INFERENCE OF 3D SHAPE FROM LINE DRAWINGS

By

SEHA KIM

A dissertation submitted to the

Graduate School–New Brunswick

Rutgers, The State University of New Jersey

In partial fulfillment of the requirements

For the degree of

Doctor of Philosophy

Graduate Program in Psychology

Written under the direction of

Jacob Feldman

And approved by

_____

_____

_____

_____

New Brunswick, New Jersey

October, 2015

ABSTRACT OF THE DISSERTATION

Inference of 3D Shape from Line Drawings

By SEHA KIM

Dissertation Director:

Jacob Feldman

Line drawings lack direct 3D depth information, yet human vision easily perceives the 3D shapes from the contours. This dissertation investigates the mechanisms underlying the 3D shape inference from 2D line drawings. Here, four psychophysical experiments and a computational model for the 3D shape inference are discussed. Experiment 1 shows that human responses in depth judgments for line drawings reflect an underlying uncertainty of the perceived 3D shape, which is based on the complex interaction of local and global depth cues propagated from the contours. The computational model estimates the posterior probability of possible 3D surfaces from the contours of a line drawing in a Bayesian framework. The comparison of the model predictions and human depth responses for the line drawings from Experiment 1 demonstrates that the model accounts for the probabilistic 3D shape interpretation of line drawings by human vision. Experiment 2 shows that the reliability of a contour segment in a line drawing as a meaningful depth cue is conditional to the complex global context. Experiments 3 and 4 show that the certainty of depth difference perceptions from partial line drawings increases as more non-local visual cues are available. The experiments and the model offer a new perspective on 3D shape perception from line drawings as an inference based on the probability over possible 3D shapes given the contour cues, providing a broader understanding on the mechanisms of human vision.

Acknowledgements

This research was funded by NIH (NEI) Grant EY021494 and the Rutgers IGERT program in Perceptual Science NSF DGE 0549115.

This dissertation took shape under the guidance of Jacob Feldman and Manish Singh. I would like to express my special appreciation for your mentorship and support.

I am grateful to my committee members, Thomas Papathomas and Ahmed Elgammal, for the brilliant comments and suggestions.

I give special thanks to Eileen Kowler for her encouragement and support.

I want to thank my fellow graduate students, the faculty members, and staff in Psychology Department and the Center for Cognitive Science. There are too many people to name here, but I could not have done this without their help and generosity.

Dedication

This dissertation is dedicated to my parents, to my sisters Wonha and Joongha, and to my brother Jiha, and to my husband Junghoon.

Table of Contents

List of Tables

List of Figures

Chapter 1

General Background

1.1 3D Shape interpretation of line drawings

Line drawings are a simple but rich domain for understanding 3D shape perception. A surprisingly rich 3D inference can be drawn from a very impoverished stimulus with sparse depth cues (Koenderink et al., 2012). Human vision can easily recover depths and surface structures from line drawings as shown in Figure 1.1. The area inside contours is not filled with direct depth cues, such as texture and shading, yet humans are very accurate at interpreting the 3D shapes (Barrow & Tenenbaum, 1981; Marr, 1982; Koenderink & van Doorn, 1982; Hoffman & Richards, 1984; Biederman, 1987; Pizlo & Salach-Golyska, 1995). An experiment has shown that even 9-month-old children can interpret line drawings without training (Hochberg & Brooks, 1962). However, there are very few studies for 3D shape interpretation from pure line drawings, compared with many studies about 3D shape inference from shading and texture (Bülthoff & Mallot, 1988; Liu & Todd, 2004; Todd et al., 1996; Koenderink et al., 1992; Koenderink & van Doorn, 1995; Koenderink, 1998; Todd & Reichel, 1990; Knill, 1998a,b; Knill & Saunders, 2003; Zimmerman et al., 1995; Norman et al., 2006; Todd, 2004; Koenderink et al., 2001).

In most areas of line drawing images, there is no information at all except the contours. Yet, human observers have a clear and strong perception of 3D structures purely on the basis of the contours (Pizlo, 2010; Koenderink et al., 2012). This implies that the depth information "propagates" from the contour in a way that allows depth judgments to be made everywhere, including the "empty" area. Depth cues are spatially transferred along the contours and also propagated over the area where no explicit local depth cues exist. This is an extremely interesting issue because it epitomizes one of

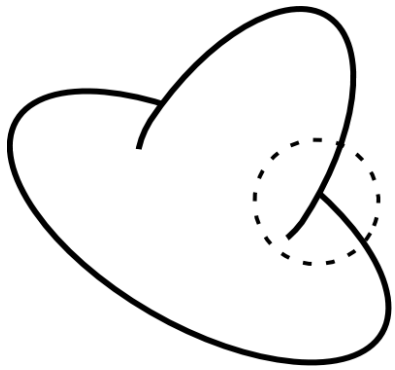


Figure 1.1: A line drawing depicting a simple form.

the central issues in perceptual organization: the integration of local and global cues. Thus, the intrinsic interest of line drawings as an example of 3D inference from the integration of local and global information promises to shed light on much more far-reaching problems in vision.

An enormous amount of work in the 1970's and 1980's tried to solve line drawing interpretation by contour labeling, which categorizes junction types and picks legitimate labels for the contours and surfaces at the junctions (Huffman, 1971; Clowes, 1971; Waltz, 1972; Mackworth, 1973; Charkravarty, 1979; Malik, 1987). These works assumed a very deterministic process aimed at finding solid models for block-like shapes that were consistent with known junction constraints. These methods were seen to be very limited in part because they do not handle violations of the assumptions well. They did not explain the surface interpretation of more general forms but focused on deterministic selection of a single 3D model consistent with the logical relations between surfaces of the solid object. The more recent tools in computer graphics construct 3D surfaces from contours (Zeleznik et al., 1996; Igarashi et al., 1999; Nealen et al., 2005; Jorge & Samavati, 2011). However, the tools also focus on creating deterministic surfaces from silhouettes and editing the surface by human users, so the tools do not explain human interpretation of contours into 3D surfaces.

A computational model to recover 3D shapes from 2D line drawings has been suggested by Pizlo and his colleagues (Li et al., 2009; Pizlo, 2010; Pizlo et al., 2010). The model uses Gestalt principle of simplicity as *a priori* constraints to recover 3D coordinates of the 2D contours in a single line drawing image for polyhedral and natural

objects. The constraints are to maintain 3D symmetry, to maximize the planarity of contours, to maximize 3D compactness, and to minimize surface area. However, the model does not explain the 3D surface percepts in the empty areas between the recovered 3D contours, and the approach is also deterministic.

A more probabilistic approach to this problem was suggested by Mamassian & Landy (1998). They suggested that local 3D shapes for the line drawings of curved surface patches—concave or convex shapes—can be estimated in a Bayesian fashion by assuming priors based on the preferences for upward surface normals, for locally convex shapes, and for surface contours (surface markings) aligned with the principal curvature lines on surfaces. In this study, I attempt to extend the similar Bayesian framework to estimate the entire shape of volumetric 3D objects by constructing a probability distribution over possible 3D surfaces, focusing on how the contour structures in line drawings, such as junctions, internal contours, and all bounding contours, interact to generate the possible 3D shapes.

1.2 The types of contours and features in line drawings

The types of contours in line drawing images that convey 3D shape information have been categorized based on different research approaches. According to Stevens (1981), contours are categorized into *occluding contours*, which outline the boundaries of objects and background, and *surface contours*, which appear due to luminance differences on the surfaces bounded by the occluding contours. Marr (1982) categorized the contours in 2D images by the origins—the discontinuities in depth, surface orientation, and surface reflectance. On the other hand, Koenderink & van Doorn (1982) and Koenderink (1984) argued that contours are generated by the discontinuities in either depth or surface orientation but not in surface reflectance, because contours refer to the visible 2D project of the *rim*—the locus along on a surface touched by the rays from a vantage point—which divides visible and invisible surfaces on the object.

In this study, Experiment 1 tests line drawings that represent occluding contours,

Experiment 2 tests the possibility of interpreting contours in line drawings as occluding contours vs. surface contours (surface markings), and Experiments 3 and 4 test occluding contours that are projected from the rim of 3D surfaces.

Among the contour features in line drawings, the termination and curvature of contours have critical information about 3D shapes. The junction labeling studies characterized contours by the terminations and cusps—junctions—and interpreted 3D structures by the rule that logically relates junctions to the 3D surfaces (Huffman, 1971; Clowes, 1971; Waltz, 1972; Mackworth, 1973; Charkravarty, 1979; Malik, 1987). Richards et al. (1987) suggested an interpretation using the curvatures of contours. The method divides smooth contours into the primitive parts—called *codons*—at the maximum, minimum, and zero curvatures on the contours, and interprets 3D structures by the rule that mathematically relates codons to the 3D surfaces. However, the rule-based methods are applicable only to very limited types of line drawings, and do not provide a computational model for the process.

In this study, the process of 3D shape inference from 2D contours is modeled in a probabilistic computation that can account for human interpretation of line drawings based on the important contour features—junctions and curvatures—in addition to the complex interaction of local and global contour structures.

1.3 Dissertation Overview

This dissertation addresses the following questions: how human vision perceives depth from line drawings, what the important 2D contour features are in the 3D shape interpretation, what decides the type of a contour in a line drawing, and how non-local cues interact with local depth cues.

This dissertation includes four psychophysical experiments and one computational model. Chapter 1 presents a psychophysical experiment of human line drawing interpretation and 3D shape perception (Experiment 1), and suggests a computation model of the 3D shape inference. The experiment investigates the interaction of local T-junction cues and non-local contour features. The model estimates a probability distribution of

possible 3D surfaces for a given line drawing. Chapter 2 describes an experiment aimed to investigate the influence of global context on interpreting contour segments in line drawings as occlusion cues or surface markings. Chapter 3 presents two experiments aimed to investigate how the amount of non-local visual information influences line drawing interpretations.

Chapter 2

Human Line Drawing Interpretation and a Bayesian Model

2.1 Experiment 1

This experiment mainly aimed to probe relative depths at various areas in line drawings, where the influence of local and global contour cues on interpreting the 3D shapes could be revealed. I placed two dots on a line drawing and asked participants "which dot appears closer." The answer informed the 3D surface perceived from the 2D line drawing. I tested line drawings that had only one critical T-junction. The perceived depths would be influenced not only by the conventional local depth cue (the T-junction) but also by global cues of the line drawings. To measure the changes induced by non-local cues, the probes were placed at various distances from the T-junction for a line drawing.

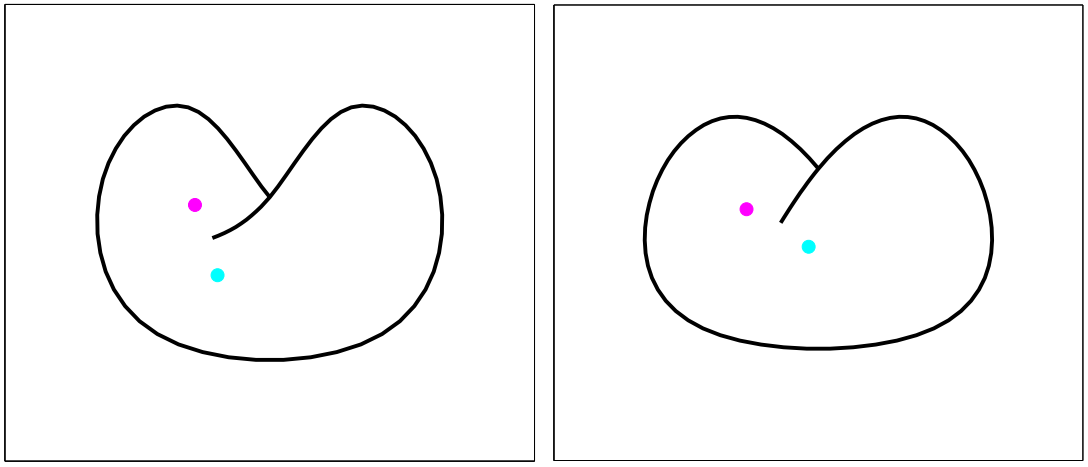
2.1.1 Methods

Participants

Subjects were 10 undergraduate students at Rutgers University who participated in the experiment for course credits.

Stimuli

The stimuli were images of 2D line drawings of black contours on white background, with two dots in different colors superimposed on the line drawing image (Figure 2.1). The subjects were asked to answer which dot appeared closer to them. The stimuli were presented on a 17 inch color monitor (1152×854 pixel at 82Hz, $27° \times 19°$ in visual



(a) Line Drawing 1 (b) Line Drawing 2

Figure 2.1: Pairwise dot comparisons on line drawings. Subjects were asked to answer "which dot appears closer" and responded by selecting the dot color. Rectangular boxes represent monitor screens.

angle) connected to a AppleG4 computer. The viewing distance was approximately 27 inches from the monitor. The line drawing stimuli subtended a maximum of 8.5° in visual angle horizontally and vertically.

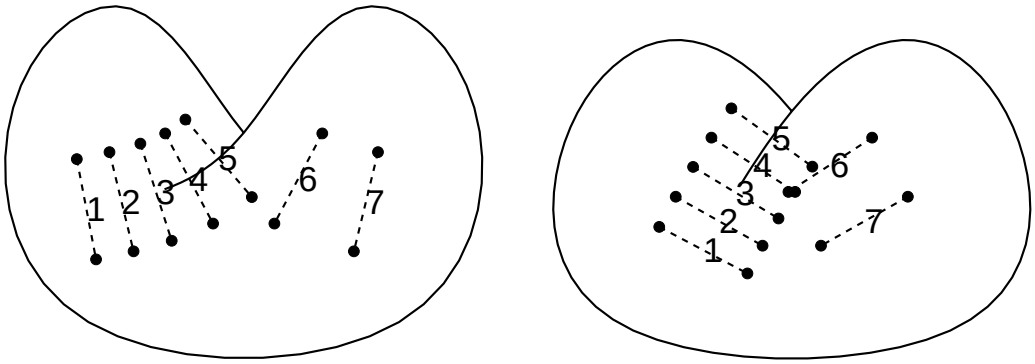
I used two types of line drawings depicting simple forms, Line Drawing 1 and Line Drawing 2 (Figure 2.1). They were not randomly picked but designed to examine the influence of local and non-local contour cues and to explain the theoretical difference between them. The bounding contour of each line drawing was made to be symmetric, having exactly the same left and right sides except for the internal contour. The symmetry allows an operational examination for the influence of local T-junction structure (in particular, the internal contour) while keeping the non-local contour structure the same. Also, the internal contours in two line drawings were made to be geometrically different; as shown in Figure 2.1, the ending of the internal contour in Drawing 1 is concave and thus consistent with a smooth surface, but the ending of the internal contour in Drawing 2 is convex and thus inconsistent with a smooth surface. The ending of a contour for a smooth 3D surface is supposed to be concave because it is made from the projection of a local saddle surface (Koenderink & van Doorn, 1982; Koenderink, 1984). Although the relations between curvatures of occluding contours and surface shapes in the two line drawings are geometrically different, the difference seems relatively subtle

and often goes unnoticed by human eyes (Koenderink & van Doorn, 1982; Koenderink, 1984). This raises the question of whether the difference will be reflected in our data.

