Science & Justice 59 (2019) 390–404

Contents lists available at ScienceDirect

Science & Justice

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

A Bayesian approach based on Kalman filter frameworks for bullet identification

H. Danandeh Hesar, S. Bigdeli, M. Ebrahimi Moghaddam*

Faculty of Computer Science and Engineering, Shahid Beheshti University G.C, Tehran, Iran

ARTICLE INFO

Keywords:
Automatic bullet identification
Cross correlation
Ensemble empirical mode decomposition
Kalman filter

ABSTRACT

When a bullet is fired from a barrel, random imperfections in the interior surface of the barrel imprint 3-D micro structures on the bullet surface that are seen as striations. Despite being random and non-stationary in nature, these striations are known to be consistently reproduced in a unique pattern on every bullet. This is a key idea in bullet identification. Common procedures in the field of automatic bullet identification include extraction of a feature profile from bullet image, profile smoothing and comparison of profiles using normalized cross correlation. Since the cross correlation based comparison is susceptible to high-frequency noise and nonlinear baseline drift, profile smoothing is a critical step in bullet identification. In previous work, we considered bullet images as nonlinear non-stationary processes and applied ensemble empirical mode decomposition (EEMD) as a preprocessing algorithm for smoothing and feature extraction. Using EEMD, each bullet average profile was decomposed into several scales known as intrinsic mode functions (IMFs). By choosing an appropriate range of scales, the resultant smoothed profile contained less high-frequency noise and no nonlinear baseline drift. But the procedure of choosing the proper number of IMFs to reduce the high-frequency noise effect was manual. This poses a problem in comparison of bullets whose images contained less or more noise in comparison to others because their useful information may be present in the corresponding discarded IMFs. Moreover, another problem arises when the bullet type changes. In this case manual inspection is needed once more to figure out which range of IMFs contain less high-frequency noise for this particular type of bullet. In this paper, we propose a novel combination of EEMD and Bayesian Kalman filter to solve these problems. First the bullet images are rotated using Radon transform. The rotated images are averaged column-wise to acquire averaged 1-D profiles. The nonlinear baseline drifts of averaged profiles are removed using EEMD algorithm. The profiles are then processed by a Kalman filter that is designed to automatically and optimally reduce the effect of high-frequency noise. Using Expectation Maximization (EM) technique, for each averaged profile, the parameters of Kalman filter are reconfigured to optimally suppress the high-frequency noise in each averaged profile. This work is the first effort that practically implements the Kalman filter for optimal denoising of firearm image profiles. In addition, we believe that Euclidean distance metric can help the normalized cross-correlation based comparison. Therefore, in this paper, we propose a comparison metric that is invariant to start and endpoints of firearm image profiles. This metric combines the prized properties of both Euclidean and normalized cross-correlation metrics in order to improve identification results. The proposed algorithm was evaluated on a database containing 180 2-D gray-scale images acquired from bullets fired from different AK-47 assault rifles. Although the proposed method needs more calculations in comparison to conventional methods, the experiments showed that it attained better results compared with the conventional methods and the previous method based on EMD in the field of automatic bullet identification.

1. Introduction

When bullets are fired from a firearm, the interior surface of barrel makes forcible contact with them creating characteristic tool marks on their surfaces called ‘striation signatures’. It is believed that these signatures are unique for each firearm. This is one of the founding theories in the field of firearm identification [1]. Since 1930, firearm examiners have compared these microscopic signatures using comparison microscopes and determined whether a pair of bullets was fired or ejected from the same firearm. Their judgments have helped relating

* Corresponding author.
E-mail address: m_moghadam@sbu.ac.ir (M.E. Moghaddam).

https://doi.org/10.1016/j.scijus.2019.02.009
Received 3 October 2018; Received in revised form 15 February 2019; Accepted 24 February 2019

recovered firearms or other firearm evidences to criminal acts. Manual comparisons of ballistic evidences are time demanding and highly subjective to firearm examiner's experience and expertise. The introduction of commercial automated ballistics identification systems such as IBIS [2] and Drugfire [3] in the early 1990s has revolutionized the firearm identification procedure as they offered much faster comparison results for very large databases and reduced the time and effort for firearm examiners. Such systems are capable of comparing a large number of ballistic evidences automatically in a short amount of time. These systems deliver a limited number of most likely candidates for examiners to inspect. This provides clearer vision for firearm examiner to decide whether the compared evidences match the subject evidence or not. The field of automatic firearm identification has a promising future with a hope that one day, firearm identifications may be accomplished or affirmed thoroughly using automated searches and matches. The striation patterns on a bullet surface are often seen as a set of straight lines with a certain angle (twist angle) to bullet's main axis. This angle is the result of circular rifling of interior surface of barrel and varies between different firearm types and brands. In the field of computerized bullet identifications, several methods utilized striation signatures on the surface of bullet as distinctive features [4–11]. These methods can be divided into two groups: 2-D based and 3-D based. 2-D based methods use 2-D images for comparison. On the other hand, 3-D based methods, use 3-D topography images for comparison. In 2-D based methods, the first step usually includes image rotation to make the main striations be perpendicular to horizontal line. After image smoothing, 1-D average profiles were extracted from rotated images using column-wise averaging. In the final step the comparison of the averaged profiles was accomplished using normalized cross-correlation metric [4,6,7,12,13]. Since normalized cross-correlation metric is sensitive to noise, profile smoothing was considered as a key step in 2-D methods. In [4], image smoothing and feature extraction were achieved using a Linear Time Invariant (LTI) Gaussian high-pass filter and top-hat morphologic transform respectively. A high-pass filter is a system which removes low-frequency components from signal and maintains its high-frequency properties and the top-hat morphologic transform is a nonlinear morphological method that uses morphological operations such as “erosion” and “dilation” in order to eliminate low-frequency components in signals and keep their high-frequency components. In [6], an LTI Gaussian band-pass filter was utilized to smooth topography images and profile extraction was accomplished using a correlation-based strategy which was able to automatically determine the areas containing valuable striae information and discard other non-informative regions. In [7,8] the preprocessing steps were similar to [6] but in order to improve matching results, a parameter named “striation density” was introduced to evaluate the areas to be averaged. In addition, in [8], instead of cross correlation metric, a criterion named ‘consecutively matching striae (CMS)’ [14] was utilized for comparison of topography measurements. Because striation patterns are random and non-stationary in nature, LTI filtering frameworks are not suitable for bullet profile smoothing. Therefore, Bigdeli et al. proposed to apply ensemble empirical mode decomposition (EEMD) for profile smoothing [10]. In [10], first, 2-D gray–scale bullet images were rotated using Radon transform. After rotation and column wise averaging, using EEMD, the 1-D averaged profiles were decomposed into several scales known as intrinsic mode functions (IMFs). By choosing a proper range of IMFs, the proposed algorithm was capable of removing nonlinear baseline drifts and high-frequency noises. Although, this approach yielded better identification results in comparison to some common 2-D based methods proposed in [4,7], the manual selection of proper modes brings some potential problems. The first problem arises when some bullet images contain less or more noise in comparison to others. In this situation, their useful information may be present in a different range of IMFs. The second problem arises when the bullet type changes. In this situation, manual inspection is needed again to figure out which range of IMFs contain less high-frequency noise for this particular type of bullet.

In this paper, we propose a novel combination of EEMD and Kalman filter to solve these problems. First the nonlinear baseline drift of averaged profiles are removed using EEMD algorithm. The profiles are then processed by a Kalman filter, designed to automatically and optimally reduce the effect of high-frequency noise. Using Expectation Maximization (EM) technique, for each averaged profile, the parameters of Kalman filter are reconfigured to optimally suppress the high-frequency noise. It was stated earlier that normalized cross correlation metric has been used by many computer-based bullet identification methods. This metric is shift (rotation) and amplitude invariant. In order to clarify the ambiguity, it is worth to mention that this metric is sensitive to amplitude change of consecutive samples. The shift invariance feature of this metric is very favorable in the field of bullet identification because most of the times, it is not clear that the starting point of bullet image corresponds to which part of bullet. The amplitude invariance feature on the other hand is not beneficial for 2-D bullet identification because most of the bullet images are digitalized using standard common image formats such as JPEG, TIFF or PNG and the amplitudes of averaged profiles contain valuable information which is neglected by normalized cross correlation metric. If we remove nonlinear baseline drifts and make normalized cross correlation metric sensitive to amplitudes, we may achieve better identification results. In this paper, in order to improve identification results, we propose a novel comparison metric that utilizes the useful features of both Euclidean and normalized cross-correlation metrics. The proposed algorithm was evaluated on a database containing 180 2-D gray-scale images acquired from fired AK-47 bullet surfaces. The experiments showed that our proposed method achieved better results compared with the previous method based on EEMD in the field of automatic bullet identification.

This paper is organized as follows: Section II provides information about the EEMD algorithm and Kalman filter and parameter estimation using EM algorithm. In section III we explain our proposed method for automatic bullet identification and section IV demonstrates the implementation results of the proposed algorithm [15,16].

2. Basic concepts

2.1. EEMD algorithm

The Ensemble Empirical Mode Decomposition (EEMD) is an improved version of Empirical Mode Decomposition (EMD) algorithm [17]. This algorithm is an empirical data-driven time-frequency decomposition method that can be used in nonlinear and non-stationary environments. EEMD is an adaptive decomposition method that breaks a given signal into multiple scales or empirical modes, known as intrinsic mode functions (IMFs) [18]. Each IMF corresponds to a certain narrow band amplitude-frequency modulation that is often linked with a particular statistical or physical process. In EEMD, “white” Gaussian noise (with different series) is added to a given signal for many times to obtain many instances of original signal. After that, EMD algorithm [17] is applied to each instance. The overall EEMD algorithm can be summarized in the following five steps:

1- For multiple N times, Add “white” Gaussian noise to original signal $x(t)$.

$$x_i(t) = x(t) + \omega_i(t) \tag{1}$$

where $\omega_i(t)$ is a “white” Gaussian noise sequence with a length equal to $x(t)$. $x_i(t)$ is the i^{th} noisy instance of $x(t)$.

2- For $i = 1, \dots, N$ do the following:
   a. Detect all local maxima and minima of signal $x_i(t)$.
   b. Connect the local maxima and estimate the upper envelope of $x_i(t)$ using cubic splines. Similarly, connect the local minima and

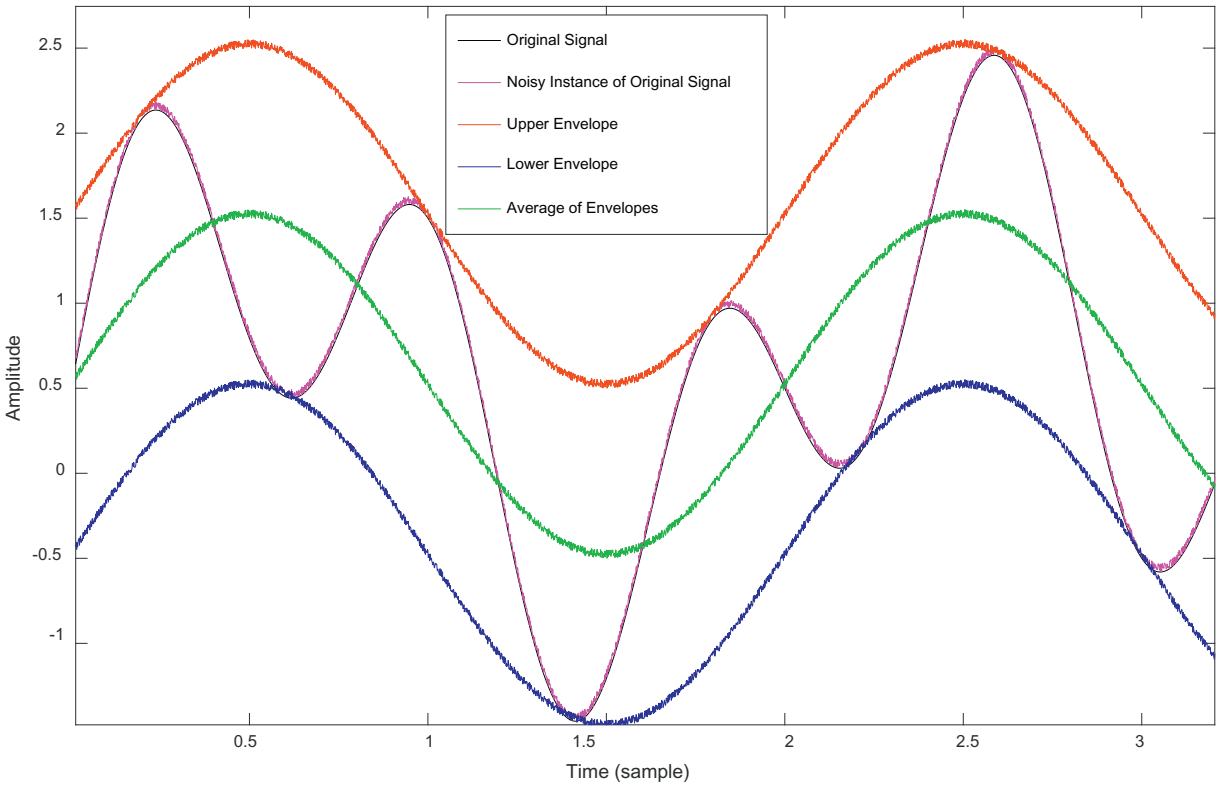


Fig. 1. Demonstration of steps1 and 2 in EEMD: (black, purple, red, blue and green lines are original signal, noisy instance of original signal, upper envelope (of noisy signal), lower envelope (of noisy signal) and the average of two envelopes respectively. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

acquire the lower envelope of $x_i(t)$ using cubic splines. Subtract the average of these two envelopes ($m_{1,\,0,\,i}$) from $x_i(t)$ to acquire the candidate for first IMF ($h_{1,\,0,\,i}$) (see Fig. 1):

$$x_i(t) - m_{1,0,i}(t) = h_{1,0,i}(t) \tag{2}$$

$h_{10,\,i}(t)$ should satisfy two conditions to be considered as an IMF. First, the number of local exterma and the number of zero crossing points in $h_{1,\,0,\,i}(t)$ must be either equal or at most differ by one. Second, the mean of its upper and lower envelopes must be equal to zero. If these conditions are not achieved, steps 'a' and 'b' are iterated over and over (also called as sifting process). In each iteration the resultant $h_{1,\,0,\,i}(t)$ is considered as the original signal ($x_i(t)$) for the next iteration. To put it in simple words, the following equation is iterated for a pre-defined number of repetitions (k) until it fulfills the two aforementioned conditions of IMF:

$$h_{1,(k-1),i} - m_{1,k,i} = h_{1,k,i} \tag{3}$$

There are some criteria for ending the sifting process. Standard Deviation (SD) is often used as termination criteria:

$$SD = \sum_{t=0}^{T} \frac{|h_{1,(k-1),i}(t) - h_{1,k,i}(t)|^2}{h_{1,(k-1),i}^2(t)} \tag{4}$$

