



저작자표시-비영리-변경금지 2.0 대한민국

이용자는 아래의 조건을 따르는 경우에 한하여 자유롭게

- 이 저작물을 복제, 배포, 전송, 전시, 공연 및 방송할 수 있습니다.

다음과 같은 조건을 따라야 합니다:



저작자표시. 귀하는 원저작자를 표시하여야 합니다.



비영리. 귀하는 이 저작물을 영리 목적으로 이용할 수 없습니다.



변경금지. 귀하는 이 저작물을 개작, 변형 또는 가공할 수 없습니다.

- 귀하는, 이 저작물의 재이용이나 배포의 경우, 이 저작물에 적용된 이용허락조건을 명확하게 나타내어야 합니다.
- 저작권자로부터 별도의 허가를 받으면 이러한 조건들은 적용되지 않습니다.

저작권법에 따른 이용자의 권리는 위의 내용에 의하여 영향을 받지 않습니다.

이것은 [이용허락규약\(Legal Code\)](#)을 이해하기 쉽게 요약한 것입니다.

[Disclaimer](#)

Ph.D. DISSERTATION

An Interactive Visual Analytics Framework
for Diagnostic Gaze Data
on Volumetric Medical Images

3차원 의료 영상 관독 시선 정보의
대화형 시각적 분석 프레임워크

FEBRUARY 2016

DEPARTMENT OF ELECTRICAL ENGINEERING &
COMPUTER SCIENCE
COLLEGE OF ENGINEERING
SEOUL NATIONAL UNIVERSITY

Hyunjoo Song

An Interactive Visual Analytics Framework
for Diagnostic Gaze Data on Volumetric Medical Images

3 차원 의료 영상 판독 시선 정보의
대화형 시각적 분석 프레임워크

지도교수 서진욱

이 논문을 박사학위논문으로 제출함

2015년 12월

서울대학교 대학원

전기·컴퓨터 공학부

송현주

송현주의 박사학위논문을 인준함

2015년 12월

위원장 : 김명수 (인)

부위원장 : 서진욱 (인)

위원 : 서봉원 (인)

위원 : 김보형 (인)

위원 : 이정진 (인)

Abstract

An Interactive Visual Analytics Framework for Diagnostic Gaze Data on Volumetric Medical Images

Hyunjoo Song

Department of Computer Science and Engineering

College of Engineering

The Graduate School

Seoul National University

We propose an interactive visual analytics framework for diagnostic gaze data on volumetric medical images. The framework is designed to compare gaze data from multiple readers with effective visualizations, which are tailored for volumetric gaze data with additional contextual information. Gaze pattern comparison is essential to understand how radiologists examine medical images and to identify factors influencing the examination. However, prior work on diagnostic gaze data using the

medical images acquired from volumetric imaging systems (e.g., computed tomography or magnetic resonance imaging) showed a number of limitations in comparative analysis. In the diagnosis, radiologists scroll through a stack of images to get a 3D cognition of organs and lesions that resulting gaze patterns contain additional depth information compared to the gaze tracking study with 2D stimuli. As a result, the additional spatial dimension aggravated the complexity on visual representation of gaze data. A recent work proposed a visualization design based on direct volume rendering (DVR) for gaze patterns in volumetric images; however, effective and comprehensive gaze pattern comparison is still challenging due to lack of interactive visualization tools for comparative gaze analysis.

In this paper, we first present an effective visual representation, and propose an interactive analytics framework for multiple volumetric gaze data. We also take the challenge integrating crucial contextual information such as pupil size and windowing (i.e., adjusting brightness and contrast of image) into the analysis process for more in-depth and ecologically valid findings. Among the interactive visualization components, a context-embedded interactive scatterplot (CIS) is especially designed to help users to examine abstract gaze data in diverse contexts by embedding medical imaging representations well-known to radiologists in it. We also present the results from case studies with chest and abdominal radiologists

**keywords : Eye tracking, Gaze visualization, Gaze pattern comparison,
Volumetric medical images, Context-embedded interactive scatterplot,
interactive temporal chart**

student number : 2011-30964

Contents

Abstract	i
Contents	iv
List of Figures	vii
List of Tables	xi
Chapter 1 Introduction	1
1.1 Background	1
1.2 Research Components	5
1.3 Radiological Practice.	6
1.4 Organization of the Dissertation	8
Chapter 2 Related Work	9
2.1 Visualization Combining 2D and 3D	9
2.2 Eye Tracking Data Visualization	14
2.3 Comparative Data Analysis	16
2.4 Gaze Analysis in the Medical field.	18

Chapter 3	GazeVis: Volumetric Gaze Data	21
3.1	Visualization of Stimuli and Gaze Data	23
3.1.1	Computation of Gaze Field	26
3.1.2	Visualization of Gaze Field	29
3.1.3	Gaze Field for Interactive Information Seeking	30
3.2	Interactions and Dynamic Queries	32
3.2.1	Interaction Design	32
3.2.2	Spatial Filtering.	34
3.2.3	Temporal Filtering.	34
3.2.4	Transfer Function Control.	36
3.2.5	Gaussian Blur Control.	38
3.3	Implementation	38
3.4	Evaluation with Radiologists	38
3.4.1	Case Study Protocol	39
3.4.2	Datasets	41
3.4.3	Apparatus.	42
3.4.4	Chest Radiologists.	42
3.4.5	Abdominal Radiologists	45
3.5	Discussion.	49
3.5.1	Spatial Data Structure and Flexibility.	49
3.5.2	Interacting with Contextual Data..	51
Chapter 4	GazeDx: Interactive Gaze Analysis Framework	53
4.1	Design Rationale.	54
4.2	Overviews for Comparative Gaze Analysis	57

4.2.1	Spatial Similarity.	57
4.2.2	Qualitative Similarity Overview.	58
4.2.3	Multi-level Temporal Overview.	60
4.3	In-depth Comparison of Gaze Patterns	65
4.3.1	Detail Views for Individual Readers.	65
4.3.2	Aggregation for Group Comparison	67
4.4	CIS: Context-embedded Interactive Scatterplot	68
4.4.1	Flexible Axis Configuration.	68
4.4.2	Focus Attention with Familiar Representations.	69
4.4.3	Scatterplot Matrix with CIS.	72
4.5	Interactive Selection and Filtering..	72
4.5.1	Selection by Freehand Drawing.	73
4.5.2	Selection by Human Anatomy	74
4.6	Implementation.	76
4.7	Case Studies.	77
4.7.1	Case Study Protocol.	78
4.7.2	Apparatus.	80
4.7.3	Case Study 1: Chest Radiologists.	81
4.7.4	Case Study 2: Abdominal Radiologists	85
4.8	Discussion	88
Chapter 5 Conclusion		91
Bibliography		94
Abstract in Korean		105

List of Figures

Figure 1.1	3D gaze plot superimposed on stacked 2D cross-sectional images.	3
Figure 1.2	Two different brightness and contrast settings for (a) lung and (b) mediastinum.	7
Figure 2.1	Partial occlusion of 3D cursor for better depth perception [81]. (a) fish in front of the cursor. (b) fish behind the cursor. (c) fish within the cursor	11
Figure 2.2	Visualizations combining 2D and 3D scatterplots for multidimensional data.	13
Figure 3.1	Multi-planar reconstruction (MPR) planes of human body. 3D human model is from [41].	22
Figure 3.2	GazeVis interface with 4 views on the left and a control panel on the right. Bright green dots represent gaze points. 4 views are axial, coronal, sagittal, and 3D views, clockwise from the upper left corner. Axial, coronal, and sagittal views have corresponding cross hairs and rectangular borders in their corresponding colors: green, yellow, and red, respectively. A control panel has (A) widgets for Gaussian filter control, (B) a range slider for temporal filtering, and (C), (D) widgets for manipulating transfer functions. In this	

figure, a user manipulates thickness of the axial plane in the coronal view. The thickened axial plane is represented as a green rectangular prism in the 3D view. Gazes inside the cube (pointed by light-green arrows) are in light green, while the gazes outside the cube are in pale green (pointed by pale-green arrows). . . . 25

Figure 3.3 (a) Gaze points on stimuli volume data (i.e., contiguous medical images). (b) Gaze field with the same x -, y -, and z -resolutions as the stimuli volume data. 28

Figure 3.4 Temporal gaze flow of a radiologist in reading a chest CT scan, revealing that he scrolled through the images from top to bottom. 35

Figure 3.5 3D rendering of lung with three-dimensionally superimposed gaze data. (a) With low opacity for the lung, gaze data within the lung are clearly visualized. (b) With high opacity for the lung, the gaze data within it are partially occluded. 37

Figure 3.6 Adjustments of Gaussian blur filter size. (a) 7 X 7 with standard deviation of 1. (b) 19 X 19 with standard deviation of 3. . . . 37

Figure 3.7 Gaze pattern difference of two experts in reading a chest CT scan. (a) Expert 1: Scroll down from top to bottom, jump back to the top, adjust the window setting, and scroll down again. (b) Expert 2: Scroll down from top to bottom, adjust the window setting, and scroll up from bottom to top. 44

Figure 3.8 Gaze pattern difference between radiologists with different expertise level. (a) Expert showing more organized pattern vertically. (b) Intermediate showing a pattern in between expert and novice. (c) Novice showing no distinctive pattern with scattered gaze points. Usability testing software, showing the task in the Usability category. Participants had to make a graph (lower graph) the same as the given ripple graph (upper graph). . . . 47

- Figure 3.9 (a) 3D rendering of lung with gaze points superimposed as independent objects. (b) 3D rendering of lung with gaze data used as attribute data, enabling gaze analyses more focused on human anatomy. 50
- Figure 3.10 Gaze data distribution in an axial plane. Gray histogram represents the distribution of all gaze points on entire slices. Green one represents the distribution of the gaze points only on the current axial image. 52
- Figure 4.1 GazeDx interface for chest normal case (Overview tab). (A) Similarity matrix showing pairwise similarity between readers computed with SSIM (Structural Similarity). (B) Correlation matrix showing correlation between each pair of dimensions. Background of each cell in both matrices is color coded by its magnitude. (C) Multi-temporal view with interactive temporal charts. (D) Spatial view showing gaze data superimposed on 2D MPR (axial, coronal, and sagittal) or 3D VR images. 59
- Figure 4.2 Semantic exploration of temporal aspect of gaze data by smooth transition of the representation of the multi-temporal view according to space availability. (a) Navigation chart. (b) Stacked bar chart. (c) Bar chart. 64
- Figure 4.3 GazeDx interface for chest normal case. (A) Scatter plot matrix. (B) CIS: context-embedded interactive scatter plot with gaze points grouped into five clusters. (C) Interactive temporal chart. (D) Aggregation pane showing aggregated gaze data (E) ROI selection pane with ‘Segmentation ROI filter’ and ‘Window preset’ filter. Interactive temporal charts in (C) and (F) show the difference in gaze pattern between scanner and driller strategy (frequently changing color in (C), scanner; relatively consistent color in (F), driller). (G) CIS with horizontal axis of pupil diameter and vertical axis of slice index, where the gaze data is color-coded by window

preset (left-shifted green cluster, lung setting; and right-shifted orange cluster, mediastinum setting). (H) CIS with embedded coronal plane image. 66

Figure 4.4 Animated context-embedded interactive scatterplot (CIS). (a) Gaze points are selected by users or filtered by a criterion in the x - z scatterplot. (b) The scatterplot gets divided into three parts, and the middle area is vertically expanded into a 3D VR image. (c) 3D VR representation rotates around the horizontal axis. (d) The x - y plane (i.e., axial image), which is invisible in the original x - z scatterplot, appears with gaze points overlaid to help users look into the data from different but related aspects in a single view. 70

Figure 4.5 Axial images embedded in CIS, showing gaze points of a reader for the chest lesion case used. (a) In the lung window setting, the gaze points are scattered largely on both the lung and the mediastinum even though the mediastinum becomes too saturated to see detail. (b) In the mediastinum window setting, the gaze points are distributed mostly on the mediastinum. 76

Figure 4.6 Axial images embedded in CIS, showing gaze points of (a) an expert and (b) a novice for the abdomen lesion case. (a) The expert’s gaze points are relatively scattered on the image even with the clearly notable lesion (large hepatic mass). (b) The novice’s gaze points are distributed mostly on the lesion. 87

Figure 5.1 Three-dimensional industrial computed tomography X-ray inspection scanning. Multiple cross-sectional images are acquired and reconstructed to 3D for inspection. The image is from [72]. 93

List of Tables

Table 3.1	DICOM files used in the case studies.	41
Table 4.1	Window presets.	62
Table 4.2	DICOM files used in the case study.	79

Chapter 1

Introduction

1.1 Background

Comparative analysis of gaze tracking data has been actively used in many research areas including the radiology field. As detection of significant clinical findings could lead to critical decisions [5], the radiologists performed such comparative gaze analyses to investigate various factors influencing the process and performance of their diagnoses, which could eventually affect the health and wellbeing of patients. Previous researches worked on finding the reasons why a person misses certain lesions and compared the difference between the radiologists with different diagnosis strategies. In the prior work, visualization played an important role in analyzing the gaze data to understand human mind behind gaze patterns. Thus, there have been efforts to develop effective visualization techniques for better understanding of the gaze data

[8], [19], [54], [66], [67]. For instance, superimposition of gaze data over visual stimuli has been shown to be effective in interpreting the data [67] and the technique has been adopted in most gaze data analysis systems as an in-context visualization [8].

Gaze data visualization is relatively easy to design when visual stimuli are simple such as a static scene or a chart; however, there are many other complicated stimuli for which much more visualization design effort is needed. For example, when the target stimuli change during a study session, e.g., a series of consecutive images shown in sequence, an effective overview of the gaze data becomes challenging to create. A conventional visualization technique of just showing the changing stimuli sequentially in a 2D image with gaze data superimposed is not effective in revealing important overall gaze patterns. Gaze analysis with radiologists who read medical images such as computed tomography (CT) and magnetic resonance imaging (MRI) images is a good example for this case. Radiologists have to examine a series of consecutive 2D cross-sectional images that compose a volume of a part of human body to reach a diagnosis. However, eye tracking studies using volumetric medical images as stimuli have seldom adopted versatile and interactive visual representations.

The stimuli used in the eye tracking study can be categorized according to their dimension [8]. Most of the previous study in the radiology field dealt with static 2D stimuli (i.e., images from X-ray) [16] and used common visualization techniques for spatial and temporal analysis: a heatmap for qualitative spatial analysis and a gaze

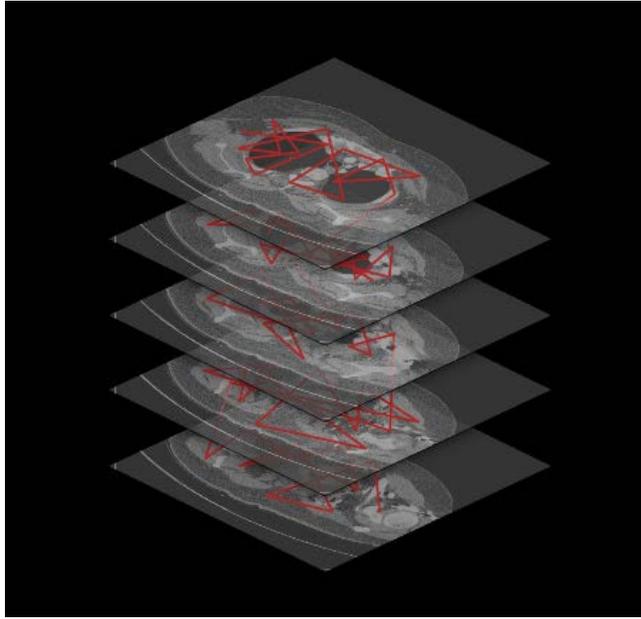


Figure 1.1. 3D gaze plot superimposed on stacked 2D cross-sectional images.

plot for analysis on overall temporal patterns [31]. For both visual representations, eye tracking data was usually aggregated into fixations and saccades using fixation filters. Then the researchers analyzed the preprocessed data using ROIs (region-of-interest) and quantitatively compared the data between readers with a number of factors such as first hit time and dwell time on an ROI. As volumetric images from CT and MRI are gaining popularity in the radiology, the importance of gaze analysis to understand the diagnoses on such images has increased. Some recent eye tracking study [7], [16], [50], [51] extended their scope to include such stimuli with innate 3D structures.

However, volumetric medical images have unique characteristics compared to well-known 3D stimulus for eye tracking study (e.g., 3D models or virtual reality

scenes). It is a stack of 2D cross-sectional images, which is shown one at a time during a diagnosis, and participants of the study (i.e., radiologists) can actively alter the shown image by wheel scrolling. Thus, the eye movements within the stimuli can provide more abundant clinical implication when treated and analyzed as 3D trajectories within 3D anatomical internal structures rather than as 2D gaze data from dynamic and active 2D scenes. Some prior work tried to resolve this problem by simply stacking the contiguous 2D cross-sectional images vertically to mimic the 3D volume while showing the abstract gaze data on each cross-sectional image using a traditional 2D gaze visualization, or gaze plot (Figure 1.1). However, this approach suffered from a severe occlusion problem wherein images above obscure ones below, failing to provide a bona fide overview of gaze data revealing the 3D nature of the cross-sectional images as a whole.

Thus, prior work mainly tried to resolve the visualization hurdles, but they still showed two limitations in comparative analysis: lack of interactivity and versatility, and neglect of contextual information. There are virtually no interactive and general-purpose tools for supporting comparative gaze analysis with volumetric medical images. Commercial gaze analysis tools support interactive and effective analysis with common stimuli, but they are incapable of managing innate 3D characteristics of volumetric medical images. Most prior work with such images came up with novel visualizations, but the visual representations were simply dedicated to show their specific findings. In addition, they were based on juxtaposition of two or more gaze

patterns in static views to make a comparison [16]. Moreover, there are many missed opportunities in comparative gaze analyses because important contextual information for diagnosis has not been properly considered in most studies. Such contextual information includes windowing information, pupil size, and distance to stimuli. For instance, radiologists perform windowing (i.e., adjustment of image brightness and contrast) to focus on different organs or lesions, but prior work controlled the experimental environment with fixed values rather than reproducing the actual diagnosis settings.

1.2 Research Components

As a remedy to this problem, we first present a 3D gaze data visualization (GazeVis). Volume rendering techniques developed in the SciVis community is employed to show the 3D large visual stimuli (i.e. organs and lesions) efficiently along with the gaze data. Visual encoding and user interaction design for the large gaze data, which are essential components of InfoVis, are adopted to enable researchers to reveal important gaze patterns more effectively that are hidden in the 3D volumetric space. However, naïve blending of selected techniques from the two visualization branches does not work that we extended traditional volume rendering to visualize the abstract information, i.e., gaze points within the volume, while delivering the depth perception of gaze points. Along with the visual representation, we designed dynamic queries to support efficient temporal and spatial filtering of gaze points within the context of the volumetric stimuli. We also report results from two case studies with 12 radiologists

following the MILCs evaluation guidelines [61], [64] as evaluation.

Then, we present an interactive gaze visualization framework (GazeDx), which is tailored for radiologists. It enables comparative analysis of gaze data from multiple readers' reviewing volumetric medical images. GazeDx is equipped with a combination of visualizations and interaction designs to support effective exploration of multiple gaze data. It also allows researchers to incorporate contextual information to cover real world environment of diagnosis. As this approach could be extended to gather more contextual information and to include them in the analysis, it will make the analysis richer in interpreting the implications of analysis results. In this work, such contextual information provided more insights to the analyst evaluating the quality of diagnosis. Ultimately, visual exploration of gaze data using the proposed framework can support hypothesis driven research in the medical field. As for evaluation, we report results from two case studies with two experienced radiologists, where they compared the gaze patterns of 14 radiologists reading two patients' CT images.

