GEOMETRIC FIGURE-GROUND CUES OVERRIDE STANDARD DEPTH FROM ACCRETION-DELETION

by

Ö. DAĞLAR TANRIKULU

A thesis submitted to the

Graduate School—New Brunswick

Rutgers, The State University of New Jersey

in partial fulfillment of the requirements

for the degree of

Master of Science

Graduate Program in Psychology

Written under the direction of

Prof. Manish Singh

and approved by

New Brunswick, New Jersey

January, 2014

ABSTRACT OF THE THESIS

Geometric figure-ground cues override standard depth from accretion-deletion

By Ö. DAĞLAR TANRIKULU

Thesis Director:

Prof. Manish Singh

Accretion/deletion, where a moving texture appears from (i.e. accretion) or disappears (i.e. deletion) at a boundary, is widely considered as a reliable cue to surface depth ordering, with the accreting or deleting surface interpreted as behind the adjoining surface. However, Froyen, Feldman, and Singh (2013) showed that when accretion-deletion occurs on both sides of a contour, some accreting-deleting regions can also be perceived as in front and as self-occluding due to rotation in 3D. In this study we ask whether geometric figure-ground cues can override the standard "depth from accretion-deletion" interpretation even when accretion-deletion takes place only on one side of a contour. We used two tasks: a relative depth task (front/back), and a motion classification task (translation/rotation). We conducted two experiments, where only one set of alternating regions contained moving texture; the other set was static. In such displays the standard accretion-deletion account would unambiguously assign farther depth to the moving regions. However, when the moving regions were convex or symmetric, they tended to be perceived as figural, and rotating in 3D (with convexity > symmetry). In the second experiment, different motion directions were given to the moving regions (hence weakening motion-based grouping) and this further weakened the traditional accretion-deletion interpretation. Our results show that the standard "depth from accretion-deletion" interpretation is overridden by geometric cues to figure-ground. When this happens, the accreting-deleting surface is perceived as self-occluding due to rotation in 3D. Overall, the results demonstrate a rich interaction between accretion-deletion, figure-ground, and structure-from-motion that is not captured by existing models of depth from motion.

Acknowledgements

This research was supported by NSF DGE 0549115 (IGERT: Interdisciplinary Training in Perceptual Science) and by NIH EY021494 (MS, JF). I would like to express my very great appreciation to Professor Manish Singh for his great mentorship throughout the whole time I have been here. I also would like to offer my special thanks to Professor Jacob Feldman, for his many helpful comments and discussions, and to Vicky Froyen, who was the person to whom I run to get help whenever I had a problem at each stage of this project. I am also grateful to Professor Thomas Papathomas for his valuable comments on this paper. I would like to thank all the members of the Visual Cognition Lab for their helpful discussions. I would also like to offer my thanks to Lorilei Alley for her help in the editing process of this paper. Lastly, I am extremely grateful to all of the faculty members of Rutgers Cognitive Science Center for creating a wonderful and an encouraging academic atmosphere that formed the basis of my motivation for this project and also for the future projects that I am going to be involved in.

Dedication

This dissertation is dedicated to my supportive parents and to my brother without whom I would never have ended on this side of the Atlantic Ocean.

Table of Contents

List of Figures

1. Geometric figure-ground cues override accretion-deletion

1.1 Introduction

One of the most crucial tasks that the visual system has to overcome is to estimate a three-dimensional layout from a two-dimensional (2D) retinal image. The first step in this task is to separate figure from background in an image which enables us to segment and define objects in a scene. The visual system uses various cues in order to obtain this figure-ground organization from a 2D projection. Figure-ground organization requires determining which regions own which contours in an image and assigning "figure" and "ground" status to those regions accordingly. The region that has the figural status is shaped by this contour while the ground region is perceptually unbounded and amodally continues behind the figural region (Nakayama, He, & Shimojo, 1995).

There are numerous cues that the visual system exploits in order to achieve figure-ground assignment. An important class of figure-ground cue is defined by geometric cues, i.e. cues involving the static geometry of the contours. Many different geometric cues that tend to promote figural status have been proposed so far, such as symmetry (Kanizsa & Gerbino, 1976), convexity (Metzger, 1953; Kanizsa & Gerbino, 1976), parallelism (Metzger, 1953), axiality and part structure (Hoffman & Singh, 1997; Froyen, Feldman, & Singh, 2010), and many others (see Wagemans et al. (2012) for a review).

Besides static cues, there are also dynamic cues to figure-ground assignment where motion provides information about the depth order. One powerful cue to depth is accretion/deletion of textured regions (Kaplan, 1969; Thompson, Mutch, & Berzins, 1985; Mutch & Thompson, 1985). When a translating texture deletes at or accretes from a boundary, it is perceived as if it was disappearing or appearing behind an occluding on the other side of the boundary. This in turn generates a vivid sense of figure-ground. This accretion/deletion of textured surfaces is often described as a self-sufficient visual cue that can unambiguously assign depth order to surfaces (Kaplan, 1969; Gibson, Kaplan, Reynolds, & Wheeler, 1969; Thompson et al., 1985; Mutch & Thompson, 1985; Niyogi, 1995; Howard & Rogers, 2002; Hegdé, Albright, & Stoner, 2004). It has even been shown to resolve the ambiguity of direction of rotation in orthogonally projected spheres. This is achieved by accretion/deletion overriding another important depth cue, motion parallax (Ono, Rogers, Ohmi, & Ono, 1988).

Even though accretion/deletion has been considered to be a sufficient cue to depth order that fully determines the ground side, there is also a possible perceptual organization where the accreting/deleting surface is figural (i.e. closer to the observer). One example of this depth order reversal is seen when the accreting-deleting surface is perceived as a rotating column. The texture disappears at the occluding boundary due to self-occlusion of a rotating 3D object. As also seen in Figure 1.1, this 3D arrengment is also compatible with an accreting-deleting texture. If the region that has textural motion is perceived in front, the deletion (or accretion) of the texture can no longer be explained by the occlusion of the static region, which is now perceived as the background. The only possible explanation for the disappearing texture becomes self-occlusion. This shows that the classical accretion/deletion cue has the possibility of producing two different percepts with opposite depth-order assignments. This possibility has occasionally been noted (Kaplan, 1969; Yonas, Craton, & Thompson, 1987; Mutch & Thompson, 1985; Thompson et al., 1985), but has not been incorporated into standard accounts of accretion/deletion.

According to the traditional accretion/deletion literature, the central region that has the textural motion in Figure 1.1 should be perceived as translating behind the static outer region. However, a 2D projection of a column rotating in depth is also compatible with the accreting-deleting region. If the region that has textural motion is perceived in front, the deletion (or accretion) of the texture can no longer be explained via occlusion by the static region, which is now perceived as the background. The only possible explanation for the disappearing texture becomes self-occlusion.

