Effects of spatial background contexts on visual object representation

（空間的背景情報が物体認知に及ぼす影響）

簡　頌恩

Abstract

Chapter 1 introduces the present thesis. Vision is important for goal-oriented action and interacting with the surrounding environment. Goal-oriented actions such as grasping and reaching rely on vision to localize objects in visual space. Visual attention is crucial for localizing visual objects precisely, because attention is concentrated on selected stimuli while ignoring others. Attended stimuli would be processed with more resources. Therefore, spatial perception at attended locations and the performance of goal-oriented action would be enhanced. However, a distracting stimulus presented briefly outside the current attentional focus induces involuntary attentional shifting. This shift of attention would induce errors in localizing objects. Many studies have shown that both preceding and succeeding peripheral transients can induce mislocalization of visual targets in different directions; however, it is not fully understood whether a single mechanism underlies the mislocalization induced by dynamic attentional shifting in visual space. Furthermore, the perceptual system integrates information from different modalities to construct the world. Therefore, peripheral transients from modalities other than vision would also induce spatial distortion in visual space. The present thesis aims to propose an integrated account on how peripheral transients affect visual localization.

In Chapter 2, I applied the experimental paradigm of attentional repulsion and attraction to explore whether a single mechanism underlies mislocalization. Two vertical lines appeared above and below the center fixation as visual targets. Visual transients were presented before/after the target as cues to attract attention. Participants were asked to judge horizontal misalignment of two vertical lines. When the preceding cues were presented, the perceived location of the visual target was shifted away from them (attentional repulsion). In contrast, the perceived location of the targets was attracted to the succeeding cues (attentional attraction). Experiments were conducted to examine (1) whether the effect sizes of the repulsion effect and the attraction effect would differ with the same experimental parameters, (2) whether the two effects would interact, and (3) whether the locus of attentional focus or the direction of dynamic attentional shift would determine the mislocalization effects. The results showed that (1) the effect size of spatial distortion was smaller when the preceding cues were presented than the succeeding cues with the same experimental parameters, (2) the two effects did not interact with each other, and (3) static attentional focus on cues induced the repulsion effect, and the dynamic attentional shift induced the attraction effect. This suggests that the repulsion and attraction effects are independent from each other.

In Chapter 3, the effect of depth perception on spatial distortion was explored. Visual cues were presented before/after the targets in different depth planes with respect to the target under both binocular and monocular stereopsis viewing. The results showed that spatial distortion was stronger when both preceding and succeeding cues were presented in the depth plane farther away from the target. However, depth modulation of preceding cues was only observed under binocular stereopsis viewing. On the other hand, depth modulation of succeeding cues was observed under both binocular and monocular stereopsis viewing. Further experiments found that depth modulation of preceding cues was observed when observers shifted attention from the fixation point to the depth plane farther away with respect to the fixation point. This suggests that attentional resources are allocated more densely to near space. The additional cost of shifting attention along depths could be compensated by the enhanced state of visual attention in near space. Therefore, larger displacement is only observed when shifting attention to far space. On the other hand, a larger attraction effect was observed with closer target-cue distance even when no depth information was provided and the physical size of the cue was the same. This implies that the attraction effect was mainly modulated by retinal coordinates. The results of the experiments described in Chapter 3 also suggest that the underlying mechanisms of the repulsion and attraction effects are different.

Chapter 4 explores how the perceptual system integrates information of sensory transients from modalities other than vision to decide the perceived location of a visual target. In this chapter, a visual object moved toward the center of the computer screen and disappeared abruptly. A brief sound was presented to observers around the moment of visual motion offset. Observers were asked to report the perceived vanishing location of the visual target. The results of experiments showed that when the sound was presented before the visual motion offset, the perceived visual offset location was shifted backward. This indicates that the mislocalization resulting from the timing of visual motion offset was attracted to the timing of the brief sound. In addition, when the sound was lateralized, the sound's spatial information influenced the perceived location of the visual motion offset. Larger spatial attraction was observed when brief sound was presented on the same side as the start position of the target. The effect of the lateralized sound was mainly noted in the spatial domain. Therefore, both temporal and spatial information of the brief sound can induce mislocalization effects on the perceived location of visual motion offset.

Chapter 5 summarizes the findings of the experiments from Chapters 2 through 4 and discusses potential applications in everyday life. Finally, possible directions for future research are suggested.

Index

Figure index

Chapter 1: Introduction

Background

Spatial distortion in visual space

The visual system processes information received from visible light to construct a visual perception of the surrounding environment. Vision is essential for human survival in everyday life, because it helps us comprehend and interpret the surrounding environment. The ability to localize an object in visual space is particularly important for visual search, decision making, and goal-oriented behaviors. For example, in order to launch an application on a mobile device, it is necessary to (1) find the icon's location on the display first and (2) move a finger to touch it. The first step relies on localizing the icon; to complete the second step, it is necessary to estimate the distance between the icon and the finger. If the icon cannot be correctly localized, it is possible that the attempt to launch the application on the mobile device would fail. However, spatial representation is not always perfect, because the brain's ability to detect and process sensory signals is limited. Thereby, the perceived location of a visual object is not always exactly the same as its physical location due to errors that occur while processing information gained through successive views.

Many studies have investigated spatial distortion in visual space when observers attempt to make eye movements. Helmholtz (1925) found that the world moved in the direction opposite passive eye movements elicited by pressing on the canthus. Furthermore, when the eye was paralyzed, the world moved in the direction of the intended eye movement. Similarly, perceived locations of visual stimuli are mislocalized if they are presented close to the onset of a saccadic eye movement. If a visual stimulus were presented in a dark space, its perceived location would be displaced in the direction of the eye movement (Cai *et al.*, 1997). However, when visual references are available after the saccadic eye movements, mislocalizations of visual stimuli result from the compression of visual space towards the target of saccadic eye movement (Ross *et al.*, 1997). These results suggest that perceptual stability relies on not only retinal images but also information on eye position provided by the oculomotor system. In addition, mislocalization of visual objects can also be observed when saccadic eye movements are absent. For example, the compression of visual space could also be observed if observers shifted attention toward visual stimuli while maintaining fixation on specific positions. This may be because shifting spatial attention can signal preparation of the motor system to shift eye position (Hamker *et al.*, 2008; Zirnsak *et al.*, 2010).

When saccadic eye movements are absent, mislocalization of visual objects can be observed in both static and dynamic displays. In the flash-lag effect, the perceived location of a moving object appears ahead of a spatially aligned flash object. In other words, the perceived location of the moving object is displaced forward in the direction of motion (Figure 1.1, top panel). There are several possible explanations for the underlying mechanism of the flash-lag effect. In the motion extrapolation hypothesis, the flash-lag effect is the result of forward displacement of the perceived location of the moving object based on analyzing its speed and trajectory (Nijhawan, 1994). The latency difference hypothesis suggests that the static flash is processed more slowly than the moving object (Purushothaman *et al.*, 1998; Whitney & Murakami, 1998; Whitney, Murakami & Cavanagh, 2000). The temporal integration hypothesis proposes that location information for the moving object within a certain temporal window is integrated to determine its perceived location (Brenner & Smeets, 2000; Brenner *et al.*, 2006; Eagleman & Sejnowski, 2000; Krekelberg & Lappe, 2000; Roulston *et al.*, 2006). The motion signals after the flash displace the object's perceived location in the direction of motion (Eagleman & Sejnowski, 2000; 2007). Although these hypotheses differ, they all reflect the need for a visual strategy to compensate the inevitable delay in processing sensory signals.

Mislocalization can also be observed in spatial short-term memory. In studies of representational momentum (RM; Freyd & Finke, 1984; Hubbard & Bharucha, 1988; Figure 1, bottom panel), the memory of a moving object's final position is displaced in the direction of motion (for reviews, see Hubbard, 2005). RM could result from the mental representation of the object's motion persisting for a brief period after the abrupt offset of visual motion (Teramoto *et al*., 2010). RM can be observed not only in continuous motion, but also in implied motion induced by frozen-action photographs (Futterweit & Beilin, 1994; Freyd, 1983; Freyd *et al.*, 1988). This suggests that RM is not only involved with low-level perceptual processing, but also high-level cognitive mechanisms. Other physical principles exhibit similar effects. Greater displacement of remembered final position is observed for descending objects than ascending objects, suggesting that the awareness of gravity influences the magnitude of this mislocalization effect. Therefore, mislocalization of visual objects could occur both in the early stage of visual processing and in the late stage, during which the brain integrates all information to form a unified spatial representation.

Both the flash-lag and RM phenomena reflect strategies used by the visual system to counteract the delay in processing sensory stimuli. However, these strategies also induce systemic errors when localizing visual objects.

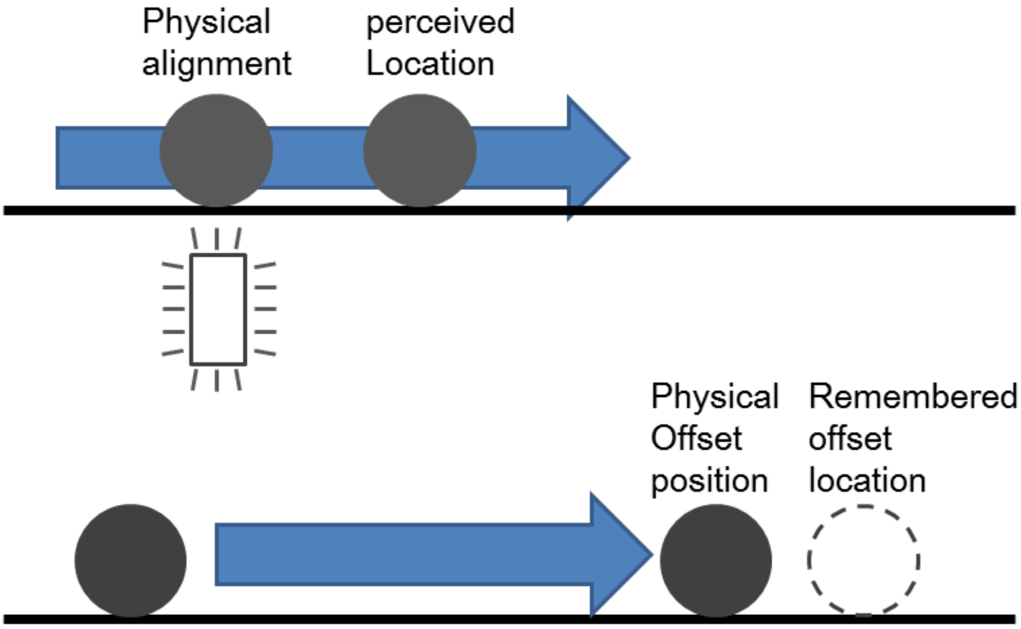


Figure 1.1 Semantic illustration of the flash-lag effect (top) and representational momentum (bottom)

Attention and spatial perception

The brain's ability and resources for processing sensory stimuli from the surrounding environment are both limited, as mentioned above. Attempting to process all sensory stimuli received from the environment would be too inefficient to support adaptive behavior. Therefore, the brain must identify and focus only on the most relevant information. Attention is the process responsible for selecting relevant information while ignoring other stimuli and allocating processing resources. In spatial perception, attention allocates processing resources to attended objects in visual space. After processing the selected objects, the information is integrated into spatial short-term memory to form a spatial representation (LaBar *et al.*, 1999).

Sensory information can be selected by both top-down and bottom-up attention. In top-down attention, observers voluntarily shift attention to stimuli relative to observers' intentions or goals; this is often labeled *goal-driven* or *endogenous attention* (Figure 1–2, top panel). In top-down attention, attentional shift is under the observers’ control. In bottom-up attention, however, attention is attracted to stimuli that are outside the current focus of attention; this is often labeled *stimulus-driven* or *exogenous attention* (Figure 1–2, bottom panel). Salient stimuli such as rapid motion or new objects in the visual field would attract attention more easily. Exogenous attention is important for adaptive behaviors, because stimuli appearing suddenly in the surrounding environment might be potentially threating (e.g., approaching predators). In bottom-up attention, the attentional shift is out of observers’ control.

As attention is meant to allocate more processing resources to the selected location in visual space, both endogenous and exogenous attention can enhance spatial perception at the selected location. In studies of endogenous attention, when observers were asked to detect a visual target as quickly as possible, reaction time was faster when a central arrow was briefly presented in the center of the display indicating the visual field in which the target would appear (Posner, 1980; Posner *et al.*, 1980). In studies of exogenous attention, accuracy and reaction time to identify a visual target were both improved when a flash was briefly presented as a visual cue in the same location as the target (Eriksen & Collins, 1969; Eriksen & Rohrbaugh, 1970; Posner, 1980; Posner *et al.*, 1980). These studies indicate that attention can enhance both spatial resolution and reaction speed at selected spatial locations.

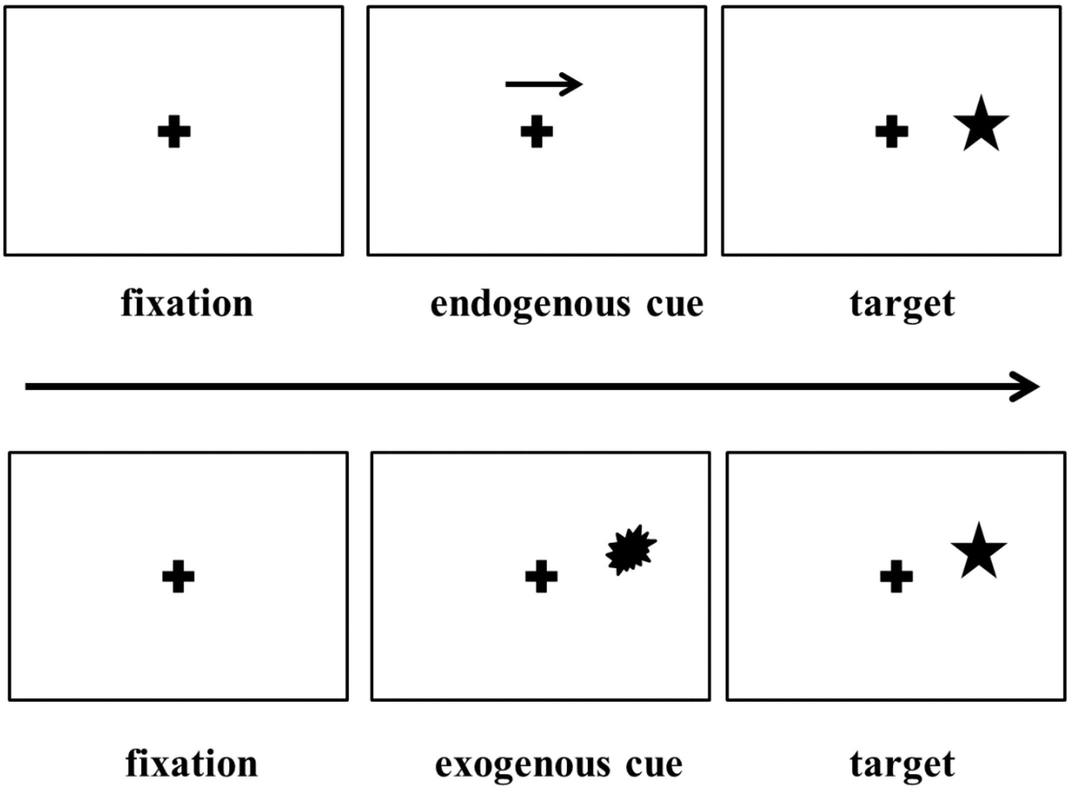


Figure 1.2 Examples of the Posner paradigm. (1) endogenous cues (top), (2) exogenous cues (bottom)

Neural mechanisms of reorienting spatial attention

Neuroimaging studies indicate that two neural networks are involved in reorienting attention (for reviews, see Corbetta *et al.*, 2008): the dorsal attention network and the ventral attention network. The dorsal attention network is involved in top-down attention by reorienting attention to stimuli based on observers' intentions and goals. It is activated by observers' expectations (Corbetta *et al.*, 2000; Hopfinger *et al.*, 2000; Kastner *et al.*, 1999; Shulman *et al.*, 1999), or by preparations of specific responses (Astafiev *et al.*, 2003; Connolly *et al.*, 2002). It is bilateral and includes the intraparietal sulcus (IPS) in the dorsal parietal cortex and the frontal eye field (FEF) in the dorsal frontal cortex. The ventral attention network is involved mostly in bottom-up attention. It is activated by the detection of salient stimuli in the visual field. Further, its activation increases along with the dorsal attention network when behaviorally relevant stimuli are detected (Corbetta *et al.*, 2000). The ventral attention network is right-hemisphere lateralized and includes the right temporal-parietal junction (RTPJ) and the right ventral frontal cortex (RVFC).

The attentional repulsion and attraction effect

Involuntary shifts of attention are often induced by salient stimuli, for example, a sudden sound, a flash, or a white rectangle against a black background. The sudden appearance of a visual transient attracts observers' attention within 200 ms, despite all efforts to attend to something else (Eriksen & Collins, 1969; Hikosaka *et al.*, 1993a, 1993b; Nakayama & Mackeben, 1989). Spatial attention would be directed to the location of the visual transient. Suzuki and Cavanagh (1997) demonstrated that involuntary attention-shifting also incurs spatial distortion around the attended space. In their study, observers were asked to maintain focus on a fixation point at the center of the display. Two disks were briefly presented as visual cues in diagonally opposed positions (top-left/bottom-right or top-right/bottom-left, Figure 1.3). Then, two vertical lines appeared above and below the fixation point as visual targets. Observers were asked to judge the horizontal misalignment of the two vertical lines. The results showed that the perceived locations of the vertical lines were displaced away from the visual cues. In other words, if the cues appeared at the top-left/bottom-right position, observers are more likely to judge that the top line was on the right side of the bottom line. They refer to this phenomenon as the *attentional repulsion effect* (Figure 1-2, top panel). In their series of experiments, they demonstrated that the repulsion effect is not caused by nonattentional effects such as apparent motion or figural aftereffects.

They use the *position-coding unit* hypothesis to explain the repulsion effect as the cost of enhanced spatial perception at the cue's location. In other words, spatial locations are represented by the overall response patterns of a population of position-coding neural units. The perceived location of visual objects is represented by the centroid of the response distribution of these units. Visual transients that suddenly appear in peripheral space inevitably attract attention toward it. The target's centroid of response distribution is skewed from the cue, the locus of attention. This might result from shrinking or recruitment of the receptive field around the focus of attention. Consequently, spatial distortion is observed.

Interestingly, when the temporal sequence of the stimuli used in Suzuki and Cavanagh's experiments is reversed, the perceived location of the visual target is displaced toward the visual cue. For example, observers are more likely to judge that the top line is on the left side of the bottom line, if the cue is presented at the top-left/bottom-right position after the target (Ono & Watanabe, 2011). This is referred as the *attentional attraction effect* (Figure 1–2, bottom panel). They also demonstrated that the attraction effect could not be explained by perceptual effects such as apparent motion, suggesting that it is also induced by involuntary attentional shift from the target to the cue. However, the position-coding unit hypothesis does not explain the attraction effect. The direction of mislocalization depends on whether the cue is presented before or after the target.

Given that the stimuli in Ono and Watanabe's experiments are similar to those in Suzuki and Cavanagh's experiments, they propose that both the repulsion and attraction effects are induced by the same underlying mechanism: the overshoot of the attentional shift (Shim & Cavanagh, 2004; Yamada *et al.*, 2008). In the repulsion effect, the overshoot of attentional shift from the cue to the target displaces the perceived location of the target in the direction of the attentional shift. In the attraction effect, the perceived location of the target is displaced in the direction of attentional shift in a short-term memory trace.

The idea of dynamic attentional shift might not be appropriate for the repulsion effect. In the attentional repulsion paradigm, observers shift attention from peripheral transients to the central target. Assuming that spatial perception is enhanced in the attended regions, spatial distortion should not be observed if the spatial attention does shift to the central region.

On the other hand, it is reasonable that dynamic shift from the central target to the peripheral transient would induce the attraction effect. When attention is attracted to the peripheral transient, the location information of the sensory transient would also be processed and remembered in the short-term memory trace. Previous research describes a *landmark attraction effect* in which the remembered location of a static target is biased toward a landmark in visual space (Bryant & Subbiah, 1994; Hubbard & Ruppel, 2000; McNamara & Diwadkar, 1997; Nelson & Chaiklin, 1980; Sadalla *et al.*, 1980; Simmering *et al.*, 2008; Spencer *et al.*, 2006). The visual system would use the location information of visual stimuli that attracts attention to modulate the location information of the target in visual short-term memory (Yamada *et al.*, 2011). Therefore, dynamic shift toward the peripheral target could bias the perceived location of the visual target in the short-term memory trace.

If the repulsion and attraction effects are induced by different underlying mechanisms, it follows that peripheral transients could induce spatial distortion in two different ways. On one hand, peripheral transients could impair information processing for the visual target. On the other hand, peripheral transients could bias the remembered position of the visual target toward themselves.

