Temporal summation of global form signals in dynamic Glass patterns

Jean-François Nankoo a,⁎, Christopher R. Madan a,⁎, Marcia L. Spetch b, Douglas R. Wylie a,b

a Department of Psychology, University of Alberta, Canada
b Neuroscience & Mental Health Institute, University of Alberta, Canada

A R T I C L E   I N F O
Article history:
Received 15 August 2014
Received in revised form 6 October 2014
Available online 20 November 2014

Keywords:
Glass pattern
Dynamic
Global form
Motion streak
Summation

A B S T R A C T
The ability to perceive complex objects in the environment requires that the visual system integrate local form information into global shapes. Glass patterns (GPs) are stimuli that are commonly used to study this integration process. GPs consist of randomly positioned dot-pairs oriented in a coherent way to create a global form. When multiple GPs are presented sequentially, observers report a percept of illusory coherent motion and have lower detection thresholds relative to a single presentation GPs. The percept of illusory motion has been attributed to the visual system interpreting the dot-pairs in GPs as motion streaks. However, it remains unclear why dynamic GPs are detected at lower thresholds than static GPs. Two main differences exist between static and dynamic GPs: (a) dynamic GPs contain multiple presentations of global form signals compared to a single presentation in static GPs and (b) dynamic GPs have a greater temporal frequency than static GPs. Here we investigated which of these two factors contributed to the heightened sensitivities for dynamic GPs. We systematically varied the number of unique GPs and the rate at which each unique frame is presented (i.e., temporal frequency). The results show that, within the range of temporal frequency used, the primary influence on detection thresholds was the number of unique frames. These results suggest that the improved detection sensitivities can be driven by a mechanism of temporal summation of global form.

© 2014 Elsevier Ltd. All rights reserved.

1. Introduction

Researchers have long established that the mammalian visual system is organized in a modular fashion, whereby different areas are specialized for processing particular types of information (Calabretta & Parisi, 2005). Consistent with this idea, form and motion information are processed by distinct neural pathways at the lower levels of the visual system (Braddick et al., 2000; Livingstone & Hubel, 1988; Milner & Goodale, 1995; Ungerleider & Mishkin, 1982; Van Essen & Gallant, 1994). However, recent psychophysical and neurophysiological studies have demonstrated interactions between the form and motion pathways (see Kourtzi, Krekelberg, & van Wezel, 2008, for review). For instance, in the phenomenon known as structure-from-motion, two-dimensional motion information provides information about the three-dimensional structure of objects (Siegel & Andersen, 1988). In a similar way, form signals have been shown to influence motion perception (Geisler, 1999). For example, Ross, Badcock, and Hayes (2000) have shown that form information constrains incoherent motion to generate the appearance of coherent global motion when multiple independently-generated Glass patterns are presented in rapid succession.

A Glass pattern is a type of static stimulus that consists of an array of randomly-positioned dot-pairs (i.e., dipoles) that are oriented in a way to provide the percept of a global shape (Fig. 1A; Glass, 1969). Glass patterns are commonly used to study how the visual system pools local orientation information to allow us to perceive the global form of objects in the environment, in the same way that random-dot stimuli are used to investigate global pooling of local motion signals (Williams & Sekuler, 1984; Wilson & Wilkinson, 1998). Ross, Badcock, and Hayes (2000) have shown that if a series of independently-generated Glass patterns, with the same global form, are shown in rapid succession, termed dynamic Glass patterns (dynamic GPs), observers perceive a salient illusion of coherent motion. They considered this to be “implied motion” and noted that their participants could not differentiate implied motion from real motion. Furthermore, Krekelberg et al. (2003) and Krekelberg, Vatakis, and Kourtzi (2005) found that cells in the prototypical motion areas of monkeys and humans (medial temporal area [MT] and medial temporal complex [MT+], respectively) do not differentiate between real motion and implied motion. Thus, the results from Krekelberg and colleagues and those...
of Ross, Badcock, and Hayes (2000) provide evidence of an interaction between form and motion, whereby the motion system of the mammalian visual system translates global form information into coherent global motion information.