1.3 Radiological Practice

Radiologists perform diagnoses with medical images obtained from various types of imaging devices. Projection radiographs, i.e., X-rays, consist of a single image per examination. More advanced imaging modalities such as computed tomography (CT) and magnetic resonance imaging (MRI) generate a large number of contiguous cross-sectional images per examination, which actually convey 3-dimensional spatial

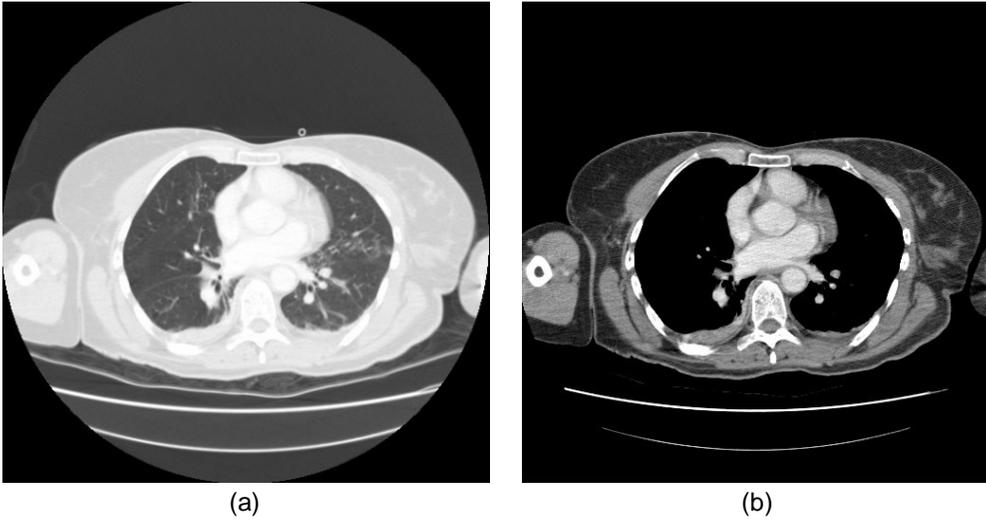


Figure 1.2. Two different brightness and contrast settings for (a) lung and (b) mediastinum.

information of human body (Figure 1.1). Radiologists scroll through the images back and forth quickly to find abnormalities in a so-called stack viewing mode where the images are logically stacked according to z -axis order and radiologists examine one slice at a time. Scrolling the images while examining each image is analogous to interactively navigating through the human body.

When radiologists scroll through the images, they often change brightness and contrast of the images to focus on different body parts. Thus, they often have to examine the same image multiple times for a full inspection. For instance, radiologists navigate to the bottom of the lung starting from neck with a certain brightness and contrast setting, and then they navigate back to the top of the lung with a different brightness and contrast setting (Figure 1.2. Two different brightness and contrast

settings for (a) lung and (b) mediastinum.). As a result, such brightness and contrast setting information could be used to determine whether one was actually looking at meaningful areas.

1.4 Organization of the Dissertation

This dissertation is organized as follows. In Chapter 2, we will describe the related works in visualization combining 2D and 3D, eye tracking data visualization, comparative data analysis, and gaze analysis in the medical field. Then, we will present 3D gaze visualization technique (Chapter 3) and gaze analysis framework (Chapter 4) followed by case study results and discussion. Chapter 5 summarizes this dissertation with plans for future work and conclusions.

Chapter 2

Related Work

In this section, we review previous work on visualization techniques combining 2D and 3D representation, followed by prior work on visual representations for eye tracking data. Then, we review previous work on comparative data analysis both from InfoVis and medical perspectives.

2.1 Visualization Combining 2D and 3D

Early works in the InfoVis field explored pros and cons of 3D visualization of abstract dataset [17], [22], [40], [57], [58], and most studies suggested that 3D visualizations have benefits only if the data has 3D spatial properties in nature and the tasks require understanding the 3D spatial structure [76]. Cockburn and McKenzie [11] compared the original 3D Data Mountain with a simple 2D redesign of it and found no significant difference in performing organization/retrieval tasks. In another study,

they even found that 3D showed lower performance than 2D in terms of task completion time for spatial memory tasks [12]. Cockburn [13] later revisited the spatial memory issue in re-running Tavanti et al.'s experiment [70] with some uncontrolled factors controlled, and found no performance difference in spatial memory tasks, which refuted the original study. Borkin et al. [9] proposed a novel 2D tree diagram representation and compared with 3D artery representations to find out that 2D visualizations outperform in terms of accuracy and efficiency. Tory and Möller [75] pointed out that interaction with 3D visualizations is difficult because of unnatural mappings between 2D mouse actions and 3D space. GazeVis differs from the conventional 3D InfoVis systems in that the abstract gaze data have inherent 3D spatial information, which could benefit from 3D visualization.

There are some work which tried to combine 2D and 3D visualization techniques in a system. Piringer et al. [52] remedied shortcomings of 2D scatterplots (e.g., overplotting problem) and 3D scatterplots (e.g., perception and interaction problem) by interactively linking the two with some extensions (Figure 2.2a). Elmqvist et al. [21] used animated 3D rotations during transitions between 2D scatterplots for multidimensional data (Figure 2.2b). While a scatterplot only displays a single aspect of the whole dataset, 3D perspective transition between scatterplots enabled users to follow transitions easily between scatterplots. Tory et al. [76] argued that a combination of 2D and 3D displays was better than exclusively 2D or 3D representation for relative position estimation tasks in a 3D space. Tominski et al. [74]

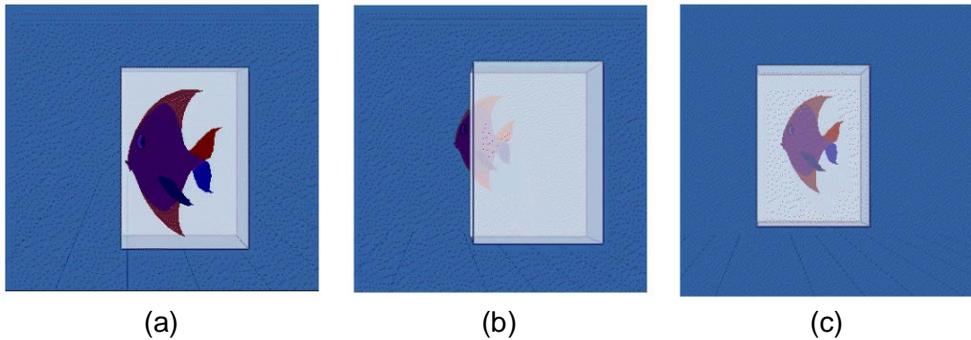
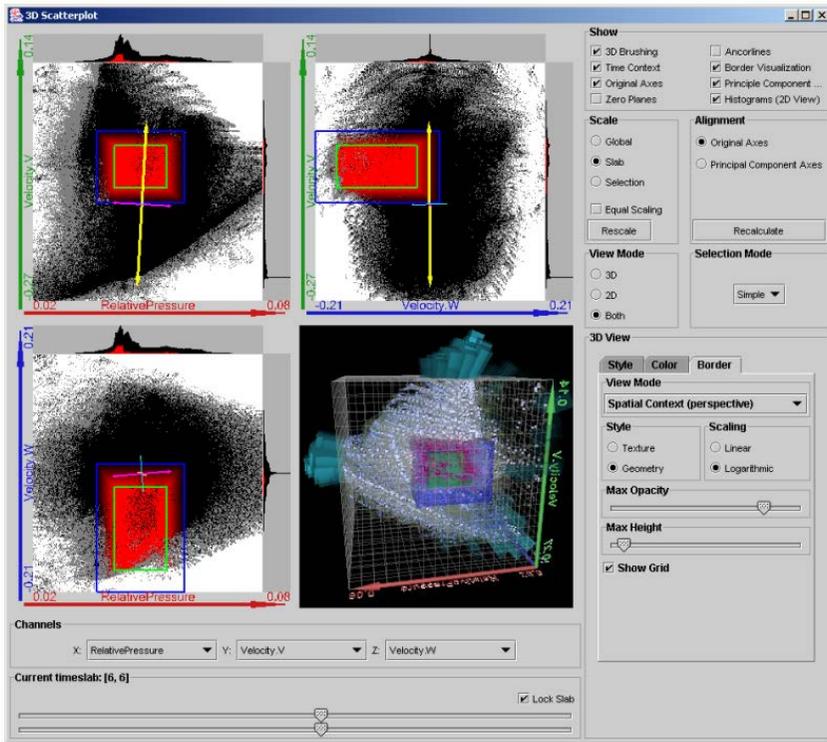


Figure 2.1. Partial occlusion of 3D cursor for better depth perception [81]. (a) fish in front of the cursor. (b) fish behind the cursor. (c) fish within the cursor.

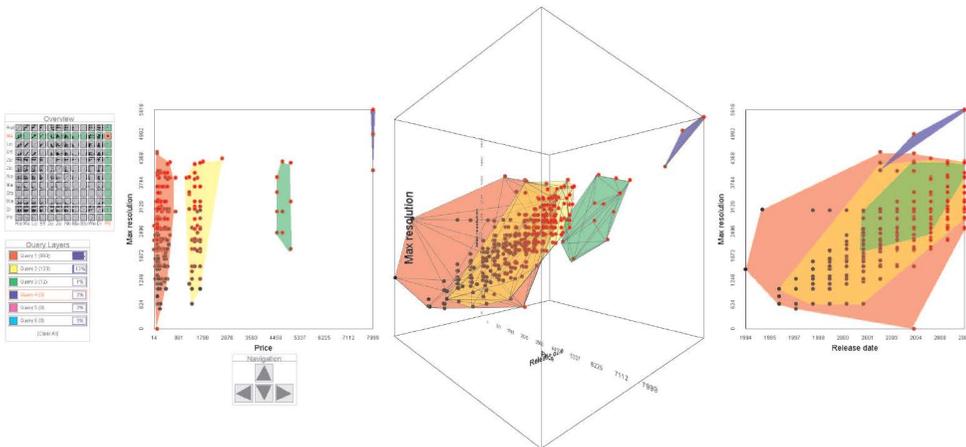
proposed a hybrid 2D/3D display for trajectory visualization with spatial, temporal and attribute components. In their work, two-dimensional representation was used for providing spatial context and showing temporal information in full detail, while multiple trajectories were stacked three-dimensionally. Amini et al. [1] compared 2D and 3D trajectory visualizations to find better performance of 3D visual representation in more complicated tasks. In GazeVis, we also tried to harmonize 2D and 3D views for more effective exploration of the gaze data providing an intuitive 3D overview and 2D detail views.

Some prior work presented design guidelines for 3D information visualization. Shneiderman [63] suggested a number of design guidelines for a better 3D interface. We followed these guidelines in designing our interface. Zhai et al. [81] reported perceptual advantage of using partial-occlusion in a 3D space. They showed that a 3D box cursor with semi-transparent sides improved the performance of target localization tasks in terms of task completion time and accuracy. As our 3D visual

representation was designed to show the gaze data in a 3D volumetric space, we also applied the semi-transparency technique to both stimuli volume data and gaze data while providing more interactive control over the semitransparency.



(a) Combined 2D and 3D scatterplots [52].



(b) Rolling the Dice [21]. Provides animated 3D transitions between 2D scatterplots for multidimensional data.

Figure 2.2. Visualizations combining 2D and 3D scatterplots for multidimensional data

2.2 Eye Tracking Data Visualization

Most prior research in gaze data visualization has developed visualization techniques for showing gaze data for a 2D static stimuli image. Among such gaze visualization techniques, heatmap and gaze plot are most widely used these days [19], and many variants of them have been proposed. There are heatmap-based visualizations that ignore the temporal order of gazes while presenting less cluttered and more compact overviews. Blignaut [10] introduced visual span in generation of a heatmap. Wooding [80] introduced a heatmap-like visualization, or fixation map built by applying a 3D Gaussian filter at each fixation location to reduce clutter and enable detailed quantitative comparisons between different gaze patterns of multiple users. Špakov et al. [69] tried to enhance the heatmap visualization by making the transparency of heatmaps adjustable.

Some gaze plot based visualizations put more emphasis on the temporal order of gazes. Lankford [36] proposed an improved gaze plot, GazeTrail where segments of a scan path are displayed in different colors according to gaze time to reveal speed of gaze movements along with fixation duration and number. Rähkä et al. [53] proposed the time plot visualization to show the temporal order of visits to area-of-interests while sacrificing the exact fixation locations. Goldberg et al. [25] extended the gaze plot (or visual scanpath representation) using time expansions, small multiples, and radial plots; and classified scanning strategies into nine categories.

Researchers also have proposed gaze visualization techniques for dynamic

stimuli. For example, Tsang et al. [77] introduced eSeeTrack to support comparison of fixation patterns on dynamic stimuli. Instead of superimposing gaze patterns on top of stimuli, eSeeTrack used a timeline and a tree visualization to show duration, frequency, and orderings of fixations. While it shows strength in comparison tasks, it is based on an assumption that fixations are automatically extracted and labeled in advance, which is a relatively strong assumption for some domains such as medical imaging. Visualization techniques for 3D virtual environments also have been introduced [18]. For example, Stellmach et al. [66], [67] introduced 3D scan paths, 3D attentional maps, and models of interest timeline view for 3D virtual environments. These techniques work only when the geometry of objects is known a priori. They do not consider the case where gaze points are not on the surface of objects but inside the objects.

In this work, we introduce a 3D gaze visual representation for more complicated dynamic stimuli which comprise a 3D volumetric space where the interior of the 3D space is of interest to users (e.g., contiguous cross-sectional medial images). We propose a 3D volumetric data structure, i.e., gaze field, as a novel representation of human attention in 3D gaze analyses, where conventional 2D fixation filters do not work well.

2.3 Comparative Data Analysis

The gaze data collected during diagnoses are multidimensional data sets with spatio-temporal dimensions. There are also additional dimensions from contextual information. Thus, we adopted and extended prior work on multidimensional data in our visualization and interaction design. Gleicher et al. [24] proposed a taxonomy that summarizes visual designs for comparison as a composition of three types of elementary groups: juxtaposition, superposition, and explicit encodings. Munzner et al. [43] used juxtaposition in their work to compare two large phylogenetic trees. In addition, GazeDx took advantage of juxtaposition to support interactive comparison between readers from spatial and temporal perspectives, and also used superposition in visualizing spatial trends of gaze data with stimuli in the background.

Regarding comparison between multiple items, matrix-based visualizations were used in numerous prior studies to support pairwise comparisons in light of Bertin's [6] introduction of permutation matrix. For example, Kim et al. [29] used a permutation matrix in ConSet to show relationships between multiple sets with a large number of elements. GazeDx took a similar approach to visualize similarity among readers in a similarity matrix and to show an overview of multidimensional relationships in scatterplot matrix views. While multiple views can help comparative data analysis, it could suffer due to a lack of screen spaces. Perlin and Fox [47] introduced semantic zooming that shows information with a different level of abstraction depending on the current zoom level. We adopted a similar approach in

showing gaze data using different visual encodings according to available screen space.

Eye tracking data, in particular, has been used as a source of comparative analysis by a number of researches. Kurzhals et al. [33] summarized visualization papers that used eye tracking. Blascheck et al. [8] proposed a taxonomy to classify the visualization techniques for eye tracking data. It includes stimulus-related (e.g., static vs. dynamic; passive vs. active; and 2D vs. 3D) and visualization-related categories (e.g., temporal, spatial, and spatio-temporal; static vs. animated; single user vs. multiple users; 2D vs. 3D; in-context vs. not in-context; and interactive vs. non-interactive). According to their classification, there are only couple of prior work on dynamic and active stimuli comparing multiple users with interactive visualizations. Kurzhals et al. [34] proposed ISeeCube that use Space-Time Cube along with timeline visualization to show AOI-based scanpaths of different viewers. It was designed for video stimuli with dynamically changing AOIs. Tsang et al. [77] introduced eSeeTrack to compare fixation patterns on dynamic 3D scene such as surgical simulation. The orderings of fixations from multiple users were explored and compared with eSeeTrack, but they were not visualized in context. Pfeiffer [48] also used 3D scene as stimuli and visualized collected gaze data using a 3D scanpath and a 3D attention volume. A scene from diagnosis with volumetric medical images is similar to video stimuli rather than 3D scene. However, its innate volumetric structure and familiarity of intended users (i.e., radiologists) with the structure distinguish the

scene from ordinary video in the eye tracking study.

Seo and Shneiderman proposed the rank-by-feature framework to support a systematic analysis of multidimensional datasets. Using the framework, available low dimensional projections are evaluated by a user-selected ranking criterion [46], [60], and the results are presented in a matrix view where each cell represents a projection and is linked to a dedicated visualization. We applied a similar analysis framework and interaction by allowing users to choose a similarity metric to evaluate relationships between factors or dimensions related to gaze patterns. In this work, we propose interactive visualizations dedicated to reveal relationships between factors especially pertaining to gaze data.

2.4 Gaze Analysis in the Medical field

In the medical field, a variety of quantitative and qualitative gaze analysis studies enriched understanding of how radiologists read various types of medical images. Kundel and Follette [31] compared the visual search patterns of experts and novices during the study of a chest radiographic image. In addition to conducting the quantitative analyses based on fixations to compare first hit time and decision errors, they juxtaposed scan paths from each reader and compared the patterns considering years of experience. Quantitative comparison using fixation and qualitative comparison with juxtaposed gaze patterns were two main research methods commonly used in numerous studies. For example, Kundel et al. [32] found that experts had a holistic perception during diagnosis by analyzing quantitative measures

such as first fixation time and the ROC index and juxtaposing multiple scan paths. Most such work focused on the analysis of a single static 2D image, which is not directly applicable to volumetric images.

Furthermore, gaze tracking studies were conducted on volumetric images. Atkins et al. [1] proposed using a static scatterplot called a navigation chart, to visualize temporal viewing sequences of a single observer with time on x -axis and slice number on y -axis. In this paper, we enhance the plot to make it more interactive and flexible while supporting semantic zooming with three different visual representations. Phillips et al. [49] used a series of cross-sectional brain MRI images in their study and showed the gaze data in two different views. The first view accumulated fixations throughout the volume into a single 2D view and the fixations on the current image are highlighted in a unique color while one scrolls through the volume in a stack viewing mode. Conversely, the second view emphasized the volumetric characteristics of the gaze data with a superimposed gaze plot on 3-dimensionally rendered image stacks. The views have two limitations for comparative gaze analysis tasks: they either discard depth information or suffer from severe occlusion and visual clutter, and one had to manually juxtapose the resulting visualizations outside the framework for inter-reader comparison.

Drew et al. [16] proposed a gaze visualization method with volumetric images based on a volume-rendering technique while characterizing the visual search of experts. They divided the image area into four quadrants and displayed a color-coded

navigation chart to identify two different visual search strategies: (1) drillers with consecutive gaze patterns in a certain quadrant and (2) scanner with gaze patterns distributed across different quadrants. While this approach effectively visualizes the gaze patterns from volumetric images, it still relied on the conventional static juxtaposition strategy for comparison. In this work, we not only overcome the limitations of existing approaches by designing visualization and interaction for more effective comparative gaze analyses, but also encourage multidimensional exploration of the data by including important contextual information in the analysis framework.

Chapter 3

GazeVis: Volumetric Gaze Data Visualization¹

GazeVis is designed to aid medical doctors in gaining insights about their gaze data which is acquired during diagnosis with contiguous cross-sectional medical images. While prior work [49] on such volumetric medical images relied on solely 2D and 3-dimensionally stacked visual representation for gaze data, the proposed visualizations suffered from either difficulty of overall gaze pattern perception or severe occlusion with scalability problem. Thus, a 3D representation is introduced in GazeVis to visualize diagnostic gaze data from volumetric images. Since the size of the gaze data gets much bigger with a series of images, we amplify cognition by adopting the visual information-seeking mantra [62]: overview first, zoom and filter, then details on

¹ A preliminary version of this chapter was published in TVCG [65].

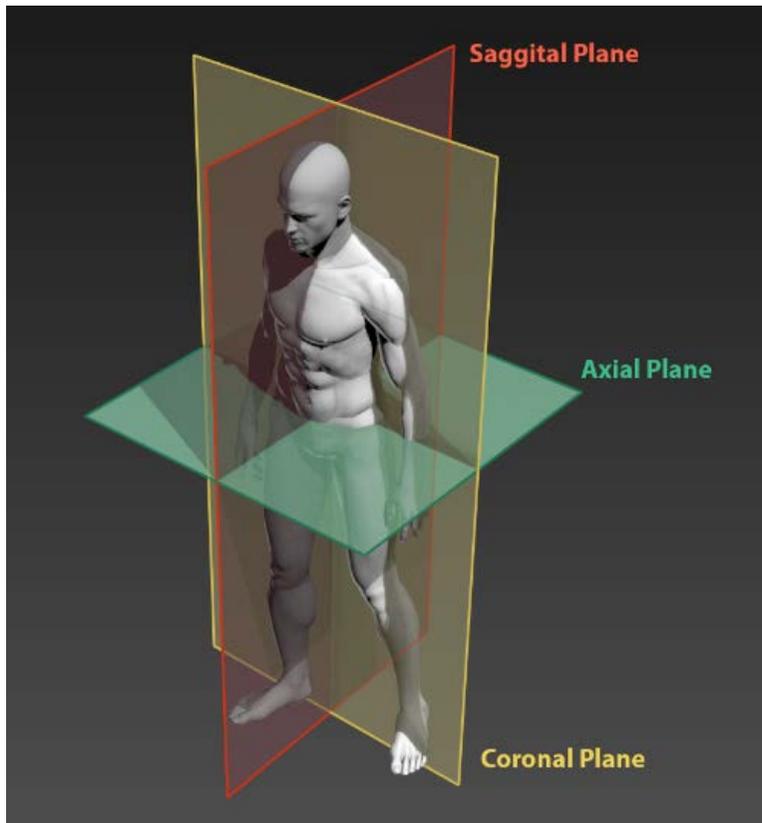


Figure 3.1. Multi-planar reconstruction (MPR) planes of human body. 3D human model is from [41].

demand. We implement volume rendering (VR or 3D) and multi-planar reconstruction (MPR) views to provide an overview and details of gaze data, respectively. We also design interactive spatial and temporal filtering techniques which will be explained in detail in this section. In 3D and MPR views, we superimpose the gaze data on the stimuli (i.e., organs and lesions) to provide users with a more intuitive spatial cue.