Design and procedure

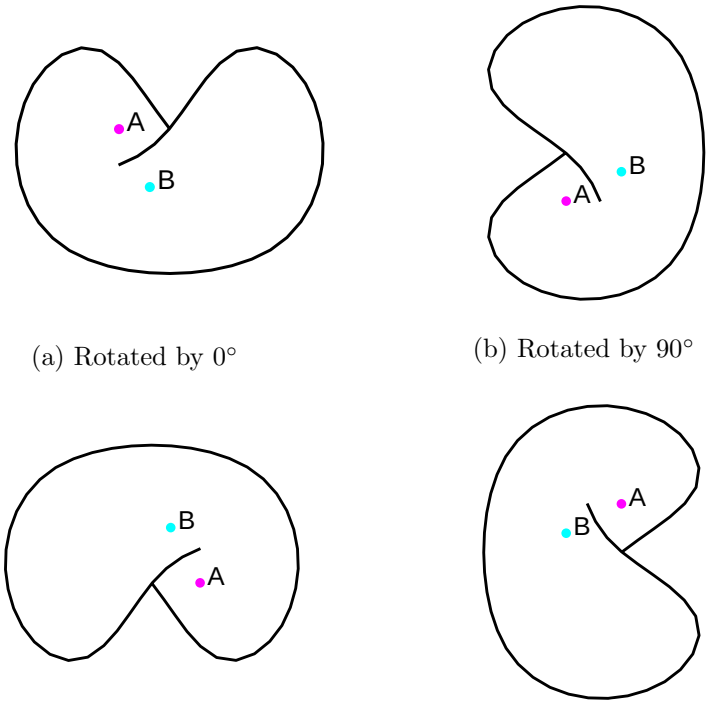
One of the main factors, manipulated in this experiment, was the location of probes with respect to the T-junction (Figure 2.2). In order to test the influence of the distances from a T-junction cue, the paired dots (from Probe Sets 1 to 5) were placed along the internal contour from farther to closer to the T-junction. Then, Probe Sets 2 and 4 were mirrored to the opposite side (between the left and right sides) where there was no internal contour, and they were Probe Sets 6 and 7, respectively. The two probe sets made it possible to compare the 3D perception either with or without the internal contour while keeping the same bounding contours, which are symmetric as explained above. Another factor was the orientation of stimuli (Figure 2.3). The middle point of two probe dots was placed at the center of screen, and then the whole image was rotated counterclockwise by one of four angles, 0°, 90°, 180°, and 270°, about the middle point. Thus, the relative locations of two probe dots on the screen were varied by the angle. The other factors were stimulus display durations (200 or 800ms), two color sets of probe dots between cyan and magenta, and two horizontally flipped images that displayed the ending of the internal contour either on the left or right side. Each condition was repeated three times; therefore, a total of 1344 trials were tested.

In each trial, after a fixation screen displayed for 1 sec, a stimulus was displayed for the duration and disappeared on the screen. Then, the subject pressed the f or j key according to the color of the dots that they perceived closer. I measured the proportion of responses "congruent" with the general interpretation of the internal contours of 3D shapes for the stimuli (*congruence rate*). For example, in the line drawings shown in Figure 2.3, the act of selecting the cyan color dot B represents a congruence choice. For Probe Sets 6 and 7, since there were no internal contours on these probes that made the general interpretation as to the local depths, the congruence rates were defined according to the congruent rates for Probe Sets 4 and 2, respectively. For example, the congruent response of Probe Set 6 is selecting the dot mirroring the congruent



(a) Line Drawing 1 (b) Line Drawing 2

Figure 2.2: Probe locations. Each pair of dots is grouped with dashed lines (the lines were not shown in the experiment). Probe Sets 1 to 5 are located along the internal contour from farther to closer to the T-junction, and Probe Sets 6 and 7 are located on the opposite side.



(a) Rotated by 0° (b) Rotated by 90°

(c) Rotated by 180° (d) Rotated by 270°

Figure 2.3: The orientations of stimuli. Stimuli are rotated counter-clockwise about the midpoint of probe dots A and B.

dot of Probe Set 4. Congruence rates at the various probe locations would reveal the differences in depth perception among the conditions.

2.1.2 Results and discussion

In ANOVA analysis, I found a significant main effect of probe location ($F(6, 54) = 2.593,\ p = 0.0279$) on congruence rates (Figure 2.4). First of all, these results support the idea that human 3D interpretation from 2D line drawing is not deterministic. If one perceived a deterministic depth difference at two probe dots, then the judgment about the closer dot would always be the same, and the consequent congruence rate should always be either one or zero. However, this was not the case. The congruence rates from the depth judgments were between zero and one, even where the probe dots were obviously separated by the internal contour. Such rates are accounted by the well-known "probability matching" phenomena. Humans tend to match the rate of choice between two alternatives in proportion to the actual probability, instead of always choosing the one with the higher likelihood in repeated binary decision tasks (Fantino & Esfandiari, 2002; Green et al., 2010). Therefore, the congruence rates reveal the probability of perceived depth difference by the two probe dots from the repeated depth judgments. The higher congruence rate means that the certainty of depth difference is higher, and the lower congruence rate means that the certainty is lower.

Secondly, the certainty for the depth perceptions at the probe sets along the internal contour (from Probe Sets 1 to 5) was influenced by the distance from the T-junction. Analyzed as a function of distance, congruence rates systematically increased as probes got closer to the junction. According to a logistic regression model, the log odds of congruent responses were positively related to the distance ($z = 5.061,\ p = 4.17e^{-07}$). This suggests that depth information from a local T-junction cue propagates both along its contour and even further inward to create the global surface interpretation. In addition, the congruence rates for Probe Sets 6 and 7 were very different from the congruence rates for the corresponding Probe Sets 4 and 2, respectively. This shows that the internal contours clearly influenced perceived surfaces while the nearby bounding contours are exactly the same. Taking these results together, perceiving 3D surfaces

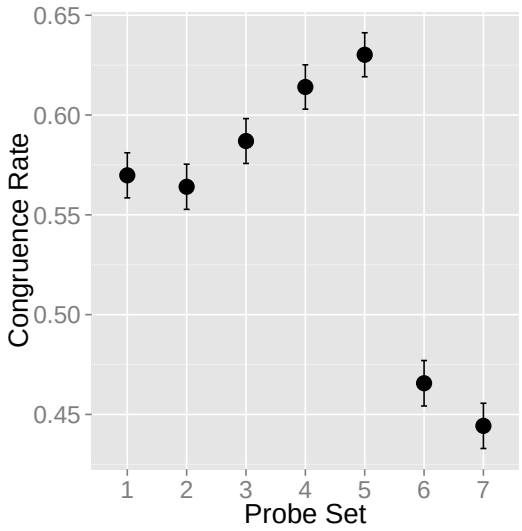


Figure 2.4: The influence of probe locations on congruence rates. Means and standard errors in each condition are shown.

from line drawings can be explained as the integration of information propagated from both local cues, such as junctions, and non-local cues, such as the global context of contours.

Thirdly, the congruence rates were influenced by the relative height of probe dots on the screen, which was changed by the different orientations of stimuli. For example, as shown in Figure 2.3, for the probe set, B is lower than A on the stimulus with 0° orientation, but B is higher than A on the stimulus with 180° orientation. The exact height differences for all different probe sets and stimuli were grouped into ten subsets with the height difference values in the 10-quantiles, and the means and standard errors of congruence rates for each subset were plotted in Figure 2.5. The negative height difference means that the "congruent" and "closer" dots were lower than the other dot on the screen. The relation between the height differences and congruence rates were analyzed by a logistic regression model, which showed that the log odds of congruent responses were negatively related to the height differences ($z = -12.720,\ p < 2e^{-16}$). The results mean that the depth difference between two dots for the same probe set was perceived higher when the "congruent" dots were lower on the screen, implying that the perceived 3D surface from a line drawing image is influenced by an even more global context, the orientation of the whole line drawing. The asymmetric depth perceptions

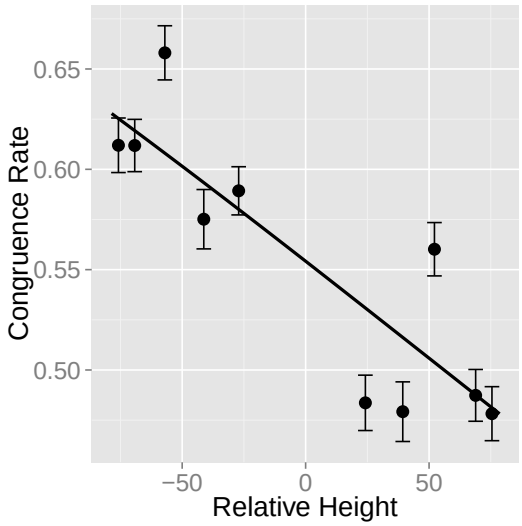


Figure 2.5: Congruence rates as a function of height difference between two probe dots. The height difference is the height of "farther" dot on the screen subtracted from the higher of "closer" dot for each probe set. Each data point shows the mean and standard error of congruence rates for each subset. The line shows a fitted curve from the logistic regression.

are consistent with the pictorial depth cue observed by Vecera et al. (2002). They found more figural perception for lower regions of display than upper regions, which leads to closer depths for lower regions than upper regions.

In addition, the two different line drawings tested here were theoretically different, but there was neither significant difference in congruence rates nor interactions (Figure 2.6). Line Drawing 1 is consistent with a smooth surface, but Line Drawing 2 is not because the internal contour ends convex (Koenderink & van Doorn, 1982; Koenderink, 1984). However, there was no distinction between the two line drawings in the depth judgments, implying that the subjects had at most limited sensitivity to this type of geometric constraint.

Taken together, these results show that the ambiguity in interpreting 3D shapes from line drawings is resolved using both local and non-local structures in the line drawings, but the perceived depth differences on the 3D surfaces are inherently probabilistic. The congruence rates represent the beliefs in the depth difference based on local and global contour cues. To capture the probabilistic nature, a model of the interpretation of 3D shapes from line drawings must incorporate a computation for a probability of possible

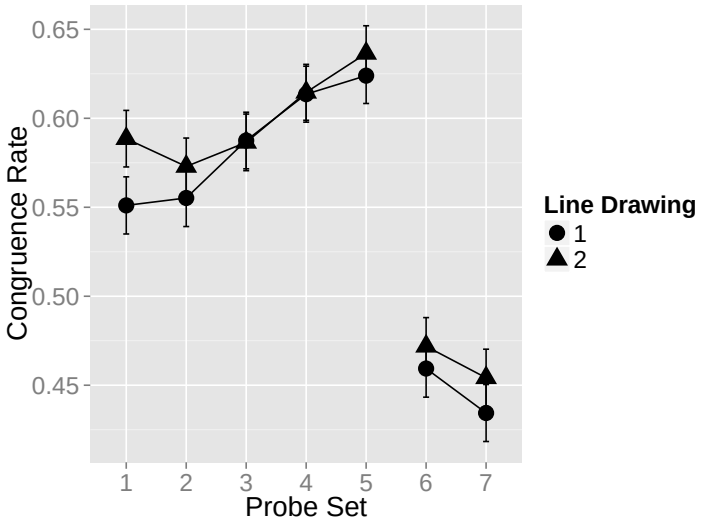


Figure 2.6: The interaction of probe locations and line drawing types. No significant difference between different type of line drawings was found.

3D shapes. The probability could provide more information about the perceived shape than one reconstructed depth map or surface as in the previous studies on pictorial surfaces (Koenderink et al., 1992; Koenderink & van Doorn, 1995; Koenderink, 1998). In the following, I propose a model by which the computation is processed.

2.2 Model

I propose a computational model of 3D shape interpretation from line drawings that accounts for (1) the probabilistic nature of human line drawing interpretation, and (2) the interaction between local and global cues in line drawings that creates depth and surface percepts. This model estimates distributions of possible 3D shapes from the contours in line drawings. Based on the distributions, probabilities of depth differences on the perceived surfaces are computed and compared with the human responses from the experiment.

This model estimates distributions of depth differences given a novel line drawing in a Bayesian framework as follows:

$$P(D|L) \propto P(L|D) \times P(D), \tag{2.1}$$

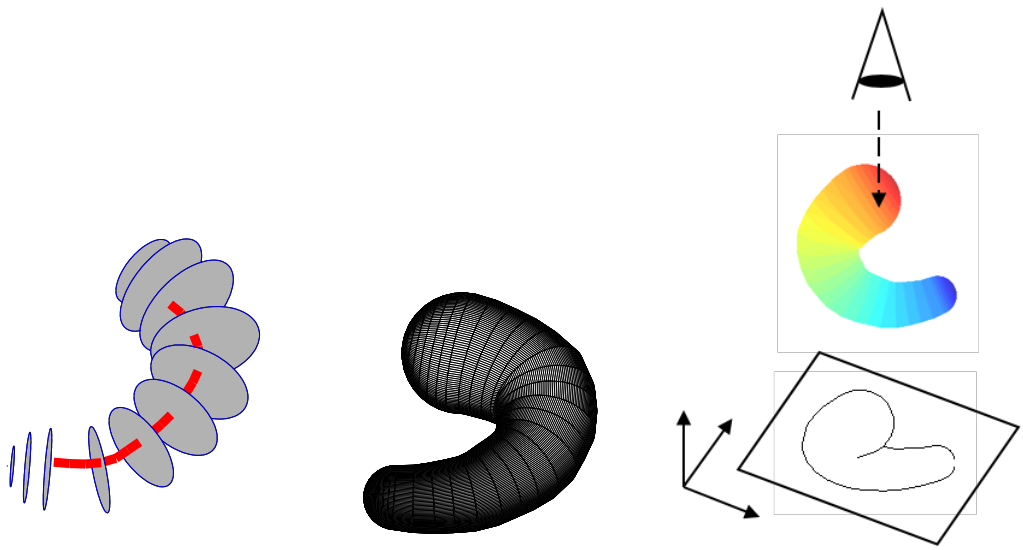
where L is a line drawing, and D is the depth difference. This model aims to formulate the probability distribution of depth difference between two dots on the pictorial surface. The posterior of depth difference for the unknown 3D surface of a line drawing, $P(D|L)$, is computed by incorporating a prior of depth difference, $P(D)$, and a likelihood of depth difference, $P(L|D)$.

