

# GazeVis: Interactive 3D Gaze Visualization for Contiguous Cross-sectional Medical Images

Hyunjoo Song, Jihye Yun, Bohyoung Kim, *Member, IEEE*, and Jinwook Seo, *Member, IEEE*

**Abstract**—Gaze visualization has been used to understand the results from gaze tracking studies in a wide range of fields. In the medical field, diagnoses of medical images have been studied with gaze tracking technology to understand how radiologists read medical images. While prior work were mainly based on diagnosis with a single image, recent work focused on diagnosis with consecutive cross-sectional medical images acquired from preoperative computed tomography (CT) or magnetic resonance imaging (MRI). In the diagnosis, radiologists scroll through a stack of images to get a 3D cognition of organs and lesions. Thus, it is important to understand radiologists' gaze patterns three dimensionally across such contiguous cross-sectional images. However, little has been done to visualize more complicated gaze patterns from the contiguous cross-sectional medical images. To address this problem, we present an interactive 3D gaze visualization tool, GazeVis, where InfoVis and SciVis techniques are harmonized to show the abstract gaze data along with a realistic 3D rendering of the visual stimuli (i.e. organs and lesions). We present case studies with 12 radiologists who use GazeVis to investigate gaze patterns of their colleagues with different levels of expertise, providing empirical evidences about the competence of our gaze visualization system.

**Index Terms**—Eye tracking, gaze visualization, volume rendering, medical images, interaction technique



## 1 INTRODUCTION

Gaze tracking has been used in many research fields such as market research, psychology research, vision research, and user research. Visualization plays an import role in analyzing the gaze data to understand human mind behind gaze patterns. There have been efforts to develop effective visualization techniques for better understanding of the gaze data [9], [27], [37], [38]. For instance, superimposition of gaze data over visual stimuli has been shown to be effective in interpreting the data [38]. The superimposition technique has been adopted in most gaze data analysis systems.

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. Especially, when the visual stimuli consist of a large

number of images, intuitive interaction techniques such as dynamic queries are also necessary to support efficient exploration of the whole gaze data. Gaze analysis with radiologists who read medical images such as CT and 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.

Understanding how radiologists read the images is a key step toward preventing diagnosis errors and training novice radiologists better. There have been many studies in the radiology field that adopted the gaze tracking technology to gain a deeper understanding of how radiologists perform a diagnosis [14], [15], [21], [22], [23], [24], [28]. An eye tracking study with mammography images helps researchers understand when and why radiologists fail to notice abnormal lesions [22]. Kundel and Follette compared gaze patterns of experts and novices in reading a static 2D chest x-ray scan [14]. These and most prior work are done with a single 2D medical image, but Phillips *et al.* presented a gaze data visualization system for a small number of consecutive 2D cross-sectional images composing a 3D volume [23]. Radiologists read the images one by one quickly scrolling through them. The system simply stacks 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 (Fig. 1). This approach suffers 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.

- Hyunjoo Song, Jihye Yun, and Jinwook Seo are with the Department of Computer Science and Engineering, Seoul National University, Seoul, Korea. E-mail: hjsong@hcil.snu.ac.kr, jhyun@cglab.snu.ac.kr, and jseo@snu.ac.kr.
- Bohyoung Kim is with the Department of Radiology, Seoul National University Bundang Hospital, Gyeonggi-do, Korea. E-mail: bhkim@snuhb.org.

Please note that all acknowledgments should be placed at the end of the paper, before the bibliography (note that corresponding authorship is not noted in affiliation box, but in acknowledgment section).

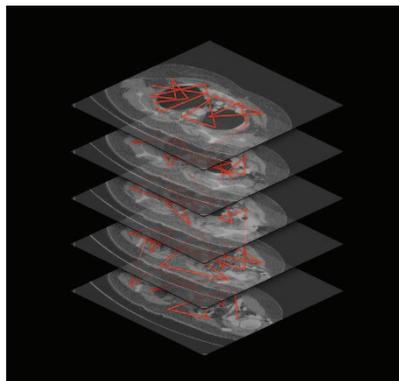


Fig. 1. Gaze plot superimposed on stacked 2D cross-sectional images.

