NOTICE: This paper was published in Measurement Science and Technology Vol. 25 Issue 05.
The final paper can be found here (<http://iopscience.iop.org/0957-0233/25/5/055105>) and the DOI number is as follows (doi:10.1088/0957-0233/25/5/055105).

KNN-classification of Metallic Targets using the Magnetic Polarisability Tensor

J Makkonen^{1,4}, L A Marsh², J Vihonen¹, A Järvi³, D W Armitage², A Visa¹ and A J Peyton²

¹ Tampere University of Technology, Department of Signal Processing,
Korkeakoulunkatu 10, P.O. Box 553, FIN-33101 Tampere, Finland

² School of Electrical and Electronic Engineering, The University of Manchester,
Manchester, M13 9PL, UK

³ Rapiscan Systems Oy, Klovinpellontie 3, Torni 2, FIN-02180 Espoo, Finland

E-mail: jarmo.makkonen@tut.fi

NOTICE: this is the author's version of a work that was accepted for publication in Measurement Science and Technology. Changes resulting from the publishing process, such as peer review, editing, corrections, structural formatting, and other quality control mechanisms may not be reflected in this document. Changes may have been made to this work since it was submitted for publication. A definitive version was subsequently published in Meas. Sci. Technol. Vol. 25 (2014), 055105 (9pp).
DIO link: [doi:10.1088/0957-0233/25/5/055105](https://doi.org/10.1088/0957-0233/25/5/055105)

⁴ To whom any correspondence should be addressed.

KNN-classification of Metallic Targets using the Magnetic Polarisability Tensor

J Makkonen¹, L A Marsh², J Vihonen¹, A Järvi³, D W Armitage², A Visa¹ and A J Peyton²

¹Tampere University of Technology, Department of Signal Processing,
Korkeakoulunkatu 10, P.O. Box 553, FIN-33101 Tampere, Finland

²School of Electrical and Electronic Engineering, The University of Manchester,
Manchester, M13 9PL, UK

³Rapiscan Systems Oy, Klovinpellontie 3, Torni 2, FIN-02180 Espoo, Finland

E-mail: jarmo.makkonen@tut.fi

Abstract. Walk-through metal detectors are used at check points for preventing personnel and passengers from carrying threatening metallic objects, such as knives and guns into a secure area. These systems are capable of detecting small metallic items, such as handcuff keys and blades, but are unable to distinguish accurately between threatening objects and innocuous items. This paper studies the extent to which a K-Nearest Neighbour -classifier can distinguish various kinds of metallic objects, such as knives, shoe shanks, belts and containers. The classifier uses features extracted from the magnetic polarisability tensor, which represents the electromagnetic properties of the object. The tests include distinguishing threatening objects from innocuous ones, classifying a set of objects into 13 classes, and distinguishing between several similar objects within an object class. A walk-through metal detection system is used as source for the test data, which consists of 835 scans and 67 objects. The results presented show a typical success rate of over 95% for recognizing threats, and over 85% for correct classification. In addition, we have shown that the system is capable of distinguishing between similar objects reliably. Overall, the method shows promise for the field of security screening and suggests the need for further research.

Keywords: Eigenvalues

1. Introduction

Walk-through metal detectors (WTMDs) are used in environments such as airports, prisons and public buildings for detection of threat objects. Their widespread use, combined with the competition between several leading manufacturers, has ensured that the current generation of technology is capable of detecting very small items such as handcuff keys and integrated circuits [1]. In the aviation industry, metallic threats are principally considered to be knives, guns, and gun parts. They are thus considerably larger

than the smallest detectable objects. Improved signal-to-noise ratios may continue to be sought by manufacturers in an attempt to increase the detectability of non-magnetic, low conductivity materials such as some stainless steels, however the most significant challenge is no longer to detect the objects, but to determine whether they present a threat or not.

It is reported that between 2000 and 2010 a total of 17 out of an estimated 310 million commercial flights were hijacked worldwide [2]. Using these numbers, it is estimated that in this ten year period the chances of a plane being hijacked was in the region of 1 in 18.2 million. In contrast to this threat, in environments such as airports, it is common for people to carry several innocuous items containing metal, such as mobile phones, jewellery, and keys. Current regulations require that travellers must remove these items prior to WTMD screening. This causes a degree of disruption and inconvenience, and requires a large number of staff to administer. Any method capable of reliably recognizing the items located on a person can remove the need for manual inspections, or even make it possible to be screened without removing metallic items. To our knowledge, there are few published studies on performing an intelligent classification apart from [3].

This paper describes a K-nearest neighbour (KNN) based classifier algorithm, first proposed by [3], and studies the extent to which it can distinguish between different objects that may typically be encountered with a WTMD. This is demonstrated with several test cases. Real walkthrough data is used for the testing. An advanced metal detection system [4] which is capable of calculating the position and magnetic polarisability tensor, $\vec{\mathbf{M}}$, of one or more metallic objects from a single walk-through scan has been used as the WTMD and the measurement system. A library of tensors produced by this system has been recorded for a variety of different objects.

