

AnaFe: Visual Analytics of Image-derived Temporal Features – Focusing on the Spleen

Ievgeniia Gutenko *Student Member, IEEE*, Konstantin Dmitriev *Student Member, IEEE*,
Arie E. Kaufman *Fellow, IEEE*, and Matthew A. Barish

Abstract— We present a novel visualization framework, *AnaFe*, targeted at observing changes in the spleen over time through multiple image-derived features. Accurate monitoring of progressive changes is crucial for diseases that result in enlargement of the organ. Our system is comprised of multiple linked views combining visualization of temporal 3D organ data, related measurements, and features. Thus it enables the observation of progression and allows for simultaneous comparison within and between the subjects. *AnaFe* offers insights into the overall distribution of robustly extracted and reproducible quantitative imaging features and their changes within the population, and also enables detailed analysis of individual cases. It performs similarity comparison of temporal series of one subject to all other series in both sick and healthy groups. We demonstrate our system through two use case scenarios on a population of 189 spleen datasets from 68 subjects with various conditions observed over time.

Index Terms—Visual Knowledge Discovery, Temporal Feature Analysis, Radiomics, Spleen, Abdominal Imaging

1 INTRODUCTION

The spleen, the largest organ in the lymphatic system, is often forgotten by laypeople but is of significant importance to clinicians. An increase in splenic size, splenomegaly, accompanies immune response in a wide range of abnormal conditions, including immunologic, hematopoietic, infectious, and storage diseases [2, 36]. The evaluation, staging and response assessment of Hodgkin and Non-Hodgkin Lymphoma incorporate the spleen in the diagnostic process and it underlines a lack of consensus on standard splenic metrics [9, 39]. Additionally, spleens vary in shape and size across the patients, further complicating the task of finding a suitable evaluation criteria.

Such variation is common among many medical sub-domains. Most medical research studies which focus on the analysis of time-varying imaging data are faced with this obstacle. The goal of these studies is to determine universally applicable metrics for use in clinical practice. However, the process of finding even an initial set of candidate metrics requires a comparison of volumetric imaging data across patients and studies over the time of disease progression.

Diseases that affect the spleen pose particular challenges since they lack the clear-cut characterization between healthy and unhealthy subjects, which is often found in other organs (e.g. kidney). Specifically:

- Splenic maladies can manifest as variations in both shape and size. Thus, domain experts who are characterizing such abnormalities often find it difficult to do so based on a single metric.
- Traditionally, splenic disease has been determined on simple measurements (width/length) or volumetric estimates. These metrics often fail to characterize disease, since unhealthy spleens can fall within normal ranges. Consequently, it is important to consider other parameters, such as shape.
- Groups of patients can exhibit patterns of disease progression (e.g. changes in organ volume, shape, or other features) over time, which are of value in identifying the efficacy of treatment regimens and the accuracy of certain measurements for characterization.

Since these challenges are based on the interplay of multi-dimensional, time-varying data, across populations, we feel that a visual analytics approach has the potential to support medical analysis of the problem domain of splenic disease.

In this paper, we propose a visual analytics tool, *AnaFe*, for the analysis of changes in the organ based on imaging-derived features. The current quality of image acquisition and analysis allows for the extraction of reproducible quantitative imaging features. We take advantage of such features derived from the field of radiomics [31] to enable an open-ended similarity search and trend exploration in splenic imaging data. Such features have several benefits: simplicity in derivation, mapping to visualization, and most importantly reproducibility and versatility across imaging types and domains.

AnaFe supports a similarity search and comparison workflow based on a set of robust radiomics features. Our proposed workflow enables a user to construct custom similarity-based queries through interaction with these features. Through a number of visual queries and a rapid visual feedback, the user is able to compare several time-variant imaging sets and their correlating features in a single overview. Driven by demands of medical research analysis, *AnaFe* combines a set of linked visualization views. Thus, researchers can concentrate on exploring and characterizing changes in data and corresponding features.

The utility of our tool is demonstrated through two case studies conducted by our collaborating radiologist on a set of 189 datasets. We begin this paper with an overview of background on the spleen and other medical visual analytics systems. Based on the target application domain we outline a set of visualization requirements for our visual analytics system and describe the resulting design decisions. Then, we illustrate the implementation of our tool and focus on the integration of various feature types and their mapping to visualization views. Section 5 is dedicated to application use-case scenarios. We conclude by offering directions for future work.

2 BACKGROUND IN APPLICATION AREA

The simplicity of existing comparison methods drives the need for visual exploration. Together with our collaborating radiologists, we assessed existing issues and defined the requirements for our system.

2.1 Splenic Measurements and Comparison

Quantitative measurements of the spleen are often based solely on the organ craniocaudal length (L), marked by the first and last axial slices of the organ in the scan, or on the linear regression of the product of width (W), thickness (T), and length (L), called the *splenic-index* [3, 14, 35]. This metric provides a very rough volume estimation of a complex shape and does not generalize well to new populations. Additional inconsistency in volume estimation originates from the selecting of

• {igutenko, kdmitriev, ari} @cs.stonybrook.edu
Computer Science Department, Stony Brook University, NY, 11794-2424
{matthew.barish} @stonybrookmedicine.edu

Manuscript received xx xxx. 201x; accepted xx xxx. 201x. Date of Publication xx xxx. 201x; date of current version xx xxx. 201x. For information on obtaining reprints of this article, please send e-mail to: reprints@ieee.org.
Digital Object Identifier: xx.xxx/TVCG.201x.xxxxxxx

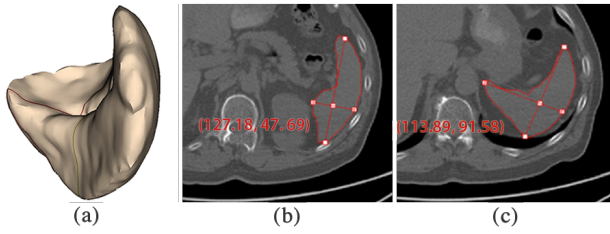


Fig. 1. Measurements of (a) the spleen's width and thickness taken on (b) the slice with maximal axial width and (c) the slice selected by an expert radiologist. In this case, the splenic volume is calculated as 368.6 and 281.2 mL for the automatic for prolate ellipsoid [47] and Rezaei [35] methods and 632.6 and 462.6 mL for the observer measurement, respectively, whereas the true volume is 381.3 mL. Selection of the slice varies significantly between observers and different heuristics. As shown here, simple measurements do not capture the shape of the organ, resulting in a large error.



Fig. 2. Bland-Altman plot of agreement between the true volumes of our data and the prolate ellipsoid volume estimates illustrates a large error for large organs. CCC is moderate: $r=0.9125$ (95% CI 0.81 - 0.93), mean= -77.3 ± 142.64). The plot shows that the estimation method does not perform well for large spleens of sick subjects, which are those that require monitoring of changes.

the axial slice for measurement. Figure 1 shows an example where a poor choice of the axial slice can result in a large estimation error. We have compared volume estimation methods and heuristics of slice selection on time-varying data [17]. Even when optimally selecting the slice, volume estimation alone does not capture complex changes in the organ. Primarily, there are two main issues (IS):

IS1 New volume estimation methods compare the concordance correlation coefficient (CCC) to the gold standard (ground truth segmentation). Figure 2 shows these results for prolate ellipsoid volume estimation [47] on our set of 189 scans. CCC (adjusted for repeated measurements [7]) describes the result as moderate correlation. However, the Bland-Altman plot indicates a large number of outliers outside the limits of agreement. Further inspection and comparison of organ variation is required.

IS2 To the best of our knowledge, there are no studies performing evaluation on multi-timepoint spleen data. Evaluating similarity in disease progression and outcomes requires more points/features of comparison due to large organ variation. Changes in imaging-derived features, other than volume, length, width, and thickness, have not been previously considered.

These issues stress the need to explore the variations among spleens and find features characterizing progressive change in a diseased organ.

2.2 Spleen Variational Anatomy and Clinical Significance

Large differences between spleens among different subjects make their comparison a challenging task. Splenic size varies with age, race, and

gender [19,26]. The size can change in response to trauma or temporary variation in splenic vascularization (splenic vein hypertension) [27]. The spleen shape forms during development in relation to neighboring structures and can vary from a slightly curved wedge to a “domed” tetrahedron. *Variational anatomy of the spleen* is imperative knowledge for surgical interventions, including splenectomy, resection of tumors, and extirpation of cysts [27]. Next, we outline processes where changes in the organ over time carries clinical significance.