Previous user studies showed that 3D interfaces/visualizations could suffer from occlusions, complicated user controls, and disorienting navigation, which lead to the degradation of users' task performance [5], [29], [39]. When designing a 3D gaze data visualization system for consecutive 2D cross-sectional images composing a 3D volume, it is relevant to take those three problems of 3D interfaces or visualizations into account as discussed in [2]. The disorientating navigation problem can be resolved to some extent when gaze points are visualized in a realistic volume rendering of stimuli. The realistic volume rendering can help users' holistic perception of the human body, which is a strong contextual landmark for navigation. The second problem of complicated user controls can also be addressed if interaction designs are based on familiar user interactions in the field. The occlusion problem can be also alleviated to some extent through intuitive manipulation of semitransparency of body parts in stimuli and gaze points in gaze data.

We believe that a better approach to the 3D gaze data visualization should come from a close collaboration between two branches of visualization: information visualization (InfoVis) and scientific visualization (SciVis). Volume rendering techniques developed in the SciVis community should be 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 also required 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. Traditional volume rendering should be extended to visualize the abstract information, i.e. gaze points within the volume while delivering the depth perception of gaze points. Dynamic queries should also be designed to support efficient temporal and spatial filtering of gaze points within the context of the volumetric stimuli.

In this paper, we present a novel interactive 3D gaze visualization system, GazeVis, for consecutive 2D cross-sectional medical images. We first explore the design space of visualizations for interactively analyzing the gaze tracking data with such medical images, and find a need for integrating InfoVis and SciVis techniques. Using

GazeVis, we also perform two case studies with 12 radiologists following the MILCs evaluation guidelines [33], [36]. Based on the case study results, we summarize pros and cons of feasible representation techniques of GazeVis. It is followed by discussion and future work.

## 2 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 information of human body (Fig. 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.

## 3 RELATED WORK

This work is a result of close collaboration between InfoVis researchers and SciVis researchers and thus related to various previous work from both communities. We summarize relevant prior work in four most related areas: 3D information visualization, real-time volume rendering, gaze visualization in general, and gaze visualization in the medical field.

### 3.1 Visualization combining 2D and 3D

Early works in the InfoVis field explored pros and cons of 3D visualization of abstract dataset [11], [19], [30], [31], 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 [45]. Cockburn and McKenzie compared the original 3D Data Mountain with a simple 2D redesign of it and found no significant difference in performing organization/retrieval tasks [4]. In another study, they even found that 3D showed lower performance than 2D in terms of task completion time for spatial memory tasks [5]. Cockburn later revisited the spatial memory issue [6] in re-running Tavanti *et al.*'s experiment [41] with some uncontrolled factors controlled, and found no performance difference in spatial memory tasks, which refuted the original study. Tory and Möller pointed out that interaction with 3D visualizations is difficult because of

unnatural mappings between 2D mouse actions and 3D space [44]. 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.* 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 [25]. Elmqvist *et al.* used animated 3D rotations during transitions between 2D scatterplots for multidimensional data [10]. 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.* 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 [45]. 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 suggested a number of design guidelines for a better 3D interface [35]. We followed these guidelines in designing our interface. Zhai *et al.* reported perceptual advantage of using partial-occlusion in a 3D space [48]. 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 GazeVis 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.

### 3.2 Volume Rendering

Direct volume rendering (DVR) is one of methods visualizing 3D volumetric data from CT or MRI scanners [8], [18], [42]. It classifies the volume data using a transfer function to map a scalar value to optical properties such as color and opacity. Among various DVR techniques, we adopt ray-casting DVR which is known for its high-quality rendering [17].

### 3.3 Gaze 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 [9], 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. Wooding introduced a heatmap-like visualization, or fixation map built by applying a 3D Gaussian filter at each fixation location [47] to reduce clutter and enable detailed quantitative comparisons between different gaze patterns of multiple users. Špakov *et al.* tried to enhance the heatmap visualization by making the transparency of heatmaps adjusta-

ble [40].

Some gaze plot based visualizations put more emphasis on the temporal order of gazes. Lankford 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 [16]. Rähkä *et al.* proposed the time plot visualization to show the temporal order of visits to area-of-interests [26] while sacrificing the exact fixation locations. Goldberg *et al.* extended the gaze plot (or visual scanpath representation) using time expansions, small multiples, and radial plots; and classified scanning strategies into nine categories [12].

Researchers also have proposed gaze visualization techniques for dynamic stimuli. For example, Tsang *et al.* introduced eSeeTrack to support comparison of fixation patterns on dynamic stimuli [46]. 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. For example, Stellmach *et al.* introduced 3D scan paths, 3D attentional maps, and models of interest timeline view for 3D virtual environments [37], [38]. 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 visualization system, GazeVis 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 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. We will further elaborate on this issue in section 4.2.