The value of termination SD is usually defined in the range [0.2, 0.3]. Once the termination criteria are satisfied for an iteration k', $h_{1,\,k',\,i}$ is considered as the first IMF of $x_i(t)$ ($IMF_{1,\,i}(t)$).

3- Obtain the first IMF of $x(t)$ by ensemble averaging of all noisy IMFs calculated in step 2:

$$\overline{IMF_1}(t) = \frac{1}{N}\sum_{i=1}^{N} IMF_{1,i}(t) \tag{5}$$

4- Once the first IMF ($\overline{IMF_1}(t)$) is acquired, calculate the residue:

$$r(t) = x(t) - \overline{IMF_1}(t) \tag{6}$$

5- Treat the residue $r(t)$ as the original signal and follow the steps 1 to 4 to obtain additional IMFs.

Once the residue has no > 3 exterma, stop the iterations.

The signal $x(t)$ can be reconstructed using a linear summation of IMFs and residue i.e.:

$$x(t) = \sum_{i=1}^{n} \overline{IMF_i}(t) + r(t) \tag{7}$$

2.2. Kalman filter

Kalman filter is a Bayesian recursive linear quadratic estimation (LQE) technique that uses a series of noisy and inaccurate measurements observed over time and provides estimates of unknown variables. The recursive sequential estimation procedure in Kalman filter is proved to be more precise than other methods based on a single measurement alone [19].The simplicity and applicability of the Kalman filter algorithm encouraged researchers to implement its variants in a wide range of applications such as signal denoising, robotics, guidance, navigation, and control of vehicles, particularly aircraft and spacecraft.

2.2.1. Definition

Kalman filter algorithm can be applied on both continuous and discrete-time systems. The discrete-time Kalman filter provides the closed form recursive solution for estimation of linear discrete-time dynamic systems, which can be described generally by the following equations:

$$\mathbf{x}_k = A_k\mathbf{x}_{k-1} + \mathbf{w}_{k-1} \tag{8.a}$$

$$\mathbf{y}_k = C_k\mathbf{x}_k + \mathbf{v}_k \tag{8.b}$$

where

- $\mathbf{x}_k$ is the state vector of the system at the time step *k*.
- $\mathbf{y}_k$ is the measurement vector at the time step *k*.
- $\mathbf{w}_{k-1}$ is the process noise vector at the time step *k-1* which models the uncertainty of state dynamics using normal "white" Gaussian noise with covariance matrix Q_{k-1} ($\mathbf{w}_{k-1}\sim N(0, Q_{k-1})$).
- $\mathbf{v}_k$ is the measurement noise vector at the time step *k* which models the uncertainty of measurement dynamics using normal "white" Gaussian noise with covariance matrix R_k ($\mathbf{v}_k\sim N(0, R_k)$).
- A_k is the transition matrix of the dynamic model at time step k.
- C_k is the measurement matrix.

From statistical point of view, the above system describes a simple first-degree Markov process. For simplicity, it can be assumed that $\mathbf{v}_k$ and $\mathbf{w}_k$ are statistically independent from each other. The model in Eq. (8) can also be equivalently expressed in probabilistic forms with the following distributions:

$$p(\mathbf{x}_k \mid \mathbf{x}_{k-1}) = N(\mathbf{x}_k; A_k\mathbf{x}_{k-1}, Q_{k-1}) \tag{9.a}$$

$$p(\mathbf{y}_k \mid \mathbf{x}_k) = N(\mathbf{y}_k; C_k\mathbf{x}_k, R_k) \tag{9.b}$$

The prior (initial) distribution for the state is assumed to be $\mathbf{x}_0\sim N(\mathbf{m}_0, P_0)$. Parameters $\mathbf{m}_0$and P_0 are chosen based on prior knowledge about the system under the study.

The Kalman filter algorithm has a two-step process. In the first step (also known as prediction step), the current state of the system is estimated (predicted) given the previous measurements. In the second step (also known as update step) the predicted state of the system is corrected (updated) given the measurement at the current time step. The overall algorithm can be summarized as follows [20,21];

a. **Prediction step**

$$\hat{\mathbf{x}}_{k|k-1} = A_k\hat{\mathbf{x}}_{k-1|k-1} \tag{10.a}$$

$$P_{k|k-1} = A_kP_{k-1|k-1}A_k^T + Q_{k-1} \tag{10.b}$$

b. **Update step**

$$\mathbf{e}_k = \mathbf{y}_k - C_k\hat{\mathbf{x}}_{k|k-1} \tag{11.a}$$

$$S_k = C_kP_{k|k-1}C_k^T + R_k \tag{11.b}$$

$$K_k = P_{k|k-1}C_k^TS_k^{-1} \tag{11.c}$$

$$\hat{\mathbf{x}}_{k|k} = \hat{\mathbf{x}}_{k|k-1} + K_k\mathbf{e}_k) \tag{11.d}$$

$$P_{k|k} = P_{k|k-1} - K_kS_kK_k^T \tag{11.e}$$

where

- $\hat{\mathbf{x}}_{k|k-1}$ and $P_{k|k-1}$ are the predicted mean (predicted (a priori) state estimate) and covariance (predicted (a priori) estimate covariance) of the state vector, respectively, at the time step *k* before seeing the measurement.
- $\hat{\mathbf{x}}_{k|k}$ and $P_{k|k}$ are the estimated mean (updated (a posteriori) state estimate) and covariance (updated (a posteriori) estimate covariance) of the state vector, respectively, at the time step *k* after seeing the measurement.
- $\mathbf{e}_k$ is the innovation or the measurement residual or prediction error at the time step *k*.
- S_k is the measurement prediction covariance at the time step *k*.
- K_k is the Kalman filter gain at the time step *k*, which tells how much the predictions should be corrected.

Despite its efficiency and simplicity, the Kalman filter suffers from one major problem. In order to obtain desirable results from Kalman filter, the building parameters of the dynamic system in Eq. (8) i.e. $\{A_k, C_k, Q_k, R_k, \mathbf{m}_0, P_0\}$ should be defined or identified correctly (especially Q_k and R_k). Practical implementation of the Kalman Filter is often challenging because in many situations, there is a little knowledge about the noise covariance matrices. The problem of choosing correct parameters for Kalman is a subject of system identification field and many researchers proposed different approaches to resolve this matter [20,22,23]. One of the most recent propositions is a robust Expectation Maximization (EM) algorithm along with Kalman filter to identify system dynamics [23].

In system identification, the marginal likelihood of the Kalman filter is computed. The marginal likelihood of Kalman filter is defined as the *probability* that a Kalman filter with a given set of parameters (prior distribution, transition and observation models, and control inputs) would generate a particular observed signal $Y_N = \{\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_k, \dots, \mathbf{y}_N\}$. In order to compute this probability, the values of the hidden state vectors are integrated over ("marginalized out"). In this way, the marginal likelihood can be computed using only the observed signal Y_k. The marginal likelihood is very beneficial to assess different parameter choices. Using chain rule, it is straightforward to compute the marginal likelihood as the product of the probability of each observation given previous observations i.e.:

$$p(Y_N) = \prod_{k=0}^{N} p(\mathbf{y}_k \mid \mathbf{y}_{k-1},\dots,\mathbf{y}_0) \tag{12}$$

Because the Kalman filter described in eqns. (10) and (11) is defined for a Markov process, all valuable and related information from previous measurements resides in the current state estimate ($\hat{\mathbf{x}}_{k|k-1}$ and $P_{k|k-1}$). Therefore, the marginal likelihood is given by:

$$\begin{aligned} p(Y_N) &= \prod_{k=0}^{N} \int p(\mathbf{y}_k \mid \mathbf{x}_k)p(\mathbf{x}_k \mid \mathbf{y}_{k-1},\dots,\mathbf{y}_0)d\mathbf{x}_k \\ &= \prod_{k=0}^{N} \int N(\mathbf{y}_k; C_k\mathbf{x}_k, R_k)N(\mathbf{x}_k; \hat{\mathbf{x}}_{k|k-1}, P_{k|k-1})d\mathbf{x}_k \\ &= \prod_{k=0}^{N} N(\mathbf{y}_k; C_k\hat{\mathbf{x}}_{k|k-1}, C_kP_{k|k-1}C_k^T + R_k) \\ &= \prod_{k=0}^{N} N(\mathbf{y}_k; C_k\hat{\mathbf{x}}_{k|k-1}, S_k) \end{aligned} \tag{13}$$

According to Eq. (13), the marginal liklihood is the product of Gaussian densities, each corresponding to the density of one observation $\mathbf{y}_k$ ($k = 1, \dots, N$), under the current filtering distribution with mean $C_k\hat{\mathbf{x}}_{k|k-1}$ and covariance S_k. The computation of Eq. (12) is very simple and can be accomplished in the update step. In order to avoid numeric errors, instead of the $p(Y_N)$ the *log* marginal likelihood $\ell_N = \log(p(Y_N))$ is preferred in practical implementations. Under the assumption that $\ell_{-1} = 0$, the *log* marginal likelihood can be recursively computed in the update step as follows [22]:

$$\ell_k = \ell_{k-1} - \frac{1}{2}(\mathbf{e}_k^TS_k^{-1}\mathbf{e}_k + \log|S_k| + N_y\log 2\pi) \tag{14}$$

where N_y is the dimension of measurement vector $\mathbf{y}_k$.

2.2.2. RTS smoother

The results of Kalman filter can be further smoothed using backward-pass algorithms such as Rauch–Tung–Striebel (RTS) smoother [21]. The RTS smoother is a fixed interval smoothing algorithm. The difference between Kalman filter and RTS smoother is that the recursion in Kalman filter moves forward and in RTS smoother backward. The RTS algorithm is comprised of two stages (passes): forward pass and backward pass. In the forward pass, a Kalman filter is applied to the system and the predicted and updated state estimates ($\hat{\mathbf{x}}_{k|k}$, $\hat{\mathbf{x}}_{k|k-1}$) along with covariance matrices ($P_{k|k}$, $P_{k|k-1}$) are saved and stacked for

all time steps for the backward pass. In the last time step, backward pass recursions produce the smoothed state estimates and covariance matrices while proceeding backwards to earlier time steps. The backward pass recursions are as follows:

$$L_k = P_{k|k}A_{k+1}^T P_{k+1|k}^{-1} \tag{15.a}$$

$$\hat{\mathbf{x}}_{k|N} = \hat{\mathbf{x}}_{k|k} + L_k(\hat{\mathbf{x}}_{k+1|N} - \hat{\mathbf{x}}_{k+1|k}) \tag{15.b}$$

$$P_{k|N} = P_{k|k} + L_k(P_{k+1|n} - P_{k+1|k})L_k^T \tag{15.c}$$

where $\hat{\mathbf{x}}_{k|N}$ and $P_{k|N}$ are the smoothed state estimate and covariance matrix generated in the backward pass at time step k respectively.

