Velocity constancy in a virtual reality environment

Hartwig K Distler, Karl R Gegenfurtner, Hendrik A H C van Veen
Max Planck Institute for Biological Cybernetics, Spemannstrasse 38, 72074 Tübingen, Germany
Michael J Hawken¶
Centre for Neural Science, Room 809, New York University, 4 Washington Place, New York, NY 10003, USA; e-mail: mjh@cns.nyu.edu
Received 15 December 1999, in revised form 2 June 2000

Abstract. During everyday life the brain is continuously integrating multiple perceptual cues in order to allow us to make decisions and to guide our actions. In this study we have used a simulated (virtual reality—VR) visual environment to investigate how cues to speed judgments are integrated. There are two sources that could be used to provide signals for velocity constancy: temporal-frequency or distance cues. However, evidence from most psychophysical studies favours temporal-frequency cues. Here we report that two depth cues that provide a relative object – object distance—disparity and motion parallax—can provide a significant input to velocity-constancy judgments, particularly when combined. This result indicates that the second mechanism can also play a significant role in generating velocity constancy. Furthermore, we show that cognitive factors, such as familiar size, can influence the perception of object speed. The results suggest that both low-level cues to spatiotemporal structure and depth, and high-level cues, such as object familiarity, are integrated by the brain during velocity estimation in real-world viewing.

1 Introduction
The advantage of the virtual reality (VR) environment is that it is more like the real visual world. Hence we can test how cues (Landy et al 1995; Norman et al 1996) that are found to be important in reduced visual environments, often employed in psychophysical studies, apply in more natural conditions. Velocity constancy is a (higher-level) perceptual ability whereby observers can perceptually equate the speeds of objects, located at different distances away from the observer (depths), that are moving at the same physical speed. Observers equate the speeds despite the fact that the objects are moving at different angular speeds on the retina. There are two main ways that this ability could be achieved: judging object speed against a reference frame (temporal frequency) or scaling retinal velocity by perceived distance between the objects (McKee and Smallman 1998). Evidence from psychophysical studies is most often interpreted as favouring the first possibility (Wallach 1939; McKee and Smallman 1998), thereby excluding depth cues, such as disparity and motion parallax, that give a relative object – object distance (Johnston et al 1994) from a role in scaling for velocity constancy.

   The ability to correctly judge the physical velocity of moving objects was shown by Brown (1931) to be invariant with the distance of the object from the observer. He called this velocity constancy. The image of a moving object has an angular velocity on the retina, so that objects moving at the same physical speed but at different distances will have different angular retinal velocities. All current models of human motion estimation are based on retinal speed exclusively (Adelson and Bergen 1985; van Santen and Sperling 1985; Watson and Ahumada 1985; Smith and Edgar 1994). Yet human observers can judge objects at different distances to be moving at the same physical speed, even though their retinal speed is different. How is this transformation from an angular to apparent physical speed accomplished?

¶ Author to whom correspondence and requests for reprints should be addressed.

There are two principal hypotheses that detail how velocity constancy could be achieved by the brain. Both of these rely on a transformation of the retinal speed into estimates of object speed. The first is the temporal-frequency hypothesis. According to the temporal-frequency hypothesis observers judge the retinal speed against a frame of reference to obtain an estimate of relative object speed. This was proposed by Wallach in his reinterpretation of Brown's experimental results (Wallach 1939). In many of the earlier studies on velocity constancy this frame would have been the boundary of the display area, such as a viewing aperture or the border of a television screen. In the experiments that we describe here using a VR environment, where there is no clear screen border, the reference frame might be the object itself. The reference frame would be determined by the relative size of the object and the temporal frequency obtained by scaling angular speed by the number of times the object traverses its own length. All these types of scaling are consistent with the use of temporal frequency of the object to determine constancy—one of the suggestions put forward by McKee and Smallman (1998) to account for constancy results. The second hypothesis that allows constancy to be achieved is to scale the angular retinal speed by the distance (depth) of the object from the observer. In the experiments that we describe in this paper we are asking observers to judge the relative speed of objects, so it is the distance between the objects (inter-object distance) that is the important variable for scaling purposes rather than the absolute distance between observer and the moving objects.

Most of the psychophysical evidence can be interpreted within a framework that uses a temporal-frequency coding strategy to achieve velocity constancy. This is exemplified by Wallach's (1939) relational hypothesis in which perceived observer–object distance (depth) does not play a role in determining velocity constancy. Both size and physical velocity interact to give a perceived velocity that tends towards velocity constancy. The size cue corresponds, at least partly, to changes in spatial frequency. The changes in spatial frequency (due to size) lead to changes in temporal frequency. A fuller account of this relationship is given in section 6. However, it should also be noted that most experiments have been conducted in rather restricted environments where a frame of reference is often a powerful cue, and it might be argued that the stimulus conditions themselves have unwittingly contributed to the strength of the temporal-frequency effect. On the other hand, some studies (Rock et al 1968; Epstein 1978; Wist et al 1976) conclude that estimates of perceived observer–object distance do play a role in determining the scaling for velocity constancy. Although some recent studies (McKee and Welch 1989; Zohary and Sittig 1993) provide evidence that an estimate of object distance is unnecessary for velocity constancy under some viewing conditions, this does not preclude a role for observer–object distance estimates in natural viewing. Therefore it is still important to consider cues, which are usually referred to as distance cues, which can influence velocity judgments, hence constancy.