The purpose of this paper is to investigate the potential of the methods proposed above using data from the WTMD portal. The portal is the first of its type and the deviation in the tensor components between different object locations and orientations can be significant. However, in theory, the eigenvalues for measured object tensors are considered to be rotation and location invariant, as discussed in section 3.2. Hence, we work under this assumption and use measurements from a single location and object orientation known to yield results with good signal to noise ratio; these results should then be generalizable for other locations and orientations.

The data library used in this paper is not exhaustive. The classes have been chosen arbitrarily, but are a representative sample of the range of objects that are commonly carried by people. Moreover, the classes contain only a sparse sample of the range of objects that might be included in a real world scenario. Finally, the overall amount of measurements is considered to be sufficient to prove whether further collection of a larger library is justifiable.

The rest of the paper is organized as follows. The following section contains background and related work to this study. The third section describes the experimental system i.e. WTMD and tensor inversion, along with the data library for the study. The

fourth section presents the method. The fifth section describes the experimental results while section six discusses the results. Finally, chapter seven concludes the paper and provides suggestions for future research. Although the focus of this paper is on aviation industry, the methods documented may be applied to other types of metal detection systems capable of yielding the magnetic polarisability tensor, e.g. detection systems for UXO and landmines.

2. Background

WTMDs have been in use since the 1970s, and the technical development has lead them to be highly accurate systems. However, there have been few attempts to add intelligence to them. Steps to this direction have been taken by Kauppila et al. [3], and Marsh et al. [4] that have made it possible to determine the polarisability tensor and the path of a metallic object passing through a real working WTMD. This is based on an inversion algorithm that is covered in section 3.2.

Pattern classification is an old research problem, and its theory has developed throughout decades. The KNN classifier is one of the simplest and most well-known algorithms for this, and countless modifications have been made to it for tackling a variety of classification problems. It is easy to implement, needs only few parameters, and it can be applied to many kinds of data. Also, the algorithm is capable of modeling arbitrary decision surfaces based on the data. However, its performance is highly dependent on the underlying example library, which has to be large enough. The large database might lead to slow classification performance and increased system memory requirements. Moreover, performance depends on the explicitly defined distance function that is used for comparing samples. [5]

KNN has been applied to landmine and UXO detection by Fails et al. and Tantum et al. [6, 7]. Compared to our work, these use distinct methods of characterizing the electromagnetic responses of the objects. Bell et al. [8] and Norton et al. [9] have used eigenvalues of the magnetic polarisability tensor for distinguishing between unexploded ordnance (UXO) and clutter, but have not used KNN for the classification. Kauppila et al. [3] have extended the idea and used a KNN classifier for classifying the eigenvalues. We use the same algorithm, but on a more advanced WTMD system.

3. The walk-through metal detector measurement system

This chapter provides background information on the measurement system, the inversion algorithm and their use for measurements. Together they provide the data for the study. The WTMD measurement system will be referred to as *the portal*. It should be noted that the method reported in this paper may be used with any system capable of consistently inverting the magnetic polarisability tensor.

3.1. The portal hardware

The WTMD described in [4] was used as a data acquisition system. This system uses sixteen coils, eight of which are transmitters which operate at different frequencies. These frequencies range from 8 kHz to 13.8 kHz and are arranged lowest to highest from floor level upwards, with a separation of approximately 700 Hz between neighbouring channels. The coil geometry, along with the portal, are shown in figure 1. The system is capable of taking 320 complex measurements per second across this detector volume and is capable of inverting the magnetic polarisability tensor to a typical accuracy of 10% within 5 seconds of a single walk-through scan.

Figure 1. Left: Coil geometry of WTMD data acquisition system. Right: The portal hardware.

3.2. Tensor inversion

The system measures the change in signal caused by objects passing through the portal space. We would like to get information on the characteristics of the object, i.e. its material, size and shape. To achieve this, we estimate what kind of an object could have caused the measured change in signal. For this purpose, we use a magnetic polarisability tensor model. It refers typically to a complex-valued symmetrical 3-by-3 matrix

$$\overset{\leftrightarrow}{\mathbf{M}}(f) = \begin{bmatrix} m_{x,x} & m_{x,y} & m_{x,z} \\ m_{y,x} & m_{y,y} & m_{y,z} \\ m_{z,x} & m_{z,y} & m_{z,z} \end{bmatrix} = \begin{bmatrix} M_{x,x} + jN_{x,x} & M_{x,y} + jN_{x,y} & M_{x,z} + jN_{x,z} \\ M_{y,x} + jN_{y,x} & M_{y,y} + jN_{y,y} & M_{y,z} + jN_{y,z} \\ M_{z,x} + jN_{z,x} & M_{z,y} + jN_{z,y} & M_{z,z} + jN_{z,z} \end{bmatrix} \quad (1)$$

which provides a point representation of the conductive and magnetic properties of an object. The tensor is symmetric such that $m_{x,y} = m_{y,x}$, $m_{x,z} = m_{z,x}$, and $m_{y,z} = m_{z,y}$ and it is possible to model the system response by