GazeVis consists of four views on the left and a control panel on the right (Figure

3.2). The four views are axial, coronal, sagittal, and 3D views, clockwise from the upper left corner. Each of the three MPR views—axial, coronal, and sagittal view—represents an orthogonal plane that divides human body (Figure 3.1). Each MPR view has cross hairs to control the location of the corresponding plane. For instance, the vertical line in the axial view corresponds to the sagittal plane. User interactions on the four views and their interactive coordination will be explained in detail in section 3.2.

On the right side of GazeVis, there are four sets of UI widgets that control parameters for gaze visualization. The widgets on the top are for manipulating parameters of Gaussian filter (explained later in section 3.2.5). Below the widgets, there is a range slider for temporal filtering (explained later in section 3.2.3). Next two sets of widgets are for manipulation of two transfer functions, one for gaze data and the other for stimuli volume data (explained later in section 3.2.4).

3.1 Visualization of Stimuli and Gaze Data

We adopted common visualization methods—DVR and MPR—for stimuli volume data (i.e., contiguous cross-sectional medical images) to provide a better and more intuitive context for gaze data, wherein the DVR is for overall context and the MPR for detailed context. Both methods use a virtually constructed 3D volume by stacking images in z -order. However, each method has different characteristics and conveys different information.

The cross-sectional medical images are three dimensionally rendered by the ray-

casting DVR while the gaze data is superimposed on it, giving the overall context for gaze data. In the DVR, a transfer function maps voxel values (i.e., CT intensities) to colors and opacities, and the mapping is called classification. By adjusting the transfer function interactively, users define the “look” of the data: determine parts to be visualized; determine parts with transparency; and determine assigned colors for each part. Through the classification, users can obtain a realistic 3D scene from gray-scale volume data. It has been previously shown that when the gaze data is overlaid on the 3D scene of stimuli, the depth cue of the gaze data can be enhanced by exploiting partial occlusion with semi-transparent visualization of the stimuli [81]. Thus, we adopted this DVR technique to provide an overview of the gaze data.

MPR, on the other hand, reconstructs a 2D image by cutting through the 3D volume in three different orthogonal planes: axial, coronal, and sagittal (Figure 3.1). While each plane shows a single cross section at a time, radiologists most frequently use these views, especially the axial view, during diagnosis. Gaze data is also superimposed on those MPR views (Figure 3.2). In GazeVis, we implemented MPR views as detail views for two reasons: (1) to provide detailed spatial information of the gaze data in a familiar way to radiologists and (2) to boost task performance in navigating and positioning exact gaze points in the 3D volumetric space.

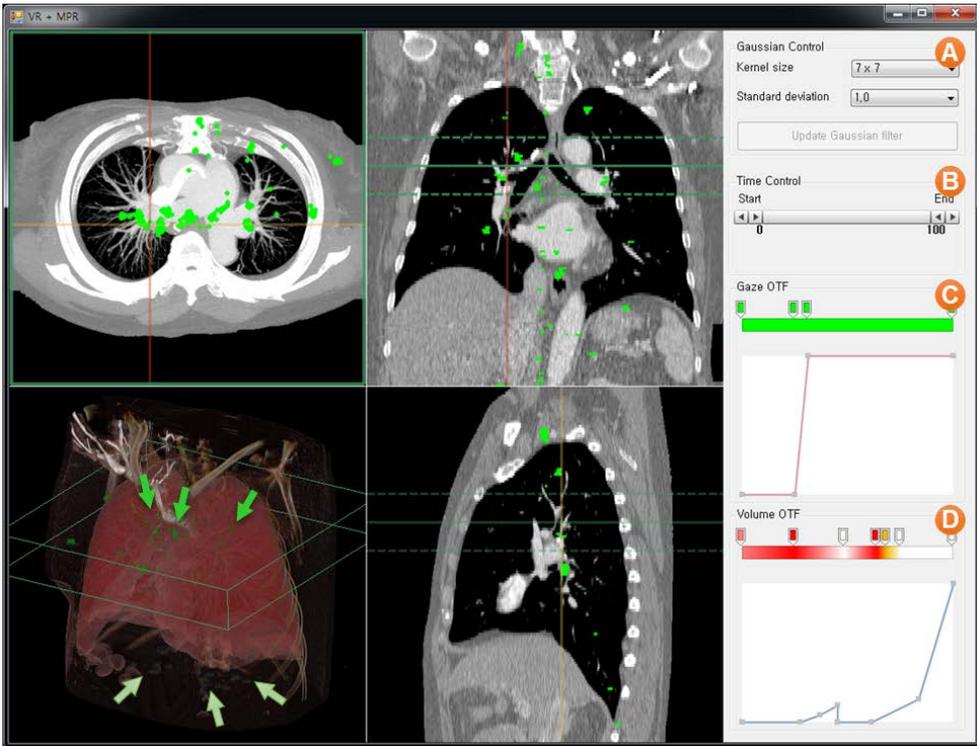


Figure 3.2. GazeVis interface with 4 views on the left and a control panel on the right. Bright green dots represent gaze points. 4 views are axial, coronal, sagittal, and 3D views, clockwise from the upper left corner. Axial, coronal, and sagittal views have corresponding cross hairs and rectangular borders in their corresponding colors: green, yellow, and red, respectively. A control panel has (A) widgets for Gaussian filter control, (B) a range slider for temporal filtering, and (C), (D) widgets for manipulating transfer functions. In this figure, a user manipulates thickness of the axial plane in the coronal view. The thickened axial plane is represented as a green rectangular prism in the 3D view. Gazes inside the cube (pointed by light-green arrows) are in light green, while the gazes outside the cube are in pale green (pointed by pale-green arrows).

3.1.1 Computation of Gaze Field

As discussed earlier, the gaze data, overlaid on the 3D scene can have the enhanced depth cue through semi-transparent rendering of stimuli. However, a simple overlaying such as blending a volume-rendered image of stimuli and a projected image of gaze data in 2D-image level should lead to a loss of depth cue. To avoid this, the gaze data has to be visualized (or rendered) simultaneously during rendering the stimuli, which enables the preservation of accurate depth order between the stimuli and gaze data.

Gaze data requires proper pre-processing for visualization. In this study, we used a 60Hz eye tracker, which generates 60 data points every second with gaze position, time stamp, pupil diameter, and additional information. Without proper pre-processing, the amount of raw eye tracking data from a modern eye tracker gets too large for human readers when the recording time gets longer [53]. Thus, most of the prior work adopted a fixation filter in the data pre-processing phase to identify fixations and saccades: Fixation is a short stop at a certain area; and saccade is a movement between fixations. They aggregate gaze points into meaningful clusters, which makes gaze plots less cluttered.

Fixation filters mostly depend on spatial and temporal information in two dimensions [59], wherein gaze points are classified as a fixation by velocity or proximity thresholds. For a static 2D scene, such traditional filters can successfully identify the fixations; however, they are not suitable when a sequence of images is

read while being scrolled up and down repeatedly. Staring at a specific position in a static 2D scene should be classified as a fixation. However, when users stare at the same location in different image slices while scrolling them, it should not be classified as a fixation. It could be thought of as a z -directional saccade. The traditional fixation filters misclassify those gaze data as a fixation, too.

To address this problem, we collect slice information, i.e. slice number as readers scroll through the image stack, along with the typical gaze information such as gaze location and time stamp. We implement a DICOM (Digital Imaging and Communication in Medicine) image viewer in GazeVis, which provides commonly-used image-viewing functionalities of commercial PACS (Picture Archiving and Communication System) such as windowing (i.e., adjustment of image contrast and brightness) and zooming. When users read the images using this viewer, it collects the gaze data directly from an eye tracker in 60 Hz by using the SDK of the eye tracker. The viewer also collects the slice numbers of images in the order they are seen. The slice numbers can be easily obtained from the DICOM header information.

As mentioned above, traditional 2D fixation filters are not appropriate for the gaze data acquired while scrolling a sequence of consecutive images. Thus, we propose a novel gaze analysis method, which adopts the heatmap approach. The method applies a 2D Gaussian blur function (equation 1) to each gaze point:

$$G(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} \quad (1)$$

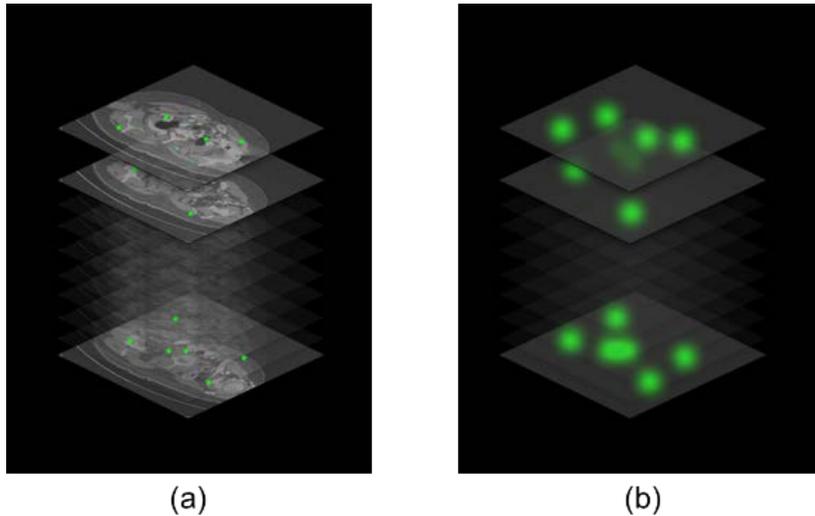


Figure 3.3. (a) Gaze points on stimuli volume data (i.e., contiguous medical images). (b) Gaze field with the same x -, y -, and z -resolutions as the stimuli volume data.

By convolving the Gaussian function to the gaze data in a given slice, we obtain the corresponding gaze-scalar image for the slice, of which the pixel has a scalar value accumulated through the Gaussian convolution. We normalized the scalar values to make the maximum 255 (i.e., one-byte precision). In this way, we obtain a 3D scalar field containing the values representing the gaze density (i.e., level of human attention), referred to as gaze field, across the entire contiguous cross-sectional medical images (Figure 3.3). This gaze field has the same resolution as the stimuli volume data.

We adopt the Gaussian blur because it is well-known to approximate the spatial receptive fields of the human visual system (i.e., a receptive field of ON-center and OFF-surround) [4]. We use the Gaussian blur as an alternative for traditional fixation

filters which are not suitable in our work. The Gaussian blur can attenuate the visibility of saccadic movements by making such saccadic movements have low intensity in the gaze field. Also, the human eyes inherently have micro-saccade movements when focusing on a specific region. In other words, there are involuntary very small eye movements even when the person is trying to stare at a single point. The Gaussian blur filter could smooth out such drift movements, as does the fixation filters for 2D gaze data. In this way, we can more appropriately simulate human attention at each focused region.

3.1.2 Visualization of Gaze Field

The 3D gaze field is also rendered using the ray-casting DVR, wherein its own transfer function maps the gaze scalar to color and opacity. The default transfer function we use for the gaze field is a ramp function which maps a higher opacity to a higher gaze scalar value. It makes the regions where eyes remained more frequently or for a longer time more apparent. We use green as the default color for the gaze field. Using this transfer function, users can adjust the look-and-feel of the gaze data interactively such as color change or opacity change.

The gaze field can be rendered simultaneously along with the stimuli medical volume data by using multi-volume DVR, which is a straightforward extension of the single-volume DVR. A ray casted from an image plane traverses both volume datasets simultaneously while sampling the intensity from the medical data and the scalar value from the gaze field at an interval and accumulating colors and opacities

evaluated via their own transfer functions. As a result, we obtain a 3D rendering of stimuli (i.e., organs and lesions) with the gaze data being three-dimensionally superimposed as small cylinder-shaped independent objects (3D view in Figure 3.2 and Figure 3.5). Such three-dimensional superimposition does not suffer from any loss in depth information, and therefore it delivers the 3D rendering of stimuli and gaze data in the accurate depth order with more enhanced depth cue. Furthermore, by adjusting the transparency either of stimuli and gaze data interactively as needed, users can be relieved from the problem of disturbing occlusion of 3D visualization to a considerable extent.

3.1.3 Gaze Field for Interactive Information Seeking

In addition to its accurate depth-ordered superimposition with adjustable transparency, the proposed gaze field has advantages in terms of supporting interactive information seeking such as dynamic queries. As the gaze field is a sort of spatial data structure which stores the gaze information at the corresponding position of the stimuli, dynamic spatial queries on gaze data can be directly supported (presented in detail later). In addition, any attribute of the gaze information, if stored in the spatial coordinates of the gaze field, can be also visualized and manipulated during volume-rendering of the gaze field.

In gaze analyses, temporal information often provides very important clues. For example, Atkins et al. proposed a novel plot, i.e., navigation chart to show the temporal viewing sequence of medical image slices, and found that radiologists

usually examined images in two distinct phases: locate phase and review phase [1]. Timeline-based gaze visualization and temporal filtering were used in some prior work [67], [77]. To support accurate interactive temporal queries in this work, all gaze time information should be stored in the gaze field, which is loaded into GPU memory during GPU-based DVR for its interactive rendering. However, because some gaze points are gazed many times, too much GPU memory may be required to store all the gaze time stamps in the gaze field. Thus, we decided to prioritize the interactivity of temporal query over its accuracy by storing only the last (i.e., most recent) time stamp for the gaze points gazed multiple times to focus more on the review phase [1]. Such a design decision was made based on the comment by an expert radiologist participating the case study (presented later) that the last time stamp is more important than others as the last gazing on a region is likely to confirm the final diagnostic decision for that region.

In the gaze field, time stamps are stored in two-byte unsigned short format, giving the time granularity of 65536 ($=2^{16}$) which amounts to about 18 minutes with the 60Hz eye tracker. Although the reading time per examination greatly varies with clinical situations and diagnostic tasks, it typically ranges from 3 to 15 minutes on average in outpatient examinations. Thus, two bytes are sufficient to hold the time stamps.

3.2 Interactions and Dynamic Queries

We designed various intuitive interactions for effective gaze data analyses in GazeVis. Three key design goals were: (1) we have to make them scalable in terms of the gaze data size; (2) we have to provide an instantaneous feedback to help researchers recognize causality; and (3) we expect them to help researchers perceive the position of gaze point in the 3D volumetric space. Real-time rendering of the visual stimuli (i.e., organs and lesions) and the gaze data is a baseline requirement to meet these design goals. Using GPU-accelerated ray-casting DVR techniques, we achieve the real-time rendering of the visual stimuli and the gaze data. In the following subsections, we present user interactions along with dynamic query interfaces.

3.2.1 Interaction Design

All four views in GazeVis (i.e., three MPR views and 3D view) are coordinated together so that user interactions in one view are reflected to the others instantaneously. GazeVis highlights the object under the cursor (e.g., a border of the view and a cross hair in the view) and all UI components representing that object in other views. For example, the cursor moves over the horizontal line in the coronal view, the corresponding cut plane, the axial view gets its border highlighted, and the rectangle or a rectangular prism in the 3D view representing the axial plane is also highlighted (Figure 3.2).

Users can drag the horizontal or vertical cross hair on an MPR view to adjust the location of the corresponding plane. For example, users drag the vertical line in the

axial view to change the sagittal plane, i.e., to change the sagittal view. One can also drag the center of cross hairs to move two corresponding planes at the same time. When one wheel-scrolls on an MPR view while the cursor not being on any cross hair, the view under the cursor changes the location of the plane.

Wheel-scrolling on either horizontal or vertical cross hair changes the thickness of the corresponding plane. For example, when one wheel-scrolls up on the horizontal line in the coronal view, the axial plane gets thicker to make the axial view show a composite (i.e., intensity-averaged) image for all the image slices within the thickness. The thickened axial plane is displayed in dashed lines on the coronal and sagittal views and in a rectangular prism in the 3D view (Figure 3.2).

One can also change the brightness or contrast of MPR images by right dragging. In the radiology field, it is known as window setting where window level denotes brightness and window width corresponds to contrast. Users can adjust the image brightness and contrast by right dragging vertically and horizontally, respectively. These are widely-used interactions in commercial medical imaging products. During the case study, we also additionally provided keyboard shortcuts for frequently used window settings.

The 3D view on the bottom left corner supports slightly different interactions. One can rotate an object in the view with left dragging, and one can zoom in or out with wheel-scroll.

3.2.2 Spatial Filtering

There are at least hundreds of contiguous cross-sectional images per CT/MRI scan for a patient with tens of thousands of gaze points scattered on them. Researchers have to explore the gaze points across the images to identify interesting gaze patterns. To support such exploration tasks, we designed a spatial filtering mechanism for the gaze data. Instead of creating a separate dynamic query widget for the spatial filtering, we integrate it into the thickness-MPR function.

When users want to focus on a subset of cross-sectional images, they can perform a spatial filtering by wheel-scrolling on a cross hair to define a satisfying range of cross-sectional images. The wheel-scrolling adjusts the selection range centered at the cross hair so that only the gaze points within the range are selected and highlighted. For example, when a user wheel-scrolls on the horizontal cross hair in the coronal view, the spatial range (and the thickness) for the axial plane is adjusted and then the gaze points within the range are accumulated and shown in the (thickened) axial view (Figure 3.2). In addition, the gaze points within the selection range are displayed in bright green, while the rest are shown in pale green in the 3D view (Figure 3.2).

3.2.3 Temporal Filtering

We also designed a temporal filtering mechanism in GazeVis using a range slider to support researchers' exploration of the gaze data based on time and order. When a specific temporal range is selected using the range slider, MPR and 3D views are

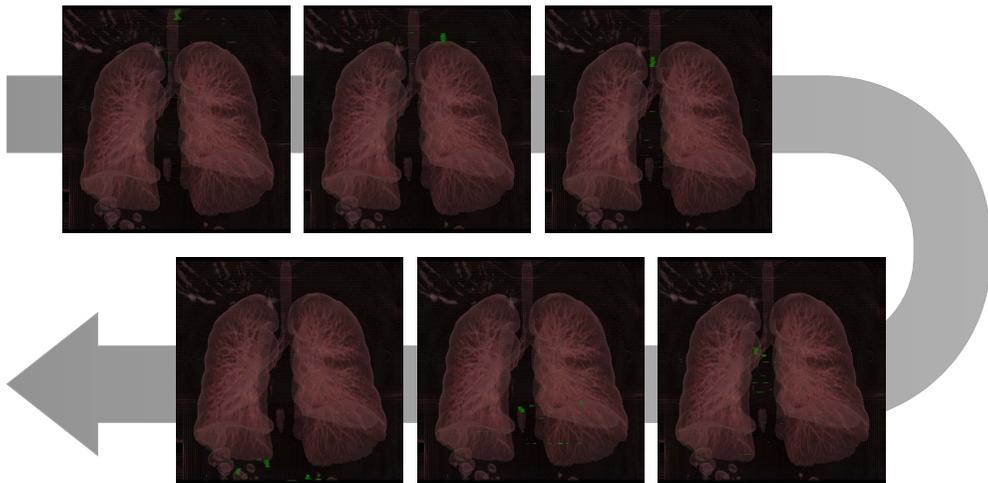


Figure 3.4. Temporal gaze flow of a radiologist in reading a chest CT scan, revealing that he scrolled through the images from top to bottom.

dynamically updated accordingly so that the gaze points outside the temporal range are hidden from all the views.

By incorporating the temporal information in the gaze field (described earlier in section 3.1.3), it is straight-forward to adjust the rendering algorithm to give an instantaneous feedback to any temporal filtering queries even on the large number of gaze points. During rendering the gaze field, only the gaze points within the temporal range are visualized (i.e., classified through the transfer function); other gaze points outside the range were simply skipped. By interactively dragging the temporal range slider left or right, users can grasp temporally changing gaze patterns even without an explicit gaze plot (Figure 3.4).

3.2.4 Transfer Function Control

The transfer function adjusts the color and opacity of anatomic regions and gaze data in the 3D view (Figure 3.5). We provided a UI widget to allow users to interactively manipulate the transfer function ((C) and (D) in Figure 3.2). As shown in (C) in Figure 3, it consists of two panels: upper one for color control and lower one for opacity control. Horizontal axis in the both panels corresponds to the intensity of the medical data (usually 12-bit data for CT and MR scan), ranging from 0 to 4095. In the color panel, the horizontal bar shows assigned colors to the corresponding intensities. One can add a color thumb with a double click, and change the assigned color. Color thumbs can also be reordered by dragging. It is necessary to use multiple colors because a single color is often not enough to distinguish a high intensity point with low opacity from a low intensity point with high opacity. In the opacity panel, the vertical axis represents assigned opacities to the corresponding intensities. One can assign opacity to a certain intensity by manipulating the control points of the opacity function. The maximum value of the vertical axis represents the opacity of 1.0, which is completely opaque, while the minimum value corresponds to 0.0, which is completely transparent. Similar transfer function is applied to the gaze field volume, with the gaze intensity ranging from 0 to 255.



