SPIE Int. Symposium on Law Enforcement Technologies, Orlando, Florida, USA, 1–5 April 2002, Proc. SPIE 4709

Strategies for the automated recognition of marks in forensic science

M. Heizmann

Institut für Mess- und Regelungstechnik, Universität Karlsruhe (TH),
Postfach 6980, D-76128 Karlsruhe, Germany

ABSTRACT

To enable the efficient comparison of striation marks in forensic science, tools for the automated detection of similarities between them are necessary. Such marks show a groove-like texture which can be considered as a "fingerprint" of the associated tool. Thus, a reliable detection of connections between different toolmarks from the identical tool can be established. In order to avoid the time-consuming visual inspection of toolmarks, automated approaches for the evaluation of marks are essential. Such approaches are commonly based on meaningful characteristics extracted from images of the marks that are to be examined. Besides of a high recognition rate, the required computation time plays an important role within the design of an adequate comparison strategy. The cross-correlation function presented in this paper provides a faithful quantitative measure to determine the degree of similarity. It is shown that appropriate modelling of the signal characteristics considerably improves the performance of methods based on the cross-correlation function. A strategy for quantitative assessment of comparison strategies is introduced. It is based on the processing of a test archive of marks and analyses the comparison results statistically. For a convenient description of the assessment results, meaningful index numbers are introduced and discussed.

Keywords: forensic science, pattern recognition, correlation methods, automated visual inspection, image processing, marks examination, striation patterns, toolmarks, firearm bullets, comparison strategies

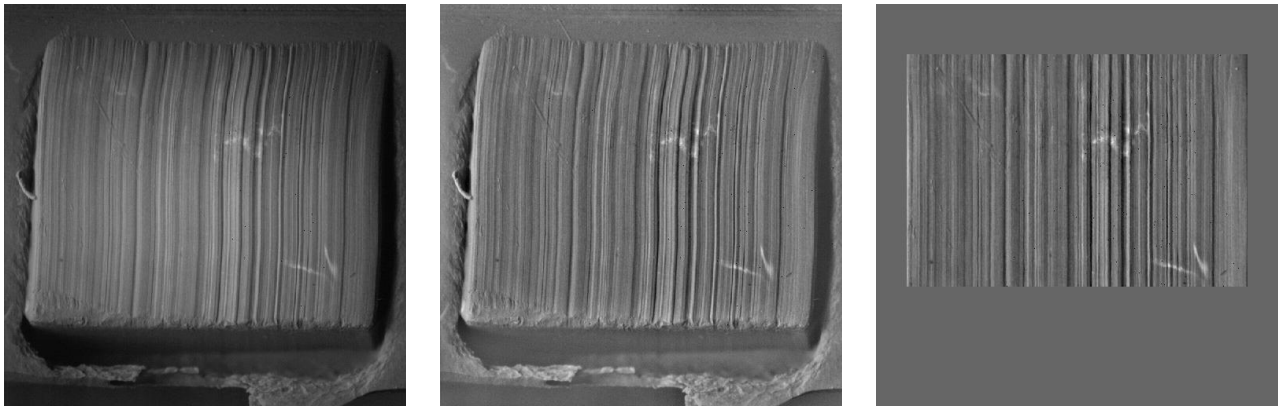
1. INTRODUCTION

To support forensic examiners in the time-consuming evaluation of marks, automated systems for the processing of data taken from the marks are essential. Promising approaches for such systems are mainly based on methods of digital image processing and automated visual inspection. In comparison to other systems which process depth information obtained from 3-D measurements, image based strategies offer considerable advantages: Firstly, appropriate illumination of marks using sophisticated lighting techniques significantly enhances the features and marks of interest. Secondly, some of the unwanted information can be simultaneously suppressed, thus resulting in more meaningful and concise data.[9] Furthermore, since forensic examiners are used to interpret images from their daily work, such a strategy simplifies the supervision of the results obtained from the computerized processing by the forensic examiner. Last but not least, the time required for the image acquisition is considerably shorter than 3-D measurements.

However, the direct comparison of gray-level images taken from the marks is a very inefficient way, since in general, the relevant information of marks can be significantly compressed. Let us consider the example of striation marks, which can be found, e.g., on the circumferential surface of firearm bullets and also as an important portion of toolmarks. Here, the representation can be traced back to the generation process of the marks.[3] A mark can be assumed to be generated by a unidimensional edge, which cuts the groove texture into the ground material. Therefore, a characteristic representation is obtained by reducing the image to a unidimensional signal; see Sect. 2.

For the task of comparing these characteristic signals, correlation methods represent a well-known and reliable approach to detect similarities.[11] They have been widely investigated in theory and successfully applied to numerous similar problems in signal and image processing. Since efficient implementations like those based on the FFT algorithm are available, they meet the requirement of fast data processing. In Sect. 3, a strategy to apply correlation methods to the

Further author information:
E-mail: Michael.Heizmann@mrt.uka.de; WWW: http://www-mrt.mach.uni-karlsruhe.de;
Telephone: +49-721-608-2338; Fax: +49-721-661874



(a) (b) (c)

Figure 1: Image processing of a toolmark image: (a) captured image; (b) homogenized image; (c) straightened image section.

characteristic signals of marks is presented. Since a thorough image acquisition may result in multiple signals, an algorithm is introduced to enhance the comparison results by suitably preprocessing these signals. That way, the performance of correlation methods is considerably increased while keeping the advantage of applying efficient implementations.

Finally, the forensic scientist is interested in a quantitative assessment of comparison strategies. Such an assessment serves not only to compare different strategies or systems, it also helps to optimize parameters of comparison strategies and to investigate the benefit of processing steps. To this end, a benchmarking test is proposed in Sect. 4 and applied to experimental results of the proposed algorithm. It yields different quantitative measures of the achieved recognition rate, which are then discussed with respect to their significance.

