Temporal Methods for Eye Movement Analysis

Walter F. Bischof, Nicola C. Anderson and Alan Kingstone

Department of Psychology, University of British Columbia, 2136 West Mall, Vancouver, BC, V6T 1Z4, Canada

Table of Contents

1. Summary and Learning Objectives

This chapter is aimed at introducing the reader to current methods for the spatial and temporal analysis of eye movements. There are four main parts. In the first part of the chapter, we introduce the relation between attention and eye movements and then review the foundations of attention research, social attention and the effect of stimulus saliency on attention. In the second part, we first review traditional characterizations of eye movements, specifically measures of fixations and saccades, and then we review the major methods for the spatial analysis of eye movements, including heat maps and area-of-interest analyses. In the third part, we introduce a new temporal analysis of eye movements, recurrence quantification analysis. We review the basic methods and measures of this analysis, and we discuss several applications. In the last part of the chapter, we provide an overview of methods for the comparison of scanpaths and show that recurrence quantification analysis can be generalized successfully to scanpath comparisons.

With the information provided in this chapter, readers will be able to understand
- The relationship between attention and eye movements
- The different methods for characterizing eye movements in the spatial and temporal domain
- The fundamentals of recurrence quantification analysis for characterizing eye movement dynamics
- The major scanpath comparison methods

2. Introduction: Attention and eye movements

Human attention is the mechanism that enables one to filter out some sensory information in the environment and sharpen the perception of other information, much as one disregards the surrounding text on a page while reading specific words (Broadbent, 1971). Typically, attentional orienting is accompanied by fast jerky eye movements - called saccades - that rapidly acquire detailed information from the environment, such as those being performed while reading the sentences in this chapter (see e.g. chapter by Hyönä & Kaakinen, this volume). These shifts in attention are called *overt* because they involve reorienting the eyes, a behaviour that can be observed by others (Findlay & Gilchrist, 2003). Interestingly, though, eye movements are not needed for shifts of attention to occur. For instance, when you reach the end of this sentence, stare at the period and then identify the word that is one line below it while keeping your eyes on the period. These shifts in attention in the absence of eye movements are called *covert* because they do not require reorienting of the eyes and thus cannot be directly observed by others (Posner, 1978). In general we look at things that interest us, and therefore overt attention and covert attention align. Therefore the focus of discussion in this chapter concerns attention, as measured by overt shifts in attention, specifically, those that are executed as saccadic eye movements.

Both forms of orienting, covert and overt, can be activated *reflexively* by external events in the environment (this is sometimes referred to as exogenous orienting) and *voluntarily* by internal goals and expectations (also called endogenous orienting) (Kingstone, 1992). Consider for example that in reading this sentence you are moving your overt attention voluntarily from one word to the next.

Alternatively, imagine that you are driving a car when a ball suddenly bounces into the road ahead of you. Initially, your attention is captured by the ball (overt reflexive orienting) but you do not track the ball for long. Rather, you look back to the area that the ball came from (overt voluntary orienting) to see if a child is about to run into the road after the ball.

In everyday life there are often competing demands on reflexive and volitional attention, and in order to respond appropriately to these competing demands, one needs to be able to coordinate attentional processes. For example, in the driving scenario described above, you would need to be able to break your reflexive attention away from the ball to search in a volitional manner for a child.

In the lab, two classic paradigms used to isolate these two processes, and their coordination, are the prosaccade task and the anti-saccade task (see Figure 1 and chapter by Pierce et al, this volume). Saccadic latency is shorter when saccades are directed towards a target (prosaccades) than away from it (anti-saccades). This difference is typically attributed to the fact that different processes are involved in prosaccade and antisaccade generation. Prosaccades are reflexive (exogenous) responses triggered by the onset of a stimulus. Antisaccades require two processes: the inhibition of a prosaccade and volitional (endogenous) programming of a saccade in the opposite direction (Olk & Kingstone, 2003).
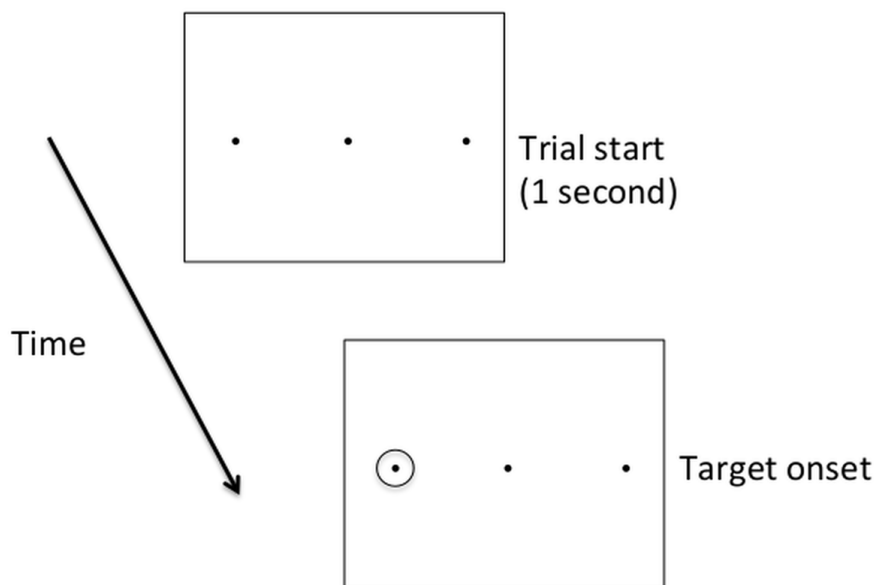

Figure 1: Example trial sequence in a prosaccade and antisaccade task. Participants are presented with a central fixation dot flanked by two possible target locations. After a delay of delay of 1 second, a target circle is presented on

the left or right. In a prosaccade task the participant is to fixate the target location as quickly as possible. In an antisaccade task the participant is to fixate the location that is mirror opposite to where the target appeared. Thus, in the trial that is illustrated, if a participant was performing a prosaccade task a left eye movement would be correct, and if a participant was performing an antisaccade task a right eye movement would be correct.

Difficulty in coordinating attentional processes can be a major source of disability in people with neurologic disorders (see Müri et al, this volume). For example, patients with frontal lobe lesions, caused by a stroke or a closed head injury, may perform normally when looking reflexively to a peripheral light that appears suddenly, as in the prosaccade task. But when asked to look away from the stimulus light, as in the antisaccade task, these patients are extremely slow to respond, and often incorrectly continue to make reflexive eye movements towards the light. It is as if the driver in our scenario above could not stop tracking the ball and thus could not begin to search for a child. Thus it is believed that frontal brain systems are crucial for generating voluntary shifts of attention and inhibiting reflexive ones.

3. Historical Annotations

3.1 Attention research: Assumptions of process stability and control

The study of human attention can be segmented into three historical stages. The first stage occurs in the late 1950s and is characterized by a rapid scientific progression propelled by the methods of traditional psychophysics and experimental psychology. The second stage appears in the mid-1970s and is driven by computational analyses that heralded the arrival of cognitive science. The third phase, which began in the mid-1980s, incorporated evidence from neuropsychology and animal neurophysiology, and more recently, brain imaging, and is subsumed by the field of cognitive neuroscience.

Each of these historical stages are grounded on two basic research assumptions. One is that human attention is controlled by processes that are stable across different situations, meaning that, for example, the processes that are studied in the lab are the same as the processes that are expressed in the real world. Second one can maximize research power by exerting experimental control and minimizing all variability in a situation save for the factor of interest. Below we describe briefly why these two assumptions may very well be invalid insofar as they fail to shed light on attention in real world situations, and how the field has responded by conducting far more complex studies that occur in real life or at least better approximate real-world situations. These more sophisticated studies have in turn demanded more advanced analytical tools, and these analytical methods -- in particular, the recurrence quantification analysis -- are the main focus of the present chapter. Before turning to these analyses, however, a brief historical review is warranted and presented.

3.2 Social attention and stimulus saliency

While the assumptions of process stability and situational control are commonly held and applied in studies of attention, adopting them comes with a degree of risk. The assumption of stability for example eliminates any need or obligation by the scientist to confirm that the factors being manipulated and measured in the lab actually express themselves in the real world. The field does of course check routinely that the effects being measured are stable within the lab environment by demanding that results in the lab be replicable. Unfortunately a result that is stable within a controlled laboratory environment does not necessarily mean that it is stable outside the lab. Indeed there are many examples within the field of human attention indicating that even minimal changes within a laboratory situation will compromise the replicability of the effect (e.g., Soto, Morein-Zamir & Kingstone., 2005; Hunt, Chapman and Kingstone, 2008; Chica et al., 2010).

It has been proposed in much greater detail elsewhere (e.g., Kingstone, Smilek & Eastwood, 2008; Risko & Kingstone, 2011) that an impoverished highly controlled experimental situation is unlikely to inform the field about the attentional processes as they are expressed in everyday real-life situations. It stands to reason then that by increasing situational complexity and reducing experimental control one will begin to better approximate the mechanisms that operate in everyday life.

This approach can be illustrated by first considering that the prevailing eye movement model of Itti and Koch (2000) assumes that where people look is determined by a 'winner take all' visual saliency map. This saliency map is generated from basic stimulus features, such as luminance, contrast and colour. These features are claimed to be combined in a biologically plausible way (based on the workings of the visual cortex) to represent the most interesting or 'salient' regions of a display, image or video (see Figure 2). Despite the mounting evidence that this model has, at best, minimal construct validity in an very narrow band of situations (e.g., Tatler, Baddeley & Gilchrist, 2005; Nystrom & Holmqvist, 2008), it is still prominent in the literature.

What is striking about this model, besides its vast popularity, is how the scenes that have been used to test the model have rarely contained any people. The real world is much more than images of landscapes, buildings, and empty rooms. The real world contains people, and much of it operates in service of the needs of people.
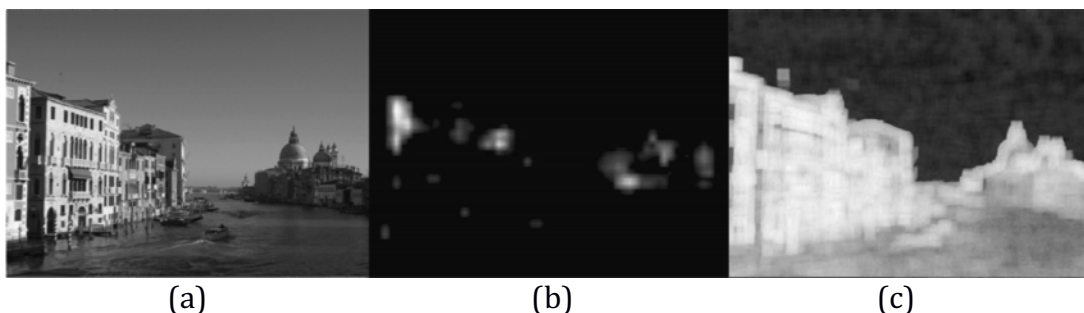

| (a) | (b) | (c) |

Figure 2. (a) Example image; (b) salience map as computed by the Saliency Toolbox in MATLAB (Walther & Koch, 2006); (c) salience map as computed by the Attention based on Information Maximization (AIM) model (Bruce & Tsotsos, 2009).

The recent eye tracking work of Birmingham and colleagues (Birmingham, Bischof, & Kingstone, 2008a, 2008b) has revealed that people are extraordinarily interested in people, in particular, the eyes of people, even when they are embedded in complex scenes. It does not appear to matter very much where the people are in the scenes, what they are wearing, or even how tiny they are represented in a scene. If there is a person somewhere in a photo, then participants are going to look at them quickly, and often -- especially their eyes.

From the perspective of the saliency model, these results are not expected because often the people in the scenes are very small and not at all visually salient (Birmingham, Bischof and Kingstone, 2009). And yet, observers quickly, consistently, and repeatedly seek them out. Thus, there seemed to be a profound bias to search out people, and in particular the eyes of individuals, in complex social scenes.

3.3 Social attention in the real world

The research conducted using social scenes, of course, pertain to simple static scenes (i.e. photographs) of people. In real life people move about, they look at each other, and they talk to one another. What happens in such a situation? Foulsham et al. (2010) asked precisely this question. In their study participants watched videos of different groups of three individuals sitting around a table discussing a hypothetical situation regarding the most important items that they would take to the moon while having their eye movements tracked. Foulsham et al. (2010) found that despite the fact that the individuals in these videos moved, talked and interacted with one another, there remained a tremendous consistency in the participants' looking behavior. Specifically, participants fixated primarily on the eyes of the people in the video (see also Cheng et al. 2013). Thus, even in this dynamic social context, participants' looking behavior evidenced a clear bias to attend to the eyes of others. Furthermore, as with Birmingham et al. (2009), these findings cannot be explained in terms of basic low-level stimulus saliency, in this case, features like visual motion and sound onsets. Foulsham and Sanderson (2013) and Coutrot and Guyader (2014) both investigated whether looks to the faces and eyes of individuals engaged in conversation were significantly affected by changes in visual saliency, or whether the audio is present or absent. In both studies, participants again view complex, dynamic scenes featuring conversation while their eye movements were recorded. Their results indicated that the addition of an audio track increased looks to the faces and eyes of the talkers, and also resulted in greater synchrony between observers when they looked at the speakers (Foulsham & Sanderson, 2013). Critically, however, whether sound was present or not, and independent of changes in low-level visual saliency (Coutrot & Guyader, 2014),

people spent most of their time looking at the faces and eyes of the individuals in the videos.