Figure 3.5. 3D rendering of lung with three-dimensionally superimposed gaze data. (a) With low opacity for the lung, gaze data within the lung are clearly visualized. (b) With high opacity for the lung, the gaze data within it are partially occluded.

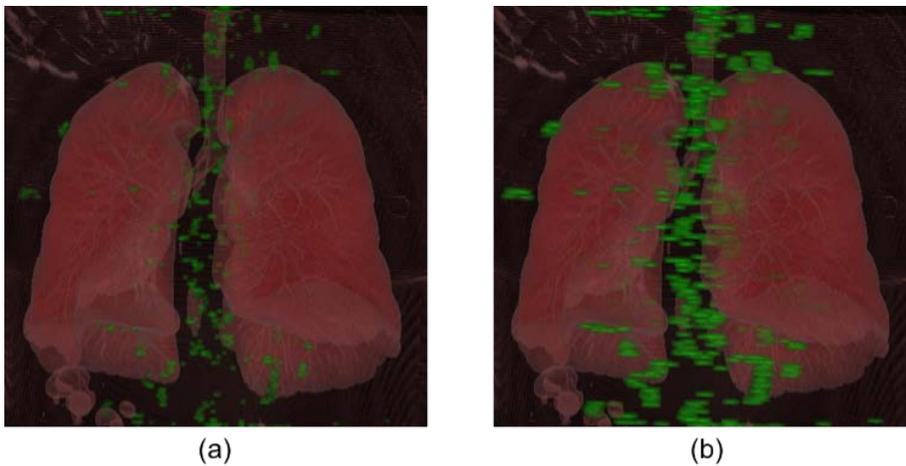


Figure 3.6. Adjustments of Gaussian blur filter size. (a) 7 X 7 with standard deviation of 1. (b) 19 X 19 with standard deviation of 3.

3.2.5 Gaussian Blur Control

We also designed UI widgets to help researchers interactively manipulate the parameters for the Gaussian blur filter. As mentioned earlier, one of the purposes of the Gaussian blur filter was to smooth out saccadic or micro-saccadic eye movements. Thus by adjusting the size of the filter support, we can change the level of abstraction in representing the human gaze. One can increase the filter size to see more abstract overview of human attention map (Figure 3.6) while attenuating fine movements. The standard deviation for the filter kernel function can control the density distribution within the filter range. In this manner, smaller standard deviations can emphasize the center of gaze, which can make more apparent each gaze point on the gaze paths.

3.3 Implementation

GazeVis is implemented with C#, C++, and WPF. We used Tobii analytics SDK [73] for connecting eye tracker and acquiring gaze data. Most of the user interface is implemented with C# and WPF, and the volume rendering was implemented using C++ with Microsoft DirectX SDK.

3.4 Evaluation with Radiologists

We adopted a case study-based evaluation method for information visualization, i.e. Multi-dimensional In-depth Long-term Case studies (MILCs) [61], [64] to evaluate the efficacy and effectiveness of GazeVis as a whole system. A comparative evaluation study was not suitable since there are few gaze visualization tools

comparable to this work and existing ones [49] have innate problems of occlusion and loss of 3D spatial information when applied to real world cases.

We conducted two case studies at a third-tier university hospital with two groups of radiologists: 6 chest radiologists and 6 abdominal radiologists. We followed the MILCs guidelines [64] and recruited the 12 domain experts for the evaluation. In early stages of the case studies, we developed rapport with the radiologists for about a year, familiarizing ourselves to medical diagnosis process and letting them know about visualization research. In later stages, we iteratively improved GazeVis according to the participants' comments and collected gaze data when they read images to reach a diagnosis for patients. The collected gaze data were analyzed by two participating expert radiologists, one from each group who has more than ten years of experience in the field.

3.4.1 Case Study Protocol

In each case study, we visited the hospital 3 times in 3 weeks. Besides the visits, we also communicated with the expert radiologists on a daily basis to help them stay in the flow by refreshing their memories of what had been done before [61]. In the first visit, we conducted a pilot study using a GazeVis prototype, with one radiologist. We first explained and demonstrated the gaze collection process to him. Afterwards we captured gaze data using a DICOM viewer of the GazeVis prototype when he read a prepared CT scan. Then, we showed his gaze data in the GazeVis prototype, and later we debriefed him to collect feedback on the GazeVis prototype to refine the tool.

Among the feedback, the radiologist complained about the lack of predefined brightness and contrast settings (i.e., window settings) which are usually provided with keyboard shortcuts in commercial DICOM viewers. When radiologists read medical images, they have to adjust the brightness and contrast of images to see more clearly a region of interest, e.g. a specific organ. In fact, we adopted a commonly used windowing interface of right-click dragging in vertical and horizontal directions. However, he pointed out two problems: one thing was that it could cause frustration as radiologists have to spend time adjusting the window setting frequently; and the other was that it could yield unintentional gaze data during manually adjusting the window setting. This could happen as one may look at a point in the screen unintentionally during the manual window adjustment even when the point is not clinically important. Thus we improved our design to support keyboard-shortcuts for frequently used window settings before the second visit.

In the second visit, we collected actual gaze data from a group of radiologists. Before the actual gaze data collection, participants had a training session where we let them use GazeVis for as long as needed to get used to the interface. The training session lasted about 5 minutes on average. Then we calibrated the eye tracker using a 9-point calibration procedure before starting actual data collection. During gaze data collection, we showed a prepared set of medical images in a stack viewing mode. Participants were asked to perform a diagnosis as if it were a real reading by scrolling up and down the images and changing the window setting. After the data collection,

we showed their gaze data in GazeVis to the radiologists, and received comments about it. Overall, it took about 10 minutes for each participant.

In the last visit, we asked the most experienced expert radiologist in each group to look into all the gaze data in GazeVis which had been collected in the second visit. We asked him to find any notable gaze patterns and compare the participants' gaze patterns using GazeVis. We also collected comments on his experience in using GazeVis for gaze analysis.

3.4.2 Datasets

We prepared six CT datasets, which are from two body parts, chest and abdomen. For each body part, three datasets were used for pilot study, training session, and in a main study. Details of the CT datasets are summarized in Table 3.1.

Table 3.1. DICOM files used in the case studies.

Modality	Body part	Usage	Resolution	# of images
CT	Chest	Pilot	512 x 512	140
CT	Chest	Training	512 x 512	141
CT	Chest	Main	512 x 512	154
CT	Abdomen	Pilot	512 x 512	166
CT	Abdomen	Training	512 x 512	149
CT	Abdomen	Main	512 x 512	150

3.4.3 Apparatus

When collecting the gaze data, we used a PC with a quad-core processor connected to a 20.8-inch Totoku medical monitor. A Tobii X60 eye tracker, which is known to have accuracy of 0.4° under ideal conditions, was used in the case study. The distance from participants to the eye tracker was approximately 65 cm on average, measured while they performed diagnosis as usual in this experiment.

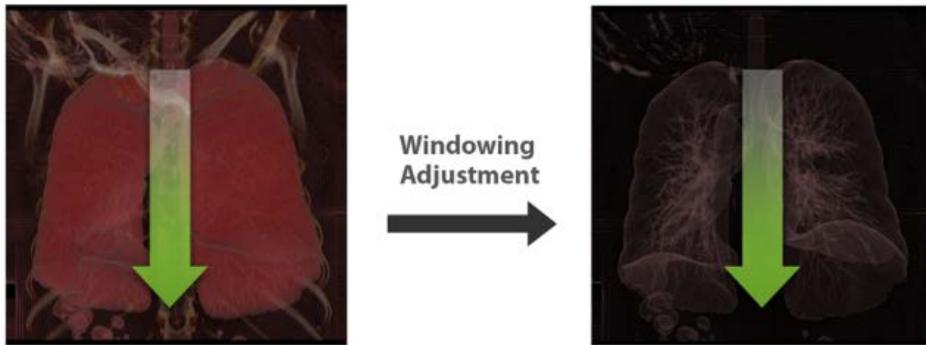
For gaze analysis, we decided to use a separate system because GazeVis adopted a color visual encoding which was not supported by traditional medical monitors, and GazeVis requires high computing power for interactive volume/gaze data visualization. We used an Intel i7 PC equipped with a 3.2 GHz Quad-core processor and 12 GB of main memory, and a 27-inch color monitor. The system was also equipped with an NVIDIA GTX 480 GPU with 1.5 GB of graphic memory. The ray-casting DVR of the medical/gaze data was all accelerated by GPU programming using Direct3D 11 graphics API with HLSL shader model 4.0.

3.4.4 Chest Radiologists

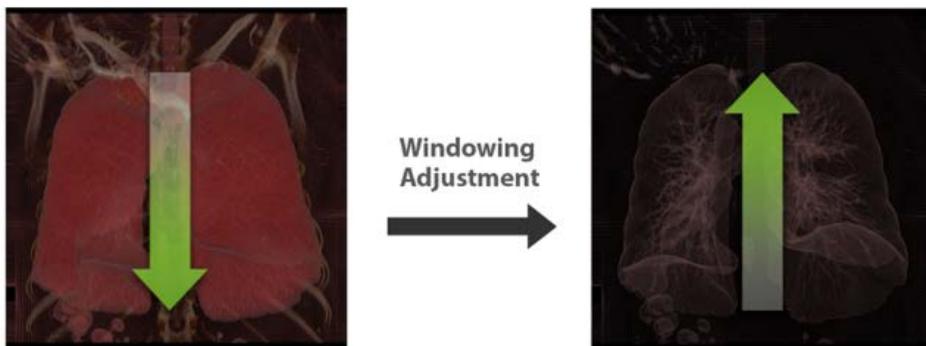
Following the case study protocol laid out in section 3.4.1, we performed the first case study with 6 chest radiologists. Two of them have more than 10 years of experience (experts), and three of them have about 3 to 5 years of experience (intermediates). One of them is a first year resident (novice). We asked them to perform diagnosis without any prior knowledge about the patient. It took about 3 minutes on average to finish diagnosis.

We asked the most experienced radiologist to look into the gaze data of the others and his own in the last visit. At first, he focused on the overall gaze pattern. Using the 3D view, he found that experts tend to have gaze points near the mediastinum, which is located at the center of each image. On the contrary, a novice made gaze points widely scattered on an image. He explained this tendency with peripheral vision, while this tendency may require further studies to generalize. As experts can detect lesions even with their peripheral vision, they tend to gaze at the center of each image. On the other hand, novices have to scan over a wider range of the image thoroughly with their foveal vision as they have relatively low confidence with their peripheral vision due to the lack of experience in diagnosing the images using the visual information detected in the peripheral vision.

Another finding was made with the range slider for temporal filtering. Chest radiologists usually perform diagnosis with two different window settings, one setting at a time to focus on different body parts: lung and mediastinum. However, the 3D view showed all gaze data as a whole in a given transfer function. The expert used the temporal filtering slider to see a group of gaze points clustered according to the gaze time. He himself, at first, read the images scrolling them down from the top to the bottom with the lung window setting; afterwards he jumped back to the top; and scrolled down to the bottom again with the mediastinum window setting as in Figure 3.7(a). While examining other radiologists' temporal gaze patterns using the temporal filtering slider, he could learn that some of his colleagues showed a different gaze



(a)



(b)

Figure 3.7. Gaze pattern difference of two experts in reading a chest CT scan. (a) Expert 1: Scroll down from top to bottom, jump back to the top, adjust the window setting, and scroll down again. (b) Expert 2: Scroll down from top to bottom, adjust the window setting, and scroll up from bottom to top.

pattern. He narrowed down the selected range, and dragged the slider to the right to navigate the gaze data in temporal order. During the playback, he noticed that another expert did not jump back to the top to scroll down, but instead, he changed the window setting to the mediastinum setting and scrolled up from the bottom, examining in the opposite direction as in Figure 3.7(b). Similar individual differences

in the navigation strategy were also reported in [1]. Comparing gaze patterns of radiologists with different expertise level, he could also notice that while the gaze patterns of experts and intermediates have a vertical cylindrical pattern, the novice did not show such a pattern but a noisily scattered gaze pattern.

After the gaze analysis with GazeVis, the expert showed great interest in conducting further formal user studies with GazeVis. He was enthusiastic about studying a gaze pattern difference between experts and novices with a larger number of participants. He was encouraged by his promising experience with GazeVis so that he strongly discussed that the results from a follow-up study with more participants could lead to important clinical implications and could be used to educate novices. He was also eager to test whether prior knowledge affects the gaze pattern. While Reed et al. [55] performed a similar study with chest X-ray images, he mentioned that GazeVis can help researchers investigate the gaze pattern with cross-sectional medical images.

3.4.5 Abdominal Radiologists

We conducted another case study with abdomen CT images while also following the protocol described in section 3.4.1. Six abdominal radiologists participated in the case study. One of them has more than 10 years of experience (expert), and three of them have about 3 to 5 years of experience (intermediates). Two of them are first year residents (novices). We asked them to perform diagnosis without any prior knowledge about the patient. It took about 3 minutes on average to finish diagnosis.

We asked the expert radiologist to analyze the collected gaze data. Before the analysis, he used the axial view to refresh his memory about the dataset. In the 3D view, he found a difference between him and novices. As in Figure 3.8(a), he showed more organized pattern compared to novices. His gaze pattern looked like a set of short vertical cylinders, implying that experts tend to fixate on the same location when he was scrolling throughout a contiguous set of images. On the other hand, there was no distinctive pattern in the novices' gaze data (Figure 3.8 (c)). In case of intermediate radiologists, they showed gaze patterns somewhat in between the expert and the novices (Figure 3.8(b)). The expert explained that this difference was attributed to their peripheral vision, which is similar to what the expert chest radiologist explained.

After comparing overall gaze patterns of the participants, he performed a spatial filtering using the thickness MPR function. He wanted to confirm a hypothesis, proposed by one of the study participants, that radiologists tend to focus on the organ boundary. He could check the hypothesis right away using a dynamic query function of GazeVis. He increased the thickness of the coronal plane in the axial view and scrolled up and down to investigate the whole images. He examined whether the gaze points were densely distributed on the boundary of organs in the coronal view. This spatial filtering led him to reject the hypothesis since there were insufficient gaze patterns that supported the hypothesis.

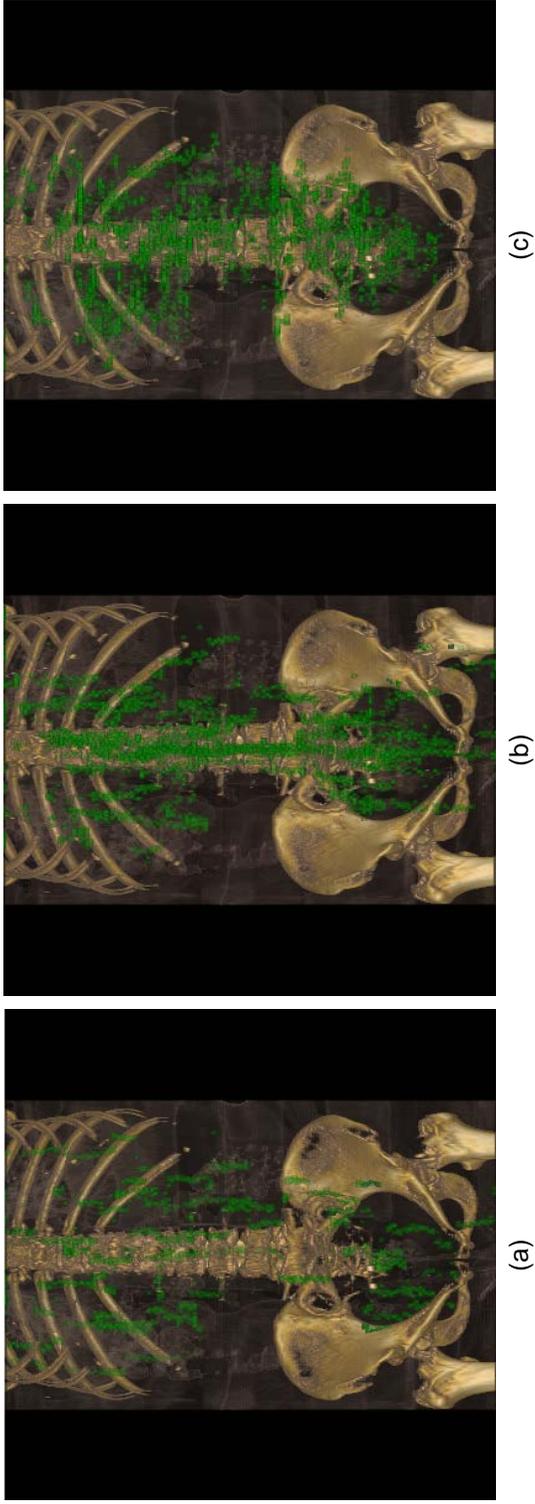


Figure 3.8. Gaze pattern difference between radiologists with different expertise level. (a) Expert showing more organized pattern vertically. (b) Intermediate showing a pattern in between expert and novice. (c) Novice showing no distinctive pattern with scattered gaze points. Usability testing software, showing the task in the Usability category. Participants had to make a graph (lower graph) the same as the given ripple graph (upper graph).

Using the range slider for temporal filtering, the expert narrowed down the range to examine the gaze data in temporal order. He recognized a pattern that the gaze of the expert or intermediates showed more vertical movements compared to the novices. It was similar to the analysis result of chest radiologists, but the length of each unit cylinder is much longer than the chest case, which might be in part due to the anatomical difference of the two body parts.

Abdominal radiologists used four different window settings during diagnosis since there are more organs in abdomen than in chest. Thus the abdominal expert also wanted to cluster the gaze points according to window setting, which was not supported in GazeVis. He commented that grouping and filtering the gaze data according to window setting can help with finding missed regions more accurately since a window setting can either emphasize or hide a specific body part. This issue will be further elaborated in section 3.5.1.

The expert also wanted to explore the gaze pattern difference among radiologists with different specialties. There are diverse groups of radiologists, including abdominal, chest, and musculoskeletal radiologists, and each group might investigate the same body part differently with different gaze patterns. We believe that this kind of exploration could be efficiently supported in GazeVis with some additional functions such as color tagging of body parts.

3.5 Discussion

We presented GazeVis, a novel interactive 3D gaze visualization tool. Chest and abdominal radiologists used GazeVis to collect their gaze data and analyzed them. Based on the lessons learned from two case studies, we present our thoughts and considerations on the further improvement of GazeVis from some related research perspectives.

3.5.1 Spatial Data Structure and Flexibility

In GazeVis, we stored gaze information in the gaze field with the same resolution of the stimuli volume data. Such a spatial data structure for gaze makes GazeVis scalable and interactive even with a large number of gaze points. It is achieved in a way that the gaze field is computed by cumulating gaze scalar values at each voxel position and the computed gaze field is then visualized and manipulated interactively by using the GPU-accelerated DVR technique.

Another notable thing for the gaze field is that it can be interpreted as an attribute data for the stimuli volume data, which adds additional information (i.e., how long or often a given position was gazed during the measurement), to the stimuli themselves (i.e., organs or lesions). We can combine the gaze field with the volume intensity of the stimuli data to improve classification: anatomic regions with larger gaze scalar value (i.e., regions gazed more frequently or longer) can be mapped to different color or opacity. In this way, the difference in gazing density becomes immediately apparent in the 3D rendering while the densely-gazed regions, which are likely more

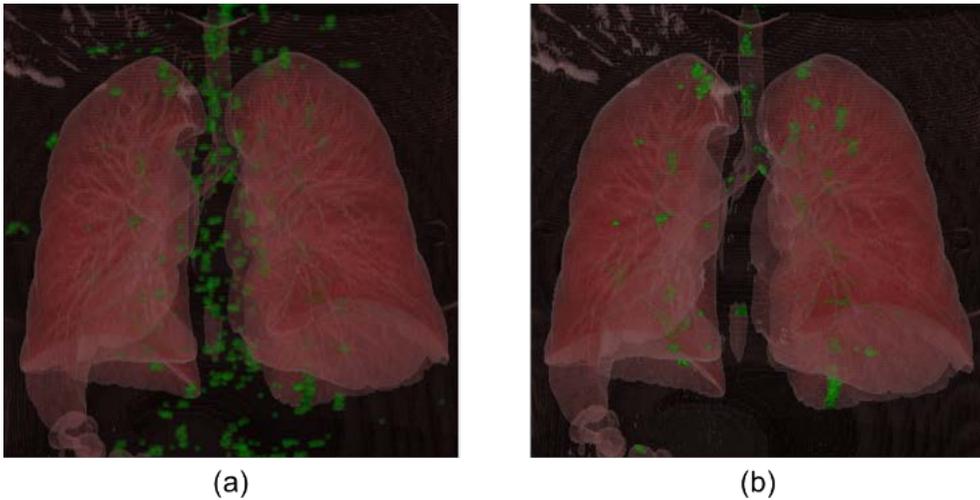


Figure 3.9. (a) 3D rendering of lung with gaze points superimposed as independent objects. (b) 3D rendering of lung with gaze data used as attribute data, enabling gaze analyses more focused on human anatomy.

important in the diagnostic reading, are more clearly marked (Figure 3.9). Such combination of the gaze field as an attribute data can be achieved by adopting the 2D transfer function approach [14], [30].