Designing experiments to combine many cues that appear natural and in the correct spatial relationship to each other and to the observer can be achieved most easily in a virtual reality environment (Distler et al 1998; van Veen et al 1998). We wished to determine what role the two types of scaling might play in an environment more naturalistic than the environment often employed in many laboratory types of psychophysical experiments. One means of obtaining a natural environment while keeping the option of manipulating the stimulus variables has recently become available as a large-scale virtual reality system (figure 1). We have taken advantage of this system to test the role of the two strategies of obtaining velocity constancy.
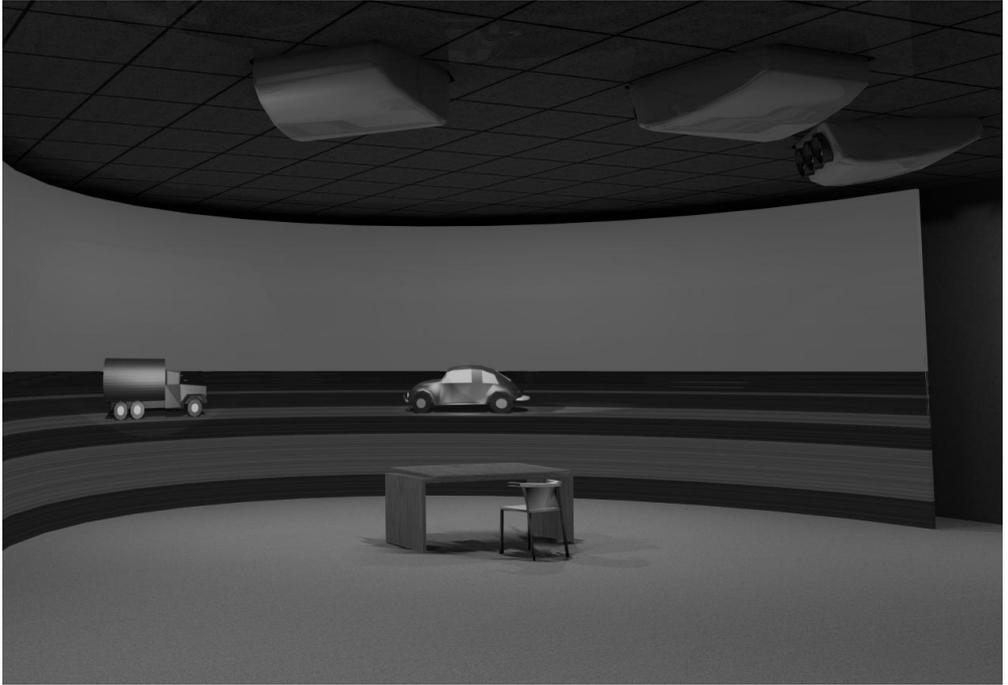
Figure 1. In the experimental setup the stimulus images were front-projected onto a large half-cylindrical projection screen (diameter: 7 m; height: 3.15 m) by means of three 3 CRT projectors. The subjects were seated facing the display, with the projected image subtending an angle of 180 deg × 50 deg in the field of view of the subjects. All experiments were conducted in a dark room where the only brightness information available originated from reflections off the projection screen. When normal projection was used, the spatial resolution of the projected image was approximately 3544 × 1024 pixels. A car and a truck are shown on the middle of the three projected roads. The projectors are shown on the roof and the position of the subject facing the projection screen would be at the table in the centre of the room.
Note: A colour version of figures 1 and 2 can be seen on the web at http://www.perceptionweb.com/perc1200/distler.html and is archived on the annual CD-ROM distributed with this issue.

2 General methods
2.1 *Setup*
The experiments were performed in a large-scale simulation environment (figure 1) at the Max Planck Institute for Biological Cybernetics (Distler et al 1998; van Veen et al 1998). A 3-pipe Silicon Graphics Onyx2 InfinitReality was used to compute the stimulus images which were then front-projected onto a large half-cylindrical projection screen (diameter: 7 m; height: 3.15 m) by means of three 3 CRT projectors. Video blending hardware (Panoram Panomaker II) was used to achieve a smooth transition between the three images. The subjects were seated in the centre of the display; therefore the projected image subtended an angle of 180 deg × 50 deg in the subjects' field of view. All experiments were conducted in a dark room where the only brightness information available originated from reflections off the projection screen. When using normal projection, the spatial resolution of the projected image was 3544 × 1024 pixels. Video refresh rate and frame rate were adjusted to 72 Hz. If the scene was projected stereoscopically, the resolution of the screen was 2835 × 768 pixels, the video refresh rate was 96 Hz (48 Hz for each eye), and the frame rate was 48 Hz. Active LCD shutterglasses (Stereographics) were used to resolve the flickering images.

