

Temporal dynamics of eye movements are related to differences in scene complexity and clutter

David W.-L. Wu

Department of Psychology, University of British Columbia,
Vancouver, British Columbia, Canada



Nicola C. Anderson

Department of Cognitive Psychology, Vrije Universiteit
Amsterdam, Amsterdam, The Netherlands



Walter F. Bischof

Department of Computing Science, University of Alberta,
Edmonton, Alberta, Canada



Alan Kingstone

Department of Psychology, University of British Columbia,
Vancouver, British Columbia, Canada



Recent research has begun to explore not just the spatial distribution of eye fixations but also the temporal dynamics of how we look at the world. In this investigation, we assess how scene characteristics contribute to these fixation dynamics. In a free-viewing task, participants viewed three scene types: fractal, landscape, and social scenes. We used a relatively new method, recurrence quantification analysis (RQA), to quantify eye movement dynamics. RQA revealed that eye movement dynamics were dependent on the scene type viewed. To understand the underlying cause for these differences we applied a technique known as fractal analysis and discovered that complexity and clutter are two scene characteristics that affect fixation dynamics, but only in scenes with meaningful content. Critically, scene primitives—revealed by saliency analysis—had no impact on performance. In addition, we explored how RQA differs from the first half of the trial to the second half, as well as the potential to investigate the precision of fixation targeting by changing RQA radius values. Collectively, our results suggest that eye movement dynamics result from top-down viewing strategies that vary according to the meaning of a scene and its associated visual complexity and clutter.

world scenes in eye movement tracking studies. These investigations often involve the use of different types of scenes, such as interior and exterior scenes (e.g., Foulsham & Underwood, 2008; Tatler & Vincent, 2008), but rarely test for any differences across scene content. Studies that do explore how scene content influences attentional guidance rely primarily on quantifying low-level scene characteristics, typically visual salience (for a review, see Tatler, Hayhoe, Land, & Ballard, 2011) or global scene properties such as gist (Oliva, 2005), to predict the pattern of eye movement fixations (e.g., Itti & Koch, 1999). While the spatial distribution of fixations has been quantified extensively by these approaches, the temporal aspects of eye movement dynamics, and their relation to scene content, have not enjoyed a similar degree of empirical examination.

Scanpath measures that do take the temporal dynamics of eye movements into account, such as string-edit distance analysis (Bunke, 1992; Levenshtein, 1966; Underwood, Foulsham, & Humphrey, 2009) or the MultiMatch method (Dewhurst et al., 2012), have been used effectively to compare two scanpaths, but they are limited in their ability to provide quantitative information about any one scanpath and thus are constrained in their ability to relate the temporal dynamics of a given scanpath to the content of a visual scene. Several recent scanpath quantification methods have been developed to address this problem using state-of-the-art analysis techniques such as coupled hidden Markov models (e.g., Cagli, Coraggio, Napo-

Introduction

The desire to understand how attention is guided in the natural complex world has increased the use of real-

Citation: Wu, D. W.-L., Anderson, N. C., Bischof, W. F., & Kingstone, A. (2014). Temporal dynamics of eye movements are related to differences in scene complexity and clutter. *Journal of Vision*, 14(9):8, 1–14, <http://www.journalofvision.org/content/14/9/8>, doi:10.1167/14.9.8.

letano, & Boccignone, 2008), reinforcement learning algorithms (Hayes, Petrov, & Sederberg, 2011), modeling using spatial point processes (e.g., Barthelmé, Trukenbrod, Engbert, & Wichmann, 2013), and related computational models (e.g., Wang, Freeman, Merriam, Hasson, & Heeger, 2012). Methods for dynamic stimuli are also beginning to be developed (Açık, Bartel, & König, 2014; Schütz, Lossin, & Kerzel, 2013). While undoubtedly useful, these methods may not be easily accessible to researchers who are not familiar with the advanced modeling or statistical techniques required.

Here, we use an analysis technique known as recurrence quantification analysis (RQA). RQA quantifies the temporal dynamics of fixations that compose a single scanpath. More specifically, it is aimed at characterizing sequences of fixations by analyzing the pattern of recurrences (i.e., the repeated fixations of the same image positions). The recurrence patterns are quantified by several measures, which are described in more detail below. The main advantage of RQA over other recent temporal analyses is that RQA analysis is readily and directly interpretable in terms of fixations and saccades, and, critically, these measures can be compared with the content of a visual scene as well as generalized across different images (Anderson, Bischof, Laidlaw, Risko, & Kingstone, 2013).

In RQA, the pattern of fixation recurrences is quantified by several measures, which provide readily interpretable measures of temporal dynamics. The following four measures have been found to be most useful: recurrence, which is the percentage of fixations that are refixations of previous positions; determinism, which is the percentage of recurrent fixations that form sequences or repeated trajectories; laminarity, which is the percentage of recurrent fixations that indicate that a particular area is being repeatedly fixated; and the center of recurrent mass (CORM), a measure of the temporal distribution of recurrences where small values indicate recurrent fixations occurring in close temporal proximity and large values indicate recurrent fixations occurring after relatively large temporal intervals. Anderson et al. (2013) found RQA measures to be sensitive to differences in general scene types (e.g., landscapes vs. interiors). They found determinism and laminarity measures to be higher for interior scenes than for exterior and landscape scenes, indicating that scene content has an effect on the temporal dynamics of eye movements. These differences were hypothesized to be due to higher levels of clutter in interior scenes compared with landscapes, but this attribution, and its implications, was never verified experimentally. The purpose of the current investigation is to expand on this proposition and examine how scene characteristics influence the temporal dynamics of fixation patterns. This was achieved by the systematic application of two

research analyses that yielded several key predictions. These are presented in turn below.

First, we investigated whether general categories of scenes (“scene types”) influence RQA measures by having participants view scenes that lacked meaningful content (fractal scenes) or possessed meaningful content (social stimuli and landscapes). Fractal images are complex images at multiple scales, with few or no easily identifiable meaningful parts, and appear abstract and cluttered. In contrast, landscape and social scene images have a lower complexity and possess readily identifiable and meaningful parts, such as clouds, trees, and people.

Second, to expand on Anderson et al.’s (2013) hypothesis that fixation dynamics are affected by differences in the clutter and complexity of scenes, we applied a fractal analysis (Mandelbrot, 1982) that enabled us to objectively quantify the clutter and complexity of the scenes. Specifically, we used the fractal dimension and lacunarity measures of the images and compared those values with the recurrence quantification of participants’ scanpaths. The fractal dimension quantifies the overall complexity of a scene, while the lacunarity measure quantifies the heterogeneity or “gappiness” of an image.

Several predictions flow from the literature as to how the temporal dynamics of eye movements may differ for social, landscape, and fractal scenes. Given previous research suggesting that scene meaning can impact the spatial characteristics of scanpaths (Underwood et al., 2009), we predicted that scene meaning would influence the temporal characteristics of scanpaths. Additionally, Anderson et al. (2013) speculated that scenes with high clutter and complexity (e.g., fractal scenes) lead to higher determinism and laminarity scores.

