly, the final LDA colour feature matrix contains 250 LDA feature vectors, with 50 feature vectors for each denomination. Thus, the average value of the 50 feature vectors for the same denomination is considered as the LDA colour feature for that particular denomination.

For the BPNN classifier used in the experiments, the MLP model and sigmoid transfer function that is typically used as the activation function in the hidden layer in a multilayer network are involved in the three-layer BPNN classifier. Sigmoid functions compress an infinite input range into a finite output range, featuring the slopes approaching zero when the inputs are getting large. When using the steepest descent to train a multilayer network with a sigmoid function, small alterations to the weights and biases can be caused by the gradients that have a very small magnitude, even if the weights and biases are far from their optimal values. To eliminate the negative influence of the magnitude of the partial derivatives, the resilient back-propagation algorithm (Rprop) is adopted in the experiments to train the network, as it determines the direction of the weight update by the sign of its derivative rather than the magnitude of the derivative. Moreover, the Rprop algorithm is much faster than the standard steepest descent algorithm, and it requires only a modest increase in memory requirements (Riedmiller & Braun, 1993), which are also the reasons for adopting the Rprop algorithm in the experiments.

To be specific, when the network weights and biases are initialised, the network is ready

for training. The default performance function of a feed-forward network, MSE, is used as the network performance function in this thesis. The simplest optimisation algorithm - gradient descent - is used to update the network weights and biases in the direction where the performance function decreases most rapidly, the negative of the gradient. The batch training mode is implemented; all the inputs in the training set are applied to the network before the weights are updated. The `trainrp()` function is used in MATLAB as the network training function to update weights and biases in accordance with the Rprop algorithm. The training parameters include *epochs*, *show*, *goal*, *time*, *min_grad*, *max_fail*, *delt_inc*, *delt_dec*, *delta0*, and *deltamax*. Since the performance of the Rprop algorithm is not very sensitive to the training parameters, these parameters are left at the default. Training stops when the maximum number of *epochs* is reached or the maximum amount of *time* is exceeded or performance is minimised to the *goal*, or the performance gradient falls below *min_grad*, or validation performance has increased over *max_fail* times since the last time it decreased. The training window appearing during training displays the constantly updated progress. The final result will be displayed on the training window at the end of each training time.

Since there are 12 combined methods where BPNN classifiers are involved, 12 different networks are correspondingly generated in the experiments. For the 12 networks, the number of nodes in the output layer is five, corresponding to the five different classes of output results – 5 NZD, 10 NZD, 20 NZD, 50 NZD and 100 NZD. However, the number of the input nodes and the hidden nodes varies depending on the properties of the combination. The number of input nodes is equal to the dimensions of the input feature vector in the network. Although hidden layers have no direct interaction with the external environment, the number of hidden nodes has a considerable impact on the final output. The trial and error method, coupled with the forward approach, is used to decide the number of nodes in the hidden layer. The trial and error method is characterised by continual attempts until success, and the forward approach starts with selecting a small number of hidden nodes and then increases the number of hidden nodes after testing the performance of the trained network (F. S. Panchal & M. Panchal,

2014). The experiments begin with two hidden nodes, followed by repeating the above procedure until the best possible performance is achieved.

Because of the random setting of the initial weights and biases of the network, given a finite number of training times, the training results will not always be the same. At the end, the best parameters of each network are ascertained after proper training times, as shown in Table 4.2 and Table 4.3; the performance of those networks will undergo a final comparison with the other combined methods.

Table 4.2 Parameters of the optimal networks when using the scanned banknote images for training

Combination No.	Number of input nodes	Number of hidden nodes	Number of output nodes	Best validation performance
7	4	4	5	1.4487e-3 at epoch 131
8	99	4	5	2.6582e-2 at epoch 218
9	4	5	5	1.9148e-3 at epoch 169
10	4	4	5	7.7004e-4 at epoch 191
11	4	4	5	7.0968e-3 at epoch 126
12	4	5	5	4.2773e-4 at epoch 140

Table 4.3 Parameters of the optimal networks when using the banknote images captured by webcam for training

Combination No.	Number of input nodes	Number of hidden nodes	Number of output nodes	Best validation performance
7	4	4	5	7.5475e-4 at epoch 554
8	96	4	5	1.9531e-2 at epoch 217
9	4	5	5	1.9297e-3 at epoch 164
10	4	4	5	5.1783e-4 at epoch 137
11	4	4	5	1.7356e-3 at epoch 190
12	4	5	5	1.70e-4 at epoch 192

4.2 Testing procedure

Each combined method is tested on a set of 50 banknotes, with 10 banknotes for each denomination. The testing condition is the indoor environment where the light produced

by ordinary incandescent bulbs is provided. In the testing stage, once the camera is activated, the 800×400 recognition window is highlighted by a yellow colour to make the testing banknote image distinguishable from the background, as shown in Figure 4.9. In the meanwhile, we place the banknote being tested in front of the camera, and smoothly rotate the testing banknote left and right without exceeding 15 degrees. The banknotes within the recognition window are captured as frames, and each frame serves as the testing image to be analysed. Subsequently, the testing image is processed in the same way as processing the training images, including pre-processing, feature extraction and dimensionality reduction. Finally, the processed banknote image being tested is then sent to the classification model generated in the training stage for denomination recognition.

During testing, taking a 5 NZD testing banknote as an example, after 10 times of recognition of a 5 NZD testing note, a five-second intermission is given to place another 5 NZD for testing within the recognition window. Thus, it is noteworthy that 50 testing samples eventually generate a total of 500 recognition results, with each testing sample generating 100 recognition results. The recognition result of each time is shown above the recognition window, highlighted by a red colour, as shown in Figure 4.9.

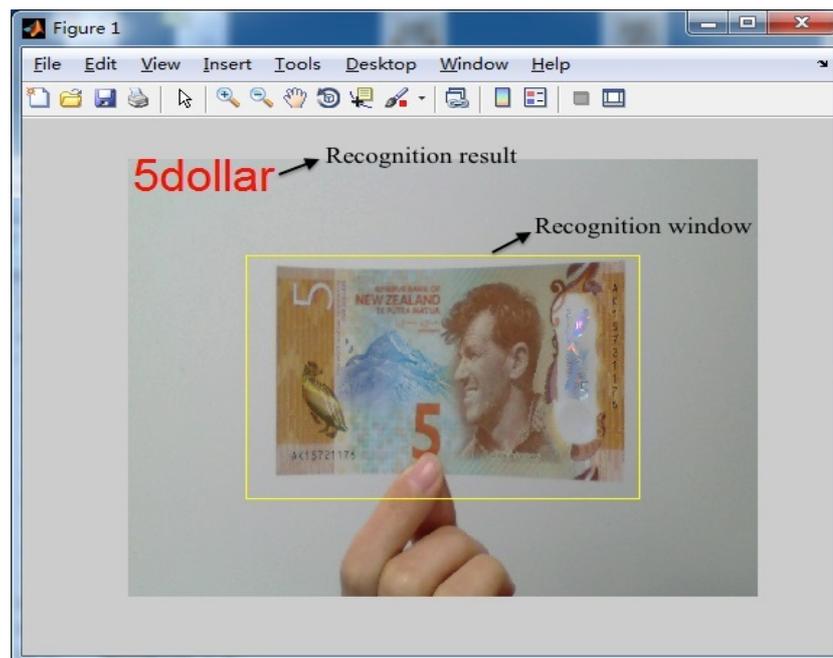


Figure 4.9 Testing window in MATLAB

Chapter 5 Findings and discussion

In this chapter, we will present the experimental results, and examine the initial hypothesis in light of the findings of this research. Similarities and differences between the current study and previous research in the field of paper currency recognition will be discussed and addressed.

5.1 Experimental results

The recognition results are displayed using a confusion matrix in MATLAB. The performance is based on the recognition results of 500 test images, with 100 test images for each class. As two training sets are involved in contrast experiments, two groups of recognition results are correspondingly generated at the end of the experiments. One group covers recognition results using the combined methods when the training set is composed of scanned banknote images, as shown in Appendix B. The other group covers recognition results using the combined methods when the training set comprises the banknote images captured by the webcam, as shown in Appendix C. The F-measure of each combined method for recognising each denomination in the two groups of training sets is shown in Table 5.1 and Table 5.2, respectively. The comparison of the average time cost of recognition using each combined method in the two groups of the training set is shown in Figure 5.1. Since each combined method costs almost the same period of time for recognition using either scanned banknote images for training or the banknote images captured by webcam for training, the mean time cost of recognition between the two groups of training sets is therefore exploited as the recognition time for the combined method, as shown in Figure 5.2. Four decimal places are reserved with the rounding calculation for the F-measure and the recognition time.

Table 5.1 The F-measure of each combined method for recognising each denomination when using scanned banknote images for training

Combination No.	Denomination class					Average
	\$5	\$10	\$20	\$50	\$100	
1	0.9049	1	0.8827	1	1	0.9575
2	0	0.4750	0.8718	0	0.7609	0.4216
3	0.9901	1	0.9899	1	1	0.9960
4	0.7663	1	0.6014	1	0.9796	0.8695
5	0.9529	1	0.9569	1	1	0.9820
6	0.6838	0.3471	1	0.9662	1	0.7994
7	0.9434	1	0.9362	1	1	0.9759
8	0.8889	0.8772	0.9091	0.8372	1	0.9025
9	0.9901	1	0.9899	1	1	0.9960
10	0.8967	1	0.8701	1	1	0.9534
11	1	0.9132	0.9071	1	1	0.9641
12	0.9390	0.9305	1	1	1	0.9739

Table 5.2 The F-measure of each combined method for recognising each denomination when using the banknote images captured by webcam for training

Combination No.	Denomination class					Average
	\$5	\$10	\$20	\$50	\$100	
1	1	0.9848	1	0.9852	1	0.9940
2	0.8729	0.9529	0.8170	0.9495	0.9744	0.9133
3	1	0.9899	1	0.9901	1	0.9960
4	0.9804	1	1	1	0.9796	0.9920
5	1	0.99804	1	1	0.9796	0.9920
6	1	1	0.8439	0.9756	0.9009	0.9441
7	0.9950	0.9899	0.9950	0.9901	1	0.9940
8	0.9100	0.9442	0.9189	0.8095	0.8368	0.8839
9	1	1	0.9950	1	0.9950	0.9980
10	0.9804	0.9796	1	1	1	0.9920
11	1	0.9804	0.9796	1	1	0.9920
12	0.9901	0.9849	1	0.9950	1	0.9940