2.2.3. Square root implementation

Kalman filter is vulnerable against round-off errors produced in the process of calculating the estimated state covariance matrices. These problems occur when the process noise covariance (Q_k) is small or the hardware in which the Kalman filter is implemented has finite word length limitations. To put it in simple words, the round-off errors cause the covariance matrices of Kalman filter to be indefinite or non-symmetric and hence incomputable in the recursive filtering process. Although nowadays computers have much more capabilities, the numerical stability of Kalman filter is still an issue in finite word length implementation of it, especially in embedded systems [20]. To address this problem, in several literatures, it was proposed to use square root filtering approaches for implementation of Kalman filters. Square root Kalman filters are numerically more stable than the conventional Kalman filters. In these types of filters, the square root forms of covariance matrices are propagated instead of their complete forms. Any positive definite matrix M has the property that it has a triangular matrix square root S which satisfies $M = S.S^T$. Theoretically, the covariance matrix of the square root filters would face less numerical problem because the condition number for the square root of a covariance matrix is the square root of the condition number of that matrix [24]. If the covariance matrices are kept in the form of S instead of M, they can never have a negative diagonal or become asymmetric. In literature, several square root filtering methods such as Cholesky decomposition and its variants, *UD*-factorization of Kalman filter [20,25], *QR* factorization of Kalman filter [24,26] have been proposed.

In this paper we adopted the $\mathcal{QR}$ factorization approach for square root implementation of Kalman filter. If matrix $E_{m\times n}$ has linearly independent columns, then it can be factored as:

$$E = \mathcal{QR} = [q_1 \ \dots \ q_m]\begin{bmatrix} \mathcal{R}_{11} & \mathcal{R}_{12} & \cdots & \mathcal{R}_{1n} \\ 0 & \mathcal{R}_{22} & \cdots & \mathcal{R}_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \mathcal{R}_{nn} \end{bmatrix} \tag{16}$$

where $\mathcal{Q}$ is unitary matrix ($\mathcal{Q}^H\mathcal{Q} = I$) with orthonormal vectors, $q_1, \dots, q_m$ i.e.:

$$\|q_i\| = 1,\ q_i^T q_j = 0 \ i \neq j \tag{17}$$

$\mathcal{R}$ is an upper triangular matrix whose diagonal elements are nonzero positive ($\mathcal{R}_{ii} > 0$). The matrix E can be a real or a complex valued matrix. In both cases, the elements of matrix $\mathcal{R}$ are always real. When E is real, $\mathcal{Q}^T\mathcal{Q} = I$ and when it is complex, $\mathcal{Q}^H\mathcal{Q} = I$ [27].

To implement $\mathcal{QR}$ factorization, first we should seek for the problematic parameters in Kalman filter and RTS smoother equations. These parameters are $P_{k|k-1}$ (in Eq. (10.b)), $P_{k|k}$ (in Eq. (11.e)) and $P_{k|n}$ (in Eq. (15.c)). To acquire a recursion which propagates the square root $(P_{k|k-1})^{1/2}$ for which $P_{k|k-1} = (P_{k|k-1})^{\frac{1}{2}}(P_{k|k-1})^{\frac{T}{2}}$ satisfies Eq. (10.b), the following $\mathcal{QR}$ is constructed:

$$\begin{bmatrix} P_{k-1|k-1}^{T/2}A_k^T \\ Q_k^{T/2} \end{bmatrix} = \mathcal{Q}\begin{bmatrix} \overline{\mathcal{R}}_{11} \\ 0 \end{bmatrix} \tag{18}$$

Now, exploiting the unitary nature of $\mathcal{Q}$, and multiplying the left side of Eq. (18) by its transpose we obtain the following:

$$\begin{aligned}\left[A_k P_{k-1|k-1}^{\frac{1}{2}}\ Q_k^{\frac{1}{2}}\right] \times \begin{bmatrix} P_{k-1|k-1}^{\frac{T}{2}}A_k^T \\ Q_k^{\frac{T}{2}} \end{bmatrix} &= [\overline{R}_{11}^T\ 0]\ \mathcal{Q}^T \times \underbrace{\mathcal{Q}}_{I}\begin{bmatrix} \overline{\mathcal{R}}_{11} \\ 0 \end{bmatrix} \\ &= [\overline{\mathcal{R}}_{11}^T\ 0] \times I \times \begin{bmatrix} \overline{\mathcal{R}}_{11} \\ 0 \end{bmatrix} = \overline{\mathcal{R}}_{11}^T\overline{\mathcal{R}}_{11}\end{aligned} \tag{19}$$

It can be easily noticed that the left side of Eq. (19) matches $A_k P_{k-1|k-1}A_k^T + Q_{k-1} = P_{k-1|k-1} = \overline{\mathcal{R}}_{11}^T\overline{\mathcal{R}}_{11}$. In other words, $(P_{k|k-1})^{\frac{1}{2}} = \overline{\mathcal{R}}_{11}^T$.

In order to acquire a recursion which propagates the square root $(P_{k|k})^{1/2}$ for which $P_{k|k} = (P_{k|k})^{\frac{1}{2}}(P_{k|k})^{\frac{T}{2}}$ satisfies Eq. (11.e), consider the following $\mathcal{QR}$ factorization:

$$\begin{bmatrix} R_k^{T/2} & 0 \\ (P_{k|k-1})^{T/2}C_k^T & (P_{k|k-1})^{T/2} \end{bmatrix} = \mathcal{Q}\begin{bmatrix} \mathcal{R}_{11} & \mathcal{R}_{12} \\ 0 & \mathcal{R}_{22} \end{bmatrix} \tag{20}$$

Using the same strategy in Eq. (19) and multiplying the left side of Eq. (20) by its transpose we obtain the following:

$$\begin{aligned}&\begin{bmatrix} R_k^{T/2} & 0 \\ (P_{k|k-1})^{T/2}C_k^T & (P_{k|k-1})^{T/2} \end{bmatrix}^T \times \begin{bmatrix} R_k^{T/2} & 0 \\ (P_{k|k-1})^{T/2}C_k^T & (P_{k|k-1})^{T/2} \end{bmatrix} \\ &= \left(\begin{bmatrix} \mathcal{R}_{11} & \mathcal{R}_{12} \\ 0 & \mathcal{R}_{22} \end{bmatrix}^T \mathcal{Q}^T\right) \times \left(\mathcal{Q}\begin{bmatrix} \mathcal{R}_{11} & \mathcal{R}_{12} \\ 0 & \mathcal{R}_{22} \end{bmatrix}\right) \\ &= \begin{bmatrix} R_k + C_k P_{k|k-1}C_k^T & C_k P_{k|k-1} \\ P_{k|k-1}C_k^T & P_{k|k-1} \end{bmatrix} = \begin{bmatrix} \mathcal{R}_{11}^T\mathcal{R}_{11} & \mathcal{R}_{11}^T\mathcal{R}_{12} \\ \mathcal{R}_{12}^T\mathcal{R}_{11} & \mathcal{R}_{12}^T\mathcal{R}_{12} + \mathcal{R}_{22}^T\mathcal{R}_{22} \end{bmatrix}\end{aligned} \tag{21}$$

Using $\mathcal{R}_{11}^{-T}\mathcal{R}_{11}^T = I$, the upper right part of (21) can be written as:

$$\mathcal{R}_{12}^T\mathcal{R}_{11} = \mathcal{R}_{12}^T\mathcal{R}_{11}^{-T}(\mathcal{R}_{11}^T\mathcal{R}_{11}) = \mathcal{R}_{12}^T\mathcal{R}_{11}^{-T}(R_k + C_k P_{k|k-1}C_k^T) = P_{k|k-1}C_k^T \tag{22}$$

From Eq. (22) it can be seen that

$$\mathcal{R}_{12}^T\mathcal{R}_{11}^{-T} = P_{k|k-1}C_k^T(R_k + C_k P_{k|k-1}C_k^T)^{-1} = P_{k|k-1}C_k^T S_k^{-1} = K_k \tag{23}$$

For the lower right part of Eq. (21), we have:

$$\mathcal{R}_{22}^T R_{22} = P_{k|k-1} - \mathcal{R}_{12}^T\mathcal{R}_{12} \tag{24}$$

Using $\mathcal{R}_{11}^{-T}\mathcal{R}_{11}^T = I$ and the result from Eqs. (23) and (24) is rewritten as follows:

$$\mathcal{R}_{22}^T\mathcal{R}_{22} = P_{k|k-1} - (\mathcal{R}_{12}^T\mathcal{R}_{11}^{-T})(\mathcal{R}_{11}^T\mathcal{R}_{12}) = P_{k|k-1} - K_k C_k P_{k|k-1} = P_{k|k} \tag{25}$$

In other words, in Eq. (20) we have $\mathcal{R}_{22}^T = P_{k|k}^{1/2}$. Using Eqns. (18) and (20), we calculate the square roots of covariance matrices required for the Kalman filter. For the smoothed covariance matrix ($P_{k|N}$) in the RTS smoother, the following $Q\mathcal{R}$ is constructed:

$$\begin{bmatrix} P_{k|k}^{T/2}A_k^T & P_{k|k}^{T/2} \\ Q^{T/2} & 0 \\ 0 & P_{k+1|N}^{T/2}L_k^T \end{bmatrix} = \mathcal{Q}\begin{bmatrix} \mathcal{R}_{11} & \mathcal{R}_{12} \\ 0 & \mathcal{R}_{22} \\ 0 & 0 \end{bmatrix} \tag{26}$$

Exploiting the same strategy in Eqs. (19) and (21) and multiplying the left side of Eq. (26) by its transpose we obtain the following:

$$\begin{bmatrix} P_{k+1|k} & A_k^T P_{k|k} \\ P_{k|k}A_k^T & P_{k|k} + L_k P_{k+1|N}L_k^T \end{bmatrix} = \begin{bmatrix} \mathcal{R}_{11}^T\mathcal{R}_{11} & \mathcal{R}_{11}^T\mathcal{R}_{12} \\ \mathcal{R}_{12}^T\mathcal{R}_{11} & \mathcal{R}_{12}^T\mathcal{R}_{12} + \mathcal{R}_{22}^T\mathcal{R}_{22} \end{bmatrix} \tag{27}$$

For $\mathcal{R}_{22}^T\mathcal{R}_{22}$ we have:

$$\begin{aligned}\mathscr{R}_{22}^T\mathscr{R}_{22} &= P_{(k|k)} + L_k P_{(k+1|N)}\, L_k^T - \mathscr{R}_{12}^T\mathscr{R}_{12} = P_{(k|k)} + L_k P_{(k+1|N)}\, L_k^T \\ &- \mathscr{R}_{12}^T(\mathscr{R}_{11}\mathscr{R}_{11}^{-1})(\underbrace{[\![(\mathscr{R})_{11}^T)^{-1}\mathscr{R}_{11}^T)}_{I}\mathscr{R}_{12} = P_{(k|k)} + L_k P_{(k+1|N)}\, L_k^T \\ &- \mathscr{R}_{12}^T\mathscr{R}_{11}\,\underbrace{[\![(\mathscr{R})_{11}^T}_{I}\mathscr{R}_{11})^{-1}\mathscr{R}_{11}^T R_{12} = P_{(k|k)} + L_k P_{(k+1|N)}\, L_k^T \\ &- P_{(k|k)A_k^T}\,[\![(P)_{(k+1|k)})^{-1}A_k P_{(k|k)} \\ &= P_{(k|k)} + L_k(P_{(k+1|N)}\,[\![-P_{(k+1|k)})L]\!]_k^T = P_{(k|N)} \end{aligned} \tag{28}$$

From Eq. (28), it can be derived that in Eq. (26), $\mathscr{R}_{22}^T = (P_{k|N})^{1/2}$. In Eq. (28), some algebraic expressions were colored red because we wanted the readers to better figure out how the final output of Eq. (28) is derived. Under each red expression, we put an underbrace $(\underbrace{\ }_{I})$ to show that the corresponding expression equals the identity matrix I and does not have any effect on the equation.

2.3. Parameter selection using expectation maximization

Before explaining the EM algorithm used in this paper, it is worth to briefly explain the maximum likelihood (ML) estimation. For a given set of probabilistic measurements $Y_N = \{\mathbf{y}_1, \ldots \mathbf{y_k}, .., \mathbf{y}_N\}$, the ML estimate of probabilistic parameter set $\boldsymbol{\theta}$ is obtained by maximizing the marginal likelihood of $\mathbf{y}_{1:N}$ with respect to $\boldsymbol{\theta}$, i.e.:

$$\hat{\theta}_{ML} = \underset{\theta}{\operatorname{argmax}}(log\ p_{\theta}(\mathbf{y}_{1:N})) \tag{29}$$

In the case of unknown linear Gaussian state-space models, the probabilistic parameter set we wish to estimate is $\boldsymbol{\theta} = \{A_k, C_k, Q_k, R_k, \mathbf{m}_0, P_0\}$.For simplicity let's suppose that A_k, C_k, Q_k andR_k are constant for all time steps i.e. $\boldsymbol{\theta} = \{A, C, Q, R, \mathbf{m}_0, P_0\}$. It was suggested in [22,23,28] that EM method can be used for ML estimation of $\boldsymbol{\theta}$ using the outputs of RTS smoother. This method is implemented in the following two steps:

1- The *E*-step: involves calculating the expectation of log likelihood $\mathscr{O}(\theta, \theta_i) = E[\log p_{\theta_i}(\mathbf{y}_{1:N}, \mathbf{x}_{1:N} \mid \mathbf{y}_{1:N})]$ given the estimated parameters of $\boldsymbol{\theta}$ at *i*th iteration. The quantity of $\mathscr{O}$ depends on the following [22,28]:

$$E[\mathbf{x}_k \mid \mathbf{y}_{1:N}] = \hat{\mathbf{x}}_{k|N} \tag{30.a}$$

$$S_{1,1,k} = E[\mathbf{x}_k\mathbf{x}_k^T \mid \mathbf{y}_{1:N}] = P_{k|N} + \hat{\mathbf{x}}_{k|N}\hat{\mathbf{x}}_{k|N}^T \tag{30.b}$$

$$S_{1,0,k} = E[\mathbf{x}_k\mathbf{x}_{k-1}^T \mid \mathbf{y}_{1:N}] = P_{k,k-1|N} + \hat{\mathbf{x}}_{k|N}\hat{\mathbf{x}}_{k-1|N}^T \tag{30.c}$$

$$S_{0,0,k} = E[\mathbf{x}_{k-1}\mathbf{x}_{k-1}^T \mid \mathbf{y}_{1:N}] = P_{k-1|N} + \hat{\mathbf{x}}_{k-1|N}\hat{\mathbf{x}}_{k-1|N}^T \tag{30.d}$$

where

$$P_{k,k-1|N} = P_{k|k}L_{k-1}^T + L_k(P_{k+1,k|N} - P_{k|k})L_{k-1}^T \tag{31}$$

which is initialized by $P_{N,\,N-1|N} = (I - K_N)P_{N-1|N-1}$ [22,28].