Finally, based on the vast evidence indicating that social stimuli have a powerful draw on the allocation of attention, in particular the face and eye regions of people (e.g., Birmingham, Bischof, & Kingstone, 2009; Birmingham & Kingstone, 2009), we reasoned that recurrence would be especially high for social scenes.

Methods

Participants

Fifty students from the University of British Columbia were given course credit or paid \$5 to participate in the present study.

Stimuli

Thirty unique images featuring fractals, landscape scenes, and social scenes (10 of each type) were



Figure 1. An example of the fractal and landscape scene types used. Examples of social scenes used can be found in Birmingham et al. (2008).

presented.¹ Fractals and landscapes were from Foulsham and Kingstone (2010) (see Figure 1 for examples), and social scenes were from Birmingham, Bischof, and Kingstone (2008). The scenes were 1024×768 pixels and corresponded to a horizontal visual angle of approximately 42° and a vertical visual angle of approximately 33° .

Apparatus

An Eyelink 1000 (SR Research, Ottawa, ON, Canada) eye-tracking system recorded participants' eye movements at 1000 Hz. Stimuli were presented to participants on a 23-in. monitor. Scenes and eye movements were also presented to the experimenter on an adjacent monitor located in the testing room.

Procedure

Participants were seated 60 cm from the computer monitor, with their heads positioned in a chin rest. Participants were told to view each photograph as they would normally look at scenes. Images were presented for 10 s. Participants viewed 30 randomly ordered images.

Data handling

Fixation sequences from each participant looking at each image were run through an RQA as outlined in detail in Anderson et al. (2013). Briefly, RQA compares each sequence of fixations in a scanpath with itself in a manner similar to autocorrelation. When a fixation is within a given radius of any other fixation in the

scanpath, that fixation is considered recurrent. A two-dimensional plot is created for each scanpath that indicates the number of recurrent fixations in a given sequence. From this plot, the measures recurrence, determinism, laminarity, and CORM are extracted. Recurrence is the percentage of fixations that are refixations of previous positions. Determinism is the percentage of recurrent fixations that form sequences or repeated trajectories. For example, if a particular sequence of fixations is repeated later in the trial, this sequence is considered deterministic. Laminarity is the percentage of recurrent fixations that indicate that a particular area is being repeatedly fixated. For example, if a person looked once at a particular region, then later in the trial looked at that same region over several fixations in a row, that sequence would be considered laminant. CORM is a measure of the temporal distribution of recurrences, where larger values indicate recurrent fixations occurring more distributed in time and smaller values indicate that most refixations are occurring closer together in time. In the present work, fixations were considered recurrent if they landed within a 64-pixel radius of another fixation. This corresponds to 2.6° visual angle and was chosen to reflect the area of foveal vision that is centered at fixation.

Results and discussion

Scene type

A one-way multivariate analysis of variance (MANOVA) was conducted to assess how scene type (landscape, fractal, social) influenced the four RQA

Measure	Landscape	Fractal	Social
Recurrence	10.63 (0.48)	13.32 (0.72)	13.41 (0.43)
Determinism	29.63 (0.86)	35.49 (1.03)	43.24 (0.71)
Laminarity	25.53 (0.83)	32.98 (0.97)	44.24 (0.69)
CORM	32.06 (0.41)	31.54 (0.50)	30.97 (0.33)
Saccade amplitude	4.93 (0.06)	4.86 (0.07)	4.60 (0.05)
Fixation duration	337.83 (11.03)	398.77 (24.47)	316.49 (12.92)

Table 1. Means for each RQA measure based on the scene type, with standard error of the mean (SEM) in parentheses. Traditional fixation measures of saccade amplitude and fixation duration (ms) are also included. $N = 50$.

variables (recurrence, determinism, laminarity, CORM). The MANOVA produced a significant effect among the three scene types on the four RQA measures, Wilk's $\lambda = 0.23$, $F(8, 190) = 25.32$, $p < 0.001$. The effect size was strong, $\eta^2 = 0.52$. See Table 1 for the descriptive statistics of each RQA measure for the three scene types.

To follow up the significant MANOVA, Bonferroni-corrected analyses of variance (ANOVA) were conducted with each dependent variable. Significant effects of scene type were found for recurrence, $F(2, 98) = 5.01$, $p = 0.009$, $\eta^2_p = 0.093$; determinism, $F(2, 98) = 53.65$, $p < 0.001$, $\eta^2_p = 0.52$; and laminarity, $F(2, 98) = 99.85$, $p < 0.001$, $\eta^2_p = 0.67$; but not for CORM, $F(2, 98) = 1.17$, $p = 0.32$, $\eta^2_p = 0.023$.

Posthoc Tukey's honest significance difference (HSD) tests with a significance level of $\alpha = 0.05$ were used to follow up the significant ANOVA. For recurrence, landscape scenes were different from fractal and social scenes, but the latter two were not different from each other. For determinism, all three scene types were significantly different from each other, with landscapes showing the lowest determinism and social scenes showing the highest. Laminarity had the same pattern of results, with landscapes having the lowest values and social scenes having the highest values. See Figure 2.

We also performed an analysis of traditional measures and found significant effects of scene type for saccade amplitude, $F(2, 98) = 7.33$, $p = 0.001$, $\eta^2 = 0.13$, and for fixation duration, $F(2, 98) = 4.46$, $p = 0.01$, $\eta^2 = 0.083$. Posthoc Tukey's HSD tests showed that, for saccade amplitude, social scenes were different from landscape and fractal scenes, but the latter two were not different from each other. For fixation duration, fractal scenes were different from landscape and social scenes, but the latter two were not different from each other.

Our results suggest that temporal dynamics of eye movements are dependent on the type of scene being viewed. Specifically, social scenes had the highest levels of determinism and laminarity, while landscapes had the lowest levels. Fractal scenes were situated between these extremes.

A high determinism value indicates repetition in gaze fixation sequences; that is, after viewing one particular region in the scene, participants tend to visit another particular region. For example, after visiting the eye region of one person, participants are likely to visit the eye region of another person, very possibly to assess the nature of the social interaction between the two individuals (Birmingham et al., 2009; Dalrymple et al., 2013). A high laminarity value suggests that there is detailed scanning in specific regions of the scene during various points in time; that is, participants frequently return to the same region in space and stay there for extended periods of time. Social scenes showed high laminarity, consistent with the fact that participants scan the facial regions repeatedly and in great detail (Birmingham & Kingstone, 2009; Laidlaw, Risko, & Kingstone, 2012; Levy, Foulsham, & Kingstone, 2013). In contrast, landscape scenes showed low determinism and laminarity, consistent with the fact that these scenes tend to encourage visual exploration (Wu, Jakobsen, Anderson, Bischof, & Kingstone, 2013).

Fractal scenes showed intermediate RQA values. We had expected that due to their high complexity and clutter, they would produce highly structured dynamics (e.g., high determinism and laminarity values). Thus, it seems that it is not complexity or clutter that is the exogenous factor driving fixation dynamics as suggested by Anderson et al. (2013). However, so far we relied on an intuitive notion of complexity and clutter. An objective analysis is provided in the next section.

