SENSITIVITY TO 3D ORIENTATION IN TEXTURED SURFACES

by

GENIVA LIU

B. A. Honors, The University of Alberta, 1997

A THESIS SUBMITTED IN PARTIAL FULFILMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF ARTS

in

THE FACULTY OF GRADUATE STUDIES

(Department of Psychology, Cognitive Systems Programme)

We accept this thesis as conforming
to the required standard

THE UNIVERSITY OF BRITISH COLUMBIA

September 2001

© Geniva Liu, 2001

In presenting this thesis in partial fulfilment of the requirements for an advanced degree at the University of British Columbia, I agree that the Library shall make it freely available for reference and study. I further agree that permission for extensive copying of this thesis for scholarly purposes may be granted by the head of my department or by his or her representatives. It is understood that copying or publication of this thesis for financial gain shall not be allowed without my written permission.

Department of Psychology, Cognitive Systems Programme

The University of British Columbia
Vancouver, Canada

Date September 12, 2001

Abstract

The present study investigated depicted 3D orientation in texture gradients as a potential feature to represent data attributes in scientific visualization applications. The task was to detect and/or localize texture elements that varied across two components of 3D orientation, theta (left-right orientation differences from y-axis) and phi (rotation about the y-axis). Experiment 1 determined observers' detection sensitivity to orientation-defined 3D targets, and revealed that detection performance for 3D targets was based on depicted 3D orientation rather than 2D image orientation. In Experiment 2, the detection task was followed by a localization via direct action (pointing to an onscreen location) or indirect action (making a spatially mapped keypress). Results showed that if targets were correctly detected, they were also localized to the visual field quadrant, with accuracy generally being higher for direct action responses. Additionally, observers showed an ability to localize targets that they failed to detect. Experiment 3 demonstrated that, compared to performance of the component tasks alone, detection and direct localization did not suffer in a dual-task context, but indirect localization did. These results support a dual visual systems theory that postulates separate processing streams that specialize in processing visual information for action or perception. Results also have practical implications for how 3D orientation can be used in visualization applications.

Table of Contents

Abstract	ii
Table of Contents	iii
List of Tables	v
List of Figures	vi
Dedication	vii
 Introduction	 1
Experiment 1	3
Method	4
Participants	5
Stimuli and Design	5
Procedure	7
Results	8
Detection Errors for 2D Textures	8
Detection Errors for 3D Texture Gradients	10
Influence of 2D versus 3D factors.	12
Discussion	13
Experiment 2	14
Method	19
Participants	19
Stimuli and Design	20
Procedure	21
Results	22
Detection errors	22
Influence of 2D versus 3D factors.	23
Sensitivity and criterion levels.	23
Localization Errors	24
Controlling for baseline differences in detection sensitivity and criterion.	26
Predicting Target Localization From Target Detection	27
Localization For Undetected Targets	27
Discussion	29
Experiment 3	34
Method	36
Participants	36
Stimuli, Design, and Procedure	36
Results	37
Single- and Dual-task Detection	37
Sensitivity and criterion levels.	38
Single- and Dual-task Localization	38
Discussion	39
General Discussion	41
Implications for scientific visualization	43
Implications for the visual perception of target identity and location	45

Concluding Remarks	47
References	49
Figure Captions.....	54

List of Tables

Table 1: Experiment 1: Partial Correlations Between Predictors Of Detection Errors in 3D Texture Gradients.....	53
--	----

List of Figures

Figure 1. Examples of 2D and 3D textures.	55
Figure 2. Detection errors according to 2D and 3D orientation	56
Figure 3. Localization array and response keys	57
Figure 4. Direct and indirect target localization errors at cell, quadrant, and hemifield resolutions.	58
Figure 5. Detection accuracy plotted against direct and indirect localization accuracy	59
Figure 6. Direct and indirect localization errors for trials where observers failed to detect the target.....	60
Figure 7. Single- and dual-task detection errors as a function of Theta and Phi Magnitude.	61
Figure 8. Direct and indirect localization errors for single- and dual-task conditions	62

to the cat

Introduction

Technological advances have allowed us to collect, examine, and store ever-increasing amounts of information. Accompanying this development is the challenge of how to present vast amounts of data such that users can extract relevant information. Visualization is defined as “the use of computer-supported, interactive, visual representations of data to amplify cognition” (Card, Mackinlay, & Shneiderman, 1999, p. 6). Examples of visualization applications are those used to represent data for medical imaging, air traffic control displays, and advanced World Wide Web searches. Scientific, or data visualization, refers to applications designed to facilitate interactive exploration of large, complex, multidimensional data sets by representing them in a visual format. To be successful, designers of such applications have endeavored to take advantage of human perceptual abilities when creating the visualization displays. For instance, Healey and his colleagues (e.g., Healey, Booth, & Enns, 1996; Healey & Enns, 1998, 1999) have used differences among visual features, such as colour and density, that can be detected rapidly and in parallel to represent data attributes. These features have been applied to visualizations of oceanographic and meteorological datasets.

A trend in visualization applications is to employ three-dimensional (3D) depictions of the information set, both because more information can be represented in a 3D scene than in a two-dimensional (2D) image and because the human visual system is equipped to rapidly interpret 3D scenes. Also, recent technological advances in the ability to render complex graphics with greater efficiency have contributed to an increase in 3D visualization applications.

The number of data attributes that can be represented in a visualization application increases both with the number of independent dimensions that are varied within a display and with the number of feature values that are possible along a single dimension. In the visualization applications prototyped by Healey & Enns (1998), the height, colour, and density of 3D texture elements were allowed to vary in order to represent variations in dataset attributes such as wind

speed, ocean temperature, etc. A potentially useful feature of the texture elements that was not varied in those applications was the 3D orientation of the texture elements. Although it has been demonstrated in previous research that element orientation is a feature that can be detected efficiently in densely textured 2D images (Sagi & Julesz, 1985b; Treisman & Gormican, 1988) and in sparsely populated displays of 3D objects (Enns & Rensink, 1990a; 1990b), there has not been a study that has examined target detection based on the orientation of image elements in dense 3D textures such as those used in the visualization applications of Healey & Enns (1998).

The first goal of this study was to examine the detection sensitivity of human observers to differences in the depicted 3D orientation of perceptual elements in densely textured arrays (called pexels by Healey & Enns, 1998). Using a simple detection task, sensitivity to a range of target-background pexel orientation differences was measured and compared in 2D textures and depicted 3D texture gradients. Results showed that orientation sensitivity between 2D textures and 3D texture gradients was quite similar. For the 3D texture gradients, sensitivity was shown to be based primarily upon depicted 3D orientation, and not 2D image-based orientation differences.

Results of Experiment 1 raised the question of how precisely observers could locate the targets that were detected in the display. Visualization applications are often used to represent information that changes across time, such as changing meteorological patterns. After searching through several displays, users may need to go back to find target information that has been previously viewed. Thus one issue was determine if observers had information about target location so that it could be used as a method to guide searches within a previously viewed dataset. An additional issue examined was whether the method in which users indicated the target location made a difference to localization performance. Currently, many applications rely on conventional input devices such as the keyboard for user interface with the application. However, recent theory and research point to the superiority of direct, visually guided action

when localizing targets (e.g., Milner & Goodale, 1995). Thus, different methods of localization, either a direct action such as pointing to the screen, or an indirect action such as using a keyboard, may be differentially sensitive to the information that users have about target location. To examine these questions in Experiment 2, observers performed the detection task of Experiment 1 followed by a localization task that required either a direct action response in the form of pointing to an onscreen grid, or an indirect action response in the form of a spatially mapped keypress.

Results of Experiment 2 showed that for any given pexel orientation, location sensitivity was quite similar to detection sensitivity. Additionally, location sensitivity was superior when the response was in the form of a direct action, pointing to the onscreen grid. These results raised the question of the extent to which asking observers to perform two visual tasks reduced performance on each of the component tasks. It has often been reported that performance of dual tasks results in poorer performance for at least one of the two tasks when compared to the dedicated performance in a single task (e.g., Shapiro, Arnell, & Raymond, 1997). Thus, it is important to understand how these dual-task performance deficits may influence the visualization of data within a given application. Experiment 3 examined this question for the detection and localization tasks used in the previous experiments. Results showed that detection and direct localization were not affected by the dual-task context.

Experiment 1

The first experiment addressed the question of human sensitivity to the orientation of pexels in densely textured arrays. Specifically, sensitivity to pexel orientation was compared in 2D textures and depicted 3D texture gradients.

Previous research using more sparsely arrayed displays has shown that visual search for drawings of simple convex blocks can be performed efficiently when target blocks differ from distractor blocks by a 90° difference in 3D orientation. This holds true even when there are no

associated target-distractor differences in 2D orientation among the elements (lines and shapes) that are used to depict the 3D blocks (Enns and Rensink, 1990a). Direct comparisons of visual search based on 2D versus 3D orientation differences indicated that orientation sensitivity was similar in the two domains (Enns and Rensink, 1990b). However, with regards to the present question, these findings have limited generalizability because they were based on tests of relatively large objects (one degree of visual angle or more), they involved rather sparse displays (no more than 15 objects in a display), and only large differences in orientation were tested (i.e., 90°).

Another previous study revealed that visual search was based in part on the apparent 3D size of cylinders drawn on a texture gradient, rather than only on the 2D image size of the cylinders (Aks & Enns, 1996). This suggests that early visual processes employ mechanisms of size constancy, but again, there are limitations in generalizing this result directly to the present question. For example, it is unclear whether size constancy mechanisms can operate based upon the apparent sizes and orientations of elements in a dense texture. If they cannot, then problems associated with crowding may impair search for pexels located toward the back of a depicted surface and problems of foreshortening may make it difficult to detect targets.

Examples of the texture arrays used in this experiment are shown in Figure 1a and 1b. Each array consisted of a 20 x 20 matrix of rectangular pexels that was either depicted in the fronto-parallel plane of the observer (2D texture) or on a depicted surface that slanted away from the observer (3D texture gradient). Targets consisted of a 2 x 2 patch of pexels that differed in their orientation from the background pexels. The main measure of orientation sensitivity was target detection error as a function of the angular difference between the orientation of target and background pexels.

Method

Participants

50 undergraduate students (36 females, 14 males; mean age = 20.8 years) participated in exchange for course credit. All reported normal or corrected-to-normal vision. 20 students participated in the 2D condition (10 students each in the 0° - 45° and the 45° - 90° background orientation group; see Design for details). 30 participated in the 3D condition. A larger number was tested in the 3D condition in order to increase the statistical power for the larger number of experimental factors analyzed in this condition.

Stimuli and Design

Participants were seated in a well-lit room, viewing a 17" AppleVision monitor (1024 x 768 pixel resolution). A chinrest maintained a constant 57 cm viewing distance. For the 2D texture condition, stimuli were presented and data were collected using VScope presentation software (Enns & Rensink, 1991, version 1.2.7) running on a 80 MHz Power Macintosh 8100. For the 3D texture condition, the experimental software was run on a 300 MHz Power Macintosh 8600.

