Detecting Terahertz Signatures Using Guided Under-determined Source Signal Separation

Luis A. Rivera, *Graduate Student Member, IEEE,* and Guilherme. N. DeSouza, *Senior Member, IEEE*
ViGIR - Vision-Guided and Intelligent Robotics Lab, University of Missouri, Columbia, MO, 65211

***Abstract*—Terahertz technology has seen many advances over the past decades. Special interest for the use of this technology is directed to the detection of illegal drugs, explosives and other hazardous materials. These materials exhibit characteristic signatures at terahertz wavelengths which may be used to identify them. In this paper we present a method for detecting the presence of chemicals and other materials using their terahertz signatures and a technique for signature recognition called Guided Under-determined Source Signal Separation (GUSSS). The method was tested using a public THz database, achieving high true positive and true negative percentages.**

I. Introduction

TERAHERTZ technology (THz) has been greatly developed over the past decades. The terahertz radiation has the ability to penetrate many common barrier materials enabling hidden objects to be seen; adequate spatial resolution for imaging or localization of threat objects due to short wavelengths; non-ionising properties; safe to use on people at modest intensities; etc. Many objects and chemicals including explosives, illegal drugs, proteins and other biological materials have characteristic spectroscopic signatures at THz wavelengths which can be used to identify them [1], [2].

In this paper we propose a framework for detecting chemical compounds using their THz signatures. Our framework relies on a technique called "Guided Under-determined Source Signal Separation" (GUSSS) [3] to separate mixed signals and determine whether or not a particular chemical is present in a mixture. This technique is derived from Independent Component Analysis (ICA), and when combined with common classification methods, it obtains high levels of accuracy [3]. Our approach aims at detecting materials even if they are mixed together with other materials. Such a framework can be used for detection of Improvised Explosive Devices (IED), for detection of hazardous or illegal substances in airports, etc.

II. Proposed Framework

Figure 1 shows a block diagram illustrating the various parts of the system and the flow of the process. We start with a collection of THz signatures, which in our case come from public databases. In our approach we use time domain signals, so we need to transform the THz signatures from the databases (IFFT block in Figure 1). Time signatures were used for training the system, for creating test signals, and for obtaining the feature characterizing sensed signals, the GUSSS ratio. The GUSSS ratio can be used as a criterion for determining whether a particular signature is present or not in the sensed signal [3]. It is calculated for training signals in the system, and it is used to classify the testing signals using a Support Vector Machine (SVM).

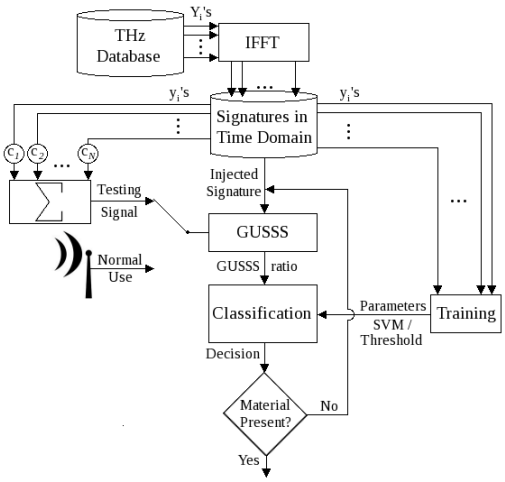

Figure 1. Proposed system.

A. Training the System

Let $Y_1, \ldots, Y_N$ be a set of THz signatures, and let one of them, say Y_p, be the signature corresponding to a particular material or chemical that we are interested in detecting in a sensed signal x. Let $y_1, \ldots, y_p, \ldots, y_N$ be the corresponding time domain signatures. These signatures are inputs for the training module shown in Figure 1. The output of this module are classification parameters later necessary for testing.

The training module uses the time signatures to create training mixed signals as the following linear combinations:

$$x = c_1 y_1 + c_2 y_2 + \cdots + c_p y_p + \cdots + c_N y_N$$

The mixing coefficients c_i represent the intensities that each individual signature y_i has in the mixed signals. In other words, the training signals simulate sensed signals we would get from a real THz sensor. So, we must create two classes of training signals. The first class corresponds to signals containing the particular signature y_p, i.e. class "present". To ensure that signature y_p is present in the mixture, the corresponding coefficient c_p has to be non-zero. So we employed a lower limit (ll) and an upper limit (ul) that satisfy $0 < ll \leq ul \leq 1$. Then, the coefficient c_p is randomly chosen from a uniform distribution defined over the interval $[ll, ul]$. The remaining coefficients are randomly chosen from a uniform distribution over the unit interval $[0, 1]$ and then normalized so that $\sum_{i \neq p} c_i = 1 - c_p$. Note that ll can be interpreted as a minimum % of signature y_p present in the mixtures.

The second class of training signals corresponds to "not present", that is, signals not containing the particular signature y_p. Therefore, the corresponding coefficient c_p is set to zero. The other coefficients are once again randomly chosen from the unit interval and normalized as before.

Next we apply the GUSSS algorithm described in [3] to all training signals. We obtain GUSSS ratios labeled "present" and "not present". Finally, we obtain the training parameters – the support vectors – that will be used later to classify the test signal mixtures.

B. Testing the System (for Classification)

The goal of the system is to detect the presence of specific materials among sensed mixtures. So, after training the system for the particular signatures that we want to investigate, we use those same signatures for testing.