3.4 Summary

To summarize, it has been found that attention paradigms that are conducted in isolation using simple non-social, carefully controlled visual scenes, a model that assumes that people look at the most salient items can explain some eye movement behaviour. However, when contrary to the classic research approach of simplification and control, participants are shown a wide variety of photos containing people (the pictures are all different) and the behaviour is unconstrained (participants are free to look wherever they wished), then it is discovered that people are primarily interested in looking at the people in the scenes, especially their eyes. These findings persisted when stimulus complexity is further increased by introducing videos that involved people moving and talking. Finally, these data provided a strong test of the Itti-Koch saliency model of human looking behaviour and found that it could not account for such behaviour in these more natural and complex displays.

To address these and similar criticisms, the Itti-Koch saliency model has been revised regularly by incorporating additional features, such as depth and motion, by adding top-down mechanisms, such as face detectors, or contextual guidance (Torralba, Oliva, Castelhano, & Henderson, 2006; Anderson, Donk, Meeter, 2016). Finally, some recent work has used machine learning to learn an optimal set of bottom-up features (e.g. Vig, Dorr, & Cox, 2014). These extensions are further discussed in the chapter by Foulsham (this volume). We also encourage the reader to consult the Saliency Benchmark website http://saliency.mit.edu for a review of recent saliency models and their performance.

4. Traditional Characterizations of Eye Movements

So far we have reviewed fundamental characteristics of attention with a focus on eye movements, and some of the basic paradigms used in attention research. We began with very simple tasks for the study of reflexive and volitional attention using, for example, the prosaccade and antisaccade tasks. We concluded that although these studies are useful and important, highly controlled experimental situations may not fully inform us about the attentional processes expressed in everyday real-life situations (see section 4.2). We then reviewed studies employing more complex tasks such as viewing of static and dynamic images depicting complex social scenes.

In the remainder of this chapter, we focus on the measurement and description of eye movements. First, we briefly review some of the traditional basic eye movement measures and note that they are simply not able to capture and reflect the complex patterns of spatial and temporal eye movement behaviours being produced. Second, we review popular methods for assessing spatial and temporal characteristics of eye movements that are more suitable for describing eye movements in viewings complex scenes. Third, we introduce recurrence quantification analysis, a method that is well suited to the description of the

temporal characteristic of eye movements in real-world situations. Finally, we review several methods of comparing sequences of eye movements.

4.1 Fundamental Measures for Eye Movements

From a psychological perspective, the most important eye movement events are fixations (see also the chapter by Alexander & Martinez-Conde, this volume) and saccades, as described earlier in the chapter. A saccade is a rapid, ballistic motion of the eyes from one point to another, while a fixation is the brief (around 200 millisecond) pause between saccades during which the most visual information is gleaned. Fixations and saccades are extracted from the raw eye-tracking data that is recorded using specialized eye-tracking equipment by applying an algorithm, or series of algorithms to the raw data. Both fixations and saccades can be described in the spatial and the temporal domains. One of the more important spatial fixation measures is the variability of fixations, or where exactly people looked, which can be used for assessing the consistency of eye movements across different observers or across repeated presentations for the same observer. The variability can be measured by determining the variance or the range of fixation positions, and it can be measured with respect to the whole stimulus area or with respect to particular regions of interest. The temporal fixation measures are based on the duration of fixations and include, for example, the average fixation duration, the distribution of fixation durations, or the total fixation duration for fixations within different regions of interest.

The spatial measures of saccadic eye movements include amplitude and direction of saccades. The former refers to the size of saccades and is typically measured by the average saccade amplitude or the distribution of saccade amplitudes for each experimental condition. The direction distribution of saccades describes how often saccades were made in each direction. The temporal measures include saccade duration, that is how much time the saccades take on average, and saccade rate, that is how frequently saccades are made. Finally, spatio-temporal saccade measures include, for example, saccade velocity and acceleration (for further details, see Holmqvist et al., 2011).

In practice, eye movement studies often use combinations of spatial and temporal fixation and saccade measures for assessing differences between experimental groups or conditions. While these measurements are undoubtedly important for characterizing eye movement behaviour, it is often difficult to explain why certain differences exist between experimental conditions or experimental groups, except in very specific circumstances. For example, the average saccade amplitudes may differ between two groups of participants, but without an analysis of eye movement dynamics and without a model of saccade generation, it would be difficult to explain why this would be the case. This, in turn, puts limits on the usefulness of these measures for characterizing eye movement behaviour with complex stimuli. For this reason, we introduce spatial and temporal eye movement measures that are suitable for such stimuli.

4.2 Spatial Analysis of Eye Movements

This section presents an overview of the predominant methods for visualizing and analyzing the spatial distribution of fixations. Figure 3a shows a scene of three people playing cards, with over 2000 fixations produced by 21 participants while viewing this scene. Each red dot represents one or more fixations at that image location. From the distribution of fixations, one can see that a large proportion of fixations landed on the persons in the scene, and in particular on their faces. The fixation plot of Figure 3a can be visualized in different ways that make areas with high fixation densities more explicit.
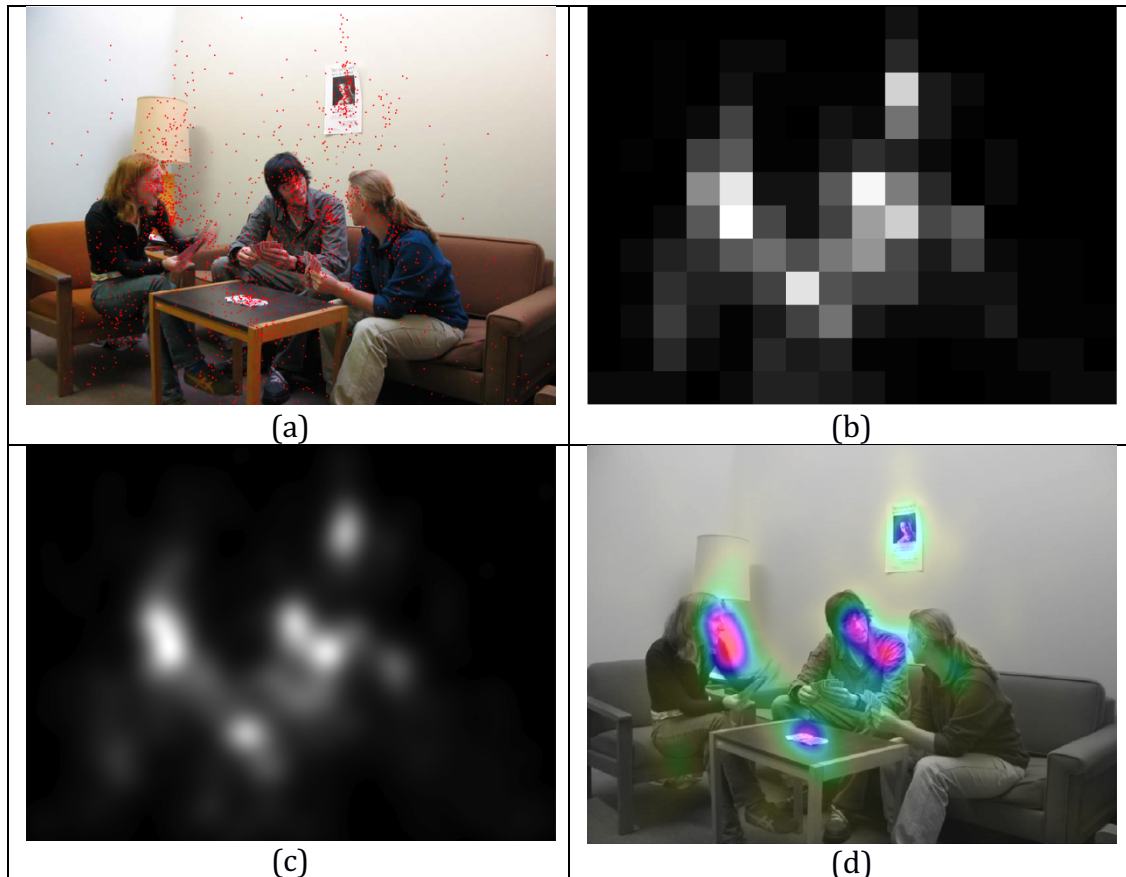

Figure 3: a) Image of a social scene with fixations overlaid as red dots. b) Gridded heat map with 16x12 cells. Grey level of each cell is proportional to the number of fixations that landed within the cell. c) Smooth heat map, where fixations have been replaced by 2D-Gaussians. The grey level at is point is proportional to the height of the heat map. d) Variation of smooth heat map where only the peaks are shown in colour with the color proportional to the height of the heat map, whereas the original image is shown in areas of low heat map values.

In the gridded heat map (see Figure 3b), the stimulus area is partitioned in a rectangular grid of square cells, in this case of 16 by 12 cells. The grey level of each cell is proportional to the number of fixations that landed within the cell. The gridded heat map makes it easy to see the image areas with the most fixations; here the two faces on the left, the card deck on the table and the picture on the wall. The visualization of fixations using gridded heat maps has a clear advantage over raw fixation maps (as in Figure 3a), and it allows to make

statistical comparisons of fixation frequency counts between different groups or experimental conditions fairly easy. On the negative side, this technique introduces artificial boundaries between cells that have no relation to the scene content.

Alternatively, the fixations can be visualized using a smooth heat map. In this method, each fixation is replaced by a 2D-Gaussian with a pre-defined standard deviation, and these Gaussians are added together, resulting in a smooth fixation map, which is often referred to as a heat map. In the case of Figure 3c, the standard deviation was chosen to be 20 pixels (with an image size of 800 by 600 pixels). There are, however, no hard and fast guidelines regarding the best choice of the standard deviation. This can often make it difficult to compare heatmaps across different experimental conditions. The heat map can also be visualized by assigning a grey level proportional to the height of the heat map (Figure 3c) or by assigning a Gaussian standard deviation based on fixation duration. Again, the heat map makes it clear which image areas were fixated the most. An interesting variation of heatmaps was created by Woodford (2014), where the heat map is overlaid over the original image, but only the peaks are shown in colour with the color proportional to the height of the heat map, whereas in regions with low values, the original image is shown (see Figure 3d).

4.2.1 Limitations of Heat Maps

Heat maps are good visualization tools and can give a quick overview of fixation patterns for a large number of participants and to easily locate fixation hotspots such as the faces in Figure 3. For several reasons, however, the analysis and the interpretation of heat maps are difficult. First, heat maps are specific to a particular stimulus layout, and hence can be used only to compare heat maps of fixation patterns obtained for the same images (or images with identical spatial layouts). Second, fixation hotspots may not be due to participants fixating a particular scene point, but could be due to the fact that participants tend to fixate the center area of stimuli more frequently than peripheral areas. This central fixation bias is illustrated in Figure 4, which shows the heat map obtained from the fixations of 21 participants looking at 10 different images. The (red) hotspots in the center indicate a strong bias towards fixating central image areas independent of particular scene contents (Tatler, 2007).
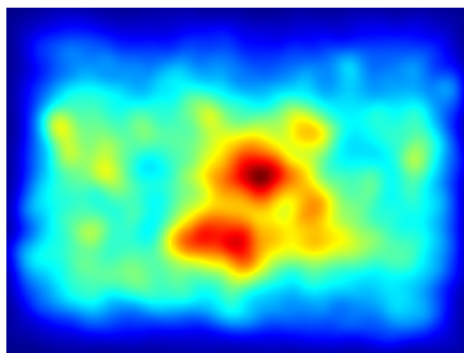
Figure 4: Heat map obtained from fixations by 21 participants looking at 10 different images. Red areas indicate regions with high fixation counts and blue indicates areas with low fixation counts.