Displays consisted of 400 rectangular elements (pexels) arrayed on an imaginary 20×20 grid (Figure 1a, 1b). For convenience in programming, the array was subdivided into 100 grid locations, each containing four pexels. Each pexel was jittered randomly from its starting position by 0 - 30% of the underlying grid size in the 2D textures and half this amount (0 - 15%) in the 3D textures. Jitter was halved in the 3D textures to compensate for the fact that jitter was more noticeable with the upright pexels. Each pexel was a rectangular volume, rendered in black (0% pixels lit) on a white background (100% pixels lit) with an aspect ratio of 0.4 (height) x 0.1 (width) x 0.05 (depth, only relevant in the 3D arrays). For the 2D textures, the entire pexel array subtended $21.5^{\circ} \times 21.5^{\circ}$ of visual angle, while each pexel subtended $.7^{\circ}$ in height. For the 3D textures, the entire display subtended 27.5° in width at the lower end of the array (depicted as

closest to the observer), 10.2° in width at the upper end of the array (depicted as farthest from the observer), and 9.1° in height; pexels depicted in the closest row of the array subtended $.9^\circ$ in height, while pexels depicted in the farthest row subtended $.3^\circ$ in height.

For the 2D textures, background pexel orientation was tested at 19 levels, ranging from 0° (horizontal) to 90° (vertical) in 5° steps. To make testing manageable with a single hour-long session, each participant received only half this range: $0^\circ - 45^\circ$ in one group and $45^\circ - 90^\circ$ in the other. Within each display, background orientation was uniform. For each level of background orientation, there were ten target pexel orientations ranging from $0^\circ - 45^\circ$ in 5° steps. The one combination in which target and background orientations were identical (e.g., 10° target, 10° background) was the array in which the target was absent. The remaining combinations all yielded a visible target that differed from the background by 5° to 45° in orientation. These combinations of background and target orientation resulted in 90 different conditions where the target was present and 10 different conditions where the target was absent. In order to encourage a uniform response bias, the total number of target-present and target-absent trials were equal within a block of trials.

For the 3D texture gradients, pexels were arrayed on a surface that was slanted 65° away from the fronto-parallel plane of the viewer and rendered using perspective projection (Figure 1b). The variation of background and target pexel orientation was similar to that in the 2D condition, with the exception that there were now two degrees of freedom for orientation rather than one. As shown in Figure 1c, 3D orientation can be described by the parameters theta and phi. Relative to an upright reference orientation, theta (θ) describes angular deviations from the upright Y-axis, whereas phi (ϕ) describes the degree of rotation about the Y-axis. For convenience, $\phi = 0^\circ$ was taken to be a pexel lying in the fronto-parallel plane of the observer, with negative phi values reflecting a rotation toward the observer and positive phi values reflecting a rotation away from the observer.

All distractor pexels in this experiment were upright ($\theta = 0^\circ$ and $\phi = 0^\circ$) on the depicted surface plane. Target pexels were selected to vary over 7 different values of θ (0° , 5° , 10° , 15° , 20° , 25° , 30°) and 11 different values of ϕ (45° , 35° , 25° , 15° , 5° , 0° , -5° , -15° , -25° , -35° , -45°), yielding 77 combinations in total. Of these, 76 were target-present conditions (either θ or ϕ not 0°) and 1 was the target-absent condition (θ and ϕ both 0°). For each of the 76 target-present conditions, one target was located randomly in each of the surface quadrants, with the restriction that no targets would occur on the boundary edge of the surface (defined as the closest row, the furthest two rows, and each side column of the underlying grid). This yielded 304 target-present displays (76 target-present combinations $\times$ 4 quadrant locations).

Procedure

Participants indicated whether or not they had detected a target by pressing keys marked “yes” and “no.” In each condition, half the observers had the ‘z’ key marked with a “yes” sticker and the ‘/’ key marked with a “no” sticker; the other half of the participants had this mapping reversed. Participants were told (a) to keep their eyes at the center of the display on each trial, (b) that half the displays contained targets, and (c) that they should respond as quickly and as accurately as possible, keeping their errors to 10% or less. In the 3D condition, the experimenter used a small rectangular box to demonstrate how targets were defined. To demonstrate the depicted position of the background pexels, the box was held upright with the width facing the observer. To demonstrate changes in θ , the box was tilted to the observer’s right, away from the Y-axis at varying angles. To demonstrate ϕ , the box was kept upright while it was rotated about the y-axis. Finally, to illustrate the effects of θ and ϕ varying together, the box was tilted to the right and rotated either away from or toward the observer.

Each trial began with a randomly chosen texture array that remained on the screen until a response had been made or, in the 3D condition, until 7 seconds had elapsed (this cut-off was introduced in the 3D condition because pilot participants were taking up to half a minute to

respond). Correct responses were followed by a '+' symbol presented at the center of the screen for 600 ms, incorrect responses by a '-' symbol, and no response by a '0' symbol. The inter-trial interval was 440 ms. At the end of each block of trials, a message displayed the percentage of errors for the most recent block of trials and participants were prompted to initiate the next block when they were ready.

The 2D condition consisted of a total of 540 trials, divided into 9 blocks of 60 trials. The 90 target-present conditions were evenly distributed among one half of the trials, while the 10 target-absent conditions were evenly distributed among the remaining half. The order of conditions was randomized across trials.

The 3D condition consisted of 608 trials, divided into 8 blocks of 76 trials. The 76 target-present conditions were evenly distributed among one half of the trials, while the target-absent conditions were evenly distributed among the remaining half. The order of conditions was randomized across trials.

Results

Detection Errors for 2D Textures

Detection errors for target-present conditions are shown in Figure 2a as a function of the angular difference between background and target pexel orientation.

In each group ($0^\circ - 45^\circ$ and $90^\circ - 45^\circ$) participants were tested with one cardinal background orientation (0° or 90°) and the same main diagonal (45°). Thus, a preliminary question concerned whether there were between-group differences in target detection when the data were paired according to deviations from cardinal orientation. That is, the 0° background in one group was paired with the 90° background in the other group, the 5° background was paired with the 85° background, and so on, until the identical 45° backgrounds were compared in each group. Target orientations were paired in the same manner for each background pair. If the factor

of Group had little influence on the data, then the detection results could be reported compactly in terms of target-background difference in orientation.

Detection errors were analyzed with a mixed-design ANOVA with one between-group factor of Group ($0^\circ - 45^\circ$, $90^\circ - 45^\circ$ background orientations) and two within-participant factors of Background Orientation (10 levels: cardinal orientation, deviations from cardinal orientation of 5° , 10° , 15° , 20° , 25° , 30° , 35° , 40° , 45°) and Target-Background Orientation Difference (0° , 5° , 10° , 15° , 20° , 25°). Target-background orientation differences greater than 25° were not included in this analysis because detection errors were almost nonexistent by that point. Detection responses of less than 100 ms were considered anticipatory responses and were excluded from analysis. The analysis revealed three main points:

1. Detection errors were very similar between groups. There was no main effect of Group $F(1, 18) = .05$, $p = .83$. There were significant interactions of Group x Background Orientation, $F(9, 162) = 6.26$, $p = .0001$, and Group x Background Orientation x Target-Background Orientation, $F(45, 810) = .04$, $p = .0001$; however, these interactions did not reflect any systematic group differences.

2. Errors on target-present trials decreased as the difference between target and background orientation increased. There was a significant main effect of Target-Background Orientation Difference, $F(5, 90) = 429.11$, $p = .0001$. For target-absent trials (orientation difference = 0°), errors were 4%, indicating that participants were unlikely to report seeing targets when there were none. As the target-background orientation difference increased to 5° , errors jumped to 82.7%, but by a target-background difference of 15° , errors had reached floor again at a rate of 5.6% errors, which is significantly less than the 10% criterion given to the participants, $F(1, 90) = 8.06$, $p < .01$.

3. Fewer detection errors were made with the cardinal background orientations of 0° and 90° . This was reflected in a main effect of Background Orientation, $F(9, 162) = 19.27$, $p = .0001$.

Simple effects revealed that cardinal background orientations resulted in significantly lower errors (12.1%) compared to all other background orientations combined (21.5%), $F(1, 162) = 77.17$, $p = .0001$. There was also a significant interaction between Background Orientation and Target-Background Orientation Difference, $F(45, 810) = 4.77$, $p = .0001$. This interaction reflected the fact that for target-background orientation differences of 0° , 15° , 20° , and 25° , errors were at floor across all background orientations, but for differences of 5° and 10° , errors varied across background orientations. This interaction also reflected the fact that for a small target-background orientation difference of 5° , fewer errors were made with cardinal backgrounds than all other backgrounds, $F(1, 810) = 11.84$, $p = .0001$.

This advantage for cardinal backgrounds replicates an effect often seen in the visual search literature (e.g., Treisman & Gormican, 1988; Wolfe, Friedman-Hill, Stewart, & O'Connell, 1992). Certain feature values appear to be privileged by the visual system, such that deviations from the privileged value are detected more readily than are same-size deviations from other values. In the present case, this is true for the cardinal orientation of horizontal (0°) and vertical (90°). Targets deviating from cardinal orientations are more readily detected than targets deviating from noncardinal orientations.

Detection Errors for 3D Texture Gradients

Detection errors for target-present conditions are shown in Figure 2b as a function of theta and phi. For target-absent trials ($\theta = 0^\circ$, $\phi = 0^\circ$) errors were 13.3%; thus, as in the 2D case, participants were unlikely to report seeing targets when there were none. Errors were analyzed with a repeated-measures ANOVA that examined within-participant factors of Theta ($0^\circ, 5^\circ, 10^\circ, 15^\circ, 20^\circ, 25^\circ, 30^\circ$), Phi ($-45^\circ, -35^\circ, -25^\circ, -15^\circ, -5^\circ, 0^\circ, 5^\circ, 15^\circ, 25^\circ, 35^\circ, 45^\circ$), and Depicted Location (near, far). Detection responses of less than 100 ms were considered anticipatory responses and were excluded from analysis. This analysis revealed four main points:

1. Detection errors decreased monotonically as theta increased, as reflected in a significant main effect of Theta, $F(6, 174) = 888.88, p = .0001$. Specifically, for target-present trials, errors averaged 85.7% when theta was 0°, 78.6% when theta was 5°, and then dropped steadily with each additional 5°, reaching 8.8% errors when theta was 20°.

2. Detection errors tended to increase as phi deviated from 0°, as reflected in a significant main effect of Phi, $F(10, 290) = 33.12, p = .0001$. Specifically, on target-present trials, errors averaged 22.8% when phi was 0° and increased to 43.1% when phi was $\pm 45^\circ$. In these displays, a foreshortened view of the pexels was generated as phi increased in absolute value. This foreshortening was associated with the increased detection errors.

3. The effects of increasing phi were evident only at intermediate values of theta. The Theta x Phi interaction was significant, $F(60, 1740) = 15.29, p = .0001$, reflecting the fact that at low (0°, 5°) and high (20°, 25°, 30°) values of theta, error rates were near ceiling and floor, respectively, across all values of phi. However, for intermediate (10°, 15°) values of theta, errors varied with phi as described in point 2 above. In other words, there was greater foreshortening of the target pexels as phi deviated from zero and errors increased accordingly.

4. Detection performance varied by target location in depth. There was a main effect of Depicted Location, $F(1, 29) = 184.63, p = .0001$, reflecting the fact that fewer detection errors were made for targets depicted as being near (28.8%) than far (41.2%). This is not surprising, given that targets depicted as farther away were pictorially smaller and more crowded, thus more difficult to resolve.