2. SIGNATURE EXTRACTION

In order to enable an automated comparison, the information about the groove texture present on the specimen has to be converted into a data record suitable for the computerized evaluation. The processing of forensically relevant marks before the actual comparison step is advisably divided in several stages; see Figs. 1 and 2:

- Firstly, images from the mark have to be acquired containing all the relevant information. This step requires much care in order to preserve the most significant features of the mark which mainly consist in faint structures. If necessary, the image recording has to be assisted by a powerful fusion strategy enabling the extraction of the relevant information from a series of images. Such strategies are of crucial importance, when noticeable illumination problems arise (e.g. with firing pin prints), but they also show significant advantages in most cases of forensically relevant marks.[8]

  In the present case, where striation patterns are the signals of interest, attention should be paid in order to apply a suitable illumination resulting in images of high contrast. A directional lighting, where the illumination is perpendicular to the grooves, obviously meets the requirements best.[3] Experimental results have shown that it is reasonable to record each mark twice, changing the directional illumination to the opposite side of the mark and thus leading to a pair of images for each mark. As an example for the image acquisition, Fig. 1(a) shows an image that was recorded from a casting of a toolmark. Here, only the image with the illumination from the right side is shown. All processing steps are however applied equivalently to the other image of the pair.

- Secondly, a preprocessing eliminates undesired parts of the recorded information. A so-called homogenization of first degree serves to suppress inhomogeneities that arise from the interaction of the illumination process with the reflectance properties of the object as well as its shape.[1] To show the effects of the homogenization of first degree, the strategy has been applied to an example of a toolmark image. Figure 1(b) illustrates that inhomogeneities resulting from the illumination have been suppressed, such that insufficiently illuminated areas of the raw image also obtain satisfactory contrast.

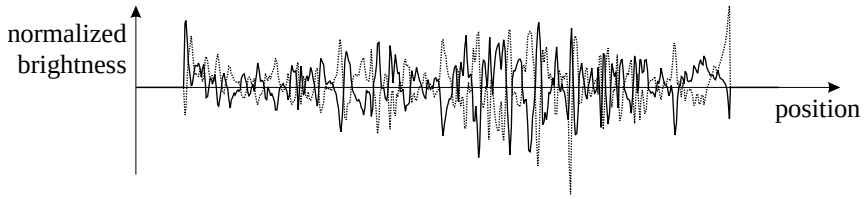


Figure 2: Projection of the straightened grooves: illumination from the left (continuous line) and the right side (dotted line).

- The next step is to extract the relevant features, such that a unidimensional representation of the mark is obtained. For this task of data reduction, model-based strategies have proven excellent suitability.[3] The strategy ensures the straightness of the possibly curved grooves. The straight grooves are then aligned in groove direction, i.e. along the columns of the image.

  A model-based straightening optimized to handle parallel grooves has been applied to the interesting part of the homogenized image for the toolmark example, yielding straight grooves; see Fig. 1(c). The projection of the gray values in the vertical direction along the grooves is depicted in Fig. 2 (continuous line). The dotted line in Fig. 2 shows the result for the same mark, when the lighting is coming from the left side. It can be clearly seen that the profiles are almost symmetric to the zero line. This fact will be exploited in the comparison step in order to improve the reliability of the detection.

- The projection obtained up to now is still not suitable for the automated comparison. Since the comparison is based on the detection of linear similarity by means of correlation techniques, but images of different marks may show non-linear distortions, a further processing of the projections is advisable.[11] Basically, the processing is based on the manipulation of the signal statistics, such that the histograms of the signals equal a given distribution. In Ref. 11, it is shown that such a non-linear transform leads to a significant improvement of the comparison results.

By means of the presented methodology, a meaningful and compact description of the mark is available. Such representations are commonly referred to as signatures. Moreover, since signatures are unidimensional signals instead of two-dimensional images, the computational expense is reduced, thus leading to reasonable computation times in the following comparison step.

3. COMPARISON OF SIGNATURES

The objective of the actual comparison procedure is to provide a reliable measure enabling a quantitative statement on the similarity of two marks. The detection of similarity is carried out on a feature level by comparing the unidimensional signatures of the marks. To this end, the application of the empirical cross-correlation function (CCF)[5]

$$k_{12}(\tau) := \tilde{q}_1(\xi) \circledast \tilde{q}_2(\xi) = \int_{-\infty}^{\infty} \tilde{q}_1(\xi)\, \tilde{q}_2(\xi+\tau)\, \mathrm{d}\xi \tag{1}$$

provides a reliable means to detect similarities between the two signals* $q_1(\xi)$ and $q_2(\xi)$, where

$$\tilde{q}_1(\xi) := \frac{q_1(\xi) - m_{q_1}}{s_{q_1}} \qquad \text{and} \qquad \tilde{q}_2(\xi) := \frac{q_2(\xi) - m_{q_2}}{s_{q_2}} \tag{2}$$

denote the signatures q_i centered around their mean values m_{q_i} and normalized by their standard deviation s_{q_i}. By means of this normalization, the CCF becomes independent of a global offset—the brightness—as well as of a global scaling parameter—the global contrast. Besides, the range of values of the CCF is then limited to $-1 \leq k_{12}(\tau) \leq 1$.

The CCF incorporates two important results that indicate a possible correspondence of the two signatures. The maximum value

$$\rho_{12} := \max\bigl\{k_{12}(\tau)\bigr\} \tag{3}$$

*Since signatures establish signals of finite energy, the integral of Eq. (1) yields a real value.

provides a quantitative measure of the similarity between both signals. The location of the maximum of the CCF $k_{12}(\tau)$ indicates the shift

$$\tau_0 := \arg\max_{\tau} \{k_{12}(\tau)\} \tag{4}$$

leading to the best possible correspondence between the two signatures. This distance can be favorably used for the visual inspection of the correspondence indicated by the maximum value of the CCF, since displaying the two marks at the appropriate position simplifies the inspection.

As mentioned in Sect. 2, each mark is illuminated from two opposite directions. The feature extraction therefore generates two signatures for each mark. These signatures differ from each other in the sequence of bright and dark edges. Although the connection of the illumination angle and the brightness is non-linear and extremely dependent on shadowing effects, one signature of the pair resembles the *negative* of the other one.

