Speed–size illusion correlates with retinal-level motion statistics

Zixin Yong
Neuroscience and Brain Disorders Program, Duke-NUS Medical School, Singapore

Po-Jang Hsieh
Neuroscience and Brain Disorders Program, Duke-NUS Medical School, Singapore

It is a common perceptual experience that smaller objects appear to move faster than larger ones when their physical speeds are the same in either the laboratory or daily life. In this study, we show that the speed–size illusion is correlated with retinal image speed distribution bias. The illusion was quantified with a two-alternative, forced choice speed comparison paradigm, and retinal image speed distributions for different image sizes were obtained by simulation. Simulation results show that smaller retinal images tend to have slower projected speed, and the retinal image speed distribution bias correlates with the strength of the speed–size illusion. Furthermore, exposure to a training movie containing unnatural motion statistics tended to modulate the illusion in a way that was consistent with the speed distribution bias. We discuss how the data could be explained by empirical ranking theory, Bayesian theory, and motion adaptation.

Introduction

The human visual system exploits information carried by the pervasive electromagnetic waves in the environment to guide behaviors. Conversely, the environment shapes and constrains the visual system through natural selection (Banks, Sprague, Schmoll, Parnell, & Love, 2015; Land & Nilsson, 2012). The intricate relationship between the environment and visual perception suggests tight links exist between the regularities of the visual environment and perception (Geisler, 2008). Studies have revealed such correspondence by analyzing statistical properties of natural scenes and comparing them to psychophysical or neurophysiological measurements (Girshick, Landy, & Simoncelli, 2011; Laughlin, 1981; Liu, Bovik, & Cormack, 2008; Yang & Purves, 2003). This approach alludes to possible strategies our visual system employs to obtain useful information (Geisler, 2008).

Visual motion perception does not always agree with physical measurements (Blakemore & Snowden, 1999; Webster, 2015). Misjudgments about speed are frequently observed in the psychophysics laboratory as well as daily situations (Leibowitz, 1985). Perceptual deviations that disagree with the immediate physical measurements of speed, however, may be explained by natural environmental statistics (Weiss, Simoncelli, & Adelson, 2002). Here we show that speed–size illusion, with which smaller objects appear to move faster than larger ones with identical physical speed, is correlated with the motion statistics at the retinal level.

It is known that the size of an object affects estimation of its speed (Mckee & Smallman, 1998). Experiments using standard monitor or virtual reality presentation showed that smaller objects and vehicles appeared to move faster (Barton, 2006; Clark, Perrone, & Isler, 2013; Distler, Gegenfurtner, Van Veen, & Hawken, 2000). However, these experiments often involved factors other than size, such as change in size (Clark et al., 2013), distance perception (Barton, 2006; Distler et al., 2000), and familiarity of objects (Distler et al., 2000). Here we quantified the speed–size illusion with a two-alternative, forced choice (2AFC) speed comparison paradigm, in which white disks translated on a black background in the fronto-parallel plane. Because there was no distance cue of any sort being presented and the change in object size was undiscernible, size was the only primary factor other than physical speed that might contribute to speed perception. Motion statistics were obtained by simulating a virtual 3-D environment in which moving objects projected onto a 2-D retina-like image plane. Smaller retinal images were shown to have right-skewed speed distributions that fell off more quickly. Last, we demonstrated that recent visual experiences embedded with unnatural speed–size statistics tended to modulate the illusion in a manner that was consistent with the retinal image speed distribution bias.

Methods

Quantifying the speed-size illusion (Experiment 1)

Eighteen (nine female, 19–31 years old) and 16 (eight female, 21–25 years old) naive participants with normal or corrected-to-normal vision took part in Experiment 1a and b, respectively. Informed consent was obtained according to the Declaration of Helsinki, and the participants were reimbursed S$10 for the 1-hr experiment. Stimuli were rendered on a Dell P170Sb monitor (75 Hz, 1280 × 1024 pixels, 23.8° × 19.2° view angle) with Psychtoolbox-3 (Kleiner et al., 2007), Palamedes v1.8.0 (Prins & Kingdon, 2009) and Matlab 2012b.

To quantify the perceived illusory speeds of objects with different sizes, we used a 2AFC paradigm to obtain the point of subjective equality (PSE; Figure 1, Supplementary Movie S1). In each trial, one reference and one test object moving to either the left or right were presented consecutively with one above and the other below the fixation dot. The presentation order of the reference and test object was randomized. Participants were required to answer which object appeared to move faster: the one above or below the fixation dot. If the reference (or test) object was chosen, the test object’s speed would increase (or decrease) by 20% of the reference’s speed in the next presentation. Ten choice reversals were required for the experiment termination of each pair of objects. The last six reversal points were averaged to determine the PSE. Participants’ head position was fixed during the experiment with a chin rest.