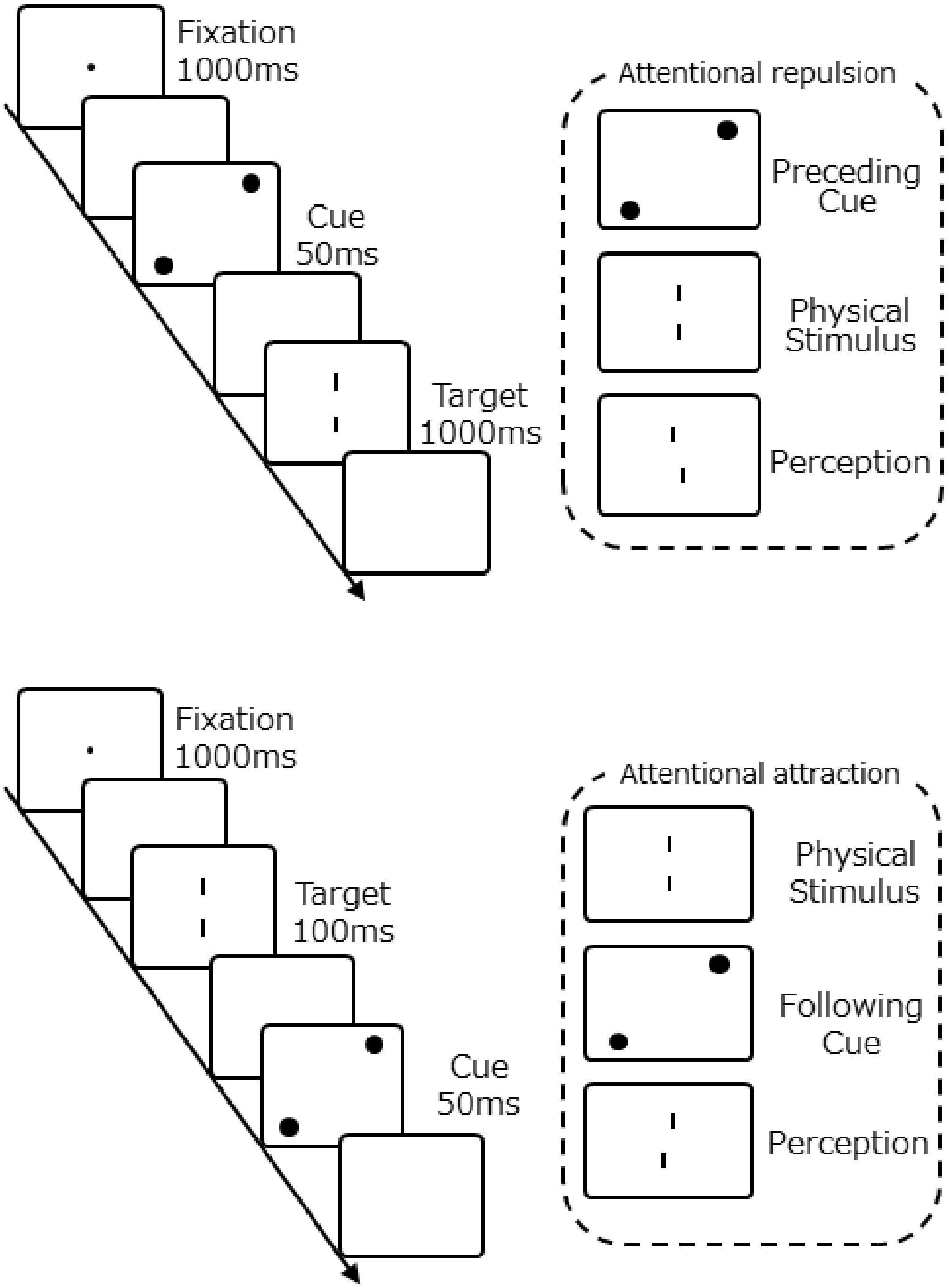


Figure 1.3 Semantic illustrations of the attentional repulsion effect (top) and the attentional attraction effect (bottom)

Crossmodal interaction in localization of visual objects

The perceptual system receives information through different yet interacting sensory modalities. The inputs from different sensory modalities interact in various ways. The perceptual system also integrates information from modalities other than vision to localize visual stimuli. Vroomen and de Gelder (2004) found that the magnitude of the flash-lag effect was reduced when a brief sound was presented before or simultaneously with the flash. Heron and colleagues (2004) demonstrated that the location of a horizontally moving object was perceptually displaced forward with respect to the direction of previous motion when a sound was presented after the actual bounce, and the perceived bounce position was shifted in the direction opposite to its previous motion when a sound was presented before the actual bounce. The multisensory effect of a brief sound could also be observed when localizing a static visual object. Arnott and Goodale (2006) demonstrated that the attentional effect could also be induced by presenting lateralized sound as peripheral cues, showing that auditory spatial information can displace the perceived positions of static visual stimuli. These findings indicate the possibility that the location of sound may affect retinotopic coding.

Two-stream hypothesis for vision

Visual perception plays an essential role in our everyday life, because it helps humans understand and interpret the surrounding environment by processing light information. Images of visual stimuli are projected onto the retina. The outputs of the retina pass the optic nerve, cross and split at the optic chiasm, then pass through the optic tract to the lateral geniculate nucleus (LGN). Information then passes to the primary visual cortex (V1) in the occipital lobe. V1 transmits information to two separate pathways, the dorsal and ventral streams. The dorsal stream begins from V1, continues through the prestriate cortex (V2) to the dorsomedial area (V6) and visual area MT (middle temporal) (V5), then stretches into the posterior parietal cortex (PPC). The dorsal stream, referred to as the “where pathway,” is associated with the visual representation of object locations. The ventral stream also begins from V1, continues to V2, then passes through the visual area V4 in the extrastriate visual cortex to the inferior temporal cortex. Associated with recognition and object representation, the ventral pathway is referred to as the “what pathway” (Mishkin *et al.*, 1983).

Vision directs action by transforming visual inputs into appropriate motor outputs (Goodale & Humphrey, 1998). It has recently been suggested that the two separate visual pathways are critical for visual perception and visually guided action, respectively (Goodale & Milner, 1992). According to the Goodale and Milner model, the ventral stream is responsible for vision-for-perception: it constructs the perceptual representation of the world. The ventral stream also provides a perceptual foundation for offline control of action and incorporating stored information into the control of current action (Goodale, 2011), while the dorsal stream is responsible for mediating on-line visually guided actions directed to visual targets (Figure 1-4).

Previous research on size-contrast illusions (e.g., the Ebbinghaus illusion) demonstrate that spatial distortions in perception do not affect the performance of visually guided actions, such as grasping, supporting the dissociation between the ventral and dorsal pathways (Aglioti *et al.*, 1995; Haffenden and Goodale, 1998, 2000; Haffenden *et al.*, 2001). However, it is also suggested that spatial distortion in visual space affects performance in visuomotor tasks when it relies on on-line control. Notably, the attentional repulsion effect could also be observed when observers were asked to judge a target’s location by both mouse pointing and guided limb localization tasks (Pratt & Turk-Browne, 2003). Furthermore, motor responses produce larger localization errors than cognitive judgments (Kerzel, 2003; Yamagishi, *et al.*, 2001).

Therefore, spatial distortion in visual space not only provides erroneous interpretations of the surrounding environment but also potentially affects the precision of actions guided by vision. Impairment of visuomotor performance is even larger when responses are made in a dynamic environment.

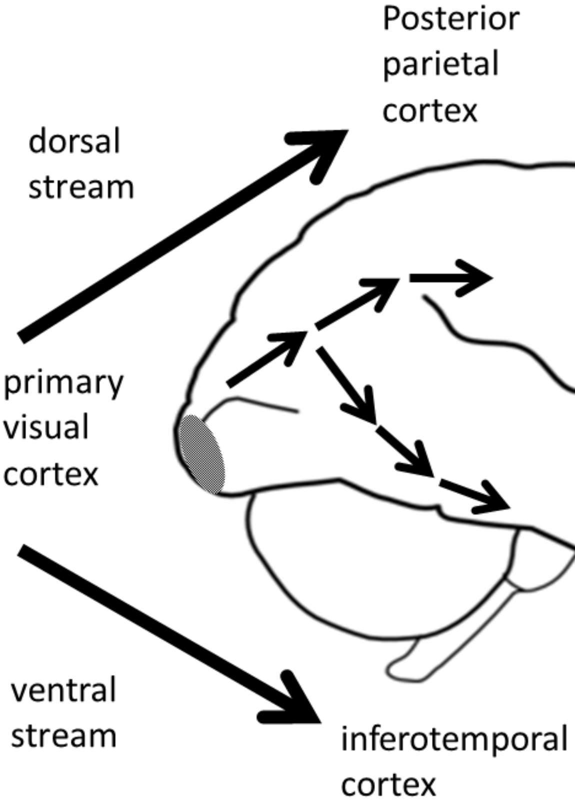


Figure 1.4 Separate visual pathway for the dorsal stream (vision-for-action) and the ventral stream (vision-for-perception)

Objective of the thesis

Previous research suggests it is possible that (1) peripheral transients induce spatial distortion by either impairing the processing of target information or biasing the perceived location of visual targets in short-term memory, and (2) the perceptual system integrates information from different modalities to construct the world. The present thesis aims to propose an integrated account of how dynamic attentional capture induced by peripheral stimuli influences the visual system in localizing visual stimuli. Localizing a target in visual space involves four steps: (1) receiving information on the visual stimuli, (2) information processing, (3) information integration, and (4) forming location representations of the visual stimuli (Figure 1-5). The image of the target is projected onto the retina and transmitted to the visual cortex. Basic visual information is processed in areas such as V1 and V5 (the MT area). Previous studies have indicated that the posterior parietal cortex (PPC) is susceptible to interruption by distracting stimuli (Knudsen, 2007) and responses to stimuli away from the target location of memory-guided saccade (Powell & Goldberg, 2000). Therefore, I propose that information necessary to form location representations is integrated in the PPC.

A sensory transient that captures attention incurs mislocalization in two aspects. It draws attention away from the visual target; thus, the attentional resources distributed to the target are reduced. This impairs performance at the information processing stage and shifts the perceived location of the visual target away from the peripheral transients (the repulsion effect). The other is that spatial information on peripheral stimuli is also processed when attention is attracted toward it. Consequently, spatial information on the peripheral stimulus is incorrectly integrated with other necessary information at the information integration stage. It biases the perceived location of the visual target toward the peripheral transient (attraction effect). Therefore, while both the repulsion and attraction effects are induced by peripheral stimuli, they occur at different stages of spatial information processing and act independently. Furthermore, the visual system combines multisensory information to form the location representation of a visual target. Therefore, the perceived location of the visual target would be biased toward peripheral transients from modalities other than vision.

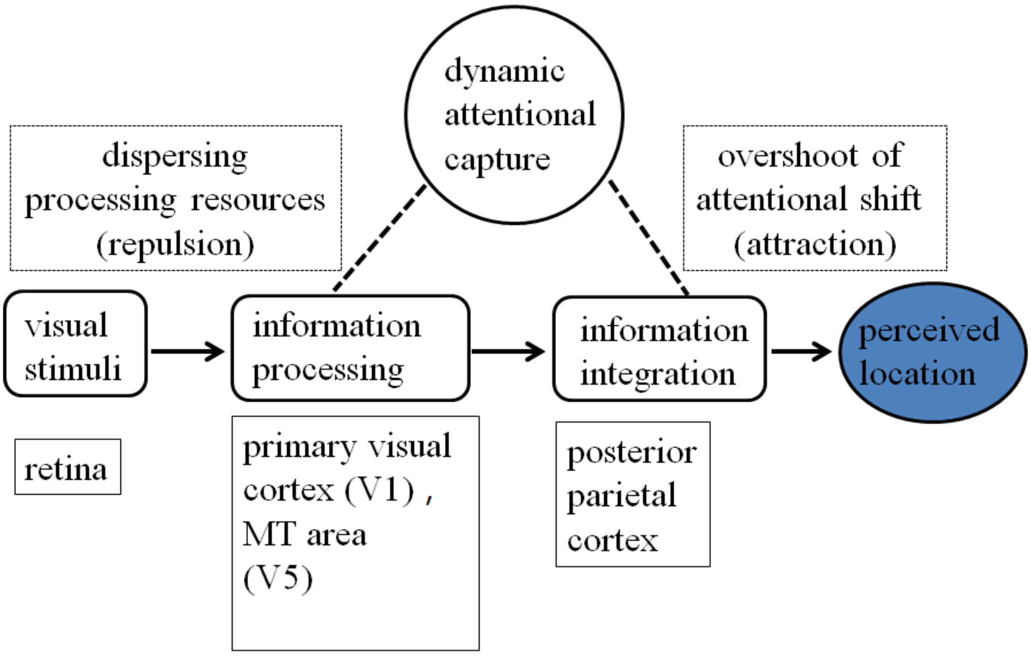


Figure 1.5 The attention account for spatial distortion

Outline of the thesis

In the present thesis, I explore how the visual system processes and integrates information to localize objects in visual space by presenting observers with distracting peripheral transients with different properties.

In the experiments presented in Chapter 2, I examined whether the repulsion and attraction effects are induced by the same or different mechanisms by examining (1) whether the two effects would interact, and (2) whether the locus of focus attention or the direction of attentional shift would determine the mislocalization effects. The results suggest that the repulsion effect is induced by the cost of perceptual enhancement at the cue’s location, and that the attraction effect is induced by the dynamic attentional shift from the target to the cue. Furthermore, our findings indicated that the two underlying mechanism did not interact with each other.

In Chapter 3, I tested whether depth perception in 3D space would influence the effects of spatial distortion by presenting peripheral cues at different depths. The results suggest that attentional shift in depth modulates the repulsion effect, but the attraction effect is mainly modulated by changes in targets’ positions on the retinal coordinate. It further supports the idea that the underlying mechanisms of the repulsion and attraction effect are different.

In Chapter 4, I examined the crossmodal effect induced by auditory transients on localizing visual objects. The results showed that a brief sound presented around the offset of a visual moving target could bias the perceived offset position toward the brief sound in the temporal domain. In addition, the stronger attentional capture induced by presenting lateralized sound to the same visual field as the target could produce a larger displacement of the perceived offset position. This implies that the multisensory effect on the localization of visual objects was related to the magnitude of dynamic attentional capture.

Finally, in Chapter 5, I summarize our empirical findings and suggest some directions for future research and possible applications in everyday life.

Chapter 2: The attentional repulsion and attraction effect are independent

Objective

When attention is directed to specific regions other than the fovea by presenting peripheral transients, the shift of visual attention improves the detection and identification of peripheral visual objects. This also results in better spatial resolution of peripheral regions (Posner & Peterson, 1990; Heemskerk *et al.,* 1996; He *et al.*, 1997; Yeshurun & Carrasco, 1998). However, this also causes systemic distortions in perceived locations around the focus of attention. Suzuki and Cavanagh (1997) showed that the perceived location of a visual target appeared to displace away from preceding visual cues. This is referred to as the attentional repulsion effect. They proposed that this effect reflects the cost of perceptual enhancement at the position of briefly presented cues. This account is supported by recent studies that suggest that attentional repulsion is caused by shifts of attention toward the cued locations, and that the magnitude of the repulsion effect depends on the cue-target distance (Arnott & Goodale, 2006; Pratt & Arnott, 2008; Kosovicheva *et al.,* 2010).

However, while Ono and Watanabe (2011) used visual stimuli similar to those used in Suzuki and Cavanagh's experiments (1997), the temporal sequence of the stimuli was reversed. In their study, two vertical lines appeared first and were followed by visual cues. They found the attentional attraction effect: if the visual cue appeared after the target, the perceived location of the target shifted toward the location of the following cue. This shows that attention had a retrospective influence on spatial perception, and that different temporal sequences of identical visual stimuli can result in spatial distortions of perceived locations in opposite directions. They proposed that both the repulsion and attraction effects result from the overshoot of the dynamic attentional shift. In the attraction effect, attentional shifts from the target to the cue shift beyond the cue’s actual location. In the repulsion effect, however, attentional shifts from the cue to the target shift beyond the target’s actual location. Previous research on representational momentum and the flash-lag effect support the idea that the mislocalization of perceived location is caused by this overshoot of attention (Yamada *et al.*, 2008; Shim & Cavanagh, 2004).

This chapter aimed to explore the unanswered question of whether a single mechanism underlies both the attentional repulsion and attraction effects. Alternatively, the attentional repulsion effect reflects the cost of enhanced spatial perception performance and the attentional attraction effect is caused by the dynamic shift of attention. To investigate this issue, this chapter examines the interaction between the repulsion and attraction effects. In Experiment 1, I examined whether effect sizes of the repulsion and attraction effects would differ with the same experimental parameters. Then, I examined whether the two effects would interact in Experiment 2. Finally, Experiment 3 examined whether the location of attentional focus or the direction of attentional shift would determine the spatial distortion effects.

Experiment 1

The purpose of Experiment 1 was to examine whether effect sizes of attentional repulsion and attraction would differ by estimating the point of subjective equality (PSE), at which the proportions of “right response” and “left response” are near equal in the experimental paradigm of attentional repulsion and attraction (Ono & Watanabe, 2011; Suzuki & Cavanagh, 1997).

Method

Observers

Fourteen paid volunteers participated. All of the observers had normal or corrected-to-normal visual acuity and were naïve to the purpose of this study.

Apparatus and stimuli

The observers viewed a 23-inch Mitsubishi CRT monitor with a refresh rate of 100 Hz at a distance of 60 cm. All stimuli were presented by using Psychtoolbox extensions for MATLAB (Brainard, 1997; Pelli, 1997). All stimuli appeared in white (69.40 cd/m^2) against a black (0.01 cd/m^2) background. The white central fixation point was 0.2° in diameter. The cue stimuli were two 1°-diameter disks presented at diagonal locations (i.e., top-left/bottom-right or top-right/bottom-left). The disks were vertically and horizontally displaced 3.5° from the fixation point. The probability of the cues appearing at either diagonal position was the same. The target stimuli were two vertical lines 2.5° above and below the fixation point. Each line was 1.0° long and 0.1° wide. The bottom line appeared just below the location of the fixation point. The top line could appear in one of eleven possible positions (Figure 2.1). The distance between possible positions was 0.1°. Five positions were located to the left of the bottom line, one was just above the center fixation, and the others were to the right of the bottom line. The leftmost/rightmost position was 0.5° away from the bottom line in the horizontal orientation. The top line was equally likely to appear in any of the eleven locations.

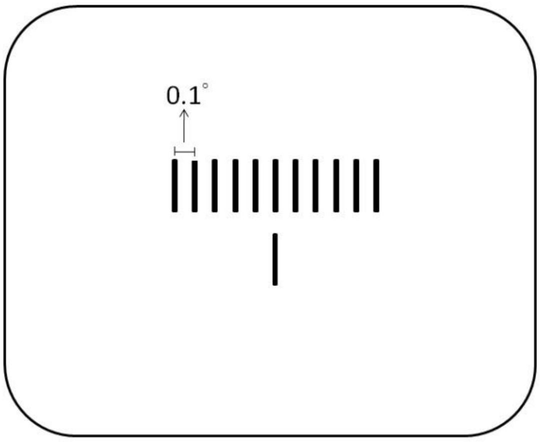


Figure 2.1 Possible locations for the top line. The top line appeared at one of 11 possible locations. The distance between possible locations was 0.1°. Five positions were located to the left of the bottom line, one was just above the center fixation, and the others were to the right of the bottom line. The leftmost/rightmost position was 0.5° away from the bottom line in horizontal orientation. The top line was equally likely to appear in any of the eleven locations.

Procedure

Observers initiated each trial by pressing the spacebar on a standard computer keyboard. Observers were instructed to maintain fixation on a fixation point that appeared for 1000 ms. After a 100-ms blank interval, the cue and target were presented in two different conditions. Under the cue-target condition, the cue was presented for 50 ms. After a 150-ms blank interval, two vertical target lines were presented for 100 ms. Under the target-cue condition, the target was presented for 100 ms first, followed by a 100-ms blank interval. The cue was then presented for 50 ms. The cue-target and target-cue SOAs (stimulus onset asynchrony) were 200 ms in all conditions (Figure 2.2) because previous research showed that the attentional repulsion and attraction effects peaked when the cue was presented around 200 ms before/after the target (Ono & Watanabe, 2011; Suzuki & Cavanagh, 1997). The two conditions were arranged in a random sequence. Observers were asked to judge whether the top line was located to the left or right of the bottom line by pressing the arrow keys on the keyboard, and instructed to choose a direction (forced-choice task) even if they perceived that the top and bottom lines were at the same horizontal position. Observers performed 44 practice trials and then completed 440 test trials.

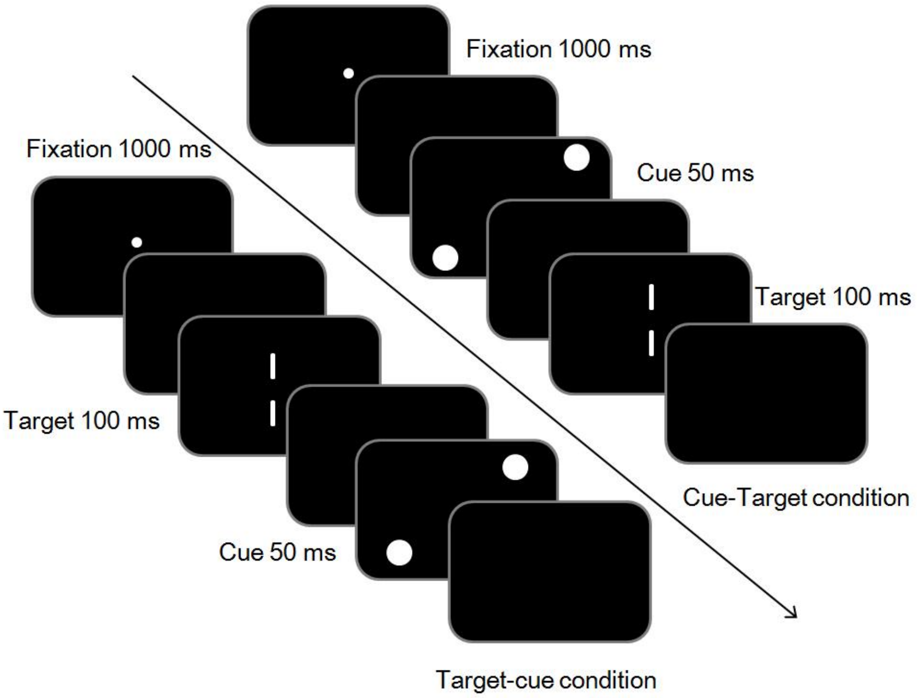


Figure 2.2. Trial events in Experiment 1. Observers were instructed to maintain fixation on a fixation point that appeared for 1000 ms. The cue and the target were presented in two conditions: [1] the cue-target condition and [2] the target-cue condition. The cue-target and the target-cue SOAs (stimulus onset asynchrony) were 200 ms in all conditions.