Recently, several studies have focused on the ambiguity caused by accretion/deletion. Kromrey, Bart, and Hegdé (2011) showed that accretion/deletion needs additional information about the occlusion border in order to unambiguously assign depth order. In their stimulus, where an enclosed region of translating random dot-texture was surrounded by random-dot texture, accretion-deletion created two possible

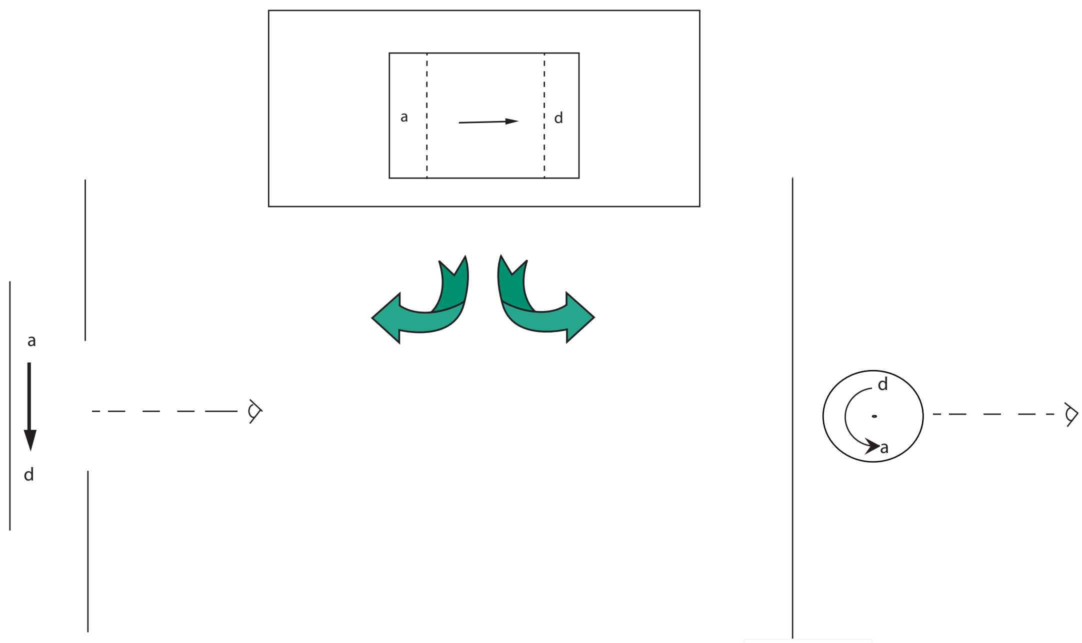

Figure 1.1: On the top panel, the frontal projection of an accreting surface is shown. Side "a" refers to the place where texture accretes and "d" refers to the side where texture deletion is happening. The two arrows indicate the overhead views of the two possible 3D arrangement with different depth-order assignments that are consistent with the frontal view of the accreting and deleting surface.

interpretation, in which the central region is either seen as in front (when surrounding texture is flickering) or farther away from the surrounding region (when it is static). According to their results, only when the delineation of the border between the center and the surround regions was made easier by segmentation cues (such as making the surrounding region static, or increasing the contrast of the dots between the two regions), the interpretation that is consistent with the traditional accretion/deletion cue (i.e. seeing the translating texture farther away) was favored.

Froyen et al. (2013) introduced a new phenomenon where 3D columns rotating in depth are perceived as a result of an interaction between the shape of the border and the accretion/deletion cue. In their experimental stimulus, the ambiguity between the two different interpretation of accretion/deletion mentioned above was increased by introducing accretion/deletion on both sides of a border. This created a bi-stable, multiple-region, figure-ground stimulus (Figure 1.2). In such a stimulus, geometric figure-ground cues (e.g. convexity) resolved the ambiguity so that the regions that are perceived as figural are also perceived as 3D volumes rotating in depth.

The studies mentioned above indicate that shape of the border interacts with accretion/deletion cue. However there aren't any studies that systematically focus on this interaction, specifically, on conditions under which this interaction causes inverted depth-order percepts. An account of this interaction and the depth-order inversion that it causes might require us to re-define the phenomenon of accretion/deletion as a cue to relative depth. In the experiments below, we examine the interaction between two geometric figure-ground cues and accretion/deletion, in which 3D structures are perceived by the subjects even though the linear dot textural motion is inconsistent with 3D structure from motion.

In our experiments, we used multiple-region figure-ground stimuli similar to the one in Figure 1.2. In our crucial experimental conditions, either the odd or the even regions are made static. This would make the stimulus similar to the classical accretion/deletion stimulus, which is supposed to unambiguously assign depth order. We also introduce symmetry or convexity to one set of regions in order to examine the interaction between these geometric cues and accretion/deletion. In this way we will

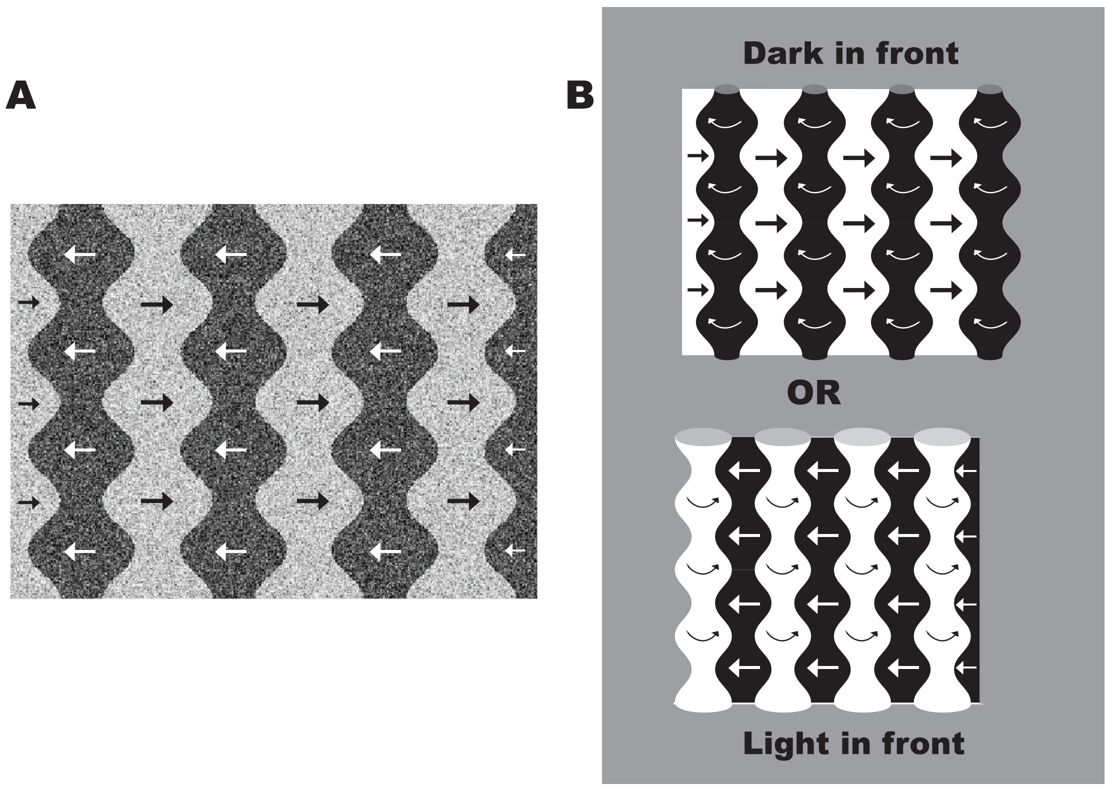

Figure 1.2: Display setup and phenomenology: A. The displays were created by adding motion in one direction to odd regions and in the other direction to even regions in classical figure-ground displays. B. This could yield one of two percepts depending on which one was perceived as figural. The black ones were perceived as rotating in front of a white background which was seen as sliding behind them, or vice versa. Adapted from: "Rotating columns: relating structure-from-motion, accretion/deletion, and figure/ground," by V. Froyen, J. Feldman, and M. Singh, 2013, Journal of Vision, 13(10)

test whether accretion/deletion is able to unambiguously assign depth-order and also examine its interaction with the shape of the border where the texture is being accreted and/or deleted.

