Attractive Flicker
Guiding Attention in Dynamic Narrative Visualizations

Manuela Waldner, Mathieu Le Muzic, Matthias Bernhard, Werner Purgathofer, *Member, IEEE*, and Ivan Viola

Abstract— Focus+context techniques provide visual guidance in visualizations by giving strong visual prominence to elements of interest while the context is suppressed. However, finding a visual feature to enhance for the focus to pop out from its context in a large dynamic scene, while leading to minimal visual deformation and subjective disturbance, is challenging. This paper proposes Attractive Flicker, a novel technique for visual guidance in dynamic narrative visualizations. We first show that flicker is a strong visual attractor in the entire visual field, without distorting, suppressing, or adding any scene elements. The novel aspect of our Attractive Flicker technique is that it consists of two signal stages: The first "orientation stage" is a short but intensive flicker stimulus to attract the attention to elements of interest. Subsequently, the intensive flicker is reduced to a minimally disturbing luminance oscillation ("engagement stage") as visual support to keep track of the focus elements. To find a good trade-off between attraction effectiveness and subjective annoyance caused by flicker, we conducted two perceptual studies to find suitable signal parameters. We showcase Attractive Flicker with the parameters obtained from the perceptual statistics in a study of molecular interactions. With Attractive Flicker, users were able to easily follow the narrative of the visualization on a large display, while the flickering of focus elements was not disturbing when observing the context.

Index Terms—Visual attention, flicker, narrative visualization

1 INTRODUCTION

Dynamic visualizations of complex phenomena or real-time events can be interactively explored by the user, but at the same time they tell a story. Examples are 3D visualizations of physiological processes, crowd simulations, multiple coordinated views of dynamically changing stock exchange data, or observation and collaboration interfaces in emergency response centers. What these examples have in common is that a lot of dynamically changing data has to be displayed concurrently. This is often achieved by using large displays, such as projection screens or multi-monitor setups. As a result, the user is confronted with a large amount of information and visual clutter. A major challenge in such a setting therefore is to guide the user's attention to the current main actors of the story while not introducing too much distraction from the remaining content or causing subjective annoyance.

Focus+context techniques make *"uneven use of graphics resources [...] to visually discriminate data-parts in focus from their context"* [14] and thereby guide the user's attention to elements of importance. Common approaches exaggerate or suppress visual features of scene elements, like hue [8], luminance [22, 7], sharpness [25], or size [10], to achieve visual guidance. However, in large cluttered scenes with dynamic elements – like those found in the examples mentioned above – modifying these features may not be sufficient to generate a popout effect to reliably attract the user's attention. Especially if visualizations exceed the user's parafoveal field of view, such static visual attraction techniques may fail without any additional guidance, like trails [16, 39] or a virtual searchlight [22].

In addition, many of the above mentioned techniques visually deform data entities and thereby increase the probability that the scene is misunderstood. For instance, fisheye distortions can lead to misjudgments of relative distances [12], and blurring of context elements

makes it hard to identify contextual details [48]. Our goal therefore is to achieve visual guidance with minimal deformation of the data to maintain scene understanding.

A class of attraction techniques that has been demonstrated to work well under various conditions are those adding dynamic changes to focus elements. We are particularly interested in attractors that change the visual appearance of focus elements over time, *i.e.*, make them flicker. While being highly effective even in dynamic scenes [34], flickering attractors in graphical user interfaces are also known to be a source of annoyance for users [11, 16].

In this paper, we explore the design space of flickering attractors to make elements of interest stand out in large, complex visualizations while adding minimal annoyance for the observer. To achieve this goal, we introduce a new class of flickering attractors: *Attractive Flicker*. What distinguishes Attractive Flicker from other visual guidance techniques is that it uses a two-stage signal: A short, but highly salient, *orientation stage* is followed by a smooth decay into a minimally disturbing *engagement stage*. We empirically determined the flicker parameters to effectively attract the user's attention in the orientation stage, and to generate low annoyance in the engagement stage, in two perceptual pilot experiments on a large display showing a dynamic scene. Finally, we demonstrate that Attractive Flicker effectively guides the user through a narrative visualization while adding low distraction in a study of large-scale molecular interactions.

2 BACKGROUND

Attractive Flicker utilizes knowledge of basic visual attention research to provide effective visual guidance. Therefore, we first summarize basics from the visual perception and cognition point of view, followed by related work on attention guidance.

2.1 Visual Attention

When observing an image, there are several factors that influence where our attention is directed. *Feature integration theory* (FIT) [41] suggests that multiple feature dimensions of an image – such as color, orientation, spatial frequency, brightness, and movement direction – are rapidly processed automatically and in parallel. This *preattentive* processing ability enables observers to catch certain properties of a scene in less than 250 milliseconds, *i.e.*, without moving the eyes to scan the scene. For instance, simple search tasks, like determining whether there is a single red dot within a number of blue dots, can be conducted preattentively (see Healey and Enns [15] for an overview).

- *Manuela Waldner, Mathieu Le Muzic, Matthias Bernhard, Werner Purgathofer, and Ivan Viola are with Vienna University of Technology. E-mail:* {*waldner*|*mathieu*|*Matthias.Bernhard*|*wp*|*viola*}*@cg.tuwien.ac.at*
- *Werner Purgathofer is with VRVis Research Center. E-mail: Purgathofer@VRVis.at.*

Only after this preattentive processing, objects can be identified by integrating the separate features with the support of focused attention. Experiments with implementations of bottom-up saliency models (*e.g.*, [20, 19]) showed that motion and flicker feature maps were stronger predictors of users' saccade directions than color, intensity, and orientation contrasts [18, 31], indicating that dynamic features are stronger visual attractors than static ones.

It is well known that bottom-up saliency is by far not the only factor influencing attention in visual scenes. *Guided search* theory by Wolfe *et al.* [47] assumes that visual attention is guided by both, bottom-up and top-down information. Both, stimulus-driven bottom-up and expectation-driven top-down processing, contribute to a common *activation map*, guiding the observer's attention in the scene.

2.2 Attention Guidance

Despite the fact that top-down attention plays a major role in visual attention, the influence of bottom-up attention should not be neglected. In psychology and education research, *selective processing* describes the phenomenon that learners are mainly able to reproduce insights encoded by perceptually salient image regions [26]. Healey and Enns [15] argue that directing the users' attention to important elements in a visualization may strengthen their engagement and, as a result, increase their insight.

To increase bottom-up saliency in regions of interest, focus+context techniques enhance certain features of focus elements or suppress features of context elements [14]. For instance, common ways to highlight elements of interest are to use distinct hues for the focus elements (commonly used for brushing and linking [8]), to decrease the luminance of context elements [22, 7], blur the context [25], or generally reduce the amount of details for context regions by simplifying the rendering style [6]. Traditional focus+context techniques apply spatial distortions to dedicate more screen space to focus elements, while suppressing the context (*e.g.*, *Fisheye* techniques [10]). A different approach is to add artificial visual cues to the scene to make a target pop out. Examples are text labels for the selected elements (*e.g.*, in *Gapminder* [1]) or trails leading the user's gaze to highlighted elements [16, 39]. Most of these examples apply substantial modifications to the scene. In many cases, this is desirable, as the attention guidance mechanism itself also serves as an aesthetic scene property (for instance, when using stylized rendering [6]). In other cases, a strong visual deformation of the context may lead to an undesirable misinterpretation of the scene.

Another challenge of visual attention guidance is to find an image feature that can be modified to create a reliable popout effect in complex scenes. As formalized by Rosenholtz [35], a focus element needs to be sufficiently different from the context with respect to the modified feature to generate a popout effect. One approach to achieve this goal is to use an additional output modality to draw the user's attention to a certain location, such as spatial audio [30, 40]. As a purely visual solution, researchers have suggested to iteratively tune static image features in regions of interest to increase the bottom-up saliency of these regions for videos [42] or volume visualizations [23]. A different approach is to pick a feature to enhance that is underrepresented in a scene. For instance, on desktop interfaces, "moticons" [4] are icons effectively drawing the user's attention by simple motions. In information visualization, oscillating movements have been used for filtering and brushing [3], to highlight connected nodes in subgraphs [44], but also to visualize multivariate data [17, 45]. It could be shown that dynamic targets are much easier to detect than static ones, especially with increasing distance from the users' fixation point [33, 4]. But what if the scene itself contains inherent motion?