Results and discussion

Figure 2.3 shows the results of Experiment 1. The value on the vertical axis represents the proportion of “right” responses for the left-diagonal cue (top-left/bottom-right) and “left” responses for the right-diagonal (top-right/bottom-left) cue in each possible position under both cue-target and target-cue conditions. Positive values on the horizontal axis indicate that the top line was at the right of the bottom line for the left-diagonal cue and at the left of the bottom line for the right-diagonal cue, and vice versa. Note that the proportion of key-press responses in the opposite direction of the diagonal cue was generally larger under the cue-target condition (open circles) than under the target-cue conditions (filled circles). The point of subjective equality (PSE), defined as the intersection of the cumulative Gaussian curves with the line marked p = 0.5, was -0.054° (dotted curve) for the cue-target condition and 0.096° (solid curve) for the target cue condition. The PSEs depicted are the mean of individual PSEs. The coefficient of determination was 0.99, yielded by the pooled data for both conditions in Experiment 1. The mean of the PSE was significantly smaller in the cue-target condition than the target-cue condition (paired t-tests: $t(13) = 7.36$, $p < .001$). The PSEs were significantly different from zero (cue-target condition, $t(13) = 4.89$, $p < .001$; target-cue condition, $t(13) = 6.18$, $p < .001$). In addition, the means of the absolute values of PSEs differed significantly between the cue-target and target-cue conditions (paired t-tests: $t(13) = 2.25$, $p < .05$). The results confirmed that the direction of spatial distortion of the target stimuli depended on the temporal sequence of the cues. In addition, they showed that the effect size of spatial distortion was smaller when preceding cues were presented.

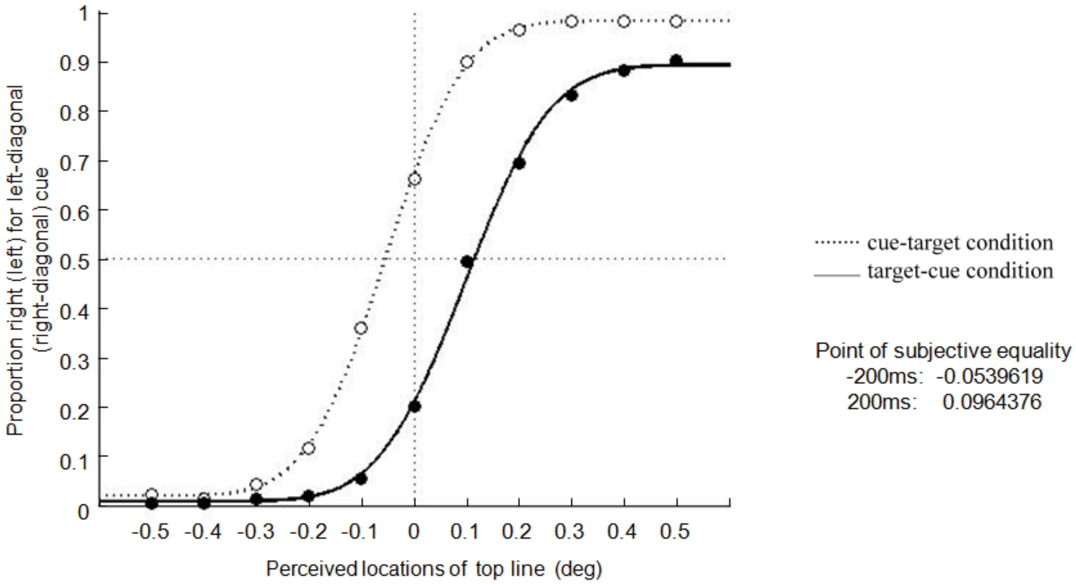


Figure 2.3 The results of Experiment 1. The vertical axis represents the proportion of “right” responses for the left-diagonal cue and “left” responses for the right-diagonal cue in each possible target position both under cue-target and target-cue conditions. Positive values on the horizontal axis indicate that the top line was to the right of the bottom line for left-diagonal cues and to the left of the bottom line for right-diagonal cues, and vice versa.

Experiment 2

The results in Experiment 1 showed that following cues induced larger spatial distortion than preceding cues in the opposite direction. Experiment 2 examined the hypothesis that the effects of preceding and following cues do not interact with each other. I measured spatial distortion when both preceding and following cues were presented. If the effects of preceding and following cues are independent, the effect size of spatial distortion would be a simple sum of repulsion and attraction effects, that is, the repulsion effect would negate the attraction effect, leaving a smaller attraction effect.

Method

Observers

Twelve paid volunteers were newly recruited and participated in Experiment 2. All of the observers had normal or corrected-to-normal visual acuity and were naïve to the purpose of this study.

Apparatus, stimuli, and procedures

The apparatus and stimuli were the same as used in Experiment 1. Observers initiated each trial by pressing the spacebar. They were instructed to maintain fixation on a fixation point that appeared for 1000 ms. After a 100-ms blank interval, the cue was presented for 50 ms. Then, after a 150-ms blank interval, the target was presented for 100 ms. After another 100-ms blank interval, the cue was presented again for 50 ms. The cue-target and target-cue SOA were both 200 ms. Each observer conducted 10 practice trials and 220 experimental trials (Figure 2.4).

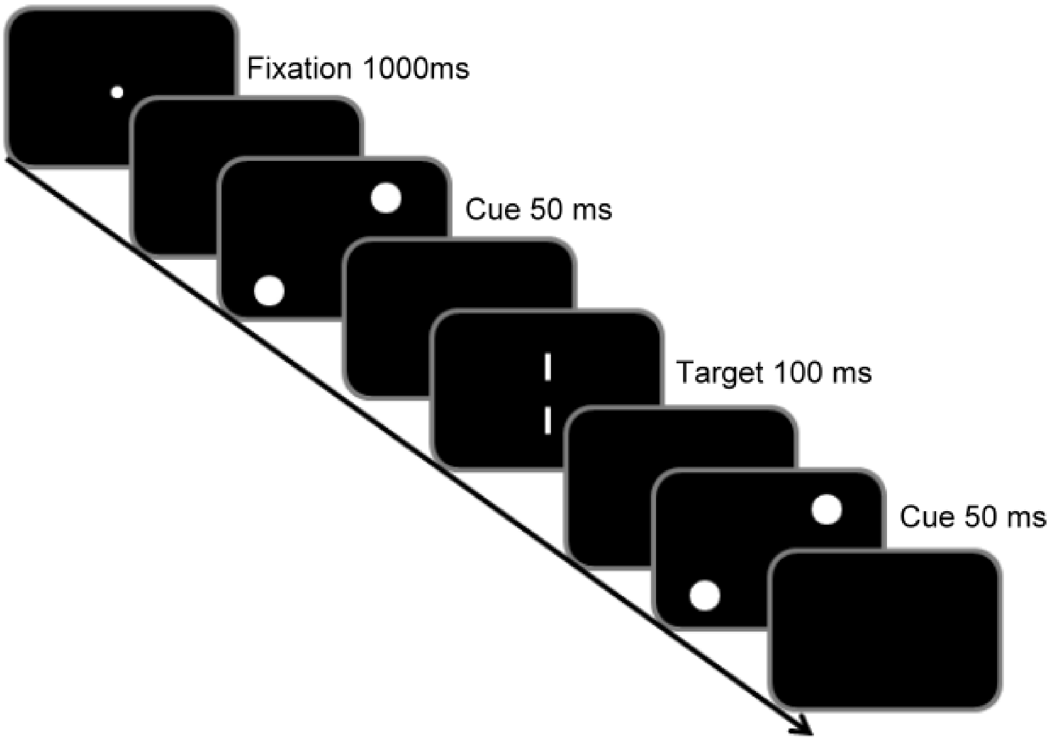


Figure 2.4 Trial events in Experiment 2.

Results

The results of Experiment 2 are shown in Figure 2.5. The mean of PSEs was 0.045°, as defined by the intersection of the cumulative Gaussian curve with the line marked P = 0.5. The PSEs are the mean of individual PSEs. The coefficient of determination yielded by the pooled data in Experiment 2 was 0.99. The effect size of the spatial distortion observed in Experiment 2 was close to the sum of the repulsion and attraction effects in Experiment 1 (-0.054° + 0.096° = 0.042°). I calculated the sum of the repulsion and attraction effects for each observer in Experiment 1 and compared them with those in Experiment 2. There was no significant difference between them (unpaired *t*-tests: *t*(22) = 0.16, *p* = .43). The results suggested that the perceived location of the target was influenced independently by both the preceding and following cues.

However, it was possible that the simple sum observed in Experiment 2 caused by each phenomenon was processed serially in one single localization mechanism without weighting. Therefore, Experiment 3 examined whether the repulsion and attraction effects were the result of different mechanisms.

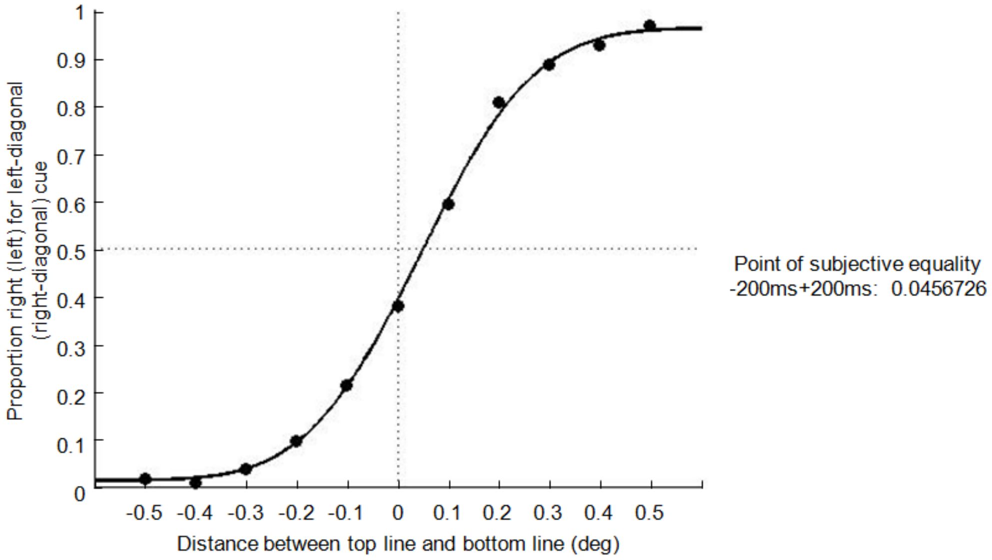


Figure 2.5 The results of Experiment 2. The mean of PSEs in Experiment 2 was 0.045°. Pooled data provided the data points and fitting curves.

Experiment 3

The results of Experiments 1 and 2 showed that the repulsion and attraction effects did not interact; rather, they simply added to each other. In Experiment 3, I presented the cue and target simultaneously in one frame, and presented the cue again before or after the target (Figure 2.6). Observers' attention would be attracted to the cues' positions in the target frame in either the cue-target or target-cue conditions. If the attraction was induced by the dynamic attentional shift from the target to the cue, the attraction effect in the target-cue condition should be attenuated in Experiment 3, because the cue would already have been processed in the target frame. The visual cues appearing in the cue frame would not induce the dynamic shift of visual attention from the target to the cue. This would attenuate the dynamic shift of attention, consequently diminishing the attraction effect. On the other hand, the repulsion effect should not be affected, as it was induced by the cost of enhanced performance of spatial perception at the cues' locations in the cue frame.

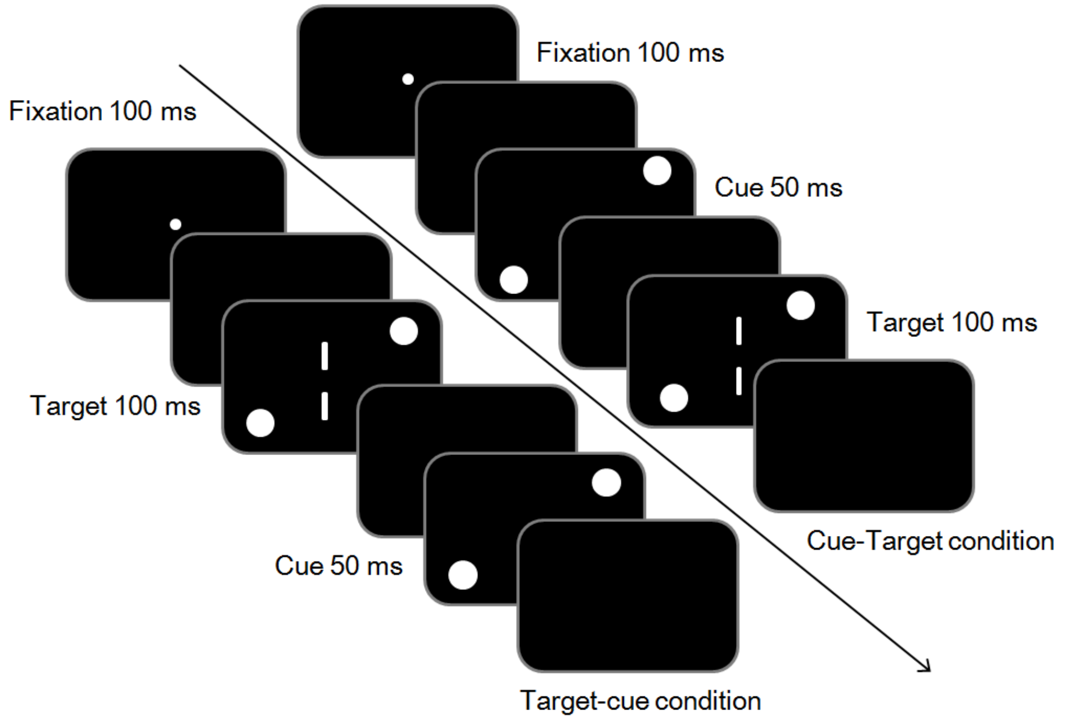


Figure 2.6 Trial events in Experiment 3.

Method

Observers

Participants in this experiment were eleven newly recruited paid volunteers. All observers had normal or corrected-to-normal visual acuity, and were naïve to the purpose of this study.

Stimuli and procedure

The top line of targets appeared in one of three locations: directly above (0°) or to the left (-0.3°) or right (+0.3°) of the bottom line. The cues appeared at the same time and for the same duration as the target lines in half of the trials (double-cue condition). In the other half, the cues appeared only once, either before or after the targets (single-cue conditions). Visual stimuli could be presented in either cue-target or target-cue order (temporal). Observers were instructed to perform the forced-choice task to judge whether the top line was located to the left or right of the bottom line. The experiment utilized a 2 × 2 within-subject design (cue-target vs. target-cue × single-cue vs. double-cue). The cue-target and target-cue SOAs were 200 ms in all conditions. Observers performed a forced-choice task to judge the perceived location of the top line. Each observer completed 10 practice trials and 240 test trials.

Results

Figure 2.7 shows the results of Experiment 3. I calculated the averaged "bias away from the cue" to estimate the magnitude of spatial distortion. The bias was computed as the mean of the proportions of "right" responses for the left diagonal cue (top-left/bottom-right) and "left" responses for the right diagonal cue (top-right/bottom-left). A positive value indicated that the perceived location of the target was away from the cue (attentional repulsion effect), and a negative value implied that the perceived location of the target was shifted toward the cue (attentional attraction effect).

The two-way ANOVA revealed significant main effects of temporal sequence (cue-target or target-cue) and cue manipulation (single- or double-cue) ($F(1,10) = 84.38$, $p < .001$; $F(1,10) = 5.78$, $p < .05$, respectively); the interaction between temporal sequence and cue manipulation was also significant ($F(1,10) = 17.57$, $p < .005$). The post-hoc tests showed that differences in spatial distortion between the single- and double-cue conditions were not significant when the cue preceded the target ($t(10) = 0.86$, $p = .41$, with Bonferroni's correction). However, when the cue followed the target frame, the magnitude of spatial distortion in the single-cue condition was significantly larger than in the double-cue condition ($t(10) = 3.18$, $p < .05$, with Bonferroni's correction). In addition, the magnitude of the attraction effect in the double-cue condition was not different from zero ($t(10) = 1.10$, $p = .29$). Thus, the simultaneous cue presented with the target eliminated the attraction effect but had no influence on the repulsion effect.

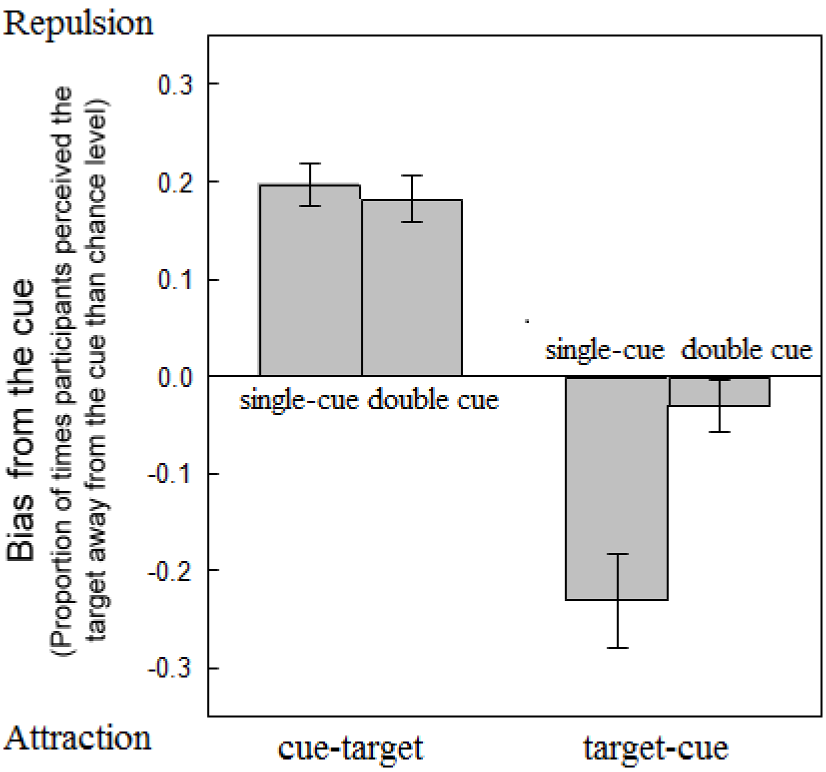


Figure 2.7 Results of Experiment 3. The positive values on the vertical axis indicate that the perceived location of the target was away from the cue (attentional repulsion effect), and negative values indicate that the perceived location of the target was shifted toward the cue (attentional attraction effect). Error bars represent the standard error of the mean.

Discussion

In the double-cue condition with the target-cue temporal sequence, visual attention was attracted to both the cue and target locations at the beginning of each trial. Thereby, observers did not have to shift attention to the cue's location in the next frame. However, in the single-cue condition with same temporal sequence, observers needed to shift attention to the cue's location in the following frame because there were no cues in the first frame; hence, the attraction effect would occur due to the dynamic shift of attention from the target to the cue in the single-cue condition (Figure 2.8). However, the repulsion effect was still observed in the double-cue condition with the cue-target temporal sequence. This could be because the brief cue attracted attention at the beginning of each trial, and the cue was still presented in the next frame. Thus, the cost of enhanced spatial perception was incurred, and persisted in the double-cue condition.

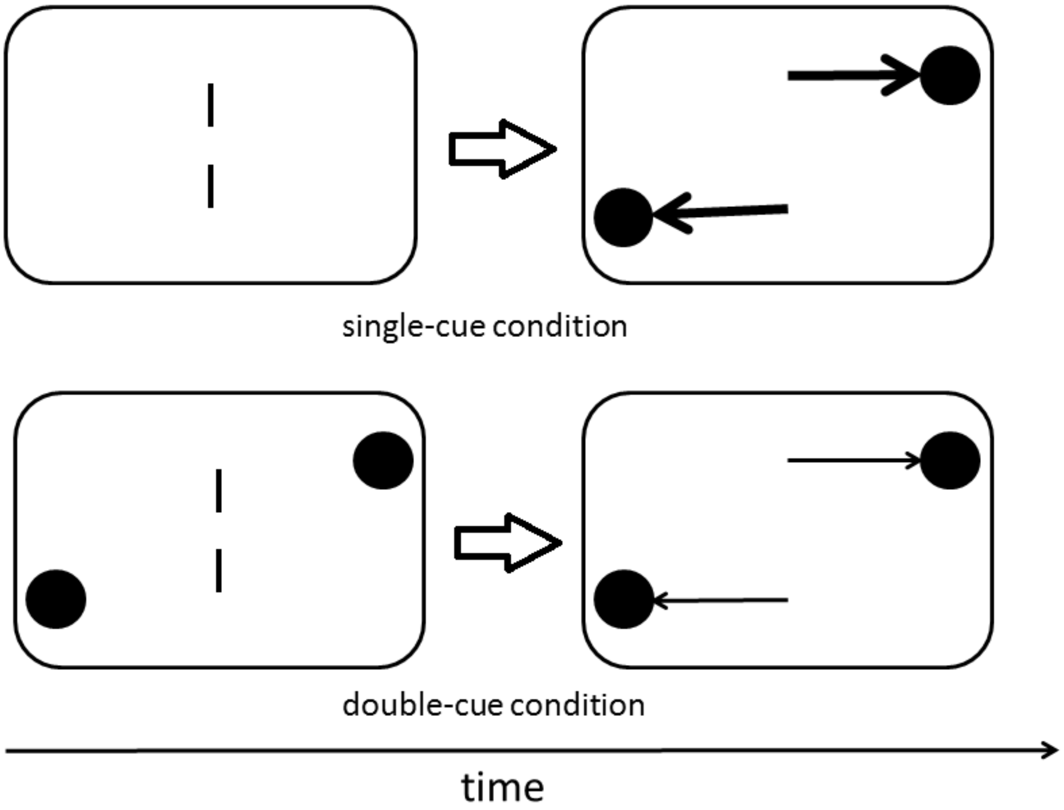


Figure 2.8 Stronger attentional shift was observed in the single-cue condition than the double-cue condition under the target-cue temporal sequence in Experiment 3. Thick and thin arrows represent strong and weak magnitudes of attentional shifts, respectively. Shifts of attention were attenuated in the double-cue condition because attention had already been attracted to the cues' location in the first presented frame.