Consistent with the advantage for “near” targets was a significant interaction between Theta and Depicted Location, $F(6, 174) = 24.46, p = .0001$. As theta increased, errors decreased more quickly for targets depicted as near than they did for targets depicted as far. Specifically, as theta increased from 0° to 15°, errors decreased by 72.5% for targets depicted as near; the same theta increase for far targets led to only a 53.8% decrease in errors. The interaction between Phi

and Depicted Location, $F(10, 290) = 1.89$, $p = .05$, had a similar pattern. As phi increased in magnitude from 0° to $\pm 45^\circ$, errors increased by 18.0% for near targets and by 22.0% for far targets. Thus, for targets depicted as closer to the observer, the benefit of increasing deviations in theta was greater, while the cost of increasing deviations in phi was lessened.

The Theta x Phi x Depicted Location interaction was significant, $F(60, 1740) = 3.98$, $p = .0001$, reflecting variations in error rate across Theta x Phi combinations; importantly, it did not reflect any cases where detection was better for targets depicted as far.

Influence of 2D versus 3D factors.

An analysis was conducted on detection errors in the 3D condition to address the question of whether target detection was influenced more by 2D or 3D factors. Mathematically, the difference in angular orientation in the picture plane (2D orientation difference) can be derived from the 3D parameters of theta, phi, and viewing angle (Stevens, 1983). However, the interesting perceptual question is whether performance is based primarily on the 2D orientation difference, or whether it is separately sensitive to 3D variables of theta and phi. If the former is true, then it suggests that despite our efforts to render a 3D environment for this task, performance is based solely on the information in the resulting 2D image. However, the latter is true, then it confirms that observers are performing target detection on the basis of a 3D mental representation.

This question was addressed through a multiple regression analysis in which detection errors on target present trials was modeled as a function of 2D Tilt (local orientation difference between target and distractors in picture plane), 3D Theta, 3D Phi, and 3D Phi Direction (rotation away from or towards viewer). Together in a simultaneous regression model, these four predictors accounted for 86% of error variance, $F(4, 71) = 109.76$, $p < .0001$. However, when the individual regression parameters were examined, only two of the factors accounted for

significant portions of variance, 3D Theta, $t(71) = -5.32$, $p < .001$, and 3D Phi, $t(71) = 2.42$, $p < .02$. Importantly, 2D Tilt was not a significant contributor to this model, $t(71) = .32$, $p = .75$.

The partial correlations for all factors are shown in Table 1. This shows that although 2D Tilt and 3D Theta are highly correlated predictors, 2D Tilt plays a negligible role in predicting target accuracy when 3D Theta and Phi are held constant. In contrast, 3D Theta and Phi together account for almost as much variance ($R^2 = .859$) in errors as all four predictors taken together ($R^2 = .861$), which is an insignificant difference, $F(2, 71) = .27$, $p > .25$. This more parsimonious two-predictor (3D Theta and Phi) model was highly significant, $F(2, 73) = 222.24$, $p < .0001$. Thus, 2D Tilt and Phi Direction were not significant predictors of detection accuracy.

The regression coefficients from this reduced model indicated that for every one degree of difference in theta between background and target, accuracy increased by 3% (regression coefficient for 3D Theta = $-.032$). This translates into a 48% decrease in errors for every 15° of Theta. Similarly, every degree of difference in phi decreased accuracy by 0.3% (regression coefficient for 3D Phi = $.003$), which translates into a 4.5% decrease for every 15° of phi.

Discussion

The goal of this first experiment was to compare sensitivity to pexel orientation in 2D textures and 3D texture gradients, as well as to extend the results of previous work on visual search to the current types of displays. The results indicated that orientation sensitivity was similar for the angular differences in orientation (tilt) in 2D textures and for the differences in object orientation (theta) in 3D texture gradients. This indicates that the conclusions derived from previous studies of visual search using sparse displays, large objects, and large orientation differences (Enns & Rensink, 1990a, 1990b; Aks & Enns, 1996) generalizes quite well to densely textured arrays of elements depicted on a 3D surface. Notably, even though 2D orientation differences by themselves were able to predict target detection rather well in the present study, and even though 2D orientation was highly correlated with the 3D measure of

orientation (θ), only θ accounted for a significant portion of the variance in detection errors when they were both entered into the same regression model. This indicates that search in the present study was based on the depicted 3D orientation of the pexels, and not only on the 2D differences in pexel orientation.

This conclusion should not be taken to indicate, however, that 2D characteristics did not have any influence on search. For instance, the finding that ϕ has a negative influence on target detection, over and above the influence of tilt, likely reflects the increased foreshortening of pexels that accompany increases in ϕ . Foreshortening is the reduction in visibility of an object characteristic that results from the particular vantage point from which an object is viewed. In everyday viewing, the height of tall buildings and mountains are routinely underestimated by viewers because of foreshortening. In the present study, large absolute values of ϕ led to a reduction in the 2D tilt associated with a given level of θ .

This result is probably also closely tied to the slant of the depicted surface in this study, which was set to 65° . This slant was chosen for the present experimental displays in order to provide maximum visibility of differences for the θ and ϕ levels tested. Thus, employing other surface slants that might increase foreshortening with θ and ϕ variations would also be expected to negatively affect target detection.

Finally, it is notable that search was generally worse for targets depicted as being further away from the viewer on the surface plane. This increase in errors could reflect a combination of factors, including decreased retinal size of the pexels, increased crowding of the pexels, and foreshortening effects that may have been exaggerated by a reduction in pexel size. These are all factors that reduce sensitivity to target orientation because of the geometry of rendering 3D environments onto 2D images.

Experiment 2

The second experiment examined the extent to which observers were able to localize targets after performing the detection task of Experiment 1. The goals were threefold: (a) to determine the resolution of target localization for a given level of target detection accuracy, (b) to compare two localization responses that differed in the degree to which they activated the 'action-oriented' dorsal visual stream, and (c) to examine localization accuracy for visual targets that observers failed to detect.

The relationship between target detection and localization is unclear. Some researchers have proposed that efficient target detection can occur independently of target localization (Treisman & Gelade, 1980; Treisman & Sato, 1990). In the Feature Integration Theory (FIT) proposed by these authors, target detection can occur because one of the feature maps (e.g., orientation, colour, shape) has been uniquely activated. This can lead to correct target detection even though the observer has not consulted the location information that is associated with the unique feature map activity. In this theory, focused attention is required to make precise spatial information about target location available to the observer. Therefore, any search which can be performed without the focusing of attention, such as a "pop-out" search, can in principle be performed without access to the location of the target. Searches that require focused attention, on the other hand, such as search for specific conjunctions of features, a specific spatial relation, or a difficult discrimination, will have to consult a mental representation of the target's location.

In support of this theory, Treisman and Gelade (1980) found that observers could accurately identify targets that were defined by a distinctive colour or shape, even when major errors were made in target localization. However, for targets defined by a conjunction of color and shape, correct detection was associated with correct or approximately correct localization.

Other researchers have proposed that target detection occurs through a process of detecting spatially localized differences among elementary visual features (Julesz, 1984; Sagi & Julesz, 1985a). According to texton theory by these authors, visual images are analyzed on the

basis of elementary shape features such as elongated blobs, blob endings, and blob intersections. A local difference in these features can result in a spatially localized region of discontinuity and correct target detection, even before the nature of the discontinuity has been identified. In this theory, target localization based on discontinuity detection would not require focused attention.

Texton theory has been supported by findings such that detection and localization of orientation-defined targets can be performed in parallel for a small number of targets (Sagi & Julesz, 1985b). However, this has not gone unchallenged. Some have pointed out that the localization task used in that study involved only a relative localization of the targets and, as such, did not test accuracy for absolute target location (e.g., Atkinson & Braddick, 1989). Others involved in this debate have reported that if a target defined by orientation is correctly detected, then it can also be localized, but only to the correct side of the visual field (Green, 1992). In a further refinement, Atkinson and Braddick reported that for masked oriented lines, targets could be localized to a fine degree of resolution for relatively longer stimulus durations, but only to a coarser spatial resolution for shorter durations.

This review of the background literature indicates that location information cannot always be assumed to accompany correct target detection. Under some circumstances it may, but under others, detection and localization appear to be independent. The first goal of this experiment was therefore to examine the relationship between detection and localization in the 3D pexel arrays. In addition to refining our understanding of the ways in which target detection and localization are related, an answer to this question would have important implications for the applied field of scientific visualization. To what extent is the correct detection of a visual target accompanied by information about the target's location? How precise is that information? In many visualization applications, displays are constantly being updated in real-time to reflect changes in the represented data. Therefore, any particular display may only be visible for a brief period of time. If the target data is no longer visible on the display, it is not clear whether

observers will have knowledge of target location that can help them guide further queries and interactions with the dataset.

A second issue explored in this experiment concerns the nature of the target localization response. Recent theory and research point to the importance of the “action system” in the visual perception of object location (Milner & Goodale, 1995). According to this theory, visual information is processed in two neurological pathways that diverge following initial processing in cortical area V1: a ventral stream in the temporal lobe is specialized for the conscious perception of objects (e.g., recognition, identification), whereas a dorsal stream involving the parietal lobe is specialized for largely unconscious actions directed toward objects (e.g., pointing, reaching, grasping). An important prediction of this theory for the present experiment is that visually guided actions originating in the dorsal stream should contain more precise spatial information than is accessible to the ventral processing stream.

Support for this theory comes from studies of the role of action in perception. For instance, memory for objects that were experienced through touch can be superior to visual memory for the same objects (Shelton & McNamara, 2001). Other research has shown that visually guided actions are not as vulnerable to illusions of extent and contrast as are direct reports of these same displays. For example, although observers report seeing the Ebbinghaus Illusion (when a target disk is surrounded by smaller disks it appears larger than when the same target disk is surrounded by larger disks), their grasp toward the target disk was scaled to the actual, and not the perceived, size of the disk (Aglioti, DeSouza, & Goodale, 1995). The same insensitivity in grip scaling has been shown for the Ponzo Illusion (two lines of equivalent length appear to be of different lengths when superimposed at different points on a set of converging lines), although grip force (Jackson & Shaw, 2000) and target pick-up velocity (Brenner & Smeets, 1996) were susceptible to the illusion. Manual aiming movements have been shown to be sensitive to object displacements even without conscious awareness of these displacements.

Specifically, when pointing toward a target disk that is displaced during the reaching movement, observers' reach compensated for the target displacement even when they could not report presence of the change (Fecteau, Chua, Franks, & Enns, 2001). Thus, even without accurate awareness of the physical nature of an object's size or position, observers can make accurate visually guided actions towards the object.

Thus, the second goal of this experiment was to compare two different methods of responding to target location following target detection: (a) direct localization by pointing to the target location on the screen, versus (b) indirect localization by using a spatially mapped keyboard. The action-system theory predicts that the direct pointing method will lead to more accurate target localization than the indirect reporting method because the precise information used to guide the hand is not readily accessible to consciousness.

A third issue concerns the possibility that a degree of target localization may be possible despite incorrect detection. In addition to this possibility being suggested by the action-system theory (i.e., that dorsal and ventral stream functioning is independent to a degree), there is empirical evidence of correct localization in the absence of conscious awareness of a target. For example, in an experiment by Klotz & Neumann (1999), shapes were backward-masked by surrounding target shapes (i.e., metacontrast masking) so that they were subjectively unreportable by observers; nonetheless, motor responses to targets was still influenced by the unseen shapes. In this case the response was a spatially mapped keypress (indirect action) to the location signaled by the target shape. Schmidt (in press) demonstrated that isoluminant colour primes that were backward-masked by the colour target influenced motor responses toward the target, even when the primes were subjectively unreportable; here the response was to point toward the target location on the screen (direct action). In other research, target regions within flickering texture patterns were rendered subjectively invisible through the superimposition of

“complementary” textures; however, observers could still report above chance levels the display quadrant in which the unseen target appeared (Kolb & Braun, 1995).