Various research groups have reported that the detection thresholds for dynamic Glass patterns are significantly lower than the detection thresholds for static Glass patterns (e.g., Burr & Ross, 2006; Or, Khuu, & Hayes, 2007). Because thresholds for real global motion are generally lower than those for global form, the lower thresholds for dynamic GPs relative to static GPs suggest that dynamic GPs are processed in a similar way to real motion. However, based on psychophysical evidence, Nankoo et al. (2012) have suggested that the decrease in thresholds with dynamic GPs is likely related to the form system (e.g., V4). Nankoo et al. (2012) measured the detection threshold for concentric, radial, spiral, horizontal and vertical static GPs, dynamic GPs, and real global motion. They showed that even though thresholds for both dynamic GPs and real motion were significantly lower than static GPs, the relative performance in each of the patterns suggests that the low thresholds of dynamic GPs and real motion are based on different mechanisms. In particular, with real motion, detection thresholds were equivalent for all patterns except for higher thresholds for spiral motion (see also Morrone, Burr, & Vaina, 1995). In contrast, with dynamic GPs, participants were best at detecting concentric and radial patterns, and worst at vertical and horizontal patterns, with spiral at an intermediate detection threshold. The relative ranking of the thresholds for dynamic GPs were identical to the relative ranking of the thresholds for static GPs (see also Wilson & Wilkinson, 1998). Nankoo et al. (2012) argue that this suggests that the decrease in threshold found in dynamic GPs is driven by the same or similar form-related processes that drive the detection of GPs, as opposed to motion-related processes.

Recently, Day and Palomares (2014) reported a negative linear relationship between temporal frequency and coherence threshold in dynamic GPs; as temporal frequency increased, threshold decreased (see also Edwards & Crane, 2007). Day and Palomares (2014) argued that their result is consistent with the idea that the dynamic GPs is processed by the ‘motion streak’ system (Ross, 2004; Ross, Badcock, & Hayes, 2000). The motion streak model is based on the finding that fast-moving objects leave a trailing blur due to temporal integration (Geisler, 1999). At high velocities, the visual system appears to utilize the form from the trailing blur (i.e., streak) to disambiguate direction information (Burr & Ross, 2002). Day and Palomares suggested that if dynamic GPs are interpreted as motion streaks by the visual system, it follows that increasing the temporal frequency would increase sensitivity. However, while Day and Palomares’ study showed the importance of temporal frequency, it does not rule out the possibility that lower detection thresholds for dynamic GPs are also due to the additional form signals present in dynamic GPs. The increase in temporal frequency also means that there is an increase in the number of unique frames presented. Thus, it is unclear whether the increased sensitivity of dynamic GPs relative to static GPs is due to the summation of multiple global form signals.

In the current study we tested the hypothesis that the lower thresholds observed for dynamic GPs are due to a summation of the form signals. Given that dynamic GPs consist of multiple independent static GPs, and thus contain multiple presentations of unique global form signals relative to static GPs, we measured the detection thresholds of our participants for static GPs (one GPs frame), dynamic GPs (12 GPs frames), and intermediate stimuli containing two, four, and six unique GPs frames, presented in different types of frame alternation sequences to also manipulate temporal frequency (see Table 1). If the lower thresholds observed for dynamic GPs are due to the summation of multiple form signals, we can expect a linear decrease in threshold as the number of unique frames increases. In addition, each GPs in dynamic GPs is presented for a short duration relative to one GPs in static GPs (i.e., temporal frequency). In order to account for this factor, we measured the thresholds for stimuli that contained blocks of unique GPs (Table 1).

2. Method

Nine adults with normal or corrected-to-normal vision participated in this study (n = 9). This sample included three of the authors, two graduate students, and four undergraduate students from the University of Alberta. All the participants were naïve to the purpose of the experiment, except for the three authors. The experiment was conducted in accord with the Code of Ethics of the World Medical Association (Declaration of Helsinki).

2.1. Apparatus

The stimuli were displayed on a 22” Viewsonic VX2268wm Full-HD LCD monitor (resolution: 1680 × 1050 pixels; refresh rate: 120 Hz). Participants were seated comfortably at a viewing distance of 45 cm to the monitor, with the center of the monitor positioned at eye-level. Participants’ head position was fixed with a chin rest. Stimuli were generated using in-house MATLAB code and presented using the Psychophysics toolbox (Brainard, 1997; Pelli, 1997).