2.3 Flicker Attractors

A low-level visual feature that is typically not prominently used in both, static and dynamic visualizations is flicker. Flicker has a couple of interesting properties that make it attractive for attention guidance in visualization. Pinto *et al.* [34] could show in an experiment that blinking targets can be effortlessly discriminated from moving distractors. When flickering in coherent phases, even differences between flicker

frequencies can be easily detected [17]. Hoffmann *et al.* [16] found that blinking window frames were more effective to guide the user's attention on large displays than static attractors like a red frame around the window. However, their results also showed that blinking window frames were considered the most annoying visual cue. In contrast, in an experiment by Bartram *et al.* [4], users found that a 1 Hz blinking of icons is the least disturbing – but also least effective – motion cue. In a similar experimental setting, Gluck *et al.* [11] discovered that a proxy icon following the cursor was more effective, but also more annoying than a ∼3 Hz blinking icon. These experiments indicate that there seems to be a trade-off between effectiveness of an attractor and its perceived annoyance. However, there is little knowledge how the two different parameters of a flicker signal – namely amplitude and frequency – interact and affect detection rate and perceived annoyance. The goal of our work is to answer this research question and to design a new class of flicker attractors based on empirical evidence.

An alternative approach to overcome the annoyance introduced by flickering attractors was presented by Bailey *et al.* [2]. Their *Subtle Gaze Direction* (SGD) technique uses 10 Hz luminance or hue flicker signals to attract the user's attention to targets in the peripheral vision. To avoid annoyance, they use an eye tracker to detect the user's current fixation point and saccade direction. Whenever the user initiates a saccade towards the flickering target, the flicker is stopped. Studies showed that SGD could reliably guide the users' gaze towards flickering targets in static images [2], as well as a sequence of image regions in narrative paintings [29], without users being aware of the flickering. We pursue a slightly different goal than subtle attention guidance techniques like SGD [2] or saliency modulation of static image features [23, 42]. Instead of subtly attracting the users' attention without their conscious knowledge, we want to provide a clearly visible, yet minimally disturbing signal the user can actively follow, to understand the narrative of a visualization. The grand challenge is to design a minimally distracting attractor that works with complex, dynamic visualizations, and does not prevent the user from simply walking up and start using the system without hardware-induced constraints or time-consuming calibration processes.

3 REQUIREMENTS FOR VISUAL GUIDANCE

When addressing the problem of attention guidance in large, dynamic visualizations, we are confronted with a number of requirements that rule out many traditional focus+context techniques:

R1: *Visual guidance should not be limited to the (para)foveal visual field. It should also work on large displays:* Humans are color-blind in the peripheral vision [32], which implies that attractors based on chromatic contrast are unsuitable for large displays. In general, dynamic cues are considered more effective visual attractors when using large displays than static ones [4, 43].

R2: *Visual guidance should also work in complex and cluttered scenes with a lot of different colors and elements changing their visual appearance or location over time:* From a neurobiological point of view, flicker is encoded in a separate feature map and therefore is less likely to interfere with other features like hue or motion direction [19, 34], which are more frequently used in visualizations than flicker. However, little is known how well a flicker signal can attract a user's attention in the peripheral field of view when the visualization itself contains a large number of heterogeneous and dynamic elements. We will address this research question in a comparative pilot experiment (Section 4) and in our first perceptual study (Section 6).

R3: *To enable visual guidance, modifications to the scene – both to the attractors and the distractors – should be minimal:* When integrating over time, the flicker feature (*e.g.*, the object's geometry or appearance) remains constant. Thereby, it does not persistently modify the scene in an indeterminable way such as, for instance, a continuous change of hue or saturation, but can still be clearly discriminated from the context.

R4: *Scene modifications should lead to minimal annoyance or irritation of the user:* Flicker has been shown to be more irritating than many static attractors [16], but less distracting than other motion cues [4]. In our second perceptual study, we will therefore explore the

trade-off between flicker effectiveness and annoyance by controlling its amplitude and frequency (Section 7).

R5: *Visual guidance should be possible without expensive or restricting hardware:* There are attention guidance techniques that presumably work well in large, complex visualizations with negligible modifications to the visual scene by using spatial audio through tracked headphones (*e.g.*, [40]) or an eye tracker to switch off a dynamic attractor on demand [2]. However, these techniques require non-commodity and potentially restricting hardware equipment. We therefore seek to explore a new purely visual attractor method that does not rely on specialized hardware and allows for walk-up interaction without prior calibration.

Driven by these considerations, we decided to thoroughly investigate flicker for guiding the observer through a dynamic narrative visualization. There are different visual features that can be oscillated to generate a flicker effect: geometric features (*e.g.*, object shape or size) and color (*e.g.*, hue and luminance). Since we were initially motivated by guidance through molecular processes (see Section 9), where preserving the geometry is crucial for an in-depth analysis of molecular structures and binding sites, we concentrated on color flicker. After informal pilot experiments, we could confirm that luminance flicker is a much stronger visual attractor than hue flicker. This is consistent with findings by Bailey *et al.* [2], and can be explained by the chromatic contrast blindness of humans in the peripheral vision [32]. We will therefore put the focus on luminance flicker in this paper.

4   COMPARATIVE PILOT EXPERIMENT

The goal of this initial pilot experiment was to verify that luminance flicker is indeed a strong visual attractor, as shown by prior work [34, 11, 16], and to investigate its effectiveness in a large, complex visualization. With this experiment, we aim to extend findings by Pinto *et al.* [34], who found that flickering targets can be easily spotted within uniformly moving distractors, by two aspects: first, an increased scene complexity by adding different colors and random object movements, and second, an extended visual field covered by the visualization. Our work also extends findings from experiments in desktop interface research (*e.g.*, [4, 11, 16]) by comparing visual guidance techniques not only on a large display, but also in a complex, dynamic scene. For this purpose, we asked users to participate in a dual-task experiment in a scene with 1000 colored and randomly moving dots on a multi-monitor setup, comparing luminance flicker to two static visual guidance techniques.

The traditional approach to analyze visual attention is to employ eye tracking. However, the sole analysis of eye movements fails to explain cognitive processes of the observer while viewing the scene [5, 15], since eye movement patterns do not necessarily affect the user's cognitive processing of the scene [7]. We therefore use other well-known metrics to quantify the amount of attention attraction, such as subjective reports and reaction times [5, 15] over eye tracking in all our experiments.

4.1   Participants

Ten participants (aged 21 to 50, 3 females) were recruited from a local university and a nearby company, including business assistants, software developers, and computer scientists.

4.2   Apparatus

We arranged three 27 inch monitors so that they subtended a visual angle of approximately 180°, as visualized in Fig. 1. We curved the monitors around the user so all pixels are equally usable, as recommended by Shupp *et al.* [38]. We used three identical BenQ BL2710 LCD monitors which were calibrated with an external monitor calibrator using a seven-color sensor. In sum, the three monitors provided 7680×1440 pixels. Users were not fixated into a chin rest and were instructed to behave naturally.

To compare attention attraction for conventional display sizes and large display setups covering also the peripheral visual field, we split the available display space into two separate regions. We used Sanders' [36] separation of *head field* (*i.e.*, where combined head-eye
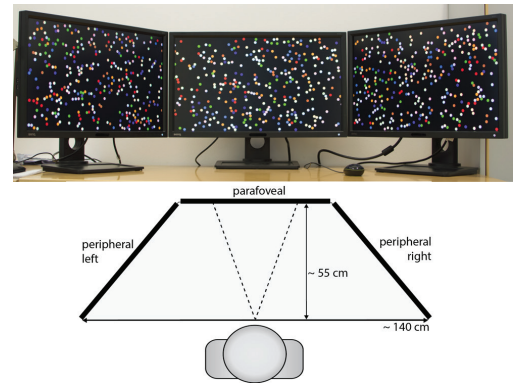

Fig. 1. Setup with three 27 inch monitors used in our experiments.

movement is required to acquire two subsequent targets) and *eye field* and *stationary field*, respectively (*i.e.*, where subsequent targets can be spotted without without head movements), to distinguish these two regions. Depending on the complexity of the scene, the *head field* starts at approximately 60°[36]. We therefore chose the two outer monitors – both starting at a subtended visual angle of $\sim 66°$– for displaying *peripheral* targets. To clearly distinguish peripheral from central targets, we defined a *(para)foveal* display region in the center of the central monitor, covering a subtended visual angle of $\sim 40°$, as indicated by the dashed lines in Fig. 1.