Textbox: Comparison of Heat Maps

Heat maps must be analyzed statistically to establish differences and similarities obtained under different experimental conditions, as well as to establish whether certain fixation hotspots are significantly higher than the rest of the heat map (for a detailed overview of heat map comparisons see Le Meur & Bacchino, 2013). The similarity of heat maps can, for example, be measured using simple correlation between two heat maps, measured over all image positions. Second, one can compare the two heat maps using the Kullback-Leibler divergence, a non-symmetric measure of the difference between two probability distributions (Kullback & Leibler, 1951). Third, the heatmaps can be compared using ROC analysis (Green & Swets, 1966) where the ROC curve expresses the correspondence between the two maps (for more details see Judd, Ehinger, Durand, & Torralba, 2009; Le Meur & Baccino, 2013).

The three methods are unproblematic for gridded heat maps because the number of fixations is independent between different cells of the grid. For smooth heat maps, however, the comparison is more complicated due to the fact that independence between nearby positions is lost due to the spatial smoothing with the Gaussian filter. This is a situation similar to the one in the analysis of fMRI results, where a smoothing is applied to the raw data (Friston, Jezzard, & Turner, 1994; Friston, et al. 1995). For eye movements, Caldara (2011) has proposed a pixelwise comparison of heat maps that takes this dependence into account. In general, it is best to rely on bootstrapping methods (Efron & Tibshirani, 1993) for the statistical analysis of heat maps.

The methods presented in this section visualize the spatial distribution of fixations of many participants in an image, and they are easy to understand. Their statistical analysis is well understood, but their interpretation is a bit more difficult. They express the spatial distribution of fixations in image coordinates without reference to specific scene content, so the analysis with respect to scene content requires additional work.

4.3 Area of Interest Analysis of Eye Movements

In order to take the content of a visual scene into account, a second group of eye movement measures is concerned with analyzing the proportion and duration of fixations in pre-defined areas of interest (AOIs). In a visual search display, one might define AOIs for each target and each distractor. In images of more complex scenes, for example, in work with social scene stimuli (e.g., Birmingham, Bischof, and Kingstone, 2008a, 2008b), the AOIs can include the bodies, faces and eye regions of the persons in the scenes (see Figure 5). An AOI analysis consists of

counting all fixations or determining the fixation proportions or average fixation durations (dwell time) for the eye movements within each AOI.

Figure 5: Areas of interest of the social scene. The color of each area has been chosen randomly.

4.3.1 Advantages of AOI analyses

AOI analyses have proven useful for assessing eye movement patterns with simple and complex images. Fixation frequencies or dwell times on different AOIs can be compared between different experimental groups or conditions and even between different stimuli provided they have the same image structure, i.e. they contain the same types of regions (for example, people, faces or eyes in social scenes). Assuming a uniform distribution of fixations across images, the number of fixations or dwell time in an AOI is proportional to the AOI area. If fixation counts or durations need to be compared between different AOIs, then it is advisable to normalize the number or duration of fixations by the AOI area and obtain the number of fixations or dwell time per unit area in each AOI.

The definition of AOIs is relatively straight-forward with static images and many eye tracking systems provide software for defining static AOIs. With dynamic stimuli, the definition of AOIs becomes more difficult. Example of such cases include movies as stimuli or with eye tracking in real-world experiments, where participants are wearing a mobile eye tracker and fixations are projected onto the recorded scenes by the eye tracking software. This is illustrated in Figure 6, where the rider changes position and size in every movie frame, leading to an enormous effort for defining the AOI rider for long frame sequences. One possible solution to this problem relies on so-called keyframing. In this technique, the relevant AOIs are defined only in certain frames (key frames), and the computer then interpolates the AOI boundaries for all the frames in between (Igarashi et al., 2005). With this technique, the effort of encoding AOIs is substantially reduced. Under certain circumstances, in particular with computer-generated stimuli, AOIs can be automatically generated with relatively little effort. In mobile eye tracking it is also possible to utilize markers (unique patterns printed on a piece of paper) to define regions of interest in the real environment. These markers can then be used to translate world-view fixation

coordinates to fixation coordinates based on the area around a marker instead (e.g., Krassner, Patera & Bulling, 2014)

Figure 6: Three frames from the opening scene of the movie "The Good, the Bad and the Ugly". The size and location of the rider change in each frame, increasing the effort of defining areas of interest.

4.3.2 Disadvantages of AOI analyses

The AOI analysis is useful for the comparison of fixation patterns for identical or structurally similar stimuli with a common set of elements. It can, however, be difficult to use AOIs to derive meaningful results if highly diverse stimuli, e.g., when a mixture of social scenes, landscapes and artificial stimuli is being used. In this case, there is no obvious way of comparing AOIs across the set of stimuli.

The sections on the spatial analysis of fixation patterns using heat maps and areas of interest have introduced several analysis methods that are very popular in the eye movement literature for analyzing the spatial allocation of attention in images. On the downside, the methods do not capture the dynamics of eye movement behaviour, unless one analyzes their development over time. This is, however, only possible with a very coarse temporal resolution. Methods more suitable for the temporal analysis of eye movements are discussed in the next section.

4.4 Temporal Analysis of Eye Movements

Temporal eye movement analysis is concerned with the dynamic aspects of eye movements, the temporal sequence of fixations and saccades as they unfold over time (i.e., the scanpath). In the context of AOI analysis, it is concerned with the transitions between different AOIs and typically focuses on AOI transitions matrices. For example, with a set of stimuli derived from an image of two people sitting at a table, the regions head1, torso1, head2, torso2, table, and chairs, could yield a transition matrix between the AOI regions as shown in Table 1.

	head1	torso1	head2	torso2	chairs	table
head1	0.1	0.4	0.3	0.1	0.1	0.0
torso1	0.2	0.2	0.1	0.2	0.2	0.1
head2	0.3	0.1	0.3	0.3	0.0	0.0
torso2	0.1	0.1	0.4	0.2	0.1	0.1
chairs	0.1	0.3	0.1	0.3	0.1	0.1
table	0.3	0.1	0.3	0.1	0.2	0.0

Table 1: Example of a transition matrix for six AOI regions

Each row in Table 1 indicates the probability of saccades from the AOI on the left to each of the AOIs. For example, if a fixation at time t is in the region head2, then the next fixation at time t+1 is in the region head3 with a probability of 0.3, in the region torso1 with a probability of 0.1, and so on. The AOI transitions in Table 1 characterize the dynamic sequence of eye movements for a given experimental condition, and comparisons to other conditions can be made by assessing the similarity of the transition matrices. It is important to point out, however, that such transition matrices capture only the overall characteristics of the transitions, but not, for example, their change over time.

In specific applications, the transition matrices have been modeled using Markov and hidden Markov models (Stark and Ellis, 1981; Holmqvist et al., 2011, Boccignone, 2014), for example, in face recognition (Chuk, Chan, & Hsiao, 2014) or participants solving items from Raven's Advanced Progressive Matrices Test (Hayes, Petrov, & Sederberg, 2011; see also Boccignone's chapter in this volume). As pointed out earlier, such AOI-based analyses are feasible only when the class of stimuli is sufficiently restricted, e.g. contain a common set of regions. Otherwise, the transition matrices cannot be compared across different stimuli. In contrast, we introduce in the next section a method for characterizing scanpaths obtained with related as well as unrelated images.

4.5 Summary

In this section we have introduced the analysis of eye movements and reviewed some common spatial and temporal methods of analysing fixations. We have discussed spatial techniques such as using heatmaps to determine where people look as well as AOI analyses that include information about the scene content. We have briefly discussed the use of transition matrices to analyze the dynamic aspects of using AOI's.

An important group of other dynamic fixation measures are the scanpath analyses that were developed in reading research. In these studies, scanpath events, such as backtracking, regression, look-ahead and return have a direct interpretation. In scene viewing, however, scanpath measures are more difficult to interpret, and there is little consensus on good measures for scanpath comparison. Almost no measures have been developed for directly quantifying a single scanpath in an image. In the sections below, we describe one recently developed method for quantifying a single scanpath, then we describe several techniques for comparing similarities between scanpaths.

5. Recurrence Quantification Analysis (RQA)

Recurrence analysis has been used successfully as a tool for describing complex dynamic systems, for example, climatological data (Marwan & Kurths, 2002), electrocardiograms (Webber & Zbilut, 2005), or postural fluctuations (Riley & Clark, 2003, Pellecchia & Shockley, 2005), which are inadequately characterized by standard methods in time series analysis (e.g., Box, Jenkins, & Reinsel, 2008).

It has also been used for describing the interplay between dynamic systems in cross-recurrence analysis, e.g., for analyzing the postural synchronization of two persons (Shockley, Baker, Richardson, & Fowler, 2007; Shockley, Santana, & Fowler, 2003; Shockley & Turvey, 2005). The fundamental idea of recurrence analysis is to analyze the temporal pattern of repeated (recurrent) events, for example, the same tidal height in tide analysis, the same waves in the ventricular cycle in electrocardiogram analysis, the same postural relation in the analysis of postural synchronization, or the same fixated locations in an image. Here we describe a simplified version of recurrence analysis based on categorical data. This simplified analysis is ideal because it allows for direct interpretation of the various recurrence measures as they apply to categorical fixation data.

5.1 Categorical Recurrence Quantification

Richardson, Dale and colleagues have generalized recurrence analysis to the analysis of categorical data and have used it for analyzing the coordination of gaze patterns between individuals (e.g., Cherubini, Nüssli, & Dillenbourg, 2010; Dale, Kirkham, & Richardson, 2011; Dale, Warlaumont, & Richardson, 2011; Richardson & Dale, 2005; Richardson, Dale, & Tomlinson, 2009; Shockley, Richardson, & Dale, 2009). For example, Dale, Warlaumont & Richardson, 2010, quantified the coordination between a speaker and a listener's eye movements as they viewed actors on a screen (see Figure 7).
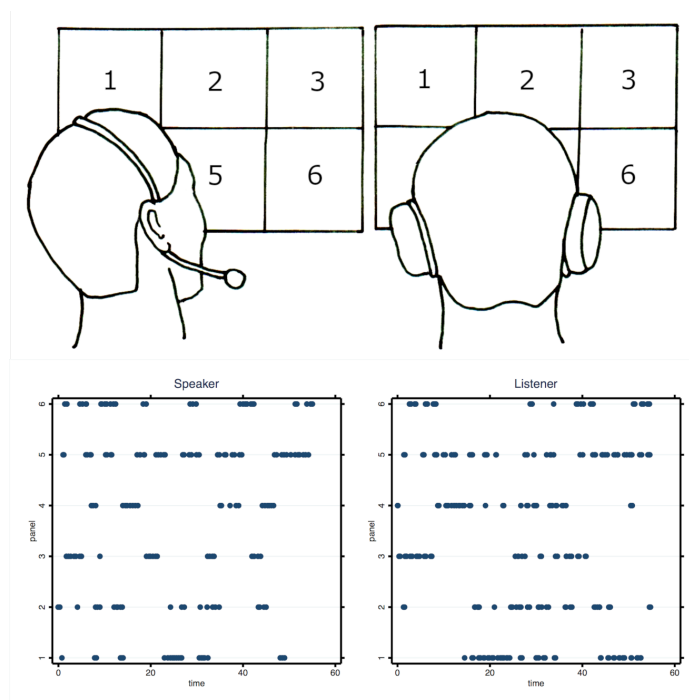

Figure 7: Dale, Warlaumont & Richardson's (2011) experiment. Upper panel: The left person is speaking about a 2 × 3 grid displaying television characters while the right person listens in. Lower panel: Each interval of the one-minute experiment produces a numeric value representing the panel that was fixated. Using these data, a cross-recurrence analysis compares the panels fixated by speaker and listener over different time lags. (Figure adapted from Dale, Warlaumont & Richardson, 2011)

In this experiment, one person (left) was talking about a particular television character while the other person (right) listened in. Pictures of characters including the one discussed were displayed in a 2x3 grid in front of the speaker and the listener. The eye movements were categorized into a grid corresponding to the six photos of the television characters. The eye movements generated by speaker and listener during a 60 second period are shown in a series of numbers 1-6 corresponding to the six grid locations. Dale et al. (2010) used a cross-recurrence analysis of these fixations and were able to show that the listener tended to follow the same fixation patterns as the speaker, with a delay of approximately 2 seconds.

5.2 Generalized Recurrence Quantification

Dale et al.'s cross-recurrence analysis can provide an overall measure of similarity across two eye movement sequences (i.e., a form of scanpath comparison that we will discuss later in this chapter). Recently, we have introduced a generalized form of Dale et al.'s categorical recurrence analysis to characterize gaze patterns of a single observer (Anderson, Bischof, Laidlaw, Risko, & Kingstone, 2013), and we were able to show that it is a very useful tool for encoding general characteristics of fixation sequences. The essential idea is to consider each fixation as one in a series throughout the image. Recurrence analysis quantifies when one fixation might overlap in space with another (one person looking at the same place twice or more times in the course of looking at a scene). In the following, we first review the fundamentals of recurrence quantification analysis (RQA) for use with categorical eye movement data, with specific consideration of fixation sequences. Then we describe and interpret the main RQA measures.