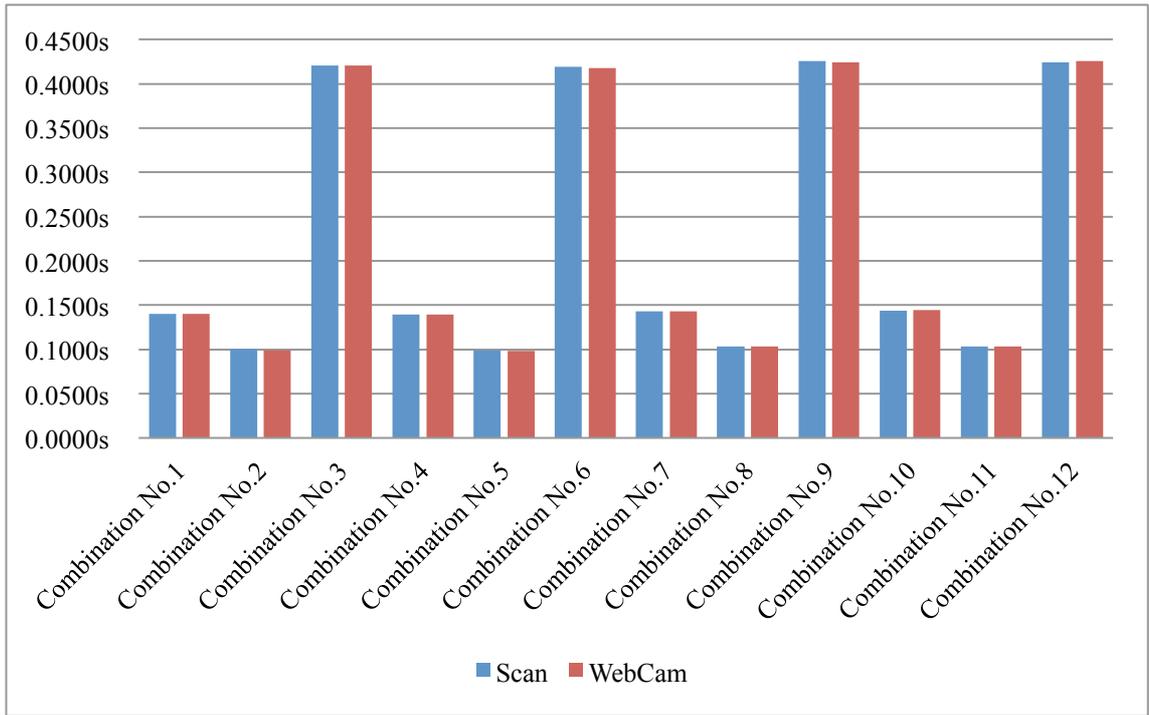


Figure 5.1 Comparison of the average time cost of recognition using each combined method in the two groups of training sets

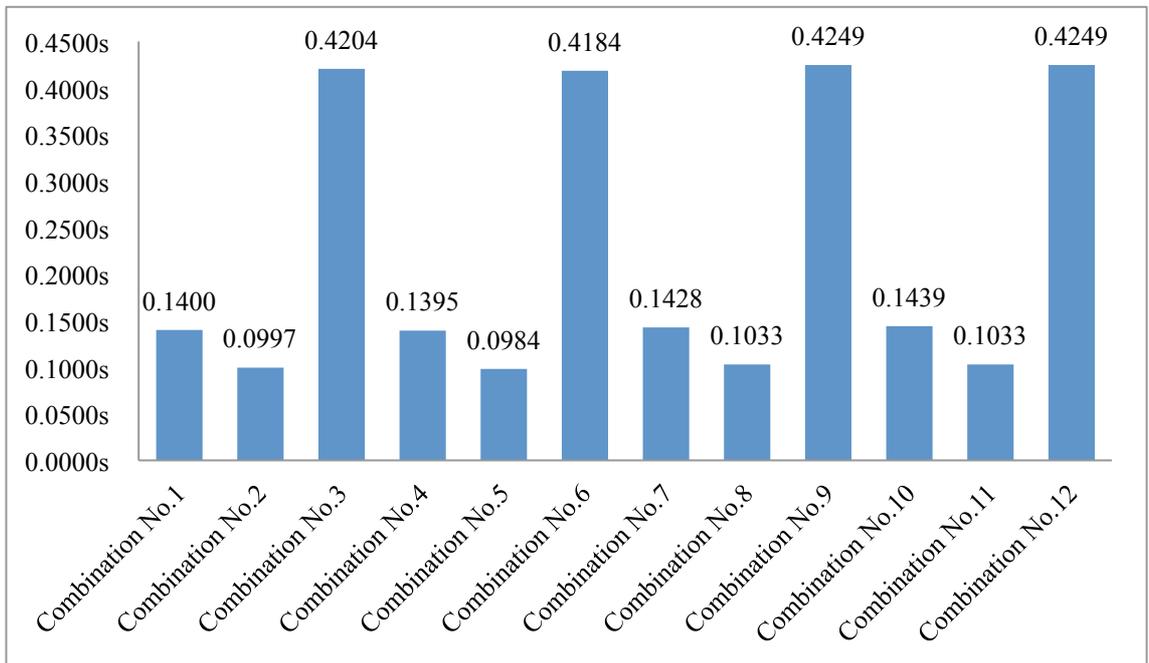


Figure 5.2 The average time cost of recognition using each combined method

5.2 Hypothesis validation

A few findings are noticed from the experiments, which have verified several initial hypotheses. Finally, the core research question is answered.

- Hypothesis 1.2 that the training set generated by webcam would yield better recognition results proves true by the experiments.

Table 5.3 The comparison of the average F-measure between the training sets

Combination No.	Training sets	
	Scan	Webcam
1	0.9575	0.9940
2	0.4216	0.9133
3	0.9960	0.9960
4	0.8695	0.9920
5	0.9820	0.9920
6	0.7994	0.9441
7	0.9759	0.9940
8	0.9025	0.8839
9	0.9960	0.9980
10	0.9534	0.9920
11	0.9641	0.9920
12	0.9739	0.9940

We have found that in the experiments, the training set composed of the banknote images captured by webcam performs better than the training set consisting of the scanned banknote images. The average F-measure of every combination is beyond 0.88 when using the banknote images captured by webcam for training, while not all the combinations can achieve this when using the scanned banknote images for training. Additionally, the average F-measure of each combination when using the banknote images captured by webcam for training is higher than that when using the scanned banknote images for training, except for the Combination No.8 (LBP_PCA_BPNN) whose average F-measure is 0.8839 when using the banknote images captured by webcam for training but is 0.9025 when using the scanned banknote images for training.

From Figure 4.1 (the comparison between the banknote images produced by scanner and the images produced by webcam), we can observe the main differences between the two groups of training images; namely, the images produced by scanner have greater clarity than the images produced by webcam, and also the majority of the images produced by webcam have a little distortion. When using the banknote images produced by webcam for training, both the training images and the images for testing are produced in the situation where the banknote samples are placed within the recognition window. The obtained training images are therefore similar to the images for testing in terms of clarity and unintended distortion, which enhances the performance of each combination with those training images as the training set. Thus, for developing an efficient banknote real-time recognition system, it is better to produce the training images in nearly the same way as generating the images for testing.

- Hypothesis 2.3 that the combined use of the colour elements and texture elements would result in better recognition results than solely adopting either colour or texture feature proves true by the experiments.

Table 5.4 The comparison of the average F-measure between the extracted features

Combinations	Extracted features					
	Colour		LBP		Colour+LBP	
	PCA	LDA	PCA	LDA	PCA	LDA
Scan_MDC	0.9575	0.8695	0.4216	0.9820	0.9960	0.7994
Scan_BPNN	0.9759	0.9534	0.9025	0.9641	0.9960	0.9739
Webcam_MDC	0.9940	0.9920	0.9133	0.9920	0.9960	0.9441
Webcam_BPNN	0.9940	0.9920	0.8839	0.9920	0.9980	0.9940

Firstly, we have found that the colour feature investigated in this thesis works better with PCA than with LDA. In both of the two groups of training sets, the Combination No.1 (Colour_PCA_MDC) outperforms the Combination No.4 (Colour_LDA_MDC), and the Combination No.7 (Colour_PCA_BPNN) performs better than the Combination No.10 (Colour_LDA_BPNN). Secondly, we have found that the LBP feature investigated in this thesis works better with LDA than PCA. In both of the two groups

of training sets, the Combination No.5 (LBP_LDA_MDC) outperforms the Combination No.2 (LBP_PCA_MDC), and the Combination No.11 (LBP_LDA_BPNN) outperforms the Combination No.8 (LBP_PCA_BPNN). Thirdly, we have found that the composite feature investigated in this research performs better with PCA than with LDA. In both of the two groups of the training sets, the Combination No.3 (Colour+LBP_PCA_MDC) performs better than the Combination No.6 (Colour+LBP_LDA_MDC), and the Combination No.9 (Colour+LBP_PCA_BPNN) outperforms the Combination No.12 (Colour+LBP_LDA_BPNN). Overall, the PCA-based composite feature is the most robust one in this thesis, as it succeeds in delivering an average F-measure of no less than 0.9960 in both of the two groups of training sets no matter working with the MDC or the BPNN.

From Figure 4.3 (colour histograms with 16 bins for the scanned banknote images) and Appendix A (colour histograms with 16 bins for the banknote images captured by webcam), we can observe that the colour features of the banknotes are unevenly distributed in the colour histograms with 16 bins, and the colour histogram of each denomination varies greatly according to the dominant colour of banknote. For example, the colour histogram of 100 NZD has a very high value in the first bin, whereas the colour histograms of the other denominations all have a low value in the first bin. Accordingly, the first bin is helpful for classifying the 100 NZD from the others. Such distinguishable bins in the colour histogram have high covariance. Since PCA features are more reliable to the original features of high covariance, such distinguishable bins therefore enable the PCA-based colour feature to be an effective feature in this research. Correspondingly, the PCA-based colour feature reaches an average F-measure of 0.9940 when using the banknote images captured by webcam for training. However, we notice that the average F-measure of the PCA-based colour feature drops when the training set is changed to the scanned banknote images. Specifically, it drops to 0.9575 with the MDC and to 0.9759 with the BPNN. The intrinsic cause of this phenomenon is the *hue* differences between the colour of the scanned banknote image and the colour of the banknote images captured by webcam. For example, the *hue* of the scanned 20 NZD

image is primarily concentrated in the sixth bin, the seventh bin, and the eighth bin, as shown in Figure 4.3. By contrast, when the banknotes are captured by webcam, the *hue* of 20 NZD image not only focuses on the sixth bin, the seventh bin, and the eighth bin, but also has the highest value in the fifth bin, as shown in Appendix A. Even though the colour feature is part of the composite feature, the PCA-based composite feature is still capable of maintaining an average F-measure of 0.9960 when using the scanned banknote images for training, which benefits from the contribution from the texture feature part of the composite feature.

- Hypothesis 3.2 that the BPNN classifier would generate better recognition results proves true by the experiments.

