

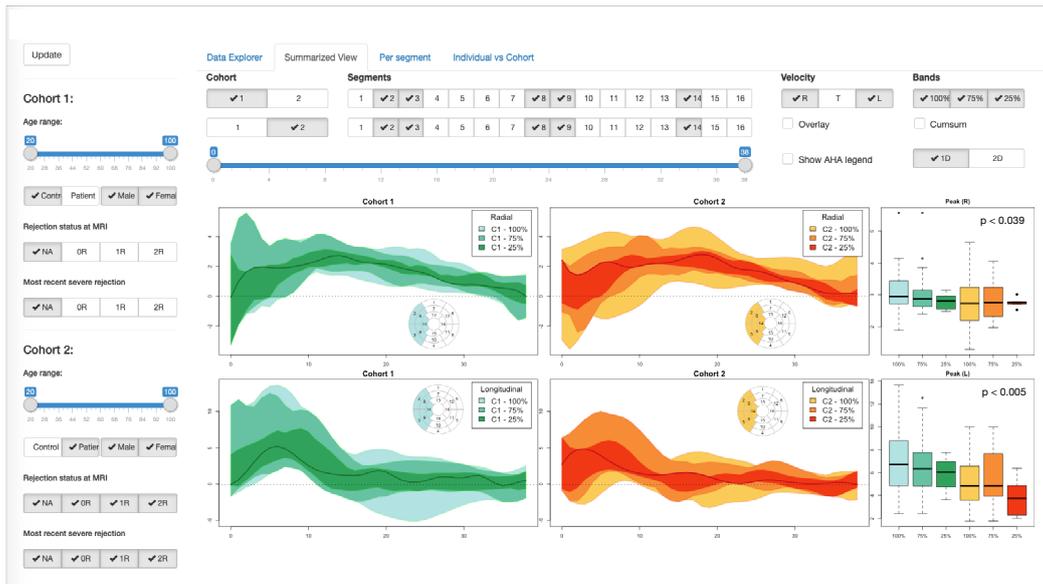
# A Visual Analytics Tool for Cohorts in Motion Data

Ali Sheharyar<sup>1,2</sup>, Alexander Ruh<sup>3</sup>, Dimitar Valkov<sup>2</sup>, Michael Markl<sup>3</sup>, Othmane Bouhali<sup>1</sup>, Lars Linsen<sup>2</sup>

<sup>1</sup>Texas A&M University at Qatar, Doha, Qatar,

<sup>2</sup>Westfälische Wilhelms-Universität Münster, Germany

<sup>3</sup>Feinberg School of Medicine, Northwestern University, Chicago, IL, USA



**Figure 1:** Comparison of radial (top) and longitudinal (bottom) velocities in all healthy volunteers (green-blue) and all patients (red-yellow). The panel to the left is used to define the cohorts, the one on top to set the visualization parameters. Differences in longitudinal velocities among the cohorts can be observed.

## Abstract

Motion data are curves over time in a 1D, 2D, or 3D space. To analyze sets of curves, machine learning methods can be applied to cluster them and detect outliers. However, often metadata or prior knowledge of the analyst drives the analysis by defining cohorts. Our goal is to provide a flexible system for comparative visual analytics of cohorts in motion data. The analyst interactively defines cohorts by filtering on metadata properties. We, then, apply machine learning and statistical methods to extract the main features of each cohort. Summarizations of these features are visually encoded using, in particular, boxplots and their extensions to functional and curve boxplots, depending on the number of selected dimensions of the space. These summarizations allow for an intuitive comparative visual analysis of cohorts in a juxtaposed or superimposed representation. Our system provides full flexibility in defining cohorts, selecting time intervals and spatial dimensions, and adjusting the aggregation level of summarizations. Comparison of an individual sample against a cohort is also supported. We demonstrate the functionality, effectiveness, and flexibility of our system by applying it to a range of diverse motion data sets.

## 1. Introduction

Motion data comprise information of positions that change over time (or their temporal derivatives) within a spatial embedding, where the spatial embedding may have up to three dimensions. According to this general, broad definition, motion data include a wide range of applications ranging from trajectories in a 2D geographical space (e.g., tracked by using GPS sensors) or 3D gestures of humans (e.g., measured by motion capturing) to deformations of

tissues (e.g., measured by using medical imaging), to just name a few prominent ones. Analytics tasks for motion data are common pattern recognition tasks such as detecting clusters and outliers or examining variability and trends. Thus, machine learning methods can be applied for pattern recognition.

Pattern recognition approaches typically treat the motion data samples as an (unorganized) set. However, in many applications, there is more knowledge available, which may be captured in the

form of metadata for each motion data sample or which may lie within the expertise of the analyst. This is most obvious in the context of medical or electronic health record data, where much additional information about the recorded persons is known including knowledge about the existence of a certain disease that is being examined. Thus, the analyst can form cohorts of motion data samples based on the metadata. The analytics tasks are then primarily concerned with comparing the motion data of two cohorts.

In this paper, we present a tool that allows for the analysis of motion data cohorts. One main strength of our system is that analysts can quickly and flexibly form any cohort for analysis and switch between any settings interactively. It is common practice in many application fields to manually define cohorts and feed them to a statistics program, where switching settings is rather cumbersome. Our system, instead, allows for quick tests of all kinds of settings within one framework. Another main strength is the visual representation of cohorts in the form of summarizations at different levels of aggregation, which allows for an intuitive comparison between cohorts at appropriate and adjustable levels of detail. Machine learning methods are embedded where useful and results are conveyed using visual representations within the analytical workflow, see Section 4.

For the design of our tool, we performed a requirement analysis with experts from different fields with different types of motion data, see Section 3. To demonstrate the functionality of our tool and its effectiveness, we apply it to those data sets, see Section 5. One driving application behind this work was the analysis of heart motion data. 3D motion data are captured using a magnetic resonance imaging method, from which the myocardium of the left ventricle is extracted, subdivided into spatial regions or segments, and tracked over time. The analysis goal is to detect spatio-temporal heart motion patterns from patients and healthy volunteers. Our system allows for analyzing cohorts formed by metadata such as disease records, age, and gender. Another driving application was the analysis of motions of virtual environment users. A warning system for collision of users with real-world objects when being immersed in a virtual environment was tested under different conditions. Respective conditions can be flexibly selected and combined to form cohorts, the motion of which are then compared with respect to the users' reaction to the warning system. Our system's capabilities and limitations are discussed in Section 7.

## 2. Background and Related Work

Motion data can formally be described as a set of curves  $\mathbf{m}_i(t) : \mathbb{R} \mapsto R$  parametrized over time  $t$ , where the range  $R$  can have up to three dimensions and represents the space, in which the curves are embedded. Typically, each curve is known at discrete time points  $t_0, \dots, t_n$  such that a curve is given by a sequence of time-space pairs  $\langle (t_0, \mathbf{m}_i(t_0)), \dots, (t_n, \mathbf{m}_i(t_n)) \rangle$ , where the time points may or may not be synchronized among all curves. Instead of positions, one may also consider their first and second derivative in the form of velocities and acceleration.