Two different tasks are given to the subjects with the same stimulus. The first task requires them to indicate whether they see the target region, which is marked with arrows, in front relative to its adjacent region, which is a classical question to measure figure-ground. The other task requires subjects to indicate whether they see a rotational or a translational motion in the target region that has textural motion. In this way, we also aim to understand the relationship between perceiving a region in front and perceiving the moving region as a 3D column rotating in depth.

1.2 Experiment 1

In this experiment, the interaction between geometric figure-ground cues and accretion/deletion cue was examined by combining them in various conditions. In the cue competition condition, the two cues are introduced to the same region such that while geometric cues (convexity and symmetry in this case) were suggesting figural status in a certain region, accretion/deletion was suggesting the opposite. In the same way, in the cue cooperation condition the two cues were introduced to different regions. There were also conditions, in order to replicate the results of Froyen et al. (2013), where accretion/deletion cues were present in every region. Subjects performed two different tasks on the same stimuli. Subjects' responses on depth order (figure-ground task) and their interpretation of the motion (rotation task: whether there is a rotational or a translational motion) were examined.

1.2.1 Method

Participants

Thirteen Rutgers University students who were naive to the purpose of the experiment participated in the study. Nine of them were paid to participate and the other four participated for course credit.

Stimuli

The stimulus for this experiment consisted of eight alternating black and white vertical stripes. The stimuli were 7.29 arc high and 9.68 arc wide. Either the odd or the even regions were given one of two configural figure-ground cues, convexity and symmetry (Figure 1.3). Convex regions were also made symmetric so that the regions could be interpreted as surface of revolution.

As in Froyen et al. (2013), the convex regions were created by using a series of half circles with random radii as a boundary and then mirroring it on the other side of the region. Symmetric contours were created by using B-spline functions with 20 control points. The control points were set so that the sum of signed curvature was kept at zero along each boundary, and the area of each region was constant. This was done separately for every convex and symmetric region in a display, so that no two regions were the same in terms of shape in a single display.

On half of the trials the odd regions were dark and the even regions were light colored, and on the other half it was vice versa (i.e. counterbalanced and crossed with other factors). The phase of the stimuli (e.g. whether the rightmost part of the display starts with a convex/symmetric or a concave/asymmetric region) was also counterbalanced and crossed with other factors by mirroring the displays about their vertical middle axis.

To these stimuli, textural motion was added as a moving random dot texture. For the dark regions, the dot texture was sampled from a beta distribution with parameters $[\alpha = 6, \beta = 2]$ that resulted in a dark texture with sparsely scattered light pixels. The light regions included random dot texture sampled from a beta distribution with parameters $[\alpha = 2, \beta = 6]$, which resulted in a light texture with sparsely scattered dark pixels. The size of a single pixel was 1.47 arcmin by 1.47 arcmin. The texture could move either to the right or to the left, and it was implemented as follows. For the rightward motion, in each frame t the texture columns [2, N] were taken from texture columns $[1, N-1]$ in frame $t-1$, and the first column in frame t was resampled in the manner described above. The implementation was the same for the leftward motion.

This procedure was repeated at a rate of 40 frames/sec, resulting in a motion with a speed of 0.98 DVA/sec.

We created three different types of displays in terms which regions contained moving texture. In one type, all regions had textural motions where odd and even regions move in opposite directions (i.e. the third column in Figure 1.3). In the second type of display, the regions where the configural cues are introduced (the convex and/or symmetric regions) were made static, and the other regions that do not have these cues had consistent motion either to the left or right (i.e. the second column in Figure 1.3). The third type of display was the opposite of the second type, where the concave and asymmetric regions contained static texture, while the convex and symmetric regions had consistent motion either to the left or right (the first column in Figure 1.3). The direction of motion was counterbalanced and crossed with other factors for all displays.

Design and Procedure

Subjects sat at 85 cm from a 21″ CRT monitor (144 Hz, 1024pxl x 768pxl) connected to a Windows XP PC. The experiment was presented using psychtoolbox on MATLAB. The experiment included two different tasks. One of them was the "figure-ground task", where subjects were asked which one of the two indicated adjoining regions was in front of the other one. The other task was the "rotation task", where subjects were asked whether they see a rotation or translation in the indicated region. In both tasks, each trial started with 800 msec of pre-mask, followed by 800 msec of the pre-mask with a fixation cross added to it. The mask was created by randomly generating frames of figure-ground displays with unbiased (in term of geometric cues) contours, and then overlaying them on top of each other.

After the mask, the experimental display with moving textures was shown for 3 seconds. In the last two seconds of these three seconds, two regions (for the figure-ground task) or one region (for the rotation task) were indicated by triangle-shaped arrows that appeared at the top and the bottom of the target region (5 pixels away from the display; see Figure 1.3 and 1.4). For the figure-ground task, the target regions were chosen among four central regions. There were three different locations that the arrows

