THE RELATIONSHIP BETWEEN FEATURES AND EDGE TYPES IN NATURAL IMAGES

A Thesis
Submitted to the Graduate Faculty
of the
North Dakota State University
of Agriculture and Applied Science

By

Jonathon Michael George

In Partial Fulfillment of the Requirements
for the Degree of
MASTER OF SCIENCE

Major Department:
Psychology

October 2011

Fargo, North Dakota

North Dakota State University
Graduate School

Title

The Relationship Between Features and Edge Types in Natural Images

By

Jonathon Michael George

The Supervisory Committee certifies that this ***disquisition*** complies with North Dakota State University's regulations and meets the accepted standards for the degree of

MASTER OF SCIENCE

ABSTRACT

George, Jonathon Michael, M.S., Department of Psychology, College of Science and Mathematics, North Dakota State University, October 2011. The Relationship Between Features and Edge Types in Natural Images. Major Professor: Dr. Mark Nawrot.

One of the most important processes in the human visual system involves detecting and understanding edges. Edges allow humans to break a visual scene up into meaningful chunks of information. Without edges, a visual scene is meaningless. As important as edges are to human visual perception, how they are detected and classified is not well understood. This study provides evidence that humans are able to classify edges into appropriate categories when enough visual information is presented but objects in the scene are not detectable. In addition, this study shows that regions of interest (ROIs) of a particular edge type can be clustered according to similarities in structure using a simple algorithm. This study examines the relationship between image features (i.e. closure, texture & repetition) and the type or cause of an edge (i.e. albedo, depth, shadow & specular) in natural visual scenes. Two groups of human subjects were used to carry out the current study; the cause estimators (CEs) and the feature experts (FEs). The CEs were asked to state the cause of an edge presented in a ROI. The FEs were asked to label specific features for the same set of ROIs as the CEs. The first analysis describes the relationship between image features and the actual cause of the edge in the ROIs presented. The second analysis describes the relationship between image features and the cause estimation provided by the CEs. This study provides evidence that closure, texture and repetition may help to inform human observers as to the cause of an edge when limited but sufficient visual information is available.

ACKNOWLEDGEMENTS

I would like to thank Dr. Mark Nawrot for his guidance, motivation and patience over the past two years. Your assistance is greatly appreciated. In addition, I would like to thank my committee; Dr. Benjamin Balas, Dr. Wendy Gordon and Dr. Lisa Montplaisir. Finally, I would like to thank Dr. Mark Brady for his expertise and guidance in the beginning stages of this process.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF EQUATIONS

INTRODUCTION

Vision is a fascinating and complex sense used by most humans to maneuver, manipulate and understand the surrounding environment. In fact, we have the remarkable ability to survive in a variety of situations without putting much cognitive effort into how exactly we see. Despite this fact, visual processing remains a sophisticated task. In addition to its intricacy, visual processing is continuous and near instantaneous. Given the speed that visual processing occurs raises many interesting questions. However, answering them requires an understanding of three primary components involved in visual processing.

The first component involved in visual processing is the environment. Our surroundings require the interaction of two things; light energy and objects. The interaction of light with objects creates an infinite number of patterns. Light is either absorbed or reflected based on the composition of each object and its surface properties. The reflected light travels through the atmosphere where it interacts with other objects, or an observer.

The second component in visual processing is the eye. Some of the reflected light from the environment travels into the eye through the pupil. The iris adjusts the pupil size based on the overall luminance of the photons entering the eye. Light travels through the lens where it is focused and projected onto the retina. Active visual processing begins by converting light energy to electrical energy through a process called transduction (Pinel, 2009). Specifically, photoreceptors that comprise a large part of the retina undergo physical changes resulting in electro-chemical signals that travel through several cell layers to the retinal ganglion cells. The ganglion cells converge to form the optic nerve that exits at the back of the eye and continues on to the brain. This process is necessary for a few reasons. First, electro-chemical signals are transmitted rapidly compared to mechanical signals. Second, the brain is heavily protected necessitating a specialized communication pathway, thus electro-chemical signaling is ideal based on its physical properties. Lastly, the physical properties of electric current allow for unique signal types leading to a greater number of signal permutations for visual discrimination.

The third component of visual processing is the brain. Once transduction occurs, electro-chemical impulses are sent to the brain. Here the signals sent from the retinal ganglion cells are decoded in order to interpret what humans understand as sight. How this occurs is complex and what much of the research in object recognition attempts to explain.

Background

One of the difficult problems in vision research is the variation in dimensionality at each level of visual processing (Poggio, Torre & Koch, 1985). The world around us is organized in three spatial dimensions and an infinite dimension of light and surface properties. However, at the level of the eye, the environmental dimensions are choked down to a two-dimensional retinal image. The retinal image is conveyed to the brain through electrical impulses where they are decoded as a three-dimensional estimation of the surrounding world.

One explanation of this phenomenon is stereopsis (Wheatstone, 1838; Poggio & Poggio, 1984). Stereopsis is the disparity created between each eye's retinal image. The brain is able to measure this difference and estimate depth relative to a point of fixation. In addition to stereopsis, motion parallax also provides an explanation for the perception of depth. Specifically, as the observer moves, the retinal image of objects in the visual scene move relative to a point of fixation (Nawrot, 2003a; Nawrot, 2003b; Nawrot & Joyce, 2006; Nawrot & Stroyan, 2009). The brain uses the motion of these objects in relation to the fixation point and other objects in the visual scene to estimate depth.

Both explanations provide substantial evidence as to how the brain interprets three-dimensional space from two-dimensional retinal images derived from the physical world. However, they do not provide an explanation of how dimensionality is inferred when the original stimulus is two-dimensional and the observer is stationary. Moreover, there are circumstances when the visual scene is not always clear. For example, say the observer is looking through a long dark tube or a peep-hole in a door. Regardless of the viewing circumstance, humans are often able to distinguish depth in visual scenes when both motion and stereopsis are absent.

Many theories have attempted to explain how humans sense and perceive the world. One of the first modern theories came from Gestalt psychology. Structuralism, a Gestalt principle, posits that perception is a phenomenon where basic pieces of information are put together to form a "whole" (Goldstein, 2002). For example, an image displayed on a computer screen is made up of pixels. Each pixel generates information pertaining to the color and luminance of a particular point in that image. Structuralism would say that the brain assembles each pixel according to its position on the screen and the whole visual scene is perceived based on the combination of all pixels of a given point relative to the other pixels that make up the image. While this approach is certainly logical, it does not address how humans are able to decipher three-dimensional space from a two-dimensional retinal image. Moreover, while many Gestalt principles hold for simple examples, they often fail when applied to real world scenarios. Regardless, Gestalt theory introduced the idea that perception is a method of assembling small pieces of information into meaningful patterns. While it is unclear whether visual perception is solely a bottom-up process, it is a helpful and productive approach to the study of visual perception.

Marr (1982) stresses the importance of modularity in object recognition, meaning objects can be broken down into smaller components that contain even more basic components. While this sounds much like structuralism, it differs in that component parts can be thought of as viewing the image at different scales. Different features are salient at different scales. At smaller scales, pixels may be important while at larger scales, features formed from pixels may be more informative. The features at each scale can be grouped and used to inform the next largest set of features.

