Forensic Science International 198 (2010) 143–149

Contents lists available at ScienceDirect

Forensic Science International

journal homepage: www.elsevier.com/locate/forsciint

Analysis of geometric moments as features for firearm identification☆

Nor Azura Md Ghani [a,1], Choong-Yeun Liong [b,*], Abdul Aziz Jemain [b,2]

[a] Center for Statistical Studies, Faculty of Computer and Mathematical Sciences, Universiti Teknologi MARA, 40450 Shah Alam, Selangor D.E., Malaysia
[b] School of Mathematical Sciences, Faculty of Science and Technology, Universiti Kebangsaan Malaysia, 43600 UKM Bangi, Selangor D.E., Malaysia

ARTICLE INFO

Article history:
Received 20 August 2009
Received in revised form 5 February 2010
Accepted 9 February 2010
Available online 7 March 2010

Keywords:
Firearm identification
Correlation analysis
Geometric moments
Feature extraction
Discriminant analysis

ABSTRACT

The task of identifying firearms from forensic ballistics specimens is exacting in crime investigation since the last two decades. Every firearm, regardless of its size, make and model, has its own unique 'fingerprint'. These fingerprints transfer when a firearm is fired to the fired bullet and cartridge case. The components that are involved in producing these unique characteristics are the firing chamber, breech face, firing pin, ejector, extractor and the rifling of the barrel. These unique characteristics are the critical features in identifying firearms. It allows investigators to decide on which particular firearm that has fired the bullet. Traditionally the comparison of ballistic evidence has been a tedious and time-consuming process requiring highly skilled examiners. Therefore, the main objective of this study is the extraction and identification of suitable features from firing pin impression of cartridge case images for firearm recognition. Some previous studies have shown that firing pin impression of cartridge case is one of the most important characteristics used for identifying an individual firearm. In this study, data are gathered using 747 cartridge case images captured from five different pistols of type 9 mm Parabellum Vektor SP1, made in South Africa. All the images of the cartridge cases are then segmented into three regions, forming three different set of images, i.e. firing pin impression image, centre of firing pin impression image and ring of firing pin impression image. Then geometric moments up to the sixth order were generated from each part of the images to form a set of numerical features. These 48 features were found to be significantly different using the MANOVA test. This high dimension of features is then reduced into only 11 significant features using correlation analysis. Classification results using cross-validation under discriminant analysis show that 96.7% of the images were classified correctly. These results demonstrate the value of geometric moments technique for producing a set of numerical features, based on which the identification of firearms are made.

1. Introduction

Over the last two decades, the identification of firearms from forensic ballistics specimens is exacting and tremendous in crime investigation. Experts in the general area of firearms identification are usually called firearms examiners. In forensic laboratories, firearms examiners identified firearm by visual examination using comparison microscopes. This is the traditional way which is used to identify a particular firearm [1,2]. In the past decade, automatic ballistics identification systems such as FIREBALL, IBIS, EVOFINDER were developed which can help investigators to link crimes by automatically finding similarities among bullet and cartridge case images [3–6]. It usually takes long time because of the huge number of firearm evidence in database and the number of firearms to be matched is usually too large. There are occasions in which the firearms examiners make some mistakes due to making judgment by visual examination. It is a large burden for authentication manually.

In order to identify the firearm efficiently, an automatic identification system is indeed much needed to narrow down the possible number of matching required quickly. In this paper, a system for firearm identification using features from cartridge case image is introduced. To the best of our knowledge, this is the first attempt to use numerical based features to classify firearms. Cartridge is one of the important clues for security staff to solve the gun file. There are many marks left on the surface of the cartridge case when a gun is fired. These characteristic markings can be recognized as a "fingerprint" for identification of the make and model of the firearm [3].

In this study, the features for the purpose of firearm identification are extracted from cartridge case images using geometric moment up to the sixth order to form a set of numerical

☆ Published for Educational Purposes only.

* Corresponding author. Tel.: +60 3 89213712/5724; fax: +60 3 89254519.
E-mail addresses: azura@tmsk.uitm.edu.my (N.A. Md Ghani), lg@ukm.my (C.-Y. Liong), azizj@ukm.my (A.A. Jemain).
[1] Tel.: +60 3 55435371; fax: +60 3 55435501.
[2] Tel.: +60 3 89213712/5724; fax: +60 3 89254519.