During the experiment, users were exposed to a scene showing 1000 randomly moving and colliding dots with 20 pixels radius each. On the central monitor, one dot subtended a visual angle of $\sim 0.96°$, which is about half of the foveal visual field with 2°. The number and size of dots was chosen to closely represent the appearance of well-known visual attention experiments (see [15] for an overview). To add complexity to such a scene, we added heterogeneity to two visual features: color and motion. The dots were colored by 12 different colors from a qualitative *ColorBrewer* [13] scheme on black background. Movement of dots was defined by random impulse forces applied routinely in random direction. The parameters of the random motion are the magnitude of the impulse and the friction coefficient. The parameters were chosen empirically so that dots would move between 21 to 170 pixels per second, which with our monitor resolution results in a speed the human visual system is most sensitive to [9]. The refresh rate of the visualization was synchronized with the monitor refresh rate of 60 Hz. Fig. 1(top) shows the scene on the three monitors.

4.3   Design

To prevent users from actively searching for a visual guidance signal, we employed a dual-task design. The primary task was to select as many red dots as possible on the central monitor. The secondary task was to press the space bar as quickly as possible when spotting a so-called "bonus-dot" appearing anywhere on the three monitors. The bonus dot was highlighted by one of three visual guidance techniques:

In the **flicker** technique, the bonus dot was signaled by a 10 Hz luminance flicker, covering the full luminance range from 0 to 100 in the CIELab space.

In the **spotlight** technique, presented by Khan *et al.* [22], the bonus dot was surrounded by a bright ring with a luminance value of 50, with twice the radius of the dot. In addition, the luminance of context dots was decreased by 75% (Fig. 2 left).

Finally, the **halo** technique used a ring identical to spotlight, but without context darkening (Fig. 2 right).

We used a 3×2 within-subjects design with the following factors: **Technique** for visual guidance, as described above and **eccentricity** of the bonus dot on the screen (either *(para)foveal* or *peripheral*, as shown in Fig. 1).

Users had to perform ten repetitions, resulting in 60 trials in total. The trials were blocked per technique, and the order of the techniques was counter-balanced. The exact spatial location of the bonus
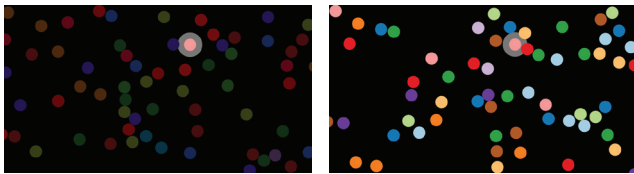
Fig. 2. Small sub-window with a spotlight (left) and halo (right).

dot within the (para)foveal or peripheral display regions, as well as the color of the dot, were picked randomly for each stimulus.

4.4 Procedure

For each trial, users had to press the Return-key to start a two-second countdown animation in the center of the screen. After this countdown, users could start with clicking red dots as their primary task. After a random period between two and six seconds, the bonus dot was highlighted. As soon as they pressed the space bar, or after a five-seconds timeout after the appearance of the highlight, the stimulus image was frozen, darkened, and all highlights were removed. Users then had to select the bonus dot with the mouse cursor, or press "n" in case they did not see it. Before the experiment, we demonstrated all three visual guidance techniques in their order of appearance and asked users to perform a training run. In the experiment, we logged *distance* (in pixels) of the user's selection to the bonus dot (or -1, if the user pressed "n") as a measure of correctness and *reaction time* between the highlight appearance and pressing space. In addition, we performed an audio-recorded post-experiment interview, where we asked people about perceived differences between the techniques.

4.5 Results and Discussion

The *success* of a trial was either 0, if the user pressed "n"or if the distance of the selection click to the actual bonus dot center was larger than 80 pixels (*i.e.*, 4 times the dot radius), or 1. The success rates were compared by a repeated measures logistic regression. Both, technique and eccentricity had a significant main effect (technique: $\chi^2(2) = 741.907, p < .001$; eccentricity: $\chi^2(1) = 901.518, p < .001$). Pairwise Bonferroni-corrected comparisons showed that halo led to a significantly lower success rate than spotlight and flicker, and that (para)foveal targets scored a higher success rate than peripheral ones (*cf.*, Fig. 3 left).

We conducted a mixed model analysis of reaction time of successful trials with technique and eccentricity as fixed factors and user as random factor. Both fixed factors are significant (technique: $F_{2,509.644} = 37.730, p < .001$; eccentricity: $F_{1,509.319} = 87.238, p < .001$), and there is also an interaction between them ($F_{2,509.438} = 13.771, p < .001$). Bonferroni-corrected post-hoc comparisons revealed that spotlight led to the fastest reaction times and halo to the slowest. (Para)foveal targets could be detected significantly faster than peripheral ones (*cf.*, Fig. 3 right).

In summary, the findings of our pilot experiment are:

First, targets in the (para)foveal region could be detected more reliably and faster than those in the periphery.

Second, a simple halo around the target was more likely to be missed than a flickering target or a bright halo in a dark scene, and also led to the slowest reaction times. In the post-experiment interview, all users stated that halo was the hardest technique.

Third, flicker had an equal success rate as spotlight but led to a slower reaction time. Users explained that with spotlight *"you just got a visual cue: now it happened!"* (User 10). Subsequently, *"you had to search around where the dot is exactly"* (User 8), *"...which was not very hard"* (User 9). In contrast, with flicker *"I don't have to search, really, because I already know in which direction the blinking was"* (User 7). However, a few users had the impression that spotlight caused a too obtrusive interruption (*e.g.*, User 5: *"It is slightly irritating, because it covers the other things."*). While this darkening did not influence the performance in our simple experimental task, this strong interruption signal may be undesirable in many narrative visualizations, where users should be able to choose whether to follow the story, or to explore the scene without visual guidance.

Our findings suggest that luminance flicker can achieve a similar performance as the spotlight technique [22] in large, dynamic visualizations, even though the modifications to the scene are much more subtle. This extends findings from prior work by showing that flicker is also an effective attractor in large, dynamic scenes. But given that users rated flicker as disturbing in prior work [16], can we find a way to make it more appealing for narrative visualizations, where observers will be exposed to visual guidance for a longer time?

5 ATTRACTIVE FLICKER

The goal of Attractive Flicker is to exploit the effectiveness of luminance flicker for visual guidance, as demonstrated in the preceding pilot experiment, while decreasing the disturbance of the signal for a long-term exposure. The basic idea is to split the signal into two distinct stages, corresponding to the two visual attention stages by human observers described by Healey and Enns [15]: In the initial *orientation stage*, a short but sufficiently strong signal effectively guides the user's attention to the focus element. In the subsequent *engagement phase*, the signal strength is smoothly reduced and remains at a constantly low level to minimize disturbance. What needs to be investigated is how long the initial orientation stage needs to be, as well as the flicker parameters for the two stages.

While the color of the focus element (given as L^*, a^*, and b^* $\in [0, 100]$) and the time t that passed since the element gained focus are defined by the narrative visualization, Attractive Flicker has two properties that define the signal's perceived saliency, as well as its annoyance: amplitude and frequency. Since Attractive Flicker is split into two stages with different signal strengths, there are six parameters that define the overall signal: peak-to-peak amplitudes A_o and A_e (both $\in [0, 100]$) for the orientation and engagement stage, respectively, the periods T_o and T_e (where T is the reciprocal of the signal's frequency f, given in milliseconds), and durations d_o of the orientation stage and d_t of a smooth transition period between orientation and engagement signal in milliseconds.

The luminance value $L^*(t)$ for a focus element subject to Attractive Flicker with a given L*a*b* color at time t is described by a wave
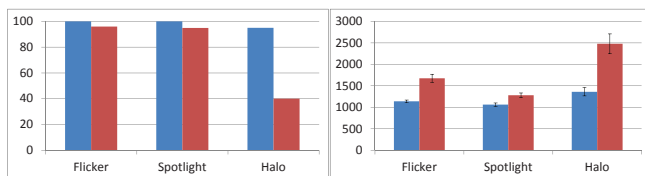

Fig. 3. Success rate in % for (para)foveal (blue) and peripheral (red) targets (left) and average reaction time in milliseconds with standard error bars (right).
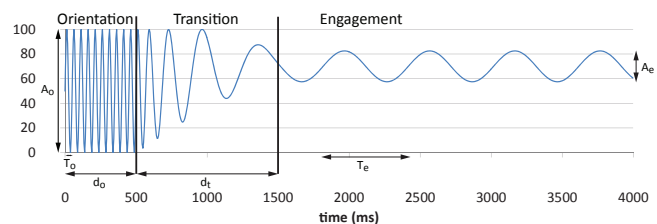

Fig. 4. Exemplary Attractive Flicker signal (four seconds) of a focus element of luminance $L^* = 70$ with parameters $A_o = 100$, $T_o = 50ms$, $A_e = 25$, $T_e = 600ms$, $d_o = 500ms$, and $d_t = 1000ms$.