Assessment of Splenomegaly. O'Reilly *et al.* [30] have studied 2,505 cases of splenomegaly accompanying lymphoma, AIDS, splenic cysts, and splanchnic vein obstruction. They found that *progressive splenic enlargement* demands diagnostic study because these patients usually have a hematologic cancer.

Observation of Lymphoma and Leukemia. In clinical trials of lymphoma and leukemia patients, the variation in spleen size is considered when determining if the patient is responding to treatment or if the disease is advancing. In non-Hodgkin's lymphoma, radioimmunotherapy treatment [39] has shown to be more successful when a decrease in splenic volume occurs in subjects suffering from splenomegaly, and a 15% decrease in this metric is considered substantial. There is an increasing interest in the use of more features describing *progressive splenic enlargement and shrinking* due to complex changes in the organ.

Liver Cirrhosis. Correlation of splenomegaly and Child-Pugh scores suggests significance of splenomegaly in the staging of liver cirrhosis [8]. The volume ratio of liver to spleen assessed with CT is significantly lower in patients who developed primary biliary cirrhosis (PBC) [29]. Hence, it could be an indicator for predicting the development of symptoms and prognostic outcomes.

Observation of *progressive changes in the spleen* is imperative in clinical trials and medical research studies aiming to improve patient prognosis and care. *Variational anatomy* can't be described by existing simple measurements. *AnaFe* fills in this gap by allowing hypothesis-free exploration of imaging-derived features, allowing the user to gain insights on changes with a purpose of further statistical evaluation.

2.3 Visual Analysis Requirements

Through close collaboration with expert radiologists, and based on the outlined issues, we have identified the following challenges:

- RQ1** *Observation of progressive changes in the spleen over time and their comparison to standard measurements.* What are the actual changes happening to the organ when its volume decreases [30] (**IS2**)? An improved understanding of changes based on multiple features could help answer the question of whether a single 1D [3] or “splenic index” based measurement can describe the change.
- RQ2** *Similarity of changes across subjects.* How do measurements and other imaging features compare across subjects in both sick and healthy populations (**IS1**)? Observation of per subject trends of disease progression and response to treatment is imperative for better treatment planning and patient monitoring, thus, it should not be limited to basic measurements.
- RQ3** *Overview of organ variation in the population.* Given the large organ variation in terms of shape and size [27] (**IS1**), a visualization of organ variation can help in anatomical understanding for surgical planning and can also be used for education purposes. Geraghty *et al.* [14] compare the need for such an overview in the form of spleen studies to the studies and databases of osteoporosis. In these studies, availability of large normative databases led to standardization of bone mineral density (BMD) measurement.
- RQ4** *User-driven feature exploration.* Imaging-derived features have allowed for comprehensive comparison in radiomics studies of liver, ovarian, and lung cancers [31]. While widely used, these features have been selected based only on quantitative methods [10]. We are using a combination of visual analytics of a largely forgotten organ through highly interactive exploration.

The visual analysis process in *AnaFe* is structured in a way which satisfies these requirements.

3 RELATED WORK IN VISUAL ANALYTICS DOMAIN

The majority of related works have addressed the visualization and exploration of several types of medical information: heterogeneous data from cohort studies (including temporal information), data from electronic health records (EHR), and visual analysis of physical and feature spaces for medical data. Our work combines the analysis of image and non-image data, but focuses on user interactions with imaging-derived features that change over time.

3.1 Interactive Visual Analysis of Heterogeneous Data

Recent visual analytics research focuses on the creation of highly interactive tools [21, 44] and data organization solutions [1, 40]. The power of such tools is in *hypothesis-free* [40, 44] exploration that allows domain experts to iterate over multiple data variables. Thus, researchers do not need to rely solely on intuition and observations from clinical practice, but can get a visual summary of multiple parameters at once.

Steenwijk *et al.* [40] have proposed a conceptual framework for heterogeneous temporal patient data organization analysis. This framework defines domains, features, mappings, and studies, and combines them into a relational database. The data-cube model [1] for cohort studies handles partially overlapping data subsets and provides higher computation efficiency in comparison to a relational database. While the proposed model accounts for the aggregation of multi-timepoint data (included in the study), there are no visualizations presented that allow direct comparison over the course of time. Their visual linking of spatial and non-spatial views is limited to viewing a single set in a 3D view at a time. Our work focuses on aligning temporal subject sequences based on imaging-derived features and allows simultaneous viewing of multiple 3D sets. Klemm *et al.* [21] have introduced an interactive visual analysis (IVA) workflow for epidemiological cohort studies targeted at domain experts. Their work supports the definition of demographics subgroups, among subjects with lower back pain, driven by spine shape clustering features. The mean representative 3D spine shape is visualized over a familiar information visualization with statistical information.

Keefe *et al.* [20] have highlighted the importance of combining a 3D data overview and 2D information visualization for multi-dimensional data analysis. They take advantage of an overview visualization of 3D small multiples of pig mandibles and related chewing traces. We find their overview display particularly helpful for rapid visual pinpointing of differences when compared to single [1] or mean shapes [21]. For our application, small multiples overview of temporal spleen data substitutes for traditional sequential analysis. Highly interactive visual analytics tools for cohort studies are particularly powerful with the addition of descriptive statistics of multiple data dimensions [44]. In this form they enable dual analysis for hypothesis generation. In our work we focus on a combination of imaging-derived features and physical objects to enable a rapid comparison of disease progression.

3.2 Visual Analytics of Physical and Feature Space

Researches have combined the visualization of physical and feature spaces for medical data. Feature space may be defined as measurements taken by clinicians or features derived from imaging data. For example, WEAVE [16] combines the visualization of measurements and an anatomical representation for cardiac simulation. Raicu [33] has summarized mining knowledge from medical imaging data based on features derived from CT. He pays particular attention to tissue classification, which includes classification of spleen tissues with lowest sensitivity and precision values due to its similar attenuation to liver.

The feature space can further be divided into higher level representations, for example models of shape, or lower level representations per voxel. Per-voxel features are explored in works by Fang *et al.* [12] and Blass *et al.* [4]. Per-voxel features enable the creation of time activity curves (TAC) which are important in the areas of nuclear medicine [12]. Tightly coupled views of per-voxel feature projections enable pattern finding and interactive segmentation from multi-dimensional data [4]. In the same domain, Raidou *et al.* [34] proposes visual analysis of tumor characterization based on dimensionality reduction techniques.

Higher level representations allow for the exploration of larger collections of data. For example, shape variation can be explored through

visualization of a given 3D shape within a projected shape space [5]. Busking *et al.* [5] introduce a framework comprised of three views. It supports the exploration of shapes over a population and individual shape progression. Statistical deformation models (SDM) (Hermann *et al.* [18]) are used to study anatomical shape covariation interactively. Caban *et al.* [6] give an overview on visualization of SDM.

For the purpose of our work we wish to avoid the complexities and clutter of per-voxel features, and not favor a particular feature type such as shape. Specifically for the spleen, shape varies significantly from subject to subject and is defined during development. Detailed features of spleen shape such as notches and lobules are of interest to clinicians. While the intra-subject variation in splenic shape is largely considered normal, the actual distribution of types of shapes is not well studied and is becoming of interest to medical researchers. In our work, we use robust and reproducible imaging features from the field of radiomics [10, 31]. The goal of radiomics is to obtain high quality, reproducible imaging features in order to provide improved information for patient management. These features represent quantitative information about statistical intensity, texture, and shape. But most importantly, they provide information about imaging data in a concise form. Parmar *et al.* [31] outline related examples of using quantitative imaging features in liver, ovarian, and lung cancers.

3.3 Interactive Visual Analysis of Temporal Patient Data

Multi-subject visualization of temporal data necessitates the organization of a wide range of information (5 W's [48]), allows one to focus on patterns and specific scenarios of events in the emergency room during patient intake [11, 28], and finds differences and similarities in temporal patient data (CoCo [25]). Finding similarities in patient cohorts is addressed in: *LifeFlow* [46], *CareFlow* [32], *DecisionFlow* [15], and *OutFlow* [45]. Not only analysis, but also construction of the cohort is a challenging question when dealing with large amounts of information [22]. It is important to emphasize that while these works focus on visualization of similarities and differences in temporal patient histories, they do not include analysis of imaging-derived features.

