Volume xx (200y), Number z, pp. 1–19

A survey of visualisation for live cell imaging

A. J. Pretorius[1], I. A. Khan[2], and R. J. Errington[2]
[1]School of Computing, University of Leeds, UK
[2]School of Medicine, Cardiff University, UK

Abstract

Live cell imaging is an important biomedical research paradigm for studying dynamic cellular behaviour. Although phenotypic data derived from images are difficult to explore and analyse, some researchers have successfully addressed this with visualisation. Nonetheless, visualisation methods for live cell imaging data have been reported in an ad hoc and fragmented fashion. This leads to a knowledge gap where it is difficult for biologists and visualisation developers to evaluate the advantages and disadvantages of different visualisation methods, and for visualisation researchers to gain an overview of existing work to identify research priorities. To address this gap, we survey existing visualisation methods for live cell imaging from a visualisation research perspective for the first time. Based on recent visualisation theory, we perform a structured qualitative analysis of visualisation methods that includes characterising the domain and data, abstracting tasks, and describing visual encoding and interaction design. Based on our survey, we identify and discuss research gaps that future work should address: the broad analytical context of live cell imaging; the importance of behavioural comparisons; links with dynamic data visualisation; the consequences of different data modalities; shortcomings in interactive support; and, in addition to analysis, the value of the presentation of phenotypic data and insights to other stakeholders.

Categories and Subject Descriptors (according to ACM CCS): I.3.8 [Computer Graphics]: Applications—Visualisation of phenotypic data derived from live cell imaging I.3.8 [Computer Graphics]: Applications—Dynamic data visualisation

1. Introduction

Biology is rapidly changing from a benchtop paradigm to a computational science where [WS07], increasingly, biologists are automating large-scale experiments to capture large collections of results as digital images [Car07]. Time-lapse microscopy, in particular, allows biologists to image live cell experiments as they progress [Jen13]. By employing image processing algorithms, the temporal dynamics of cells can then be derived. This approach has been applied, for example, to study the effects of genes on cell division in human cancer cells [NWH*10].

Visualisation offers a way to analyse and explore the data obtained from live cell imaging. Yet, existing results are dispersed amongst diverse papers, many published in the biomedical domain. Because visualisation methods are often research theme specific, this restricts their use to specific research groups. Moreover, users of these methods typically seek to advance a particular niche in the biological sciences and not the state of the art in visualisation research. As a result, despite the major impact of live cell imaging on biology and the potential of visualisation to assist in the analysis and communication of data derived form live cell imaging, there is a knowledge gap between the biology and visualisation research communities. This makes it difficult for biologists and visualisation developers to judge the suitability of different visualisation methods, and for visualisation researchers to gain an overview of existing work to identify research priorities.

For reasons outlined here, we anticipate an increase in the importance of visualisation for live cell imaging and in future collaboration on this topic between biologists, visualisation developers, and visualisation researchers. First, live cell imaging is recognised as a major growth area in bioscience research [HGO10], and a number of ambitious research consortia have been established. This includes CellCognition, MitoCheck, and the Systems Microscopy Consortium [HSF*10, Mit, Sys]. Second, data analysis is increasingly highlighted as a bottleneck [XFP*15], as substantial support for imaging is now in place. There are at least four microscope companies that offer live cell imaging platforms [Lei, Nik, Oly, Zei], several that offer other live cell

imaging instruments [Bec, Pha, Son, Tes], and a number of automated live cell imaging solutions [GE ,Mol,Per]. Third, visualisation has been singled out as having great potential for facilitating analysis of live cell imaging data [WSB*10]. We therefore conclude that it is critical to address the knowledge gap noted above.

To this end, we present for the first time a survey of visualisation for live cell imaging. In doing so, we contribute a structured analysis to critically assess past results to inform future work. The paper is organised as follows. We provide a concise description of live cell imaging in Section 2. Next, in Section 3, we describe our analytical approach. Based on this, Section 4 contains the results of a systematic analysis of visualisation methods for live cell imaging. This includes reports on domain and data characterisation (Section 4.1), task abstraction (Section 4.2), and visual encoding and interaction design (Section 4.3). We discuss the implications of our analysis, including a number of research gaps, in Section 5 and conclude with Section 6.

2. Live cell imaging

To provide context for our survey, we now provide a brief overview of the live cell imaging paradigm. An exhaustive account is beyond the scope of this paper, and we refer the interested reader to detailed descriptions in [IJS*07, Jen13, WS07].

Live cell imaging involves recording images of microtitre plates (also known as multi-well plates) containing a grid of small depressions called wells [IJS*07]. Every well contains a small volume of the biology being studied (live cell culture, tissue, or organism). Pipetting is used to add chemical compounds to some of the wells while others are left untreated to serve as experimental controls. The key technologies are time-lapse microscopy and subsequent image processing [WS07]. Time-lapse microscopy captures the dynamic behaviour of cells by recording image sequences at the appropriate temporal and spatial resolution. Because manual inspection is tedious, image processing algorithms are used to derive rich descriptive phenotypic data from the images to allow for downstream analysis. A phenotype is an observable characteristic such as a physiological or behavioural property.

Figure 1 summarises the four main steps of live cell imaging: cell and compound preparation, image acquisition, image processing, and data analysis and exploration [Car07, VLH*06]. We describe these below:

1. *Cell and compound preparation.* The plate is prepared for imaging. Cell quality must be ensured, for example, by maintaining environmental conditions [Jen13].
2. *Image acquisition.* The choice of imaging technique depends on the aims of a study. The most popular technique is fluorescence microscopy where fluorescent markers enable the identification of specific structures by using specific light wavelengths [LC05]. Fluorescent protein genes, in particular green fluorescent protein (GFP), are spliced into DNA close to the region that codes a target protein [CT10]. When the target is activated, so is the fluorescent marker. For example, Figure 2 shows a sequence of time-lapse images (images were originally captured at 15 minute intervals, but selected frames are shown to highlight important phases of the cell cycle). Just before cell division, Cyclin B1 protein translocate to the nucleus and become activated. This gives rise to an increased fluorescence intensity in the fifth and sixth frames. 2D images are most common, but techniques like confocal and light-sheet microscopy [KSWS08], enable reconstruction of 3D volumes from image stacks. An exhaustive discussion of imaging is beyond the scope of this paper and we refer the interested reader to surveys in [LC05, CT10].
3. *Image processing.* This step typically involves image preparation, including illumination correction; image segmentation, where objects such as cells are identified by separating them from the background; and feature extraction, where phenotypes of interest are quantified [WS07,CJL*06]. Phenotypes often include summary measures (such as cell counts), descriptions of morphology (cell shape, area, or texture) and fluorescence intensity (which captures protein expression). Image processing produces a multidimensional data set where a vector of features is associated with every detected object. For example, in Figure 2, by first segmenting cells and then considering their fluorescence intensity, it is possible to derive the consecutive cell cycle phases that the cell marked by the black arrowhead goes through before division (M1-, G1-, S-, G2-, Pro-, and Metaphase). Feature extraction may be followed by tracking, where objects are linked across consecutive images [JLM*08]. This yields time series that describe, for example, the trajectory of a cell. Sophisticated approaches combine tracking with event detection [ARG*06], where key events like cell division and cell death are identified to produce cell lineages: hierarchical descriptions of the genealogy of a population of cells [GLHR09]. Figure 2 shows how the cell indexed by a black arrowhead changes its shape and position and, from the sixth to seventh frames, divides into two cells.
4. *Data analysis and exploration.* Generally, there are two approaches for analysing phenotypic data obtained from live cell imaging [Car07]. First, hypothesis-driven analysis tests specific theories by quantifying and analysing relevant features. Second, with explorative analysis, there is no hypothesis, but a general interest in any relevant phenotypic changes that result from perturbations. A perturbation is a change in a cell's environment brought on, for example, through disease or drug treatment. Explorative analysis typically requires the computation of multiple features that are rich enough to capture many phenotypes [WS07].

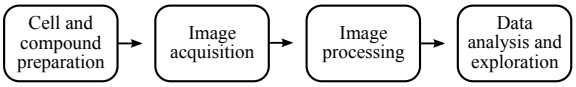


Figure 1: *The time-lapse microscopy workflow (adapted from [Car07, VLH*06]).*

A common goal of live cell imaging is to test and study the influence of perturbations, such as chemical compounds (also called small molecules), on cell phenotypes. In drug discovery, for example, hit selection involves screening many compounds to identify those with a desired effect [WS07]. Live cell imaging also facilitates more in-depth analysis, for example, to understand the action mechanisms of perturbations such as anti-cancer agents [KHC*07], or to unravel the mechanisms of disease [VLH*06]. As a result, it is also an important paradigm for fundamental cell biology research.

Based on the aims of a live cell study, an existing experimental procedure is adopted or a novel one is developed. Such an experiment design, or protocol, specifies compound concentrations, timings, and so forth, and is referred to as an assay [IJS*07]. In practice, however, experiments themselves are often called assays or screens.

3. Approach

To inform and structure our analysis of visualisation methods for data derived from live cell imaging, we draw on recent visualisation theory. We outline this work below.

3.1. Visualisation theory

The visualisation research community considers design studies as valuable sources of insight into domain problems and visualisation design solutions [Mey13]. This partly serves as our motivation for analysing and transferring the knowledge captured in the existing, but fragmented work on visualisation for live cell imaging. Further motivation is the drive toward systemisation and theoretical reflection on visualisation research. We are inspired by Sedlmair et al.'s work on design study methodology and their emphasis, in addition to visual design, on data and task abstraction [SMM12]. Unlike them, however, our objective is not to propose a comprehensive procedure for conducting design studies, but to structure and analyse the problem and design spaces of visualisation methods for a particular domain: live cell imaging. As we describe in Section 5.3, to some extent, our work also relates to dynamic data visualisation more generally.

We employ Munzner's nested model for capturing design decisions [Mun09, MSQM13]. It conceptualises the visualisation design space as an interdependent chain of consecutive design components. From high- to low-level, the components are:

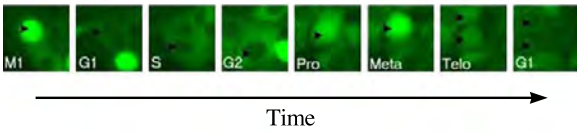


Figure 2: *A sequence of images obtained by time-lapse microscopy showing different phases of the cell cycle. The image sequence also shows spatial orientation, morphological changes, and the expression profile of green fluorescent protein controlled through the cyclin B1 promoter protein (in green). A tracked cell is marked with a black arrowhead (the phases of its cell cycle are labeled in white). From the sixth to seventh frames the tracked cell divides into two daughter cells. Images were originally captured at 15 minute intervals, but here selected frames are shown to emphasise important phases of the tracked cell's cell cycle.*

1. Domain and data characterisation
2. Task abstraction
3. Visual encoding and interaction design
4. Algorithm design

We will use the first three of these components to structure our analysis of visualisation methods for data derived from live cell imaging. As prescribed by Munzner, we will apply them in the above order. Algorithm design, the fourth component, is beyond the scope of this paper and we encourage the interested reader to consult the relevant cited papers for further details.

Although Munzner's model emphasises task abstraction, in fact, it does not provide guidance for abstracting tasks. To address this, we use Brehmer and Munzner's task typology for task abstraction [BM13]. This typology offers three advantages. First, by prescribing a controlled vocabulary, it structures analysis (we use **bold** type for prescribed terms). Second, it bridges the gap between low- and high-level descriptions, unlike other taxonomies that focus either on low-level tasks (for example, [AES05]), or high-level intents (for example, [LS10]). Third, it disambiguates the ends (why?) and the means (how?) of tasks, while providing a way to capture important contextual detail (what?).

We also considered other recent approaches to visualisation task abstraction, in particular that of Roth [Rot13], and that of Schultz et al. [SNHS13]. We found, however, that the former's focus on cartographic interaction primitives and the latter's reliance on a formal abstract grammar make them less suitable for our objectives.