2.2. Stimuli and design

Each stimulus was presented for a total duration of 200.00 ms (12 frames, 60 Hz image update rate). Each GPs subtended a visual angle of 10.7° (diameter of aperture), and each square dot within the stimulus subtended 0.04° × 0.04°. The density of dots within each pattern was set at 6% and the dot separation was 0.25°. The dipoles were oriented to generate a percept of vertical structure (Fig. 1). We chose vertical GPs because Nankoo et al. (2012) have previously shown that the improvement in the detection threshold between static GPs and dynamic GPs is largest for vertical patterns relative to other orientations such as concentric or horizontal, and thus would provide us with the greatest statistical sensitivity for the current study.

A temporal two-alternative forced-choice design was used, whereby the participants were presented with two consecutive patterns; one pattern that contained form signals (i.e., GPs) and one that contained a noise pattern (i.e., randomly-oriented dipoles). The participants’ task was to identify which pattern contained the signal. The order of the signal stimulus and the noise stimulus were identical to the relative ranking of the thresholds for static GPs, the relative performance in each of the patterns suggests that the decrease in threshold found in dynamic GPs and real motion are generally lower than those for global form. The lower thresholds observed for dynamic GPs relative to static GPs are due to the summation of multiple global form signals. In this method, coherence (the % of dipoles aligned in the pattern) was systematically increased or decreased depending on the participant’s performance. In each trial, a psychometric function is fit to all the data collected, and an estimate of the threshold is derived.

2.3. Presentation sequence

As shown in Table 1, the number of unique GPs (i.e., unique frames) used was 2, 4, and 6, in addition to the static and dynamic GPs condition (i.e., 1 and 12 unique frames, respectively). The unique frames were presented in two ways. In one presentation format, the unique frames were presented in an alternating sequence whereby a sequence of unique frames was repeated for a total of 12 frames per stimulus. For example, in patterns with two unique frames (“A” and “B”), the pattern would consist of a repeating...
sequence of the two unique frames (i.e., frame A → frame B → frame A → frame B; Table 1) until 12 frames have been presented. For these alternating sequences, the temporal frequency was kept at 60 Hz given that the image is updated with a new frame every 16.67 ms. In the other presentation format, the unique frames were presented in a blocked sequence where we manipulated the rate at which each unique frame was presented (i.e., temporal frequency). Relative to the alternating sequences, with the blocked sequences all identical frames are shown consecutively (i.e., a block) before the next block of a new unique frame is shown (e.g., frame A → frame A → frame B → frame B; Table 1). The block arrangement and number of unique frames resulted in temporal frequencies of 20 Hz, 30 Hz, and 40 Hz.

### 2.4. Procedure

Prior to each session, the participants were reminded of the global pattern that they were to detect (vertical). On each trial, the

![Fig. 1. Panel A shows examples of vertical and random Glass patterns with 100%, 50%, and 0% coherence. Panel B shows an illustration of a single trial.](image)

**Table 1**

Details on each condition. Each condition name consists of two numbers corresponding to the number of unique frames and the temporal frequency used. For the frame sequence, each unique letter represents an independent Glass pattern, and each letter position represents a frame at 60 Hz. Frames denoted with sequential letters have no relation (e.g., frames A and B are independent).