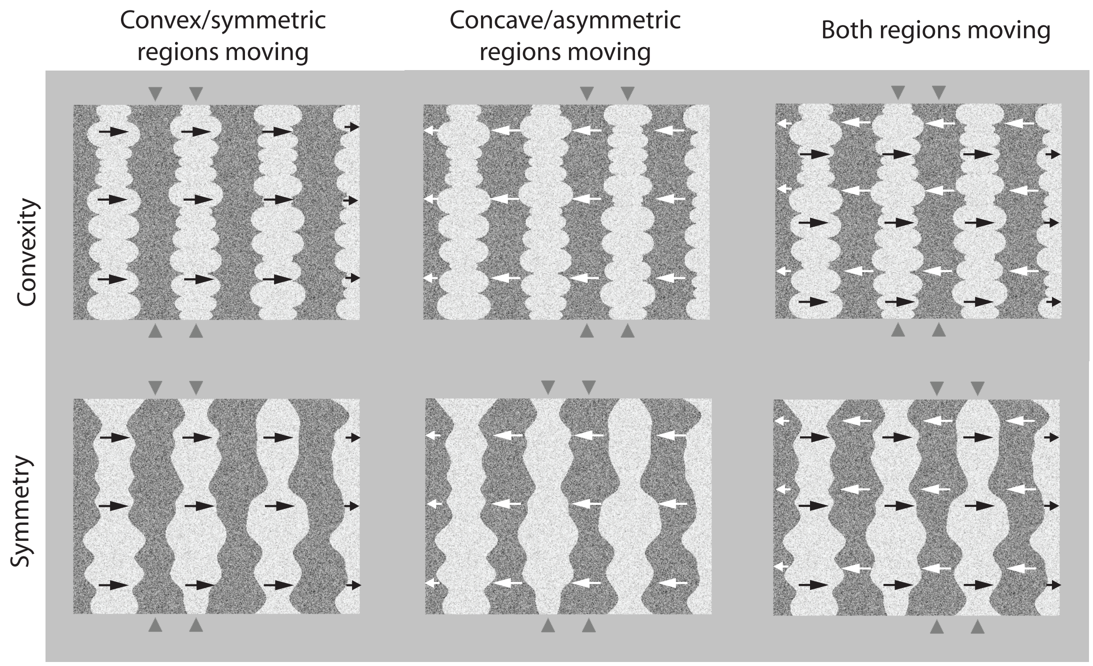

Figure 1.3: The six conditions used in the figure-ground task of Experiment 1

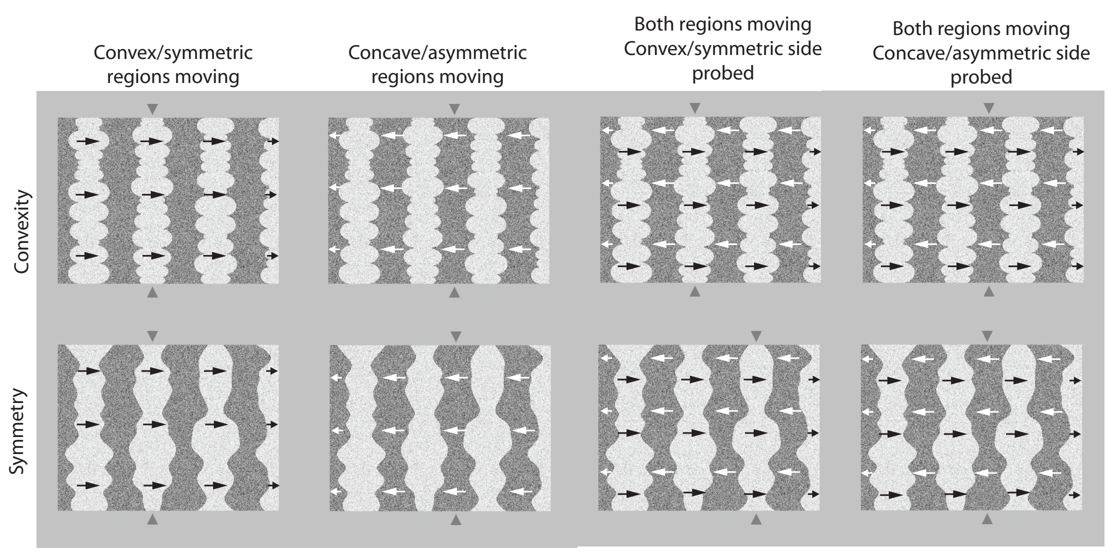

Figure 1.4: The eight conditions used in the rotation task of Experiment 1

could appear (upper row of Figure 1.3). For the rotation task, the question (whether rotation or translation is perceived) was asked regarding a single region. There were only three different locations where the arrow could appear, which were the regions that have textural motion. At the beginning of each trial, the exact location of the arrows was randomly determined for that trial only.

After 3 seconds (1 second without arrows, and 2 seconds with arrows), the subjects were presented with a post-mask that was identical to the pre-mask for a minimum of 800ms. Once this post-mask was presented, the subjects were asked the experimental question with respect to their current task. The subjects responded by means of keys on the keyboard.

For the figure-ground task, subjects ran 96 experimental trials split into two blocks, i.e. 2(geometric cues: convexity/symmetry) x 3(motion in: convex/symmetrical region; in concave/asymmetric regions; in both regions) x 2(luminance: dark/bright) x 2(phase) x 2(direction of motion: right/left) x 2(repetition). For the rotation task, subjects ran 128 experimental trials split into two blocks. The experimental conditions were the same for the motion interpretation task, except that the motion condition included four levels (rather than three levels seen on Figure 1.3). The reason for this extra level is that for the condition where all regions on the display have motion, either the convex/symmetric or the concave/asymmetric region could be probed (the last two columns in Figure 1.4). Including the two tasks, there were a total of 224 experimental trials. All conditions were counterbalanced for each subject, and trials were randomized for each subject separately. The order in which the subjects received the two tasks was also counterbalanced across subjects. Before the experimental trials began, 16 practice trials were run in order to acquaint the subject with the displays and the tasks. It took approximately 50 minutes for subjects to complete the whole experiment.

1.2.2 Results and Discussion

Figure 1.5 shows the results plotted as the proportion of times subjects reported seeing the specified regions as in front (for the figure-ground task) or as rotating (for the rotation task). In what follows, we focus on the two essential factors, i.e. the geometric cue, and the location of motion. Other factors were found not to yield any main effect.

The first row in Figure 1.5 shows the results for the figure-ground task, and the second row is for the rotation task. The graphs in the left column are for the conditions where either the convex/symmetric or the concave/asymmetric regions contained textural motion, whereas the other regions are static. These conditions correspond to the first two columns depicted in Figures 1.3 and 1.4. The graphs in the right column of Figure 1.5 are for the conditions where every region has textural motion, where odd and even regions are moving in opposite directions. These conditions correspond to the rightmost column in Figure 1.3, and the last two columns of Figure 1.4. The graphs in the left column depict the results in the crucial conditions regarding our research questions. The graphs in the right column can be considered a replication of Froyen et al. (2013).

Even though there are a total of six experimental conditions for the figure-ground task (see Figure 1.3), there are eight different bars in the graphs for that task. In the top-right graph in Figure 1.5, the red bars show the proportion of times the convex/symmetric regions were seen as figural. The turquoise bars show the proportion of times the concave/asymmetric regions were seen as figural. The turquoise bars are calculated by subtracting the proportions shown by the red bars from 1. This is because if the convex/symmetric region is not seen as figural, then the adjacent concave/asymmetric region is seen as figural. The turquoise bars are added in order to make the comparison between the two tasks easier. In this way, each experimental condition in the rotation task has a corresponding condition in the figure-ground task.

T-test analyses were performed (for both task responses) in order to see whether the proportions reported on Figure 1.5 were significantly different than 0.5, i.e. chance level. The proportions that were significantly different than chance level were shown with their corresponding p-values via star symbols on Figure 1.5. (For all the significant differences obtained: $t_{max} = 30.99$, $t_{min} = 3.3$, $df = 12$, $p < 0.05$. The maximum and minimum values were reported in absolute values).