3.2. Methodology

It has been convincingly argued that qualitative analysis supports a holistic understanding by considering visualisation in its context of use [IZCC08]. This is the approach we take by

Approach	Papers
Spatial embedding	[CBI*08, FHWL12, GBBS09, GME03, HLLK09, PM07]
Space-time cubes	[MSD06, MWV*03, TBM*99]
Temporal plots	[GLHR09, KHC*07, KLC*11, MWV*03, PKE15, SHT*12, SMC*06, WHN*09]
Aggregate visualisations	[BGS13, DTW*15, CJL*06, PKE15, SHT*12, JKW*08]
Dimension reduction	[BSB*11, HWKT09, SBB*12]
Lineage diagrams	[ARG*06, CBI*08, ENS09, FHWL12, GLHR09, KHC*07, PKE15, SMKM10, WWB*14, WWR*11]

Table 1: *Six classes of visualisation methods for live cell imaging with representative papers.*

qualitatively analysing published research that describe visualisation methods used in a live cell imaging context. We identified relevant work by reviewing the visualisation and high-throughput screening literature. We found some applicable papers in the visualisation literature, but most were published at biology outlets.

A shortlist of 76 papers was compiled with input from biologists at the Broad Institute and Cardiff University. From these, we identified 28 papers particularly relevant to this study (see Table 1). Our criterion for inclusion was not technical novelty, though some methods are, but the application of visualisation to support real-world analysis for live cell imaging. Some methods, such as temporal plots, are very popular and we picked papers with the most emphasis on visual analysis. We cannot claim to be 100% comprehensive, but believe these papers to be a representative sample of visualisation methods for live cell imaging. For qualitative studies, validity must also be considered [IZCC08]. For this reason, and to inject rigour and counter bias, we base our analysis on the well-motivated theoretical models and frameworks described in Section 3.1 (as opposed to devising our own). As we show in Section 4, papers were systematically coded in terms of these frameworks: domain and data characterisation, task abstraction, and visualisation and interaction design.

Finally, we note that papers from biology outlets emphasise advances in the biological sciences and not visualisation research. In this paper we present this work to a visualisation audience for the first time. To a biology audience, who may be familiar with subsets of these methods from a pragmatic point of view, we present for the first time a broader systematic analysis from a visualisation perspective.

4. Results

By applying the approach outlined in Section 3, we identified six classes of visualisation methods for live cell imaging: spatial embedding, space-time cubes, temporal plots, aggregate visualisations, dimension reduction, and lineage diagrams. Table 1 cites relevant papers for every class, Figure 3 provides diagrammatic summaries of different approaches within each class, and Figures 4–9 show selected examples of implemented systems. Due to space constraints and copyright restrictions, image reproductions of implementations for every individual technique is unfeasible. For this we refer the interested reader to the relevant cited papers.

As we show below, the classes of methods differ in three respects. First, from a data perspective, their emphases range from the positions of objects in the field of observation to abstract data derived during post-processing. Second, from a task perspective, they support different modes of analysis emphasising, for example, temporal changes, aggregate behaviour, or descendant relationships. Third, from a visual design perspective they emphasise different data properties to support different tasks. The sections below may be read in serial, to compare different methods with respect to parts of the analytical framework we use (see Section 3.1), or in parallel, to consider different aspects of particular methods.

4.1. Domain and data characterisation

In this section, we characterise the users and objectives supported by the six classes of visualisation methods outlined above. We also describe the data that these classes of methods cater for.

4.1.1. Users

Visualisation methods for live cell imaging are used by algorithm developers, biochemists, bioinformaticians, cell and molecular biologists, and geneticists.

4.1.2. Objectives

The objective common to all visualisation methods is to understand dynamic cellular behaviour and the influence of perturbations (for example, drug treatment or disease) on such activity. Sometimes users also want to judge the quality of post-processing algorithms. Below, we discuss specific user objectives in more detail.

Spatial embedding. At a basic level, users are interested in cell growth, cell movement (motility), cell shape (morphology), and cell reproduction (proliferation) as captured by phenotypes visible in the acquired images [GME03, GBBS09, HLLK09, PM07]. Further objectives include understanding the influence of cell positions and migration, that is, their impact on cell fate, tissue formation, organs, and organisms [CBI*08, FHWL12]. Spatial embeddings also

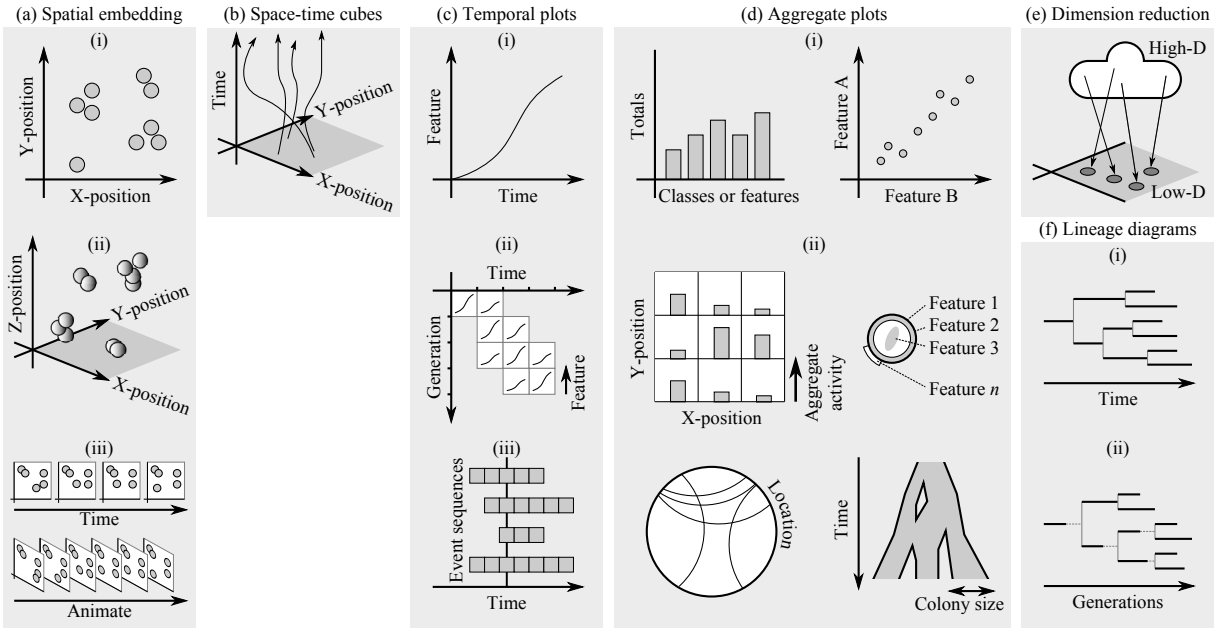


Figure 3: *Six classes of visualisation methods for live cell imaging data. (a) Spatial embedding, where cells and other structures are visualised in the 2D or 3D field of observation. (b) Space-time cubes, where cell position is mapped to the x- and y-axis and time is mapped to the z-axis. (c) Temporal plots include (i) plotting a derived feature as a function of time, (ii) dividing elapsed time into intervals (x-axis) and considering feature behaviour at different generational distances from a progenitor cell (y-axis), and (iii) showing phenotypic event sequences (x-axis) associated with different genes (y-axis). (d) Aggregate plots include (i) standard visualisations of relationships between features such as histograms and scatterplots. (ii) Custom visualisations have also been developed, for example, by showing aggregate levels of activity for discrete spatial regions; by using glyphs that encode data features with visual attributes including radius, line width, orientation, and so forth; by positioning proteins on a circle and encoding protein location changes with arcs; or by showing how an entire cell colony splits and merges over time. (e) Dimension reduction, where data are considered as vectors in high-dimensional space and mapped to 2D with low-dimensional projection techniques. (f) Lineage diagrams, which show the proliferation of cells as a branching tree structure, typically oriented left-to-right or top-to-bottom. Cell tracks can be aligned (i) by elapsed time or (ii) by successive generations. Some tools combine some of these approaches.*

serve as analytical grounding, enabling users to interpret their data in a way that intuitively represents the spatial locations where activity occurred.

Space-time cubes. With this approach, users most often want to confirm and present the results of image processing algorithms [MSD06, TBM*99]. Occasionally, they also want to confirm and present observations made with other modes of analysis, including quantitative analysis [MWV*03]. We will show how space-time cubes attempt to support this by incorporating space and time in a single representation.

Temporal plots. User objectives share an emphasis on understanding cellular behaviour specifically related to time. This includes analyses of temporal effects of perturbations on cell development [GLHR09, SHT*12, WHN*09], and temporal function of sub-cellular structures [SMC*06, MWV*03, WHN*09]. There are also users who aim to refine and validate simulation models of temporal cellular behaviour [KHC*07, KLC*11], or study the temporal properties of successive generations of proliferating cells [PKE15].

Aggregate visualisations. Users want to understand cellular behaviour at an aggregate level (versus single-cell studies). Typically, this takes the form of cellular responses to perturbations at whole-experiment or sub-population level [CJL*06, BGS13, JKW*08, PKE15], or at regular temporal intervals [DTW*15]. Occasionally, users study higher-level biological organisations, such as cell colonies [SHT*12].

Dimension reduction. Some users ultimately strive for automated statistical analyses of live cell imaging results. To achieve this, they want to investigate results of image classification algorithms to assess their quality [HWKT09]. Other users are interested in the influence of multiple perturbations on cellular behaviour, specifically to explore the effects of different, but structurally similar chemical com-

pounds [BSB*11, SBB*12]. As we will show, projecting high-dimensional data points to 2D offers a way to address these objectives.

Lineage diagrams. With lineage diagrams, the objective is to understand the intra- and cross-generational behaviour of proliferating cells. This includes temporal development, divisional history, key cellular events (such as cell division and cell death), and how these aspects are interrelated [GLHR09, CBI*08, ENS09, PKE15]. This assists users in studying the processes of cell differentiation into specialised cell types [FHWL12, WWB*14, WWR*11]. To reliably derive this kind of data, users also want to validate and correct image processing results [ARG*06, WWB*14, WWR*11], and validate models of behaviour of cells and their progeny (that is, their descendants) [KHC*07, SMKM10].

4.1.3. Data

For all methods, data originate from images obtained with time-lapse microscopy, typically fluorescent microscopy (see Figure 2 and Section 2), and vary in the degree and nature of post-processing applied. This often leads to multimodal data that include, for example, temporal and multivariate components. In a few instances, simulated data are also considered. Below, we consider in more detail the data that the six classes of visualisation methods we identified were designed for.

Spatial embedding. Imaging is performed at cellular or sub-cellular resolution with sub-cellular structures, such as nuclei [GBBS09], or whole cells [HLLK09, PM07], marked by fluorescent proteins. Cells may be imaged at different light excitation wavelengths to record different fluorescent labels [HLLK09]. Such sequences of "raw" time-lapse images are transformed to obtain the following data: temporal sequences of 2D fluorescent intensity maps, each containing measured fluorescence intensities for every pixel in the corresponding image [GBBS09]; temporal sequences of 2D images or 3D volumes [CBI*08, FHWL12]; or animated movies constructed by sequentially animating through the time-lapse images [PM07].

Space-time cubes. Consecutive images are processed to produce time series that contain the spatial coordinates of cells or sub-cellular structures [MSD06, MWV*03, TBM*99]. Additional measures are often derived and included, for example, fluorescence and luminosity values.