Time-oriented data exist in many types. An exhaustive overview is provided by Aigner et al. [AMST11]. The existing visualization approaches are classified with respect to the data representation

(frame of reference, variables), time representation (arrangement, primitives), and visualization approach (mapping, dimensionality). Motion data classifies as *spatial* in terms of the frame of reference and *univariate* in terms of variables (when neglecting metadata). Time arrangement may be *linear* or *cyclic* and time primitives are usually *instant*. The dimensionality of the visualization approach may be *2D* or *3D* depending on the range  $R$  of the motion data. A *static* mapping shall be preferred to not unnecessarily burden the cognitive load for the user. Aigner et al. also discuss clustering as an unsupervised machine learning method embedded into the visual analytics workflow, where clustering was based on time series similarity only.

The main motivation for this work is that motion data in many applications can be augmented with metadata, which may be explicitly stored as global variables for each motion data sample (i.e., curve) or may be implicitly given by the knowledge of the analyst. Such global connotations of the motion data samples are used by the analysts to form cohorts of samples with some interesting properties. Hence, the data at hand can be described as labeled data, where each cohort is represented by one label. However, such a static labeling does not allow for a flexible analysis.

All facets of visual analytics of movement are discussed in the book by Andrienko et al. [AAB\*13]. Their definition of movement data is close to our motion data definition, but they mainly considered trajectories in a geospatial context, which is one instance of what we want to consider. They discuss visual analytics based on space, time, and movers, but the analysis of cohorts built on metadata was beyond the scope of their book.

Andrienko et al. [AAB\*13] also discuss summarizations of trajectories. This topic relates to edge bundling in graphs [HvW09]. Edge bundling effectively reduces visual clutter in line-based drawings, when exact positions of edges may be slightly altered. In case this is not desired, one should rather use summarizations by drawing bands. Such bands rely on statistical properties. For example, the idea of boxplots can be generalized to functional data leading to *functional boxplots* [SG11] or even to curves moving freely in a 2D or 3D space and being parametrized over time leading to *curve boxplots* [MWK14]. These computations are based on band depth. Thus, one can compute the median and bands that include a given percentage of all curves, e.g., the equivalent to the original boxplots' interquartile range. Curve boxplots have been successfully applied to convey uncertainty in trajectories of tornadoes captured by simulation ensembles. Ferstl et al. [FKRW17] propose to use clustering of such trajectories in case they diverge and split into groups. Each cluster is then represented by an own band. Our tool adopts these concepts by displaying bands that relate to functional boxplots, in case of plotting motion over a time axis, and to curve boxplots, in case of plotting curves within spatial dimensions parametrized over time. We also couple the visual output to clustering results, among others.

Closest to our system are visual analytics approaches for electronic health/medical record data. Some approaches only address cross-sectional (i.e., non-temporal) data [AOH\*14, SG15], while others consider longitudinal studies [WLM\*14, DVGG\*17, BSM\*15], i.e., functional data over time, which is a special case of motion data (for a 1D range  $R$ ). The listed approaches for lon-

longitudinal studies do not allow for interactive flexible definition of cohorts from metadata and their comparative analysis with summarizations at different levels of aggregation, which is our main focus. Flexible cohort definition is supported by Zhang et al. [ZGP12], but they represented the temporal aspect of the longitudinal studies only as discrete events over time in a Sankey diagram, which is quite different from the motion data we are dealing with. Rogers et al. [RSN\*19] developed a visual analytics tool that allows dynamic definition of patient cohorts using demographic information and other parameters to analyze the patient-reported changes (outcome scores) in physical functions following various spinal procedures. They aggregated the scores to visualize the interquartile range with a line representing the median. This approach is closest to our work, but we are dealing with the more general case of motion data with multiple dimensions and diverse applications.

### 3. Requirement Analysis

We performed a requirement analysis for our system by conducting interviews with potential users having different backgrounds (medical imaging/radiology experts vs. human-computer interaction experts) addressing motion data of rather different types (motion extracted from dynamical medical imaging vs. tracked motion of humans). Seven experts from medical imaging and one expert from human-computer interaction participated in identifying the system requirements.

**User motions in immersive virtual environments (IVE).** For testing a new set-up or system for users operating in immersive virtual environments, human-computer interaction experts conduct controlled experiments, where the motion (e.g., walking behavior) of users under different test conditions is captured. During the interviews, main analytical tasks were identified as follows:

(V1) *Analyze all motions of a particular user under all test conditions.* Such an analysis allows for detecting users as outliers or strange behavior patterns. Such outliers may be due to users incorrectly executing the experiment. Individual users may exhibit generally a strange behavior or only under a group of test conditions.

(V2) *Analyze motion of a single user against all others for a specific condition.* Such an analysis will reveal experiments as outliers in particular conditions. Such outliers may be due to some measuring errors and, consequently, are subject to removal for further analysis steps.

(V3) *Compare motion of all users under different conditions.* Here, cohorts are formed by filtering experiments according to one or multiple experimental settings. In the latter case, one should be able to flexibly combine settings of any parameters. Such analyses allow for the investigation which (combination of) parameters influenced the motion behavior (most severely).

(V4) *Compare motion of different user groups.* Users may be grouped into different cohorts according to meta-data such as gender, age, level of expertise, visual deficiencies, etc., or combinations thereof. Analyses tasks are then to investigate whether there are any significant differences among those groups for each of the experimental conditions.

**Motion from medical imaging.** Dynamic medical imaging methods capture motion implicitly by differences in images over time. Segmentation and tracking algorithms need to be applied to detect

a tissue or organ of interest and track over time. The goal is to compare the motion patterns of patients and healthy people. Individual tasks can be described as:

(M1) *Analyze motion variability within a cohort.* Such an analysis is essential, e.g., to understand how much variation is normal within a cohort of healthy people. Similarly, the degree of variation among patients is of relevance.

(M2) *Compare motion of individual persons to cohort.* Such an analysis is core when trying to investigate, if the motion pattern of an individual person matches that of a cohort of healthy people or that of a cohort of patients with pathology.

(M3) *Detect spatio-temporal location of motion anomaly.* For patients with pathology it is of interest to detect where in the scanned tissue motion anomalies occur and when those anomalies occur within the scanned time interval.

(M4) *Analyze severeness of anomalies.* The goal is to quantify how much the motion pattern of a patient with pathology deviates from a cohort of healthy people.

(M5) *Detect common motion patterns for pathological cases.* If possible, one would try to detect common deviations in the motion pattern for patients against healthy people, i.e., one would compare the two cohorts against each other.

**System requirements:** When comparing the analysis requirements for the two scenarios, we observe a large overlap. Almost all requirements occur in similar form in both lists. When abstracting from the application scenario and assuming that a cohort can also exist of a single motion sample, we can formulate the following analysis tasks:

(T1) *Detect spatio-temporal motion patterns and outliers (V1,M3).*

(T2) *Analyze motion variability within cohort (M1).*

(T3) *Compare single motion sample to cohort (V2,M2).*

(T4) *Compare multiple cohorts to detect/quantify motion differences (V3,V4,M4,M5).*

## 4. Visual Analytics System

The analytical workflow within our system is based on an iterative scheme of two alternating steps. The first step is that of defining cohorts. Any knowledge about the data samples can be used to define cohorts. Commonly, one is interested in comparing two cohorts, which includes the special case of comparing a single data sample against a cohort. This comparison is performed in the second step when the defined cohorts are being analyzed. For the analysis, statistical methods, machine-learning approaches, and visual summarization methods are combined. Based on the analysis outcome, the cohorts can be altered in the next iteration, or completely new cohorts can be selected. Thus, many tests between cohorts can be quickly defined and analyzed in an interactive setting.