Example: To visualize this correspondence in a simple case, the groove texture is modelled by a surface with a harmonic profile

$$z(\mathbf{x}) = A \cdot \cos x\,, \quad \mathbf{x} = (x,\, y)^{\mathrm{T}}\,, \tag{5}$$

which is observed by a camera in the direction of the surface normal.[9] Figure 3 shows the cross section of the surface with the x, z-plane. The illumination is chosen as directional light perpendicular to the groove direction with the illumination vector

$$\mathbf{l}_{\mathrm{i}} = (-\sin\theta_{\mathrm{i}},\, 0,\, -\cos\theta_{\mathrm{i}})^{\mathrm{T}}\,, \tag{6}$$

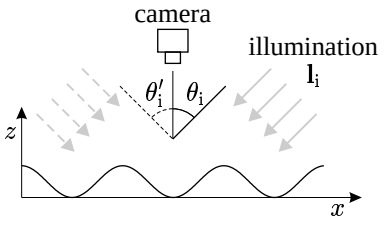


Figure 3: Harmonic test profile.

where θ_{i} denotes the elevation angle of the incident light coming from the right side; see Fig. 3, continuous arrows. In the absence of shading, the surface radiance $L_{\mathrm{r,right}}(x)$ seen by the camera for a perfectly diffuse reflecting (Lambertian) surface—this assumption is well suitable for dull materials like casting materials—can be written as

$$L_{\mathrm{r,right}}(x) \propto \frac{A \sin\theta_{\mathrm{i}} \sin x + \cos\theta_{\mathrm{i}}}{\sqrt{1 + A^2 \sin^2 x}}\,. \tag{7}$$

When the lighting is switched to the opposite direction (see Fig. 3, dashed arrows), the surface radiance changes to

$$L_{\mathrm{r,left}}(x) \propto \frac{-A \sin\theta'_{\mathrm{i}} \sin x + \cos\theta'_{\mathrm{i}}}{\sqrt{1 + A^2 \sin^2 x}}\,. \tag{8}$$

From Eqs. (7) and (8) follows that

$$L_{\mathrm{r,right}}(x) \cdot \sqrt{1 + A^2 \sin^2 x} - \cos\theta_{\mathrm{i}} \quad \propto \quad -\left(L_{\mathrm{r,left}}(x) \cdot \sqrt{1 + A^2 \sin^2 x} - \cos\theta'_{\mathrm{i}}\right)\,, \tag{9}$$

which indicates that there is a *qualitative* connection between $L_{\mathrm{r,right}}(x)$ and $-L_{\mathrm{r,left}}(x)$. Figure 4 depicts the surface radiance of the harmonic test texture in case of light coming from the right (continuous line) and the inverted radiance in case of light coming from the left side (dashed line). The qualitative connection between them is clearly recognizable. $\diamond$

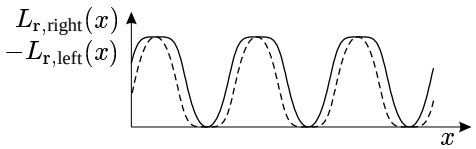


Figure 4. Calculated surface radiances of the harmonic test profile of Fig. 3 for the illumination with directional light from different directions.

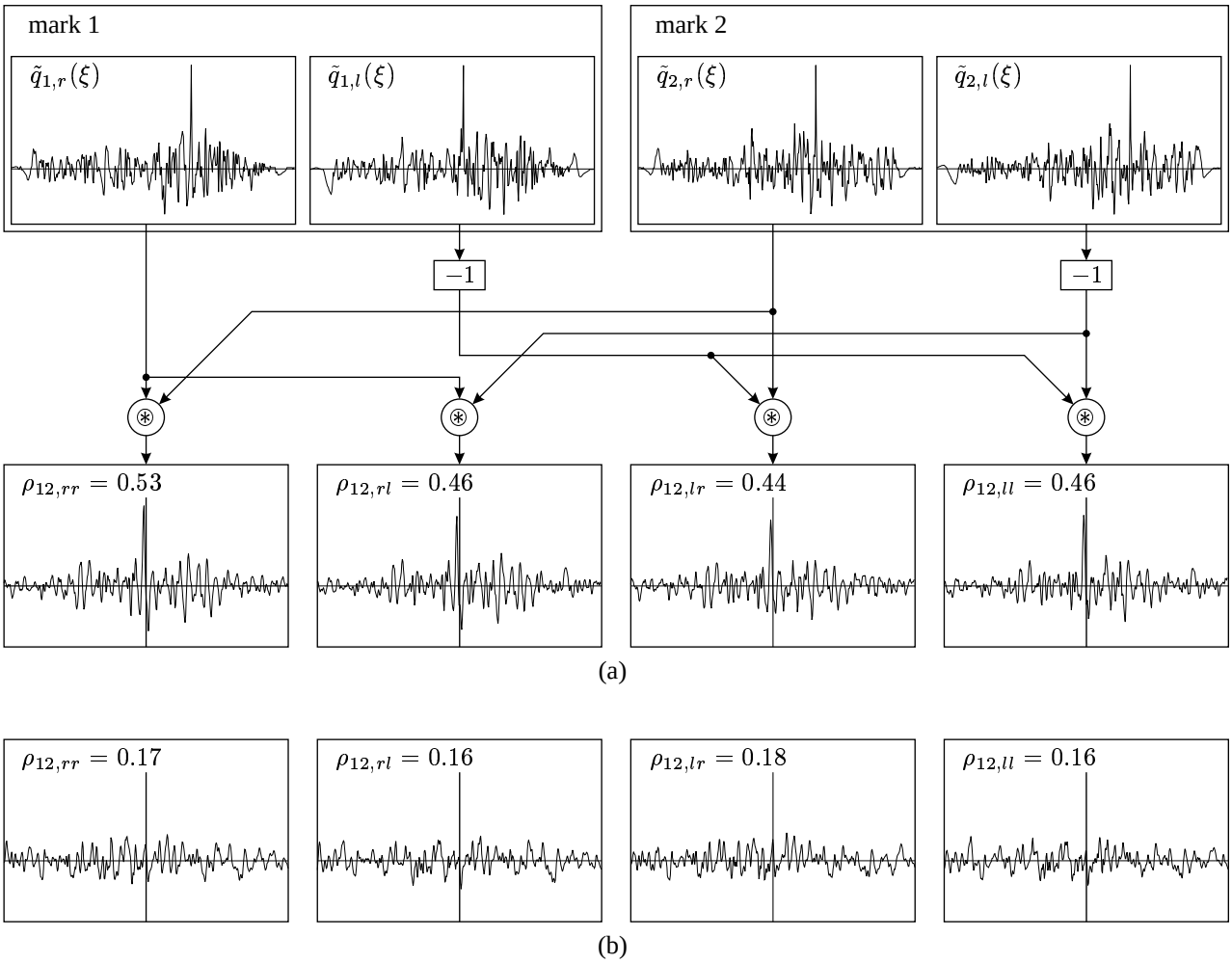


Figure 5. Comparison results: (a) signatures of toolmarks belonging to the same tool (upper row); resulting cross-correlation function (lower row), the thin vertical line indicates $\tau = 0$; (b) cross-correlation functions for toolmarks from different tools.