### 3.4 Gaze Analysis in Medical Field

In the radiology field, gaze tracking techniques have been widely used with static 2D medical images. Most gaze analysis studies in this field used conventional gaze plot and heatmap. Kundel and Follette compared visual search patterns between experts and novices during diagnosis of a chest radiographic image [14]. Reed *et al.* investigated the effect of prior knowledge in reading chest radiographic images [28]. Mammographic images were also used by a number of researchers. Nodine *et al.* conducted an eye tracking study to determine whether unreported breast cancers during the diagnosis received sufficient visual attention [22]. Kundel *et al.* found that experts read a mammographic image with a holistic perception [15]. All these work focused on analysis of a single static image at a time and thus their approaches are not directly applicable to analyzing contiguous cross-sectional images as a whole.

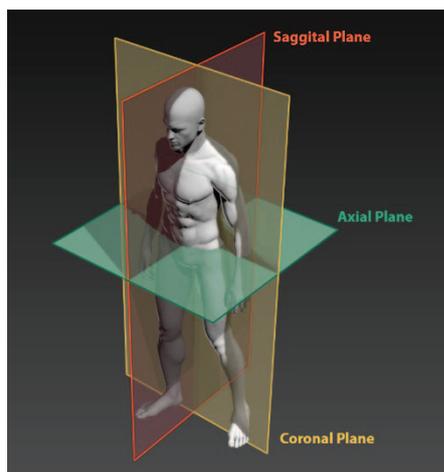


Fig. 2. Multi-planar reconstruction (MPR) planes of human body. 3D human model is from [20].

Most relevant to this work is Phillips *et al.*'s study with a series of cross-sectional brain MRI images [23]. They showed the gaze data in two different views. The first view accumulated all fixations throughout whole images on a 2D view. When one scrolls through the images like as in a stack viewing mode, fixations associated with the current image were highlighted in a different color. Since this view discards depth information when showing fixations, it is difficult for users to perceive the 3D nature of the gaze data in this view. Another view showed a 3D visualization of the images stack and superimposed gaze plots on it as shown in Fig. 1. It showed the gaze data on a 3D space, but this approach suffers from occlusion and visual clutter. Later, Phillips *et al.* compared the traditional stack viewing mode and a virtual colonoscopy mode, using CT colonography images [24]. They made a virtual colonoscopy video with a virtually constructed 3D scene using the cross-sectional images. Then they asked participants to watch the video, and collected gaze data from them. Using the two kinds of viewing modes, they compared fixation characteristics taking expertise of participants into account.

In this work, we propose intuitive representations of gaze data using volume rendering techniques while alleviating innate problems of conventional 3D visualization by adopting enhanced techniques: transparency adjustment, interactive coordination between 2D and 3D views, and rich user interactions such as spatial and temporal dynamic queries.

## 4 GAZEVIS

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. Since the size of gaze data gets much bigger with a series of images, we amplify cognition by adopting the visual information-seeking mantra [34]: overview first, zoom and filter, then details on 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 (Fig. 3). 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 (Fig. 2). 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 5.

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 5.5). Below the widgets, there is a range slider for temporal filtering (explained later in section 5.3). 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 5.4).

### 4.1 Visualization of Stimuli Volume Data

We adopted common visualization methods for stimuli volume data (i.e. contiguous cross-sectional medical images) - DVR and MPR 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: which parts are visualized? which parts are transparent? and which parts have which color. 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 [48]. 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 (Fig. 2). 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 (Fig. 3). In GazeVis, we implemented MPR views as detail views for two reasons: (1) to provide