Marr (1982) posits that the geometric shape, light reflectance properties, the position of the viewer relative to the image and the lighting itself are the four factors responsible for image intensity changes. These intensity changes can be grouped with similar intensity changes within the image based on image features that are analogous in structure and position. This grouping is referred to as the primal sketch. Essentially, the primal sketch gives a crude skeleton-like map of the visual scene. The primal sketch gives rise to more complex grouping that leads to construction of the observed surfaces in an image and is referred to as the two-and-a-half-dimensional sketch. These two representations can be used to inform the viewer of the relationship between surfaces in an image. Interactions between surfaces result in boundaries or edges. Finally, the combination of the primal sketch and two-and-a-half-dimensional sketch lead to a 3-dimensional interpretation of the surrounding world.

edge. *Shadow edges* are created by obstructing light from interacting with part of a surface. This is the case when a tree located between sunlight and the surrounding grass obstructs the illumination of some part of the grass. Finally, *specular edges* are created when light is reflected from a surface, creating a bright spot visible on the surface of the reflecting object. This is often observed when looking at glossy surfaces.

Humans perform well at distinguishing between different types of edges when enough information is present in the visual scene (Brady & Padmanabhan, 2007). Specifically, human observers are able to distinguish between albedo, depth, shadow and specular edges presented in two-dimensional edge profiles when the visual scene is unrecognizable and both motion parallax and stereopsis do not provide useful information. For example, when participants are given a very small viewing aperture (i.e. 9 minutes of arc) they perform at chance levels when distinguishing between the four edge types. However, when a larger viewing radius is presented (i.e. 134 minutes of arc), participants are able to distinguish between edge types quite accurately even when objects in the scene are indistinguishable. This implies that humans are using some kind of information to parse the visual scene. This study posits there exists a relationship between combinations of specific image features and specific edge types in natural visual scenes.

Features

Our surroundings require the interaction of two things; light energy and objects. Light is either absorbed or reflected based on the composition of each object and its surface properties. Reflected light then travels through the atmosphere where it interacts with other objects, or an observer. While these interactions can create a seemingly infinite number of patterns, the visual system is able to impose a small range of three-dimensional solutions based on regularities found in nature (Richards, Koenderink & Hoffman, 1987). It is these patterns along with the constraints imposed by the visual system that create several features that play a role in object recognition. While many are important for different aspects of visual processing, this study examines three features thought to be important in identification and classification of edges.

The first important feature for edge classification is called *closure*. Closure occurs when an edge creates a closed boundary within the viewing area. Closure by itself can define a surface boundary within a two-dimensional image. If an image contains a closed boundary, the area inside of a closed boundary can be brighter or darker than the area outside of the boundary. The variation in contrast between the inner and outer portion of a closed boundary may inform the viewer as to the type of boundary present in the visual scene.

The second feature important for edge classification is *texture*. Texture is defined as a surface property orthogonal to an edge that cannot be modeled by a quadratic function. For example, a luminance change that occurs gradually from dark to light would not be deemed a texture. However, an abrupt change from dark to light (e.g. the change between two blocks in a black and white checkerboard pattern) would be considered a texture if the change occurs more than twice within a specified boundary. Textures help define surfaces and may assist the viewer in detecting and interpreting edges in a visual scene.

The third feature important for edge classification is *repetition*. Repetition occurs when a portion of the edge is created by an object or surface property that repeats throughout the viewing aperture. Repeating elements can provide pattern information that is not present in texture alone. For example, imagine looking through a chain link fence that occludes the environment behind it. The occluded environment may or may not be textured but the chain link provides additional pattern information regardless of the presence of a texture. *Appendix A* provides further explanation of these three image features.

The current study seeks to determine whether specific combinations of texture, closure and repetition can be used as cues in differentiating four different edge types. (i.e. albedo, depth, shadow and specular edges) in natural visual scenes. If a relationship exists between combinations of the above-mentioned features and specific edge types, it may indicate that humans use this information to help interpret the surrounding environment.

PRELIMINARY STUDIES

Two preliminary studies were conducted prior to the current study. The first preliminary study was intended to see whether or not human subjects could accurately identify different types of edges when varying degrees of limited information was available. The second preliminary study was intended to see if similarities in image structure were indicative of similarly caused edges. Each study serves as a motivating factor for the current study.

Preliminary Study 1

Given the importance of edge detection in human visual processing, we needed to see how accurately humans were at estimating the cause of a particular edge. Moreover, we needed to see whether or not performance varied when the amount of visual information was limited and object recognition was not possible. If humans are able to correctly identify the cause of an edge (i.e. albedo, depth, shadow or specular) when object recognition is not possible, then something in the image (e.g. features) must help inform their decision. The first preliminary study was designed to see whether or not human subjects can accurately determine the type of edge presented in an image when visual information necessary for object recognition is insufficient.

Apparatus

A standard computer with Windows XP operating system was used for all stimulus presentation purposes. A basic computer keyboard and mouse served as the input devices for participant responses. All stimuli were presented using MATLab 2007a. All response data was recorded in MATLab and written as a Microsoft Excel 2003 edition spreadsheet with an .xls file extension.

All stimuli were presented as stereographs on a 24-inch LCD Hewlett Packard computer monitor. A stereoscope was mounted to the monitor at a viewing distance of 41.7 cm. The stereoscope serves two important functions. First, stereoscopic images allow for better separation of the visual stimuli from the background that they are presented on. Second, this helps to ensure that a constant viewing distance is maintained throughout the experiment.

Stimuli

Images were taken from a database created by our lab for the purpose of both the preliminary and current studies. A Canon EOS 1Ds Mark III digital SLR camera in RAW format was used to photograph all images in the database. Camera settings (aperture = f/5.6, focal length = 17mm, focal distance = 41.7cm) matched human depth of field under full sunlight conditions (pupil diameter = 3mm) at a viewing distance of 41.7 cm. In order to control for the types of edges present in each image, backdrops were manipulated and shadows were cast on the object of interest. Illumination was also manipulated to move or eliminate shadows and specularities. Shadow images were compared to the same image without a cast shadow (i.e. the object interfering with the light was removed) to make sure that shadows were clearly visible in the image. Images were converted to true luminance values and displayed on a calibrated computer monitor at luminance levels proportional to those from the original outdoor scene. In total, 480 regions of interest (ROIs) were created from 11 different objects displayed in various conditions to adequately provide samples of all 4 edge types. These 480 ROIs were divided into 5 equal groups of 96 patches. Each of the 5 groups is defined by the size of the ROI context radii (9, 17, 33, 65 and 134 minutes of arc). To clarify, imagine you have a photograph and a blank sheet of paper of the same size. By cutting a hole in the paper and placing the paper over the picture, only a portion of the photo can be seen through the hole in the paper. Essentially, this experiment achieves the same result by only presenting a small, circular section or ROI of each image. Finally, each group contained an equal number of edge types (24 albedo, 24 depth, 24 shadow and 24 specular). Figure 1 shows a single image as it would appear at each aperture size.

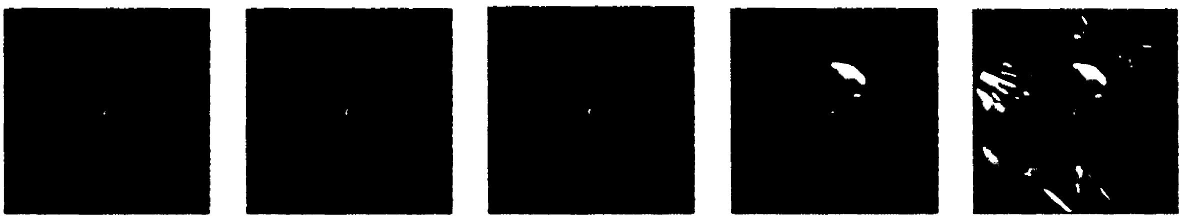

Figure 1. Example of aperture size. An example of each of the 5 aperture sizes used in the experiment. Aperture 1 is the image on the far left. Aperture 5 is the image on the far right.