Complexity and clutter

Fractal analysis (Mandelbrot, 1982) is widely used in many fields to measure complexity of images—for example, in cell biology (Smith, Lange, & Marks, 1996), physiology (Landini, Murray, & Mission, 1995), landscape textures (Plotnick, Gardner, & O'Neill, 1993), and satellite imagery (Malhi & Román-Cuesta, 2008). In this analysis, two measures are most suitable for characterizing images: fractal dimension and lacunarity. Fractal dimension is a measure of complexity that describes how systematic pattern details change over different scales of resolution. A higher

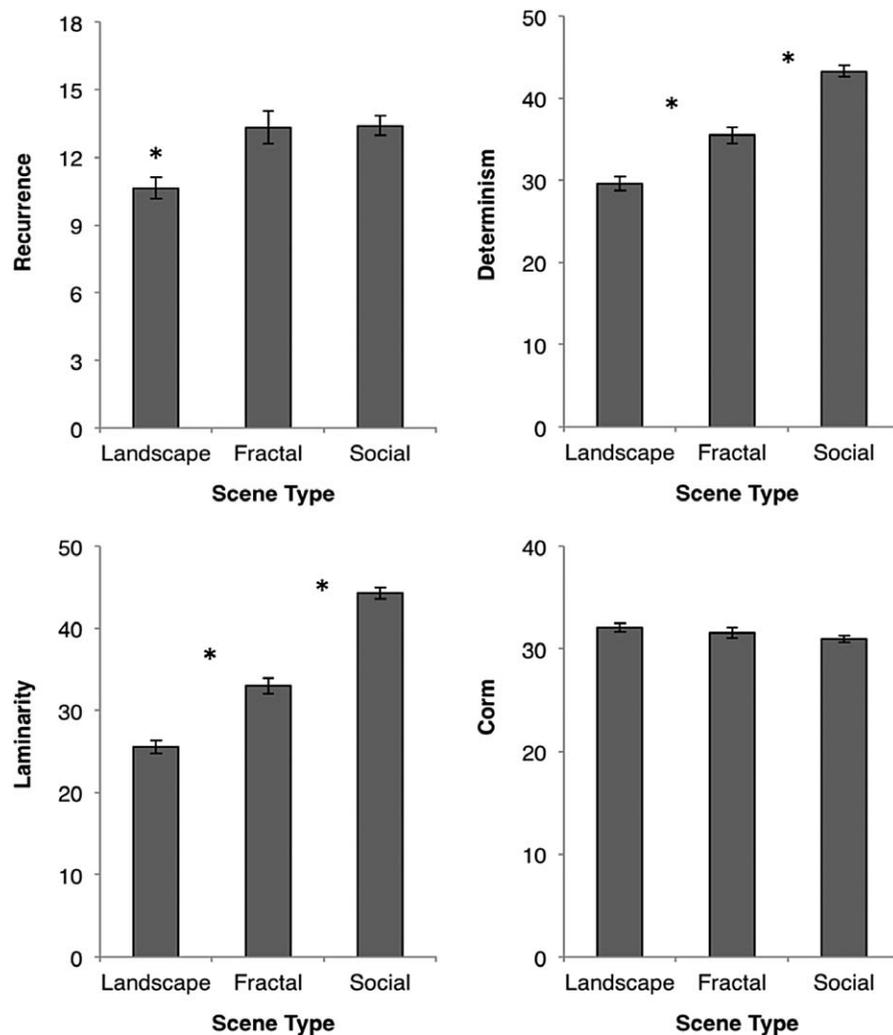


Figure 2. RQA values as a function of scene type. $N = 50$; asterisk (*) denotes significant differences at $p < 0.05$.

fractal dimension is found for more complex images. Lacunarity describes the distribution of the gap sizes in images. Lacunarity can thus be thought of as a measure of “gappiness” and heterogeneity. High lacunarity values correspond to less clutter and greater heterogeneity in the image. For example, an image with a large patch of open sky would have a high lacunarity value. See Figure 3 for examples.

In order to measure the fractal dimension and lacunarity of the stimuli, a standard box counting analysis was performed on gray-scaled scenes that were used in Experiment 1 applying the FracLac plugin (version 2.5wb126) for the program ImageJ (version 1.46r) (Karperien, 2012). We used all recommended default values except that we restricted our box counting analysis to six grid orientations, which yield six different pixel starting points for the box counting analysis. Increasing the number of grid orientations did not significantly change the fractal dimension or lacunarity values. We correlated the fractal dimension and lacunarity measures for each of

the 30 scenes with the corresponding RQA values averaged across all 50 participants. Fractal dimension and lacunarity values averaged within each scene type are reported in Table 2.

We find that higher fractal dimension of the scene (e.g., high complexity) is related negatively to RQA measures. The correlation values are shown in Table 3, and Figure 4 plots the fractal dimension values for the RQA measures across the different scene types. From Figure 4, we see that the fractal dimension for fractal scenes is quite homogenous and not nearly as strongly related to RQA values as the other two scene types. Indeed, if fractal scenes are excluded, all the correlations become stronger, ranging from $r = -0.40$ for CORM to $r = -0.68$ for laminarity, with $p < 0.001$ for all. Thus, it appears that a high level of scene complexity relates to the dynamics of eye movement behavior. Specifically, fixation sequences become less structured when images are of a higher complexity.

For lacunarity, we find that, in general, higher lacunarity values (less clutter) of a scene correspond to



Figure 3. The top row shows the two landscape scenes with the lowest and highest fractal dimension values, while the bottom row shows the two landscape scenes with the lowest and highest lacunarity values. Top left has a fractal dimension of 1.31, and top right has a fractal dimension of 1.57. Bottom left has a lacunarity value of 0.14, and bottom right has a lacunarity value of 0.64.

higher RQA values. The correlation values are shown in Table 4, and Figure 5 plots the values for recurrence and determinism measures. Similar to the pattern of results for fractal dimension, fractal scenes show the weakest relation.² When excluded, correlations increased substantially, ranging from $r = 0.52$ for determinism and laminarity to $r = 0.71$ for recurrence, with $p < 0.02$ for all.

Using fractal analysis we quantified complexity and clutter to determine whether these scene characteristics influence fixation dynamics. We found a clear division between scenes with meaningful content (landscapes

Fractal analysis	Landscape	Fractal	Social
Fractal dimension	1.47 (0.08)	1.62 (0.07)	1.38 (0.09)
Lacunarity	0.33 (0.14)	0.06 (0.03)	0.38 (0.22)

Table 2. Means for fractal dimension based on the scene type, with standard deviations in parentheses. $N = 10$.

and social scenes) and scenes with meaningless content (fractals). For scenes with meaningful content, contrary to the speculation of Anderson et al. (2013), we found that highly complex and cluttered scenes produced less structured eye movement dynamics. This may be because participants viewing scenes with meaningful content, driven by their curiosity, adopt exploratory task sets (Risko, Anderson, Lanthier, & Kingstone, 2012). Essentially, a highly complex and cluttered scene

RQA measure	Pearson R	p -value
Recurrence	-0.24	0.20
Determinism	-0.45	0.012
Laminarity	-0.48	0.007
CORM	-0.48	0.008

Table 3. Correlation values for RQA measures and the fractal dimension of scenes viewed. $N = 30$. Bolded values are significant correlations at $p < 0.05$.

