Visual Analytics of Single Cell Microscopy Data Using a Collaborative Immersive Environment

Author:
Lock, John; Filonik, Daniel; Lawther, Robert; Pather, Nalini; Gaus, Katherina; Kenderdine, Sarah; Bednarz, Tomasz

Publication details:
International Conference on Virtual Reality Continuum and its Applications in Industry (VRCAI '18)
pp. 1 - 4
9781450360876 (ISBN)

Event details:
International Conference on Virtual Reality Continuum and its Applications in Industry (VRCAI '18)
Hachioji, Japan
2018-12-02 - 2018-12-03

Publication Date:
2018-12-02

Publisher DOI:
https://doi.org/10.1145/3284398.3284412

Visual Analytics of Single Cell Microscopy Data Using a Collaborative Immersive Environment

John G. Lock
UNSW, School of Medical Sciences
Sydney, NSW, Australia
john.lock@unsw.edu.au

Daniel Filonik
UNSW Art & Design, CSIRO Data61
Sydney, NSW, Australia
d.filonik@unsw.edu.au

Robert Lawther
UNSW Art & Design
Sydney, NSW, Australia
r.lawther@unsw.edu.au

Nalini Pather
UNSW School of Medical Sciences
Sydney, NSW, Australia
n.pather@unsw.edu.au

Katharina Gaus
UNSW School of Medical Sciences
Sydney, NSW, Australia
k.gaus@unsw.edu.au

Sarah Kenderdine
EPFL
Lousanne, Switzerland
sarah.kenderdine@epfl.ch

Tomasz Bednarz
UNSW Art & Design, CSIRO Data61
Sydney, NSW, Australia
t.bednarz@unsw.edu.au

ABSTRACT

Understanding complex physiological processes demands the integration of diverse insights derived from visual and quantitative analysis of bio-image data, such as microscopy images. This process is currently constrained by disconnects between methods for interpreting data, as well as by language barriers that hamper the necessary cross-disciplinary collaborations. Using immersive analytics, we leveraged bespoke immersive visualizations to integrate bio-images and derived quantitative data, enabling deeper comprehension and seamless interaction with multi-dimensional cellular information. We designed and developed a visualization platform that combines time-lapse confocal microscopy recordings of cancer cell motility with image-derived quantitative data spanning 52 parameters. The integrated data representations enable rapid, intuitive interpretation, bridging the divide between bio-images and quantitative information. Moreover, the immersive visualization environment promotes collaborative data interrogation, supporting vital cross-disciplinary collaborations capable of deriving transformative insights from rapidly emerging bio-image big data.

CCS CONCEPTS

• **Human-centered computing** → **Visual analytics**; *Virtual reality*; • **Applied computing** → **Systems biology**;

KEYWORDS

High-Performance Visualization, Visual and Immersive Analytics, Confocal Microscopy, Systems Microscopy

VRCAI '18, December 2–3, 2018, Hachioji, Japan

ACM ISBN 978-1-4503-6087-6/18/12…$15.00
https://doi.org/10.1145/3284398.3284412

ACM Reference Format:
John G. Lock, Daniel Filonik, Robert Lawther, Nalini Pather, Katharina Gaus, Sarah Kenderdine, and Tomasz Bednarz. 2018. Visual Analytics of Single Cell Microscopy Data Using a Collaborative Immersive Environment. In *International Conference on Virtual Reality Continuum and its Applications in Industry (VRCAI '18), December 2–3, 2018, Hachioji, Japan.* ACM, New York, NY, USA, 4 pages. https://doi.org/10.1145/3284398.3284412

1 BACKGROUND INTRODUCTION

A revolution is occurring in the life sciences; systems microscopy integrates automated cellular imaging with computational image quantification and statistical analyses to reveal how complex processes underlying human physiology and disease emerge in time and space1. Yet converting this bio-image 'big data' into actionable insights is challenging. This reflects a disconnect between the direct visual perception of information in images and the interpretation of multiple numerical features extracted from those images. This disconnect coincides with deeper fissures at the boundaries between biological and analytical sciences, where differences in language and expertise hamper inter-disciplinary collaboration.

We have utilized an immersive, interactive visualization environment to bridge the gap between bio-images and image-derived quantitative data. This facilitates the formation of new insights and a more comprehensive understanding of the multiple forms of data from the synthesis of information in images and image-derived quantitative data, while also improving communication and collaboration between biological and analytical disciplines.