The reference object subtended 2° visual angle and moved at one of the five possible speeds (7°–35°/s, 7°/s increments). The sizes of the test objects were 1°, 1.5°, 2.5°, and 3°. For each pair of objects (5 speeds × 4 test sizes), the test object’s initial speed was 40% or 160% of the reference’s speed (low/high initial speed). In total, there were 40 distinct pairs of objects. The experiment was divided into two blocks, one with the 1° and 2.5° test objects and the other with the 1.5° and 3° test objects. Half of the participants finished the 1° and 2.5° block first. Trials within each block were mixed and randomly presented. In Experiment 1a, the reference and test objects’ centers were equidistant from the middle horizontal line where the fixation dot was located. In Experiment 1b, the objects’ inner edges were equidistant from the horizontal line.

One participant in Experiment 1b was excluded from the analysis due to the same reason.

Simulation

We obtained approximation of retinal image speed distributions for different image sizes by simulation. As described in previous studies (Wojtach, Sung, & Purves, 2009; Wojtach, Sung, Truong, & Purves, 2008), a virtual sphere with 3-D objects in translational motion was used to simulate the retinal image formation process (Figure 2). Objects were initiated outside a frustum and projected onto a 50° × 50° image plane once they entered the frustum. The virtual environment’s size was defined in arbitrary units such that the image plane measured 50 × 50 units and the sphere radius was 230 units.

The speed of the objects ranged from 0.1 to 150 units/s, and the size ranged from 0.1 to 6.0 units, following uniform distribution. The speed of the objects would give rise to the maximum image speed of 150°/s, roughly the limit of human perception (Burr & Ross, 1986). The sizes of objects were restricted to the relatively small range to generate projected images with size relevant to the psychophysical experiment (1°–3°, Experiment 1). As discussed in the Results section,
The speed–size illusion was quantified before, immediately after, and 10 min after the training session with the same procedure as in Experiment 1 except that only 1° and 3° objects were tested against a 2° reference object moving at 21°/s. Between the first and second post-training tests, participants remained in the laboratory. During the training session, participants watched a 500-s movie containing 30 moving white disks whose speed and size changed constantly (Supplementary Movie S2 and S3 for Experiment 2a and b, respectively). The objects followed laws of reflection when they hit the screen boundaries.

In Experiment 2a, the objects’ size changed between 0.5° and 3.5°, and speed changed between 7°/s and 35°/s. Their size and speed were inversely correlated in a linear fashion (i.e., an object moved at 35°/s when it was 0.5° in size and moved at 7°/s when it was 3.5° in size). In Experiment 2b, the same size and speed ranges applied, but they were positively correlated. Objects were initiated with random sizes, moving directions, and locations within the display. When the movie began, half of the randomly selected objects started to enlarge, and the other half shrunk at the rate of ~0.72°/s (0.5 pixel/frame) while their speeds changed accordingly. When the objects became maximum or minimum in size, the change reversed.

The speed–size distribution in Experiment 2a’s training movie was against the distribution observed in the simulation that mimicked natural statistics, in which smaller objects tended to move slower. It was therefore an “unnatural” distribution because it violated the trend in the environment. The reverse was true for the “natural” distribution in Experiment 2b. To keep participants engaged in viewing, one randomly selected object would become brighter about every 3 s on average, and the participants were required to detect
that object by pressing a key as quickly and accurately as possible.

# Results

## Quantifying the speed–size illusion

As shown in Figure 3A, larger objects possessed higher speeds when they were deemed to move equally fast as smaller ones across the tested spectrum of size and speed. Conversely, smaller objects appeared to move faster when their actual speeds were the same as larger objects. The speed ratios between test objects and the 2° reference (representing the effect size of the illusion) across the five speeds are shown in Figure 3B.

A two-way, repeated-measures ANOVA test examining the effect of size and speed on the ratios revealed a significant effect of size, $F(3, 51) = 133.77$, $p < 10^{-10}$, and interaction between size and speed, $F(12, 204) = 2.13$, $p = 0.017$, but no significant effect of speed, $F(4, 68) = 0.19$, $p = 0.94$. Post hoc paired-sample $t$ tests confirmed the main effect that larger objects had higher ratios when compared to any other speeds: 7°/s and 14°/s, $t(71) = 5.1$, $p = 2.7 \times 10^{-6}$; 14°/s and 21°/s, $t(71) = 0.1$, $p = 0.92$; 21°/s and 28°/s, $t(71) = 1.0$, $p = 0.30$; 28°/s and 35°/s, $t(71) = 0.7$, $p = 0.48$.

