

Configurable point rasterization for large scatterplots



Tereza Kulichová | Supervisor: doc. RNDr. David Hoksza, Ph.D.

Thesis goals

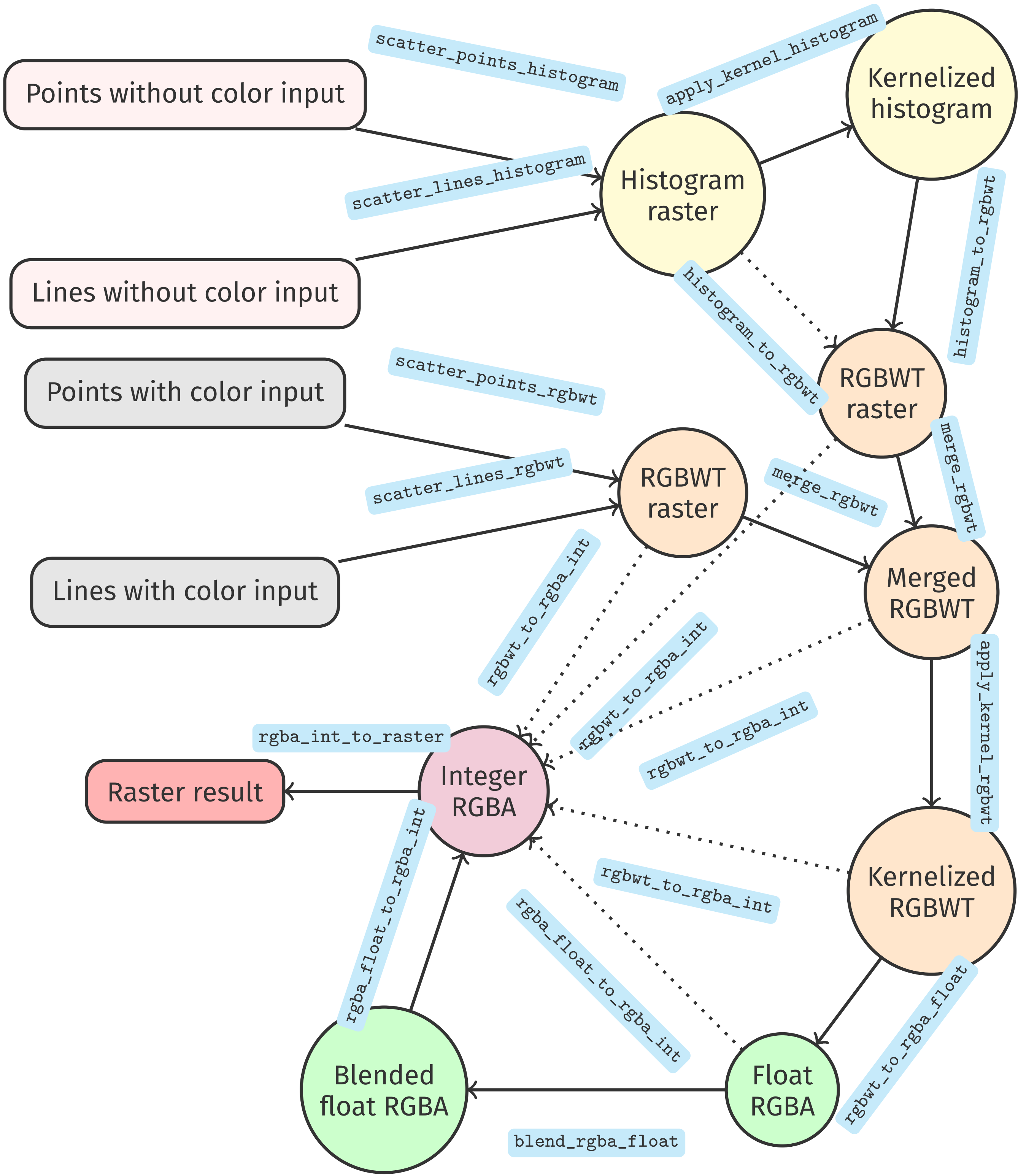
Scattermore is a simple R package used for scatterplot visualizations. The goal of the thesis is to expand `scattermore` to support new plotting functionality while retaining the performance.

- **High-customizable API.** We assign each subtask to a single method to ensure more variation in output for the user.
- **Reduction of overplotting.** We introduce the new RGBWT data format, which combines images equally.
- **Fast plotting.** We implement the computationally expensive operation in C language.

Approach

- Design a specific pixel format that supports order-independent alpha-blending.
- Try some acceleration possibilities (SIMD, multithreading).
- Package and document everything properly to make `scattermore` easier to use.