$$\rho(\mathbf{p}) = \mathbf{H}_t^T(\mathbf{p}) \overset{\leftrightarrow}{\mathbf{M}} \mathbf{H}_r(\mathbf{p}) \quad (2)$$

given that the object center position vector $\mathbf{p} = [x \ y \ z]^T$, the transmitter and receiver coil magnetic field vectors $\mathbf{H}_t$ and $\mathbf{H}_r$, and the object tensor $\overset{\leftrightarrow}{\mathbf{M}}$ are known. Consequently, in our WTMD application, $\overset{\leftrightarrow}{\mathbf{M}}$ and $\mathbf{p}$ are unknown and need to be estimated based on the system response and the model (2). This leads to an optimization problem to find the tensor $\overset{\leftrightarrow}{\mathbf{M}}$ that produces an approximate response ρ , minimizing the function

$$F = ||\rho_{\text{mes}} - \rho||_2^2. \quad (3)$$

where ρ_{mes} represents the measured system responses. This is solved with the Levenberg-Marquardt algorithm.

The reliability of the inversion, i.e. how well can the model fit the measured signal, can be estimated from the residual value, r . The residual is calculated by taking the L2-norm of the difference between the actual measurements, ρ_{mes} , and the forward response as a function of the inverted tensor and 3D coordinates, ρ , and dividing this value by the L2-norm of the inverted measurements, $\hat{\rho}$; this is defined in (4). The residual value can be also seen as a confidence value for the tensor.

$$r = \frac{\sqrt{\sum_{k=1}^n |\rho_k - \rho_{\text{mes}[k]}|^2}}{\sqrt{\sum_{k=1}^n |\rho_{\text{mes}[k]}|^2}} \quad (4)$$

3.3. Test library production

Test data were recorded by repeatedly conducting walk-through scans. One object was carried through the portal at a time, keeping the orientation and position constant. The position was approximately in the center of the portal horizontally, and at chest height vertically. The orientation was as follows. For a flat rectangular object, the object was kept flat against the chest, in such a way that the longest edge of the object was aligned horizontally. The data recorded for each walk-through consists of the estimated tensor value, $\overset{\leftrightarrow}{\mathbf{M}}$, the residual of the inversion algorithm, r , and the estimated coordinates for the path of the object, $\mathbf{P}$.

Since there were several problems to be studied, two test libraries were recorded. The first one, *Library A*, contains knives, metallic containers (different tins), shoe shanks, and a model gun with phantom bullet cases. The knives consist of six hunting knives

Table 1. The contents of *Library B*. Numbers in brackets represent the total number of measurements per class. The number of scans per object varies, but there is a minimum of five scans per object.

1. <i>Knives (200)</i> 14 Knives	2. <i>Model guns / bullet cases (180)</i> 8 Model guns 2 Sets of phantom bullet cases	3. <i>Scissors (30)</i> 1 Pair of nail scissors 1 Pair of office scissors
4. <i>Shoe shanks (90)</i> 6 Shoe shanks	5. <i>Containers (70)</i> 5 Containers	6. <i>Belts (25)</i> 4 Belts
7. <i>Wristwatches (20)</i> 3 Wristwatches	8. <i>Keys (20)</i> 4 Sets of keys	9. <i>Lighters (20)</i> 2 Lighters
10. <i>Jewellery (30)</i> 1 Bracelet 1 Necklace	11. <i>Coins (75)</i> 9 Sets of coins	12. <i>Artificial joints (15)</i> 1 Hip joint
13. <i>Mobile electronics (60)</i> 3 Mobile phones 2 MP3 players		

(HK) and four Swiss army knives (SK, Swiss knife). The hunting knives that we used are Opinel brand, while the Swiss army knives were of various brands. The model gun is a phantom of a Glock 17 [10], with a ferrous steel barrel and slide and polymer handle and magazine. Also, there is a set of phantom bullet cases made of brass. For the model gun with phantom bullet cases, five different objects, i.e. using a varying number of phantom bullet cases inside the model gun, were defined and 20 measurements conducted per object. For the rest of the objects; 10 knives, six containers, and six shoe shanks; 10 measurements each were conducted. The total size of *Library A* is 320 scans for 27 objects. The contents of *Library A* are shown in Table 2.

A broader range of objects was used for *Library B*. It contains *Library A* as a whole, but the additional measurements performed increased the total size to 835 scans for 67 objects. The contents of the library are shown Table 1. Some items are shown in Figure 2.

4. Classification algorithm

The method is based on the classification algorithm presented in [3]. For the rest of the paper, this is referred to as the *classifier*, whereas the whole method including tensor inversion is referred to as the *method*. The K-Nearest Neighbour classifier (KNN) is a simple algorithm that assumes that the training data is clustered using a certain feature space, that a similarity metric that is derived from the distance between samples, is enough to find the correct class for the sample. The algorithm firstly calculates the features for

Table 2. *Library A* contents, excluding the model gun and phantom bullet cases. *HK* stands for hunting knife, *SK* for Swiss knife, *SS* for shoe shank, *CS* for carbon steel, and *StS* for stainless steel. The dimensions are the blade lengths for the hunting knives, and *height x width* for the shoe shanks as they are thin objects.