Table 5.5 The comparison of the average F-measure between the classifiers

Combinations	Classifiers	
	MDC	BPNN
Scan_Colour_PCA	0.9575	0.9759
Scan_LBP_PCA	0.4216	0.9025
Scan_Colour+LBP_PCA	0.9960	0.9960
Scan_Colour_LDA	0.8695	0.9534
Scan_LBP_LDA	0.9820	0.9641
Scan_Colour+LBP_LDA	0.7994	0.9739
Webcam_Colour_PCA	0.9940	0.9940
Webcam_LBP_PCA	0.9133	0.8839
Webcam_Colour+LBP_PCA	0.9960	0.9980
Webcam_Colour_LDA	0.9920	0.9920
Webcam_LBP_LDA	0.9920	0.9920
Webcam_Colour+LBP_LDA	0.9441	0.9940

We have found that when using the scanned banknote images for training, the performance of the BPNN classifier is better than the MDC, except for the Combination No.11 (LBP_LDA_BPNN) whose average F-measure is lower than the Combination No.5 (LBP_LDA_MDC), and the Combination No.9 (Colour+LBP_PCA_BPNN) whose average F-measure is equal to the Combination No.3 (Colour+LBP_PCA_MDC). Specially, when using the scanned banknote images for training, the Combination No.7

(Colour_PCA_BPNN) outperforms the Combination No.1 (Colour_PCA_MDC); the Combination No.8 (LBP_PCA_BPNN) is superior to the Combination No.2 (LBP_PCA_MDC); the Combination No.10 (Colour_LDA_BPNN) is superior to the Combination No.4 (Colour_LDA_MDC); the Combination No.12 (Colour+LBP_LDA_BPNN) performs better than the Combination No.6 (Colour+LBP_LDA_MDC). In addition, we have found that when using the banknote images captured by webcam for training, the performance of the BPNN classifier is better than the MDC, except for the Combination No.8 (LBP_PCA_BPNN) whose average F-measure is lower than the Combination No.2 (LBP_PCA_MDC), the Combination No.7 (Colour_PCA_BPNN) whose average F-measure is equal to the Combination No.1 (Colour_PCA_MDC), the Combination No.10 (Colour_LDA_BPNN) whose average F-measure equals to the Combination No.4 (Colour_LDA_MDC), and the Combination No.11 (LBP_LDA_BPNN) whose F-measure is equal to the Combination No.5 (LBP_LDA_MDC). Specially, when using the banknote images captured by webcam for training, the Combination No.12 (Colour+LBP_LDA_BPNN) outperforms the Combination No.6 (Colour+LBP_LDA_MDC); the Combination No.9 (Colour+LBP_PCA_BPNN) is superior to the Combination No.12 (Colour+LBP_PCA_MDC). Moreover, for the PCA-based LBP feature of the scanned banknotes, i.e. the Combination No.2 (LBP_PCA_MDC) and the Combination No.8 (LBP_PCA_BPNN) in Table 5.1, the MDC cannot recognise 5 NZD or 50 NZD, while the BPNN is able to recognise all denominations with an average F-measure of 0.9025. Thus, overall, the BPNN classifier has better performance than the MDC in the experiments.

The MDC has more influence on the elements of a feature vector that has a large magnitude in a change, which means the MDC is only good at dealing with certain features. By contrast, BPNN classifiers treat every element in an input feature vector without bias or favour, assigning each input node (element) a random weight initially, so each element has an equivalent degree of influence on the network. This enables BPNN to handle a variety of features. This is the reason why the BPNN outperforms the MDC in the experiments.

- The core research question – how to achieve real-time recognition of the new Series 7 New Zealand banknotes under camera using digital image processing?

As far as the 24 contrast experiments in this research concerned, the training set composed of the banknote images captured by webcam results in better recognition results than the training set made up of the scanned banknote images. This suggests that the best way to produce training images is to simulate the testing environment as much as possible, for achieving real-time recognition. As for the 12 combined methods, the highest average F-measure is 0.9980 obtained by the Combination No.9 (Colour+LBP_PCA_BPNN) when using the banknote images captured by webcam for training. Moreover, the Combination No.9 (Colour+LBP_PCA_BPNN) is still able to achieve an average F-measure of 0.9960 when using the scanned banknote images for training. Thus, the Combination No.9 (Colour+LBP_PCA_BPNN) is the overall winner in this research, with the average time taken for each banknote to be recognised is 0.4249s.

5.3 Discussion

The most noteworthy finding in this research is the robustness of the PCA-based composite feature containing the colour and texture elements that works very well with either the MDC or the BPNN for Series 7 New Zealand banknote real-time recognition. This finding supports the opinion that a perfect feature extractor would produce a representation that makes the job of the classifier of little account (Duda et al., 2012). Thus, it is critical to extract proper features for developing an efficient banknote real-time recognition system.

A histogram is a straightforward approach to represent image features. It is widely used as a solution to image classification problems, by virtue of its translation, rotation, and scaling invariant properties. As the most direct visual feature, colour has always been one of the main features extensively utilised in banknote classification, and especially, the most common representation of the colour feature is a colour histogram. A colour histogram is simply a statistical histogram showing the proportion of pixels in a

particular colour to the pixels in the other colours in the whole image. In this thesis, an HSV colour space-based colour histogram is adopted. The motivation to adopt HSV rather than other colour space like RGB or YUV is that HSV colour space has the advantage of being closer to the human conceptual understanding of colours, and can separate chromatic and achromatic components (Pathrabe & Karmore, 2011). More importantly for paper currency in most countries, its design is based on the rule that human eyes are supposed to be able to distinguish different denominations by features including colour. Correspondingly, the presumption that HSV is the appropriate colour space for investigating the colour features of Series 7 New Zealand banknotes has been confirmed by our experiments. The colour feature based on HSV individually investigated in this research reaches an average F-measure of no less than 0.9920 when the training set is composed of the banknote images captured by using a webcam, either working with the MDC or the BPNN. In some other banknote recognition studies, HSV colour space is appropriately utilised to extract a colour feature of banknotes, which succeeds in obtaining a satisfactory result (Aasma & Asma, 2016; Pathrabe & Karmore, 2011; Zeggeye & Assabie, 2016). As well, HSV colour space is proved to be efficient in the selection of discriminative colours of banknotes as well. García-Lamont, Cervantes, López and Rodríguez (2013) performed contrast experiments on the colour feature selection of Mexican banknotes in two different colour spaces – RGB and HSV, with the outcome that the selected colour feature in the HSV space is able to enhance the recognition rate.

The chosen colour space needs to be quantised to construct colour histograms with low dimensions. The level of colour space quantisation depends on the specific subject to study. Basically, the computational complexity and storage space are increased non-linearly with the increasing number of quantised colours, but excessive colour quantisation reduces the visual quality of an image (Zeng & Zhou, 2008). One simple colour quantisation method is uniform quantisation of each colour channel for every pixel. However, uniform quantisation not only disregards the interdependency among pixels but also takes no notice of any actual colour distributions in a given image set. As

far as HSV colour space is concerned, as the *hue* (H) is more important in the human visual system than the *saturation* (S) and the *value* (V) components (Jeong, Won, & Gray, 2004), it is reasonable to assign more bins to the *hue* than to the other components. Such a non-uniform quantisation approach is found in many image retrieval-related studies, where the *hue* is quantised into 16 bins, the *saturation* is quantised into four bins, and the *value* is quantised into four bins (Peng et al., 2012; Zhao et al., 2007). In the light of the effectiveness of the 16:4:4 colour quantisation approach, we take such an approach into account for constructing a promising colour histogram. On the other hand, since the HSV colour space is driven by the human vision system in a sense that humans describe colour by means of the *hue* and the *saturation* (Suhasini, Krishna, & Krishna, 2016), the *value* is ignored in this research. Thus, the unique non-uniform colour quantisation in HSV is piloted at the initial stage of the experiments, where the *hue* is divided into 16 partitions, and the *saturation* is split into four partitions. From the pilot test, the *saturation* is observed as the unnecessary component for an efficient colour histogram that is peculiar to Series 7 New Zealand banknotes. Moreover, the *hue* represents the most significant characteristic of a colour (Suhasini et al., 2016), and is insensitive to the change in illumination and camera direction (Jain & Johari, 2016). Eventually, only the *hue* in HSV is adopted for representing the single colour feature of a banknote for recognition, ensuring that the quantised colour components constitute histogram bins effectively.

Accordingly, in the combinations where the colour feature is to be investigated individually in this research, i.e. the Combination No.1 (Colour_PCA_MDC), the Combination No.4 (Colour_LDA_MDC), the Combination No.7 (Colour_PCA_BPNN) and the Combination No.10 (Colour_LDA_BPNN), colour feature is represented by the colour histogram with 16 bins in HSV where the *hue* is quantised into 16 bins while getting rid of the *saturation* and the *value*. To our best knowledge, it is the first time such an HSV colour space quantisation approach, presented in this thesis, has applied to the field of banknote recognition. Other researchers make use of HSV colour space for banknote feature extraction in a different way where no quantisation involved, such as

perceiving the mean value of the *hue*, the *saturation*, and the *value* separately as the predominant colour feature of banknotes (Aasma & Asma, 2016), solely considering the mean value of the *hue* as the colour feature of paper currency (Murthy, Kurumathur, & Reddy, 2016; Zeggeye & Assabie, 2016).

Colour histogram indicates the global spread of colours regarding the banknote image, but it does not pay attention to the information on the spatial distribution of the image. To avoid the lack of notice on spatial information of banknotes, the texture depicted in the banknote image is also taken into consideration in this research. As for the texture feature of the obverse side of Series 7 New Zealand banknotes, there is the same design style but different content among banknotes of different denominations. To be specific, the images of key characters – Sir Edmund Hillary on 5 NZD, Kate Sheppard on 10 NZD, Queen Elizabeth II on 20 NZD, Sir Āpirana Ngata on 50 NZD, Lord Rutherford of Nelson on 100 NZD - are located in the same area of the banknotes; the background images - Aoraki/Mount Cook on 5 NZD, white camellia flowers on 10 NZD, New Zealand Parliament Buildings on 20 NZD, Porourangi Meeting House on 50 NZD and Nobel Prize medal on 100 NZD - are located in the same area of the banknotes; the images of colour-changing birds – Hoiho (yellow-eyed penguin) on 5 NZD, Whio (blue duck) on 10 NZD, Kārearea (New Zealand falcon) on 20 NZD, Kōkako (blue wattled crow) on 50 NZD and Mohua (yellowhead) on 100 NZD.

The LBP operator provides a good way of exploiting the spatial characteristics of the banknote image where each pixel is compared with its neighbours to generate its LBP code. Accordingly, the LBP histogram is used to describe the texture features of Series 7 New Zealand banknotes. In particular, considering the powerfulness of uniform LBP on reflecting the texture feature of an image (Ojala et al., 2002), uniform LBP based on a neighbour set of eight members on a circle with a radius of one pixel is adopted to create an efficient uniform LBP histogram with fewer bins. Moreover, for the purpose of highlighting the detailed texture information of Series 7 New Zealand banknotes, when investigating the uniform LBP feature of the banknote, we divide the banknote

image into 32 equal-sized blocks to construct the block uniform LBP histogram with 1888 (32×59) bins for the banknote. In our study, the LDA-based LBP feature described by block uniform LBP histograms perform well, with an average F-measure of no less than 0.9616 when using the scanned banknote images for training and an average F-measure of 0.9920 when using the images captured by webcam for training. A similar block-LBP method is discovered in the study undertaken by Guo et al. (2010), where a banknote image was divided into 300 (30×10) equal-sized blocks and 256 kinds of LBP codes were evenly separated into eight bins, thereby forming the final LBP feature vector of 9600 ($256/8 \times 300$) dimensions for each banknote. Although uniform LBP was not employed in their study, the recognition rate was improved by employing the block-LBP method instead of employing the traditional LBP operator (Guo et al., 2010).

