Package 'peakPick'

December 4, 2015

Type Package
Title Peak Picking Methods Inspired by Biological Data
Version 0.11
Date 2015-12-04
Author Paal O. Westermark
Maintainer Paal O. Westermark <pal-olof.westermark@charite.de></pal-olof.westermark@charite.de>
Description Biologically inspired methods for detecting peaks in one-dimensional data, such as time series or genomics data. The algorithms were originally designed by Weber, Ramachandran, and Henikoff, see documentation.
License MIT + file LICENSE
Imports matrixStats, parallel
LazyData TRUE
Suggests knitr, rmarkdown
VignetteBuilder knitr
RoxygenNote 5.0.1
NeedsCompilation no
Repository CRAN
Date/Publication 2015-12-04 15:40:34
R topics documented:
detect.spikes
Index

2 detect.spikes

detect.spikes	Detects spikes in series of numbers.

Description

This algorithm detects spikes rising above a user-specified number of standard deviations numbers in a certain window. Use this algorithm to detect short spikes rather than smooth bumps in series of numbers. Please refer to the paper by Weber et al. for more details.

Usage

```
detect.spikes(mat, roi, winlen, spike.min.sd = 3, mc.cores = 1,
  verbose = FALSE)
```

Arguments

mat	matrix of series with series organized columnwise. The algorithm treats each column separately.
roi	vector of two integers (c(min, max)) defining positions in all series (rows in mat) to consider for spike detection, used together with winlen. Must lie within the interval [2, nrow(mat) - 1]. Will be coerced to integers.
winlen	integer defining the window of positions to consider for mean and sem estimation for each series. Each estimation limits itself to the position and a plus/minus winlen positions large window. Thus, winlen must not be chosen larger than that the windows fit within mat, given the roi. I.e. roi[1] - winlen >=1 AND roi[length(roi)] + winlen <= nrow(mat). Will be coerced to an integer.
spike.min.sd	numeric minimum number of standard deviations for a spike to rise above the mean in order to be considered for a spike call and to be excluded from the mean estimation of each subsequent iteration of the spike calling algorithm
mc.cores	the number of cores do perform this calculation
verbose	Boolean indicating the number of new peaks detected with each iteration. The algorithm stops as soon as this number does not sink anymore. Turn this on if running into problems.

Value

boolean matrix corresponding to mat, representing spike positions.

References

Weber, C.M., Ramachandran, S., and Henikoff, S. (2014). Nucleosomes are context-specific, H2A.Z-modulated barriers to RNA polymerase. Molecular Cell 53, 819-830.

peakpick 3

peakpick	Identifies smooth peaks in series of numbers.

Description

This algorithm detects peaks, smooth bumps in series of numbers. This algorithm should not be used for series containing brief spikes. Consider filtering/smoothing your data before using this algorithm. Please refer to the paper by Weber et al. for more details.

Usage

```
peakpick(mat, neighlim, deriv.lim = 0.04, peak.min.sd = 0.5,
   peak.npos = 10L, mc.cores = 1)
```

Arguments

mat	matrix of series with series organized columnwise
neighlim	integer limit for how far apart peaks must be. Peak pairs closer than or equal to neighlim to each other have the lesser peak eliminated.
deriv.lim	numeric upper limit for the estimatied derivative for a point to be considered for a peak call
peak.min.sd	numeric minimum number of standard deviations for a peak to rise above the mean of its immediate vicinity in order to be considered for a peak call
peak.npos	integer peak standard deviations and means will be estimated plus/minus npos positions from peak
mc.cores	the number of cores to perform this computation

Value

boolean matrix with dimensions of mat representing peaks

References

Weber, C.M., Ramachandran, S., and Henikoff, S. (2014). Nucleosomes are context-specific, H2A.Z-modulated barriers to RNA polymerase. Molecular Cell 53, 819-830.

Index

detect.spikes, 2
peakpick, 3