<table>
<thead>
<tr>
<th>Condition</th>
<th>Frame sequence</th>
<th>Number of unique frames</th>
<th>Temporal frequency (Hz)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Static (1–0)</td>
<td>AAAAAAAAAAAAA</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Dynamic (12–60)</td>
<td>ABCDEFGHIJKL</td>
<td>12</td>
<td>60</td>
</tr>
<tr>
<td>2–10</td>
<td>AAAAAAAAAAAA</td>
<td>2</td>
<td>10</td>
</tr>
<tr>
<td>2–20</td>
<td>AAAAAAAAAAAA</td>
<td>2</td>
<td>20</td>
</tr>
<tr>
<td>2–60</td>
<td>AAAAAAAAAAAA</td>
<td>2</td>
<td>60</td>
</tr>
<tr>
<td>4–20</td>
<td>AAAAAAAAAAAA</td>
<td>4</td>
<td>20</td>
</tr>
<tr>
<td>4–60</td>
<td>AAAAAAAAAAAA</td>
<td>4</td>
<td>60</td>
</tr>
<tr>
<td>6–30</td>
<td>AAAAAAAAAAAA</td>
<td>6</td>
<td>30</td>
</tr>
<tr>
<td>6–60</td>
<td>AAAAAAAAAAAA</td>
<td>6</td>
<td>60</td>
</tr>
</tbody>
</table>

Fig. 1. Panel A shows examples of vertical and random Glass patterns with 100%, 50%, and 0% coherence. Panel B shows an illustration of a single trial.
participants were presented sequentially with two stimuli, one containing the pattern signals of varying coherence (based on the QUEST estimates), and one containing only randomly-oriented dipoles (i.e., 0% coherence). The stimuli were presented centrally, and were temporally separated by 500 ms (Fig. 1B). A message then prompted the participant to press the key “A” if the pattern containing the signal was presented first, or press the key “L” if the pattern containing the signal was the second stimulus presented. No feedback was provided. After a 2 s inter-trial interval, the next trial began.

Testing was carried out over three sessions, all conducted within 1 week. Each session consisted of all eight conditions presented in a random order. Participants completed 45 trials per condition, which yielded a total of 405 trials per session.

2.5. Data analysis

The detection thresholds were determined by a maximum likelihood procedure using the QUEST adaptive staircase procedure (Watson & Pelli, 1983). In the QUEST procedure, the participant’s psychometric function is assumed to follow a Weibull distribution (Weibull, 1951) and coherence levels are based on responses in previous trials.

To more conclusively measure the amount of unique variance explained by each factor, we statistically tested the relationship between both the number of unique frames and temporal frequency with the detection threshold, controlling for the other factor, using a partial correlation analysis. These partial correlations were conducted within each subject and then aggregated using Fisher’s transform (see Corey, Dunlap, & Burke, 1998). Both correlations reported in the results section are partial correlations aggregated using Fisher’s transform.

To improve the reliability of our detection threshold estimates, we had participants complete three sessions. Our task required considerable effort to maintain attention, and occasional lapses in attention could decrease the accuracy of an adaptive method from converging on the participant’s actual detection threshold. For each participant we therefore used the two estimates for each condition that had the lowest SD, and we averaged these two estimates for our measure of each participant’s detection threshold in each condition. Because our selection was based on SD and not on threshold, and because it was applied equally to all conditions, it would not distort the comparison between conditions.

All statistical analyses were conducted using MATLAB (The MathWorks Inc., Natick, MA). Effects were considered significant based on an alpha level of .05.

3. Results

Fig. 2 shows the detection thresholds in terms of coherence level for all nine conditions. As with previous studies (e.g., Nankoo et al., 2012), participants’ thresholds were significantly lower at detecting dynamic GPs (mean = 21%; SD = 0.10%) compared to static GPs (mean = 36%; SD = 0.08%), t(8) = 6.89, p < .001. A repeated-measures ANOVA revealed a significant effect across the remaining conditions (F(6,48) = 4.52, p = .001; Fig. 2). As visible in Fig. 3A, the detection thresholds of the remaining conditions decreased as a function of the number of unique frames. Indeed, the partial correlations confirmed a significant correlation between the number of unique frames and the detection thresholds [r_{p-pop}(8) = −.44, p = .043]. In contrast, as shown in Fig. 3B, the correlation between temporal frequency and detection threshold was not significant [r_{p-pop}(8) = −.27, p = .13]. Thus, the results suggest that the addition of new form signals (i.e., more unique frames) in dynamic Glass patterns is the dominant factor in the detection advantage for dynamic versus static Glass patterns.