In many studies, both colour and texture features are extracted as a composite feature of banknotes for recognition. In our research, the composite feature containing colour and texture elements of Series 7 New Zealand banknotes is also analysed for recognition, which is described by the colour-LBP histogram with 123 bins. Specifically, for each banknote, the colour part of the composite feature is represented by the colour histogram with 64 bins in an HSV colour space where the *hue* is quantised into 16 bins, the *saturation* is quantised into four bins, and the *value* is discarded, while the LBP part of the composite feature is represented by the LBP histogram with 59 bins. By doing so, the dimensionality imbalance between the colour histogram part and the LBP histogram part in the colour-LBP histogram is minimised, so that the colour part and LBP part are treated fairly, having almost the same impact on the composite feature. In the experiments, the PCA-based composite feature shows its robustness on real-time banknote recognition, achieving an average F-measure of no less than 0.996 in the two groups of training sets. It is important to mention that in this study it is assumed that there are no illumination variations during experimentation. In other words, the images of paper currency are captured under the same illumination condition, for both training and testing phases.

The effectiveness of the composite feature containing color and texture elements has also been reported in other banknote recognition studies. In Chambers' (2012) study, the two-part feature vector composed of both colour and texture features showed significant improvement for recognition accuracy, in comparison with individually adopting the colour or the texture feature; the composite feature delivered the optimal performance at 98.6% recognition accuracy when using the feed-forward neural network as the classifier trained by Bayesian back-propagation regulation learning algorithm. Likewise, García-Lamont et al. (2012) also demonstrated that by combining colour and texture features of banknotes to analyse, the recognition performance was improved as compared with single colour or texture analysis; the composite feature obtained 98.95% recognition rate when working with LVQ networks.

Compared with the composite feature suggested by Chambers (2012) and the composite feature presented by García-Lamont et al. (2012), the composite feature investigated in our research has its highlights. In Chambers' study, the color part of the composite feature only concerns the intensity histogram based on grey-scale image, whereas in our research, the color part of the composite feature is described by the color histogram based on a colorful image, which is truly representative of the color information of a banknote. Different from the research undertaken by García-Lamont et al. (2012) where the color feature part is extracted based on RGB color space, the HSV color space is adopted to characterise the color of the banknote in our research. Additionally, García-Lamont et al. (2012) employed traditional LBP operator for extracting the texture feature part of the composite feature; by contrast, we take advantage of the powerfulness of uniform LBP to present the texture feature of the banknote. The combined use of uniform LBP and PCA effectively reduces the dimensionality of the composite feature vector, thereby speeding up the recognition time to 0.4249s per note, which is faster than the approach proposed by García-Lamont et al. (2012) that requires 0.9641s for recognising a banknote.

In summary, on the one hand, although the colour feature is reported to be incapable of

providing sufficient classification information (Hassanpour & Farahabadi, 2009; Solymár et al., 2011), in our research, by making the most of HSV colour space, the PCA-based colour feature is able to reach an average F-measure of no less than 0.9575 in the two groups of the training sets, no matter working with the MDC or the BPNN. On the other hand, it is meaningful to notice that the PCA-based composite feature containing both the colour and texture elements of the banknote is competent in obtaining an average F-measure of no less than 0.9960 in the two groups of training sets, no matter working with the MDC or the BPNN. Thus, the composite feature along with the BPNN as the classifier is proposed for approaching Series 7 New Zealand banknote real-time recognition. Since the composite feature investigated in this research takes account of both colour and texture features of the banknote, it could potentially serve as a fundamental solution to the real-time recognition of other paper currencies. In particular, when the colour feature is not enough for discriminating banknotes, such as the banknotes of different denominations having the same colour, the texture feature components enable the composite feature to be a very promising feature to be exploited for real-time recognition tasks.

Chapter 6 Conclusion

This chapter will conclude the thesis by summarising the current research, underlining the novelty of the research, highlighting the significance of the research, and finally identifying its limitations and future research directions.

6.1 Summary

It is not easy to answer how to achieve real-time recognition of the new Series 7 New Zealand banknotes under camera using digital image processing. It depends on multiple factors, such as constitution of the training set, features to be extracted and classifiers to be employed. Based on the reviewed literature on banknote recognition, this thesis presents empirical approaches for Series 7 New Zealand banknote real-time recognition. Two hundred and fifty banknote samples of different denominations are used for training, including 50 each of 5 NZD, 10 NZD, 20 NZD, 50 NZD and 100 NZD. Two groups of training sets are generated by processing the training samples in different ways. One training set is composed of the banknote images produced by scanning banknote samples, and the other training set consists of the banknote images generated through webcam capturing banknote samples. Three kinds of features of the banknotes, namely, colour feature, texture feature, and composite feature containing both the colour and texture elements, are studied, along with two classifiers - the MDC and the BPNN. Effective dimensionality reduction algorithms, PCA and LDA, are utilised in the experiments to build up computing speed, thereby adapting to real-time recognition. A total of 24 contrast experiments have been conducted using various combined methods. Fifty banknote samples participate in the testing stage, with 10 banknote samples for each denomination. The experimental results are analysed to investigate the relationship between the experimental subject and various factors.

The results suggest that the training sets should be generated in the way that simulates the testing environment to the maximum extent, to achieve better recognition results. With a high-recognition rate as a prerequisite for evaluating the performance of the

combined methods investigated in this research, the time cost of recognition is also calculated as a secondary performance indicator for real-time recognition. Finally, the PCA-based composite feature along with the BPNN is the overall winner that yields an average F-measure of 0.9960 when using the scanned banknote images for training and 0.9980 when using the banknote images captured by webcam for training, with an average time taken of 0.4249 seconds for each note to be recognised, and is therefore proposed to deal with Series 7 New Zealand banknote real-time recognition. Specifically, for the composite feature, the colour part is described by the HSV colour histogram with 64 bins where the *hue* is quantised into 16 intervals, the *saturation* is quantised into four intervals, and the *value* is discarded, while the LBP part is described by the uniform LBP histogram with 59 bins. PCA is then used to reduce the dimensionality of the composite feature vector, meanwhile retaining 95% of the information of the composite feature. Finally, a three-layer back-propagation neural network with sigmoid activation function trained by the Rprop algorithm is employed for recognition, which is comprised of four input nodes corresponding to the four dimensions of the PCA-based composite feature, five hidden nodes and five output nodes corresponding to the five different denominations \$5, \$10, \$20, \$50 and \$100.

6.2 Novelty of the research

To the best of our knowledge, it is the first time that the HSV colour quantisation approach presented in this thesis has appeared in the field of banknote recognition. In addition, no researchers have applied the block uniform LBP histogram to paper currency recognition before us. Moreover, the approach of extracting the composite features of banknotes proposed in this thesis is also the innovation point of the research, since it appears in the field of banknote recognition for the first time. On the other hand, most research work in the banknote recognition field has concentrated on the static recognition approach where the images of banknotes being tested are obtained and stored in a folder before the testing stage, and the images are read one by one for recognition during the testing stage. By contrast, this research project makes use of a webcam to capture the images of banknotes for testing to achieve banknote real-time

recognition.

6.3 Significance of the research

Paper currency-related transaction machines, such as ATMs and self-service payment kiosks, are supposed to have a classification function in their system for determining the correct denomination of banknotes at the time of a transaction. The outcome of this research project fills the gap that has hitherto existed in banknote recognition literature – the recognition of Series 7 New Zealand banknotes. The proposed method can be considered to apply to the banknote transaction machines to achieve the real-time recognition of the denominations of Series 7 New Zealand banknotes.

This research project is also potentially beneficial to blind or visually impaired people. Currency is the most common medium of exchange in human society. However, for certain people, such as the blind or the visually impaired, how do they identify different face values of currency for normal usage? Under this scenario, the camera-based classification approaches for Series 7 New Zealand banknotes presented in this thesis can help with the future development of a prototype helping the blind to recognise the denominations of Series 7 New Zealand banknotes.

6.4 Limitation and future work

Although an effective method for Series 7 New Zealand banknote real-time recognition is proposed in this thesis, a few limitations in this research should be pointed out. The primary limitation is with sample size. Strictly speaking, the number of banknote samples collected for the experiments is still relatively small if the proposed method is to be put to practical use in the market. In addition, Series 7 New Zealand banknotes \$5 and \$10 were released in October 2015, and the remaining three denominations were released in May 2016, indicating that the banknotes collected for experimentation are relatively new. As a result of the lack of worn or wrinkled banknote samples, the proposed method may not be able to precisely recognise worn-out Series 7 New

Zealand banknotes. In future, the research could be improved in the following aspects so it is capable of working in a more complex context.

Firstly, more samples of Series 7 New Zealand paper currency would be used in future, so as to increase the size of both training and testing sets. More training samples are helpful for neural network classifiers to train a better network, and more statistical data can be obtained from the experiments for analysis when more testing samples are provided. As well, as time goes by, the worn or wrinkled Series 7 New Zealand banknotes would appear in market circulation. Those banknotes need to be collected at that time for study, to equip the current system with the ability to recognise their denominations.

Secondly, a banknote detection module would be added to auto-locate the banknote under a camera. On this basis, rotation-invariant LBP can be employed to achieve rotation invariant banknote recognition. Thus, not only would the conditions for the placement of banknotes under a camera become less strict, but also the dimensionality of the LBP feature vector would be reduced, thereby lowering computational cost. The original LPB operator generates 256 kinds of different output values corresponding to different binary patterns, and an LBP histogram of an image will be a 256-dimensional vector, which is quite a large data volume. By applying uniform LBP, the dimensionality of its vector can be reduced from 256 to 59, assigning each uniform pattern to a separate bin and all non-uniform patterns to a single bin. By applying rotation invariant LBP, 256 patterns can be simplified to 36 rotation-invariant patterns, thereby reducing the dimensionality of its vector to 36. It is worth noting that the dimensionality of its vector can be reduced to nine by using the combined use of uniform patterns and rotation invariant descriptors, choosing nine uniform patterns out of the 36 rotation invariant patterns and merging the remaining 27 under the miscellaneous label.

Thirdly, the proposed approach in this thesis potentially paves the way for

multi-currency real-time recognition, as taken both colour and texture features of the banknote into consideration. In future, other paper currencies would be involved in research project in order to validate the general applicability of the proposed approach.