0379-0738/$ – see front matter © 2010 Elsevier Ireland Ltd. All rights reserved.
doi:10.1016/j.forsciint.2010.02.011

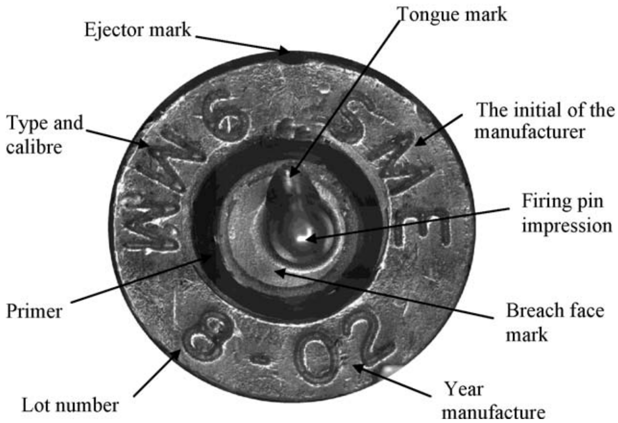

Fig. 1. Head of cartridge case and its individual characteristics.

features. Geometric moment features were used for the classification of the cartridge case images into the respective classes using discriminant analysis. Besides this introduction, Section 2 describes the forensic ballistics features and Section 3 explains the data collection process. Section 4 gives the details on the segmentation process which extracts the regions of interest from the cartridge case images, and the data collection. Section 5 describes the features used and the feature extraction process. Moments are presented in Section 6 and feature selection is discussed in Section 7. Results and discussion on the features produced and the classification results are presented in Section 8. Conclusion and suggestion for further work are also given.

2. Forensic ballistics features

Forensic ballistics feature are formed in the process of loading in the bullet, shooting and throwing out the bullet. It has two types of tiny markings, called class and individual characteristics. The class characteristics of firearms that relate to the bullets fired from them includes the caliber of the firearm and the rifling pattern contained in the barrel of the firearm. The individual characteristics of firearm include firing pin impression, breach face marks, extractor marks, ejector marks and others [7]. Several prominent individual characteristics and head stamps are as shown in Fig. 1.

In this study, we focus on the firing pin impression because from the literature, unlike the other features such as the breach face marks or the ejector marks, the firing pin impression is always robust [4,8–10]. The firing pin impression is resulted by the force of the firing pin during the shooting and it appears like a small cave, and varies according to firearm. This information includes the position, radius, depth, shape of the firing pin mark, tongue mark and tiny mark on the surface of the firing pin. Firearm manufacturers have produced a variety of shapes for firing pins. Some examples of firing pin impression which can be classified by shape are bar pin mark, circular pin mark, double pin mark, square pin mark, rectangular pin mark and ringer pin mark [5]. The shape of the firing pin is also important for individual characteristics [9]. Therefore firing pin impression is an important factor for firearm identification.

3. Data collection

A total of five pistols of Parabellum Vektor SP1 9 mm model, made in South Africa, were used in this research work. The choice of pistols were made based on two criteria: the pistols should be relatively of the same age and capable of producing firing pin impressions that are 'similar', i.e. images that can hardly be distinguished even by an expert examiner. The pistols that were chosen by a forensic ballistics expert from the Royal Malaysian Police were labeled as Pistol A, Pistol B, Pistol C, Pistol D and Pistol E as in the sequence of the choice made in order to differentiate between the pistols. The pistols are relatively new and are cleaned weekly. The bullets used were all of the same lot of the year 2002, i.e. being manufactured at the same date. A total of 150 bullets have been fired continuously from each of the five pistols in order to identify the variation on the firing pin impressions produced due to heat and speed of firing, if any. During the session, three cartridge cases were thrown far away, and hence failed to be identified. The missing cartridge cases are one from Pistol D and two from Pistol E. Therefore, there are 150 cartridge cases each for Pistol A, B and C; 149 cartridge cases for Pistol D and 148 cartridge cases for Pistol E, and hence a total of 747 cartridge cases were recorded and used for this study.

The firing mechanism of this weapon is center-fire mechanism and the shape of the firing pin impression is a circular pin mark. In other words, the firing pin impression is of a center-fire cartridge. The cartridge case image was captured by using the CONDOR System. CONDOR is a commercial automatic system equipped with built-in lighting and scanner to capture the image of a cartridge case or a bullet. Hence all the images are captured in a well controlled lighting condition. In other words, the images captured are not affected by lighting problem.

4. Image segmentation

Image segmentation is one of the most important steps leading to the analysis of processed image data. Image segmentation aims to separate the object(s) of interest from the unwanted background information and is often the first step in any automated vision based system [11]. Fig. 2 shows the major components of the proposed firearm identification system.

The firing pin impression needs to be captured, preprocessed, segmented and enhanced for further processing as it is hardly differentiable by the naked eyes. Here, segmentation is the task to separate the firing pin impression area from the cartridge case image, as well as to divide the image into regions of interest that may enhance the particular features of interest. Referring to Fig. 3 which presents the segmented regions of a cartridge case image, the aim of segmentation is to produce images (b), (c) and (d) from image (a). More details in Fig. 3 are given later.

As there are 747 images from five different pistols, and the bullets were shots continuously, there are slight differences in terms of the regions that form the firing pin impressions in the images although all are circular in form. A circle with a predetermined radius has been shown to be an adequate mask for segmenting the circular firing pin impression region for each image. For this research, the radius of the mask for each of the images has been determined manually based on the intensity values bordering the firing pin impression region. These values were then stored into a file for the segmentation processing using our own MATLAB program [12].

