HAL
open science

# Genuage: visualize and analyze multidimensional single-molecule point cloud data in virtual reality

Thomas Blanc, Mohamed El Beheiry, Clément Caporal, Jean-Baptiste Masson, Bassam Hajj

## To cite this version:

Thomas Blanc, Mohamed El Beheiry, Clément Caporal, Jean-Baptiste Masson, Bassam Hajj. Genuage: visualize and analyze multidimensional single-molecule point cloud data in virtual reality. Nature Methods, 2020, 17 (11), pp.1100 - 1102. 10.1038/s41592-020-0946-1 . hal-02989365

**HAL Id: hal-02989365**

**https://hal.science/hal-02989365v1**

Submitted on 4 Dec 2020

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Genuage: visualizing and analyzing multidimensional point cloud data in virtual reality

Thomas Blanc[1,2,T], Mohamed El Beheiry[1,2,3,T], Clément Caporal[1,2], Jean-Baptiste Masson[3,*,†], Bassam Hajj[1,2,*,†]

[1] Laboratoire Physico-Chimie, Institut Curie, PSL, Research University, CNRS UMR168, 75005, Paris, France

[2] Sorbonne Universités, UPMC Univ Paris 06, 75005 Paris, France

[3] Decision and Bayesian Computation, CNRS USR 3756, Department of Computational Biology and Neuroscience. Institut Pasteur, Paris, France

*Email: jean-baptiste.masson@pasteur.fr, bassam.hajj@curie.fr

T and †: contributed equally

**Abstract:**

The quantity of experimentally recorded point cloud data, such generated in single-molecule experiments, is increasing continuously in both size and dimension. Gaining an intuitive understanding of the data and facilitating multi-dimensional data analysis remains a challenge. It is especially challenging when static distribution properties are not predictive of dynamical properties. Here, we report a new open-source software platform, Genuage, that enables the easy perception, interaction and analysis of complex multidimensional point cloud datasets by leveraging virtual reality. We illustrate the benefit of the Genuage with examples of three-dimensional static and dynamic localization microscopy datasets, as well as some synthetic datasets. Genuage has a large breadth of usage modes, due to its compatibility with arbitrary multidimensional data types extending beyond the single-molecule research community.

**Main text:**

A consistent research trend today is to design and leverage experiments that generate massive databases. The complexity of data is equally increasing, raising a major concern on data presentation and interpretation, especially in early phase research. Several fields share the common feature of generating multidimensional point cloud data such as single-molecule localization microscopy [1], flow cytometry, transcriptomics and astrophysics to name a few. To tackle the interpretation of high-dimensional data, it has become common to seek low-dimensional embeddings of the N-dimensional data (N >> 3); typically, in to mostly two dimensions. In this format, it becomes possible to extract exploitable visual graphs [2], although at the price of lower fidelity to the ground truth and the concealment of interesting features.

A demonstrative example of this trend is found in single-molecule (SM) localization microscopy, where a single experiment can result in millions of individual 3D localizations [3]. New imaging approaches have furthermore extended localizations to encapsulate information complimentary to the position. These include temporal information (i.e. individual molecule trajectories) and molecular orientation [4], which notably grants new insights into how molecular arrangements influence global cellular architecture and function.

While automation and reproducibility of SM experiments remain a sought-after goal by many research groups, the fundamental challenge of extracting meaningful quantitative information from a single experiment persists. For example, extracting shapes from sparse SM point clouds, identifying a region of interest in a new experiment or asserting a domain boundary are complex tasks that resist standardization. Exploiting any geometric information stored in SM localizations is nontrivial; linking it to the evolution of dynamic properties (e.g. diffusion or drift)

is particularly difficult as the statistical properties of point density may not be directly linked to their dynamical properties [5].

Recently, virtual reality (VR) has been re-introduced as a technology of interest for a vast number of applications due to the availability of low-cost consumer headsets and developer kits. The technology has already shown great promise for scientific research [6,7]. Our conviction is that there is an immediate dual opportunity for VR in scientific research:

- Advanced representation and easy navigation of data to allow information to be disseminated rapidly
- Human-in-the-loop data treatment that mixes data interaction in VR to automated algorithms

