Comparative visual analytics for multi-modal single-cell and spatial omics data

Basu, S.

Comparative Visual Analytics for Multi-Modal Single-Cell and Spatial Omics Data

Soumyadeep Basu

Colophon

About the cover:

The front cover of this thesis features a human brain outline forming a border. Within this outline, a node-link diagram connects multiple nodes, each color-coded to correspond to a specific brain region. This visualization reflects the brain-related data examined in this work. Inside the brain outline, icons representing twenty-five mammalian species highlight the comparative focus of the thesis, which involves analyzing brain datasets across different species. Additional icons depicting charts and visual analytics tools, such as graphs and boards, emphasize the development and application of visual analytic methods throughout the research.

The back cover presents an abstract representation of a DNA strand, symbolizing the analysis of gene expression data across species.

Comparative Visual Analytics for Multi-Modal Single-Cell and Spatial Omics Data
Soumyadeep Basu

Thesis layout & cover designed by Soumyadeep Basu
with design suggestions from Thomas Kroes.

Comparative Visual Analytics for Multi-Modal Single-Cell and Spatial Omics Data

Proefschrift

ter verkrijging van
de graad van doctor aan de Universiteit Leiden,
op gezag van rector magnificus prof.dr. S. de Rijcke,
volgens besluit van het college voor promoties
te verdedigen op dinsdag 15 september 2026
klokke 10:00 uur

door

Soumyadeep Basu
geboren te Kolkata, West-Bengalen, India
in 1993

Promotor:	Prof. Dr. B. Lelieveldt
Co-promotor:	Dr. T. Höllt *Delft University of Technology, The Netherlands*
Leden promotiecommissie:	Prof. Dr. J. Batenburg *Leiden University, The Netherlands*
	Dr. O. Basak *University Medical Center Utrecht, The Netherlands*
	Dr. K. Buehler *Vienna Research Center for Visual Computing, Austria*
Secretaris:	Prof. Dr. O. Meijer

The research presented in this thesis was conducted in the Netherlands at the Division of Image Processing (LKEB), Department of Radiology at Leiden University Medical Center, and the Computer Graphics and Visualization (CGV) group within the Department of Intelligent Systems, Faculty of Electrical Engineering, Mathematics and Computer Science, at Delft University of Technology. It was financially supported by the NWO TTW project 3DOMICS (NWO project number 17126) and the NWO Gravitation project BRAINSCAPES: A Roadmap from Neurogenetics to Neurobiology (NWO project number 024.004.012).

"Arise, awake, and stop not till the goal is reached."

SWAMI VIVEKANANDA

Dedicated to my parents

RUPASRI BASU AND SATYAJIT BASU

CONTENTS

1 General Introduction **1**
- 1.1 Background . . . 1
- 1.2 Visualization Challenges in Large-Scale Brain Cell Atlases . . . 1
- 1.3 Web-Based, Desktop-Based, and Script-Based Computational Approaches . . . 2
- 1.4 Towards Interactive Visual Analytics for Comparative Analysis . . . 3
- 1.5 Thesis contributions and outline . . . 4

2 Cytosplore Simian Viewer **7**
- 2.1 Introduction . . . 8
- 2.2 Biological Background . . . 9
- 2.3 Related Work . . . 10
- 2.4 Domain Abstraction . . . 11
  - 2.4.1 Data . . . 12
  - 2.4.2 Analysis Goals . . . 14
  - 2.4.3 Tasks . . . 14
- 2.5 Cytosplore Simian Viewer . . . 16
  - 2.5.1 Heatmap View . . . 17
  - 2.5.2 Scatterplot View . . . 19
  - 2.5.3 Differential Gene Expression View . . . 19
  - 2.5.4 Gene Expression Comparison View . . . 19
- 2.6 Use Cases . . . 21
  - 2.6.1 Analysis of Laminar Cell Distribution . . . 21
  - 2.6.2 Analysis of Gene Expression Patterns . . . 23
- 2.7 Discussion and Conclusion . . . 24

3 Cytosplore EvoViewer **27**
- 3.1 Introduction . . . 28
- 3.2 Biological Background . . . 30
- 3.3 Related Work . . . 32
  - 3.3.1 Comparative Visualization . . . 33
- 3.4 Domain Analysis . . . 33
  - 3.4.1 Design Process . . . 33
  - 3.4.2 Data . . . 34
  - 3.4.3 Domain Questions . . . 35

3.4.4 Tasks . . . 36
3.5 Cytosplore EvoViewer . . . 36
3.5.1 Workflow . . . 37
3.5.2 Cell Type Hierarchy . . . 37
3.5.3 Cell Abundance and Gene Identification . . . 38
3.5.4 Single Cell Embedding . . . 41
3.5.5 Phylogenetic Tree . . . 42
3.6 Validation . . . 43
3.6.1 Case Study . . . 43
3.6.2 Participants and Setup . . . 46
3.6.3 User Study . . . 47
3.7 Design Reflections . . . 51
3.8 Conclusion . . . 52

4 Cytosplore Gradient Surfer **53**
4.1 Introduction . . . 54
4.2 Results . . . 55
4.2.1 Interactive probing of spatial directions reveals gene expression gradients . . . 57
4.2.2 Cross-dataset projection is rotationally invariant and highly robust against protocol, cell subsampling, and feature subsampling . . . 61
4.2.3 Cell-type-specific gradient decomposition reveals intra-type spatial variation . . . 62
4.2.4 Cross-dataset projection and correlation analysis reveal conserved and divergent spatial gene expression gradients . . . 65
4.3 Discussion . . . 67
4.4 Limitations and Future Work . . . 69

5 Cytosplore Viewer **71**
5.1 Introduction . . . 72
5.2 Comparative brain atlas analysis . . . 73
5.3 Impact through Cytosplore Viewer . . . 74
5.3.1 Supporting Comparative Analysis with Visual Analytics . . . 74
5.3.2 Facilitating pairwise cross-species comparison . . . 75
5.3.3 Scaling cross-species comparison to phylogenies . . . 77
5.3.4 Exploring gene expression gradients across species . . . 79
5.3.5 Investigating the coordination between regulatory landscapes and gene expression . . . 81

5.4 Discussion 83

6 Discussion **85**

Supplementary Information **91**
S1: Supplementary for Cytosplore Gradient Surfer (Chapter 4) 91
S1.1: Technical validation 91

References **95**

Summary **103**

Samenvatting **105**

List of publications **107**

Acknowledgements **109**

Curriculum Vitae **111**