5.3 Formal Definition of Recurrence

Consider a sequence of N fixations f_i, i=1,…,N, with each f_i characterized by its spatial coordinates. Two fixations are considered recurrent, if they are close together. "Closeness" can be defined in several ways, but in general, one can define recurrence r_{ij} as

$$r_{ij} = \begin{cases} 1, \wedge \; d(f_i, f_j) \leq \rho \\ 0, \wedge \; otherwise \end{cases} \tag{1}$$

where d is some distance metric (usually Euclidian distance), and ρ is a given radius, i.e., two fixations are considered recurrent if they are within a distance ρ of each other. Guidelines for selecting a value for ρ are introduced further below.

5.4 Recurrence Plot

Recurrence can be represented in a recurrence diagram, which plots recurrences of a fixation sequence over all possible time lags. The essential starting point of a recurrence analysis is drawing this plot. While it is not strictly a necessary step, all of the recurrence measures are based on patterns that emerge from this plot. The plot is drawn as follows: If fixations i and j are recurrent (i.e. if r_{ij} = 1), then a dot is plotted at position i,j. All fixations are recurrent with themselves (since

d(f_i, f_i)=0), hence all elements on the major diagonal – the line of incidence - are recurring. Furthermore, since distance metrics are symmetric (i.e., d(f_i, f_j) = d(f_j, f_i)) recurrence plots are also symmetric. This is illustrated in Figure 8. Figure 8a shows a landscape image with a scanpath consisting of 30 fixations, with repeated fixations mainly in the cloud formation. The recurrence plot for the fixation scanpath is shown in Figure 8b, with each recurrence indicated by a red dot. A recurrence plot can be generated for each sequence of fixations, e.g. for each experimental trial.
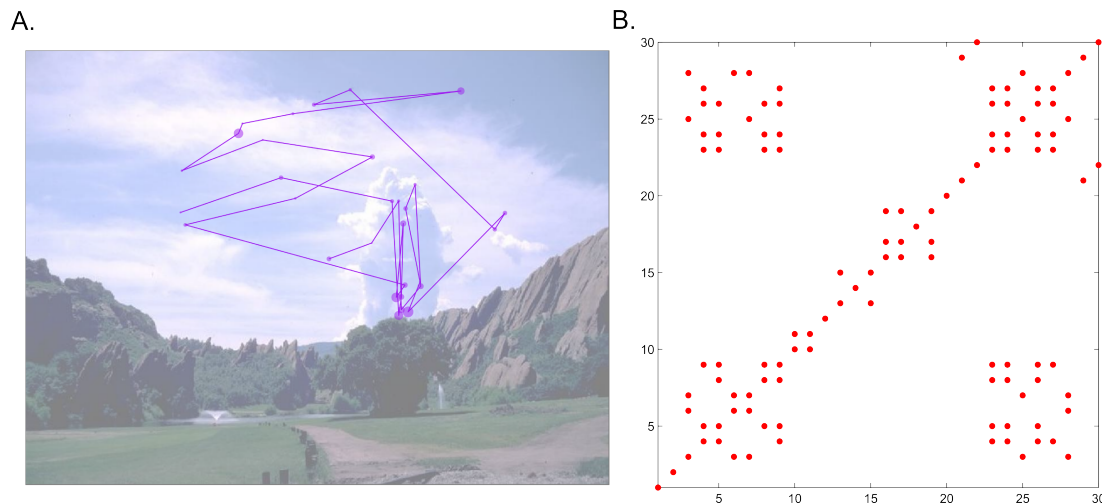
A.

B.

Figure 8: A) Image of a landscape with the scanpath produced by one participant overlaid (the size of the circle at each fixation point represents the duration of the fixation); B) recurrence plot of the scanpath in A.

5.5 Recurrence Quantification Measures

The recurrence diagram provides a useful visual representation of the recurrence patterns for a fixation sequence, but it must be complemented by a recurrence quantification analysis for comparison across different fixation sequences, e.g., across different trials, participants and experimental conditions. Here, we describe a subset of RQA measures, those that are particularly useful for the analysis of fixation sequences (see Webber & Zbilut, 2005 and Marwan et al., 2002 for a complete list of measures). All of these measures describe certain patterns that emerge in the recurrence plot. For example, the recurrence measure itself is simply a percentage of the total number of possible recurrent points. Determinism is the percentage of recurrent points that form a diagonal line on the plot, while laminarity is the percentage of vertical and horizontal lines. Given the symmetry of the recurrence diagram, these quantitative measures are usually extracted from the upper triangle of the recurrence diagram, excluding the line of incidence, which does not add any additional information (recall that the line of incidence indicates that each fixation is recurrent with itself).

First, we give some useful definitions: Given a fixation sequence of length N, f_i, i=1,…,N, let R be the sum of recurrences in the upper triangle of the recurrence

diagram, i.e., $R = \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} r_{ij}$. Let D_L be the set of diagonals lines, H_L the set of horizontal, and V_L the set of vertical lines, all in the upper triangle, and all with a length of at least L, and let $|\cdot|$ denote cardinality.

5.5.1 Recurrence

The <u>recurrence</u> measure is defined as

$$\text{REC} = 100 \frac{2R}{N(N-1)} \qquad (2)$$

For a sequence of N fixations Recurrence represents the percentage of recurrent fixations (i.e. it indicates how often an observer re-fixates previously fixated image positions). As fixations are plotted sequentially, the larger the distance between a recurrent point and the main diagonal, the larger the time interval (in number of fixations) between the original fixation and the re-fixation.

Figure 9 provides an illustration of recurrence. Fixations 6, 7, 25 and 28 in Figure 9 are all close-by, more specifically within a radius of ρ of fixation 3, hence these fixations are recurrent.
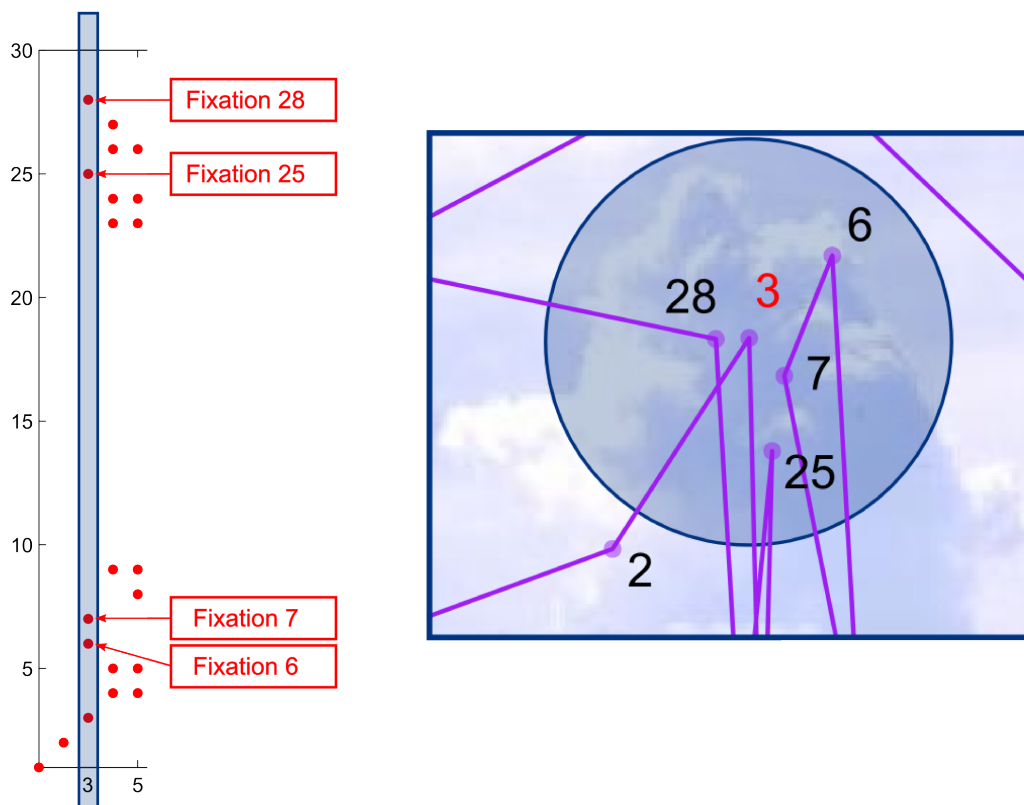

Figure 9: Detailed view of the recurrence plot and the fixation plot shown in Figure 8. Fixations 6, 7, 25 and 28 are within the radius ρ of Fixation 3, as indicated by the blue circle, producing the shown recurrences in the recurrence plot.

18

5.5.2 Determinism

The <u>determinism</u> measure is defined as

$$\text{DET} = 100\frac{|D_L|}{R} \qquad (3)$$

Determinism measures the proportion of recurrent points forming diagonal lines in the recurrence plot and represents repeating fixation sequences in the recurrence diagram. This may represent two areas of a scene where one fixation is more likely to follow another. For example, if a person looks first at one person, then another in that same order twice in a trial. It is a small section of a scanpath repeated within a trial. The minimum line length of diagonal line elements is typically set to L=2. The length of the diagonal line element reflects the number of fixations making up the repeated scanpath, and the distance from the diagonal reflects the time (in numbers of fixations) since the scanpath was first followed.

Figure 10 provides an illustration of determinism. Fixations 3 and 4 as well as fixations 25 and 26 follow the same path, defining a deterministic recurrence.
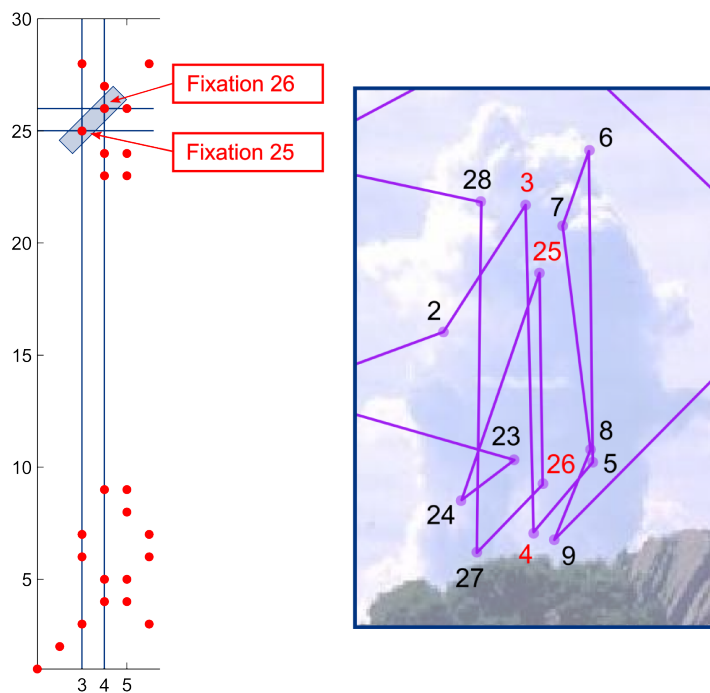

Figure 10: Detailed view of the recurrence plot and the fixation plot shown in Figure 8. Fixations 3 and 4 as well as fixations 25 and 26 follow the same path. This creates a diagonal line on the recurrence plot and defines a deterministic recurrence.

5.5.3 Laminarity

19

The <u>laminarity</u> measure is defined as

$$\text{LAM} = 100\frac{|H_L| + |V_L|}{2R} \qquad (4)$$

Vertical lines represent areas that were fixated first in a single fixation and then re-scanned in detail over consecutive fixations at a later time (e.g., several fixations later), and horizontal lines represent areas that were first scanned in detail and then re-fixated briefly later in time. Again, we set the minimum line lengths of vertical and horizontal lines to L=2. We find that the recurrence diagrams sometimes contain recurrence clusters (with horizontal and vertical lines), indicating detailed scanning of an area and nearby locations. Laminarity indicates that specific areas of a scene are repeatedly fixated, for example, when an observer returns to an interesting area of the scene to scan it in more detail.

An example of laminarity is shown with fixations in Figure 9. The position fixated in fixation 3 is refixated in a more detailed inspection in fixations 6 and 7.

5.5.4 Center of Recurrence Mass

The <u>center of recurrence mass</u> (corm) is defined as the distance of the center of gravity of recurrent points from the line of incidence, normalized such that the maximum possible value is 100.

$$\text{CORM} = 100\frac{\sum_{i=1}^{N-1}\sum_{j=i+1}^{N}(j - i)r_{ij}}{(N - 1)R} \qquad (5)$$

This measure indicates approximately where in time most of the recurrent points are situated. Small corm values indicate that re-fixations tend to occur close in time, i.e. most recurrent points are close to the line of incidence. For example, if an observer sequentially scans three particular areas of a scene in detail and never returns to those areas later in the trial, most of the recurrent points would fall close to the line of incidence. This would be represented by a small corm value. In contrast, large corm values indicate that re-fixations tend to occur widely separated in time, i.e. most recurrent points are close to the upper left and lower right corners of the recurrence diagram. This occurs, for example, when an observer re-fixates only one scene area, once at the beginning and once at the end of the fixation sequence, but not in between.