2 IMMERSIVE VISUALIZATION SYSTEM

Design of all components and experimental investigations on visualization techniques used in this case were carried out at the Expanded Perception and Interaction Centre (EPICentre[1]), at the UNSW Art & Design Faculty, see Figure 1.

EPICentre is a pioneering visualization facility that builds on decades of UNSW expertise in the design of interactive virtual environments and applications. It forges new ground in integrated

[1] http://epicentre.matters.today

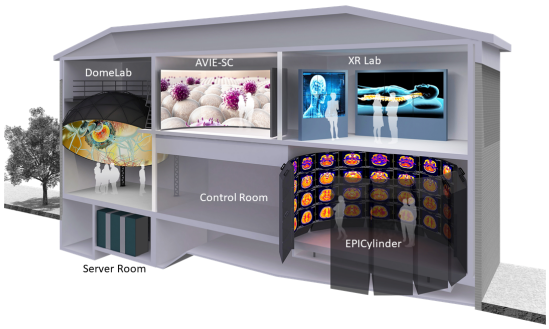

Figure 1: *EPICentre* visualization laboratories.

thinking (artistic and scientific) to facilitate understanding of complex datasets and ultra-scale imagery. EPICentre promotes cross-connection of visualization with applied computational simulations, artificial intelligence, and creativity in arts and science.

Figure 2: *EPICylinder* collaborative visualization system.

For this use case we employed the EPICylinder (see Figure 2), which constitutes a total of nearly 120 million pixels in stereoscopic 3D (26880×4320@120Hz). This is achieved by $56 \times 60''$ display cubes, assembled in a 4×14 matrix with 1-2mm edge-to-edge bezels. Visualizations are driven by a 28 nodes cluster with Xeon E5-2650 and 28 NVIDIA Quadro M6000. All visualizations were developed using Unity3D engine, running in cluster-based mode.

3 USE CASE: SYSTEM MICROSCOPY OF CANCER CELL MOTILITY

While the underlying visualization platform is easily generalisable, we developed it for the purpose of analysing confocal microscopy-based time-lapse recordings of individual cancer cell motility (imaged at the Live Cell Imaging Unit, Department of Biosciences and Nutrition, Karolinska Institute, Sweden [Shafqat-Abbasi et al. 2016]). Novel image-derived quantitative data measuring 52 cell features were subsequently extracted using Cell Profiler software [Carpenter et al. 2006] – as listed in Table 1.

Cancer cell motility drives metastasis, thereby causing more than 90% of cancer mortality [Chaffer and Weinberg 2011]. Efforts to understand and ultimately block cancer cell motility are currently hindered by several key challenges that recur broadly across the life sciences [Lock and Strömblad 2010]. Cell motility behaviour emerges from complex and dynamic patterns of molecular organization. Large numbers (10s - 100s) of quantitative features are therefore required to characterize behaviour and organization for each cell, meaning that quantitative data are complex and high-dimensional [Hernández-Varas et al. 2015; Kiss et al. 2015; Kowalewski et al. 2015; Shafqat-Abbasi et al. 2016]. Cancer cells are heterogeneous, both between cells and within cells as they change over time [Friedl and Wolf 2009; Shafqat-Abbasi et al. 2016]. This demands capacity to visualize and define both the sub-populations of cells and their evolution over time [Carpenter et al. 2006]. Given high-dimensional data defining various cellular sub-populations, understanding which features distinguish cell sub-populations and how these features correlate across cell populations demands the ability to interactively sort and explore both images and derived numerical data.

4 INTERACTION DESIGN

The interaction design takes advantage of the unique characteristics of immersive environments. All the interactions were designed around the use of a fully tracked pair of shutter glasses and controller. As a result, it is possible to navigate the virtual 3D environment in a natural fashion.

The visual representation itself is based on a conventional scatterplot, which was refined and tailored to the specific characteristics of the data set. The visual encodings consist of position, color, and shape. Special consideration was given to presentation in an immersive environment, using modern graphics programming techniques to enable real-time, smooth interactions, as well as enhancing the aesthetic appearance. Furthermore, the basic representation was augmented with special-purpose interactive features.