General discussion

Summary

The results of Experiment 1 yielded effect sizes of the repulsion and attraction effects of -0.054° and 0.096°, respectively. The results of Experiment 2 showed that the effect size of the attraction effect was close to the simple summation of the repulsion and attraction effects in Experiment 1, when the cue was presented both before and after the target. The results of Experiment 3 indicated that the simultaneous cue at the timing of the target diminished the attraction effect, but had no influence on the repulsion effect. This suggests that the underlying mechanism of the repulsion effect is different from that of the attraction effect. The cost of enhanced spatial perception and the dynamic shift of attention independently distort spatial metrics.

The position-coding unit hypothesis for the repulsion effect

Comparing the double-cue to the single-cue condition with the cue-target temporal sequence in Experiment 3, the magnitude of attentional shift would be larger in the single-cue condition, because only the target attracted attention at the moment of target presentation. If the repulsion effect resulted from the dynamic attentional shift, this attenuated attentional shift should also diminish the repulsion effect. This is not consistent with the results of Experiment 3. However, the cost of enhanced spatial perception at the cue’s location should be the same in both the single- and double-cue conditions under the cue-target temporal order, because the cue was presented at the beginning of each trial. Thus, the results of the experiments in Chapter 2 supported the position-coding hypothesis, which proposes that the repulsion effect is the cost of a general mechanism that operates to enhance perception at attended locations (Suzuki & Cavanagh, 1997). The perceived location is represented by the centroid of the distribution of the position-coding units. Attention is directed to the location where peripheral cues are presented. This shift of attentional focus would cause the centroid of distribution to shift in the opposite direction, against the peripheral cue. Suzuki and Cavanaugh (1997) proposed that this mislocalization might be due to surround suppression with receptive field recruitment, or receptive field shrinkage toward the cue’s location. Receptive field shrinkage predicts that the perceived location is always repelled from the cue. Conversely, the hypothesis of surround suppression with receptive field recruitment also predicts the attraction effect when the cue-target distance is closer. However, the repulsion effect did not change to the attraction effect when the cue-target distance was less than 20-30 min (visual angle). Recent research had indicated that visual attention shifts to the center of visual cues in the repulsion effect (Kosovicheva *et al.*, 2010), supporting the hypothesis of receptive field shrinkage.

Possible mechanism for the attraction effect

Ono and Watanabe (2011) proposed that the attraction effect resulting from the dynamic attentional shift from the target to the cue shifted beyond the cues’ actual location, and that the repulsion effect resulted from the attentional shift from the cue to the target. The results in Chapter 2 support the assertion that the attraction effect results from the dynamic shift of attention, but the repulsion effect was not induced by this shift.

A compression of visual space toward a saccade target has been observed in previous research using brief flashed stimuli (Ross *et al.*, 1997). Other studies found that covert shifts of attention can also cause compression of visual space in the direction of attentional shift (Hamker *et al.*, 2008; Zirnsak *et al.*, 2010) because convert shifts of attention can be taken as the motor plan to saccade eye movements. The compression of visual space is one of the possible mechanisms underlying the attraction effect, if this effect can retrospectively influence spatial representations. In other words, when observers shift attention from targets to cues, visual space is compressed toward the cues, consequently inducing the attraction effect. However, there is no evidence that the compression of visual space could be retrospective.

Apparent motion could be an alternative explanation for the attraction effect. Sequential presentation of static objects in different positions could induce apparent motion. In experiments containing following cues, motion signals were directed toward the peripheral cues. According to the motion-biasing model, perceived locations would be biased in the direction of motion to counter neural processing delays (Eagleman & Sejnowski, 2007). Therefore, the perceived target locations are shifted toward the cues. The diminished attraction effect in the double-cue condition under the target-cue temporal order in Experiment 3 could be considered evidence that the quality of apparent motion was impaired by the simultaneous cues. Therefore, it is possible that the different mechanisms that induce the repulsion and attraction effects are static attention focused at the cue's location and apparent motion. However, Ono and Watanabe's studies showed that the attraction effect was observed only when observers paid attention to specified cues if the left and right diagonal cues were presented simultaneously. In this case, even if the sequential presentation of static stimuli could produce motion signals, the direction of apparent motion is not restricted toward the cues. Therefore, the attraction effect cannot be explained by apparent motion. This implies that the location information for spatial representation of the attended cue is required to induce the attraction effect.

Finally, previous research has shown that the perceived locations of visual objects are shifted toward a landmark in the visual field (Hubbard & Ruppel, 2000; Sheth & Shimojo, 2001). The effect of a landmark results from the short-term memory trace bias that could be explained by attention (Eagleman & Sejnowski, 2007; Kerzel, 2002; Yamada *et al.*, 2011).

The location of the visual object is preserved in visual short-term memory. A landmark that attracts attention, biases the location information of objects toward it in spatial working memory. Thus, objects' perceived locations are attracted toward the landmark. The following cue might be used as a landmark to modulate the target's location information by attracting attention toward it. Therefore, the target was mislocalized in the direction of attentional shift in spatial working memory. When the attentional shift is stronger, the spatial distortion is larger.

Possible neural mechanisms for the repulsion and attraction effects

One recent study found that the repulsion effect could still be observed even when visual stimuli were only presented to one eye (DiGiacomo & Pratt, 2012). However, they found that there was no repulsion effect when the cue was presented only to one eye and the target was presented only to the opposite eye. This suggested that the repulsion effect was produced before binocular vision dominated. Therefore, the repulsion effect might result from modulating the receptive field in monocular cells in early visual processing prior to (or in) V1.

In contrast, other research has indicated that visual fields in the primate middle temporal area dynamically shift in the direction of the attentional shift, increasing the selectivity of visual representations within or across the visual area (Womelsdorf *et al*., 2006). This might be related to how attention shifts between stimuli, as well, as working memory represents objects of attention, and the prefrontal cortex (PFC) and posterior parietal cortex (PPC) are relative to spatial working memory (Knudsen, 2007). The PFC and PPC are also possible brain areas responsible for integrating information on attended objects to form spatial representations in spatial working memory.

Chapter 3: Spatial distortion in 3D space

Objective

Depth perception is important for humans because our surrounding environment exists in three dimensions. It is essential to estimate the distance between objects and our own selves. It is also necessary to guide action to interact with the surrounding environment, for example, grasping a pen on a desk. As depth perception relies on not only retinal images but also binocular disparity and pictorial information, it is possible that stimuli including depth information might bring additional effects when spatial distortion is induced in visual space. Given that the localizing mechanism is closely related to spatial attention, the attentional repulsion effect could also be observed in action. More specifically, the pointing responses of a computer-mouse localization task or guided-limb localization tasks would be biased away from the preceding visual cues (Pratt & Turk-Browne, 2003). If depth information of visual stimuli does cause additional effects of spatial distortion in visual space, this might also be reflected in the performance of goal-oriented action. For this reason, it is also important to explore how depth information affects the allocation of spatial attention and attentional shift.

Previous research on spatial attention in 3D space has focused on attentional shift in depth. Attention can be directed toward specific depths. This has been demonstrated by studies in which the reaction time to a target was faster when the target was presented at a cued depth than an uncued depth (Downing & Pinker 1985; Gawryszewski *et al.*, 1987; Han *et al.*, 2005; Previc, 1998). For example, Atchley and colleagues (1997) demonstrated the cost in reaction time of shifting attention from a location in one depth plane to a location in another depth plane by presenting a binocular stereopsis display. In their study, observers were cued to one of four locations at which a target might appear. Half of the locations were in the near-depth plane, and the others were in the far-depth plane. The results showed that reaction time was longer for switching attention between different depth planes than within a single depth plane. This indicates that attention can be shifted in 3D space. In addition, the spatial cueing effect in depth can also be observed in experiments that used pictorial information as depth cues. For example, Han and colleagues (2005) used 2D occlusion cues to provide depth information and found that the spatial cueing effect was stronger when the cues and targets were presented at different depths than when presented at the same depth.

However, while attentional shift in 3D space has been well examined, the cost of attentional shift (i.e., spatial distortion) in 3D space has not yet been fully studied. In Chapter 3, I applied the attentional repulsion and attraction experimental paradigm to investigate this issue. More specifically, I examined the magnitude of spatial distortion while changing the depth planes of peripheral cues in the experimental paradigm of attentional repulsion and attraction. If a shift of attention toward a peripheral cue distorted the spatial metrics around the target irrespective of the depth plane of the cue, this would indicate that the repulsion and/or attraction effects are mainly dependent on the shift of attention on the retinal coordinate. On the other hand, if the magnitude of the repulsion and attraction effects depends on the depth plane in which the cues and fixation are presented, it would suggest that the attentional shift in 3D space would be reflected in the repulsion and attraction effects.

Experiment 4

In Experiment 4, I aimed to examine whether peripheral cues presented in different depth planes would influence the repulsion and attraction effects by using a 3D stimulus presentation setup. Peripheral cues were presented in either the nearer, same, or farther depth plane from the target while keeping the fixation at the same depth as the target. In addition, I generated stimuli with (Experiment 4A) and without binocular disparity (Experiment 4B), leaving the retinal images (i.e., pictorial cues) unchanged.

Method

Observers

Seventeen paid observers were newly recruited to participate in Experiment 4A, and thirteen paid observers participated in 4B. All of the observers had normal or corrected-to-normal visual acuity and were naïve to the purpose of this study.

Apparatus and stimuli

Visual stimuli were presented by a 3D projector (Sight3D, Solidyray) on a large screen (172 cm in width and 130 cm in height) at a viewing distance of 200 cm. Observers viewed the visual stimuli through a time-sequential 3D shutter (Nvidia 3D Vision 2). The refresh rate of the display was 120 Hz (i.e., 60 Hz per eye). The visual stimuli were presented by the Psychtoolbox extensions for the MATLAB operating environment using OpenGL 3D graphics (Brainard, 1997; Pelli, 1997). In Experiment 4A, the images for the left and right eyes were generated with binocular disparity. The inter-pupil distance was fixed at 6.5 cm. In Experiment 4B, the inter-pupil distance was set to zero; hence, the images for the left and right eye were identical. Visual stimuli appeared in white (3.59 and 55.67 cd/m^2 for the target and the cue, respectively) against a black (0.91 cd/m^2) background that contained white (4.42 cd/m^2) virtual grids on the left and right walls, floor, and ceiling. The horizontal distance from the white center fixation point (2 cm in diameter) to the walls was 40 cm. The vertical distance from the center fixation point to the floor and ceiling was 40 cm. The interval between each grid was 20 cm. Grids were drawn in the range of 160 cm in front and 160 cm behind the fixation in depth. The target stimuli were two vertical lines 5.9 cm above and below the fixation point. Each line was 5 cm in length and 0.5 cm in width. The lower line appeared just below the location of the fixation point. The top line of the targets appeared in one of three locations: directly above, to the left (-1 cm), or to the right (1 cm) of the bottom line. The top line was equally likely to appear in any of the three locations. The fixation point and the target stimuli were presented at a distance of 200 cm from the observers (i.e., depth of the screen). The cue spheres were 10 cm in diameter in physical size, and were displaced 20 cm in the vertical and horizontal directions from the fixation point. The depth of the cue was either 120 cm (near-depth-cue condition), 200 cm (same-depth-cue condition), or 280 cm (far-depth-cue condition). As the visual stimuli were generated by using a 3D virtual-reality technique, the retinal size and the eccentricity of the cue depended on the depth of the cue (Figure 3.1). Viewing angle and the horizontal/vertical eccentricity of the cue sphere were 2.15° and 4° in the far-depth-cue condition, 2.86° and 5.72° in the same-depth-cue condition, and 4.58° and 8.86° in the near-depth-cue condition, respectively (Figure 3.1).

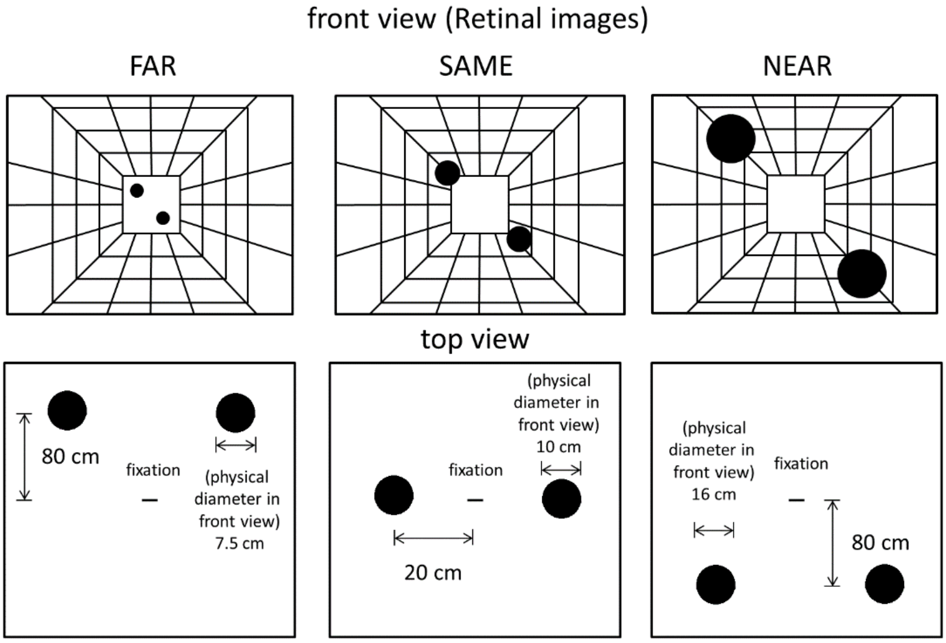


Figure 3.1 Visual cues presented in the near-, same-, and far-depth planes in Experiment 4.

Procedures

Observers initiated each trial by pressing the spacebar on the keyboard. The fixation point appeared for 1000 ms, and observers were instructed to focus on it throughout the experiment. After a 100-ms blank interval, visual stimuli were presented in either the cue-target or target-cue temporal sequence. The cue-target and target-cue SOAs were 200 ms in all conditions. Observers were instructed to perform a forced-choice task to judge whether the top line was located to the left or right of the bottom line. The experiment utilized a 2 × 3 within-subject design (cue-target vs. target-cue × near-depth-cue vs. same-depth-cue vs. far-depth-cue). Each observer completed 10 practice trials and 360 test trials (2 cue configuration × 3 cue depth × 3 possible target location × 2 temporal sequence × 10 repetitions) (Figure 3.2).

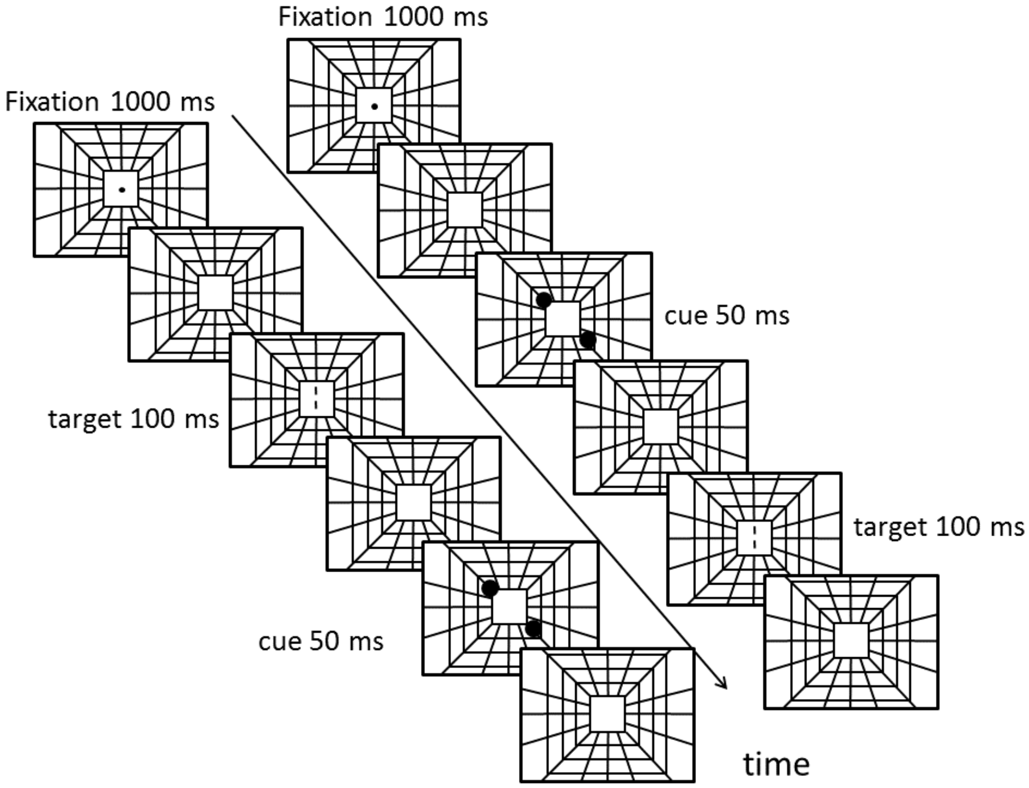


Figure 3.2 Trial events in Experiment 4.

Results

The top and bottom panels of Figure 3.3 show the results of Experiments 4A and 4B, respectively. I calculated the average “bias away from the cue” to estimate the magnitude of spatial distortion, as in Experiment 3 in Chapter 2.

Experiment 4A

The two-way ANOVA revealed a significant main effect of temporal sequence ($F(1,16) = 101.503$, $p < .001$). The main effect of the cue’s depth plane was not significant ($F(2,32) = 0.552$, $p = 0.581$). However, the interaction between the temporal sequence and the cue’s depth plane was significant ($F(2,32) = 10.23$, $p <.001$). The effects of the cue’s depth plane were significant under both the cue-target and target-cue temporal sequences ($F(2,64) =$

4.039; *F*(2,64) = 7.101, respectively, both *ps* < .05). Multiple comparisons (Ryan's method; Ryan, 1960) showed that the magnitude of the repulsion effect was larger in the far-depth-cue condition than the near-depth-cue condition under the cue-target temporal sequence, *ps* < .05. Under the target-cue temporal sequence, the attraction effect was larger in the far-depth-cue condition than the same- and near-depth-cue conditions, *ps* < .05, but there was no significant difference between the same- and near-depth-cue conditions.

Experiment 4B

In Experiment 4B, the main effect of temporal sequence was significant (*F*(1,12) = 61.713, *p* < .001). The main effect of the depth plane of the cue was not significant (*F*(2,24) = 1.917, *p* = .169), but the interaction between the temporal sequence and the depth plane was significant (*F*(2,24) = 5.636, *p* < .01). The effect of the cue's depth plane was significant only under the target-cue temporal order (*F*(2,48) = 6.249, *p* < .005). Multiple comparisons (Ryan's method; Ryan, 1960) showed that the magnitude of the attraction effect was larger in the far-depth-cue condition than the same- and near-depth-cue conditions under the target-cue temporal sequence, *ps* < .05.

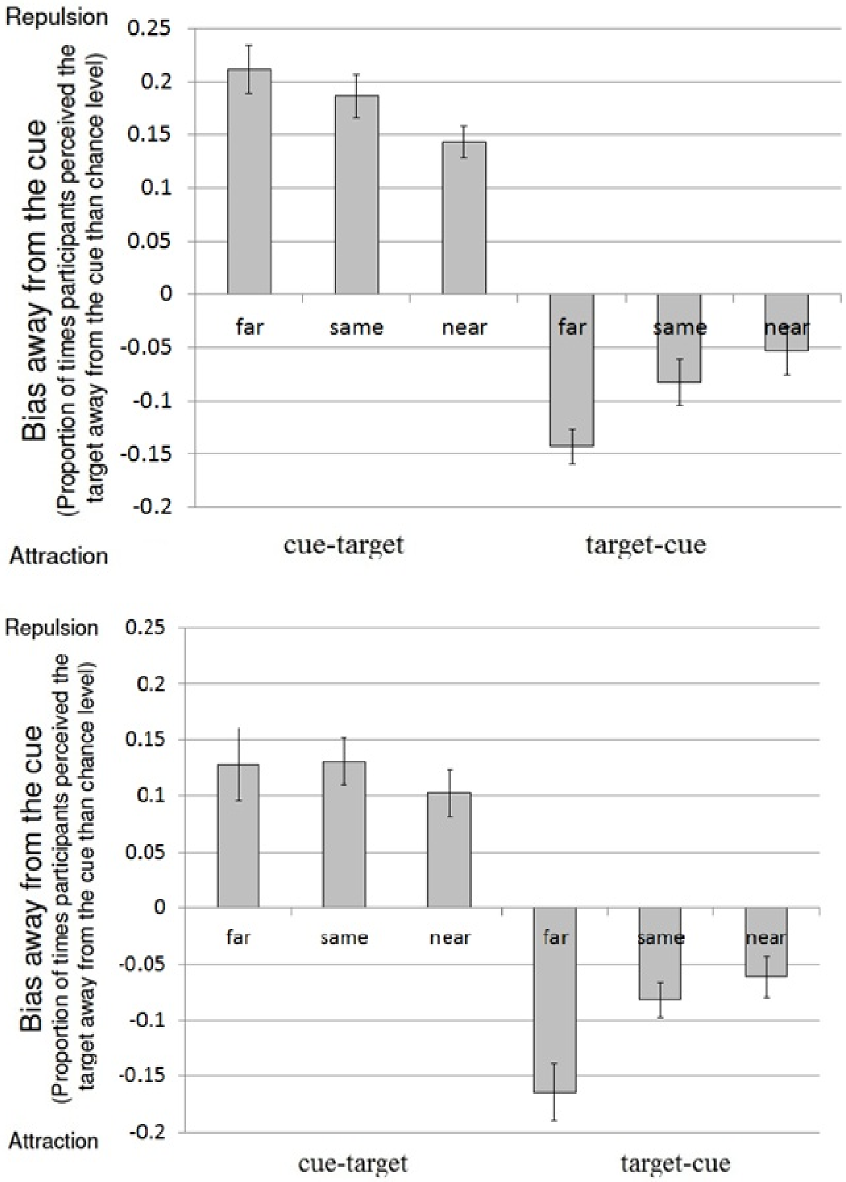


Figure 3.3 The results of Experiments 4A (top) and 4B (bottom). The positive and negative values on the vertical axis indicate the magnitude of the repulsion and attraction effects, respectively. Error bars represent within-observers SEMs (Loftus & Masson, 1994; Cousineau, 2005) for each presentation.

