Package ‘sabarsi’

August 8, 2019

Type Package

Title Background Removal and Spectrum Identification for SERS Data

Version 0.1.0

Description Implements a new approach 'SABARSI' described in Wang et al., ``A Statistical Approach of Background Removal and Spectrum Identification for SERS Data'' (Unpublished). Sabarsi forms a pipeline for SERS (surface-enhanced Raman scattering) data analysis including background removal, signal detection, signal integration, and cross-experiment comparison. The background removal algorithm, the very first step of SERS data analysis, takes into account the change of background shape.

Depends R (>= 3.5.0)

Suggests knitr, rmarkdown (>= 1.13)

Imports stats (>= 3.5.0)

License GPL-3

Encoding UTF-8

LazyData true

VignetteBuilder knitr

RoxygenNote 6.1.1

NeedsCompilation no

Author Li Jun [cre],
Wang Chuanqi [aut]

Maintainer Li Jun `<jun.li@nd.edu>`

Repository CRAN

Date/Publication 2019-08-08 12:30:02 UTC

R topics documented:

---

background_removal *Perform background removal on the whole SERS spectrum data set. Divide the SERS spectrum data into time-frequency blocks and remove background locally.*

Description

Perform background removal on the whole SERS spectrum data set. Divide the SERS spectrum data into time-frequency blocks and remove background locally.

Usage

```
background_removal(x, q = 0.4, w.chan = 50, w.time = 50)
```

Arguments

`x`	A p*n data matrix. There are n SERS spectra with p frequency channels.
`q`	A number taking value between 0 and 1. 100*q is the quantile that SABARSI uses to calculate the spectrum strength.
`w.chan`	The window size for the frequency domain. The default value of `w.chan` is 50.
`w.time`	The window size for the time domain. The default value of `w.time` is 50.

Value

A p*n data matrix, `xr`, of background-removed spectra.

Examples

```
data(SERS)
x <- SERS$R1   ## x is the matrix of SERS spectra
xr <- background_removal(x) ## xr is the matrix of background removed spectra
```

---

detect_sig *Calculate the pvalues and false discovery rates (FDRs) for a background-removed spectrum*

Description

Calculate the pvalues and false discovery rates (FDRs) for a background-removed spectrum

Usage

```
detect_sig(xr)
```

Arguments

`xr`	A p-length vector that represents a background-removed spectrum with p frequency channels.

Value

A list containing a vector of pvalues, `pvals`, and a vector of FDRs, `fdrs`.

`merge_signals` *Obtain the set of signature signals Merge groups of concatenated signals and give the time indices of signature signals.*

Description

Obtain the set of signature signals Merge groups of concatenated signals and give the time indices of signature signals.

Usage

```
merge_signals(xr, object, t.tol = 4, cor.tol = 0.5)
```

Arguments

`xr`	A p-length vector that represents a background-removed spectrum with p frequency channels.
`object`	An object obtained from `signal_detection`
`t.tol`	A positive integer, which is the tolerance of time difference when comparing two signals.
`cor.tol`	A number between 0 and 1. Two signals is considered to be similar enough if their correlation is higher than `cor.tol`.

Value

A vector recording the time indices of signature signals.

Examples

```
## xr is the matrix of background removed spectra.
res <- signal_detection(xr) ## detect the signals in xr
tim.index.ss <- merge_signals(xr = xr, object = res)  ## the set of signature signals.
```

`SERS` *A real SERS spectrum data set in*

Description

This data contains the SERS spectra from two technical replicates.

Usage

```
data(SERS)
```

Format

An object containing the following variables:

- `R1` A data matrix of SERS spectra in replicate 1.
- `R2` A data matrix of SERS spectra in replicate 2.

Details

This data contains the SERS spectra from two technical replicates. Each replicate contains spectra with 500 frequency channels from 500 time points.

References

Wang et. al "A Statistical Approach of Background Removal and Spectrum Identification for SERS Data"

Examples

```
data(SERS)
x <- list()
x[[1]] <- SERS$R1
x[[2]] <- SERS$R2
```

`shift_match` *Match signals from two experiments. For each signal in the first experiment,* `shift.match` *function finds the best matched signal in the second experiment. This function takes the potential frequency shifts into consideration for similarity measurement.*

Description

Match signals from two experiments. For each signal in the first experiment, `shift.match` function finds the best matched signal in the second experiment. This function takes the potential frequency shifts into consideration for similarity measurement.

Usage

```
shift_match(xra, xrb, ta, tb)
```

Arguments

`xra`	A p*n data matrix of background-removed spectra in the first experiment.
`xrb`	A p*n data matrix of background-removed spectra in the second experiment.
`ta`	A vector of time indices of signals in the first experiment.
`tb`	A vector of time indices of signals in the second experiment.

Value

A list containing the time indices of signals in the first experiment, `ta`, and the time indices of corresponding time indices in the second experiment, as well as the correlation of each match pairs, `corra`.

`signal_detection` *Detect signals in background-removed spectra*

Description

Detect signals in background-removed spectra

Usage

```
signal_detection(xr, fdr.c = 0.01, stren.c = 200, wid.c = 5)
```

Arguments

`xr`	A p*n data matrix of background-removed spectra, where n is the number of spectra, and p is the number of frequency channels.
`fdr.c`	A number between 0 and 1, which is the cutoff for FDR.
`stren.c`	A positive number for the cutoff of the strength of a signal.
`wid.c`	A positive number for the cutoff of the width of a signal.

Value

A list containing the indices of spectra with detected signals, `tim.index`, and a matrix recording the peaks of signals, `peaks`.

Examples

```
res <- signal_detection(xr)
# xr is the matrix of background removed spectra
# xr can be obtained by background_removal function

head(res$tim.index)  ## check the first few time indices of signals
```

Index