Textbox: RQA and Fixation Duration

Fixation duration can be an important indicator of processing during fixation (Holmqvist et al., 2011, pp. 377ff). RQA can be generalized to take fixation durations into account (Anderson et al., 2013).

Given a fixation sequence f_i, i = 1, . . . , N, and the associated vector of fixation durations t_i, i = 1, . . . , N, one can redefine recurrence $r_{ij}{}^t$ as

$$r_{ij}^t = \begin{cases} t_i + t_j, & d(f_i, f_j) \le \rho \\ 0, & otherwise \end{cases} \qquad (1)$$

with the (Euclidian) distance metric d and the radius ρ. With the modified recurrence definition of $r_{ij}{}^t$, the RQA measures have to be renormalized. Let $R = \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} r_{ij}^t$, and $T = \sum_{i=1}^{N} t_i$. Then the revised definitions for *REC, DET, LAM and CORM* are as follows.

$$REC^t = 100 \frac{2R^t}{(N-1)T}$$

$$DET^t = \frac{100}{R^t} \sum_{(i,j) \in D_L} r_{ij}^t$$

$$LAM^t = \frac{100}{2R^t} \left(\sum_{(i,j) \in H_L} r_{ij}^t + \sum_{(i,j) \in V_L} r_{ij}^t \right)$$

$$CORM^t = 100 \frac{\sum_{i=1}^{N-1} \sum_{j=i+1}^{N} (j-i) r_{ij}^t}{(N-1)^2 T}$$

In summary, the recurrence and corm measures capture the global temporal structure of fixation sequences. They measure how many times given scene areas are re-fixated (recurrence) and whether these re-fixations occur close in time or widely separated in a trial (corm). In contrast, determinism and laminarity measure the finer temporal structure of fixation sequences. Specifically, they indicate sequences of fixations that are repeated (determinism) and points at which detailed inspections of an image area are occurring (laminarity). These measures can then be compared across different types of images, experimental contexts and participants to assess the dynamic structure of eye movements.

5.6 Selection of the Recurrence Radius

As explained in section 6.3, two fixations are considered recurrent if they are within a distance ρ of each other, with the radius ρ being a free parameter. The number of recurrences is directly related to the radius. As the radius ρ approaches zero, (off-diagonal) recurrences approach zero, and as ρ approaches the image size, recurrences approach 100%. The dependence of recurrence on

radius leads to the obvious question of how an appropriate radius for recurrence analysis should be selected.

Webber and Zbilut (2005) suggested several guidelines for selecting the proper radius, including the selection of a radius such that percentage of recurrences remains low, for example about 1-2 percent. In the case of eye movements, one can apply more content-oriented criteria. For example, fixations can be considered as recurring if their foveal areas overlap, using a radius size of 1–2 degrees of visual angle. This is discussed further by Anderson et al. (2013).

5.7 Statistical Analyses of Recurrence Measures

In this section, we discuss the distribution of RQA measures as well as their correlations. This is important when RQA measures are used to compare and discriminate between experimental conditions and groups. Figure 11 shows the histograms of the measures recurrence, determinism, laminarity and corm obtained with 104 participants viewing 1872 images, each for a duration of 10 seconds. As the histograms show, the RQA measures are distributed more or less symmetrically, with the exception of Recurrence, permitting the use of analyses of variance for their statistical analysis.
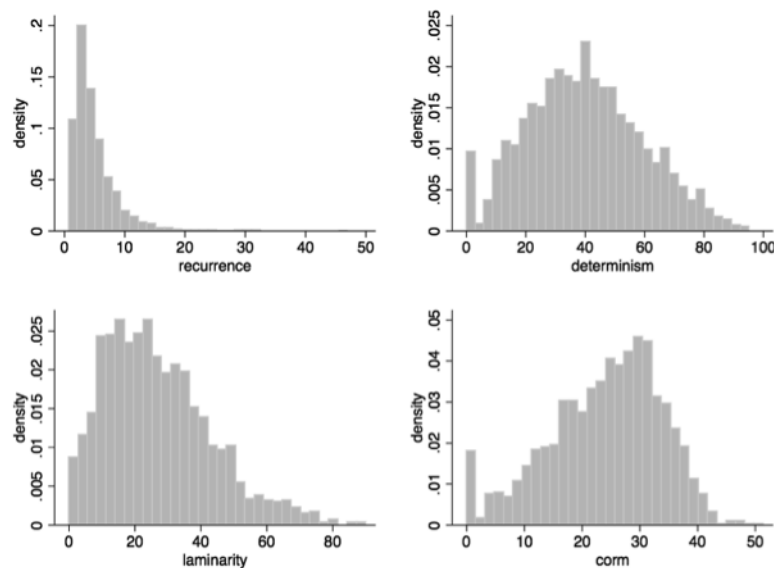

Figure 11: Histograms of the measures recurrence, determinism, laminarity and corm obtained with 104 participants viewing 1872 images, each for a duration of 10 seconds.

Figure 12 illustrates scatter diagrams and correlations between RQA measures, obtained for the same group of participants and the same images as in Figure 11. As the Figure shows, the correlations between RQA measures vary substantially. While the correlation between laminarity and corm is essentially zero, other correlations, e.g. between recurrence and laminarity are relatively high. Such dependencies between the RQA measures are not surprising, given that they are

defined and selected to be easily interpretable, rather than being independent of each other. Indeed, laminarity and determinism are very closely related to recurrence as they are themselves percentages of recurrent points.
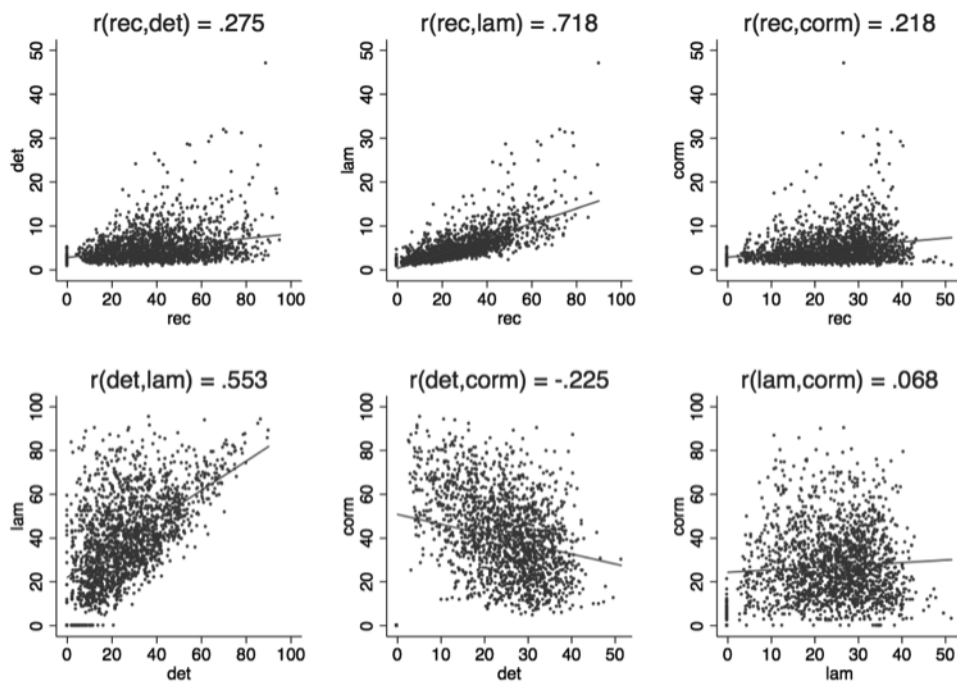

Figure 12: Correlations between RQA measures. Each panel shows a scatterplot for a pair of measures, the regression line and the correlation value.

For this reason, it is useful to consider several or all RQA measures for comparing and discriminating different groups of participants or different experimental conditions. In other words, it is advisable to take all RQA measures into account and use discriminant analyses for distinguishing different scanpath patterns, a multivariate analysis of variance, analysis of variance, or canonical correlation in regression-type analysis. This is further described below.

5.8 Discrimination of Gaze Patterns using RQA

The previous sections showed that RQA is a useful method for discriminating eye movement patterns based purely on dynamic characteristics. Each of the RQA measures used (recurrence, determinism, laminarity and corm) can discriminate to a certain extent between experimental groups or different groups of participants. At the same time, however, the measures are not independent, as the correlations between RQA measures in the previous section showed. To elaborate on this aspect, we evaluated how well the RQA measures discriminate between eye movement patterns of participants who viewed social scenes under natural viewing conditions (Full-View condition) and participants who view the same scenes through a gaze-contingent window of limited spatial extent (Restricted-View condition).

Figure 13 shows histograms for the measures recurrence, determinism, laminarity and corm, for the Full-View and the Restricted-View conditions. For clarity of presentation, the histograms were smoothed using a Gaussian smoothing filter with half-width = 2 percent or $\sigma = 0.849$ percent of the total range. An inspection of the four histograms shows that each measure can discriminate to some extent between the two groups, but, at the same time, there is a substantial overlap that varies between RQA measures. Consequently, classification accuracy of the discrimination between the two groups varies accordingly, with accuracy for the recurrence measure at 75%, for the determinism measure of 88.6%, for laminarity at 77.7%, and for corm at 70.8%. In other words, each of the measures can discriminate between the two experimental groups, but discrimination performance is far from perfect.
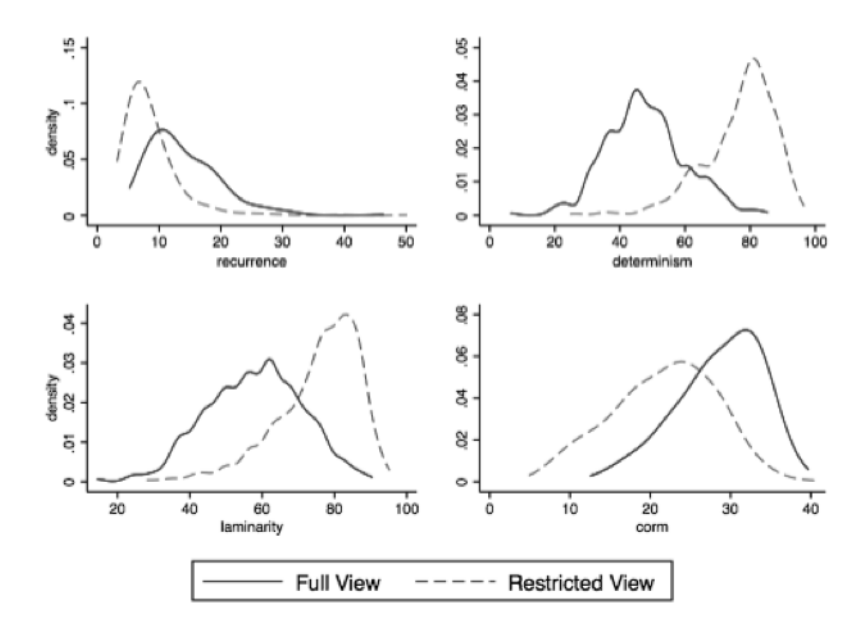

Figure 13: Smoothed histograms for the measures recurrence, determinism, laminarity and corm for the full-view (natural viewing) condition and the Restricted-View (viewing through a small, gaze-contingent window). The histograms were smoothed with a Gaussian filter with half-width = 2 percent.

Discrimination between the eye movement patterns of the two groups can be improved substantially using combinations of RQA measures. This is possible because the measures are somewhat, but definitely not perfectly correlated. For example, using the measures recurrence + laminarity + corm improves discrimination accuracy to 85.5%, and the discrimination accuracy using all four measures is 94.5%, a substantial improvement over the use of single measures. The results show clearly that the RQA measures are sensitive to, and useful for discriminating gaze patterns under Full-View and Restricted-View conditions.

Textbox: Potential Limitations of the RQA Method

1. RQA can measure the characteristics of re-fixations, but it cannot measure idiosyncratic eye movement characteristics, such as saccadic distances, angles, and so on.

The purpose of RQA is to capture the dynamic aspects of scanpaths and should be used <u>in addition</u> to other measures (e.g. fixation and saccade characteristics), not replace them.