The images were segmented into three parts using the MATLAB program, i.e. into firing pin impression (whole image), centre of the firing pin impression (centre image) and ring image as shown in Fig. 3. The ring image is actually the difference between the images of the firing pin impression and the centre of the firing pin impression. In other words, the ring image is obtained by eliminating the centre of the firing pin impression from the firing pin image. The reason why the images are partitioned into such regions is to enhance the features residing in those regions. The significance of the features extracted from these regions and the eventual classification performance are

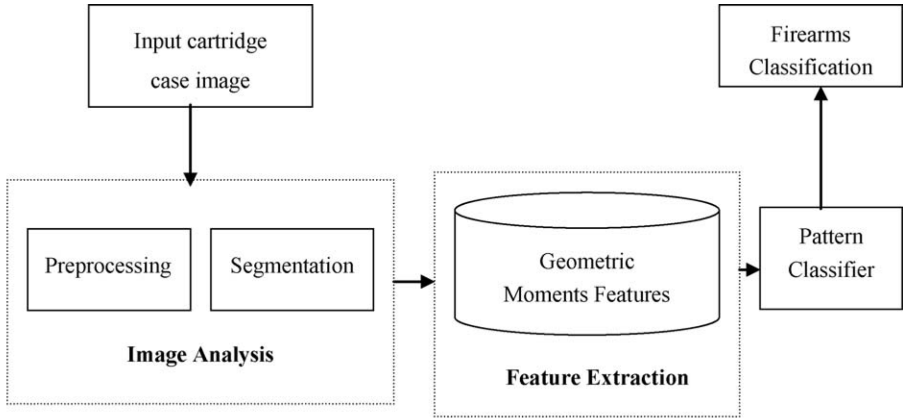

Fig. 2. The major components of the firearm identification system.

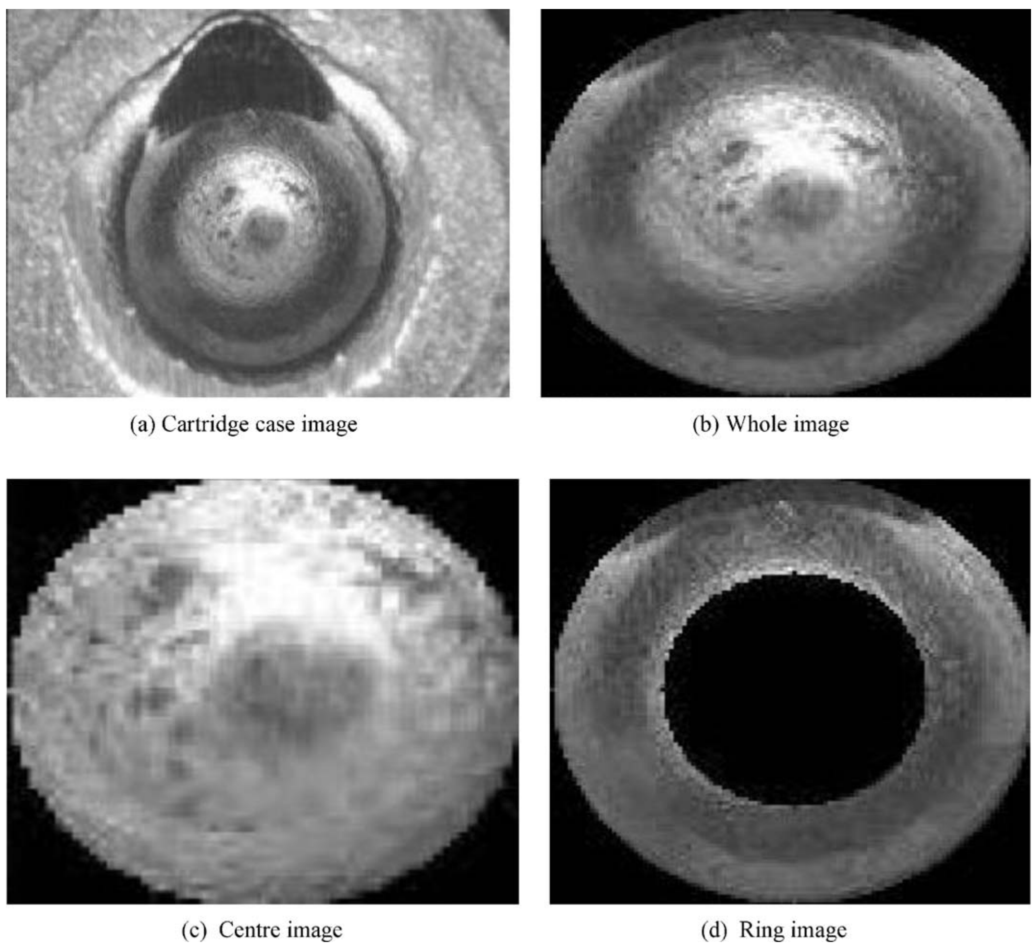

Fig. 3. Extracted regions of a cartridge case image.

the measuring criteria to the importance of such segmentation strategy.

