

DETECTION AND CLASSIFICATION OF EXPLOSIVE SUBSTANCES IN MULTI-SPECTRAL IMAGE SEQUENCES USING LINEAR SUBSPACE MATCHING

Maria Axelsson*, Ola Friman, Ida Johansson, Markus Nordberg, and Henric Östmark

Swedish Defence Research Agency, Sweden

ABSTRACT

Fast detection and analysis of dangerous substances from longer distances is highly desired in many security applications. Imaging Raman spectroscopy is a novel multi-spectral imaging technique designed for stand-off screening and detection of explosive substances. In this paper we present a method for detection and classification of explosive substances in multi-spectral image sequences from imaging Raman spectroscopy using linear subspace matching. Our approach uses limited spectral information and is computationally efficient, which enables fast screening of interesting areas. The performance of the method is evaluated on real stand-off measurements from a demonstrator system. We show that the method can detect and classify substances with high accuracy.

Index Terms— Explosives, detection, Raman spectroscopy, multi-spectral imaging

1. INTRODUCTION AND RELATED WORK

There is a need to detect explosives, e.g., improvised explosive devices, in real environments at stand-off distances. The time to detect a deployed explosive device is often limited. However, detection of trace particles from production of the explosives can extend the available time to detect the threat before deployment. Detection of small amounts of particles on a surface among other unknown substances at stand-off distance is a challenging problem, but laser-based spectroscopy techniques, e.g., Raman spectroscopy, has shown promising results [1, 2].

Raman spectroscopy is a commonly used laser-based technique to analyze substances. A laser spot illuminates the target area and part of the scattered laser light is shifted in wavelength depending on the vibrational modes of the molecules in the target substance [3]. The shifted light is collected by a lens and passed to a spectrometer which records a densely sampled spectrum. The spectrum has peaks at the wavelengths of the returning shifted light which depends on the substance. This spectrum of the Raman scattered light

*The Swedish Contingencies Agency (MSB) is gratefully acknowledged for funding this study. Maria Axelsson is also partly funded by the Swedish Innovation Board (VINNOVA).

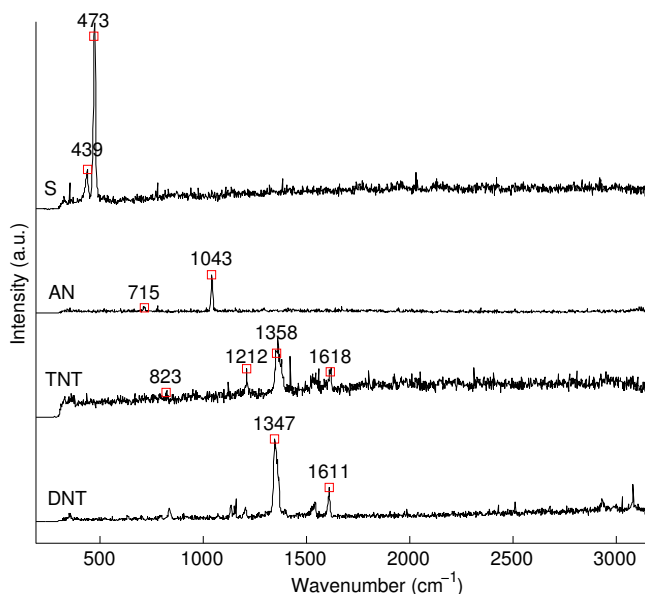


Fig. 1. Example stand-off Raman spectra for Sulfur (S), Ammonium Nitrate (AN), TNT, and DNT. The spectra are vertically offset for visibility and the Raman peaks used for detection and classification are marked.

provides a signature which can be used to classify the target substance.

Imaging Raman is a novel sensor technique based on Raman spectroscopy. It is designed to detect explosive substances at stand-off distance and meet the need to detect small particles [4, 5]. In contrast to conventional Raman spectroscopy [3, 6], which measure a single full spectrum from an area where spectra from object and background are mixed, an imaging Raman system uses a high resolution gated intensified charge-coupled device (ICCD) camera to resolve the target area and a tunable liquid-crystal filter to select the wavelength to image in each frame. The high-resolution image of the target area enables detection of small amounts of substance, such as trace particles, since the Raman scattering from any background substances is reduced in the pixels which contain the target substance. A multi-spectral image

sequence acquired by imaging Raman has two spatial and one spectral dimension. The spectral dimension is imaged sequentially and the ICCD is exposed to the light from one wavelength at the time. Due to the limited imaging time in real situations, the spectral dimension needs to be sparsely sampled. Therefore, all methods for detection and classification of substances must limit the required spectral information to enable real-time analysis. Figure 1 shows example spectra from imaging Raman which are densely sampled.