2- The *M*-step: involves re-estimation of model parameters by maximizing the $\mathscr{O}$ over $\boldsymbol{\theta}$ using partial derivatives of $\mathscr{O}$ and setting them to zero. Solving these equations yields the updated parameters (in the *i*th iteration) as follows:

$$A^{(i)} = \left(\sum_{k=2}^N S_{1,0,k}\right)\left(\sum_{k=2}^N S_{0,0,k}\right)^{-1} \tag{32.a}$$

$$Q^{(i)} = \frac{1}{N-1}\left(\sum_{k=2}^N S_{1,1,k} - A^{(i)}\left(\sum_{k=2}^N S_{1,0,k}\right)^T\right) \tag{32.b}$$

$$C^{(i)} = \left(\sum_{k=1}^N \mathbf{y}_k\hat{\mathbf{x}}_{k|N}^T\right)\left(\sum_{k=1}^N P_{k|N}\right)^{-1} \tag{33.c}$$

$$R^{(i)} = \frac{1}{N}\left(\sum_{k=1}^N \left(\mathbf{y}_k - C^{(i)}\hat{\mathbf{x}}_{k|N}\right)(\mathbf{y}_k - C^{(i)}\hat{\mathbf{x}}_{k|N})^T - C^{(i)}P_{k|N}(C^{(i)})^T\right) \tag{32.d}$$

$$P_0^{(i)} = P_{1|N} \tag{32.e}$$

$$\mathbf{m}_0^{(i)} = \hat{\mathbf{x}}_{1|N} \tag{32.f}$$

The above equations provide the formal EM algorithm specification but it can be implemented in a more robust and numerically efficient way [23]. It is known that the EM algorithm is an iterative process and it is always required to have positive semi-definite $Q^{(i)}$ and $R^{(i)}$ to use the Kalman filter. But because of the presence of subtraction and inverse operations in eqns. (32.a-d), it is possible that after a few iterations, the outputs of EM algorithm don't meet the Kalman filter requirements. In order to decrease such possibility, in this paper, we adopt the method proposed in [23] which implements a numerically robust EM algorithm using square-root filtering strategy. This method is summarized in the following steps:

1- Let $i = 0$ and initialize estimates at $\hat{\theta}_0 = \{A, C, Q, R, \mathbf{m}_0, P_0\}$
2- Using the system specification $\hat{\theta}_i = \{A, C, Q, R, \mathbf{m}_0, P_0\}$ compute for $k = 1, \ldots, N$, the sequences $\{\hat{\mathbf{x}}_{k|k}\}$, $\{P_{k|k-1}{}^{1/2}\}$ and $\{P_{k|k}{}^{1/2}\}$ using eqns. (11.d), (18) and (20) respectively.
3- Compute for $k = 1, \ldots, N$, the smoothed sequences $\{\hat{\mathbf{x}}_{k|N}\}$, $\{P_{k|N}{}^{1/2}\}$ using eqns. (15.b) and (26) respectively.
4- Compute the following matrices

$$E[\mathbf{x}_k\mathbf{x}_k^T \mid \mathbf{y}_{1:N}] = P_{k|N}^{\frac{1}{2}}\left(P_{k|N}^{\frac{1}{2}}\right)^{\boldsymbol{T}} + \hat{\mathbf{x}}_{k|N}\hat{\mathbf{x}}_{k|N}^T \tag{33.a}$$

$$E[\mathbf{x}_k\mathbf{x}_{k-1}^T \mid \mathbf{y}_{1:N}] = P_{k,k-1|N} + \hat{\mathbf{x}}_{k|N}\hat{\mathbf{x}}_{k-1|N}^T = \left(P_{k|k}^{\frac{1}{2}}\right)\left(P_{k|k}^{\frac{1}{2}}\right)^T L_{k-1}^T + L_k(P_{k+1,k|N}\left(P_{k|k}^{\frac{1}{2}}\right)\left(P_{k|k}^{\frac{1}{2}}\right)^T)L_{k-1}{}^T \tag{33.b}$$

$$E[\mathbf{y}_k\mathbf{x}_k^T \mid \mathbf{y}_{1:N}] = \mathbf{y}_k\hat{\mathbf{x}}_{k|N}^T \tag{33.c}$$

5- Using the matrices from step 4, compute the following matrices:

$$\Phi \triangleq \frac{1}{N}\sum_{k=1}^N E[\boldsymbol{\xi}_k\boldsymbol{\xi}_k^T \mid \mathbf{y}_{1:N}] \tag{34.a}$$

$$\Psi \triangleq \frac{1}{N}\sum_{k=1}^N E[\boldsymbol{\xi}_k\mathbf{x}_k^T \mid \mathbf{y}_{1:N}] \tag{34.b}$$

$$\Sigma \triangleq \frac{1}{N}\sum_{k=1}^N E[\mathbf{x}_k\mathbf{x}_k^T \mid \mathbf{y}_{1:N}] \tag{34.c}$$

where $\boldsymbol{\xi}_k{}^T = [\mathbf{x}_{k+1}{}^T, \mathbf{y}_k{}^T]$.

6- Maximize $\mathscr{O}(\theta, \theta_i)$ with respect to $\boldsymbol{\theta}$ to acquire a new estimate $\boldsymbol{\theta}_{i+1}$ as follows:

(a) Compute new estimates of *A, C* and D via using *Eq.* (35) using pivoting and Gaussian elimination.

$$\Gamma = \begin{bmatrix} A \\ C \end{bmatrix} = \Psi\Sigma^{-1} \tag{35}$$

(b) Use the modified Cholesky factorization proposed in [29] to calculate new estimates of *Q* and *R* as follows:

$$\left(\begin{bmatrix} \Sigma & \Psi^T \\ \Psi & \Phi \end{bmatrix}^{\frac{1}{2}}\right)^T = \begin{bmatrix} \mathscr{R}_{11} & \mathscr{R}_{12} \\ 0 & \mathscr{R}_{22} \end{bmatrix} \tag{36.a}$$

$$\Pi = \Phi - \Psi\Sigma^{-1}\Psi^T = \begin{bmatrix} Q & 0 \\ 0 & R \end{bmatrix} = R_{22}^T R_{22} \tag{36.b}$$

(c) Set $\mathbf{m}_0 = \hat{\mathbf{x}}_{1|N}$ and $P_0 = P_{1|N}^{\frac{1}{2}}\left(P_{1|N}^{\frac{1}{2}}\right)^T$ where $\hat{\mathbf{x}}_{1|N}$ and $P_{1|N}^{\frac{1}{2}}$ are computed using eqns. (11.d) and (20) in the *E*-step.

The above steps are repeated for several iterations until it satisfies a certain stopping criterion. This criterion can be defined based on marginal likelihood (ℓ_k) defined in Eq. (14) in the following way:

$$\frac{|\ell_{N,\boldsymbol{\theta}_i} - \ell_{N,\boldsymbol{\theta}_{i-1}}|}{|\ell_{N,\boldsymbol{\theta}_{i-1}}|} < \varepsilon \tag{37}$$

where $\varepsilon > 0$ is a threshold defined by user, $\ell_{N,\ \boldsymbol{\theta}_i}$ and $\ell_{N,\ \boldsymbol{\theta}_{i-1}}$ are marginal likelihoods acquired using $\boldsymbol{\theta}_i$ and $\boldsymbol{\theta}_{i-1}$ in the last time step respectively.

3. Proposed method and materials

In this section we describe the proposed method for identification of bullets. Our approach for 1-D profile acquisition is explained first. The next subsection provides the preprocessing method applied to profiles using EEMD algorithm and Kalman filter and the last subsection explains the novel comparison step. The block diagram of overall algorithm is shown in Fig. 2.

3.1. Image rotation and profile acquisition

A typical 2-D gray-level bullet surface image contains striation-like twisted micro-structures. The angle of twist is due to barrel rifling. It is common to reduce the number of samples without the loss of useful information when comparing 2-D gray-level bullet images. This is achieved via image rotation. In a 2-D bullet image, if two sequential rows are observed carefully, it can be seen that one row has similar patterns with respect to the other row if it is shifted circularly a few pixels forwards or backwards. The number of these pixels depends on the twist angle. If a bullet image is rotated based on this angle of twist, the strong striations become perpendicular to horizontal axis. The rotated image can be averaged column-wise without the loss of useful information to produce a 1-D average profile. The first step in our proposed method is bullet image rotation and 1-D profile extraction. In this paper, Radon transform is used along with Canny edge detector to find the twist angle. The Radon transform is an integral transform described as the projection of the image pixels along a radial line oriented at a specific angle. The output of the Radon transform has the highest value near the twist angle of a straight line. Hence by defining a suitable angle range (0–180°) and a proper angle resolution (0.1°) an accurate detection of twist angle is accomplished by finding the extrema of this transform. In Fig. 3, the proposed image rotation procedure is illustrated for a typical 2-D gray-scale bullet image. First a Canny edge detector extracts the striation patterns from image (Fig. 3(b)). Then a Radon transform is applied to the binary edge image (Fig. 3(c)). In Fig. 3(c), the brighter areas indicate the high density of lines in a certain angle. In other words, the output of the Radon transform has the highest amplitude near the twist angle of strong striations. Based on this idea and using a proper angle resolution (0.1°) and an appropriate angle range ($0–180^\circ$), the twist angle is accurately determined. Finally, the bullet image is rotated based on the found twist angle (Fig. 3(d)). Using column wise averaging a 1-D average profile is extracted from the rotated image (Fig. 3(e)).

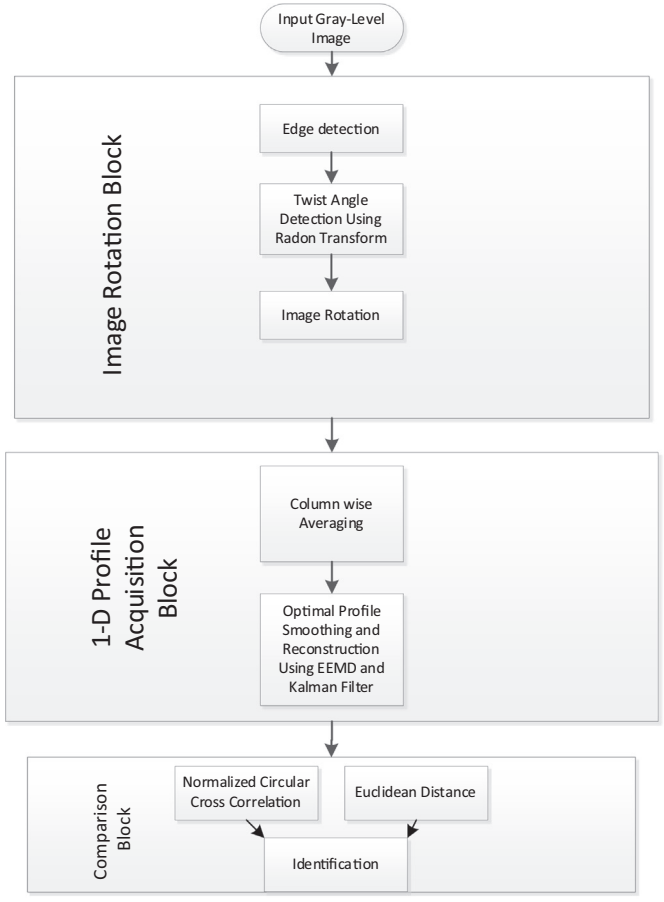


Fig. 2. Block diagram of the proposed method for automatic bullet identification

3.2. Profile smoothing

It can be seen in Fig. 3(e) that the average profile is an oscillatory nonlinear signal. The oscillations are the results of optical density changes, location of striations, lighting conditions, noise. The electronic hardware and circuits inside image recording devices are responsible for converting analog optical reflections from lens to digital electrical signals and then again converting these signals to digital images. The process of analog to digital conversion includes the quantization procedure. The quantized signal is different from the original analog signal due to an error called the quantization error. The noise in average profile is the result of this quantization error, interferences from other light sources and imperfections in the manufacturing process of image recording hardware. Some of these imperfections cannot be avoided. For example, making a perfect lens is almost physically impossible.