(a) (para)foveal: $\beta_0 = -16.348, \beta_1 = 3.058, \beta_2 = 0.69$

(b) peripheral: $\beta_0 = -17.335, \beta_1 = 2.375, \beta_2 = 1.022$

(c) measurements

(Para)foveal

	0	25	50	75	100
250	0	0,05	0,45	0,4	0,85
500	0	0,25	0,5	0,85	0,9
1000	0	0,1	0,65	0,85	0,85
1500	0	0,15	0,6	0,8	1
2000	0,05	0,15	0,7	0,95	1

Peripheral

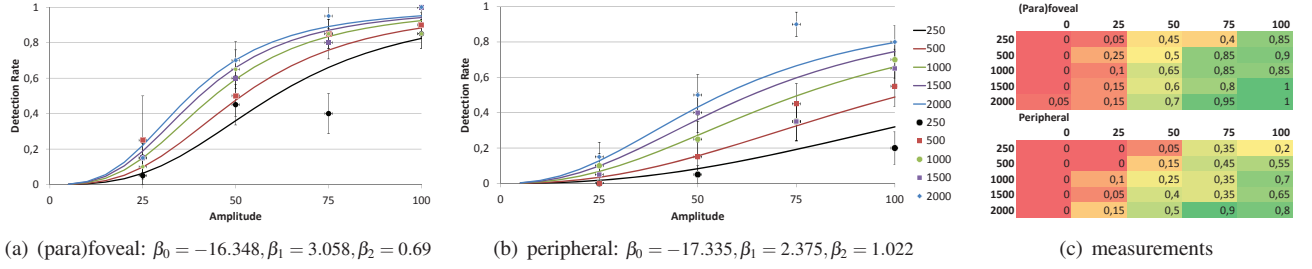|	0	25	50	75	100
250	0	0	0,05	0,35	0,2
500	0	0	0,15	0,45	0,55
1000	0	0,1	0,25	0,35	0,7
1500	0	0,05	0,4	0,35	0,65
2000	0	0,15	0,5	0,9	0,8

Fig. 5. Models predicting the detection rate for (a) (para)foveal and (b) peripheral targets, along with the estimated coefficients for Equation 3. Separate lines show the predicted detection rates per tested duration. Dots show the measured samples with standard error bars. (c) Average detection rates for (para)foveal (top) and peripheral targets (bottom), with durations as rows and amplitudes as columns.

function:

$$L^*(t) = L^{*\prime}(t) + \frac{A(t)}{2} sin\left(2\pi \int_0^t \frac{1}{T(\tau)} d\tau\right), \quad (1)$$

where $A(t)$ is the peak-to-peak amplitude and $T(\tau)$ is the period at time $\tau \leq t$. In the transition period, the amplitude $A(t)$ and the period $T(t)$ with $d_o \leq t < d_o + d_t$ are linearly interpolated between A_o and A_e and T_o and T_e, respectively. The center of the wave amplitude $L^{*\prime}$ at a given time t is defined as:

$$L^{*\prime}(t) = \begin{cases} min\{L^* + \frac{A(t)}{2}, 100\} - \frac{A(t)}{2}, & \text{if } L^* \geq 50 \\ max\{L^* - \frac{A(t)}{2}, 0\} + \frac{A(t)}{2}, & \text{otherwise,} \end{cases} \quad (2)$$

which assures that all colors have the same flicker amplitude, irrespective of the element's luminance. Fig. 4 illustrates an exemplary Attractive Flicker signal. Candidates for the signal parameters A_o, T_o, A_e, T_e, d_o, and d_t were determined in two perceptual studies, presented in the following sections.

6 PERCEPTUAL STUDY 1: ATTENTION ATTRACTION

The goal of the first perceptual study was to empirically determine the minimally required flicker amplitude A_o and duration d_o to reliably guide the user's attention to both, (para)foveal and peripheral display regions, in the orientation stage. For this purpose, users were asked to participate in a visual search task for a flickering target on a large display containing 1000 randomly moving dots. By systematically varying the flicker amplitude of the target and the stimulus duration, our aim was to find candidates for these parameters so that the detection rate is above a pre-defined threshold of 90%. The hardware setup and scene settings were identical as in the comparative pilot study (Section 4).

6.1 Participants

Ten participants (aged 25 to 37, one female) were recruited from a local university, including research scientists in computer science and technical assistants.

6.2 Design

We employed a $2 \times 5 \times 5$ within-subjects design with the following factors:

Eccentricity of the target (either *(para)foveal* or *peripheral*, as described in Section 4),

amplitude A of the luminance modulation (as described in Equation 1), defined by the peak-to-peak modulation range between 0 and 100 in CIELab space (0, 25, 50, 75, 100, where 0-values represented target-absent trials), and

duration d of the stimulus presentation in milliseconds (250, 500, 1000, 1500, 2000).

The duration values were chosen so that the shortest duration was slightly longer than a minimum dwell period (≥ 200 ms) and that the maximum tested duration was clearly longer than saccadic movements

followed by head movements (≥ 500 ms) [43]. Amplitude values covered a range from borderline visibility determined in an informal pilot test ($A = 25$) to the full possible luminance amplitude ($A = 100$).

Users had to perform two repetitions, resulting in 100 trials in total, where the order of appearance of the 50 different stimuli was randomized. In each stimulus, there was exactly one blinking dot — or none for target-absent trials. Flicker frequency was fixed to 20 Hz, since informal pilot tests indicated that slow blinking was generally much harder to detect (see also [4, 11]).

6.3 Procedure

Before each trial, the user pressed the Return-key to start a two-second countdown animation in the center of the screen. Users were instructed to fixate the countdown animation until the stimulus would appear. Each stimulus was shown according to the duration factor, *i.e.*, between 250 and 2000 milliseconds. After this duration, the image was frozen and darkened. Users then had to mark the approximate area where they perceived some flicker before the stimulus end with the mouse cursor, or press "n" in case they did not perceive any flicker. We logged *detection* (0 or 1) and, if the flickering dot was detected, *accuracy* by the distance of the user's indicated flicker location and the actual position of the target in pixels.

6.4 Results

We first performed a Wilcoxon Signed-Rank test to compare average detection rates between (para)foveal and peripheral display regions. We found a significant difference for eccentricity ($Z = -2.803, p = .005$), with the average detection rate being 43% lower in peripheral display regions than in the (para)foveal region. All subsequent analyses were performed separately for (para)foveal and peripheral regions.

To analyze the effect of flicker amplitude A and signal duration d on the detection rate, we performed a binary logistic regression with logarithmic A and d as predictors:

$$dr(A, d) = \frac{exp(\beta_0 + \beta_1 ln(A) + \beta_2 ln(d))}{1 + exp(\beta_0 + \beta_1 ln(A) + \beta_2 ln(d))} = \frac{1}{1 + e^{-\beta_0} A^{-\beta_1} d^{-\beta_2}}. \quad (3)$$

The estimated coefficients and the resulting predictions are shown in Fig. 5. The Nagelkerke-R^2 goodness-of-fit measures indicate moderate (.472 for the (para)foveal model) and modest (.35 for the peripheral model) fits of model to data, respectively.

As indicated in Fig. 5, for the (para)foveal display regions, our model predicts a detection rate of $\geq 90\%$ with flicker duration $d \geq 650$ and amplitude $A = 100$. In peripheral display regions, the maximum predicted detection rate for the maximum flicker duration $d = 2000$ and amplitude $A = 100$ is slightly below 80%.

Accuracy of target location estimation was fairly high, with 74.28% of all detected targets being directly selected. For those detected dots that were not accurately selected, the average distance to the target was 146.3 pixels. For (para)foveal display regions, users could accurately determine the location of all detected targets for all trials with duration $d \geq 1000$, independently of the amplitude. For peripheral targets,
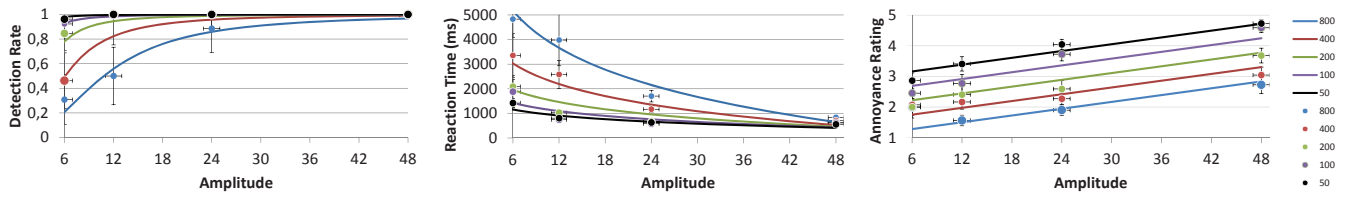
Fig. 6. Models predicting the (left) detection rate, (middle) reaction time, and (right) annoyance. Separate lines show the predicted models per tested periods. Dots show the measured samples with standard error bars.

only 67.8% of targets could be accurately spotted, while the average distance to the remaining targets was 180.7 pixels.