Discussion

Depth modulation of the repulsion effect

The results of Experiment 4 showed that the repulsion effect was larger when the cue was presented in the far depth plane if the stimuli were generated with binocular disparity. However, the repulsion effect was the same in all depth conditions when the stimuli were generated without binocular disparity. This suggests that binocular disparity played an important role in the depth modulation of the repulsion effect. Binocular disparity provides important depth information for visual objects. Previous research reported that the spatial cuing effect in depth increased with increased binocular disparity (He & Nakayama, 1995). Therefore, it is possible that the observed depth modulation of the repulsion effect resulted from the shift of attention in depth defined by binocular disparity. In addition, the depth modulation of the repulsion effect was likely not induced by the pictorial cues (i.e., the cue's size or eccentricity of the retinal images), because the effect of the cues' depth plane was not observed in Experiment 4B, in which the pictorial cues' changes depended on the depth of the cue.

The results in Chapter 2 suggest that the repulsion effect reflects the cost of enhanced spatial perception at the peripheral cues' locations. In Experiment 4, observers were asked to focus on the fixation point. It is possible that the enhanced repulsion effect reflected a larger cost of shifting attention to enhance the spatial perception of the cue's location behind the fixation point, leading to a larger repulsion effect. I manipulated the depth of the fixation point in Experiment 5 to further investigate whether the shift of attention from the fixation point to the farther cues was responsible for this depth modulation.

Is there depth modulation on the attraction effect?

The attraction effect was larger in the far-depth-cue conditions in both Experiments 4A and 4B. This implies that the enhanced magnitude of the attraction effect might result from the changes of pictorial information depending on the depth plane of the cue. In Experiment 4, both the size and eccentricity of the cues changed when the cues were presented in different depth planes. Previous research indicated that the attraction effect was induced when the cue's location information modulated the perceived position of the target (Yamada *et al.*, 2011). Hence, it is possible that the observed depth modulation resulted from the changes in eccentricity that depended on the depth plane of the cue. In Experiment 4, when the cues were presented in the far-depth-cue condition, the distance of retinal coordinates between the cues and the target was less than in the same- and near-depth-cue conditions. Therefore, it is possible that the attraction effect would be stronger if the target-cue distance was smaller, even if no depth information was available. I conducted Experiment 6 to further investigate this possibility and examine the effect of target-cue distance on the attraction effect.

Experiment 5

In Experiment 4, a larger repulsion effect was observed when preceding cues were presented in the depth plane farther away with respect to the target and the fixation. In Experiment 5, I manipulated the depth plane of the fixation point as between-subject factors while keeping the parameters of the cues and targets the same as in Experiment 4A. In Experiment 5, the fixation point was presented (1) before the target in Experiment 5A, (2) in the same depth plane as the target in Experiment 5B, and (3) behind the target in Experiment 5C to test the hypothesis that the repulsion effect would be enhanced when the cues were presented in the depth plane behind the fixation. Therefore, I expected that (1) a larger repulsion effect would be observed when the cues were presented in the same depth plane as or behind the target in Experiment 5A, (2) a larger repulsion effect would be observed only when the cues were presented behind the target in Experiment 5B, or (3) no depth modulation would be observed wherever the cues were presented.

Method

Observers

Fifteen paid observers were newly recruited to participate in Experiments 5A, 5B, and 5C. All observers had normal or corrected-to-normal visual acuity and were naïve to the purpose of this study.

Apparatus, stimuli, and procedures

The apparatus and stimuli were identical to the cue-target condition in Experiment 4A except that the depth plane of the fixation was presented at a viewing distance from the observers of (1) 120 cm in Experiment 5A (near-depth-fixation condition), (2) 200 cm in Experiment 5B (same-depth-fixation condition), and (3) 280 cm in Experiment 5C (far-depth-fixation). In Experiment 5, the experimental conditions of far-, same-, and near-depth-cue were applied to the targets.

Each observer completed 10 practice trials and 180 test trials (2 cue configurations × 3 cue depths × 3 possible target locations × 10 repetitions) in each of the experiments.

Results and discussion

The results of Experiment 5 are shown in Figure 3.4. The two-way (depth plane of fixation and cue as between-subjects and within-subjects factors, respectively) mixed-design ANOVA revealed that the main effect of the cue's depth plane and the interaction between cue depth plane and fixation were significant ($F(2,84) = 5.262$, $p < .01$; $F(4,84) = 5.664$, $p < .005$).

There was an effect of the depth plane of the fixation on the same-depth-cue conditions ($F(2,126) = 4.098$, $p < .01$). Multiple comparisons (Ryan's method; Ryan, 1960) showed that the magnitude of the repulsion effect was larger in the near-depth-fixation (Experiment 5A) than in the same- (Experiment 5B) and far-depth-fixations (Experiment 5C) when the cues were presented in the same-depth-cue conditions, *ps* < .05.

There were also significant effects of the cues' depth planes on the near- (Experiment 5A) and same-depth-fixation (Experiment 5B) conditions (*F*(2,84) = 8.861, *p* < .001; *F*(2,84) = 6.947, *p* < .005). In the same-depth-fixation condition (Experiment 5B), multiple comparisons (Ryan's method; Ryan, 1960) showed that the magnitude of the repulsion effect was larger in the far-depth-cue condition than the same- and near-depth-cue conditions, *ps* < .05. In the near-depth-fixation (Experiment 5A), the magnitude of the repulsion effect was larger in the far- and same-depth-cue conditions than in the near-depth-cue condition, *ps* < .05. In addition, there was no effect of cue plane on the far-depth-fixation condition (Experiment 5C, *F*(2,84) = 0.783, *p* = 0.460).

The results of Experiment 5 indicate that a larger repulsion effect was observed when the cues were presented in the depth plane farther away with respect to the fixation point (Experiment 5A and 5B), while no depth modulation was observed when the fixation point was presented in the same depth plane with the farthest cues.

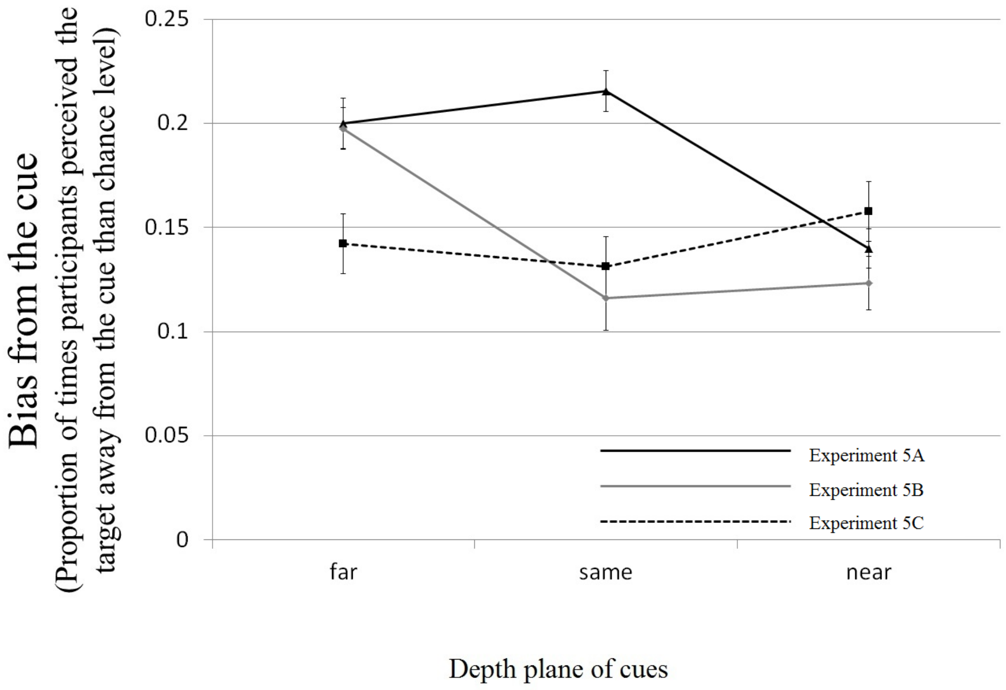


Figure 3.4 The results of Experiment 5. The X-axis represents the experimental conditions of the depth plane of the cues. The Y-axis represents the magnitude of the repulsion effect. Error bars represent within-observers SEMs (Loftus & Masson, 1994; Cousineau, 2005) for each presentation.

Experiment 6

In Experiment 4, larger attraction effects were observed in the far-depth-cue conditions under both binocular and monocular stereopsis viewing. Thus, it is possible that a larger attraction effect resulted from changes in pictorial information on the retinal images depending on the depth plane of the cues. The eccentricity of the cue in the far-depth-cue condition was smaller, resulting in a smaller distance between the target and the cue. In Experiment 6, I manipulated target-cue distance to examine the hypothesis that larger attraction effects in Experiment 4 resulted from closer target-cue distances. Furthermore, the grids used in Experiment 4 to provide depth information were removed, and the physical sizes of the cues were the same in all conditions of target-cue distance. If the larger attraction effect in Experiment 4 resulted from the closer target-cue distance in the far-depth-cue condition, I expected that a larger attraction effect should be observed in experimental conditions with closer target-cue distance in Experiment 6.

Method

Observers

Fifteen paid volunteers were newly recruited to participate in Experiment 6. All observers had normal or corrected-to-normal visual acuity and were naïve to the purpose of this study.

Apparatus, stimuli, and procedures

The apparatus and stimuli were identical to those used in Experiment 3 in Chapter 2 except for the presented positions of the disks (visual cues). In Experiment 6, the distance between the targets and the cues was manipulated. There were eight conditions of target-cue distances: (1) 0.5°, (2) 1°, (3) 2°, (4) 4°, (5) 6°, (6) 8°, (7) 10°, and (8) 12° (Figure 3.5). The temporal sequence of the presented visual stimuli was identical to the target-cue condition in Experiment 4. Observers were asked to perform the forced-choice task. After 10 training trials, each observer completed 480 test trials (2 cue configurations × 3 target positions × 8 target-cue distances × 10 repetitions).

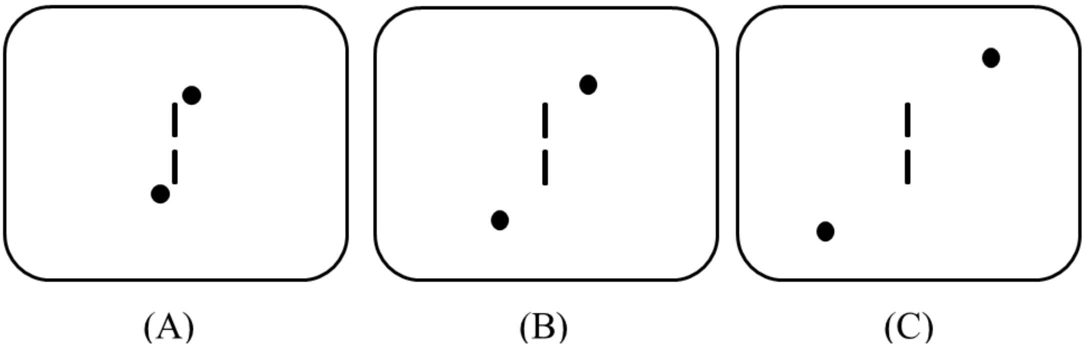


Figure 3.5 Examples of possible cue-target distances in Experiment 6. (A) 0.5° (B) 4° (C) 12°

Results and discussion

Figure 3.6 shows the results of Experiment 6. The magnitude of the attraction effect was estimated by calculating the "bias away from the cue." The vertical axis represents the "bias away from the cue," and the horizontal axis represents the experimental conditions of target-cue distance. The repeated-measures one-way ANOVA revealed a significant main effect of target-cue distance ($F(7,98) = 9.839$, $p < .001$), indicating that the magnitude of the attraction effect tended to be larger when the target-cue distance was smaller. Multiple comparisons (Ryan's method; Ryan, 1960) showed that (1) the magnitude of the attraction effect with a 0.5° target-cue distance was larger than with 4°, 6°, 8°, 10°, and 12° target-cue distances, (2) the magnitude of the attraction effect with a 1° target-cue distance was larger than with 4°, 6°, 8°, 10°, and 12° target-cue distances, and (3) the magnitude of the attraction effect with a 2° target-cue distance was larger than with 10° and 12° target-cue distances. This indicated that the magnitude of the attraction effect was larger when the target-cue distance was smaller.

The results of Experiment 6 indicated that the attraction effect was larger when the target-cue distance was smaller, even when there were no pictorial cues to provide depth information (i.e., the grids in Experiment 4). This supports the hypothesis that the depth modulation of the attraction effect observed in Experiment 4 actually resulted from changes in eccentricity of the cue depending on the depth plane on the retinal coordinate. Therefore, I observed changes in the magnitude of the attraction effect under both monocular and binocular stereopsis viewing in Experiments 4A and 4B.

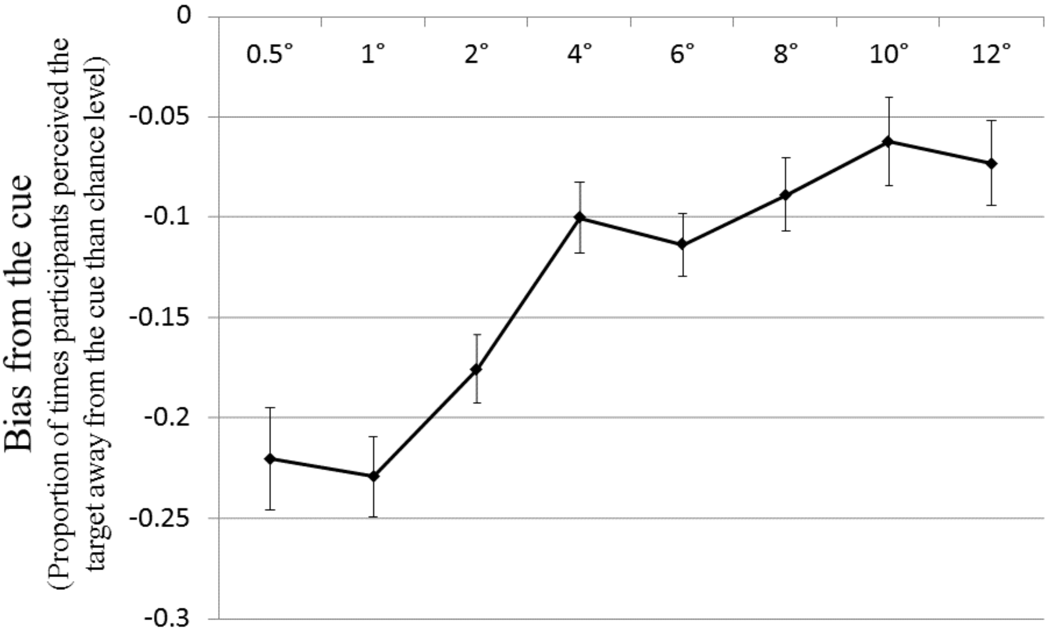


Figure 3.6 The results of Experiment 6. The horizontal axis represents experimental conditions of target-cue distance in degree of visual angles, and the Y-axis represents the magnitude of the repulsion effect. Error bars represent within-observer SEMs (Loftus & Masson, 1994; Cousineau, 2005) for each presentation.

General discussion

Summary

Chapter 3 aimed to examine the effect of attentional shift in depth on spatial distortion by using the attentional repulsion and attraction experimental paradigm. The experiments in

Chapter 3 found that larger spatial distortion was observed when peripheral visual cues were presented in the farther depth plane (Experiment 4). For the repulsion effect, depth modulation resulted when observers shifted attention from the fixation point to cues in the depth plane farther away with respect to the fixation point. This suggested that the depth information for the peripheral stimuli had an additional effect on spatial distortion (Experiment 5). However, a larger attraction effect was observed with smaller target-cue distances even when there was no depth information (i.e., the grids in Experiment 4 and 5) and/or the cues were the same physical size (Experiment 6).

Attentional shift in depth modulates the repulsion effect

In the experiments in this chapter that presented preceding cues, observers were instructed to fixate on the central fixation point, thereby paying attention to it. Then, a brief peripheral cue appeared, and the observer's attention would shift involuntarily toward the cue, which was then followed by the central target. In Experiments 4 and 5, attention shifted from central to peripheral on the 2D retinal coordinate, while in 3D space, attention shifted farther, stayed at the same depth, or shifted nearer. Experiment 5 demonstrated that attentional shift toward a brief cue in a depth plane farther away from the fixation point led to a larger repulsion effect, and that attentional shift from the cues to the target did not influence the magnitude of repulsion effects (Figure 3.7).

One possible explanation is based on three assumptions: a reduction of attentional resources associated with the attentional shift in depth (i.e., attention shifting in depth is more demanding than that within the same depth; Downing & Pinker 1985; Gawryszewski *et al.*, 1987; Han *et al.*, 2005; Previc, 1998; Atchley *et al.*, 1997), an enhanced attentional process for the space closer than the fixation (Chen *et al.*, 2012; Miura *et al.*, 2002; Andersen & Kramer 1993; Gawryszewski *et al.*, 1987), and a residual effect of such enhanced attentional status.

According to Figure 3.7, attentional shifts between the fixation and cues in the same depth plane occurred for the same-depth condition in Experiment 5B, the near-depth condition in Experiment 5A, and the far-depth condition in Experiment 5C. All of the other conditions included attentional shifts between the fixation and cues at different depth planes, which might require more attention, leaving less attention available for later target localization. However, this might be compensated by the enhanced attention status caused by the attentional shift to the space closer than the fixation (Chen *et al.,* 2012; Miura *et al.,* 2002; Andersen & Kramer, 1993; Gawryszewski *et al.,* 1987), perhaps because of a transient need for elevated alertness that enables better and/or faster responses to visual stimuli closer to the observer. This explanation might also be related to the differential functional representations between peripersonal and extrapersonal space (Previc, 1998; Maringellia *et al.*, 2001).

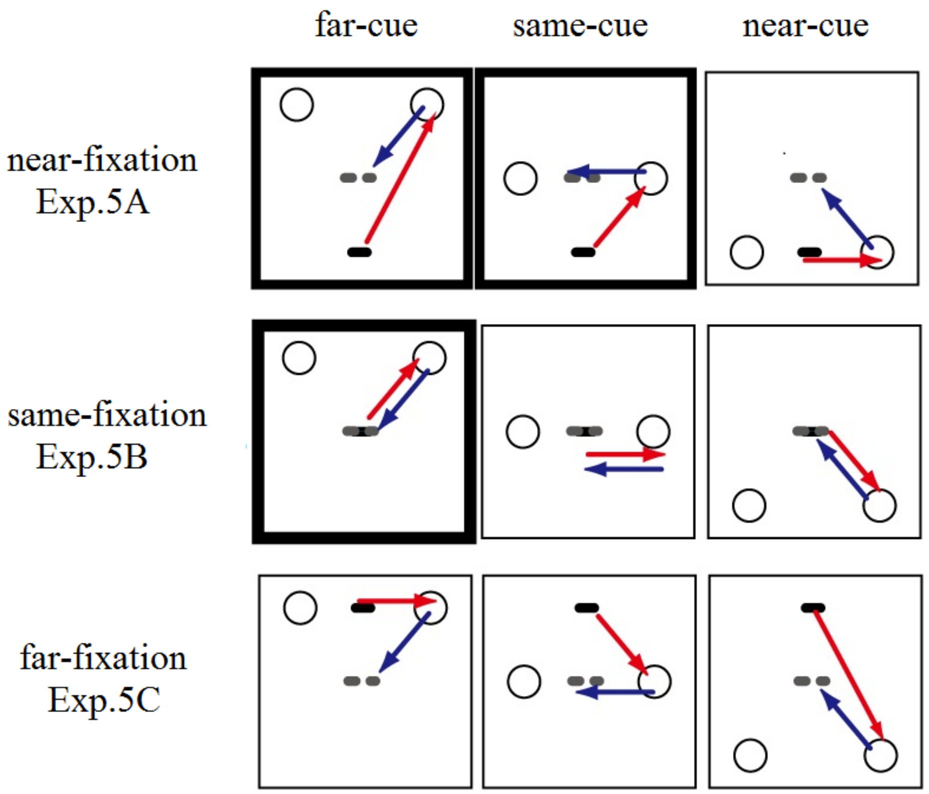


Figure 3.7 Expected shifts of attention in Experiments 5A, 5B, and 5C. Single bars indicate the fixation point. The two circles represent the cues. The two separate lines indicate the target bars. Conditions under which the enhanced repulsion effect was observed are outlined with heavy borders. Note that only the condition with attentional shift from the fixation to the far-depth planes led to the enhanced effect, while attentional shift from the cue to the target did not affect the results.

Magnitude of the attraction effect modulated by the target-cue distance

The results of Experiment 6 suggest that a larger attraction effect would be observed when the target-cue distance was smaller. It is notable that the effect of target-cue distance was observed even when the grids providing depth information were removed. This suggests that the changes in magnitude of the attraction effect observed in Experiment 4 might not result from depth perception. Further, the physical size of the cues in Experiment 6 was the same across all target-cue distance conditions. This also implies that the retinal size of the cues was not responsible for the enhanced attraction effect.

Previous research indicated that the magnitude of the repulsion effect depended on cue-target distance in 2D space (Kosovicheva *et al*., 2010). However, they found that the average magnitude of the repulsion effect peaked at a cue-target distance of about 2°; the repulsion effect attenuated when the cue-target distance was below 2°. This differs from the tendency of the attraction effect observed in Experiment 6. This also supports the idea that the attraction effect is induced by the location information of the brief cues that modulate the location information of the target. In other words, when a visual transient flashed near the location of the visual target, the remembered location of the target might be confounded by that of the visual transient. Moreover, cues with closer target-cue distance might bias the target's perceived location to a larger degree because of the closer spatial relation between the visual target and cue.