Temporal plots. Similar to space-time cubes, image data are processed to produce temporal data sets. These include: time series of multiple single-cell measures or of population-level aggregate measures [MWV*03, SMC*06]; and cell lineages that integrate time series into hierarchical descriptions of cellular descendent relationships, including key cellular events (such as cell division and cell death) [KHC*07, GLHR09, KLC*11]. Occasionally, measures of variation are included with time series [SHT*12]. More complex transformations can also be performed. This includes: grouping cells by binning elapsed time into discrete intervals and considering cells' generational distance from a progenitor cell [KLC*11, PKE15]; and time-resolved records of cellular responses to perturbations, which are further processed to produce sequences of typical cellular events (referred to as event-order summaries) [WHN*09]. Simulated data obtained from models of cellular behaviour are sometimes used [GLHR09, KLC*11].

Aggregate visualisations. Most often, aggregation relies on images grouped by experiment. Sometimes it starts during image acquisition by employing imaging techniques that optimise the identification of spatially clustered cells (for example, phase contrast imaging) [SHT*12]. Image processing can also be used to yield multi-dimensional data associated with individual cells (which are aggregated during visualisation) [CJL*06, JKW*08], or groups of cells. These are typically measures of size, shape, and texture, but may include assay-specific properties such as protein abundance. Image processing may also yield time series containing the locations of cellular events [PKE15], sometimes with additional derived measures that describe motility, morphology, and associated uncertainty [DTW*15]. At the per-cell level, it is possible to associate additional information, such as protein localisation, through semi-automated markup of cells. Further processing may then be performed to produce higher-level summary data, such as a network where nodes represent proteins (with associated up- or down regulation) and directed links represent changes in protein localisation corresponding to different stress conditions [BGS13]. For cell colonies, directed networks can be derived to represent their temporal splitting and merging behaviour [SHT*12].

Dimension reduction. This approach relies on post-processing to augment image data with a multidimensional vector of features, for example, by augmenting images with statistical measures [HWKT09]. A different and innovative approach is to consider the "chemical space" of a library of small molecules that cells had been perturbed with. Here, every molecule is represented as a multidimensional vector that captures its structural detail and is linked with corresponding cellular behaviour, such as protein activity levels [BSB*11, SBB*12].

Lineage diagrams. Image processing is used to identify cells and key cellular events, and to link cells across time [ARG*06, ENS09, KHC*07, PKE15, SMKM10]. These results are integrated to produce cell lineages: hierarchical structures that capture the developmental history of a population of cells from a common progenitor cell. This description includes key cellular events such as cell division (mitosis) and cell death. Due to the complexity and challenges of the algorithmic processing required for lineage construction, and to minimise uncertainty, this approach may include manual curation [KHC*07, WWB*14, WWR*11], or simulation based on mathematical models of cellular

behaviour [GLHR09]. Lineages may be linked to objects, such as cells, identified in the original images [CBI*08, FHWL12].

4.2. Task abstraction

We now consider the tasks supported by the six classes of visualisation methods we identified. We first consider the ends of the tasks (why?), then the means (how?), and then consider additional contextual details (what?).

4.2.1. Why?

In this section we consider tasks related to user objectives. At a high level, all approaches share **verification** and **discovery** tasks [BM13], which align with the two general approaches for analysing data obtained through live cell imaging (see Section 2, Step 4) [Car07]. All methods also aim to **present** findings. This is achieved by different combinations of lower level tasks, described below.

Spatial embedding. Users want to **discover** patterns of behaviour, cell division, and cell specialisation into anatomical structures [CBI*08, FHWL12]. To achieve this they must **locate** sub-cellular, cellular, or spatiotemporal patterns of behaviour by **exploring** representations of fluorescence intensity maps [GBBS09], images [HLLK09], volumes [CBI*08], trajectories [FHWL12], or animated movies [PM07]. At the lowest level, this implies **identification** and **comparison** of regions of similar colour and intensity in sub-cellular structures [GBBS09], cells [CBI*08, HLLK09, PM07], or classes of similar spatiotemporal behaviour (migration characteristics, size, shape) [FHWL12].

Space-time cubes. Although space-time cubes are primarily used for **presenting** results [MSD06, MWV*03, TBM*99], to a limited extent users are also interested in **verification** of movement trajectories and classes of movement [MSD06, TBM*99]. Users **locate** movement patterns and **locate** classes of behaviour by **identifying** and **comparing** trajectories.

Temporal plots. Users want to **explore** temporal behaviour to **locate** behavioural similarities and differences [MWV*03, KLC*11, SHT*12, WHN*09], and classes of behaviour [GLHR09, SMC*06]. If known in advance, behavioural classes are **looked up** [KHC*07, PKE15]. At a low level, users **identify** and **compare** scenarios (for example, different perturbarions) and classes of behaviour. In some cases, users also **identify** and **compare** key cellular events [MWV*03], or event sequences [WHN*09].

Aggregate visualisations. User objectives require **exploring** population-level relationships between derived features, **location** of sub-populations of cells where features take certain values, and **location** of feature thresholds where cell properties change [CJL*06, JKW*08]. To achieve this, the following is required: **comparison** of distributions of features; **identification** of outliers, minima, and maxima in feature distributions; and **comparison** of the underlying "raw" image data. **Exploration** of aggregate data (for example, at the proteome or cell colony-level) helps users interpret upstream experimentation and data analysis to **locate** classes of behaviour [BGS13, SHT*12]. Users **identify** behaviour corresponding to different experimental treatments, **identify** classes of behaviour, and **compare** classes that result from different treatments or conditions [BGS13, DTW*15, PKE15, SHT*12]. This may include **comparison** of associated derived measures [DTW*15].

Dimension reduction. Users aim to **locate** clusters of data items and classes identified by upstream classification algorithms to **explore** relationships between them [PM07]. In addition to **locating** clusters and outliers, they must also be **compared**. Thorough analysis of chemical spaces requires additional support to **browse** and **lookup** molecules by activity level and for **exploring** relationships between activity levels and compounds [BSB*11, SBB*12]. This involves **identification** of activity levels, and **identification** and **comparison** of similar compounds, their structure, and activity levels.

Lineage diagrams. Cell lineage analysis requires **lookup** of cellular events, in particular cell death and differentiation events, and **exploring** the influence of cell differentiation [KHC*07, ARG*06, GLHR09, ENS09, SMKM10]. Users also **lookup** successive generations, branching patterns, temporal distances, and (algorithmically) identified classes of behaviour [FHWL12, PKE15]. This involves **identification** of colour-coded branches and branching points within and across generations of progeny. To support the need for **exploring** relationships between lineages and to **identify** errors in upstream lineage construction, lineage diagrams may be combined with other visualisations such as spatial embeddings [CBI*08, WWB*14, WWR*11] and temporal plots [PKE15]. These further enable **identification** and **comparison** of relevant spatio-temporal behaviour across lineages.

4.2.2. How?

Here we explore tasks related to the means by which users analyse data. Exploratory **navigation** is common to all approaches and lower-level analysis is supported by different combinations of tasks, typically **filtering** and **selection**. Visualisations are very often **recorded** to illustrate findings.

Spatial embedding. Users **navigate** spatial representations of the original field of observation, 2D images, or 3D volumes for behaviour that confirm or contradict expectations [GME03, GBBS09]. Comparisons are made by considering images along a temporal axis, sometimes across different imaging modalities [HLLK09]. Users view animations or spatial trajectories embedded in 2D or 3D to analyse spatial data with respect to time [PM07]. They **navigate** animations

linearly. Users often **select** trajectories, or paths within trajectories, to view associated data. Visualisations are **annotated** by choosing different attributes [CBI*08], or precomputed behavioural classes [FHWL12], which are mapped to sub-cellular structures, cells or trajectories.

Space-time cubes. Users **navigate** trajectories to confirm expected behaviour [MSD06, MWV*03]. To avoid occlusion, they usually **filter** the data [TBM*99]. This is achieved by **selecting** a subset of time series to consider based on classes identified by classification algorithms, or by **selecting** intervals of associated features to exclude. To consider different data properties, visual attributes such as colour are **changed** by **selecting** data attributes to encode [TBM*99].

Temporal plots. Users **navigate** plots to confirm expected temporal behaviour [GLHR09,MWV*03]. They can **change** the type of plot or the features that are visualised [KHC*07, SHT*12]. Visual attributes, like colour, are **changed** by selecting data attributes to encode [SMC*06]. Temporal intervals can also be **changed**. In some cases, users **derive** and visualise summary statistics to **aggregate** the data further [KLC*11, PKE15]. With event-order maps (which show typical cellular event sequences), users **change** the layout by aligning event sequences by a specified event type [WHN*09].

Aggregate visualisations. Users **change** encodings by altering the plot type [CJL*06, JKW*08]. The data granularity can also be **changed** by viewing data per experiment, for subsets of images, or for subsets of cells [DTW*15, PKE15]. For selected data elements, users sometimes **import** underlying images or additional associated information. Users can change the focus to subsets of data by **selection** [BGS13], or can **select** different data attributes to visually encode [SHT*12]. Data are **filtered** by specifying feature values or intervals to remove. To capture their insights, users **annotate** views with thresholds, where behavioural changes are visually observed, or **derive** statistical summaries [JKW*08].

Dimension reduction. By **encoding** the results of precomputed clusters, users can **annotate** views with contextual cues [HWKT09]. When available, users also consider linked views, such as aggregate visualisations [BSB*11, SBB*12]. They **select** data in these views to **change** the main view. Users also **filter** data based on associated features by specifying values or intervals to remove. Often they **select** data points to **import** associated information such as high resolution images or chemical descriptions of molecules.

Lineage diagrams. The descendant hierarchy is **navigated** with emphasis on its temporal, generational, and branching structure [KHC*07, SMKM10]. Users **change** the event types, or other features, that are encoded by colour or texture [ENS09, GLHR09]. They **select** branching points to **change** the visualisation, for example, by showing subsets of data [ARG*06]. They also **select** layouts that either emphasise elapsed time or successive generations of cells within lineages [PKE15]. **Filtering** is performed by specifying feature values or class membership to exclude [FHWL12]. If provided, users **select** time points on lineage diagrams to view corresponding time points in linked views (such as spatial embeddings) [CBI*08]. To correct errors, some tools allow users to **change** the lineage structure or image processing parameters [WWB*14, WWR*11].

4.2.3. What?

Below, we provide contextual detail related to user tasks. In all cases, **input** are the data described in Section 4.1.3 and **output** are small collections of selected images or visualisations for inclusion in presentations, reports, and papers. Visual analysis is often performed interchangeably with other modes of analysis, for example, statistical procedures. Based on intermediate findings, users may also run additional experiments [HLLK09].

Spatial embedding. Users sometimes augment their visual analysis with other visualisation techniques, for example, temporal plots of features obtained from image processing [HLLK09]. Increasingly, visualisation tools combine spatial representations with other visualisation methods such as lineage diagrams [CBI*08, FHWL12].

Space-time cubes. Visual analysis may be preceded by quantitative analysis, for example, to classify trajectories and select a smaller subset of trajectories to visualise. Space-time cubes are usually employed in a supportive role during more rigorous analysis such as quantitative and statistical assessments [MSD06].

Temporal plots. Visual analysis is often performed in parallel to statistical analysis, while conducting further experiments [WHN*09].

Aggregate visualisations. In addition to screen captures, outputs include subsets of data elements, grouped by filtering criteria, and statistical summaries [JKW*08]. This may serve as input to further analysis. In recent work, the spatial field of observation serves as a frame of reference for showing aggregate visualisations [DTW*15, PKE15].

Dimension reduction. The multidimensional data associated with images originate from statistical analysis or from features derived during image processing. The availability of standardised graphical representations of chemical compounds (structural formula) have been leveraged by annotating visualisations with these [BSB*11, SBB*12].

Lineage diagrams. The analysis of lineage diagrams is often preceded by in-depth quantitative analysis [KHC*07], or modelling [GLHR09]. As noted, users may also refer to other visual representations while interpreting lineage diagrams, in particular, spatial embedding [ENS09], or temporal plots [PKE15].