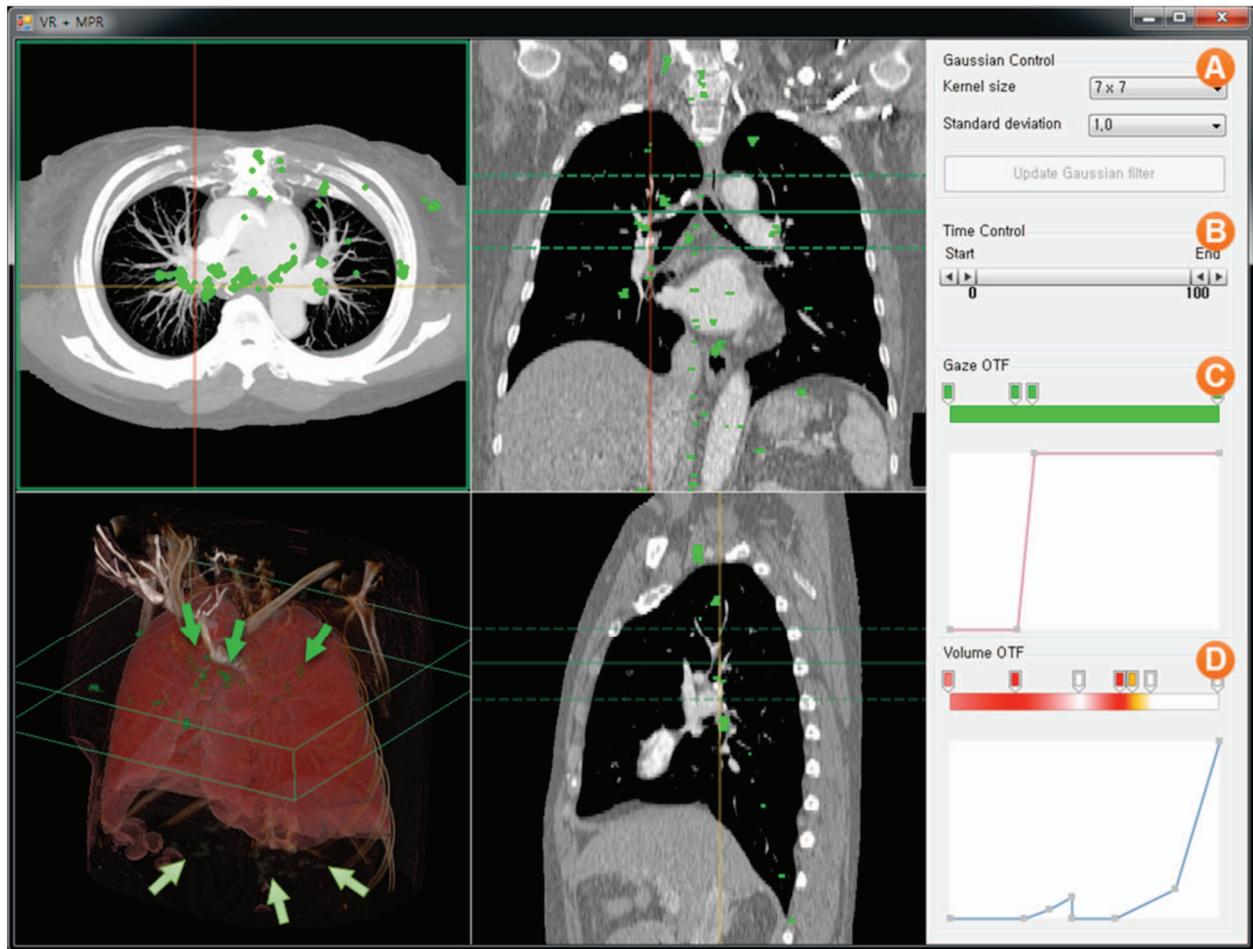


Fig. 3. GazeVis interface with 4 views on the left and a control panel on the right. Bright green dots represent gaze points. 4 views are axial, coronal, sagittal, and 3D views, clockwise from the upper left corner. Axial, coronal, and sagittal views have corresponding cross hairs and rectangular borders 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).

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.

#### 4.2 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.

In this paper, we propose a representation method for the gaze data, referred to gaze field, which stores the gaze data in a scalar field whose value, at a position, represents the magnitude of fixation duration at the position. The gaze field is rendered three dimensionally along with cross-sectional medical images, delivering the accurate depth order and thus more enhanced depth cue.

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 [26]. 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 [32], 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 fixa-

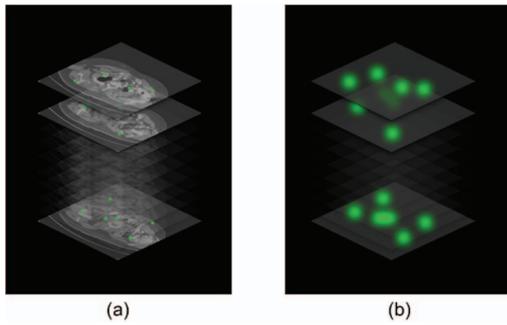


Fig. 4. (a) Gaze points on stimuli volume data (i.e. contiguous medical images). (b) Gaze field with the same  $x$ -,  $y$ -, and  $z$ -resolutions as the stimuli volume data.

tion. 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 (eq.1) to each gaze point:

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

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 (Fig. 4). 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) [3]. 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.