2.2 Stimuli

The basic condition was a textured ground plane and two vehicles moving at different distances, with a simulated eye height of 1.5 m (figure 2). Each trial was initiated by the presentation of a fixation mark (stop sign) whose simulated vehicle – observer distance was equal to the mean distance of two vehicles that were presented subsequently. 250 ms after fixation onset, two vehicles appeared, offset to the left and right of fixation.
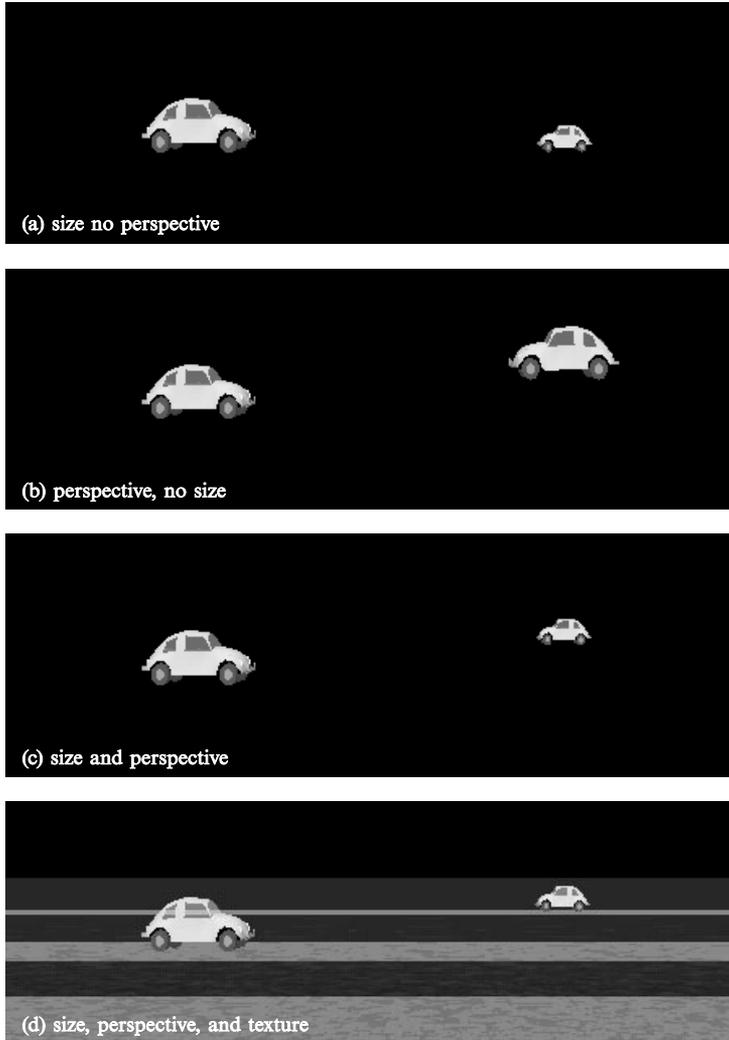
Figure 2. The contribution of different depth cues investigated in experiment 1: (a) size; (b) perspective-viewing height; (c) size and perspective-viewing height; and (d) size, perspective-viewing height, and texture gradient. The ground plane contained three roads, which ran perpendicular to the observer's line of sight. The roads were 5 m wide and were displayed at simulated vehicle – observer distances of 11.7 m, 20.9 m, and 40.5 m. If not noted otherwise, the subjects viewed the scene from the view of a simulated observer whose eye height was adjusted to 1.5 m, the perspective-viewing height. The basic model of the vehicle we used in the experiments was a 3-D computer graphic model of a VW beetle. The dimensions of the vehicle were 1.2 m × 3.1 m × 1.2 m (width × length × height). At the simulated vehicle – observer distance of 20.9 m for the standard vehicle, the vehicle subtended an angle of 8.4 deg × 3.3 deg (width × height) in the observer's field of view. The respective angles subtended by the test vehicles moving at different distances were 14.8 deg × 5.9 deg (distance 11.7 m), 8.4 deg × 3.3 deg (20.9 m), and 4.4 deg × 1.7 deg (40.5 m).

The fixation mark disappeared after a further 250 ms and one of the vehicles started moving. Then after a random delay (100 ms $\leqslant \Delta t \leqslant$ 250 ms) the second vehicle started to move. Both vehicles moved towards the centre of the display, then both cars disappeared 1 s after the onset of the second car's motion. In a 2AFC paradigm subjects indicated which vehicle appeared to be moving faster. Since an acceleration effect at motion onset could have been responsible for incorrect velocity judgments, subjects were instructed to delay making their decision until the vehicles were moving at a constant velocity. Subjects were instructed to fixate the stop sign or its prior position throughout the trial. One vehicle was the standard vehicle, always moving at a simulated observer–vehicle distance of 20.9 m and a constant velocity of 3.0 m s^{-1}. The standard vehicle was initially randomly positioned at the left or right of the centre of the display. The influence of vehicle distance on perceived velocity was studied by using three distances of the test vehicle. The simulated vehicle–observer distances of the test vehicle were 11.7 m, 20.9 m, and 40.5 m; its velocity was adjusted by means of an adaptive staircase procedure (Levitt 1970) to determine the point of subjective equality (PSE). Six staircase reversals were required to finish one experimental block.

Several modifications were introduced to the experimental procedure to ensure that subjects did not use the vehicle's start and end position in combination with time elapsed to judge and/or compare the vehicle's physical velocity: (i) A random spatial offset was added to one vehicle's start position. The random offset was balanced between standard and test vehicle to prevent observers from predicting the offset direction. (ii) The motion onsets of the vehicles were delayed with respect to each other. The delay was random (0.1 s $\leqslant \Delta t \leqslant$ 0.25 s) and evenly distributed between standard and test vehicle. Measured from the onset of the second vehicle's motion, both vehicles were visible for 1 s. (iii) Neither vehicle crossed the centre of the screen. Even more important, no overlap of the front parts of the vehicles in the image plane was observed. (iv) Before the vehicles actually started moving, the fixation mark, which potentially could serve as a reference, was removed. (v) Throughout the experiments the car's wheels did not turn to prevent subjects from comparing the number of wheel revolutions.

2.3 Subjects
Three (experiment 1), four (experiment 2), or eight (experiment 3) subjects with normal or corrected-to-normal vision participated in the experiments. In general, before they performed their first experiment, subjects were given two blocks of practice trials to get familiar with the task.

3 Experiment 1: Cues to velocity constancy
It is well established that relative object size influences velocity perception (Brown 1931; Epstein and Cody 1980; Zohary and Sittig 1993). In the current study, our aim was to investigate how object size interacts with distance cues in promoting velocity constancy under naturalistic viewing conditions. While earlier studies tended to investigate a limited number of distance cues in isolation, here the first experiment was designed to investigate the influence on velocity judgments of the combination of a number of distance cues. In the VR environment, these cues lead to a compelling depth sensation. In addition, combinations of distance cues are used to evaluate the importance of cue combination (Richards 1985; Bülthoff and Mallot 1988; Rogers and Collett 1989; Tittle and Braunstein 1991; Johnston et al 1993, 1994) in the perception of object velocity.

The first experiment was designed to study a set of cues that gave the observer information about the relative distance between the two cars. The cues were perspective size, texture of the ground plane, observer viewing height, disparity, and motion parallax. Since distance cues in isolation are thought to be relatively ineffective, a combination of relative distance cues was used to better approximate natural viewing conditions.