4.3. Visual encoding and interaction design

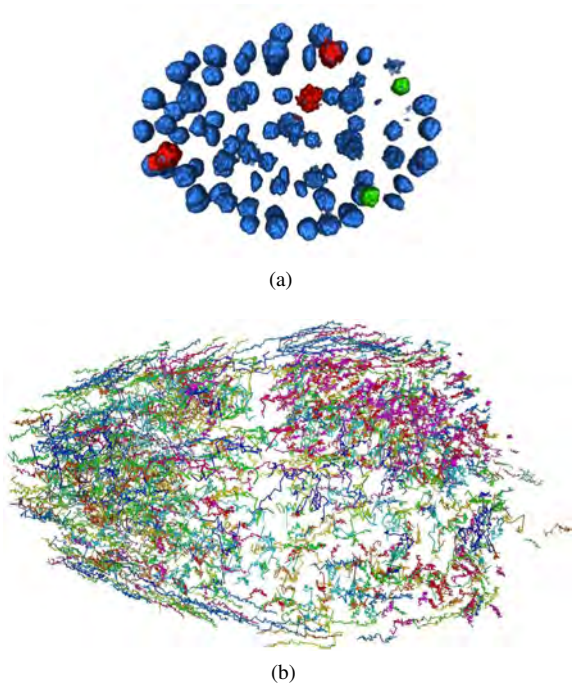
We now consider visual encodings and interaction design, which vary widely. Visual encoding and interaction design are a function of data properties and the characteristics of user tasks described above (see Sections 4.1 and 4.2) [Mun09, PvW09].

Spatial embedding. Whether 2D or 3D, spatial embeddings use position within the original field of observation as the primary visual mapping (see Figure 3(a)(i) and Figure 3(a)(ii)). Pure image-based techniques do not contain further visual encoding (Figure 2), with the exception of superimposing fluorescence intensity onto images with a colour map [GBBS09]. Images may be displayed to emphasise their sequential nature (Figure 3(a)(iii)), typically by positioning them along a horizontal axis [HLLK09]. To assist users in comparing different imaging modalities, different tracks of images that share a temporal axis, may be shown along the vertical axis. Another approach is to present images as frames in an animated movie (Figure 3(a)(iii)) [PM07].

More advanced approaches embed visualisations of sub-cellular and cellular objects, and of cellular events identified during image analysis in the 2D or 3D space of observation. This enables derived attributes to be mapped onto objects within the spatial frame of reference (see Figure 4(a)) [CBI*08]. The temporal dimension is sometimes visualised by showing object trajectories, for example of cells, directly in the plane or volume (Figure 4(b)) [FHWL12].

For pure image-based approaches, interaction support ranges from none to rudimentary, while animations are typically viewed with standard video viewers that offer functionality such as play, pause, rewind, and forward. More advanced approaches offer spatial embeddings linked to other views, for example, lineage diagrams [CBI*08]. Apart from capabilities such as panning and zooming, interaction methods like brushing enable users to highlight and select objects across views. Using standard widgets, users can specify subsets of objects to filter out, for example, by indicating a specific temporal interval.

Space-time cubes. The time series derived from high-throughput screening represent the trajectories of cells or sub-cellular structures as a temporal sequence of x- and y-positions. Space-time cubes encode these trajectories as curves in 3D by mapping the positions to the x- and y-axis and by mapping time to the z-axis (see Figure 3(b)). Additional features associated with the time series, such as measures of displacement, are often encoded on the curves with colour (see Figure 5) [MSD06, MWV*03, TBM*99]. Standard interactive capabilities are typically offered for panning, rotation, and zooming. Occlusion is often a problem and to address this, users can remove trajectories from the visualisation by using filtering techniques.



(a)

(b)

Figure 4: *Examples of spatial embedding. (a) Visualisation of cells at their spatial locations within an embryo. User-selected cells have been highlighted. Image reproduced from [CBI*08], with permission from the authors and SPIE. ©2008 Society of Photo Optical Instrumentation Engineers. (b) Visualisation of cell trajectories within their spatial field of observation. Trajectories with similar geometric properties and orientation are shown in the same colour. Image courtesy of Jens Fangerau, also see [FHWL12].*

Temporal plots. The most common approach is to encode time series of derived features as a function of time with bar or line charts [GLHR09, MWV*03, SMC*06] (see Figure 3(c)(i)). Time series may also be derived from more complex data, for example, the totals of cellular events per generation may be obtained from cell lineages [KHC*07]. When charts encode an aggregate quantity, they are sometimes augmented by encoding associated variation with error bars [PKE15], or an envelope around the curve (see Figure 6) [SHT*12].

More complex visualisations have also been developed. For example, a matrix can be defined by mapping elapsed time, divided into discrete intervals, to the x-axis and by mapping to the y-axis the generational distance of cells from a progenitor cell [KLC*11]. The intersections of these axes define different stages in an experiment (see Figure 3(c)(ii)). This approach has been used to compare plots of distributions of cells of real experiments to simulated data at each of these experimental stages. Another approach uses event-order maps to encode sequences of typical cellular events

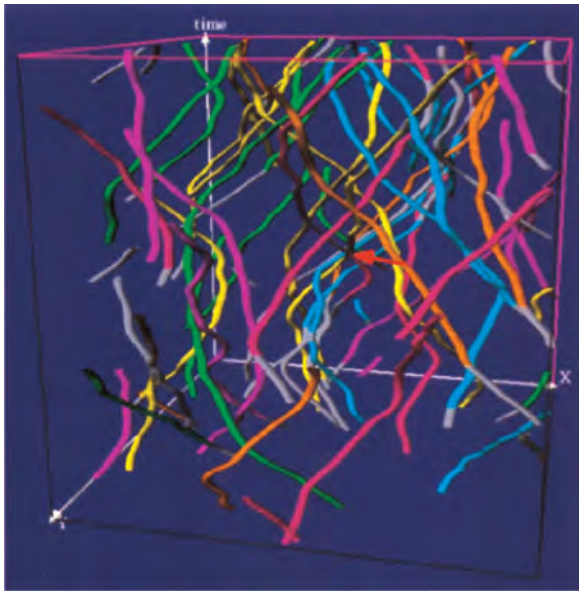


Figure 5: *An example of a space-time cube. Cell trajectories that have been identified by an object tracking algorithm are visualised by embedding them within a volume defined by x- and y-position, and time. Image reproduced from [TBM*99], with permission from the authors and PNAS. ©1999 The National Academy of Sciences.*

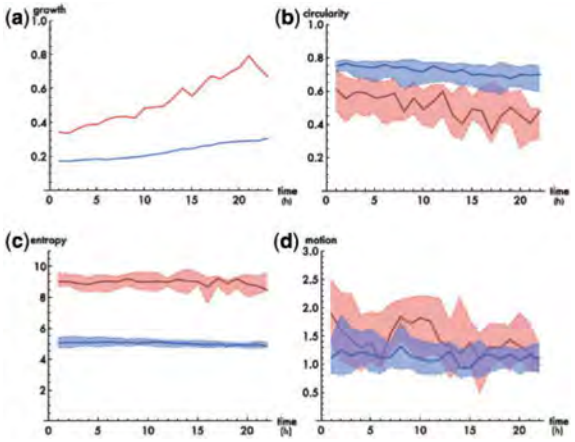


Figure 6: *Examples of temporal plots. Different derived measures that characterise cell colony development are visualised as a function of time. This is a rare example of also visualising variation, here as an envelope around each curve (with the exception of the first plot). Image reproduced from [SHT*12], with permission from the authors and under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/3.0).*

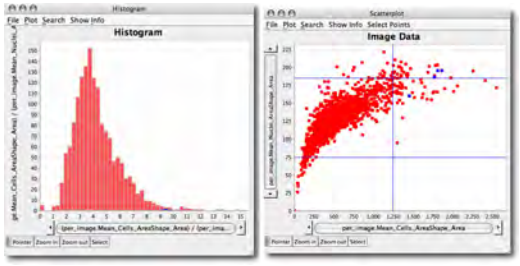
associated with different genes [WHN*09]. An event sequence is encoded by a horizontal strip of colour swatches that represent event types. Genes are mapped to the y-axis with event sequence strips shown at corresponding vertical positions (Figure 3(c)(iii)).

The reviewed literature does not report on the degree of interactive support for bar and line charts, but it is straightforward to extend them with standard interactive capabilities (brushing, selection, and filtering). The matrix-based representation of experimental stages provides limited interaction by allowing users to generate additional summary plots [KLC*11]. By enabling users to centre event-order strips around different event types, event-order maps support interactive exploration of the genetic impact of event patterns relative to different even types [WHN*09].

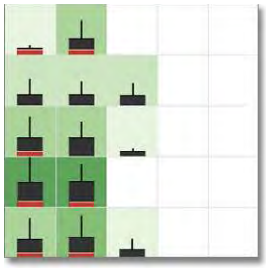
Aggregate visualisations. One approach is to visualise aggregates using well-known techniques that support flexible analysis (see Figure 3(d)(i)). These include [CJL*06, JKW*08]: histograms to show the distribution of values over predefined intervals or bins, for example, the distribution of cells by their DNA content (Figure 7(a)); scatterplots to consider relationships between pairs of derived measures, for example, overall cell area × cell nucleus area (Figure 7(a)); parallel coordinate plots to consider an arbitrary number of measures [Ins85]; and density plots where, for a pair of measures, the number of data points that map to a particular coordinate are colour-coded.

A second general approach for aggregate data is to develop custom visual encodings (see Figure 3(d)(ii)). Some researchers divide the spatial field of observation into a small number of discrete regions and summarise levels of activity, including the variation from average behaviour, for each region (Figure 7(b)) [PKE15]. This is achieved by mapping activity levels to heat maps and the height of bars shown in each region. Glyphs can also be used to summarise a number of derived features, typically describing motility or morphology, for different classes of cells (for example, for different treatment conditions) [DTW*15]. Within the spatial field of observation, such glyphs can be combined with cellular trajectories to summarise derived features for predefined temporal intervals.

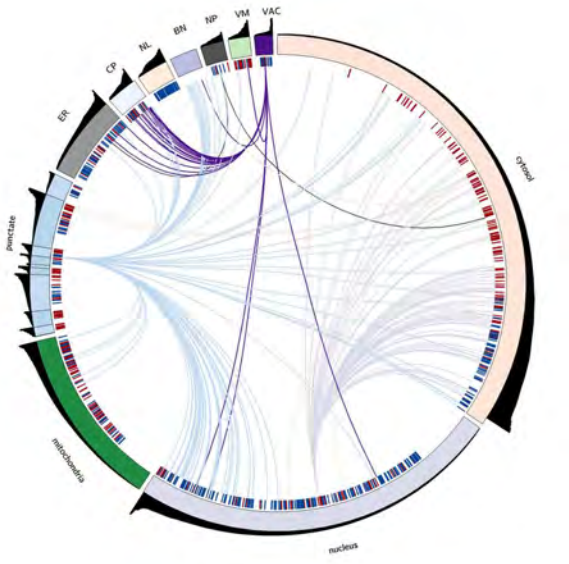
Two further methods were developed to visualise networks that summarise aggregate cellular and sub-cellular behaviour. To visualise the changes in a proteome under stress conditions, a circular network graph is used (see Figure 7(c)) [BGS13]. Nodes encode proteins and are positioned on the circumference of a circle while changes in protein location (the links in the underlying network) are encoded with directed arcs between nodes. Protein up- and down-regulation are encoded by the height of bars juxtaposed with the nodes on the circumference. As another example, a stream-like visualisation, oriented top-to-bottom, was developed to encode a network that describes the evolution of a cell colony (Figure 3(d)(ii)) [SHT*12]. Here the



(a)



(b)



(c)

Figure 7: *Examples of aggregate visualisations. (a) A histogram shows the distribution of cells by cell area, and a scatterplot shows the relationship between mean cell area (x-axis) and mean nucleus area (y-axis). Images reproduced from [JKW*08], with permission from the authors and under the terms of the Create Commons Attribution License (http://creativecommons.org/licenses/by/2.0). (b) Aggregate levels of activity, and variation from it, are shown for discrete sub-regions of the spatial field of observation. Image reproduced from [PKE15], with permission from the authors and John Wiley & Sons Ltd. ©2015 The Authors and Computer Graphics Forum, ©2015 The Eurographics Association and John Wiley & Sons Ltd. (c) Proteins are mapped to the circumference of a circle and arcs encode changes in protein locations for a particular permutation (or stress condition). Image courtesy of Maya Schuldiner, also see [BGS13].*

forking and joining of streams encode the splits and mergers of a colony, while additional features are encoded by colour.