Twenty-five participants known as cause estimators (CE) viewed 96 image ROIs at one of five apertures. Participants were given definitions of each edge type (i.e. Albedo, depth, shadow or specular) and shown a photograph with examples of each of these types of edges. This helped to ensure that participants understood the four different edge types. In addition, this provided both visual and verbal examples of each edge type to help reduce the possibility of misinterpretation of the definitions. Each participant in the CE group rated 96 patches resulting in a total of 5 observations per ROI for each of the 5 aperture sizes.

Procedure

Each ROI was displayed on a background of random black and white dots with a mean luminance equal to that of the mean luminance of the ROI being displayed. Image ROIs were presented as stereographs and viewed through a stereoscope at a fixed viewing distance of 41.7 cm. Each ROI contained an edge centered in the aperture marked by a 1 pixel red dot that could be toggled on and off by pressing the space bar on a standard computer keyboard. Stimuli remained on the screen until the participant gave a response. *Appendix B* gives the experimental protocol used in the first preliminary study.

For each trial, the four edge types of interest (i.e. albedo, depth, shadow, specularity) were presented at the bottom of the monitor below the ROI. Participants selected the edge type that they thought was present in the ROI displayed on the screen by clicking the appropriately labeled icon. Upon selecting the edge type, a message appeared asking them to press the space bar to continue on to the next ROI. Participant responses were recorded in a spreadsheet for future analysis.

Preliminary Study 1 Results

Results were measured as proportion of total correct responses given for each edge type at each aperture size. Participants performed at chance levels at the two smallest apertures for classifying edges formed from depth discontinuities, changes in albedo and cast shadows. However, participants were able to distinguish contrast from specular reflections even at the smallest apertures, pSpecular =.47 or 47% correct at aperture 1 and pSpecular =.63 or 63% correct at aperture 2. The third aperture yielded slightly better classification for all edge types with pAlbedo and pShadow =.39 or 39% correct, pDepth = .53 or 53% correct and pSpecular = .68 or 68% correct. At aperture 4, albedo edges were correctly identified 59% of the time, depth edges 67%, shadow edges 52% and specular edges 90%. Finally, at the largest aperture, albedo edges were correctly identified 83% of the time, depth edges 87%, and shadow and specular edges were both correctly identified 93% of the time.

While percent correct does give us an idea of participant performance, it does not tell us about any response bias. For example, a participant could score 100% correct for albedo ROIs, but may have responded to every ROI as being albedo. In order to test for biases, individual cause estimation data was combined at each aperture and average *sensitivity*, *specificity* and *d'* scores were computed at each of the 5 apertures for each of the 4 cause types. *Sensitivity* is the proportion of correctly classified ROIs of a certain cause to total number of ROIs that are of said cause. *Sensitivity* gives the ability of the cause estimators to correctly identify ROIs of each cause. *Specificity* is the proportion of non-matching ROIs that were correctly identified as not being of a particular cause to total number of non-matching ROIs. *Specificity* gives the ability of the cause estimators to correctly identify ROIs as *not* being of a particular cause. D-prime (*d'*) scores are normalized values that give the difference between the true positive and false positive rates. Greater *d'* scores are indicative of the participants ability to accurately discriminate between one type of edge and all of the others. Table 1 gives *sensitivity, specificity* and *d'* scores for the average cause estimator at each aperture for each edge type. Figure 2 gives a graphical presentation of the *d'* scores for each edge type at each aperture.

Preliminary Study 2

The second preliminary study examined the images used in preliminary study 1 to determine whether image ROIs with similar structure and appearance contained the same type of edge. The relationship between images was determined using two different clustering algorithms. A distance (given in pixel luminance value) is calculated for each comparison. Mean sum of squares error (*MSSE*) values were calculated for each of the ROIs in all of the apertures. Each pixel in a given ROI is compared to the corresponding pixel value of an ROI of interest. Distances were computed in pixel luminance values with smaller values indicating more similarity between two ROIs and larger values indicating less similarity between two ROIs. Put simply, differences between pixel values are used as a measure of similarity at each point in the image and combined to give a single score that represents how closely one ROI is to another. Two identical ROIs would yield a value of zero. *MSSE* was calculated using equation 1 where n is the total number of pixels in the ROI of interest, r is a single pixel in the comparison ROI at location (i, j) and s is a single pixel in the ROI of interest located at the same coordinate (i, j) as the comparison ROI.

$$MSSE = \frac{\sum_n (r_{ij} - s_{ij})^2}{n} \tag{1}$$

In order to optimize the distance, each patch was rotated 1 degree at a time and compared to the ROI of interest until it had rotated a total of 360 degrees. A total of 360 comparisons were made for each patch relative to the patch of interest. Each ROI was used as the comparison patch and distances were calculated relative to each ROI of interest. The smallest distance from a given patch to the seed was used to cluster all patches.

Larger aperture ROIs were included in clusters of smaller aperture ROIs. The *MSSE* for larger patches were calculated based on the area of the aperture size used for comparison. Thus any part of the patch outside the area of interest was ignored in the

Table 1. *Sensitivity, Specificity and d' for Average Cause Estimator at Each Aperture*

[illegible]	[illegible]	Sensitivity	Specificity	d'
Aperture 1	**Albedo**	0.26	0.76	0.05
	Depth	0.34	0.77	0.34
	Shadow	0.38	0.76	0.41
	Specular	0.45	0.85	0.92
		Sensitivity	Specificity	d'
Aperture 2	**Albedo**	0.27	0.78	0.14
	Depth	0.31	0.78	0.27
	Shadow	0.40	0.75	0.43
	Specular	0.66	0.90	*1.69
		Sensitivity	Specificity	d'
Aperture 3	**Albedo**	0.37	0.73	0.24
	Depth	0.38	0.80	0.53
	Shadow	0.39	0.85	0.79
	Specular	0.64	0.88	**1.55
		Sensitivity	Specificity	d'
Aperture 4	**Albedo**	0.68	0.86	**1.57
	Depth	0.54	0.93	*1.71
	Shadow	0.79	0.84	*1.86
	Specular	0.86	0.97	*2.52
		Sensitivity	Specificity	d'
Aperture 5	**Albedo**	0.93	0.82	*2.43
	Depth	0.73	0.98	*2.62
	Shadow	0.80	0.98	*2.88
	Specular	0.80	0.98	*2.84

Note. Table 1 shows *sensitivity*, *specificity* and *d'* values for the average cause estimator at each aperture for each cause. The single asterisk (*) represents significant *d'* values, $p < .05$. The double asterisk (**) indicates marginal significance.

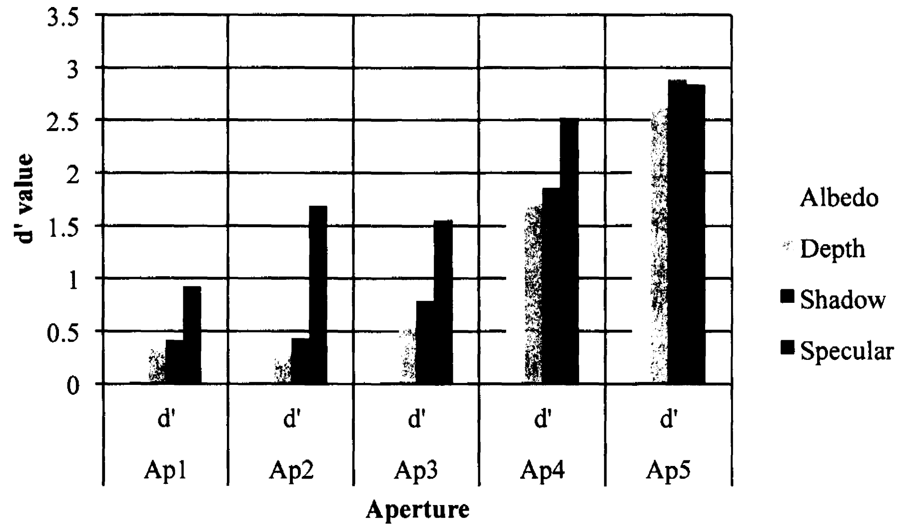

Figure 2. D' values as a function of aperture and edge type. Graph of d' for each edge type as a function of aperture size

computation of distance. For each of the apertures, 9, 17, 33, 65 and 134 minutes of arc, distance matrices were sizes [480,479], [384,383], [288,287], [192,191] and [96,95] respectively. The matrices were organized so that each ROI was used as an exemplar once and each row represented the cluster of patches as they related to a single exemplar. All of the other patches were organized according to how closely they resembled the exemplar based on the computed distance value. For example, in the smallest aperture, row 1 of the matrix represented patch 1. 479 columns (1 for each patch being compared to the exemplar) were organized according to how well they matched the comparison patch based on the *MSSE* algorithm. This was done for all patches resulting in a matrix with 480 rows (each exemplar) and 479 columns (each comparison patch).