The gaze field has another advantage that it can store any kind of contextual information as long as the GPU memory can accommodate it. For example, the gaze field can store windowing values at each gaze point. Most of prior gaze analyses mainly focused on the location of the gaze; however, as discussed earlier, radiologists often read a single examination with a couple of window settings for diagnosing different lesions. Chest radiologists examine the chest wall and overall lung structure under the mediastinum and lung window settings, respectively. If the windowing value, stored in the gaze field, is shown at each gaze point with an appropriate visual

encoding, the 3D overview can provide researchers with a richer context. They could easily discriminate anatomic regions mainly examined in different window settings. In addition, they could easily notice inadvertently missed or unnecessarily gazed regions in some window setting.

3.5.2 Interacting with Contextual Data

Adding contextual information in the gaze field requires further work on interaction design for interactive exploration of such information. For example, with the current version of GazeVis, an expert chest radiologist in the case study had to playback the gaze data several times with a narrowed temporal range slider to perform gaze analyses regarding the windowing setting. Those analyses could be supported more efficiently by using a dynamic query interface for interactively selecting a range of windowing values or for choosing frequently used windowing values.

Future research is also needed to develop visualizations for helping users find spatial or temporal gaze patterns in a more comprehensive way. Visualizations to show marginal distributions of gaze data in each MPR plane can provide users with more contextual overviews. For example, if a histogram of gaze data is attached to a side of an MPR plane as shown in Figure 3.10, users can immediately check the gaze distribution of the current slice (in green) against that of entire slices (in gray). This feature has been implemented in GazeVis after the case studies. When other contextual data such as pupil size and windowing values are shown in the histogram alongside the gaze data on an MPR plane, it could help users to get even more

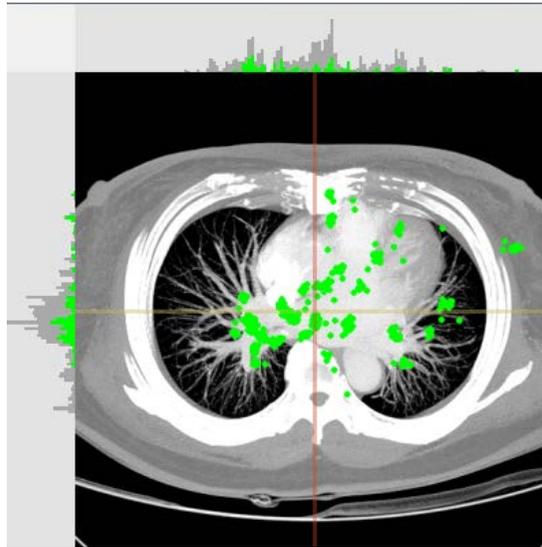


Figure 3.10. Gaze data distribution in an axial plane. Gray histogram represents the distribution of all gaze points on entire slices. Green one represents the distribution of the gaze points only on the current axial image.

valuable insights.

Visualizations to show the temporal exploration sequence of 2D image slices could reveal interesting information about individual variability in exploration. The navigation chart [1] (a plot of image slice number against time) is a good example, which could unveil not only the exploration sequence but also the speed of exploration. It could be incorporated into GazeVis as a part of a scented widget for the temporal query or as a separate interactive visualization.

Chapter 4

GazeDx: Interactive Gaze Analysis Framework

As a stimulus of an eye tracking study, a stack of volumetric medical images introduces both a threat and an opportunity for visual analysis: conventional fixation filter is inapplicable to additional z -dimension; but intended users (i.e., radiologists) are familiar to the volumetric structure of stimuli. In order to overcome the threat, three-dimensional gaze data is preprocessed into a volume (i.e., gaze field) [65]. Each gaze point goes through two-dimensional Gaussian filter, and is accumulated to each corresponding cross-sectional plane. The resulting planes are then stacked to form a volume for visualization. As for the design of the visual representations, we actively utilized the intended users' familiarity with the stimulus as an opportunity by adopting a number of common visual representations in the medical field.

We adopted the visual information-seeking mantra [62] in comparative analysis on multiple gaze data. GazeDx first shows the overviews of the gaze data from multiple readers and allows users to select a small set of readers for detailed comparison. It then enables users to investigate the multidimensional aspects of individual readers' gaze data. Following the mantra, we designed GazeDx to have two main tabs: Overview and Comparison view. Users begin their gaze analysis in the overview tab, where they can check the overall similarity relationships among all readers in a similarity matrix view based on a rank-by-feature framework [60]. They can also examine the overall spatial and temporal exploration patterns of readers using small multiples. Once they set their target readers for detailed comparative analysis in the overview tab, they move on to the comparison view tab, where they can compare the gaze patterns of the selected readers in greater detail, taking into account important contextual information.

4.1 Design Rationale

To achieve the main design goal, we first identified design requirements for GazeDx by conducting interviews with radiologists during the case studies (described in section 4.7) and by reviewing prior work on gaze analysis in both radiology and visualization field. Two expert radiologists (one chest and one abdominal), who also participated in the case studies, especially took part in the review of previous works from the radiology field. They also helped establishing fundamental design requirements of the analysis tool by sharing their experience in the field. We designed

GazeDx to meet the requirements. In this section, we describe our rationale behind the visualization and interaction design for GazeDx.

(1) Support comparative visualization of multiple gaze data

Finding similar patterns in the gaze data is one of the common analysis task, and usually achieved by comparing data from multiple participants [8]. Such comparative analysis was also performed in the medical field to identify the differences in gaze patterns between radiologists [16], [32], and then to explore the sources of the differences [31], [42], [55]. However, prior work from the radiology field shows weakness in interactive comparison between multiple gaze data. Thus, considering its importance, it is essential to support visual exploration of multiple gaze data. Therefore, we utilized small multiples design and the rank-by-feature framework in GazeDx.

(2) Reduce cognitive load on radiologists

As GazeDx aims to help radiologists analyze gaze data from clinical diagnosis, we limited intended users of the tool to radiologists and utilized familiarity with the stimuli in providing in-context visualizations [8]. After the interview with the expert radiologists, we selected a number of visual representations familiar to the users and adopted them: 2D multi-planar reformation (MPR) images (axial, coronal and sagittal images) and 3D volume rendering (VR) image. We used the visualizations to provide important clinical contexts during gaze analysis, mainly by superimposing abstract gaze data over them. It alleviated the complexity

coming from three-dimensional space recognition of gaze data.

(3) **Support interactive temporal analysis**

Temporal information related to gaze data has played a pivotal role in clinical gaze tracking studies as a metric to measure diagnosis efficiency [42], inferring strategies of scanning [16], and defining the phase within a single diagnosis [49]. Moreover, volumetric images can be thought of as dynamic stimuli in gaze tracking studies because the scene (as stimulus) changes over time as radiologists consistently scroll through a long stack of images. Thus, we incorporated visual representations, such as interactive temporal charts and the multi-temporal pane, to show temporal information related to gaze data in our tool.

(4) **Provide flexible exploration of gaze data**

In this study, we used a 60 Hz gaze tracker that records 60 gaze points every second. Hence, gaze tracking data for a five-minute diagnosis contains up to 18,000 gaze points with multiple dimensions dedicated to additional contextual information as well as spatio-temporal data. The experts demanded abundant interactivity during the analysis of such data rather than a static representation. To support such flexible exploration of the data, we provided not only multi-level visual summaries and aggregations, but also interactive selection and filtering capabilities. Specifically, we introduce a novel filtering technique based on segmentation of clinically meaningful structures.

4.2 Overviews for Comparative Gaze Analysis

4.2.1 Spatial Similarity

Once users load the gaze data of all readers in GazeDx, they can check how similar they are in terms of examination patterns based on eye gaze. To facilitate the overall exploration of inter-reader similarity, we took an approach based on the rank-by-feature framework [61]. Upon selecting a similarity ranking criteria, a color-coded similarity matrix shows a succinct overview of similarity relations among numerous different readers (Figure 4.1A). Each row (and column) of the matrix represents each reader, and each cell represents a similarity value between a corresponding pair of readers. A diagonal cell shows important contextual information, i.e., the expertise level for the corresponding reader (Exp for expert, e.g., R1 and R2 in Figure 4.1; Int for intermediate, e.g., R3 and R4 in Figure 4.1; and Nov for novice, e.g., R5, R6 and R7 in Figure 4.1). The background of each cell other than the diagonal cells is color-coded in gray scale according to its similarity value for improved visual perception of the similarity between the corresponding readers.

GazeDx provides SSIM (Structural SIMilarity) as a similarity measure for 3D gaze volume data. SSIM is a perceptual image quality metric that takes two images (an original image and its distorted version) as input and measures the quality (or fidelity) of the distorted image [79]. Each aggregated gaze volume (i.e., gaze field [65]), which is a stack of slices, was handled as a concatenated image to apply the metric. Compared to other candidate measures, SSIM is a commonly used perceptual

metrics due to its comparable or superior performance to other complex metrics simulating the human visual system even with its considerably lower computational expense. This metric attempts to calculate difference in structure between two images.

- SSIM:
$$\frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)} \quad (2)$$

Where μ_x and μ_y are the average intensity of gaze data; σ_x^2 and σ_y^2 are the variance of gaze data for two readers; σ_{xy} is the covariance of gaze data for two readers; and c_1 and c_2 are the constants from $(k_1L)^2$ and $(k_2L)^2$ respectively, in which L is the dynamic range of the pixel-values (255 in 8-bit precision), and k_1 and k_2 are 0.01 and 0.03, respectively. During the design process, we provided several other similarity measures as ranking criteria (e.g., sum of squared difference, peak-signal-to-noise-ratio, and normalized-cross-correlation), but SSIM was the most consistent with the perception of participating radiologists.

Users can select a group of readers in the similarity matrix view (Figure 4.1A) and quickly perform a correlation analysis on the selected readers in the correlation matrix view (Figure 4.1B). The correlation analysis can help users perform more focused analyses in the subsequent comparative gaze analysis in the comparison tab by identifying important contextual information.

4.2.2 Qualitative Similarity Overview

Along with the quantitative overview in the similarity matrix view, users can check the overall gaze patterns of multiple readers using small multiples of familiar spatial

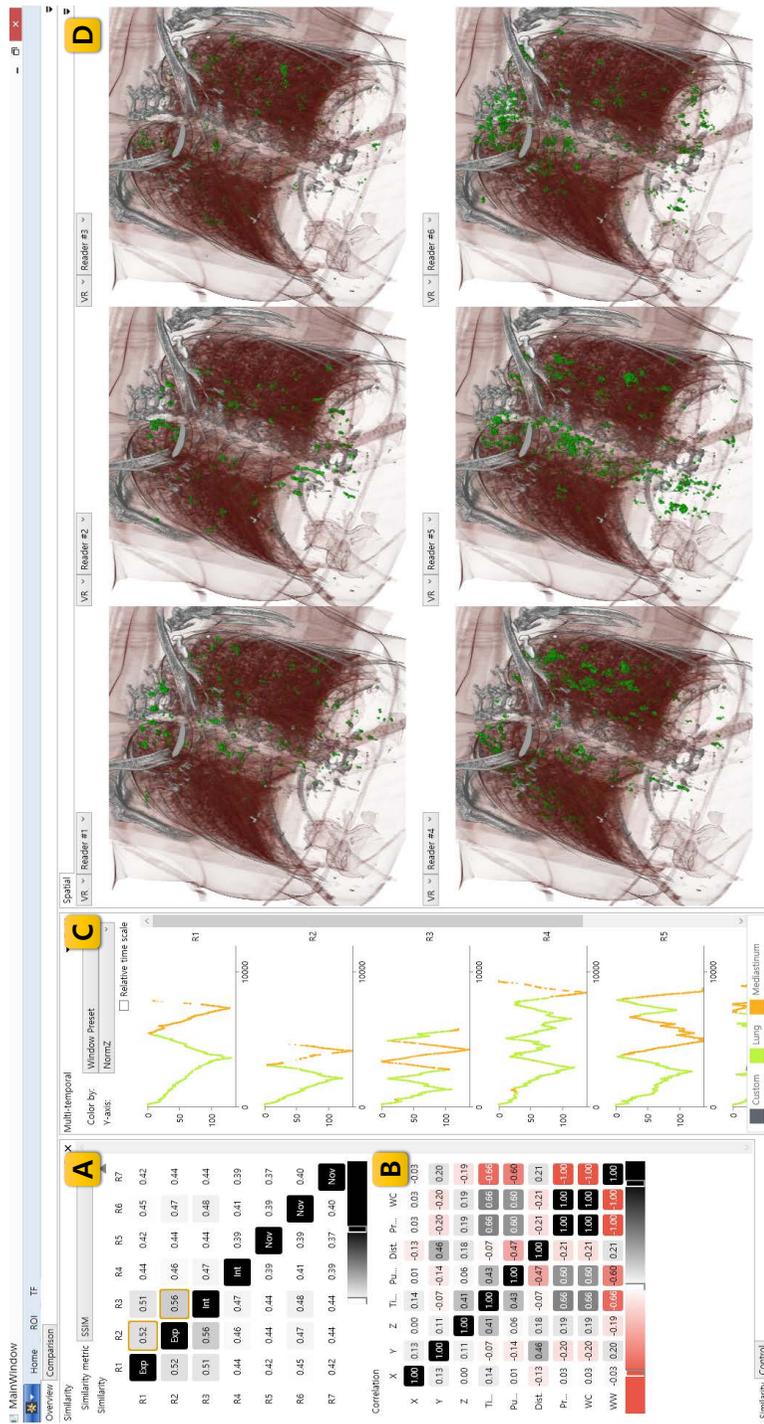


Figure 4.1. GazeDx interface for chest normal case (Overview tab). (A) Similarity matrix showing pairwise similarity between readers computed with SSIM (Structural Similarity). (B) Correlation matrix showing correlation between each pair of dimensions. Background of each cell in both matrices is color coded by its magnitude. (C) Multi-temporal view with interactive temporal charts. (D) Spatial view showing gaze data superimposed on 2D MPR (axial, coronal, and sagittal) or 3D VR images

representations. In GazeDx, we adopt a number of spatial representations for gaze pattern. It is based on the observation that the superimposition of gaze data over visual stimuli is effective for interpretation [66]. Thus, we visualize the gaze data using conventional 2D and 3D representations in medical field, namely 2D MPR images of axial, coronal and sagittal images, and 3D VR images. Using these conventional representations, users (radiologists) can qualitatively compare the gaze data of multiple readers as if they were reviewing CT images in their clinical practices. The spatial representation, which shows the 2D and 3D scenes, is presented as small multiples in the overview tab (Figure 4.1D). For each view, users can determine whose gaze data is displayed in which image representation among the four (axial, coronal, sagittal, and 3D VR) by using left and right combo boxes at the top of each view. The gaze data is superimposed on the 2D MPR images and/or on the 3D VR image using the gaze field approach of GazeVis [65]. GazeVis stores the gaze data in a scalar field of 8-bit precision whose value, at each position, represents the magnitude of fixation duration at the position.

4.2.3 Multi-level Temporal Overview

The temporal dimension is yet another important perspective that could reveal meaningful differences between readers [1]. In GazeDx, users can check the overall temporal gaze patterns of multiple readers using small multiples of a temporal chart (Figure 4.1C), i.e., navigation chart [1]. The original navigation chart is a type of scatterplot that plots the index of gazed slice (mapped to y -axis) over time (mapped

to x -axis). We generalized the original navigation chart, enriched the visual encoding options, and designed user interactions to make an interactive version of the navigation chart, which we call the “interactive temporal chart.” It is enhanced to support both absolute and relative time scales as the time it takes to make a diagnosis differs between readers. In the absolute time scale, the tick below the chart shows actual elapsed time in seconds. Users can easily compare the difference in diagnostic duration (who finished the image review quickly or slowly). In addition, one can use the relative time scale to examine the similarity in and difference between the overall temporal patterns of gaze for multiple readers, regardless of duration. Compared to the original navigation chart, it is improved in two ways: (1) configurable y -axis and (2) color-coding capability with categorical values. As an example, when setting pupil size to the y -axis of the plot, one can not only observe the pupil size over time, but one can, when color-coding the plot by window setting, also inspect whether the image brightness and contrast influence the temporal trend of pupil size in a single view.

In the multi-temporal view, a number of categorical attributes, including expertise level, selection state by a user, and windowing information can be used for color-coding. The windowing information consists of two continuous values each for image brightness and image contrast. However, we categorized the values into several groups depending on the body part. Radiologists usually select a window setting from a set of window settings pre-defined in the PACS (picture archiving and

Table 4.1. Window presets

Body part	Window setting	Window level (HU)	Window width (HU)	Shortcut (hotkey)
Chest	Lung	-600	1500	F1
	Mediastinum	30	400	F2
Abdomen	Soft tissue	45	440	F1
	Bone	500	4000	F2
	Lung	-500	2000	F3
	Liver	130	160	F4

communication system) diagnostic workstation to review the corresponding organs closely. There are typically two window settings for the chest (lung and mediastinum settings) and four window settings for the abdomen (soft tissue, bone, lung, and liver settings) (Table 4.1). Moreover, radiologists tend to customize the window value further as needed by dragging the right mouse vertically (adjusting the window level to change the image brightness) or horizontally (adjusting the window width to change the image contrast).

GazeDx supports semantic level-of-detail exploration of the temporal aspect of gaze data with three visual representations: (1) navigation chart, (2) stacked bar chart, and (3) bar chart (Figure 4.2). The interactive temporal chart transforms its view from one representation to another depending on users' analysis needs. The first representation (a navigation chart) is used to show temporal gaze patterns of a reader as an overview from the temporal aspect. In this representation, one can compare

readers' overall gaze pattern, in terms of slice number, pupil diameter, or distance to the monitor. As the space becomes smaller, information mapped on the y -axis is neglected and the change of categorical contextual information over time is plotted in a stacked bar chart. With this representation, one compares the readers based on the changes of categorical contextual information over time. When the available space becomes even smaller, gaze data is aggregated into a number of groups depending on selected contextual information. Consequently, the proportion of each group in each gaze data is visualized using a bar chart. Thus one can effectively check the distribution of the selected categorical contextual information.



Figure 4.2. Semantic exploration of temporal aspect of gaze data by smooth transition of the representation of the multi-temporal view according to space availability. (a) Navigation chart. (b) Stacked bar chart. (c) Bar chart.

4.3 In-depth Comparison of Gaze Patterns

After finishing examining overviews and overall comparison in the overview tab, users can perform in-depth comparisons with a set of selected readers in the comparison tab while taking into account important contextual information (Figure 4.3). In this section, we describe the comparison tab, which consists of several vertical panes: the aggregation pane – the left most pane for showing aggregation results and ROI (Region of Interest) selection, and individual view panes – other panes for showing gaze data for the selected individual readers with each user in a separate pane.