4 SYSTEM DESIGN AND IMPLEMENTATION

AnaFe is a framework for analytical exploration of 3D organ data and associated medical-imaging derived features. We focus on parts (c)-(f) in Figure 3: deriving the imaging features and their visual analysis in combination with non-image features from the study data.

4.1 Input Data

Our application consumes several types of data that stems from the medical research study of spleen variation. Each study is comprised of multiple subjects observed over several visits.

Subject Information. Each study contains information with respect to gender, age, and disease status of the subject (sick or healthy).

Imaging-derived Features. Feature vectors for each study are derived from the original DICOM data after applying segmentation masks. We describe these features in detail in Section 4.2. The effect of segmentation on quantifiable radiomics features has been previously evaluated [31]. For the purpose of unbiased evaluation we used features derived only from manually segmented data.

3D Surface Mesh Pre-processing. *AnaFe* requires a 3D mesh for each organ for rendering, which is obtained in two steps.

Mesh generation. From the manually segmented data we derive a 3D surface mesh of the organ via marching cubes [24]. The surface geometry is then post-processed with Taubin smoothing [41].

Mesh registration. For comparison of several organs' surfaces, we align the meshes using the Iterative Closest Point (ICP) algorithm [23]. Next, meshes are registered using non-rigid ICP to obtain correspondences between the vertices of consecutive timepoint meshes. This type of registration has not been previously used for the spleen. We have selected these methods based on the preliminary evaluation of their performance and ready availability.

Thus, any study of a group of subjects with a series of repeated measurements (dependent variables) over time (independent variable), that contains similar information can be analyzed by *AnaFe*.

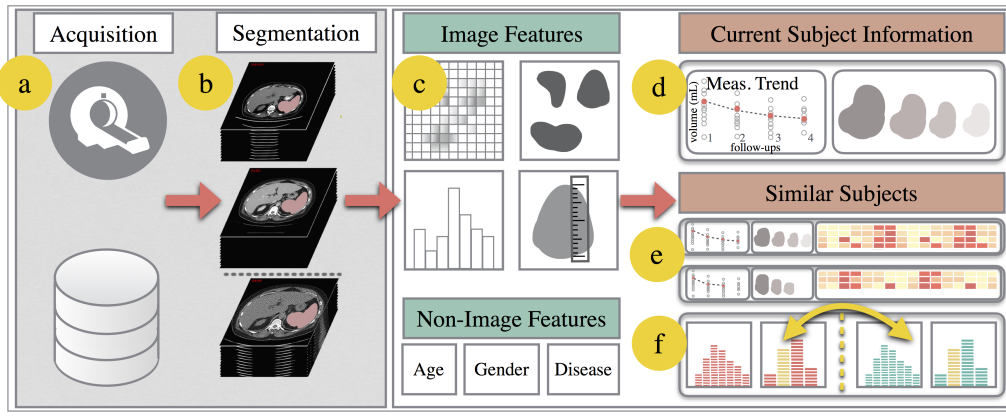


Fig. 3. An overview of the steps required for temporal feature analysis by *AnaFe*. The data is sourced from follow-up CT scans (a) that have been segmented (b). In this work, we first derive robust image features from the field of radiomics (c): texture, shape, intensity, and measurements. These features, along with other study information (non-image features), are used by our web-based application *AnaFe*. After selection of the current subject (d), the application performs a search for subjects with similar progression. Search results filtering is performed by brushing a selection over the distribution charts (f) of average features and average change in progression (interchangeable views shown with a double yellow arrow.)

4.2 Feature Types

In clinical practice determining the spleen size and response to treatment is often a qualitative rather than quantitative process. For the purpose of quantitative comparison of inter-subject similarity of organ change along with basic measurements, we integrate a set of imaging-derived features. With our application and other potential domains in mind, we group these features into the following four categories:

- **Measurements.** In practice, volume, craniocaudal length, width, and thickness serve as primary comparisons for spleens. Therefore, we define them as a separate category. For many other applications, measurements can be defined as descriptors of shape.
- **Shape.** Shape descriptors characterize spleen shapes, which include elliptical, triangular, wedge, and tetrahedron [36].
- **Intensity.** In CT scans, the spleen is often described as an organ of homogeneous density [36]. Thus, intensity is not extremely important for this application. We include this feature group for the purpose of extensibility to other domains.
- **Texture.** Texture features are based primarily on a gray-level co-occurrence matrix (GLCM). Such features describe cluster prominence and can be used for tumor characterization. Similar to intensity, this feature is included for extensibility.

The detailed definitions of these features can be found in the related literature [31]. In addition, we provide information on the features used in our case studies in Appendix A. These imaging-derived features, along with study information (gender, age, and disease status), comprise the full feature vector for each dataset.

4.3 Similarity Comparison

Comparing changes across subjects (RQ2) requires analyzing data from multiple timepoints. The number of visits varies from subject to subject based on the condition and progression of the disease. It was noted by our medical collaborators that the time difference between the two consecutive patient visits is not of essence for this comparison. Varying duration of treatment and the subject's initial conditions can affect the quantitative difference in the measurement/features. These compounding variables about the subject's original state were not available during the current study. Regardless of the velocity of change in the organ's condition, it must be detected. The focus of our study is visual identification and correlation of predictors of change among the large number of features (measurements and imaging-based) for future analysis and statistical testing.

We obtain a vector of imaging-derived features for each subject's visit. For similarity computation purposes, these features are normalized to the [0, 1] range based on maximum and minimum values. The

original value of each feature is stored for display purposes (in labels and tooltips). Thus, each subject is described by several feature vectors over multiple visits. Similarity comparison between two studies is performed by computation of two measures: *cosine similarity* of feature vectors and *dynamic time warping* (DTW) of time-series.

Cosine similarity is defined as the similarity between two feature vectors for a given timepoint. Specifically, it is an angle between two feature vectors. In our application, we use a special case of weighted cosine similarity. Initial feature weights are assigned to be equal and can be changed by the user in the process of similarity-query construction as described in Section 4.5. Based on the weight, some features can be fully excluded from the comparison. For example, texture and intensity of the spleen can be described as homogeneous, and thus the user can fully exclude such features if not found to be meaningful.

Time-series similarity is defined by the DTW distance measure [37, 38] algorithm with the above described cosine similarity as a distance cost at each point. DTW distance can be determined for sequences of unequal lengths and does not account for the difference in time between two points, which satisfies our requirements. As we have mentioned earlier, in the series of subject's scans the difference between two timepoints can be ignored. For the datasets where this assumption does not hold, two possible variations can be considered. In a simple case of evenly spaced events, Euclidean distance between time series with modification for missing data can be used. For irregular time series, variations of DTW or model-based search methods can be used. Additionally, DTW has been previously applied to the analysis of medical data with incomplete series [42].

AnaFe performs a similarity comparison of the selected multi-timepoint study to all other studies loaded in the system. The results of the similarity search update two synchronized views.

4.4 Visualization Design

AnaFe implements our collaborators' requirements through a number of highly interactive linked views described in this section:

- **Demographics Overview (DO)** - an overview of basic demographic information (age, gender, and disease types).
- **3D Small Multiples Objects over Time (SMO)** - an overview visualization of the organ progression via 3D small multiples that simplifies direct comparison of the organs as per RQ2 and RQ3.
- **Feature Distribution Overview (FDO)** - a visualization of several groups of quantitative imaging features via familiar information visualization plots with interactive capabilities as per RQ4.
- **3D Object Detail (OD)** - a detailed progression view of 3D organ mesh shapes highlighting the changes between timepoints as an expansion to RQ1.

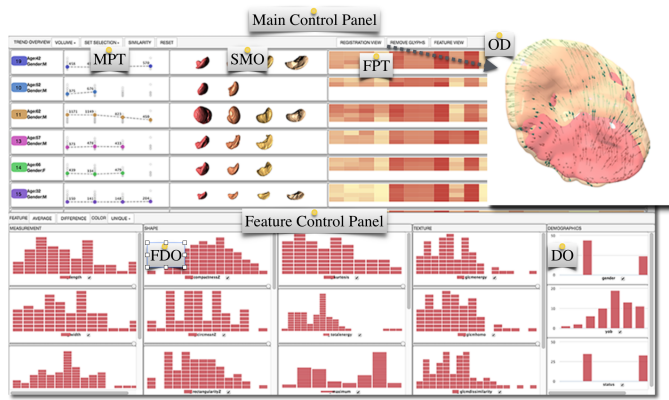


Fig. 4. An application layout showing locations of all views, as described in Section 4.4. The arrow indicates interchangeable views.