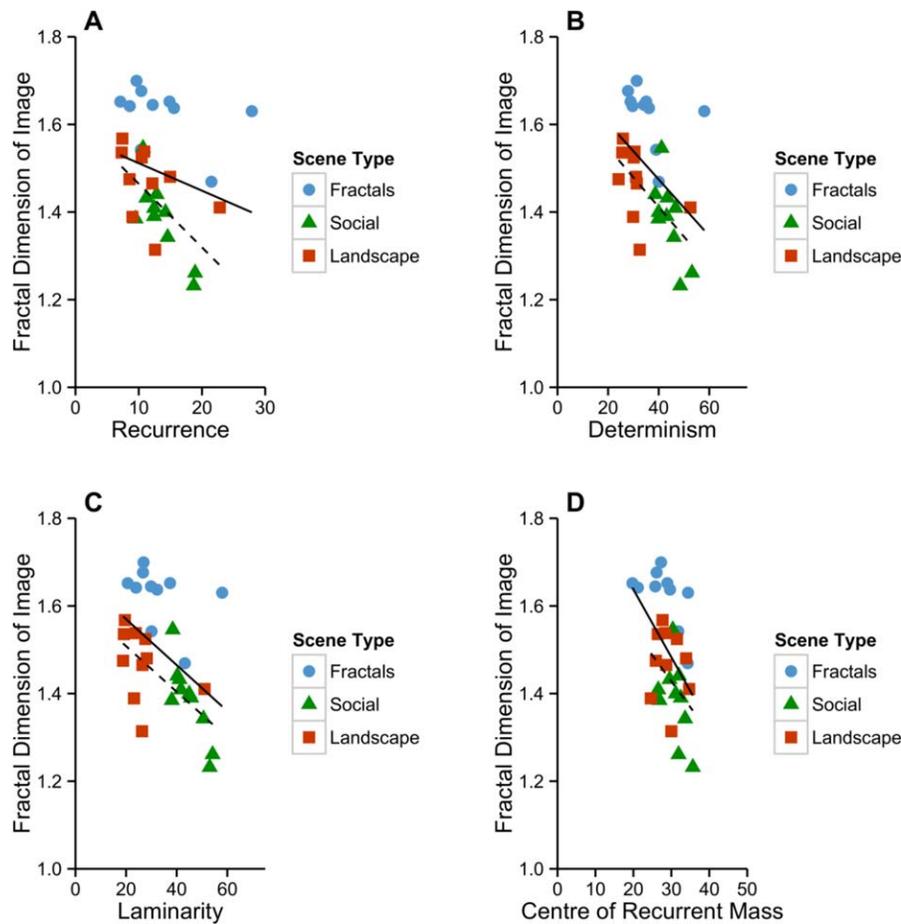


Figure 4. A, B, C, and D show plots of scene fractal dimensions with recurrence, determinism, laminarity, and CORM, respectively. Solid lines represent regression lines when all three scene types are taken into account, while dashed lines represent regression lines when fractals are removed.

may promote relatively random exploration, whereas a scene with low complexity and clutter may produce structured exploration revolving around a limited number of objects.

Saliency

It is important to note that the role of complexity and clutter appears to be specific to scenes with meaningful content (i.e., landscapes and social stimuli). When the scenes are devoid of meaning, as in the case of fractals,

RQA measure	Pearson <i>R</i>	<i>p</i> -value
Recurrence	0.27	0.15
Determinism	0.39	0.034
Laminarity	0.40	0.028
CORM	0.44	0.015

Table 4. Correlation values for RQA measures and the lacunarity value of scenes viewed. *N* = 30. Bolded values are significant correlations at *p* < 0.05.

complexity and clutter appear to have little effect on eye movement dynamics. This suggests that complexity and clutter are not merely covariates of low-level saliency features in a scene, whereby a wider distribution or a greater number of salient peaks would imply less clutter, and therefore less structured fixation dynamics.

To test this idea directly, saliency maps were computed using Saliency Toolbox (Walther, 2012). Salient peaks were defined as locations where the saliency value was at least 50% of the maximum saliency in the scene. We calculated the area spanned by these peaks (i.e., their convex hulls; see Figure 6). Results indicated that neither the distribution nor the number of peaks correlated with any of the four RQA measures (all *p* > 0.36) or with either of the two fractal analysis measures (all *p* > 0.13). See Table 5 for the descriptive statistics.

Time course analysis

Thus far, we have determined that complexity and clutter influence the temporal dynamics of scanning

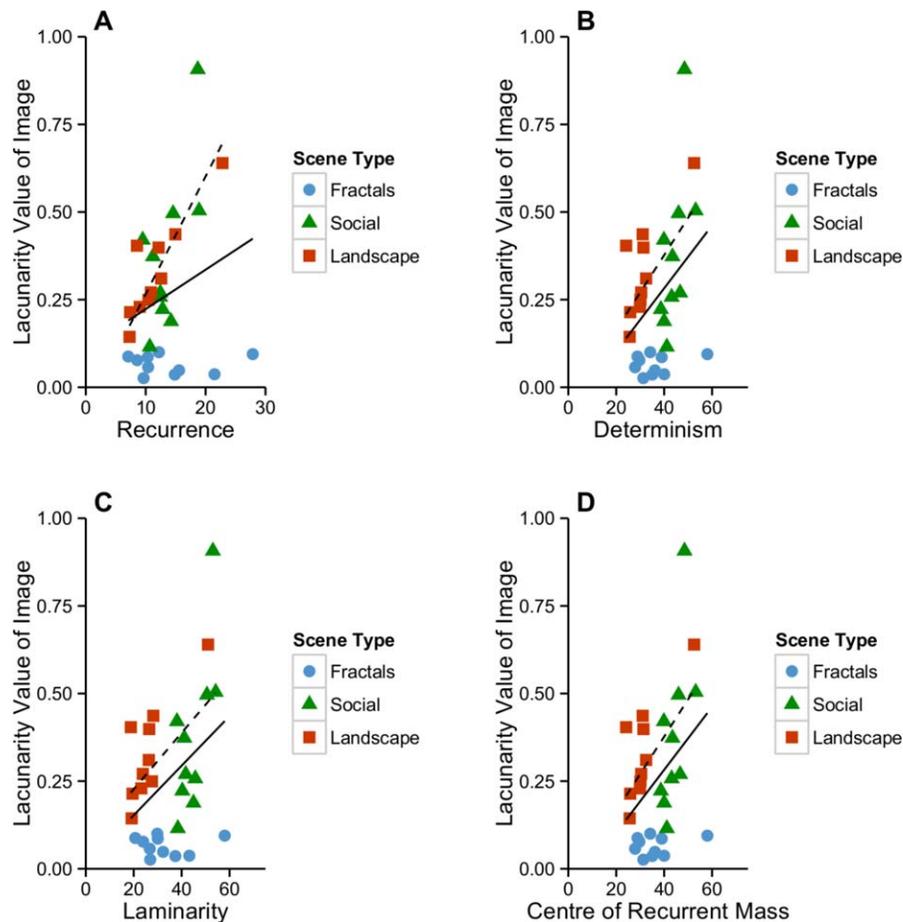


Figure 5. A, B, C, and D show plots of scene lacunarity values with recurrence, determinism, laminarity, and CORM, respectively. Solid lines represent regression lines when all three scene types are taken into account, while dashed lines represent regression lines when fractals are removed.