The third goal of this experiment was therefore to test for the possibility of correct target localization following detection failure. The background literature implies that for targets that participants fail to detect, some degree of localization may still be possible, with both direct and indirect types of action. To address this question, localization accuracy was examined for trials on which observers failed to correctly detect the target. If an advantage for visually guided movements holds for target localization in visual search, it should be possible to take advantage of this ability for the purposes of scientific visualization. Specifically, if direct action (i.e., pointing) reveals more accurate spatial information about targets than a less direct action (i.e., keypress), this implies that direct input devices, such as touch screens, are better suited for visualization tasks that involve localization.

In Experiment 2, the 3D pexel detection task from Experiment 1 was followed immediately by a localization task in which a 4 x 4 matrix of possible locations was presented (Figure 3a). This matrix remained on view until a response was made. Location sensitivity was indexed by plotting localization accuracy as a function of the orientation difference between target and background pexels. Localization errors were also examined at three levels of resolution (cell, quadrant and hemifield) in order to assess the spatial resolution of the location information.

Method

Participants

52 undergraduate students (42 females, 11 male; mean age = 19.6 years) students participated in exchange for course credit. All reported normal or corrected-to-normal vision. 26 students participated in the direction localization condition (14 in the positive phi condition and 12 in the negative phi condition; see Design for details). 26 students participated in the indirect

localization conditions (13 participants in the positive phi condition and 13 in the negative phi condition).

Stimuli and Design

Participants were seated in a well-lit room. A chinrest maintained a constant 50 cm viewing distance from the monitor. Stimuli were presented using VScope presentation software (Enns & Rensink, 1991, version 1.3) running on either a 300 MHz Power Macintosh 8600 or a 233 MHz iMac G3. In the direct localization condition, a sham video camera was placed to the side of the participant and focused on the monitor as an incentive for participants to indicate their location responses first by pointing directly to the screen.

For the detection task, which came first on each trial, the displays were 3D pexel arrays with the same specifications as used in Experiment 1 (Figure 1b). For the subsequent localization task on target-present trials, the display was a black-and-white outline drawing of the surface plane upon which the pexels were located (Figure 3a). This display was rendered using the same perspective projection as in Experiment 1. The plane was divided into a 4 x 4 grid that subdivided the 64 possible target locations into 16 regions. Each region was labeled with a letter that ranged from 'A' to 'P,' starting from the top-left region and following left to right until the bottom-right region. This grid was surrounded by an empty frame that corresponded to the boundary edge where no targets were presented.

To restrict the length of experimental sessions to one hour or less, the 7 theta and 11 phi values used in Experiment 1 were split into two separate conditions that were run with different participants. In the positive phi group, target pexels varied over seven values of theta (0°, 5°, 10°, 15°, 20°, 25°, 30°) and six values of phi (0°, 5°, 15°, 25°, 35°, 45°), yielding 42 combinations in total. Of these, 41 were target-present conditions (either theta or phi not 0°) and 1 was the target-absent condition (theta and phi both 0°). For each of the 41 target-present conditions, one target was located randomly in each of the surface quadrants, with the restriction that no targets would

occur on the boundary edge of the surface (defined as the closest row, the furthest row, and each side column of the underlying grid). This yielded 164 target-present trials (41 target-present combinations x 4 quadrant locations). The same specifications were used in the negative phi condition, where the seven values of theta were combined with six negative values of phi (0° , -5° , -15° , -25° , -35° , -45°).

Procedure

Instructions for the detection task were the same as in Experiment 1, with additional instructions for the localization task. The 3D pexel array was presented for 7 seconds or until the participant responded, whichever came first. Target-present displays were always followed by the location array, which remained on screen until the participant responded. Target-absent displays were always followed by a message prompting participants to press the space bar to continue onto the next trial. Importantly, the presentation of the location array was contingent upon the presentation of a target-present display, and not on the participant's response. Thus, participants were instructed to use the presence of the localization array as feedback regarding their detection response, with the location array indicating a target-present display and the space bar prompt indicating a target-absent display.

When the location array appeared, participants indicated a target location either directly or indirectly, depending on their group assignment. If they had responded "no" on a target-present trial (a miss in the detection task), they were instructed to make their best guess of target location.

In the direct response group, participants used their dominant hand to point to the grid square that they thought contained the target in the preceding 3D array. Following this pointing response, they used the keyboard to record the letter that corresponded to the grid square they had just pointed to. The keyboard was labeled with the 16 letters from the location array (Figure 3b), spatially corresponding to the locations of the letters on the screen as close as possible. Keys

not covered with letter labels were covered with blank labels to reduce the possibility of responding with nondesignated keys.

In the indirect response group, the procedure was the same except that the pointing response was omitted. Localization responses were made only by pressing the labeled key that corresponded to the depicted target location. Thus, this response required less action than the direct response.

Each of the four between-group conditions (direct localization-positive phi, direct localization-negative phi, indirect localization-positive phi, indirect localization-negative phi) consisted of 328 trials divided into 8 blocks of 41 trials. The 41 target-present conditions were evenly distributed among one-half the trials, while target-absent conditions were evenly distributed among the remaining half. The order of conditions was randomized across trials.

Results

Detection errors

Target detection errors were analyzed with a mixed-design ANOVA that included two within-participant factors of Theta (0° , 5° , 10° , 15° , 20° , 25° , 30°) and Phi Magnitude (0° , 5° , 15° , 25° , 35° , 45°) and two between-group factors of Phi Direction (positive or negative) and Action Type (direct or indirect). Detection responses of less than 100 ms were considered anticipatory responses and were coded as errors. This analysis revealed three main points:

1. The influence of Theta and Phi Magnitude on detection errors closely resembled that for 3D textures in Experiment 1 (average target-present errors = 37.3%, target-absent errors = 14.8%; significant effects of: Theta, $F(6, 294) = 781.54$, $p = .0001$, Phi Magnitude, $F(5, 245) = 107.24$, $p = .0001$, and Theta x Phi Magnitude, $F(30, 1470) = 34.09$, $p = .0001$).

2. The main effect of Phi Direction was significant, $F(1, 49) = 4.63$, $p = .04$, with fewer detection errors made for targets with negative phi (35.4%) than for those with positive phi (39.2%). There were also significant interactions between Theta x Phi Direction, $F(6, 294) =$

13.26, $p = .0001$, Phi Magnitude x Phi Direction, $F(5, 245) = 2.91$, $p = .01$, and Theta x Phi Magnitude x Phi Direction, $F(30, 1470) = 5.33$, $p = .0001$. All of these effects appeared to stem from the slightly greater sensitivity to negative than to positive phi differences, with the interactions arising from floor and ceiling effects in accuracy at the extreme values of theta and phi.

3. There was a significant Theta x Action Type interaction, $F(6, 294) = 2.22$, $p = .04$. For theta 15°, 20°, and 25°, the direct action group showed better detection than the indirect action group. There were no other significant effects of Action Type.

Influence of 2D versus 3D factors.

To confirm the Experiment 1 finding that 3D Theta and Phi were the main factors necessary to account for target detection, detection errors in the present experiment were examined as a function of 2D Tilt, 3D Theta, 3D Phi, and 3D Phi Direction, just as they had been in Experiment 1. In a simultaneous regression model, the four predictors together accounted for 86% of detection variance, $F(4, 77) = 119.99$, $p < .0001$. Only two factors accounted for significant portions of variance, 3D Theta, $t(77) = -4.02$, $p = .0001$, and 3D Phi, $t(77) = 2.26$, $p = .03$. These two predictors alone accounted for almost as much variance ($R^2 = .858$, $F(2, 79) = 238.87$, $p < .0001$) as all four predictors together, ($R^2 = .862$), which was an insignificant difference, $F(2, 77) = .06$, $p > .25$. These results therefore replicate the finding in Experiment 1 that detection performance was based primarily on the depicted 3D orientation of the pexels, and not 2D orientation.

Sensitivity and criterion levels.

A signal detection analysis was used to examine detection discriminability separately from any detection biases that observers may have had. This seemed particularly important because there were some group differences in detection sensitivity between the two types of action and subsequent analyses showed that these groups differed in location accuracy as well.

Detection sensitivity (d') and response criterion ($\log \beta$) were therefore calculated for each participant. These were analyzed with an ANOVA examining the between-groups factor of Action Type (direct or indirect). The analyses revealed two findings:

1. Observers were significantly more sensitive in the direct group ($d' = 1.57$) than in the indirect group ($d' = 1.29$), $F(1, 50) = 10.24$, $p = .002$. This greater sensitivity of the direct action group is consistent with the better performance seen in this group when detection errors were examined as a function of theta. However, one possibility is that this greater sensitivity came about because response demands of the direction localization task elicited a more liberal response bias. This possibility was examined with an analysis of criterion levels.

2. The analysis of beta revealed that observers were significantly more conservative (less liberal) in the direct group ($\beta = 2.25$, $\log \beta = .30$) than the indirect group ($\beta = 1.68$, $\log \beta = .20$), $F(1, 50) = 4.28$, $p = .04$, ANOVA performed on $\log \beta$. This confirmed that the greater sensitivity of observers in the direct group was not the consequence of a more liberal bias. The direct action group was both more sensitive and more conservative than the indirect action group. The question that remains is whether this reflected inherent group differences despite random assignment of participants, or whether it reflected the differential influence of the action requirements of the task. This question was addressed in other analyses that follow.

Localization Errors

Target localization responses were coded using three levels of resolution: cell (only the correct cell in the 4 x 4 matrix was counted as correct), quadrant (any one of the four cells in the correct quadrant were counted as correct), and hemifield errors (any of the eight cells in the correct hemifield were counted as correct). For each resolution of response, chance levels of error were calculated as $1 - (\text{chance probability of correct response})$. Chance error rates were therefore 93.8% for the cell level of resolution, 75% for the quadrant level, and 50% for the

hemifield level. Target localization errors are shown in Figure 4 for each resolution as a function of theta and phi. Errors are shown separately for direct and indirect responses.

Location errors were analyzed with a mixed-design ANOVA that included two within-participant factors of Theta (0° , 5° , 10° , 15° , 20° , 25° , 30°) and Phi Magnitude (0° , 5° , 15° , 25° , 35° , 45°) and two between-group factors of Phi Direction (positive or negative) and Action type (direct or indirect). This analysis was conducted for each of the cell, quadrant, and hemifield resolutions. Because there were no data points for the target absent trials (i.e., combination of $\text{Theta} = 0^\circ$ and $\text{Phi} = 0^\circ$), and the design was incomplete without these cells, the rate of chance location errors was used in these cells of the design. This analysis revealed three main points:

1. Localization errors generally resembled detection errors with respect to the influence of each of the factors. Errors decreased with increases in Theta, $F(6, 294) = 301.74$ (cell), 556.18 (quadrant), 338.56 (hemifield); errors increased with increases in Phi Magnitude, $F(5, 245) = 52.63$ (cell), 50.37 (quadrant), 8.18 (hemifield); the influence of Phi Magnitude was most apparent at intermediate values of Theta (floor and ceiling effects served to attenuate these effects at more extreme values of Theta), significant Theta x Phi Magnitude interaction, $F(30, 1470) = 29.96$ (cell), 13.68 (quadrant), 6.52 (hemifield); $p = .0001$ for all reported effects.