5. Feature extraction

A feature is a region of interest on the surface of a part. The ultimate aim in a large number of image processing applications is to extract important features from image data, from which a description or understanding of the scene can be provided [13]. Feature extraction refers to a step in image processing where measurements or observations are processed to find attributes that can be used for classification [14]. Feature extraction can generate large data sets but it can also introduce a large amount of noise [15]. If the features extracted are carefully chosen, it is expected that the features set will extract the relevant information from the input data in order to perform the desired task using this reduced representation instead of the full size input. If many features are extracted and very little or no significant information is available, it is practically impossible to perform any evaluation. It needs to be noted that finding the right features is generally the most demanding task in designing a pattern classification system as there are no rules to follow [16]. One generally has to depend on one's knowledge of the problem domain and backed by a careful study of the data. In other words, it is more of a heuristics approach rather than procedural.

In this work, the segmented images are gray level images, i.e. images with intensity values in the range of [0, 255]. The geometric moment features are generated based on the image intensities which were computed using MATLAB program. In other words, these features consider two-dimensional information of the image, i.e. the image intensity value and the location of the value are used in the computation of the features.

6. Moments

The mathematical concept of moments has been around for many years and has been utilized in many fields ranging from mechanics and statistics to pattern recognition and image understanding. Describing images with moments means that global properties of the image are used rather than just the local properties [17]. Historically, the first significant work considering moments for pattern recognition was introduced by Hu [18].

6.1. General moment definition

A general definition of moment functions ϕ_{pq} of order $(p+q)$, of an image intensity function $f(x, y)$ can be defined as follows:

$$\phi_{pq} = \int_y \int_x \psi_{pq}(x,y)\, f(x,y)\, dx\, dy \quad p, q = 0, 1, 2, 3, \ldots \tag{1}$$

where $\psi_{pq}(x, y)$ is the moment weighting kernel. For a digital image, Eq. (1) needs to be expressed in the discrete form as follows:

$$\phi_{pq} = \sum_x \sum_y \psi_{pq}(x,y)\, f(x,y) \quad p, q = 0, 1, 2, 3, \ldots \tag{2}$$

where (x, y) are the pixel coordinates of $f(x, y)$. The basis functions may have a range of useful properties that may be passed onto the moments. For examples, the orthogonality property of the basis function is passed onto the moments. Thus, non-orthogonal basis function result in non-orthogonal moments and orthogonal basis functions result in orthogonal moments. Geometric moments are non-orthogonal moments, whereas Legendre moments are orthogonal moments [13,19].

6.2. Geometric moments

Geometric moments are the most popular types of moments and have been frequently used for a number of image processing tasks. Geometric moment computations on images can be easily performed and implemented, as compared to other moments with complex kernel functions. Geometric moments are the simplest among moment functions, with the kernel function defined as a product of the pixel coordinates, i.e. $x^p y^q$. The two-dimensional geometric moment of order $(p+q)$ for an $N \times M$ discrete image is given by:

$$m_{pq} = \sum_{y=1}^{M} \sum_{x=1}^{N} x^p y^q f(x,y) \quad p, q = 0, 1, 2, 3, \ldots. \tag{3}$$

Note that even though the basis set $x^p y^q$ is complete, it is not orthogonal [13,19], and hence geometric moments are non-orthogonal moments. The uniqueness theorem states that the moment set m_{pq} is unique for a given image function $f(x, y)$ [20]. In addition, the existence theorem states that the moments of all orders exist [21]. These theorems give rise to the reconstruction property of moments and the applicability of moments for shape representation.

From the equation, it is clear that the moments rise rapidly as the order of the moments and the coordinates increase. Hence, in this work, the geometric moment is scaled into the range of $[-1, 1]$. The scaling does not cause any information lost but gives the desired effect of where the higher order moments will, in general, have progressively smaller magnitudes. This will ensure that the feature values are within a smaller range which is useful for the classification task later [22].

The image intensity function $f(x, y)$ are gray level images, i.e. images with intensity values in the range of [0, 255]. Therefore, with the scaling, the geometric moment of order $(p+q)$ for an $N \times N$ two-dimensional discrete image is given as:

$$m_{pq} = \sum_{y=1}^{N} \sum_{x=1}^{N} \left(\frac{2x}{N} - 1\right)^p \left(\frac{2y}{N} - 1\right)^q \frac{1}{255} f(x,y)$$
$$p, q = 0, 1, 2, 3, \ldots.$$

Let $u = x/N$ and $v = y/N$, the above equation can be rewritten as:

$$m_{pq} = \sum_{v=1/N}^{1} \sum_{u=1/N}^{1} (2u-1)^p (2v-1)^q \frac{1}{255} f(uN, vN)$$
$$p, q = 0, 1, 2, 3, \ldots. \tag{4}$$

Eq. (4), with the desirable property explained above, is the equation used to compute the features from the regions of interest of the cartridge case images.