This means that the comparison of two marks should consider both signatures of a mark, such that the signature obtained from a certain illumination—e.g., when the light is coming from the left side—is always inverted. Thus, a comparison of two marks comprises four correlations, where each signature $\tilde{q}_{1,\alpha}$, $\alpha \in \{r, l\}$ of the first mark is compared with each signature $\tilde{q}_{2,\beta}$, $\beta \in \{r, l\}$ of the other one. The effective measure of similarity is then obtained to the maximum value of the respective maximum values of the CCFs:

$$\rho_{12,\text{eff}} := \max_{\alpha,\beta} \left\{ \rho_{12,\alpha\beta} \right\} \tag{10}$$

with

$$\rho_{12,\alpha\beta} := \begin{cases} \max\left\{ \tilde{q}_{1,\alpha} \circledast \tilde{q}_{2,\beta} \right\}, & \alpha = r\,, \quad \beta = r\,, \\ \max\left\{ \tilde{q}_{1,\alpha} \circledast (-\tilde{q}_{2,\beta}) \right\}, & \alpha = r\,, \quad \beta = l\,, \\ \max\left\{ (-\tilde{q}_{1,\alpha}) \circledast \tilde{q}_{2,\beta} \right\}, & \alpha = l\,, \quad \beta = r\,, \\ \max\left\{ \tilde{q}_{1,\alpha} \circledast \tilde{q}_{2,\beta} \right\}, & \alpha = l\,, \quad \beta = l\,. \end{cases} \tag{11}$$

Figure 5(a) shows the comparison of two toolmarks originated from the same tool. Each toolmark is represented by a pair of signatures with different illumination (upper row). The signatures are then processed with the strategy of Eq. (11).

The resulting cross-correlation functions are depicted in the lower row of Fig. 5(a). All of them show pronounced maxima at the same shift distance. Moreover, the absolute values of the maxima lie within the same range. In contrast, typical results for the correlation functions of toolmarks from different tools are shown in Fig. 5(b). In such cases, the maxima are much lower and less distinct than in the case of toolmarks from the same tools. As a consequence, the range of the absolute values supplies a reliable tool for the detection of correspondences of marks.

To apply the proposed method to digital data like the signatures, the calculation of the CCF according to Eq. (1) can be implemented efficiently in the frequency domain based on the FFT algorithm.[2] However, an important property of the discrete Fourier transform is the inherent periodic continuation of the signals. In the case that signatures from firearm bullets are evaluated, the periodic continuation is a desired feature, since the circumferential surface results in a cyclic signature. In contrast, when toolmarks are examined, the periodic continuation lacks any physical reasoning. One possible approach is to pad the signature to its double length by appending zero values for the latter case.[2] An important advantage of the FFT algorithm is its low computational effort, which meets the requirement of a fast comparison algorithm. The calculation time for a correlation of two signals containing 512 samples is presently about 1 ms on a standard PC, which enables a complete search in a database containing 15,000 marks in one minute.

However, the comparison of toolmarks by means of the CCF requires that their relative orientations are identical. In the case of striation marks on firearm bullets, the orientation is clearly defined by the shape of the bullet. In general, when striation patterns caused by toolmarks are examined, such a unique indication is not available. The simple application of the CCF to the signatures with the accidental orientation during the image acquisition may thus yield poor identification results. Consequently, the comparison should be extended, such that the determination of the CCF by means of Eq. (1) is also done with one signal inverted in the ξ-direction.

By means of the comparison strategy proposed above, the measure $\rho_{12,\text{eff}}$ is obtained which quantifies the similarity between two single marks. In practice, however, a common task is to investigate the connection of a newly found mark with a given archive of marks. Consequently, a series of comparisons has to be performed where the new exhibit is tested with each mark of the database. Based on these individual comparisons, a hitlist can be generated by sorting the database in the most probable order, where marks with the largest measure of similarity are placed on top of the list.

In this context, it should be pointed out that the algorithms for the generation of the hitlist do not decide whether or not a real hit can be found in the archive. The minimum required similarity for assuming two marks to stem from an identical source still remains in the responsibility of the examiner.

4. ASSESSMENT OF COMPARISON STRATEGIES

For the task of assessing a comparison strategy, a quantitative measure is required which faithfully describes the performance of the strategy. The ordering of hits in the hitlist generated by the comparison strategy can be used for this purpose. It is obviously desirable to achieve an accumulation of true hits on the first positions of the hitlist. In the ideal case, all n existing hits are concentrated at the first n positions of the hitlist. Therefore, an assessment on the basis of a measure of concentration with a dedicated marks database for test purposes can be useful and is discussed in this section.

4.1. Concentration Curve

Several approaches are known in statistics literature for the evaluation of concentration in a set of data. A particularly comprehensive way to represent concentration is obtained by plotting it in a so-called concentration curve.