### 4.1. Defining Cohorts

Cohorts are groups of data samples that share a common characteristic. Cohort definition is often based on some hypotheses. Thus, the cohorts are defined with respect to differences in some properties with the expectation that the data samples belonging to different cohorts show differences in the observed data. A typical workflow is to group data samples manually and load the manually defined cohorts to some analysis tool. Thus, the overhead of creating

cohorts is rather high. Our approach is to allow the user to interactively define and refine cohorts and test hypotheses immediately for selected data attributes.

Often a hypothesis is tested, i.e., the analyst defines a control and a test cohort that differ in one tested variable. However, it is also common to test many variables and adjust thresholds. This flexible adjustment of cohorts is what our system supports. To support the analyst in forming cohorts in a more exploratory fashion, we provide summary statistics in the form of histograms and bar charts of the metadata to give an overview of the entire data set or of already selected cohorts. A spreadsheet explorer allows for seeing all data samples assigned to a cohort including all their metadata. A manual assignment of individual data samples to cohorts is also supported by our system.

In terms of user interaction, all cohort definitions and adjustments are made via a graphical user interface, see left panel in Figure 1. Here, we use different widgets depending on the type of the metadata variable. Numerical values are mapped to a two-ended slider widget, which allows for the selection of any interval within the given range. Categorical (and ordinal) values are mapped to a checkbox widget, where each category of the categorical variable is mapped to an own checkbox. Using checkboxes (rather than radio buttons) allows for grouping data samples from different categories into one cohort (e.g., merging patients with different symptoms for comparison against a healthy cohort).

#### 4.2. Visual Summarizations

Having defined the cohorts, the goal is to compare the observed motion data for the cohorts. For the comparison of cohorts, data of all data samples belonging to a cohort need to be aggregated and respective summarizations of the data need to be generated. To answer the comparative analysis tasks coined in Section 3, our system is required to allow for respective interactive queries, to perform respective data transformations, and to provide respective visual encodings. The interactive queries define the data aggregation level as well as what needs to be aggregated and how. The data transformations should extract the necessary information from the data. The visual encodings shall convey the extracted information in an effective and intuitive way.

**Interactive queries** define, which facets of the motion data shall be analyzed. To fulfill Task (T1), we need to define which time intervals shall be analyzed, which spatial dimensions shall be analyzed, and/or which curve characteristics shall be considered. Our graphical user interface allows for such queries. As time is a numerical variable, we follow the design decisions from above and use a two-ended slider to select time intervals. Motion data may have up to three spatial dimensions. We use a checkbox widget for each of the dimensions to select which dimensions shall be included in the current analysis. While it is often of interest to observe changes over time, sometimes specific features of the curves shall be extracted and compared. Such features are based on aggregation over time, where the aggregation function can be chosen. Currently, our system supports computations of mean, median, minimum, and maximum over all time steps. Other characteristics could be easily incorporated, if desired. Finally, motion data may be acquired in

different spatial regions and it may be of interest to compare spatial regions. While the granularity of the spatial regions is due to the data acquisition step, during the analysis one may define larger regions by spatial aggregation, which is supported by our graphical user interface by providing checkbox widgets. In summary, interactive queries are set, which, in combination with the cohort definition, lead to aggregation over data samples, time, and space as well as to selection of time intervals, spatial regions, and the motion's spatial dimensions.

**Data transformations** can be used to extract information from the raw motion data. To support Tasks (T1,T3,T4), we apply machine-learning methods. Unsupervised learning methods can compute clusters of motion data samples and detect outliers. To apply such methods, we have to define a similarity measure between samples. Assuming that the motion data samples are synchronized, i.e., they have the same starting point and the same temporal sampling, we can use a Euclidean distance metric to compute motion data similarities. Otherwise, a time-warping approach should be applied [Mül07]. Given the pairwise distances of the motion data samples, we can apply unsupervised learning. A hierarchical clustering approach is most suitable in this context, since it only requires distance values as input and is not assuming spatial sample distributions such as for spatial kernels or spatial binning. We decided to use the hierarchical clustering with average linkage, because average linkage is supposed to represent a natural compromise between the single and complete linkage measures and provide a more accurate evaluation of the distance between clusters [YR15]. The hierarchical clustering results in a binary tree, which is pruned using a dynamic branch cutting approach that identifies the clusters based on the tree shape instead of cutting at a fixed height [LZH07]. After pruning, the cluster tree splits into a forest, where each tree of the forest represents a cluster or an outlier. More precisely, if a tree consists of multiple data samples, it represents a cluster, if it only contains a single data sample, it represents an outlier. The hierarchical clustering outcome documents how homogeneous a cohort is within itself (e.g., to detect “false positives”). Next, we want to relate the cohort to data samples of the other cohort to examine how similar data samples of the other cohort are to the given cohort (e.g., to detect “false negatives”). Thus, we tentatively add each sample of the other cohort to this cohort and re-compute the hierarchical clustering. The outcome of this step tells us whether the sample of the other cohort would be an outlier within the given cohort or whether it would fit one of the clusters of the given cohort.

**Visual encodings** shall convey the data aggregation and data transformation results in an effective and intuitive way. We have to develop a visual encoding for each aggregation level. The simplest design choice is for the visualization of features aggregated over time: This is a single numerical value for each motion data sample. These numerical values need to be summarized for the entire cohort. Well-established methods exist: Statistical methods are used to compute the median, inter-quartile range, minimum-maximum span, and outliers of each cohort. These statistical values are commonly visually encoded using a boxplot representation. Boxplots support Task (T2) in conjunction with Task (T4). Figure 1 shows examples to the lower right. In addition to the visual representation

of the boxplot, we automatically compute p-values for statistical significance and display them with the boxplots.

Next, we want to consider the visual encoding of motion data over a single spatial dimension. When aggregating over the cohort, a summarization of curve plots over the time axis needs to be supported. Plotting all curves of the cohort does not scale well with the size of the cohort and would eventually lead to substantial visual clutter. A bundling strategy [HvW09] would reduce visual clutter, but would alter the curves, which is not acceptable in this context. Plotting only the mean or the median would to some extent address Task (T1), but it would clearly neither fulfill Task (T2) on conveying the variability (or uncertainty) nor Task (T4) on quantifying differences. Hence, we decided to render functional boxplots [SG11], which is a generalization of the common boxplot. It is computed using band depth and renders the median curve as well as nested bands representing envelopes of chosen percentages of all curves. For a comparative visualization of two cohorts, the boxplots of the two cohorts may be shown in a juxtaposed or superimposed fashion [GAW\*11]. Switching between the two concepts is easily possible. Moreover, we can choose which of the nested bands are shown. By default we use bands containing 25%, 75%, and 100% of the curves and no outlier visualization. Juxtaposed views allow for the rendering of more bands, while superimposed views get harder to interpret with increasing number of bands, cf. Figures 11 and 12.