Furthermore, some cue combinations with conflicting information were employed to begin to appreciate the strength of different cues.

3.1 Methods

In these experiments we are using a complex visual scene to simulate natural viewing conditions. The size variable can contribute to both the relative-depth information in the scene and to the temporal frequency of the moving object. The other cues—texture, viewing height, disparity, and motion parallax—contribute to the depth alone. The relative strengths of different variables compared to the size cue are difficult to assess as it is not yet established whether there is a strong contribution of depth to constancy nor whether we have chosen values for the cues that give rise to a maximal depth effect. The range of values for some of the cues is relatively well constrained by natural viewing conditions. Viewing height, which simulates the height of the observers' eyes above the ground plane, was chosen to be 0 or 1.5 m, the difference between lying on the ground and standing. Disparity was chosen to be correct for the viewing conditions; therefore it is either present or absent, although we do not have an independent measure of the contribution of disparity to distance estimations per se under these viewing conditions. Simulated observer motion to give the motion parallax cue was chosen to be a velocity of 1 m s^{-1}, a speed corresponding to a brisk walk. Thus the combination of viewing height, simulated forward motion, and disparity is supposed to represent an averaged-height person with binocular vision walking towards the cars at normal pace. The texture information was added to provide a minimal delineation of the area between the cars without adding further objects to the scene, such as spatial objects, for example trees, that would clearly add explicit frames of reference for judgments of temporal frequency.

The stimulus for these experiments was a pair of vehicles presented in a simulated natural visual environment (figures 1 and 2). To measure the degree of velocity constancy, the pair of cars, one of which was the test (right-hand vehicle in figure 2) and the other the standard, moved in opposite directions on the simulated roadways and the observers' task was to judge which vehicle was moving faster. The results of the experiments give the relative velocity of the test vehicles at the point of subjective equality (PSE) as a function of their simulated vehicle–observer distance (figure 3a). The relative velocity, v_r, is computed as $v_r = v_s / v_t$, where v_s is the velocity of the standard vehicle and v_t is the velocity of the test vehicle at the PSE.

If subjects show perfect velocity constancy, meaning that they compare the physical velocities of the vehicles, the relative velocity at the PSE independent of the test vehicle's distance should always be 1.0 (figure 3a, dotted line). However, if subjects compare the angular velocities, the PSEs are 1.79 (viewing distance $d = 11.7$ m), 1.0 ($d = 20.9$ m), 0.52 ($d = 40.5$ m) (figure 3a, solid line). An apparent velocity greater than 1.0 means that the velocity of the test vehicle is overestimated compared to the velocity of the standard. Figure 3b shows example data, collected at seven different viewing distances. When both standard and test vehicle are moving at the same observer–object distance, the relative velocity, independent of the condition, is approximately 1.0. When the test vehicle is closer than the standard vehicle, the relative velocity is usually greater than 1.0, subjects overestimate the physical velocity of the near vehicle. If the test vehicle is farther away than the standard vehicle, the relative velocity is usually less than 1.0; the subjects underestimate the physical velocity of the far vehicle. To illustrate the influence of viewing condition on velocity constancy we computed a factor quantifying the degree of velocity constancy (figure 3). A linear regression was fitted to the data points on a logarithmic scale. The slope of the resulting line provides an estimate of the degree of velocity constancy. A slope of 0 indicates perfect constancy, and a slope of -1 indicates a complete lack of constancy. The slope was turned into a
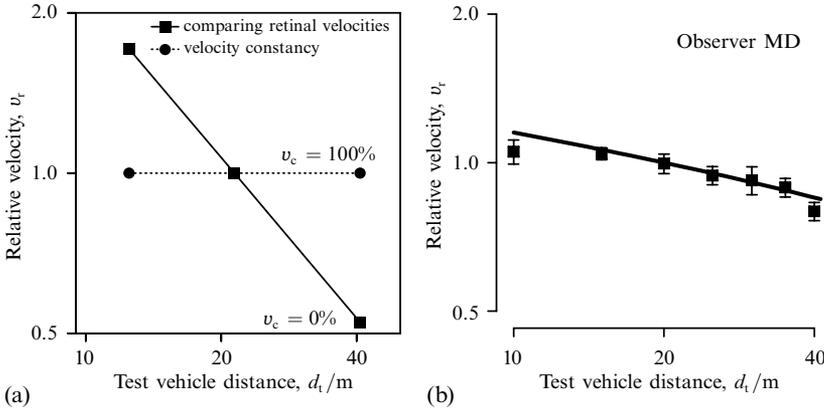
Figure 3. (a) Theoretical relationship between retinal and physical velocities. Assuming the relationship between perceived relative velocity and perceived relative distance is linear, the relationship can be expressed as: $v_r = v_s/v_t = 1 + (1 - v_c)\log(d_s/d_t)$, where v_c is a measure of velocity constancy, v_t is the test car's physical velocity at match, v_s is the standard car's physical velocity, d_t is the simulated observer – object distance of the test car, d_s is the standard car's distance. If the observer compares retinal velocities, then $v_c = 0$ (squares, solid line); if the observer compares the physical velocities (velocity constancy), $v_c = 1$ (circles, dashed line). (b) Example data from an actual experiment. Seven different distances between 10 m and 40 m were used. The solid line shows the best-fitting linear regression through the data points. It has a slope of -0.1825, which corresponds to a velocity constancy factor of 81.75%.

constancy factor v_c by putting $v_c = 100 \times (1 + \text{slope})$. The constancy factor will be 0 if observers report the angular velocities. The constancy factor value will be 100% if physical velocities of the vehicles are used.