To address the scientific challenges outlined above, our visualization environment provides multiple distinct yet integrated representations of underlying biological data. Each representation exposes new information in the data, while integration of these differing views connects information. Insights are therefore additive, creating a comprehensive understanding of the data rather than a series of isolated findings. Below, we highlight several such data representations, noting the insights they provide and how they are connected.

4.1 Dimension Reduction (Cell Attributes)

The higher dimensional data set comprises a total of 52 distinct measurements of cell attributes obtained through microscopic image segmentation, along with subsequent image processing and computer vision. Therefore, to gain an overview of heterogeneity across all recorded data, the user can view cell observations (i.e. every cell at every time point) arranged in abstract three-dimensional visualization space resulting from several dimension reduction techniques, including: Principal Component Analysis (PCA), Multidimensional Scaling (MDS) and t-Distributed Stochastic Neighbour Embedding (t-SNE).

To assist users in comparing the results of different algorithms, we seamlessly interpolate between the results when transitioning between them. The continuous transition aids the perception of the position change of clusters of cells in the final visualization space. During the transformation different clusters may emerge, depending on the characteristics of the respective dimension reduction algorithms. The immersive nature of the visualization environment entices users to view these transformations from various perspectives within the data. Each cell observation may be represented either as a sphere or by the original time-lapse microscopy recording, as well as by colour coded to define unique cell identities (across all time points) or relative levels of any one of the 52 recorded features.

The direct embedding of images into the dimension reduction space, combined with colour-coding by individual feature values, ensures seamless integration of a) images with b) 52-dimensional image-derived data and c) univariate feature trends. This aids intuitive interpretation by manner of visual exploration. Correlation in feature trends across cell populations can also be assessed, either pairwise, by mirroring the dimension reduction representation with two distinct feature value colour-codings, or by presenting 52 views simultaneously, where each view is colour-coded based on one of the recorded features.

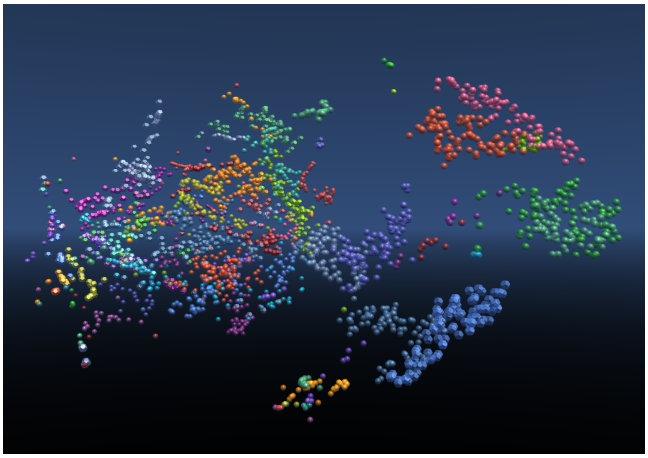

Figure 3: Cell Attributes.

4.2 Temporal Dimension (Cell Lifelines)

The full data set not only contains 52 distinct measurements for each individual cell, but also tracks how these attributes change over time. In the main visualization, observations belonging to a single cell can be identified by the color and shape encoding of the marks. For the purpose of exploring the temporal dimension and heterogeneity in cells over time, the application provides a dedicated view that is triggered by selecting a group of cell observations. This view prominently features the cell images as they were captured by the microscope. It is possible to cycle through the individual images to observe how the cell moves and changes shape on the dish. Additionally, the observations are connected by a line geometry, indicating the progression of their attributes through time in the abstract visualization space.

Scrolling through time-points then links changes in position within the dimension reduction space to visual differences in cell organization. This intuitively integrates variability in cell morphology/behavior, including in response to experimental perturbations (e.g. drugs), translated into visual or quantitative changes.

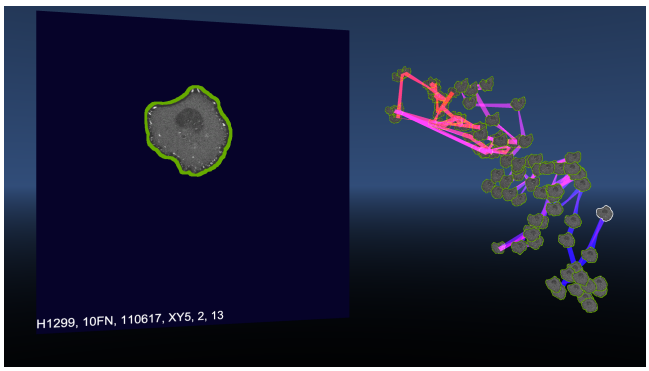

Figure 4: Cell Lifelines.