When not restricting motion to 1D, but allowing for curves parametrized over time in a 2D space, functional plots are no longer applicable. However, we can apply the concept of curve boxplots [MWK14]. They assume synchronized time curves, which allows for a computation of band depths analogously to functional boxplots. The curves may exhibit any shape and may even be self-intersecting. Like functional boxplots, curve boxplots can be rendered juxtaposed or superimposed with adjustable selection of bands, see Figures 2 and 3. Curve plots also generalize to 3D curves, but the renderings impose occlusion and perception problems. Thus, we decided to not render 3D curve boxplots, but instead show different combinations of 2D boxplots in case of 3D curves, see Figure 2.

The outcome of an unsupervised learning step can be visually conveyed at multiple levels. To examine how homogeneous a cohort is (cf. Task (T2)), the unsupervised learning result is conveyed in the form of band plots. Each cluster is rendered in the form of an own functional/curve boxplot (using a different color for each boxplot). Outliers are shown as individual curves. Figure 6 shows an unsupervised learning result using functional boxplots. In case one wants to have an overview of several clustering results (e.g., for multiple spatial regions), we provide a more aggregated view by displaying the clustering result in a cluster heatmap. Each row represents one clustering result, where outliers are shown in red color and clusters by other colors. We decided to use a heat map, as it uses a space-efficient (space-filling) tabular structure and color coding is an intuitive and effective (pre-attentive) mapping. Thus, it scales sufficiently well to larger numbers of motion data samples (columns) and clustering results (rows). The accompanying video shows an example.

When testing how individual data samples of another cohort fit

the given cohort (cf. Task (T3)), we make the same (consistent) design choices. If we are examining one individual data sample of the other cohort (for one spatial region), we use the functional/curve boxplots, where the individual data sample is always shown as a curve. If this curve matches one of the clusters, it is shown in the respective color of that cluster. If it does not match any cluster, it is classified as an outlier and rendered using a red color. If multiple (or all) data samples of the other cohort are tested against the given cohort, an embedding of all those curves into the band visualization would lead to visual clutter. Instead, we apply, again, a heatmap encoding: We are interested in observing which data samples of the other cohort is an outlier for the given cohort and may perform the outlyingness test for one, two, or three motion dimensions. Figure 5 shows a respective cohort heatmap: The columns represent the motion data samples and the rows represent the spatial regions. The outlyingness is tested for all three motion directions. For each motion data sample, each spatial region, and each motion direction we count how often the data sample was on outlier. This count (0, 1, 2, or 3) is visually encoded by a 4-step color map, where darker colors represent higher counts. Hence, one can quickly obtain an overview of which data samples of the other cohort do not match the given cohort. Moreover, one can observe quickly for which spatial regions differences occur. To further ease the readability of the heatmap, we perform a sorting of rows and columns. We independently take all rows and all columns, and feed both individually to a hierarchical clustering algorithm. The clustering trees for rows and columns are depicted to the left and the top of the heatmap. Rows and columns are reordered according to the clustering outcome. Thus, data samples with similar outlyingness patterns over spatial regions are now shown in neighboring columns and spatial regions with similar outlyingness patterns over data samples are now shown in neighboring rows.

### 4.3. Analytical Workflow

In this section, we want to discuss the analytical workflow and the interplay of all components discussed above. After loading the data, the analyst defines the cohorts using the panel to the left, see Figure 4. Initial cohort definition is usually triggered by some hypothesis that the analyst has in mind. If not, the analyst can obtain an overview of the metadata distributions using histograms or bar-charts. After having defined the two cohorts that are to be tested against each other, the analyst selects which part of the motion data shall be analyzed using the visualization control panel on top. The analyst can select which motion dimensions to be investigated. If multiple are selected, he/she chooses whether to use multiple 1D function boxplots or 2D curve boxplots. If multiple plots are selected, they are shown in multiple rows in the visualization panel. Within each row, the analyst decides whether the plots of the two cohorts shall be shown juxtaposed or superimposed and may adjust which and how many bands are shown. The analyst may also decide to see features aggregated over time. Those are shown as boxplots in each row, where the colors of those boxplots match the colors of the bands, see Figure 4. With this setting, the analyst can quickly look at different spatial regions by selecting them individually in the visualization control panel or may select multiple spatial regions at the same time, which triggers an aggregation of those spatial regions used for all visualizations.

The analyst may adjust the selected cohorts at any time. To do so, the analyst may switch to the cluster heatmap (see accompanying video) to see whether there are some outliers in the cohort or whether the cohort can be split into multiple clusters. The analyst may decide to alter the cohorts accordingly. Moreover, the analyst may want to relate individual data samples of one cohort to the other using the cohort heatmap (as in Figure 5). Starting from these overviews, the analyst can drill down on selected data samples, selected time steps, selected spatial regions, or selected motion directions, i.e., transitioning to less and less aggregated views.

#### 4.4. Implementation

The tool has been implemented as a web application in the programming language R using the Shiny framework. It uses Shiny's reactive programming framework to compartmentalize and cache computationally expensive operations, effectively avoiding re-computations during the interactive session. It is fully cross-platform and can be deployed either locally or on a remote web server. The current implementation depends on several packages, including (i) *roahd* to compute the outliers and order statistics of the multivariate functional data; (ii) *gplots* for plotting an enhanced version of a heat map; (iii) *dtwclust* for hierarchical clustering and comparing time series, and (iv) *r-sp* along with (v) *r-rgdal* to implement the curve boxplot function. The functional and curve boxplotting functions are extension of the *fbplot* function from the *fa* R package.

#### 5. Application Scenarios

We apply our tool in scenarios for motion data of quite different domains, i.e., the ones used for the requirement analysis in Section 3. The accompanying videos show the usage of our tool in interactive sessions. To demonstrate the general applicability of our tool, the videos show a third scenario of analyzing motion capture data from 19 daily and sports activities [DKT17].