In conclusion, the findings in Chapter 3 suggest that dynamic attentional shifts in depth were reflected in the repulsion effect, while the attraction effect was mainly dependent on the location information on the retinal coordinate. As in Chapter 2, the findings in Chapter 3 also suggest that the underlying mechanism of the attraction effect is different from that of the repulsion effect.

Chapter 4: Crossmodal interaction on spatial distortion

Objective

The findings presented in Chapters 2 and 3 suggest that peripheral transients induce mislocalizations of visual objects either by impairing processing of the target's location representation, or because location information of peripheral transients biases the perceived location of visual targets in short-term memory traces. The latter case occurs because the perceptual system integrates information from both the target and the peripheral transient to decide the perceived target's location.

However, the perceptual system receives information through different modalities. The perceptual world is constructed by integrating information from different modalities. Therefore, sensory transients from modalities other than vision could also induce mislocalization of visual objects as they could also induce involuntary attentional shifting. In this chapter, I aimed to explore whether the dynamic attentional capture induced by auditory transients could also induce mislocalization of visual targets.

Continuous sound could modulate representational momentum

Studies have demonstrated that the perceived location or remembered final position of a moving object is consistently biased in the forward direction of motion (i.e., the flash-lag effect and representational momentum). Studies of representational momentum (RM) have demonstrated that the final perceived position of a moving object is mislocalized in the forward direction of motion (Freyd & Finke, 1984; Hubbard & Bharucha, 1988). RM results from the mental representation of an object's motion persisting for a brief period after abrupt visual motion offset (Teramoto *et al*., 2010). Recently, Teramoto and colleagues (2010) found that the magnitude of RM is influenced by a continuous sound accompanying a moving visual object. They showed that RM is enhanced when the sound terminates after the offset of visual motion, but reduced when the sound terminates before visual motion offset.

Auditory transients could modulate motion perception

Several previous studies have also investigated how visual motion perception is modulated by an auditory transient. In the flash-lag effect, the perceived position of a moving object appears to be relatively ahead of a physically aligned flash (e.g., Maus & Nijhawan, 2009; Nijhawan, 1994; Watanabe & Yokoi, 2006, 2007, 2008). This phenomenon is the result when the visual representation of moving objects shifts forward spatially to counteract neural system delays in processing sensory signals. Vroomen and de Gelder (2004) showed that the magnitude of the flash-lag effect is reduced when an auditory transient is presented before or simultaneously with the flash. In addition, Heron and colleagues (2004) demonstrated that the perceived bounce position of a horizontally moving object that changes direction against a virtual vertical surface is perceptually modulated by a sound presented before/after the actual bounce event.

Sound’s spatial information could modulate visual motion perception

Recently, Teramoto *et al.* (2012) showed that when visual apparent motion in conjunction with a sound delivered alternately from two loudspeakers aligned horizontally or vertically, observers reported that the direction of visual apparent motion was consistent with the direction of sound alternation or lay between that direction and the actual direction of the visual apparent motion. This suggests that auditory spatial information could also modulate the perception of visual motion, especially in the peripheral visual field. More specifically, enhanced audiovisual interaction would be observed when the visual target and the auditory transients are presented from close spatial locations, as previous research has indicated that the audiovisual dynamic capture effect is larger when the auditory and visual stimuli are presented in the same hemifield (Meyer *et al.*, 2005; Soto-Faraco *et al*., 2002; Spence, 2007).

In Chapter 4, based on these findings, I presented an auditory transient around the moment of visual motion offset and asked observers to report the perceived location of the visual motion offset in Experiment 7. In Experiment 8, I tested whether auditory spatial information would influence the perceived location of visual motion offset by presenting a lateralized auditory transient to the observers on the same or opposite side as the visual target. I expected to observe a larger auditory effect when the lateralized sound was presented on the same side with the visual target, because same-side sound would induce stronger attentional capture. After affirmative results were obtained in both experiments, I examined whether the auditory effects were caused by distortion in the perceived timing of the offset of the visual moving object (Experiment 9).

Experiment 7

In Experiment 7, I examined the possibility that the timing of a transient auditory signal would affect the perceived offset position of a visual moving object. Such an effect would demonstrate that a continuous auditory stimulus during visual motion is not necessary to alter the perceived visual offset position. I conducted Experiments 7A and 7B. The visual target appeared in the left visual field and moved rightward (Experiment 7A), or in the right visual field and moved leftward (Experiment 7B). Then, the visual target disappeared near the center of the display. An auditory transient was presented around the time of the visual motion offset of the visual target.

Method

Observers

Sixteen paid volunteers were recruited to participate in Experiments 7A and 7B. All observers had normal or corrected-to-normal visual and auditory acuity, and were naïve to the purpose of this study.

Apparatus and stimuli

Apparatus were the same as those used in Chapter 2. The background of the display was divided horizontally into two parts. The upper part was gray (40° × 10.5°, 7.85 cd/m^2) and the lower part was black (40° × 19.5°, 0.03 cd/m^2). A white fixation cross (1° × 1°, 61.27 cd/m^2) was presented at the center of the lower background (Figure 4.1).

The visual stimulus was a black disk (1° in diameter) that appeared at the bottom of the gray background, 15° to the left in Experiment 7A and to the right in Experiment 7B. It moved at a constant speed of 15°/s. The disk disappeared when its center was at the midpoint or randomly jittered from the midpoint by ± 0.3°. The auditory stimulus was a transient auditory signal with a 1000 Hz pure tone without onset or offset intensity ramps, presented via headphones to both ears for 10 ms. The approximate range of sound pressure level was 60–65 dB. The sound was presented 120, 80, or 40 ms before visual motion offset, simultaneously with the visual offset (0 ms), or 40, 80, or 120 ms after visual offset. The sound might also be absent during the trial as a control condition.

Procedure

Observers initiated each trial by pressing the spacebar. The black disk appeared and remained stationary at the initial position for 500 ms. Observers were asked to track the disk while keeping their eyes on the fixation cross. After the initial stationary period, the disk moved at a constant speed of 15°/s and then disappeared around the midpoint of the screen. A mouse cursor then appeared 1° above the fixation 200 ms after the disappearance of the disk. Observers were instructed to move the mouse cursor and click the mouse button at the position where the target disappeared (Figure 4.1).

Observers performed 10 practice trials to familiarize themselves with the position judgment task. Then, they completed 240 test trials (8 sound conditions × 3 visual offset positions × 10 repetitions). Trials were randomized in all conditions.

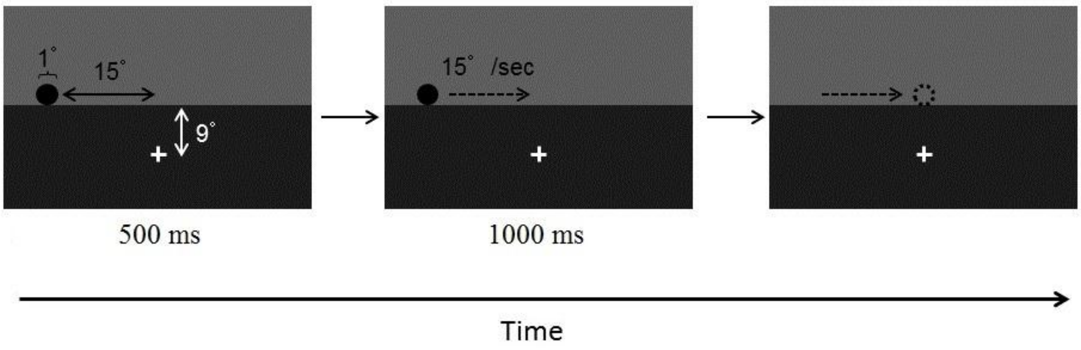


Figure 4.1 Trial events in Experiment 7.

Results

I calculated the average deviation of the perceived visual offset position from the physical visual offset position for each sound condition. Figure 4.2 shows the combined results of Experiments 7A and 7B. The horizontal axis represents different sound conditions. The vertical axis represents the perceived deviation from physical visual offset position. A negative value on the Y-axis means that the perceived visual offset was behind the actual visual offset position.

The two-way mixed-design ANOVA (between-subjects factor: visual field of the start position; within-subjects factor: the timing of the auditory signal) revealed a significant main effect of the timing of the auditory signal ($F(7,210) = 36.261$, $p < .001$), while the main effect of the visual field of the start position was not significant ($F(1,30) = 0.499$, $p = .485$). The interaction between the start position and the sound timing was not significant ($F(7,210) = 0.48$, $p = .849$). Overall, the results suggested that the earlier the auditory signal was presented, the farther backward and away the perceived visual offset position was shifted.

Next, the cell means of the perceived visual offset position were compared against zero to test whether there were significant mislocalizations from the actual position in each condition. In Experiment 7A, the -120, -80, and -40 ms conditions significantly differed from zero ($t(15) = 5.569$, $t(15) = 5.88$, and $t(15) = 5.89$, respectively, all $p < .006$, with Bonferroni correction). In Experiment 7B, the -120 and -80 ms conditions significantly differed from zero ($t(15) = 6.57$ and $t(15) = 5.27$, respectively, all p <. 006, with Bonferroni correction). Thus, it was confirmed that when the brief sound was presented before the physical offset of the visual target, the perceived visual offset location tended to be perceived as behind the actual offset location. In contrast, no significant spatial distortion was found in the 0, 40, 80, and 120 ms conditions, implying that the auditory transient had no effect when presented after or at the same time as the visual motion offset.

In addition, the cell means of each condition were compared to the cell mean of the silent condition. In Experiment 7A, the perceived visual offset positions in the -120, -80, and -40 ms conditions differed from those in the silent condition ($t(15) = 4.46$, $t(15) = 4.23$, and $t(15) = 3.34$, respectively; all $p < .007$, with Bonferroni correction). In Experiment 7B, the perceived visual offset positions in the -120, -80, and -40 ms conditions differed from those in the silent condition ($t(15) = 5.24$, $t(15) = 5.01$, and $t(15) = 3.30$, respectively; all $p < .007$, with *Bonferroni cor*rection). Furthermore, the perceived visual offset position in the silent condition did not differ from conditions in which the auditory transient was presented after the actual visual offset in either Experiment 7A (*t*(15) < 1.11, *p* > .05) or 7B (*t*(15) < 1.05, *p* > .05).

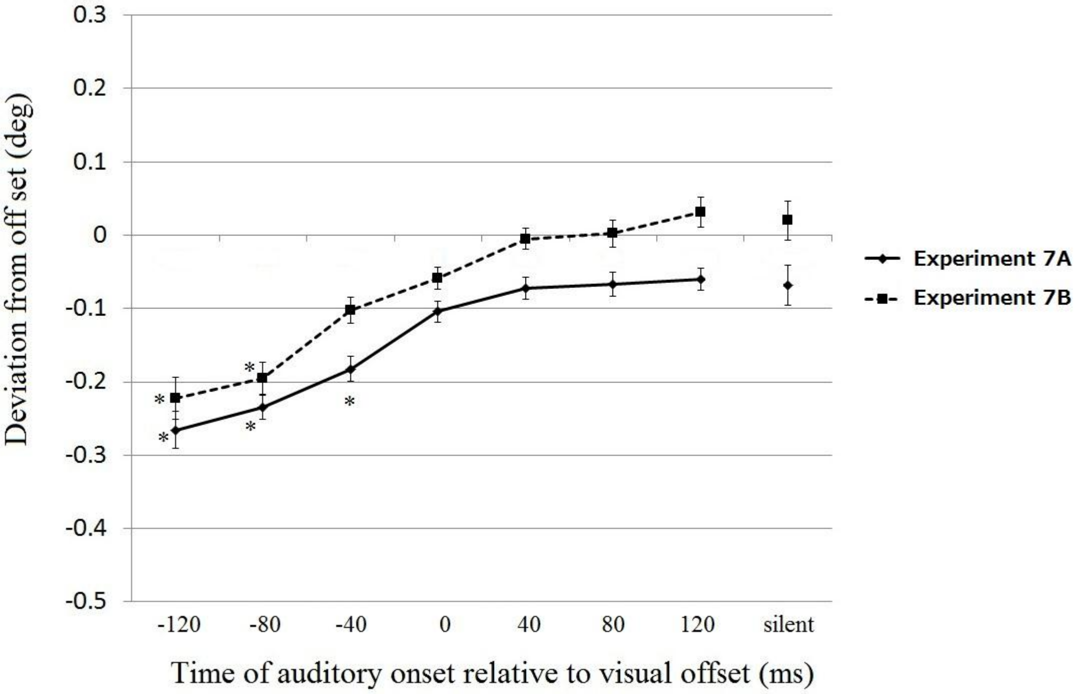


Figure 4.2 The results of Experiments 7A and 7B. The horizontal axis represents the experimental conditions for the timing of the auditory signal. The vertical axis represents the perceived deviation from the physical visual offset position in degrees of visual angle. Error bars represent within-observers SEMs (Loftus & Masson, 1994; Cousineau, 2005) for each presentation. Data points with an asterisk (*) indicate that the perceived positions differed from zero.

Discussion

The results of Experiment 7 showed that the perceived offset location was attracted toward the timing of presentation of the auditory transient. However, the perceived offset location of the visual target in the silent condition did not differ from conditions in which the auditory transient was presented after the physical offset. Therefore, in Experiment 7, displacement of perceived visual motion offset was observed when the auditory transient was presented before the physical visual offset.

The lack of RM in Experiment 7 might be partially attributable to the shorter delay between the target offset and the appearance of the mouse cursor. Kerzel *et al.* (2001) showed that RM was larger with longer delays between the target and the probe used to estimate the perceived location of the visual motion offset. In the experiments in Chapter 4, the delay was 200 ms, while the delay used in previous research was often longer than 500 ms.

Experiment 8

In Experiment 8, I tested whether the spatial contingency between the auditory transients and the visual events would modulate the auditory influence on the perceived offset position of the visual motion. I presented a lateralized auditory transient to either the left or right ear with the same visual stimuli used in Experiment 7. The visual target appeared in the left visual field and moved rightward in Experiment 8A. In Experiment 8B, the visual target appeared in the right visual field and moved leftward. I expected that larger displacement of perceived visual motion offset would be observed when the lateralized auditory transient was presented in the same hemifield in which the visual target originated.

Method

Observers

Fifteen paid volunteers were newly recruited to participate in Experiments 8A and 8B. All observers had normal or corrected-to-normal visual and auditory acuity, and were naïve to the purpose of this study.

Apparatus, stimuli, and procedures

The apparatus and the visual and auditory stimuli in Experiment 8 were identical to those used in Experiment 7 except for the following points: (1) the auditory transients were presented in the left ear in half of the trials and the right ear in the other half of the trials; (2) since I did not find differences between conditions in which the auditory transient was presented after the physical visual offset and the silent condition in Experiment 7, I did not include the silent condition in Experiment 8. Observers completed 10 training trials and 420 test trials (2 sound positions × 3 visual offset positions × 7 timings of the auditory transient × 10 repetitions). Trials of all conditions were randomly ordered such that in each trial the auditory transient could be presented to the same or opposite side as the start position of the visual target.

Results

The top and bottom panels of Figure 4.3 show the results of Experiments 8A and 8B. A three-way mixed-design ANOVA (between-subjects factor: the visual field of the start position; within-subjects factors: the sound contingency and the timing of the auditory transients) was conducted. The sound contingency signifies whether the auditory transient was presented to the same or opposite side as the originating position of the visual target. The main effects of auditory timing ($F(6,168) = 77.48$, $p < .001$) and sound contingency ($F(1,28) = 43.526$, $p < .001$) were both significant. The main effect of start position approached the significant level ($F(1,28) = 4.172$, $p = 0.052$). There were no significant interactions effects (visual field × sound contingency, $F(1,28) = 0.049$; visual field × auditory timing, $F(6,168) = 0.52$; sound contingency × auditory timing, $F(6,168) = 1.540$; visual field × sound contingency × auditory timing, $F(6,168) = 1.010$; all $p > .05$). The results implied that the timing of the auditory transient induced mislocalization of the perceived visual offset position, consistent with the results of Experiment 7. Moreover, when the auditory transient was presented in the same hemifield as the visual target's start position, the audiovisual interaction was enhanced (i.e., more backward mislocalization).

Experiment 8A

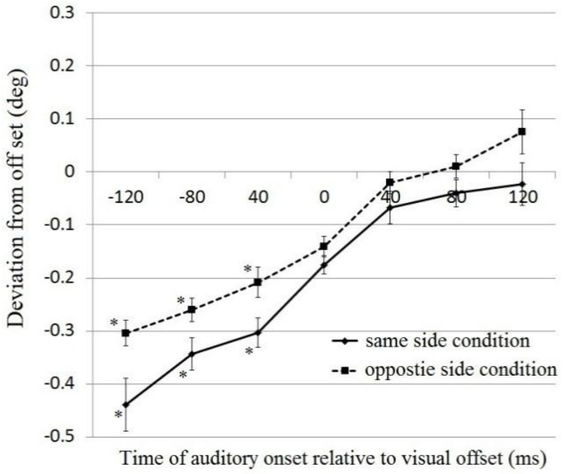
The cell means of the perceived visual offset position in Experiment 8A were compared against zero to test whether there were significant spatial distortions from the actual visual offset position in each condition. When the auditory transient was presented on the same side from which the visual target appeared, the perceived visual offset positions significantly differed from zero in the -120, -80, and -40 ms conditions ($t(14) = 5.02$, $t(14) = 4.25$, and $t(14) = 4.09$, respectively; all $p<.007$, with *Bonferroni correction*). When the auditory signal was presented in the hemifield opposite that from which the target appeared, the perceived visual offset positions significantly differed from zero in the -120, -80, and -40 ms conditions ($t(14) = 4.72$, $t(14) = 4.04$, and $t(14) = 4.02$, respectively; all $p < .007$, with *Bonferroni correction*).

The cell means in Experiment 8A were also compared between the -120, -80, and -40 ms conditions in which significant spatial distortions were observed. When the auditory transient was presented on the same side as the visual target origination, the displacement of perceived visual motion offset was larger ($t(14) = 2.44$ and $t(14) = 2.51$ for the -120 and -80 ms conditions, respectively, both $p<.007$, with *Bonferroni correction*).

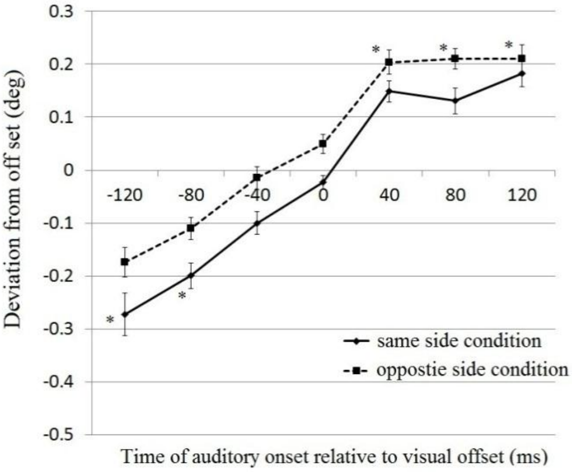
Experiment 8B

The results of Experiment 8B differed from Experiment 8A when comparing the cell means against zero. In Experiment 8B, RM was observed in the 40, 80, and 120 ms conditions when the auditory transient was presented on the side opposite to the visual target origination ($t(14) = 3.45$, $t(14) = 3.27$, and $t(14) = 3.51$ for the 40, 80 and 120 ms conditions, respectively, all $p < .007$, with *Bonferroni correction*). However, there were no significant differences among these three cell means (all $p > .017$, with *Bonferroni correction*). This suggests that when the auditory transient was presented to the opposite side, it attracted the offset position of the visual target, which resulted in larger forward displacement.

Conversely, significant backward displacements were observed in the -120 and -80 ms conditions when the auditory transient was presented to the same side as the visual target ($t(14) = 3.25$ and $t(14) = 3.21$, respectively; all $p < .007$, with *Bonferroni correction)*, but the difference between these conditions was not significant ($t(14) = 1.91$, $p = 0.15$). Nevertheless, significant backward shift was observed when the sound was presented on the same side as the visual target.



Experiment 8A



Experiment 8B

Figure 4.3 The results of Experiments 8A (top) and 8B (bottom). The horizontal and vertical axes represent, respectively, the different sound timings and the perceived deviation from the physical visual offset position in the degrees of visual angle. A negative value on the Y-axis means that the perceived offset position was behind the physical visual offset position. Error bars represent within-observers SEMs (Loftus & Masson, 1994; Cousineau, 2005) for each presentation. Data points with an asterisk (*) indicate that the perceived positions differed from zero.

Discussion

The results in Experiment 8B showed significant forward displacement of the visual motion offset (RM). It is possible that RM was more evident in Experiment 8B. Previous research has shown that the attentional mechanisms in the left hemisphere tend to distribute attention resources within the right visual field, while the attentional mechanisms in the right hemisphere distribute attentional resources across both left and right visual fields. Therefore, there may be a slight bias of spatial attention favoring the right visual field (Mesulam, 1999).

However, similar to the results of Experiment 8A, the auditory transient presented before the physical visual offset exhibited a net effect of RM, and a process induced by the auditory transient decreased RM or even induced backward displacement of the perceived visual offset position in some conditions. This effect was stronger when the auditory transient was presented on the side from which the visual target appeared. The results of Experiment 8 imply that a transient auditory signal closely associated with the visual offset also influences the perceived visual offset from the physical visual offset position.