The likelihood function, $P(L|D)$, is defined from the relations between line drawings and the 3D shapes that give rise to the line drawings, assuming a generative process of the 2D line drawings from the 3D shapes. It is difficult to compute this likelihood function analytically, so it is computed numerically by a Monte Carlo simulation, which uses random line drawings and 3D shapes from the following generative model.

2.2.1 The inflation model

This generative model creates 3D shapes and line drawings in a stochastic way. The shapes in this model represent simple forms with smooth surfaces of natural 3D objects, such as animals' torsos or fruits. The random 3D shapes are generated by inflating surfaces from random skeletons, and the random line drawings are generated by projecting the inflated surfaces from random viewpoints. This procedure is summarized in Figure 2.7.

Firstly, random skeletons are generated. To produce perceptually continuous and smooth random curves for skeletons, I connect line segments which are close to being collinear (Field et al., 1993; Geisler et al., 2001; Feldman, 2001). The angle differences between two consecutive segments on skeleton polygons are sampled from a zero-mean Gaussian. For simplicity, in this model, all skeletons are planar. Secondly, a skeleton is inflated to have circular cross-sections at each vertex, thus creating a volumetric shape similar to the *generalized cones* with circular cross-sections (Marr, 1977). The thickness of the inflated surface is linearly changing along the skeleton with the parameters randomly sampled as well. Then, random 2D line drawings are projected from the inflated 3D shapes via orthographic projections from random viewpoints. The contour is made where a surface is turning away (Barrow & Tenenbaum, 1981). The marginally visible *rim* points on the inflated surface are projected on the image plane, and the points on



(a) Skeleton and circular cross-section

(b) Inflated surface

(c) Extracting a line drawing

Figure 2.7: A schematic diagram of the inflation model. (a) 3D surfaces are inflated from smoothly curved skeletons, having circular cross-sections. (b) The inflated 3D shape is represented by rectangular meshes with the surface normals on the center of each mesh. (c) Line drawing contours are extracted via orthographic projections.

the image are connected to make the contours of a line drawing.

In this way, many random line drawings with the corresponding 3D shapes are generated. The examples of random line drawings used in this study are shown in Figure 2.8. The relations of the 3D shapes and line drawings are tabulated in a database, from which the likelihood of depth difference given a new line drawing, $P(L|D)$, will be estimated. The relations and the representations of the line drawings and 3D shapes in the database will be explained in the following section.

2.2.2 The database of 2D contours and 3D surfaces

The inflation model generates random inflated shapes and projects them as random line drawings. The contour segments on the line drawings are related to the local 3D surfaces on the inflated shapes. The contour segments and local surface patches are quantified and tabulated to be compared and to be combined later when probable 3D shapes are queried for a novel contour.

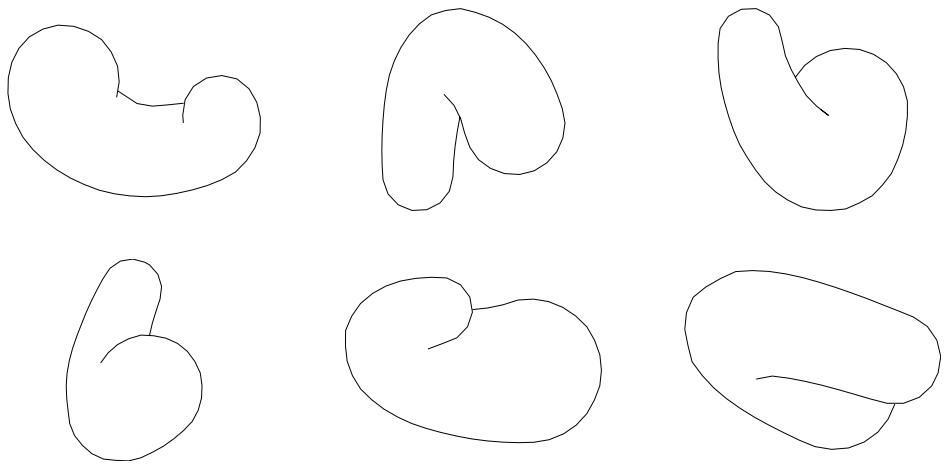


Figure 2.8: Examples of the line drawing extracted from the inflated surfaces.

The 2D contours of line drawings are quantified by *turning angles*—the angle differences between two consecutive edges along contours—which approximate curvatures for discrete curves (Figure 2.9). Then, each contour point is represented by a contour segment, which is the visual cue around the contour point in a "receptive field." Each contour segment is converted into a vector of turning angles along the segment, as shown in Figure 2.10. The size of a turning angle vector—the number of elements of the vector—represents the scope of visual cues at a contour point. The longer turning angle vector stands for the more global features of the contour, thus the larger receptive field around a contour point.

In addition to the turning angle patterns, local T-junction structures on contours are used to characterize the contour segments. T-junctions suggest occlusion structure near them; thus the head surface is figural and closer than the stem surfaces as shown in Figure 2.11. In this model, this figure/ground cue is implicitly characterized by defining the left side of a contour as figural (when traversed counterclockwise) and by differentiating the stem and internal contour segments from other ordinary contour segments. Such local T-junction information accompanying with turning angle vectors are tabulated to characterize 2D contour features and are used as a key of this database to look up possible 3D shapes.

While the contours are divide into contour segments, the inflated 3D surfaces are

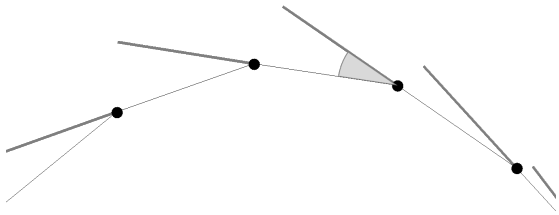
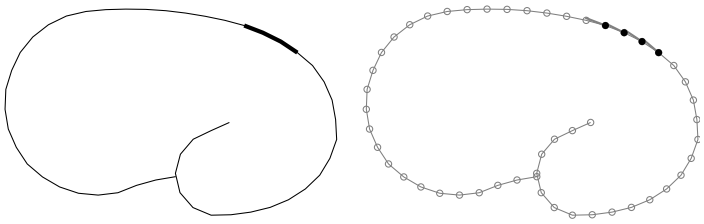
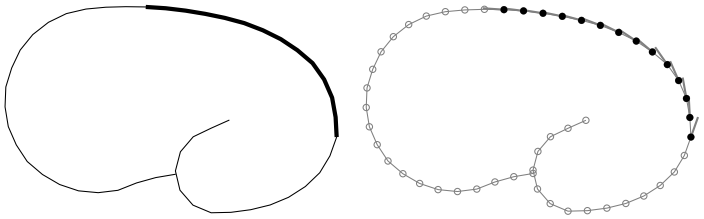


Figure 2.9: A turning angle is the angle difference between two adjacent edges at a point, the current and the previous edges along the contour counterclockwise.



(a) More-local contour cue: a shorter contour segment and turning angle vector



(b) More-global contour cue: a longer contour segment and turning angle vector

Figure 2.10: The local and global features of contour segments (bold lines) are represented by the size of turning angle vectors (black dots). The scope of cues is quantitatively characterized by the size of a turning angle vector.

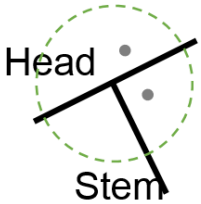
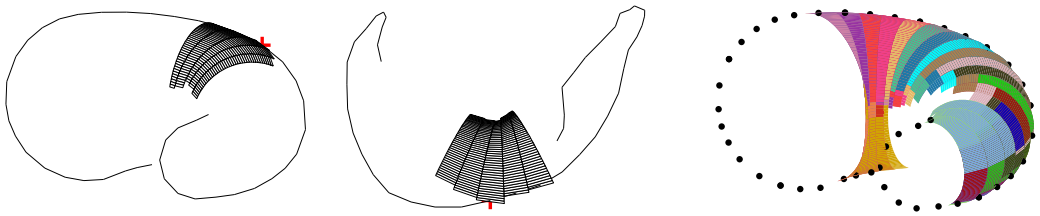


Figure 2.11: T-junction as a local depth cue. The surface owning the head contour is a figure, and the depth is closer. The surface owning the stem contour is ground, and the depth is farther.

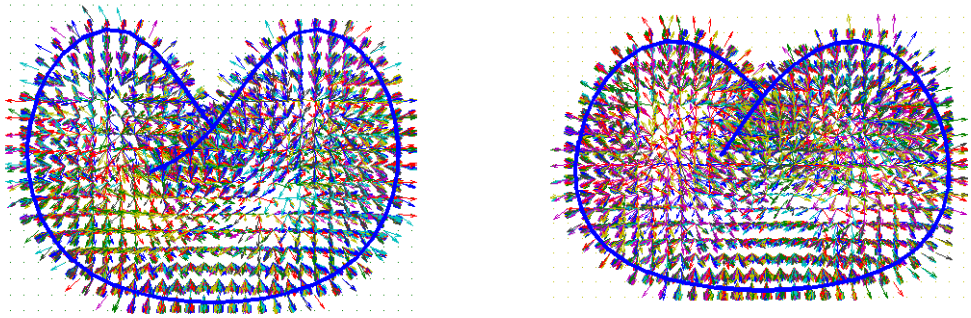


(a) local surface patch (b) Patches from a surface

Figure 2.12: Local surface patch. (a) A local surface patch at rim point (+ symbols) is seen from two different viewpoints. (b) Local surface patches are extracted at each rim point (black dots) on an inflated 3D shape.

divided into local 3D surface patches. The 3D surfaces are represented by surface orientations as in many other studies on 3D shape perception (Kent, 1983; Todd et al., 1996; Phillips et al., 1997; Todd, 2004; Norman et al., 2006). The main reason for choosing surface orientations instead of depths is that there is no ground truth zero for depth in this model. The locations of observers are not defined; therefore, the depths cannot be directly compared across the differently inflated shapes. However, surface orientations can be compared with each other because the angles are clearly determined by the direction of the line of sight. A local surface patch is defined at each rim point, as shown in Figure 2.12, and it is represented by surface orientation information (the surface normals).

While each contour segment represents the contour cues at each contour point, each surface patch represents the 3D shape at each rim point that gives rise to the contour point; therefore, contour segments and local surface patches are related one by one. The relations are saved in a database, from which possible local surface patches are searched. In this study, a total of 1550 relations between the contour segments and local surface patches are saved from 31 random line drawings. When a novel contour segment is given, the 10 most similar contour segments in the database are picked, and the corresponding local surface patches are retrieved. Euclidean distance on the space of turning angle vectors is applied to measure the similarity among contour segments (cosine similarity).



(a) Line Drawing 1 (b) Line Drawing 2

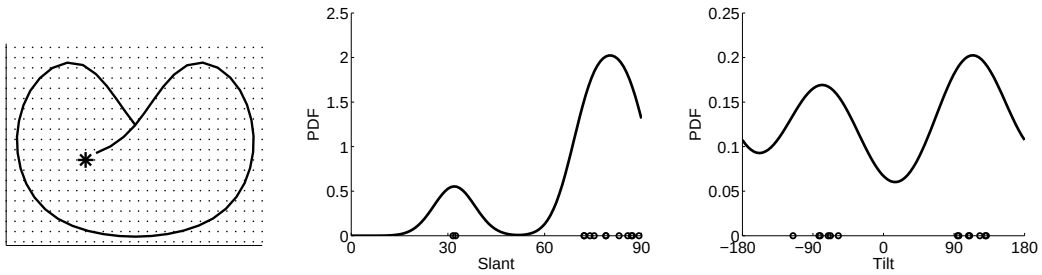
Figure 2.13: Surface normals samples (arrows) at grid points on line drawing images. All probable local surface patches retrieved from contour comparisons are overlaid, and all the normal vectors on the patches are aggregated.

2.2.3 The distribution of 3D surface orientation

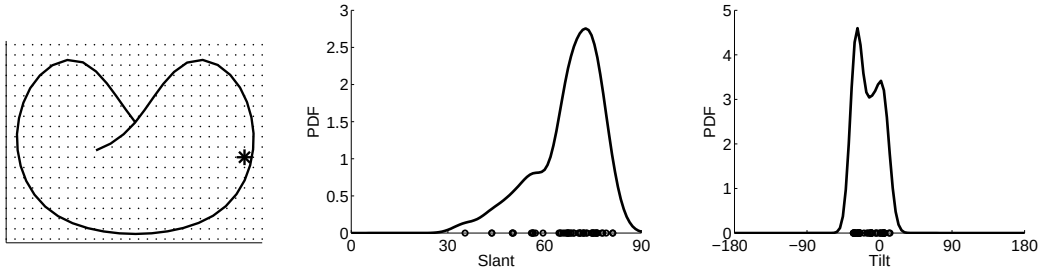
Given a new line drawing, the model looks for contour segments in the database which are similar to the contour segments on the line drawing, and it returns the probable local surface patches. Then, all retrieved surface patches along the whole line drawing contours are integrated. The surface orientation samples on the patches are aggregated to estimate distributions of 3D surface orientation.

Figure 2.13 shows the retrieved samples of surface orientations—normal vectors—at each grid point on the line drawing images. Note that the count of samples are varied from one grid point to the other. The number is dependent on the distance of a grid point from the contours. The closer grid points have a higher number of samples, and thus more information, while the farther grid points have a smaller number of samples, and thus less information. This is consistent with the spatial decaying of depth information from contours inward as discussed in the experiment. Based on the normal vector samples, distributions of 3D surface orientations at each grid point are computed. When there is no sample, the distribution is assumed to be uniform.

Estimating the distributions of 3D surface orientations, two mathematically independent elements of surface orientation, slant (the angles between the normal vectors



(a) Distributions at a near point to the internal contour



(b) Distributions at a farther point from the internal contour

Figure 2.14: Distributions of 3D surface orientations. Distributions of slant and tilt are estimated from the orientation samples (small circles on x-axes) at each grid point. (a) The distributions at a near point (* symbols) to the internal contour are wider and having clear two peaks, consistent with the possibility of surface orientations at the point. (b) The distributions at a farther point are narrower.

and line of sight) and tilt (the angles of normal vectors on the image plane), are analyzed separately because they are psychophysically different as well. The different sensitivities for slant and tilt that were observed in various studies imply two distinctive mechanisms for each in perceptual tasks (Stevens, 1983; Kent, 1983; Christou et al., 1996; Knill, 1998b; Koenderink et al., 1992; Mamassian & Landy, 1998; Norman et al., 2006). Thus, at each grid point, a distribution of slant and a distribution of tilt are separately estimated. The distributions are estimated in a non-parametric way, using *kernel density estimation* (also known as *Parzen windows*) with Gaussian kernel functions (Parzen, 1962). For tilt distributions, I used a periodic kernel density estimation module developed by Bylan Muir (Biozentrum, University of Basel, Switzerland). In consequence, there are two distributions at each grid point on a line drawing image. Figure 2.14 shows the examples of the distributions at two different grid points. Based on the distributions, the aimed likelihood function of depth difference is computed.