4.3.1 Detail Views for Individual Readers

Each reader selected in the overview tab has a dedicated pane for the in-depth examination of multidimensional aspects of the corresponding gaze data. Users can also change the reader for a pane by selecting a different reader in a combo box in the pane. Each pane consists of three main views: a scatterplot matrix to show an overview of all scatterplots (Figure 4.3A), a context-embedded interactive scatterplot to investigate a selected scatterplot in detail (Figure 4.3B), and an interactive temporal chart (Figure 4.3C). Although users can investigate the temporal aspect of the gaze data by setting the horizontal axis as time in a scatterplot, we decided to show an interactive temporal chart as a separate visualization for further analyzing the gaze patterns of an individual reader. We made this decision because temporal analysis is one of the key analytical tasks in gaze-based clinical studies [1]. The interactive

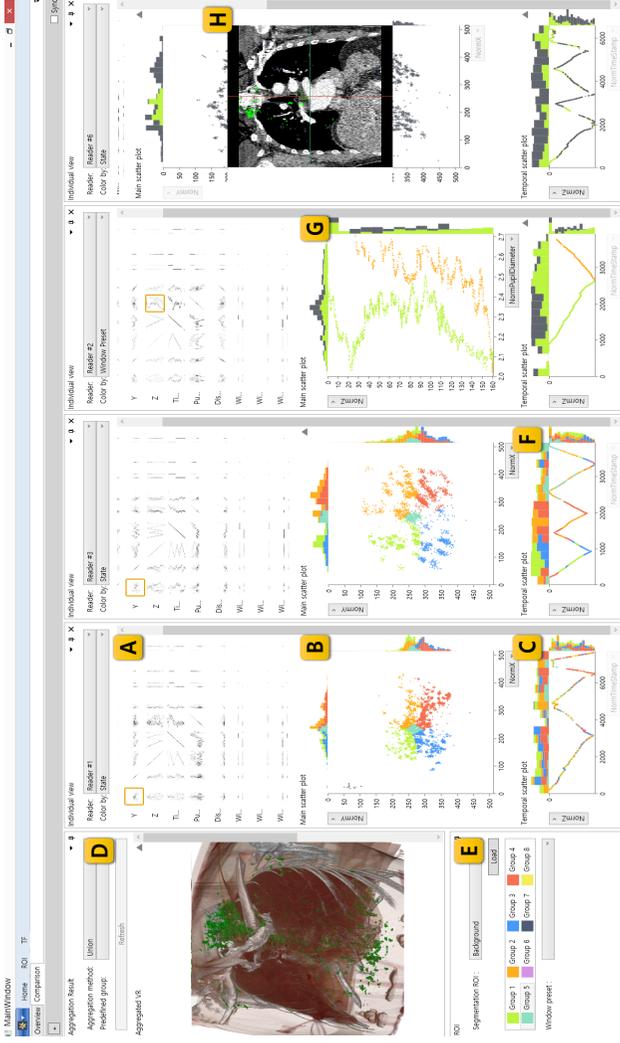


Figure 4.3. GazeDx interface for chest normal case. (A) Scatter plot matrix. (B) CIS: context-embedded interactive scatter plot with gaze points grouped into five clusters. (C) Interactive temporal chart. (D) Aggregation pane showing aggregated gaze data (E) ROI selection pane with ‘Segmentation ROI filter’ and ‘Window preset’ filter. Interactive temporal charts in (C) and (F) show the difference in gaze pattern between scanner and driller strategy (frequently changing color in (C), scanner; relatively consistent color in (F), driller). (G) CIS with horizontal axis of pupil diameter and vertical axis of slice index, where the gaze data is color-coded by window preset (left-shifted green cluster, lung setting; and right-shifted orange cluster, mediastinum setting). (H) CIS with embedded coronal plane image.

temporal chart is further enhanced from the interactive temporal chart used in the overview tab by attaching histograms to the top and the right side of the chart (in the same way as we do for the interactive scatterplot in section 4.4). We made this design decision to support users' analytical needs to view a higher-level summary of the data while examining the detail in the interactive temporal chart. Top and right histograms help reveal distributions of gaze points across temporal and vertical axes respectively.

4.3.2 Aggregation for Group Comparison

Researchers compared the gaze patterns of multiple readers after stratifying the readers by a contextual categorical variable, such as expertise level or stress level in numerous studies [31], [42], [55]. To facilitate such analytical needs, GazeDx provides users with an aggregated overview for comparison between groups (for example, experts vs. novices) (Figure 4.3D). Users can interactively define a new group of readers by selecting the readers for the individual view panes. In addition, a user can select a predefined group based on the expertise level using the combo box at the top of the aggregation pane. Upon selecting a group from the combo box, individual view panes are filled with readers that correspond to the selected expertise level. GazeDx builds a new aggregated collection of gaze data for the corresponding readers using one of the four types of aggregation options: union, intersection, average, and max. The aggregated data is visualized as 3D VR representation in the aggregation pane. Union and intersection are set operations that result in a new aggregated gaze data with binary values of either 0 or 255 (minimum and maximum

gaze values in 8-bit precision). The average and max operations result in a new gaze data with the average and maximum gaze intensity values respectively. Users can obtain full gaze coverage of a group of readers efficiently using the union option, or the commonly gazed region of a specific group using the intersection option.

4.4 CIS: Context-embedded Interactive Scatterplot

4.4.1 Flexible Axis Configuration

In our framework, we included a number of contextual information in each gaze point to form a collection of multidimensional data points. In order to support visual exploration from diverse perspectives of such data, we adopted a scatterplot with configurable axes. Depending on the property of a selected axis, the range and origin of the axis change to make the scatterplot more intuitive and understandable to users. For spatial variables x and y , and contextual information, such as pupil size and distance-to-monitor, the origin is placed at the bottom-left of the scatterplot, as is usual in the typical Cartesian coordinate system. However, a spatial variable z (for slice index) is placed in an inverted direction when mapped to the vertical axis of the scatterplot (z value ascending downward along the vertical axis). We made this design decision because radiologists are familiar with the inverted mapping for the z -axis in practice. Unlike abstract contextual information, gaze points from volumetric medical images are mapped naturally to the human body, and thus it is natural to place gaze points on the first slice for the top of the body with the smallest z value at the top.

4.4.2 Focus Attention with Familiar Representations

The scatterplot is extended by embedding medical imaging representations well known to radiologists. Even though a scatterplot is an effective visualization to reveal bivariate relationships, it is not intuitive for radiologists to connect the points in the scatterplot to the anatomical structure of the human body, which is of greater clinical significance. Therefore, we embedded contexts that are more meaningful and intuitive into the scatter-plot by integrating familiar medical imaging representations. When a group of gaze points is selected in the scatterplot (as described in section 4.5.1), a user can right-click on it to investigate it in a 2D MPR representation embedded in the scatterplot.

To make this transition easy to follow, we designed a multi-stage animation. Assume gaze points are selected by users or filtered by a criterion (detailed in section 4.5) in the scatterplot with spatial variables x and z mapped to its horizontal and vertical axes, respectively. If a user right click on any area in the x - z scatterplot (Figure 4.4a), the scatterplot is divided into three parts: upper area, middle area with selected gaze points, and remaining lower area. Then the middle area is expanded vertically, transforming itself into a 3D VR image where the selected gaze points are placed on a 3D representation of a patient's body that carries greater clinical meaning (Figure 4.4b). Afterward, the 3D VR representation with the gaze points rotates around the horizontal axis to show the x - y plane of the volumetric images (Figure 4.4c). Finally, the x - y plane (axial image) appears with gaze points overlaid

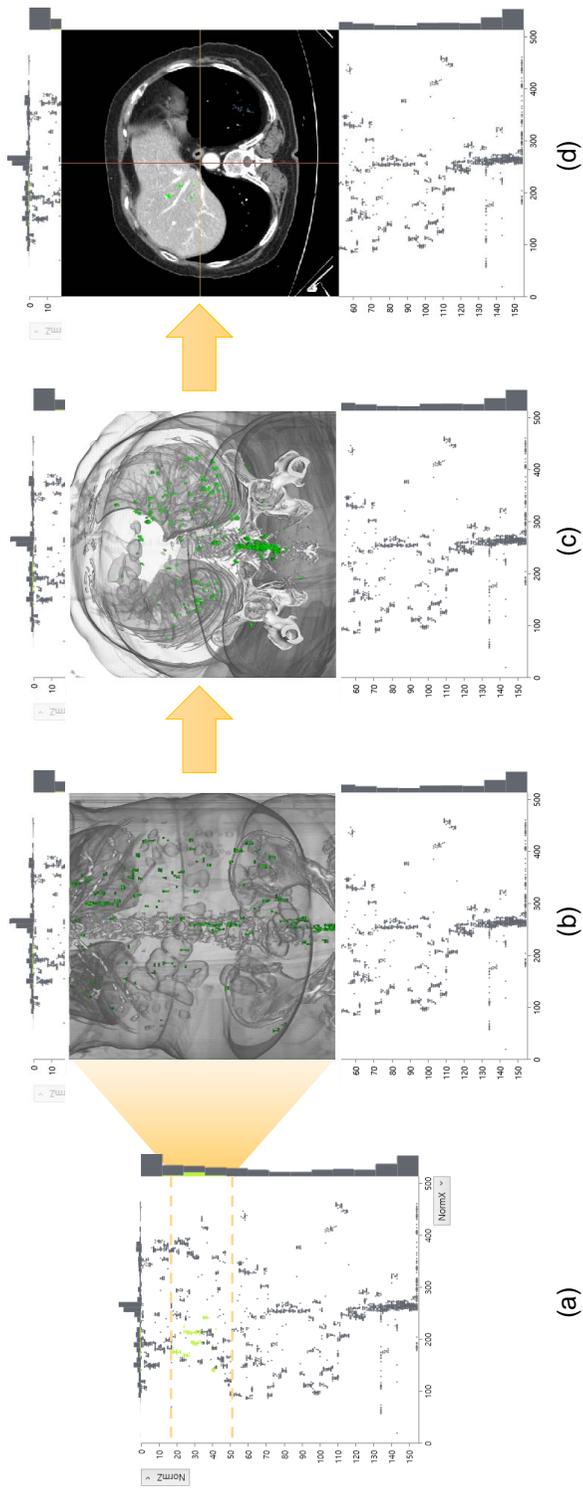


Figure 4.4. Animated context-embedded interactive scatterplot (CIS). (a) Gaze points are selected by users or filtered by a criterion in the x-z scatterplot. (b) The scatterplot gets divided into three parts, and the middle area is vertically expanded into a 3D VR image. (c) 3D VR representation rotates around the horizontal axis. (d) The x-y plane (i.e., axial image), which is invisible in the original x-z scatterplot, appears with gaze points overlaid to help users look into the data from different but related aspects in a single view.

(Figure 4.4d). The user can scroll up and down x - y plane images by using wheel scrolling. In this way, gaze points projected on the same position in the original x - z scatterplot appear and can be examined in the x - y plane image, enabling user to compare the selected (or filtered) gaze data in a single space (i.e., CIS) in terms of all three spatial dimensions. After spatial comparison as needed, one can close the embedded 2D plane image by right click on any area not the 2D image in CIS.

We intentionally employed the 3D VR representation during transformation to take advantage of the distinct characteristics of gaze tracking data: innate 3D structures of volumetric images. Because radiologists use 3D VR and MPR representations in their everyday practice, they can not only follow the transformation easily, but also better understand the gaze pattern in a more familiar and clinically meaningful context. In addition, by embedding a familiar visual representation within the scatterplot, users can investigate the gaze patterns in a different context without changing the focus of their attention to other views. To some degree, CIS enables GazeDx to use the screen space with greater efficiency so users can compare the gaze patterns of multiple readers in a limited screen space.

We further improve CIS by attaching two histograms on the top and right sides of the scatterplot to show the distribution of gaze points across horizontal and vertical dimensions, respectively. These histograms can help ease the over-plotting problem of gaze points when numerous regions of interest are located closely, which was observed in the case study. To increase the robustness of the histogram, we

determined the size of bins based on the Freedman-Diaconis rule [23].

4.4.3 Scatterplot Matrix with CIS

As the number of dimensions of the gaze data increases, one must perform a geometrically increasing number of pairwise comparisons to review the gaze data for each individual reader fully. We introduced a scatterplot matrix to resolve this issue. One can explore the data space by alternating the axis in the dynamic scatterplot, but one must spend a considerable amount of time to cover all possible pairs of axes completely. To boost the exploration, GazeDx employs a scatterplot matrix for each individual reader (Figure 4.3A). The rows and columns of the matrix stem from the dimensions of the gaze data and each cell in the matrix contains a thumbnail image of a scatterplot with dimensions corresponding to the combination of the row and column. One can click on a specific cell to see the corresponding scatterplot in the aforementioned CIS. The selection can be synchronized across all views for other readers to encourage comparison across readers.

4.5 Interactive Selection and Filtering

With the aforementioned visual components, one can select a small number of readers of interest to narrow down the exploration space. However, the number of gaze points for each reader is still over ten thousand even for a short three-minute diagnosis; examining them all at once easily exceeds one's perception capability. Moreover, the inclusion of contextual information in gaze analysis increases the number of

dimensions. To support the flexible exploration of such large amounts of multidimensional gaze data from diverse perspectives, GazeDx provides interactive selection and filtering capabilities to help users efficiently select and compare groups of gaze points in which they are interested. Our review of existing gaze analysis systems and gaze-based clinical research led us to (1) a selection interaction using freehand drawing, which is we believe is the most flexible selection interaction based on direct manipulation and (2) a filtering interaction based on human anatomy, which is a novel interaction to filter gaze points within an anatomical structure of interest.

4.5.1 Selection by Freehand Drawing

GazeDx supports the interactive selection of gaze data with free-drawn boundary drawing in CIS. Users can select multiple groups of gaze points using this interaction. Upon drawing a boundary, selected gaze points within the boundary are highlighted in a specific color while two histograms on the top and right sides of CIS show distributions of the selected gaze cluster along with the overall distribution (Figure 4.3C and Figure 4.3E). To support multiple view coordination, the selected gaze cluster is highlighted in that specific color in the interactive temporal chart below CIS. The selection persists even after changing the axis of CIS. Thus, one can investigate gaze points of interest from diverse perspectives by taking advantage of the configurable axes feature in CIS. To improve the consistency of interaction, identical querying interaction is supported in the interactive temporal chart below CIS, and the selection is always synchronized between CIS and the interactive temporal chart.

When one selects multiple gaze clusters by drawing multiple boundaries, he/she can assign a color to an individual gaze cluster from a list box on the leftmost pane (Figure 4.3H). The assigned color applies to the two histograms on the top and right sides of CIS and to the interactive temporal chart below. The multiple view coordination between CIS and the interactive temporal chart makes it possible for users to comparatively analyze gaze clusters comparatively from multidimensional perspectives, including spatial and temporal perspectives. For example, that radiologists' diagnostic process can be divided into two phases: skimming and verification [1]. When a radiologist diagnoses a volumetric CT or MR image, he/she usually makes a decision after scrolling up and down the whole image several times, wherein the last scroll is regarded as the verification phase and the prior scrolls are regarded as the skimming phase. If one draws a boundary surrounding the gaze points for the last scroll and draws another boundary surrounding the remaining gaze points in the interactive temporal chart, he/she can compare the gaze points in those two gaze clusters in the upper CIS from spatial and multidimensional perspectives while changing its horizontal and/or vertical axes.

4.5.2 Selection by Human Anatomy

It is common in gaze-based radiology studies to investigate whether readers' gaze points are in a region of interest that is usually an organ or a lesion. However, it is extremely labor intensive, if not impossible, to define such a region of interests in an existing gaze analysis system. We introduce a novel filtering technique in GazeDx

that utilizes segmentation results. We extract an organ or a lesion of interest (for example, lung, liver, metastasis, and nodule) a priori by using an external segmentation tool. Using these segmentation results, GazeDx supports clinically meaningful spatial filtering. If one selects a segmentation result using the combo box of 'Segmentation ROI' (Figure 4.3H), GazeDx shows only the gaze points within the segmented result to help users examine the gaze patterns confined to the corresponding organ or lesion.

Such an anatomical spatial filtering can support multidimensional exploration of gaze data when examined using configurable axes of CIS for the contextual information. Combined with the windowing information which is important contextual information affecting the visibility of stimuli (organs or lesions), anatomical spatial filtering could enable gaze comparison in terms of diagnostic performance or behavior. For instance, in the lung (window) setting for chest scans, the structures inside lungs are clearly visible, whereas the mediastinum (region between the right and left lung) and soft tissue are greatly saturated to a point where they become almost indistinguishable (Figure 4.5a). In the mediastinum (window) setting, the mediastinum and soft tissues are visible; the lung is mostly black (Figure 4.5b). By combining the anatomical query with the windowing information, one can easily identify spatially mismatching gaze points. For example, by selecting 'Segmentation ROI' of the lung, one can easily find gaze points on lungs with the

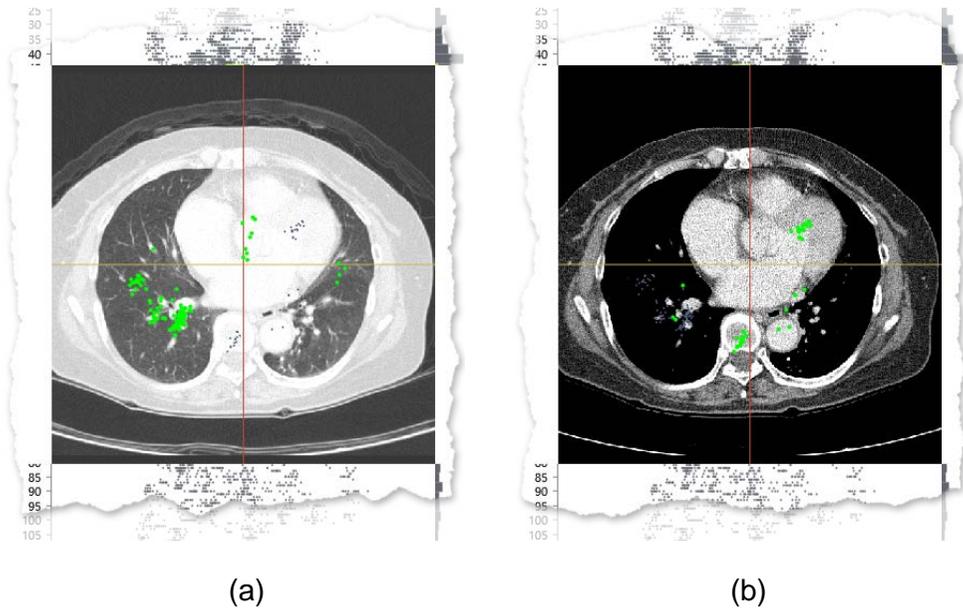


Figure 4.5. Axial images embedded in CIS, showing gaze points of a reader for the chest lesion case used. (a) In the lung window setting, the gaze points are scattered largely on both the lung and the mediastinum even though the mediastinum becomes too saturated to see detail. (b) In the mediastinum window setting, the gaze points are distributed mostly on the mediastinum.

mediastinum setting, which indicates that a reader gazed at the mostly invisible lungs. Such gazing can be understood as useless gazing or as the result of other human factors, regarding the diagnostic performance or behavior.

4.6 Implementation

GazeDx is implemented with C# and WPF using Microsoft DirectX SDK and SharpDX API. Scenes using DirectX technology are drawn at 30fps. We implemented a separate C# application for collecting gaze data and related contextual information during diagnosis. In addition to the gaze data, the application saves the index of the

currently shown slice, windowing information, pupil diameter, and the distance from the monitor, by using Tobii analytics SDK. The application is capable of displaying DICOM (Digital Imaging and Communication in Medicine) images as commercial PACS with zooming, panning, and full windowing functionalities, including shortcut keys for predefined window settings.

4.7 Case Studies

We followed an evaluation method for information visualization, namely Multi-dimensional In-depth Long-term Case studies (MILCs) [61], [64] to evaluate the efficacy and effectiveness of GazeDx as an interactive visualization framework for comparatively analyzing multiple gaze data from diagnoses with volumetric images. A comparative evaluation study was not feasible because there were no comparable gaze analysis systems that could handle large multidimensional gaze data from multiple readers. We conducted two case studies, each with a different body part, at a university hospital. We acquired gaze data from 14 radiologists (seven chest radiologists and seven abdominal radiologists). We collected their gaze data while each radiologist read two patients' CT images. After collecting gaze data, the two most experienced radiologists among the 14 performed comparative gaze analysis on the collected gaze patterns of their colleagues (including their own) with GazeDx. We used a modified pair analytics method [35] for our case studies, where the main designer of GazeDx took the wheel and the two radiologists directed the analysis using their expertise in the field.

4.7.1 Case Study Protocol

We conducted two case studies following a protocol with three phases: preparation, gaze collection and evaluation of GazeDx. In this section, we describe the protocol in detail.

(1) Phase I: Preparation of Case Study

Phase I included daily meetings with two radiologists (one chest and one abdominal) to determine task allocation and datasets for gaze acquisition. This was done to construct an experimental setting that would best resemble radiologists' workspace and diagnostic settings. Based on feedbacks from our expert radiologists, we developed an independent gaze collection application, which is similar to the PACS system in terms of functionality and appearance; however, we enhanced the functionality of PACS to include the collection of contextual information such as index number of gazed slice. These features not inherent in the conventional PACS application. Task was refined to enable readers to use a set of predefined window settings when collecting gaze data. Feedback from our radiologists suggested that, in clinical practice, radiologists quickly adjust window values using hotkeys (Table 4.1). As such, identical hotkeys for window value selection were incorporated into the gaze collection application as well as in GazeDx. Task design also included preparation of a short written report after diagnosis, again to mimic actual clinical practice settings closely.

Table 4.2. DICOM files used in the case study

Body part	Usage	Dimension	# of images	Remarks
Chest	Training	512×512	154	Normal
	Main study	512×512	161	Normal
	Main study	512×512	137	Lesions (multiple nodules)
Abdomen	Training	512×512	150	Normal
	Main study	512×512	155	Normal
	Main study	512×512	110	Lesions (huge hepatic mass)

Three datasets were chosen by radiologists for each body part: one for the training session; a second without lesions; and a third with notable lesions (Table 4.2), in order to allow for gaze comparison between normal and abnormal cases. Under the guidance of our expert radiologists we determined a desired focus of interest within a particular window setting. The delineated focus was subsequently segmented into regions of interests (ROIs) according to provided feedback. Segmentation included lung and surrounding structures for chest CT data, and lung, liver and surrounding structures for abdominal CT data using proper segmentation algorithms[27], [38].

