Package ‘autothresholdr’

December 13, 2023

Type Package

Title An R Port of the 'ImageJ' Plugin 'Auto Threshold'

Version 1.4.2

Maintainer Rory Nolan <rorynoolan@gmail.com>

Description Algorithms for automatically finding appropriate thresholds for numerical data, with special functions for thresholding images. Provides the 'ImageJ' 'Auto Threshold' plugin functionality to R users. See <https://imagej.net/plugins/auto-threshold> and Landini et al. (2017) <DOI:10.1111/jmi.12474>.

License GPL-3


BugReports https://github.com/rorynolan/autothresholdr/issues

Depends R (>= 3.5)

Imports checkmate (>= 1.9.3), ijtiff (>= 2.2), magrittr (>= 1.5), purrr, Rcpp (>= 1.0.11), rlang (>= 0.3.3), stats, strex (>= 1.4.1), stringr (>= 1.4)

Suggests covr, dplyr, ggplot2, knitr, rmarkdown, spelling, styler (>= 1.3.2), testthat (>= 3.0), utils

LinkingTo Rcpp (>= 1.0.11)

VignetteBuilder knitr

Config/testthat/edition 3

Encoding UTF-8

Language en-US

RoxygenNote 7.2.3

NeedsCompilation yes

Author Rory Nolan [aut, cre, trl] (<https://orcid.org/0000-0002-5239-4043>), Luis Alvarez [ctb] (<https://orcid.org/0000-0003-1316-1906>), Sergi Padilla-PARRA [ctb, ths] (<https://orcid.org/0000-0002-8010-9481>), Gabriel Landini [ctb, cph] (<https://orcid.org/0000-0002-9689-0989>)
Description

Usage

```r
auto_thresh(
  int_arr,
  method,
  ignore_black = FALSE,
  ignore_white = FALSE,
  ignore_na = FALSE
)
```

```r
auto_thresh_mask(
  int_arr,
  method,
  ignore_black = FALSE,
  ignore_white = FALSE,
  ignore_na = FALSE
)
```

```r
auto_thresh_apply_mask(
  int_arr,
  method,
  fail = NA,
)
ignore_black = FALSE,
ignore_white = FALSE,
ignore_na = FALSE
)

mask(
  int_arr,
  method,
  ignore_black = FALSE,
  ignore_white = FALSE,
  ignore_na = FALSE
)

apply_mask(
  int_arr,
  method,
  fail = NA,
  ignore_black = FALSE,
  ignore_white = FALSE,
  ignore_na = FALSE
)

Arguments

int_arr An array (or vector) of non-negative integers.

method The name of the thresholding method you wish to use. The available methods are "IJDefault", "Huang", "Huang2", "Intermodes", "IsoData", "Li", "MaxEntropy", "Mean", "MinErrorI", "Minimum", "Moments", "Otsu", "Percentile", "RenyiEntropy", "Shanbhag", "Triangle" and "Yen". Partial matching is performed i.e. method = "h" is enough to get you "Huang" and method = "in" is enough to get you "Intermodes". To perform manual thresholding (where you set the threshold yourself), supply the threshold here as a number e.g. method = 3; so note that this would not select the third method in the above list of methods.

ignore_black Ignore black pixels/elements (zeros) when performing the thresholding?

ignore_white Ignore white pixels when performing the thresholding? If set to TRUE, the function makes a good guess as to what the white (saturated) value would be (see 'Details'). If this is set to a number, all pixels with value greater than or equal to that number are ignored.

ignore_na This should be TRUE if NAs in int_arr should be ignored or FALSE if you want the presence of NAs in int_arr to throw an error.

fail When using auto_thresh_apply_mask(), to what value do you wish to set the pixels which fail to exceed the threshold? fail = 'saturate' sets them to saturated value (see "Details"). fail = 'zero' sets them to zero. You can also specify directly here a natural number (must be between 0 and 2^16 - 1) to use.
Details

- Values greater than or equal to the found threshold pass the thresholding and values less than the threshold fail the thresholding.
- For ignore_white = TRUE, if the maximum value in the array is one of $2^{8}-1$, $2^{12}-1$, $2^{16}-1$ or $2^{32}-1$, then those max values are ignored. That’s because they’re the white values in 8, 12, 16 and 32-bit images respectively (and these are the common image bit sizes to work with). This guesswork has to be done because R does not know how many bits the image was on disk. This guess is very unlikely to be wrong, and if it is, the consequences are negligible anyway. If you’re very concerned, then just specify the white value as an integer in this ignore_white argument.
- If you have set ignore_black = TRUE and/or ignore_white = TRUE but you are still getting error/warning messages telling you to try them, then your chosen method is not working for the given array, so you should try a different method.
- For a given array, if all values are less than $2^{8}$, saturated value is $2^{8} - 1$, otherwise, if all values are less than $2^{16}$, the saturated value is $2^{16} - 1$, otherwise the saturated value is $2^{32}-1$.
- For the auto_thresh() function, if you pass int_arr as a data frame with column names value and n, that’s the same as passing an integer array having n entries of each value. For this form of int_arr, ignore_white and ignore_black are irrelevant.

Value

auto_thresh() returns an object of class th containing the threshold value. Pixels exceeding this threshold pass the thresholding, pixels at or below this level fail.

auto_thresh_mask() returns an object of class masked_arr which is a binarized version of the input, with a value of TRUE at points which exceed the threshold and FALSE at those which do not.

auto_thresh_apply_mask() returns and object of class threshed_arr which is the original input masked by the threshold, i.e. all points not exceeding the threshold are set to a user-defined value (default NA).

mask() is the same as auto_thresh_mask() and apply_mask() is the same as auto_thresh_apply_mask().

Acknowledgements

Gabriel Landini coded all of these functions in Java. These java functions were then translated to C++.

References


Zack GW, Rogers WE, Latt SA (1977), "Automatic measurement of sister chromatid exchange frequency", J. Histochem. Cytochem. 25 (7): 74153, PMID 70454


Examples