Experiment 9

The results of Experiment 7 suggest that the perceived offset position of a visual moving object shifts backward when a transient auditory signal is presented before the physical visual offset. In addition, I observed larger backward displacement when the auditory signal was presented earlier. I interpreted this finding to mean that the perceived timing of visual motion offset is attracted toward the timing of the presentation of transient sound, which results in a decreased magnitude of RM and induces backward displacement. Experiment 8 showed that the perceived visual offset position exhibits a larger shift induced by the spatial information relative to the visual target when the transient auditory signal is presented to the same side as the visual field from which the moving object originates. It can be argued that a sound presented on the same side as the visual object might be heard earlier (perhaps because attention might be biased toward the side where the visual object appeared) and consequently shift the perceived visual offset position backward more strongly (i.e., the effect is temporal). Alternatively, the spatial information of the sound relative to the visual target might shift the perceived visual offset position toward the side of the auditory signal without influencing the timing judgment (i.e., a spatial attraction of the visual offset by the auditory signal). Experiment 9 was conducted to examine whether the relative timing between visual and auditory events differs when the sound is presented on the same or opposite side as the visual object. Although RM was observed only in Experiment 8B, the results of Experiments 8A and 8B were similar. For this reason, I used the same visual and auditory stimuli as in Experiment 8A, but I asked participants to perform a temporal-order judgment task.

Method

Observers

Fifteen paid volunteers were newly recruited to participate in Experiment 9. All observers had normal or corrected-to-normal visual and auditory acuity and were naïve to the purpose of this study.

Apparatus, stimuli, and procedures

The apparatus and stimuli were identical to those used in Experiment 8A. Observers were asked to focus on the fixation cross and observe the moving disk. The auditory transients were presented -120, -80, or -40 ms before the visual offset, synchronous with the visual offset, or 40, 80, or 120 ms after the visual offset. Observers were asked to judge whether the auditory transient was presented before or after the offset of the moving disk. After 10 training trials, observers completed 420 experimental trials (2 sound positions × 3 visual offset positions × 7 sound timings × 10 repetitions). Trials in all conditions were randomly ordered.

Results and discussion

Figure 4.4 shows the results of Experiment 9. The vertical axis represents the proportion of "target disappeared first" responses. The horizontal axis represents the timing of the sound presentation. The two-way repeated measures ANOVA revealed a significant main effect of sound timing ($F(6,84) = 18.214$, $p < .001$), while the main effect of sound contingency was not significant ($F(1,14) = 0.90$, $p = .358$). No interaction was observed between sound timing and sound contingency ($F(6,84) = 0.735$, $p = .623$). Thus, the proportion of "target disappeared first" responses increased with the delay of the auditory transient; more importantly, the proportion of these responses did not differ between the same-side and opposite-side sound conditions.

The results of Experiment 9 suggest that the spatial information of the auditory transient did not affect the judgment of relative timing between the auditory and visual events. Therefore, enhanced displacement induced by the sound from the same visual field as the visual target in Experiment 8 resulted from the spatial information of the sound relative to the visual target. This produced larger spatial attraction to the visual offset location. The effect of the sound's spatial information did not interact with that of the sound's temporal information.

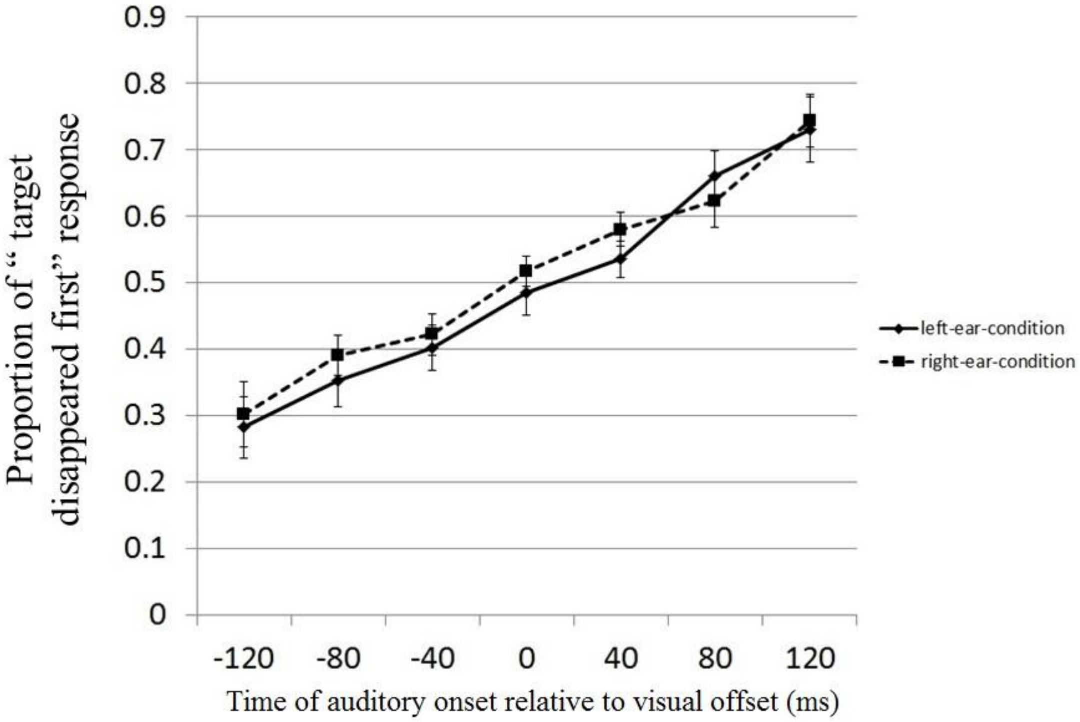


Figure 4.4 The results of Experiment 9. The horizontal and vertical axes represent the different sound presentation conditions and the proportion of “target disappeared first” responses, respectively. Error bars represent within-observers SEMs (Loftus & Masson, 1994; Cousineau, 2005) for each presentation.

General discussion

Summary

The experiment in Chapter 4 showed that the perceptual system combines auditory information to decide the perceived location of a visual target. In Experiment 7, I found that an auditory transient presented before visual motion offset shifted the perceived visual offset location backward as if it truncated the visual trajectory. In Experiment 8, when the sound was lateralized, its spatial information (on the same or opposite side as the visual target) influenced the perceived visual offset location; the visual offset position tended to be attracted toward the side of the sound presentation. In Experiment 9, the results showed that the spatial information of the lateralized sound did not influence the judgment of visual offset timing, implying that the effect of lateralized sound in Experiment 8 occurred mainly in the spatial domain (Figure 4.5).

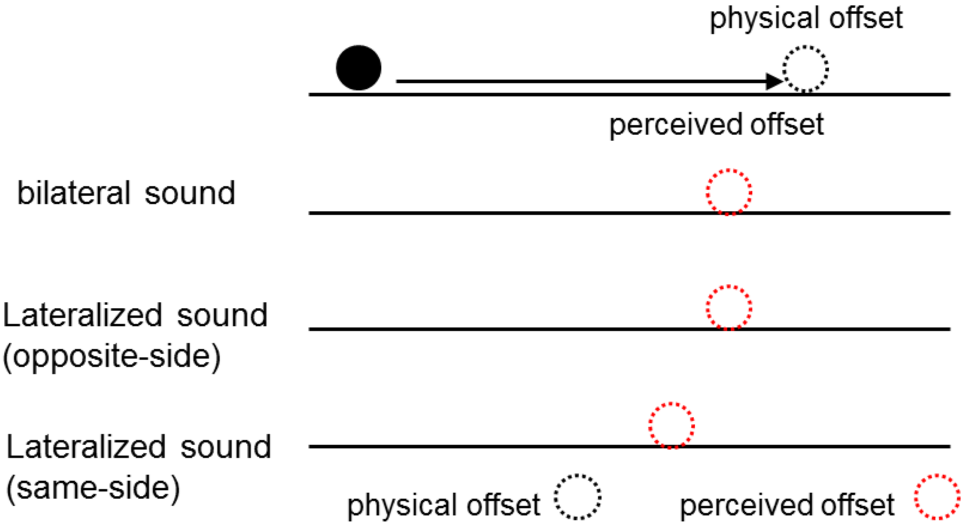


Figure 4.5 Spatial attraction induced by bilateral or lateralized auditory transients.

Crossmodal interaction in spatial representation

The results of experiments in Chapter 4 showed that the auditory transient influenced the perceived visual offset location through both spatial and temporal processes. Temporal information of the auditory transient influenced the perceived offset timing of visual objects as if it truncated the visual trajectory. However, when the auditory transient occurred in the same hemifield as the visual target, enhanced backward displacement was observed relative to when the auditory transient occurred in the hemifield opposite to the visual target.

The auditory system is generally superior to the visual system in terms of the temporal perception, and the visual system is generally superior to the auditory system in term of spatial perception. Therefore, vision can provide more accurate spatial information, while the audition can provide more accurate temporal information. The modality precision hypothesis suggests that the modality with highest precision with regard to the required task tends to be dominant in multimodal interactions (Shipley, 1964; Spence & Squire, 2003; Welch & Warren, 1980, 1986). In Experiment 7, I found that the perceived visual offset location was shifted backward when the auditory transient was present before the visual offset. This implies that the perceived timing of visual motion offset was attracted to the presentation timing of the auditory transient. Consequently, the perceived location of the visual offset was shifted backward. This is consistent with the auditory superiority for temporal perception (e.g., temporal ventriloquism; Vroomen & de Gelder, 2004). On the other hand, the results of Experiment 8 in this chapter also suggest that the effect of the lateralized sound was spatial rather than temporal. This finding cannot be explained by the modality precision hypothesis. There seem to exist significant spatial effects from audition to vision, particularly when blurred, poorly localized visual stimuli are presented (Alais & Burr, 2004). Recent research has demonstrated that the spatial aspects of sound can modulate visual motion perception (Hidaka *et al.*, 2011; Teramoto *et al.*, 2012), suggesting that visual and auditory modalities influence each other in motion processing. Thus, taken together, the results presented in Chapter 4 indicate that auditory information influences visual perception via both temporal and spatial processes.

Auditory transients might modulate RM

Maus and Nijhawan (2009) proposed a dual-process model to explain differences between how the visual system processes the positions of abruptly vanishing objects and gradually disappearing objects. The first process calculates the location of a moving object in the near future by analyzing its speed and trajectory. When the moving object disappears abruptly, the second process modulates the RM. This modulation mechanism relies on accurate spatial information provided by the transient of the abrupt offset of the moving object. A stronger transient leads to more accurate localization of the moving object because it aids location information by employing the retinal off-transient to win the competition for perceptual awareness. The findings in Chapter 4 could be interpreted to mean that the modulation mechanism relies not only on visual information provided by the retinal off-transient, but also on information provided by an auditory transient temporally and spatially close to the visual motion offset. If the auditory transient is firmly associated with the visual motion offset, the neural system also uses temporal and spatial information provided by the auditory transient to modulate possible overshoots. The present thesis suggests the possibility that the visual system integrates auditory information presented before and after the offset of visual motion.

Nevertheless, signals from different sensory modalities are not combined indiscriminately. I observed the backward displacement mainly when an auditory signal was presented 120 or 80 ms before the actual visual offset. However, I observed that the spatial information of an auditory signal modulated RM only when sound was presented 80 ms after physical visual offset (Experiment 8B). This might imply that the temporal window during which the visual system integrates auditory information is approximately 100 ms before and after visual motion offset. This is consistent with the temporal window of sound-induced illusory flash (Shams *et al.*, 2002) and multisensory integration in superior colliculus neurons in the mammalian brain (Meredith *et al.*, 1987).

Conclusion

In conclusion, the findings in this chapter suggest that an auditory transient presented before or after the offset of physical motion of a visual stimulus can modulate the perceived visual offset position. The magnitude of the backward or forward shift depends on the spatial relationship between the auditory and visual stimuli.

Given that the remembered location of visual motion offset is persistent in mental representation in short-term memory trace, the displacement of the perceived visual motion offset results from the bias in short-term memory trace. This is consistent with the proposed attention account in the present thesis. In other words, the perceived localization of the target is biased toward the peripheral transient if the perceptual system also integrates information from the peripheral transient to decide the perceived location of the visual target.

Chapter 5: General Discussion

Summary

The present thesis aimed to explore how dynamic attentional capture induced by peripheral stimuli affects the localization of visual stimuli. Peripheral sensory transients might shift perceived locations of visual targets either away (the repulsion effect) or toward (the attraction effect) them. The empirical findings can be summarized as follows: (1) The effect size of the attraction effect was larger than that of the repulsion effect with the same physical properties. (2) The preceding visual cues that induced the repulsion effect did not interact with succeeding cues that induced the attraction effect. (3) Attenuated attentional shift between the target and the cue diminished the attraction effect but had no influence on the repulsion effect. (4) The repulsion effect was larger when the cue was presented in the depth plane farther away with respect to the fixation point observers were instructed to attend to. (5) The attraction effect was larger when the distance between the target and the cue on the retinal coordinate was closer. (6) An auditory transient could shift perceived offset location of a visual target toward it. (7) Lateralized sound from the same side as the visual target produced larger displacement of the perceived offset position.

These findings are consistent with the integrated account proposed in Chapter 1. Dynamic attentional capture induced by sensory transients produced mislocalization of visual targets in two different aspects. On one hand, the peripheral transient would attract attention toward itself. Then, processing resources are allocated to processing its spatial representation. This would enhance spatial perception at the location of the peripheral transient. However, processing resources available for the following visual target would be reduced. Information processing for its location representation is impaired, leading to a systematic error that shifts the perceived location away from the preceding sensory transients. On the other hand, the visual system integrates multisensory information from both the visual target and contextual background to form the location representation of the visual target. Therefore, information on the peripheral transient might also be integrated. This would bias the target's perceived location toward the peripheral transient in the short-term memory trace.

Possible engineering applications

The findings in Chapter 3 suggest that there is an asymmetry in the allocation of spatial attention in 3D space. Observers tend to allocate more spatial attention to space before the fixation plane than space farther away from the fixation plane in a 3D environment. This is because of the most important behaviors from an ecological perspective is to rapidly redirect attention to potentially threatening or rewarding stimuli that unexpectedly invade the space close to us. This could be applied to engineering fields in presenting the contents of virtual 3D environments.

The visual environment appears to be stable and of uniform clearness. However, eye, head, and body movements lead to continuously changing retinal images. Stable spatial representation is achieved as the visual system constructs scene representations by integrating information across successive views (Hayhoe *et al.*, 1991). This suggests that not every part of a scene is processed equally. Only the attended portions of a visual scene are continuously processed. Spatial attention is needed for enhanced spatial perception to construct clear visual representations. The densely allocated spatial attention in space before the fixation plane suggests that information to construct 3D perception is mainly perceived from near space across successive views. Therefore, only details in near space before the fixation plane are continuously processed and spatial information in far space might be perceived with low resolution.

When constructing a virtual 3D environment, developers should focus on providing and updating information in detail in near space. Thus, vivid 3D perception could be effectively achieved. On the other hand, it might be not necessary to update every detail at every moment in far space. Details of the contents of far space could be presented in detail and updated when observers need to shift their attention toward it. This may also help conserve system resources when presenting a virtual 3D environment, or reduce the cost of constructing a larger-scale virtual 3D environment. For example, it might be helpful to save the system resources when presenting scenes in 3D games to provide strong 3D perception for players. In addition, film makers of 3D movies can save production cost of making 3D scenes by only providing detailed information in the specific part of virtual 3D visual space.

Possible directions for future research

The present thesis explored the effect of dynamic attentional capture induced by peripheral stimuli on the localization of visual stimuli. However, there are still some possible avenues of investigation for future research.

Spatial distortion effects in people with developmental disorders

The present thesis explores how dynamic attentional capture by peripheral transients induces mislocalization of visual objects. Observers in all experiments had no difficulties in controlling attention. People with developmental disorders have different attention mechanisms from healthy people (Dakin & Frith, 2005). For example, people with attention deficit hyperactivity disorder (ADHD) have difficulty maintaining sustained attention (Barkley, 1997), and people with autism spectrum disorders have difficulties in processing information from both temporal and spatial contexts (O'Riordan *et al.*, 2001). It is unclear whether peripheral transients would induce the same spatial distortion effect in individuals with these disorders. Studies of the effect of sensory transients on people with developmental disorders would elucidate how they process and integrate information to construct spatial representations. This would also provide further insight into the attention mechanisms of people with developmental disorders.

Evidence for neural mechanisms of mislocalization effects

The present thesis proposed an integrated account of how peripheral transients induce mislocalization of visual stimuli and the corresponding functional areas of the brain. Direct evidence from neuroimaging studies for different mislocalization effects could provide further evidence for the proposed account. Previous studies (Pratt & Turk-Browne, 2003; DiGiacomo & Pratt. 2012) and the present thesis all suggest that the repulsion effect occurs in the early stage of visual processing, before (1) the visual pathway splits into the ventral and dorsal pathways for perception and action (Goodale & Milner, 1992) and (2) binocular vision dominates. Future research should focus on brain regions reacting to peripheral transients in the primary visual cortex. On the other hand, if the attraction effect can be explained by the erroneous bias of short-term memory trace (e.g., Sheth & Shimojo, 2001; Yamada *et al.*, 2008), possible neural mechanisms for the attraction effect should be related to spatial working memory. Spatial working memory is highly related to the posterior parietal cortex (PPC). It contains neurons with delay-period activity when localizing stimuli (Constantinidis & Wang, 2004); further, the PPC is more susceptible to interruption by distracting stimuli (Knudsen, 2007). In particular, neurons in the lateral intraparietal area (LIP) exhibited an enhanced response to visually distracting stimuli away from the target location of a memory-guided saccade (Powell & Goldberg, 2000). Future research should explore whether the PPC is responsible for the attraction effect.

Mislocalization in depth

Third, although peripheral stimuli containing depth information is employed in Chapter 3, I examined the mislocalization of the target within the same depth plane. In other words, I did not examine whether depth information of sensory transients could induce mislocalization in depth. Given that attentional shift can be directed to specific depths (Downing & Pinker 1985; Gawryszewski *et al.,* 1987; Han *et al.,* 2005; Previc, 1998) and there is a cost to reaction time in shifting attention in depth (Atchley *et al.*, 1997; Han *et al.*, 2005), it is possible that mislocalization in depth could be observed, as previous research has shown that attention could distort visual space in 2D space. Moreover, it would be interesting to investigate how depth information is integrated into spatial representations and maintained in short-term memory. In 3D space, attention allocation is more dense in near space (Mimura *et al.*, 2002), and attention is the key to integrating and maintaining information perceived from the surroundings in consciousness (Mark & Rock, 1998; Simons & Chabris, 1999). It is also possible that information in near space would be processed more elaborately when the spatial representation is formed and retained longer. It would be interesting to investigate how depth information is processed in spatial working memory.

Crossmodal interaction in depth perception

Finally, it would also be interesting to investigate the roles of depth information provided by auditory stimuli in forming 3D spatial representations. In the present thesis, lateralized sound presented via headphones was used to provide spatial information for brief sounds. The auditory travel time to observers was negligible because of the short distance, and it did not provide depth information. The depth information of sound can also be provided by loudness, the energy ratio of direct and reverberant sound, and binaural information. In particular, the ratio of direct-to-reverberant energy provides the strongest depth cue with distant sound in an enclosed environment (Bronkhorst & Houtgast, 1999). Generally, the sound of a distant audiovisual event would be perceived later due to the slow speed of sound relative to light. Therefore, a delay of sound relative to visual stimuli is necessary for audiovisual temporal alignment. Research has shown that if information is available on the ratio of the direct-to-reverberant energy of distant sound, the auditory lag that results in perceived alignment increased as the simulated auditory source distance increased (Alais & Carlile, 2005). This suggests that the brain can compensate for the fact that the auditory stimuli will arrive at the observer later than the corresponding visual stimuli. In addition, this effect could be modulated by attention. It may be also possible that the perceived depth of an audiovisual event in 3D space could be modulated by the depth information of sound, such as the ratio of direct-to-reverberant energy, especially when the audiovisual event is presented far away from the observer.

Remarks

The present thesis demonstrated how dynamic attentional capture induced spatial distortion, and discussed its implications. Knowledge about cognitive processing is important for everyday life, because human errors often result from errors or overload in the cognitive system. Humans use vision to guide action. Understanding the working of vision and its limitations also helps us to behave adaptively and avoid possible human errors. In this respect, the present thesis has advanced our knowledge of how attention modulates spatial vision. In the academic field, this thesis can suggest future directions for physiological studies to explore how the brain constructs the perceptual world, and illustrate possible underlying mechanisms of processing spatial information.

References

Aglioti, S., DeSouza, J. F., & Goodale, M. A. (1995). Size-contrast illusions deceive the eye but not the hand. *Current Biology, 5*(6), 679-685.

Alais, D., & Burr, D. (2004). The ventriloquist effect results from near-optimal bimodal integration. *Current biology, 14*(3), 257-262.

Alais, D., & Carlile, S. (2005). Synchronizing to real events: Subjective audiovisual alignment scales with perceived auditory depth and speed of sound. *Proceedings of the National Academy of Sciences of the United States of America, 102*(6), 2244-2247.

Arnott, S. R., & Goodale, M. A. (2006). Distorting visual space with sound. *Vision research, 46*(10), 1553-1558.

Astafiev, S. V., Shulman, G. L., Stanley, C. M., Snyder, A. Z., Van Essen, D. C., & Corbetta, M. (2003). Functional organization of human intraparietal and frontal cortex for attending, looking, and pointing. *The Journal of Neuroscience*, *23*(11), 4689-4699.

Atchley, P., Kramer, A. F., Andersen, G. J., & Theeuwes, J. (1997). Spatial cuing in a stereoscopic display: Evidence for a “depth-aware” attentional focus. *Psychonomic Bulletin & Review*, *4*(4), 524-529.

Barkley, R. A. (1997). Behavioral inhibition, sustained attention, and executive functions: constructing a unifying theory of ADHD. *Psychological bulletin*, *121*(1), 65.

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial vision, 10*(4), 433-436.

Brenner, E., & Smeets, J. B. (2000). Motion extrapolation is not responsible for the flash–lag effect. *Vision research*, *40*(13), 1645-1648.

Brenner, E., van Beers, R. J., Rotman, G., & Smeets, J. B. (2006). The role of uncertainty in the systematic spatial distortion of moving objects. *Journal of Experimental Psychology: Human Perception and Performance, 32*(4), 811-825.