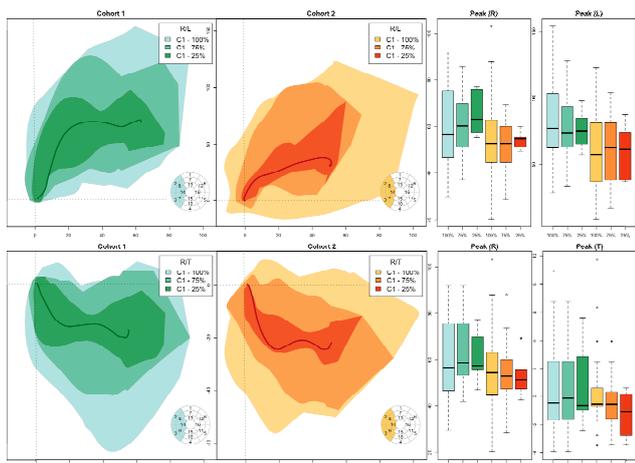
**Myocardial Motion.** Heart motion data are acquired using a magnetic resonance imaging (MRI) method known as tissue-phase mapping (TPM) [FJS\*09], which is a phase-contrast technique to measure 3D regional myocardial velocities. Data from 22 patients ( $50 \pm 13$  yrs, 14 male) [SRB\*18] and 18 healthy volunteers ( $49 \pm 14$  yrs, 12 male) [LCB\*16] were obtained, i.e., we have 40 data motion samples. All patients had gone through the heart transplant procedure within one year before the MRI scan. The MR images are segmented to extract the myocardium of the left ventricle (LV) for up to one cardiac cycle [LCB\*16]. The analysis of LV wall motion is crucial in understanding the cardiac function and diagnosis of ventricular diseases [YWT\*17]. The myocardium is further subdivided into spatial regions (referred to as segments) according to the guidelines from the American Heart Association (AHA) [CWD\*02]. The acquired velocities are aggregated over each segment and transformed to a cylindrical coordinate system reflecting radial, circumferential, and longitudinal motion. The motion data are synchronized for the systolic phase. Due to incomplete heart cycles for several data samples (and missing heart rate information), a synchronization of the diastolic phase was not possible such that we focus on the systole in this study.

For each person, metadata are provided including gender and age, whether they are patients or healthy volunteers, and in case of patients the number and severity of acute cellular rejection (ACR) episodes [SWF\*05]. The metadata are mapped to the respective widgets in the cohort definition panel, where age is considered the only numerical value. Figure 4 shows the instance of our tool for heart motion analysis. There are two specific extensions for this application scenario. First, the motion is given in the form of velocities not positions, i.e., vectors not points. To create motion curves in a 2D space, 2D points  $\mathbf{m}_i(t_j)$  are created by accumulating the velocities up to each time point  $t_j$ . Second, the medical experts are used to see the spatial regions (segments) laid out according to the AHA model. Our functional and curve boxplots are enhanced by insets of these radial plots, where we highlight, which of the segments were selected to create the plot, see Figure 4.

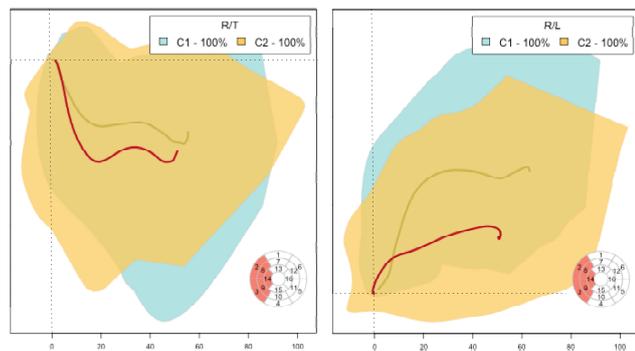
*Defining Cohorts:* As mentioned above, the initial definition of cohorts is triggered by some hypotheses or analysis tasks. In this scenario, the obvious hypothesis is that the cohort of patients should differ from the cohort of healthy volunteers (H1). However, patients may have different anomalies. In particular, one may hypothesize that cohorts of patients with different ACR rejection severity may differ from each other (H2). If H2 were true, then it may be worth testing individual patients against the healthy cohort, assuming that the individual patients differ from the cohort of healthy volunteers (H3). Moreover, there are other factors that may influence the analysis substantially. For example, one may hypothesize that cohorts of different age differ substantially (H4). Finally, it may also be interesting to compare regional heart motion in different segments, e.g., within healthy people, as there may also be substantial differences (H5).

*Results:* To test Hypothesis H1 we formed Cohort 1 of all healthy volunteers (green-blue) and Cohort 2 of all patients (red-yellow), see Figure 1 (left), and compared their motion data. For example, in Figure 1 (top), we aggregate motion data over all segments to the left of the AHA model to analyze radial (R) and longitudinal (L) velocities in juxtaposed functional boxplots using three nested bands. We also extract peak velocities of each velocity direction shown in the boxplots next to the functional boxplots. From the boxplots we can observe a difference in the longitudinal velocities, which is statistically significant, i.e., the hypothesis is confirmed. In the functional boxplots, we can observe when the peak velocities occur and how the shapes of the curves compare. Switching from 1D functional boxplots to 2D curve boxplots, the results for radial-longitudinal (R/L) velocities and radial-twisting (R/T) are shown in Figure 2. The lower longitudinal velocities (lower values in vertical axis) for patients can clearly be observed in the curve boxplots. To see the discrepancy in longitudinal velocities even better, the curve boxplots can also be shown superimposed as in Figure 3, where the number of rendered bands is reduced to one per cohort for less visual complexity.

To test how homogeneous the patient cohort is, we further subdivide this cohort in a cohort with moderate rejection status (1R or 2R) and no rejection status (0R), cf. Hypothesis H2. Figure 4 shows respective functional boxplots. We again encounter differences in the longitudinal velocities ( $p < 0.005$  for peak velocities).

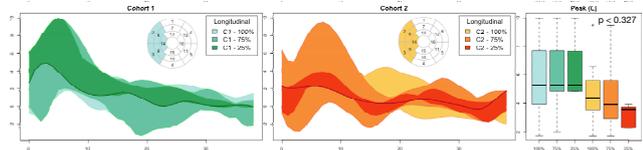


**Figure 2:** (left) Juxtaposed curve boxplots for 2D heart motion analysis of healthy volunteer cohort vs. patient cohort (aggregated over selected segments). (right) Boxplots of the respective peak velocities per velocity dimension exhibit significant differences in longitudinal velocities ( $p < 0.005$  for peak longitudinal velocities).



**Figure 3:** Superimposed curve boxplots for 2D heart motion analysis of healthy volunteer cohort vs. patient cohort (aggregated over selected segments). Differences in longitudinal direction (vertical axis in right plot) can be observed.

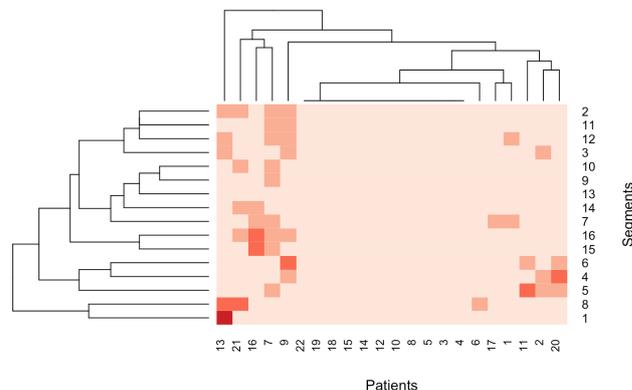
So, there are indeed differences within the patient cohort and the hypothesis can be confirmed.



**Figure 4:** Comparison of patients with moderate (red-yellow) and no (green-blue) rejection status exhibits difference in longitudinal velocities. Statistically significant peak velocities are observed.

With this observation, it is worth investigating how individual patients compare against the healthy cohort (Hypothesis H3). We take all 2D combinations of velocity dimensions (R/L, R/T, and L/T) and generate the respective curves for all healthy volunteers.