For the figure-ground task responses, a multilevel logistic regression showed

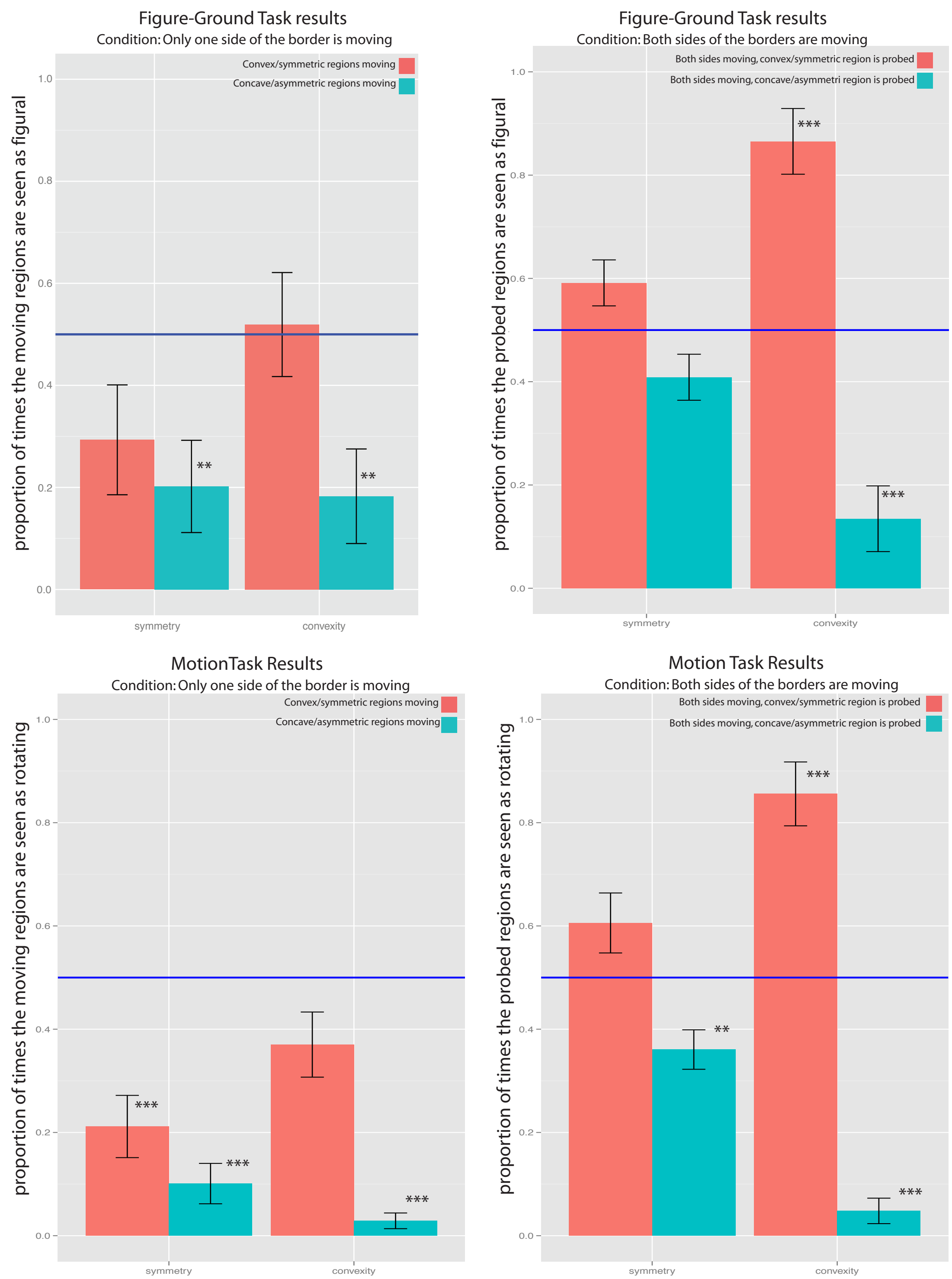

Figure 1.5: Results of Experiment 1. Error-bars represent $+/-1SE$ as computed between subjects. The blue line shows the chance level, i.e. where the proportion equals to 0.5. The stars (*) represent the proportions that are significantly different than chance level. The number of stars indicate the significance level: '***' $p < 0.001$; '**' $p < 0.01$; '*' $p < 0.05$

significant main effects of both geometric cue and location of motion, when compared to an unconditional means model (containing only an intercept) by means of a likelihood test (Geometric cue: $LR = 39$, $df = 3$, $p < 0.001$; location of motion: $LR = 553.7$, $df = 7$, $p < 0.001$). It was found that the regions that contained the geometric cues were more likely to be seen as figural when the geometric cue was convexity compared to when it was symmetry. This was expected for two reasons. First, it is already known that symmetry is a relatively weak cue compared to convexity (Kanizsa & Gerbino, 1976; Froyen, Tanrikulu, Singh, & Feldman, 2012, 2013). Secondly, convex regions were also symmetric, which causes convex regions to exert extra influence. Tukey pairwise comparisons, done between the three different conditions of the location of motion, revealed the folowing effects. When regions on both sides of a boundary had motion (i.e. Column 3 in Figure 1.3), subjects were more likely to see the indicated moving region as figural compared to the condition where only convex/symmetric regions were moving (i.e. Column 1 in Figure 1.3) ($p < 0.05$), and also to the condition where only concave/asymmetric regions were moving (i.e. Column 2 in Figure 1.3) ($p < 0.001$). Subjects were also more likely to see the moving region as figural in the condition where only the convex/symmetric regions were moving, compared to the condition where only concave/asymmetric regions were moving. ($p < 0.001$). The regression analysis also revealed a significant interaction effect for the location of motion and geometric cue. It was observed by comparing a model that includes geometric cue and location of motion factors as fixed effects to a model that also includes the two factors and plus the interaction term as fixed effects ($LR = 33.52$, $df = 13$, $p < 0.01$). The interaction is seen when the effect of location of motion in the convexity condition is compared to its effect in the symmetry condition. Whether placing accretion/deletion onto the convex/symmetric region or onto the concave/asymmetric region makes a big difference when the geometric cue is convexity, but not that much when it is symmetry.

We also found a significant interaction effect between the color factor (i.e. whether the target region is dark or light) and the location of motion, by comparing a model that includes color and location of motion as fixed effects to a model that also includes the two factors and plus the interaction term as fixed effects. ($LR = 26.2$,

$df = 13$, $p < 0.05$) Further investigation reveled that this interaction occurs because of the difference between the condition where both sides of the border had motion and the conditions where only one side had motion. When just one side of the border had motion, whether the moving region is light or dark did not have that much influence on subject's responses. However, when there was motion on both sides of the border, there appears an influence of color (i.e. when the indicated region is light, the proportion of the time the indicated region is seen as figural was 0.65, whereas it was 0.80 when it was dark). This result is consistent with our earlier studies (Froyen et al., 2012). When cues to figure-ground are not enough to solve the figure-ground separation, subjects' biases toward different colors manifest themselves and influence their responses. There is no structured effect of color on figure-ground responses; while some people have a bias towards dark regions, others have a bias towards light regions. This is also what we have observed in this experiment. Introducing motion on both sides of the border increased ambiguity for the figure-ground interpretation, and as a result subjects' color biases had an influence on their responses for this condition. In order to avoid any possible confounding effect of the color factor, the logistic regression analyses reported above were repeated, but this time we controlled for color. This was done by including the color factor both into the "null" and "alternative" models in all the likelihood ratio tests, then by checking whether the geometric cue and location of motion introduce a significant expansion to a model that already includes the color factor. The results obtained after controlling for color were fairly similar to ones reported above.