behavior in scenes with meaningful content. However, how scene content may differentially impact these dynamics has yet to be explored. Different scenes, and the meanings they imbue, may influence viewing strategy. For example, social scenes may initiate unique scan patterns that help one to understand the social relationships in that scene (Birmingham et al., 2008), whereas landscape scenes may tend to encourage exploration or curiosity (Wu et al., 2013). To investigate this question, we provide below some analyses on how the time course (first half vs. last half of viewing time) can interact with scene type.

We investigated how RQA values may differ between the first and the second half of viewing time. In landscape scenes, if the viewing strategy is largely exploratory, one would expect RQA values to be similar in both halves of the trial, with refixations occurring on a steady basis throughout. However, in social scenes, we expected rapid fixations to the eyes of the people in the scene as a means for participants to understand the social relation of the people in the scene before scanning other contextual elements to build a

narrative. Coupled with the semireflexive tendency to target eyes (Laidlaw et al., 2012), we suspected more structured viewing, with determinism and laminarity values greater in the first half than in the second half of the trial.

A factorial MANOVA was conducted to assess the effects of scene type (landscape, fractal, social) and interval (first half, second half) on the four RQA variables (recurrence, determinism, laminarity, CORM). The MANOVA showed significant effects of scene type, Wilk's $\lambda = 0.27$, $F(8, 190) = 21.81$, $p < 0.001$, $\eta^2 = 0.48$; interval, Wilk's $\lambda = 0.28$, $F(4, 46) = 29.31$, $p < 0.001$, $\eta^2 = 0.72$; and interaction of scene type and interval, Wilk's $\lambda = 0.83$, $F(8, 190) = 2.25$, $p < 0.05$, $\eta^2 = 0.04$. Follow-up ANOVA showed significant main effects of scene type (replicating those reported above; therefore, they will not be discussed here).

For determinism, we found a significant interaction between scene type and time course, $F(2, 98) = 5.31$, $p = 0.007$, $\eta^2 = 0.098$. This interaction is driven by the fact that, when viewing fractal scenes, determinism increases significantly over time, Bonferroni-corrected

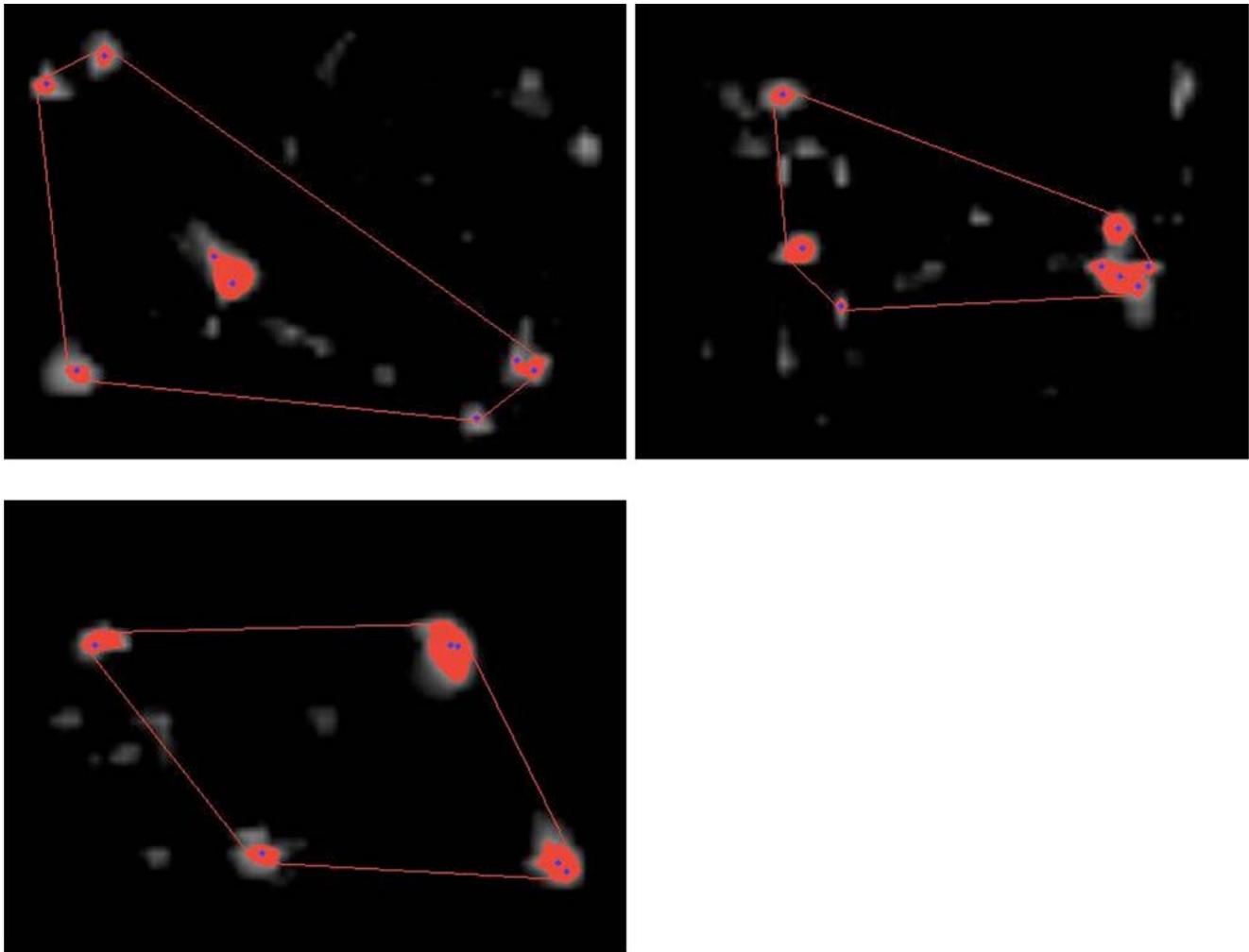


Figure 6. Examples of saliency maps from each scene type (top left = fractal, top right = landscape, bottom left = social). Large red dots represent areas where saliency was greater than 50% of the maximum saliency in the scene. Blue dots (found in the center of the large red dots) represent saliency peaks. Red lines represent the perimeter of the convex hull.

pairwise $t(49) = 3.17$, $p = 0.02$. Time course failed to play a significant role for landscape or social scenes.