2. RQA analyzes scanpaths under the assumption that there is always recurrence.

This is not quite correct. First, the recurrence measure can always be computed and is zero when there is no recurrence. It is, however, correct that the other measures, determinism, laminarity, and corm are undefined if there is no recurrence. Second, one can always compute the RQA measures with larger recurrence radii ρ (see equation 1) and possibly with a whole range of recurrence radii. As the recurrence radius ρ increases, so does the number of recurrences (see section 6.4 and Anderson et al., 2013).

3. RQA cannot be applied if there is a single fixation.

For this extreme case, it is indeed true that RQA cannot be applied, but the same is true for many other measures including those describing saccade characteristics.

4. What are the computational limitations of RQA?

The number of cells in the recurrence plot increases with the square of the number of fixations n in a scan path. Consequently, computing RQA measures becomes computationally more and more expensive as n increases. In our own work, we have rarely used RQA for scanpaths with more than 100 fixations.

5.9 Summary

In this section we have presented a new technique, RQA, for quantifying the characteristics of a single scanpath. This is a useful measure that compliments other spatial and temporal measures of eye movement characteristics. We have described recurrence in general, several other useful measures associated with RQA, and how to use RQA to discriminate gaze patterns across groups of participants or stimuli. In the next section we will describe methods to compare between two scanpaths, a technique that is useful for describing how two eye movement patterns are similar to each other.

6. Comparison of Scanpaths

Eye movements unfold over time, hence one can also examine the inter-relationship between sequences of eye movements. In his seminal work, Yarbus' (1967) noticed that observers displayed similar scan patterns in viewings of Repin's painting "The Unexpected Visitor" and concluded that "observers differ in the way they think and, therefore, differ also to some extent in the way they look at things" (p. 192). A brief inspection of these scan patterns reveals that they are complex and non-random, and they contain sequences of repeated fixations. Noton and Stark (1971) noticed that observers tend to show similar scan patterns during encoding and later recognition of images. According to their "Scanpath Theory", the sequence of fixations during the first viewing of a stimulus is stored in memory as a spatial model, and stimulus recognition is facilitated through observers following the same scanpath during repeated exposures to the same image. These early observations were made informally by visual inspection, but later research has aimed at quantifying the similarity of scanpaths, for the same observer at different time points or when solving different tasks, or between different subjects.

In the following sections, we describe the scanpath comparison methods that have been introduced in the literature (see Anderson, Anderson, Kingston, & Bischof, 2014). In each case, we give a short description, and the reader is advised to consult the original publications for further details.

6.1 Edit distance

One successful way for comparing scanpaths is based on the string edit distance (Bunke, 1992; Levenshtein, 1966; Wagner & Fischer, 1974), which is used to measure the dissimilarity of character strings. In this method, a sequence of transformations (insertions, deletions, and substitutions), is used to transform one string into the other and their similarity is represented as the number of transformation steps between the two strings. This method has been adapted for comparing the similarity of scanpaths (Brandt & Stark, 1997; Foulsham & Kingstone, 2013; Foulsham & Underwood, 2008; Harding & Bloj, 2010; Underwood, Foulsham, & Humphrey, 2009). To achieve this, a grid is overlaid on an image, and each cell in the grid is assigned a unique character. Fixation sequences are then transformed into a sequence of characters by replacing the fixation with the character corresponding to the grid cell a fixation falls in. With this approach, scanpaths are being represented by strings of characters, and the dissimilarity of two scanpaths can then be represented by the number of transformations required to convert the string corresponding to the first scanpath to the string corresponding to the second scanpath.

The string edit distance method has been very popular in early scanpath comparison work (e.g., Brandt & Stark, 1997) and has been used subsequently in a variety of experimental contexts (e.g., Harding & Bloj, 2010; Underwood et al., 2009). This is an advantage for researchers wishing to directly compare results to these earlier studies. The main advantage of the string edit measure, however, lies in the fact that it captures the intuitive notion of the distance between two scanpaths (i.e., their dissimilarity) in a simple and straightforward way.

Several criticisms have been raised against the use of edit distance for scanpath comparisons. First, as described in section 5.2 for gridded heatmaps, the grid is defined independently of image content. It may thus be too coarse in regions of interest while being too fine in other regions. Second, two fixations may be considered different even when they are close together, namely if they fall on either side of a grid line. Some variants of the string edit distance have been developed to address these problems. For instance, assigning characters to pre-defined areas of interest allow the researcher to add semantic information to the quantization process (Josephson & Holmes, 2002; West, Haake, Rozanski, & Karn, 2006), but the definition of regions of interest can be time-consuming.

6.2 ScanMatch

Cristino et al. (2010) proposed a generalized scanpath comparison method that addresses many of the deficiencies of the string edit distance method. Their generalization aligns eye movement sequences based on the Needleman-Wunsch algorithm (Needleman & Wunsch, 1970), which is used in bioinformatics to compare DNA sequences. In their method, scanpaths are spatially and temporally binned and then recoded to create a sequence of letters that retains fixation location, duration, and sequence information. The two character sequences are compared by maximizing the similarity score computed from a substitution matrix, which in turn provides the score for all letter pair substitutions, and includes a penalty for gaps. Critically, the substitution matrix can encode information about the relationship between specific regions of interest, thus providing the opportunity to include semantic information in the similarity measure.

A major advantage of the ScanMatch method is that it can take into account spatial, temporal and sequential similarity in the comparison of scanpaths. In addition, semantic information can be easily added using the substitution matrix. One disadvantage of this method is that it suffers from the quantization issues inherent to any measure using grids or regions of interest.

6.3 Sample-based Measures

Shepherd et al. (2010) introduced several measures for assessing the similarity of two scanpaths, which are described in the following subsections. For each of the measures, the eye positions are first resampled uniformly in time (at 60Hz), and truncated to the length of the shorter sequence. These measures introduced below are sample-based in the sense that they do not require pre-processing of eye-tracking data into discrete fixation-saccade sequences, as is usually the case in eye movement analyses.

6.3.1 Fixation overlap

Fixation overlap is a measure of the similarity of two scanpaths in space and over time. To this effect, two gaze samples are considered overlapping if they are within a predefined radius. This measure is extremely sensitive to differences in absolute timing between two scanpaths, but is slightly less sensitive to differences in position (due to the use of the radius). Given these sensitivities, it

is reasonable to expect this measure to perform similarly to the ScanMatch measure, which is also sensitive to the spatial and temporal similarities between two scanpaths.

6.3.2 Temporal Correlation

Shepherd et al. (2010) also introduced temporal correlation (see also Hasson, Yang, Vallines, Heeger & Rubin (2008)) as a measure of the similarity between scanpaths. For two scanpaths f and g, the temporal correlation is defined as the average of the correlation between the x-coordinates of f and g and between the y-coordinates of f and g.

This measure is very sensitive to temporal and spatial differences between the two scanpaths. The sensitivity to temporal differences can be advantageous when timing is important, e.g., when the stimuli change over time, such as in videos. The correlation measure is also sensitive to small differences in fixation positions, given that there is no spatial quantization of the fixations. A significant advantage of this method is its use of the straightforward and readily interpretable correlation analysis. This measure is more sensitive to similarities in position than the fixation overlap method, while also taking sequential information into account. However, this strong spatial-temporal sensitivity may be less robust to noisy data than other measures that are grid-based or rely on a radius-based definition of fixation proximity.

6.3.3 Gaze shift

Shepherd and colleagues' (2010) gaze shift measure assesses how similar the saccade times and amplitudes are between two scanpaths. Gaze shift is computed as the correlation between the absolute values of the first derivative of each scanpath and is computed in the same manner as the temporal correlation, but using the first derivative instead of the position.

To smooth the scanpaths and compute their derivatives, each scanpath is convolved with the derivative of a Gaussian filter. Gaze shift is sensitive to the amplitude of the saccade as well as its temporal location, and it reflects how similar two scanpaths are in terms of the sequence of large and small saccades. This captures some aspects of a global viewing strategy, as subjects who produce small saccades within a localized region have very different scanpaths than subjects who produce large saccades within the entire visible area. This is also useful for comparing dynamic stimuli (e.g., video) to assess how subjects respond to temporal changes in the scene.

6.4 Linear Distance

Mannan, Ruddock, and Wooding (1995) analyzed the overall similarity of two scanpaths by computing the linear distances between the fixations of the first scanpath and the nearest neighbour fixations of the second scanpath, as well as the linear distances between the fixations of the second scanpath and the nearest neighbour fixations of the first scanpath. These distances are averaged and normalized against randomly generated scanpath sequences. Mannan et al.'s

method was further developed by Mathot, Cristino, Gilchrist, and Theeuwes (2012).

The major advantage of the linear distance method is that it does not need to be quantized like, for example, the string edit distance method. It simply compares each fixation of one scanpath with the fixations of another scanpath in terms of their spatial similarity. However, by comparing only nearest neighbour fixations in terms of distance, this method ignores sequential information. To address some of these issues, Mannan et al.'s (1995) method was modified by Henderson, Brockmole, Castelhano, & Mack (2007) to enforce a one-to-one mapping between two scanpaths, provided that they have the same length. The results for the two methods are very similar (Foulsham & Underwood, 2008), which is likely due to the fact that Mannan et al. average the distances from the first to the second and from the second to the first scanpath, hence clusters of fixations in one scanpath are averaged out.

6.5 Scasim

Scasim is a scanpath comparison technique developed initially for use in analysing eye movement patterns while reading (von der Malsburg & Vasishth, 2011). It uses similar logic to the string edit and Levenshtein distance metrics, however, it does not require discretization of fixations into regions of interest and it takes fixation duration into account. It compares both the duration and spatial location between fixations by adding or subtracting durations dependent on their distance apart. One unique advantage of Scasim is that it allows the user to specify what sort of cost spatial distance between fixations might have. While by default, this is related to the drop-off of visual acuity from the fovea, this cost can be changed depending on whether spatial distance is more or less important to the hypothesis in question.

6.6 MultiMatch

Recently, Jarodzka, Holmqvist, and Nyström (2010), Dewhurst et al. (2012) and Foulsham et al. (2012) introduced the MultiMatch method for comparing scanpaths. The MultiMatch methods consists of five separate measures that capture the similarity between different characteristics of scanpaths, namely shape, direction, length, position and duration. Computation of each MultiMatch measure begins with scanpath simplification, which involves combining iteratively successive fixations if they are within a given distance or within a given directional threshold of each other. This simplification process aids in reducing the complexity of the scanpaths while preserving the spatial and temporal structure.

Following this simplification, scanpaths are aligned based on their shape using a dynamic programming approach. The alignment is computed by optimizing the vector difference between the scanpaths (note, however, that scanpaths may be aligned on any number of dimensions in MultiMatch). This alignment reduces the comparison's sensitivity to small temporal or spatial temporal variations and allows the algorithm to find the best possible match between the pair of

scanpaths. All subsequent similarity measures are computed on these simplified, aligned scanpaths.

6.6.1 Vector Similarity

The MultiMatch vector similarity measure is computed as the vector difference between aligned saccade pairs, normalized by the screen diagonal and averaged over scanpaths. This measure is sensitive to spatial differences in fixation positions without relying on pre-defined quantization. It is a measure of the overall similarity in shape between two fixation-saccade sequences.

6.6.2 Length Similarity

MultiMatch length similarity is computed as the absolute difference in the amplitude of aligned saccade vectors, normalized by the screen diagonal and averaged over scanpaths. This measure is sensitive to saccade amplitude only, not to the direction, location or the duration of the fixations.

6.6.3 Direction Similarity

MultiMatch direction similarity is computed as the angular difference between aligned saccades, normalized by π and averaged over scanpaths. This measure is sensitive to saccade direction only, but not to amplitude or absolute fixation location.

6.6.4 Position Similarity

MultiMatch position similarity is computed as the Euclidean distances between aligned fixations, normalized by the screen diagonal and averaged over scanpaths. This measure is sensitive to both saccade amplitudes and directions.

6.6.5 Duration Similarity

MultiMatch duration similarity is computed as the absolute difference in fixation durations of aligned fixations, normalized by the maximum duration and averaged over scanpaths. This measure is insensitive to fixation position or saccade amplitude.

The main advantage of the MultiMatch method is that it provides several measures to choose from for assessing scanpath similarity, and each measure on its own captures a unique component of scanpath similarity. Given the multiplicity of measures, it remains, however, difficult to assess which measure, or which set of measures, is most applicable in a given scenario. Furthermore, because each scanpath is initially simplified it is also not clear how robust each measure is to scanpath variations.

6.7 Cross-Recurrence Quantification Analysis

The recurrence quantification analysis introduced in section 6 can be generalized for the comparison of scanpaths in a cross-recurrence analysis. To this effect, we have generalized the RQA measures for the comparison of scanpaths, and we now introduce these generalized measures.