The suite of plots offered by CellProfiler Analyst provides well-designed interactive support [JKW*08]. Users can interactively generate new plots, change the plot type and data attributes to encode. Standard interaction methods such as brushing, selection, filtering, and drill-down are provided. For the custom visualisations, the level of interaction support is less clear [BGS13, DTW*15, SHT*12], although it should be relatively simple to extend these methods with standard interactive capabilities.

Dimension reduction. The point of departure is to consider feature vectors associated with the data as points in high-dimensional space. This may include statistical properties that describe images, features that quantify properties of cells, or vectors that capture the structure of molecules [HWKT09, BSB*11, SBB*12]. Dimension reduction techniques, such as multidimensional scaling, principal component analysis, or Sammon mapping are then used to project these points onto the 2D plane such that they are drawn near each other when their corresponding vectors are proximate in high-dimensional space [Fod02] (see Figures 3(e)(a)). The corresponding data elements, images or representations of molecules, are rendered at these locations (Figure 8(a)). Associated features, such as precomputed classes of images or activity levels of molecules, are mapped to visual attributes such as colour and size.

Not all implementations of dimension reduction offer flexible interaction. However, through well-designed interaction supported by linked views, HiTSEE greatly enhances the analysis process (see Figure 8(b)) [BSB*11, SBB*12]. Users can directly interact with a list of molecules, sorted by activity level, to select a subset of molecules to visualise. From here, they are able to browse and compare structurally similar molecules. The designers have leveraged the well-understood semantics of standardised graphical representations used in chemistry, by allowing users to annotate visualisations with chemical diagrams.

Lineage diagrams. Cell lineages are typically visualised with directed node-link diagrams that place the root at the left, with the lineage hierarchy displayed from left-to-right (see Figure 3(f)(i)) [KHC*07, GLHR09, PKE15], or the root at the top, with the lineage hierarchy displayed top-to-bottom [ARG*06, ENS09, SMKM10]. With this approach,



(a)



(b)

Figure 8: *Examples of dimension reduction. (a) Images that have been annotated with derived measures that characterise sub-cellular localisation are projected to 2D. This results in images with similar characteristics being positioned near each other. Image reproduced from [HWKT09], with permission from the authors and under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/2.0). (b) Dimension reduction is also used in combination with other visualisations. Here, in the centre, small molecules are projected to 2D based on descriptions of their chemical structure to support users in studying the effect of structurally similar compounds on cellular behaviour. This approach is interesting in that, instead of visualising cells, it visualises the "chemical space" of a library of molecules that cells had been perturbed with. Image reproduced from [SBB*12], with permission from the authors and under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/2.0).*

branching points encode cell division events, leading to the birth of two daughter cells, and leaves encode cell death events or the termination of the experiment. Branch lengths encode the elapsed time between cell birth and cell division or cell death. A layout that aligns branches per generation, as opposed to absolute elapsed time, has also been developed (Figure 3(f)(ii)) [PKE15]. Typically, branches may be colour-coded to represent associated features derived during image processing. Users can pick different derived features to map to the diagrams [GLHR09,ENS09]. Such approaches are often used for static analysis or for the presentation of findings.

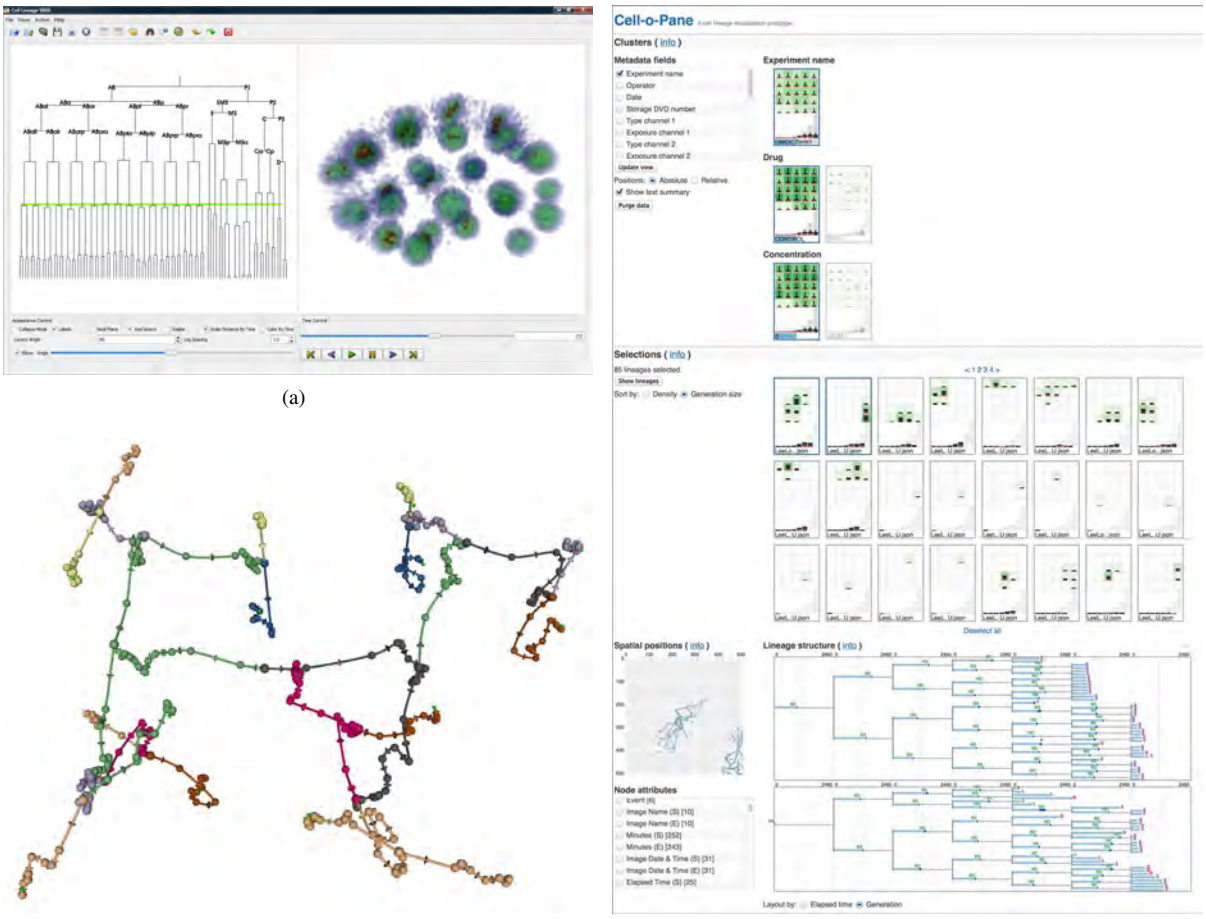
Tools like ProgeniTRAK provide limited interaction capabilities for querying features associated with cellular events [KHC*07]. Recent work emphasises interactive analysis more. This is achieved by making linked views a fundamental part of the design, for example, by linking lineage diagrams with spatial embeddings (see Figure 9(a)) [CBI*08], or temporal plots [PKE15]. Brushing enables users to correlate the structural and spatial dimensions of the data. As tools like LEVER and LEVER-3D illustrate, this can be combined with the ability to correct lineages and image processing parameters on the fly [WWB*14,WWR*11].

Another approach analyses tracks in the lineage structure to cluster paths into similar behavioural classes [FHWL12]. Going further, the approach taken with Cell-o-pane enables users to group lineages according to associated metadata (which typically describe experimental parameters) to identify and investigate particular scenarios [PKE15]. Both methods enable users to compare and analyse the spatiotemporal and structure characteristics across multiple lineages. With the former, lineage diagrams are shown in a 3D spatial embedding of the data (Figure 9(b)). The latter aligns lineages by elapsed time or by generation, and the user is further supported by linked spatial and temporal plots (Figure 9(c)).

5. Discussion

Visualisation is typically advocated as an effective and efficient approach to obtain insight from data. This is achieved by exploiting the power of the human perceptual system [CMS99, Nor06]. Indeed, the work surveyed above shows that visualisation already plays an important part in the analysis of phenotypic data obtained from live cell imaging to understand the biological processes that these data sets capture.

As we argued in Section 1, interest in and work on visualisation for live cell imaging data is likely to expand. By providing an overview of all approaches that have been developed to date, our analysis will enable biologists, visualisation developers, and visualisation researchers to evaluate the advantages and disadvantages of these methods. This will allow them to make more informed decisions on appropriate visualisation methods for their data and tasks. However, for visualisation to optimally support the analysis of data derived from live cell imaging, our analysis also highlights a number of remaining research gaps. We discuss these below.



(a) (b) (c)

Figure 9: *Examples of lineage diagrams. (a) At the left, descendent relationships between cells are visualised as a tree. The layout is from top-to-bottom and by elapsed time. This is combined with a linked view at the right, which visualises the cells in a spatial embedding. Image reproduced from [CBI*08], with permission from the authors and SPIE. ©2008 Society of Photo Optical Instrumentation Engineers. (b) It is also possible to directly embed the cell lineage structure in the spatial field of observation. Image courtesy of Jens Fangerau, also see [FHWL12]. (c) Lineage diagrams can be combined with multiple linked views to support flexible analysis of cellular behaviour, including spatial, temporal, and descendent relationships. At the bottom, to facilitate comparison, multiple lineage diagrams are shown left-to-right along the same x-axis, and aligned by generation. Image reproduced from [PKE15], with permission from the authors and John Wiley & Sons Ltd. ©2015 The Authors and Computer Graphics Forum, ©2015 The Eurographics Association and John Wiley & Sons Ltd.*

5.1. Broad analytical context

In the majority of cases considered, visualisation is interspersed with other modes of analysis. This includes deriving quantitative features, statistical analysis, modelling, simulation, and machine learning. Due to the rigour required of biomedical research, insights and hypotheses from visual analysis are often not considered sufficient and are further investigated with rigorous qualitative analysis (for example, [HLLK09,SMC*06]). Users often employ several visual and non-visual methods as and where needs arise. In addition, there is a trend of distributing data capture and analysis efforts across multiple institutions [WSB*10].

This raises the issue of provenance, or how successions of data transformations, analysis workflows, and conclusions are captured to ensure reproducibility and to enable validation. Past efforts include using metadata to describe the history of data resources [SPG05]. Visual reasoning requires more than data provenance, however, and the visualisation research community has responded by enabling users to maintain visual interaction histories, to annotate representations, and to share visualisations [HS12].

In a live cell imaging context it is not clear how visualisation systems should be designed to cater for a fragmented and non-cohesive analysis environment. In particular, how can provenance be guaranteed in such a context? One strategy, which is promoted by the visual analytics community [KKEM10], is a much tighter integration of visualisation with non-visual analysis methods. Indeed, from our analysis we conclude that when objectives and tasks can be well-defined in advance, this approach has been shown to work [FHWL12, WHN*09, JKW*08, WWB*14, WWR*11]. In general, however, it imposes significant technical and practical challenges to move beyond visualisation methods targeted at specific narrowly scoped problems [WSB*10].