The strong peaks and valleys in the average profile correspond to start and end points of land impressions. The average profile has also a nonlinear baseline drift which doesn't correspond to striations and therefore should be suppressed in cross correlation-based comparisons [10]. Other than nonlinear baseline drifts, the high frequency noises should be removed from average profiles in order to reach acceptable results. In [4,6,7,12,13], in order to suppress high frequency noises, LTI filtering frameworks were used. But because of the arbitrary nature of striations as random unique micro structures caused by random imperfections in the rifling process, the LTI filtering frameworks may throw away useful information along with noise in bullet average profiles [10,30]. In [10], it was proposed to use EEMD as nonlinear non-stationary filtering framework to denoise bullet average profiles. In this approach, first the average profiles are decomposed into several IMFs (see Fig. 4). Each IMF corresponds to certain frequency band or oscillation. It can be seen in Fig. 4 that, the first IMFs are very oscillatory where high frequency noises exist. The last IMFs are related to slower oscillations where nonlinear baseline drift is present. In [10], by choosing a proper range of IMFs, the average profile was smoothed in a way that both high frequency low amplitude noise-related fluctuations

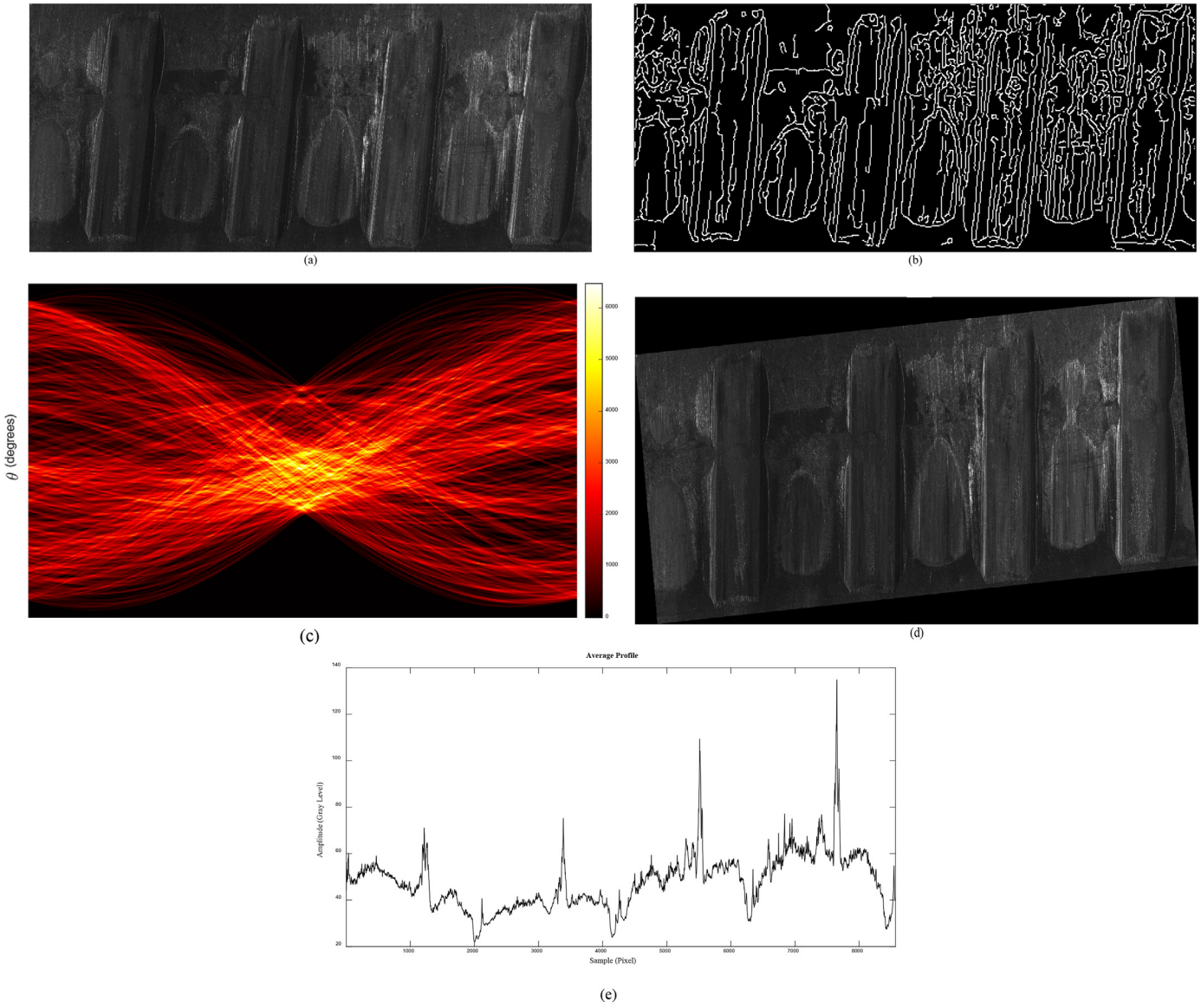


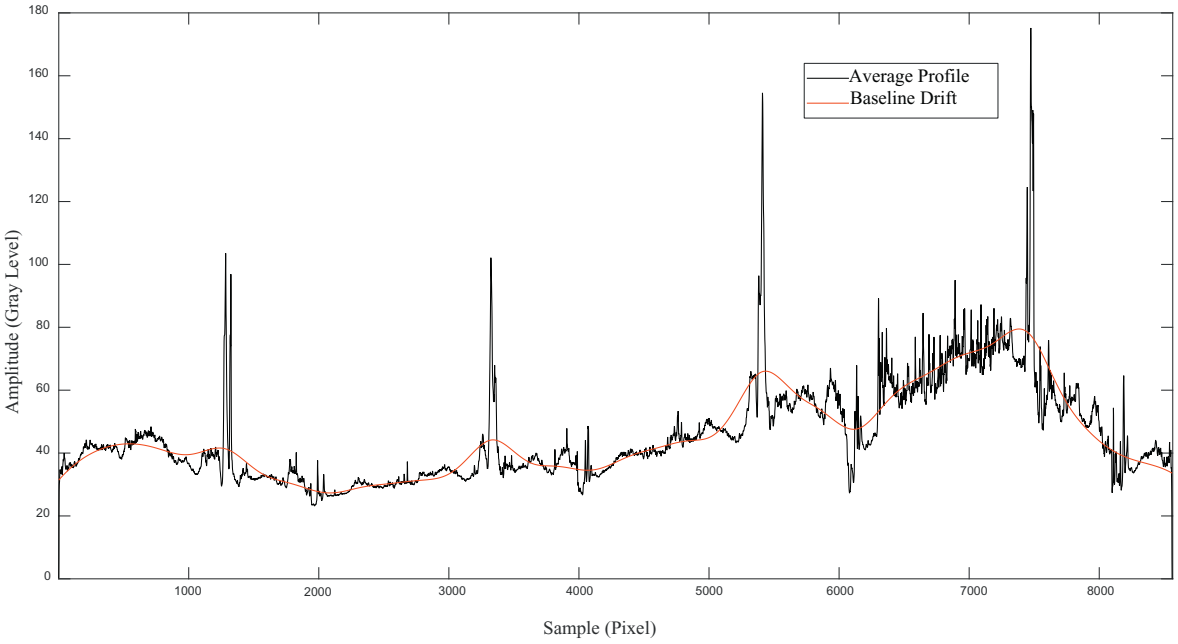
Fig. 3. Proposed approach for image rotation and profile extraction (a) A typical 2-D gray-level bullet surface image (b) edge detection using Canny edge detector (c) Result of Radon transform applied on edge image (d) image rotation (e) Average profile of rotated bullet image

and nonlinear baseline drifts were removed. The resultant smoothed profile was claimed to contain components that are primarily related to striations. But the procedure of choosing the proper number of IMFs to reduce the high-frequency noise effect was manual. This poses a problem in comparison of bullets whose images contain less or more noise in comparison to others because their useful information may be present in the corresponding discarded IMFs. Moreover, another problem arises when the bullet type changes. In this situation, visual inspection is needed once again to figure out which range of IMFs contains less high-frequency noise for this particular type of bullet.

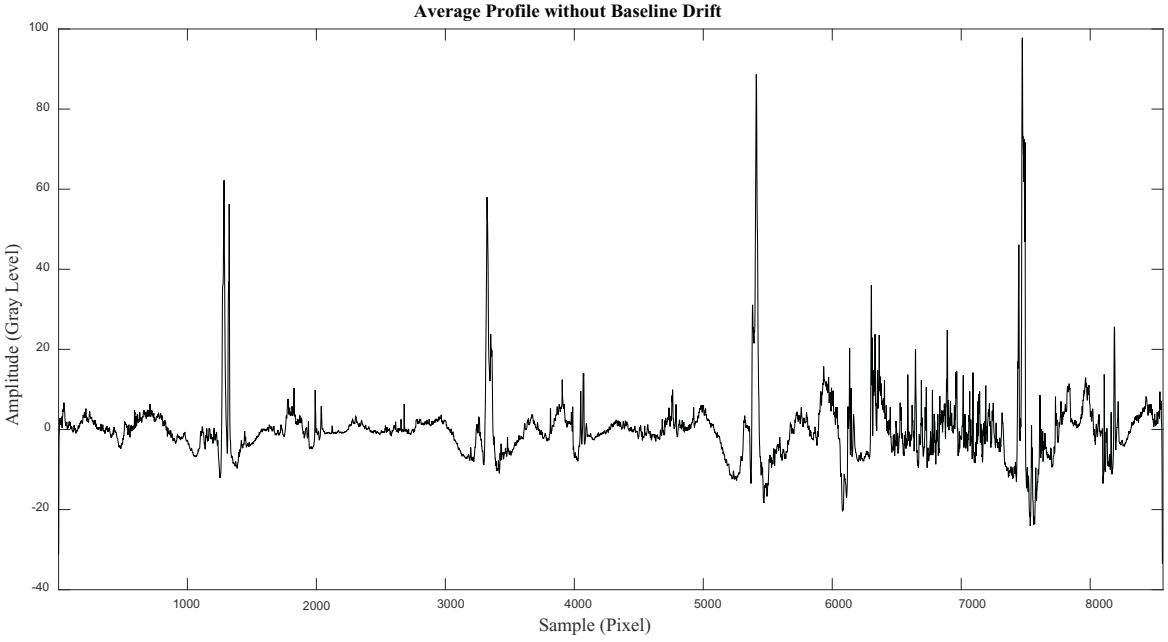
In order to overcome the aforementioned problems, in this paper, we propose to use Kalman filter along with EEMD to automatically remove high frequency noises from average profiles. Fig. 4 illustrates our approach for a typical bullet average profile. In our method, first we remove the nonlinear baseline drift from average profile using EEMD (Fig. 4). The parameters of Kalman filter are identified using EM algorithm explained in previous section. It is worth to mention that the parameter selection procedure is exclusive for each average profile. In other words, for each average profile, a different Kalman filter is employed to suppress high-frequency noises. In this way, the profile smoothing process is independent to bullet type or noise density variations. After parameter identification, the average profile is smoothed using the RTS smoother. Fig. 5 shows the output of RTS smoother for a segment in an average bullet profile. Using RTS smoother, valuable information is conserved while the noisy or insignificant information which may mislead the identification process are removed. By comparing Fig. 5(a) and (b), it can be realized that small noisy fluctuations are suppressed in the output of RTS smoother. Although, the RTS smoother seems to suppress few weak fluctuations from the signal in Fig. 4(b), in next sections we will see that this process has a positive effect in correlation-based bullet identification.