### 4.3 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 Fig. 3 and Fig. 6). 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.

### 4.4 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 [38], [46]. 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.

## 5 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.

### 5.1 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 (Fig. 3).

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 sag-

ittal 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 (Fig. 3).

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.

### 5.2 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 (Fig. 3). 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 (Fig. 3).

### 5.3 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,

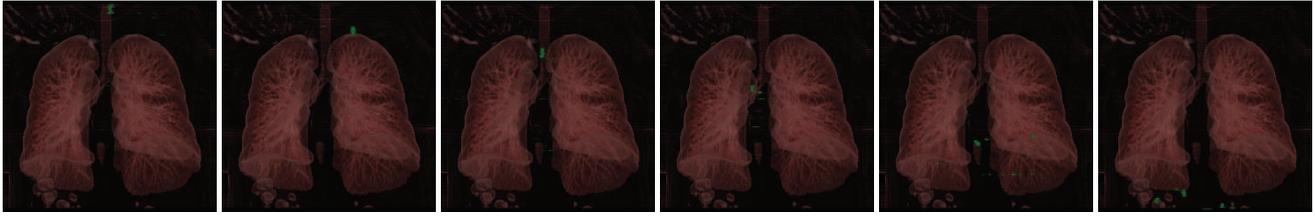


Fig. 5. Temporal gaze flow of a radiologist in reading a chest CT scan, revealing that he scrolled through the images from top to bottom.

MPR and 3D views are 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 4.4), it is straightforward 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 (Fig. 5).

#### 5.4 Transfer Function Control

The transfer function adjusts the color and opacity of anatomic regions and gaze data in the 3D view (Fig. 6). We provided a UI widget to allow users to interactively manipulate the transfer function ((C) and (D) in Fig. 3). As shown in (C) in Fig. 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 opaci-

ty 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.

#### 5.5 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 (Fig. 7) 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.

### 6 IMPLEMENTATION

GazeVis is implemented with C#, C++, and WPF. We used Tobii analytics SDK [43] 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.

### 7 EVALUATION OF GAZEVIS WITH RADIOLOGISTS

We adopted a case study-based evaluation method for information visualization, i.e. Multi-dimensional In-depth Long-term Case studies (MILCs) [33], [36] to evaluate the

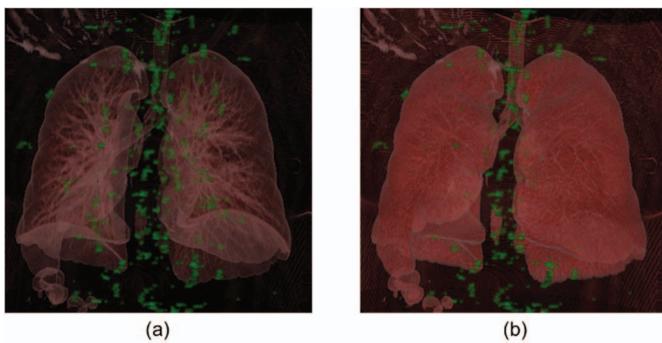


Fig. 6. 3D rendering of lung with three-dimensionally superimposed gaze data. (a) With low opacity for the lung, gaze data within the lung are clearly visualized. (b) With high opacity for the lung, the gaze data within it are partially occluded.

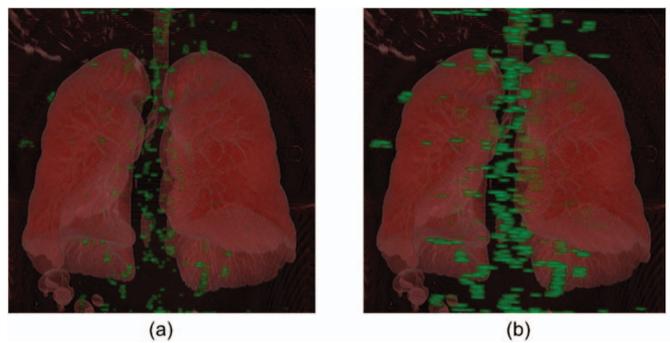


Fig. 7. Adjustments of Gaussian blur filter size. (a) 7 X 7 with standard deviation of 1. (b) 19 X 19 with standard deviation of 3.

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 [23] 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 [36] 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.

### 7.1 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 [33]. 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

TABLE 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

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.