In the literature, only a few methods have been proposed for the analysis of data from imaging Raman. In [4], two methods are applied for detection of explosives. The first method calculates the sample correlation coefficient for the spectrum in each pixel and a reference spectrum for each substance. This approach only takes shape into account and not absolute signal level, which makes the method sensitive to both the signal level and noise in the background. The correlation image for each substance can be thresholded to obtain a classification result. The second method is a linear regression of each measured spectrum, using reference spectra from all substances and a constant spectrum as regressors. The constant spectrum is used as a model of the background. The classification is given by the substance with the largest regression coefficient. This method is sensitive to the scaling of the reference spectra and to the length of the spectra for substances with only few peaks.

In this paper we present a novel approach for detection and classification of explosive substances in multi-spectral image sequences from imaging Raman spectroscopy using linear subspace matching. Similar techniques for spectral matching have been used in medical imaging and hyper spectral imaging applications [7, 8, 9]. The proposed method is evaluated on real image data from imaging Raman spectroscopy.

The paper is structured as follows. The method is presented in Section 2, experiments and results on real data are presented in the evaluation in Section 3, and the paper is concluded in Section 4 with a discussion and conclusions.

2. METHOD

Two features are extracted from the Raman spectrum in each pixel: the peak to noise ratio (Section 2.1) and a shape feature based on spectral matching using linear regression (Section 2.2). The features are used for detection (Section 2.3) and classification (Section 2.4) of measured spectra. The detector and classifier are trained using images with marked areas for the respective classes.

2.1. Peak to noise ratio

The peak to noise ratio (PNR) is defined as the maximum of the measured spectrum divided by the mean noise level in the image (calculated as the mean value for a region in the back-

ground). The noise level can be calculated beforehand using the same imaging parameters and the peak value is then easily calculated for each measured spectrum during the imaging.

2.2. Shape matching

A spectral subspace model is constructed for each substance using training data from imaging Raman obtained at the peak locations in the respective spectral signatures. The peak locations can be obtained from, e.g., reference spectra measured with high spectral resolution using conventional Raman.

For each substance i , a matrix \mathbf{B}_i is constructed, with the training spectra in columns. The mean is subtracted from each spectrum and the matrix is decomposed using Singular Value Decomposition (SVD) [10]

$$\mathbf{B}_i = \mathbf{U}_i \mathbf{\Sigma}_i \mathbf{V}_i. \quad (1)$$

The first n vectors of \mathbf{U}_i are used as the subspace model \mathbf{X}_i . Two vectors are often enough to describe typical shape variations of each substance. Note the difference to the regression in [4], where one linear model is used for all substances.

The training of the subspace models is very simple, which makes it fast. This is useful if new substances must be added. The peak locations for all substances can be used in all models if there are few substances. If many substances are used, the spectral support for each model can be reduced to reduce the influence from background noise and the influence from other unknown substances.

To obtain the shape feature, the sparse spectrum in each spatial position of the measured multi-spectral image is matched to all models using linear subspace matching. The measured spectrum in a pixel, \mathbf{y} , is modeled as

$$\mathbf{y} = \mathbf{X}_i \boldsymbol{\beta}_i + \boldsymbol{\varepsilon}_i \quad (2)$$

using each spectral subspace model \mathbf{X}_i . The components of the coefficient vector $\boldsymbol{\beta}_i$ are identified using regular least squares and the fit of the measured spectrum to the model is measured using the score s_i

$$s_i = \frac{\|\mathbf{X}_i \boldsymbol{\beta}_i\|^2}{\|\mathbf{y}\|^2}, \quad (3)$$

which is the shape feature for each substance.