4.3 Outlier Detection (Cell Variance)

One of the common issues that were discovered by viewing the data set in the immersive environment were anomalous measurements due to errors in the image processing algorithm. The visual inspection of the spatial arrangement of marks in the abstract visualization space was well-suited to detect outliers. This allows us to improve the processing pipelines, but also could possibly be used to detect unique cell behaviors reacting to applied drugs.

To further support this use case, an interactive feature was implemented that allows collapsing all the observations of a cell into a single mark. The mark geometry used in this instance is an axis-aligned cuboid, and its extent along each axis represents the variance of the data along said dimension. As a result, it is possible to quickly compare the variance of observations among different cells. This is a powerful method to perform quality control, and to identify cells experiencing unusually high or low temporal heterogeneity.

This feature significantly reduces the visual noise, and could potentially support much larger data sets, as we are expecting to receive in future iterations of the study.

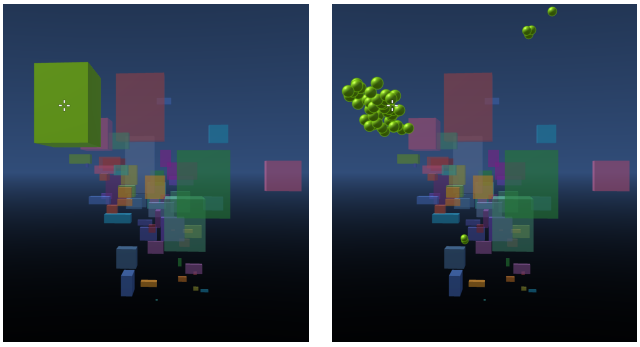

Figure 5: Cell Variance a) collapsed and b) expanded.

4.4 Linked Brushing (Cell Comparison)

Finally, the application also provides an option to dynamically select and inspect subsets of the complete data set. The desired selection is made using a brushing mechanism. The 3D tracking information of the controller is used to determine location of the marks along the pointer direction. The controller can be freely moved in space, adding cell observations to the selection in a painterly fashion whenever one of the main buttons is held down.

Up to two independent selections can be made simultaneously with different brushes, triggering a direct comparison. When needed, the comparison can be performed along each of the underlying 52 measurements with the aid of a secondary linked parallel coordinates view. This rapidly reveals how cell subsets differ in each feature, and what multivariate 'phenotypic fingerprint' defines subpopulations differences more systemically.

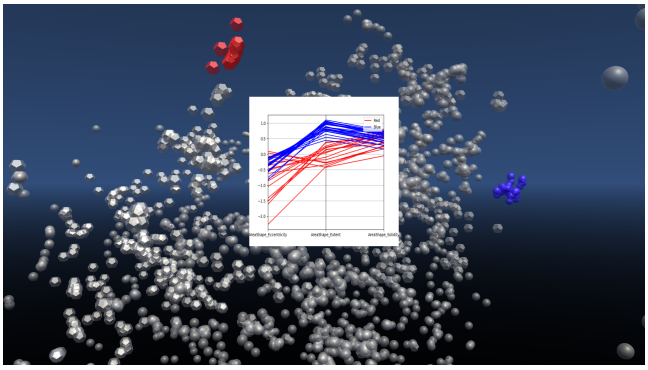

Figure 6: Interactive comparison using parallel coordinates.

5 CONCLUSIONS

While the availability of large scale display and tracking technologies is growing, more research is needed to evaluate the effectiveness of immersive environments for specific visualization purposes. This project presents one such practical application of an immersive environment for inspecting single cell microscopy data.

For life science researchers, such visualization capabilities can accelerate the derivation of critical knowledge from bio-image big data. Not simply a matter of exploring more data faster, the integration of multiple, complementary data representations ensures that complex insights can be intuitively 'constructed' on several synergistic perspectives. This has real potential to enhance rates of progress in life science research, including in clinical areas related to cancer diagnosis and therapy development, as described herein.