Before moving on to the computation of the likelihood, I present a test for the

consistency of the distributions. I estimated the distributions of 3D surface orientations for some simple line drawings, which are shown in Figure 2.15, and the Line Drawing 1 and 2, which are tested in the experiment. Then, assuming that these distributions are the posteriors of surface orientations, the values of *Maximum a Posterior* (MAP) were picked at each grid point on the image plane. The MAP slants, tilts and surface normals are presented in Figure 2.15, validating that they are perceptually correct: the estimates are consistent with human interpretation of these line drawings. For example, the MAP slants along the contours are near 90°, estimating correctly the rim on surface; the MAP normals along the internal contours are the opposite directions, illustrating two surfaces are facing each other.

2.2.4 Comparison with human response

In this section, this model computes the posterior of depth difference, $P(D|L)$, which is computed by combining the likelihood and the prior. From the posterior, the probability of depth difference at the probe dots on the probable 3D surfaces can be predicted. The probability will be compared with the congruence rates from the experiment.

The likelihood of depth difference at any two points on the line drawing is computed from the distributions of 3D surface orientations estimated in the previous section. The depth difference can be calculated by taking the integral of surface orientation because slant is mathematically the gradient of depth. Thus, the likelihood of depth difference, $P(L|D)$, is estimated as the integration of the slant distributions along a straight path connecting the two probe dots on the image.

Then, the prior for depth difference, $P(D)$, is considered by taking known human biases in depth perception. One well-known bias is slant underestimation. The perceived slants of surfaces are smaller than the actual slants, meaning that slanted surfaces are perceived closer to the fronto-parallel plane (Perrone, 1981; Mamassian & Kersten, 1996; Christou et al., 1996; Todd & Perotti, 1999; van Ee et al., 2003). Consequently, the perceived depth differences on the surfaces get smaller, which makes a non-uniform prior of depth difference. The distribution of the prior would be higher for smaller depth differences near zero and lower for larger depth differences far from zero, such

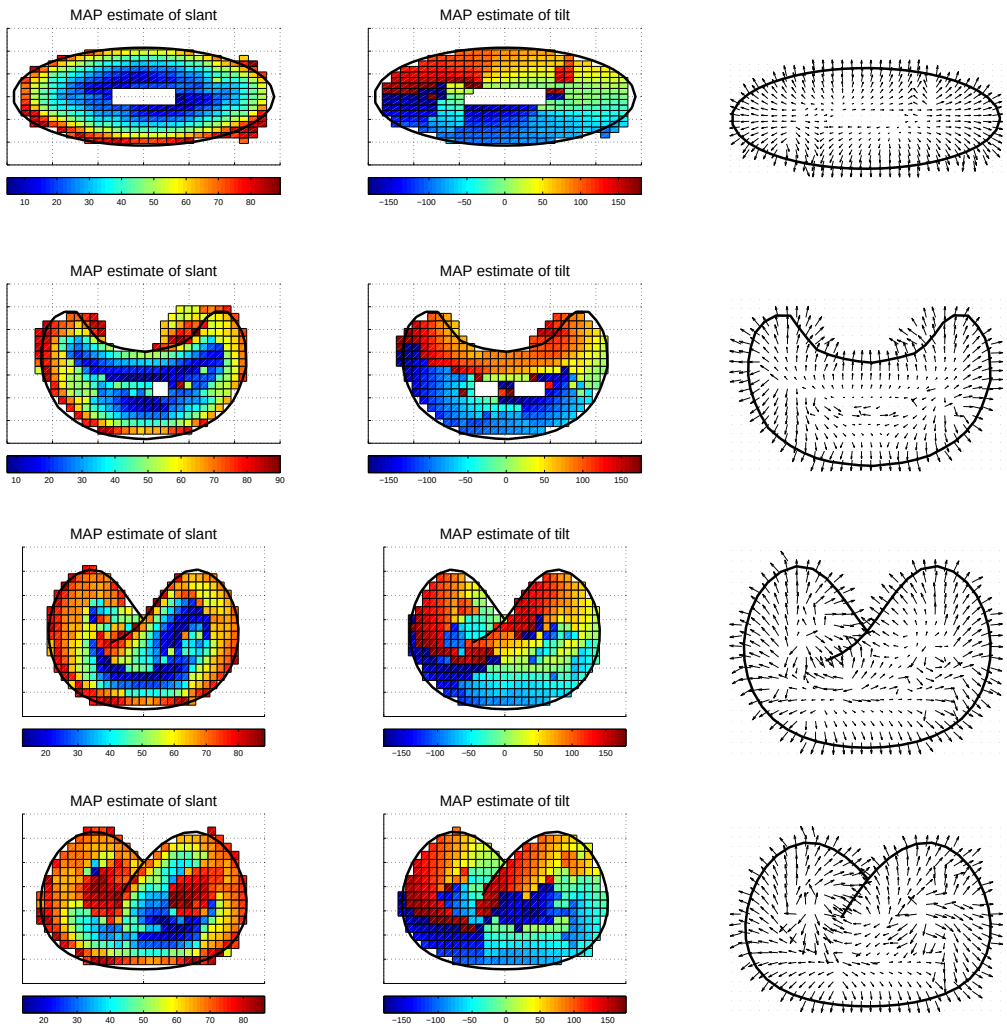


Figure 2.15: MAP estimates for slants (left), tilts (center), and surface normals (right) from the distribution of 3D surface orientations.

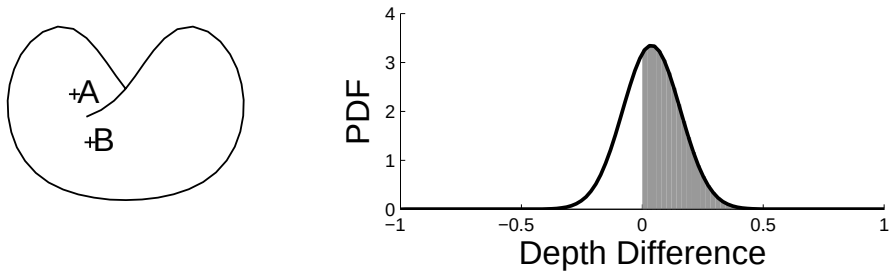
as a Gaussian. Therefore, to model this bias, a Gaussian distribution is selected whose standard deviation reflects the strength of slant underestimation. For example, the smaller standard deviation represents the stronger slant underestimation, and thus less depth difference on the surface.

The other bias is lower-region figure in figure/ground organization. The lower region is more likely to be perceived as a figure than ground, and thus to be perceived closer to the viewer (Vecera et al., 2002). This bias is also observed in experiment 1, in which the perceived depth difference was higher when the "closer" dot was located lower on the image relative to the other dot. Due to this bias, the perceived slants become skewed upward from the fronto-parallel plane, which can be modeled by a Gaussian with a non-zero mean. The mean value reflects the strength of lower-region figure bias. For example, the larger positive mean represents the stronger asymmetry of depth difference. Thus, the prior of depth difference, $P(D)$, is summarized with the pair of a mean and a standard deviation of a Gaussian.

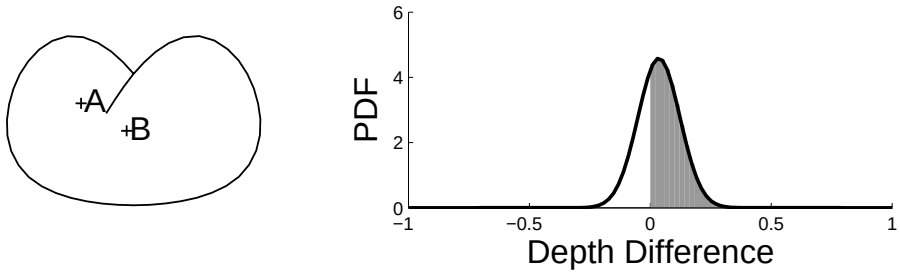
These biases could be modeled directly in terms of surface orientation as well, but for mathematical simplicity I choose to model them in terms of depth difference for the following reasons. Firstly, the slant underestimation can be modeled as a bias of slant solely, but applying the same slant priors at all grid points repeatedly is computationally less efficient than applying the prior of depth difference at once, while both results are mathematically the same. Secondly, the closer lower-region bias can be modeled with both slant and tilt, but this bias means that the slant and tilt are not independent, which requires a complicated bivariate distribution to model this. Therefore, modeling this bias in terms of depth difference is much simpler, which involves only one variable.

Finally, for two probe dots on a line drawing, the prior of depth difference, $P(D)$, is combined with the likelihood of depth difference, $P(L|D)$, to generate the posterior of depth difference, $P(D|L)$. Examples of the posteriors for a probe set on Line Drawing 1 and Line Drawing 2 are shown in Figure 2.16. In this test, the maximum-likelihood prior parameters were selected for each line drawing.

From the posterior of depth difference, the probability of depth difference for each probe set—P(B is closer than A)—is computed. The predicted probabilities across the

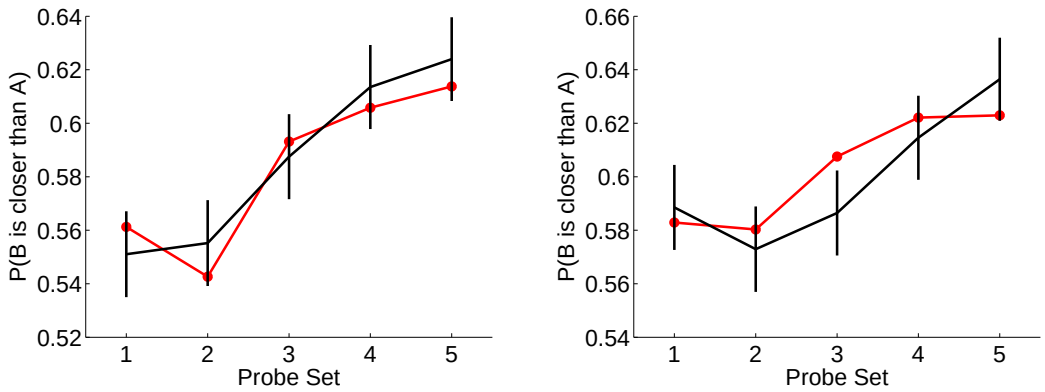


(a) Depth differerence for Probe Set 3 on Line Drawing 1



(b) Depth difference for Probe Set 3 on Line Drawing 2

Figure 2.16: The posteriors of depth difference between two probe dots A and B. The range of depth difference is scaled so that the distance between two probe dots is one, and positive depth difference means the depth at B from the observer is smaller than the depth at A, thus B is closer. Therefore, P(B is closer than A) is the area under the curve of the posterior (shades).



(a) Line Drawing 1 (b) Line Drawing 2

Figure 2.17: Comparison between P(B is closer than A)s predicted from this model (red dots) and the congruence rates with the standard errors from the experiment (black lines).

probe sets for each line drawing were compared with the congruence rates from the experiment (Figure 2.17). The certainties of depth difference perceptions and congruent rates were consistent, showing human depth judgments on the line drawings were accounted for by this model. Such comparisons confirm that this computational model explains human interpretation of 3D shapes from line drawings.

So far all the probabilities of depth differences are computed with a fixed scope of contour cues. The following analysis further investigates the role of non-locality of contour cues on 3D shape inference with this model.

2.2.5 The influence of local and global cues

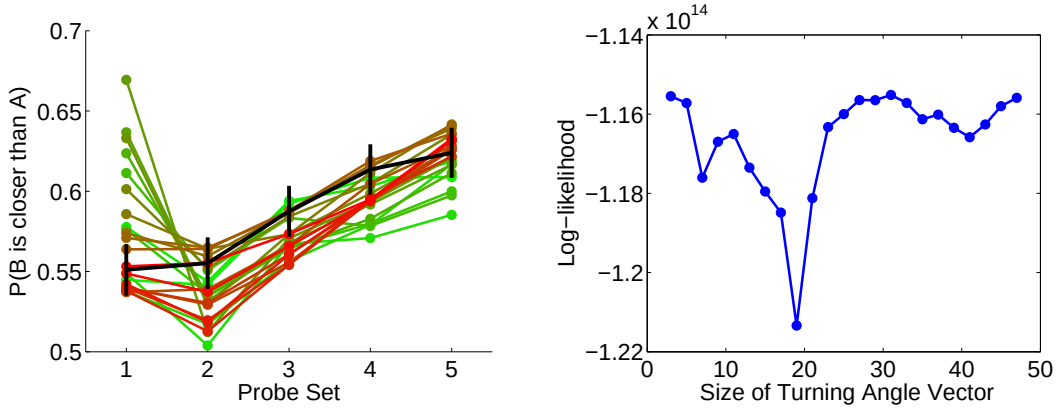
As suggested above, estimation of local surface normals is based on a combination of local and global cues, which interact in potentially complex ways. In this section, I look more closely at how the degree of locality of each shape cue (that is, the size of the neighborhood over which evidence is integrated) influences subjects judgments.

As discussed about the database for 2D contour features, the scope of contour cues is represented by the size of turning angle vectors. The shorter turning angle vector represents the more-local contour cues, and the longer turning angle vector represents the more-global contour cues. At each size of turning angle vectors, the probabilities of depth difference—P(B is closer than A)—are predicted and compared

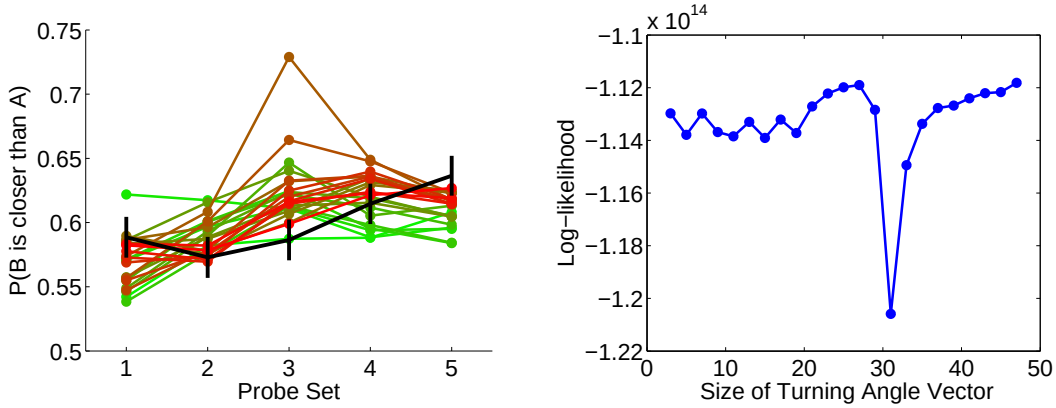
with the congruence rates from the experiment (Figure 2.18).