We also examined the possibility that this effect was driven by the varying distances between fixation dot and the objects’ inner edge because larger objects had some part being closer to the fixation dot in Experiment 1a, and it might engage more foveal vision (Tynan & Sekuler, 1982). Hence, the edge-fixation distance was matched for all testing objects in Experiment 1b. Similar to Experiment 1a, larger objects had higher speeds when they were perceived to be equally fast as smaller ones (Figure 4A). A two-way, repeated-measures ANOVA test examining the effect of size and speed on the ratios revealed a significant effect of size, $F(3, 42) = 91.85$, $p < 10^{-10}$, and speed, $F(4, 56) = 2.56$, $p = 0.048$, and an interaction between them, $F(12, 168) = 4.23$, $p = 7.9 \times 10^{-6}$. Post hoc paired-sample $t$ tests confirmed the main effect that larger objects had higher ratios: 1° and 1.5°, $t(74) = -5.5$, $p = 5.0 \times 10^{-7}$; 1.5° and 2.5°, $t(74) = -8.3$, $p = 3.5 \times 10^{-12}$; 2.5° and 3°, $t(74) = -3.8$, $p = 3.2 \times 10^{-4}$), and the mixed effects of speed, $p = 0.09$, 0.012, 0.0062, and 0.013 for 7°/s (and 14°/s, 21°/s, 28°/s, 35°/s); $p = 0.11$, 0.048, and 0.27 for 14°/s (and 21°/s, 28°/s, 35°/s); $p = 0.93$ and 0.73 for 21°/s (and 28°/s, 35°/s); and $p = 0.63$ for 28°/s and 35°/s. The results therefore confirmed that the speed misperception did not originate from different edge-fixation distances.
Simulation

The speed probability density function for five image sizes (1°–3°) are plotted in Figure 5. In both the general 3-D environment (Figure 5A and C) and biological data–inspired model (Figure 5B and D), the distribution for smaller images appeared to fall off more quickly, reflecting the fact that small images often resulted from objects in far distance and, hence, also had smaller projected speeds. The overall projected speed in the second simulation was lower because biological motions are relatively slow. Figure 5C and D shows the speed probability for image sizes from 0.1° to 5°, in which the pixel intensity represented probability. The subfigures above had the probability normalized to the most probable speed’s probability within each individual size (0.1° resolution), and the subfigures below had the probability normalized to the overall maximum. The distributions were highly skewed to low speed and small size.

Because real-world motion statistics are not well defined, we have assumed uniform distributions for the objects’ speed and size in our simulation. The result that smaller images’ speed distribution fell off more quickly, however, is mostly determined by the geometrical transformation from 3-D to 2-D motion rather than the specific parameters being used. To see the point, we can compare the speed distributions between two retinal image sizes α and β, where α is smaller than β (Figure 6). α results from projections of objects from size x to z, and β results from projections of objects from size y to z, where x < y < z, and z is the maximum object size in the environment. Because α is smaller, the majority of the sources that give rise to α (objects from size y to z) will tend to be further away when comparing to the sources that give rise to β, and therefore, these objects’ projected image speeds are also lower for α. Assuming further that objects are homogeneously distributed in the environment, then images of size α will be mostly generated by objects in the [y, z] range rather than the [x, y] range because the number of occurrences for any object size is proportional to the distance from viewpoint. Smaller image speeds (relative to β’s) will then dominate α’s speed distribution. Thus, the speed distribution for α will always fall off more quickly comparing to β unless objects in the [x, y] range are much more abundant and much faster than objects in the [y, z] range. Taken together, smaller objects’ speed distributions falling off quickly is likely the natural result of geometrical transformation.

Training

In Experiment 2a and b, the speed–size illusion persisted before and after the training session. As
shown in Figure 7, 1° objects had lower image speed, and 3° objects had higher image speed when they were deemed to be equally fast as the 2° reference. The illusory speed perception was reduced, t(23) = -2.1, p = 0.043 for 1° object and 2° reference, t(23) = 2.3, p = 0.03 for 3° object and 2° reference, paired-sample t test, after unnatural speed–size distribution training in Experiment 2a (Figure 7A and C, Post 1) when participants viewed slow-moving large objects and fast-moving small objects. Conversely, the illusion did not attenuate after training with natural speed–size distribution in Experiment 2b (Figure 7B and D, Post 1). It was observed that 10 min after the training session, the illusory effect was neither significantly different from that before the training, nor significantly different from the first post-test (Figure 7, Post 2), suggesting the
existence of some residue effect. Note that the reported \( p \) values were not corrected for multiple comparisons (number of comparisons = 6).