Consider two fixation sequences f and g that have the same lengths. For sequences of unequal length, the longer sequence is truncated. Within these sequences, two fixations f_i and g_j are cross-recurrent if they match or are close together, i.e., if their distance is below a given threshold. One can define cross-recurrence c_{ij} as

$$c_{ij} = \begin{cases} 1, \wedge\ d(f_i, g_j) \leq \rho \\ 0, \wedge\ otherwise \end{cases} \qquad (1)$$

where d is distance metric and ρ is a given radius, as in the definition of recurrence. Cross-recurrence can be represented in a cross-recurrence diagram, which plots cross-recurrences of the fixation sequences over all possible time lags. If fixations f_i and g_j are recurrent (i.e. if $c_{ij} = 1$), then a dot is plotted at position i,j.

This is illustrated in Figure 14, which shows an example from a study in which participants looked at many images, each for about 10 seconds, and later saw the same images again among many new ones. The black lines in Figure 14a indicate the first scanpath, the red lines the second scanpath, and the black circles indicate cross-recurrences, i.e., fixations of one scanpath that were close to fixations of the other scanpath.
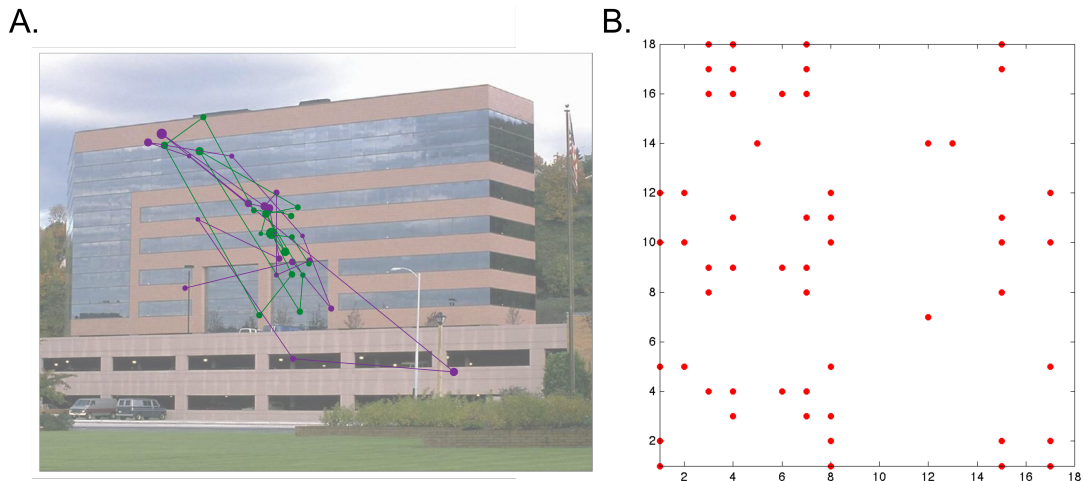
A.

B.

Figure 14: a) Image of building with two scanpaths (purple and orange) produced by the same participant. b) Corresponding cross-recurrence diagram.

The resulting cross-recurrence diagram is shown in Figure 14b, in which the fixations of the first scanpath are shown along the x-axis and the fixations of the second scanpath along the y-axis. In contrast to the recurrence diagram shown in Figure 8, the cross-recurrence diagram is not symmetric. Second, there is no line

of incidence in the cross-recurrence diagram. For these reasons, the RQA measures have to be generalized for the application in cross-recurrence analysis.

6.7.1 Cross-Recurrence

The cross-recurrence measure of two fixation sequences represents the percentage of cross-recurrent fixations, i.e. the percentage of fixations that match between the two fixation sequences. The more similar two fixation sequences, the higher the number of cross-recurrent points on the plot. It is invariant to differences in the order of the fixations as fixations are considered recurrent only if they overlap in position. Given that cross-recurrence quantifies similarity in position only, it is most similar to the linear distance measure and the MultiMatch position measure.

6.7.2 Determinism

The determinism measure encodes the percentage of cross-recurrent points that form diagonal lines in the cross-recurrence plot and represents the percentage of fixation trajectories common to both fixation sequences. That is, determinism quantifies the overlap of a specific sequence of fixations, preserving their sequential information. An advantage of this measure is that it provides unique information about the type of similarity between two scanpaths. Although two scanpaths may be quite dissimilar in their overall shape or fixation positions, this measure can show whether certain smaller sequences of those scanpaths are shared.

6.7.3 Laminarity

The laminarity measure is a measure of repeated fixations on a particular region that are common to both scanpaths. Laminarity is closely related to determinism. If both laminarity and determinism are high, then in both scanpaths fixations tend to cluster on one or a few particular locations and remain there across several fixations. If laminarity is high, but determinism is low, then it quantifies the number of locations that were fixated in detail in one of the fixation sequences, but only fixated briefly in the other fixation sequence. It is a measure of the clustering of fixations across two sequences.

6.7.4 Center of Recurrence Mass

The center of recurrence mass (corm) is defined as the distance of the center of gravity of recurrences from the main diagonal in a recurrence plot. The corm measure indicates the dominant lag of cross-recurrences. Small corm values indicate that the same fixations in both fixation sequences tend to occur close in time, whereas large corm values indicate that cross-recurrences tend to occur with either a large positive or negative lag. This is a measure of whether one scanpath leads (with positive lag) or follows (with negative lag) its paired

scanpath. Their overall similarity in shape or position may be different, but offset, such that one sequence proceeds in a particular trajectory, and the other follows the same trajectory only later on in time (e.g., a few fixations later). If there is no specific prediction about whether one scanpath leads or follows the other, the absolute value of the corm value can be used rather than averaging over positive and negative values.

In summary, cross-recurrence has been shown to be a natural extension of recurrence. While recurrence is used to characterize individual scanpaths, cross-recurrence is used to characterize the similarity of two different scanpaths. One major advantage of recurrence and cross-recurrence analysis lies in the fact that the same measures, recurrence, determinism, laminarity and corm can be applied to both situations. As a note of caution, however, it must be emphasized that the measures have to be interpreted differently in the two cases.

6.8 Summary

In this section, we have reviewed several common methods for assessing the similarity between scanpaths, and by extension, differences between scanpaths. These various methods all have their strengths and weaknesses, but all provide unique information regarding the similarity between two scanpath sequences. Comparing scanpaths is useful, as like RQA (described in Section 6), they preserve and quantify the temporal characteristics of eye movement behavior. These methods can be used for comparing the similarities and differences in eye movements between or among different observers.

7. General Summary

In the preceding sections, we first reviewed traditional measures for characterizing eye movements, starting with basic fixation and saccade measures. We then considered spatial eye movement analyses with a focus on heat maps and area of interest analyses. Finally, we examined popular temporal analyses of eye movements. Taken together, the reviews showed that the spatial analyses were not capturing the dynamic characteristics of eye movements, whereas the temporal analyses were applicable only in restricted circumstances.

We introduced recurrence quantification analysis (RQA) as a method for measuring the dynamic characteristics of eye movements. Although the analysis may appear to some as rather complicated, RQA can readily and accurately be conceptualized as an analysis of the temporal pattern of refixations. Critically, the RQA measures we introduced (recurrence, determinism, laminarity, and corm) capture important and interpretable aspects of this pattern. We showed for instance that RQA is suitable for discriminating eye movement patterns independent of the spatial structure of stimuli by focusing on temporal aspects exclusively. For this reason, it is ideally suited for the analysis of eye movements in more complex and dynamic situations.

Finally, we discussed current approaches to the comparison of scanpaths (edit distance, sample-based measures, linear distance, Scasim, ScanMatch, and

MultiMatch), and showed that a simple generalization of RQA, that is, cross-RQA is well suited for the spatial comparison of scanpaths (i.e., how closely the individual fixations of two scanpaths overlap) while also capturing aspects of their temporal nature (e.g., their sequential information). Collectively then, RQA and cross-RQA, represents two powerful analysis techniques for future studies of eye movement behaviour, both within controlled laboratory settings and less controlled, more complex, natural environments. More generally. as eye movement analysis moves into the real world, techniques that capture the temporal nature of eye movements will no doubt prove useful in quantifying this more complex behavior.

8. Suggested Readings

Anderson, N. C., Bischof, W. F., Laidlaw, K. E., Risko, E. F., & Kingstone, A. (2013). Recurrence quantification analysis of eye movements. Behavior research methods, 45(3), 842–856.

This paper introduces recurrence quantification analysis for the analysis of eye movements. Much of the material in section 6 is based on this paper.

Anderson, N. C., Anderson, F., Kingston, A., & Bischof, W. F. (2014). A comparison of scanpath comparison methods. Behavior Research Methods, DOI 10.3758/s13428-014-0550-3.

This paper reviews and compares most of the recent scanpath comparison methods. Much of the material in section 7 is based on this paper.

Dale, R., Warlaumont, A. S., & Richardson, D. C. (2011b). Nominal cross recurrence as a generalized lag sequential analysis for behavioral streams. International Journal of Bifurcation and Chaos, 21, 1153–1161. doi:10.1142/S0218127411028970.

Recurrence analysis can be generalized to categorical data. This paper introduces categorical cross-recurrence analysis for analyzing the coordination of gaze patterns between individuals (see section 6).

Holmqvist, K., Nyström, M., Andersson, R., Dewhurst, R., Jarodzka, H., & van de Weijer, J. (2011). Eye tracking: A comprehensive guide to methods and measures. Oxford, U.K.: Oxford University Press.

This handbook is a comprehensive guide to the methods and measures for eye tracking. We recommend that researchers in the area of eye movement analysis consult this handbook. A new edition is scheduled for 2016 or 2017.

Marwan, N., & Kurths, J. (2002). Nonlinear analysis of bivariate data with cross recurrence plots. Physics Letters A, 302, 299–307.

This paper introduces recurrence analysis as a tool for describing complex dynamic systems. The works of Dale et al. (2011b) and of Anderson et al. (2013) are applications of, and extensions to this paper.

9. Abbreviations

The following abbreviations are used in this chapter:

AOI	Area of interest
CORM	Center of recurrence mass
DET	Determinism
LAM	Laminarity
LDA	Linear discriminant analysis
REC	Recurrence
RQA	Recurrence quantification analysis

10. Questions to Students

a. What are the fundamental differences between grid-based and heat map methods for the spatial analysis of eye movements?
b. What are the characteristics of stimuli for which area-of-interest analyses are useful, and when are they less useful or not at all useful?
c. What are the fundamental advantages of recurrence analysis over the other spatial and temporal methods presented in this chapter?
d. What aspect of the recurrence patterns are captured by the measures recurrence, determinism, laminarity and corm?

11. Acknowledgements

12. Bibliography

Anderson, N. C., Bischof, W. F., Laidlaw, K. E., Risko, E. F., & Kingstone, A. (2013). Recurrence quantification analysis of eye movements. *Behavior Research Methods, 45*(3), 842–856.

Anderson, N. C., Anderson, F., Kingston, A., & Bischof, W. F. (2014). A comparison of scanpath comparison methods. *Behavior Research Methods*, DOI 10.3758/s13428-014-0550-3.

Anderson, N. C., Donk, M., & Meeter, M. (2016). The influence of a scene preview on eye movement behavior in natural scenes. Psychonomic Bulletin &

Review, DOI 10.3758/s13423-016-1035-4.

Birmingham, E., Bischof, W. F., & Kingstone, A. (2008a). Gaze selection in complex social scenes. *Visual Cognition, 16*, 341– 355. doi:10.1080/13506280701434532

Birmingham, E., Bischof, W. F., & Kingstone, A. (2008b). Social attention and real-world scenes: The roles of action, competition and social content. *Quarterly Journal of Experimental Psychology, 61*, 986–998. doi:10.1080/17470210701410375.

Birmingham, E., Bischof, W. F., & Kingstone, A. (2009). Saliency does not account for fixations to eyes within social scenes. *Vision Research, 49*, 2992-3000.

Box, G. E. P., Jenkins, G. M., & Reinsel, G. C. (2008). *Time series analysis* (4th ed.). New York, NY: Wiley.

Brandt, S. A., & Stark, L. W. (1997). Spontaneous eye movements during visual imagery reflect the content of the visual scene. *Journal of Cognitive Neuroscience, 9*, 27–38. doi:10.1162/jocn.1997.9.1.27

Broadbent, D. E. (1971). *Decision and stress.* London: Academic Press.

Bunke, H. (1992). Recent advances in string matching. In H. Bunke (Ed.), *Advances in structural and syntactic pattern recognition* (pp. 107–116). Singapore: World Scientific.

Caldara, R., & Miellet, S. (2011). iMap: A Novel Method for Statistical Fixation Mapping of Eye Movement data. *Behavior Research Methods, 43*(3), 864-878.