7. Feature selection

One of the problems often arises in pattern recognition applications is the curse of dimensionality when data is relatively limited versus the dimension of the data. The easiest way to reduce the input dimension is to select some inputs and discard the rest. This method is called feature selection. For pattern recognition applications, feature selection is essential in that it is able to reduce execution time, and improve the accuracy and the efficiency of the classification tasks [23].

Besides that, large numbers of features are often impractical because it can harm the performance of classification. Therefore, the primary aim of a pattern recognition application is to determine the set of features most useful and relevant for the identification task. This is achieved by maximizing the retention of class discriminatory information and reducing the class commonality information [24]. Therefore a feature selection process based on correlation analysis and statistical tests were adopted and are discussed in the next section.

8. Results and discussion

Based on the literature [21], higher order moments are generally highly sensitive to the effect of noise. Often moments up the fourth order are sufficient for many recognition tasks. In this study, geometric moments up to the sixth order were computed from the 747 firing pin impression images of the five pistols for feature selection investigation. A total of 48 features were extracted from each of the images. The list of the features and the notations used are given in Table 1.

The sample means of the 48 features, computed over all the images for each of the pistols, are given in Table 2. For examples, the value of MP_{00} for Pistol A, 18,700, is the average of the 150 MP_{00} values computed from the whole image of the cartridge cases for Pistol A. So are for Pistols B and C. For Pistol D the value is an average from 149 images, and for Pistol E the value is computed from 148 images. The number of images are less for Pistols D and E because three cartridge cases were missing as explained in Section 3.

Table 1
List of geometric moment features.

No	Firing pin impression image features	Notation	No	Firing pin impression image features	Notation
1.	M_{00} whole	MP_{00}	25.	M_{20} centre	MT_{20}
2.	M_{01} whole	MP_{01}	26.	M_{21} centre	MT_{21}
3.	M_{02} whole	MP_{02}	27.	M_{22} centre	MT_{22}
4.	M_{03} whole	MP_{03}	28.	M_{23} centre	MT_{23}
5.	M_{10} whole	MP_{10}	29.	M_{30} centre	MT_{30}
6.	M_{11} whole	MP_{11}	30.	M_{31} centre	MT_{31}
7.	M_{12} whole	MP_{12}	31.	M_{32} centre	MT_{32}
8.	M_{13} whole	MP_{13}	32.	M_{33} centre	MT_{33}
9.	M_{20} whole	MP_{20}	33.	M_{00} ring	MC_{00}
10.	M_{21} whole	MP_{21}	34.	M_{01} ring	MC_{01}
11.	M_{22} whole	MP_{22}	35.	M_{02} ring	MC_{02}
12.	M_{23} whole	MP_{23}	36.	M_{03} ring	MC_{03}
13.	M_{30} whole	MP_{30}	37.	M_{10} ring	MC_{10}
14.	M_{31} whole	MP_{31}	38.	M_{11} ring	MC_{11}
15.	M_{32} whole	MP_{32}	39.	M_{12} ring	MC_{12}
16.	M_{33} whole	MP_{33}	40.	M_{13} ring	MC_{13}
17.	M_{00} centre	MT_{00}	41.	M_{20} ring	MC_{20}
18.	M_{01} centre	MT_{01}	42.	M_{21} ring	MC_{21}
19.	M_{02} centre	MT_{02}	43.	M_{22} ring	MC_{22}
20.	M_{03} centre	MT_{03}	44.	M_{23} ring	MC_{23}
21.	M_{10} centre	MT_{10}	45.	M_{30} ring	MC_{30}
22.	M_{11} centre	MT_{11}	46.	M_{31} ring	MC_{31}
23.	M_{12} centre	MT_{12}	47.	M_{32} ring	MC_{32}
24.	M_{13} centre	MT_{13}	48.	M_{33} ring	MC_{33}

From the table, a careful inspection will show that the patterns of the sample means of the 48 features are different across the five pistols. For examples, MP_{00} is lowest (18,400) for Pistol E, and increasing (18,700) for Pistol A, Pistol C (19,500), Pistol D (20,800) until it reaches the highest level (21,500) for Pistol B. Similar pattern are seen for the other features in Table 2 where the values do vary in significant amount among the five pistols.

In order to establish a more significant proof, a statistical test would be the right tool to determine if these geometric moment features are significantly different among the pistols. Here, multivariate analysis of variance (MANOVA) will be applied. MANOVA [25] is an extension of analysis of variance to examine several measurements simultaneously. The purpose of MANOVA is to test whether the vectors of means for the two or more groups are sampled from the same sampling distribu-

Table 2
Means of the 48 geometric moment features extracted from the pistols and *p*-value based on MANOVA test.