**Discussion**

Our results show that the size of an object affects the perception of its speed. In particular, smaller objects appeared to move faster in translational motion. This phenomenon is correlated with the observation that smaller retinal images have speed distributions fall off more quickly, which is likely the natural result of 3-D to 2-D transformation—distant objects generate smaller retinal image size and speed. Viewing a movie containing unnatural speed–size statistics tended to modulate the illusion in a way that is consistent with the speed distribution bias. These results demonstrated that speed perception is correlated with retinal image statistics although how and why they were linked is not immediately clear. There are a few theories that might bridge this gap, including empirical ranking theory, Bayesian theory, and visual adaptation (Soon, Dubey, Ananyev, & Hsieh, 2017). We discuss their relevance to our results below.

The empirical ranking theory proposes that inference of physical property is not the goal of the visual system; rather, visual experience is defined by the trial-and-error evolutionary history of the visual system’s interaction with the world (Purves, Morgenstern, & Wojtach, 2015; Purves, Wojtach, & Lotto, 2011). It postulates that the perceptual quality of a retinal stimulus (e.g., length) is a function of the relative frequency of occurrence of the relevant parameter in accumulated past experiences of a given visual system (Howe & Purves, 2005); perceptual response could be estimated from the cumulative distribution function (CDF) of the parameter. This postulation is similar to the observations based on efficient coding theory. For instance, the contrast response functions of fly compound eye interneurons (Laughlin, 1981) and macaque LGN M cells (Clatworthy, Chirimuuta, Lauritzen, & Tolhurst, 2003) approximate the CDFs of natural scene contrast levels, which is the most efficient way to code contrast with a limited dynamic range.

If we plot the CDF of speed from the speed distribution functions in Figure 5A, the speed perception of different sizes will be a function of their percentile rank of the image speed (Figure 8A). As shown in Figure 8A, smaller images occupy a higher percentile rank for any image speed. According to the empirical ranking theory, higher percentile rank indicates higher perceived speed (Wojtach et al., 2009), which is consistent with the result. A prediction of Experiment 1’s result can be made by getting the image speeds for different sizes at the same percentile rank. It is observed that the speed–size illusion is predicted by the theory based on the image statistics (Figure 8B). In Experiment 2a, the speed–size distribution in the training movie should reduce the difference between the CDFs because small images are biased to higher speed. The training should therefore lead to a reduced illusory effect, according to the theory. This was likely the case immediately after the training (Figure 7A, Pre-Post 1). The fact that this reduction disappeared after 10 min
Figure 7. Speed ratios between 1° or 3° objects and 2° reference when they were perceived to be equally fast. (A, B) Bars within the 25th and 75th percentile boxes represent median. (C, D) Each circle/diamond represents data from one participant. Data from the same participant are connected by lines. Short horizontal bars represent group means. (A, C) Unnatural speed–size distribution training. The illusion reduced immediately after the training, \( t(23) = -2.1, p = 0.043 \) for 1° object; \( t(23) = 2.3, p = 0.030 \) for 3° object; paired-sample \( t \) test between Pre and Post 1, and recovered 10 min after the training, \( t(23) = -0.6, p = 0.57 \) for 1° object; \( t(23) = 1.6, p = 0.12 \) for 3° object; paired-sample \( t \) test between Pre and Post 2. (B, D) Natural speed–size distribution training. The illusion persisted both immediately after the training, \( t(22) = 1.3, p = 0.21 \) for 1° object; \( t(22) = -1.3, p = 0.22 \) for 3° object; paired-sample \( t \) test between Pre and Post 1, and 10 min after the training, \( t(22) = 0.5, p = 0.65 \) for 1° object; \( t(22) = -1.4, p = 0.18 \) for 3° object; paired-sample \( t \) test between Pre and Post 2. Post 1 and Post 2 ratios were not significantly different for both trainings. Unnatural distribution: \( t(23) = -1.7, p = 0.10 \) for 1° object; \( t(23) = -0.03, p = 0.98 \) for 3° object. Natural distribution: \( t(22) = 0.7, p = 0.51 \) for 1° object; \( t(22) = -0.06, p = 0.96 \) for 3° object. Pre: Test before training. Post 1: Test immediately after training. Post 2: Test 10 mins after training. \( p \) values were uncorrected for multiple comparisons.
(Figure 7, Pre-Post 2) suggests that temporary changes in the motion statistics (training movie in this case) are not sufficient to induce long-term changes in the visual system’s internalized empirical rankings, which are supposed to be shaped in the evolutionary time scale. In summary, the correlation between motion statistics and speed perception is consistent with the empirical ranking theory.