In addition to the determinism findings, we found a significant interaction in laminarity, $F(2, 98) = 3.58$, $p = 0.03$, $\eta^2 = 0.068$, driven by a significant decrease when viewing social scenes, $t(49) = -3.32$, $p = 0.014$. For CORM values, there was both a main effect of time interval, $F(1, 49) = 71.02$, $p < 0.001$, $\eta^2 = 0.59$, and a significant interaction, $F(2, 98) = 3.80$, $p = 0.023$, $\eta^2 = 0.072$. Both landscape, $t(49) = -4.47$, $p < 0.001$, and

Measure	Landscape	Fractal	Social
Saliency peaks (M)	5.0 (0.87)	6.7 (1.77)	4.3 (0.72)
Area of convex hull	0.09 (0.03)	0.19 (0.07)	0.11 (0.03)

Table 5. Means for each saliency measure based on the scene type, with SEM in parentheses. The area of convex hull is expressed as a proportion of the image area.

fractals, $t(49) = -6.56$, $p < 0.001$, showed decreasing CORM values over time. We found no main effect of time course or interactions with scene type with recurrence. See Figure 7.

Our time course analysis showed that, in social scenes, laminarity significantly decreased over time relative to fractal and landscape scenes. While not significant, determinism values in social scenes were also the only ones that tended to decrease over time. This is consistent with our hypothesis that less structured viewing occurs as time passes, supporting the idea that while the eyes in images may capture attention initially, other factors may modulate this capture over time (Wu, Bischof, Anderson, Jakobsen, & Kingstone, 2014). In landscape scenes, we found determinism and laminarity values to be constant over time. Again, these findings support our idea that landscape scenes promote an exploratory viewing strategy, where there are relatively low RQA values and

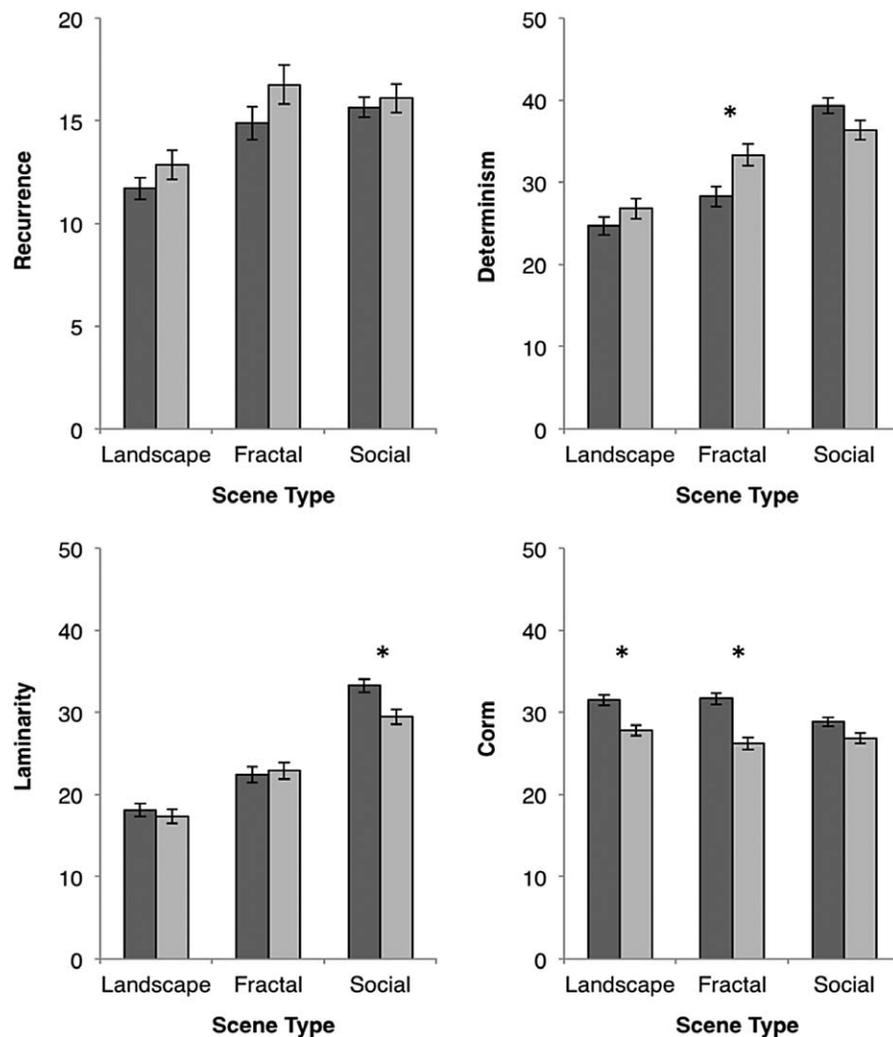


Figure 7. RQA values as a function of time course and scene type. Dark bars represent RQA values in the first half of the trial. Light bars represent RQA values in the last half of the trial. $N = 50$; asterisk (*) denotes significant differences at $p < 0.05$.

little change in structured viewing. In fractal and landscape scenes we found that CORM values decreased over time. These results suggest that recurrences tend to occur closer in time during the first half of the trial compared with the second half. One possible explanation for this result is that participants may tend to establish key points of interest during the beginning of the trial and be willing to explore the scenes more in the later part of the trial only once these key points of interest have been established.

Recurrence radius

In the analyses above, fixations were considered recurrent if they landed within a radius of 64 pixels (2.6° of visual angle) of another fixation. As this radius increases, the proportion of recurrences increases to a maximum of 100%, and as the radius decreases, the

proportion of recurrences decreases to zero (see Anderson et al., 2013). The rate at which recurrence changes with radius can reveal interesting aspects of viewing strategy, which, in turn may depend on the scene type. For example, if the rate of recurrence change is smaller in one scene type, then this suggests that fixation targeting is more fine grained and focused on key objects (e.g., eyes). On the other hand, if the rate of recurrence change is larger, then fixations may be targeting a broader area of interest (e.g., an interestingly shaped cloud). Thus, an analysis of recurrence change with radius can provide another line of evidence to point at the viewing strategies participants engage in.

As Figure 8 shows, recurrence for fractal and social scenes increases more at low radius sizes relative to landscape scenes. This result supports the hypothesis that fixation in social and fractal scenes may be targeted toward fine-grained, small objects, whereas

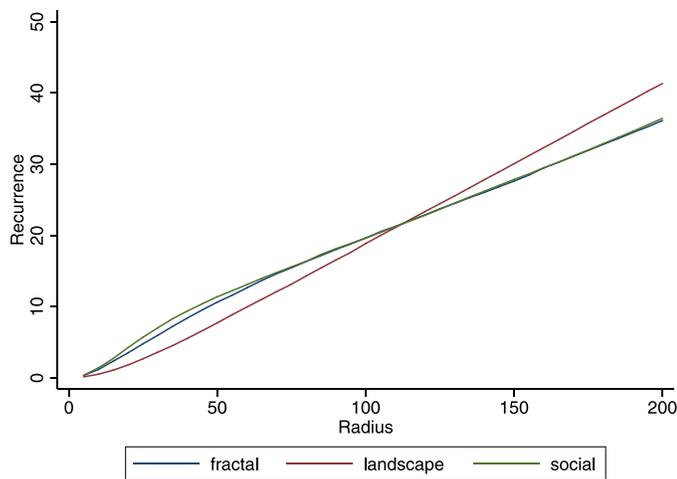


Figure 8. Recurrence plotted as a function of changing radius size and scene type.

fixation in nature scenes may be targeted toward larger objects or regions. Although testing this hypothesis is beyond the scope of the present work, the graph suggests that the rate of recurrence change dependent on the radius size may be a fruitful measure for future research.