This comparison shows that as the cue scope increases, the predictions converge on the experimental data. To quantitatively examine the quality of fit at each cue scope to the data, the log-likelihood is computed assuming that each depth judgment at a probe is a Bernoulli procedure with the predicted probability, P(B is closer than A). The log-likelihoods for Line Drawing 1 and Line Drawing 2 have similar patterns. They are fluctuating for the shorter turning angle vectors and then converging to an asymptote for the longer turning angle vectors. Such patterns signify that global features of contour play important roles in interpreting 3D surface and making decisions about the depths, but the global context helps only up to a point where the influence is saturated.

Just before approaching the asymptote, both log-likelihoods make a deep valley although the scales at which the valley occurs are different for the two line drawings. The deep valley for Drawing 2 happens later than for Drawing 1, and then both go to the relatively stable state. This implies that Drawing 2 needs more-global features to interpret it properly. This difference might be caused by the distinction in the geometrical properties of the two drawings. The contour of Line Drawing 1 is theoretically consistent with a smooth-surface interpretation, but the contour of Drawing 2 is not because its internal contour ends convex. As discussed in Methods, contours of a smooth surface must end concave (Koenderink & van Doorn, 1982; Koenderink, 1984). This suggests that interpreting the surface from Line Drawing 2 may be more dependent on the global contour features because its local contour features are conflicting with a natural smooth surface. Considering humans generally do not discriminate this difference easily, the difference observed here can suggest there is a possibility that human vision selects an appropriate level of cue scope to infer the reasonable surfaces from varied contours even when they are geometrically inconsistent with a smooth form. However, this issue clearly needs further exploration with a wide range of stimuli before any strong conclusions can be drawn.



(a) The influence of cue scopes for Line Drawing 1



(b) The influence of cue scopes for Line Drawing 2

Figure 2.18: The influence of cue scopes. The predictions of congruence rate from this model, P(B is closer than A), are plotted in color varied from green to red as the size of turning angle vector increases, which are compared with the congruence rates measured in the experiment (black line). The log-likelihoods of predictions at each cue scope level are plotted as a function of the cue scope.

2.3 General Discussion

To understand the mechanism of 3D shape inference from line drawings, I tested a psychophysical experiment on the interpretation of line drawings and suggested a computational model that explains the results. From the experiment, I concluded that human depth judgments reflect the underlying probability of possible 3D shapes induced by a line drawing. The possible shapes are perceived as the result of complex interaction between local and non-local contour cues. In this complex interaction, the depth information is propagated from the contour cues, and as a result, the influence of local T-junction cue is decayed as the distance from the local cue increases. I modeled the inference of 3D shapes for a line drawing based on the interaction of contour cues in a Bayesian framework. In this probabilistic model, the probability distribution over possible 3D shapes is estimated for the line drawing, and perceptual decisions about the line drawing are made based on the distribution. The model predicts the probability of depth difference between the two dots on a line drawing from the posterior distribution of depth difference, which represents the underlying probability of possible 3D interpretations. The predictions from this model were consistent with human responses for the line drawings in the experiment, which confirms this model provides a computational framework for 3D shape perception from line drawings.

The study of perceiving 3D shapes from line drawings is especially interesting. Line drawing images have very limited depth cues compared with other types of images with shading and texture because the only available depth cues in line drawings are the contours. While most of line drawing images are empty without any local cues for surface depth at all, humans naturally interpret line drawings as 3D surfaces by filling in the blank areas without difficulty. This shows that the contours in line drawing images connote adequate information to produce the 3D percept (Koenderink et al., 2012). Pizlo et al. (2010) have also emphasized that contour is essential to describe 3D shapes, and the contours in a line drawing are sufficient to produce the percept of a 3D object with clearly defined surfaces.

Traditionally, local junction geometry has been considered a conclusive and virtually

deterministic cue to 3D surface structure (Huffman, 1971; Clowes, 1971; Waltz, 1972; Mackworth, 1973; Charkravarty, 1979; Malik, 1987), but this approach has not been directly compared with human 3D interpretation of line drawings. Koenderink et al. (2012) measured the 3D surfaces perceived from Picasso's line drawings, using gauge figures to reconstruct 3D surface maps for the drawings, but this approach is also deterministic. In the current study, I devised a simpler set of line drawings for smooth objects and picked natural and easier visual tasks about the line drawings to examine the role of contour cues—the depth comparison on the perceived 3D surfaces induced by the contour cues. The depth judgments were tested by probing the depth order between pairwise dots that were superimposed on each line drawing. The probes were placed along the internal contour of a T-junction with varied distances from the junction to examine the influence of the local junction cue. Human observers were asked to indicate which dot appeared closer for each probe, and I measured the congruence rate, the rate of picking the "closer" dot. The rate revealed the underlying probability of the perceived depth difference, and the uncertainty for the depth difference systemically increased as the distance of probes from the T-junction increased.

The result from the experiment indicates that the influence of local T-junction cue is not deterministic, contrary to the deterministic role of the T-junction cue in traditional line drawing interpretations (Huffman, 1971; Clowes, 1971; Waltz, 1972; Mackworth, 1973; Charkravarty, 1979; Malik, 1987). The influence is rather probabilistic. The influence of T-junction cue spatially propagated while the strength decayed. The propagation here focuses on spatial spreading than temporal spreading as in the method suggested by (Weiss, 1997), and the spatial spreading is also different from the propagation of contours suggested by Tse (2002), because what propagates here is not the contour itself but the depth information from the contour. The spatial dependency of the strength implies that the local depth cue is coupled with non-local cues to create the possible global 3D percepts. The influence of local and non-local contour cues are combined and fill inward the empty areas, and the complex interaction generates not a deterministic interpretation for the 3D surface. Such influence of global context is not new in visual perception. Koenderink et al. (2012) have pointed out that local cues in

interpretation of a line drawing are "overridden" by context. The interaction of local and global information is also an important topic in perceptual grouping. Therefore, I suggested a model that simulates the complex interaction of local and global cues for the probabilistic 3D shape inference from the contours of a line drawing.

I selected a Bayesian framework to model the 3D shape inference. The method of Bayesian inference has been applied to model various human behaviors in perception and cognition (Mamassian & Landy, 1998; Yuille & Bulthoff, 1994; Mamassian et al., 2002; Kersten & Yuille, 2003; Kersten et al., 2004; Mamassian, 2006; Yuille & Kersten, 2006; Maloney & Mamassian, 2009; Ma, 2012). In this Bayesian framework, the likelihood function and the prior are combined to compute the posterior on which the perceptual decisions are based. This framework provides quantitative measures to represent and to analyze the uncertainty of 3D shape interpretation observed in the experiment, by using the probability distributions. The distributions represent the ambiguity of 3D surfaces propagated from the depth cues in contours. Also, the probabilistic representation of 3D shapes provides richer information about the quantity of belief for the interpreted 3D shapes instead of a single specific surface orientation or depth map. In addition, this probabilistic representation for shape information is feasible for biological systems, which can encode the distributions by the population of neurons (Zemel et al., 1998; Pouget et al., 2000). Moreover, the Bayesian framework accommodates to incorporate human biases in depth perception as the form of prior probability distribution.

The likelihood of depth differences on perceived 3D surfaces is computed by assuming a generative model—the inflation model. The inflation model creates random line drawings from random 3D surfaces that are inflated from skeletons, and the relations of many random line drawings and the corresponding 3D surfaces are kept in a database. The database keeps the information of local 3D surfaces for 2D contours with a specific cue scope, and the likelihood is numerically computed from the database. Such an approach has several advantages. Firstly, this method avoids complex computing to get the analytical solution of the likelihood function by using random samples of the 3D surfaces retrieved from the database to estimated the function. Secondly, this

method defines the likelihood at different scopes of 2D contour cues in a easier way by using the size of the turning angle vector to represent the scope of a contour cue. In addition, any existing natural bias in the distribution of surface orientations—such as any natural distributions over slant and tilt for the type of shapes generated by the inflation model—is already incorporated in the database, so no correction is needed. Note that the existing human biases in slant and tilt perception are separately applied as the prior.

The predictions for the probability of depth difference from the model were consistent with human responses from the experiment, showing that our model reasonably explains human 3D shape interpretation from line drawings based on the estimated posteriors. This Bayesian model accounts for the inference of 3D shape information from 2D line drawing contours, including surface orientations and depths in the empty areas without other direct depth cues at the locations, and encodes the underlying uncertainty contributed to the propagation of indirect depth cues as the posterior probability distributions, which can convey richer information for the later cognitive processes to make decisions about the 3D shapes. Also, this model provides a good quantitative tool to examine the complex influence between local and global cues on shape inference, which is a central challenge in perceptual grouping. The analysis for the varied contour cue scope suggests that human vision may choose the appropriate scope of contour cues to estimate the global 3D surface, considering the local and non-local features without awareness, and this computational model suggest the mechanism.

2.4 Conclusion

This chapter investigated 3D shape inference from line drawings. Line drawings representing simple forms were picked to test human interpretation about the pictorial surfaces. The experiment revealed that human depth judgments based on the perceived surface were not deterministic but rather probabilistic; the judgment rates were matched with the probability that encodes the uncertainty in perceiving 3D shapes. The perceived 3D surfaces can be explained as the result of propagation of depth information from both local cues and global cues in line drawing contour structures.

The computational model was suggested for this 3D shape interpretation in a Bayesian framework, which encodes the probabilistic nature of estimating 3D surfaces, incorporates human biases in 3D shape perception, and provides a tool for simulating the 3D shape inference with respect to various factors, such as contour cue scopes.

Chapter 3

Perception of Contour Types: Depth Cues vs. Surface Markings

Contours in line drawing images convey 3D shape information, but the interpretation is dependent on the global context in the line drawing images. Contours in images can be caused by various discontinuities in depth, surface orientation, and surface color on the surface that gives rise the contours (Stevens, 1981; Barrow & Tenenbaum, 1981; Marr, 1982; Koenderink & van Doorn, 1982; Koenderink, 1984; Knill, 1992; Mamassian & Landy, 1998; Phillips et al., 2003). The contours are interpreted as either occluding contours, which represent depth structures and provide direct information about the shapes of surfaces, or as surface markings, which represent the changes in surface reflectance and provide constrained information about the shape. For example, at an intersection of two contours (a T-junction), the internal contour can be interpreted either as an occluding contour or as a surface marking. Deciding the reliability of a contour as a depth cue is not trivial process, but is a complex inference based on global context with non-local contour features in line drawings.

3.1 Experiment 2

The purpose of this experiment was to examine the influence of global context on interpreting a contour segment as an occlusion cue or a surface marking. Based on the original line drawings used in Experiment 1 (Line Drawing 1 and Line Drawing 2), a contour segment is added or deleted to change the junction structures, and the change would make a different interpretation of the 3D surfaces and depths. I used the same pair-wise depth comparisons for the same line drawings as in Experiment 1.

3.1.1 Methods

Participants

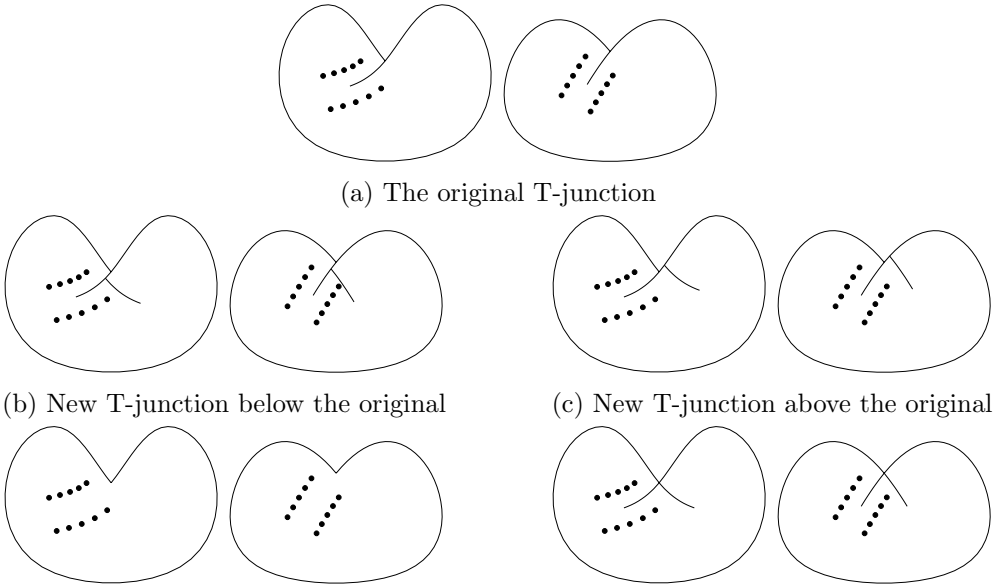
Sixteen Rutgers undergraduate students participated in this experiment to get course credits.

Stimuli

Based on the same line drawings used as in Experiment 1, four other types of junctions were made (Figure 3.1). A contour segment that had the same shape with the original internal contour of the original T-junction was added either below or above the original junction on the opposite side of the original stem contour, as shown in Figure 3.1 (b) and (c). The addition generated another T-junction that made conflicting depth relations with the relations that were created from the original T-junction as an occlusion cue. In addition, removing the segment from the original T-junction made an L-junction, and adding the segment on the other side at the original T-junction made an X-junction, as shown in Figure 3.1 (d) and (e). Both L- and X-junctions provide very ambiguous depth information. The 3D interpretations were measured by probe sets, which were the pairs of two color dots along the original internal contour, as was similar in Experiment 1.

Design and Procedure

The locations of probes and the types of junctions were the main factors manipulated in this experiment. Five probe locations (Probe Sets 1 to 5) were tested from farther to closer to the original junction along the internal contour, and five types of junctions were tested, as shown in Figure 3.1. Two different types of base line drawings (Line Drawing 1 and 2) were used. Each line drawing image was horizontally flipped left or right, and the color of the two probe dots were changed between cyan and magenta. The orientation of the whole stimuli was fixed as shown in Figure 3.1. Each stimulus was displayed for 300, 750, or 1200 ms. Each condition was repeated twice; therefore, a total of 1200 trials were tested. Participants were asked to indicate which dot appeared closer to them, and the congruence rates were measured.



(a) The original T-junction

(b) New T-junction below the original

(c) New T-junction above the original

(d) L-junction

(e) X-junction

Figure 3.1: Five junction types for the set of Line Drawing 1 (left) and Line Drawing 2 (right). Based on the original line drawing with one T-junction (a), a contour segment that had the same shape and length with the internal contour of the T-junction was added near the junction (b and c), or the internal contour was deleted or added to generate various types of junction (d and e). The dots show the locations of probe sets.

3.1.2 Results and Discussion

Firstly, the influence of probe locations was significant according to an ANOVA analysis ($F(4, 60) = 4.532,\ p = 0.00289$), and the log odds of congruent responses were positively related to the distance of probes from the original T-junction according to a logistic regression ($z = 6.073,\ p = 1.25e^{-09}$) (Figure 3.2). Such results confirm again that the interaction of local T-junction cue and non-local cues influenced the depth difference perception that was observed in Experiment 1.