Another way to address the challenge of broad analytical context would be to develop ontology-aware visualisation tools. Ontologies, such as the hugely successful Gene Ontology, are widely used in the biosciences to structure knowledge [The00]. An ontology provides a controlled vocabulary for a particular field of inquiry, encourages assimilation and transfer of new insight into existing knowledge precepts, and enables effective querying and interoperability between different databases [Bar03]. In the first instance, there is scope for developing visualisation methods that conform to particular ontologies. A more ambitious goal would be to develop methods that cater for an arbitrary ontology by conforming to existing ontology standards.

Ways of dealing with its broad analytical context could facilitate a step-change in visualisation for live cell imaging. In this respect, more design studies that explicitly characterise objectives and data, describe tasks, and link these to visual and interaction design decisions would help shed light on both the problem and the design space [BM13, Mey13, Mun09, SMM12]. In this respect, the recent founding of the Symposium on Biological Data Visualization and its emphasis on design studies is a welcome development [Sym].

5.2. Behavioural comparisons

All user tasks relate to studying and understanding cellular behaviour. This includes analysis of phenotypic responses to perturbations such as drugs, disease, or other experimental conditions [BSB*11, BGS13, CJL*06, DTW*15, GLHR09, HLLK09, JKW*08, KHC*07, PKE15, SBB*12, SMKM10]; and analysis of cellular development including cell specialisation, spatial, and temporal function of cellular and sub-cellular structures [ARG*06, CBI*08, GBBS09, PM07, MWV*03, SHT*12, SMC*06, WHN*09, ENS09, WWB*14]. To achieve this, users are often interested in modelling and algorithmic extraction of behaviour, and validation of the results [FHWL12, GME03, MSD06, TBM*99, KLC*11, HWKT09, WWR*11].

There is a persistent interest in behaviour as a function of time and, consequently, many visualisation methods emphasise temporal patterns. Three classes of methods explicitly map data to a temporal axis (*space-time cubes*, *temporal plots*, and *lineage diagrams*; see Figure 3(b), (c), and (f)). The latter caters for users who study temporal behaviour in the context of successive generations of cells. Aggregate behaviour is also important and multivariate data visualisation methods such as scatterplots, histograms, parallel coordinate plots, density plots, and glyphs are used to analyse relationships between derived features independently of time (*aggregate visualisations*; see Figure 3(d)).

Other approaches for visualising behaviour include animation [PM07], and dimension reduction. These approaches may not be optimal. Research has shown that animation often leads to interpretation errors [RFF*08], and that viewing 3D space on a 2D interface introduces perceptual challenges [War01]. With dimension reduction [HWKT09], the domain semantics of the high-dimensional distances that these methods preserve are unclear and it is difficult for users to relate visual patterns to data characteristics. However, by clearly defining the meaning of proximity, for example as structural similarity of chemical compounds, and by providing linked views that support a set of well-defined tasks, Bertini et al. and Strobelt. et al. show how this approach can support meaningful analysis [BSB*11, SBB*12].

Our survey also shows that comparison is an important and recurring task type. To facilitate comparison, most of the methods that we reviewed use superposition, for example, by showing multiple time series plots on a shared system of axes [SMC*06], or by mapping multiple cell lineages to the same spatial embedding [FHWL12]. However, as the sizes of live cell imaging data sets increase, this approach is likely to face scalability problems, for example, Gleicher et al. point out the associated challenges of visual clutter and occlusion [GAW*11].

In addition to superposition, Gleicher et al. identify two other approaches to facilitate visual comparison tasks: juxtaposition and the explicit encoding of relationships between data elements. Some of the reviewed work takes this approach, for example, to show aggregated spatial and temporal activity [PKE15], or to enable users to compare glyphs of aggregated derived measures side-by-side [DTW*15]. Nonetheless, there is a need for more targeted studies of comparison tasks and further exploration of the design space to support such tasks within the context of live cell imaging. As an example of critically reconsidering design conventions, the designers of HiTSEE question the suitability to biological data of the dominant design mantra of converging from high-level overviews of entire datasets to subsets [Shn96]. They convincingly show that users are sometimes better served by the ability to target and explore small neighbourhoods within large data spaces [BSB*11, SBB*12].

5.3. Dynamic data visualisation

There is a clear relationship between visualising behaviour captured by live cell imaging data and the more general chal-

lenge of dynamic data visualisation, which is concerned with visualising data where one or more variables change over time. A comprehensive discussion of the topic is beyond the scope of this paper, but a few observations are relevant. First, we argue that the analysis in this paper offers a valuable, albeit application-specific, contribution to the literature on dynamic data visualisation.

Second, recent work by Bach et al. offers a useful approach for reflecting on the design space of dynamic data visualisation [BDA*14]. The authors model the visualisation of dynamic temporal data as operations on a conceptual space-time cube (unlike our discussion in Section 4, the authors stress that this work is not about space-time cubes as a visualisation method, per se). They show how different ways of visualising spatial or abstract data that involve a temporal dimension, correspond to slices through, compressions of, or other transformations of space-time. In the spirit of Bach et al. and others [PvW09], we recommend that biologists, visualisation developers, and visualisation researchers working with live cell imaging data critically consider the following. What aspects of their data are being emphasised by visualisations? And, importantly, what characteristics are being suppressed?

Third, from a research perspective, live cell imaging offers both a valuable test bed for dynamic data visualisation and stands to gain from this body of work. As we wrote in Section 1, the aim of this paper is to survey existing visualisation methods for live cell imaging data. However, a plethora of visualisation methods have been developed for dynamic data visualisation [AMM*07, AMST11]. There is great scope for investigating if and how these approaches carry over to live cell imaging. For example, to pick up on a previous point, what are the circumstances (if any) under which animation adds value [TMB02]? Also, visualisation for movement data is a burgeoning field of research [AA13]. Although this work focuses on geographical space and single object trajectories, do some of the principles that have been developed also apply to live cell imaging data? Much more work is needed to investigate these and other open questions in a live cell imaging context.

5.4. Data modality, uncertainty, and curation

The cases considered in this paper show that, in a live cell imaging context, visualisation supports effective analysis for different data modalities and different levels of abstraction. This includes "raw" images obtained during image acquisition and derived data obtained during subsequent image processing: spatio-temporal data, time series, multidimensional feature vectors, and more complex structures such as cell lineages.

Yet, there is limited flexibility for the visual analysis of different data modalities and abstraction levels in an integrated fashion. Most of the approaches considered provide representations of a particular data modality at a particular level-of-detail. As Amar and Stasko argue, this lack of flexibility is likely to hinder users' ability to move beyond simplistic and static cognitive models of the phenomena their data describe [AS05]. Notable exceptions are CellProfiler Analyst, which enables users to consider data aggregates, select subsets, and view underlying images, and approaches that enable users to combine visual analysis of structural abstractions with spatial representations [CBI*08, FHWL12, PKE15, WWB*14, WWR*11].

A related issue is that of data uncertainty. With a few exceptions [DTW*15, PKE15, SHT*12], none of the cases reviewed explicitly deal with data uncertainty or variation. Most existing visualisation tools for live cell imaging data, like visualisation methods more generally, are designed with the assumption that data capture clean causal relationships and, hence, do not reveal uncertainty [AS05, SLSR10]. This is remarkable, given the complex workflows of data acquisition and post-processing typically encountered and the importance of quality control in high-throughput screening [BFHC12].

According to Saraiya et al. users in the biological sciences adopt a critical stance towards automation and prefer to have the ability to organise, review, and correct results [SND05]. As an example from live cell imaging, deriving cell lineages from sequences of images relies on accurate detection of cellular events. This is a complex and error-prone undertaking, which often requires manual intervention [KHC*07]. Well-designed visual curation tools have the potential to improve the effectiveness and efficiency of such approaches, which are currently cumbersome. By providing examples of how this can be achieved, Wait et al. and Winter et al. make a good case for further work in this direction [WWB*14, WWR*11].

5.5. Interactive support

Visual representations communicate meaning to users and this relay can be optimised by providing a feedback loop from the user back to the visualisation [War04]. Consequently, a central tenet of visualisation research is to enable users to become active participants in analysis through interaction with, and real-time updates of, visualisations [CMS99].

Many biologists visualise their data using standard image viewers or movie players [PM07], volume visualisation software [GME03], or plotting tools [SMC*06]. This implies that to select data subsets or different data attributes, new visualisations have to be generated from scratch, which curtails interactive exploration. Some approaches, like ProgeniTRAK [KHC*07], event-order maps [WHN*09], and protein-localisation networks [BGS13], offer limited interaction capabilities and in this way start supporting rudimentary interrogative analysis.

In contrast, our review of papers that describe fully interactive tools like Cell-o-pane [PKE15], CellProfiler Analyst [JKW*08], HitSEE [BSB*11, SBB*12], LEVER [WWR*11], and LEVER-3D [WWB*14], show that these support complex explorative analysis workflows across multiple coordinated views. They generate new insights that would be very challenging without interactive visualisation. This suggests that properly designed interaction support has the potential to dramatically improve the effectiveness and efficiency of visualisation for live cell imaging. Yet, many papers aimed at biomedical researchers do not thoroughly discuss interaction (for example, [ARG*06, GLHR09, GBBS09, HLLK09, MWV*03, SHT*12]). This highlights a missed opportunity for the biology research community to reflect more on the very important role that interaction can play.

From a visualisation perspective, there is also a case to be made for critically reevaluating conventional interaction strategies for live cell imaging, and high-throughput screening, in general. In particular, the point raised earlier about reconsidering entrenched interaction strategies also applies here (see Section 5.2, above).

5.6. Presentation and narration

Our analysis revealed that a major objective of visualisation for live cell imaging, in addition to supporting analysis, is to present and communicate findings to wider audiences. This includes the transfer of data and knowledge between peers, communication across disciplinary boundaries, and the generation of diverse research outputs. In fact, the output from nearly every case considered have been visualisations to be included in publications, reports, and supplementary material. Moreover, and related to the discussion in Section 5.1, the results that users wish to present visually are just as likely to originate from visual analysis as from non-visual investigation (for example, quantitative analysis, statistical tests, or machine learning).

The shift from a focus purely on analysis (the visualisation research community's traditional preoccupation) to one that also includes explicit communication raises important questions about the presentation of data and about fitting visualisations into cohesive narratives [KM13]. For a start, it makes sense to consider the consumers of such communicative visualisations as users in their own right. This has implications for the design of visualisation methods. The most suitable representation for analysis is not necessarily the most suitable for communicating results where, for example, interactivity is usually absent [MLF*12]. Possible solutions to the challenge of effectively presenting findings graphically may be addressed by publishing visualisations as scripted walkthroughs [HS12], but then raises questions about standard protocols for reporting results.

In general, the relationship between visualisation for analysis versus visualisation for presentation and narration is under-researched. Visualisation for live cell imaging, in particular, would benefit greatly from a better understanding of this issue.

6. Conclusion

Live cell imaging is an important emerging paradigm for biomedical research where automated high-content or high-throughput experiments, image capture, and image processing are routinely used to produce rich phenotypic data sets. The goals of achieving best practice and of implementing standards for data analysis through a multi-disciplinary endeavour, including interactive data visualisation, suggests an ambitious future for this area.

Visualisation can play an important role in the interrogation of phenotypic data derived from live cell imaging, but results have been reported in an ad hoc and fragmented fashion. Consequently, there is a knowledge gap between the biology and visualisation research communities. We have argued that it is critical to address this gap. First, this will enable biologists and visualisation developers to evaluate the advantages and disadvantages of different visualisation methods that may be applicable to specific users, objectives, and data. Second, it will allow visualisation researchers to obtain an overview of existing work, identify shortcomings in current research results, and prioritise their efforts to address these.

We have addressed the knowledge gap by surveying how visualisation is used in a live cell imaging context. Using recent theoretical frameworks and typologies, we analysed current approaches toward visualisation for live cell imaging. We identified six classes of approach: spatial embedding, space-time cubes, temporal plots, aggregate visualisations, dimension reduction, and lineage diagrams. Based on our analysis, we also identified six priorities for further work aimed at visualisation for live cell imaging: the broad analytical context of analysis; the recurring importance of behavioural comparisons; links with dynamic data visualisation; the consequences of different data modalities and scale, including managing uncertainty and curation; current shortcomings of interactive support; and the significance of the presentation and narration of results.