- *Measurement Progression over Time (MPT)* - an overall progression *trend* for the patient based on measurement metrics (volume, length, width, and thickness) (**RQ1**).
- *Feature Progression over Time (FPT)* - a heatmap-style visualization of feature progression over time (**RQ4**).

The layout of the application (Figure 4) is broken down into four main views and two control panels (main and feature selection). From the main control panel, the user can select a multi-subject study to be loaded from the back-end. The *DO*, *SMO*, *FDO*, and *MPT* views will be populated based on the number of subjects and the computed feature distributions as described in Section 4.4.5. The *OD* and *FPT* views alternate based on the user’s selection in the control panel. *MPT*, *SMO*, and *FPT* (if selected) comprise a single table/list style scrollable view. The implemented capability of “virtual” scrolling allows for rendering of only currently visible items. This implementation alleviates the bottleneck of simultaneously handling multiple WebGL canvases and contexts for 3D small multiples rendering.

The selected study is displayed at the top of the table list and highlighted accordingly. The similarity comparison is performed based on the user defined query through filtering options in the *FDO*. We describe each of the views below.

4.4.1 Demographics Overview (DO)

The demographics overview consists of a bar chart visualization showing the distribution of gender, age, and disease status (sick or healthy) in the population. Spleen volumetry studies analyze the distribution of the organ’s volume and other measurements (length, width, thickness) in the population for different demographic groups. Through interactions with this view, the user can filter which sets are shown and analyzed. By clicking on the column in the bar chart (or by brushing a selection over a range of columns), the user can select a category or a range of values. For example, only sick patients in the age group of 40-80 can be selected for analysis.

4.4.2 Measurement Progression over Time (MPT)

In clinical practice, the response to treatment or disease progression for the spleen is evaluated quantitatively based on length or volume of the organ [9, 35]. Similarly, in many applications, tumor growth is measured by a single diameter or volume. The progression trend (MPT view) should show an overall increase or decrease in this metric over a series of the subject’s follow-up visits.

This view is implemented as a scatterplot (Figure 5) that shows the trajectory of the measurement for a given subject at each period of time (color filled markers). The dashed line connects two consecutive measurements into the progression trend per subject. In order to enable fast comparison of the trend of one subject with respect to others, this chart shows distribution of measurements across the population (grey empty markers). All measurements of one subject are highlighted on mouse interaction, thus providing trajectories for comparison.

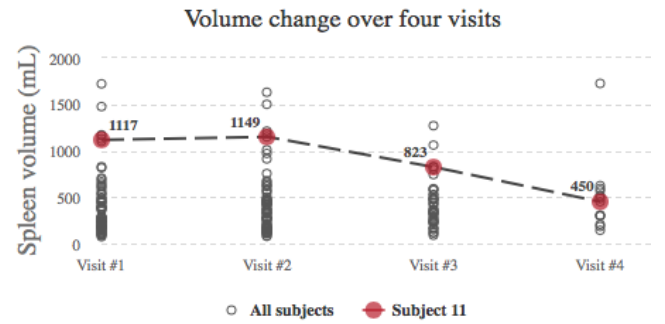


Fig. 5. An example scatterplot of Measurement Progression over Time (MPT). Red markers show measurements for a given subject (#11) at each follow-up visit. The dashed line connects the measurements showing the progression trend for this given subject. Grey markers show measurements within the population at each visit.

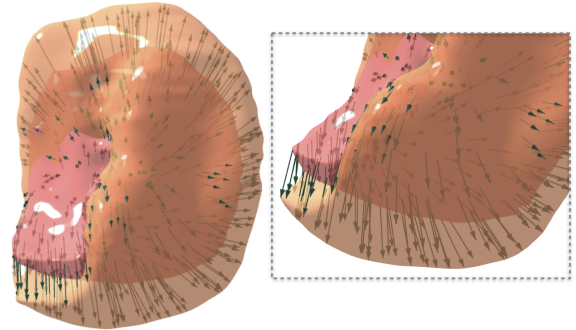


Fig. 6. An example of the OD view showing two spleens rigidly aligned. Small black arrows indicate the direction of change and are obtained from the result of non-rigid registration.

As previously discussed [17], there is a lack of consensus on which measurements are used to compare disease progression. This panel allows the user to define a trend metric, e.g. splenic volume, and visually compare the change to the *SMO* and the *OD* views with the 3D rendering of the organ. Evaluating and understanding the manner in which spleens change in size, for example shrink faster in length or in width, can characterize a disease progression or treatment. In this way, an abnormal spleen can be characterized even when it falls within the normal range for splenic size.

4.4.3 3D Small Multiples Objects over Time (SMO)

In a traditional clinical setting, only a few imaging results can be rendered simultaneously side-by-side. In such a scenario, the user can compare the change only for a single patient. To compare organ change among several subjects, we employ Tufte’s principle of small multiples [43]. Keefe *et al.* [20] have used a similar approach for the visualization of 3D pig mandibles.

In our 3D small multiple views, the organ model is shown sequentially for each timepoint. It allows the user to quickly compare multiple organ models between several subjects (Figure 9 (1)). The time-series 3D object overview visualization allows individual rotation of each model for comparison from multiple angles. Furthermore, each of the organs can be selected with a double click for detailed investigation in the *OD*.

The *SMO* and *MPT* are positioned side-by-side in the form of the table/list view. In this way the measurement scatterplot provides the required trend visualization when change is not immediately visible on the 3D object. The 3D *SMO* view serves as a 3D thumbnail preview of the organ shapes. Correlating a trend metric from the *MPT* view (e.g. volume, length, or width) to the actual shapes can be done without context switching to multiple images. Thus, we combine both scientific and information visualization for analytics purposes.

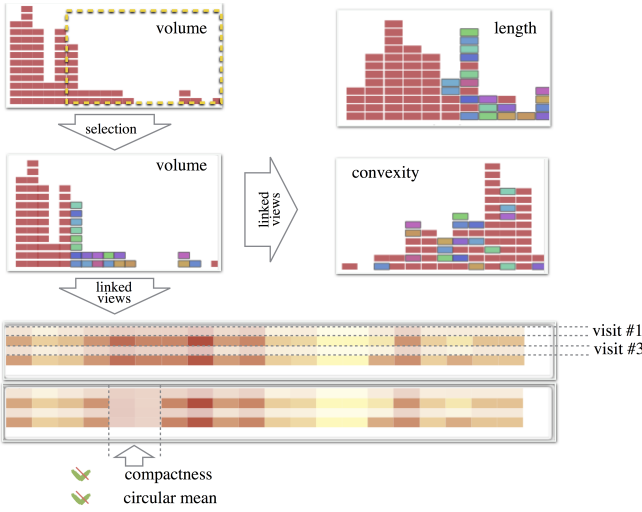


Fig. 7. An example of linking via brushing support between “volume,” “length,” and “convexity” in the FDO view. By brushing in the “volume” chart visualization, the user can compare how selected subjects are distributed in other charts. Subjects and timepoints outside the selected range are removed from the similarity comparison (row). If the feature is fully deselected via the checkbox, the respective column in the FPT heatmap is dimmed out.

4.4.4 3D Object Detail (OD)

Traditionally, organ progression observation and analysis is limited to statistical 1D measurements. Further investigation of 1D measurements and correlated spatial changes requires manual browsing through 2D slices of DICOM data for each dataset. Our collaborators have noted that this comparison demands substantial context-switching between several applications (statistical packages, DICOM viewers, etc.), incurring a significant time and effort overhead.

AnaFe renders the 3D surface mesh of the organ, thus enabling close investigation of local changes in the organ. Based on rigid alignment (ICP), organs from consecutive timepoints are aligned. Non-rigid registration (non-rigid ICP) provides point-by-point correspondences to render arrow glyphs. This type of representation can show the direction of the greatest change in the spleen (Figure 6). From the semi-transparent rendering view the user can see whether the decrease was uniform in all dimensions or whether one dimension increased/decreased more rapidly. In many organs, a disease is defined in terms of distortion of their shape. Diseases that uniformly enlarge the organ without shape changes tend to fall into one group, while diseases that fundamentally change the shape tend to belong to another category. For the spleen, this view helps the user to understand whether the change is uniform or irregular.

4.4.5 Feature Distribution Overview (FDO)

Histograms are common to visualize density distribution in the data. Feature visualization of multi-timepoint data should provide an understanding of distribution of average feature values and relative changes for each subject. The number of histogram bins and their size have to be chosen carefully to depict the overall density. While a basic histogram can provide an answer with respect to the overall distribution of values, it does not identify individual entities that contribute to each bin.