The concentration curve is generated by plotting the ratio of finding a true hit as a function of the portion of the database to be examined; see Fig. 6(a). Thus, it indicates for each relative position of the sorted archive the cumulative probability of finding the hit when the archive is searched up to this position. Since the concentration curve represents a cumulative distribution function, it is a monotonously growing function, starting from $(0/0)$—no hit is found before the first position is considered—and ending in $(100\%/100\%)$—all hits are found after searching the whole database. In case that no concentration of hits is present—i.e. an unsorted database—, the expected concentration curve meets the bisector. The better a strategy of sorting the database—automated or by experts—performs, the farther the curve leaves the bisector to the upper left corner. In Fig. 6(b), the meaning of the concentration curve is demonstrated with three examples showing typical cases:[10]

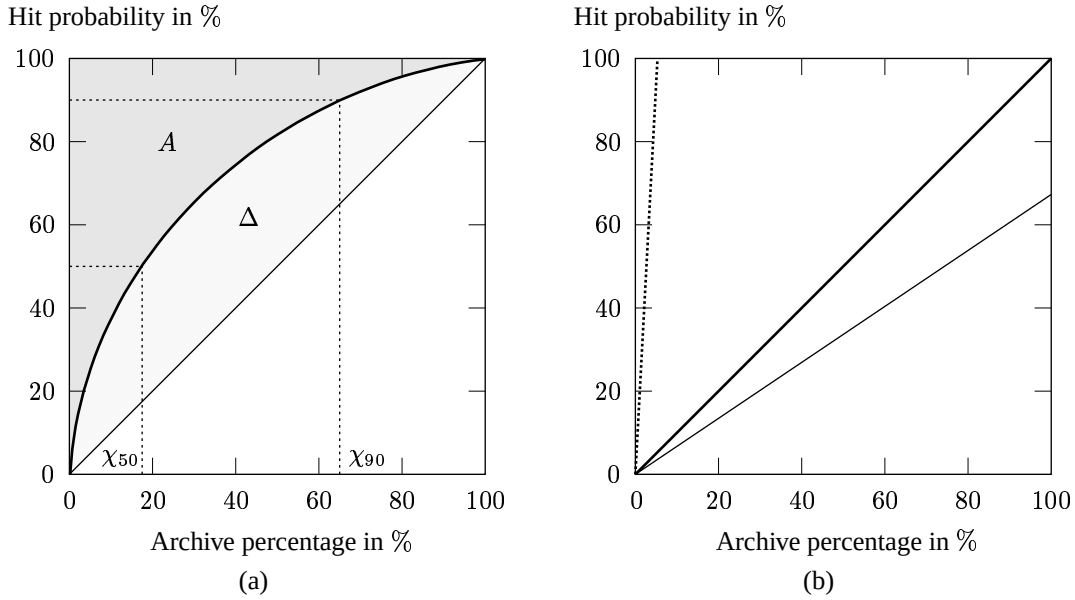


Figure 6. Concentration curves: (a) assessment principle; (b) concentration curves for an ideal examiner (bold line), real examiner (thin line), and for an examiner searching a database sorted by an ideal automated system (dotted line).

The *ideal forensic examiner*—represented by the bold line—makes no errors in his search and thus does not oversee any existing correspondence. When he looks at the database containing one real hit, the cumulative probability of finding the hit grows exactly with the portion of the database already searched. After the entire archive has been examined, the hit is reliably identified, which leads to a hit probability of 100%. Thus, the typical concentration curve of the ideal forensic examiner coincides with the diagonal.

In reality, however, the *real forensic examiner*—represented by the thin line—makes mistakes in his comparisons. Consequently, he oversees existing correspondences, which makes it less probable in comparison to the ideal forensic examiner that an existing hit is found at a certain archive percentage. Even after the whole archive has been examined, it is still uncertain whether an existing hit is recognized, i.e. the hit probability at the end of the archive is less than 100%. Although the actual error rate of the real examiner is unknown, it can be assumed that the imperfect examination performed by the human expert leads to a significantly lower hit probability.

The *ideal automated system*—represented by the dotted line—leads to a perfect sorting of the archive, where the matching mark is always placed at the first position of the hitlist now holding the place of the unsorted archive. Since the first entry in the hitlist already reveals the actual hit, the amount of specimens in the database to be searched by the forensic examiner is reduced to 1. Consequently, the concentration curve meets 100% already at the archive percentage representing one entry.

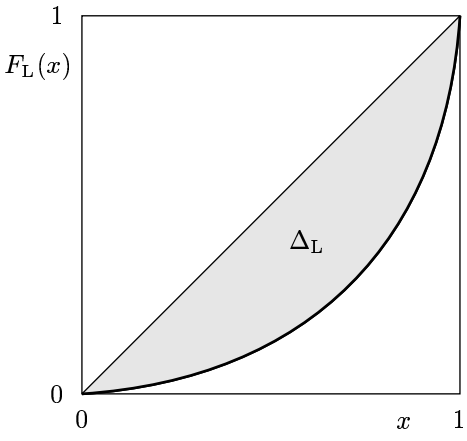


Figure 7: Lorenz curve.

The concentration curve used for the assessment of comparison methodologies in this paper resembles to some extent the *Lorenz curve*.[4] The Lorenz curve is based on a list $\mathcal{L} = \{x_1, \ldots, x_N\}$ sorted according to the feature characteristic $f_{\mathrm{L}}(.)$ in *ascending* order:

$$f_{\mathrm{L}}(x_1) \leq f_{\mathrm{L}}(x_i) \leq f_{\mathrm{L}}(x_N)\,; \quad 1 < i < N\,, \quad \sum_i f_{\mathrm{L}}(x_i) = 1\,. \tag{12}$$

As a consequence, the Lorenz curve is always convex to the abscissae; see Fig. 7. This characteristic is easy to prove for the continuous case. Here, the Lorenz curve can be described by the cumulative distribution

$$F_{\mathrm{L}}(x) = \int_0^x f_{\mathrm{L}}(\xi)\,\mathrm{d}\xi \tag{13}$$

with the continuous frequency distribution $f_{\mathrm{L}}(x) \geq 0$ where

$$\frac{\mathrm{d}f_{\mathrm{L}}(x)}{\mathrm{d}x} \geq 0 \tag{14}$$

per definition. Since

$$\frac{\mathrm{d}F_{\mathrm{L}}(x)}{\mathrm{d}x} = f_{\mathrm{L}}(x) \geq 0 \quad \text{and} \quad \frac{\mathrm{d}^2F_{\mathrm{L}}(x)}{\mathrm{d}x^2} = \frac{\mathrm{d}f_{\mathrm{L}}(x)}{\mathrm{d}x} \geq 0\,, \tag{15}$$

the angle between the Lorenz curve and the abscissae is always positive and acute, and the angle increases with increasing x.