Last, but not the least, deep learning, the most cutting-edge technology in machine learning would be a prospective trend for future research.

References

- Aasma, S., & Asma, N. (2016). Indian Paper Currency Denomination Recognition By Image Processing. *International Journal for Innovations in Engineering, Science and Management*, 4(4).
- Ahangaryan, F. P., Mohammadpour, T., & Kianisarkaleh, A. (2012). Persian banknote recognition using wavelet and neural network. In *International Conference on Computer Science and Electronics Engineering*, (Vol. 3, pp. 679-684). IEEE.
- Ali, A., & Manzoor, M. (2013). Recognition System for Pakistani Paper Currency. *World Applied Sciences Journal*, 28(12), 2069-2075.
- Althafiri, E., Sarfraz, M., & Alfarras, M. (2012). Bahraini paper currency recognition. *Journal of Advanced Computer Science and Technology Research*, 2(2), 104-115.
- Bender, K. (2006). *Moneymakers*. Berlin: John Wiley and Sons.
- Bezdek, J. C. (2013). *Pattern recognition with fuzzy objective function algorithms*. Springer Science & Business Media.
- Bharkad, A. A. S. S. (2013). Survey of currency recognition system using image processing. *International Journal of Computational Engineering Research*, 03(7).
- Bishop, C. M. (2006). *Pattern recognition and machine learning*: springer.
- Bow, S. T. (Ed.). (2002). *Pattern recognition and image preprocessing*. CRC Press.
- Chae, S. H., Kim, J. K., & Pan, S. B. (2009). A study on the Korean banknote recognition using RGB and UV information. *Communication and Networking*, 477-484.
- Chakraborty, T., Nalawade, N., Manjre, A., Sarawgi, A., & Chaudhari, P. P. (2016). Review of various image processing techniques for currency note authentication. *Int. J. Comput. Eng. Res. Trends*, 3(3), 119-122.
- Chambers, J. (2012). *Digital currency forensics, Masters Thesis, Auckland University of Technology*.
- Chapelle, O., Zien, A., & Schölkopf, B. (Eds.). (2006). *Semi-supervised learning*. Cambridge, MA: MIT Press.

- Cherkassky, V., Friedman, J. H., & Wechsler, H. (Eds.). (2012). *From statistics to neural networks: theory and pattern recognition applications* (Vol. 136). Springer Science & Business Media.
- Choi, E., Lee, J., & Yoon, J. (2006). Feature extraction for bank note classification using wavelet transform. In *18th International Conference on Pattern Recognition*, (Vol 2, pp. 934-937). IEEE.
- Cios, K. J., Swiniarski, R. W., Pedrycz, W., & Kurgan, L. A. (2007). Feature Extraction and Selection Methods. In *Data Mining* (pp. 133-233). Springer US.
- Cristea, P. D. (2009). Application of neural networks in image processing and visualization. In *GeoSpatial Visual Analytics* (pp. 59-71). Springer Netherlands.
- De Sa, J. M. (2012). *Pattern recognition: concepts, methods and applications*. Springer Science & Business Media.
- Devi, O. R., Prasad, J., Rao, B., Babu, M., Kumar, L., Prakash, P., Rani, K. (2016). Survey on paper currency recognition system. *Int. J. Emerg. Trends Technol. Comput. Sci*, 5(2), 105-108.
- Dittimi, T. V., Hmood, A. K., & Suen, C. Y. (2017). Multi-class SVM based gradient feature for banknote recognition. In *IEEE International Conference on Industrial Technology*, (pp. 1030-1035). IEEE.
- Duda, R. O., Hart, P. E., & Stork, D. G. (2012). *Pattern classification*. John Wiley & Sons.
- Fukunaga, K. (2013). *Introduction to statistical pattern recognition*. Academic press.
- García-Lamont, F., Cervantes, J., & López, A. (2012). Recognition of Mexican banknotes via their color and texture features. *Expert Systems with Applications*, 39(10), 9651-9660.
- García-Lamont, F., Cervantes, J., López, A., & Rodríguez, L. (2013). Classification of Mexican Paper Currency Denomination by Extracting Their Discriminative Colors. In *Mexican International Conference on Artificial Intelligence* (pp. 403-412). Berlin: Springer, Heidelberg.
- Grijalva, F., Rodriguez, J. C., Larco, J., & Orozco, L. (2010). Smartphone recognition of the US banknotes' denomination, for visually impaired people. In

- ANDESCON, 2010 IEEE* (pp. 1-6). IEEE.
- Guillaumin, M., Verbeek, J., & Schmid, C. (2010). Multimodal semi-supervised learning for image classification. In *23rd IEEE Conference on Computer Vision & Pattern Recognition* (pp. 902-909). IEEE Computer Society.
- Gunaratna, D. A. K. S., Kodikara, N. D., & Premaratne, H. L. (2008). ANN based currency recognition system using compressed gray scale and application for Sri Lankan currency notes-SLCRec. *Proceedings of World academy of science, engineering and technology*, 35, 235-240.
- Guo, J., Zhao, Y., & Cai, A. (2010). A reliable method for paper currency recognition based on LBP. In *2nd IEEE International Conference on Network Infrastructure and Digital Content*, (pp. 359-363). IEEE.
- Guyon, I., & Elisseeff, A. (2006). An introduction to feature extraction. In *Feature extraction* (pp. 1-25). Springer Berlin Heidelberg.
- Hassanpour, H., & Farahabadi, P. M. (2009). Using Hidden Markov Models for paper currency recognition. *Expert Systems with Applications*, 36(6), 10105-10111.
- Hassanpour, H., Yaseri, A., & Ardeshiri, G. (2007). Feature extraction for paper currency recognition. In *9th International Symposium on Signal Processing and Its Applications*, (pp. 1-4). IEEE.
- He, J. B., Zhang, H. M., Liang, J., Jin, O., & Li, X. (2015). Paper Currency Denomination Recognition Based on GA and SVM. In *Chinese Conference on Image and Graphics Technologies* (pp. 366-374). Berlin: Springer, Heidelberg.
- Hinton, G. E. (2002). How neural networks learn from experience. In T. A. Polk & C. M. Seifert (Eds.), *Cognitive modeling* (pp.181-195). Cambridge, MA: MIT Press.
- Hinwood, A., Preston, P., Suaning, G. J., & Lovell, N. H. (2006). Bank note recognition for the vision impaired. *Australasian Physical & Engineering Science in Medicine*, 29(2), 229-233.
- Hlaing, K. N. N., & Gopalakrishnan, A. K. (2016). Myanmar paper currency recognition using GLCM and k-NN. In *Second Asian Conference on Defence Technology (ACDT)*, (pp. 67-72). IEEE.
- Jahangir, N., & Chowdhury, A. R. (2007). Bangladeshi banknote recognition by neural

- network with axis symmetrical masks. In *10th International Conference on Computer and Information Technology*, (pp. 1-5). IEEE.
- Jain, A. K., Duin, R. P. W., & Mao, J. (2000). Statistical pattern recognition: A review. *IEEE Transactions on pattern analysis and machine intelligence*, 22(1), 4-37.
- Jain, R., & Johari, P. K. (2016). An improved approach of CBIR using Color based HSV quantization and shape based edge detection algorithm. In *IEEE International Conference on Recent Trends in Electronics, Information & Communication Technology (RTEICT)*, (pp. 1970-1975). IEEE.
- Jankowski, N., & Grabczewski, K. (2006). Learning machines. In *Feature Extraction* (pp. 29-64). Berlin: Springer, Heidelberg.
- Jeong, S., Won, C. S., & Gray, R. M. (2004). Image retrieval using color histograms generated by Gauss mixture vector quantization. *Computer Vision and Image Understanding*, 94(1), 44-66.
- Kamal, S., Chawla, S. S., Goel, N., & Raman, B. (2015). Feature extraction and identification of Indian currency notes. In *Fifth National Conference on Computer Vision, Pattern Recognition, Image Processing and Graphics*, (pp. 1-4). IEEE.
- Khashman, A., & Sekeroglu, B. (2005). Multi-banknote identification using a single neural network. In *International Conference on Advanced Concepts for Intelligent Vision Systems* (pp. 123-129). Berlin: Springer, Heidelberg.
- Kurita, T. (2014). Principal component analysis (PCA). In *Computer Vision* (pp. 636-639). US: Springer.
- Kwak, N., & Choi, C. H. (2002). Input feature selection for classification problems. *IEEE Transactions on Neural Networks*, 13(1), 143-159.
- Lamsal, S., & Shakya, A. (2015). Counterfeit Paper Banknote Identification Based on Color and Texture. In *Proceedings of IOE Graduate Conference* (pp. 160-168).
- Langwasser, K. (2014). Designing New Zealand's new banknote series. *Reserve Bank of New Zealand Bulletin*, 77, 1-9.
- LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, 521(7553), 436-444.

- Lee, J. K., Jeon, S. G., & Kim, I. H. (2004). Distinctive point extraction and recognition algorithm for various kinds of euro banknotes. *International Journal of Control, Automation, and Systems*, 2(2), 201-206.
- Li, S., Liu, C., & Wang, Y. (Eds.). (2014). *6th Chinese Conference on Pattern Recognition*, (Vol. 483). Springer.
- Liu, H., & Motoda, H. (1998). Less is more. In *Feature Extraction, Construction and Selection* (pp. 3-12). Springer US.
- Liu, H., & Motoda, H. (Eds.). (2007). *Computational methods of feature selection*. CRC Press.
- Luo, D., Ding, C. H., & Huang, H. (2011). Linear Discriminant Analysis: New Formulations and Overfit Analysis. In *AAAI*.
- Maestro, B. (1993). *The story of money*. Houghton Mifflin Harcourt.
- Marsland, S. (2015). *Machine learning: an algorithmic perspective*. CRC press.
- McCulloch, W. S., & Pitts, W. (1943). A logical calculus of the ideas immanent in nervous activity. *The bulletin of mathematical biophysics*, 5(4), 115-133.
- Murthy, S., Kurumathur, J., & Reddy, B. R. (2016). Design and implementation of paper currency recognition with counterfeit detection. In *Online International Conference on Green Engineering and Technologies (IC-GET)*, (pp. 1-6). IEEE.
- Nanni, L., Lumini, A., & Brahmam, S. (2012). Survey on LBP based texture descriptors for image classification. *Expert Systems with Applications*, 39(3), 3634-3641.
- Ni'am, B. A. A. A., Faisal, M., & Arif, Y. M. (2014). Identification of Nominal Value and Authenticity of Rupiah Using Support Vector Machine. *IEESE International Journal of Science and Technology*, 3(1), 13-20.
- Ojala, T., Pietikainen, M., & Harwood, D. (1994). Performance evaluation of texture measures with classification based on Kullback discrimination of distributions. In *Proceedings of the 12th IAPR International Conference on Pattern Recognition*, (Vol. 1, pp. 582-585). IEEE.
- Ojala, T., Pietikainen, M., & Maenpaa, T. (2002). Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Transactions on pattern analysis and machine intelligence*, 24(7), 971-987.