2. There was some influence of Phi Direction on localization errors. Although there was no main effect of Phi Direction at any resolution, $F(1, 49) = .07$ (cell), 1.71 (quadrant), 1.05 (hemifield), $p > .20$, there was a significant interaction between Theta x Phi Direction at all three resolutions, $F(6, 294) = 2.39$, $p = .03$ (cell), $F(6, 294) = 4.43$, $p = .0003$ (quadrant), $F(6, 294) = 4.36$, $p = .0003$; between Phi Magnitude x Phi Direction at the quadrant level, $F(5, 245) = 3.06$, $p = .01$; and between Theta x Phi Magnitude x Phi Direction at all three resolutions $F(30, 1470) = 2.51$ (cell), 3.73 (quadrant), 2.28 (hemifield), $p = .0001$. As with detection errors, these differences appeared to reflect generally greater sensitivity to negative directions of Phi, interacting with floor and ceiling effects to produce the statistically significant interactions.

3. There was a consistent advantage for the direct response at all resolutions of localization, $F(1, 49) = 17.69$, $p = .0001$ (cell), $F(1, 49) = 13.49$, $p = .0006$ (quadrant), $F(1, 49) = 11.98$, $p = .001$ (hemifield). Fewer errors were made for direct versus indirect action at the cell (57% vs. 66.4%), quadrant (32.9% vs. 39.3%), and hemifield (19.4% vs. 24.3%) resolutions.

The Theta x Action Type interaction was also significant at all resolutions of localization, $F(6, 294) = 3.50$, $p = .002$ (cell), $F(6, 294) = 3.24$, $p = .004$ (quadrant), $F(6, 294) = 3.85$, $p = .001$ (hemifield), as was the Theta x Phi Magnitude x Action Type interaction for cell localization errors, $F(30, 1470) = .05$, $p = .006$. Again, these interactions seemed to reflect nothing more than an attenuation of the effect of action type as the result of ceiling effects in the most difficult discrimination conditions (i.e., small values of theta and large values of phi).

Controlling for baseline differences in detection sensitivity and criterion.

Because the analysis of detection sensitivity revealed baseline differences in target detectability in the two action groups, it raises the possibility that these differences reflected inherent group differences rather than differences that arose as a result of the action requirements of the task. To address this question, participants were sampled from both the direct and indirect action groups in order to equate these groups in target detection prior to an analysis of their location sensitivity. Beginning with the 26 participants in each group, participants with the highest d' in the direct group and the lowest d' in the indirect group were excluded, until d' -prime and criterion levels were equated. After 7 participants had been deleted from each group, the mean d' was 1.44 for each group; however, β was higher for the indirect ($\beta = 1.82$) than the direct ($\beta = 1.73$) localization group. To equate for β as well, examination of the d' and β levels for each participant revealed that excluding the participant with the 9th highest d' instead of the 7th highest resulted in a dataset with 19 participants per group, with $d' = 1.44$ for each group, and $\beta = 1.81$ for the direct and $\beta = 1.82$ for the indirect localization group.

Localization errors for these remaining participants were then examined with a mixed-design ANOVA involving the two within-participant factors of Theta (0°, 5°, 10°, 15°, 20°, 25°, 30°) and Phi Magnitude (0°, 5°, 15°, 25°, 35°, 45°) and two between-group factors of Phi Direction (positive or negative) and Action type (direct or indirect). The general pattern of results remained largely unchanged. Most importantly, the main effect of Action Type was significant at the cell, $F(1, 35) = 6.59$, $p = .02$, and quadrant levels, $F(1, 35) = 4.44$, $p = .04$, and approached significance at the hemifield level, $F(1, 35) = 3.22$, $p = .08$.

Predicting Target Localization From Target Detection

Detection accuracy was plotted against cell, quadrant, and hemifield localization accuracy for both direct and indirect localization groups (Figure 5) in order to examine the extent to which localization performance could be predicted on the basis of the detection response. Each plotted data point represented the mean accuracy for each of the 41 target-present Theta x Phi Magnitude combinations. The equation for the line of best fit is shown below each plot. These analysis revealed three main points:

1. When detection was nearly perfect (e.g., 98%, indirect response), localization accuracy was not as high at the cell resolution (e.g., only 57.9 %) , but was near perfect at the quadrant and hemifield resolutions (e.g., 91.6% and 99.3%, respectively). This indicates that location accuracy is very high, even at the best of times, to only a quadrant level of resolution.
2. For all resolutions and types of localization, the relationship between detection and localization accuracy was fit very well by a linear equation: R^2 ranged from .79 to .97.
3. For all levels of location resolution, the Y-intercept for the regression line was greater for the direct action response, again reflecting greater localization accuracy for that mode of response.

Localization For Undetected Targets

To examine the ability to localize targets that went unnoticed in the detection phase of the experiment, location errors were examined for the trials where observers failed to detect the target (misses). 37.1% of target-present trials met this criterion. Localization errors for these trials are shown in Figure 6 for cell, quadrant, and hemifield resolutions.

Mean errors for each combination of Theta x Action Type were analyzed in two ways. First, each mean error rate was compared to the appropriate chance localization levels using a one-tailed one-sample t-test. Second, a mixed-design ANOVA examined the within-participants factor of Theta (0° , 5° , 10° , 15° , 20° , 25° , 30°) and a between-groups factor of Action Type (direct or indirect). The same approach was used to analyze errors for each of Phi Magnitude x Action Type, with the ANOVA including a within-participants factor of Phi Magnitude (0° , 5° , 15° , 25° , 35° , 45°) and a between-groups factor of Action Type (direct or indirect). All t-tests and ANOVA analyses were repeated for errors at the cell, quadrant, and hemifield resolution. Localization errors for Theta x Phi Magnitude x Action Type were not analyzed because there were insufficient data points to complete the analysis. These analysis revealed two main points:

1. At all levels of resolution, there were conditions in which observers demonstrated significant levels of correct localization even after failing to detect the target. At the cell resolution, t-tests revealed that localization performance was significantly ($p \leq .05$) better than chance in the direct action group for theta of 5° , 10° and 15° , and phi magnitude of 0° , 15° , 25° , 35° , and 45° ; and in the indirect action group for theta of 15° , 20° and 25° , and phi magnitude of 15° , and 25° .

At the quadrant resolution, localization performance was significantly better than chance in the direct action group for theta of 5° , 10° , 15° , and 20° , and phi magnitude of 0° , 5° , 15° , 25° , 35° , and 45° ; and in the indirect action group for theta of 15° , 20° and 25° , and phi magnitude of 25° .

At the hemifield resolution, localization performance was significantly better than chance in the direct action group for theta of 5°, 10° and 15°, and phi magnitude of 0°, 25°, 35°, and 45°; and in the indirect action group for theta of 15°.

Thus, across values of theta, the direct action group could localize undetected targets to a significant degree for small values of theta (5° and 10°) when the indirect group could not. In other words, for targets that are difficult to detect (i.e., small theta), observers can directly localize the target to a degree even when the target is missed. For an intermediate value of theta (15°), both groups could significantly localize undetected targets better than chance. For larger values of theta (20°, and 25°), the indirect action group could significantly localize undetected targets while the direct action group could not. Previous analyses of detection errors revealed that fewest errors were made with large values of phi. Thus, few trials for which missed targets had large values of theta were available for analysis, so there is insufficient data to support any speculation on whether there is actually an advantage for indirect localization in such cases.

2. As suggested in point 1, under some circumstances, the direct action group demonstrated fewer localization errors than the indirect action group after failure to detect the target. The Phi Magnitude x Action Type ANOVA revealed that the main effect of Action Type approached significance at the cell resolution, $F(1, 50) = 3.52$, $p = .07$, and reached significance at the quadrant, $F(1, 50) = 8.04$, $p = .007$, and hemifield $F(1, 50) = 5.21$, $p = .03$, resolutions. This reflected a consistent advantage for the direct over the indirect action group. The Theta x Action Type ANOVA revealed no significant effects. These results indicate that the advantage for direct action holds even when the target is not detected; however, the advantage is only seen for orientation values that participants are least able to detect.

Discussion

The three questions addressed in this experiment included (a) the relationship between detection and localization accuracy, (b) the impact of action type (direct or indirect) on localization errors, and (c) the ability to localize targets in the absence of their correct detection.

Results showed a strong relationship between detection and localization accuracy for the orientation-defined 3D pexels. When detection was poor, localization was also poor at all resolutions of localization. However, when detection was accurate, localization was quite accurate at the quadrant level of resolution. In the context of theories of target detection, these results lie in between the extremes of FIT and texton theory. Recall that FIT predicts a weak or nonexistent relationship between detection and localization (Treisman & Gelade, 1980), whereas texton theory predicts a strong relationship between the two, such that if a target is detected, then it can also be precisely localized in space (Sagi & Julesz, 1985a). The results for 3D pexels are closer to texton theory than FIT. In particular, for targets that are detected with high accuracy, localization is extremely accurate at the quadrant level of resolution, but not at the cell level.

For purposes of scientific visualization, these results provide a method of predicting localization accuracy given any level of detection accuracy. Within a visualization application, if target data is no longer visible, observers will have an approximate, if not precise level of information regarding target location to guide interactions with the data. However, the ability to locate where a target appeared is likely a function of the duration of the pexel display that contains the target. Atkinson and Braddick (1989) demonstrated that for orientation-defined targets, the longer the duration of the target display, the more accurate the localization of the target. In the present experiments, the duration of the pexel displays was self-paced up to a maximum of 7 seconds before the target display disappeared. With shorter display durations, the ability to locate target pexels would likely become less precise.

With respect to the impact of response action type on localization errors, results showed a consistent advantage for direct over indirect localization responses. Observers were more

accurate at localizing targets when the response involved pointing to the screen than when the response was only a spatially-mapped keyboard response. According to theories that postulate that the visual control of actions is dependent on different mechanisms than those for visual perception, actions that originate in the dorsal stream should reveal more precise spatial information than actions originating in the ventral stream (Milner & Goodale, 1995). In support of this, the present results showed that the direct pointing response revealed more precise spatial information than the indirect keypress response.

It should be noted that both the direct and indirect actions in the present experiment involved a type of motor response, pointing to the screen and pressing a key. However, the direct response involved acting on the space in which the target was depicted, whereas the spatially mapped keypress involved acting on an indirect representation of the target space. As well, the direct action advantage was not due to any mapping differences between the onscreen localization grid and the keyboard localization grid; in both direct and indirect response conditions, the localization grid was displayed and the keypress response was made. The only difference between the conditions was that, in the direct action condition, observers had the additional requirement of pointing to the localization grid before making the keypress response. Thus, the advantage for the direct response was related to the action of pointing to the screen.

One premise of the dual visual system theories is that dorsal stream processing is specialized to act online, that is, act on objects that are present in the environment. Only the ventral stream can guide actions towards objects that are no longer present. In the current experiment, observers could make the direct localization response as soon as the localization grid appeared. Thus, the target was no longer onscreen, but there was a minimal delay between target offset and pointing onset. However, given the direct action advantage, the delay was short enough that dorsal stream information was still available to guide the direct action response. If the direct action advantage was based upon dorsal stream processing, then one prediction is that

the greater the delay between the target display and the localization response, the less likely that the localization response can be based on dorsal stream information. Instead, even a direct localization response may be based on ventral stream information after a sufficient delay, because the response will then be based on a stored representation of target location.