The same multilevel logistic regression analysis was also done for the rotation task responses. The analysis revealed a significant main effect for the location of motion, when compared to an unconditional means model ($LR = 580.25$, $df = 12$, $p < 0.001$), but it did not yield a significant effect of the geometric cue. However, there was a significant interaction effect between the two factors which is revealed by comparing a model that includes the location of motion factor as the fixed effect to a model that includes the location of motion factor plus the interaction term between it and the geometric cue ($LR = 178.5$, $df = 30$, $p < 0.001$). Tukey pairwise comparisons were done between the four different conditions of location of motion. The order of the four conditions, in terms of the proportion of times the moving/indicated region is perceived as rotating is as follows: The highest proportion was obtained in the condition where both sides of the border were moving while the convex/symmetric region was indicated. The cue-competition condition (i.e. only convex/symmetric region is moving) followed it as the second. In the third position there was the condition where both sides were moving and the concave/asymmetric region was indicated. The lowest proportion was obtained in the cue-cooperation condition where only the concave/asymmetric region was moving. All the pair-wise comparisons were significantly different from each other with $p < 0.001$, except the comparison between the cue-competition condition and the condition where both regions were moving while the concave/asymmetric region is indicated. The interaction effect between the location of motion and the geometric cue can be seen when the effect of the the location of motion is examined in different conditions of the geometric cue. For example, if the bottom-left graph is examined, it can be seen that when symmetry is used as the geometric cue, the location of motion did not significantly influence the responses of subjects. However, when convexity is used, this difference becomes significant (i.e. the difference in proportion between the cue competition condition and the cue cooperation condition when the geometric cue is convexity) ($p < 0.001$).

When the graphs in the left column of Figure 1.5 are examined, it can be seen that even in the displays that are taken as unambiguous by traditional accounts of accretion/deletion (i.e. one side of each border is static), the shape of the border combines with the accretion/deletion cue for figure-ground interpretation. Both geometric cues combine with accretion/deletion in various degrees. As it is seen from the second red bar in the top-left graph of Figure 1.5, convexity cancels out the effect of accretion/deletion when the two cues are made to compete against each other. The figure-ground responses of the subjects are at chance levels in this cue competition condition. The symmetry cue also causes the regions that contains accretion/deletion to be perceived as figural on a certain proportion of the trials.

The responses of subjects obtained from the experimental conditions where both sides of a border had motion (the graphs in the right column of Figure 1.5) are very similar to the responses obtained by Froyen et al. (2013). The subjects were more biased to see the moving region as figural and also as rotating when both sides of a border had motion. This was expected since introducing accretion/deletion to both sides of a border creates more ambiguity in the display. Geometric cues resolve this ambiguity and that is why the effect of geometric cues is increased in these conditions.

The response patterns obtained from the figure-ground task were fairly similar to the responses obtained from the rotation task. A regression analysis was done for each subject where the predictor variable was the proportion of the time the moving/indicated region was seen as in front for each experimental condition given to a subject in the figure-ground task, whereas the predicted variable was the proportion of the times the moving/indicated region was seen as rotating for each experimental condition in the rotation task. As a result each subject had eight data points on his/her scatter plot of the predictor and the predicted values. The regression analysis showed that for 11 subjects (out of 13) the proportions obtained from the figure-ground task was a significant predictor of the proportions obtained from the rotation task. (For all the subjects a significant regression result was found: $R^2_{max} = 0.92$, $R^2_{min} = 0.60$, $F(1,9)_{max} = 69.43$, $F(1,9)_{min} = 8.96$, $p < 0.05$). This high correlation suggests that perception of rotation in these displays is causally connected to the figure-ground interpretation. For example, the response patterns suggest that if the moving side was interpreted as figural, it is highly probable that it would also be perceived as a 3D column rotating in depth.

In the displays where one side of the border was static, there was one additional factor that we thought might be influencing subjects' responses. All the regions that have textural motion were moving in the same direction and at the same speed. This common motion present in the displays creates a grouping effect, where all the moving regions are more likely to be perceived as being grouped together in the background. Since this grouping effect favors the interpretation consistent with accretion/deletion (in its traditional definition), it creates an unfair cue competition condition for the geometric cues. In such cases, geometric cues were competing against two distinct cues; accretion/deletion and this grouping effect. In order to eliminate this extra factor from the displays, incoherent motion is used in Experiment 2.

1.3 Experiment 2

In the first experiment, in the conditions where motion was introduced to every other region, the moving regions were all translating in the same direction. This common motion might have resulted in a grouping effect which may have biased subjects to see these moving regions as being grouped in the background. In the cue-competition condition, such a grouping effect would be competing against the geometric cues while it would be cooperating with the accretion/deletion cue. In order to eliminate such an unfair competition between accretion/deletion and geometric cues, textural motion was made incoherent in the second experiment. Only the cue-competition condition was used in this experiment. In other words, the geometric and accretion/deletion cues were introduced to the same regions. On some trials, the textural motions in different areas were made incoherent by alternating the direction of motion in the moving regions such that there would be no motion-based grouping.

1.3.1 Method

Participants

Eight Rutgers University students, naive to the purpose of the experiment, were paid to participate..

Stimuli

The stimulus was generated in exactly the same manner as in Experiment 1, except that only four different conditions were used in this experiment (See Figure 1.6). The only difference was the addition of the incoherent motion condition, in which each convex/symmetric region was moving in opposite directions. In the coherent motion condition, all the convex/symmetric regions were moving in the same direction (as in Experiment 1). All the concave/asymmetric regions were kept static

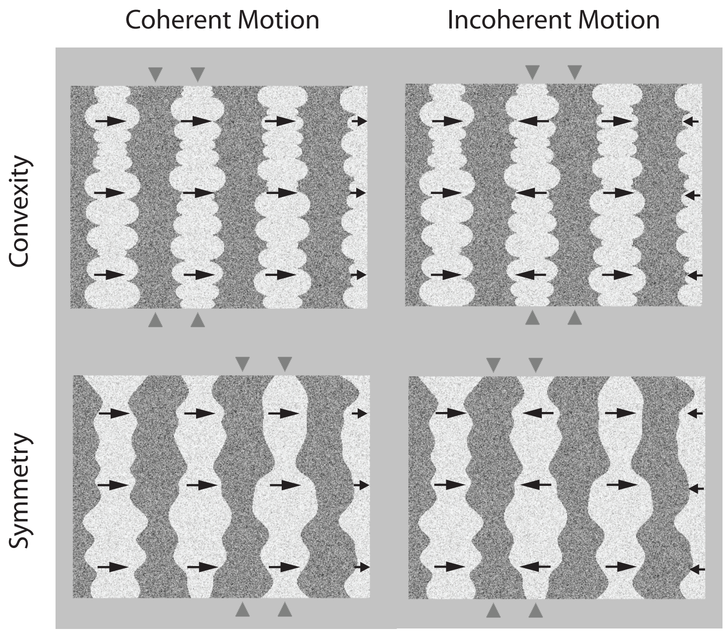

Figure 1.6: The stimuli used in the different conditions of Experiment 2

Design and Procedure

The procedure was exactly the same as in Experiment 1. In order to eliminate any possible influence due to the particular shape of the contours, we used two different contours for each geometric cue. For each task subjects ran 128 experimental trials split into two blocks, i.e. 2(geometric cues: convexity/symmetry) x 2(motion: incoherent / coherent) x 2(luminance: dark/bright) x 2(phase) x 2(direction of motion: right/left) x 2(Shape: two different contours for each geometric cue) x 2(repetition). Including the two tasks, there were a total of 256 experimental trials. It took not more than an hour for each subjects to complete the experiment.