Preliminary Study 2 Results

Clusters were evaluated on how accurately the ROIs were sorted based on the actual cause of the edge in the region of interest. *A* values were computed by taking the number of ROIs that matched the cause of the ROI of interest at a given point (*i, j*) within the cluster ($m_{i,1\ldots i,j}$), added to the number of ROIs that did not match the cause of the ROI of interest that were correctly excluded from the cluster ($d_{i,j+1\ldots i,n}$), and dividing the sum by the total number of ROIs at that given aperture (n). *A* values were computed at each point in the cluster. *Amax* values represent the maximum proportion of correctly included ROIs to correctly excluded ROIs for each cluster. The equation for *Amax* is shown in equation 2.

$$Amax = \max_{aperture}\left(\frac{\sum m_{i,1\ldots i,j} + \sum d_{i,j+1\ldots i,n}}{n}\right) \qquad (2)$$

Cluster sizes were determined by *Amax* values. Clusters with *Amax* values greater than or equal to .75 were considered significant. Significant clusters ranged in size from 12 ROIs to 108 ROIs at aperture 1. Significant clusters at aperture 2 ranged in size from 12 ROIs to 86 ROIs. Significant clusters at aperture 3 ranged in size from 12 ROIs to 43 ROIs. Significant clusters at aperture 4 ranged in size from 12 ROIs to 21 ROIs. Finally, significant clusters at aperture 5 ranged in size from 12 ROIs to 23 ROIs.

The most diagnostic clusters were found to have *specular edges* in the region of interest at all apertures except at the largest aperture where *shadow edges* were slightly more accurate. *Amax* values for significant clusters ranged from .75 to .86 with the greatest *Amax* values at the smallest aperture. While this may seem counterintuitive, the smallest aperture has the largest number of patches. While the ratio of each patch type is the same at each aperture, the potential for achieving greater *Amax* values is greater for the smallest aperture size. A second measure (distance in pixel luminance value), of how well ROIs were clustered was used to determine how closely the patches in each cluster resembled the patch of interest. Each distance is measured by the absolute value of the total difference in pixel luminance value between the ROI of interest and a given patch. When the distance is factored in, a much greater distance between ROIs is seen at the smallest aperture relative to the largest aperture. For example, the smallest distance from the exemplar patch to the last patch in the cluster for aperture 1 is 2267 while the largest distance is 13719. At the largest aperture, the distance ranges from 100 to 949. This indicates that while a greater number of correctly identified ROIs (compared to misidentified ROIs) was observed at smaller apertures, the distance these ROIs was from the comparison ROI was much greater. Thus, at larger apertures the ROIs are much closer to the exemplar than at smaller apertures. Positive predictive power (the ratio of accurately classified patches to total number of patches classified as being of same cause as the exemplar) ranged from .72 to .84 (M=.79). Negative predictive power (the ratio of accurately excluded patches to the total number of excluded patches) ranged from .85 to .87 (M=.86).

Results indicate that the MSSE clustering algorithm performs similarly to human subjects, especially at small apertures. Both human subjects and our computational analysis show that specular edges are the most accurately classified. The results of each of the preliminary studies seem to indicate that there is some characteristic inherent in natural images (especially when specular edges are present) that assist in our ability to segment objects based on the type of edge present in a visual scene.

Discussion of Preliminary Studies

Preliminary study 1 indicates that individuals are able to determine the cause of a specified edge when sufficient visual information is present in a portion of a natural image.

This holds true even when the information present in the image is not sufficient to identify specific objects contained in the ROI. Moreover, the *MSSE* clustering algorithm used in the second preliminary study performs similarly to human participants. The results provide evidence that some aspect of these image parts is important for establishing the cause of an edge (i.e. albedo, depth, shadow or specular reflection) when limited information is present.

Current Study

Given the results of each of the preliminary studies, it is probable that the specific features that make up a particular edge are indicative of its cause. This study seeks to determine whether or not a relationship exists between image features (i.e. texture, closure and repetition) and edge type (albedo, depth, shadow and specular) in natural visual scenes. More specifically, we examine the relationship between combinations of the three image features and the four edge types. Four predictions were made regarding the relationship between feature combinations and edge type. Prediction 1 is that ROIs rated as having no texture or a single texture will be rated as being *albedo* edges regardless of closure and repetition ratings. Prediction 2 is that the presence of one or two distinct textures will be indicative of a *depth* edge regardless of closure rating and in the absence of repetition. Prediction 3 is that ROIs rated as having no texture, a closure rating of either no closure or closure with a darker interior luminance (i.e. closed and dark) and no repetition will be *shadow* edges. Finally, Prediction 4 is a special case of prediction 1. Prediction 4 is that ROIs rated as having no texture or a single texture and the presence of closure with brighter interior luminance (i.e. closed and bright) will be highly indicative of *specular* edges, regardless of repetition ratings.

METHODS

Participants

Two different groups of participants were used in the current experiment. This was necessary to prevent response bias due to learned associations between edge types and image features. For example, if a participant knows that they are viewing a specular edge, their response regarding texture may be based on what they know about specular reflections rather than what they observe in the image. To simplify the data collection process, the data from the 25 cause estimators (CE) from preliminary study 1 was used in part of the final analysis for the current study. Informed consent was obtained for all participants in the current study. Appendix 3 is the consent form given to each participant.

A second group of 7 participants known as the feature experts (FE) was recruited for the second set of data used in the current study. Each FE viewed the same stimuli as those in the CE group, but were not asked to identify the cause of the edge, rather they were asked to identify whether or not specific features were present in each ROI. The current study used a graphical user interface (GUI) to present and record participant responses. Each participant in the FE group was trained as to how each of the image features is defined and how to use the GUI. Participants were given a list of definitions for all of the image features and shown examples of each. During training, the FE group was given a set of practice ROIs and given feedback regarding the correctness of their feature ratings. Finally a definition sheet was available for each participant to refer to as necessary during stimulus presentation. *Appendix A* gives the list of definitions that was given to each participant to use during training and experimentation.

Stimuli

The stimuli used in each of the preliminary studies were presented in the same fashion as preliminary study 1. This is important for a few key reasons. First, the data collected in the first preliminary study were used in part of the final analysis. Second the stimuli contain both the types of edges and features thought to be important in edge classification. Third the image database constructed for the first two studies contains natural images taken in a well-controlled environment and provides sufficient visual stimuli to test the hypothesis. Lastly, the processes used to construct the necessary stimuli and experimentation are quite time consuming and efficiency is greatly enhanced by using the images already collected.