7 PERCEPTUAL STUDY 2: MINIMALLY DISTURBING FLICKER

The goal of the second perceptual study was to investigate the trade-off between flicker effectiveness and subjective annoyance in the engagement stage by investigating two parameters: flicker amplitude A_e and period T_e. In a visual search task, we aimed to find promising amplitude and period ranges by evaluating a cost-benefit function which trades off a high detection rate (benefit) against increased reaction time and annoyance (cost).

7.1 Participants

Eleven participants (aged 23 to 33, four females) were recruited from a local university. All participants were master students or postgraduates of computer science.

7.2 Apparatus

The stimuli were presented in a 300×300 pixels window on the central monitor of Fig. 1. The size of this window was chosen so that the search radius would correspond to the average distance to targets that could not be uniquely identified in the first perceptual study (~ 150 pixels). The window contained 12 randomly moving and colliding dots of same size, colors, and movement speed as in the previous studies.

7.3 Design

We employed a 5×5 within-subjects design with the following factors:
Amplitude A of the luminance modulation, defined by the peak-to-peak modulation range between 0 and 100 in CIELab space (0, 6, 12, 24, 48, where 0-values represented target-absent trials), and
period T of the flicker signal in milliseconds (800, 400, 200, 100, 50). The period value 50 corresponds to the flicker frequency $f = 20$ Hz used in the first perceptual study.

Users had to perform two repetitions, resulting in 50 trials in total, where the presentation order of the 25 stimuli was randomized. Color and spatial location of the target dot within the window was picked randomly.

7.4 Procedure

As in the previous perceptual study, the users could initiate a new trial by hitting the Return-key, followed by a two-second countdown visualization. Then, the stimulus was shown for a maximum duration of 10 seconds. Users were asked to hit the space key as soon as they spotted a flickering dot. After hitting the space key or reaching the time-out, the stimulus was frozen and darkened, and users had to pick the previously flickering dot with the mouse cursor, or press "n", if they did not see any flickering. If a dot was selected, the stimulus was restarted and users were asked to watch the scene until they felt confident to rate the annoyance of the signal with a value between 1 (not annoying at all) and 5 (extremely annoying) on the numpad. We instructed the users that annoyance should be rated with the assumption that the signal would be shown to them for a longer time, following a methodology by Bartram *et al.* [4]. For each trial, we logged target *detection* (0 or 1) and, if the flickering dot was detected, *reaction time*, the subjective *annoyance* rating, and *correctness* by evaluating the distance of

the picked dot location to the actual target dot center, where all picked locations ≤ 25 pixels from the target's center were counted as correct.

7.5 Results

To find a good trade-off between detection, annoyance, and reaction time, we defined a cost-benefit function. We optimized the cost-benefit subject to the desired constraint that the detection rate is ≥ 0.9, according to the reliability threshold we used in the first perceptual study:

$$cb(A,T) = \begin{cases} \frac{dr(A,T)}{an(A,T) \cdot rt(A,T)}, & \text{if } dr(A,T) \geq 0.9, \\ 0, & \text{otherwise,} \end{cases}$$

where $dr(A,T)$, $rt(A,T)$, and $an(A,T)$ are predicting the detection rate, reaction time, and annoyance, respectively, as a function of A and T, as described below:

The detection rate $dr(A,T)$ is described by a binary logistic regression, as given in Equation 3. The Nagelkerke-R^2 goodness-of-fit measure reveals that the model fits moderately well to the data (.51). The coefficients ($\beta_0 = 7.192, \beta_1 = 2.268, \beta_2 = -1.884$) indicate that amplitude is a slightly more reliable predictor of the detection rate than the signal period. Our model predicts that for all our tested amplitudes $A \geq 6$, we could achieve 90% detection rate for periods $T \leq 100$. For amplitudes $A \geq 30$, all periods lead to a detection rate above 90%. Fig. 6 (left) visualizes the detection rate model, as well as the experiment samples.

Of all detected targets, 92.2% were correct. The analyses of reaction times and annoyance were performed only for the correctly detected targets:

To model the reaction time, the best-fitting model ($R^2 = .426$) was achieved by a regression, with the following estimated coefficients:

$$rt(A,T) = 1317.423 - 238.143 \cdot ln(A) + 9.682 \cdot T - 2.42 \cdot ln(A) \cdot T.$$

As visualized in Fig. 6 (middle), the reaction time decreases with logarithmically increasing amplitude and linearly decreasing period. Also, there is an interaction between amplitude and period. While reaction times strongly deviate for small amplitudes depending on the frequency, they all converge to around 500 milliseconds for the highest measured amplitude $A = 48$.

Annoyance was modeled by a regression with the following best-fitting model ($R^2 = .359$) and estimated coefficients:

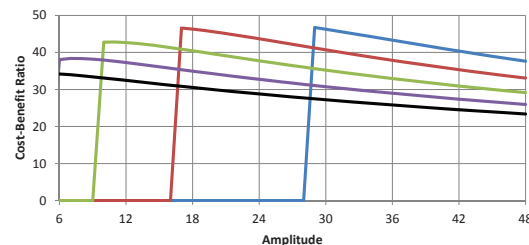$$an(A,T) = 5.598 + 0.037 \cdot A - 0.679 \cdot ln(T).$$

Fig. 7. Cost-benefit function for flicker amplitudes and tested periods (separate lines); higher values are better.

The model predicts that annoyance increases with linearly increasing amplitude and logarithmically decreasing period (*cf.*, Fig. 6 right).

From observing Fig. 7, we can determine that the maximum of the cost-benefit function is reached for periods between $T = 400$ and $T = 800$ with amplitudes between $A = 17$ and $A = 29$.

8   DISCUSSION OF PERCEPTUAL STUDY RESULTS

Like the initial pilot experiment, our first perceptual study showed that the detection of a flickering target is more reliable in (para)foveal display regions as compared to peripheral regions. But even in the (para)foveal region, only scene elements flickering with the highest possible luminance amplitude of $A = 100$ could reliably attract the visual attention of the observer within 500 milliseconds in our complex scene. It seems that with lower amplitudes, the popout effect of flickering dots in our scene was too weak to be perceived pre-attentively. In peripheral display regions, we could not reach our desired 90% detection rate within our tested stimulus durations of $d \leq 2000$ milliseconds. This implies that luminance flicker alone may not be sufficient for reliably attracting the attention to peripheral display regions within the first two seconds of stimulus onset. In the future, it will therefore be necessary to investigate different flicker features, or combinations of features, to guide the attention more effectively to these peripheral areas.

If detected, users could point at target locations fairly accurately, even after short flicker durations. Unsurprisingly, the accuracy increased with the stimulus duration. On the peripheral monitors, users often could not indicate the exact target, but were able to identify a close area around it. This is sufficient, since our two-stage signal should take care that the user's gaze will be guided to the correct target within this small search window after the initial orientation stage.

For the small search window tested in the second experiment, we found that a much smaller amplitude and lower frequency is sufficient to be able to reliably detect a target. However, with increasing detection rate and shorter reaction times, subjective annoyance is also increasing. In other words: the more effective a signal, the more disturbing it is perceived. This is consistent with prior experiments [4, 11, 16]. Our results indicate that a good trade-off between detection reliability and annoyance can be found for the low frequency and medium amplitude ranges tested in our experiment.

Our results provide some guidance for the choice of suitable amplitude and frequency candidates for both flicker stages, as well as the minimum duration of the orientation stage. However, there are several factors in a visualization that will influence these candidate parameters:

First, while we varied the dots' colors and motion directions in our experiments, they all had the same size, each stimulus contained the same number and density of dots, and the distance of the user to the three displays was quite homogeneous. It can be expected that the size of the target has a strong influence on the minimally required amplitude and frequency parameters for reliable detection. Similarly, the much higher detection rates in the second perceptual study suggest that the size of the search window and the number of context elements, respectively, play a major role in the perceived flicker strength. Systematically exploring how the interplay between focus and context size influences the flicker parameters will be an important future work.

Second, the dots in our scene were moving in random directions, but with homogeneous velocities. Variations of movements speeds are likely to also influence the popout effect of flickering targets in the orientation stage.