### 7.2 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 1.

### 7.3 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.

### 7.4 Chest Radiologists

Following the case study protocol laid out in section 7.1, we performed the first case study with 6 chest radiolo-

gists. 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 Fig. 8(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 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 Fig. 8(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

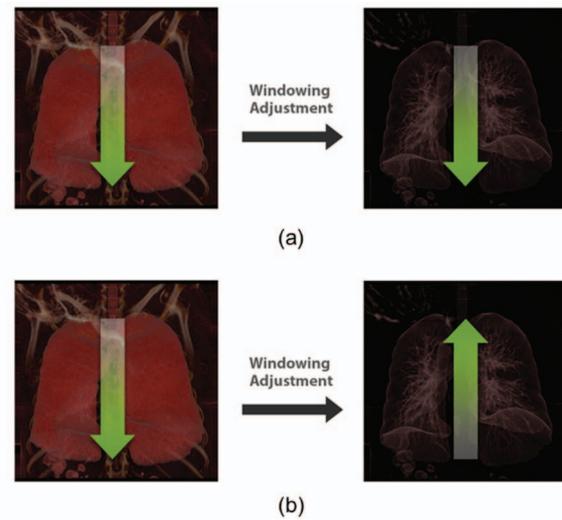


Fig. 8. Gaze pattern difference of two experts in reading a chest CT scan. (a) Expert 1: Scroll down from top to bottom, jump back to the top, adjust the window setting, and scroll down again. (b) Expert 2: Scroll down from top to bottom, adjust the window setting, and scroll up from bottom to top.

gaze pattern. While Reed *et al.* performed a similar study with chest x-ray images [28], he mentioned that GazeVis can help researchers investigate the gaze pattern with cross-sectional medical images.

### 7.5 Abdominal Radiologists

We conducted another case study with abdomen CT images while also following the protocol described in section 7.1. 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 Fig. 9(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 (Fig. 9(c)). In case of intermediate radiologists, they showed gaze patterns somewhat in between the expert and the novices (Fig. 9(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

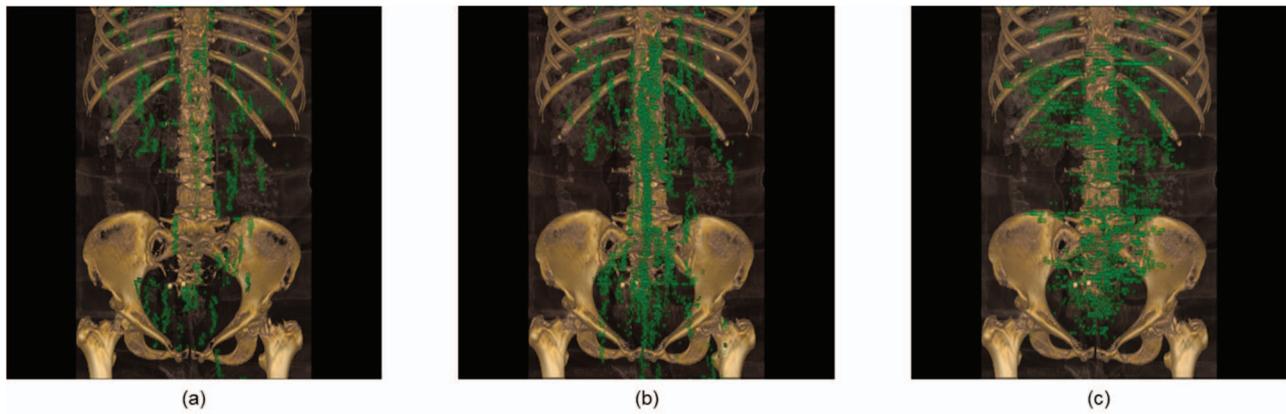


Fig. 9. Gaze pattern difference between radiologists with different expertise level. (a) Expert showing more organized pattern vertically. (b) Intermediate showing a pattern in between expert and novice. (c) Novice showing no distinctive pattern with scattered gaze points.

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.

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 8.1.

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.

## 8 DISCUSSION AND FUTURE WORK

In this paper, 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.

### 8.1 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 important in the diagnostic reading, are more clearly marked (Fig. 10). Such combination of the gaze field as an attribute data can be achieved by adopting the 2D transfer function approach [7], [13].

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.

