Accepted Manuscript

Title: Classification of THz pulse signals using two-dimensional cross-correlation feature extraction and non-linear classifiers

Author: Siuly Xiaoxia Yin Sillas Hadjiloucas Yanchun Zhang

PII: S0169-2607(15)30121-8
DOI: http://dx.doi.org/doi:10.1016/j.cmpb.2016.01.017
Reference: COMM 4070

To appear in: *Computer Methods and Programs in Biomedicine*

Received date: 19-8-2015
Revised date: 20-1-2016
Accepted date: 21-1-2016

Please cite this article as: Siuly, X. Yin, S. Hadjiloucas, Y. Zhang, Classification of THz pulse signals using two-dimensional cross-correlation feature extraction and non-linear classifiers, *Computer Methods and Programs in Biomedicine* (2016), http://dx.doi.org/10.1016/j.cmpb.2016.01.017

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

ACCEPTED MANUSCRIPT

Classification of THz pulse signals using two-dimensional cross-correlation feature extraction and non-linear classifiers

Siuly[1], Xiaoxia Yin[1], Sillas Hadjiloucas[2] and Yanchun Zhang[1]

[1]Centre for Applied Informatics, College of Engineering & Science

Victoria University, Melbourne, Australia

[2]School of Systems Engineering, University of Reading, Reading RG6 6AY, UK

siuly.siuly@vu.edu.au; xiaoxia.yin@vu.edu.au; s.hadjiloucas@reading.ac.uk; yanchun.zhang@vu.edu.au

Abstract.

This work provides a performance comparison of four different machine learning classifiers: multinomial logistic regression with ridge estimators (MLR) classifier, k-nearest neighbours (KNN), support vector machine (SVM) and naïve Bayes (NB) as applied to terahertz (THz) transient time domain sequences associated with pixelated images of different powder samples. The six substances considered, although have similar optical properties, their complex insertion loss at the THz part of the spectrum is significantly different because of differences in both their frequency dependent THz extinction coefficient as well as differences in their refractive index and scattering properties. As scattering can be unquantifiable in many spectroscopic experiments, classification solely on differences in complex insertion loss can be inconclusive. The problem is addressed using two-dimensional (2-D) cross-correlations between background and sample interferograms, these ensure good noise suppression of the datasets and provide a range of statistical features that are subsequently used as inputs to the above classifiers. A cross-validation procedure is adopted to assess the performance of the classifiers. Firstly the measurements related to samples that had thicknesses of 2mm were classified, then samples at thicknesses of 4 mm, and after that 3 mm were classified and the success rate and consistency of each classifier was recorded. In addition, mixtures having thicknesses of 2 and 4 mm as well as mixtures of 2, 3 and 4 mm were presented simultaneously to all classifiers. This approach provided further cross-validation of the classification consistency of each algorithm. The results confirm the superiority in classification accuracy and robustness of the MLR (least accuracy 88.24%) and KNN (least accuracy 90.19%) algorithms which consistently outperformed the SVM (least accuracy 74.51%) and NB (least accuracy 56.86%) classifiers for the same number of feature vectors across all studies. The work establishes a general methodology for assessing the performance of other hyperspectral dataset classifiers on the basis of 2-D cross-correlations in far-infrared spectroscopy or other parts of the electromagnetic spectrum. It also advances the wider proliferation of automated THz imaging systems across new application areas e.g., biomedical imaging, industrial processing and quality control where interpretation of hyperspectral images is still under development.

Key-words: Terahertz spectroscopy; 2-D cross-correlation; Multinomial logistic regression classifier; K-nearest neighbours; Support vector machine; Naïve Bayes.

1. Introduction

Accepted Manuscript

Download English Version:

https://daneshyari.com/en/article/6891431

Download Persian Version:

https://daneshyari.com/article/6891431

Daneshyari.com