Object	Material	Dim (mm)	Object	Material	Dim (mm)
HK1	CS	80	SK1	StS	90 x 26 x 17
HK2	CS	85	SK2	StS	85 x 27 x 19
HK3	CS	90	SK3	StS	90 x 26 x 17
HK4	CS	100	SK4	StS	58 x 19 x 9
HK5	CS	120	Tobacco tin (C1)	Unknown	100 x 76 x 29
HK6	StS	85	Cigarette box (C2)	Unknown	96 x 81 x 15
SS1	Unknown	100 x 23	Lighter box (C3)	Unknown	97 x 46 x 20
SS2	Unknown	110 x 12	Pocket mirror (C4)	Unknown	92 x 60 x 9
SS3	Unknown	110 x 15	Credit card holder (C5)	Unknown	93 x 55 x 6
SS4	Unknown	120 x 12	Zippo (C6)	Unknown	57 x 30 x 10
SS5	Unknown	128 x 14			
SS6	Unknown	135 x 15

Figure 2. Some objects from Library B. (a) NIJ test object, phantom stainless steel gun, and the model gun with some phantom bullet cases (b) Shoe shanks (c) Artificial hip joint.

each data point, i.e. tensor $\overset{\leftrightarrow}{\mathbf{M}}$. These are the eigenvalues of the tensor, given by

$$\boldsymbol{\lambda} = [\lambda_1 \ \lambda_2 \ \lambda_3] \quad (5)$$

They are a rotation-invariant representation of $\overset{\leftrightarrow}{\mathbf{M}}$. This is shown by the following short proof.

Let $\mathbf{R}$ be any rotation matrix that satisfies the conditions $\mathbf{R}^T = \mathbf{R}^{-1}$, $\det(\mathbf{R}) = 1$ and $\mathbf{R}\mathbf{R}^T = \mathbf{I}$, where $\mathbf{I}$ is the identity matrix. For any 3 by 3 matrix, a rotated version $\overset{\leftrightarrow}{\mathbf{M}}_R$ of $\overset{\leftrightarrow}{\mathbf{M}}$ is as follows:

$$\overset{\leftrightarrow}{\mathbf{M}}_R = \mathbf{R} \overset{\leftrightarrow}{\mathbf{M}} \mathbf{R}^{-1} \quad (6)$$

Consider an eigenvalue λ of $\overset{\leftrightarrow}{\mathbf{M}}_R$ and the corresponding eigenvector $\mathbf{v}$. By definition,

$\overset{\leftrightarrow}{\mathbf{M}}_R \mathbf{v} = \lambda \mathbf{v}$, given that $\mathbf{v} \neq 0$. Noting that $I\mathbf{v} = \mathbf{v}$, we may write $(\overset{\leftrightarrow}{\mathbf{M}}_R - \lambda I)\mathbf{v} = 0$. Evaluation of the characteristic polynomial yields

$$\begin{aligned} \det(\overset{\leftrightarrow}{\mathbf{M}}_R - \lambda I) &= \det(\mathbf{R} \overset{\leftrightarrow}{\mathbf{M}} \mathbf{R}^{-1} - \lambda I) = \det(\mathbf{R} \overset{\leftrightarrow}{\mathbf{M}} \mathbf{R}^{-1} - \mathbf{R} \lambda \mathbf{R}^{-1}) \\ &= \det(\mathbf{R}(\overset{\leftrightarrow}{\mathbf{M}} - \lambda I)\mathbf{R}^{-1}) = \det(\overset{\leftrightarrow}{\mathbf{M}} - \lambda I). \end{aligned}$$

Therefore, the eigenvalues of $\overset{\leftrightarrow}{\mathbf{M}}$ and $\overset{\leftrightarrow}{\mathbf{M}}_R$ are the same. Hence, the eigenvalues are rotation invariant.

As $\overset{\leftrightarrow}{\mathbf{M}}$ is complex its eigenvalues form a vector containing three complex values. These three values are then sorted in order of increasing *magnitude*, $||\lambda_k||$. In a 2D complex plane, it is the distance from the origin of the eigenvalue. The magnitude is given by

$$||\lambda_k|| = \sqrt{\lambda_k \cdot \bar{\lambda}_k} \quad (7)$$

where $\bar{\lambda}_k$ is the complex conjugate of λ_k . The eigenvalues are sorted because the distance calculation uses a pairwise comparison between them, and therefore the sorting acts like a guess to minimize the distance to the other objects in the class. The eigenvalues can be seen to exist in a polar coordinate system, and therefore in addition to the magnitude, the *angle* $\phi(\lambda_k)$ of the eigenvalue, as seen from the origin, is given by

$$\phi(\lambda_k) = \varphi = \text{atan2}(\text{Re}(\lambda_k), \text{Im}(\lambda_k)) \quad (8)$$

where atan2 is the arctangent function, and Re and Im are the real and imaginary parts of the eigenvalue, respectively.