Briand, K. A., & Klein, R. M. (1987). Is Posner's" beam" the same as Treisman's" glue"?: On the relation between visual orienting and feature integration theory. *Journal of Experimental Psychology: Human Perception and Performance, 13*(2), 228-241.

Bronkhorst, A. W., & Houtgast, T. (1999). Auditory distance perception in rooms. *Nature, 397*(6719), 517-520.

Cai, R.H., Pouget A., Schlag-Rey, M., & Schlag, J. (1997). Perceived geometrical relationships affected by eyemovement signals. *Nature, 386,* 601-604.

Chen, Q., Weidner, R., Vossel, S., Weiss, P. H., & Fink, G. R. (2012). Neural mechanisms of attentional reorienting in three-dimensional space. *The Journal of Neuroscience, 32*(39), 13352-13362.

Connolly, J. D., Goodale, M. A., Menon, R. S., & Munoz, D. P. (2002). Human fMRI evidence for the neural correlates of preparatory set. *Nature neuroscience, 5*(12), 1345-1352.

Constantinidis, C., & Wang, X.-J. (2004). A neural circuit basis for spatial working memory. *The Neuroscientist, 10*(6), 553-565.

Corbetta, M., Kincade, J. M., Ollinger, J. M., McAvoy, M. P., & Shulman, G. L. (2000). Voluntary orienting is dissociated from target detection in human posterior parietal cortex. *Nature neuroscience, 3*(3), 292-297.

Corbetta, M., Patel, G., & Shulman, G. L. (2008). The reorienting system of the human brain: from environment to theory of mind. *Neuron, 58*(3), 306-324.

Cousineau, D. (2005). Confidence intervals in within-subject designs: A simpler solution to Loftus and Masson’s method. *Tutorials in Quantitative Methods for Psychology, 1*(1), 42-45.

Dakin, S., & Frith, U. (2005). Vagaries of visual perception in autism. *Neuron, 48*(3), 497-507.

DiGiacomo, A., & Pratt, J. (2012). Misperceiving space following shifts of attention: Determining the locus of the attentional repulsion effect. *Vision Research, 64,* 35-41.

Downing, C., & Pinker, S. (1985). The spatial structure of visual attention. In M. I. Posner & O. S. M. Martin (Eds.), *Attention and performance XI* (pp. 171–187). Hillsdale: Erlbaum.

Eagleman, D. M., & Sejnowski, T. J. (2000). Motion integration and postdiction in visual awareness. *Science, 287*(5460), 2036-2038.

Eagleman, D. M., & Sejnowski, T. J. (2007). Motion signals bias localization judgments: A unified explanation for the flash-lag, flash-drag, flash-jump, and Frohlich illusions. *Journal of vision, 7*(4):3, 1-12.

Eriksen, C. W., & Collins, J. F. (1969). Temporal course of selective attention. *Journal of experimental psychology, 80*(2p1), 254-261.

Eriksen, C. W., & Rohrbaugh, J. W. (1970). Some factors determining efficiency of selective attention. *The American Journal of Psychology*, 330-342.

Freyd, J. J. (1983). The mental representation of movement when static stimuli are viewed. *Perception & Psychophysics, 33*(6), 575-581.

Freyd, J. J., & Finke, R. A. (1985). A velocity effect for representational momentum. *Bulletin of the Psychonomic Society, 23*(6), 443-446.

Freyd, J. J., Pantzer, T. M., & Cheng, J. L. (1988). Representing statics as forces in equilibrium. *Journal of Experimental Psychology: General, 117*(4), 395-407.

Gawryszewski, L. d. G., Riggio, L., Rizzolatti, G., & Umiltá, C. (1987). Movements of attention in the three spatial dimensions and the meaning of “neutral” cues. *Neuropsychologia, 25*(1), 19-29.

Goodale, M. A. (2011). Transforming vision into action. *Vision research, 51*(13), 1567-1587.

Goodale, M. A., & Keith Humphrey, G. (1998). The objects of action and perception. *Cognition, 67*(1), 181-207.

Goodale, M. A., & Milner, A. D. (1992). Separate visual pathways for perception and action. *Trends in neurosciences, 15*(1), 20-25.

Grimes, J. (1996). On the failure to detect changes in scenes across saccades. In A kins K. (Ed.), *Perception (Vancouver Studies in Cognitive Science*, Vol. 5, pp.89-110). New York: Oxford University Press.

Haffenden, A. M., & Goodale, M. A. (1998). The effect of pictorial illusion on prehension and perception. *Journal of Cognitive Neuroscience, 10*(1), 122-136.

Haffenden, A. M., & Goodale, M. A. (2000). Independent effects of pictorial displays on perception and action. *Vision Research, 40*(10), 1597-1607.

Haffenden, A. M., Schiff, K. C., & Goodale, M. A. (2001). The dissociation between perception and action in the Ebbinghaus illusion: Nonillusory effects of pictorial cues on grasp. *Current Biology, 11*(3), 177-181.

Hamker, F. H., Zirnsak, M., Calow, D., & Lappe, M. (2008). The peri-saccadic perception of objects and space. *PLoS computational biology, 4*(2), e31.

Han, S., Wan, X., & Humphreys, G. W. (2005). Shifts of spatial attention in perceived 3-D space. *The Quarterly Journal of Experimental Psychology, 58*(4), 753-764.

Hayhoe, M., Lachter, J., & Feldman, J. (1991). Integration of form across saccadic eye movements.*Perception,20*(3), 393-402.

He, S., Cavanagh, P., & Intriligator, J. (1997). Attentional resolution. *Trends in cognitive sciences, 1*(3), 115-121.

He, Z. J., & Nakayama, K. (1995). Visual attention to surfaces in three-dimensional space. *Proceedings of the National Academy of Sciences, 92*(24), 11155-11159.

Heemskerk, J., DiNardo, S., & Kostriken, R. (1996). Attentional resolution and the locus of visual awareness. *Nature, 383*, 26.

Helmholtz, H. V. (1925). *Treatise on physiological optics.* (J.P.C Southall, Ed. and Trans.). New York: Optical Society of America.

Henderson, J. M., & Hollingworth, A. (1999). The role of fixation position in detecting scene changes across saccades. *Psychological Science, 10*(5), 438-443.

Heron, J., Whitaker, D., & McGraw, P. V. (2004). Sensory uncertainty governs the extent of audio-visual interaction. *Vision research, 44*(25), 2875-2884.

Hidaka, S., Teramoto, W., Sugita, Y., Manaka, Y., Sakamoto, S., & Suzuki, Y. (2011). Auditory motion information drives visual motion perception. *PloS one, 6*(3), e17499.

Hikosaka, O., Miyauchi, S., & Shimojo, S. (1993a). Focal visual attention produces illusory temporal order and motion sensation. *Vision research, 33*(9), 1219-1240.

Hikosaka, O., Miyauchi, S., & Shimojo, S. (1993b). Voluntary and stimulus-induced attention detected as motion sensation. *Perception, 22*, 517-526.

Hopfinger, J., Buonocore, M., & Mangun, G. (2000). The neural mechanisms of top-down attentional control. *Nature neuroscience, 3*(3), 284-291.

Hubbard, T. L. (2005). Representational momentum and related displacements in spatial memory: A review of the findings. *Psychonomic Bulletin & Review, 12*(5), 822-851.

Hubbard, T. L., & Bharucha, J. J. (1988). Judged displacement in apparent vertical and horizontal motion. *Perception & Psychophysics, 44*(3), 211-221.

Hubbard, T. L., & Ruppel, S. E. (2000). Spatial memory averaging, the landmark attraction effect, and representational gravity. *Psychological research, 64*(1), 41-55.

Kastner, S., Pinsk, M. A., De Weerd, P., Desimone, R., & Ungerleider, L. G. (1999). Increased activity in human visual cortex during directed attention in the absence of visual stimulation. *Neuron, 22*(4), 751-761.

Kelly, D. (1971a). Theory of flicker and transient responses, I. Uniform fields. *Journal of the Optical Society of America, 61*(4), 537-546.

Kelly, D. (1971b). Theory of flicker and transient responses, II. Counterphase gratings. *Journal of the Optical Society of America, 61*(5), 632-640.

Kelly, D. (1972). Adaptation effects on spatio-temporal sine-wave thresholds. *Vision research, 12*(1), 89-101.

Kerzel, D. (2002). Memory for the position of stationary objects: Disentangling foveal bias and memory averaging. *Vision research, 42*(2), 159-167.

Kerzel, D. (2003). Mental extrapolation of target position is strongest with weak motion signals and motor responses. *Vision Research, 43*(25), 2623-2635.

Kerzel, D., Jordan, J. S., & Müsseler, J. (2001). The role of perception in the mislocalization of the final position of a moving target. *Journal of Experimental Psychology: Human Perception and Performance, 27*(4), 829-840.

Knudsen, E. I. (2007). Fundamental components of attention. *Annu. Rev. Neurosci., 30*, 57-78.

Kosovicheva, A. A., Fortenbaugh, F. C., & Robertson, L. C. (2010). Where does attention go when it moves?: Spatial properties and locus of the attentional repulsion effect. *Journal of Vision, 10*(12):33, 1-13.

Krekelberg, B., & Lappe, M. (2000). A model of the perceived relative positions of moving objects based upon a slow averaging process. *Vision research, 40*(2), 201-215.

LaBar, K. S., Gitelman, D. R., Parrish, T. B., & Mesulam, M. (1999). Neuroanatomic overlap of working memory and spatial attention networks: a functional MRI comparison within subjects. *Neuroimage, 10*(6), 695-704.

Levin, D. T., & Simons, D. J. (1997). Failure to detect changes to attended objects in motion pictures. *Psychonomic Bulletin & Review, 4*(4), 501-506.

Loftus, G. R., & Masson, M. E. (1994). Using confidence intervals in within-subject designs. *Psychonomic Bulletin & Review, 1*(4), 476-490.

MacDonald, J., & McGurk, H. (1978). Visual influences on speech perception processes. *Perception & Psychophysics, 24*(3), 253-257.

Mack, A., & Rock, I. (1998). *Inattentional blindness*: The MIT Press.

Mack, A., Tang, B., Tuma, R., Kahn, S., & Rock, I. (1992). Perceptual organization and attention. *Cognitive Psychology, 24*(4), 475-501.

Maus, G. W., & Nijhawan, R. (2009). Going, going, gone: localizing abrupt offsets of moving objects. *Journal of Experimental Psychology: Human Perception and Performance, 35*(3), 611-626.

Mcnamara, T. P., & Diwadkar, V. A. (1997). Symmetry and asymmetry of human spatial memory. *Cognitive Psychology, 34*(2), 160-190.

Meredith, M. A., Nemitz, J. W., & Stein, B. E. (1987). Determinants of multisensory integration in superior colliculus neurons. I. Temporal factors. *The Journal of neuroscience, 7*(10), 3215-3229.

Mesulam, M.-M. (1999). Spatial attention and neglect: parietal, frontal and cingulate contributions to the mental representation and attentional targeting of salient extrapersonal events. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences, 354*(1387), 1325-1346.

Meyer, G. F., Wuerger, S. M., Röhrbein, F., & Zetzsche, C. (2005). Low-level integration of auditory and visual motion signals requires spatial co-localisation. *Experimental Brain Research, 166*(3-4), 538-547.

Mishkin, M., Ungerleider, L. G., & Macko, K. A. (1983). Object vision and spatial vision: two cortical pathways.*Trends in neurosciences*,*6*, 414-417.

Miura, T., Shinohara, K., & Kanda, K. (2002). Shift of attention in depth in a semi‐realistic setting. *Japanese Psychological Research, 44*(3), 124-133.

Moore, C. M., & Egeth, H. (1997). Perception without attention: evidence of grouping under conditions of inattention. *Journal of Experimental Psychology: Human Perception and Performance, 23*(2), 339-352.

Most, S. B., Scholl, B. J., Clifford, E. R., & Simons, D. J. (2005). What you see is what you set: sustained inattentional blindness and the capture of awareness. *Psychological review, 112*(1), 217-242.

Most, S. B., Simons, D. J., Scholl, B. J., Jimenez, R., Clifford, E., & Chabris, C. F. (2001). How not to be seen: The contribution of similarity and selective ignoring to sustained inattentional blindness. *Psychological Science, 12*(1), 9-17.

Nakayama, K., & Mackeben, M. (1989). Sustained and transient components of focal visual attention. *Vision research, 29*(11), 1631-1647.

Neisser, U. (1979). The control of information pickup in selective looking. In A.D. Pick(Ed.), *Perception and its development: A tribute to Eleanor J. Gibson* (pp. 201-219). Hillsdale, NJ: Erlbaum.

Nelson, T. O., & Chaiklin, S. (1980). Immediate memory for spatial location. *Journal of Experimental Psychology: Human Learning and Memory, 6*(5), 529-545.

Newby, E. A., & Rock, I. (1998). Inattentional blindness as a function of proximity to the focus of attention. *Perception, 27*, 1025-1040.

Nijhawan, R. (1994). Motion extrapolation in catching. *Nature*, 370(6487), 256-257.

Ono, F., & Watanabe, K. (2011). Attention can retrospectively distort visual space. *Psychological science, 22*(4), 472-477.

O'Riordan, M. A., Plaisted, K. C., Driver, J., & Baron-Cohen, S. (2001). Superior visual search in autism. *Journal of Experimental Psychology: Human Perception and Performance, 27*(3), 719.

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial vision, 10*(4), 437-442.

Posner, M. I. (1980). Orienting of attention. *Quarterly journal of experimental psychology, 32*(1), 3-25.

Posner, M. I., Snyder, C. R., & Davidson, B. J. (1980). Attention and the detection of signals. *Journal of experimental psychology: General*, *109*(2), 160-174.

Powell, K. D., & Goldberg, M. E. (2000). Response of neurons in the lateral intraparietal area to a distractor flashed during the delay period of a memory-guided saccade. *Journal of Neurophysiology, 84*(1), 301-310.

Pratt, J., & Arnott, S. R. (2008). Modulating the attentional repulsion effect. *Acta psychologica, 127*(1), 137-145.

Pratt, J., & Turk-Browne, N. B. (2003). The attentional repulsion effect in perception and action. *Experimental Brain Research, 152*(3), 376-382.

Previc, F. H. (1998). The neuropsychology of 3-D space. *Psychological bulletin, 124*(2), 123.

Prinzmetal, W., Presti, D. E., & Posner, M. I. (1986). Does attention affect visual feature integration? *Journal of Experimental Psychology: Human Perception and Performance, 12*(3), 361-369.

Purushothaman, G., Patel, S. S., Bedell, H. E., & Ogmen, H. (1998). Moving ahead through differential visual latency. *Nature, 396*(6710), 424-424.

Rensink, R. A., Kevin O'Regan, J., & Clark, J. J. (2000). On the failure to detect changes in scenes across brief interruptions. *Visual Cognition, 7*(1-3), 127-145.

Rensink, R. A., O'Regan, J. K., & Clark, J. J. (1997). To see or not to see: The need for attention to perceive changes in scenes. *Psychological Science, 8*(5), 368-373.

Rock, I., Linnett, C. M., Grant, P., & Mack, A. (1992). Perception without attention: Results of a new method. *Cognitive Psychology, 24*(4), 502-534.

Ross, J., Morrone, M. C., & Burr, D. C. (1997). Compression of visual space before saccades. *Nature, 386*(6625), 598-601.

Roulston, B. W., Self, M. W., & Zeki, S. (2006). Perceptual compression of space through position integration. *Proceedings of the Royal Society B: Biological Sciences, 273*(1600), 2507-2512.

Ryan, T. H. (1960). Significance tests for multiple comparison of proportions, variances, and other statistics. *Psychological bulletin, 57*(4), 318-328.

Sadalla, E. K., Burroughs, W. J., & Staplin, L. J. (1980). Reference points in spatial cognition. *Journal of Experimental Psychology: Human Learning and Memory, 6*(5), 516-528.

Shams, L., Kamitani, Y., & Shimojo, S. (2002). Visual illusion induced by sound. *Cognitive Brain Research, 14*(1), 147-152.

Sheth, B. R., & Shimojo, S. (2001). Compression of space in visual memory. *Vision research, 41*(3), 329-341.

Shim, W. M., & Cavanagh, P. (2004). The motion-induced position shift depends on the perceived direction of bistable quartet motion. *Vision research, 44*(20), 2393-2401.

Shipley, T. (1964). Auditory flutter-driving of visual flicker. *Science, 145*(3638), 1328-1330.

Shulman, G. L., Ollinger, J. M., Akbudak, E., Conturo, T. E., Snyder, A. Z., Petersen, S. E., et al. (1999). Areas involved in encoding and applying directional expectations to moving objects. *The Journal of Neuroscience, 19*(21), 9480-9496.

Simmering, V. R., Peterson, C., Darling, W., & Spencer, J. P. (2008). Location memory biases reveal the challenges of coordinating visual and kinesthetic reference frames. *Experimental Brain Research, 184*(2), 165-178.

Simons, D. J., & Chabris, C. F. (1999). Gorillas in our midst: Sustained inattentional blindness for dynamic events. *Perception-London, 28*(9), 1059-1074.

Simons, D. J., & Levin, D. T. (1997). Change blindness. *Trends in cognitive sciences, 1*(7), 261-267.

Simons, D. J., & Levin, D. T. (1998). Failure to detect changes to people during a real-world interaction. *Psychonomic Bulletin & Review, 5*(4), 644-649.

Simons, D. J., & Rensink, R. A. (2005). Change blindness: Past, present, and future. *Trends in cognitive sciences, 9*(1), 16-20.

Soto-Faraco, S., Lyons, J., Gazzaniga, M., Spence, C., & Kingstone, A. (2002). The ventriloquist in motion: Illusory capture of dynamic information across sensory modalities. *Cognitive Brain Research, 14*(1), 139-146.

Spence, C. (2007). Audiovisual multisensory integration. *Acoustical science and technology, 28*(2), 61-70.

Spence, C., & Squire, S. (2003). Multisensory integration: maintaining the perception of synchrony. *Current Biology, 13*(13), R519-R521.

Spencer, J. P., Simmering, V. R., & Schutte, A. R. (2006). Toward a formal theory of flexible spatial behavior: geometric category biases generalize across pointing and verbal response types. *Journal of Experimental Psychology: Human Perception and Performance, 32*(2), 473-490.

Suzuki, S., & Cavanagh, P. (1997). Focused attention distorts visual space: an attentional repulsion effect. *Journal of Experimental Psychology: Human Perception and Performance, 23*(2), 443-463.

Teramoto, W., Hidaka, S., Gyoba, J., & Suzuki, Y. (2010). Auditory temporal cues can modulate visual representational momentum. *Attention, Perception, & Psychophysics, 72*(8), 2215-2226.

Teramoto, W., Hidaka, S., Sugita, Y., Sakamoto, S., Gyoba, J., Iwaya, Y., et al. (2012). Sounds can alter the perceived direction of a moving visual object. *Journal of Vision, 12*(3):11, 1-12.

Treisman, A. (1985). Preattentive processing in vision. *Computer vision, graphics, and image processing, 31*(2), 156-177.

Treisman, A., & Paterson, R. (1984). Emergent features, attention, and object perception. *Journal of Experimental Psychology: Human Perception and Performance, 10*(1), 12-31.

Treisman, A., & Schmidt, H. (1982). Illusory conjunctions in the perception of objects. *Cognitive psychology, 14*(1), 107-141.

Treisman, A., & Souther, J. (1985). Search asymmetry: a diagnostic for preattentive processing of separable features. *Journal of Experimental Psychology: General, 114*(3), 285-310.

Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive psychology, 12*(1), 97-136.

Vroomen, J., & de Gelder, B. (2004). Temporal ventriloquism: sound modulates the flash-lag effect. *Journal of Experimental Psychology: Human Perception and Performance, 30*(3), 513-518.

Watanabe, K., & Yokoi, K. (2006). Object-based anisotropies in the flash-lag effect. *Psychological Science, 17*(8), 728-735.

Watanabe, K., & Yokoi, K. (2007). Object-based anisotropic mislocalization by retinotopic motion signals. *Vision research, 47*(12), 1662-1667.

Watanabe, K., & Yokoi, K. (2008). Dynamic distortion of visual position representation around moving objects. *Journal of Vision, 8*(3):13, 1-11.

Welch, R., & Warren, D. (1986). Intersensory interactions. *Handbook of perception and human performance., 1*, 25-21.

Welch, R. B., & Warren, D. H. (1980). Immediate perceptual response to intersensory discrepancy. *Psychological bulletin, 88*(3), 638-667.

Whitney, D., & Murakami, I. (1998). Latency difference, not spatial extrapolation. *Nature neuroscience, 1*(8), 656-657.

Whitney, D., Murakami, I., & Cavanagh, P. (2000). Illusory spatial offset of a flash relative to a moving stimulus is caused by differential latencies for moving and flashed stimuli. *Vision research, 40*(2), 137-149.

Womelsdorf, T., Anton-Erxleben, K., Pieper, F., & Treue, S. (2006). Dynamic shifts of visual receptive fields in cortical area MT by spatial attention. *Nature neuroscience, 9*(9), 1156-1160.

Yamada, Y., Kawabe, T., & Miura, K. (2008). Mislocalization of a target toward subjective contours: Attentional modulation of location signals. *Psychological research, 72*(3), 273-280.

Yamada, Y., Miura, K., & Kawabe, T. (2011). Temporal course of position shift for a peripheral target. *Journal of Vision, 11*(6):6, 1-12.

Yamagishi, N., Anderson, S. J., & Ashida, H. (2001). Evidence for dissociation between the perceptual and visuomotor systems in humans. *Proceedings of the Royal Society of London. Series B: Biological Sciences, 268*(1470), 973-977.

Yeshurun, Y., & Carrasco, M. (1998). Attention improves or impairs visual performance by enhancing spatial resolution. *Nature, 396*(6706), 72-75.

Zirnsak, M., Lappe, M., & Hamker, F. H. (2010). The spatial distribution of receptive field changes in a model of peri-saccadic perception: predictive remapping and shifts towards the saccade target. *Vision research, 50*(14), 1328-1337.