With the above ideas in mind, we construct a custom feature histogram view (Figure 8) that shows individual entity contributions. In our histogram the bin width is computed based on the Freedman-Diaconis rule [13]. For a dataset of 68 subjects, this resulted in up to 20-40 values per each bin for a given feature. The relatively small bin size allows us to display each entity individually as a rectangle of the respective width, and with a height computed based on the chart dimensions. The chart’s width and height are equal for all feature types and are computed by the application based on the screen size. In our feature histogram, rectangles can be colored using one of the color schemes: individually for each subject, disease status (healthy - green, sick - purple), gender (male - blue, female - pink), or

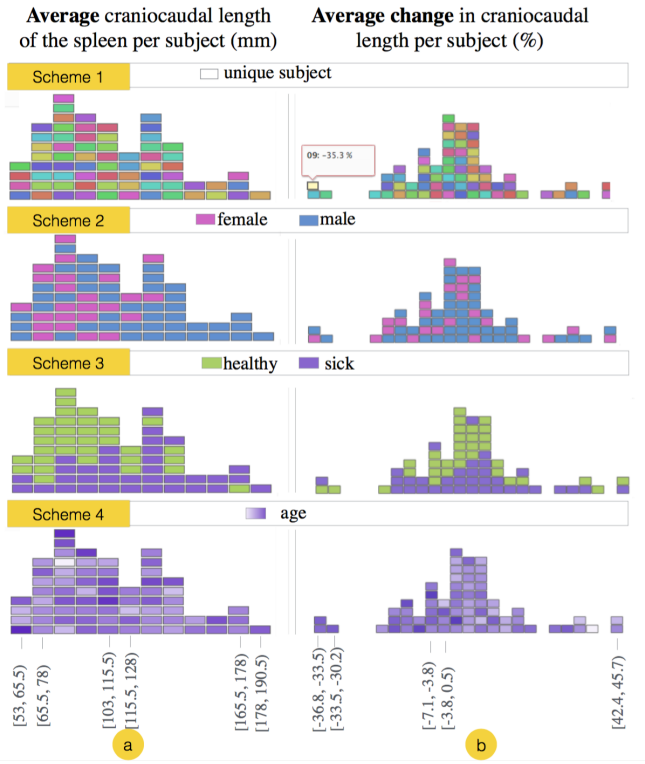


Fig. 8. An example of feature histogram charts for an average craniocaudal length of the spleen per subject, and average relative change over time. The exact value of the feature is available in the tooltip when hovering over a subject’s rectangle. The color scheme of the chart is selected by the user and is applied to each feature histogram. The figure presents 4 available color schemes: unique by subject, by gender, by disease status, and by age.

based on age (white to dark blue). *AnaFe* allows for color selection based on one of these specified color schemes and switching between two summarized types of feature distribution visualization as required. The features can be summarized by average value per subject and average relative change between consecutive timepoints.

The FDO view is one of the key visualization components of the system. It provides a global overview without eliminating individual per subject values. Globally, this view provides insight into the distribution of features. Based on interaction (hovering over a rectangle), the user can investigate into which range a given subject belongs for each feature (the corresponding subject is highlighted in all views). In clinical research, the expert can specify population groups to be included in the similarity comparison. For instance, she might want to compare all the “long” spleens of sick subjects that were undergoing treatment. By brushing over the “length” FDO chart and over the “disease status” DO chart, the population is filtered and will be used for comparison.

4.4.6 Feature Progression over Time (FPT)

FPT is a heatmap visualization of feature vectors for each subject’s visit (Figures 7 and 9 (3)). The FPT view is not immediately visible and can be shown on demand through selection on the control panel. The row number of within the chart indicates the visit, and the cell is a value of the feature at a given point in time. Cell color is determined based on the normalized value of the feature (the data value is mapped to interval from 0 to 1). The darkest red identifies the highest value of the feature (normalized to 1) and yellow indicates lowest value (normalized to 0). The original value can be viewed by hovering over the cell. Through this color coding, the FPT view shows the change of each individual feature for the subject over time in a single view. For instance, measurements such as volume, length, width, and thickness of the organ (shown in the first group in Figure 9) and quantitative radiomics shape features can be surveyed in parallel across a treatment regimen.

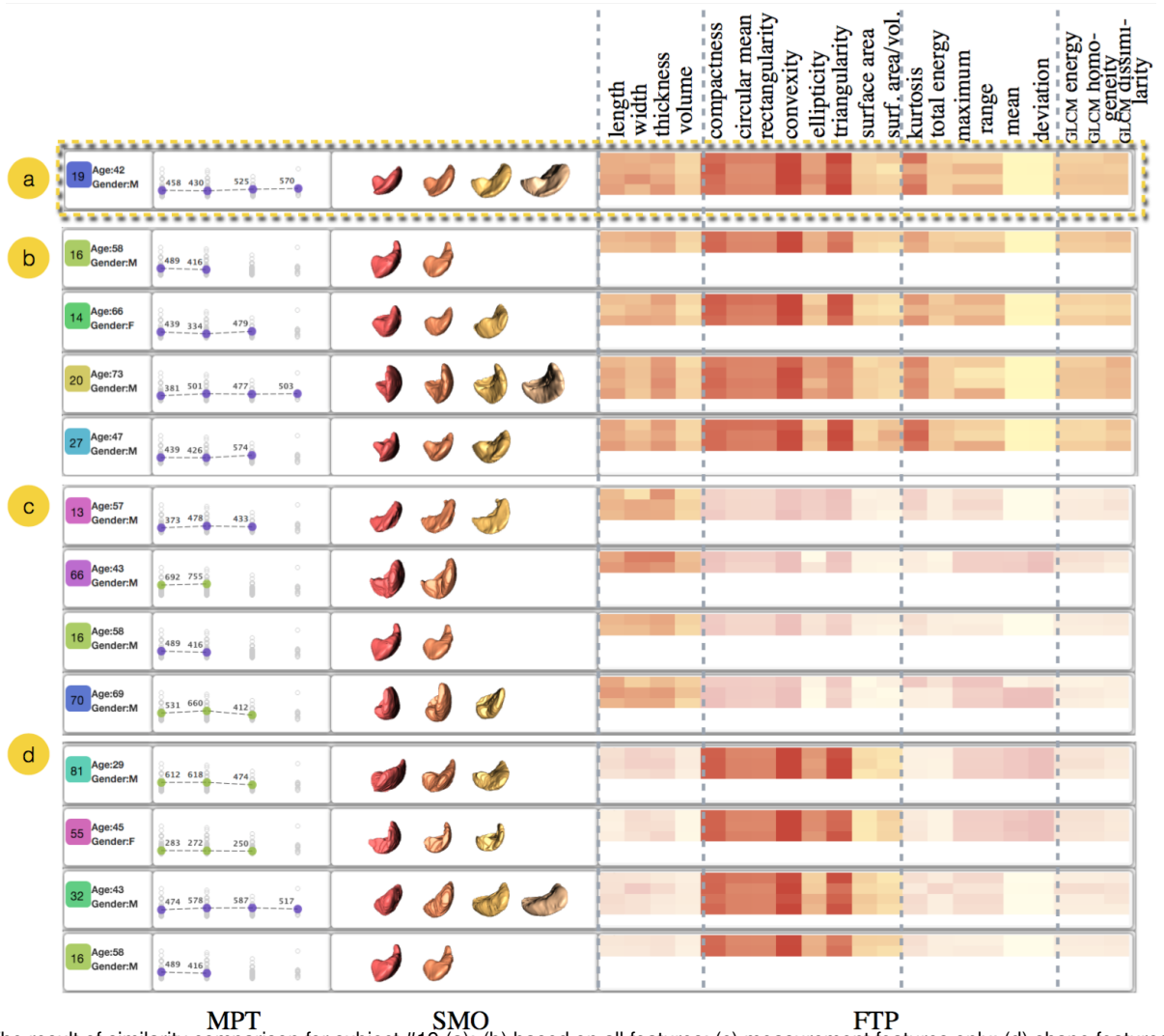


Fig. 9. The result of similarity comparison for subject #19 (a): (b) based on all features; (c) measurement features only; (d) shape features only.

4.5 Visual Queries and Interaction

Our visual query interface is comprised of multiple feature distribution histogram charts. We have identified 4 groups of imaging-derived features. In this application we currently use up to 8 features in each group with the possibility of adding more features in the future.