Pistol	Comparison M_{00}				Comparison M_{01}				Comparison M_{02}				Comparison M_{03}			
	MP_{00}	MT_{00}	MC_{00}	*p*-Value	MP_{01}	MT_{01}	MC_{01}	*p*-Value	MP_{02}	MT_{02}	MC_{02}	*p*-Value	MP_{03}	MT_{03}	MC_{03}	*p*-Value
A	18,700	4940	13,800		−343	−75	−303		4230	1040	3970		−106	−16	−101	
B	21,500	5920	15,600		−388	−172	−301		4840	1240	4530		−108	−71	−96	
C	19,500	5550	14,000	0.000	−287	−101	−236	0.000	4460	1210	4160	0.000	−842	−35	−77	0.000
D	20,800	5770	15,000		−245	−112	−195		4690	1190	4400		−572	−33	−54	
E	18,400	5180	13,200		−264	−66	−235		4150	1150	3870		−859	−31	−83	

Pistol	Comparison M_{10}				Comparison M_{11}				Comparison M_{12}				Comparison M_{13}			
	MP_{10}	MT_{10}	MC_{10}	*p*-Value	MP_{11}	MT_{11}	MC_{11}	*p*-Value	MP_{12}	MT_{12}	MC_{12}	*p*-Value	MP_{13}	MT_{13}	MC_{13}	*p*-Value
A	−1350	−362	−1170		200	7	196		−213	−51	−206		57	3	56	
B	−1050	−61	−1030		155	−6	155		−174	−26	−172		52	30	51	
C	−1060	−296	−917	0.000	106	8	103	0.000	−160	−49	−153	0.000	36	4	35	0.000
D	−9860	−213	−885		148	16	145		−159	−41	−155		45	5	44	
E	−1170	−493	−928		64	10	62		−167	−76	−157		19	3	19	

Pistol	Comparison M_{20}				Comparison M_{21}				Comparison M_{22}				Comparison M_{23}			
	MP_{20}	MT_{20}	MC_{20}	*p*-Value	MP_{21}	MT_{21}	MC_{21}	*p*-Value	MP_{22}	MT_{22}	MC_{22}	*p*-Value	MP_{23}	MT_{23}	MC_{23}	*p*-Value
A	4900	1150	4620		−62	−15	−59		742	171	732		−16	−2.4	−15	
B	5370	1370	5030		−50	−19	−47		823	202	810		−12	−7.1	−11	
C	4710	1220	4410	0.000	−27	−17	−24	0.000	740	192	729	0.000	−5	−5.1	−4	0.000
D	5120	1310	4800		−20	−15	−19		793	198	780		−3	−3.9	−3	
E	4550	1200	4260		−29	−4	−29		702	187	690		−7	−2.3	−7	

Pistol	Comparison M_{30}				Comparison M_{31}				Comparison M_{32}				Comparison M_{33}			
	MP_{30}	MT_{30}	MC_{30}	*p*-Value	MP_{31}	MT_{31}	MC_{31}	*p*-Value	MP_{32}	MT_{32}	MC_{32}	*p*-Value	MP_{33}	MT_{33}	MC_{33}	*p*-Value
A	−470	−144	−450		83	6.4	82		−64	−16	−64		19	1.4	18	
B	−419	−109	−408		61	3.7	60		−51	−13	−51		16	1.6	15	
C	−350	−140	−331	0.000	43	5.6	42	0.000	−42	−15	−42	0.000	11	1.5	11	0.000
D	−341	−121	−327		60	7.9	60		−43	−13	−43		15	1.7	15	
E	−364	−187	−342		25	−1.5	26		−45	−23	−45		6	0.2	6	

Table 3
The 11 best features and the tolerance values.

Image features	Notation	Tolerance values
M_{03} whole	MP_{03}	0.810
M_{12} whole	MP_{12}	0.261
M_{20} whole	MP_{20}	0.259
M_{01} centre	MT_{00}	0.195
M_{02} centre	MT_{02}	0.258
M_{10} centre	MT_{10}	0.289
M_{11} centre	MT_{11}	0.711
M_{12} centre	MT_{12}	0.313
M_{21} centre	MT_{21}	0.192
M_{10} ring	MC_{10}	0.256
M_{11} ring	MC_{11}	0.813

tion. From Table 2, each comparison indicates that, on average, the sets of features have a highly significant difference ($p < 0.0005$) across the five pistols [26]. This confirms that the sets of features are significantly different among the five pistols.

Combination of the features would form vectors of 48-feature that are different significantly. Even though the features are significantly difference, but it is very likely that a few of the features are affected by noise, redundant or correlated [27]. In order to solve this problem, feature selection will be performed on the sets of features. Feature selection aims to reduce the feature dimension and eliminate redundancy such that the resulting set of features is small but still discriminate sufficiently. Dimension reduction will also contribute towards faster learning phase of the classification algorithm and a more efficient classification system [28].

A good feature subset is one that contains features that are highly correlated with the class, yet uncorrelated with each other [28]. Therefore, we evaluate the best features based on the correlation perspective. Furthermore, one of the most critical assumptions for discriminant analysis that need to be met is multicollinearity. Collinearity is the association, measured as the correlation, between two independent variables. Multicollinearity refers to the correlation among three or more independent variables. Multicollinearity, measured in terms of tolerance, denotes that two or more independents variables are highly correlated, so that one variable can be highly explained or predicted by the other variable(s) and thus it adds little to the explanatory power of the entire set [25]. In other words, one of the variables can be dropped. In this study, the features should not have multicollinearity which indicates redundant features and decreases statistical efficiency.