To test this hypothesis, a delay between target offset and the initiation of the localization response should be introduced in future experiments. There may be a point at which direct and indirect responses reveal equivalent amounts of information about target location. In other words, there may be a critical period after target offset in which dorsal stream information is available to guide precise localization responses. In the context of visualization applications, the answer to this question is important in order to understand how long the user can be expected to have information about target location after the target is no longer visible.

Another practical implication of the direct localization advantage is that users of visualization applications should be better able to reveal precise spatial information about target location if their responses are directed to the screen on which the target information appeared. Current technologies rely heavily on input devices such as a mouse or a keyboard, which have an indirect spatial mapping to the onscreen display space. However, results suggest that users will be more accurate with interactions that involve target localization if they act directly on the display space. This implies that input devices such as touch screens or any gestural input that involves acting directly upon the display space will result in more accurate and efficient interactions with the dataset.

In addition to demonstrating the level of localization precision associated with a given level of detection performance, results also indicated that observers could localize targets after failure to detect the targets. It is possible that this is a form of visual information influencing behaviour without the observer's awareness. There have been previous instances in the literature where localization responses have been guided by visual information that observers cannot

consciously report; some researchers have even called this a form of blindsight in normal observers (Kolb & Braun, 1995). The term 'blindsight' is normally used to describe a phenomenon in which patients who have sustained damage to the visual cortex have an impaired ability to consciously experience visual stimuli, but can still significantly discriminate the same stimuli through their responses (Pöppel, Held, & Frost, 1973; Weiskrantz, Warrington, Sanders, & Marshall, 1974). Goodale, Milner, and colleagues have described the behaviour a patient, DF, who cannot report awareness of forms and shapes, yet can accurately direct actions towards such objects (e.g., Goodale, Milner, Jakobson, & Carey, 1991). They have hypothesized that patient DF has ventral stream damage, but her dorsal stream is intact, which allows her to make visually guided actions. In other words, even when ventral stream information is unavailable, visual information is being processed in the dorsal stream where it is unavailable for conscious report, but still able guide actions towards objects.

The proposal that the present demonstration of localization without awareness is a form of 'blindsight' in normal observers should be made cautiously. In the present experiment, healthy observers failed to report the presence of oriented targets, but their physical actions (i.e., localization responses) were guided by the same targets that they could not report. In some instances, the direct action localizations were more accurate than indirect action; this advantage was particularly in evidence for target orientations that were difficult to detect. One possibility is that the direct localization group was more conservative for the detection task. Results showed that the direct action group was more conservative, as well as being more sensitive, in the detection task. It is difficult to discern from the present data whether sensitivity and criterion differences between the direct and indirect localization groups contributed towards the direct localization advantage for undetected targets. To resolve this issue, the current experiment should be replicated with Action Type (direct or indirect) as a within-participants factor, instead of a between-groups factor as it was here. If the direct localization advantage for undetected

targets is still in evidence with Action Type as a within-participants factor, this would be consistent with the idea that the visual information that guides direct action is independent of the information that supports conscious awareness.

Another limitation of the current experiment is that there are few trials that contribute to the analysis of localization performance when observers fail to detect the target. To address this, target detection and localization should be further differentiated by decreasing detection performance while leaving localization performance intact. This could be accomplished by replicating the current detection-then-localization experiment by using a shorter exposure duration for the detection array, perhaps even followed by a backward mask. If, under these circumstances, observers can still make correct localization responses when they fail to detect the target, then this would be a more convincing demonstration of “blindsight” in normal observers. A further avenue of research would be to examine whether there are other sources of information, such as unconscious eye movements, that observers are able to use to guide correct localization responses after a detection failure.

Experiment 3

An important question not addressed by the previous experiments is the extent to which the dual tasks of detecting and localizing a target is cognitively more effortful than the component tasks of either only detecting or localizing the same target. Many studies of dual-task performance have found that there is a cost in having to coordinate responses to two different questions concerning the same visual display. For example, even simple orientation detection tasks show a cost when performed concurrently with an attention-demanding task, as in the attentional blink procedure (Di Lollo, Kawahara, Zuvic, & Visser, in press; Joseph, Chun & Nakayama, 1997). Usually this dual-task cost is reflected in reduced performance in the second task rather than in the first task performed (e.g., Braun & Sagi, 1990; Di Lollo et al.). Di Lollo et al. hypothesize that this cost is an effect of the time needed to reconfigure the visual system to

perform the second task, particularly if the second task is more difficult. However, if the second task demands relatively few resources, then there will be a smaller cost for the second task.

Accordingly, there have been reports that two tasks can both be performed without any apparent cognitive costs. For example, detection or coarse localization of an orientation-defined texture element does not suffer as a secondary task when performed after a primary task of detection or coarse localization, orientation discrimination (Braun & Sagi, 1990), or letter identification (Braun & Sagi, 1991). Such an absence of a costs for the secondary task has been reported even though the primary tasks of orientation discrimination or letter identification are shown to be relatively demanding tasks. This suggests that the detection and coarse localization of oriented texture elements are not attentionally demanding tasks.

Thus, previous research implies that a second task in a dual-task context will suffer if necessary cognitive resources have been taken up by the first task. In the present study, if the localization task of Experiment 2 is cognitively demanding, then localization in a dual-task context will be worse than localization alone. However, if the localization task demands relatively few resources, then localization alone should be equivalent to localization in a dual-task context. Consistent with this, Braun & Sagi (1990, 1991) found no costs for a second task of coarse localization of oriented target elements within a dense 2D texture of Gabor elements. One question is whether the this lack of dual-task costs will extend to the present displays of oriented pexels in a dense, depicted 3D texture.

Another question is whether there are differences in the resources required to execute direct and indirect localization responses. As demonstrated in Experiment 2, direct localization actions were consistently more accurate than indirect actions. In the framework of action system theories (Milner & Goodale, 1995), the additional information used to guide the more precise direct localization responses is processed in the dorsal stream, separately from the information processed in the ventral stream. Therefore, if direct and indirect localization responses involve

some level of independent processing in different visual streams, it is possible that there will be differential costs for direct and indirect responses when performed after target detection. Costs would depend on which stream is involved in processing for the target detection task. Given that the first task in the present experiments is a detection response based on visual information that is consciously experienced, the detection task is likely performed in part by the ventral stream. The same would be true for the indirect localization response. However, the direct localization response may be based on information in the dorsal visual pathway that does not require access to the ventral stream. Thus, direct localization after detection may be spared some of the cognitive effort of indirect localization and result in smaller dual-task costs.

The purpose of this experiment was therefore to examine whether performance on the direct and indirect localization tasks of Experiment 2 were influenced by whether or not a detection response was required first. Performance in the dual-task conditions of Experiment 2 were therefore compared with three single-task control conditions: (a) detection alone (b) indirect localization alone and (c) direct localization alone.

Method

Participants

31 undergraduate (21 females, 10 male; mean age = 21.1 years) students participated in exchange for course credit. All reported normal or corrected-to-normal vision. 11 students participated in the detection-alone condition. 20 students participated in the localization-alone conditions (10 students each in the direct and indirect localization groups).

Stimuli, Design, and Procedure

Testing conditions and apparatus were the same as in Experiment 2. Detection and localization displays were the 3D texture arrays used in Experiment 2. All 7 values of theta (0°, 5°, 10°, 15°, 20°, 25°, 30°) and 11 values of phi (45°, 35°, 25°, 15°, 5°, 0°, -5°, -15°, -25°, -35°, -

45°) tested in Experiment 2 were tested in a single hour-long session for detection- and localization-alone conditions.

In the detection-alone condition, the pexel array remained on the screen until a response had been made, or seven seconds had elapsed. Correct detection responses were followed by a '+' symbol at the center of the screen, incorrect responses by a '-' symbol, and no response by a '0' symbol. At the end of each block of trials, a message displayed the percentage of errors for the most recent block of trials and participants were prompted to initiate the next block when ready. The localization array was never presented. Instructions were the same as in Experiment 2, except that the localization instructions were excluded.

The detection-alone condition consisted of 608 trials, divided into 8 blocks of 76 trials. The 76 target-present conditions were evenly distributed among one half of the trials, while the target-absent conditions were evenly distributed among the remaining half. The order of conditions was randomized across trials.

In the direct and indirect localization-alone conditions, a pexel array was presented for seven seconds during which the participant did not respond. Because localization was the only task, every pexel array contained a target; this was followed by a localization array. As in Experiment 2, no feedback was provided for localization responses. Instructions were the same as in Experiment 2, except that the detection instructions were excluded.

The direct and indirect localization-alone conditions each consisted of 304 trials, divided into 8 blocks of 38 trials. All trials contained a target-present display. The order of conditions was randomized across trials.

Results

Single- and Dual-task Detection

Target detection errors for both single- and dual-task contexts are shown in Figure 7 as a function of theta and phi.

For single-task detection, overall error rate was 37.7% for target-present trials and 15.1% for target-absent trials; this is very similar to the dual-task detection performance in Experiment 2. Target detection errors were analyzed with a mixed-design ANOVA that included one between-groups factor of Task Context (single or dual) and two within-participant factors of Theta (0° , 5° , 10° , 15° , 20° , 25° , 30°) and Phi Magnitude (0° , 5° , 15° , 25° , 35° , 45°). Detection responses of less than 100 ms were considered anticipatory responses and were coded as errors. This analysis revealed no main effect of Task Context on detection performance, $F(1, 61) = .03$, $p = .86$. Task Context also did not interact significantly with any of the within-participants factors. Thus, detection performance was not adversely affected by the dual-task context.

Sensitivity and criterion levels.

To examine the influence of single- versus dual-task contexts on detection performance, sensitivity and criterion levels for the dual-task detection performance in Experiment 2 were compared to the same for single-task detection performance in the present experiment. Sensitivity and criterion levels were each analyzed with an ANOVA that had a between-groups factor of Task Context (single or dual). Single-task sensitivity ($d' = 1.51$) did not differ significantly from dual-task sensitivity ($d' = 1.43$), $F(1, 61) = .47$, $p = .5$. Single-task criterion levels ($\beta = 2.35$) did not differ significantly from dual-task criterion levels ($\beta = 1.97$), $F(1, 61) = .96$, $p = .3$. Thus, detection sensitivity was not affected by whether or not it was followed by the localization task.

Single- and Dual-task Localization

Direct and indirect localization errors for both single- and dual-task contexts are shown in Figure 8 as a function of theta and phi.

Target localization errors were analyzed with a mixed-design ANOVA containing two within-participants factors of Theta (0° , 5° , 10° , 15° , 20° , 25° , 30°) and Phi (0° , 5° , 15° , 25° , 35° ,

45°) and two between-group factors of Task Context (single or dual) and Action Type (direct or indirect). This analysis revealed two main points.

1. Indirect localization performance suffered in the dual-task context. There was a significant main effect of Task Context for cell, $F(1, 68) = 7.25$, $p = .009$ and quadrant, $F(1, 68) = 6.63$, $p = .01$, levels of localization. When averaged over Action Type, errors for dual-task localization were consistently higher than errors for single-task localization. This cost for dual-task localization, however, was caused only by higher errors in dual-task indirect localization. Compared to the single-task context, there was no cost for direct localization performance in the dual-task context.