Our analysis suggests that, by focusing on these aspects, visualisation designers from both the visualisation and biology communities will be able to design more effective and efficient visualisation methods for data derived from live cell imaging. Moreover, for members of the visualisation research community, the work presented in this paper should serve as a valuable domain characterisation and a critical overview of existing approaches toward visualisation for live cell imaging, many of which were reported in biology research outlets.

Finally, we argue that work such as that presented here is important to enable systematic and theory-based reflection

on visualisation research and to guide future work. This will become increasingly important as more visualisation development occurs outside the confines of the visualisation research community.

Acknowledgements

We thank: Mark Bray and Anne Carpenter (Broad Institute of MIT and Harvard), Roy Ruddle (University of Leeds), and Nick White (Cardiff University) for valuable input; the authors of several cited papers for permission to reproduce their figures (in particular, Anne Carpenter, Jens Fangerau, Nico Scherf, and Maya Schuldiner for generous assistance); and the anonymous reviewers for their constructive feedback on an earlier draft of this paper. Pretorius was supported by a Leverhulme Early Career Fellowship (ECF2012-071), and Khan and Errington were supported by a BBSRC Sparking Impact Award (BB/E012574/1).

References

[AA13] ANDRIENKO N., ANDRIENKO G.: Visual analytics of movement: an overview of methods, tools and procedures. *Information Visualization 12*, 1 (2013), 3–24. 15

[AES05] AMAR R., EAGAN J., STASKO J.: Low-level components of analytic activity in information visualization. In *Proceedings of the IEEE Symposium on Information Visualization* (2005), pp. 111–117. 3

[AMM*07] AIGNER W., MIKSCH S., MÜLLER W., SCHUMANN H., TOMINSKI C.: Visualizing time-oriented data–a systematic view. *Computers & Graphics 31*, 3 (2007), 401–409. 15

[AMST11] AIGNER W., MIKSCH S., SCHUMANN S., TOMINSKI C.: *Visualization of Time-Oriented Data*. Springer Verlag, 2011. 15

[ARG*06] AL-KOFAHI O., RADKE R. J., GODERIE S. K., SHEN Q., TEMPLE S., ROYSAM B.: Automated cell lineage construction: a rapid method to analyze clonal development established with murine neural progenitor cells. *Cell Cycle 5*, 3 (2006), 327–335. 2, 4, 6, 7, 8, 11, 14, 16

[AS05] AMAR R. A., STASKO J. T.: Knowledge precepts for design and evaluation of information visualizations. *IEEE Transactions on Visualization and Computer Graphics 11*, 4 (2005), 432–442. 15

[Bar03] BARD J.: Ontologies: formalising biological knowledge for bioinformatics. *BioEssays 25*, 5 (2003), 501–506. 14

[BDA*14] BACH B., DRAGICEVIC P., ARCHAMBAULT D., HURTER C., CARPENDALE S.: A review of temporal data visualizations based on space-time cube operations. In *Proceedigns of the EG/VGTC Conference on Visualisation* (2014), pp. 23–41. 15

[Bec] BECTON, DICKINSON AND COMPANY: Company website. http://www.bd.com. Last accessed 15 July 2015. 2

[BFHC12] BRAY M. A., FRASER A. N., HASAKA T. P., CARPENTER A. E.: Workflow and metrics for image quality control in large-scale high-content screens. *Journal of Biomolecular Screening 17*, 2 (2012), 266–274. 15

[BGS13] BREKER M., GYMREK M., SCHULDINER M.: A novel single-cell screening platform reveals proteome plasticity during yeast stress responses. *The Journal of Cell Biology 200*, 6 (2013), 839–850. 4, 5, 6, 7, 8, 10, 11, 14, 15

[BM13] BREHMER M., MUNZNER T.: A multi-level typology of abstract visualization tasks. *IEEE Transactions on Visualization and Computer Graphics 19*, 12 (2013), 2376–2385. 3, 7, 14

[BSB*11] BERTINI E., STROBELT H., BRAUN J., DEUSSEN O., GROTH U., MAYER T. U., MERHOF D.: HiTSEE: A visualization tool for hit selection and analysis in high-throughput screening experiments. In *Proceedings of the IEEE Symposium on Biological Data Visualization* (2011), pp. 95–102. 4, 6, 7, 8, 11, 14, 16

[Car07] CARPENTER A. E.: Image-based chemical screening. *Nature Chemical Biology 3*, 8 (2007), 461–465. 1, 2, 3, 7

[CBI*08] CEDILNIK A., BAUMES J., IBANEZ L., MEGASON S., WYLIE B.: Integration of information and volume visualization for analysis of cell lineage and gene expression during embryogenesis. In *Proceedings of SPIE Visualization and Data Analysis* (2008). 4, 6, 7, 8, 9, 12, 13, 14, 15

[CJL*06] CARPENTER A. E., JONES T. R., LAMPRECHT M. R., CLARKE C., KANG I. H., FRIMAN O., GUERTIN D. A., CHANG J. H., LINDQUIST R. A., MOFFAT J., GOLLAND P., SABATINI D. M.: CellProfiler: image analysis software for identifying and quantifying cell phenotypes. *Genome Biology 7* (2006), R100. 2, 4, 5, 6, 7, 8, 10, 14

[CMS99] CARD S. K., MACKINLAY J. D., SHNEIDERMAN B. (Eds.): *Readings in Information Visualization: Using Vision to Think*. Morgan Kaufmann, San Francisco, CA, USA, 1999. 12, 15

[CT10] CRIVAT G., TARASKA J. W.: Imaging proteins inside cells with fluorescent tags. *Trends in Biotechnology 30*, 1 (2010), 8–16. 2

[DTW*15] DUFFY B., THIYAGALINGAM J., WALTON S., SMITH D., TREFETHEN A., KIRKMAN-BROWN J., GAFFNEY E., CHEN M.: Glyph-based video visualization for semen analysis. *IEEE Transactions on Visualization and Computer Graphics 21*, 8 (2015), 980–993. 4, 5, 6, 7, 8, 10, 11, 14, 15

[ENS09] EILKEN H. M., NISHIKAWA S., SCHROEDER T.: Continuous single-cell imaging of blood generation from haemogenic endothelium. *Nature 457*, 7231 (2009), 896–900. 4, 6, 7, 8, 11, 12, 14

[FHWL12] FANGERAU J., HOCKENDORF B., WITTBRODT J., LEITTE H.: Similarity analysis of cell movements in video microscopy. In *Proceedings of the IEEE Symposium on Biological Data Visualization* (2012), pp. 69–76. 4, 6, 7, 8, 9, 12, 13, 14, 15

[Fod02] FODOR I. K.: *A survey of dimension reduction techniques*. Tech. Rep. UCRL-ID-148494, Lawrence Livermore National Laboratory, 2002. 11

[GAW*11] GLEICHER M., ALBERS D., WALKER R., JUSUFI I., HANSEN C. D., ROBERTS J. C.: Visual comparison for information visualization. *Information Visualization 10*, 4 (2011), 289–309. 14

[GBBS09] GORDON J. L., BUGULISKIS J. S., BUSKE P. J., SIBLEY L. D.: Actin-like protein 1 (ALP1) is a component of dynamic, high molecular weight complexes in toxoplasma gondii. *Cytoskeleton 67*, 1 (2009), 23–31. 4, 6, 7, 9, 14, 16

[GE] GE HEALTHCARE LIFE SCIENCES: Company website. http://www.gelifesciences.com. Last accessed 15 July 2015. 2

[GLHR09] GLAUCHE I., LORENZ R., HASENCLEVER D., ROEDER I.: A novel view on stem cell development: analysing the shape of cellular genealogies. *Cell Proliferation 42*, 2 (2009), 248–263. 2, 4, 5, 6, 7, 8, 9, 11, 12, 14, 16

[GME03] GERLICH D., MATTES J., EILS R.: Quantitative motion analysis and visualization of cellular structures. *Methods 29*, 1 (2003), 3–13. 4, 7, 14, 15

[HGO10] HOULE D., GOVINDARAJU D. R., OMHOLT S.: Phenomics: the next challenge. *Nature Reviews Genetics 11*, 12 (2010), 855–866. 1

[HLLK09] HSU Y. Y., LIU Y. N., LU W. W., KUNG S. H.: Visualizing and quantifying the differential cleavages of the eukaryotic translation initiation factors eIF4GI and eIF4GII in the enterovirus-infected cell. *Biotechnology and Bioengineering 104*, 6 (2009), 1142–1152. 4, 6, 7, 8, 9, 13, 14, 16

[HS12] HEER J., SHNEIDERMAN B.: Interactive dynamics for visual analysis. *Communications of the ACM 55*, 4 (2012), 45–54. 13, 16

[HSF*10] HELD M., SCHMITZ M. H. A., FISCHER B., WALTER T., NEUMANN B., OLMA M. H., PETER M., ELLENBERG J., GERLICH D. W.: CellCognition: time-resolved phenotype annotation in high-throughput live cell imaging. *Nature Methods 7*, 9 (2010), 747–754. 1

[HWKT09] HAMILTON N. A., WANG J. T. H., KERR M. C., TEASDALE R. D.: Statistical and visual differentiation of subcellular imaging. *BMC Bioinformatics 10* (2009), 94. 4, 5, 6, 8, 11, 12, 14

[IJS*07] INGLESE J., JOHNSON R. L., SIMEONOV A., XIA M., ZHENG W., AUSTIN C. P., AULD D. S.: High-throughput screening assays for the identification of chemical probes. *Nature Chemical Biology 3* (2007), 466–479. 2, 3

[Ins85] INSELBERG A.: The plane with parallel coordinates. *The Visual Computer 1*, 2 (1985), 69–91. 10

[IZCC08] ISENBERG P., ZUK T., COLLINS C., CARPENDALE S.: Grounded evaluation of information visualizations. In *Proceedings of the Workshop on BEyond Time and Errors: Novel evaLuation Methods for Information Visualization* (2008), pp. 6:1–6:8. 3, 4

[Jen13] JENSEN E. C.: Overview of live-cell imaging: requirements and methods used. *The Anatomical Record 296*, 1 (2013), 1–8. 1, 2

[JKW*08] JONES T. R., KANG I. H., WHEELER D. B., LINDQUIST R. A., PAPALLO A., SABATINI D. M., GOLLAND P., CARPENTER A. E.: CellProfiler Analyst: data exploration and analysis software for complex image-based screens. *BMC Bioinformatics 9* (2008), 482. 4, 5, 6, 7, 8, 10, 11, 14, 16

[JLM*08] JAQAMAN K., LOERKE D., METTLEN M., KUWATA H., GRINSTEIN S., SCHMID S. L., DANUSER G.: Robust single-particle tracking in live-cell time-lapse sequences. *Nature Methods 5* (2008), 695–702. 2

[KHC*07] KHAN I. A., HUSEMANN P., CAMPBELL L., WHITE N. S., WHITE R. J., SMITH P. J., ERRINGTON R. J.: ProgeniDB: a novel cell lineage database for generation associated phenotypic behavior in cell-based assays. *Cell Cycle 6*, 7 (2007), 868–874. 3, 4, 5, 6, 7, 8, 9, 11, 12, 14, 15

[KKEM10] KEIM D., KOHLHAMMER J., ELLIS G., MANSMANN F.: *Mastering the Information Age: Solving Problems with Visual Analytics*. VisMaster, 2010. 14

[KLC*11] KHAN I. A., LUPI M., CAMPBELL L., CHAPPELL S. C., BROWN M. R., WILTSHIRE M., SMITH P. J., UBEZIO P., ERRINGTON R. J.: Interoperability of time series cytometric data: a cross platform approach for modeling tumor heterogeneity. *Cytometry A 79*, 3 (2011), 214–226. 4, 5, 6, 7, 8, 9, 10, 14