Secondly, there was a significant difference in the congruence rates with the original T-junction only and X-junction ($p < 0.03$), but there were no significant differences among the other junction types, from an analysis of a pairwise test. The facts that the additional T-junctions did not make differences in the depth perception with the original T-junction only line drawings, and that there was no distinction between the above and below T-junctions imply that the additional junctions were interpreted not as a meaningful depth cue but rather a noise. Therefore, the stem contours were

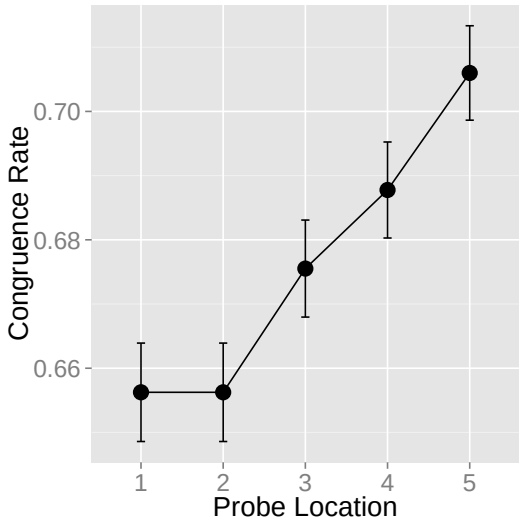


Figure 3.2: The influence of probe locations on congruence rates. The data points show the means of congruence response, and error bars show the standard errors.

considered as surface markings instead of occluding cues.

Interestingly, the congruence rates for the line drawings with L- and X-junctions were not constant although the surfaces perceived from L- and X-junctions were expected to be "flat," because both types of junctions do not provide depth information from the contour structures. However, the congruence rates with L- and X-junctions reflected the certainty of depth difference perceptions, which varied according to the location of probes, showing that the line drawings were interpreted as 3D surfaces and the depth differences were perceived. The 3D shape interpretation from L-junctions can be explained that subjects could learn the pattern of the probe locations and internal contours from the trials with T-junctions, and they assumed the internal contours that were not existing in L-junctions. Similarly, the contour that separated two probe dots in an X-junction was assumed to be the internal contour of a T-junction, representing depth relations, and the other contour was assumed to be a surface marking.

Taken together, the results show that the interpretation of contours as an occlusion cue or a surface marking is not deterministic but conditional based on the global context. Not all T-junctions are interpreted as depth cues. When two T-junctions are competing as depth cues, one is considered as a legitimate depth cue, but the other is dismissed as an informative depth cue based on the global context. Therefore, the dismissed junction

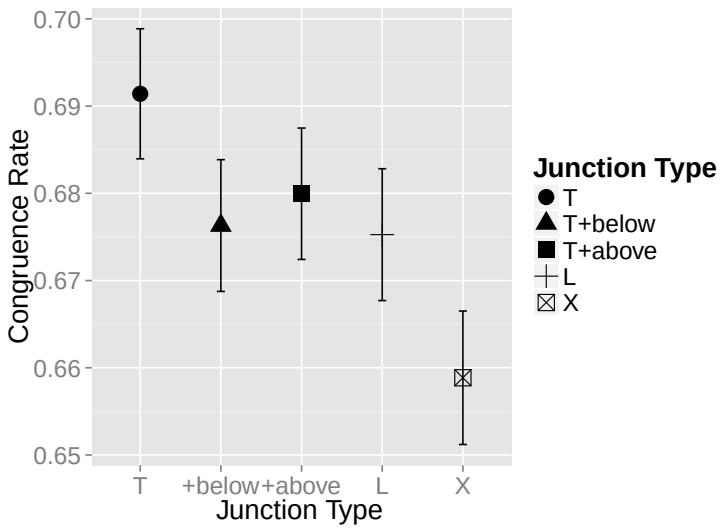


Figure 3.3: The influence of junction types on congruence rates.

contours are interpreted as surface markings. On the other hand, junctions that do not convey depth information, such as L- and X-junctions, can be interpreted as depth cues based on the context; as the L- and X-junctions created similar 3D shape perceptions as the T-junctions in this experiment.

Additionally, there was a significant difference in the responses between two types of line drawings (Line Drawing 1 and Line Drawing 2) ($F(1, 15) = 6.062,\ p = 0.0264$). Such a result is different from Experiment 1 in which there was no distinction between Line Drawing 1 and Line Drawing 2. There can be an influence of the shape of head contours—convex head or concave head—that made the differences in the current experiment, which tested line drawings with "noisy" contours. Convexity of a contour is a well known figural cue in figure/ground organization (Kanizsa & Gerbino, 1976; Pomerantz & Kubovy, 1986; Driver & Baylis, 1995). Furthermore, Burge et al. (2010) argued that the convexity is a depth cue that provides information about absolute depth differences as disparity does in natural viewing. The local geometry of contours—the curvature of head contours of T-junctions—will be tested in the upcoming experiments in the next chapter.

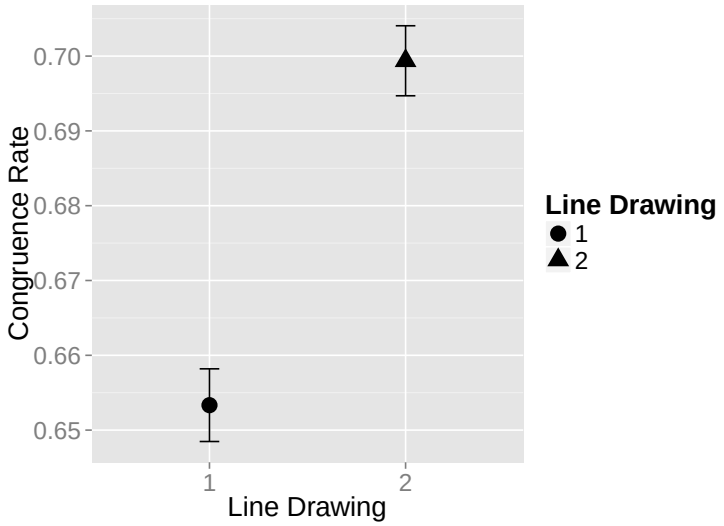


Figure 3.4: The influence of line drawing types on congruence rates.

3.2 Conclusion

Contours in line drawings are differently interpreted according to the global context in the line drawing images. The contours are interpreted as either occluding contours, which represent depth structures and provide direct information about the shapes of surfaces (Marr, 1977), or as surface markings, which represent the changes in surface reflectance and provide constrained information about the shape (Knill, 1992). For example, at an intersection of two contours (a T-junction), the internal contour can be interpreted either as an occluding contour or as a surface marking. Deciding the reliability of a contour as a depth cue is not trivial process, but is a complex inference based on global context with non-local contour features in line drawings.

The purpose of this experiment was to examine the influence of global context on interpreting a contour segment as an occlusion cue or a surface marking. Based on the original line drawings used in Experiment 1 (Line Drawing 1 and Line Drawing 2), a contour segment is added or deleted to change the junction structures, and the change would make a different interpretation of the 3D surfaces and depths. I used the same pair-wise depth comparisons from the same line drawings as in Experiment 1.

In this chapter, the experiment that tested the influence of global context to disambiguate the role of junction contours in line drawing images. The interpretation of

contours of junctions in line drawings are dependent on the global context. Not all the T-junctions are interpreted as depth cues that convey depth information, and contours of junctions that are not depth cues can be interpreted as depth cues.

Chapter 4

Line Drawings Viewed Through an Aperture: Depth Judgments and the Quantity of Information Available

Interpretations of line drawings are influenced by non-local visual cues. Experiment 1 tested the complex interaction between local and non-local cues that contributes to the uncertainty of the interpretations. In the following two experiments, the amount of non-local visual cues are varied, and the influence on perception of depth difference is examined. Experiment 3 explores how the amount of visual information at a T-junction influences the depth judgments near the junction, and Experiment 4 examines the relation of the amount of visual information and the depth interpretation with more details.

4.1 Experiment 3

The aim of this experiment is to examine the influence of non-local visual information in line drawings on the 3D shape interpretation. Parts of line drawing images are presented via circular windows, which are concentric circular apertures, showing the same T-junction at the centers and the neighboring contours. The area of visible neighboring contours is varied according to the size of windows. The change in the size of non-local visual cues would influence the interpretation of 3D surfaces and perception of depth differences.

4.1.1 Methods

Participants

Thirteen Rutgers undergraduate students participated in this experiment to get course credits.

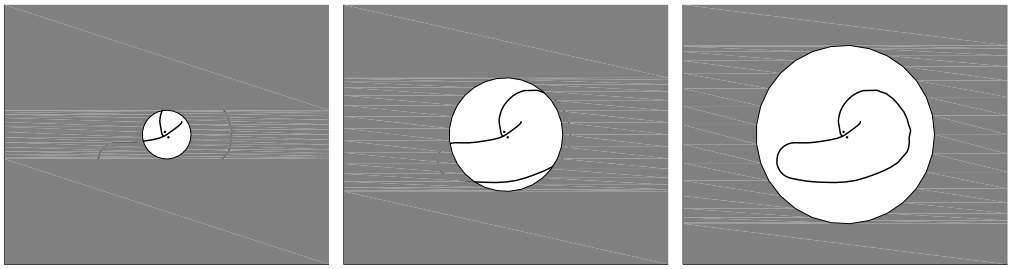
Stimuli

Line drawings were generated from the generative model explained in section 2.2 (Model), which inflates random 3D surfaces from curved skeletons and projects the surfaces into 2D contours from random viewpoints. The line drawings were black contours on white background, and the size of the whole line drawing contours subtended a maximum of 8.5° in visual angle horizontally and vertically. Each line drawing was seen through a circular window, which was a hole on the black area covering the whole screen; therefore, only part of contours were visible (Figure 4.1). The same pair-wise dots as in the previous experiments were used to probe the depth difference perceived from line drawings.

Design and Procedure

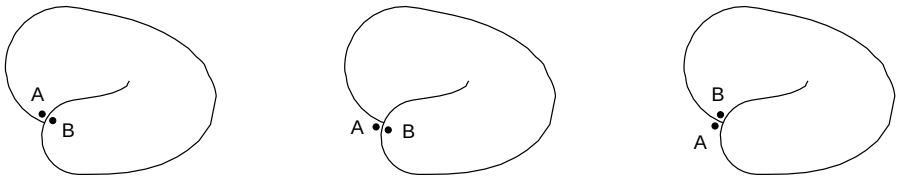
The three main factors tested in this experiment were window sizes, probe locations, and head shapes. The diameters for small, medium, and large windows were approximately 3.3°, 8.0°, 12.7° in visual angle, respectively (Figure 4.1). Probes were put on one of three contours of a T-junction (Figure 4.2): the internal contour (part of the head inside), the head contour (part of the head on the bounding contour side), or the stem contour. The centers of circular windows were positioned at the midpoint of each probe. Eight different line drawings having only one T-junction were used. Half of them had convex head contours, and the other half had concave head contours (Figure 4.3).

Each line drawing image was horizontally flipped in random, and rotated by a random angle. Each stimulus was displayed for 400ms or 800ms. The probe dot color was flipped between cyan and magenta. Each condition was repeated three times; thus, a total of 864 trials were tested.



(a) Small window (b) Medium window (c) Large window

Figure 4.1: Three window sizes.



(a) Internal probe (b) Head probe (c) Stem probe

Figure 4.2: Three probe locations.

(a) Convex head (b) Concave head

Figure 4.3: Two types of head contours.

Window Size	small	medium
medium	< 0.00001	-
large	< 0.00001	not sig.

Table 4.1: The p-values from a pairwise t-test for window size.

Before the trials, subjects were presented with line drawing images of natural objects, such as fruits, vegetables, and animals, which were selected from the study by Snodgrass & Vanderwart (1980). On the images, color probe dots were put on each line drawing, and subjects were asked to verbally answer which dot appeared closer. They got the feedback of the "correct" answers that were consistent with the ordinary interpretation of line drawings. Then, the subjects were explained that they would observe similar line drawings in the experiment. In trials, the blocks with small, medium, and large windows were tested in this order, to prevent subjects from memorizing the line drawings. The procedure was the same as in the previous experiment, and the same congruence rates were measured.

4.1.2 Results and Discussion

Firstly, the influence of window sizes on congruence rates was significant in an ANOVA analysis ($F(2, 24) = 22.01$, $p = 3.73e^{-06}$) (Figure 4.4). According to a further pairwise t-test, the rates for smaller windows were significantly different from the rates for medium and largest windows, but the rates for medium and large windows were not significantly different (Table 4.1). The results showed that the 3D shape interpretations of line drawings were qualitatively distinctive between with and without the non-local cues in line drawings, although the local T-junction was always given. Considering that the large window showed the whole line drawing contours, the partial contours in the medium windows provided commensurable information to interpret the 3D shapes.

Secondly, the influence of probe locations was significant in the ANOVA analysis ($F(2, 24) = 5.129$, $p = 0.014$) (Figure 4.5). According to a further pairwise t-test, the congruence rates at stem probes were significantly different from the congruence rates both at internal and head probes, but the rates at internal probes and at head probes were not significantly different (Table 4.2). Also, the rates were higher at internal

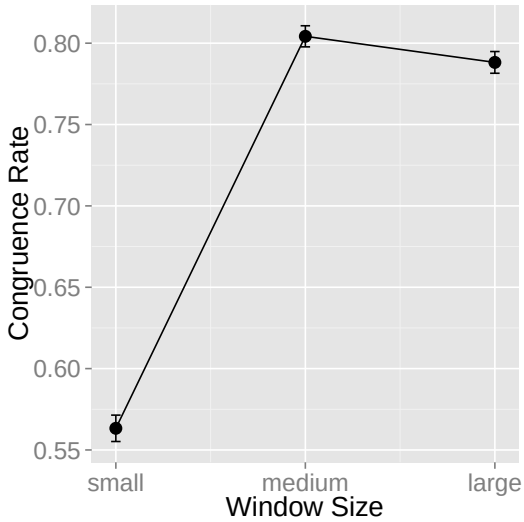


Figure 4.4: The influence of window sizes on congruence rates.

Probe Location	internal	head
head	not sig.	-
stem	< 0.00001	< 0.00001

Table 4.2: The p-values from a pairwise t-test for probe location.

and head probes than at stem probes. The depth differences across the different probe locations could not be directly compared because the depth of ground in the line drawing images was not defined. However, the same probability of depth differences at internal and head probes was consistent with the fact that both internal and head probes were, in fact, dots along the same head contours of the T-junctions.

Thirdly, the influence of head shapes—convex or concave—was not significant.