Moreover, these research enhancements are further buttressed by collaboration-supporting features of the immersive visualization environment. By forming intuitive links between images and image-derived quantitative data, all within a collaborative physical space, biological and statistical experts can establish tangible connections between familiar and unfamiliar data forms – thereby extending and combining their expertise. Again, accelerating a critical and often rate-limiting process, that of inter-disciplinary communication, this capacity also has strong potential to promote rapid progress in both fundamental and clinical research.

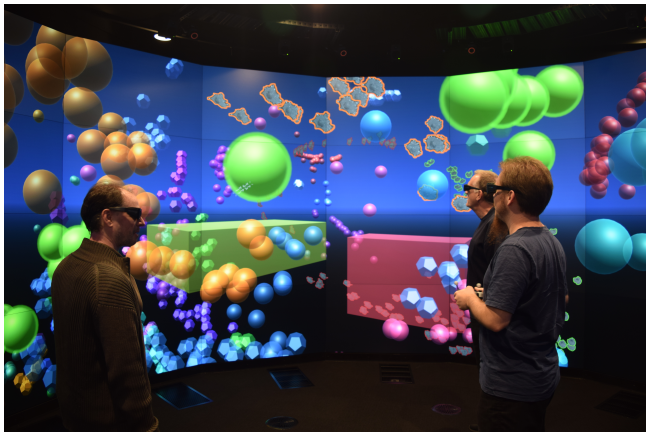

Figure 7: Users collaboratively analysing cell data.

A CELL PARAMETERS

Table 1: Summary of cell parameters.

Shape	Intensity (PAX/LifeAct)	Tracking
Compactness	Intensity	Distance
Eccentricity	– Integrated	– Integrated
Euler Number	– Max/Min	– Traveled
Extent	– Mean/Median/Std	Displacement
Feret Diameter	– Mean Absolute Deviation	Lifetime
– Max/Min	– Upper/Lower Quartile	Linearity
Form Factor	Intensity Edge	Trajectory
Length	– Integrated	– X/Y
– Major/Minor Axis	– Max/Min	
Orientation	– Mean/Std	
Perimeter	Mass Displacement	
Radius		
– Max		
– Mean/Median		
Solidity		

REFERENCES

Anne E Carpenter, Thouis R Jones, Michael R Lamprecht, Colin Clarke, In Han Kang, Ola Friman, David A Guertin, Joo Han Chang, Robert A Lindquist, Jason Moffat, and others. 2006. CellProfiler: image analysis software for identifying and quantifying cell phenotypes. *Genome biology* 7, 10 (2006), R100.

Christine L Chaffer and Robert A Weinberg. 2011. A perspective on cancer cell metastasis. *Science* 331, 6024 (2011), 1559–1564.

Peter Friedl and Katarina Wolf. 2009. Plasticity of cell migration: a multiscale tuning model. *The Journal of cell biology* (2009), jcb–200909003.

Pablo Hernández-Varas, Ulrich Berge, John G Lock, and Staffan Strömblad. 2015. A plastic relationship between vinculin-mediated tension and adhesion complex area defines adhesion size and lifetime. *Nature communications* 6 (2015), 7524.

Alexa Kiss, Xiaowei Gong, Jacob M Kowalewski, Hamdah Shafqat-Abbasi, Staffan Strömblad, and John G Lock. 2015. Non-monotonic cellular responses to heterogeneity in talin protein expression-level. *Integrative Biology* 7, 10 (2015), 1171–1185.

Jacob M Kowalewski, Hamdah Shafqat-Abbasi, Mehrdad Jafari-Mamaghani, Bereket Endrias Ganebo, Xiaowei Gong, Staffan Strömblad, and John G Lock. 2015. Disentangling membrane dynamics and cell migration; differential influences of f-actin and cell-matrix adhesions. *PloS one* 10, 8 (2015), e0135204.

John G Lock and Staffan Strömblad. 2010. Systems microscopy: an emerging strategy for the life sciences. *Experimental cell research* 316, 8 (2010), 1438–1444.

Hamdah Shafqat-Abbasi, Jacob M Kowalewski, Alexa Kiss, Xiaowei Gong, Pablo Hernandez-Varas, Ulrich Berge, Mehrdad Jafari-Mamaghani, John G Lock, and Staffan Strömblad. 2016. An analysis toolbox to explore mesenchymal migration heterogeneity reveals adaptive switching between distinct modes. *Elife* 5 (2016), e11384.