Then, we tentatively add each of the patient data sets and test whether they are outliers. The amount of outlyingness of all patients in all segments for all curve plots is summarized in the cohort heatmap shown in Figure 5. We observe that the block of eleven patients (with IDs 22, 19, 18, 15, 14, 12, 10, 8, 5, 3, and 4) do not have outlying velocities in any of the segments. Hence, they match the healthy cohort well. For other patients, instead, outliers are reported in several segments.

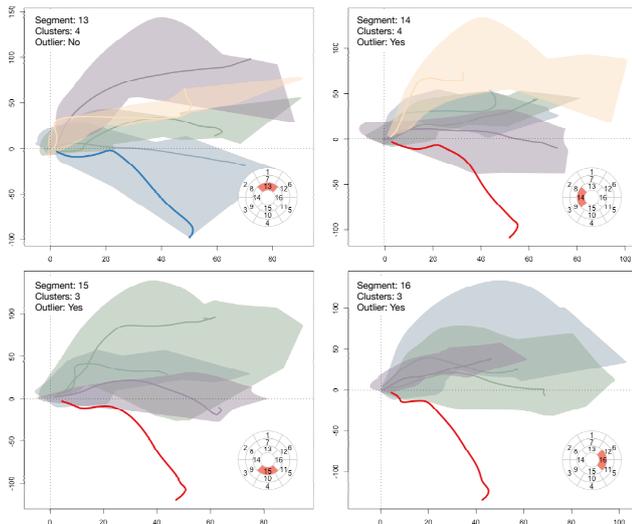


**Figure 5:** Cohort heatmap exhibits for each segment how often patients are considered outliers with respect to a healthy cohort. Order of patients and segments is based on hierarchical clustering as indicated by the cluster trees.

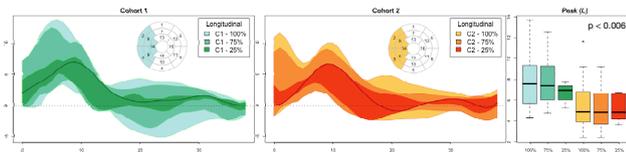
We select patient with ID 16, which has been rated as outlier in multiple segments. Figure 6 shows the clustering result when tentatively adding this patient to the healthy cohort for the segments that belong to the apical slice (Segments 13-16). We know from Figure 5 that the patient has been reported an outlier for Segments 14-16. We observe in the boxplots that for Segment 13 the patient falls into the blue cluster (patient curve is rendered blue), while for Segments 14-16 its curve is clearly different from any of the clusters (patient curve is rendered red).

One may wonder why so many patients did not exhibit any outlyingness. One factor that may play an important role is the age of the people (Hypothesis H4) [FJS\* 10]. For this analysis, we wanted to exclude any factors from a disease and used healthy volunteers only. We split the group of healthy volunteers in two groups, the ones younger than 50 years and the ones older. We again consider the longitudinal velocities for the segments on the left of the AHA model. As expected, we observe in Figure 7 that the velocities are generally lower for older people (red-yellow), which can be seen in both the functional boxplots and the boxplots for the peak velocities. The latter exhibit a statistically significant difference, which confirms the hypothesis.

To test Hypothesis H5 we select the cohort of healthy people and create functional plots where segments to the left of the AHA model (green-blue) are compared to segments to the right of the AHA model (red-yellow). Looking at the juxtaposed functional boxplots in Figure 8, we observe that there is lower radial motion (lower peak velocities) and less variability (narrower bands) of the myocardium in the left segments at the beginning of the heart cycle. The peak velocities exhibit a statistically significant difference ( $p < 0.009$ ), which confirms the hypothesis.

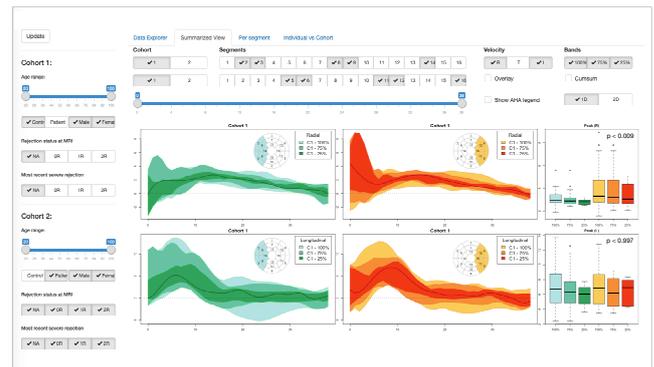


**Figure 6:** Comparing single patient against healthy cohort in curve boxplots: Patient with ID 16 is an outlier (red curve) for segments 14-16, but matches the blue cluster (blue curve) for Segment 13.



**Figure 7:** Analyzing the role of age for healthy volunteers. Cohorts of people younger (green-blue) and older (red-yellow) than 50 years are compared using juxtaposed functional boxplots and boxplots of respective peak velocities. Statistically significant differences in peak longitudinal velocities ( $p < 0.006$ ) are observed.

**Walking Motion in IVE.** In immersive virtual environments (IVE), users are able to experience the computer-generated imagery by replacing their visual and auditory perception with virtual stimuli. While wearing the VR headset, users are blind to the world around them. Therefore, the location of real objects in the surrounding is crucial for the physical safety. A stand-alone hardware device was developed containing distance sensors to detect potential obstacles and haptic devices mounted on the VR headset to alert the user by means of vibro-tactile stimulation [VL19]. The user's head position and orientation were tracked with the Lighthouse position tracking system. An experiment was conducted with the goal to test different sensor-actuator mapping approaches and functions in terms of reaction time, collision avoidance, and safety distance. For the experiment, a within-design system was used, where each participant went through all test conditions in two trials. Three test conditions were defined as: (1) mapping method (distance control only or speed control in addition), (2) the transfer function (step, linear, linear-step, exponential, or exponential-step), and (3) the distance to the virtual wall (3m, 3.5m, 4m, and 4.5m). 20 users (18 male and 2 female, age 20 - 27) were tasked to walk along the straight line while wearing the VR headset and to pick virtual objects in the IVE without colliding with the invisible virtual



**Figure 8:** Comparing regional differences in heart motion for healthy cohort using juxtaposed functional boxplots and boxplots of peak velocities.

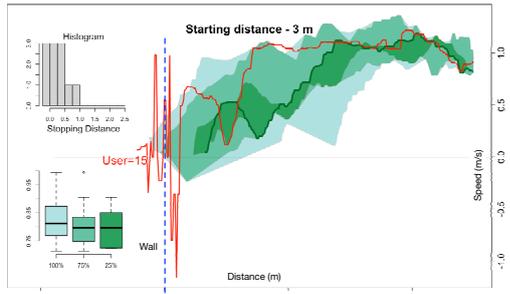
wall [VL19]. The motion data to be analyzed are that of the walking behavior of the users, which is captured in the form of speed values over distance to the virtual wall. Note that this is not a function so that curve boxplots need to be applied.

**Defining cohorts:** Metadata about the subjects including gender, age, height, and eye-correction status were provided. However, the main tasks here were to compare different test conditions, i.e., to investigate cohorts that are formed by choosing different (1) mapping methods, (2) transfer functions, and (3) distances to the virtual wall. Hypotheses were that there would be differences for (1) and (2), but not for (3).