The procedure of the analysis performed is as follows. First, the correlation among the feature elements was calculated. Since the data have been obtained by using five different pistols of the same make and model, it is not surprising to see that some of the features are highly correlated. Therefore, before classification of pistols can be made based on the features, the problem of multicollinearity among the features has to be studied. If there is a problem of multicollinearity, we cannot ascribe the unique reduction in the variation among the pistols as contributed by a particular feature. According to Wulder [29], multicollinearity occurred when features are highly correlated, i.e. with a correlation value of 0.90 and above. Multicollinearity in discriminant analysis is identified when the tolerance value for a feature is less than 0.10. However, one should not attempt to interpret an analysis with a multicollinearity problem until it is resolved by removing or combining the problematic variables [30]. Therefore, combinations of problematic variables were also examined to identify the significant and practical set of features.

Finally, 11 features are chosen and the classification results show that 96.7% of all the cartridge case images were classified correctly using cross-validation. The 11 best features and the respective tolerance values are given in Table 3.

The tolerance values for all of the features are larger than 0.10 and so there is no multicollinearity problem. Hence, all the features are independent to each other. This assumption is important for discriminant analysis and must be met. Subsequently, discriminant analysis [25,31] is used to determine how good the features discriminate the five pistols.

From Table 3, 6 out of 11, or 54.5% of the best features are derived from centre image of the firing pin impression. This shows the importance and the positive effects of partitioning the firing pin image into the different regions. The list also shows that the centre image is the most robust compared to the images of the other regions.

The outcomes of classification of the firing pin impression images using cross-validation classification in SPSS are as shown in the confusion matrix in Table 4. The cross-validation classification was carried out using the leave-one-out procedure. Cross-validation has been employed in order to enable all of the available data to be utilized for training while still give an unbiased estimate of the generalization capabilities of the resulting classifier [32].

Table 4 shows that each of the pistols has achieved a high degree of classification accuracy as the classification rates higher than 95%. The overall classification rate of 96.7% $[(144 + 144 + 145 + 144 + 145) \times 100/747]$ can be considered as very good as many studied in pattern recognition found that classification rates above 70% as acceptable [28].

Kappa coefficient [33] was computed to assess the accuracy in prediction of the group membership. The output was kappa equal to 0.968 which indicates that satisfactory accuracy is

Table 4
Frequency of correct classification and rates (in %) based on the 11 selected geometric moments for classifying the five classes of pistol using discriminant analysis.

		Predicted class					
		Pistol A	Pistol B	Pistol C	Pistol D	Pistol E	Total
Actual class	Pistol A	144 (96.0)	0 (0)	3 (2.0)	2 (1.3)	1 (0.7)	150
	Pistol B	0 (0)	144 (96.0)	0 (0)	6 (4.0)	0 (0)	150
	Pistol C	0 (0)	0 (0)	145 (96.7)	2 (1.3)	3 (2.0)	150
	Pistol D	1 (0.7)	0 (0)	4 (2.7)	144 (96.6)	0 (0)	149
	Pistol E	0 (0)	0 (0)	1 (0.7)	2 (1.4)	145 (98.0)	148

achieved in the prediction. Thus, this shows that the eleven best features chosen from the geometric moments features are good features for firearm recognition, and that numerical based features show good potential toward efficient firearms identification.

9. Further work

We propose that Legendre orthogonal moments could also be used to extract new features. It is believed that Legendre orthogonal moments, due to its orthogonal property, will produce a better set of features in terms of size and uniqueness. As expected, this will not only ease the feature selection process but should also produce at least as good a classification rate as those found using the geometric moments.

The work suggested in this paper could further be enhanced by investigating the robustness of the results obtained based on the different choices of the size of the radius of the circular mask. Based on the classification rates obtained one can decide on the optimum size of the circular mask.

Acknowledgements

We would like to extend our gratitude and appreciation to DSP Abdul Rahman bin Kassim and his colleagues at the Ballistics Section, Forensic Laboratory of the Royal Malaysian Police, for providing the equipment and assistance in support of this research work.