Each object is classified by using (9) to compare the Euclidean distance, D_i between the objects sorted eigenvalues and all of the samples in the library, where the subscript i refers to the library index. The vector of distances is sorted in order of smallest to largest values to produce $\overset{\leftrightarrow}{\mathbf{M}}$, and the first K values are selected. The calculation of the Euclidean distance in the case of eigenvalues is shown in Figure 3. The classification outcome, ζ , is calculated by taking the statistical mode of the class of each of the K nearest distances as shown in (10); often K is an odd number to reduce the chances of even numbers of neighbours belonging to different, most popular classes.

$$D(\vec{\lambda}_a, \vec{\lambda}_b) = \sqrt{(\vec{\lambda}_{a,1} - \vec{\lambda}_{b,1})^2 + (\vec{\lambda}_{a,2} - \vec{\lambda}_{b,2})^2 + (\vec{\lambda}_{a,3} - \vec{\lambda}_{b,3})^2} \quad (9)$$

$$\zeta = \text{Mo}(\text{Class}\{S_1, \dots, S_k\}) \quad (10)$$

where Mo is the statistical mode.

5. Test cases and results

We investigate the performance of the algorithm by using four different scenarios. The first one is the most typical case for a WTMD: distinguishing between threatening objects

Figure 3. Principle of KNN Euclidean distance calculation based on eigenvalues. The distance here is shorter to class *B*, and is $D = d(B, 1) + d(B, 2) + d(B, 3)$. The X-axis represents the real component of the eigenvalue, whereas the y-axis represents the imaginary component.

(threat, T) and innocuous, i.e. non-threatening, objects (non-threat, T'). The second case is classifying objects to 13 object classes. The third one studies distinguishing knives from shoe shanks. The final one is a more fine-grained analysis, as we investigate the ability to distinguish a variety of knives from one another. Also, the same is done for shoe shanks and the model gun and phantom bullet cases.

The evaluations are carried out using a *leave one sample out* (LOO) cross validation (CV), and a 4-fold cross validation over 100 iterations. For the folds, three scenarios are used: 25% of the samples are used as test data and 75% are used as the library (training) data (25/75-CV); 50% - 50% (50/50-CV); and 75% test - 25% library (75/25-CV). The results quoted are the averages of all iterations.

Accuracy refers to the percentage of the samples that were correctly classified. Recall describes the percentage of the samples of a specific class that were correctly classified. Normalized accuracy (Norm. acc) is used for eliminating the effect of varying sample numbers between the classes. It is defined as the average recall over all the classes.

5.1. Test case 1 (TC1): Threats vs. innocuous objects

For the usefulness of the algorithm, its capability of separating threats and non-threats is crucial. Generally, threats are weapons and other dangerous objects. These are sharp items such as knives and scissors; and guns, their parts and bullet cases. Moreover, we consider metallic containers threats because they can conceal the aforementioned objects from the scanner, as the magnetic field does not enter them. Hence, if one was to put a knife in a metallic box, the knife may not be seen by a scanner. It should be noted that this threat definition differs from earlier studies [3], and therefore, comparing the results

Table 3. Results for TC1 and TC2. T recall (%) indicates how many percent of the threat objects were correctly classified. Class. (%) is the overall accuracy that indicates how many percent of all samples were correctly classified into 13 classes. Norm. (%) is the normalized accuracy that indicates the average recall over all 13 classes.

Library	K	CV	T vs. T' (%)	T recall (%)	Class. (%)	Norm. (%)
B	1	LOO	97.6	97.9	94.7	91.1
B	1	25/75-CV	97.5	97.8	94.5	90.8
B	1	50/50-CV	97.0	97.2	93.3	89.3
B	1	75/25-CV	95.4	95.8	90.1	85.0
B	3	LOO	97.0	97.3	93.3	89.2
B	5	LOO	96.3	95.3	91.1	86.6
B, $r < 0.35$	1	LOO	99.8	99.5	99.1	94.6
B, $r < 0.50$	1	LOO	98.6	99.1	97.1	94.2

is not straightforward.

We have divided Library B into threats and non-threats by tagging guns and parts, knives, scissors, and containers as threats. All remaining items are tagged non-threats. Hence, the library for this test case is as follows: 480 threats and 355 non-threats. For further testing, we have created additional libraries by removing all samples from the library that have a residual r greater than a threshold value. We call these libraries $B, r < \text{Threshold}$, where *Threshold* is replaced by the actual value.

The NIJ [11] defines general requirements for WTMD performance [12]. False negatives, i.e. classifying an object as a non-threat when it is really a threat, are not acceptable in the application area. Therefore, in official tests, any commercial WTMD must recognize all introduced threat objects correctly while maintaining a low false alarm rate. According to the NIJ, the false positive rate can be up to 25%. This means that one out of four non-threats can be classified as a threat. The false positives are not considered to be dangerous; however they cause unnecessary time delays as there is a subsequent need for manual inspection.

The results for TC1 are shown in Table 3. The threat recall scores (T recall (%)) mean that one object of class T in 20 to 200 would be falsely classified as T' . Furthermore, the high overall accuracies (Class. (%)) gained indicate a much lower false alarm rate than required by NIJ.