4. Discussion

In the current study, we investigated the mechanisms behind the detection of dynamic GPs by measuring the detection thresh-
olds for vertical GPs as a function of both temporal frequency and the number of unique form signals. Detection thresholds were higher for the static GPs, which contained only one unique frame and a temporal frequency of 1 Hz, than for the dynamic GPs, which contained 12 unique frames and a temporal frequency of 60 Hz. These results are consistent with previous studies (e.g., Nankoo et al., 2012), and served as a control to allow us to examine the intermediate levels of both factors. Similar to Day and Palomares (2014), we found that an increase in temporal frequency lowered detection thresholds, but we also found that the contribution of this factor was minor relative to the contribution of unique global form signals. Therefore, our results suggest that the addition of unique form signals is a dominant factor for the improved sensitivity in dynamic GPs.

Some researchers have suggested that the illusory coherent motion perceived with dynamic GPs is the result of activation of the ‘motion streak’ sensors (Burr & Ross, 2002; Geisler, 1999; Ross, Badcock, & Hayes, 2000). At high velocities, movement of an object across the retina leaves a trail of blur that is parallel to the axis of motion. Geisler (1999) has shown that the visual system is able to use the streak, a form signal, to aid motion direction detection (Mather et al., 2013). Geisler’s model of motion streak suggests that motion signals are summated with parallel orientation signals at early cortical levels. Evidence from imaging studies supports the existence of a motion streak system at the level of V2 (Aphthorpe et al., 2013). Given that the dipoles in Glass patterns are known to activate orientation detectors in V1/V2 (Smith, Bair, & Movshon, 2002; Smith, Kohn, & Movshon, 2007), it has been suggested that the perceived coherent motion and the lower thresholds for dynamic GPs, relative to static GPs, are the result of the motion streak sensors interpreting the dipoles as streaks (Day & Palomares, 2014; Ross, Badcock, & Hayes, 2000). Even though it is plausible that the illusory coherent motion perceived with dynamic GPs is mediated by the motion streak system, current evidence does not exclude alternative hypotheses for the reduction in threshold relative to static GPs.

A major difference between static GPs and dynamic GPs is that dynamic GPs consist of multiple presentations of the global form signals relative to the single signal present in static GPs. Our result indicated that the increased sensitivity is primarily due to the additional form signals present in dynamic GPs. It is known that under certain circumstances, stimulus repetition results in lower neural activity in various brain regions and more importantly, improves performance (i.e., priming; Grill-Spector, Henson, & Martin, 2006). Dynamic GPs consist of multiple presentations of static GPs, and can be interpreted as a case of stimulus repetition. Bar and Biederman (1998) provided evidence of priming in V4 using stimuli presented at 65 ms. Given that current evidence suggests that Glass patterns are processed at the intermediate levels of visual processing (e.g., V4), it is possible to improve the detection thresholds of dynamic GPs relative to static GPs is a consequence of the multiple instances of global form signals akin to priming mechanisms. This hypothesis would suggest that the global form information is present in the stimulus, the lower the threshold will be. Our findings support this prediction, suggesting a potential temporal summation of global form signals.

5. Conclusions

Studies have shown that dipoles within Glass patterns activate orientation detectors at the level of V1 and V2, and it has been argued that the motion streak sensors also utilize information from orientation detectors within these areas to disambiguate motion direction (Smith, Bair, & Movshon, 2002; Smith, Kohn & Movshon, 2007). For these reasons, the percept of illusory coherent motion with dynamic GPs has been attributed to the motion streak system. However, the motion streak hypothesis does not adequately explain the lower thresholds of dynamic GPs relative to static GPs, and the differences between dynamic GPs and real motion (Nankoo et al., 2012). Although our results do not exclude the contribution of a motion-based mechanism (Day & Palomares, 2014), we showed that the amount of unique global form signals is an important factor in reducing thresholds in dynamic Glass patterns. It is likely that the global form signals are summated in a similar way to the proposed “snapshot” neuron model of biological motion that summates form information across frames (Giese & Poggio, 2003).

Acknowledgments

This research was supported by grants from the National Science and Engineering Research Council (NSERC) of Canada to M.L.S. and D.R.W., and by an NSERC Alexander Graham Bell Canada Graduate Scholarship (Doctoral level) to C.R.M.

References