3.3. Comparison metrics

In the field of automatic bullet identification, many methods employed the normalized cross-correlation (NCC) metric for comparison of bullet profiles. Although this metric is susceptible to noise and nonlinear baseline drifts, it can be used for comparison of both stationary and non-stationary signals. The normalized cross-correlation between two discrete signals is defined as [6]:



(a)



(b)

Fig. 4. Proposed method for baseline drift removal (a) original average profile and its baseline drift extracted using EEMD (b) average profile without baseline drift.

$$NCC(r, l_i, \tau) = \frac{\sum_k ((r(k) - \bar{r}) \times (l_i(k+\tau) - \bar{l}_i))}{\sqrt{\sum_k (r(k) - \bar{r})^2} \sqrt{\sum_k ((l_i(k+\tau) - \bar{l}_i)^2}} \quad (38)$$

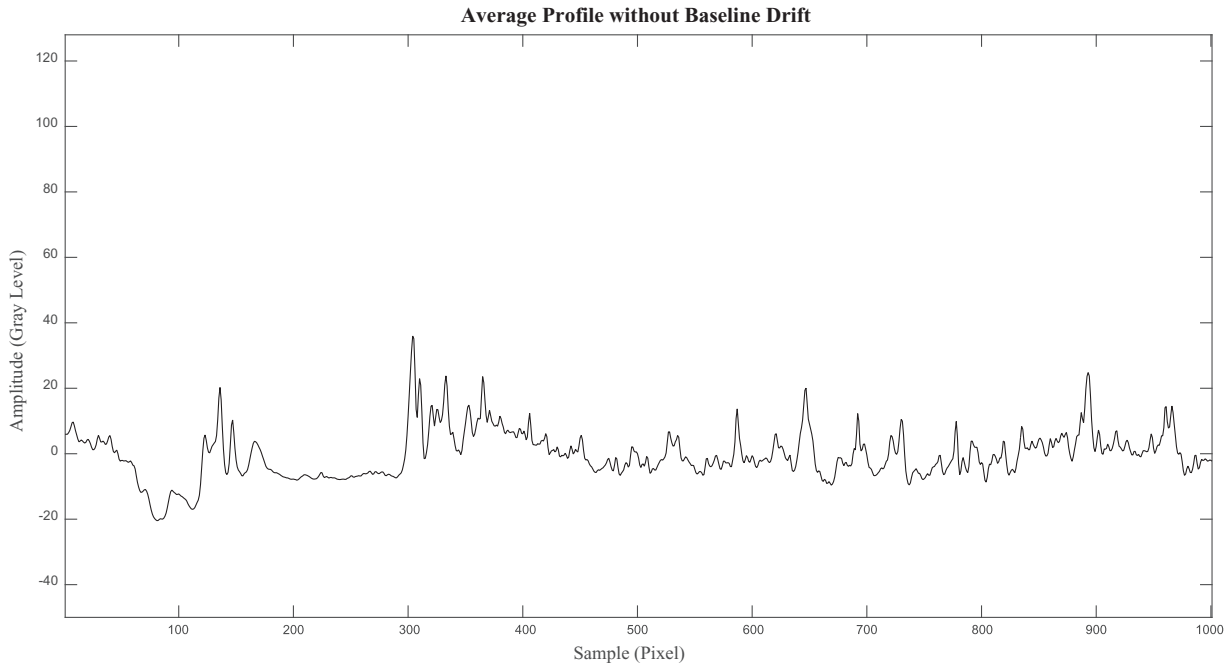
where r represents the reference profile; l_i (i = 1,2, …, n) represents one of the correlated profiles; $\bar{r}$ and $\bar{l}_i$ are the mean value of r and l_i. The location of the maximum of the correlation function $NCC_{max}(r, l_i, \tau)$, indicates the shift τ_i leading to the best possible correlation between the two correlated profiles and its amplitude represents the similarity between them [6]. If the amplitude is 1, the two profiles are perfectly matched and if it is 0, they are completely dissimilar. NNC in Eq. (38) can be rewritten as follows:

$$NCC(r, l_i, \tau) = \frac{(r(k) - \bar{r}) \circledast (l_i(-k) - \bar{l}_i)}{\sqrt{\sum_k (r(k) - \bar{r})^2} \sqrt{\sum_k ((l_i(k+\tau) - \bar{l}_i)^2}} \quad (39)$$

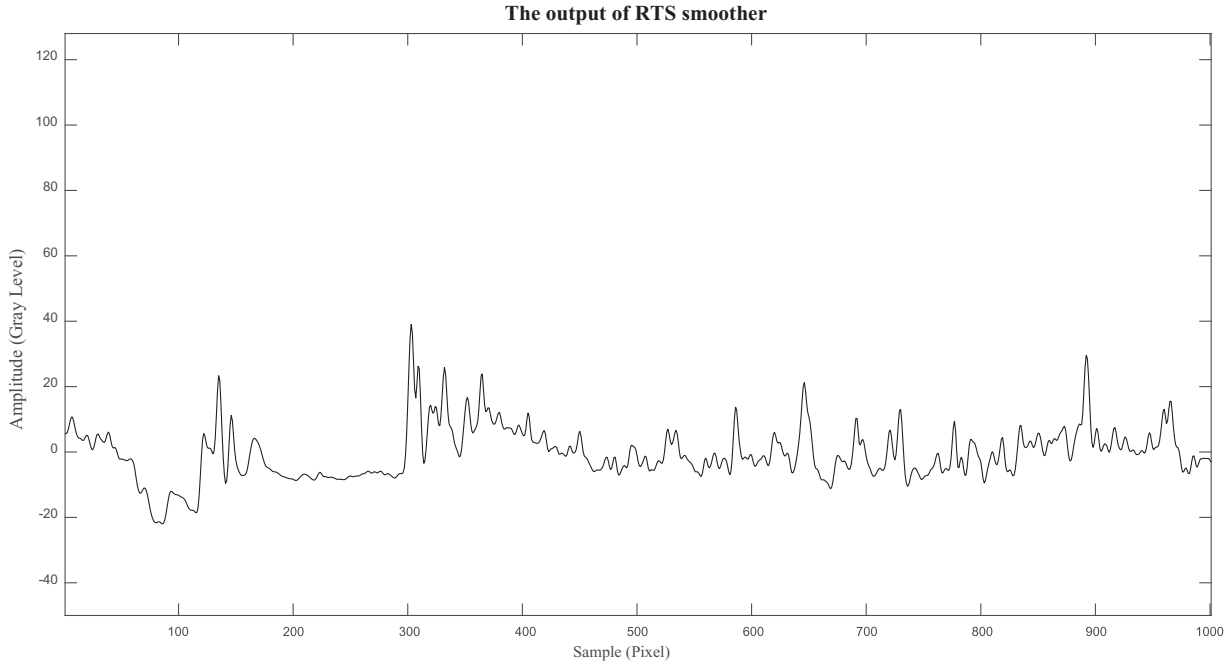
where ⊛ denotes the convolution. As is well-known, the convolution of two discrete signals $x(k)$ and $y(k)$ (both having length n) is calculated as follows:

$$(x \circledast y)(\tau) = \sum_{k=1}^{n} x(k) y(\tau - k) \quad (40)$$

Although NNC can be used efficiently for comparison of segmented bullet surfaces, it should be noted that Eq. (38) should be modified for comparison of average profiles extracted from images containing whole



(a)



(b)

Fig. 5. The proposed approach for high-frequency noise removal (a) a segment from an average bullet profile (b) the output of RTS smoother

bullet surfaces. The reason lies in the procedure of image acquisition of bullet surfaces. If the image acquisition device is set to capture a particular area of bullet surface (e.g. land impression), it is clear that the start and endpoint of the captured image corresponds to start and endpoint of that area. Therefore, using the time shift operation in Eq. (38), NCC can be fully implemented for comparison. However, if the image acquisition device is set to capture the whole cylindrical surface of bullet, most of the times, it is not clear that the start and endpoint of image corresponds to which part of bullet unless the firearm expert manually defines it before image acquisition. In other words, in images taken from whole bullet surface, we have circular time shifts instead of normal time shifts and because of using normal convolution in Eq. (38), NCC cannot yield reliable results in the cases where circular time shifts are involved. For better understanding of the aforementioned comments, a simple example is explained here. Suppose that we acquired two images from a bullet surface but, the start points of images differ from each other. For simplicity, the average profile of the whole bullet surface is assumed to be a simple sine function defined in range$[0, 2\pi]$.

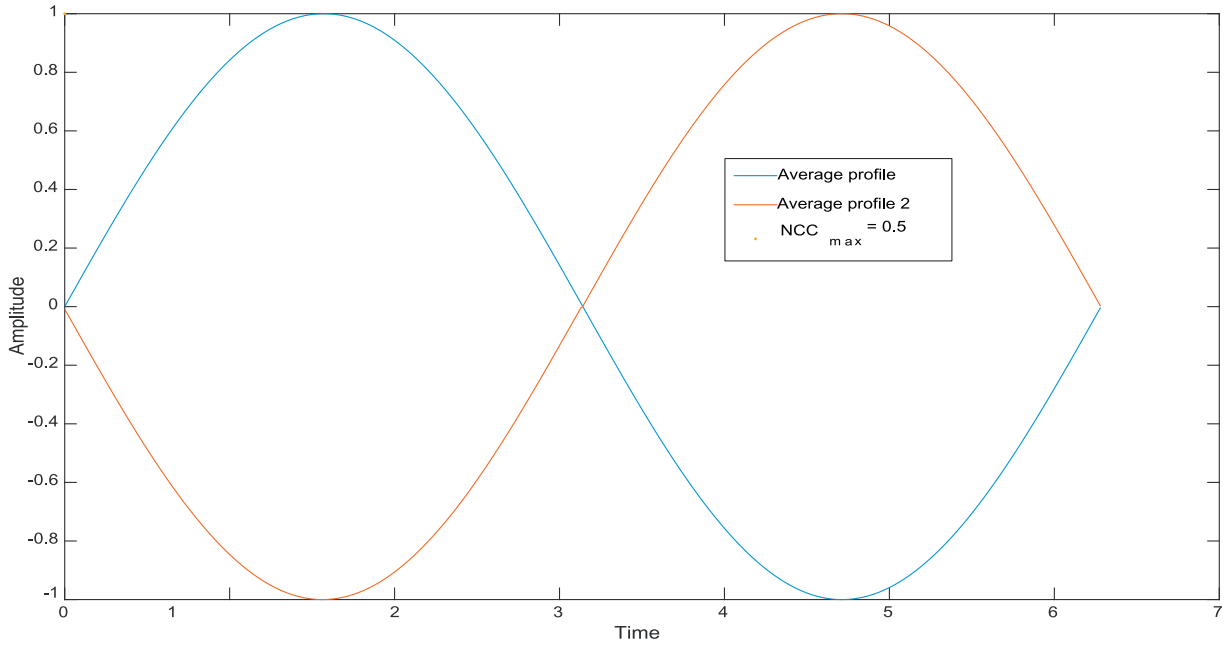


Fig. 6. An example of the NCC problem in comparison of whole bullet surfaces.

Also suppose that the average profiles of these two images correspond to the two signals shown in Fig. 6. As can be seen in this figure in this figure, the two signals are completely identical to each other if one of them is shifted circularly by $\tau = \pi$. But if we use (38) for these two signals, the value for NCC_{max} will be 0.5. This problem causes the identification results to be misleading, unreal and unreliable. To solve this problem, we propose to use circular convolution in Eq. (38). The circular convolution of two discrete signals x(k) and y(k) (both having length n) is calculated as follows:

$$(x \circledast_n y)(\tau) = \sum_{k=1}^{n} x(k)(y(\tau - k))_n \tag{41}$$

where $\circledast_n$ denotes the circular convolution operator and $()_n$ denotes the circular shift operator. Assume that a given signal $y(k)$ (with a length of n samples), is wrapped around a cylinder having n notches. In circular shifting, $(y(k-m))_n$ corresponds to rotating the cylinder m notches clockwise. Using Eqns. (41), (38) is modified as follows:

$$NCCC(r, l_i, \tau) = \frac{(r(k) - \bar{r}) \circledast_n (l_i(-k) - \bar{l}_i)}{\sqrt{\sum_k (r(k) - \bar{r})^2} \sqrt{\sum_k ((l_i(k+\tau) - \bar{l}_i)^2}} \tag{42}$$

where *NCCC* stands for normalized circular cross correlation. The NCCC metric in Eq. (42) is invariant to circular shifting and thus it can be used effectively in comparison of profiles of whole bullet surfaces.

To improve the NCCC metric in comparison of bullet profiles, we propose to use Euclidean distance along with NCCC. In order to demonstrate the importance of Euclidean distance, a simple example is demonstrated in Fig. 7. In this figure, for simplicity in understanding, three simple sine signals are represented as three bullet average profiles namely 'average profile 1', 'average profile 2' and 'average profile 3'. As can be seen in Fig. 7, 'average profile 2' is identical to 'average profile 1' if it is shifted circularly by π. Although the amplitude of 'average profile 3' is different from 'average profile 1', because of the normalization procedure in Eq. (42), $NCCC_{max}$ for 'average profile is 1. This means that if the comparison algorithm solely depends on $NCCC_{max}$, it is possible that non-related average profiles get high score in the comparison procedure. Not to mention that NCC_{max} suffers the same problem too because of the normalization procedure in Eq. (38). To solve this problem, we propose to use the following metric for comparison:

$$WNCCC(r, l_i, \tau) = \frac{NCCC(r, l_i, \tau)}{|r(k) - (l_i(k+\tau))_n|} \tag{43}$$

Where WNCCC stands for weighted normalized circular cross correlation. As can be seen in Eq. (43), the Euclidean distance, weighs the score of NCCC and makes it more reliable. Using WNCCC, we are not worried about the normalization or circular shifting procedures anymore. This metric can be effectively applied in comparison of images containing whole bullet surfaces.