Next, given a test signal mixture, we want to determine if a certain material (i.e. a certain signature) is present in the mixture or not. As before, we assume that the test signal mixture is a linear combination of arbitrary time signatures. As Figure 1 shows, one of the inputs of the GUSSS module is the test signal mixture. The other input of this module is the signature of the material to be investigated. In other words, if we want to determine if the particular material P_1 with signature y_{p_1} is present in the signal mixture, then we would inject a copy of y_{p_1} in the GUSSS module.

The GUSSS ratio obtained from the test signal is then input to an SVM classification function along with the learned SVM parameters. This function determines whether the given ratio (and therefore the corresponding signature) belongs to class "present" or class "not present". Despite the decision reached by the system for material P_1, different materials P_i can be sequentially tested for their presence or absence in the mixture.

III. Experiments and Results

The experiments that we conducted rely on THz signatures obtained from two public databases [4], [5]. We ran 18 tests, representing different combinations of materials from the databases. For each of the 18 tests, we selected one material to be the target material – i.e. the one to be detected. We used its signature and the signatures of 7 additional materials to train the system and obtain the training parameters. Then, we created testing mixtures using the same 8 THz signatures plus 2 additional ones. We introduce those two new THz signatures to simulate situations when the target material is mixed with materials for which no training had been provided – i.e. these 2 extra materials had never been seen by the system before the tests.

After selecting the materials for each of the 18 tests, we created 9 different sub-cases representing different potential concentrations of the target material in the total mixture. We did that by randomly selecting the concentration of that target material between a certain lower limit, ll, and an upper limit, ul. Finally, for each of the 9 sub-cases in each test, we created 500 mixture signals for training and another 1000 for testing. Half of the mixtures corresponded to cases when the target material is present, and the other half to cases when the material is not present. The mixtures were created as linear combinations of all 10 THz signatures in the test.

Figure 2 illustrates the behavior of the GUSSS ratios as the lower limit ll varies from 0.1 to 0.9. The figure depicts in red and blue the distributions of the GUSSS ratios corresponding respectively to the cases when the particular material is present and when it is not present (log-scale). The left-hand figure shows the best material – i.e. the material for which the classification rate was the highest. The right-hand figure shows a more typical material. From the figures, we can notice that the GUSSS ratios for the cases when the material's signature is not present remain approximately within the same range, independent of ll – i.e. the averages of the logarithm of those GUSSS ratios are approximately constant throughout the plot. However, for the cases when the signature is present in the mixtures, the GUSSS ratios are smaller, and their values decrease further with the increase in the value of ll – i.e. with increases in the concentration of the particular material in the mixture. Note how the separation between the set of "not present" cases and the set of "present" cases increases as ll increases, as expected.

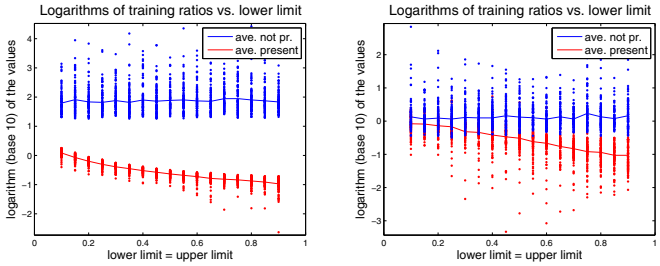

Figure 2. Logarithm of GUSSS ratios vs. lower limit (ll) of a particular target material concentration %. Red: ratios when the target material is present; Blue, when not present. The continuous lines go through the averages of the log-GUSSS ratios. Left: Best case. Right: Average case.

Table I
Overall average percentages of True Positives, True Negatives and Correct Classification rates (18 experiments).

$[ll,\ ul]$	TP	TN	CC
[0.10, 0.20]	67.1	71.9	69.5
[0.20, 0.30]	77.1	79.3	78.2
[0.30, 0.40]	85.3	85.1	85.2
[0.40, 0.50]	88.9	92.0	90.5
[0.50, 0.60]	93.5	94.4	93.9
[0.60, 0.70]	95.1	96.8	96.0
[0.70, 0.80]	96.4	97.8	97.1
[0.80, 0.90]	97.4	98.0	97.7
[0.90, 0.95]	97.5	98.3	97.9

The complete set of tests and more details about the materials used can be found at http://vigir.missouri.edu/THz_GUSSS.htm

IV. Conclusion

This paper introduced a framework for detecting the presence of particular chemicals or materials in a mixture. We used THz signatures of various materials obtained from public databases for our simulations. Our method presented very good results for detecting a variety of such materials under various assumptions on the percentage of the target material in the scene.

References

[1] I. Hosako, N. Sekine, M. Patrashin, S. Saito, K. Fukunaga, Y. Kasai, P. Baron, T. Seta, J. Mendrok, S. Ochiai, and H. Yasuda, "At the dawn of a new era in terahertz technology," *Proceedings of the IEEE*, vol. 95, no. 8, pp. 1611–1623, Aug. 2007.

[2] M. C. Kemp, "Detecting hidden objects: Security imaging using millimetre-waves and terahertz," in *IEEE Conference on Advanced Video and Signal Based Surveillance*, Sept. 2007, pp. 7–9.

[3] L. A. Rivera and G. N. DeSouza, "A power wheelchair controlled using hand gestures, a single sEMG sensor, and guided under-determined source signal separation," in *4th IEEE RAS & EMBS International Conference on Biomedical Robotics*, Jun. 2012, pp. 1535–1540, BioRob2012, Rome, Italy.

[4] "Terahertz database." [Online]. Available: http://thzdb.org

[5] "Thz database web," tera-phtonics Laboratory, RIKEN Sendai. [Online]. Available: http://www.riken.jp/THzdatabase/