- Panchal, F. S., & Panchal, M. (2014). Review on methods of selecting number of hidden nodes in artificial neural network. *International Journal of Computer Science and Mobile Computing*, 3(11), 455-464.
- Pathrabe, T., & Karmore, S. (2011). A novel approach of embedded system for Indian paper currency recognition. *Int. J. Comput. Trends Technol*, 1, 152-156.
- Peng, X., Zhu, Y., & Lin, Z. (2012). New method of image retrieval based on interest image region blocks and comprehensive features. *Wuhan University Journal of Natural Sciences*, 17(1), 55-60.
- Pham, T. D., Park, Y. H., Kwon, S. Y., Park, K. R., Jeong, D. S., & Yoon, S. (2016). Efficient banknote recognition based on selection of discriminative regions with one-dimensional visible-light line sensor. *Sensors*, 16(3), 328.
- Pietikäinen, M., Ojala, T., & Xu, Z. (2000). Rotation-invariant texture classification using feature distributions. *Pattern Recognition*, 33(1), 43-52.
- Pratt, W. K. (2013). *Introduction to digital image processing*. CRC Press.
- Preston, K., & Carvalko, J. R. (1972). On determining optimum simple Golay marking transformations for binary image processing. *IEEE Transactions on Computers*, 21(12), 1430-1433.
- Rahman, S., Banik, P., & Naha, S. (2014). LDA based paper currency recognition system using edge histogram descriptor. In *17th International Conference on Computer and Information Technology (ICCIT)*, (pp. 326-331). IEEE.
- Reserve Bank of New Zealand. (n.d.). *Security features of New Zealand's banknotes*. Retrieved from <http://www.rbnz.govt.nz/notes-and-coins/notes/security-features-of-new-zealands-banknotes>
- Riedmiller, M., & Braun, H. (1993). A direct adaptive method for faster backpropagation learning: The RPROP algorithm. In *IEEE International Conference on Neural Networks*, (pp. 586-591). IEEE.
- Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1985). *Learning internal representations by error propagation*. California Univ San Diego La Jolla Inst for Cognitive Science.
- Samarasinghe, S. (2006). *Neural networks for applied sciences and engineering: from*

fundamentals to complex pattern recognition. CRC Press.

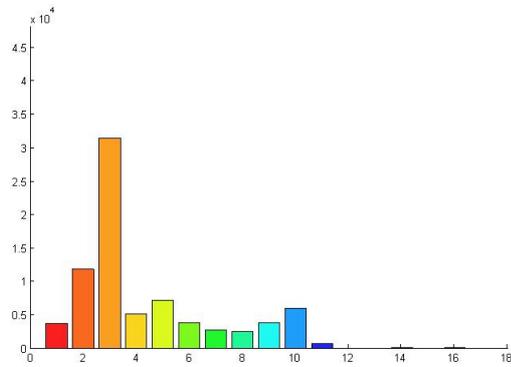
- Sarfraz, M. (2015). An Intelligent Paper Currency Recognition System. *Procedia Computer Science*, 65, 538-545.
- Sargano, A. B., Sarfraz, M., & Haq, N. (2014). An intelligent system for paper currency recognition with robust features. *Journal of Intelligent & Fuzzy Systems*, 27(4), 1905-1913.
- Sawant, K., & More, C. (2016). Currency Recognition Using Image Processing and Minimum Distance Classifier Technique. *International Journal of Advanced Engineering Research and Science*, 3(9), 1-8.
- Schalkoff, R. J. (1992). *Pattern recognition*. John Wiley & Sons, Inc..
- Schmidhuber, J. (2015). Deep learning in neural networks: An overview. *Neural networks*, 61, 85-117.
- Semary, N. A., Fadl, S. M., Essa, M. S., & Gad, A. F. (2015). Currency recognition system for visually impaired: Egyptian banknote as a study case. In *5th International Conference on Information & Communication Technology and Accessibility*, (pp. 1-6). IEEE.
- Sethi, I. K., & Jain, A. K. (Eds.). (2014). *Artificial neural networks and statistical pattern recognition: old and new connections* (Vol. 11). Elsevier.
- Smith, J. R., & Chang, S. F. (1996). Tools and Techniques for Color Image Retrieval. In *Storage and Retrieval for Image and Video Databases (SPIE)* (Vol. 2670, pp. 2-7).
- Smith, L. I. (2002). A tutorial on principal components analysis. *Cornell University, USA*, 51(52), 65.
- Sokolova, M., & Lapalme, G. (2009). A systematic analysis of performance measures for classification tasks. *Information Processing & Management*, 45(4), 427-437.
- Solymár, Z., Stubendek, A., Radványi, M., & Karacs, K. (2011, August). Banknote recognition for visually impaired. In *20th European Conference on Circuit Theory and Design*, (pp. 841-844). IEEE.
- Sonka, M., Hlavac, V., & Boyle, R. (2014). *Image processing, analysis, and machine vision*. Cengage Learning.

- Suhasini, P. S., Krishna, K. S. R., & Krishna, I. M. (2017). Content based Image Retrieval based on Different Global and Local Color Histogram Methods: A Survey. *Journal of The Institution of Engineers: Series B*, 98(1), 129-135.
- Takeda, F., Nishikage, T., & Omatu, S. (1999). Banknote recognition by means of optimized masks, neural networks and genetic algorithms. *Engineering Applications of Artificial Intelligence*, 12(2), 175-184.
- Takeda, F., & Omatu, S. (1995). High speed paper currency recognition by neural networks. *IEEE Transactions on Neural Networks*, 6(1), 73-77.
- Tarnoff, B. (2011). *Moneymakers: the wicked lives and surprising adventures of three notorious counterfeiters*. New York: The Penguin Press HC.
- Umbaugh, S. E. (2005). *Computer imaging: digital image analysis and processing*. CRC press.
- Vishnu, R., & Omman, B. (2014). Principal component analysis on Indian currency recognition. In *International Conference on Computer and Communication Technology (ICCCT)*, (pp. 291-296). IEEE.
- Wang, L., & He, D. C. (1990). Texture classification using texture spectrum. *Pattern Recognition*, 23(8), 905-910.
- Wang, X., & Tang, X. (2004). Experimental study on multiple LDA classifier combination for high dimensional data classification. *Multiple Classifier Systems*, 3077, 344-353.
- Yan, W. Q., & Chambers, J. (2013). An empirical approach for digital currency forensics. In *IEEE International Symposium on Circuits and Systems (ISCAS)*, (pp. 2988-2991). IEEE.
- Yan, W. Q., Chambers, J., & Garhwal, A. (2015). An empirical approach for currency identification. *Multimedia Tools and Applications*, 74(13), 4723-4733.
- Ye, F., Shi, Z., & Shi, Z. (2009). A comparative study of PCA, LDA and Kernel LDA for image classification. In *International Symposium on Ubiquitous Virtual Reality*, (pp. 51-54). IEEE.
- Youn, S., Choi, E., Baek, Y., & Lee, C. (2015). Efficient multi-currency classification of CIS banknotes. *Neurocomputing*, 156, 22-32.

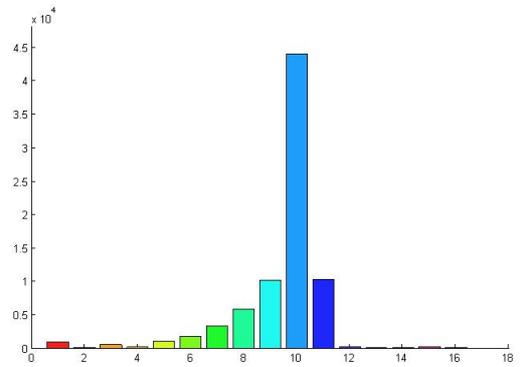
- Zeggeye, J. F., & Assabie, Y. (2016). Automatic recognition and counterfeit detection of Ethiopian paper currency. *International Journal of Image, Graphics and Signal Processing*, 8(2), 28.
- Zeng, Z. Y., & Zhou, L. H. (2008). A novel region-based image retrieval technique using spacial color distribution feature. In *2nd International Symposium on Systems and Control in Aerospace and Astronautics*, (pp. 1-5). IEEE.
- Zhang, E. H., Jiang, B., Duan, J. H., & Bian, Z. Z. (2003). Research on paper currency recognition by neural networks. In *International Conference on Machine Learning and Cybernetics*, (Vol. 4, pp. 2193-2197). IEEE.
- Zhao, J., Yan, D. M., & Zhang, Y. K. (2007). Two-layer image retrieval method based on wavelet and local color spatial features. In *International Conference on Wavelet Analysis and Pattern Recognition*, (Vol. 1, pp. 254-259). IEEE.
- Zhu, X. (2011). Semi-supervised learning. In *Encyclopedia of machine learning* (pp. 892-897). Springer US.
- Zhu, X., & Goldberg, A. B. (2009). Introduction to semi-supervised learning. *Synthesis lectures on artificial intelligence and machine learning*, 3(1), 1-130.

Appendices

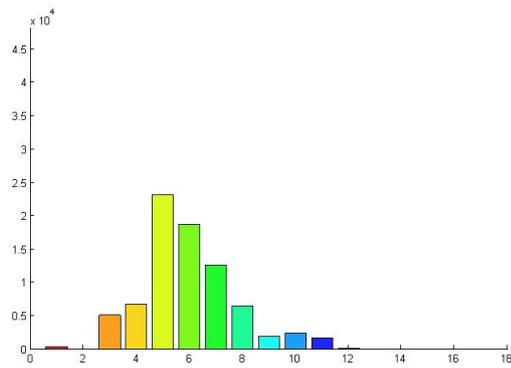
Appendix A: Colour histograms with 16 bins for the banknote images captured by webcam (the displayed colour of each column determined by the median value of each partition range)



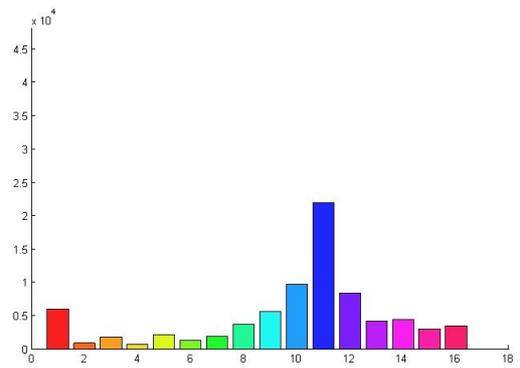
(a) 5 NZD



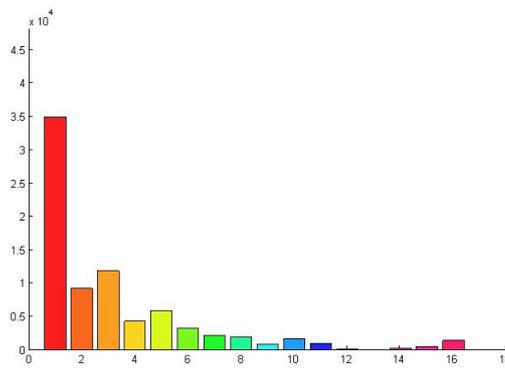
(b) 10 NZD



(c) 20 NZD



(d) 50 NZD



(e) 100 NZD

Appendix B: Recognition results when using the scanned banknote images for training