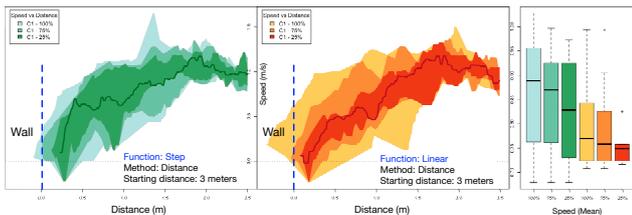
**Results:** First, we formed cohorts and analyzed their homogeneity to detect outliers. For example, Figure 9 shows a comparison of one subject (red) against all other subjects of a cohort. We observe that this subject has a rather unusual walking behavior including going back and forth indicated by alternating positive and negative speed values. The analyst may decide to exclude such strange cases in further analysis steps.

To evaluate the effect of different transfer functions on the walking behavior, curve boxplots of all five conditions are compared. Figure 10 shows curve boxplots (100%, 75%, and 25% bands) for the step and linear transfer functions ("distance control only" mapping method, 3m distance) and boxplots of average user speed. One can observe that users walked faster towards the wall with the step transfer function and mostly stopped before reaching it. For the linear transfer function multiple users "collided" with the wall, but on the other hand stopped closer to the wall, i.e., they did not leave unnecessary safety space. We also observe that speeds were generally higher for the step transfer function. Similarly we compare the two mapping methods (distance control only vs. speed control in addition) in Figure 11 (linear transfer function, 3m distance). We observe that with the distance-only method (left) most of the users stopped close to the wall (a few collided), while with the distance-plus-speed method (mid) the subjects stopped much earlier, i.e., rather far from the wall, none of the users could reach the wall. We see that the subjects walked quite fast initially and then stopped quite abruptly, which also leads to a higher average speed shown by the boxplot (right). The last test condition was that of the initial

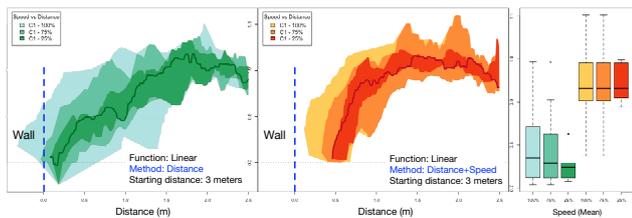
distance to the virtual wall. Figure 12 compares the experiments for 3m vs. 4.5m (linear transfer function, distance-and-speed control method) using overlaid curve boxplots (100% bands). We observe that most of the users that started walking from 4.5m distance stopped much farther away from the wall and that their speed was on average lower. Hence, we conclude that all three test variables actually led to differences in the walking behavior, which was expected for the mapping method and the transfer function, but not for the initial wall distance. More details on the analysis within this application scenario are given in the accompanying video.



**Figure 9:** Comparing a single subject against others within a cohort exhibits a clearly different pattern for the subject with UserID 15 who was stepping back and forth multiple times.



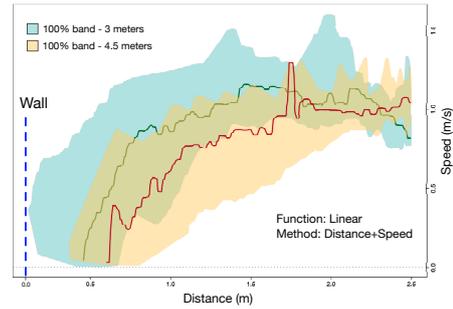
**Figure 10:** Comparing the effect of step (left) and linear (middle) transfer functions using juxtaposed curve boxplots of walking motion in the IVE and boxplots of average speed. The blue vertical line represents the invisible virtual wall.



**Figure 11:** Comparing the effect of mapping method for distance only (left) vs. distance and speed (middle) using juxtaposed curve boxplots of walking motion in the IVE.

### 6. Domain Expert Evaluation

Our tool was demonstrated to one expert from the field of medical imaging and another from the field of human-computer interaction (showing the respective application scenario only). Both experts have many years of experience in their field and had earlier been involved to perform the requirement analysis.



**Figure 12:** Comparing the effect of starting distance to the wall using overlaid curve boxplots (yellow 4.5m, blue 3m).

The medical expert confirmed that all of the analytical tasks M1-M5 (Section 3) can be performed using our tool. He confirmed that “as of now we do not have a visualization system to analyze our data. We use scripts to read in the data, then process them, and display the results in individual figures.” He liked the “user-friendly user interface with lots of possibilities for analysis”. He rated the usefulness of the tool as 3 on a scale from -5 (worst) to 5 (best). He commented that “the web-based graphical user interface is an easy way to get an overview of the data and to easily try different analysis approaches. It is also convenient to select different cohorts”. The user proposed to include the diastolic heart function in the analysis as it is as important to study as systolic function. Note that the tool allows studying both systolic and diastolic functions but due to incomplete motion data samples, it was not possible to incorporate the diastolic function in the analysis. The expert further wished to have a feature to export the figures from the tool to some vector graphics format.

The human-computer interaction expert mentioned that motion tracking data are commonly recorded in VR-experiments, but “the recorded data are rarely evaluated in depth”, as they are “missing an appropriate tool for this”. The user shared his experience from a recent experiment, where he needed to check manually every single recording, just to identify some outliers. He shared that he struggled to identify different behavioral patterns in the recordings. The expert said that our “tool is indeed interesting” and he can “easily imagine a lot of applications” in his field. He verified that the tool supports all requirements (V1-V4) mentioned in Section 3. He shared his desire to incorporate visualizations similar to RadViz [DCFMM10], where he would pick a few important speed profiles and color the nodes/points by different variables to see, if there is some relation among them. We plan to look into this in future work.

### 7. Discussion and Conclusion

We presented an approach for comparative visual analytics of motion data. The main components were the efficient and intuitive designing of cohorts based on metadata and the comparative visualizations based on aggregated data at multiple levels of granularity, which applied statistical methods and machine learning methods for aggregations. Using quite different application scenarios, we have shown that our system is generally and flexibly applicable to motion data analysis. Cohorts can be defined quickly to test hy-

potheses or to adjust them interactively by restricting to some subsets or changing thresholds. However, our system also allows for exploratory analysis, where the analyst has no particular hypothesis in mind, but tests different cohorts. Due to the interactivity of our system, many cohort settings can be tested and adjusted quickly. The domain experts whom we collaborated with reported that with our tool they can perform many comparative tests that they would typically neglect due to the large overhead. The medical experts also reported that they commonly restrict the analysis to derived values such as peak velocity and do not perform a proper analysis of the motion data. One action item for future work is that in our current implementation the mapping from metadata to interaction widget is not yet fully automated. We plan to make our source code publicly available afterwards.