Chica, A. B., Klein, R. M., Rafal, R. D., & Hopfinger, J. B. (2010). Endogenous saccade preparation does not produce inhibition of return: Failure to replicate Rafal, Calabresi, Brennan, & Sciolto (1989). *Journal of Experimental Psychology: Human Perception and Performance*, *36*(5), 1193.

Cheng, J. T., Tracy, J. L., Foulsham, T., Kingstone, A., & Henrich, J. (2013). Two ways to the top: Evidence that dominance and prestige are distinct yet viable avenues to social rank and influence. *Journal of Personality and Social Psychology*, *104*(1), 103.

Cherubini, M., Nüssli, M.-A., & Dillenbourg, P. (2010). This is it! Indicating and looking in collaborative work at distance. *Journal of Eye Movement Research, 3*(3), 1–20.

Chuk, T., Chan, A. B., & Hsiao, J. H. (2014). Understanding eye movements in face recognition using hidden Markov models. *Journal of Vision, 14*(11):8, 1-14.

Coutrot, A., & Guyader, N. (2014). How saliency, faces, and sound influence gaze in dynamic social scenes. *Journal of Vision*, *14*(8), 5.

Cristino, F., Mathôt, S., Theeuwes, J., & Gilchrist, I. D. (2010). ScanMatch: A novel method for comparing fixation sequences. *Behavior Research Methods, 42*, 692–700. doi:10.3758/BRM.42.3.692

Dale, R., Kirkham, N. Z., & Richardson, D. C. (2011a). The dynamics of reference and shared visual attention. *Frontiers in Psychology, 2*(355), 1–11. doi:10.3389/fpsyg.2011.00355

Dale, R., Warlaumont, A. S., & Richardson, D. C. (2011b). Nominal cross recurrence as a generalized lag sequential analysis for behavioral streams. *International Journal of Bifurcation and Chaos, 21*, 1153–1161. doi:10.1142/S0218127411028970

Efron, B., & Tibshirani, R. J. (1993). *An introduction to the bootstrap*. New York, NY: Chapman & Hall.

Findlay, J. M., & Gilchrist, I. D. (2003). *Active vision: The psychology of looking and seeing*. Oxford: Oxford University Press.

Foulsham, T., Cheng, J. T., Tracy, J. L., Henrich, J., & Kingstone, A. (2010). Gaze allocation in a dynamic situation: Effects of social status and speaking. *Cognition*, *117*(3), 319-331.

Foulsham, T., & Kingstone, A. (2013). Fixation-dependent memory for natural scenes: An experimental test of scanpath theory. *Journal of Experimental Psychology: General*. doi:10.1037/a0028227

Foulsham, T., & Sanderson, L. A. (2013). Look who's talking? Sound changes gaze behaviour in a dynamic social scene. *Visual Cognition*, *21*(7), 922-944.

Foulsham, T., & Underwood, G. (2008). What can saliency models predict about eye movements? Spatial and sequential aspects of fixations during encoding and recognition. *Journal of Vision, 8*(2)(6), 1–17. doi:10.1167/8.2.6

Friston, K. J., Jezzard, P. and Turner, R. (1994). Analysis of Functional MRI Time-Series, *Human Brain Mapping, 1*, 153-171.

Friston, K. J., Holmes, A. P., Poline, J.-B., Grasby, P. J., Williams, S. C. R., Frackowiak, R. S. J. and Turner, R. (1995). Analysis of fMRI Time-Series Revisited, *Neuroimage, 2*, 45-53.

Harding, G., & Bloj, M. (2010). Real and predicted influence of image manipulations on eye movements during scene recognition. *Journal of Vision, 10*(2)(8), 1–17. doi:10.1167/10.2.8.

Hayes, T. R., Petrov, A. A., & Sederberg, P. B. (2011). A novel method for analyzing sequential eye movements reveals strategic influence on Raven's Advanced Progressive Matrices. *Journal of Vision, 11*, 1-11

Hasson, U., Yang, E., Vallines, I., Heeger, D. and Rubin, N. (2008). A hierarchy of

temporal receptive windows in human cortex. *Journal of Neuroscience, 28*, 2539-50.

Henderson, J. M., Brockmole, J. R., Castelhano, M. S., & Mack, M. L. (2007). Visual saliency does not account for eye movements during visual search in real-world scenes. In R. van Gompel, M. Fischer, W. Murray, & R. W. Hill (Eds.), *Eye movements: A window on mind and brain* (pp. 537–562). Amsterdam, The Netherlands: Elsevier.

Holmqvist, K., Nyström, M., Andersson, R., Dewhurst, R., Jarodzka, H., & van de Weijer, J. (2011). *Eye tracking: A comprehensive guide to methods and measures.* Oxford, U.K.: Oxford University Press.

Hunt, A. R., Chapman, C. S., & Kingstone, A. (2008). Taking a long look at action and time perception. *Journal of Experimental Psychology: Human Perception and Performance*, *34*(1), 125.

Igarashi, T., Moscovich, T., & Hughes, J.F. (2005). Spatial keyframing for performance-driven animatoin. *Proceedings of the 2005 ACM SIGGRAPH/Eurographics symposium on computer animation*, pp. 107-115, New York: ACM

Itti, L., & Koch, C. (2000). A saliency-based search mechanism for overt and covert shifts of visual attention. *Vision Research, 40*, 1489–1506.

Jarodzka, H., Holmqvist, K., & Nyström, M. (2010). A vector-based, multidimensional scanpath similarity measure. Paper presented at the *Proceedings of the 2010 Symposium on Eye-Tracking Research and Applications.*

Josephson, S. & Holmes, M. E. (2002). Visual attention to repeated images on the World-Wide Web: Another look at scanpath theory. *Behavior Research Methods, Instruments, & Computers, 34*(4), 539-548.

Judd, T., Ehinger, K., Durand, F., & Torralba, A. (2009). Learning to predict where humans look. Paper presented at the *IEEE International Conference on Computer Vision* (ICCV), Kyoto, Japan.

Kingstone, A. (1992). Combining expectancies. *The Quarterly Journal of Experimental Psychology*, *44A*, 69-104.

Kingstone, A., Smilek, D., & Eastwood, J. D. (2008). Cognitive ethology: A new approach for studying human cognition. *British Journal of Psychology*, *99*(3), 317-340.

Krassner, M., Patera, W., Bulling, A. (2014) Pupil: An Open Source Platform for Pervasive Eye Tracking and Mobile Gaze-based Interaction. *Computer Vision and Pattern Recognition*, arXiv:1405.0006

Le Meur, O. and Baccino, T. (2013). Methods for comparing scanpaths and saliency maps: strengths and weaknesses. *Behavior Research Methods, 45*,

p. 251-266

Levenshtein, V. I. (1966). Binary codes capable of correcting deletions, insertions, and reversals. *Soviet Physics – Doklady, 10*, 707–710.

Levy, J., Foulsham, T., & Kingstone, A. (2012). Monsters are people too. *Biology letters, Feb 23;9*(1):20120850. doi: 10.1098/rsbl.2012.0850.

Mannan, S., Ruddock, K. H., & Wooding, D. S. (1995). Automatic control of saccadic eye movements made in visual inspection of briefly presented 2-D images. *Spatial Vision, 9*, 363–386. doi:10.1163/156856895X00052

Marwan, N., & Kurths, J. (2002). Nonlinear analysis of bivariate data with cross recurrence plots. *Physics Letters A, 302*, 299–307.

Mathôt, S., Cristino, F., Gilchrist, I. D., & Theeuwes, J. (2012). A simple way to estimate similarity between pairs of eye movement sequences. *Journal of Eye Movement Research, 5*(1).

Needleman, S. B., & Wunsch, C. D. (1970). A general method applicable to the search for similarities in the amino acid sequence of two proteins. *Journal of Molecular Biology, 48*(3), 443–453.

Noton, D., & Stark, L. (1971). Scanpaths in eye movements during pattern perception. *Science, 171*, 308–311. doi:10.1126/science.171.3968.308

Olk, B., & Kingstone, A. (2003). Why are antisaccades slower than prosaccades? A novel finding using a new paradigm. *Neuroreport*, *14*, 151-155.

Pellecchia, G. L., & Shockley, K. (2005). Cross recurrence quantifica- tion of interpersonal postural activity. In M. A. Riley & G. C. Van Orden (Eds.), *Tutorials in contemporary nonlinear methods for the behavioral sciences* (pp. 95–141). Washington, DC: National Science Foundation, Directorate for Social, Behavioral & Economic Sciences. Retrieved August 25, 2011, from www.nsf.gov/ sbe/bcs/pac/nmbs/nmbs.jsp

Posner, M.I. (1978). *Chronometric explorations of mind.* Hillsdale, NJ: Erlbaum.

Richardson, D. C., & Dale, R. (2005). Looking to understand: The coupling between speakers' and listeners' eye movements and its relationship to discourse comprehension. *Cognitive Science, 29*, 1045–1060.

Richardson, D. C., Dale, R., & Tomlinson, J. M. (2009). Conversation, gaze coordination, and beliefs about visual context. *Cognitive Science, 33*, 1468–1482. doi:10.1111/j.1551-6709.2009.01057.x

Riley, M. A., & Clark, S. (2003). Recurrence analysis of human postural sway during the sensory organization test. *Neuroscience Letters, 342*, 45–48.

Risko, E. F., & Kingstone, A. (2011). Eyes wide shut: implied social presence, eye tracking and attention. *Attention, Perception, & Psychophysics*, *73*(2), 291-

296.

Shepherd, S. V., Steckenfinger, S. A., Hasson, U., & Ghasanfar, A. A. (2010). Human–monkey gaze correlations reveal convergent and divergent patterns of movie viewing. *Current Biology, 20*, 649–656.

Shockley, K., & Turvey, M. T. (2005). Encoding and retrieval during bimanual rhythmic coordination. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31*(5), 980.

Shockley, K., Santana, M.-V., & Fowler, C. A. (2003). Mutual interper- sonal postural constraints are involved in cooperative conversation. *Journal of Experimental Psychology: Human Perception and Performance, 29*(2), 326.

Shockley, K., Baker, A. A., Richardson, M. J., & Fowler, C. A. (2007). Articulatory constraints on interpersonal postural coordination. *Journal of Experimental Psychology: Human Perception and Performance, 33*(1), 201.

Shockley, K., Richardson, D. C., & Dale, R. (2009). Conversation and coordinative structures. Topics in Cognitive Science, 1, 305–319. Shockley, K., Santana, M.-V., & Fowler, C. A. (2003). Mutual interpersonal postural constraints are involved in cooperative conversation. *Journal of Experimental Psychology. Human Perception and Performance, 29*, 326–332. doi:10.1037/0096-1523.29.2.326

Soto-Faraco, S., Morein-Zamir, S., & Kingstone, A. (2005). On audiovisual spatial synergy: The fragility of the phenomenon. *Perception & psychophysics*, *67*(3), 444-457.

Stark, L. W. & Ellis, S. R. (1981) Scanpaths revisited: cognitive models direct active looking. In *Eye movements: cognition and visual perception* (pp. 193-226), D. F. Fisher, et al., (Eds.), Hillsdale, NJ: Lawrence Erlbaum Associates.

Tatler, B. W. (2007). The central fixation bias in scene viewing: Selecting an optimal viewing position independently of motor biases and image feature distributions. *Journal of vision*, *7*(14), 4-4.

Torralba, A., Oliva, A., Castelhano, M. S., & Henderson, J. M. (2006). Contextual Guidance of Eye Movements and Attention in Real-WorldScenes: The Role of Global Features in Object Search. Psychological Review, 113(4), 766-786.

Underwood, G., Foulsham, T., & Humphrey, K. (2009). Saliency and scan patterns in the inspection of real-world scenes: Eye movements during encoding and recognition. *Visual Cognition, 17*, 812–834.

Vig, E., Dorr, M., & Cox, D. (2014). Large-scale optimization of hierarchical features for saliency prediction in natural images. The IEEE Conference on Pattern Recognition, Columbus Ohio, June 24-27.

von der Malsburg, T., & Vasishth, S. (2011). What is the scanpath signature of syntactic reanalysis? *Journal of Memory and Language, 65*(2), 109-127.

Wagner, R. A., & Fischer, M. J. (1974). The string-to-string correction problem. *Journal of the Association for Computing Machinery, 21*, 168–178.

Webber, C. L., Jr., & Zbilut, P. J. (2005). Recurrence quantification analysis of nonlinear dynamical systems. In M. A. Riley & G. C. Van Orden (Eds.), *Tutorials in contemporary nonlinear methods for the behavioral sciences* (pp. 26–94). Washington, DC: National Science Foundation, Directorate for Social, Behavioral & Economic Sciences. Retrieved August 25, 2011, from www.nsf.gov/sbe/bcs/ pac/nmbs/nmbs.jsp

Yarbus, A. (1967). *Eye movements and vision*. New York, NY: Plenum Press.