2.3. Detection

Detection of a substance is based on the PNR and the corresponding shape feature for the substance. An example of the PNR and the shape feature values for four substances (sulfur, ammonium nitrate, TNT, and DNT) are shown in Figure 2. In the detection, the PNR value is thresholded at a level corresponding to 99.9% correctly classified background pixels in training data. A detection map is calculated for each substance based on the corresponding shape feature in all detected pixels.

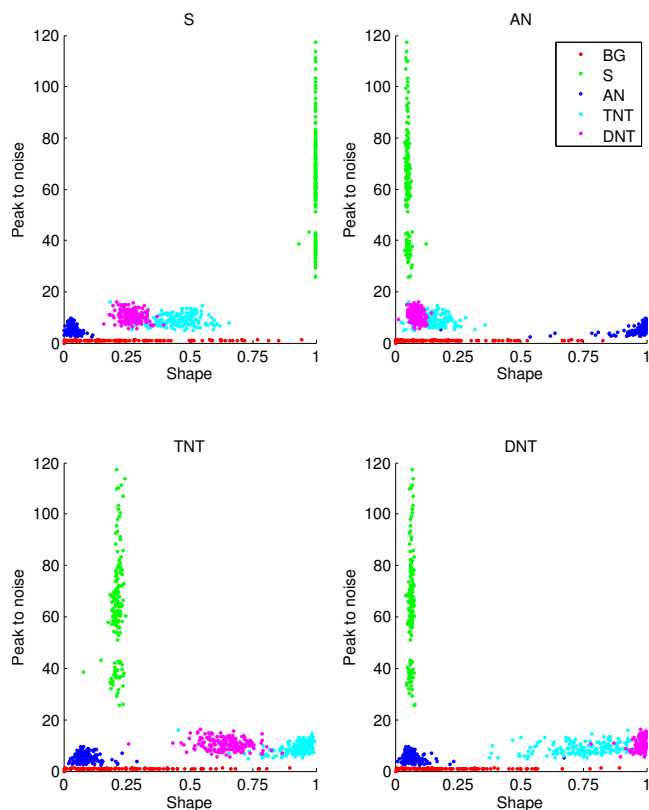


Fig. 2. An example of the PNR and shape feature values for the four substances and background (BG) in dataset *D2*.

2.4. Classification

Classification of the substances is computed per pixel in the image by finding the class with the maximum shape score above a threshold in the pixel using the detection maps for all classes.

Spatial smoothing of the result can be applied after the classification step if small particles are not of interest. The method for detection and classification is computationally efficient since it only needs to solve one least squares problem per model for all spectra in the image and calculate simple scores and find maximal values. All calculations can be applied per pixel.

3. EVALUATION

The accuracy of the method is evaluated using real datasets with ground truth. The method is also evaluated on images from a scenario with three small pieces of substance imaged through a car window from approximately 10 meters distance.

Measurements from a demonstrator system in a lab environment are used to evaluate the method. The demonstrator system has a gated ICCD camera with a sensor of 1024×1024 pixels. At 10 meters distance the imaged

Table 1. Confusion matrix for the classification of dataset *D1* and *D2* using five classes: background (BG) and the four modeled substances S, AN, TNT, and DNT.

<i>D1</i>		Predicted label				
		BG	S	AN	TNT	DNT
True label	BG	2742	1	0	1	0
	S	0	1196	0	0	0
	AN	0	0	1331	0	0
	TNT	0	1	0	1456	4
	DNT	0	0	0	0	1219
<i>D2</i>		Predicted label				
		BG	S	AN	TNT	DNT
True label	BG	6120	19	4	0	1
	S	0	7636	0	0	0
	AN	21	0	8602	0	13
	TNT	4	10	0	7674	244
	DNT	12	0	0	44	8279

surface area is 25×25 mm. Two datasets for accuracy evaluation were collected at approximately 10 meters stand-off. Samples from sulfur (S), ammonium nitrate (AN), 2,4,6-trinitrotoluene (TNT), and 2,4-dinitrotoluene (DNT) were used.