3.4. Data

In this study, the images were acquired using a Dino-Lite microscope (model AD7013MTL). This device delivers 10× to 92× zoom magnification with extended working distance and a high resolution 5.0MP (Mega Pixel) microscope camera. The encasing of this microscope was an aluminum alloy shell which offers a good safety measure in demanding work environments. Other features of this microscope are 8 On/off white LED lights, a Micro Touch image capture, measuring software and removable nozzles including a polarization nozzle for avoiding undesirable reflection lights from metal surfaces. The pole stand in this microscope is a machine-controlled rotating pole with a 360° rotation capability and provides up to 8″ of vertical working distance. The microscope can capture RGB color images with a resolution of 1280 × 1240 pixels. In order to precisely acquire the image of a single bullet surface, each time a single land or groove engraved area was captured and stored. The stored images were then carefully positioned next to each other to comprise the whole bullet image. For ease of use, the color images were transformed to 2-D gray-scale images. The gray-scale images were stored in a database for experiments and comparison. For this study, the bullet image database included 180 2-D gray-level images from 90 different classes. The bullets were test fired in laboratory from 90 different AK-47 assault rifles. All bullets were full metal jacket (FMJ) with a caliber of 7.62x39mm, and mean weight of 123 grain from BBM brand and had 4 LEAs. The visual comparison of some images in this database was difficult because there were scratched or twisted land marks or uneven illuminations in different areas of some images.

It is worth mentioning that this database was the same database used in [10] for bullet identification.

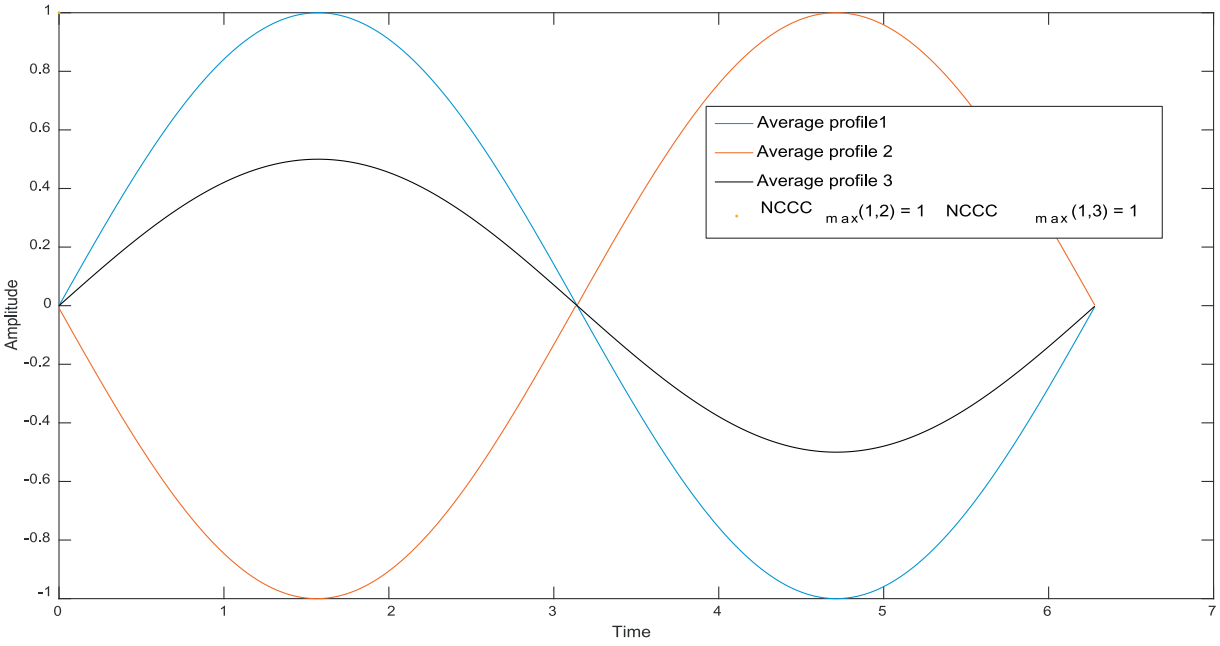


Fig. 7. An example of the NCCC problem caused by normalization procedure in comparison bullet surfaces.

4. Experiments and results

4.1. Experimental procedure

In order to evaluate the effects of the propositions (specifically NCCC, WNCCC and Kalman smoothing) in the previous section, we conducted several experiments. In each experiment, a specific algorithm was evaluated on our database. Each algorithm consisted of different combinations of propositions suggested in the previous section. These algorithms are demonstrated in Table 1. It should be noted that 'Algorithm_1' in Table 1 is the method used in [10] that proved to outperform conventional automatic bullet identification methods proposed in [4,7]. As a reminder for the readers, in Fig. 8, we represented the performance comparison results of our previously proposed method in [10] ('Algorithm_1') with the methods proposed in [4,7]. As can be seen in this figure, the algorithms proposed in [4,7] are beaten by 'Algorithm_1'. Therefore in this paper, we didn't compare these two conventional methods with our new algorithm.

By looking at Table 1, it can be easily figured that 'Algorithm_5' is actually our proposed method for automatic bullet identification in this paper. Some may ask why the range of IMfs for construction of smoothed profiles is fixed in 'Algorithm_1', 'Algorithm_2' and 'Algorithm_3'. Well, it was shown in [10] that using this range yields the best identification results on our database. It is worth to mention that based on the experiments in [10], for our database, 9 IMFs are enough to represent each profile signal and the number of sifting iterations was chosen to be 200 to ensure achieving better oscillatory components separation. In addition, the Standard Deviation criterion for terminating the sifting process is chosen to be 0.2.

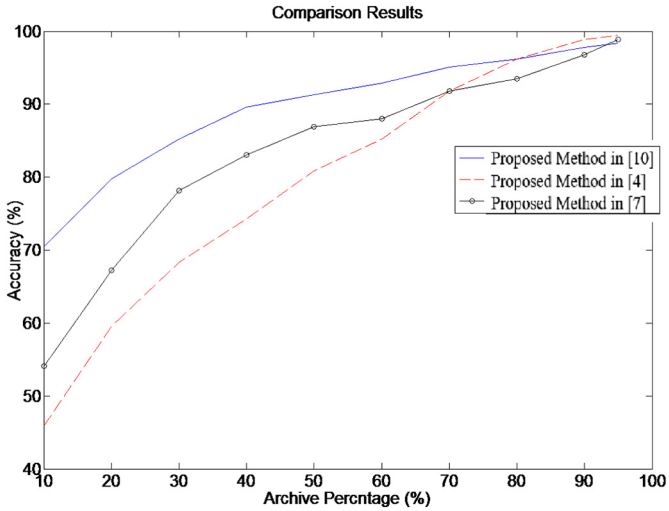


Fig. 8. Comparison results of the methods proposed in [10] ('Algorithm_1'), [4,7].

By comparing different algorithms presented in Table 1 with each other, we can determine whether a certain proposition improves the identification results or not. For example, by comparing the identification results of 'Algorithm_1', 'Algorithm_2' and 'Algorithm_3' we can decide which comparison metric (NCC or WNCCC or NCCC) is best for

Table 1
Benchmark algorithms for evaluation of proposed suggestions for automatic bullet identification.

Methods	Propositions		
	Image rotation for profile averaging	**Profile smoothing**	**Comparison metric**
'Algorithm_1' [10]	Canny edge detection & Radon transform	EEMD (IMfs:3rd-8th)	NCC
'Algorithm_2'	Canny edge detection & Radon transform	EEMD (IMfs:3rd-8th)	WNCCC
'Algorithm_3'	Canny edge detection & Radon transform	EEMD (IMfs:3rd-8th)	NCCC
'Algorithm_4'	Canny edge detection & Radon transform	EEMD (IMfs:1st-8th) & Kalman smoothing	NCCC
'Algorithm_5'	Canny edge detection & Radon transform	EEMD (IMfs:1st-8th) & Kalman smoothing	WNCCC

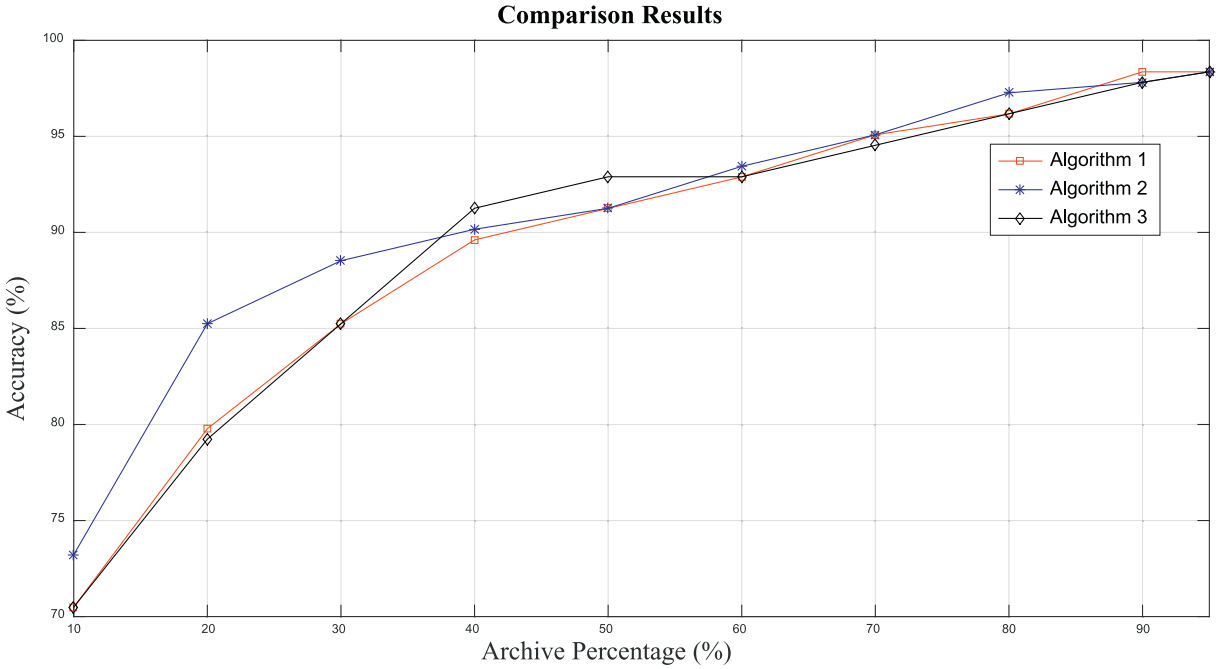


Fig. 9. Comparison results of 'Algorithm_1', 'Algorithm_2' and 'Algorithm_3'.

identification of images of whole bullet surfaces. The comparison of 'Algorithm_4' and 'Algorithm_5' can determine whether using WNCCC is better than NCCC or not. By comparing the identification results of 'Algorithm_2', 'Algorithm_5' the effect of using the proposed Kalman smoothing method can be illuminated.

4.2. Comparison of the proposed method

It is worth to mention that all of the algorithms mentioned in Table 1 were implemented using Matlab 2015Ra on a 4-core computer at 3 GHz with 4 GBs of RAM capable of parallel processing. The effectiveness of a particular identification algorithm can be defined by the probability of finding a match (hit) for different sizes of database. If these probabilities were plotted against different sizes of databases, a curve is visualized which is called the concentration curve [4]. If the area above this curve is small, it demonstrates the algorithm's efficiency for firearm identification and vice versa. In Fig. 9, we compared the performances of 'Algorithm_1', 'Algorithm_2' and 'Algorithm_3' for different sizes of database. It can be seen that 'Algorithm_2' has outperformed the other two algorithms. For example, for a hit list of 10%, 'Algorithm_2' reached 73.22% accuracy whereas the other two reached no > 70.49% accuracy. It can also be noticed from Fig. 8 that 'Algorithm_2' has an accuracy of over 85.25% for the archive percentage of 20%. However, 'Algorithm_1', and 'Algorithm_3' reached this accuracy for the archive percentage of 30%. Although, 'Algorithm_3' excels 'Algorithm_2' for archive percentages of 40% and 50%, the overall area above the concentration curve of 'Algorithm_2' is smaller than the area above the curve of 'Algorithm_3'. This fact implies that the overall performance of 'Algorithm_2' is better than 'Algorithm_3'. Because the three algorithms in Fig. 9 used similar profile smoothing process, the outperformance of 'Algorithm_2' in this figure means that the WNCCC metric is more reliable in comparison to NCC and NCCC.

The reliability of WNCCC over NCCC can also be seen in Fig. 10. In this figure, the performances of 'Algorithm_4' and 'Algorithm_5' are compared with each other. It can be seen from this figure that although both algorithms use the same profile smoothing technique, 'Algorithm_4' outclasses 'Algorithm_5' because of using WNCCC. In Figs. 8 and 9, the effectiveness of WNCCC is investigated and the results from these figures shows that the proposed WNCCC comparison metric is more reliable and consistent than conventional NCC in identification of images containing whole bullet surfaces. Another interesting fact that can be understood by looking at both Figs. 9 and 10 is that our proposed algorithm ('Algorithm_5') has overall the best accuracy for different sizes of hit lists and it could reach the accuracy of 77.49% for archive percentage of 10% and accuracy of over 90% for archive percentage of 35%. Needless to say that the previously proposed method in [10], yield the accuracy of 70.49% for archive percentage of 10% and barely overpassed the accuracy of 90% after archive percentage of 40%. 5% less archive percentage means a lot for experts and saves their time and energy while searching for matches in very large databases.