5.2. Test Case 2 (TC2): Object classification

We have divided Library B into 13 classes of objects, as shown in Table 1. As in 5.1, additional libraries called $B, r < \text{Threshold}$ are also used. The classifier was run with the same test setups as in TC1.

The results for all tests are shown in Table 3. The normalized confusion matrix for the case *Library B*, $K=1$, *LOO* is shown in Figure 4. It should be noted that object

classes *Keys* and *Artificial Joints* have ceased to exist in library *B*, $r < 0.35$, since these classes yield high residual scores. Therefore, the result is for 11 classes instead of 13. The lowest overall accuracy was 85%. On the other hand, the best accuracy of around 99% demonstrates great potential.

Figure 4. Normalized confusion matrix for *Library B*, $K=1$, *LOO*. Each row represents an object class. The colours show to which object classes the classifier has assigned the samples. Correctly classified samples are on the diagonal.

5.3. Test Case 3 (TC3): Knives vs. shoe shanks

Shoe shanks present a difficult problem for WTMDs, since they have a similar shape and material as some knives. Furthermore, many people do not realize that their shoes could contain metal and therefore will trigger the WTMD to alarm. In addition, taking one's shoes off and putting them back on is considered to be a time consuming and a frustrating operation. Some airports, though, enforce this in order to reduce false alarms.

To study this problem, we set up a library of six shoe shanks and 10 knives, 10 measurements per object. Hence, the library contains only two classes. Figure 5 shows the eigenvalues for the used test data. It can be seen that the angles of the eigenvalues differ between the two object groups.

The classification algorithm was run with the same cross validations as in the previous experiments, but a *leave one object out* cross validation was added. This means that there were no samples of the object to be recognized in the KNN library, and therefore the problem is much harder for the classifier. This gives us an idea whether the classifier can make generalizations of the two classes by using a limited number of objects in the library. The results are shown in Table 4. The results demonstrate that the classifier can distinguish between knives and shoe shanks with an accuracy greater than 93% in all cases.

Table 4. Results for TC3

K	CV	Accuracy (%)	Knife recall (%)
1	LOO	100	100
1	25/75-CV	100	100
1	50/50-CV	100	100
1	75/25-CV	99.5	99.3
1	Leave one object out	93.8	90.0

Figure 5. Eigenvalues for knives and shoe shanks. The circles represent knives, while the dots are shoe shanks. The light shaded circles are the 10 knife samples that were misclassified in the *leave one object out* case.**Table 5.** Results for TC4

K	CV	Knives (%)	Shoe shanks (%)	Containers (%)	Model gun/bullet cases (%)
1	LOO	99.0	93.0	100	98.0
1	25/75-CV	99.0	92.8	99.1	96.9
1	50/50-CV	99.0	91.3	98.1	95.2
1	75/25-CV	99.0	88.0	94.2	91.7

5.4. Test Case 4 (TC4): Resolution studies

In order to get an idea of the resolution of the method, we have classified different groups of similar objects in Library A: knives, shoe shanks, metallic containers, and the model gun with phantom bullet cases. The classification results of the test cases are shown in Table 5.

5.4.1. Knife distinction There are 10 different knives in Library A, and they differ by size, shape and material. Many of the knives are very similar, typical difference in length

being 1 cm between nearest examples. The hunting knives have a thin metallic blade and a wooden handle, supported by a metallic ring. The Swiss knives are thick pieces of metal containing many small blades and tools, in addition to a larger blade that is approximately the length of the knife as a whole. The details for the knives are shown in Table 2.

Figure 6. Eigenvalues of hunting knives. The dots represent the middle and smallest eigenvalues, and the different symbols the largest ones.

The eigenvalues for the six different hunting knives are shown in Fig. 6. The knives HK1-HK5 are made of the same carbon steel material. This can be clearly seen from the fact that their largest eigenvalues seem to have the same angle. The magnitude of the sample cluster centre gives an indication of the length of the blade of the particular knife. However, the stainless steel knife, HK6, has a distinct angle. The blade of the knife is of the same length as the HK2, but the magnitudes of the two knife cluster centres differ significantly. This suggests that only when sample clusters have the same angle can assumptions of the object size be made.

Figure 7 shows the eigenvalues of the Swiss knives. It can be seen that the eigenvalues follow the same trend as the ones for the hunting knives, but that the magnitude of the middle eigenvalue component cluster centre is different. This is due to the thickness of the knives. Clearly, SK3 is the largest, and SK4 the smallest one. The proportions are somewhat related to those seen in Table 2. These observations suggest that assumptions on object material and dimensions might be made based on the eigenvalues.

All test cases for the knife distinction yielded an excellent 99% accuracy, as shown in Table 5.

5.4.2. Shoe shank distinction Shoe shanks are seemingly made of more or less the same material, and their dimensions and shape tend to be similar. The dimensions of the shoe

Figure 7. Eigenvalues for the Swiss army knives.

shanks in Library A are shown in Table 2, and their largest eigenvalue components are shown in Figure 8; the two other components are small and clustered together. It can be seen that although some of the clusters overlap, they seem mostly separable. The classification accuracy was an acceptable 88% - 93%. The results for this test are shown in Table 5.