```r
img_location <- system.file("extdata", "eg.tif", package = "autothresholdr")
img <- ijtiff::read_tif(img_location)
auto_thresh(img, "huang")
img_value_count <- magrittr::set_names(as.data.frame(table(img)),
c("value", "n"))
print(head(img_value_count))
auto_thresh(img_value_count, "Huang")
auto_thresh(img, "tri")
auto_thresh(img, "Otsu")
auto_thresh(img, 9)
mask <- auto_thresh_mask(img, "huang")
```
mean_stack_thresh

\[ \text{ijtiff::display(mask[, , 1, 1])} \]
\[ \text{masked <- auto_thresh_apply_mask(img, "huang")} \]
\[ \text{ijtiff::display(masked[, , 1, 1])} \]
\[ \text{masked <- auto_thresh_apply_mask(img, 25)} \]
\[ \text{ijtiff::display(masked[, , 1, 1])} \]

---

masked_arr

*Masked array class.*

Description

A *mask* of an array with respect to a given threshold is found by taking the original array and setting all elements falling below the threshold to FALSE and the others to TRUE. An object of class masked_arr has the attribute thresh detailing the threshold value that was applied.

Usage

masked_arr(arr, thresh)

Arguments

*arr*  
An array of logicals (the mask).

*thresh*  
The threshold. Either a scalar or an object of class th.

Value

An object of class masked_arr.

---

mean_stack_thresh

*Threshold every image frame in an image stack based on their mean.*

Description

An *ijtiff_img* is a 4-dimensional array indexed by img[y, x, channel, frame]. For each channel (which consists of a stack of frames), this function finds a threshold based on the sum all of the frames, uses this to create a mask and then applies this mask to every frame in the stack (so for a given pillar in the image stack, either all the pixels therein are thresholded away or all are untouched, where pillar x,y of channel ch is img[y, x, ch, ]).
**Usage**