With this motivation, we present Genuage, a VR-compatible open-source software platform compatible with any form of point cloud data, although we emphasize here its application to SM experiments. In complement to the approach described in ref [8] focusing on super-resolution, we extend the use of VR to dynamic data and higher dimensions. The design and architecture of Genuage were built around a simple basic concept that we see as essential in all scientific VR software: it is a 3D visualization vehicle. Specifically, VR serves to facilitate volumetric data interpretation by the user and can influence actions to take (i.e. human-in-the-loop data treatment) but it is an uncomfortable technology during extended uses. Genuage handles this issue with a dual visualization interface: a desktop mode for data loading and examination on a normal computer screen, and a VR mode for efficient data visualization, interaction and analysis. The desktop mode derives its general principles from the ViSP software by being ergonomic for users, allowing key parameters to be set while avoiding wearing the headset for long periods of time [9]. We emphasize that while progress in VR is rendering the technology more accessible, the experience remains far

from comfortable and efficiency in processing the data is essential. The main characteristics of Genuage are summarized hereafter.

(i) Genuage features an efficient means for loading and presenting in VR millions of data points within seconds (Supplementary Movie1). Point cloud data can be overlaid with other high-dimensional information such as molecular orientation, particle velocity, trajectories and color code in a robust and smooth way (Figure 1).

(ii) The second main feature of Genuage, is a set of quantification tools specific to the VR mode. In VR, actions and measurements can be performed efficiently and precisely by the user in 3D (distances, angles, counting, local density calculator and histogram profiler) (Figure 1). In addition, the quantification set includes a selection tool for highlighting regions of interest for further analysis or for export (Figure 1, Supplementary movies 2-4 ). An automatically generated JSON metadata file records all the performed tasks for future sessions with Genuage and for post-treatment with other software. We provide a simple MATLAB script to illustrate how to read the JSON file and access the metadata.

(iii) Genuage communicates with common software for data analysis such as Python and MATLAB, to exchange data sets after treatment (e.g. to be visualized in Genuage) or for further analysis (e.g. after selection and thresholding in Genuage). As such, applying already established quantification algorithms requires minimal effort in readapting the codes for specific programming language.

(iv) In addition, Genuage offers a Bayesian analysis package derived from the InferenceMAP software [10] for live analysis of dynamic data in VR. In Figure 1 (Supplementary movie 5) we show an example of the trajectories of diffusing single molecules in the nucleus. The

Suppelementary movie 5 shows the user interacting with 4D localization data with Genuage. The diffusion coefficient and forces are calculated on the selected trajectory or region of interest.

Genuage is an open-source software under a BSD license. Updated versions of the software will be regularly uploaded to here. We provide a detailed user manual describing the desktop and VR tools. While the examples presented in this paper are of single-molecule datasets, Genuage is adapted to other forms of multidimensional point cloud data generated synthetically or from other experiments. Supplementary videos show various examples of case studies and data manipulation in VR.

**Acknowledgments:** We acknowledge funding from Fondation pour la recherche médicale – FRM– DEI20151234398 (B.H.), Agence National de la recherche ANR-19-CE42-0003-01, Labex Celtisphybio (ANR-10-LBX-0038), and Institut Curie (B.H.). We recognize the support by France-BioImaging infrastructure Grant ANR-10-INBS-04 Investments for the future (B.H.). We acknowledge the financial support of Agence pour la Recherche sur le Cancer (ARC Foundation) ARC (B.H.) and DIM ELICIT (B.H.). We acknowledge funding from the Pasteur Institute (J.B.M), the sponsorships of CRPCEN, Gilead Science and foundation EDF (J.B.M), the ANR-17-CE23-0016 TRamWAy (J.B.M), the INCEPTION project (PIA/ANR- 16-CONV-0005, OG) (J.B.M), the programme d'investissement d'avenir supported by L'Agence Nationale de la Recherche ANR-19-P3IA-0001 Institut 3IA Prairie (J.B.M).

**Author contributions**

TB and MEB coded the software. JBM and BH coded the analysis routines. BH provided SR and single particle tracking data. JBM and BH conceived the idea, directed the project and wrote the paper. All authors read the paper, contributed and approved the content.

## References


1 Betzig, E. *et al.* Imaging intracellular fluorescent proteins at nanometer resolution. *Science* **313**, 1642-1645, doi:10.1126/science.1127344 (2006).

2 Stuart, T. *et al.* Comprehensive Integration of Single-Cell Data. *Cell* **177**, 1888-1902 e1821, doi:10.1016/j.cell.2019.05.031 (2019).

3 Hajj, B., El Beheiry, M., Izeddin, I., Darzacq, X. & Dahan, M. Accessing the third dimension in localization-based super-resolution microscopy. *Physical chemistry chemical physics : PCCP* **16**, 16340-16348, doi:10.1039/c4cp01380h (2014).