Another framework of visual perception that links to natural scene statistics is based on Bayes’ theorem (Kersten, Mamassian, & Yuille, 2004). Visual perception in this regard is viewed as a process of probabilistic inference based on current observation and prior knowledge (i.e., prior and likelihood function). Previous works suggest that speed prior is skewed to low speed, and higher measurement noise leads to stronger bias toward the prior (Stocker & Simoncelli, 2006; Weiss et al., 2002). In this account, the speed–size illusion would arise from lower measurement noise for smaller objects, which would then be less biased toward low speed. However, large objects are likely to have less noise because more evidence about speed is available at the edge, especially in the context of Experiment 1, in which objects had high contrast on the black background (Figure 1).

Another Bayesian explanation is that small objects have speed distributions (i.e., prior) skewed to the higher end comparing to large objects, and hence, the perceived small object motions are biased to higher speed. Although it is true that smaller objects move faster in some cases (e.g., a hunting cheetah is faster than a walking elephant), this is not true in general (e.g., a bicycle is slower than a car). Validating this Bayesian explanation would nevertheless require the speed distribution of objects in the 3-D world to be determined in future experiments. Our simulation relied on generic 3-D world speed and size distributions and, therefore, could not provide evidence for this account.

The third Bayesian explanation is more specific to the speed–size effect manifested in Experiment 1 but not the effect in general. An observer might implicitly assume that the white disks seen during the experiments were identical objects at different distances. In this scenario, a smaller disk was most likely the object at further distance. Accordingly, it should also possess higher physical speed given the same image speed as a larger disk. An observer who is trying to infer the object’s physical speed instead of the disk’s image speed would then be biased by the likelihood function $P(\text{Physical speed}|\text{Image speed})$ under the constraint of identical objects. This explanation does not apply to the speed–size effect in other cases as the effect often happens between different objects (Clark et al., 2013).

The results of Experiment 2 do not support the conventional Bayesian explanation. Adams, Graf, and Ernst (2004) showed that the light-from-above prior...
Visual adaptation could also lead to misjudgment of speed and motion direction (Clifford & Ibbotson, 2002). It is conceivable that the constant presence of a certain adaptor will cause an illusory perception to last indefinitely. Hietanen, Crowder, and Ibbotson (2008) showed that adaptation to low speed and test at high speed would increase perceived speed and vice versa. It is then possible that our visual system adapts to the statistically lower speed of small objects, and the speed perception of small objects is recalibrated relative to that of large objects such that small objects appear faster. This explanation requires speed adaptation to be size-specific, similar to the orientation-specific color adaptation seen in the McCollough effect (McCollough, 1965). That is, the adaptation to small objects’ relative low speed and large objects’ relative high speed leads to separate calibration of small and large objects’ speed perception. This concept of constant adaptation is markedly different from conventional adaptation, which is a temporary effect. It is more akin to the concept of Bayesian prior working in the opposite way.

Experiment 2 falls within the time range of conventional adaptation. The training movies worked essentially as adaptors, and the post-training tests gauged the adaptation effect. Experiment 2a showed adaptation results (Figure 7A) similar to Hietanen et al. (2008), in which adaptation to low speed and test at high speed increased perceived speed and vice versa. Although only significant before correction, the adaptation effect for both low and high speed agreed with Hietanen et al. immediately after the training and then attenuated in minutes. Experiment 2b, however, did not show a significant adaptation effect. The major difference between Hietanen et al.’s and our experiment was that the adaptors in our experiment were a mixture of objects with varying speed and size rather than a fixed size. Specifically, fast and slow adaptors were combined in a single adaptation session, and speed was correlated with size. The results therefore suggest that speed adaptation could be size-specific.

**Conclusion**

We showed that the speed-size illusion is correlated with retinal-level speed–size distribution bias. Exposure to biased motion statistics tended to modulate the illusion in the manner that was consistent with the retinal image speed distribution bias. These results could be consistent with empirical ranking theory, Bayesian theory, and motion adaptation.

**Keywords:** illusion, speed perception, empirical ranking theory, Bayesian, adaptation

**Acknowledgments**

Commercial relationships: none.
Corresponding author: Po-Jang Hsieh.
Email: pojang.hsieh@duke-nus.edu.sg.
Address: Neuroscience and Brain Disorders Program, Duke-NUS Medical School, Singapore.

**References**