2. The size of the dual-task cost varied across Theta x Phi Magnitude combinations, as reflected in a significant Theta x Phi Magnitude x Task Context interaction at the fine localization level, $F(30, 2040) = 1.85$, $p = .004$. However, only 1 (theta = 0°, phi = 5°) of these 41 Theta x Phi Magnitude combinations reflected a case in which direct localization errors were higher than indirect localization errors. This supported the finding (point 1) that there was a consistent cost for indirect localization in a dual-task context. Task Context did not interact significantly with any other factors at any level of resolution.

Discussion

The goal of the present experiment was examine whether the dual tasks of target detection and localization in Experiment 2 imposed greater demands on performance compared with completing each of the tasks alone.

Comparison of detection errors, sensitivity, and criterion levels in single- and dual-task contexts revealed no differences between detection alone and detection followed by localization. This lack of dual-task cost for detection is not surprising in light of previous research that has shown that dual-task costs are not usually reflected in lower performance in the first task (Braun

& Sagi, 1990; Di Lollo et al., in press). When dual tasks reduce performance levels, this is usually most noticeable in the performance of the task that is performed second.

Comparison of localization errors in single- and dual-task contexts revealed an accuracy cost to localization when it followed detection, but only for the indirect localization response (spatially mapped keypress). In other words, performance for dual-task indirect localization was worse than for single-task indirect localization, but performance for dual-task direct localization was the same as for single-task direct localization. However, it should be noted that at the cell resolution of localization, there was a trend, though not significant, for single-task direct localization to be more accurate than dual-task direct localization (Figure 8). It is possible that there was an insufficient number of participants in the single-task localization condition to reveal a significant advantage of single-task over dual-task direct localization. Note that there were 10 participants each in the single-task direct and indirect localization conditions, but 26 participants each in the dual-task direct and indirect localization conditions. Thus, this advantage for direct localization in the single-task situation may become statistically significant if the statistical power of the experiment were increased. However, this does not detract from the main finding that the dual-task cost is reduced when the localization response is a direct action. While single-task direct localization was not significantly better than dual-task direct localization, single-task indirect localization was significantly better than dual-task indirect localization, even with the small number of participants in the indirect localization-alone condition. Thus, any dual-task cost of localizing after detecting a target is much attenuated for the direct localization response. Why is direct localization “easier,” or more accurate, than an indirect localization following detection? One possibility, outlined in the introduction, is that direct localization in the form of pointing to the screen does not require the same processing resources as those required of indirect localization.

Within the framework of Milner and Goodale's (1995) dual visual systems theory, it has been proposed that each system may have its own pool of resources available for processing visual information with little interference from the other system; however, very few studies exist to address this possibility (Norman, in press). In the present experiment, the detection task was merely a keypress response indicating target presence or absence, which is characteristic of the conscious perception functions attributed to the ventral stream. Thus, if the detection task in the present experiment is a ventral stream task, then resources may be unavailable to optimally perform a second ventral stream task of indirect localization, but may be available to optimally perform a dorsal stream task of direct localization. This may account for the present dual-task cost to indirect, but not direct localization.

One approach that may be useful in future studies seeking further support for the hypothesis that detection and direct localization share fewer cognitive resources than detection and indirect localization is to vary the resources available for each response more systematically. For example, if the detection and localization responses both depended on pointing (e.g., pointing to a sample of the seen target below the display before pointing to its screen location) there should be greater dual-task costs than in the present study. Conversely, making target localization rely even less on the dorsal stream (e.g., by requiring a vocal response) should also increase the dual-task costs of detection and localization.

General Discussion

The present study was designed to investigate human sensitivity to depicted 3D orientation of rectangular elements in dense texture arrays. Experiment 1 mapped out detection sensitivity to 3D orientation as a function of systematic variations in theta (left-right angular differences around upright Y-axis) and phi (rotation about the Y-axis). The results indicated that for 2D and 3D textures, a 15° difference in theta and 15° difference in angular orientation, respectively, were about equally effective in permitting accurate and effortless target detection.

The results of this experiment also indicated that target detection for the 3D textures was based on depicted 3D orientation, not 2D image orientation. This suggests that observers used a 3D mental representation of these displays and behaved as though the surfaces were receding in depth. At the very least, the results show that observers were not treating these depictions as flat texture gradients in the picture plane. How richly all aspects of the 3D world are represented in performance of this task is a question for future research.

Experiment 2 showed that response patterns for target localization, made immediately after the screen was erased, followed response patterns for target detection very closely. When targets were correctly detected, they were also accurately localized to a spatial resolution corresponding to a visual field quadrant. This is better spatial resolution than is predicted by some theories of target localization (e.g., Treisman & Gelade, 1980; Treisman & Sato, 1990). A second important result of Experiment 2 was that a localization response that involved pointing directly to the screen, before pointing to a spatially corresponding grid location below the screen, resulted in higher accuracy than simply pointing to the spatially corresponding grid location. This is evidence that some visually-guided actions can be used to enhance perceptual performance. A third finding of Experiment 2 was that even when observers failed to detect a target, they still had some information about target location as revealed through their localization responses made after the display had been erased.

Experiment 3 examined the extent to which performing two tasks together (detection followed by localization) resulted in poorer performance for either of the two tasks in comparison to performing the component tasks on their own. The results showed that the first task, target detection, was not at all influenced by subsequent requirements to localize the target. This is consistent with previous research on dual-task performance, which suggests that usually it is the second of two tasks that suffers (Braun & Sagi, 1990, Di Lollo et al., in press). Whether the second task, target localization, was influenced by a prior detection task depended on whether

the response involved a direct localization response (pointing to the screen before pointing to a corresponding grid location) or only an indirect response (pointing to a corresponding grid location). Only the indirect localization response showed significant costs in the dual-task condition. This suggests that detection and direct localization share fewer cognitive resources than do detection and indirect localization.

Implications for scientific visualization

The results of this study can be applied directly to the purposes of data visualization in a number of ways. First, the finding that human orientation discrimination is based on the parameters of 3D orientation, and not only 2D differences in angular orientation, confirms the hypothesis that more information can be conveyed to an observer by using an underlying 3D model than by using only a 2D model. This has several important implications, one of which is that using a 3D model increases the potential dimensionality of the data to be visualized. Instead of being restricted to a 2D space (x, y) in which angular orientation can vary from $0 - 90^\circ$, a 3D space can be used (x, y, z) and orientation can vary over both θ and ϕ . Furthermore, the same information that, in a 2D rendering, would occupy the entire screen, can now be presented in a foreshortened perspective view that still leaves room above the horizon and along the sides of the texture gradient for other visual information.

Second, these results indicate that in order to be readily detectable, borders between neighboring texture regions that are based on 3D orientation should involve θ differences of at least 15° . This means that there are effectively about seven readily discriminable steps of θ between 0° and 90° that can be used to code pexel orientation. This effective range will be reduced to some extent by the simultaneous variation of ϕ , but so long as the larger magnitudes of ϕ are avoided ($>15^\circ$), seven different values of θ will correspond to readily detectable differences in the underlying data.

Third, these results indicate that the localization of detected targets is accurate at least to the visual quadrant. This will be adequate for many purposes but it should be kept in mind that if an application depends on more precise visual localization, additional processing time will be required.

Fourth, the advantages found in this study for direct over indirect localization responses point to the importance of taking the response mode of the observer into account. These results remind us that the contents of the observer's visual experience is not independent of their goals in the visually-guided tasks they are undertaking. For some applications, the use of a touch screen may therefore contribute to more accurate perceptions, and do so with a smaller cognitive load than would be expected if responses are made through a more abstract spatial mapping. Given these preliminary results, it will be important in future research to more closely examine the relationship between the accuracy of visual information and the response mode. The larger lesson is that whenever the observer is confronted with dual tasks in a visualization application, it is important to try to combine tasks that share as few cognitive resources with each other as possible. The distinction between human vision for perception versus that for action seems to be a natural division, one that can be exploited effectively if it is understood by the designer of an application.

Several limitations of the present study should also be noted so that future research may be directed at the important questions still left unanswered. For instance, the present study examined target detection and localization involving a single target region against an otherwise homogeneous background of elements. They did not address the influence of background heterogeneity, either in the range of values that were used in the task-relevant dimension (3D orientation) or in the dimensions that were permitted to vary on non-task relevant properties (e.g., color, height, density, etc). Previous research has examined this question in displays

consisting of 2D textures (Healey, Booth, & Enns, 1996; Healey & Enns, 1998, 1999), but it remains to be seen how closely these results will generalize to the 3D environment.

A second limitation of the present study is that the displays were rendered in only two levels of color (black and white). In other words, these results were obtained using only the bare minimum required to depict the third dimension, and it is an open question whether the use of visually richer techniques would improve performance. The use of shading and shadows would permit the impact of different lighting models, surface reflectance functions, and surface shape models to be explored. The use of motion parallax and stereovision would permit the influence of stronger 3D cues to be explored.

Implications for the visual perception of target identity and location

The results of this study also are relevant to an understanding of how target identity and location are processed. As discussed in the introduction, there are two primary theories regarding the relationship between target detection and localization. Feature integration theory (Treisman & Gelade, 1980; Treisman & Sato, 1990) states that attention is necessary for correct target localization, but not necessary for the correct detection of distinctive targets. Texton theory (Julesz, 1984; Sagi & Julesz, 1985a), on the other hand, proposes that both target identity and location information can be available in the absence of attention, provided that the target is sufficiently distinctive in its texton composition. The results of the present study showed that when targets were correctly detected, they could also be localized to a fairly high level of accuracy (visual field quadrant). Furthermore, target detection and localization accuracy were tightly coupled. This suggests that target identity is not as separable from target location as is implied by feature integration theory.

However, it should also be acknowledged that a strong test of this hypothesis would require a more systematic exploration of the role of attention than was conducted in this study. The predictions of feature integration theory with respect to separability of target detection and

localization apply to features that can be detected preattentively, meaning without, or at least prior to, a focusing of attention to a specific region of the display. While this condition was almost surely met for the most distinctive targets that were tested (e.g., more than 20° of difference in theta), there was no independent measure of attentional load in the present study. The analysis of dual task performance in Experiment 3 strongly suggested that combining target detection with target localization produced very little task interference at all levels of target distinctiveness, but future studies should test a wider range of attentional manipulations (e.g., variations in set size) and secondary tasks that require more attention than the localization response (e.g., the attentional blink procedure).

Clearly the most intriguing result that should be followed up in future research is the finding that observers could correctly localize targets after failing to detect them. This finding strongly implies that some target location information is available not only without the focus of attention to a specific region of the display, but also without the awareness of the observer. This finding is not easy to accommodate within either feature integration theory or texton theory. Neither of these theories provide a way for accurate target localization to occur in the absence of correct target detection.

However, another way to look at this result is from the perspective of Milner & Goodale's (1995) theory of two visual systems. According to this theory, information about target identity is processed primarily in the ventral stream, which is specialized for the conscious perception of objects. Depending on task demands, responses concerning target location are guided by different streams. A task that involves acting directly on objects for the purposes of reaching, grasping, or locomotion will be guided by the dorsal stream, whereas a task that involves acting on the conscious representation of objects should be guided by the ventral stream. While the degree of interplay between the dorsal and ventral stream is still under considerable debate and research, each stream can function independently to some extent, as

evidenced by patients who have damage to the ventral stream but have spared dorsal stream functions (e.g., Goodale et al., 1991), and by patients who have damage to the dorsal stream but have spared ventral stream functions (e.g., Perenin & Vighetto, 1988).