In addition, the influence of durations for display was significant ($F(1, 12) = 6.087,\ p = 0.0297$), and the congruence rates were higher for the longer durations (Figure 4.6). The result is in agreement with the fact that the longer display provided more time to understand the contour structures in line drawings, leading to the higher certainties in depth difference perceptions for the line drawings.

Also, there was a significant interaction between the repetition of trials and size of windows ($F(4, 48) = 5.534,\ p = 0.000956$), showing that the repetition increased the congruence rates only when the window sizes were small (Figure 4.7). Such a result suggests that the line drawings shown through the medium and large windows provided

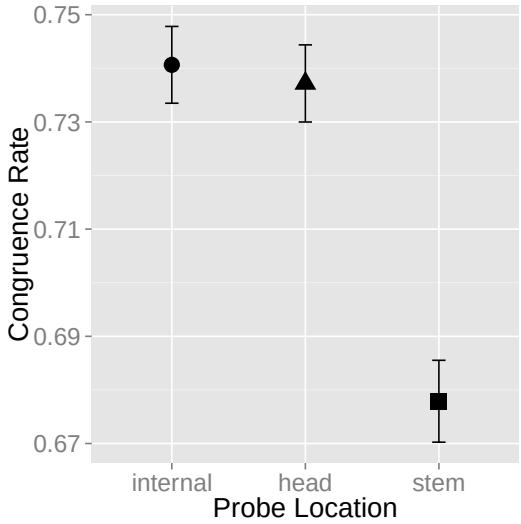


Figure 4.5: The influence of probe locations.

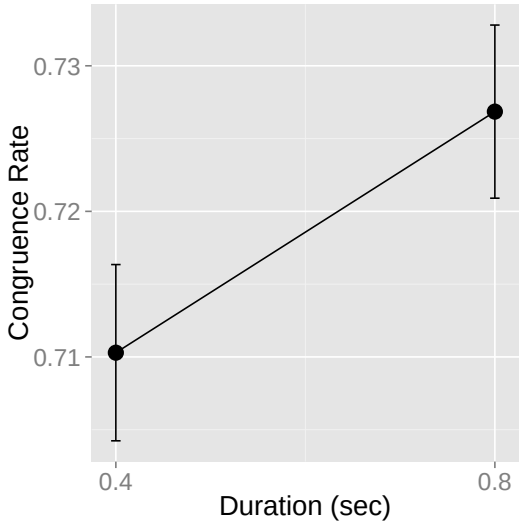


Figure 4.6: The influence of durations of display.

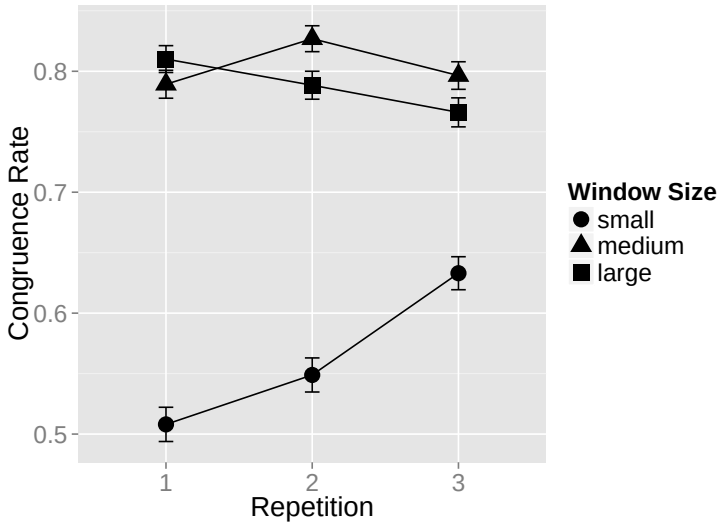


Figure 4.7: The interaction between repetition and window sizes on congruence rates.

sufficient visual cues while the line drawings shown through the small windows provided deficient visual cues. The subjects could learn from the partial visual cues in the small windows by repetition to interpret the 3D shapes and to be more certain about the perceived depth differences.

4.1.3 Conclusion

In this experiment, the same line drawings were shown through circular windows with varying sizes, and the perceived 3D shapes were measured by pairwise depth comparison tasks to examine the influence of non-local visual cues. The amount of non-local visual cues made the difference in 3D shapes and depth perception. Different window sizes allowing different extent of visible contour structures changed the certainties of depth difference perceptions on the same line drawings. In addition, there were learning effects; the certainty of depth difference perceptions got higher as the line drawings were displayed longer and repeated more. Yet, the learning from repetition was also limited only when the amount of visual information was small. All of these results imply that the line drawings shown through the small windows and through the medium or large windows were categorically different. This shows that the depth information in the small windows was not sufficient, but the information in the medium windows was

sufficient to interpret the visible contour features into the same 3D shapes with the full information in the large windows. There can be a jump in 3D shape interpretation between the small and medium windows due to the qualitative difference of depth information, or there can be a gradual increase, but it could not be examined in this experiment. In the next experiment, the window sizes will be smaller and more numbers of sizes will be used to quantitatively analyze the changes of 3D shape interpretations by the amount of visual information in the windows.

4.2 Experiment 4

The previous experiment confirmed that the amount of visual information made qualitative differences in 3D shape interpretation of line drawings, using circular windows with varying sizes that showed partial contours of line drawings. In this experiment, the influence of the amount of visual information is examined more closely, using smaller windows with more numbers of sizes than Experiment 3 for the same stimuli as used in Experiment 3, to investigate a quantitative relation between the amount of information and the depth judgment.

4.2.1 Methods

Participants

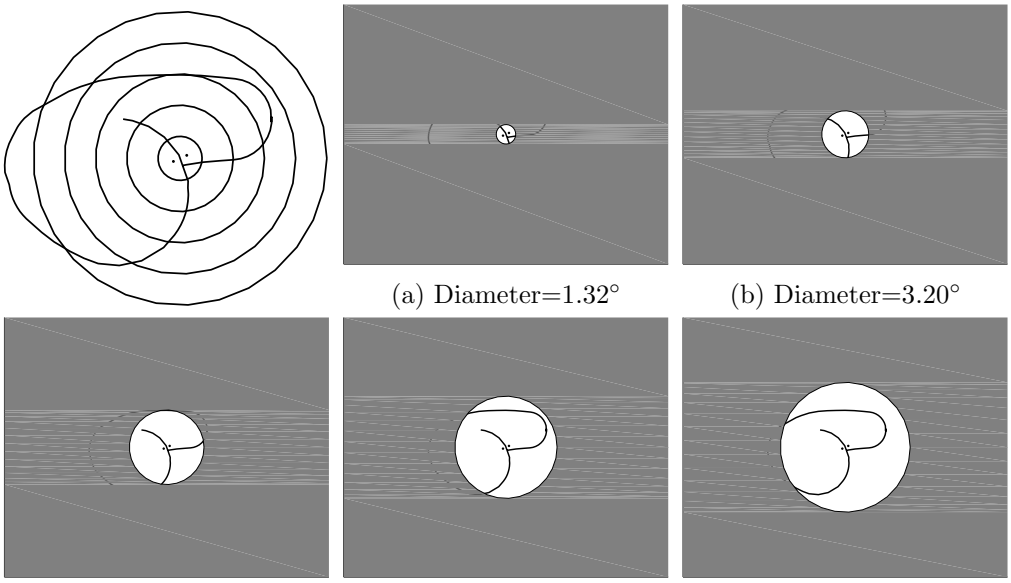
Thirteen Rutgers undergraduate students participated in this experiment to get course credits.

Stimuli

The same stimuli were used as in Experiment 3.

Design and Procedure

The same design with Experiment 3 was used except the duration of display, which was fixed to 800 ms in this experiment, and the sizes of circular windows, which were five different sizes. The diameters of the windows were 1.32°, 3.20°, 5.08°, 6.96°, 8.84°



(a) Diameter=1.32° (b) Diameter=3.20°

(c) Diameter=5.08° (d) Diameter=6.96° (e) Diameter=8.84°

Figure 4.8: Five window sizes.

in visual angle, respectively. (Figure 4.8). The smallest window in this design was smaller than the small window in Experiment 3, and the largest window was a little bit larger than the medium window in Experiment 3. Only partial contours were visible across the different window sizes. The same pairwise probe dots were positioned at one of three places on the internal, head, and stem contours of a T-junction (Figure 4.2) as in Experiment 3. The same eight line drawings were used (Figure 4.3) as in Experiment 3. Two colors were switched for the probe dots. Each line drawing image was horizontally flipped in random, and rotated by a random angle. Each condition was repeated three times; therefore, a total of 720 trials were tested. The five blocks were tested in the order of window sizes from smallest to largest. The same congruence rates were measured.

4.2.2 Results and Discussion

Firstly, the effect of window sizes on congruence rates was significant ($F(4, 48) = 6.126,\ p = 0.000458$), which confirmed again the influence from the different amount of non-local visual cues as observed in Experiment 3 (Figure 4.9). In addition, according to a logistic regression, the congruence rates increased as the window sizes increased

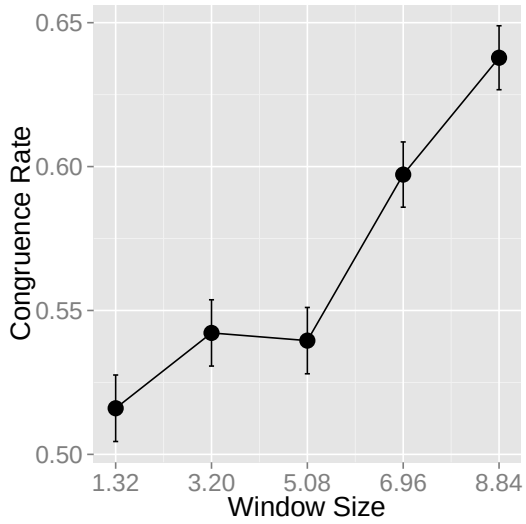


Figure 4.9: The influence of window sizes on congruence rates.

Window Size	1.32	3.20	5.08	6.96
3.20	not sig.	-	-	-
5.08	not sig.	not sig.	-	-
6.96	< 0.00001	< 0.005	< 0.005	-
8.84	< 0.00001	< 0.00001	< 0.00001	< 0.05

Table 4.3: The p-values from a pairwise t-test for window size.

($z = 8.559,\ p < 2e^{-16}$). Furthermore, according to a pairwise t-test, the congruence rates for the line drawings shown through the two larger windows were significantly different from all the others, while the congruence rates for the line drawings shown through the three smaller windows were not significantly different with each other (Table 4.3). Therefore, the certainties of depth difference perceptions for the line drawings shown through three smaller windows were qualitatively equivalent. Considering the results in Experiment 3, in which the congruence rates for the two larger windows were saturated to be qualitatively equivalent, the current results imply that the congruence rates increased not linearly but abruptly as the window sizes increased. There could be a jump in 3D shape interpretation and depth perception from the partial contours of a line drawing when critical visual cues, such as the endpoint of an internal contour, were shown.

Interestingly, there was a marginally significant interaction between the location of probes and size of windows ($F(8, 96) = 1.885,\ p = 0.0712$) (Figure 4.10). While the

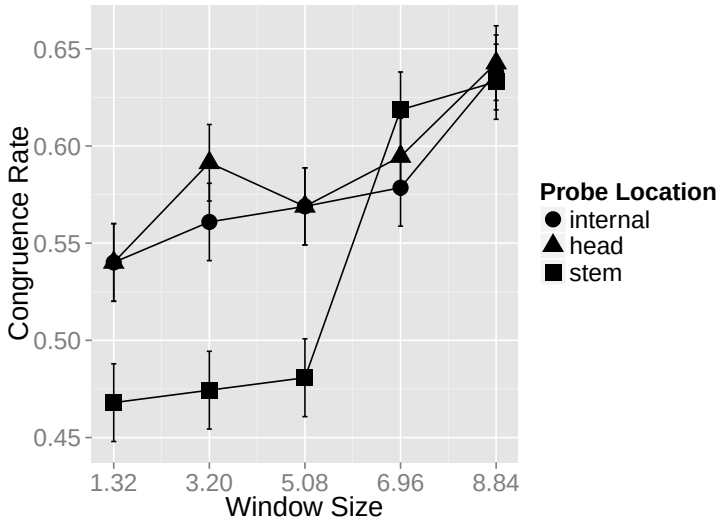


Figure 4.10: The interaction between the size of windows and location of probes.

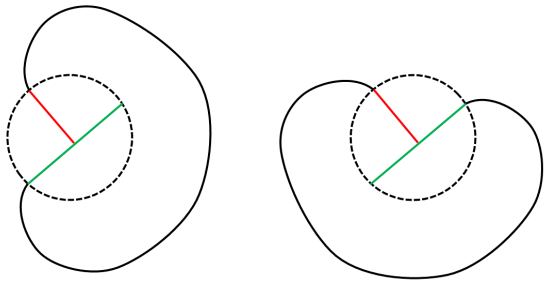


Figure 4.11: The ambiguity in interpreting surfaces from a stem contour when shown through an aperture. The surface relation along a stem contour for the same T-junction can be interpreted in two opposite ways.

congruence rates at internal and head probes gradually increased, the congruence rates at stem probes made a sudden jump as the window sizes increased. This result was consistent with the ambiguity of stem contours. Surface relations along a stem contour of a T-junction are undetermined until the information of the endpoint of internal contour is available, as illustrated in Figure 4.11. Such a result implies that important non-local cues, such as the endpoints of internal contours in T-junctions, are necessary to interpret local T-junction cues.

In addition, the influence of the location of probe dots on the display screen was examined. The height differences between two probe dots for each probe set were

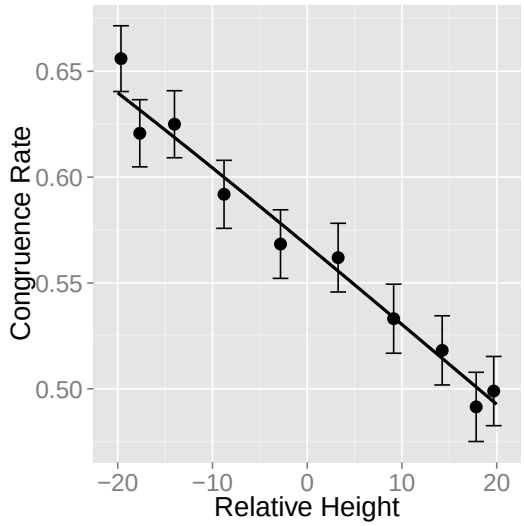


Figure 4.12: Congruence rates as a function of height difference between two probe dots. The height difference is the height of "farther" dot on the screen subtracted from the higher of "closer" dot for each probe set. Each data point shows the mean and standard error of congruence rates for each group. The line shows a fitted curve from the logistic regression.

computed, and the relation of congruence rates with the height differences was analyzed. To visualize the congruence rates as a function of the height difference, the height differences were grouped into ten groups with the same number of elements, and the means and standard errors of congruence rates for each group were plotted in Figure 4.12. According to a logistic regression, the congruence rates systematically decreased as the height differences increased and as the "closer" dot was located higher than the "farther" dot on the screen ($z = -10.348$, $p < 2e^{-16}$). The result shows that the certainty of depth difference perceptions were higher as the "closer" dot was located on the lower region, which is consistent with the lower-region figure bias that makes the lower region perceived as closer.