```r
mean_stack_thresh(
  img,            
  method,         
  fail = NA,      
  ignore_black = FALSE, 
  ignore_white = FALSE, 
  ignore_na = FALSE
)
```

**Arguments**

- `img` A 4-dimensional array in the style of an `ijiff_img` (indexed by `img[y, x, channel, frame]`) or a 3-dimensional array which is a single channel of an `ijiff_img` (indexed by `img[y, x, frame]`).
- `method` The name of the thresholding method you wish to use. The available methods are "IJDefault", "Huang", "Huang2", "Intermodes", "IsoData", "Li", "MaxEntropy", "Mean", "MinErrorI", "Minimum", "Moments", "Otsu", "Percentile", "RenyiEntropy", "Shanbhag", "Triangle" and "Yen". Partial matching is performed i.e. `method = "h"` is enough to get you "Huang" and `method = "in"` is enough to get you "Intermodes". To perform manual thresholding (where you set the threshold yourself), supply the threshold here as a number e.g. `method = 3.8` (so note that this would not select the third method in the above list of methods). This manual threshold will then be used to threshold the sum stack to create a 2D mask and then this mask will be applied to all frames in the stack. If you want a different method for each channel, specify this parameter as a vector or list, one element per channel.
- `fail` When using `auto_thresh_apply_mask()`, to what value do you wish to set the pixels which fail to exceed the threshold? `fail = 'saturate'` sets them to saturated value (see 'Details'). `fail = 'zero'` sets them to zero. You can also specify directly here a natural number (must be between 0 and `2^16 - 1`) to use.
- `ignore_black` Ignore black pixels/elements (zeros) when performing the thresholding?
- `ignore_white` Ignore white pixels when performing the thresholding? If set to TRUE, the function makes a good guess as to what the white (saturated) value would be (see 'Details'). If this is set to a number, all pixels with value greater than or equal to that number are ignored.
- `ignore_na` This should be TRUE if NAs in `int_arr` should be ignored or FALSE if you want the presence of NAs in `int_arr` to throw an error.

**Details**

It’s called `mean_stack_thresh()` and not `sum_stack_thresh()` because its easier for people to visualize the mean of an image series than to visualize the sum, but for the sake of this procedure, both are equivalent, except for the fact that the thresholding routine invoked inside this function prefers integers, which we get by using a sum but not by using a mean.

- Values greater than or equal to the found threshold **pass** the thresholding and values less than the threshold **fail** the thresholding.
• For `ignore_white = TRUE`, if the maximum value in the array is one of $2^8 - 1$, $2^{16} - 1$ or $2^{32} - 1$, then those max values are ignored. That’s because they’re the white values in 8, 16 and 32-bit images respectively (and these are the common image bit sizes to work with). This guesswork has to be done because R does not know how many bits the image was on disk. This guess is very unlikely to be wrong, and if it is, the consequences are negligible anyway. If you’re very concerned, then just specify the white value as an integer in this `ignore_white` argument.

• If you have set `ignore_black = TRUE` and/or `ignore_white = TRUE` but you are still getting error/warning messages telling you to try them, then your chosen method is not working for the given array, so you should try a different method.

• For a given array, if all values are less than $2^8$, saturated value is $2^8 - 1$, otherwise, saturated value is $2^{16} - 1$.

Value

An object of class `stack_threshed_img` which is the thresholded image (an array in the style of an `ijtiff_img`). Pillars not exceeding the threshold are set to the `fail` value (default NA).

Examples

```r
img <- ijtiff::read_tif(system.file("extdata", "50.tif", 
   package = "autothresholdr") )
ijtiff::display(img[, , 1, 1])
img_thresh_mask <- mean_stack_thresh(img, "Otsu")
ijtiff::display(img_thresh_mask[, , 1, 1])
ijtiff::display(img[, , 1, 1])
img_thresh_mask <- mean_stack_thresh(img, "Huang")
ijtiff::display(img_thresh_mask[, , 1, 1])
```

---

`med_stack_thresh`  
Threshold every image frame in a stack based on their median.

Description

An `ijtiff_img` is a 4-dimensional array indexed by `img[y, x, channel, frame]`. For each channel (which consists of a stack of frames), this function finds a threshold based on all of the frames, then takes the median of all the frames in the stack image, uses this to create a mask with the found threshold and then applies this mask to every frame in the stack (so for a given pillar in the image stack, either all the pixels therein are thresholded away or all are untouched, where pillar `x, y` of channel `ch` is `img[y, x, ch, ]`).
med_stack_thresh

Usage

med_stack_thresh(
  img,
  method,
  fail = NA,
  ignore_black = FALSE,
  ignore_white = FALSE,
  ignore_na = FALSE
)

Arguments

img
  A 3-dimensional array (the image stack, possibly a time-series of images) where
  the \( n \)th slice is the \( n \)th image in the stack.

method
  The name of the thresholding method you wish to use. The available meth-
  ods are "IJDefault", "Huang", "Huang2", "Intermodes", "IsoData", "Li",
  "RenyiEntropy", "Shanbhag", "Triangle" and "Yen". Partial matching is
  performed i.e. method = "h" is enough to get you "Huang" and method = "in"
  is enough to get you "Intermodes". To perform manual thresholding (where
  you set the threshold yourself), supply the threshold here as a number e.g.
  method = 3 (so note that this would not select the third method in the above list
  of methods). This manual threshold will then be used to threshold the median
  stack to create a 2D mask and then this mask will be applied to all frames in the
  stack. If you want a different method for each channel, specify this parameter
  as a vector or list, one element per channel.

fail
  When using auto_thresh_apply_mask(), to what value do you wish to set
  the pixels which fail to exceed the threshold? fail = 'saturate' sets them to
  saturated value (see 'Details'). fail = 'zero' sets them to zero. You can also
  specify directly here a natural number (must be between \( 0 \) and \( 2^{32} - 1 \)) to use.

ignore_black
  Ignore black pixels/elements (zeros) when performing the thresholding?

ignore_white
  Ignore white pixels when performing the thresholding? If set to TRUE, the func-
  tion makes a good guess as to what the white (saturated) value would be (see
  'Details').

ignore_na
  This should be TRUE if NAs in int_arr should be ignored or FALSE if you want
  the presence of NAs in int_arr to throw an error.

Details

- Values greater than or equal to the found threshold pass the thresholding and values less than
  the threshold fail the thresholding.
- For ignore_white = TRUE, if the maximum value in the array is one of \( 2^8-1, 2^16-1 \) or \( 2^32-1 \), then those max values are ignored. That’s because they’re the white values in 8, 16
  and 32-bit images respectively (and these are the common image bit sizes to work with). This
  guesswork has to be done because R does not know how many bits the image was on disk.
  This guess is very unlikely to be wrong, and if it is, the consequences are negligible anyway.
If you’re very concerned, then just specify the white value as an integer in this `ignore_white` argument.

- If you have set `ignore_black = TRUE` and/or `ignore_white = TRUE` but you are still getting error/warning messages telling you to try them, then your chosen method is not working for the given array, so you should try a different method.
- For a given array, if all values are less than $2^8$, saturated value is $2^8 - 1$, otherwise, saturated value is $2^{16} - 1$.

Value

An object of class `stack_threshed_img` which is the thresholded image (an array in the style of an `ijtiff_img`). Pillars not exceeding the threshold are set to the `fail` value (default NA).

Examples