Third, due to the task descriptions and the high number of repetitions, users' top-down attention was tuned to detect flicker, which is an inherent property of classic target detection experiments (*cf.*, [15]). However, when exploring a guided narrative visualization, we assume that users will also learn to expect a certain repeating popup feature.

Forth, the relative annoyance ratings of the second perceptual study only relate to the signal itself, not to the signal in context of a real task. In real scenarios, the annoyance caused by an attractor also depends on its perceived usefulness. Therefore, researchers have recommended to adjust the attractor strength to its interruption importance [11]. In the case of Attractive Flicker, the duration of the orientation stage, as well as amplitude and frequency of the engagement stage could be adjusted to the predicted degree of interest of the target.

Finally, we experimented with amplitude and frequency for luminance flicker. Even though we expect similar results for parameters of alternative flicker features, more empirical evidence is required in this regard.

9   EXPERIMENT: MOLECULAR STORIES

We demonstrate the effectiveness of Attractive Flicker in an exemplary showcase of molecular visualization. According to Jenkison and McGill [21], a more complex representation of a molecular scene — showcasing a lot of details like random motion or molecular crowding — offers a better understanding of the visualized process than a more simplistic abstraction. This strongly motivates the use of guided narrative storytelling in molecular dynamic simulations with a very detailed molecular scenery, while the observer is visually guided through the narrative of a biochemical process.

Complex systems like these are often simulated by agent-based models. In such systems, events take place punctually at distinct times and locations. To be able to grasp the visualization's narrative, the user needs visual guidance to track the ongoing process in focus, but also requires basic understanding of the embedding surroundings.

As an example scenario of the molecular machinery, we simulate and visualize a simplified version of the *nicotinamide adenine dinucleotide* (NAD) synthesis process in the *salvage pathway*. NAD is a molecule present in the cells of all living organisms, mostly used in electron transfer reactions.

To assess the effectiveness of Attractive Flicker in such a scenario, we compare it to two baseline conditions: to the unmodified visualization and a radical context masking, where only the elements in focus are being displayed. We have two hypotheses: First, we hypothesize that using Attractive Flicker, users will be able to follow the visualization's narrative as well as if only the focus elements were displayed without any context. Second, we hypothesize that Attractive Flicker will be minimally distracting, so perception of contextual information will be as good as with the unmodified scene without any visual guidance.

9.1   Participants

Twelve paid participants (aged 23 to 32, four females), including computer scientists, students, and teachers of different disciplines, participated in the study.

9.2   Apparatus

The experiment was conducted on the three-monitor setup shown in Fig. 1. In a 7680×1440 pixels full-screen window, we displayed a 3D molecular scene containing four different types of enzymes and ten different types of metabolites of the NAD synthesis process. Since metabolites were quite small and had similar shapes, we used an adapted qualitative color-coding from *ColorBrewer* [13] to distinguish them. Enzymes were rendered in grey. We set the quantities of all but two metabolites to 100 elements per scene. One metabolite quantity was set to 800, one reaction product was set to 0 initially. In total, the scene contained ~ 2000 elements (*cf.*, Fig. 8).

In each trial, the simulation triggered a chain of three reactions out of the salvage pathway, which was replayed four times. The first reaction started approximately five seconds after the trial initiated. Between each reaction, there was a break of 0 to 15 seconds. Since metabolites and enzymes were subject to random forces, the duration of the reactions varied. In total, each trial scene was displayed for four to five minutes.

To ensure the visibility of reactions, the depth of the volume was limited to avoid occlusions of enzymes. Whenever an enzyme was selected as reaction catalyst, we rotated the enzyme so its binding site was facing towards the user and the docking metabolites were clearly visible.
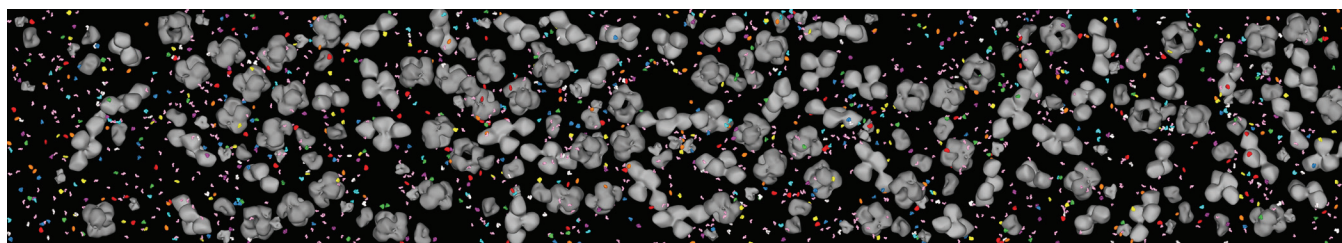
Fig. 8. Screenshot of the dynamic visualization showing the molecular machinery with four different enzymes and ten different metabolites.

9.3 Design

Since we were interested in the users' insights generated from the displayed visualization, we evaluated this experiment by means of content questionnaires. We designed two classes of questions: The *focus questions* targeted towards the chain of reactions. For all three reactions in each trial, users had to identify the metabolites acting as substrates and the resulting products. In addition, they had to note the order of the three reactions. The *context questions* addressed overall scene properties, irrespective of the reactions taking place. For each trial, users had to find the metabolite with most and least instances in the scene, respectively. Since the molecules were presented purely visually without any textual labels, users had to select the correct metabolites by picking their color from a list of all 10 metabolite colors used in the scene, printed on the questionnaires.

We employed a within-subjects design with three visual guidance conditions:

In the **Normal** condition, we showed the unmodified scene without any visual hints supporting the detection of ongoing reactions.

In the **Attractive Flicker** condition, the material color of the reaction partners (*i.e.*, the enzyme and one or two substrates) was oscillated using our Attractive Flicker technique. According to the findings in our perceptual studies, we displayed the orientation phase with full amplitude $A_o = 100$ and a short period $T_o = 50$ for 500 milliseconds. For the subsequent engagement period, we used an amplitude of $A_e = 25$ and a period of $T_e = 600$. The two signals were connected by a transition period of 1500 milliseconds.

In the **No Context** condition, we displayed only the reaction partners and completely masked the remaining scene. This can be compared to an extreme *spotlight* effect (*cf.*, Section 4). Mind that we did not issue the context questions for the *No Context* condition.

The presentation order of the three conditions was counterbalanced. In addition, we constructed two scenes with slight variations of the salvage pathway, and changed the color assignments to the metabolites for each scene. The assignment of these three resulting scenes to the three visual guidance conditions was also counterbalanced. While the color-coding and quantities of the metabolites, as well as the chain of reactions, were pre-defined for each scene, the spatial locations of the reactions on the display were picked randomly.

9.4 Procedure

Before starting the first trial, we showed the user a short demonstration scene in the *No Context* condition to explain the visual appearance of the individual reactants, as well as how a reaction would look like. Before each trial, users were issued a short textual description of the current visual guidance condition. We avoided any mentioning of the term "flicker" in these scene descriptions, and consistently used the term "highlighting" instead, to avoid tuning the top-down attention of participants to flickering elements. Then, users could first conduct a test run with a training data set. After the experiment, we asked users to rate the task difficulty for both, focus and context questions, for each visual guidance technique on a five-point Likert scale. Finally, we gathered qualitative feedback in an audio-recorded semi-structured interview by asking for more details about the subjective questionnaire ratings and whether users had any additional comments.

9.5 Results

We will report the correctness of the answers for focus questions, as well as their perceived difficulty, for all three guidance conditions. Results for the context questions and their perceived difficulty will only be compared between *Attractive Flicker* and the *Normal* scene.

9.5.1 Focus Questions

To determine the correctness of the noted reactions, we counted the number of correctly identified reactants, where a maximum of four points could be scored for each reaction. A correct sequence was scored when two subsequent reactions were reported in their correct order, which leads to a maximum value of two correct sequences for the three reactions per trial.

We performed a one-way ANOVA with Bonferroni-corrected post-hoc comparisons on correctness of reaction partners across the three visual guidance conditions. We found a significant difference between the three conditions ($F_{1.113,12.245} = 580.833, p < .001, \eta^2 = .981$). Post-hoc comparisons revealed that both, *No Context* and *Attractive Flicker* (both 99.3% correctness), yielded significantly better results than *Normal* (6.9% correctness). Similarly, we found a significant difference for the correctness of reaction sequences ($F_{2,22} = 209.0, p < .001, \eta^2 = .95$). *Attractive Flicker* (87.5% correctness) led to a significantly better result than the *Normal* scene. In the *No Context* condition, all sequences were reported correctly, while in the *Normal* scene, no correct sequence was found.