Query construction. The feature query interface allows the user to specify three types of information required for the comparison of time-series similarity: inclusion/exclusion of each feature, weight of feature contribution, and range of feature interval. The user can *select a feature* (include/exclude) via a checkbox in the legend in order for the feature to be included. The *feature weight* can be specified via a slider located next to each feature histogram. *Feature range selection*, brushing selection over a block of the histogram, allows the user to select a range of values for each feature. If a histogram bin is selected, all sets within that range will be included in the similarity comparison. Mouse hover *links* all feature values for a given subject on all histograms.

Query results. A similarity query compares the currently selected study (displayed first) with every other multi-timepoint study available in the dataset. If a value of a feature falls outside the range selected by brushing, and leaves only a single timepoint per subject, the subject will be compared based on that one point. After the user completes the feature selection, she can request the similarity results. Results are displayed in all table/list views: subject information, MPT, and SMO.

4.6 Implementation

The *AnaFe* front-end is implemented using standards-compliant Javascript, CSS, HTML5, and WebGL and utilizing Bootstrap¹ layouts. The core application scaffolding is driven by Angular.js². The visualizations are generated using three.js³ and Highcharts.js⁴. The core web application functionality is implemented on top of Node.js⁵, with Sails.js⁶. Data pre-processing and derivation of all imaging features is handled in Python utilizing scikit-image⁷.

5 APPLICATION

In this section, first, we describe the data collection used. Then we present two usage scenarios, showcasing our application in the process of analyzing a collection of splenic imaging datasets, and obtain related domain expert feedback. For both of the scenarios we show our system by providing an overview of the data, allowing the user to zoom, filter, and examine example results and details on demand. Finally, we present an example of *AnaFe*'s extensibility to other organs and medical imaging types.

¹<http://getbootstrap.com>

²<https://angularjs.org>

³<http://threejs.org>

⁴<http://www.highcharts.com>

⁵<https://nodejs.org/>

⁶<http://sailsjs.org>

⁷<http://scikit-image.org>

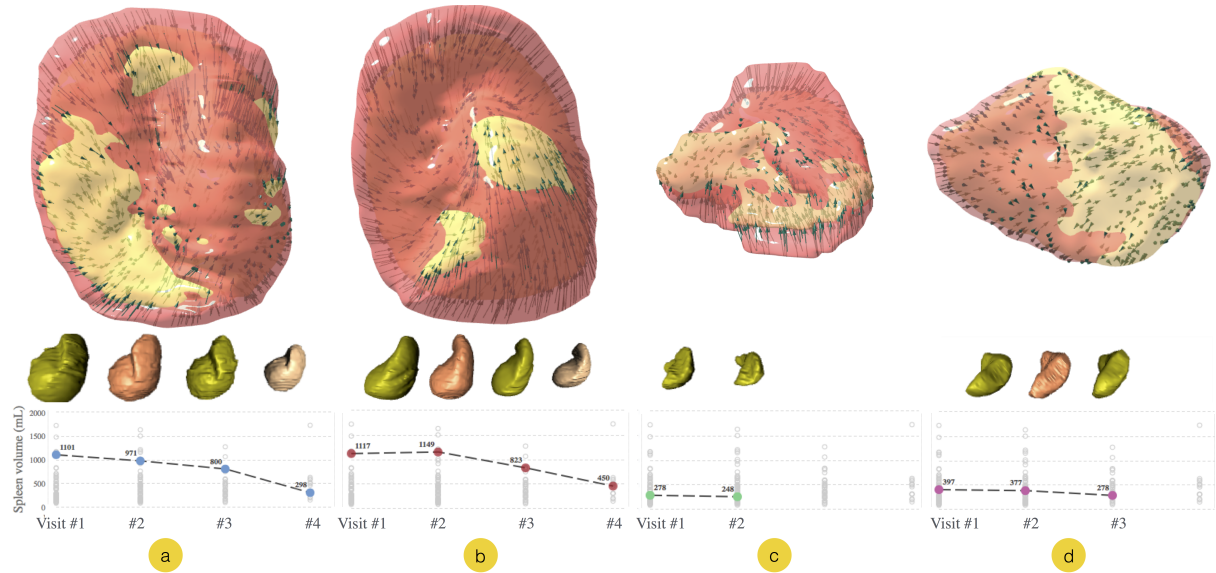


Fig. 10. An example of OD, SMO, and MPT views for 4 subjects. Case (b) was found to be the most similar to case (a) in the sick subject data. Both show the same decrease in organ volume. Cases (c) and (d) were found to be most similar among healthy subjects. While the spleen (c) decreased in volume, (d) remained almost unchanged.

5.1 Spleen Data

The data used in our study was collected retrospectively over a period of three years. This data includes two groups of subjects: those with *no known causes of splenomegaly or diseases* affecting the spleen (healthy), and those with *lymphoma or leukemia*, diseases known to affect the organ size, that were undergoing treatment (sick). Only subjects with two or more abdominal CT scans over a three year period were included. This study was approved by our IRB. The DICOM image data was de-identified and stored in a secure off-line database.

A total of 33 healthy subjects with up to 3 follow-ups over the given time period were included, resulting in 79 healthy datasets. For “sick” subjects, we included data with up to 4 follow-up scans over the course of treatment / disease observation, resulting in 35 subjects with a total of 110 datasets. The data included 24 female subjects and 44 male.

For each study, manual tracing of the splenic contour was performed on each axial slice in a blinded fashion under the supervision of a fellow-trained attending radiologist, providing a ground truth segmentation. This work was performed using the Alice software⁸. While the time between the follow up scans varied from one week to over a year, it was noted by our collaborators that only the difference between measurements was important, and the time between visits was not of high significance, as explained earlier.

5.2 Case 1: Similarity Comparison

In this section we demonstrate the similarity comparison feature for multiple datasets and features. We focus on the usability and interactivity of the system by providing step-by-step scenarios of usage.

Data selection. First, the user should select a multi-subject study and select a set for comparison. By default, all of the study parameters and imaging-derived features will be included in the search. *AnaFe* provides full customization of the query parameters to the user.

Visual queries. The search for similar shapes can be narrowed down through *AnaFe*’s visual query interface in the FDO view. Based on the selection of search parameters (types of measurements and features), some timepoints will be excluded from comparison. Once the query selection is complete, the entire time series for a particular subject is compared to that of all other subjects. The similarity is computed in real-time (for 189 spleens), and the result is displayed immediately in the list/table views. Initially, the preview of temporal sets in the 3D SMO view provides immediate feedback to the user.

Example of findings. Figure 9(a) shows an example of the four most similar sets as compared to subject #19. First, similarity comparison was performed using all available features (Figure 9(b)), next based on

diameters and volume of the organ (c), and finally based only on shape features (c). The color of the marker in the MPT scatterplot indicates disease status of the subject (with green being healthy). In cases (c) and (d) the system has identified healthy subjects #66, #70, #81, and #55 as similar. In the first case, with all features selected, the system has identified only diseased subjects. Similarity comparison in this case did not include demographics information and was based solely on imaging features.

Zooming-in on the result. Next, the user can select the organs by double clicking on them for zoomed-in investigation in the 3D OD view. Figure 10 shows rendering of temporal scans for four subjects. All measurements (volume, width, length, thickness) were used for comparison query. 3D meshes of the organ are all registered sequentially. The reviewer can inspect the direction of change. For example, for sick subjects (a) and (b) in Figure 10, change happens in all dimensions. In case (c), some change has occurred in the healthy subject as well. For such a small organ volume, this change may indicate the variation in hydration status. However, for most of the healthy subjects found in our data, there is no change between consecutive visits.

5.3 Case 2: Interactive Feature Exploration

Radiomics uses medical imaging features for the prediction of diagnostic outcomes [10, 31]. In the examined applications these features provide a non-invasive way of quantifying and monitoring tumors. We use this comprehensive quantification of imaging to find similarities in progression of disease and response to treatment in the spleen data. However, there haven’t been any previous attempts at visualizing these features for detailed user examination.

Feature overview. *AnaFe* provides an overview of all features in the FDO view, thus creating a second scenario for the view usability. From this view, we can see that, as expected, the mean and standard deviation of intensity communicate homogeneous density properties of the organ on the CT scan. Thus, these features might not be discriminative enough to show differences between subjects.