Although the Lorenz curve is characterized by a very similar definition to the concentration curve, the ordering of the abscissae elements reveals an important difference: Since the entries of the hitlist—the abscissae elements—are sorted by the comparison methodology, which is a posteriori information in the sense of Bayesian statistics, but the concentration curve represents the real—a priori—correspondences, the precondition of a monotonously decreasing—or increasing in the case of the Lorenz curve—frequency distribution is in general not valid. Nevertheless, some index numbers known from the Lorenz curve are also applicable and useful in the present case.

4.2. Index Numbers

Since the overall performance of a comparison methodology is of particular interest, index numbers describing the curve in its entirety are of great interest. Such a description is favorably obtained by processing the integral of the curve. This consideration coincides with the evaluation of areas in the diagram of the concentration curve.

From statistics literature concerned with the Lorenz curve, *Gini's coefficient* C_{G} is widely known as

$$C_{\mathrm{G}} := 2 \cdot \Delta_{\mathrm{L}}\,, \tag{16}$$

where Δ_{L} denotes the area between the diagonal and the Lorenz curve; see Fig. 7. Since the overall area of the Lorenz diagram equals 1, Gini's coefficient obviously lies within the range of $0 \leq C_{\mathrm{G}} \leq 1$ for real Lorenz curves, increasing with rising concentration. For the application to the concentration curves in the present case, the area Δ has to be generalized. The area Δ is here defined by

$$\Delta := \int_0^1 F(x)\,\mathrm{d}x - \frac{1}{2}\,, \tag{17}$$

denoting the *signed* area between the concentration curve and the diagonal; see Fig. 6(a), where

$$F(x) = \int_0^x f(\xi)\,\mathrm{d}\xi \tag{18}$$

is the cumulative distribution of $f(x)$ given by the sorting of the archive. Hence, C_{G} takes values $-1 \leq C_{\mathrm{G}} \leq 1$.

The *Herfindahl index*[†] is a common measure of concentration, which is applied in several definitions. One usual implementation uses the integral over the squared frequencies, which is defined in the continuous case by means of the frequency distribution $f(x)$ to

$$C_{\mathrm{H}} := \int_0^1 f^2(x)\,\mathrm{d}x\,. \tag{19}$$

[†]The Herfindahl index is closely connected with the Herfindahl-Hirschman-Index (HHI), which is used e.g. by the U.S. Department of Justice and associated authorities to measure market concentration for purposes of antitrust enforcement.

However, since the Herfindahl index is defined with the squared frequency without a statement of where the respective frequency occurs, no decision can be made whether the accumulation of hits is obtained at the desired place, i.e. on top of the hitlist. This disadvantageous behaviour will be seen in an example shown below.

The most common index known for the evaluation of concentration curves in forensic science is the *Rosenbluth index* C_{R}. It is defined by

$$C_{\mathrm{R}} := \frac{1}{2A}\,, \tag{20}$$

where A represents the area above the concentration curve; see Fig. 6(a). In case of a continuous frequency distribution $f(x)$, the Rosenbluth index can be calculated by means of

$$A = \int_0^1 \left[1 - F(x)\right]\,\mathrm{d}x\,. \tag{21}$$

The area A also represents the average portion of the archive to be examined in order to find an existing correspondence. Hence, indices based on the evaluation of this area—or the counterpart to it like Gini's coefficient—can be intuitively interpreted.

Example: Three remarkable cases will now serve to illustrate the meaning of the indices: In the first example, a missing concentration of hits is considered such that the hits are *equally distributed* over the archive. Hence, the cumulative distribution and the respective continuous frequency distribution are given by

$$F_{\mathrm{I}}(x) := x \quad \text{and} \quad f_{\mathrm{I}}(x) := \frac{\mathrm{d}F_{\mathrm{I}}(x)}{\mathrm{d}x} = 1\,; \quad x \in [0;1]\ . \tag{22}$$

Equations (16), (19) and (20) yield

$$C_{\mathrm{G,I}} = 0\,, \quad C_{\mathrm{H,I}} = 1 \quad \text{and} \quad C_{\mathrm{R,I}} = 1\,. \tag{23}$$

When the hits are *ideally concentrated* on top of the hitlist, the certainty of finding the hit is reached already at the beginning of the hitlist. In the continuous case, the resulting probability functions may be modelled with the limits

$$F_{\mathrm{II}}(x) := \lim_{\epsilon\to 0^+} H(x-\epsilon) \quad \text{and} \quad f_{\mathrm{II}}(x) := \lim_{\epsilon\to 0^+} \delta(x-\epsilon)\,; \quad x \in [0;1]\ , \tag{24}$$

where

$$H(x) := \begin{cases} 1\,, & x \geq 0\,, \\ 0\,, & x < 0 \end{cases} \tag{25}$$

denotes the Heaviside step function and $\delta(x)$ is known as the Dirac impulse with the properties[6]

$$\delta(x) := 0\,; \quad x \neq 0 \quad \text{and} \quad \int_{-\infty}^{\infty} \delta(x)\,\mathrm{d}x := 1\,. \tag{26}$$

Gini's index yields a real value of $C_{\mathrm{G,II}} = 1$. Since the Dirac impulse may also be defined by the limit

$$\delta(x) := \lim_{\omega\to\infty} \frac{\sin \omega x}{\pi x}\,, \tag{27}$$

the integration of the squared Dirac impulse yields

$$\int_{-\infty}^{\infty} \delta^2(x)\,\mathrm{d}x = \lim_{\omega\to\infty} \int_{-\infty}^{\infty} \frac{\sin^2 \omega x}{(\pi x)^2}\,\mathrm{d}x = \lim_{\omega\to\infty} \frac{\omega}{\pi}\,. \tag{28}$$

Thus, the Herfindahl index tends to infinity. The Rosenbluth index is obtained to

$$C_{\mathrm{R,II}} = \lim_{\epsilon\to 0^+} \frac{1}{2\epsilon}\,, \tag{29}$$

Table 1: Comparison of common concentration indices with test distributions.

	equally distributed	ideally concentrated	inversely concentrated
Gini	0	1	-1
Herfindahl	1	∞	∞
Rosenbluth	1	∞	$\frac{1}{2}$

tending also to infinity.