(2) Phase II: Gaze Collection

Before beginning Phase II, participants had a training session in order to become accustomed to the interface. Training sessions took less than 5 minutes for all participants owing to the application's similarity to the conventional PACS

system. Additionally, we calibrated the eye tracker using a 9-point calibration in preparation for the main gaze collection phase. During the gaze collection phase, participants were asked to review images for diagnosis, then write a brief report outlining their findings. Time to completion for this task was approximately 15 minutes (including training session) on average. The gaze collection application recorded both gaze data and contextual information.

(3) **Phase III: Evaluation of GazeDx**

In Phase III, we introduced GazeDx to two chosen expert radiologists. We used a modified pair analytics [35] method whereby the main developer is the visual analytics expert and the radiologists are subject matter experts. During the evaluation phase our radiologists quickly became accustomed to GazeDx and were able to manipulate the application without need for assistance. The developer met with the radiologists four times during the evaluation phase in order to further refine the application. We asked each expert to compare the collected gaze data of seven readers using GazeDx and to look for any notable similarities or differences between both readers and datasets. Additionally, diagnostic comments made by the expert radiologists regarding notable gaze patterns using GazeDx were collected.

4.7.2 Apparatus

For gaze data collection, we used a PC with a quad-core processor connected to a 20-inch Barco medical monitor. A Tobii X60 eye tracker with 0.4° optimum accuracy

was used. The distance from participants to the eye tracker was approximately 65 cm. For gaze comparison using GazeDx, we used an Intel i5 3.4 GHz quad core PC equipped with 8 GB of main memory, and NVIDIA GTX 760 GPU with 2 GB of memory.

4.7.3 Case Study 1: Chest Radiologists

Following the case study protocol described above, we conducted a case study with seven chest radiologists: two with more than 16 years' experience (expert; R1 and R2), two had about four to six years' experience (intermediate; R3 and R4), and three in first to second years of residency (novice; R5, R6, and R7). All physicians were asked to examine radiologic chest images for the purposes of diagnosis (Table 4.2). Following data collection, one of the senior chest radiologists was asked to compare the gaze data using GazeDx.

Using the images without any notable lesion (i.e., the normal case), the radiologist first examined the overview using 3D VR images in the spatial view (Figure 4.1D). However, as the number of gaze points for each reader was approximately 7,000 points, it was difficult to find notable differences in the 3D VR images. Instead, he used the similarity view. Here, readers were labelled as R1 to R7 to reflect level of expertise in descending order of expertise (Figure 4.1A), and he found that the expert radiologists had a more similar gaze pattern (dark grey) than the novices (light grey).

Next, he used the multi-temporal view for more detail (Figure 4.1C). With slice

index mapped to the y -axis in the interactive temporal charts, he determined that the novice radiologists tended to use the image scroll feature more frequently than the expert radiologists did. Moreover, he noted a visible difference in gaze patterns between experts and novices using the temporal charts. Here, the experts' gaze patterns formed a straight line while the novices exhibited a large number of gaze fluctuations within a short time period (temporal charts in Figure 4.1C). This indicates that the experts had scrolled through the images at a relatively constant speed whereas the novices had spent varying amounts of time on each image. This was hypothesized by our expert radiologist as a difference in use of peripheral vision and diagnostic confidence. During diagnosis experts are able to use their peripheral vision effectively, gaining wider coverage with a single glance at an image. In contrast, novices have less experience using their peripheral vision and less confident in their decision-making. As a result, they tend to reexamine images many times before making a diagnosis.

He also found a tendency among all readers to exhibit stiff angle with the mediastinal setting in the multi-temporal view (Figure 4.1C). Based on the window preset and using color coding, there was a commonality in scrolling speed (orange for mediastinum vs. green for lung). This was explained through human anatomy: the mediastinum is a small region with relatively fewer inner organs, and so it is relatively easier to review than other areas, regardless of the expertise level of the reader.

Following analysis of the overall gaze data, our expert radiologist moved to the

comparison tab to investigate individual gaze patterns. In order to determine viewing style, he used CIS with spatial variables x and y mapped to its horizontal and vertical axes. Next, he divided the plotting area into five regions each with distinct colors: upper-left, upper-right, lower-right, lower-left, and central regions so as to determine whether readers are scanners or drillers [16] (Figure 4.3C and Figure 4.3E). Gaze points in the interactive temporal chart below were painted with the same colors. Here, he could easily discern a difference in viewing styles between those who scroll images while spreading their gaze in a wider area in a single image (R1) (Figure 4.3) and those who scroll images while focusing their gaze on a localized area (R3) (Figure 4.3F).

Gaze data was further examined for notable similarities and differences in pupil size amongst readers, since there is a known correlation between pupil diameter and task difficulty. To look for a correlation between pupil diameter and human anatomy, our expert radiologist first used the scatterplot matrix and selected the cell corresponding to the pair of slice indices and pupil diameter (see orange rectangle above Figure 4.3). Next, he color-coded the gaze data in CIS with window preset. He noted two distinct clusters with notable valley on both lung and mediastinal settings (left-shifted green cluster, lung setting; and right-shifted orange cluster, mediastinum setting, respectively) (Figure 4.3). This was explained as a generally brighter illumination in lung setting (Figure 4.3) resulting in smaller pupil diameter. The two distinct valleys were due to human anatomy: those areas with inherent bone structures

or internal organs appear brighter than areas of tissue only, seen in identical window settings (see Figure 4.3H). As a result, it revealed that further analysis is required to derive correlation between cognitive load and pupil diameter in this experimental setup.

Our expert then examined his own gaze pattern in the lesion case. In particular, he focused on the windowing information affecting the stimuli visibility in the images. In the individual view pane he expanded an axial plane image in CIS by drawing freehand and selecting most of his gaze data. He looked at his gaze pattern across both lung and mediastinal images by changing window preset filter. Interestingly, he found a notable difference in gaze pattern between mediastinal and lung images: in the mediastinal setting, his gaze was directed solely at the mediastinum; however, in the lung setting his gaze was concentrated on both lung and mediastinum (Figure 4.3). This is of interest since the mediastinum becomes saturated in the lung setting, resulting in difficulty distinguishing detail in this area. He explained this particular gaze pattern as the visual ability to see both areas of interest while focusing one's gaze on a midpoint, in this case the mediastinum. On completion of this task, our expert concluded that GazeDx's context-embedded nature is useful for examining anatomic structures within a scatterplot without switching views. By focusing on one view, the gaze pattern can be delineated in close detail without the reader's attention being shifted.

4.7.4 Case Study 2: Abdominal Radiologists

Seven abdominal radiologists participated in this case study including two with more than 10 years' experience (expert; R1 and R2), two with approximately four years' experience (intermediate; R3 and R4) and three in their first to third year of residency (novice; R5, R6, and R7). They were asked to perform diagnoses using two sets of CT images. Gaze data was collected after which we asked one expert abdominal radiologist to compare the gaze data using GazeDx.

Gaze comparison was first conducted on the normal case. The expert radiologist first tried qualitative comparison using spatial views in the overview tab, which revealed an overall uniformity between readers in the similarity view. Next, he examined the multi-temporal view. This showed a uniformity of time spent to completion between expert readers and a similarity between all readers in terms of gaze pattern with the exception of R4. On examination of R4's gaze pattern the expert radiologist determined that R4 had not applied the liver window setting during diagnosis. This is unusual as the liver is a standard part of abdominal radiology. By shortening the width of the multi-temporal view and checking the bar chart representation (Figure 4.2c) using the mouse, he discovered that R4's bar chart representation did not include a bar corresponding to the liver setting. This indicates that R4 did not apply the liver setting. To further investigate R4's gaze pattern, the expert moved to the comparison tab and selected R4 in an individual view pane. In the scatterplot matrix, he selected a cell with z and x dimensions to examine the gaze

points spatially. He then applied the ‘Segmentation ROI’ filter for the liver and right-clicked on CIS to expand the axial plane image so he could examine the gaze points closely. This application revealed that R4 had, in fact, reviewed the liver, but not using the conventional window setting.

During review of the overview tab, there was interest in a group comparison based on expertise level; in particular, the indices of gazed slices over time. To examine this our expert used the comparison tab to conduct a comparative analysis on expertise level. First, he selected the expert group from the aggregation pane. Next, he selected the corresponding cell in the scatterplot matrix. From the histogram on the right side of the interactive temporal chart below CIS, he noticed that the experts’ gaze was directed to the upper abdomen. Using the comparison tab to review intermediates’ and novices’ gaze pattern, he noted a similarity between all users. This was explained by the anatomical structure of the human body, where the upper abdomen houses a number of major organs including the pancreas, duodenum, common bile duct, and common hepatic artery. Thus, readers tended to train their gaze more on the upper abdomen than the lower abdomen during diagnosis.

Gaze comparison was next performed on the case with notable lesions. The expert radiologist examined the similarity matrix and multi-temporal view, and then moved to the comparison tab. He was interested in the gaze difference between experts and novices when diagnosing an apparent lesion – in this case, a large hepatic mass. To analyze this difference, he prepared four individual view panes for R1, R2,

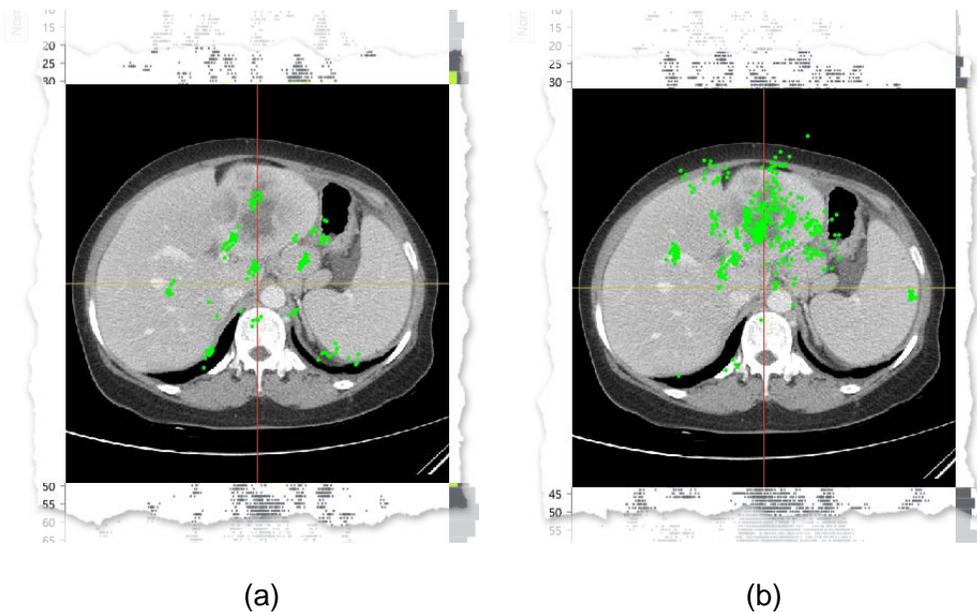


Figure 4.6. Axial images embedded in CIS, showing gaze points of (a) an expert and (b) a novice for the abdomen lesion case. (a) The expert's gaze points are relatively scattered on the image even with the clearly notable lesion (large hepatic mass). (b) The novice's gaze points are distributed mostly on the lesion.

R5, and R6. Then he used qualitative comparison after expanding the axial images in all CISs by scrolling systematically through the images. He found that experts tended to an evenly distributed gaze pattern throughout the abdomen despite the overt presence of the hepatic mass (Figure 4.6a). Conversely, novices tended to a concentration of gaze on the mass itself, with less attention paid to other abdominal organs (Figure 4.6b). This was explained as a difference in diagnostic confidence such that experts can quickly determine an overt area of concern while continuing to look for less overt anomalies.

4.8 Discussion

GazeDx has been developed to enhance comparative gaze analysis with combination of conventional visual representations and context-embedded visualization. The goal of the case study was to determine this combination could be an effective enhancement to former visualizations used for medical diagnosis of volumetric images. In a perspective on medical imaging, the preliminary result from our case study showed potential as an interactive exploration tool. The experts were able to explore the data, build a hypothesis based on the comparison result, and validate the hypothesis with the tool. As they were familiar to the adopted medical visual representations (i.e., axial, coronal, sagittal and 3D VR), they were able to easily determine the clinical implications with the combination of interactive visualizations. We speculate that the inclusion of familiar spatial visualization alleviated the complexity of three-dimensional cognition of data.

Along with the visualization, GazeDx explored a novel ROI-based filtering technique by utilizing anatomical segmentation results. Using the filtering results, one of the experts successfully identified whether one of the reader actually paid attention to a certain organ. The technique could be especially useful, as a typical dynamic stimulus often requires burdensome annotation to identify ROIs in each individual scene. Moreover, as the segmented ROIs are from anatomical structures, it could benefit the gaze analysis in the radiology field, which requires medical context. We expect a similar segmentation approach to identify ROIs in the

volumetric data could enrich the gaze analysis by providing more meaningful context.

While the expert radiologists actively used most of the components, they infrequently used similarity matrix (Figure 4.1A) and correlation matrix (Figure 4.1B) compared to the others. As each matrix showed overall comparison results, they were seldom used after gaining insights over the dataset. They also showed limitation by using aggregated values for each pair that they were unable to show partial similarities. Thus, it would be meaningful to extend the components to update the values upon selection of ROIs. Then, it could show partial comparison results within specific organs or within certain time spans as in ISeeCube [34].

The expert radiologists suggested that GazeDx could be an effective teaching tool in a medical tutoring environment by visualizing and comparing the gaze pattern of the experts as suggested by Vitak et al. [78]. Specifically, gaze data from multiple expert radiologists could be aggregated to form a model from which novices and trainees could be guided to read radiology images systematically and effectively. In addition, trainee gaze data could be collected and, using similarity measures, compared to the expert model, to act as a visual monitor of trainees' thoroughness in examining images. Further, GazeDx could be useful for general quality assurance, of particular interest in medical imaging where accuracy and completeness can determine vital medical treatment for patients.

As noted previously, one of the premises of eye tracking studies on medical diagnosis is that readers often identify lesions using their foveal vision. However,

previous studies have noted the significance of peripheral vision during diagnosis [20], and instances of this were noted during our case study. There were a number of instances of gaze patterns occurring in less diagnostically meaningful areas in a given window setting. According to our expert radiologist these gaze points occurred when readers were using their peripheral vision. However, as prior work on chest X-rays [78] reported that the difference in scanning strategies was the main cause of different scanpaths, whether a reader actually used peripheral vision should be further investigated. If we could distinguish when readers use their peripheral vision during diagnosis, we could adopt existing computation models of peripheral vision into GazeDx. Still, knowing when peripheral vision is in action during visual exploration is a challenging problem.

Chapter 5

Conclusion

In this dissertation, we presented a 3D gaze visualization technique called GazeVis and an interactive analytics framework called GazeDx to support gaze pattern comparisons for volumetric medical images. We designed visual representations and interactions to compare multiple gaze data effectively, focusing on multidimensional aspect of the data while incorporating clinically relevant contextual information (especially windowing information) into the analysis process. The visual components were tightly coupled to enable more effective visual exploration. The resulting framework showed efficacy in both explorative and comparative data analysis and in incorporation of real world diagnosis environment. An enhanced scatterplot (namely CIS) and interactive temporal chart played meaningful roles in achieving the design goal.

For GazeVis, case studies with twelve chest and abdominal radiologists revealed that the differences in expertise level and preferred diagnosis strategy of radiologists

led to significant differences in 3D gaze patterns. In the case studies with GazeDx, which were conducted with two expert radiologists, showed that the framework has the potential to be an effective exploratory data analysis tool for eye tracking studies in medicine. Moreover, we gained meaningful ideas that could guide our design improvements for GazeDx. The embedded familiar medical image representation provided users with useful additional contexts regarding the scatterplot in a CIS view, but the user interaction on the embedded representation could be further enriched. For example, when users rotate the representation with mouse interaction, we could even change the axes of the surrounding scatterplot to match the rotated representation. However, such interaction should be restricted to the cases where the gulf of evaluation is not wide.

While the proposed framework showed feasibility in analyzing and comparing multiple gaze data, there is a limitation that the analyses were mostly relied on users' manual manipulation. As the framework supports a number of visualizations with diversity of functionalities, some analyses might require excessive interaction to accomplish intended comparison (e.g., manipulation of multiple filters to analyze gaze points within lungs using a proper windowing setting). Such limitation could become worse when one wants to repeatedly perform an identical analysis over multiple data. As a remedy to this problem, we could provide automated analysis process for common cases. The processes for automation could be either identified from additional case studies with the radiologists, or an analyst could record previously performed analysis process as a macro for repetition.

The proposed framework can be applied to other domain. For instance, industrial

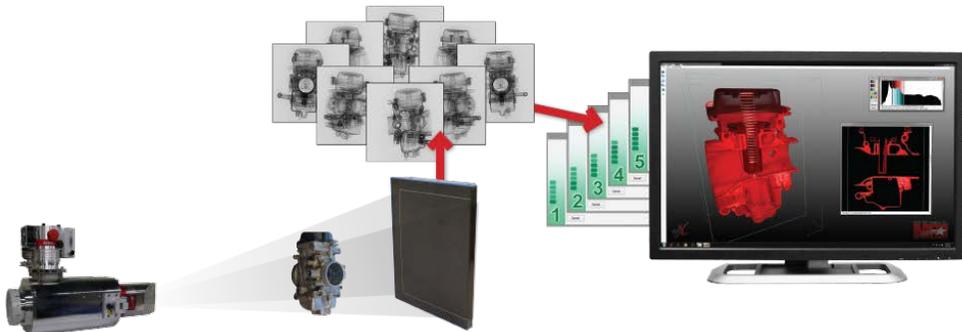


Figure 5.1. Three-dimensional industrial computed tomography X-ray inspection scanning. Multiple cross-sectional images are acquired and reconstructed to 3D for inspection. The image is from [72].

computed tomography (CT) scanning results in cross-sectional images that resembles volumetric medical images (Figure 5.1). Then, the resulting images are used for failure analysis and assembly verification that requires intensive visual search and there are several prior researches [26], [28] that worked on improving the quality of the process with novel visualization techniques. Thus, the eye tracking analysis with such volumetric stimuli can benefit from the framework with its similar characteristics that one would be able to explore multiple eye tracking data with closely linked visual representations with embedded context. As in the medical field, the analysis results using the framework could be used for both improving the quality of product inspection and educating complicated inspection techniques, which are hard to explain verbally, to the novices.

Bibliography

- [1] F. Amini, S. Rufiange, Z. Hossain, Q. Ventura, P. Irani, and M.J. McGuffin, “The Impact of Interactivity on Comprehending 2D and 3D Visualizations of Movement Data,” *IEEE Trans. on Visualization and Computer Graphics*, vol. 21, no. 1, pp. 122-135, 2015.
- [2] M.S. Atkins, J. Fernquist, A.E. Kirkpatrick, and B.B. Forster, “Evaluating Interaction Techniques for Stack Mode Viewing,” *Journal of Digital Imaging*, vol. 22, no. 4, pp. 369-382, 2009.
- [3] M.S. Atkins, A.E. Kirkpatrick, A. Knight, and B. Forster, “Evaluating User Interfaces for Stack Mode Viewing,” *Proc. SPIE-Medical Imaging 2007*, vol. 6515, pp. 65150A-65150A-10, 2007.
- [4] M. Basu, “Gaussian-based Edge-detection Methods—A Survey,” *IEEE Trans. on Systems, Man, and Cybernetics, Part C: Applications and Reviews*, vol. 32, no. 3, pp.252-260, 2002.
- [5] C.A. Beam, E. Krupinski, H.L. Kundel, E.A. Sickles, and R.F. Wagner, “The Place of Medical Image Perception in 21st-century Health Care,” *Journal of the American College of Radiology*, vol. 3, no. 6, pp. 409-412, 2006.
- [6] J. Bertin, *Graphics and Graphic Information-Processing*. Berlin: Walter De Gruyter, 1981.