Figure 8. Largest eigenvalues for the shoe shanks. The middle and smallest eigenvalues are clustered together near the origin since the objects are thin and narrow strips of metal.

5.4.3. Container distinction As it can be seen from Figure 9, the eigenvalues give information on the material and the physical dimensions of the objects. This is in line with the observations of the other tests presented in this section.

The tobacco tin (C1) is the largest object in size and consequently the magnitudes of its eigenvalues are the largest of the group. They form three clusters, the centres

Figure 9. Eigenvalues for the containers.

of which have large magnitudes. The cigarette tin (C2), on the other hand, is slightly smaller in two dimensions and significantly smaller in the other and that is mirrored in the eigenvalues. There are two clusters of the two objects that are close to each other and one cluster each that are separated. It is possible to see that the objects are made of a similar material, because the angles of the cluster centres are similar for both objects.

The rectangular markers representing the credit card holder (C5) suggest similar results. The cluster centres have the same angles as the aforementioned objects, suggesting a similar metal being used to construct it. However, in this case, one cluster is close to the clusters of the other two objects, while the other two clusters are much closer to the origin. This indicates that one dimension of the objects is close to the two other objects, and the two other dimensions are significantly smaller.

The zippo (C6), shown with pentagram markers, is the smallest object of the group, since all the eigenvalues are grouped near the origin, i.e. have small magnitude. The pocket mirror (C4) and the lighter box (C3) eigenvalues cluster centres have different angles, compared to the rest of the objects. Moreover, despite having one dimension approximately as large as the largest object, the cluster centre magnitudes are all significantly smaller than the ones of C1 and C2. This suggests that the magnitude is only comparable along the same angle when comparing object dimensions.

The results for the LOO cross validation yields a 100% accuracy of classifying the containers. Despite the physical similarity of the objects, this is hardly a surprise, given the clear separability between the eigenvalue clusters shown in Fig. 9.

5.4.4. Model gun with phantom bullet cases distinction For testing the algorithms capability of distinguishing small differences in objects, we use the steel model gun with a plastic handle. The model gun body (barrel) should yield a certain strong signal, whereas

the phantom bullet cases should look distinct in the feature space. We try to see whether the number of phantom bullet cases within the handle can be derived from the eigenvalues only. The test cases that we consider are: empty (no phantom bullet cases), 3, 8, and full 17 phantom bullet cases. For each of these cases, 20 measurements are taken in one location.

Figure 10. Eigenvalues for the model gun and phantom bullet cases. A set of clusters is zoomed in to show the similarity in shape between the phantom bullet case eigenvalue clusters and the others. The bullet signal appears to be summed on the phantom gun -signal

The eigenvalues for the model gun and phantom bullet cases are shown in Fig. 10. Clearly, the electromagnetic properties of the phantom bullet cases (shown with $\circ$ -markers) differ significantly from the other cases. The largest eigenvalues of the model gun cases seem to be clustered to around the same value. This is due to the large tensor component caused by the steel barrel of the model gun. However, the 3 phantom bullet cases situation is already some distance from the others in terms of the middle eigenvalue, and even further away in terms of the smallest eigenvalue. The other model gun cases are separated only in terms of the smallest eigenvalue. This suggests that the separation of the 3 phantom bullet cases case is due to a positioning issue of the phantom bullet cases, rather than the amount of phantom bullet case metal in the object. In the latter case, the samples for the empty model gun case and the 3 phantom bullet cases situation should be closer to one another.

The results shown in Table 5 indicate that the cases can be classified with over 91% accuracy. This is encouraging for the KNN algorithm, as the separability of the cases is not clear based on Figure 10.

6. Discussion

The results of TC1 suggest that threats and non-threats can be effectively distinguished by the method. However, the recall scores were inadequate for a real life application since they are not considered to be sufficiently accurate. Only when high residual values were removed from the library, and therefore some problematic items disappeared from the library altogether, were they close to 100% and therefore could pass commercial testing. This, however, suggests that the method is capable of reaching this level of accuracy if the underlying hardware can yield consistent, accurate data. Different cross-validations demonstrate that having only a few measured samples of the object in the library signifies a drop in accuracy. Increasing the value of K has a negative effect on accuracy scores.

TC2 shows that the method can distinguish between 13 classes at a very good accuracy, with the normalized values varying between 85% and 91%. This is not much lower than the scores in TC1, where there were only two classes. The results behave similarly to TC1 as a function of library size, the value of parameter K , and the quality of the tensors used. The confusion matrix shows that some classes are more difficult to recognize than others; these include belts, keys, and coins.

TC3 and TC4 suggest that it is possible to distinguish between intuitively similar objects. For all of the test scenarios, very high accuracy scores were recorded. Furthermore, the figures of the eigenvalues show clearly that the classes are indeed separable by a simple classifier. Moreover, especially the hunting knife comparison has suggested that object material and size can be yielded from the eigenvalues, suggesting there is more work needed on the separability of different materials, shapes and sizes of metallic objects.