3.2 Results

When perspective size (S), texture (T), viewing height (H), disparity (D), and motion-parallax (P) cues are combined in the VR visual environment, perfect constancy is obtained (figures 4a and 4b, column labelled SHTDP). All of these factors contribute, to some extent, to velocity constancy but the major contribution comes from size and the combination of disparity and motion parallax. In order to determine the effects of the cues, either singly or in combination, we either removed or added them to the basic size condition described next. In the basic size condition two cars are moving on the roads on the ground plane at different distances from the observer with a simulated eye height of 1.5 m with the three roads and textured background visible (figure 2d). For the disparity condition, the basic size condition was presented stereoscopically with the vergence point at 3.5 m, putting the vergence point in the plane of the projection screen. For the motion-parallax condition the simulated observer was moving forward at 1.0 m s^{-1} in the basic size condition. However, since the stimuli were shown for only about 1 s, this motion had little effect on the actual disparities.

In the basic size condition (SHT, figure 2d), when observers were asked to judge the speed of the car at 40 m distance relative to the standard car at 20 m, they systematically underestimated the physical speed of the distant car (figure 4a, column SHT, right versus middle bars). Conversely, a comparison of the speeds of the near car, at 10 m, to the standard car, at 20 m, leads to an overestimate of the speed of the near car (figure 4a, column SHT, left versus middle bars); resulting in a constancy factor of about 85% (figure 4b, column SHT). Removing the size cue results in a significant reduction ($p < 0.01$, t-test) in the constancy factor of 25 units (figures 4a and 4b, column HT). Removing texture (T) and viewing height (H) cues produced small but consistent reductions in constancy (figures 4a and 4b, columns SH and ST respectively).
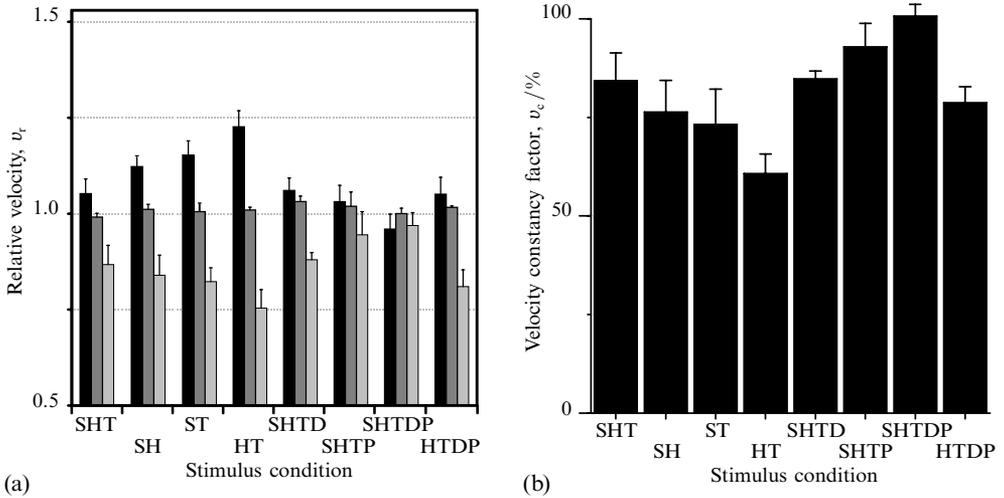
Figure 4. (a) The relative velocity of the test vehicle at the PSE as a function of the viewing condition and its distance. The different distances of the test vehicle in each condition are designated with different shading: 11.7 m (black), 20.9 m (mid-grey), 40.5 m (light grey). (b) The values of the constancy factor for each viewing condition. The results constitute the mean of three subjects (4 iterations per subject). The error bars correspond to one standard error of the mean. The viewing conditions are designated by the letters under each bar graph as follows: D = disparity, H = perspective-viewing height, P = motion parallax, S = perspective size, T = texture.

Next, two cues that can be thought of as primarily giving additional distance information—disparity and motion parallax—were added to the basic size condition. When subjects viewed the stimulus stereoscopically (figures 4a and 4b, column SHTD), there was no improvement in constancy judgments. The constancy factor remained at about 85% (figure 4b, column SHTD). When the observer was simulated as moving forward, thus providing the observer with distance information from motion parallax, there was a small but nonsignificant increase in the constancy factor (figures 4a and 4b, column SHTP). However, when disparity and motion-parallax information was given in combination, velocity constancy was significantly improved ($p < 0.05$, t-test). In this condition, vehicles at both near and far distances are judged to be moving at the same speed as the vehicle at the intermediate distance (figure 4a, column SHTDP) resulting in a constancy factor of 100% (figure 4b, column SHTDP). In the final condition in this series, the size cue was removed while keeping all the other cues in place (figure 4a, column HTDP). This resulted in a significant reduction (t-test, $p < 0.01$) of 21 constancy factor units (figure 4b, column HTDP).

4 Experiment 2: Effects of size

In the previous experiment it was shown that the presence or absence of the size cue had a major effect on perceived velocity in the VR environment. This new experiment was designed to determine whether size per se of the test vehicle is an important factor.

4.1 Methods

Three different sizes of the test vehicles (relative to the size of the standard vehicle) were used at each distance: the test vehicle was 50%, 100%, and 200% the size of the standard vehicle. The standard vehicle was of normal size and moving at a simulated distance of 20.9 m (velocity: 3.0 m s^{-1}). As in the first experiment, the test vehicles were moving at simulated vehicle–observer distances of 11.7 m, 20.9 m, and 40.5 m.