Procedure

Stimuli for the feature expert (FE) group were presented in the same manner as in preliminary study 1. Each FE viewed 96 ROIs at each of the 5 apertures for a total of 480 ROIs. Because the task for the FEs takes a significant amount of time to complete, participants were given a two week window to complete the experiment. This was done in order to give participants ample time to complete the task and to avoid participant fatigue and apathy. Participants were able to save their responses at any point in the experiment and continue at another time.

A graphical user interface (GUI) with 7 different response categories was added below the stimulus for participants to record their responses. In order to avoid potential distraction an icon with the words "HIDE CONTROLS" was displayed for participants to click in order to eliminate the GUI under the stimuli until the participant was ready to respond. A button with the word "NEXT" was presented within the GUI so that the individual could proceed onto the next stimulus at any time during the experiment. A second button with the word "PREVIOUS" was also displayed so participants could go back to an ROI they previously skipped or desired to change a particular response. Finally, a button with the word "DONE" was displayed so the participant could exit the experiment at any time. Upon clicking the "DONE" button, any data entered was saved so that the participant could continue where they left off upon returning to finish the task.

The first response category in the current experiment was *closure*. Participants were instructed to click a box next to the words "YES" or "NO" depending on whether or not the specified edge forms a closed boundary within the viewing aperture. A "YES" response prompted a second response category called "Brighter inside?" and the response alternatives "YES" and "NO." Each participant was to respond to the category based on the overall luminance observed in the area inside the closed boundary as it compares to the overall luminance observed in the area immediately surrounding it. If the response to *closure* was "YES," a third response category asked that participants trace the contour of the closed boundary. Participants were instructed to use the mouse to trace the contour of the closed boundary they had observed. This was done to make sure that the participant correctly identified the edge in question.

The fourth feature response category was *texture*. It is important to reiterate the meaning of *texture* in the proposed experiment. Participants were instructed on and given a sheet of definitions regarding each of the response categories. Moreover, *texture* was loosely defined in order to avoid biasing participant responses. One could argue that an area with a constant luminance value has a smooth texture, but for the purpose of this experiment, subjects were instructed to consider this a non-texture. Furthermore, a single *texture* that is found on both sides of the specified contour was also considered a non-texture or "0" response. The idea was to find out whether *texture* is a factor in the formation of an edge and if so, whether "1" or "2" distinct *textures* were involved. Participants were asked to specify whether or not there were "0," "1," or "2" different *textures* abutting the specified boundary. This was done by checking a box next to the number of distinct *textures* involved in the formation of the edge presented (i.e. 0, 1 or 2).

The fifth response category is termed *repetition* and was used to determine whether or not some element involved in the formation of a particular boundary was observed in other areas of the ROI outside of the area where the edge was formed. Participants checked a box next to "YES" or "NO" as to whether or not an element involved in the creation of the edge in question repeated throughout the viewing aperture. Moreover, the repeating elements need not be identical, rather similar in form, texture and luminance. For example, if the ROI being viewed was a portion of a zebras stripes and the edge specified was formed at the point of interaction of a black and white stripe, the appropriate response would be "YES" as long as there was at least one portion of another black or white stripe present in the viewing aperture that appeared to be part of a group. This category allows for the recognition of larger patterns within a given image.

The sixth and seventh response categories were used to determine whether or not participants recognized an object in a given ROI. In the sixth category, participants were asked to check a box next to "YES" or "NO" regarding whether or not they recognized a particular object in the ROI being presented. If the individual checks the box next to "YES" the seventh category is highlighted and participants were asked to type in the name or a description of the object that they observed in the viewing window. The purpose of this response category is to eliminate patches that are subject to top down effects due to known factors about the object they have observed. For example, assume the ROI is taken from a picture of a water bottle sitting on a desk and the edge in question is that which is formed by the interaction of the bottom of the bottle and the desk. If a participant checks the "YES" response box, and specifies that they see a cup sitting on a table, their response will be influenced by what they know about cups sitting on tables rather than visual observation alone. Learning effects exist whether or not the participant correctly guesses the objects in the scene. Responses of "YES" to question six were excluded from the final analysis for all participants to eliminate potential learning effects.

RESULTS

The goal of the current study was to determine whether or not a relationship exists between specific combinations of image features (i.e. closure, texture and repetition) and the cause of an edge in a visual scene. For example, two distinct textures, on adjacent sides of an edge may be indicative of a depth discontinuity edge but not the remaining three categories. Similarly, a closed boundary with one texture may indicate a specular edge but not a depth edge. An albedo edge may have one distinct texture but no closure. Finally, shadow edges may have a continuous texture on both sides of an edge thus texture differences may not play a role in discerning a shadow edge from any other edge.

In order to determine whether such relationships exist between image features and edge type, we categorized image feature combinations using the numbers 1 through 18 to represent all of the feature combinations. Similarly edge types for each region of interest (ROI) were labeled using the digits 1 through 4. Frequency graphs were used to show each of the 4 edge types as a function of each of the 18 feature combinations. Comparisons were made between different edge types based on combinations of image features.

Data

All of the feature data were organized as a set of matrices. A total of 35 matrices were produced; one for each of the seven feature experts (FEs) at each of the five apertures. Each of the matrices contains at most 96 rows and 3 columns. Each row represents a single ROI. Column 1 of each matrix contains the FEs closure ratings for each of the ROIs for a given aperture. Closure was coded using the numbers 0, 1 and 2. A closure rating of 0 indicates that the edge being rated has an open boundary within the viewing aperture. A closure rating of 1 indicates that the edge creates a closed boundary within the viewing aperture and that the area inside the boundary is brighter than the area outside of the boundary. Finally, a closure rating of 2 indicates that the edge creates a closed boundary within the viewing aperture and that the area inside the boundary is darker than the area outside of the boundary.

Column 2 of each matrix contains the FEs texture ratings for each of the ROIs for a given aperture. Texture was coded using the numbers 0, 1 and 2. A texture rating of 0 indicates that no texture was observed or that a single continuous texture was located on both sides of the edge. A texture rating of 1 indicates that there was a single texture on one side of the edge but no texture on the opposing side of the edge. Finally, a texture rating of 2 indicates that there were two distinct textures present; one on each side of the edge. Colum 3 of each matrix contains the FEs repetition ratings for each of the ROIs for a given aperture. Repetition was coded using the numbers 0 and 1. A repetition rating of 0 indicates that no part of the edge was created by an element that repeats throughout the aperture. A repetition rating of 1 indicates that the edge was formed from some part of an element that repeats more than once within the viewing aperture. A rating for each feature was done for each of the ROIs. The result was a matrix where each row represents a particular ROI with 3 separate ratings; one for each of the three image features of interest (i.e. closure, texture and repetition). Table 2 provides an example and explanation of the FE data set.

In addition to the FE data, the actual cause (ground truth) of the edge in each ROI was used in part of the analysis. The data for ground truth is organized in a single column of digits (column vector). There are 5 column vectors in all; one for each of the 5 aperture sizes. Each column contains at most 96 elements that correspond to a particular ROI. An

ROI with a ground truth of Albedo =1, Depth =2, Shadow = 3 and Specular = 4. Table 3 provides an example and explanation of the ground truth data set for a single aperture.

Table 2. *Sample Feature Expert Data Matrix*

	$S1_{feature}$			$S2_{feature}$			...			$S6_{feature}$			$S7_{feature}$		
ROI	Cl	Txt	Rpt	Cl	Txt	Rpt	Cl	Txt	Rpt	Cl	Txt	Rpt	Cl	Txt	Rpt
1	{0,1,2}	{0,1,2}	{0,1}	...	...	...	...	...	...	...	...	...	{0,1,2}	{0,1,2}	{0,1}
2	...	...	...												
.	...	...	...												
.	...	...	...												
.	...	...	...												
95	...	...	...												
96	{0,1,2}	{0,1,2}	{0,1}												

Note. Sample data table for the feature expert data for each region of interest at a single aperture from each of the 7 feature experts. Each row represents a single ROI. Each set of 3 columns represents one of the 7 feature experts ($Sn_{feature}$). Each of the 3 columns under a particular feature expert represents the closure (Cl), texture (Txt) and repetition (Rpt) ratings for that subject.