New API & functionality overview



Deployed user-friendly documentation

scattermore **1.1**

ReferenceArticles

scattermore

CRAN 1.0

downloads 30K/month

R-CMD-check passing

codecov 87%

Scatterplots with more datapoints. If you want to plot bazillions of points without much waiting, use this.

If you want to report the usage of `scattermore` within a scientific project, you may want to refer to it systematically – `scattermore` has been peer-reviewed as a part of a larger package for interactive cytometry data analysis (`ShinySOM`). Links here: [PubMed 32049322](#), [OUP](#), [doi:10.1093/bioinformatics/btaa091](#)

Installation

- from CRAN repositories (recommended): `install.packages('scattermore')`
- from GitHub (development version): `devtools::install_github('exaexa/scattermore')`

Links

[View on CRAN](#)

License

GPL (>= 3)

Citation

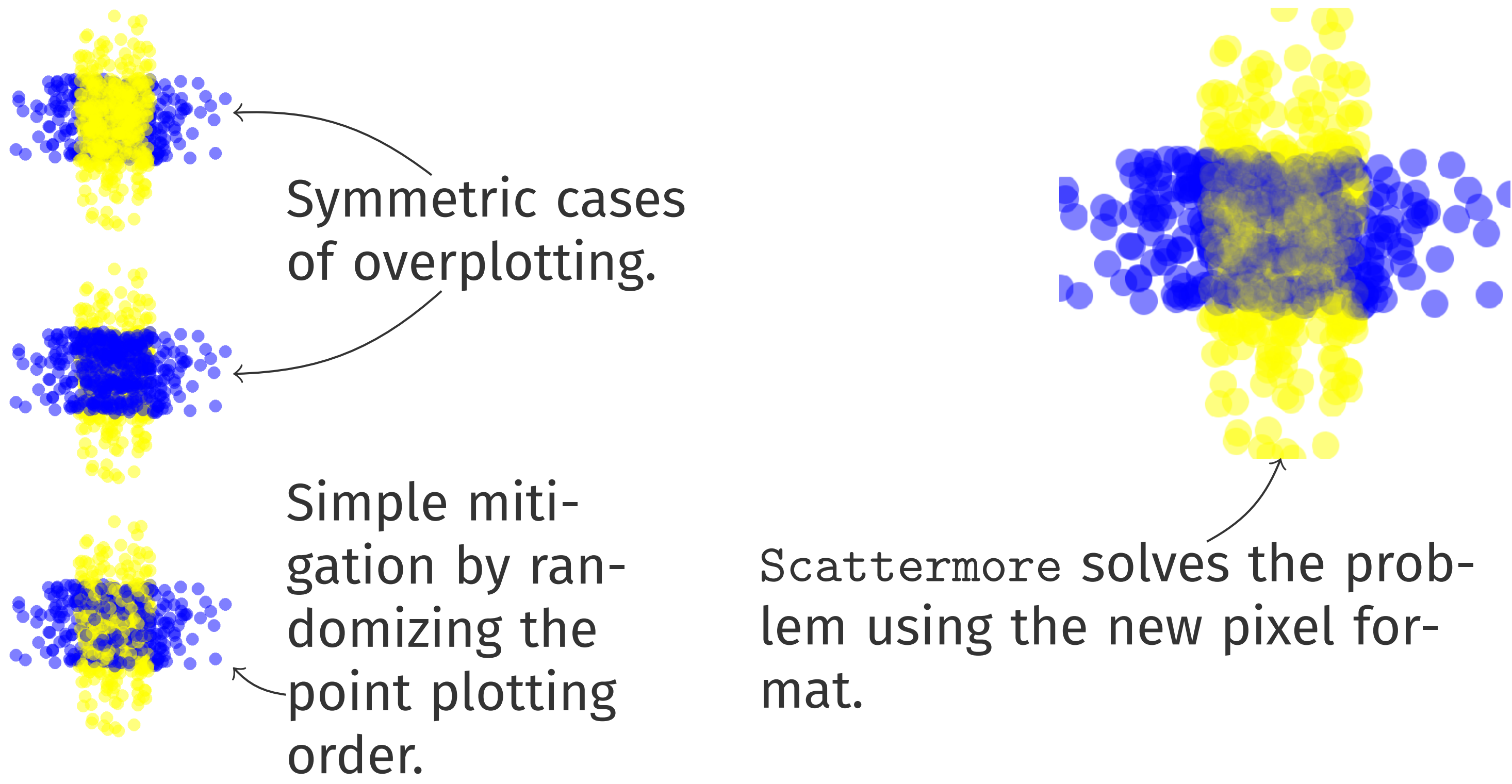
[Citing scattermore](#)