In the first dataset, *D1*, four multi-spectral image sequences, one for each substance, were imaged. The samples are round discs of about 10 mm in diameter which are placed in holes in an aluminum plate. In the second dataset, *D2*, six image sequences of different test plates were imaged with all substances in each multi-spectral image. The samples were placed in holes with a diameter of 2 mm in an aluminum plate. This dataset has much smaller pieces of the substances and it was imaged four months after the first. All images were collected at different times with the instrument, which also incorporate some variation in the system calibration parameters. Ground truth to evaluate the accuracy of the presented method was obtained by manually marking and labeling the interior of each substance in the images. Each spectral image used in the experiments for training and evaluation was median filtered using a small kernel to reduce influence from spurious noise. It is assumed that the area of a substance particle in the image is larger than the kernel size.

Measurements from the strongest peaks were selected for each substance to construct the models. In total ten peaks were selected for the models, see Figure 1. All peak locations were used to model all substances. If only one substance is of interest, a few measurements in the background must also be added for reference. The subspace models were constructed using a training set from dataset *D1*.

First the accuracy of the detection and classification method is evaluated on the marked evaluation set from *D1*. The evaluation set consists of 7951 measured spectra from the five classes, the four substances and the background. The

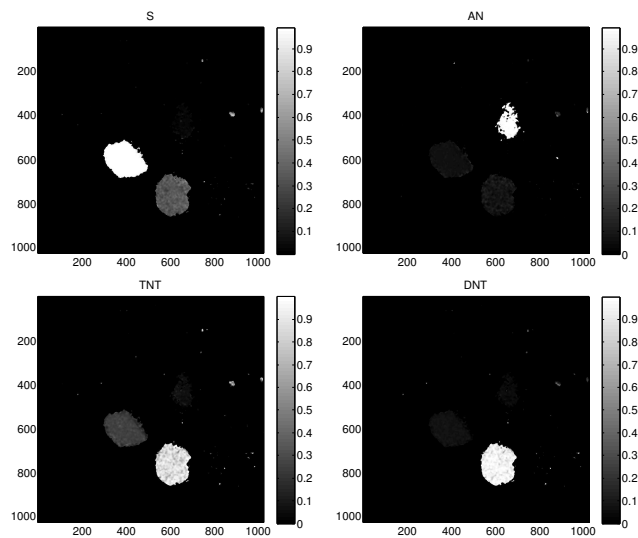


Fig. 3. Detection maps of substances on the steering wheel of a car. In each image the shape feature is shown for all the detected pixels for the respective substances. The corresponding classification using the shape threshold 0.7 is shown in Figure 4 (bottom right).

overall classification accuracy is 99.91%. However, even if the training and evaluation data are different the datasets are very similar. Table 1 shows the corresponding confusion matrix. The shape threshold is 0.5 in the experiments.

To investigate the variability, the model which was constructed using dataset *D1* is also used to classify the substances in dataset *D2*. This evaluation set consists of 38683 measured spectra from the five classes. The overall classification accuracy for all plates is 99.01% and Table 1 shows the corresponding confusion matrix. Some mixing of DNT and TNT is introduced, as can be expected, since they have similar molecular structure and share some of the Raman peaks.

Figure 3 and 4 show an example of detection and classification from a more realistic outdoor scenario where three substances, sulfur, ammonium nitrate, and DNT, have been placed on the steering wheel of a car and imaged at stand-off distance of about 10 meters through the side window of the car. Figure 3 shows the detection maps for all modeled substances. Figure 4 shows images of the experiment setup and the corresponding classification result. The subspace models trained on dataset *D1* was used with the shape threshold 0.7. In addition, only eight wavenumbers were used (since 823 and 1347 cm^{-1} were not available). All three substances are detected, but ammonium nitrate is more difficult to detect since the signal is only slightly higher than the noise level. In general, the signal to noise ratio (SNR) can be increased using longer integration time which will further improve the classification result, but increase the required imaging time.