- [7] R. Bertram, L. Helle, J.K. Kaakinen, and E. Svedström, “The Effect of Expertise on Eye Movement Behaviour in Medical Image Perception,” *PloS one*, vol. 8, no. 6, e66169, 2013.
- [8] T. Blascheck, K. Kurzhals, M. Raschke, M. Burch, D. Weiskopf, and T. Ertl, “State-of-the-Art of Visualization for Eye Tracking Data,” *Proc. of EuroVis 2014 State-of-the-Art Reports*, 2014.
- [9] M.A. Borkin, K.Z. Gajos, A. Peters, D. Mitsouras, S. Melchionna, F.J. Rybicki, C.L. Feldman, and H. Pfister, “Evaluation of Artery Visualizations for Heart Disease Diagnosis,” *IEEE Trans. on Visualization and Computer Graphics*, vol. 17, no. 12, pp. 2479-2488, 2011.
- [10] P. Blignaut, “Visual Span and Other Parameters for the Generation of Heatmaps,” *Proc. ACM Symp. Eye Tracking Research and Applications (ETRA '10)*, pp. 125-128, 2010.
- [11] A. Cockburn, and B. McKenzie, “3D or not 3D?: Evaluating the Effect of the Third Dimension in a Document Management System,” *Proc. ACM SIGCHI Conf. Human Factors in Computing Systems (CHI '01)*, pp. 434-441, 2001.
- [12] A. Cockburn, and B. McKenzie, “Evaluating the Effectiveness of Spatial Memory in 2D and 3D Physical and Virtual Environments,” *Proc. ACM SIGCHI Conf. Human Factors in Computing Systems (CHI '02)*, pp. 203-210, 2002.
- [13] A. Cockburn, “Revisiting 2D vs 3D Implications on Spatial Memory,” *Proc. Conf. Australasian User Interface (AUIC '04)*, pp. 25-31, 2004.
- [14] C.D. Correa, and K.-L. Ma, “Size-based Transfer Functions: A New Volume Exploration Technique,” *IEEE Trans. on Visualization and Computer Graphics*, vol. 14, no. 6, pp. 1380-1387, 2008.
- [15] R.A. Drebin, L. Carpenter, and P. Hanrahan, “Volume Rendering,” *Proc. ACM*

- SIGGRAPH Conf. Computer Graphics and Interactive Techniques (SIGGRAPH '88), pp. 67-72, 1988.
- [16] T. Drew, M.L. Vo, A. Owl, F. Jacobson, S.E. Seltzer, and J.M. Wolfe, "Scanners and Drillers: Characterizing Expert Visual Search Through Volumetric Images," *Journal of Vision*, vol. 13, no. 10, article 3, 2013.
- [17] S. Dübel, M. Röhlig, H. Schumann, and M. Trapp, "2D and 3D Presentation of Spatial Data: A Systematic Review," *Proc. IEEE VIS International Workshop on 3DVis (3DVis '14)*, pp. 11-18, 2014.
- [18] A.T. Duchowski, E. Medlin, N. Cournia, A. Gramopadhye, B. Melloy, and S. Nair, "3D Eye Movement Analysis for VR Visual Inspection Training," *Proc. ACM Symp. Eye Tracking Research and Applications (ETRA '02)*, pp. 103-110, 2002.
- [19] A.T. Duchowski, M.M. Price, M. Meyer, and P. Orero, "Aggregate Gaze Visualization with Real-time Heatmaps," *Proc. ACM Symp. Eye Tracking Research and Applications (ETRA '12)*, pp. 13-20, 2012.
- [20] S.M. Ellis, X. Hu, L. Dempere-Marco, G.Z. Yang, A.U. Wells, and D.M. Hansell, "Thin-Section CT of the Lungs: Eye-tracking Analysis of the Visual Approach to Reading Tiled and Stacked Display Formats," *European Journal of Radiology*, vol. 59, no. 2, pp. 257-264, 2006.
- [21] N. Elmqvist, P. Dragicevic, and J.-D. Fekete, "Rolling the Dice: Multidimensional Visual Exploration using Scatterplot Matrix Navigation," *IEEE Trans. on Visualization and Computer Graphics*, vol. 14, no. 6, pp. 1141-1148, 2008.
- [22] K.M. Fairchild, S.E. Poltrock, and G.W. Furnas, "Semnet: Three-dimensional Graphic Representations of Large Knowledge Bases," *Cognitive Science and its Applications for Human-Computer Interaction*, R. Guindon, eds.,

- Lawrence Erlbaum Associates, pp. 201-233, 1988.
- [23] D. Freedman and P. Diaconis, "On the Histogram as a Density Estimator: L2 Theory," *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, vol. 57, no. 4, pp. 453-476, 1981.
- [24] M. Gleicher, D. Albers, R. Walker, I. Jusufi, C.D. Hansen, and J.C. Roberts, "Visual Comparison for Information Visualization," *Information Visualization*, vol. 10, no. 4, pp. 289-309, 2011.
- [25] J.H. Goldberg, and J.I. Helfman, "Visual Scanpath Representation," *Proc. ACM Symp. Eye Tracking Research and Applications (ETRA '10)*, pp. 203-210, 2010.
- [26] M. Hadwiger, L. Fritz, C. Rezk-Salama, T. Höllt, G. Geier, and T. Pabel, "Interactive Volume Exploration for Feature Detection and Quantification in Industrial CT Data," *IEEE Trans. On Visualization and Computer Graphics*, vol. 14, no. 6, pp. 1507-1514, 2008.
- [27] H. Hong, J. Lee, and Y. Yim, "Automatic Lung Nodule Matching on Sequential CT Images," *Computers in Biology and Medicine*, vol. 38, no. 5, pp. 623-634, 2008.
- [28] R. Huang, K.-L. Ma, P. McCormick, and W. Ward, "Visualizing Industrial CT Volume Data for Nondestructive Testing Applications," *Proc. IEEE Conf. Visualization (VIS '03)*, pp. 547-554, 2003.
- [29] B. Kim, B. Lee, and J. Seo, "Visualizing Set Concordance with Permutation Matrix and Fan Diagram," *Interacting with Computers*, vol. 19, no. 5-6, pp. 630-643, 2007.
- [30] J. Kniss, G. Kindlmann, and C. Hansen, "Multidimensional Transfer Functions for Interactive Volume Rendering," *IEEE Trans. on Visualization and Computer Graphics*, vol. 8, no. 3, pp. 270-285, 2002.

- [31] H.L. Kundel, P.S. La Follette, Jr, "Visual Search Patterns and Experience with Radiological Images," *Radiology*, vol. 103, no. 3 pp. 523-528, 1972.
- [32] H.L. Kundel, C.F. Nodine, E.F. Conant, and S.P. Weinstein, "Holistic Component of Image Perception in Mammogram Interpretation: Gaze-tracking Study," *Radiology*, vol. 242, no. 2, pp. 396-402, 2007.
- [33] K. Kurzhals, B. Fisher, M. Burch, and D. Weiskopf, "Evaluating Visual Analytics with Eye Tracking," *Proc. Workshop on Beyond Time and Errors: Novel Evaluation Methods for Visualization (BELIV '14)*, pp. 61-69, 2014.
- [34] K. Kurzhals, F. Heimerl, and D. Weiskopf, "ISeeCube: Visual Analysis of Gaze Data for Video," *Proc. ACM Symp. Eye Tracking Research and Applications (ETRA '14)*, pp. 43-50, 2014.
- [35] B.C. Kwon, B. Fisher, and J.S. Yi, "Visual Analytic Roadblocks for Novice Investigators," *IEEE Symp. Visual analytics Science and Technology (VAST '11)*, pp. 3-11, 2011.
- [36] C. Lankford, "Gazetracker: Software Designed to Facilitate Eye Movement Analysis," *Proc. ACM Symp. Eye Tracking Research and Applications (ETRA '00)*, pp. 51-55, 2000.
- [37] B. Lee, J. Yun, J. Seo, B. Shim, Y. Shin, and B. Kim, "Fast High-quality Volume Ray Casting with Virtual Samplings," *IEEE Trans. on Visualization and Computer Graphics*, vol. 16, no. 6, pp. 1525-1532, 2010.
- [38] J. Lee, N. Kim, H. Lee, J.B. Seo, H.J. Won, Y.M. Shin, Y.G. Shin, and S.-H. Kim, "Efficient Liver Segmentation Using a Level-Set Method with Optimal Detection of the Initial Liver Boundary from Level-Set Speed Images," *Computer Methods and Programs in Biomedicine*, Vol. 88, no. 1, pp. 26-38, 2007.
- [39] M. Levoy, "Display of Surfaces from Volume Data," *IEEE Computer Graphics*

and Applications, vol. 8, no. 3, pp. 29-37, 1988.

- [40] J.D. Mackinlay, G.G. Robertson, and S.K. Card, "Perspective Wall: Detail and Context Smoothly Integrated," Proc. ACM SIGCHI Conf. Human Factors in Computing Systems (CHI '91), pp. 173-176, 1991.
- [41] Male Figure Realistic Standing Human, http://artist-3d.com/free_3d_models/dnm/model_disp.php?uid=107, 2007.
- [42] D. Manning, S. Ethell, T. Donovan, and T. Crawford, "How Do Radiologists Do It? The Influence of Experience and Training on Searching for Chest Nodules," Radiography, vol. 12, no. 2, pp. 134-142, 2006.
- [43] T. Munzner, F. Guimbretière, S. Tasiran, L. Zhang, and Y. Zhou, "TreeJuxtaposer: Scalable Tree Comparison using Focus+Context with Guaranteed Visibility," ACM Trans. on Graphics (Proc. SIGGRAPH 2003), vol. 22, no. 3, pp. 453-462, 2003.
- [44] C.F. Nodine, H.L. Kundel, C. Mello-Thoms, S.P. Weinstein, S.G. Orel, D.C. Sullivan, and E.F. Conant, "How Experience and Training Influence Mammography Expertise," Academic Radiology, vol. 6, no. 10, pp. 575-585, 1999.
- [45] C.F. Nodine, C. Mello-Thoms, S.P. Weinstein, H.L. Kundel, E.F. Conant, R.E. Heller-Savoy, S.E. Rowlings, and J.A. Birnbaum, "Blinded Review of Retrospectively Visible Unreported Breast Cancers: An Eye-position Analysis," Radiology, vol. 221, no. 1, pp. 122-129, 2001.
- [46] A. Perer, and B. Shneiderman, "Balancing Systematic and Flexible Exploration of Social Networks," IEEE Trans. on Visualization and Computer Graphics, vol. 12, no. 5, pp. 693-700, 2006.
- [47] K. Perlin, and D. Fox, "Pad: An Alternative Approach to the Computer Interface," Proc. ACM SIGGRAPH Conf. Computer Graphics and Interactive

- Techniques (SIGGRAPH '93), pp. 57-64, 1993.
- [48] T. Pfeiffer, "Measuring and Visualizing Attention in Space with 3D Attention Volumes," Proc. ACM Symp. Eye Tracking Research and Applications (ETRA '12), pp. 29-36, 2012.
- [49] P. Phillips, D. Manning, T. Donovan, T. Crawford, and S. Higham, "A Software Framework for Diagnostic Medical Image Perception with Feedback, and a Novel Perception Visualization Technique," Proc. SPIE-Medical Imaging 2005, vol. 5749, pp. 572-580, 2005.
- [50] P. Phillips, D. Manning, T. Crawford, D. Burling, C.-L. Tam, and A. Taylor, "Searching in Axial and 3D CT Visualizations," Proc. SPIE, vol. 6917, pp. 69171J-69171J -8, 2008.
- [51] P. Phillips, D. Boone, S. Mallett, S.A. Taylor, D.G. Altman, D. Manning, A. Gale, and S. Halligan, "Method for Tracking Eye Gaze during Interpretation of Endoluminal 3D CT Colonography: Technical Description and Proposed Metrics for Analysis," Radiology, vol. 267, no. 3, pp. 924-931, 2013.
- [52] H. Piringer, R. Kosara, and H. Hauser, "Interactive Focus+Context Visualization with Linked 2D/3D Scatterplots," Proc. IEEE Conf. Coordinated and Multiple Views in Exploratory Visualization (CMV '04), pp. 49-60, 2004.
- [53] K.-J. R  ih  , A. Aula, P. Majaranta, H. Rantala, and K. Koivunen, "Static Visualization of Temporal Eye-tracking Data," Human-Computer Interaction – Interact 2005, M.F. Costabile and F. Patern  , eds., Springer Berlin Heidelberg, pp. 946-949, 2005.
- [54] R. Ramloll, C. Trepagnier, M. Sebrechts, and J. Beedasy, "Gaze Data Visualization Tools: Opportunities and Challenges," Proc. IEEE Conf. Information Visualisation (IV '04), pp. 173-180, 2004.
- [55] W.M. Reed, J.T. Ryan, M.F. McEntee, M.G. Evanoff, and P.C. Brennan, "The

- Effect of Abnormality-prevalence Expectation on Expert Observer Performance and Visual Search,” *Radiology*, vol. 258, no. 3, pp. 938-943, 2011.
- [56] K. Ridsen, M.P. Czerwinski, T. Munzner, and D.B. Cook, “An Initial Examination of Ease of Use for 2D and 3D Information Visualizations of Web Content,” *International Journal of Human-Computer Studies*, vol. 53, no. 5, pp. 695-714, 2000.
- [57] G.G. Robertson, J.D. Mackinlay, and S.K. Card, “Cone Trees: Animated 3D Visualizations of Hierarchical Information,” *Proc. ACM SIGCHI Conf. Human Factors in Computing Systems (CHI '91)*, pp. 189-194, 1991.
- [58] G.G. Robertson, M. Czerwinski, K. Larson, D.C. Robbins, D. Thiel, and M. van Dantzich, “Data Mountain: Using Spatial Memory for Document Management,” *Proc. ACM Symp. User Interface Software and Technology (UIST '98)*, pp. 153-162, 1998.
- [59] D.D. Salvucci, and J.H. Goldberg, “Identifying Fixations and Saccades in Eye-tracking Protocols,” *Proc. ACM Symp. Eye Tracking Research and Applications (ETRA '00)*, pp. 71-78, 2000.
- [60] J. Seo, and B. Shneiderman, “A Rank-by-Feature Framework for Interactive Exploration of Multidimensional Data,” *Information Visualization*, vol. 4, no. 2, pp. 99-113, 2005.
- [61] J. Seo, and B. Shneiderman, “Knowledge Discovery in High-dimensional Data: Case Studies and a User Survey for the Rank-by-Feature Framework,” *IEEE Trans. on Visualization and Computer Graphics*, vol. 12, no. 3, pp. 311-322, 2006.
- [62] B. Shneiderman, “The Eyes Have It: A Task by Data Type Taxonomy for Information Visualizations,” *Proc. IEEE Symp. Visual Languages (VL '96)*,

pp. 336-343, 1996.

- [63] B. Shneiderman, "Why not Make Interfaces Better than 3D Reality," *IEEE Computer Graphics and Applications*, vol. 23, no. 6, pp. 12-15, 2003.
- [64] B. Shneiderman, and C. Plaisant, "Strategies for Evaluating Information Visualization Tools: Multi-dimensional In-depth, Long-term Case Studies," *Proc. AVI Workshop BEyond Time and Errors: Novel Evaluation Methods for Information Visualization (BELIV '06)*, pp. 38-43, 2006.
- [65] H. Song, J. Yun, B. Kim, and J. Seo, "GazeVis: Interactive 3D Gaze Visualization for Contiguous Cross-Sectional Medical Images," *IEEE Transactions on Visualization and Computer Graphics*, vol. 20, no. 5, pp. 726-739, 2014.
- [66] S. Stellmach, L. Nacke, and R. Dachsel, "3D Attentional Maps: Aggregated Gaze Visualizations in Three-dimensional Virtual Environments," *Proc. ACM Conf. Advanced Visual Interfaces (AVI '10)*, pp. 345-348, 2010.
- [67] S. Stellmach, L. Nacke, and R. Dachsel, "Advanced Gaze Visualizations for Three-dimensional Virtual Environments," *Proc. ACM Symp. Eye Tracking Research and Applications (ETRA '10)*, pp. 109-112, 2010.
- [68] A. Sutcliffe and U. Patel, "3D or Not 3D: Is It Nobler in the Mind?," *People and Computers: Proc. British Human-Computer Interaction Conference*, pp. 79-94, 1996.
- [69] O. Špakov, and D. Miniotas, "Visualization of Eye Gaze Data using Heat maps," *Electronics and Electrical Engineering*, vol. 2, no. 74, pp. 55-58, 2007.
- [70] M. Tavanti, and M. Lind, "2D vs 3D, Implications on Spatial Memory," *Proc. IEEE Symp. Information Visualization (INFOVIS '01)*, pp. 139-145, 2001.
- [71] A.C. Telea, *Data Visualization: Principles and Practice*. A K Peters/CRC Press, 2007.

- [72] Three-dimensional Industrial Computed Tomography X-ray Inspection, <http://xrayinspectionsservice.com/3d-services/>.
- [73] Tobii Eye Tracking Analytics SDK, <http://www.tobii.com/en/eye-tracking-research/global/products/software/tobii-analytics-software-development-kit/>.
- [74] C. Tominski, H. Schumann, G. Andrienko, and N. Andrienko, "Stacking-Based Visualization of Trajectory Attribute Data," *IEEE Trans. on Visualization and Computer Graphics*, vol. 18, no. 12, pp. 2565-2574, 2012.
- [75] M. Tory, and T. Möller, "Human Factors in Visualization Research," *IEEE Trans. on Visualization and Computer Graphics*, vol. 10, no. 1, pp. 72-84, 2004.
- [76] M. Tory, A.E. Kirkpatrick, M.S. Atkins, and T. Möller, "Visualization Task Performance with 2D, 3D, and Combination Displays," *IEEE Trans. on Visualization and Computer Graphics*, vol. 12, no. 1, pp. 2-13, 2006.
- [77] H.Y. Tsang, M. Tory, and C. Swindells, "eSeeTrack—Visualizing Sequential Fixation Patterns," *IEEE Trans. on Visualization and Computer Graphics*, vol. 16, no. 6, pp. 953-962, 2010.
- [78] S.A. Vitak, J.E. Ingram, A.T. Duchowski, S. Ellis, and A.K. Gramopadhye, "Gaze-Augmented Think-Aloud as an Aid to Learning," *Proc. ACM SIGCHI Conf. Human factors in Computing Systems (SIGCHI '12)*, pp. 2991-3000, 2012.
- [79] Z. Wang, A.C. Bovik, H.R. Sheikh, and E.P. Simoncelli, "Image Quality Assessment: From Error Visibility to Structural Similarity," *IEEE Trans. on Image Processing*, vol. 13, no. 4, pp. 600-612, 2004.
- [80] D.S. Wooding, "Fixation Maps: Quantifying Eye-movement Traces," *Proc. ACM Symp. Eye Tracking Research and Applications (ETRA '02)*, pp. 31-36, 2002.

- [81] S. Zhai, W. Buxton, and P. Milgram, "The Partial-occlusion Effect: Utilizing Semitransparency in 3D Human-computer Interaction," *ACM Trans. on Computer-Human Interaction*, vol. 3, no. 3, pp. 254-284, 1996.

요약

본 연구에서는 3차원 의료 영상 판독 시선 정보에 대한 대화형 시각적 분석 프레임워크를 제안한다. 본 프레임워크는 복수의 사용자로부터 획득한 시선 정보를 비교 분석하기 위한 목적으로 디자인되었으며, 특히 추가적인 맥락적 정보를 포함한 3차원 시선 정보를 위해 특화되어 있다. 시선 정보의 비교는 방사선 전문의가 어떻게 의료 영상을 판독하는지 이해하고 이에 영향을 미치는 요인들을 확인하는데 있어 필수적이다. 그러나 CT나 MRI로부터 획득한 3차원 영상에 대해 진행되었던 선행 연구들은 비교적 분석에 있어서 몇 가지 한계점을 내포하고 있다. 판독과정에서 방사선 전문의들은 수직 방향으로 적층된 영상들을 스크롤을 통해 왔다 갔다 하면서 장기와 병변에 대한 3차원적인 인지를 하게 되기에, 이 과정에서 수집되는 시선 패턴은 일반적인 2차원 영상을 사용하는 실험과 비교했을 때 추가적인 깊이 정보를 갖는다. 이에 따라 추가적인 공간 정보로 인하여 시선 정보의 가시화 복잡도는 보다 높아지게 된다. 최근 연구에서는 direct volume rendering (DVR) 기법을 활용한 시각화 디자인을 제안했으나, 이 역시 효율적이고 종합적인 시선 패턴 비교에 있어서는 비교 목적의 대화형 시각화 도구의 부재로 인해 한계점을 나타냈다.

이에 본 논문에서는 3차원 시선 정보에 대한 효율적 가시화 기법을 제시하고, 이를 활용한 대화형 분석 프레임워크를 제안한다. 이와 함께 주요한 맥락적 정보인 동공의 크기나 windowing (영상의 밝기와 대비의 조정) 과정에 대한 정보를 분석 과정에 포함시킴으로써 보다 심도 있고 생태적으로 올바른 발견을 지원하고자 한다. 이를 위한 대화형 시각화 요소 중에서 특히 context-embedded interactive scatterplot (CIS)의 경우 사용자가 추상적 시선 정보를 분석하는데 있어서 방사선 전문의들에

게 익숙한 의학적인 시각화 방법을 내포함으로써 보다 용이하게 진행할 수 있도록 했다. 아울러 본 논문에서는 해당 기법들에 대한 흥부 및 복부 전문의들과의 사례 연구 결과를 보고한다.

주요어 : 시선 추적, 시선 정보 가시화, 시선 패턴 비교, 3 차원 의료 영상, 맥락 내포형 산점도 (Context-embedded scatterplot), 대화형 시간 차트 (Interactive temporal chart)

학 번 : 2011-30964