Developers

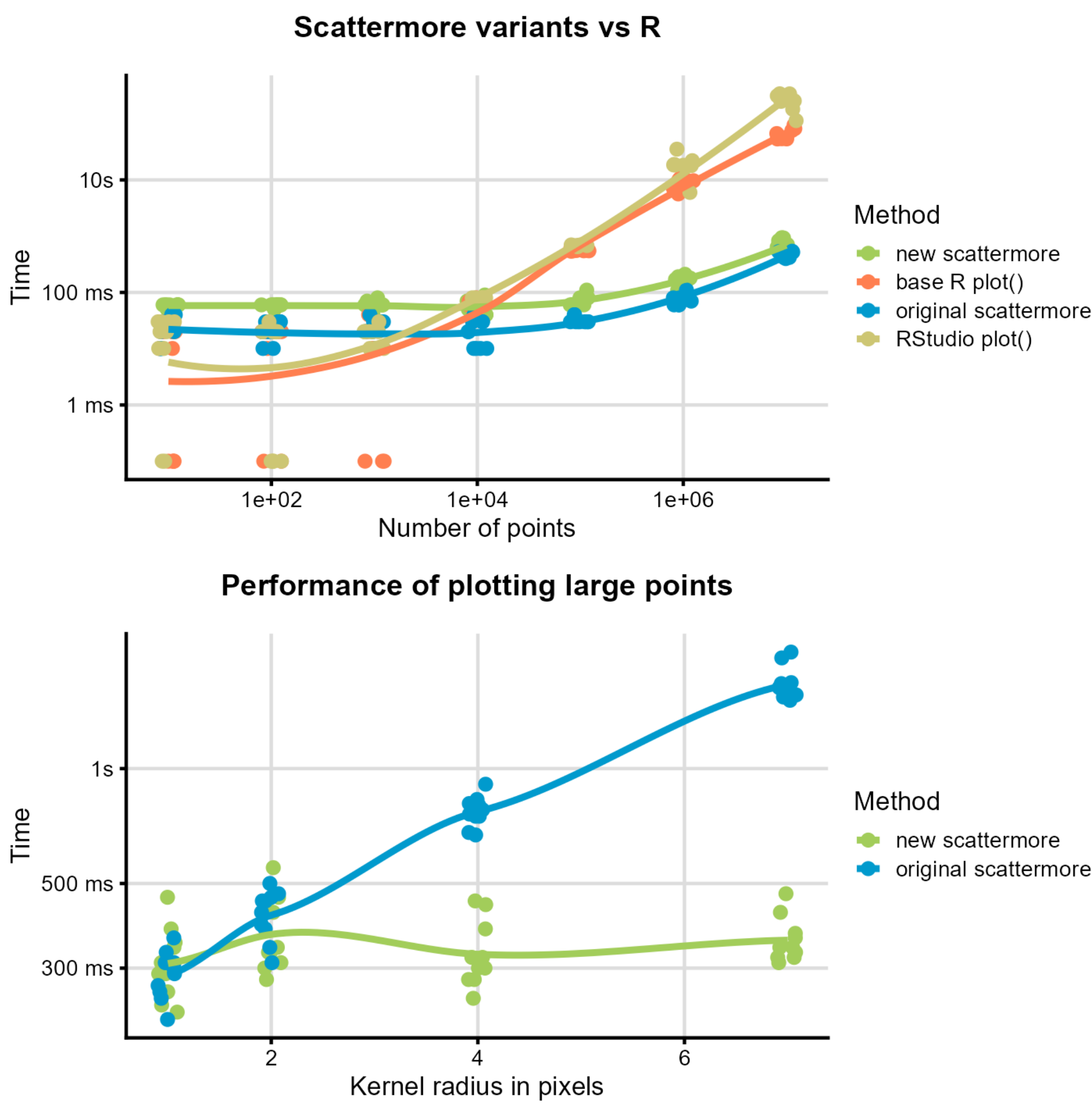
Tereza Kulichova
Author

Mirek Kratochvil
Author, maintainer

Result: RGBWT avoids overplotting artifacts



Result: Performance improvements



Result: Complex plotting



Combined plotting of more complicated primitives utilizing functionalities in the redesigned API.

Conclusions

- Scattermore enables users to blur images, blend or merge them according to their preferences and, at the same time, sufficiently fast.
- It is especially useful in the cytometric community for interactive analysis of large biological data.
- Scattermore is available on GitHub as a development version and as R's official CRAN package version.