4.2 Results

Vehicle size exerts a strong effect on the perceived velocity (figure 5). The smaller test vehicle (squares in figure 5) is perceived as moving faster than the standard-sized test car (circles in figure 5) at all viewing distances ($p < 0.01$, two tailed t-test). Although there is no significant difference between the 100% and 200% size conditions ($p > 0.05$) there is a consistent trend for the vehicles in the 200% size condition (figure 5, triangles) to be perceived as moving slower than the standard-sized vehicles (figure 5, circles). In general, size causes a constant offset in the perceived velocity of the vehicles, while the slopes of lines connecting the relative velocities for one size of the test vehicle are quite similar. Thus it is important to distinguish the relative effects of size across distance with the effects of size at the same observer–vehicle distance.
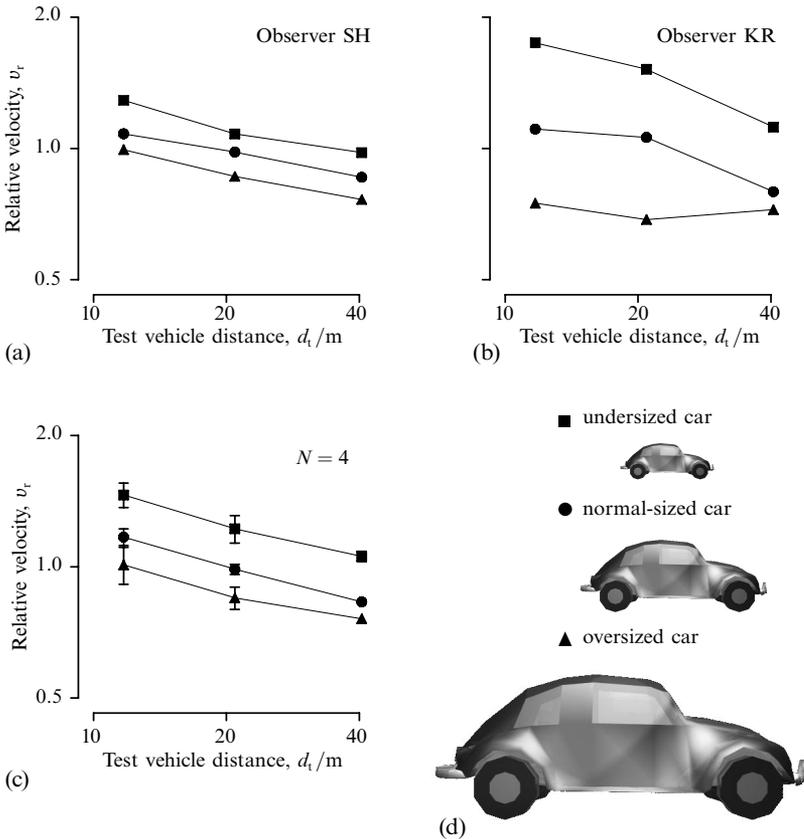
Figure 5. The effect of vehicle size on perceived velocity: the relative velocity of the test vehicle at the PSE as a function of the test vehicle's distance and size. (a) Data from observer SH. (b) Data from observer KR. (c) The mean of four subjects (3 iterations per subject). The error bars in each graph correspond to one standard error of the mean. The different symbols in the graph indicate different sizes of test vehicles: 50% (undersized car) squares, 100% (normal-sized car) circles, 200% (oversized car) triangles, as shown in (d).

5 Experiment 3: Vehicle familiarity

Are the effects of size a result of changes in the spatiotemporal structure of the stimulus or do higher-level cognitive factors also play a role in determining velocity judgments? The use of the VR environment makes it relatively easy to explore the role of familiarity or prior knowledge on perceptual judgments. In the previous experiment, subjects were required to compare the velocity of undersized and oversized cars, car sizes that are generally unfamiliar. If subjects had to compare the velocity of a normal-sized

car, an oversized car, and a truck having the same size as the oversized car, it might be predicted, on the basis of the results of size scaling seen in figure 5, that the perceived velocity of the oversized car and truck would be lower than that of the normal-sized car, since the latter is smaller than the former. The perceived velocity of the oversized car and the truck would be equal, since the two vehicles are of equal size. On the other hand, if subjects include prior knowledge about the natural size, type, and velocity of the particular vehicle into the judgment of the velocity of the vehicles, the truck should be perceived as moving faster than the oversized car. The reason for this is that we are familiar with the dimensions of the truck and how the truck's relative position in the environment changes when it is moving at a certain velocity. We do not have the same prior knowledge for the oversized car. Thus, for the oversized car we have to rely predominantly on perceptual input, whereas in the case of the truck prior knowledge also influences the perception of its velocity towards that of the smaller vehicles.

5.1 Methods

The standard vehicle was a normal-sized VW beetle (figure 6, squares) moving at 3.0 m s^{-1} at a simulated vehicle – observer distance of 20.9 m. Three different test vehicles
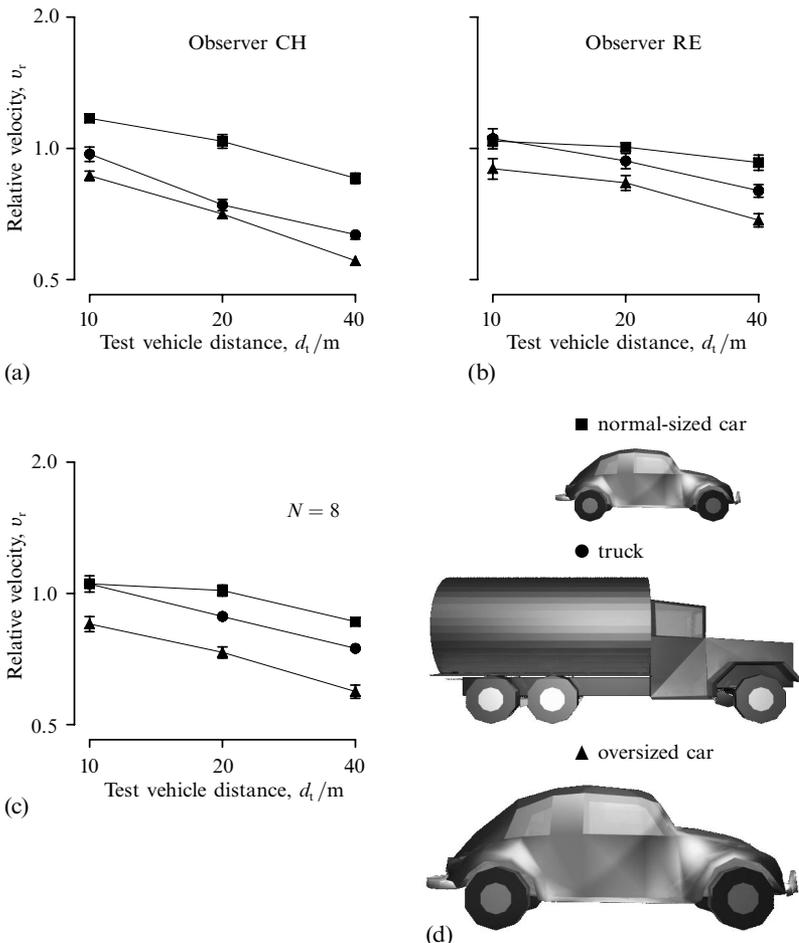
Figure 6. The effects of familiarity on perceived velocity. Relative velocity of the test vehicle at the PSE as a function of distance, size, and type of test vehicle: normal-sized car (squares); oversized car, 280% of normal car (triangles); and truck the same size as the oversized car (circles). (a) Data from observer CH. (b) Data from observer RE. (c) The mean of eight subjects (5 iterations per subject); the error bars correspond to one standard error of the mean. (d) The different test vehicles.