Table 3. *Sample Ground Truth Data Matrix*

ROI	$\textbf{Cause}_{ground\ truth}$
1	{1,2,3,4,}
2	...
.	...
.	...
.	...
95	...
96	{1,2,3,4,}

Note. Sample data table for the actual cause (ground truth) data for each region of interest at a single aperture. Each row represents a single ROI. Each entry in the single column represents a certain cause; Albedo = 1, Depth = 2, Shadow = 3, Specular = 4.

Lastly, the data set from preliminary study 1 was used in a second analysis. Recall that this data set contains human estimation of the cause of each edge in the ROIs. Five cause estimators (CEs) viewed all ROIs at each aperture. A total of 5 matrices with 96 rows and 5 columns were obtained during preliminary study 1. Each row represents a single ROI at a particular aperture size. Each column corresponds to a single cause estimator. Finally, each cell represents the cause estimation for a single ROI from a single cause estimator. The CE data uses the same coding scheme as the ground truth data (Albedo =1, Depth =2, Shadow = 3 and Specular = 4). Table 4 provides an example and explanation of the cause estimation data set.

Table 4. *Sample Cause Estimation Data Matrix*

ROI	$S1_{cause}$	$S2_{cause}$	$S3_{cause}$	$S4_{cause}$	$S5_{cause}$
1	{1,2,3,4,}	...	...	...	{1,2,3,4,}
2	...				
	...				
	...				
	...				
95	...				
[illegible]	{1,2,3,4,}				

Note. Sample data table for the cause estimation data for each region of interest for each of the 5 cause estimators at a single aperture. Each row represents a single ROI. Each column represents data for a single participant (Sn_{cause}). Each entry represents a certain cause; Albedo = 1, Depth = 2, Shadow = 3, Specular = 4.

As previously mentioned, the feature expert group was asked whether or not they recognize an object in each of the ROIs and, if so, to describe or name the object that was observed. The purpose of these questions was to eliminate the potential influence that knowledge about a particular object may have on their response. For example, if an observer views a ROI and writes that they recognize an animal, it is probable that their responses regarding closure, texture and repetition was influenced by what they know about animals or about a specific animal. In the cases that an observer reported seeing an object, the data for that ROI was taken out of the final analysis for all observers.

For the smallest 2 aperture sizes the data for all 96 patches at each size was retained. For the third aperture data from one ROI was omitted leaving 95 ROIs. At aperture 4 the data from 7 ROIs was omitted leaving the data for 89 ROIs. Finally, at aperture 5, the data from 40 ROIs was omitted leaving data for the 56 remaining ROIs.

The remaining feature expert data was re-coded into single dimensional, categorical variables to simplify interpretation. Each feature rating was given a number that represented a particular combination of closure, texture and repeating elements from each feature expert for each ROI. Given that there are 3 possible responses for closure, 3 for texture and 2 for repetition, a total of 18 different feature combinations were obtained. For example; an ROI that was rated as having a 0 for closure (no closure), a 0 for texture (no texture or a single texture on both sides of the specified edge) and a 0 for repetition (no repeating element) was classified as a 1. Table 5 gives the individual feature ratings and a corresponding feature combination code.

Feature Expert Analysis

Analysis was conducted on the feature expert data to determine consistency between individual feature ratings. Percentage of agreement was determined for each of the 3 features at each aperture. The number of ROIs where 5 or more of the 7 feature experts gave the same rating, for a particular feature were added together and the sum was divided by the total number of ROIs rated at a given aperture. Figures 3, 4 and 5 give the percentage of agreement as a function of aperture for closure, texture and repetition respectively.

Table 5. *Feature Combination Code Key*

Feature Combination	1	2	3	4	5	6	7	8	9
CLOSURE	0	0	0	1	1	1	2	2	2
TEXTURE	0	1	2	0	1	2	0	1	2
REPETITION	0	0	0	0	0	0	0	0	0

Feature Combination	10	11	12	13	14	15	16	17	18
CLOSURE	0	0	0	1	1	1	2	2	2
TEXTURE	0	1	2	0	1	2	0	1	2
REPETITION	1	1	1	1	1	1	1	1	1

CLOSURE	TEXTURE	REPETITION
0 = no closure	0 = no texture	0 = no repeating element
1 = closed and bright	1 = 1 texture	1 = repeating element
2 = closed and dark	2 = 2 textures	

Note: The row labeled feature combination gives a nominal designation to the combination of closure, texture and repetition ratings listed in the column associated with each of the 18 feature combinations. The associated key gives the numerical values and associated meanings for each of the 3 features; closure, texture and repetition.

Analysis1

Descriptive statistical methods were used to examine the relationship between combinations of closure, texture and repeating elements and the four different edge types using the actual cause or ground truth data. Information obtained from these graphs tells whether or not human observers should use image features to interpret the actual cause of edges in natural visual scenes. This is true because we are examining the relationship between image features and the actual cause of the edge in each ROI. The frequency of occurrence of each edge type is plotted as a function of each feature combination (described in Table 5).

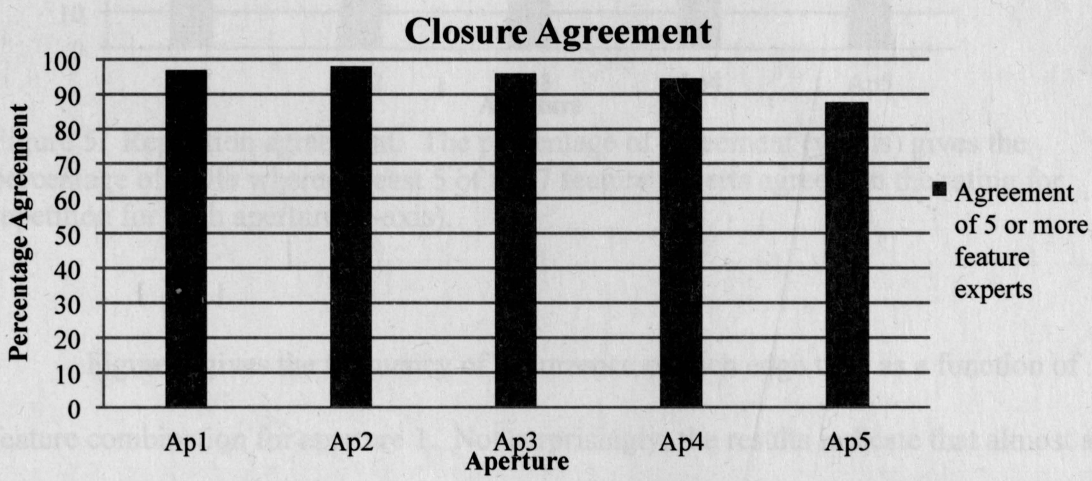

Figure 3. Closure agreement. The percentage of agreement (y-axis) gives the percentage of ROIs where at least 5 of the 7 feature experts agreed on the rating for closure for each aperture (x-axis).