4 Valades Cruz, C. A. *et al.* Quantitative nanoscale imaging of orientational order in biological filaments by polarized superresolution microscopy. *Proceedings of the National Academy of Sciences of the United States of America* **113**, E820-828, doi:10.1073/pnas.1516811113 (2016).

5 El Beheiry, M. *et al.* A Primer on the Bayesian Approach to High-Density Single-Molecule Trajectories Analysis. *Biophysical journal* **110**, 1209-1215, doi:10.1016/j.bpj.2016.01.018 (2016).

6 Balo, A. R., Wang, M. & Ernst, O. P. Accessible virtual reality of biomolecular structural models using the Autodesk Molecule Viewer. *Nat Methods* **14**, 1122 (2017).

7 El Beheiry, M. *et al.* Virtual Reality: Beyond Visualization. *Journal of molecular biology* **431**, 1315-1321, doi:10.1016/j.jmb.2019.01.033 (2019).

8 www.biorxiv.org/content/10.1101/2020.01.20.912733v1.

9 El Beheiry, M. & Dahan, M. ViSP: representing single-particle localizations in three dimensions. *Nat Methods* **10**, 689-690, doi:10.1038/nmeth.2566 (2013).

10 El Beheiry, M., Dahan, M. & Masson, J. B. InferenceMAP: mapping of single-molecule dynamics with Bayesian inference. *Nat Methods* **12**, 594-595, doi:10.1038/nmeth.3441 (2015).

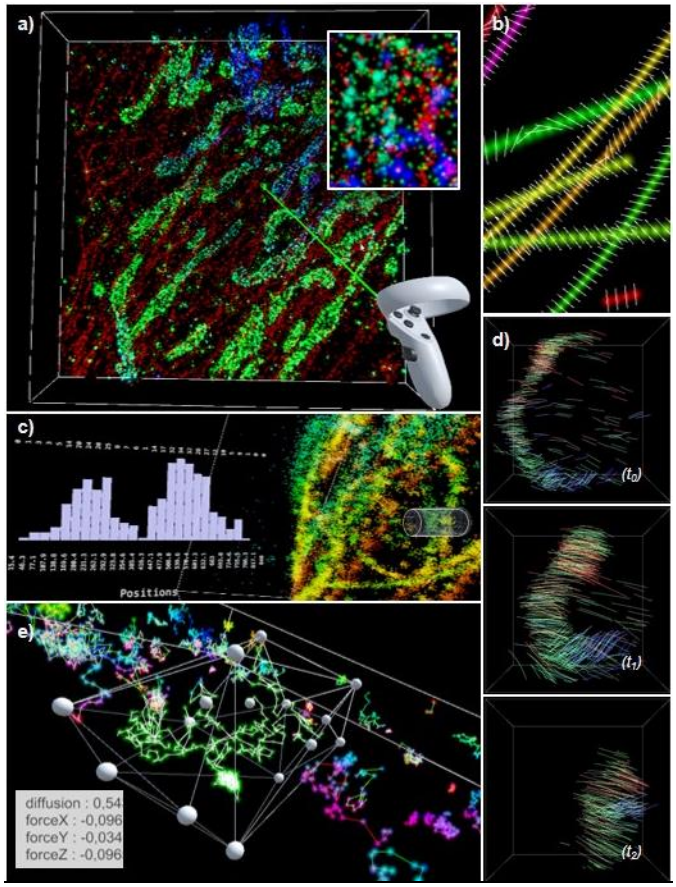


**Figure 1:** An overview of Genuage platform: a) Visualization of 2 colors 3D super-resolution point cloud data in virtual reality. The data corresponds to 2 colors SR images of mitochondria and tubulin in Hela cells. (inset) Points are presented by Gaussian spheres with transparencies. b) 3D point cloud data of simulated data points located on fibers are overlaid with additional information on 3D molecular orientation. c) VR facilitates measurements and building histograms of molecular distributions in 3D along a cylindrical segment. The cylinder crosses two overlapping tubulin fibers. d) Point cloud data with temporal information are explored. e) Retrieving local dynamical properties by applying built-in algorithms on selected 3D data points. Beads were injected in the nucleus of U2OS cells and tracked in 3D using multifocus microscope. The units of the calculated diffusion coefficient and force are data-dependent. Here the units are in $\mu m^2/s$, and $pN$ respectively.