This material and size information could be used for enhancing the classifier performance by using heuristics. Furthermore, there are more possibilities for doing this. For example, positional information derived from the inversion algorithm could be used for distinction between shoe shanks and other objects. Shoes are always near the floor level. Therefore, objects at chest level should not be classified as shoe shanks. Similarly, if an object was to be classified as a wristwatch then statistically it is very unlikely that it would be around the feet. On the other hand, knives could be hidden in boots or socks, and therefore making any assumptions like this should be done with great consideration. We have left this for future research.

The most obvious limitation of the KNN, the need for a large set of examples in the library, has so far prevented us from getting a realistic estimate on the true performance of a WTMD that uses our method. The worst case scenario for the classifier, and a realistic one, is that the actual test object to be classified is not in the training library. There are two ways of overcoming such a situation. The first one is to have a classifier capable of forming a generalization of the classes that could fit unseen objects reliably into them. The *leave one object out* cross validation that we performed in TC3 suggested that the classifier can cope with the scenario. The second, and probably easier way, is to attach

a certain level of trust into the classifications. Thus, in cases of uncertainty, a manual inspection could be conducted to determine the cause of the alarm. Moreover, this could allow the new previously unknown object to be added to the underlying library, hence enabling online learning.

7. Conclusions and future work

The results shown in this paper demonstrate that the algorithm is capable of classifying targets consistently and with a typical normalized accuracy of over 95%, and a recall value in excess of 99%. This is considered to be sufficient to suggest that this improvement shows great promise to the field of inductive metal detection. The tests have shown that similar objects of slightly different sizes can be distinguished with accuracies typically in the range of 85-100%.

However, there are a number of improvements that could be made to this algorithm and there is further research which can be conducted which the authors expect would improve the results shown here. The observations made on the eigenvalues of the hunting knives, along with similar results in TC4, suggest that information on object sizes and materials might be derived from this data. This indicates that some further heuristics could be applied to the classification algorithm. We aim to study the separability of different materials, shapes, and sizes of objects to get an understanding on the limitations of these heuristics. Also, a much broader library is needed for testing the algorithms capability to make a generalization of the classes. This library should contain samples of different object locations and orientations to enable testing of heuristics that are based on location information given by the inversion algorithm.

References

- [1] Rapiscan metor 6s product specification. www.rapiscan.com, August 2013.
- [2] T Bodington. Is the perceived threat of terrorism greater than the actual threat? <http://www.tsvc.lincoln.ac.uk/papers/submission.php?paper=118>, February 2011.
- [3] J. Kauppila, T. Ala-Kleemola, J. Vihonen, J. Jylhä, M. Ruotsalainen, A. Järvi, and A. Visa. Classification of items in a walk-through metal detector using time series of eigenvalues of the polarizability tensor. In *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, volume 7303, May 2009.
- [4] L. A. Marsh, C. Ktistis, A. Järvi, D. W. Armitage, and A. J. Peyton. Three-dimensional object location and inversion of the magnetic polarizability tensor at a single frequency using a walk-through metal detector. *Measurement Science and Technology*, 24(4):045102, April 2013.
- [5] Richard O. Duda, Peter E. Hart, and David G. Stork. *Pattern Classification (2nd Edition)*. Wiley-Interscience, 2000.
- [6] Eric B. Fails, Peter A. Torrione, Waymond R. Scott, Jr., and Leslie M. Collins. Performance of a four parameter model for modeling landmine signatures in frequency domain wideband electromagnetic induction detection systems. volume 6553, pages 65530D–65530D–8, 2007.
- [7] Stacy L. Tantom, Kenneth A. Colwell, Waymond R. Scott, Peter A. Torrione, Leslie M. Collins, and Kenneth D. Morton. Sparse model inversion and processing of spatial frequency-domain

- electromagnetic induction sensor array data for improved landmine discrimination. volume 8709, pages 87091E–87091E–8, 2013.
- [8] Thomas H. Bell, Bruce J. Barrow, and Jonathan T. Miller. Subsurface discrimination using electromagnetic induction sensors. *IEEE T. Geoscience and Remote Sensing*, 39(6):1286–1293, 2001.
 - [9] Stephen J. Norton, I.J. Won, and Ernesto R. Cespedes. Ordnance/clutter discrimination based on target eigenvalue analysis. *Subsurface Sensing Technologies and Applications*, 2(3):285–298, 2001.
 - [10] Glock 17 product specification. <http://us.glock.com/products/model/g17>, August 2013.
 - [11] The national institute of justice (nij). <http://www.nij.gov/>, June 2013.
 - [12] The national institute of justice (nij) -standard 0601.02: Walk-through metal detectors for use in concealed weapon and contraband detection. <https://www.ncjrs.gov/pdffiles1/nij/193510.pdf>, June 2013.

Acknowledgements

This work was supported by the Academy of Finland under the project Magnetic Induction and Data Analyses for Sophisticated Sensing, grant no. 251777.