The last example shows the situation of *inversely concentrated* hits. Here, the probability functions are

$$F_{\mathrm{III}}(x) := \lim_{\epsilon \to 0^+} H(x + \epsilon - 1) \quad \text{and} \quad f_{\mathrm{III}}(x) := \lim_{\epsilon \to 0^+} \delta(x + \epsilon - 1)\,; \quad x \in [0; 1]\ , \tag{30}$$

resulting in

$$C_{\mathrm{G,III}} = -1 \quad \text{and} \quad C_{\mathrm{R,III}} = \frac{1}{2}\,. \tag{31}$$

The Herfindahl index behaves exactly the same as in the second example, since Eq. (28) holds independently of the argument of $\delta(.)$.

Table 1 summarizes the results achieved for the examples. Gini's coefficient and the Rosenbluth index both show suitable behaviours for the assessment of concentration curves. As mentioned above, it is true that the Herfindahl index indicates a concentration, but it does not take the location of the concentration into consideration. Hence, Gini's coefficient and the Rosenbluth index should be preferred for the purpose of assessing comparison strategies. $\diamond$

Other methods to assess the performance of comparison methodologies are also commonly used in practice.[9] For example, the German forensic scientist use the measures χ_{50} and χ_{90} which specify the portion of the archive to be searched in order to obtain the hit with a probability of 50% and 90%, respectively; see Fig. 6(a). In addition, the measures p_1 and p_5 are used which describe the probability of finding a hit at the first position and within the first five positions, respectively. Especially the last-mentioned measures strongly depend on the archive size, thus making these measures less reliable than the global indices of Gini and Rosenbluth.

4.3. Experimental Results

The application of the index numbers introduced in Sect. 4.2 to assess a comparison methodology is favorably done by means of an archive of test marks, where the correspondences between the marks are known a priori. Such an archive consists in the case of toolmarks of a set $\mathcal{A}$ of marks from different tools

$$\mathcal{A} = \bigcup_q \mathcal{T}_q\,, \tag{32}$$

where $\mathcal{T}_q$ denotes the set of marks produced by the tool q.

To determine the concentration curve for the present task, each mark $i \in \mathcal{T}_q \in \mathcal{A}$ of the given test archive is considered as an exhibit. The remainder of the archive represents the actual database. Hence, the respective exhibit i is compared with all marks $j \in \mathcal{A}$, $j \neq i$, resulting in a hitlist for each mark. The hitlist itself can be described by a function $k = L_i(r)$ which indicates for each position r in the hitlist—also referred to as rank—the corresponding mark k. By eliminating all marks which originate from the identical tool q as the mark i except one, the modified hitlist $k = \tilde{L}_i(r)$ is obtained which now contains exactly one true hit. Based on this modified hitlist, the absolute and the relative rank of the mark $k \in \mathcal{T}_q$ are found to

$$r_{i,\mathrm{abs}}(k) := \tilde{L}_i^{-1}(k) \quad \text{and} \quad r_{i,\mathrm{rel}}(k) := \frac{\tilde{L}_i^{-1}(k)}{|\mathcal{A}| - |\mathcal{T}_q| + 1}\,. \tag{33}$$

The empirical frequency distribution $f(r)$ can be calculated with the density $\tilde{f}(r)$ by means of

$$f(r) := \frac{\tilde{f}(r)}{Z} \quad \text{and} \quad \tilde{f}(r) := \sum_{\mathcal{T}_q} \sum_{\substack{i,k \in \mathcal{T}_q \\ i \neq k}} \delta_r^{r_{i,\text{rel}}(k)} \delta\big(r - r_{i,\text{rel}}(k)\big)\,, \tag{34}$$

where

$$\delta_a^b = \begin{cases} 1\,, & a = b\,, \\ 0\,, & a \neq b \end{cases} \tag{35}$$

denotes the Kronecker symbol, and the total number of combinations of marks from the same tool is given by

$$\begin{aligned} Z &:= \int_0^1 \tilde{f}(\xi)\,\mathrm{d}\xi \\ &= \sum_{\mathcal{T}_q} 2\binom{|\mathcal{T}_q|}{2} = \sum_{\mathcal{T}_q} |\mathcal{T}_q|\,\big(|\mathcal{T}_q| - 1\big)\,. \end{aligned} \tag{36}$$

Equation (34) performs a normalization of the density $\tilde{f}(r)$ to the value of 1. The multiplication with the Dirac impulse $\delta(.)$ in Eq. (34) provides for real values of Z in the integration of Eq. (36); see the definition in Eq. (26). The empirical cumulative distribution and the indices based on it are then given by Eq. (18). An alternative way to Eq. (21) to compute the area A is given by

$$A := \frac{1}{Z} \sum_{\mathcal{T}_q} \sum_{\substack{i,k \in \mathcal{T}_q \\ i \neq k}} r_{i,\text{rel}}(k)\,, \tag{37}$$

since A also represents the average portion of the archive that has to be searched until the hit is found.

The measures χ_{50} and χ_{90} can be formulated by inverting the empirical cumulative distribution:

$$\chi = F^{-1}\big(F(\chi)\big)\,. \tag{38}$$

The measures p_1 and p_5 can be calculated similarly to Eq. (34), but with a summation over the absolute rank $r_{i,\text{abs}}(k)$ by means of

$$p_n = \frac{1}{Z} \sum_{\mathcal{T}_q} \sum_{\substack{i,k \in \mathcal{T}_q \\ i \neq k}} \sum_{r=1}^{n} \delta_r^{r_{i,\text{abs}}(k)}\,. \tag{39}$$

The proposed assessment strategy has been applied to test the comparison methodology presented in Sect. 3. To this end, a specific archive of toolmarks has been provided by a German forensic institute. Signatures were generated from image series taken from each mark using the techniques described in Refs. 3, 7, 9 and stored in a database.