References

[1] G. Burrard, The Identification of Firearms and Forensic Ballistics, Herbert Jenkins, London, 1951.
[2] B.J. Heard, Handbook of Firearms and Ballistics: Examining and Interpreting Forensic Evidence, John Wiley & Sons, Chichester, 1997.
[3] C.L. Smith, D. Li, Intelligent imaging of forensic ballistics specimens for ID, in: Proceedings of the IEEE Congress on Image and Signal Processing, 2008, pp. 37–41.
[4] Z. Geradts, J. Bijhold, R. Hermsen, F. Murtagh, Image matching algorithms for breech face marks and firing pins in a database of spent cartridge cases of firearms, Forensic Science International 119 (1) (2001) 97–106.
[5] C.L. Smith, J.M. Cross, Optical imaging techniques for ballistics specimens to identify firearms, in: Proceedings of the IEEE International Carnahan Conference on Security Technology, 1995, pp. 275–289.
[6] Automated firearms identification, 2008, http://en.wikipedia.org/wiki/Automated_firearms_identifications.
[7] V.J.M. Di Maio, Gunshot Wounds: Practical Aspects of Firearms, Ballistics and Forensic Techniques, second edition, CRC Press, Boca Raton, 1999.
[8] L.-P. Xin, J. Zhou, G. Rong, A cartridge identification system for firearms authentication, in: Proceedings of the IEEE Signal Processing, 2000, pp. 1405–1408.
[9] J. Zhou, L.-P. Xin, D.-S. Gao, C.-S. Zhang, D. Zhang, Automated identification for firearms authentication, in: Proceedings of the IEEE Computer Society Conference, vol. 1, 2001, pp. 749–754.
[10] D. Li, A new approach for firearm identification with hierarchical neural networks based on cartridge case images, in: Proceedings of the 5th IEEE International Conference on Cognitive Informatics (ICCI'06), 2006, pp. 923–928.
[11] M. Sonka, V. Hlavac, R. Boyle, Image Processing, Analysis and Machine Vision, PWS Publishing, Pacific Grove, 1999.
[12] R.C. Gonzales, R.E. Wood, S.L. Eddins, Digital Image Processing Using MATLAB, Pearson Prentice-Hall, Upper Saddle River, NJ, 2004.
[13] A.K. Jain, Fundamentals of Digital Image Processing, Prentice-Hall, Upper Saddle River, NJ, 1989.
[14] C.Y. Liong, Investigation of Vision Based System For Passenger Position Recognition, PhD thesis, Applied Mathematics and Computing Group, School of Mechanical Engineering, Cranfield University, 2001.
[15] N. Paulsson, E. Larsson, F. Winquist, Extraction and selection of parameters for evaluation of breath alcohol measurement with an electronic nose, Sensors and Actuators 84 (2000) 87–197.
[16] R.A. Haddad, T.W. Parsons, Digital Signal Processing: Theory, Applications and Hardware, W.H. Freeman and Co., New York, 1991.
[17] G. Faouzi, D. Stephane, D. Sami, M. Rim, Reconstructing with geometrics moments, International Conference on Machine Intelligence (ACIDCA–ICMI'2005), http://www.fresnel.fr/perso/derrode/publi/Acidca2005.pdf, 2005.
[18] M.-K. Hu, Pattern recognition by moment invariants, Proceeding of the IRE 49 (1961) 1428.
[19] M.R. Teague, Image analysis via the general theory of moments, Journal of Optical Society of America 70 (8) (1980) 920–930.
[20] M.-K. Hu, Visual pattern recognition by moment invariants, Proceeding of the IRE Transactions on Information Theory 8 (2) (1962) 179–187.
[21] R. Mukundan, K.R. Ramakrishnan, Moment Functions in Image Analysis: Theory and Applications, World Scientific, Singapore, 1998.
[22] T. Masters, Neural, Novel and Hybrid Algorithms for Time Series Prediction, John Wiley & Sons, New York, 1995.
[23] J. Liang, S. Yang, A. Winstanley, Invariant optimal feature selection: a distance discriminant and feature ranking based solution, Pattern Recognition 41 (5) (2008) 1429–1439.
[24] H.C. Andrews, Introduction to Mathematical Techniques in Pattern Recognition, John Wiley & Sons, New York, 1972.
[25] J.F. Hair, W.C. Black, J.B. Babin, R.E. Anderson, R.L. Tatham, Multivariate Data Analysis, sixth edition, Pearson Prentice-Hall, Upper Saddle River, NJ, 2006.
[26] S.J. Coakes, L. Steed, P. Dzidic, SPSS Version 16.0 for Windows: Analysis without Anguish, John Wiley & Sons, Milton, 2009.
[27] R.W. Swiniarski, A. Skowron, Rough set methods in feature selection and recognition, Pattern Recognition 24 (2002) 833–849.
[28] M.A. Hall, Correlation based Feature Selection for Machine Learning. PhD thesis, Department of Computer Science, University of Waikato, Hamilton, New Zealand, 1999.
[29] M. Wulder, Multivariate statistics: multicollinearity and singularity, http://www.pfc.cfs.nrcan.gc.ca/profiles/wulder/mvstats/multicol_e.html, 2006.
[30] J.A. Schwab, Data analysis and computers. II. Strategy for complete discriminant analysis, http://www.utexas.edu/courses/schwab/sw388r7/SolvingProblems/DA_CompleteProblems.ppt, 2005.
[31] R.A. Johnson, D.W. Wichern, Applied Multivariate Statistical Analysis, fifth edition, Prentice-Hall, Upper Saddle River, NJ, 2002
[32] C.M. Bishop, Neural Network For Pattern Recognition, Oxford University Press, New York, 1995.
[33] R.A. Yaffee, Common correlation and reliability analysis with SPSS for windows, http://www.nyu.edu/its/statistics/Docs/correlate.html, 2003.