The post-experiment questionnaire was evaluated using a Friedman test with Bonferroni-corrected Wilcoxon Signed-Rank post-hoc comparisons. The ease of finding reactions was rated significantly differently for the three conditions ($\chi^2(2) = 21.565, p < .001$): The average score for *No Context* of 4.92 was better than for *Attractive Flicker* with an average score of 4.08 ($Z = -2.486, p = .013$). *Normal* reached the lowest average score with 1.17, which is significantly different from both, *No Context* ($Z = -3.217, p = .001$) and *Attractive Flicker* ($Z = -3.108, p = .002$).

In the post-experiment interview, all users stated that finding reactions in the *Normal* scene was *"not or close to impossible"*. They also reported several reasons why finding and tracking reactions was slightly harder using *Attractive Flicker* than *No Context*: distraction and occlusions caused by surrounding molecules, too subtle flickering of the small metabolites, and difficulties when re-acquiring the reaction partners after switching the attention to the questionnaire. Therefore, three users suggested to add pulsating halos around the focus metabolites to put more visual emphasis especially on small targets.

These results suggest that our first hypothesis is supported: Users could answer questions about the story in focus as accurately as with a scene completely masking the context. However, it was less demanding for the users to follow the story with complete context masking.

9.5.2 Context Questions

For the quantity estimations, users could score one point for each trial for both, the highest and lowest quantity estimation. We compared the estimations using a Wilcoxon Signed-Rank test. We neither found a significant difference for the highest quantity estimations ($Z = -1.000, p = .317$ with 100% correctness for *Normal* and 91.7%

correctness for *Attractive Flicker*), nor for the lowest quantity estimations ($Z = -.577, p = .564$, with 91.6% correctness for *Normal* and 83.3% correctness for *Attractive Flicker*).

This is also reflected in the post-experiment questionnaires. A Wilcoxon Signed-Rank test revealed that users rated the ease of finding the highest and lowest concentrations equally easy ($Z = -.378, p = 0.705$), with an average score of 4.25 for the *Normal* scene and 4.17 for the *Attractive Flicker* scene. In the interview, eight users reported that the flickering did not distract them when estimating the quantities (for instance, User 11: *"The flickering, I can simply ignore."*). No user reported that the flickering reactions were disturbing when observing the scene context.

These findings confirm our second hypothesis: Attractive Flicker adds negligible distraction when concentrating on the visualization's context.

10 DISCUSSION

With these experimental results, we will first discuss the design of Attractive Flicker, followed by an analysis of additional use cases and open research questions.

10.1 Two-Stage Signal

We proposed a two-stage signal to decrease the strong attention guidance effect of flicker to a persistent signal with significantly lower subjective disturbance. We showcased this approach with luminance flicker to guide the user through a dynamic narrative visualization on a large display. However, this proposed two-stage design is not limited to luminance flicker. The same principle can be applied for all different sort of visual attractors that increase the visual emphasis of the focus object. For instance, when applying the two-stage concept to halos around focus elements, a large and clearly visible bright highlight around the focus object will dynamically shrink to a small, subtle halo in the engagement stage. Similarly as demonstrated for luminance flicker in this paper, the required parameters for the orientation and engagement stages (*e.g.*, radius and brightness difference for the two-stage halo), need to be explored empirically.

In our first perceptual study, we explored how long the orientation stage needs to last for the flicker attractor to be reliably perceivable and applied our findings in a case study of molecular stories. Although users could easily follow the story when guided by Attractive Flicker, some users reported that they were afraid to miss the initial signal in some situations. It seems that, even though the orientation stage signal can be perceived reliably, users sometimes would like to have a longer orientation stage in realistic scenarios. Another potential way to set a useful orientation stage duration could be a combination of Attractive Flicker with an eye tracking approach (as proposed by Bailey *et al.* [2]) or even a coarse head tracking technology using an off-the-shelf webcam. This way, the transition phase would be initiated only after the user directs the gaze towards the highlighted target.

10.2 Narrative Visualizations and Other Use Cases

The field of narrative visualization has generated increased interest in recent years [46, 37, 27]. According to Segel and Heer [37], *highlighting* to guide the user's attention to elements of interest is one of three *visual narrative tactics*, besides providing the user with an overview (*visual structuring*) and *transition guidance* to change the visualization viewpoint or representation without disorienting the user. These tactics can co-exist in parallel. For instance, animated transitions can resolve occlusions, while visual highlighting guides the user's gaze to the most relevant regions while preserving the context. The larger the scene, the more important highlighting is to guide the user's attention through the narrative.

We demonstrated the usefulness of Attractive Flicker as highlight technique for large, dynamic scenes. In our use case, observers could as accurately follow the narrative of a story as if only the story without any context was being displayed. Notably, our users consistently reported that Attractive Flicker was easy to ignore when observing the context of the scene. Since only for the initial orientation stage we chose amplitude and frequency parameters high enough to generate

a popout effect on the large display, the potential source of distraction was limited to less than a second. However, when guiding a user through a smaller or static visualization, static attractors (like, for instance, a halo as in the comparative pilot experiment) are probably sufficiently distinct to generate a popout effect and will therefore cause a lower annoyance. On the other hand, when dealing with volume data, simply flickering a region's luminance is not possible to make it stand out, due to potential occlusions. In such a scenario, a visual guidance technique can be implemented by utilizing and extending the concept of temporal transfer functions as proposed in the scientific visualization literature [28].

Narrative visualization is not only used to passively communicate a pre-defined story, as in our use case, but it is also an important tool to guide interactive exploration of complex information. For instance, when constantly analyzing a large amount of real-time data, such as during emergency management or stock broking, Attractive Flicker can help to make new and important data visually stand out in a minimally obtrusive way. Similarly, Attractive Flicker could serve as indicator of visual changes caused by user's brushing actions in large-scale multiple coordinated views. For this application area, it will be necessary to explore the effectiveness of multiple synchronous Attractive Flicker targets, potentially with different strengths matched to the predicted degrees of interest of the elements.

Outside of the visualization field, educational animations, for instance, are in general controversial since learners have difficulties drawing their attention to semantically relevant regions without visual guidance [26, 24]. Traditionally, attention guidance also plays an importance role in desktop interfaces to effectively deliver notifications (*cf.*, [4, 11, 16]) and advertisement. These application areas could also benefit from a technique that initially guides the attention for a short period of time, before decaying into a less distracting persistent state.

11 CONCLUSION

We presented Attractive Flicker — a visual attention guidance technique for large, complex visualizations based on luminance oscillation of focus elements. The novelty of our technique is that the flicker signal is split into two stages, inspired by perceptual attention stages of human observers. We could empirically determine that the initial orientation stage does not need to be longer than 500 milliseconds — *i.e.*, below the average duration of a coordinated eye-head movement [43] — but very salient to reliably attract the user's attention. We also found that small targets in the periphery cannot fully reliably attract the user's attention with luminance flickering alone. For the subsequent engagement stage, we observed that an amplitude covering one quarter of the luminance range and a frequency below 2 Hz represents a good trade-off between signal effectiveness and subjective annoyance. In a showcase scenario of molecular interactions on a large display, we could verify that Attractive Flicker indeed adds minimal distraction, but is still a highly effective technique to guide the user through a narrative visualization.

In the future, we want to systematically investigate more scene parameters potentially influencing the effectiveness of Attractive Flicker, such as focus and context size and heterogeneity of scene motion or object shapes. In addition, we plan to extend our proposed two-stage model to other classes of visual guidance techniques, and explore combinations of flickering and static techniques.

REFERENCES

[1] Gapminder: Unveiling the beauty of statistics for a fact based world view. *http://www.gapminder.org/*, 2014.

[2] R. Bailey, A. McNamara, N. Sudarsanam, and C. Grimm. Subtle gaze direction. *ACM Trans. Graph.*, 28(4):1–22, Sept. 2009.

[3] L. Bartram and C. Ware. Filtering and brushing with motion. *Information Visualization*, 1(1):66–79, Mar. 2002.

[4] L. Bartram, C. Ware, and T. Calvert. Moticons:: detection, distraction and task. *International Journal of Human-Computer Studies*, 58(5):515–545, May 2003.

[5] A. Borji and L. Itti. State-of-the-art in visual attention modeling. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 35(1):185–207, 2013.

[6] F. Cole, D. DeCarlo, A. Finkelstein, K. Kin, K. Morley, and A. Santella. Directing gaze in 3d models with stylized focus. In *Proceedings of the 17th Eurographics Conference on Rendering Techniques*, EGSR'06, pages 377–387. Eurographics Association, 2006.