Confusion Matrix

Output Class	\$5	\$10	\$20	\$50	\$100	
\$5	100 20.00%	0 0.00%	21 4.20%	0 0.00%	0 0.00%	82.64% 17.36%
\$10	0 0.00%	100 20.00%	0 0.00%	0 0.00%	0 0.00%	100% 0.00%
\$20	0 0.00%	0 0.00%	79 15.80%	0 0.00%	0 0.00%	100% 0.00%
\$50	0 0.00%	0 0.00%	0 0.00%	100 20.00%	0 0.00%	100% 0.00%
\$100	0 0.00%	0 0.00%	0 0.00%	0 0.00%	100 20.00%	100% 0.00%
	100% 0.00%	100% 0.00%	79.00% 21.00%	100% 0.00%	100% 0.00%	95.80% 4.20%
Target Class	\$5	\$10	\$20	\$50	\$100	

(a) Combination No.1

Confusion Matrix

Output Class	\$5	\$10	\$20	\$50	\$100	
\$5	0 0.00%	0 0.00%	0 0.00%	0 0.00%	0 0.00%	NaN% NaN%
\$10	90 18.00%	100 20.00%	15 3.00%	86 17.20%	30 6.00%	31.15% 68.85%
\$20	10 2.00%	0 0.00%	85 17.00%	0 0.00%	0 0.00%	89.47% 10.53%
\$50	0 0.00%	0 0.00%	0 0.00%	0 0.00%	0 0.00%	NaN% NaN%
\$100	0 0.00%	0 0.00%	0 0.00%	14 2.80%	70 14.00%	83.33% 16.67%
	0.00% 100%	100% 0.00%	85.00% 15.00%	0.00% 100%	70.00% 30.00%	51.00% 49.00%
Target Class	\$5	\$10	\$20	\$50	\$100	

(b) Combination No.2

Confusion Matrix

Output Class	\$5	\$10	\$20	\$50	\$100	
\$5	100 20.00%	0 0.00%	2 0.40%	0 0.00%	0 0.00%	98.04% 1.96%
\$10	0 0.00%	100 20.00%	0 0.00%	0 0.00%	0 0.00%	100% 0.00%
\$20	0 0.00%	0 0.00%	98 19.60%	0 0.00%	0 0.00%	100% 0.00%
\$50	0 0.00%	0 0.00%	0 0.00%	100 20.00%	0 0.00%	100% 0.00%
\$100	0 0.00%	0 0.00%	0 0.00%	0 0.00%	100 20.00%	100% 0.00%
	100% 0.00%	100% 0.00%	98.00% 2.00%	100% 0.00%	100% 0.00%	99.60% 0.40%
Target Class	\$5	\$10	\$20	\$50	\$100	

(c) Combination No.3

Confusion Matrix

Output Class	\$5	\$10	\$20	\$50	\$100	
\$5	100 20.00%	0 0.00%	57 11.40%	0 0.00%	4 0.80%	62.11% 37.89%
\$10	0 0.00%	100 20.00%	0 0.00%	0 0.00%	0 0.00%	100% 0.00%
\$20	0 0.00%	0 0.00%	43 8.60%	0 0.00%	0 0.00%	100% 0.00%
\$50	0 0.00%	0 0.00%	0 0.00%	100 20.00%	0 0.00%	100% 0.00%
\$100	0 0.00%	0 0.00%	0 0.00%	0 0.00%	96 19.20%	100% 0.00%
	100% 0.00%	100% 0.00%	43.00% 57.00%	100% 0.00%	96.00% 4.00%	87.80% 12.20%
Target Class	\$5	\$10	\$20	\$50	\$100	

(d) Combination No.4

Confusion Matrix

Output Class	\$5	\$10	\$20	\$50	\$100	
\$5	91 18.20%	0 0.00%	0 0.00%	0 0.00%	0 0.00%	100% 0.00%
\$10	0 0.00%	100 20.00%	0 0.00%	0 0.00%	0 0.00%	100% 0.00%
\$20	9 1.80%	0 0.00%	100 20.00%	0 0.00%	0 0.00%	91.74% 8.26%
\$50	0 0.00%	0 0.00%	0 0.00%	100 20.00%	0 0.00%	100% 0.00%
\$100	0 0.00%	0 0.00%	0 0.00%	0 0.00%	100 20.00%	100% 0.00%
	91.00% 9.00%	100% 0.00%	100% 0.00%	100% 0.00%	100% 0.00%	98.20% 1.80%
Target Class	\$5	\$10	\$20	\$50	\$100	

(e) Combination No.5

Confusion Matrix

Output Class	\$5	\$10	\$20	\$50	\$100	
\$5	93 18.60%	79 15.80%	0 0.00%	0 0.00%	0 0.00%	54.07% 45.93%
\$10	0 0.00%	21 4.20%	0 0.00%	0 0.00%	0 0.00%	100% 0.00%
\$20	0 0.00%	0 0.00%	100 20.00%	0 0.00%	0 0.00%	100% 0.00%
\$50	7 1.40%	0 0.00%	0 0.00%	100 20.00%	0 0.00%	93.46% 6.54%
\$100	0 0.00%	0 0.00%	0 0.00%	0 0.00%	100 20.00%	100% 0.00%
	93.00% 7.00%	21.00% 79.00%	100% 0.00%	100% 0.00%	100% 0.00%	82.80% 17.20%
Target Class	\$5	\$10	\$20	\$50	\$100	

(f) Combination No.6

Confusion Matrix

	\$5	\$10	\$20	\$50	\$100	
\$5	100 20.00%	0 0.00%	12 2.40%	0 0.00%	0 0.00%	89.29% 10.71%
\$10	0 0.00%	100 20.00%	0 0.00%	0 0.00%	0 0.00%	100% 0.00%
\$20	0 0.00%	0 0.00%	88 17.60%	0 0.00%	0 0.00%	100% 0.00%
\$50	0 0.00%	0 0.00%	0 0.00%	100 20.00%	0 0.00%	100% 0.00%
\$100	0 0.00%	0 0.00%	0 0.00%	0 0.00%	100 20.00%	100% 0.00%
	100% 0.00%	100% 0.00%	88.00% 12.00%	100% 0.00%	100% 0.00%	97.60% 2.40%
	\$5	\$10	\$20	\$50	\$100	
	Target Class					

(g) Combination No.7

Confusion Matrix

	\$5	\$10	\$20	\$50	\$100	
\$5	80 16.00%	0 0.00%	0 0.00%	0 0.00%	0 0.00%	100% 0.00%
\$10	2 0.40%	100 20.00%	0 0.00%	26 5.20%	0 0.00%	78.13% 21.88%
\$20	18 3.60%	0 0.00%	100 20.00%	2 0.40%	0 0.00%	83.33% 16.67%
\$50	0 0.00%	0 0.00%	0 0.00%	72 14.40%	0 0.00%	100% 0.00%
\$100	0 0.00%	0 0.00%	0 0.00%	0 0.00%	100 20.00%	100% 0.00%
	80.00% 20.00%	100% 0.00%	100% 0.00%	72.00% 28.00%	100% 0.00%	90.40% 9.60%
	\$5	\$10	\$20	\$50	\$100	
	Target Class					

(h) Combination No.8

Confusion Matrix

	\$5	\$10	\$20	\$50	\$100	
\$5	100 20.00%	0 0.00%	2 0.40%	0 0.00%	0 0.00%	98.04% 1.96%
\$10	0 0.00%	100 20.00%	0 0.00%	0 0.00%	0 0.00%	100% 0.00%
\$20	0 0.00%	0 0.00%	98 19.60%	0 0.00%	0 0.00%	100% 0.00%
\$50	0 0.00%	0 0.00%	0 0.00%	100 20.00%	0 0.00%	100% 0.00%
\$100	0 0.00%	0 0.00%	0 0.00%	0 0.00%	100 20.00%	100% 0.00%
	100% 0.00%	100% 0.00%	98.00% 2.00%	100% 0.00%	100% 0.00%	99.60% 0.40%
	\$5	\$10	\$20	\$50	\$100	
	Target Class					

(i) Combination No.9

Confusion Matrix

	\$5	\$10	\$20	\$50	\$100	
\$5	100 20.00%	0 0.00%	23 4.60%	0 0.00%	0 0.00%	81.30% 18.70%
\$10	0 0.00%	100 20.00%	0 0.00%	0 0.00%	0 0.00%	100% 0.00%
\$20	0 0.00%	0 0.00%	77 15.40%	0 0.00%	0 0.00%	100% 0.00%
\$50	0 0.00%	0 0.00%	0 0.00%	100 20.00%	0 0.00%	100% 0.00%
\$100	0 0.00%	0 0.00%	0 0.00%	0 0.00%	100 20.00%	100% 0.00%
	100% 0.00%	100% 0.00%	77.00% 23.00%	100% 0.00%	100% 0.00%	95.40% 4.60%
	\$5	\$10	\$20	\$50	\$100	
	Target Class					

(j) Combination No.10

Confusion Matrix

	\$5	\$10	\$20	\$50	\$100	
\$5	100 20.00%	0 0.00%	0 0.00%	0 0.00%	0 0.00%	100% 0.00%
\$10	0 0.00%	100 20.00%	19 3.80%	0 0.00%	0 0.00%	84.03% 15.97%
\$20	0 0.00%	0 0.00%	81 16.20%	0 0.00%	0 0.00%	100% 0.00%
\$50	0 0.00%	0 0.00%	0 0.00%	100 20.00%	0 0.00%	100% 0.00%
\$100	0 0.00%	0 0.00%	0 0.00%	0 0.00%	100 20.00%	100% 0.00%
	100% 0.00%	100% 0.00%	81.00% 19.00%	100% 0.00%	100% 0.00%	96.20% 3.80%
	\$5	\$10	\$20	\$50	\$100	
	Target Class					

(k) Combination No.11

Confusion Matrix

	\$5	\$10	\$20	\$50	\$100	
\$5	100 20.00%	13 2.60%	0 0.00%	0 0.00%	0 0.00%	88.50% 11.50%
\$10	0 0.00%	87 17.40%	0 0.00%	0 0.00%	0 0.00%	100% 0.00%
\$20	0 0.00%	0 0.00%	100 20.00%	0 0.00%	0 0.00%	100% 0.00%
\$50	0 0.00%	0 0.00%	0 0.00%	100 20.00%	0 0.00%	100% 0.00%
\$100	0 0.00%	0 0.00%	0 0.00%	0 0.00%	100 20.00%	100% 0.00%
	100% 0.00%	87.00% 13.00%	100% 0.00%	100% 0.00%	100% 0.00%	97.40% 2.60%
	\$5	\$10	\$20	\$50	\$100	
	Target Class					