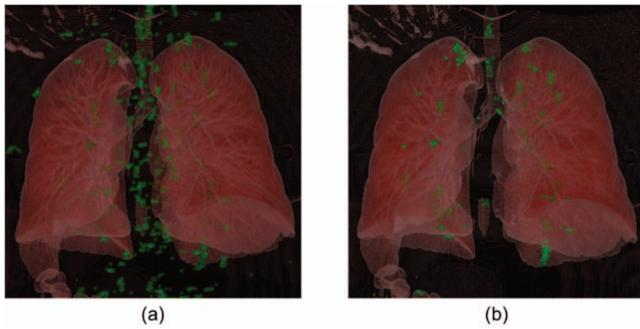


Fig. 10. (a) 3D rendering of lung with gaze points superimposed as independent objects. (b) 3D rendering of lung with gaze data used as attribute data, enabling gaze analyses more focused on human anatomy.

### 8.2 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 Fig. 11, 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 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.

### 8.3 Improving Ecological Validity

GazeVis can be made more ecologically practical in several directions. First, instead of capturing gaze data in cases where a radiologist examines only a single image stack on a 2D view, it could support more complicated cases. In practice, radiologists often use multiple views (i.e. various combinations of 2D + 3D views) to review multiple series of image stacks together (e.g. multiphased cardiac CT scans). Gaze pattern analyses for those

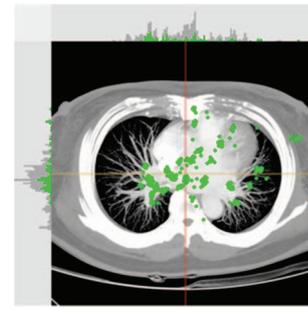


Fig. 11. Gaze data distribution in an axial plane. Gray histogram represents the distribution of all gaze points on entire slices. Green one represents the distribution of the gaze points only on the current axial image.

cases will pose many interesting challenges in visualization and interaction design, and the effort to meet those challenges will make gaze visualization tools more practically relevant.

In the case studies, the two most experienced expert radiologists, who analyzed the gaze patterns of each group of radiologists later, emphasized the importance of peripheral vision. They attributed the difference in gaze pattern between radiologists with different expertise level (i.e. organized pattern for experts and scattered pattern for novices) to the peripheral vision. They explained that the experts maintain their focus at the center of the area of interest in order to use their peripheral vision actively and that the novices who are not so experienced in reading the images using their peripheral vision should take a busy look at every position in the image not to miss any region.

Considering such importance of the peripheral vision in radiological reading, it might be meaningful to visualize the areas covered by the peripheral vision surrounding the foveal fixations. Radiologists exhibit individually different level of spatial sensitivity of the peripheral vision. Therefore, for more accurate comparison of gaze patterns between observers, it is necessary to take into account the individual difference in the spatial sensitivity of each observer's peripheral vision.

## 9 CONCLUSION

We presented GazeVis, an interactive 3D gaze visualization tool, for contiguous cross-sectional medical images that compose a 3D volume. The tight coupling of 2D and 3D views enabled users to grasp overall 3D gaze patterns along with the detailed spatial context constructed from the stimuli images. We also introduced the gaze field as an efficient representation of gaze data to support dynamic spatial and temporal filtering of gaze data. Two case studies with 12 chest and abdominal radiologists revealed that differences in expertise level and preferred diagnosis strategy of radiologists led to significant differences in 3D gaze patterns.

### ACKNOWLEDGMENT

We thank Kyoung Ho Lee, Tae Jung Kim, and other radiologists in SNU Bundang Hospital for their active partici-

pation in our case studies. This work was supported by the National Research Foundation of Korea (NRF) grants funded by the Korea government (MSIP) (No. 2011-0030813, NRF-2013R1A1A3006706), by the grant No. 02-2013-055 from the SNUBH Research Fund, and by the Interdisciplinary Research Initiatives Program funded by College of Engineering and College of Medicine, Seoul National University (2013). The ICT at Seoul National University provided research facilities for this study. Jinwook Seo and Bohyoung Kim are the corresponding authors.