were studied: a normal-sized VW beetle (figure 6, squares), an oversized VW beetle (280% the size of the normal-sized VW beetle; figure 6, triangles), and a truck with the same size as the oversized VW beetle (figure 6, circles). The test vehicles were moving at distances of 11.7 m, 20.9 m, and 40.5 m.

5.2 Results

Good velocity constancy was obtained for the normal-sized vehicle in the current experiment (figure 6, squares). Increasing the size of the test vehicle (oversized VW beetle) significantly decreased (t-test, $p < 0.01$) its perceived velocity (figure 6c, triangles). This result confirms the observations of the previous experiment. For the individual observers (CH and RE, figures 6a and 6b) and for the results combined over all eight observers (figure 6c) who participated in this experiment, the truck is perceived to be moving at an intermediate speed. Its perceived velocity was significantly higher than that of the oversized VW beetle (figure 6c, triangles; t-test, $p < 0.01$), and lower than that of the normal-sized VW beetle. This indicates effects of both the spatiotemporal structure, as well as the familiarity of the objects.

6 General discussion

In the simulated visual environment, multiple cues are combined by the brain to transform retinal angular speeds into perceived relative physical speeds. When cues to size, distance, and motion are combined in a manner that mimics viewing in the natural environment, then almost perfect velocity constancy is attained. In this simulated natural environment it is found that cues to inter-object distance play a significant role in promoting velocity constancy. Furthermore, top–down influences, such as object familiarity, can also play a significant role in determining constancy.

It should be noted that observers need to estimate only the relative speed between the cars to achieve constancy, not the absolute, veridical, speed of the cars. This is in fact the case with almost all studies of velocity constancy: the measure is the speed of one object compared to another. Therefore, what is required in terms of scaling by distance is the inter-object distance not the observer–object distance. In this regard our descriptions of the effects of cues to distance refer to inter-object distance and those to speed are relative object speed, as the observer is only required to make a match of the relative physical speeds of objects at different distances to achieve constancy.

There is clear evidence that the combination of disparity and motion parallax is more than the linear sum of their individual effects, a phenomenon that is sometimes referred to as promotion (Johnston et al 1994). In studies that have sought to determine whether there is a role played by viewing distance in velocity constancy it is possible that the cues—such as knowledge of the visual layout and lack of motion parallax to promote disparity—conflicting with the depth cues provided by disparity alone were sufficient to preclude an accurate estimate of inter-object distance, and consequently velocity constancy was not attained (Zohary and Sittig 1993). It remains an intriguing question whether the combination of disparity and motion parallax in a reduced visual environment will lead to velocity constancy.

Relative object size and familiar size can contribute to velocity constancy in addition to the contribution of relative object–observer distance. At least part of the object-size cues effect is likely to be acting through a temporal-frequency-dependent coding scheme. Size can determine temporal-frequency encoding because changes in size change the spatial-frequency content of the object. Since $f_t = v \times f_s$, where f_t is the temporal frequency, f_s is the spatial frequency, and v is the angular velocity, when v remains constant and f_s increases, which happens when an object is made smaller (figure 4), then f_t increases as does perceived velocity. An alternative way of expressing the temporal-frequency coding idea is to consider that the intrinsic size scale is set by the object

itself; thus an object travelling at a fixed velocity will traverse a certain number of its own lengths per unit time. An object of twice the size, travelling at the same fixed velocity, will traverse only half the number of its own lengths per unit time. Both schemes of temporal-frequency coding would account for the size effects that are observed at a constant viewing distance (figure 5). Size can also act via scaling the perceived inter-object distance. However, the most parsimonious explanation for size effects, especially those of the size-scaling experiments (figure 5), is through the temporal-frequency encoding route. This conclusion is supported by the finding that a change in vehicle size triggered a change in perceived velocity, which was independent of the vehicle's distance (the curves in figure 5 have a constant offset). This means that traversed distance per se cannot completely account for the subjects' judgments of the physical velocity of the vehicles, since the traversed distance is independent of object size at a fixed distance. Furthermore, although the constancy-factor scale is continuous but not necessarily linear it is instructive that perspective size adds the same number of constancy units in the absence or in the presence of disparity and motion-parallax cues (figure 4b).