Feature selection. The user can unselect these features from the similarity comparison, and the corresponding columns will be dimmed out on the FPT heatmap (Figure 7). The user can inspect relationships between features of different subjects by brushing the selection over the FDO histograms.

Example of findings. As shown in Figure 7, organs with larger volumes also have larger craniocaudal length, which is a well known relationship for the spleen. However, relationships between other features for this organ are largely unknown. *AnaFe* creates a unique opportunity for the user to inspect the derived features simultaneously and select particular ones for further statistical testing.

⁸Paraxel Informatics; Waltham, MA

5.4 Extensibility to Other Domains

AnaFe was developed with the goal of exploring changes in temporal imaging-derived features, which does not limit its application to CT spleen data. To show extensibility of our application, we use observations of prostate cancer on Magnetic Resonance Imaging (MRI). Prostate cancer is the most common malignancy and second most common cause of cancer-related mortality in men. Similar to splenic increase in size, enlargement in the prostate can be palpable and is used in primary staging. Thus, the size of the organ is one of the useful indicators that can be obtained from the segmented organ imaging.

We use temporal prostate data from the Prostate MR Image Database⁹. Manual segmentation was available only for two data time-points per subject, and demographics information was not available. Additionally, no disease status was available for these subjects. The study by domain experts is required to evaluate results of the analysis. Figure 11 shows an example of *AnaFe* with prostate data.

5.5 Expert feedback

Two of our collaborating radiologists inspected the application and its features. We have recorded their feedback and the most important points they made about the usability of the tool.

Linking measurements and 3D data over time: MPT, SMO, and OD views. The MPT view for craniocaudal length and volume (well known measurements) allows the researcher to compare their trends rapidly and gain immediate visual feedback. It becomes particularly powerful in combination with 3D rendering of the organ sequence (SMO). The overlay rendering of aligned mesh surfaces (3D OD view) provides a useful picture of how the organ has changed between patient visits. In the software that is currently used to perform medical research on the spleen, neither registration of the organ, nor 3D rendering of overlays are implemented. Thus, the MPT, SMO, and OD views, when used jointly, *provide the missing links for the current analytical workflow*, and allow for substantial time savings when comparing multiple organs. As a potential improvement, our collaborators indicated that some sort of on-mesh annotation of the areas of major change in shape or size within the 3D view would be very useful.

Feature distribution and characterization of the data: FDO view. The FDO view was selected as the most interesting feature due to its ability to analyze multiple shape features simultaneously. As was noted by our collaborators, *differentiation between sick and healthy organs* based on a set of features was unclear. The FDO chart provided visualization of distribution and ranges of each feature, as well as immediate visual feedback (based on the color of the category) if the separation between the values based on the category exists. For instance, individual color coding by subject has shown an *already well known correlation of the organ's craniocaudal length and volume*. That is for spleens with a large volume, the length is also large. For some of the shape features, differentiation between sick and healthy subjects was very visible. Using this knowledge, and the subject's diagnostic information (currently not available in the anonymized data), the clinician could investigate sources of such differences, for example, the effect of portal vein hypertension on the organ shape. This is an interesting direction for future research: analysis of statistical significance of shape features in predicting organ disease status, and predicting normality of diseased organs and abnormality of healthy ones.

In summary, the ability to visualize a large number of datasets simultaneously has significantly improved over routine comparison that has previously required context switching between different applications and was limited to comparison of only a few parameters in static charts (IS1). Most importantly, *AnaFe* has provided a unique look at multiple snapshots of spleen data and its imaging-derived features changing over time (IS2).

6 CONCLUSIONS AND FUTURE WORK

In this paper, we described a visual analytics framework for the exploration of large collections of medical imaging datasets, with a focus on disease progression accompanied by progressive spleen enlargement. Through multiple linked views we allow the user (researcher or the



Fig. 11. An example of *AnaFe* use with prostate data. User interaction with the system is displayed in the average (top) and change FDO views (bottom). Hovering the mouse over the marker (subject #73) in the MPT view (a) links to all features for this subject in the FDO view. Selected subjects in the view (c) are highlighted across all FDO charts. Not selected subjects (b) are dimmed out in the FPT view.

radiologist) to interactively explore multiple imaging-derived features. Meanwhile, a 3D mesh view allows the user to examine changes in the organ more closely and see the relation of the traditionally used 1D parameters to the full organ view. Currently, our application supports the set of most common robust and reproducible imaging features. These features describe organ intensity, shape, texture, and measurements. Our application allows for observation of trends over time to determine similarity in disease progression and outcomes. We have currently explored only a limited set of radiomics features, and were already able to find their usability for the spleen. It is also in our interest to pursue a time-dependent comparison in the future with additional parameters from other tests considered (beyond CT imaging), which requires acquisition of more data and additional IRB approval.

ACKNOWLEDGMENTS

We would like to thank our collaborating radiologists Dr. Matthew Barish and Kevin Baker for defining research issues, manual segmenting the data, and evaluating our prototype. This research has been partially supported by NSF grants CNS-0959979, IIP1069147 and CNS-1302246, the Center of Excellence for Wireless and Information Technology (CEWIT), and the Marcus Foundation.

REFERENCES

- [1] P. Angelelli, S. Oeltze, J. Haasz, C. Turkay, E. Hodneland, A. Lundervold, A. J. Lundervold, H. Hause, and B. Preim. Interactive visual analysis of heterogeneous cohort study data. *IEEE Comput Graph Appl*, 1(5):70–82, Jan. 2014.
- [2] J. Barcroft and J. G. Stephens. Observations upon the size of the spleen. *J Physiol*, 64(1):1–22, 1927.
- [3] A. S. Bezerra, G. DiIppolito, S. Faintuch, J. Szejnfeld, and M. Ahmed. Determination of splenomegaly by CT: Is There a Place for a Single Measurement? *AJR Am J Roentgenol*, 184(5):1510–1513, 2005.