4.2.3 Conclusion

In this experiment, the scope of non-locality was varied with the sizes of windows that showed the partial contours of line drawings. The quantitative relation between the congruence rates and the amount of non-local visual cues in the windows showed that the depth difference was more certain as more global visual information was available. In

particular, the increase in the certainty of depth difference perceptions from smaller to larger amounts of non-local information was abrupt, implying that there was a moment that the critical non-local visual cues were given. In addition, the relative height of probe dots on the screen influenced the depth responses, which was consistent with the closer lower-region bias. These results also confirm that global contex interacts with local visual information to create the 3D shape interpretation.

Chapter 5

General Conclusion

In contrast to the large amount of research on 3D surface shape based on local cues to surface orientation, there has been relatively little exploration of the 3D shape perception from line drawings. In this dissertation, I investigated the mechanism of 3D shape interpretation from line drawings with four psychophysical experiments and a computational model. From the experiments, I found that the interpretation of surfaces and depths of a line drawing was probabilistic; the repeated depth judgments about a line drawing reflected the underlying uncertainty of the 3D surfaces based on the interaction of local and global contour structures. In addition, I found that the interpretation of a contour as a depth cue was dependent on the more global context, such as learning from experimental context. Furthermore, the 3D shape interpretations from partial line drawings were systematically influenced by the amount of visual information. The inference of 3D shape from 2D line drawing was modeled in a Bayesian way. The model estimates the posterior of possible 3D shapes given the contour features of a line drawing—the junctions and curvatures of contours—by combining a likelihood from a generative model and a prior from biases in depth perception. Additionally, the model provides a tool for understanding the complex interaction of local and global visual features by accommodating the global effect of contours on estimating 3D shapes. The findings from the experiments and the model in this dissertation will advance the knowledge of 3D shape perception.

Bibliography

Barrow, H. G., & Tenenbaum, J. M. (1981). Interpreting line drawings as three-dimensional surfaces. *Artificial Intelligence*, *17*, 75–116.

Biederman, I. (1987). Recognition-by-components: A theory of human image understanding. *Psychological Review*, *94*(2), 115–147.

Bülthoff, H. H., & Mallot, H. A. (1988). Integration of depth modules: stereo and shading. *Journal of Optical Society of America A*, *5*(10), 1749–1758.

Burge, J., Fowlkes, C. C., & Banks, M. S. (2010). Natural-scene statistics predict how the figure-ground cue of convexity affects human depth perception. *The Journal of Neuroscience*, *30*(21), 7269–7280.

Charkravarty, I. (1979). A genealized line and junction labeling scheme with applications to scene analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *PAMI-1*, 202–205.

Christou, C., Koenderink, J. J., & van Doorn, A. J. (1996). Surface gradients, contours and the perception of surface attitude in images of complex scenes. *Perception*, *25*, 701–713.

Clowes, M. B. (1971). On seeing things. *Artificial Intelligence*, *2*(1), 79–116.

Driver, J., & Baylis, C. C. (1995). One-sided edge assignment in vision: 2. part decomposition, shape description, and attention to objects. *Current Directions in Psychological Science*, (pp. 201–206).

Fantino, E., & Esfandiari, A. (2002). Probabilty matching: Encouraging optimal responding in humans. *Canadian Journal of Experimental Psychology*, *56*(1), 58–63.

Feldman, J. (2001). Bayesian contour integration. *Perception and Psychophysics*, *63*(7), 1171–1182.

Field, D. J., Hayes, A., & Hess, R. F. (1993). Contour integration by the human visual system: Evidence for a local "association field". *Vision Research*, *33*(2), 173–193.

Geisler, W., Perry, J., Super, B., & Gallogly, D. (2001). Edge co-occurrence in natural images predits contour grouping performance. *Vision Research*, *41*, 711–724.

Green, C. S., Benson, C., Kersten, D., & Schorater, P. (2010). Alternations in choice behavior by manipulations of world model. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(37), 1–6.

Hochberg, J., & Brooks, V. (1962). Pictorial recognition as an unlearned ability: A study of one child's performance. *American Journal of Psychology*, *76*, 624–628.

Hoffman, D. D., & Richards, W. A. (1984). Parts of recognition. *Cognition*, *18*, 65–96.

Huffman, D. A. (1971). Impossible objects as nonsense sentences. In B. Meltzer, & D. Miche (Eds.) *Machine Intelligence*, chap. 19, (pp. 295–323). Edinburgh: Edinburgh Univertisy Press.

Igarashi, T., Matsuoka, S., & Tanaka, H. (1999). Teddy: A sketching interface for 3d freeform design. In *SIGGRAPH*.

Jorge, J., & Samavati, F. (Eds.) (2011). *Sketch-based Interfaces and Modeling*. Springer London.

Kanizsa, G., & Gerbino, W. (1976). Convexity and symmetry in figure-ground organization. In M. Henle (Ed.) *Vision and Artifact*, (pp. 25–32). New York, NY 10003: Springer Publishing Company, Inc.

Kent (1983). Slant-tilt: The visual encoding of surface orientation. *Biological Cybernetics*, *46*, 183–195.

Kersten, D., Mamassian, P., & Yuille, A. L. (2004). Object perception as bayesian inference. *Annual Review of Psychology*, *55*, 271–304.

Kersten, D., & Yuille, A. (2003). Bayesian models of object perception. *Current Opinion in Neurobiology*, *13*, 1–9.

Knill, D. C. (1992). Perception of surface contours and surface shape: from computation to psychophysics. *Journal of Optical Society of America A*, *9*(9), 1449–1464.

Knill, D. C. (1998a). Discrimination of planar surface slant from texture: human and ideal observer compared. *Vision Research*, *38*, 1638–1711.

Knill, D. C. (1998b). Surface orientation from texture: ideal oobserver, generic observer and the information content of texture cues. *Vision Research*, *38*, 1655–1682.

Knill, D. C., & Saunders, J. A. (2003). Do humans optimally integrate stereo and texture information for judgments of surface slant? *Vision Research*, *43*, 2539–2558.

Koenderink, J. J. (1984). What does the occluding contour tells us about solid shape? *Perception*, *13*, 321–330.

Koenderink, J. J. (1998). Pictorial relief. *Phil. Trans. R. Soc. Lond. A*, *356*, 1071–1086.

Koenderink, J. J., van Doorn, A., & Wagemans, J. (2012). Picasso in the mind's eye of the beholder: Three-dimensional filling-in of ambiguous line drawings2. *Cognition*, *125*, 394–412.

Koenderink, J. J., & van Doorn, A. J. (1982). The shape of smooth objects and the way contours end. *Perception*, *11*, 129–137.

Koenderink, J. J., & van Doorn, A. J. (1995). Relief: pictorial and otherwise. *Image and Vision Computing*, *13*(5), 321–334.

Koenderink, J. J., van Doorn, A. J., & Kappers, A. M. L. (1992). Surface perception in pictures. *Perception & Psychophysics*, *52*(5), 487–496.

Koenderink, J. J., van Doorn, A. J., & Kappers, A. M. L. (2001). Ambiguity and the 'mental eye' in pictorial relief. *Perception*, *30*, 431–448.

Li, Y., Pizlo, Z., & Steinman, R. M. (2009). A computational model that recovers the 3d shape of an object from a single 2d retinal representation. *Vision Research*, *49*, 979–991.

Liu, B., & Todd, J. T. (2004). Perceptual biases in the interpretation of 3d shape from shading. *Vision Research*, *44*, 2135–2145.

Ma, W. J. (2012). Organizing probabilistc models of perception. *Trends in Cognitive Science*, *15*(10), 511–518.

Mackworth, A. K. (1973). Interpreting pictures of polyhedral scenes. *Artificial Intelligence*, *4*, 121–137.

Malik, J. (1987). Interpreting line drawings of curved object. *International Journal of Computer Vision*, *1*, 73–103.

Maloney, L. T., & Mamassian, P. (2009). Bayesian decision theory as a model of human visual perception: Testing bayesian transfer. *Visual Neuroscience*, *26*, 147–155.

Mamassian, P. (2006). Bayesian inference of form and shape. *Progress in Brain Research*, *154*, 265–270.

Mamassian, P., & Kersten, D. (1996). Illumination, shading and the perception of local orientation. *Vision Research*, *36*(15), 2351–2367.

Mamassian, P., Landy, M., & Maloney, L. T. (2002). Bayesian modelling of visual perception. In R. P. N. Rao, B. A. Olshousen, & M. S. Lewicki (Eds.) *Probabilisitc Models of the Brain: Perception and Neural Function*, (pp. 13–36). MIT Press.

Mamassian, P., & Landy, M. S. (1998). Observer biases in 3d interpretation of line drawings. *Vision Research*, *38*, 2817–2832.

Marr, D. (1977). Analysis of occluding contour. *Proceedings of the Royal Society B*, *197*, 441–475.

Marr, D. (1982). *Vision: A Computational Investigation into the Human Representationand Processing of Visual Information*. San Francisco, California: W. H. Freeman.

Nealen, A., Sorkine, O., Alexa, M., & Cohen-Or, D. (2005). A sketch-based interface for detailed-preserving mesh editing. In *ACM Transactions on Graphics - Proceedings of ACM SIGGRAPH 2005*, vol. 24, (pp. 1142–1147).

Norman, J. F., Todd, J. T., Norman, H. F., Clayton, A. M., & McBride, T. R. (2006). Visual discrimination of local surface structure: Slant, tilt, and curvedess. *Vision Research*, *46*, 1057–1069.

Parzen, E. (1962). On estimation of a probability density function and mode. *Annals of Mathematical Statistics*, *33*(3), 1065–1076.

Perrone, J. A. (1981). *Visual Slant Underestimation*. Ph.D. thesis, University of Canterbury.

Phillips, F., Todd, J. T., Koenderink, J. J., & Kappers, A. M. L. (1997). Perceptual localization of surface position. *Journal of Experimental Psychology: Human Perception and Performance*, *23*(5), 1481–1492.

Phillips, F., Todd, J. T., Koenderink, J. J., & Kappers, A. M. L. (2003). Perceptual representation of visual surfaces. *Perception & Psychophysics*, *65*(5), 747–762.

Pizlo, Z. (2010). *3D Shape: Its Unique Place in Visual Perception*. The MIT Press.

Pizlo, Z., & Salach-Golyska, M. (1995). 3-d shape perception. *Perception & Psychophysics*, *57*(5), 692–714.

Pizlo, Z., Sawada, T., Li, Y., Kropatsch, W. G., & Steinman, R. M. (2010). New approach to the perception of 3d shape based on veridicality, complexity, symmetry and volume. *Visioin Research*, *50*, 1–11.

Pomerantz, J. R., & Kubovy, M. (1986). Simplicity and likelihood principles. In K. R. Boff, L. Kaufman, & J. P. Thomas (Eds.) *Handbook of perception and human performance*, vol. 2, chap. 36 Theoretical Approaches to Perceptual Organization, (pp. 36–1–36–44). New York: John Wiley and Sons.

Pouget, A., Dayan, P., & Zemel, R. (2000). Information processing with population codes. *Neuroscience*, *1*, 125–132.

Richards, W. A., Koenderink, J. J., & Hoffman, D. D. (1987). Inferring three-dimensional shapes from two-dimensional silhouettes. *Journal of Optical Society of America*, *4*, 1168–1175.

Snodgrass, J. C., & Vanderwart, M. (1980). A standardized set of 260 pictures: Norms for name agreement, image agreement, familiarity, and visual complexity. *Journal of Experimental Psychology*, *6*(2), 174–215.

Stevens, K. A. (1981). The visual interpretation of surface contours. *Artificial Intelligence*, *17*, 47–73.

Stevens, K. A. (1983). Surface tilt (the direction of slant): A neglected psychophysical variable. *Perception & Psychophysics*, *33*(3), 241–250.

Todd, J. T. (2004). The visual perception of 3d shape. *Trends in Cognitive Science*, *8*(3), 115–121.

Todd, J. T., Koenderink, J. J., van Doorn, A. J., & Kappers, A. M. L. (1996). Effects of changing viewing conditions on the perceived structure of smoothly curved surfaces. *Journal of Experimental Psychology: Human Perception and Performance*, *22*(3), 695–706.

Todd, J. T., & Perotti, V. J. (1999). The visual perception of surface orientation from optical motion. *Perception & Psychophysics*, *61*(8), 1577–1589.

Todd, J. T., & Reichel, F. D. (1990). Visual perception of smoothly curved surfaces from double-projected contour patterns. *Journal of Experimental Psychology: Human Perception and Performance*, *16*(3), 665–674.

Tse, P. U. (2002). A contour propagation approach to surface filling-in and volume formation. *Psychological Review*, *109*(1), 91–115.

van Ee, R., Adams, W. J., & Mamassian, P. (2003). Bayesian modeling of cue integration: bistability in stereoscopic slant perception. *Journal of the Optical Society of America A*, *20*(7), 1398–1406.

Vecera, S. P., Vogel, E. K., & Woodman, G. F. (2002). Lower region: A new cue for figure-ground assignment. *Journal of Experimental Psychology: General*, *131*, 194–205.

Waltz, D. L. (1972). Generating semantic descriptions from drawings of scenes with shadows. Tech. Rept. AI-TR-271, Massachusetts Institute of Technology, Cambridge, MA.

Weiss, Y. (1997). Interpreting images by propagating bayeisan beliefs. In M. J. M.C. Mozer, & T. Petsche (Eds.) *Advances in Neural Information Processing Systems*, vol. 9, (pp. 908–915).

Yuille, A., & Kersten, D. (2006). Vision as bayesian inference: Analysis by synthesis? *Trends in Cognitive Sciences*, *10*(7), 301–308.

Yuille, A. L., & Bulthoff, H. H. (1994). Bayeisan decision theory and psychophysics. In D. C. Knill, & W. Richards (Eds.) *Perception as Bayesian Inference*, (pp. 123–161). New York: Cambridge University Press.

Zeleznik, R. C., Herndon, K. P., & Hughes, J. F. (1996). Sketch: An interface for sketching 3d scenes. In *Proceedings of the ACM SIGGRAPH Conference on Computer Graphics*, (pp. 163–170).

Zemel, R. S., Dayan, P., & Alexandre (1998). Probabilistic interpretation of population codes. *Neural Computation*, *10*, 403–430.

Zimmerman, G. L., Legge, G. E., & Cavanagh, P. (1995). Pictorial depth cues: a new slant. *Journal of the Optical Society of America A*, *12*(1), 17–26.