In contrast to some recent studies (McKee and Welch 1989; Zohary and Sittig 1993), we found that relative distance cues could make a significant contribution towards velocity constancy. Motion parallax alone, and particularly when combined with disparity cues, yielded a significant improvement in velocity constancy (figures 4a and 4b). Although there are a number of situations where depth judgments are not improved with the combination of cues (McKee and Smallman 1998), these are mainly in situations where distances are relatively small. In the experiments reported here, the observer–object distances (10 m – 40 m) and the object–object distances (10 m – 20 m) are relatively large in comparison to those in previous studies. Other cues that give an overall impression of depth to the simulated VR scene—texture and road perspective—seem to play a rather minor role in scaling velocity judgments (figures 4a and 4b). Similarly, relatively weak effects of texture are also found in shape-judgment tasks (Johnston et al 1993).

Even though the truck, used as one of the objects to test the effects of familiarity, is slightly larger than the oversized VW beetle, it is consistently perceived as moving faster than the latter (figure 6). The results of the size experiment (figure 5), which showed that larger vehicles are perceived as moving slower, cannot account for this result. Therefore, either some aspect of the truck's appearance, such as the smaller wheel size relative to the oversized car, or prior knowledge about the relationship between the truck and the environment, must be responsible for this result. We currently favour the familiarity interpretation of this result which implies that velocity constancy is not just a bottom–up process driven by sensory information. Instead, prior knowledge about the very nature of the moving objects is included in the processing of the sensory information, as had been suggested also by Hershenson and Samuels (1999). Just as some aspects of size constancy can be affected by experience or familiarity (Kaufman 1982), so we can see the same type of effect in velocity-constancy judgments.

References
Adelson E H, Bergen J R, 1985 "Spatiotemporal energy models for the perception of motion" *Journal of the Optical Society of America A* **2** 284 – 299
Brown J F, 1931 "The visual perception of velocity" *Psychologische Forschung* **14** 199 – 232
Bülthoff H H, Mallot H, 1988 "Integration of depth modules: stereo and shading" *Journal of the Optical Society of America A* **5** 1749 – 1758

Distler H K, Veen H A H C van, Braun S J, Heinz W, Franz M O, Bülthoff H H, 1998 "Navigation in real and virtual environments: Judging orientation and distance in a large-scale landscape", in *Virtual Environment '98: Proceedings of the Eurographics Workshop in Stuttgart Germany, 16 – 18 June 1998* Eds M Göbel, J Landauer, M Walper, U Lang (Vienna: Springer) pp 124 – 133

Epstein W, 1978 "Two factors in the perception of velocity at a distance" *Perception & Psychophysics* **24** 105 – 114

Epstein W, Cody W J, 1980 "Perception of relative velocity: A revision of the hypothesis of relational determination" *Perception* **9** 47 – 60

Hershenson M, Samuels S M, 1999 "An airplane illusion: apparent velocity determined by apparent distance" *Perception* **28** 433 – 436

Johnston E B, Cumming B G, Landy M S, 1994 "Integration of stereopsis and motion shape cues" *Vision Research* **34** 2259 – 2275

Johnston E B, Cumming B G, Parker A J, 1993 "Integration of depth modules: stereopsis and texture" *Vision Research* **33** 813 – 826

Kaufman L, 1982 *Sight and Mind: an Introduction to Visual Perception* (New York: Oxford University Press)

Landy M S, Maloney L T, Johnston E B, Young M, 1995 "Measurement and modeling of depth cue combination: In defense of weak fusion" *Vision Research* **35** 389 – 412

Levitt H, 1970 "Transformed up – down methods in psychoacoustics" *Journal of the Acoustical Society of America* **49** 467 – 477

McKee S P, Smallman H S, 1998 "Size and speed constancy", in *The Perceptual Constancies* Eds J Walsh, J J Kulikowski (Cambridge: Cambridge University Press) pp 373 – 408

McKee S P, Welch L, 1989 "Is there a constancy for perceived velocity?" *Vision Research* **29** 553 – 561

Norman J F, Todd T J, Perotti V J, Tittle J S, 1996 "The visual perception of 3-D length" *Journal of Experimental Psychology: Human Perception and Performance* **22** 173 – 186

Richards W, 1985 "Structure from stereo and motion" *Journal of the Optical Society of America A* **2** 343 – 349

Rock I, Hill A L, Fineman M, 1968 "Speed constancy as a function of size constancy" *Perception & Psychophysics* **4** 37 – 40

Rogers B J, Collett T S, 1989 "The appearances of surfaces specified by motion parallax and binocular disparity" *Quarterly Journal of Experimental Psychology A* **41** 697 – 717

Santen J P H van, Sperling G, 1985 "Elaborated Reichardt detectors" *Journal of the Optical Society of America A* **2** 300 – 321

Smith A T, Edgar G K, 1994 "Antagonistic comparison of temporal frequency filter outputs as a basis for speed perception" *Vision Research* **34** 253 – 265

Tittle J S, Braunstein M L, 1991 "Shape perception from binocular disparity and structure-from-motion", in *Sensor Fusion III: 3-D Perception and Recognition* Ed. P S Schenker *Proceedings of the SPIE* **1383** 225 – 234

Veen H A H C van, Distler H K, Braun S J, Bülthoff H H, 1998 "Navigating through a virtual city: Using virtual reality technology to study human action and perception" *Future Generation Computer Systems* **14** 231 – 242

Wallach H, 1939 "On the constancy of visual speed" *Psychophysical Review* **46** 541 – 552

Watson A B, Ahumada A J, 1985 "Model of human visual-motion sensing" *Journal of the Optical Society of America A* **2** 322 – 342

Wist E R, Diener H C, Dichgans J, 1976 "Motion constancy dependent upon perceived distance and the spatial frequency of the stimulus pattern" *Perception & Psychophysics* **19** 485 – 491

Zohary E, Sittig A C, 1993 "Mechanisms of velocity constancy" *Vision Research* **33** 2467 – 2478

p