## REFERENCES

- [1] 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.
- [2] 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.
- [3] 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.
- [4] 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.
- [5] 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.
- [6] A. Cockburn, "Revisiting 2D vs 3D Implications on Spatial Memory," *Proc. Conf. Australasian User Interface (AUIC '04)*, pp. 25-31, 2004.
- [7] 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.
- [8] 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.
- [9] 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.
- [10] N. Elmqvist, P. Dragicevic, and J.-D. Fekete, "Rolling the Dice: Multi-dimensional Visual Exploration using Scatterplot Matrix Navigation," *IEEE Trans. on Visualization and Computer Graphics*, vol. 14, no. 6, pp. 1141-1148, 2008.
- [11] 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.
- [12] J.H. Goldberg, and J.I. Helfman, "Visual Scanpath Representation," *Proc. ACM. Symp. Eye Tracking Research and Applications (ETRA '10)*, pp. 203-210, 2010.
- [13] 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.
- [14] 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.
- [15] 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.
- [16] C. Lankford, "Gazetracker: Software Designed to Facilitate Eye Movement Analysis," *Proc. ACM Symp. Eye Tracking Research and Applications (ETRA '00)*, pp. 51-55, 2000.
- [17] 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.
- [18] M. Levoy, "Display of Surfaces from Volume Data," *IEEE Computer Graphics and Applications*, vol. 8, no. 3, pp. 29-37, 1988.
- [19] 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.
- [20] Male Figure Realistic Standing Human, [http://artist-3d.com/free\\_3d\\_models/dnm/model\\_disp.php?uid=107](http://artist-3d.com/free_3d_models/dnm/model_disp.php?uid=107), 2007.
- [21] 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.
- [22] 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.
- [23] 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.
- [24] 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. 6917J-6917J-8, 2008.
- [25] 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.
- [26] K.-J. Rähkä, A. Aula, P. Majoranta, H. Rantala, and K. Koivunen, "Static Visualization of Temporal Eye-tracking Data," *Human-Computer Interaction - Interact 2005*, M.F. Costabile and F. Paternò, eds., Springer Berlin Heidelberg, pp. 946-949, 2005.
- [27] 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.
- [28] 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.
- [29] K. Ridsden, 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.
- [30] 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.
- [31] 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.
- [32] 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.
- [33] 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.
- [34] 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.
- [35] B. Shneiderman, "Why not Make Interfaces Better than 3D Reality," *IEEE Computer Graphics and Applications*, vol. 23, no. 6, pp. 12-15, 2003.
- [36] 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.
- [37] 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.
- [38] 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.
- [39] 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.
- [40] 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.
- [41] M. Tavanti, and M. Lind, "2D vs 3D, Implications on Spatial Memory," *Proc. IEEE Symp. Information Visualization (INFOVIS '01)*, pp. 139-145, 2001.
- [42] A.C. Telea, *Data Visualization: Principles and Practice*. A K Peters/CRC Press, 2007.
- [43] Tobii Eye Tracking Analytics SDK, <http://www.tobii.com/en/eye-tracking-research/global/products/software/tobii-analytics-software-development-kit/>.
- [44] 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.
- [45] 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.
- [46] H. 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.
- [47] D.S. Wooding, "Fixation Maps: Quantifying Eye-movement Traces," *Proc. ACM Symp. Eye Tracking Research and Applications (ETRA '02)*, pp. 31-36, 2002.
- [48] 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.



Hyunjoon Song received the B.S. degree in computer science and engineering, and the M.S. degree in electrical engineering and computer science from Seoul National University, Seoul, Korea, in 2009, and 2011, respectively. He is currently working toward the Ph.D. degree in the Department of Computer Science and Engineering, Seoul National University, Seoul, Korea. His research interests include HCI, gaze tracking, and information visualization.



Jihye Yun received the B.S. degree in media engineering from Soongsil University, Seoul, Korea, and the M.S. degree in computer science and engineering from Seoul National University, Seoul, Korea in 2005 and 2011, respectively. She is currently working toward the Ph.D. degree in the Department of Computer Science and Engineering, Seoul National University, Seoul, Korea. Her current research interests include computer graphics, volume visualization, and medical imaging.



Bohyoung Kim received the B.S. and M.S. degrees in computer science from Seoul National University, Seoul, Korea, and the Ph.D. degree in computer science and engineering from the same university in 2001. She is currently an Assistant Professor in the Department of Radiology, Seoul National University Bundang Hospital, Korea. Her research interests include computer graphics, volume visualization, medical imaging, and information visualization.



Jinwook Seo received the B.S. and M.S. degrees in computer science from Seoul National University, Seoul, Korea, and the Ph.D. degree in computer science from the University of Maryland, College Park, MD, in 1995, 1997, and 2005, respectively. He is currently an Associate Professor in the Department of Computer Science and Engineering, Seoul National University, where he is also the Director of the Human-Computer Interaction Laboratory. His research interests include HCI, information visualization, and biomedical informatics.