(l) Combination No.12

Appendix C: Recognition results when using the banknote images captured by webcam for training

Confusion Matrix

Output Class	\$5	\$10	\$20	\$50	\$100	
\$5	100 20.00%	0 0.00%	0 0.00%	0 0.00%	0 0.00%	100% 0.00%
\$10	0 0.00%	97 19.40%	0 0.00%	0 0.00%	0 0.00%	100% 0.00%
\$20	0 0.00%	0 0.00%	100 20.00%	0 0.00%	0 0.00%	100% 0.00%
\$50	0 0.00%	3 0.60%	0 0.00%	100 20.00%	0 0.00%	97.09% 2.91%
\$100	0 0.00%	0 0.00%	0 0.00%	0 0.00%	100 20.00%	100% 0.00%
	100% 0.00%	97.00% 3.00%	100% 0.00%	100% 0.00%	100% 0.00%	99.40% 0.60%
	\$5	\$10	\$20	\$50	\$100	

Target Class

(a) Combination No.1

Confusion Matrix

Output Class	\$5	\$10	\$20	\$50	\$100	
\$5	79 15.80%	2 0.40%	0 0.00%	0 0.00%	0 0.00%	97.53% 2.47%
\$10	0 0.00%	91 18.20%	0 0.00%	0 0.00%	0 0.00%	100% 0.00%
\$20	21 4.20%	7 1.40%	96 19.20%	6 1.20%	5 1.00%	71.11% 28.89%
\$50	0 0.00%	0 0.00%	4 0.80%	94 18.80%	0 0.00%	95.92% 4.08%
\$100	0 0.00%	0 0.00%	0 0.00%	0 0.00%	95 19.00%	100% 0.00%
	79.00% 21.00%	91.00% 9.00%	96.00% 4.00%	94.00% 6.00%	95.00% 5.00%	91.00% 9.00%
	\$5	\$10	\$20	\$50	\$100	

Target Class

(b) Combination No.2

Confusion Matrix

Output Class	\$5	\$10	\$20	\$50	\$100	
\$5	100 20.00%	0 0.00%	0 0.00%	0 0.00%	0 0.00%	100% 0.00%
\$10	0 0.00%	98 19.60%	0 0.00%	0 0.00%	0 0.00%	100% 0.00%
\$20	0 0.00%	0 0.00%	100 20.00%	0 0.00%	0 0.00%	100% 0.00%
\$50	0 0.00%	2 0.40%	0 0.00%	100 20.00%	0 0.00%	98.04% 1.96%
\$100	0 0.00%	0 0.00%	0 0.00%	0 0.00%	100 20.00%	100% 0.00%
	100% 0.00%	98.00% 2.00%	100% 0.00%	100% 0.00%	100% 0.00%	99.60% 0.40%
	\$5	\$10	\$20	\$50	\$100	

Target Class

(c) Combination No.3

Confusion Matrix

Output Class	\$5	\$10	\$20	\$50	\$100	
\$5	100 20.00%	0 0.00%	0 0.00%	0 0.00%	4 0.80%	96.15% 3.85%
\$10	0 0.00%	100 20.00%	0 0.00%	0 0.00%	0 0.00%	100% 0.00%
\$20	0 0.00%	0 0.00%	100 20.00%	0 0.00%	0 0.00%	100% 0.00%
\$50	0 0.00%	0 0.00%	0 0.00%	100 20.00%	0 0.00%	100% 0.00%
\$100	0 0.00%	0 0.00%	0 0.00%	0 0.00%	96 19.20%	100% 0.00%
	100% 0.00%	100% 0.00%	100% 0.00%	100% 0.00%	96.00% 4.00%	99.20% 0.80%
	\$5	\$10	\$20	\$50	\$100	

Target Class

(d) Combination No.4

Confusion Matrix

Output Class	\$5	\$10	\$20	\$50	\$100	
\$5	100 20.00%	0 0.00%	0 0.00%	0 0.00%	0 0.00%	100% 0.00%
\$10	0 0.00%	100 20.00%	0 0.00%	0 0.00%	4 0.80%	96.15% 3.85%
\$20	0 0.00%	0 0.00%	100 20.00%	0 0.00%	0 0.00%	100% 0.00%
\$50	0 0.00%	0 0.00%	0 0.00%	100 20.00%	0 0.00%	100% 0.00%
\$100	0 0.00%	0 0.00%	0 0.00%	0 0.00%	96 19.20%	100% 0.00%
	100% 0.00%	100% 0.00%	100% 0.00%	100% 0.00%	96.00% 4.00%	99.20% 0.80%
	\$5	\$10	\$20	\$50	\$100	

Target Class

(e) Combination No.5

Confusion Matrix

Output Class	\$5	\$10	\$20	\$50	\$100	
\$5	100 20.00%	0 0.00%	0 0.00%	0 0.00%	0 0.00%	100% 0.00%
\$10	0 0.00%	100 20.00%	0 0.00%	0 0.00%	0 0.00%	100% 0.00%
\$20	0 0.00%	0 0.00%	73 14.60%	0 0.00%	0 0.00%	100% 0.00%
\$50	0 0.00%	0 0.00%	5 1.00%	100 20.00%	0 0.00%	95.24% 4.76%
\$100	0 0.00%	0 0.00%	22 4.40%	0 0.00%	100 20.00%	81.97% 18.03%
	100% 0.00%	100% 0.00%	73.00% 27.00%	100% 0.00%	100% 0.00%	94.60% 5.40%
	\$5	\$10	\$20	\$50	\$100	

Target Class

(f) Combination No.6

Confusion Matrix

Output Class	\$5	\$10	\$20	\$50	\$100	
\$5	100 20.00%	0 0.00%	1 0.20%	0 0.00%	0 0.00%	99.01% 0.99%
\$10	0 0.00%	98 19.60%	0 0.00%	0 0.00%	0 0.00%	100% 0.00%
\$20	0 0.00%	0 0.00%	99 19.80%	0 0.00%	0 0.00%	100% 0.00%
\$50	0 0.00%	2 0.40%	0 0.00%	100 20.00%	0 0.00%	98.04% 1.96%
\$100	0 0.00%	0 0.00%	0 0.00%	0 0.00%	100 20.00%	100% 0.00%
	100% 0.00%	98.00% 2.00%	99.00% 1.00%	100% 0.00%	100% 0.00%	99.40% 0.60%
	\$5	\$10	\$20	\$50	\$100	
	Target Class					

(g) Combination No.7

Confusion Matrix

Output Class	\$5	\$10	\$20	\$50	\$100	
\$5	96 19.20%	7 1.40%	3 0.60%	5 1.00%	0 0.00%	86.49% 13.51%
\$10	4 0.80%	93 18.60%	0 0.00%	0 0.00%	0 0.00%	95.88% 4.12%
\$20	0 0.00%	0 0.00%	85 17.00%	0 0.00%	0 0.00%	100% 0.00%
\$50	0 0.00%	0 0.00%	0 0.00%	68 13.60%	0 0.00%	100% 0.00%
\$100	0 0.00%	0 0.00%	12 2.40%	27 5.40%	100 20.00%	71.94% 28.06%
	96.00% 4.00%	93.00% 7.00%	85.00% 15.00%	68.00% 32.00%	100% 0.00%	88.40% 11.60%
	\$5	\$10	\$20	\$50	\$100	
	Target Class					

(h) Combination No.8

Confusion Matrix

Output Class	\$5	\$10	\$20	\$50	\$100	
\$5	100 20.00%	0 0.00%	0 0.00%	0 0.00%	0 0.00%	100% 0.00%
\$10	0 0.00%	100 20.00%	0 0.00%	0 0.00%	0 0.00%	100% 0.00%
\$20	0 0.00%	0 0.00%	99 19.80%	0 0.00%	0 0.00%	100% 0.00%
\$50	0 0.00%	0 0.00%	0 0.00%	100 20.00%	0 0.00%	100% 0.00%
\$100	0 0.00%	0 0.00%	1 0.20%	0 0.00%	100 20.00%	99.01% 0.99%
	100% 0.00%	100% 0.00%	99.00% 1.00%	100% 0.00%	100% 0.00%	99.80% 0.20%
	\$5	\$10	\$20	\$50	\$100	
	Target Class					

(i) Combination No.9

Confusion Matrix

Output Class	\$5	\$10	\$20	\$50	\$100	
\$5	100 20.00%	4 0.80%	0 0.00%	0 0.00%	0 0.00%	96.15% 3.85%
\$10	0 0.00%	96 19.20%	0 0.00%	0 0.00%	0 0.00%	100% 0.00%
\$20	0 0.00%	0 0.00%	100 20.00%	0 0.00%	0 0.00%	100% 0.00%
\$50	0 0.00%	0 0.00%	0 0.00%	100 20.00%	0 0.00%	100% 0.00%
\$100	0 0.00%	0 0.00%	0 0.00%	0 0.00%	100 20.00%	100% 0.00%
	100% 0.00%	96.00% 4.00%	100% 0.00%	100% 0.00%	100% 0.00%	99.20% 0.80%
	\$5	\$10	\$20	\$50	\$100	
	Target Class					

(j) Combination No.10

Confusion Matrix

Output Class	\$5	\$10	\$20	\$50	\$100	
\$5	100 20.00%	0 0.00%	0 0.00%	0 0.00%	0 0.00%	100% 0.00%
\$10	0 0.00%	100 20.00%	4 0.80%	0 0.00%	0 0.00%	96.15% 3.85%
\$20	0 0.00%	0 0.00%	96 19.20%	0 0.00%	0 0.00%	100% 0.00%
\$50	0 0.00%	0 0.00%	0 0.00%	100 20.00%	0 0.00%	100% 0.00%
\$100	0 0.00%	0 0.00%	0 0.00%	0 0.00%	100 20.00%	100% 0.00%
	100% 0.00%	100% 0.00%	96.00% 4.00%	100% 0.00%	100% 0.00%	99.20% 0.80%
	\$5	\$10	\$20	\$50	\$100	
	Target Class					

(k) Combination No.11

Confusion Matrix

Output Class	\$5	\$10	\$20	\$50	\$100	
\$5	100 20.00%	2 0.40%	0 0.00%	0 0.00%	0 0.00%	98.04% 1.96%
\$10	0 0.00%	98 19.60%	0 0.00%	1 0.20%	0 0.00%	98.99% 1.01%
\$20	0 0.00%	0 0.00%	100 20.00%	0 0.00%	0 0.00%	100% 0.00%
\$50	0 0.00%	0 0.00%	0 0.00%	99 19.80%	0 0.00%	100% 0.00%
\$100	0 0.00%	0 0.00%	0 0.00%	0 0.00%	100 20.00%	100% 0.00%
	100% 0.00%	98.00% 2.00%	100% 0.00%	99.00% 1.00%	100% 0.00%	99.40% 0.60%
	\$5	\$10	\$20	\$50	\$100	
	Target Class					

(l) Combination No.12