Another experiment was made to investigate the effectiveness of the proposed profile smoothing method. In this experiment, the performances of 'Algorithm_2' and 'Algorithm_5' were compared with each other. Since both algorithms used the same comparison metric (WNCCC), this comparison could determine whether our proposed profile smoothing method improves the identification accuracy or not. In Fig. 11, the concentration curves of the two above mentioned algorithms are plotted for different sizes of hit lists. As can be seen in this figure, the proposed method ('Algorithm_5') demonstrates much better results in comparison to the other method. For example, for archive percentage of 10%, 'Algorithm_5' reached the accuracy of 77.49% which shows over 4% improvement in comparison to 'Algorithm_2' and 7% improvement compared with 'Algorithm_1'. These noticeable improvements for small archive percentages indicate that our algorithm provides more reliable results when we are forced to search for a match in a limited number of candidates in the database. In other words, the proposed method helps firearm experts to save time and effort in finding true matches from a collection of large number of candidates. The comparison results presented in Figs. 9–11 confirm our previous statements in which we claimed that using Kalman smoothing instead of manually choosing the IMFs, preserves the useful information of averaged profiles and improves the identification accuracy.

Another fact that can be observed from the curves in Figs. 9–11 is the variability in performances of algorithms as a function of archive percentages. This variability is due to unwanted variations in illumination in the process of image acquisition. These variations caused some image profiles to be unidentifiable even for large archive percentages. We did not remove these profiles from the database in order

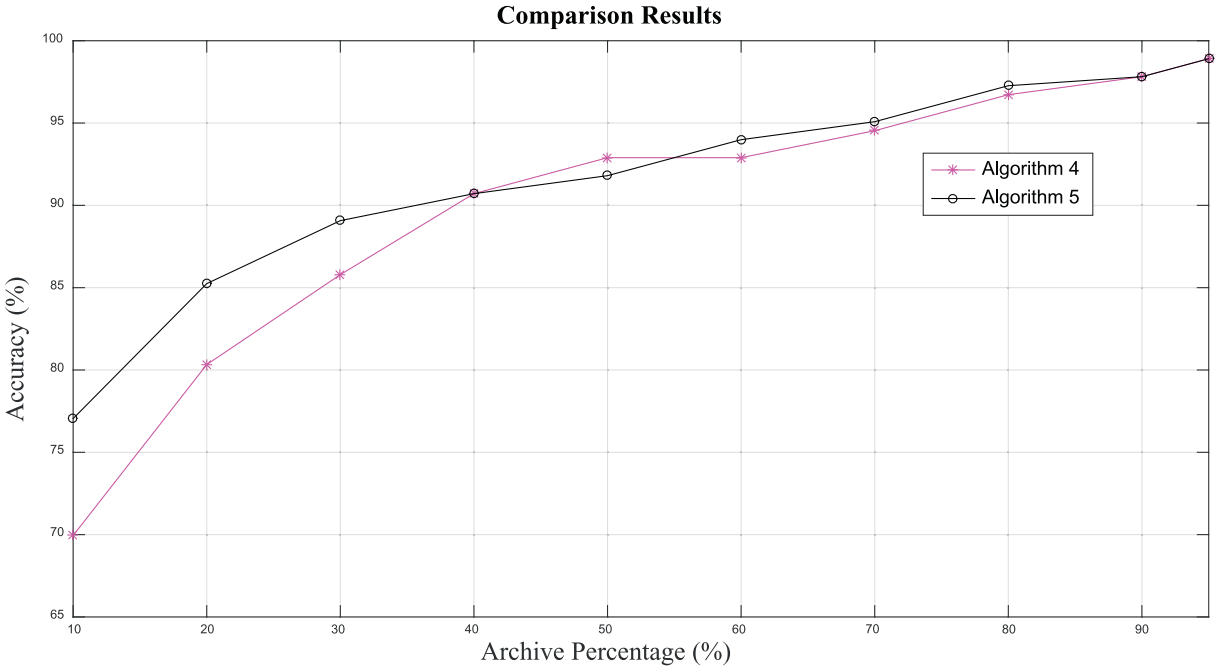


Fig. 10. Comparison results of 'Algorithm_4' and 'Algorithm_5'.

to simulate a realistic situation when amateur users predictably acquire images in non-standard conditions. Unfortunately the number of firearms in our laboratory is limited to 91 AK-47 assault rifles and we couldn't provide a larger database to acquire smoother curves in Figs. 9–11.

5. Conclusion

In this paper, we proposed a novel method which is based on EEMD and Bayesian Kalman filter for bullet identification. In the field of firearm identification, this is the first effort that reveals the benefits of using a statistical signal processing technique i.e. the Kalman filter for optimal denoising of firearm image profiles. Moreover, in order to improve the identification accuracy, we proposed a novel circular shift invariant metric called WNCCC which proved to be effective in comparison to other conventional metrics such NCCC and NCC. The WNCCC metric is independent to the start and endpoints of image profiles and uses the characteristics of both NCCC and Euclidean metrics for comparison. In our method, the nonlinear baseline drifts of averaged bullet image profiles are removed using EEMD algorithm but unlike the previous study, the high frequency noise in bullet average profiles are automatically removed using an adaptive Kalman filter. The experiments in this paper showed that the proposed algorithm outperforms previous techniques in identification of images containing whole bullet

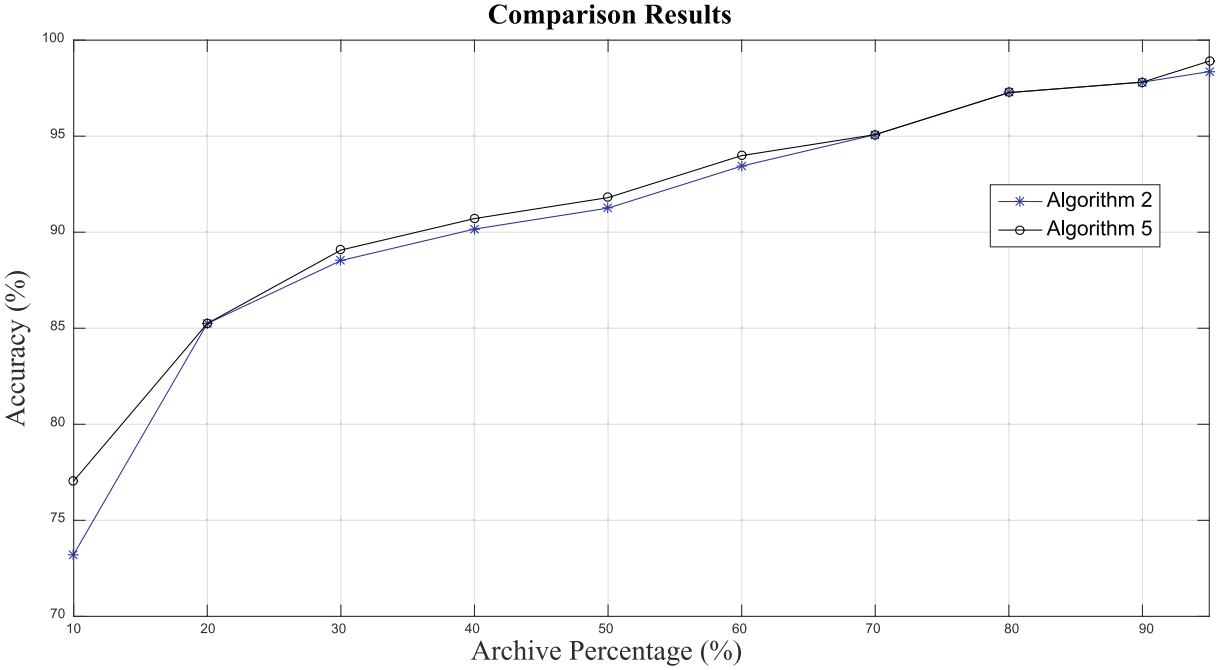


Fig. 11. Comparison results of 'Algorithm_2' and 'Algorithm_5'.

surfaces.

Conflict of interest

Authors of this article declare that he have no conflict of interest.

References

[1] R.G. Nichols, Defending the scientific foundations of the firearms and tool mark identification discipline: responding to recent challenges, J. Forensic Sci. 52 (3) (2007) 586–594.
[2] A.A. Braga, G.L. Pierce, Linking crime guns: the impact of ballistics imaging technology on the productivity of the Boston Police Department's Ballistics Unit, J. Forensic Sci. 49 (4) (2004) 1–6.
[3] R.W. Sibert, Drugfire: revolutionizing forensic firearms identification and providing the foundation for a national firearms identification network, Crime Lab. Digest. 21 (4) (1994) 63–67.
[4] F.P. León, Automated comparison of firearm bullets, Forensic Sci. Int. 156 (1) (2006) 40–50.
[5] W. Chu, J. Song, T. Vorburger, S. Ballou, Striation density for predicting the identifiability of fired bullets with automated inspection systems, J. Forensic Sci. 55 (5) (2010) 1222–1226.
[6] W. Chu, J. Song, T. Vorburger, J. Yen, S. Ballou, B. Bachrach, Pilot study of automated bullet signature identification based on topography measurements and correlations, J. Forensic Sci. 55 (2) (2010) 341–347.
[7] W. Chu, J. Song, T.V. Vorburger, R. Thompson, R. Silver, Selecting valid correlation areas for automated bullet identification system based on striation detection, J. Res. Nat. Inst. Stand. Technol. 116 (3) (2011) 647–653.
[8] W. Chu, R.M. Thompson, J. Song, T.V. Vorburger, Automatic identification of bullet signatures based on consecutive matching striae (CMS) criteria, Forensic Sci. Int. 231 (1) (2013) 137–141.
[9] J. Lu, S.-H. Wu, K.-C. Yang, M. Xia, Automated bullet identification based on striation feature using 3D laser color scanner, Optik Int. J. Light Electron Optics 125 (10) (2014) 2270–2273.
[10] S. Bigdeli, H. Danandeh, M.E. Moghaddam, A correlation based bullet identification method using empirical mode decomposition, Forensic Sci. Int. 278 (2017) 351–360.
[11] E. Hare, H. Hofmann, A. Carriquiry, Automatic matching of bullet land impressions, Ann. Appl. Stat. 11 (4) (2017) 2332–2356.
[12] B. Bachrach, Development of a 3D-based automated firearms evidence comparison system, J. Forensic Sci. 47 (6) (2002) 1253–1264.
[13] M. Heizmann, Automated comparison of striation marks with the system GE/2, Proc. AeroSense (2002) 80–91.
[14] A. Biasotti, J. Murdock, Firearms and toolmark identification: legal issues and scientific status, Modern Sci. Evid. Law Sci. Expert Testimony (1997) 124–151.
[15] Z.J. 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 Sci. Int. 119 (1) (2001) 97–106.
[16] J.B.Z. Geradts, R. Hermsen, F. Murtagh, Matching algorithms using wavelet transforms for a database of spent cartridge cases of firearms, Proc. SPIE 4232 (2001) 545–552.
[17] N.E. Huang, Z. Shen, S.R. Long, M.C. Wu, H.H. Shih, Q. Zheng, et al., The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis, *Proc. Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 1998, pp. 903–995.
[18] Z. Wu, N.E. Huang, Ensemble empirical mode decomposition: a noise-assisted data analysis method, Adv. Adapt. Data Anal. 1 (1) (2009) 1–41.
[19] P. Zarchan, Progress In Astronautics and Aeronautics: Fundamentals of Kalman Filtering: A Practical Approach, vol. 208, Aiaa, 2005.
[20] D. Simon, Optimal state estimation: Kalman, H infinity, and nonlinear approaches, John Wiley & Sons, 2006.
[21] J. Hartikainen, A. Solin, S. Särkkä, Optimal filtering with Kalman filters and smoothers, vol. 16, *Department of Biomedica Engineering and Computational Sciences, Aalto University School of Science, Greater Helsinki, Finland*, 2011.
[22] R.H. Shumway, D.S. Stoffer, Time series analysis and its applications, Stud. Informa. Control 9 (4) (2000) 375–376.
[23] S. Gibson, B. Ninness, Robust maximum-likelihood estimation of multivariable dynamic systems, Automatica 41 (10) (2005) 1667–1682.
[24] T. Kailath, A.H. Sayed, B. Hassibi, Linear estimation, vol. 1, Prentice Hall Upper Saddle River, NJ, 2000.
[25] G. Girija, J. Raol, R.A. Raj, S. Kashyap, Tracking filter and multi-sensor data fusion, Sadhana 25 (2) (2000) 159–167.
[26] M.L. Psiaki, Square-root information filtering and fixed-interval smoothing with singularities, Automatica 35 (7) (1999) 1323–1331.
[27] G.H. Golub, C.F. Van Loan, Matrix computations, vol. 3, JHU Press, 2012.
[28] Z. Ghahramani, G.E. Hinton, Parameter Estimation for Linear Dynamical Systems, Technical Report CRG-TR-96-2, University of Toronto, Dept. of Computer Science, 1996.
[29] S.J. Wright, Modified Cholesky factorizations in interior-point algorithms for linear programming, SIAM J. Optim. 9 (4) (1999) 1159–1191.
[30] T.V. Vorburger, J. Song, N. Petraco, Topography measurements and applications in ballistics and tool mark identifications, Surf. Topogr. Metrol. Prop. 4 (1) (2015) 1–52.