[7] B. B. de Koning, H. K. Tabbers, R. M. Rikers, and F. Paas. Attention guidance in learning from a complex animation: Seeing is understanding? *Learning and Instruction*, 20(2):111–122, Apr. 2010.

[8] H. Doleisch, M. Gasser, and H. Hauser. Interactive feature specification for focus+context visualization of complex simulation data. In *Proceedings of the symposium on Data visualisation 2003*, VISSYM '03, page 239248, Aire-la-Ville, Switzerland, Switzerland, 2003. Eurographics Association.

[9] E. N. Dzhafarov, R. Sekuler, and J. Allik. Detection of changes in speed and direction of motion: Reaction time analysis. *Perception & Psychophysics*, 54(6):733–750, Nov. 1993.

[10] G. W. Furnas. Generalized fisheye views. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, CHI '86, pages 16–23. ACM, 1986.

[11] J. Gluck, A. Bunt, and J. McGrenere. Matching attentional draw with utility in interruption. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '07, pages 41–50. ACM, 2007.

[12] C. Gutwin and C. Fedak. A comparison of fisheye lenses for interactive layout tasks. In *Proceedings of the 2004 Graphics Interface Conference*, GI '04, pages 213–220. Canadian Human-Computer Communications Society, 2004.

[13] M. Harrower and C. A. Brewer. ColorBrewer.org: An online tool for selecting colour schemes for maps. *The Cartographic Journal*, 40(1):27–37, June 2003.

[14] H. Hauser. Generalizing focus+context visualization. In *Scientific Visualization: The Visual Extraction of Knowledge from Data*, Mathematics and Visualization, pages 305–327. Springer Berlin Heidelberg, 2006.

[15] C. G. Healey and J. Enns. Attention and visual memory in visualization and computer graphics. *IEEE Transactions on Visualization and Computer Graphics*, 18(7):1170–1188, 2012.

[16] R. Hoffmann, P. Baudisch, and D. S. Weld. Evaluating visual cues for window switching on large screens. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '08, pages 929–938. ACM, 2008.

[17] D. Huber and C. G. Healey. Visualizing data with motion. In *IEEE Visualization, 2005. VIS 05*, pages 527–534, 2005.

[18] L. Itti. Quantifying the contribution of low-level saliency to human eye movements in dynamic scenes. *Visual Cognition*, 12(6):1093–1123, 2005.

[19] L. Itti, N. Dhavale, and F. Pighin. Realistic avatar eye and head animation using a neurobiological model of visual attention. In *Optical Science and Technology, SPIE's 48th Annual Meeting*, volume 5200, pages 64–78. International Society for Optics and Photonics, 2004.

[20] L. Itti, C. Koch, and E. Niebur. A model of saliency-based visual attention for rapid scene analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(11):1254–1259, 1998.

[21] J. Jenkinson and G. McGill. Visualizing protein interactions and dynamics: Evolving a visual language for molecular animation. *CBE-Life Sciences Education*, 11(1):103–110, Mar. 2012.

[22] A. Khan, J. Matejka, G. Fitzmaurice, and G. Kurtenbach. Spotlight: Directing users' attention on large displays. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '05, pages 791–798. ACM, 2005.

[23] Y. Kim and A. Varshney. Saliency-guided enhancement for volume visualization. *IEEE Transactions on Visualization and Computer Graphics*, 12(5):925–932, Sept. 2006.

[24] B. B. d. Koning, H. K. Tabbers, R. M. J. P. Rikers, and F. Paas. Towards a framework for attention cueing in instructional animations: Guidelines for research and design. *Educational Psychology Review*, 21(2):113–140, June 2009.

[25] R. Kosara, S. Miksch, and H. Hauser. Focus+context taken literally. *IEEE Computer Graphics and Applications*, 22(1):22–29, Jan. 2002.

[26] R. K. Lowe. Animation and learning: selective processing of information in dynamic graphics. *Learning and Instruction*, 13(2):157–176, Apr. 2003.

[27] K.-L. Ma, I. Liao, J. Frazier, H. Hauser, and H.-N. Kostis. Scientific storytelling using visualization. *IEEE Computer Graphics and Applications*, 32(1):12–19, Jan. 2012.

[28] R. Maciejewski, I. Woo, W. Chen, and D. Ebert. Structuring feature space: A non-parametric method for volumetric transfer function generation. *IEEE Transactions on Visualization and Computer Graphics*, 15(6):1473–1480, Nov. 2009.

[29] A. McNamara, T. Booth, S. Sridharan, S. Caffey, C. Grimm, and R. Bailey. Directing gaze in narrative art. In *Proceedings of the ACM Symposium on Applied Perception*, SAP '12, pages 63–70. ACM, 2012.

[30] R. Minghim and A. R. Forrest. An illustrated analysis of sonification for scientific visualisation. In *Proceedings of the 6th Conference on Visualization '95*, pages 110–117. IEEE Computer Society, 1995.

[31] P. K. Mital, T. J. Smith, R. L. Hill, and J. M. Henderson. Clustering of gaze during dynamic scene viewing is predicted by motion. *Cognitive Computation*, 3(1):5–24, Mar. 2011.

[32] J. Moreland and A. Cruz. Colour perception with the peripheral retina. *Optica Acta: International Journal of Optics*, 6(2):117–151, 1959.

[33] H. E. Petersen and D. J. Dugas. The relative importance of contrast and motion in visual detection. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 14(3):207–216, June 1972.

[34] Y. Pinto, C. N. L. Olivers, and J. Theeuwes. Selecting from dynamic environments: Attention distinguishes between blinking and moving. *Perception & Psychophysics*, 70(1):166–178, Jan. 2008.

[35] R. Rosenholtz. A simple saliency model predicts a number of motion popout phenomena. *Vision Research*, 39(19):3157–3163, Oct. 1999.

[36] A. F. Sanders. Some aspects of the selective process in the functional visual field. *Ergonomics*, 13(1):101–117, 1970.

[37] E. Segel and J. Heer. Narrative visualization: Telling stories with data. *IEEE Transactions on Visualization and Computer Graphics*, 16(6):1139–1148, 2010.

[38] L. Shupp, R. Ball, B. Yost, J. Booker, and C. North. Evaluation of viewport size and curvature of large, high-resolution displays. In *Proceedings of Graphics Interface 2006*, pages 123–130. Canadian Information Processing Society, 2006.

[39] M. Steinberger, M. Waldner, A. Lex, M. Streit, and D. Schmalstieg. Context-preserving visual links. *IEEE TVCG 2011*, 17(12):2249–2258, 2011.

[40] V. Sundareswaran, K. Wang, S. Chen, R. Behringer, J. McGee, C. Tam, and P. Zahorik. 3d audio augmented reality: Implementation and experiments. In *Proceedings of the 2Nd IEEE/ACM International Symposium on Mixed and Augmented Reality*, pages 296–297. IEEE Computer Society, 2003.

[41] A. M. Treisman and G. Gelade. A feature-integration theory of attention. *Cognitive Psychology*, 12(1):97–136, Jan. 1980.

[42] E. E. Veas, E. Mendez, S. K. Feiner, and D. Schmalstieg. Directing attention and influencing memory with visual saliency modulation. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '11, pages 1471–1480. ACM, 2011.

[43] C. Ware. *Information Visualization: Perception for Design*. Elsevier, May 2012.

[44] C. Ware and R. Bobrow. Motion to support rapid interactive queries on nodelink diagrams. *ACM Trans. Appl. Percept.*, 1(1):318, July 2004.

[45] C. Ware and R. Bobrow. Motion coding for pattern detection. In *Proceedings of the 3rd Symposium on Applied Perception in Graphics and Visualization*, APGV '06, pages 107–110. ACM, 2006.

[46] M. Wohlfart and H. Hauser. Story telling for presentation in volume visualization. In *Proceedings of the 9th Joint Eurographics / IEEE VGTC Conference on Visualization*, EUROVIS'07, page 9198, Aire-la-Ville, Switzerland, Switzerland, 2007. Eurographics Association.

[47] J. M. Wolfe, K. R. Cave, and S. L. Franzel. Guided search: An alternative to the feature integration model for visual search. *Journal of Experimental Psychology: Human Perception and Performance*, 15(3):419–433, 1989.

[48] S. Zhai, J. Wright, T. Selker, and S.-A. Kelin. Graphical means of directing users attention in the visual interface. In *Human-Computer Interaction INTERACT 97*, IFIP The International Federation for Information Processing, pages 59–66. Springer US, Jan. 1997.