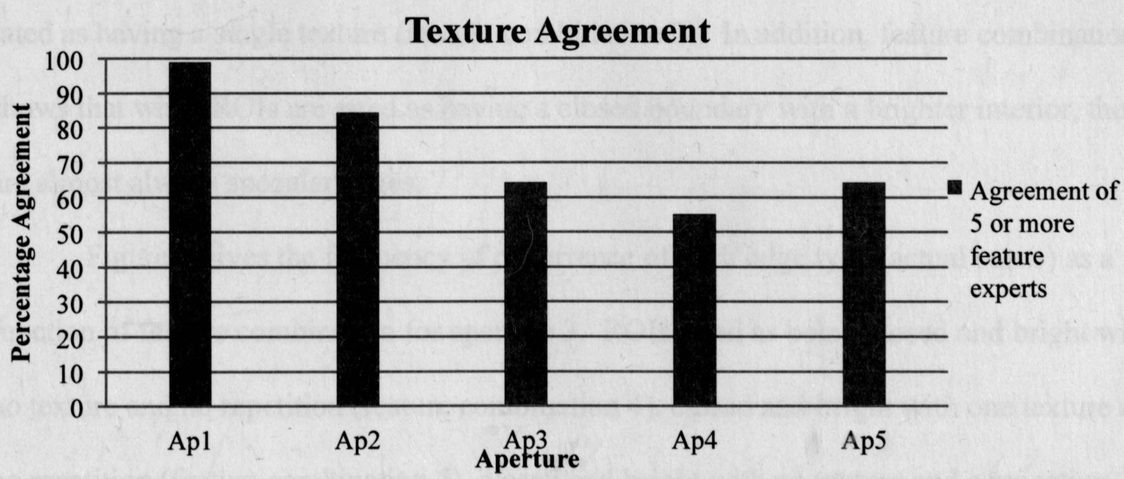

Figure 4. Texture agreement. The percentage of agreement (y-axis) gives the percentage of ROIs where at least 5 of the 7 feature experts agreed on the rating for texture for each aperture (x-axis).

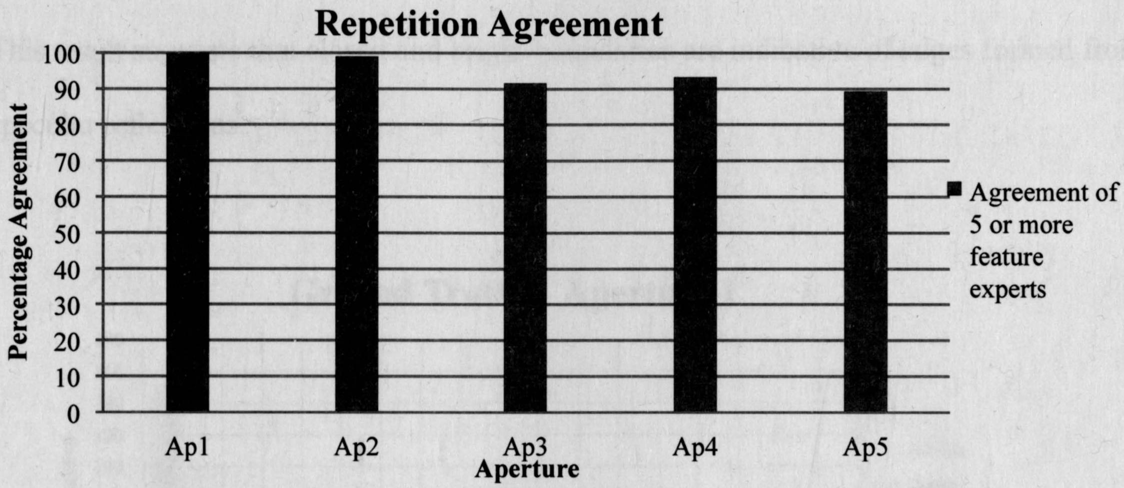

Figure 5. Repetition agreement. The percentage of agreement (y-axis) gives the percentage of ROIs where at least 5 of the 7 feature experts agreed on the rating for repetition for each aperture (x-axis).

Figure 6 gives the frequency of occurrence of each edge type as a function of feature combination for aperture 1. Not surprisingly, the results indicate that almost all of the ROIs were reported as having no closure, no texture and no repeating elements. This is likely due to the minimal amount of information available due to the small aperture size. Similar results are shown in figure 7 for aperture 2. However, there were several ROIs rated as having a single texture (feature combination 2). In addition, feature combination 4 shows that when ROIs are rated as having a closed boundary with a brighter interior, they are almost always specular edges.

Figure 8 gives the frequency of occurrence of each edge type (actual cause) as a function of feature combination for aperture 3. ROIs rated as being closed and bright with no texture and no repetition (feature combination 4), closed and bright with one texture and no repetition (feature combination 5), closed and bright with no texture and a repeating element (feature combination 13) and finally, closed and bright with one texture and a repeating element (feature combination 14) all seem to indicate that the edge is specular. This result suggests that closed and bright boundaries are indicative of edges formed from specular reflections.

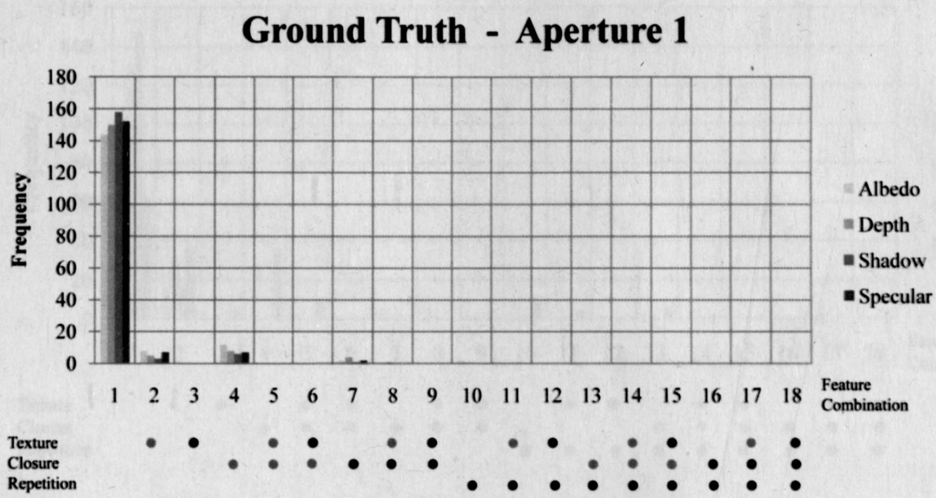

Figure 6. Ground truth aperture 1. Shown is the frequency of each ground truth cause (y-axis) plotted as a function of feature combination (x-axis) for all feature experts at aperture size 1. The dots below the x-axis represent the features associated with each feature combination. For each feature, no dot represents the absence of that feature, a gray dot represents a rating of 1 and a black dot represents a rating of 2.

ROIs rated as having no closure, a single texture and no repeating elements (feature combination 2), indicate that the edge is likely a depth or shadow edge. This makes sense for depth edges because it is quite conceivable to have a depth discontinuity between a textured and non-textured surface. While it is possible to have a shadow edge with a single texture, it is not nearly as likely. Recall that in our definition of texture, a single texture that is present on both sides of a single boundary should be rated as a non-texture. This is because we are interested in whether or not texture itself is responsible for the creation of an edge. If a shadow edge is formed from a single texture by our definition, the texture would have to abut the edge and terminate at that boundary of the shadow.