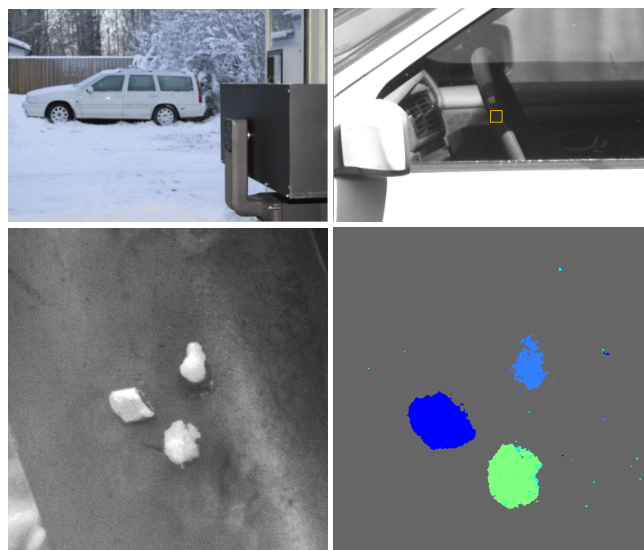


Fig. 4. Stand-off detection and classification of substances placed on the steering wheel of a car. The classified substances are sulfur (blue), ammonium nitrate (light blue), and DNT (green).

4. DISCUSSION AND CONCLUSION

We have presented a novel approach to detection and classification of explosive substances in multi-spectral image sequences from imaging Raman using linear subspace matching. The method is computationally efficient since it uses only limited spectral information. This also makes the image acquisition faster and enables the sensor to move to the next target area faster. We have shown that the method can be used to recognize substances in new unseen images with an overall classification accuracy of 99.01%. Only a few images collected at the wavelengths of the strongest peaks from each substance are needed to efficiently detect and classify all substances. We have also shown that the method provides good classification performance in real scenarios. Both the training and classification using the linear subspace method are simple and new classes can be added easily.

The method can be varied for different use cases. Depending on the required imaging situation, e.g., stand-off distance or available scan time, different scanning approaches can be used. For example a few representative peaks can be used for each substance in a first screening. Then, if a small group of pixels is detected with high accuracy as a known explosive or another substance of interest the area can be scanned again using additional peaks and a longer integration time.

In future work we intend to further investigate the sensitivity to noise and the needed variability using larger datasets with several substances. An investigation of the detection limits for different substances and the required imaging time is also of interest.

5. REFERENCES

- [1] S. Wallin, A. Pettersson, H. Östmark, and A. Hobro, "Laser-based standoff detection of explosives: a critical review," *Analytical and Bioanalytical Chemistry*, vol. 395, pp. 259–274, 2009.
- [2] A. Pettersson, I. Johansson, S. Wallin, M. Nordberg, and H. Östmark, "Near real-time standoff detection of explosives in a realistic outdoor environment at 55m distance," *Propellants, Explosives, Pyrotechnics*, vol. 34, no. 4, pp. 297–306, 2009.
- [3] D. J. Gardiner, *Practical Raman Spectroscopy*, New York: Springer-Verlag, 1989.
- [4] H. Östmark, M. Nordberg, and T. E. Carlsson, "Stand-off detection of explosives particles by multispectral imaging raman spectroscopy," *Applied Optics*, vol. 50, no. 28, pp. 5592–5599, 2011.
- [5] E. Ceco, M. Nordberg, A. Ehlerding, and H. Östmark, "The detection limit of imaging Raman spectroscopy for 2,4,6-TNT, 2,4-DNT and RDX," in *Proc. SPIE 8546*, 2012.
- [6] N.R. Butt, M. Nilsson, A. Jakobsson, M. Nordberg, A. Pettersson, S. Wallin, and H. Östmark, "Classification of raman spectra to detect hidden explosives," *Geoscience and Remote Sensing Letters, IEEE*, vol. 8, no. 3, pp. 517–521, may 2011.
- [7] O. Friman, M. Borga, P. Lundberg, and H. Knutsson, "Adaptive analysis of fMRI data," *NeuroImage*, vol. 19, no. 3, pp. 837–845, 2003.
- [8] C. Peña Ortega and M. Vélez-Reyes, "Comparison of basis-vector selection methods for structural modeling of hyperspectral imagery," in *Proc. SPIE 7457*, 2009.
- [9] G. Healey and D. Slater, "Models and methods for automated material identification in hyperspectral imagery acquired under unknown illumination and atmospheric conditions," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 37, no. 6, pp. 2706–2717, nov 1999.
- [10] L. L. Sharf, *Statistical Signal Processing, Detection, Estimation, and Time Series Analysis*, Reading MA: Addison-Wesley, 1991.