[KM13] KOSARA R., MACKINLAY J.: Storytelling: the next step for visualization. *Computer 46*, 5 (2013), 44–50. 16

[KSWS08] KELLER P. J., SCHMIDT A. D., WITTBRODT J., STELZER E. H.: Reconstruction of zebrafish early embryonic development by scanned light sheet microscopy. *Science 322*, 5904 (2008), 1065–1069. 2

[LC05] LICHTMAN J. W., CONCHELLO J. A.: Fluorescence microscopy. *Nature Methods 2*, 12 (2005), 910–919. 2

[Lei] LEICA MICROSYSTEMS: Company website. http://www.leica-microsystems.com. Last accessed 15 July 2015. 1

[LS10] LIU Z., STASKO J.: Mental models, visual reasoning and interaction in information visualization: a top-down perspective. *IEEE Transactions on Visualization and Computer Graphics 16*, 6 (2010), 999–1008. 3

[Mey13] MEYER M.: Designing visualizations for biological data. *Leonardo 46*, 3 (2013), 270–271. 3, 14

[Mit] MITOCHECK: MitoCheck project website. http://www.mitocheck.org. Last accessed 15 July 2015. 1

[MLF*12] MA K. L., LIAO I., FRAZIER J., HAUSER H., KOSTIS H. N.: Scientific storytelling using visualization. *IEEE Computer Graphics and Applications 32*, 1 (2012), 12–19. 16

[Mol] MOLECULAR DEVICES: Company website. http://www.moleculardevices.com. Last accessed 15 July 2015. 2

[MSD06] MEIJERING E., SMAL I., DANUSER G.: Tracking in molecular bioimaging. *IEEE Signal Processing Magazine 23*, 3 (2006), 46–53. 4, 5, 6, 7, 8, 9, 14

[MSQM13] MEYER M., SEDLMAIR M., QUINAN P. S., MUNZNER T.: The nested blocks and guidelines model. *Information Visualization* (2013). 3

[Mun09] MUNZNER T.: A nested model for visualization design and validation. *IEEE Transactions on Visualization and Computer Graphics 15*, 6 (2009), 1077–2626. 3, 9, 14

[MWV*03] MOLENAAR C., WIESMEIJER K., VERWOERD N. P., KHAZEN S., EILS R., TANKE H. J., DIRKS R. W.: Visualizing telomere dynamics in living mammalian cells using pna probes. *EMBO Journal 22*, 24 (2003), 6631–6641. 4, 5, 6, 7, 8, 9, 14, 16

[Nik] NIKON INSTRUMENTS: Company website. http://www.nikoninstruments.com. Last accessed 15 July 2015. 1

[Nor06] NORTH C.: Toward measuring visualization insight. *IEEE Computer Graphics and Applications 26*, 3 (2006), 6–9. 12

[NWH*10] NEUMANN B., WALTER T., HÉRICHÉ J. K., BULKESCHER J., ERFLE H., CONRAD C., ROGERS P., POSER I., HELD M., LIEBEL U., CETIN C., SIECKMANN F., PAU G., KABBE R., WÜNSCHE A., SATAGOPAM V., SCHMITZ M. H. A., CHAPUIS C., GERLICH D. W., SCHNEIDER R., EILS R., HUBER W., PETERS J. M., HYMAN A. A., DURBIN R., PEPPERKOK R., ELLENBERG J.: Phenotypic profiling of the human genome by time-lapse microscopy reveals cell division genes. *Nature 464* (2010), 721–727. 1

[Oly] OLYMPUS CORPORATION LIFE SCIENCES: Company website. http://www.olympus-lifescience.com/. Last accessed 15 July 2015. 1

[Per] PERKINELMER: Company website. http://www.perkinelmer.com. Last accessed 15 July 2015. 2

[Pha] PHASE FOCUS: Company website. http://www.phasefocus.com. Last accessed 15 July 2015. 2

[PKE15] PRETORIUS A. J., KHAN I. A., ERRINGTON R. J.: Cell lineage visualisation. *Computer Graphics Forum 34*, 3 (2015), 22–33. 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16

[PM07] PRASAD M., MONTELL D. J.: Cellular and molecular mechanisms of border cell migration analyzed using time-lapse live-cell imaging. *Developmental Cell 2*, 6 (2007), 997–1005. 4, 6, 7, 9, 14, 15

[PvW09] PRETORIUS A. J., VAN WIJK J. J.: What does the user want to see? What do the data want to be? *Information Visualization 8*, 3 (2009), 153–166. 9, 15

[RFF*08] ROBERTSON G., FERNANDEZ R., FISHER D., LEE B., STASKO J.: Effectiveness of animation in trend visualization. *IEEE Transactions on Visualization and Computer Graphics 14*, 6 (2008), 1325–1332. 14

[Rot13] ROTH R. E.: An empirically-derived taxonomy of interaction primitives for interactive cartography and geovisualization. *IEEE Transactions on Visualization and Computer Graphics 19*, 2 (2013), 2356–2365. 3

[SBB*12] STROBELT H., BERTINI E., BRAUN J., DEUSSEN O., GROTH U., MAYER T. U., MERHOF D.: HiTSEE KNIME: a visualization tool for hit selection and analysis in high-throughput screening experiments for the KNIME platform. *BMC Bioinformatics 13*, S-8 (2012), S4. 4, 6, 7, 8, 11, 12, 14, 16

[Shn96] SHNEIDERMAN B.: The eyes have it: a task by data type taxonomy for information visualizations. In *Proceedings of the IEEE Symposium on Visual Languages* (1996), pp. 336–343. 14

[SHT*12] SCHERF N., HERBERG M., THIERBACH K., ZERJATKE T., KALKAN T., HUMPHREYS P., SMITH A., GLAUCHE I., ROEDER I.: Imaging, quantification and visualization of spatio-temporal patterning in mESC colonies under different culture conditions. *Bioinformatics 28*, 18 (2012), i556–i561. 4, 5, 6, 7, 8, 9, 10, 11, 14, 15, 16

[SLSR10] SKEELS M., LEE B., SMITH G., ROBERTSON G.: Revealing uncertainty for information visualization. In *Proceedings of the Working Conference on Advanced Visual Interfaces* (2010), pp. 376–379. 15

[SMC*06] SIGAL A., MILO R., COHEN A., GEVA-ZATORSKY N., KLEIN Y., ALALUF I., SWERDLIN N., PERZOV N., DANON T., LIRON Y., RAVEH T., CARPENTER A. E., LAHAV G., ALON U.: Dynamic proteomics in individual human cells uncovers widespread cell-cycle dependence of nuclear proteins. *Nature Methods 3*, 7 (2006), 525–531. 4, 5, 6, 7, 8, 9, 13, 14, 15

[SMKM10] SANTIAGO-MOZOS R., KHAN I. A., MADDEN M. G.: Revealing the origin and nature of drug resistance of dynamic tumour systems. *International Journal of Knowledge Discovery in Bioinformatics 1*, 4 (2010), 26–53. 4, 6, 7, 8, 11, 14

[SMM12] SEDLMAIR M., MEYER M. D., , MUNZNER T.: Design study methodology: reflections from the trenches and the stacks. *IEEE Transactions on Visualisation and Computer Graphics 18*, 12 (2012), 2431–2440. 3, 14

[SND05] SARAIYA P., NORTH C., DUCA K.: Visualizing biological pathways: requirements analysis, systems evaluation and research agenda. *Information Visualization 4*, 3 (2005), 191–205. 15

[SNHS13] SCHULZ H. J., NOCKE T., HEITZLER M., SCHUMANN H.: A design space of visualization tasks. *IEEE Transactions on Visualization and Computer Graphics 19*, 12 (2013), 2366–2375. 3

[Son] SONY CORPORATION LIFE SCIENCES: Company website. http://www.sony.net/Products/LifeScience/. Last accessed 15 July 2015. 2

[SPG05] SIMMHAN Y. L., PLALE B., GANNON D.: A survey of data provenance in e-science. *ACM SIGMOD Record 34*, 3 (2005), 31–36. 13

[Sym] SYMPOSIUM ON BIOLOGICAL DATA VISUALIZATION: Symposium website. http://www.biovis.net. Last accessed 20 April 2015. 14

[Sys] SYSTEMS MICROSCOPY NETWORK OF EXCELLENCE: Project website. http://www.systemsmicroscopy.eu. Last accessed 15 July 2015. 1

[TBM*99] TVARUSKÓ W., BENTELE M., MISTELI T., RUDOLF R., KAETHER C., SPECTOR D. L., GERDES H. H., EILS R.: Time-resolved analysis and visualization of dynamic processes in living cells. *PNAS Journal 96*, 14 (1999), 7950–7955. 4, 5, 6, 7, 8, 9, 10, 14

[Tes] TESCAN: Company website. http://www.tescan.com. Last accessed 15 July 2015. 2

[The00] THE GENE ONTOLOGY CONSORTIUM: Gene ontology: tool for the unification of biology. *Nature Genetics 25*, 1 (2000), 25–29. 14

[TMB02] TVERSKY B., MORRISON J. B., BETRANCOURT M.: Animation: can it facilitate? *International Journal of Human-Computer Studies 57*, 4 (2002), 247–262. 15

[VLH*06] VAISBERG E. A., LENZI D., HANSEN R. L., KEON B. H., FINER J. T.: An infrastructure for high-throughput microscopy: instrumentation, informatics, and integration. *Methods in Enzymology 414* (2006), 484–512. 2, 3

[War01] WARE C.: Designing with a 2 1/2-D attitude. *Information Design Journal 10*, 3 (2001), 258–265. 14

[War04] WARE C.: *Information Visualization: Perception for Design*. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 2004. 15

[WHN*09] WALTER T., HELD M., NEUMANN B., HÉRICHÉ J. K., CONRAD C., PEPPERKOK R., ELLENBERG J.: Automatic identification and clustering of chromosome phenotypes in a genome wide RNAi screen by time-lapse imaging. *Journal of Structural Biology 170*, 1 (2009), 1–9. 4, 5, 6, 7, 8, 10, 14, 15

[WS07] WOLLMAN R., STUURMAN N.: High throughput microscopy: from raw images to discoveries. *Journal of Cell Science 120*, 21 (2007), 3715–3722. 1, 2, 3

[WSB*10] WALTER T., SHATTUCK D. W., BALDOCK R., BASTIN M. E., CARPENTER A. E., DUCE S., ELLENBERG J., FRASER A., HAMILTON N., PIEPER S., RAGAN M. A., SCHNEIDER J. E., TOMANCAK P., HÉRICHÉ J. K.: Visualization of image data from cells to organisms. *Nature Methods 7* (2010), S26–S41. 2, 13, 14

[WWB*14] WAIT E., WINTER M., BJORNSSON C., KOKOVAY E., WANG Y., GODERIE S., TEMPLE S., COHEN A. R.: Visualization and correction of automated segmentation, tracking and lineaging from 5-d stem cell image sequences. *BMC Bioinformatics 3*, 15 (2014), 328. 4, 6, 7, 8, 12, 14, 15, 16

[WWR*11] WINTER M., WAIT E., ROYSAM B., GODERIE S. K., ALI R. A. N., KOKOVAY E., TEMPLE S., COHEN A. R.: Vertebrate neural stem cell segmentation, tracking and lineaging with validation and editing. *Nature Protocols 6* (2011), 1942–1952. 4, 6, 7, 8, 12, 14, 15, 16

[XFP*15] XU Z., FALLET E., PAOLETTI C., FEHRMANN S., CHARVIN G., TEIXEIRA M. T.: Two routes to senescence revealed by real-time analysis of telomerase-negative single lineages. *Nature Communications 6*, 7680 (2015). 1

[Zei] ZEISS: Company website. http://www.zeiss.com/microscopy/. Last accessed 15 July 2015. 1