1.3.2 Results and Discussion

Figure 1.7 shows the results plotted as the proportion of times subjects reported seeing the moving regions in front (for the figure-ground task) or as rotating (for the rotation task). Responses were analyzed for the two essential factors, i.e. the geometric cue, and motion type (coherent vs. incoherent). Except the geometric cue, motion type and color factors, all other factors were found not to yield any main nor interaction effect.

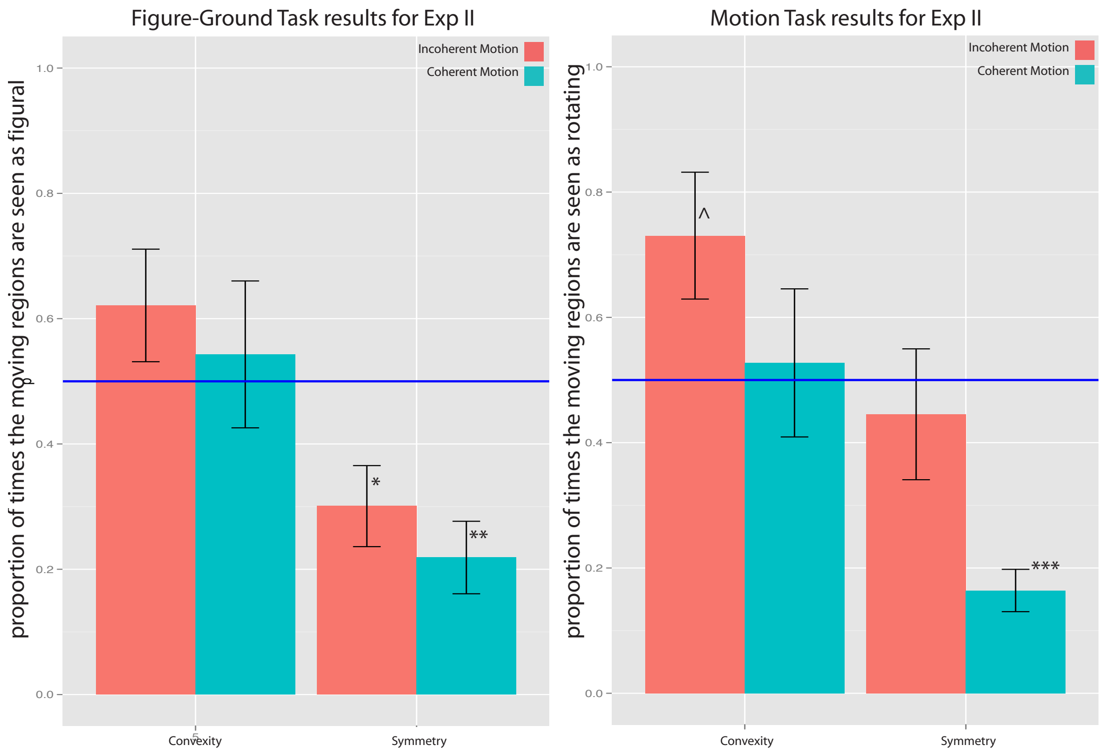

Figure 1.7: Results of Experiment 2. Error-bars represent $+/-1SE$ as computed between subjects. The blue line shows the chance level, i.e. where the proportion equals to 0.5. The stars (*) represent the proportions that are significantly different than chance level. The number of stars indicate the significance level: '***' $p < 0.001$; '**' $p < 0.01$; '*' $p < 0.05$; ' $\wedge$ ' $p = 0.0569$

t-test analysis was performed in order to see whether the proportions reported on Figure 1.7 are significantly different than 0.5, i.e. the chance level. The proportions that were significantly different than chance level are shown with their corresponding p values via star symbols on Figure 1.7. (Among all the significant differences obtained: $t_{max} = 9.94$, $t_{min} = 3.08$, $df = 12$, $p < 0.05$) (The maximum and minimum values were taken in their absolute values)

To analyze the responses of the subjects, a multilevel logistic regression was performed. For the figure-ground task responses, the geometric cue, motion type and the color factors yielded significant main effects, when compared to an unconditional means model (containing only an intercept) by means of a likelihood test (Geometric cue: $LR = 162.88$, $df = 3$, $p < 0.001$; Motion type: $LR = 12.95$, $df = 3$, $p < 0.01$; Color: $LR = 87.76$, $df = 3$, $p < 0.001$). There was no significant interaction between the two factors. As it can be seen from the graph on the left in Figure 1.7, subjects were again more likely to see the moving regions as figural when they were symmetric, compared to the cases when these regions were just symmetric. We found that the proportion of times people saw the moving region as figural was significantly higher when the motion was incoherent compared to when it was coherent. This is consistent with our prediction that motion grouping acts as an additional cue for relative-depth perception in our displays. Regarding the color factor, subjects in general were more biased towards seen the light regions as figural compared to dark regions. However, when individual responses were examined, there was not any systematic effect of color. Among eight subjects six of them has this light-color bias, whereas the other two has dark-color bias.In order to avoid any possible confounding effect of the color factor, the logistic regression analyses reported above were repeated, but this time we controlled for color. This was again done by including the color factor both into the "null" and "alternative" models in all the likelihood ratio tests, then by checking whether the geometric cue and motion type factor introduce a significant expansion to a model that already includes the color factor. The results obtained after controlling for color were not different (in terms of significant main effects) than what is reported above.

For the rotation task responses, the multilevel logistic regression yielded the same significant main effects (Geometric cue: $LR = 207.53$, $df = 3$, $p < 0.001$; Motion type: $LR = 93.72$, $df = 3$, $p < 0.001$). In the same way, subjects were more likely to see the moving regions as figural when they were convex and symmetric compared to when they were just symmetric. Subjects were also more likely to see the moving regions as rotating when the motion was incoherent instead of coherent. Color factor was also found to yield a significant effect ($LR = 54.28$, $df = 3$, $p < 0.001$). The pattern of responses regarding the color factor was similar to the one obtained for the figure-ground task. Among eight subjects six of them had this light-color bias, whereas the other two had dark-color bias. (When we controlled for color, the two factors, geometric cue and motion type still yielded significant main effects).

The responses obtained from the coherent condition of the figure-ground task were similar to the results obtained from the first experiment. The subjects were more likely to see the moving regions as figural when the motion was incoherent, compared to when it was coherent. This confirms that making the motion coherent (as it was in Experiment 1) introduces an additional factor that goes against the geometric cues. It can also be seen that incoherent motion had a greater influence on the rotation task than on the figure-ground task. Subjects' judgments on the rotation task were more sensitive to the motion coherence manipulation compared to their judgments on figure-ground assignment. (A regression analysis between the responses obtained from the two tasks could not be done for this experiment because the number of experimental conditions used was four, which is very low for a regression analysis). Overall, the significant difference observed between the responses of the coherent and incoherent conditions shows that eliminating the grouping effect by making the motion incoherent considerably alters subjects' responses.