General discussion

The aim of the present study was to expand the field's interest in how attention is allocated spatially in the natural world while encompassing the temporal dynamics of unconstrained sequences of eye movement behavior. We utilized recent RQA methodology as it enables one to measure the temporal dynamics of one or more scanpaths and to compare these scanpaths both with the content of an individual scene and across scenes of different types.

We took as our starting point the hypothesis that RQA measures would be sensitive to changes in scanning behaviors that emerge as a function of different scene types (Anderson et al., 2013). This prediction was confirmed. We found that certain scene characteristics correlated with the temporal dynamics of eye movements. Social scenes produced the most structured viewing, whereas landscape scenes produced the least structured viewing. Complexity and clutter were scene characteristics that were highly related to fixation dynamics. Greater complexity and clutter in scenes created less temporally structured eye movements as revealed by the RQA measures. Moreover, neither these image characteristics nor fixation dynamics were related to traditional low-level scene measures as defined by a scene's saliency map.

The present work demonstrates that intuitive notions of scene complexity and clutter can be objectively quantified as fractal dimension and lacunarity values. When scenes also have meaningful content, these measures appear to be related to the temporal dynamics of eye movements. In social scenes, eye movements are more structured, with more repeated fixations and scanpath segments. We believe that this relates to the fact that these scenes have a few areas of very high interest to an observer, encouraging the repeated fixations in the face area and repeated scanpaths between the people of the scene (as observed by Birmingham et al., 2008). The measurement of fractal dimension and lacunarity allows for an objective quantification of this intuitive explanation. The importance of this point should not be underestimated. Anderson et al.'s (2013) intuition was that higher laminarity and determinism scores were due to higher levels of clutter in the scene. The present data indicate that precisely the opposite relationship exists.

Nor did Anderson et al. (2013) predict the critical role stimulus meaning plays in eye movement dynamics with respect to the factors of complexity and clutter. Complexity and clutter impact eye movement dynamics only when the scenes possess meaningful content. When the scenes are meaningless (e.g., fractals), it is difficult to assess whether complexity and clutter relate to RQA measures.

Our interpretation is that these scene characteristics interact with the viewing strategy adopted by the individual. In landscape scenes, which produce reduced levels of determinism and laminarity—and therefore less repetition or perseveration on a given object—the strategy may tend to be exploratory (Risko et al., 2012). This viewing strategy was supported by our radius results, which showed that fixation targeting in landscape scenes was less fine than that in social or fractal scenes. In social scenes, which produce high determinism and laminarity, the strategy tends to focus on repeated viewing of a few fine-grained objects; for example, repeatedly looking at, or between, people in the scene in order to understand their social relationship (Birmingham et al., 2008). Over time, the capture of the eyes may lessen as participants begin to explore other areas of the scene. Our finding that laminarity values decreased over time supports this theory. In fractal scenes, the determinism and laminarity measures fall between the values for the landscape and social scenes (see Figure 2) and are not strongly related to the complexity and clutter of the image. One potential explanation is that the meaningless fractal scenes produce a viewing strategy that is unrelated to the visual features examined in this study. Although fractal images are complex and abstract, they do contain regions where

patterns tend to converge (see, e.g., Figure 1). These points may serve as frames of reference for inspection (Stainer, Scott-Brown, & Tatler, 2013; Wade, 1992) of an otherwise highly complex and cluttered scene devoid of meaning. The frames of reference idea may also support our radius and time course findings, which showed that participants appeared to be captured by fine-grained objects and that recurrences became more deterministic over time, with more repeated scanpath segments in the second half of viewing time. The temporal dynamics produced by these scenes may therefore be driven by these frames of reference rather than by the meaning or complexity of the scene.

It has been noted that a potential limitation of the free-viewing paradigm is that it may not reflect a default mode of attention allocation (e.g., Tatler et al., 2011). Thus, participants may engage in a viewing strategy that is unknown to the experimenter. Our findings suggest that RQA can be a useful tool for detecting, describing, and defining such viewing strategies, as we find substantial differences between the RQA values between different scene types.

We hope that the present study serves as a catalyst for future studies to use novel analysis techniques to investigate the temporal dynamics and processes that drive eye movement behaviors. The use of these techniques is only in its infancy (Anderson et al., 2013; Barthelmé et al., 2013; Cagli et al., 2008; Hayes et al., 2011; Wang et al., 2012), but the present study serves as a useful demonstration of the power of this approach. We found large differences in eye movement dynamics depending on the type of scene viewed. Moreover, we discovered that complexity and clutter are two scene characteristics that are highly related to fixation dynamics, but only if the scene content is meaningful. Finally, we revealed that scene primitives (stimulus saliency) was not responsible for our findings. We suggest that eye movement dynamics result from top-down viewing strategies that vary according to the meaning of a scene and its associated visual complexity and clutter.

Keywords: eye movements, recurrence quantification analysis, fractal analysis, salience, complexity

Acknowledgments

This work was funded by grants from Natural Sciences and Engineering Research Council of Canada (WFB, AK) and the Social Sciences and Humanities Research Council (AK).

Commercial relationships: none.

Corresponding author: David W.-L. Wu.

Email: david.wl.wu@gmail.com.

Address: Department of Psychology, University of British Columbia, Vancouver, British Columbia, Canada.

Footnotes

¹To access the full stimuli set, please contact the corresponding author.

²We also tested correlations when landscape scenes and social scenes were separated. We found that in social scenes, there were no significant correlations between CORM and fractal dimension or lacunarity, or between determinism and lacunarity. In landscape scenes, there were no significant correlations between any RQA measure and fractal dimension, but all were correlated with lacunarity. Linear regressions with scene type as a discrete predictor variable and fractal dimension and lacunarity as continuous predictor variables also found differential effects based on scene type. Social scenes were a significant predictor for recurrence, $t(29) = -2.52$, $p = 0.02$, whereas landscape scenes were a significant predictor for determinism, $t(29) = -2.96$, $p = 0.01$, and lacunarity, $t(29) = -2.64$, $p = 0.02$. We note, however, that these findings should not be taken as anything more than suggestive, as the number of images in each scene type is relatively low.