Within this framework, the finding of correct direct localization in the absence of correct detection fits in well with the idea of two visual pathways that can carry out their functions independently. At the same time, the finding of correct indirect localization without correct detection does not fit as clearly into this theory, and is also not adequately explained by FIT or texton theory. One possibility is the indirect localization task of making a spatially mapped keypress is not entirely reliant upon ventral stream processing. Because this response requires an observer to direct an action towards a spatial representation of the target space, some dorsal stream processing might be involved in guiding the action. It is possible that the situations in which FIT and texton theory could apply are limited to ventral stream processing. To examine this possibility in future studies, the direct and indirect localization tasks could be compared to a localization response that involves no spatially directed action, such as a vocal response, to see if there is still evidence of correct localization after detection failure.

Concluding Remarks

The present set of experiments investigated human sensitivity to depicted 3D orientation in texture elements. Experiment 1 mapped out detection sensitivity to 3D orientation and revealed that detection performance for 3D targets was based on depicted 3D orientation, and not 2D image orientation. Experiment 2 showed that patterns of localization performance closely followed that for detection performance; if targets could be detected, they could also be localized to the correct visual field quadrant. Importantly, the direct localization response of pointing to an onscreen grid revealed more precise location information than an indirect response. An intriguing finding was that even when observers failed to detect a target, they still had a degree of information about target location as revealed through both direct and indirect localization

responses. Experiment 3 demonstrated that, compared to performance for the single-task context of either detection or localization alone, detection and localization did not suffer in a dual-task context, but indirect localization did. These results support the theoretical position of two separate visual streams that are specialized to process visual information for the purposes of perception or action. From a practical standpoint, these results have important implications for how 3D orientation can be used to advantage in the design of applications.

References

- Aglioti, S., DeSouza, J. F. X., & Goodale, M. A. (1995). Size-contrast illusions deceive the eye but not the hand. Current Biology, *5*, 679-685.
- Aks, D. & Enns, J. T. (1996). Visual search for size is influenced by a background texture gradient. Journal of Experimental Psychology: Human Perception and Performance, *22*, 1467-1481.
- Atkinson, J., & Braddick, O. J. (1989). "Where" and "what" in visual search. Perception, *18*, 181-189.
- Braun, J. & Sagi, D. (1990). Vision outside the focus of attention. Perception & Psychophysics, *48*, 45-58.
- Braun, J. & Sagi, D. (1991). Texture-based tasks are little affected by second tasks requiring peripheral or central attentive fixation. Perception, *20*, 483-500.
- Brenner, E., & Smeets, J. B. J. (1996). Size illusion influences how we lift but not how we grasp an object. Experimental Brain Research, *111*, 473-476.
- Card, S. K., Mackinlay, J. D., and Shneiderman, B. (Eds.). (1999). Readings in information visualization: Using vision to think. San Francisco, CA: Morgan Kaufmann.
- Di Lollo, D., Kawahara, J., Zuvic, S. M., & Visser, T. A. W. (in press). The preattentive emperor has no clothes: A dynamic redressing. Journal of Experimental Psychology: General, *130*.
- Enns, J. T., & Rensink, R. A. (1990a). Influence of Scene-Based Properties on Visual Search. Science, *247*, 721-723.
- Enns, J. T., & Rensink, R. A. (1990b). Sensitivity to three-dimensional orientation in visual search. Psychological Science, *1*, 323-326.
- Enns, J. T. & Rensink, R. A. (1991). VSearch Color: Full-color visual search experiments on the Macintosh II. Behavior Research Methods, Instruments, & Computers, *23*, 265 – 272.

Fecteau, J. H., Chua, R., Franks, I., & Enns, J. T. (2001). Visual awareness and the on-line modification of action. Canadian Journal of Experimental Psychology, *55*, 106-112.

Green, M. (1992). Visual search: Detection, identification, localization. Perception, *21*, 765-777.

Goodale, M. A., Milner, A. D., Jakobson, L. S., & Carey, D. P. (1991). A neurological dissociation between perceiving objects and grasping them. Nature, *349*, 154-156.

Healey, C. G., & Booth, K. S., & Enns, J. T. (1996). High-speed visual estimation using preattentive processing. ACM Transactions on Computer-Human Interaction, *3*, 107-135.

Healey, C. G., & Enns, J. T. (1998). Building perceptual textures to visualize multidimensional datasets. Proceedings of IEEE Visualization '98, 111-118.

Healey, C. G., & Enns, J. T. (1999). Large datasets at a glance: Combining pexels and colors in scientific visualization. IEEE Transactions on Visualization and Computer Graphics, *5*, 145-167.

Jackson, S. R., & Shaw, A. (2000). The Ponzo Illusion affects grip-force but not grip-aperture scaling during prehension movements. Journal of Experimental Psychology: Human Perception and Performance, *26*, 418-423.

Joseph, J. S., Chun, M. M., & Nakayama, K. (1997). Attentional requirements in a 'preattentive' feature search task. Nature, *387*, 805-807.

Julesz, B. (1984). A Brief Outline of the Texton Theory of Human Vision. Trends in Neuroscience, *7*, 41-45.

Klotz, W., & Neumann, O. (1999). Motor activation without conscious discrimination in metacontrast masking. Journal of Experimental Psychology: Human Perception and Performance, *25*, 976-992.

Kolb, F. C., & Braun, J. (1995). Blindsight in normal observers. Nature, *377*, 336-338.

- Milner, A. D., & Goodale, M. A. (1995). The visual brain in action. Oxford: Oxford University Press.
- Norman, J. (in press). Two visual systems and two theories of perception: An attempt to reconcile the constructivist and ecological approaches. Behavioral and Brain Sciences, 24.
- Perenin, M. -T., & Vighetto, A. (1988). Optic ataxia: A specific disruption in visuomotor mechanisms. I. Different aspects of the deficit in reaching for objects. Brain, 111, 643-674.
- Pöppel, E. , Held, R. & Frost, D. (1973). Residual function after brain wounds involving the central visual pathways in man. Nature, 243, 295-296.
- Sagi, D, & Julesz, B. (1985a). Detection versus discrimination of visual orientation. Perception, 14, 619-628.
- Sagi, D, & Julesz, B. (1985b). "What" and "where" in vision. Science, 228, 1217-1219.
- Schmidt, T. (in press). The finger in flight: Real-time motor control by visually masked color stimuli. Psychological Science.
- Shapiro, K. L., Arnell, K. M., & Raymond, J. E. (1997). The attentional blink. Trends in Cognitive Science, 1, 291-296.
- Shelton, A. L., & McNamara, T. P. (2001). Visual memories from nonvisual experiences. Psychological Science, 12, 343-347.
- Stevens, K. A. (1983). Slant-tilt: The visual encoding of surface orientation. Biological Cybernetics, 46, 183-195.
- Treisman, A., & Gelade, G. (1980). A feature-integration theory of attention. Cognitive Psychology, 12, 97-136.
- Treisman, A., & Gormican, S. (1988). Feature analysis in early vision: Evidence from search asymmetries. Psychological Review 95, 15-48.
- Treisman, A. & Sato, S. (1990). Conjunction search revisited. Journal of Experimental Psychology: Human Perception and Performance, 16, 459-478.

Weiskrantz, L., Warrington, E. K., Sanders, M. D. & Marshall, J. (1974). Visual capacity in the hemianopic field following a restricted cortical ablation. Brain, 97, 709-728.

Wolfe, J. M., Friedman-Hill, S. R., Stewart, M. I., & O'Connell, K. M. (1992). The role of categorization in visual search for orientation. Journal of Experimental Psychology: Human Perception and Performance, 18, 34-49.

Table 1

Experiment 1: Partial Correlations Between Predictors Of Detection Errors in 3D TextureGradients

Predictor	1	2	3	4	5
1. Error rate	--	.037	-.534	.276	.114
2. 2D Tilt		--	.840	-.526	-.454
3. 3D Theta			--	.566	.430
4. 3D Phi				--	-.264
5. 3D Phi Direction					--

Figure Captions

Figure 1. (a) Example of 2D texture with target element present. (b) Example of 3D texture gradient with target element present. (c) Components of depicted 3D orientation. Theta (θ) describes left-right angular differences around the upright Y-axis. Phi (ϕ) describes rotation about the Y-axis.

Figure 2. (a) Detection errors on target-present 2D textures, plotted as a function of angular difference between background and target pexel orientation. (b) Detection errors on target-present 3D texture gradients, plotted as a function of 3D Theta and 3D Phi.

Figure 3. (a) Matrix of possible target locations presented for both direct and indirect localization responses. (b) Response keys for indirect localization. Keys are spatially mapped to the localization array of Figure 3a.

Figure 4. Direct and indirect target localization errors at cell, quadrant, and hemifield resolutions. Errors are shown as a function of (a) Theta (b) Phi Magnitude. Chance levels of localization for each resolution are indicated by the dashed line.

Figure 5. Detection accuracy plotted against direct and indirect localization accuracy for (a) cell, (b) quadrant, and (c) hemifield resolutions of localization. Equations for line of best fit are shown beneath each plot.

Figure 6. Direct and indirect localization errors for trials where observers failed to detect the target. Errors at cell, quadrant, and hemifield resolutions are shown as a function of (a) Theta (b) Phi Magnitude. Chance levels of localization for each resolution are indicated by the dashed line.

Figure 7. Single- and dual-task detection errors as a function of (a) Theta (b) Phi Magnitude.

Figure 8. Direct and indirect localization errors for single- and dual-task conditions, plotted as a function of (a) Theta (b) Phi Magnitude. Chance levels of localization for each resolution are indicated by the dashed line.

(a)

(b)

(c)

Figure 1. (a) Example of 2D texture with target element present. (b) Example of 3D texture gradient with target element present. (c) Components of depicted 3D orientation. Theta (θ) describes left-right angular differences around the upright Y-axis. Phi (ϕ) describes rotation about the Y-axis.

(a)

(b)

Figure 2. (a) Detection errors on target-present 2D textures, plotted as a function of angular difference between background and target pixel orientation. (b) Detection errors on target-present 3D texture gradients, plotted as a function of 3D Theta and 3D Phi.

(a)

(b)

Figure 3. (a) Matrix of possible target locations presented for both direct and indirect localization responses. (b) Response keys for indirect localization. Keys are spatially mapped to the localization array of Figure 3a.

(a)

(b)

Figure 4. Direct and indirect target localization errors at cell, quadrant, and hemifield resolutions.

Errors are shown as a function of (a) Theta (b) Phi Magnitude. Chance levels of localization for each resolution are indicated by the dashed line.

Figure 5. Detection accuracy plotted against direct and indirect localization accuracy for (a) cell, (b) quadrant, and (c) hemifield resolutions of localization. Equations for line of best fit are shown beneath each plot.

(a)

(b)

Figure 6. Direct and indirect localization errors for trials where observers failed to detect the target. Errors at cell, quadrant, and hemifield resolutions are shown as a function of (a) Theta (b) Phi Magnitude. Chance levels of localization for each resolution are indicated by the dashed line.

(a)

(b)

Figure 7. Single- and dual-task detection errors as a function of (a) Theta (b) Phi Magnitude.

(a)

(b)

Figure 8. Direct and indirect localization errors for single- and dual-task conditions, plotted as a function of (a) Theta (b) Phi Magnitude. Chance levels of localization for each resolution are indicated by the dashed line.