1.4 General Discussion

Traditionally accretion/deletion is considered a sufficient cue to depth-order such that it unambiguously assigns figural and ground status to image regions in dynamic 2D images. However, our results show that accretion/deletion is just one of many cues that combines with other visual information (i.e. geometric cues) to determine relative depth. Froyen et al. (2013) have shown that when an accretion/deletion cue is introduced on both sides of a border, an ambiguity about depth-order is created. Their results showed that this ambiguity can be resolved by introducing geometric figure-ground cues. However, in the cue-competition condition of the current experiments, only one side of each border had accreting-deleting texture, whereas the other side was static. Such a stimulus would be considered unambiguous in terms of traditional accounts of accretion/deletion and the moving texture would be predicted to be perceived as further away. However, as the results show, introducing geometric cues strongly modulates the perception of relative depth and layered surface structure.

Moreover, the perception of 3D columns rotating in depth is observed even in the cue-competition condition and in spite of the fact that the dot texture motion is linear, which is technically inconsistent with 3D rotation.

In general, responses obtained from the rotation task were fairly consistent with the responses obtained from the figure-ground task. Such similarities between the response patterns of the two different tasks suggest that the two judgments are strongly intertwined. In classical accretion/deletion stimuli seen in the literature, the static region is seen in front and interpreted as the region that occludes the disappearing texture. However in our stimuli, when geometric cues are able to assign figural status to the moving region, the visual system is confronted with evidence that the translating texture is in front and therefore infers that the region is a 2D projection of 3D rotating column, hence the accreting and deleting texture is explained by dynamic self-occlusion due to rotation.

In both Experiment 1 and 2, the motion manipulations (either changing the location of motion, or altering the coherence of the textural motion) had a greater influence on the responses in the rotation task than in the figure-ground task. This can be observed on all the bar graphs presented above. The response differences between the different motion conditions (the differences between the red and turquoise bars) were much higher in the rotation task than in the figure-ground task. This suggests that the rotation task responses were more sensitive to motion manipulations. We would argue that the question asked for the rotation task (i.e. whether the subjects see a rotational motion in the moving region) may be considered a more reliable and indirect method for measuring figure-ground perception, rather than directly asking about figural status. Individual differences were also observed among the participants, especially in the cue-competition condition. This can also be seen from the large standard error bars on the top-left graph in Figure 1.5. Hildreth and Royden (2011) showed that there are individual differences in the way people combine accretion/deletion and binocular disparity cues in a depth-order task. While some subjects give more weight to the accretion/deletion cue, other subjects give more weight to binocular disparity in their depth-order judgments. Consistent with the study by Hildreth and Royden (2011), our results also suggest that there are individual differences in the way people use the accretion/deletion cue for depth-order interpretation.

Our results suggest a host of unanswered questions. First of all, they require us to re-consider the status of accretion/deletion as an unambiguous cue to figure-ground. One important issue that needs to be clarified is what visual information exactly accretion/deletion conveys depending on the geometric properties of the border. For example, in our visual stimuli, it seems that once the accretion and deletion of the texture is explained by the visual system as a self-occlusion due to rotation, the accretion/deletion no longer functions as the figure-ground cue that indicates the occluded surface. An account has to be proposed that can incorporate that kind of interactions of accretion/deletion with the geometric properties of the border. Another important point is how all of these findings connect with structure-from-motion process. In our stimuli, the perception of 3D columns rotating in depth is observed even though the dot texture motion is linear, which is technically inconsistent with 3D rotation (Ullman, 1979). In order to understand the relationship between figure-ground perception and perception of 3D columns rotating in depth, more studies should be performed. Future studies might include running experiments where the motion profile of the moving texture is manipulated and experiments with large numbers of trials within individual subjects . It is also important to gain more knowledge about how accretion/deletion and geometric cues combine to give the perception of figure-ground segregation. For that a new study can be done by applying gradual changes to both geometric and accretion/deletion cues in order to come up with a cue combination model.

Bibliography

Froyen, V., Feldman, J., & Singh, M. (2010). A bayesian framework for figure-ground interpretation. In J. Lafferty, C. K. I. Williams, J. Shawe-Taylor, R. S. Zemel & A. Culotta (Eds.), *Advances in neural information processing systems 23* (Vol. 631-639).

Froyen, V., Tanrikulu, Ö. D., Singh, M., & Feldman, J. (2012). Stereoslant: a novel method for measuring figure-ground. *Poster presented at the annual meeting of the Vision Science Society, Naples, FL*.

Froyen, V., Feldman, J., & Singh, M. (2013). Rotating columns: relating structure-from-motion, accretion/deletion, and figure/ground. *Journal of Vision*, *13*(10).

Gibson, J., Kaplan, G., Reynolds, H., & Wheeler, K. (1969). The change from visible to invisible: a study of optical transitions. *Perception & Psychophysics*, *5*, 113–116.

Hegdé, J., Albright, T., & Stoner, G. (2004). Second-order motion conveys depth-order information. *Journal of Vision*, *4*, 838–842.

Hildreth, E., & Royden, C. (2011). Integrating multiple cues to depth order at object boundaries. *Attention, Perception, & Psychophysics*, 1–18.

Hoffman, D., & Singh, M. (1997). Salience of visual parts. *Cognition*, *63*(1), 29–78.

Howard, I., & Rogers, B. (2002). *Seeing in depth vol.2 depth perception*. I. Porteous, Toronto.

Kanizsa, G, & Gerbino, W. (1976). Vision and artifact. In M. Henle (Ed.). (Pp. 25–32). New York: Springer.

Kaplan, G. (1969). Kinetic disruption of optical texture: the perception of depth at an edge. *Attention, Perception, & Psychophysics*, *6*(4), 193–198.

Kromrey, S., Bart, E., & Hegdé, J. (2011). What the 'moonwalk' illusion reveals about the perception of relative depth from motion. *PLoS ONE*, *6*(6).

Metzger, F. (1953). *Gesetze des sehens*. Frankfurt-am-Main: Waldemar Kramer.

Mutch, K., & Thompson, W. (1985). Analysis of accretion and deletion at boundaries in dynamic scenes. *Pattern Analysis and Machine Intelligence, IEEE Transactions*, *2*, 133–138.

Nakayama, K., He, Z., & Shimojo, S. (1995). Visual surface representation: a critical link between lower-level and higher level vision. In S. Kosslyn & D. Osherson (Eds.), *Invitation to cognitive science* (pp. 1–70). M.I.T. Press.

Niyogi, S. (1995). *Detecting kinetic occlusion*. Massachusetts Institute of Technology, Cambridge, MA.

Ono, H., Rogers, B., Ohmi, M., & Ono, M. (1988). Dynamic occlusion and motion parallax in depth perception. *Perception*, *17*(2), 255–266.

Thompson, W., Mutch, K, & Berzins, V. (1985). Dynamical occlusion analysis in optical flow fields. *Pattern Analysis and Machine Intelligence, IEEE Transactions*, *4*, 374–383.

Ullman, S. (1979). The interpretation of structure from motion. In *Proceedings of the royal society of london* (Vol. 203, *1153*, pp. 405–426). Series B, Biological Sciences. The Royal Society.

Wagemans, J., Elder, J. H., Kubovy, M., Palmer, S. E., Peterson, M. A., Singh, M., & von der Heydt, R. (2012). A century of gestalt psychology in visual perception: 1. perceptual grouping and figure-ground organization. *Psychological Bulletin*, *in press*.

Yonas, A., Craton, L., & Thompson, W. (1987). Relative motion: kinetic information for the order of depth at an edge. *Attention, Perception, & Psychophysics*, *41*(1), 53–59.