References

- Açık, A., Bartel, A., & König, P. (2014). Real and implied motion at the center of gaze. *Journal of Vision*, *14*(1):2, 1–19, <http://www.journalofvision.org/content/14/1/2>, doi:10.1167/14.1.2. [PubMed] [Article]
- Anderson, N. C., Bischof, W. F., Laidlaw, K. E. W., Risko, E. F., & Kingstone, A. (2013). Recurrence quantification analysis of eye movements. *Behavior Research Methods*, *45*, 842–856.
- Barthelmé, S., Trukenbrod, H., Engbert, R., & Wichmann, F. (2013). Modeling fixation locations using spatial point processes. *Journal of Vision*, *13*(12):1, 1–34, <http://www.journalofvision.org/content/13/12/1>, doi:10.1167/13.12.1. [PubMed] [Article]
- Birmingham, E., Bischof, W. F., & Kingstone, A. (2008). Gaze selection in complex social scenes. *Visual Cognition*, *16*, 341–355.
- Birmingham, E., Bischof, W. F., & Kingstone, A. (2009). Saliency does not account for fixations to

- eyes within social scenes. *Vision Research*, *49*, 2992–3000.
- Birmingham, E., & Kingstone, A. (2009). Human social attention: A new look at past, present and future investigations. *The Year in Cognitive Neuroscience 2009: NY Academy of Sciences*, *1156*, 118–140.
- 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.
- Cagli, R. C., Coraggio, P., Napoletano, P., & Boccignone, G. (2008). What the draughtsman's hand tells the draughtsman's eye: A sensorimotor account of drawing. *International Journal of Pattern Recognition and Artificial Intelligence*, *22*, 1015–1029.
- Dalrymple, K., Gray, A., Perler, B., Birmingham, E., Bischof, W. F., Barton, J., & Kingstone, A. (2013). Eyeing the eyes in social scenes: Evidence for top-down control of stimulus selection in simultanagnosia. *Cognitive Neuropsychology*, *30*, 25–40.
- Dewhurst, R., Nyström, M., Jarodzka, H., Foulsham, T., Johansson, R., & Holmqvist, K. (2012). It depends on how you look at it: Scanpath comparison in multiple dimensions with MultiMatch, a vector-based approach. *Behavioral Research Methods*, *44*, 1079–1100.
- Foulsham, T., & Kingstone, A. (2010). Asymmetries in the direction of saccades during perception of scenes and fractals: Effects of image type and image features. *Vision Research*, *50*, 779–795.
- 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, <http://www.journalofvision.org/content/8/2/6>, doi:10.1167/8.2.6. [PubMed] [Article]
- 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*(10):10, 1–11, <http://www.journalofvision.org/content/11/10/10>, doi:10.1167/11.10.10. [PubMed] [Article]
- Itti, L., & Koch, C. (1999). A comparison of feature combination strategies for saliency-based visual attention systems. In B. E. Rogowitz and T. N. Pappas (Eds.), *SPIE human vision and electronic imaging IV* (pp. 473–482). Bellingham, WA: SPIE Society of Photo-optical Instrumentation Engineers.
- Karperien, A. (2012). FracLac for ImageJ (Version 2.5). <http://rsb.info.nih.gov/ij/plugins/fraclac/FLHelp/Introduction.htm>.
- Laidlaw, K. E. W., Risko, E. F., & Kingstone, A. (2012). A new look at social attention: Orienting to the eyes is not (entirely) under volitional control. *Journal of Experimental Psychology: Human Perception and Performance*, *38*, 1132–1143.
- Landini, G., Murray, P. I., & Mission, G. P. (1995). Local connected fractal dimensions and lacunarity analyses of 60 degrees fluorescein angiograms. *Investigative Ophthalmology and Visual Science*, *36*(13), 2740–2755, <http://www.iovs.org/content/36/13/2749>. [PubMed] [Article]
- 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. (2013). Monsters are people too. *Biology Letters*, *9*, 20120850, doi: 10.1098/rsbl.2012.0850.
- Malhi, Y., & Román-Cuesta, R. M. (2008). Analysis of lacunarity and scales of spatial homogeneity in IKONOS images of Amazonian tropical forest canopies. *Remote Sensing of Environment*, *112*, 2074–2087.
- Mandelbrot, B. B. (1982). *The fractal geometry of nature*. New York: Holt.
- Oliva, A. (2005). Gist of the scene. In L. Itti, G. Rees, & J. K. Tsotsos (Eds.), *Neurobiology of attention* (pp. 251–256). Burlington, MA: Elsevier Academic Press.
- Plotnick, R. E., Gardner, R. H., & O'Neill, R. V. (1993). Lacunarity indices as measures of landscape texture. *Landscape Ecology*, *8*, 201–211.
- Risko, E. F., Anderson, N. C., Lanthier, S., & Kingstone, A. (2012). Curious eyes: Individual differences in personality predict eye movement behavior in scene-viewing. *Cognition*, *122*, 86–90.
- Schütz, A. C., Lossin, F., & Kerzel, D. (2013). Temporal stimulus properties that attract gaze to the periphery and repel gaze from fixation. *Journal of Vision*, *13*(5):6, 1–17, <http://www.journalofvision.org/content/13/5/6>, doi:10.1167/13.5.6. [PubMed] [Article]
- Smith, T. G., Jr., Lange, G. D., & Marks, W. B. (1996). Fractal methods and results in cellular morphology—Dimensions, lacunarity and multifractals. *Journal of Neuroscience Methods*, *69*, 123–136.
- Stainer, M. J., Scott-Brown, K. C., & Tatler, B. W. (2013). Behavioral biases when viewing multiplexed scenes: Scene structure and frames of reference for inspection. *Frontiers in Psychology*, *4*, 624, doi:10.3389/fpsyg.2013.00624.

- Tatler, B. W., Hayhoe, M. M., Land, M. F., & Ballard, D. H. (2011). Eye guidance in natural vision: Reinterpreting salience. *Journal of Vision*, *11*(5):5, 1–23, <http://www.journalofvision.org/content/11/5/5>, doi:10.1167/11.5.5. [PubMed] [Article]
- Tatler, B. W., & Vincent, B. T. (2008). Systematic tendencies in scene viewing. *Journal of Eye Movement Research*, *2*, 1–18.
- 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.
- Wade, N. J. (1992). The representation of orientation in vision. *Australian Journal of Psychology*, *44*, 139–145.
- Walther, D. (2012). The Saliency Toolbox (Version 2.2). <http://www.saliencytoolbox.net>.
- Wang, H. X., Freeman, J., Merriam, E. P., Hasson, U., & Heeger, D. J. (2012). Temporal eye movement strategies during naturalistic viewing. *Journal of Vision*, *12*(1):16, 1–27, <http://www.journalofvision.org/content/12/1/16>, doi:10.1167/12.1.16. [PubMed] [Article]
- Wu, D. W.-L., Bischof, W. F., Anderson, N. C., Jakobsen, T., & Kingstone, A. (2014). The influence of personality on social attention. *Personality and Individual Differences*, *60*, 25–29.
- Wu, D. W.-L., Jakobsen, T., Anderson, N. C., Bischof, W. F., & Kingstone, A. (2013). Why we should not forget about the non-social world: Subjective preferences, exploratory eye-movements, and individual differences. *Proceedings of the 35th Annual Meeting of the Cognitive Science Society*, 3801–3806.