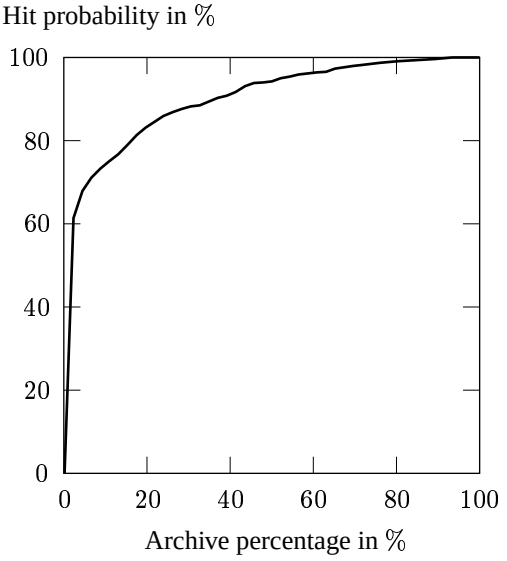


Figure 8. Concentration curve obtained from experimental data with the proposed strategies.

The resulting concentration curve obtained by integrating Eq. (34) is depicted in Fig. 8. It is clearly visible that the strategy leads to a significant concentration of hits on the first positions of the hitlists, which can also be seen from the index numbers in Table 2. On average, only 11.24% of a hitlist has to be examined to find the hit, which is equivalent to the area A; see Sect. 4.2. Gini's coefficient and the Rosenbluth index are computed as $C_\text{G} = 77.52\%$ and $C_\text{R} = 4.45$, respectively. The Herfindahl index is obtained as $C_\text{H} = 17.73$, which indicates—with the restrictions mentioned in Sect. 4.2—a significant concentration of hits. In order to find 50% of the hits, it is sufficient to test on the average $\chi_{50} = 2.17\%$ of the hitlist, which actually equals the first position of the hitlist. If 90% of the correspondences are to be found, $\chi_{90} = 36.96\%$—almost only one third—of the hitlist have to be checked. At the first position of the hitlist, a hit is already found with a probability of $p_1 = 61.34\%$. Within the first five positions, the probability of finding a correspondence increases to $p_5 = 75.00\%$.

Table 2: Assessment results for the proposed comparison strategy.

A	C_{G}	C_{R}	C_{H}	χ_{50}	χ_{90}	p_1	p_5
11.24%	77.52%	4.45	17.73	2.17%	36.96%	61.34%	75.00%

5. CONCLUSIONS

In this paper, a method for the automated comparison and recognition of striation marks has been presented. This method is mainly intended to process striated toolmarks, but since marks on firearm bullets are very similar, it is also well applicable to such cases. Based on gray-level images taken under different illumination conditions, characteristic features are extracted, which are commonly referred to as signatures. The extracted signatures are then stored in an archive of marks.

For the actual comparison step, these signatures serve as compact representations of the marks, thus enabling efficient and fast comparison algorithms. The detection of similarities between the signatures is done by correlation methods, which have proven to provide for meaningful measures of similarity in reasonable computation times. In order to enhance the recognition rate, a strategy to handle images taken under different illumination conditions has been presented.

For the important task of assessing comparison strategies, a benchmark has been proposed. The assessment is based on a synthetical test archive of marks, which is evaluated by the strategy to be assessed. By means of a statistical analysis of the achieved comparison results, quantitative measures on the performance of comparison strategies are provided. Meaningful index numbers are obtained, thus resulting in concise statements. The presented methodology for the automated recognition of marks has been assessed by means of this test with a database of toolmarks. The result proves the good performance of the presented comparison strategy.

ACKNOWLEDGMENTS

The author would like to thank Dr. Fernando Puente León and Prof. Christoph Stiller for valuable discussions on the subjects of this paper. He also acknowledges the BKA as well as the LKA Berlin PTU for their practical support.

REFERENCES

1. J. Beyerer and F. Puente León, "Suppression of inhomogeneities in images of textured surfaces," *Optical Engineering* **36** (1), pp. 85–93, 1997.
2. E.O. Brigham, *The Fast Fourier Transform and its Applications*, Prentice-Hall, Englewood Cliffs, New Jersey, 1988.
3. M. Heizmann and F. Puente León, "Model-based analysis of striation patterns in forensic science," In: *Enabling Technologies for Law Enforcement and Security*, S.K. Bramble, E.M. Carapezza, and L.I. Rudin (eds.), Proceedings of SPIE **4232**, pp. 533–544, 2001.
4. M. Kendall and A. Stuart, *The Advanced Theory of Statistics*, Griffin & Co., London, 1977.
5. A. Papoulis, *Probability, Random Variables, and Stochastic Processes*, McGraw-Hill, New York, 1991.
6. A. Papoulis, *The Fourier Integral and its Applications*, McGraw-Hill, New York, 1994.
7. F. Puente León, "Enhanced imaging by fusion of illumination series," In: *Sensors, Sensor Systems, and Sensor Data Processing*, O. Loffeld (ed.), Proceedings of SPIE **3100**, pp. 297–308, 1997.
8. F. Puente León and J. Beyerer, "Active vision and sensor fusion for inspection of metallic survaces," In: *Intelligent Robots and Computer Vision XVI: Algorithms, Techniques, Active Vision, and Materials Handling*, D.P. Casasent (ed.), Proceedings of SPIE **3208**, pp. 394–405, 1997.
9. F. Puente León, *Automatische Identifikation von Schußwaffen*, VDI Verlag, Düsseldorf, 1999.
10. F. Puente León and J. Beyerer, "Automatic comparison of striation information on firearm bullets," In: *Intelligent Robots and Computer Vision XVIII: Algorithms, Techniques, and Active Vision*, D.P. Casasent (ed.), Proceedings of SPIE **3837**, pp. 266–277, 1999.
11. F. Puente León and M. Heizmann, "Strategies to detect non-linear similarities by means of correlation methods," In: *Intelligent Robots and Computer Vision XX: Algorithms, Techniques, and Active Vision*, D.P. Casasent and E.L. Hall (eds.), Proceedings of SPIE **4572**, pp. 513–524, 2001.