## References

- [AAB\*13] ANDRIENKO G., ANDRIENKO N., BAK P., KEIM D., WROBEL S.: *Visual Analytics of Movement*. Springer Publishing Company, Incorporated, 2013. 2
- [AMST11] AIGNER W., MIKSCH S., SCHUMANN H., TOMINSKI C.: *Visualization of Time-Oriented Data*, 1st ed. Springer Publishing Company, Incorporated, 2011. 2
- [AOH\*14] ANGELELLI P., OELTZE S., HAASZ J., TURKAY C., HODNELAND E., LUNDERVOLD A., LUNDERVOLD A. J., PREIM B., HAUSER H.: Interactive visual analysis of heterogeneous cohort-study data. *IEEE Computer Graphics and Applications* 34, 5 (Sept 2014), 70–82. 2
- [BSM\*15] BERNARD J., SESSLER D., MAY T., SCHLOMM T., PEHRKE D., KOHLHAMMER J.: A visual-interactive system for prostate cancer cohort analysis. *IEEE Computer Graphics and Applications* 35, 3 (May 2015), 44–55. 2
- [CWD\*02] CERQUEIRA M. D., WEISSMAN N. J., DILSIZIAN V., JACOBS A. K., KAUL S., LASKEY W. K., PENNELL D. J., RUMBERGER J. A., RYAN T., ET AL.: Standardized myocardial segmentation and nomenclature for tomographic imaging of the heart: a statement for healthcare professionals from the cardiac imaging committee of the council on clinical cardiology of the american heart association. *Circulation* 105, 4 (2002), 539–542. 6
- [DCFMFM10] DI CARO L., FRIAS-MARTINEZ V., FRIAS-MARTINEZ E.: Analyzing the role of dimension arrangement for data visualization in radviz. In *Pacific-Asia Conference on Knowledge Discovery and Data Mining* (2010), Springer, pp. 125–132. 9
- [DKT17] DHEERU D., KARRA TANISKIDOU E.: UCI machine learning repository, 2017. URL: <http://archive.ics.uci.edu/ml>. 6
- [DVGG\*17] DIAZ D., VILLEGAS J., GUERRA-GOMEZ J. A., CHARPAK N., HERNANDEZ J. T.: Visual tools for the exploration of growth data in a cohort of kangaroo infants during their first year of life. In *2017 IEEE Workshop on Visual Analytics in Healthcare (VAHC)* (Oct 2017), pp. 9–16. 2
- [FJS\*09] FÖLL D., JUNG B., STAEHLE F., SCHILLI E., BODE C., HENNIG J., MARKL M.: Visualization of multidirectional regional left ventricular dynamics by high-temporal-resolution tissue phase mapping. *Journal of Magnetic Resonance Imaging: An Official Journal of the International Society for Magnetic Resonance in Medicine* 29, 5 (2009), 1043–1052. 6
- [FJS\*10] FÖLL D., JUNG B., SCHILLI E., STAEHLE F., GEIBEL A., HENNIG J., BODE C., MARKL M.: Magnetic resonance tissue phase mapping of myocardial motion: new insight in age and gender. *Circulation: Cardiovascular Imaging* 3, 1 (2010), 54–64. 7
- [FKRW17] FERSTL F., KANZLER M., RAUTENHAUS M., WESTERMANN R.: Time-hierarchical clustering and visualization of weather forecast ensembles. *IEEE Transactions on Visualization and Computer Graphics* 23, 1 (Jan 2017), 831–840. 2
- [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. 5
- [HvW09] HOLTEN D., VAN WIJK J. J.: Force-directed edge bundling for graph visualization. In *Proceedings of the 11th Eurographics / IEEE - VGTC Conference on Visualization* (Chichester, UK, 2009), EuroVis'09, The Eurographs Association & John Wiley & Sons, Ltd., pp. 983–998. 2, 5
- [LCB\*16] LIN K., CHOWDHARY V., BENZULY K. H., YANCY C. W., LOMASNEY J. W., RIGOLIN V. H., ANDERSON A. S., WILCOX J., CARR J., MARKL M.: Reproducibility and observer variability of tissue phase mapping for the quantification of regional myocardial velocities. *The international journal of cardiovascular imaging* 32, 8 (2016), 1227–1234. 6
- [LZH07] LANGFELDERA P., ZHANG B., HORVATHA S.: Dynamic tree cut: in-depth description, tests and applications. *November 22* (2007), 2007. 4
- [Mül07] MÜLLER M.: Dynamic time warping. *Information retrieval for music and motion* (2007), 69–84. 4
- [MWK14] MIRZARGAR M., WHITAKER R. T., KIRBY R. M.: Curve boxplot: Generalization of boxplot for ensembles of curves. *IEEE Transactions on Visualization & Computer Graphics* 20, 12 (Dec. 2014), 2654–2663. 2, 5
- [RSN\*19] ROGERS J., SPINA N., NEESE A., HESS R., BRODKE D., LEX A.: Composer: Visual cohort analysis of patient outcomes. *Applied clinical informatics* 10, 02 (2019), 278–285. 3
- [SG11] SUN Y., GENTON M. G.: Functional boxplots. *Journal of Computational and Graphical Statistics* 20, 2 (2011), 316–334. 2, 5
- [SG15] SAMUELS L. R., GREEVY R. A.: Visual pruner: Visually guided cohort selection for observational studies. In *2015 IEEE Conference on Visual Analytics Science and Technology (VAST)* (Oct 2015), pp. 215–216. 2
- [SRB\*18] SARNARI R., RUH A., BLAISDELL J., DOLAN R. S., LIN K., GHAFOURIAN K., WILCOX J. E., KHAN S. S., VOROVICH E. E., RICH J. D., ANDERSON A. S., YANCY C. W., CARR J. C., MARKL M.: Altered biventricular myocardial velocities in patients during year 1 after heart transplantation. In *Proceedings of the 30th Annual Meeting of the Society for Magnetic Resonance Angiography (SMRA)* (2018). 6
- [SWF\*05] STEWART S., WINTERS G. L., FISHBEIN M. C., TAZELAAR H. D., KOBASHIGAWA J., ABRAMS J., ANDERSEN C. B., ANGELINI A., BERRY G. J., BURKE M. M., ET AL.: Revision of the 1990 working formulation for the standardization of nomenclature in the diagnosis of heart rejection. *The Journal of heart and lung transplantation* 24, 11 (2005), 1710–1720. 6
- [VL19] VALKOV D., LINSEN L.: Vibro-tactile feedback for real-world awareness in immersive virtual environments. In *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)* (March 2019), pp. 340–349. 8
- [WLM\*14] WANG C., LI J., MA K., HUANG C., LI Y.: A visual analysis approach to cohort study of electronic patient records. In *2014 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)* (Nov 2014), pp. 521–528. 2
- [YR15] YIM O., RAMDEEN K. T.: Hierarchical cluster analysis: Comparison of three linkage measures and application to psychological data. *Tutorials in Quantitative Methods for Psychology* 11, 1 (2015), 8–21. 4
- [YWT\*17] YANG D., WU P., TAN C., POHL K. M., AXEL L., METAXAS D.: 3d motion modeling and reconstruction of left ventricle wall in cardiac mri. In *International Conference on Functional Imaging and Modeling of the Heart* (2017), Springer, pp. 481–492. 6
- [ZGP12] ZHANG Z., GOTZ D., PERER A.: Interactive visual patient cohort analysis. In *Proceedings of IEEE VisWeek Workshop on Visual Analytics in Healthcare* (2012), Seattle, WA USA. 3