⁹<http://prostatemrimagedatabase.com>

- [4] J. Blaas, C. P. Botha, and F. H. Post. Interactive visualization of multi-field medical data using linked physical and feature-space views. *EuroVis*, pages 123–130, 2007.
- [5] S. Busking, C. P. Botha, and F. H. Post. Dynamic multi-view exploration of shape spaces. *Comput Graph Forum*, 29(3):973–982, Aug. 2010.
- [6] J. J. Caban, P. Rheingans, and T. Yoo. An evaluation of visualization techniques to illustrate statistical deformation models. *Comput Graph Forum*, 30(3):821–830, June 2011.
- [7] J. L. Carrasco, B. R. Phillips, J. Puig-Martinez, T. S. King, and V. M. Chinchilli. Estimation of the concordance correlation coefficient for repeated measures using SAS and R. *Comput Methods Programs Biomed*, 109(3):293–304, Mar. 2013.
- [8] X.-l. Chen, T.-w. Chen, X.-m. Zhang, Z.-l. Li, N.-l. Zeng, T. Li, D. Wang, J. Li, Z.-j. Fang, H. Li, J. Chen, J. Liu, G.-h. Xu, J. Ren, J.-l. Wu, and C.-p. Li. Quantitative assessment of the presence and severity of cirrhosis in patients with Hepatitis B using right liver lobe volume and spleen size measured at MRI. *PLoS ONE*, 9(3):e89973–6, 2014.
- [9] B. D. Cheson, R. I. Fisher, S. F. Barrington, F. Cavalli, L. H. Schwartz, E. Zucca, and T. A. Lister. Recommendations for initial evaluation, staging, and response assessment of Hodgkin and Non-Hodgkin Lymphoma: The Lugano classification. *J Clin Oncol*, 32(27):3059–3067, Sept. 2014.
- [10] C. Christin, H. C. J. Hoefsloot, A. K. Smilde, B. Hoekman, F. Suits, R. Bischoff, and P. Horvatovich. A critical assessment of feature selection methods for biomarker discovery in clinical proteomics. *Mol Cell Proteomics*, 12(1):263–276, Jan. 2013.
- [11] J. A. Fails, A. Karlson, L. Shahamat, and B. Shneiderman. A visual interface for multivariate temporal data: finding patterns of events across multiple histories. *IEEE S Vis Anal*, pages 167–174, 2006.
- [12] Z. Fang, T. Möller, G. Hamarneh, and A. Celler. Visualization and exploration of time-varying medical image data sets. *ACM Proceedings of Graphics Interface*, pages 281–288, 2007.
- [13] D. Freedman and P. Diaconis. On the histogram as a density estimator: L2 theory. *Probab. Theory Related Fields*, 57(4):453–476, 1981.
- [14] E. M. Geraghty, J. M. Boone, J. P. McGahan, and K. Jain. Normal organ volume assessment from abdominal CT. *Abdom Imaging*, 29(4):1–11, 2004.
- [15] D. Gotz and H. Stavropoulos. Decisionflow: visual analytics for high-dimensional temporal event sequence data. *IEEE Trans Vis Comput Graphics*, 20(12):1783–1792, 2014.
- [16] D. L. Gresh, B. E. Rogowitz, R. L. Winslow, D. F. Scollan, and C. K. Yung. WEAVE: A system for visually linking 3-D and statistical visualizations, applied to cardiac simulation and measurement data. *Proceedings of the Conference on Visualization*, pages 489–492, 2000.
- [17] I. Gutenko, H. Peng, X. Gu, M. Barish, and A. Kaufman. Maximal area and conformal welding heuristics for optimal slice selection in splenic volume estimation. *Proceedings SPIE Medical Imaging*, 9785:97853V–97853V–8, 2016.
- [18] M. Hermann, A. C. Schunke, T. Schultz, and R. Klein. A visual analytics approach to study anatomic covariation. *IEEE Pac Vis Symp*, pages 161–168, 2014.
- [19] Y. S. Kaneko, Junichi, Y. Matsui, T. Ohkubo, and M. Makuuchi. Normal splenic volume in adults by computed tomography. *Hepatogastroenterology*, 49(48):1726–1727, 2001.
- [20] D. F. Keefe, M. Ewert, W. Ribarsky, and R. Chang. Interactive coordinated multiple-view visualization of biomechanical motion data. *IEEE Trans Vis Comput Graphics*, 15(6):1383–1390, 2009.
- [21] P. Klemm, S. Oeltze-Jafra, K. Lawonn, K. Hegenscheid, H. Volzke, and B. Preim. Interactive visual analysis of image-centric cohort study data. *IEEE Trans Vis Comput Graphics*, 20(12):1673–1682, 2015.
- [22] J. Krause, A. Perer, and H. Stavropoulos. Supporting iterative cohort construction with visual temporal queries. *IEEE Trans Vis Comput Graphics*, 22(1):91–100, 2016.
- [23] H. Li, R. W. Sumner, and M. Pauly. Global correspondence optimization for non-rigid registration of depth scans. *Computer Graphics Forum (Proc. SGP 08)*, 27(5), July 2008.
- [24] W. E. Lorensen and H. E. Cline. Marching cubes: a high resolution 3D surface construction algorithm. *ACM SIGGRAPH Computer Graphics*, 21(4):163–169, 1987.
- [25] S. Malik, F. Du, M. Monroe, E. Onukwugha, C. Plaisant, and B. Shneiderman. Cohort comparison of event sequences with balanced integration of visual analytics and statistics. *Proceedings 20th International Conference on Intelligent User Interfaces*, pages 38–49, 2015.
- [26] J. M. Meier, A. Alavi, S. Iruvuri, S. Alzeair, R. Parker, M. Houseni, M. Hernandez-Pampaloni, A. Monge, and D. A. Torigian. Assessment of age-related changes in abdominal organ structure and function with computed tomography and positron emission tomography. *Semin Nucl Med*, 37(3):154–172, 2007.
- [27] N. A. Michels. The variational anatomy of the spleen and splenic artery. *Am J Anat*, 70(1):21–72, 1942.
- [28] M. Monroe, R. Lan, H. Lee, C. Plaisant, and B. Shneiderman. Temporal event sequence simplification. *IEEE Trans Vis Comput Graphics*, 19(12):2227–2236, 2013.
- [29] Y. Murata, M. Abe, Y. Hiasa, N. Azemoto, T. Kumagi, S. Furukawa, B. Matsuura, K. Michitaka, N. Horiike, and M. Onji. Liver/spleen volume ratio as a predictor of prognosis in primary biliary cirrhosis. *J Gastroenterol*, 43(8):632–636, 2008.
- [30] R. A. O'Reilly. Splenomegaly in 2,505 patients at a large university medical center from 1913 to 1995. 1963 to 1995: 449 patients. *Western J Med*, 169(2):88, 1998.
- [31] C. Parmar, E. Rios Velazquez, R. Leijenaar, M. Jermoumi, S. Carvalho, R. H. Mak, S. Mitra, B. U. Shankar, R. Kikinis, B. Haibe-Kains, P. Lambin, and H. J. W. L. Aerts. Robust radiomics feature quantification using semiautomatic volumetric segmentation. *PLoS ONE*, 9(7), July 2014.
- [32] A. Perer and D. Gotz. Data-driven exploration of care plans for patients. *CHI Extended Abstracts on Human Factors in Computing Systems*, pages 439–444, 2013.
- [33] D. S. Raicu. *Data Min. Knowl. Discov.*, chapter Mining knowledge in computer tomography image databases, pages 487–508. Springer London, 2007.
- [34] R. G. Raidou, U. A. Van Der Heide, C. V. Dinh, G. Ghobadi, J. F. Kallehauge, M. Breeuwer, and A. Vilanova. Visual analytics for the exploration of tumor tissue characterization. *Comput Graph Forum*, 34(3):11–20, Apr. 2015.
- [35] P. Rezaei, S. M. Tochetto, M. S. Galizia, and V. Yaghmai. Splenic volume model constructed from standardized one-dimensional MDCT measurements. *AJR Am J Roentgenol*, 196(2):367–372, 2011.
- [36] F. Robertson, P. Leander, and O. Ekberg. Radiology of the spleen. *Eur Radiol*, 11:80–95, 2001.
- [37] H. Sakoe and S. Chiba. Dynamic programming algorithm optimization for spoken word recognition. *IEEE Trans Sig Process*, 26(1):43–49, 1978.
- [38] J. Serrà and J. L. Arcos. An empirical evaluation of similarity measures for time series classification. *Knowledge-Based Systems*, 67:305–314, 2014.
- [39] S. Shen, G. L. DeNardo, A. Yuan, C. Hartmann-Siantar, R. T. O'Donnell, and S. J. DeNardo. Splenic volume change and nodal tumor response in Non-Hodgkin's lymphoma patients after radioimmunotherapy using radiolabeled Lym-1 antibody. *Cancer Biother Radiopharm*, 20(6), 2005.
- [40] M. D. Steenwijk, J. Milles, M. A. van Buchem, J. H. C. Reiber, and C. P. Botha. Integrated visual analysis for heterogeneous datasets in cohort studies. *IEEE VisWeek Workshop on Visual Analytics in Health Care*, pages 1–8, Oct. 2010.
- [41] G. Taubin. A signal processing approach to fair surface design. *Proceedings 22nd Annual Conference on Computer Graphics and Interactive Techniques*, pages 351–358, 1995.
- [42] P. Tormene, T. Giorgino, S. Quaglini, and M. Stefanelli. Matching incomplete time series with dynamic time warping: an algorithm and an application to post-stroke rehabilitation. *Artif Intell Med*, 45(1):11–34, 2009.
- [43] E. Tufte. Envisioning information. *Optom Vis Sci*, 68(4):322–324, 1991.
- [44] C. Turkay, A. Lundervold, A. J. Lundervold, and H. Hauser. Hypothesis generation by interactive visual exploration of heterogeneous medical data. *Human-Computer Interaction and Knowledge Discovery in Complex, Unstructured, Big Data*, pages 1–12, 2013.
- [45] K. Wongsuphasawat and D. Gotz. Exploring flow, factors, and outcomes of temporal event sequences with the outflow visualization. *IEEE Trans Vis Comput Graphics*, 18(12):2659–2668, 2012.
- [46] K. Wongsuphasawat, J. A. Guerra Gómez, C. Plaisant, T. D. Wang, M. Taieb-Maimon, and B. Shneiderman. Lifeflow: visualizing an overview of event sequences. *Proc SIGCHI Conf Hum Factor Comput Syst*, pages 1747–1756, 2011.
- [47] E. M. Yetter, K. B. Acosta, M. C. Olson, and K. Blundell. Estimating splenic volume: sonographic measurements correlated with Helical CT determination. *AJR Am J Roentgenol*, pages 1615–1620, 2003.
- [48] Z. Zhang, B. Wang, F. Ahmed, R. Zhao, A. Viccellio, and K. Mueller. The Five W's for information visualization with application to healthcare informatics. *IEEE Trans Vis Comput Graphics*, 19(11):1895–1910, 2013.