```r
img <- ijtiff::read_tif(system.file("extdata", "50.tif", package = "autothresholdr")
ijtiff::display(img[, 1, 1])
img_thresh_mask <- med_stack_thresh(img, "Otsu")
ijtiff::display(img_thresh_mask[, 1, 1])
ijtiff::display(img[, 1, 1])
img_thresh_mask <- med_stack_thresh(img, "Triangle")
ijtiff::display(img_thresh_mask[, 1, 1])
```

Description

A stack-thresholded array is an array which has had stack-thresholding applied to it. See `mean_stack_thresh()`.

It has 3 necessary attributes:

- `thresh` is the threshold that was applied. This is either a number or an object of class `th`. Values in the original array which were less than this value are deemed to have failed the thresholding.
- `fail_value` is the value to which elements of the array which failed the thresholding were set. This could be something like 0 or NA.
- `stack_thresh_method` details which stacked-thresholding method was employed; this is either "mean" or "median".

Usage

```r
stack_threshed_img(img, thresh, fail_value, stack_thresh_method)
```
Arguments

- **img**: A 4-dimensional array in the style of an `ijtiff_img` (indexed by `img[y, x, channel, frame]`) or a 3-dimensional array which is a single channel of an `ijtiff_img` (indexed by `img[y, x, frame]`).
- **thresh**: The threshold that was used. Either a number or an object of class `th`.
- **fail_value**: The value to which elements of the array which failed the thresholding were set.
- **stack_thresh_method**: This must be set to either "mean" or "median" to tell which stacked-thresholding method was employed.

Value

An object of class `stack_threshed_img`.

See Also

`th`, `mean_stack_thresh()`, `med_stack_thresh()`.

---

`th`  
*Automatically found threshold class.*

Description

A threshold found automatically via `auto_thresh()`. It is a number (the value of the threshold) with 4 attributes:

- `ignore_black` is TRUE if black values were ignored during the thresholding and FALSE otherwise.
- `ignore_white` is TRUE if white values were ignored during the thresholding and FALSE otherwise.
- `ignore_na` is TRUE if NAs were ignored during the thresholding and FALSE otherwise.
- `autothresh_method` details which automatic thresholding method was used.

Usage

`th(thresh, ignore_black, ignore_white, ignore_na, autothresh_method)`

Arguments

- **thresh**: A scalar. The threshold.
- **ignore_black**: TRUE if black values were ignored during the thresholding and FALSE otherwise.
- **ignore_white**: TRUE if white values were ignored during the thresholding and FALSE otherwise.
- **ignore_na**: TRUE if NA values were ignored during the thresholding and FALSE otherwise.
- **autothresh_method**: The name of the automatic thresholding method used.
A thresholded array is an array which has had a threshold applied to it. It has an attribute `thresh` which is the threshold that was applied which can be a number or an object of class `th`.

**Usage**

```r
threshed_arr(arr, thresh)
```

**Arguments**

- `arr` The thresholded array (not the original array).
- `thresh` The threshold that was used. Either a number or an object of class `th`.

**Details**

The term 'array' is used loosely here in that vectors and matrices qualify as arrays.

**Value**

An object of class `threshed_arr`.

**See Also**

`stack_threshed_img`, `apply_mask()`.
Index

apply_mask (auto_thresh), 2
apply_mask(), 12
auto_thresh, 2
auto_thresh(), 4, 11
auto_thresh_apply_mask (auto_thresh), 2
auto_thresh_mask (auto_thresh), 2

ijtiff_img, 6–8, 10, 11

mask (auto_thresh), 2
masked_arr, 4, 6, 6
mean_stack_thresh, 6
mean_stack_thresh(), 10, 11
med_stack_thresh, 8
med_stack_thresh(), 11

stack_threshed_img, 8, 10, 10, 12

th, 4, 6, 10, 11, 11, 12
threshed_arr, 4, 11, 12, 12