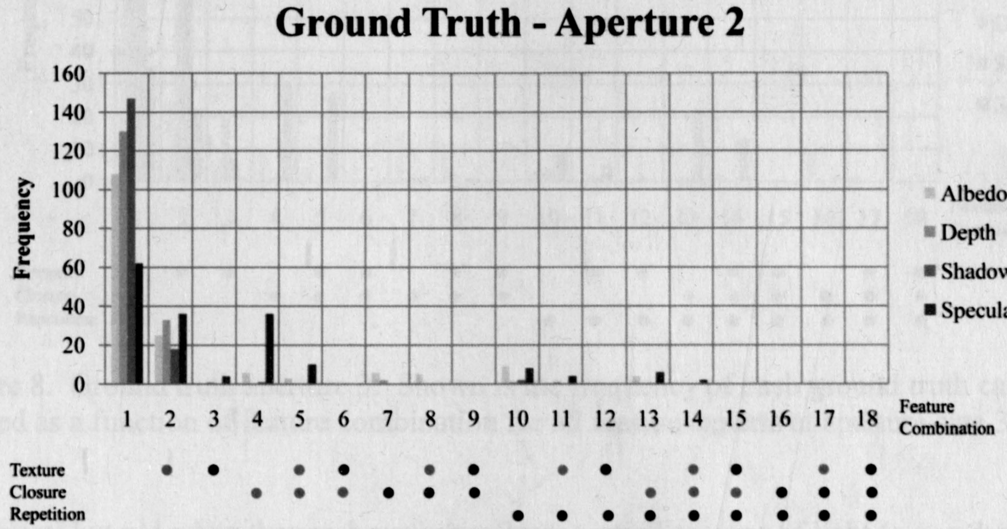

Figure 7. Ground truth aperture 2. Shown is the frequency of each ground truth cause plotted as a function of feature combination for all feature experts at aperture size 2.

Though only a small number of ROIs were rated as having closed boundaries with darker interior luminance, it is of note that this feature is almost always associated with albedo edges. This is true for ROIs rated as having no texture and a single texture (feature combinations 7 and 8) as well as no texture and a repeating element (feature combination 16). This makes sense in the non-textured cases as an albedo edge indicates a change in the surface reflectance of an object. For example, the dot of a lower case *i* printed on a piece of paper in black ink creates a closed boundary with darker interior luminance. For the single texture cases, the interpretation is not so clear. It is possible that a subtle texture may be visible when a large amount of light is reflected (e.g. a brightly painted portion of painted surface). It is also possible that the closed and dark boundary is part of the overall

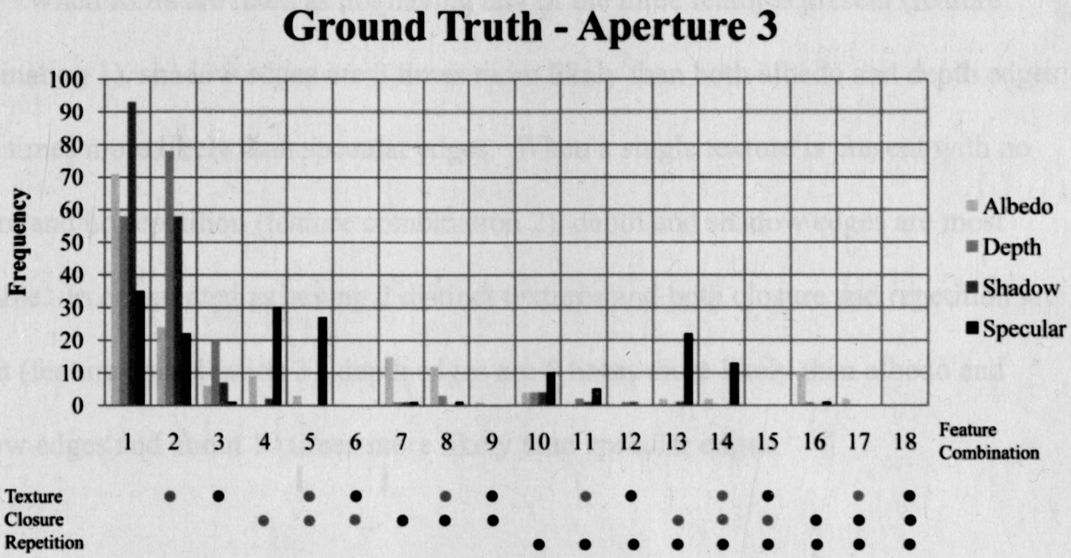

Figure 8. Ground truth aperture 3. Shown is the frequency of each ground truth cause plotted as a function of feature combination for all feature experts at aperture size 3.

an object) but not when the same surface reflects a small amount of light (e.g. a dark gray texture of the surface that, when viewed at a smaller aperture size, appears to be a separate area of the same surface.

At apertures 4 and 5, ROIs rated as having closed and bright boundaries with no texture or a single texture are almost exclusively specular edges. This is further support for the predictions for specular edges. Albedo edges are most probable in ROIs with no closure, no texture and the presence of a repeating element (feature combination 10) and in cases with closed and dark boundaries with either no texture or one texture (feature combinations 16 and 17 respectively). This indicates that repetition, as well as closed and dark boundaries may be important predictors of albedo edges. Figure 9 shows the frequency of each edge type as a function of feature combination for all feature experts at aperture 4. Similarly, figure 10 shows the frequency of each edge type as a function of feature combination for all feature experts at aperture 5.

When ROIs are rated as not having any of the three features present (feature combination 1), shadow edges are 3 times more likely than both albedo and depth edges and 6 times more likely than specular edges. When a single texture is present with no closure and no repetition (feature combination 2), depth and shadow edges are most probable. In cases rated as having 2 distinct textures and both closure and repetition are absent (feature combination 3), depth edges are 5 times more likely than albedo and shadow edges and about 10 times more likely than specular edges.

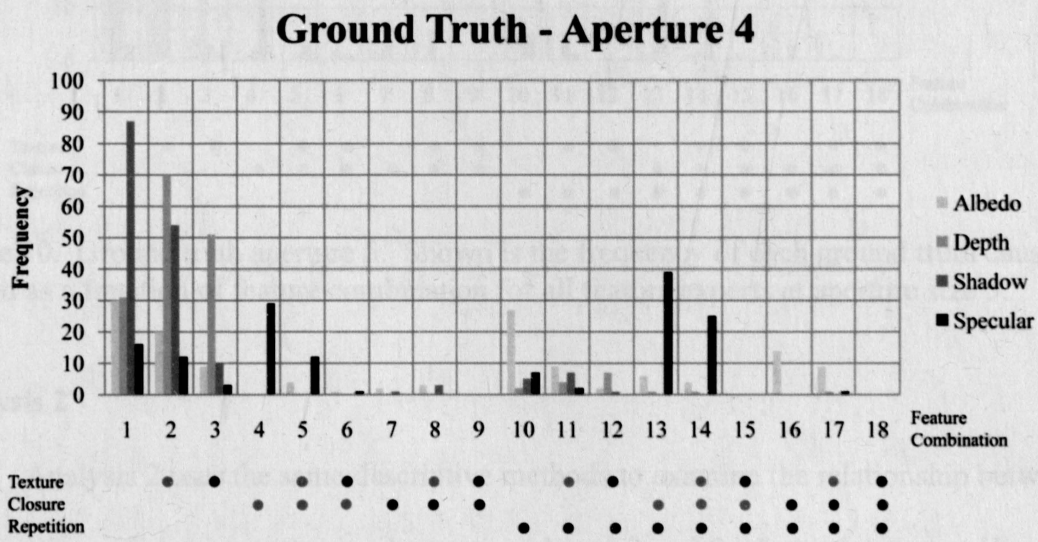

Figure 9. Ground truth aperture 4. Shown is the frequency of each ground truth cause plotted as a function of feature combination for all feature experts at aperture size 4.

It is clear that while these features may help disambiguate different edge types, they are certainly not the only image features involved. Even at the largest apertures a significant number of the ROIs used in this study were rated as lacking all three of the image features of interest. Although a lack of features could be informative in some cases, the results of this study indicate that more information is needed to disambiguate edge types at the scales used. Moreover, this study suggests that scale selection plays an important role in object recognition (McDermott, 2004).

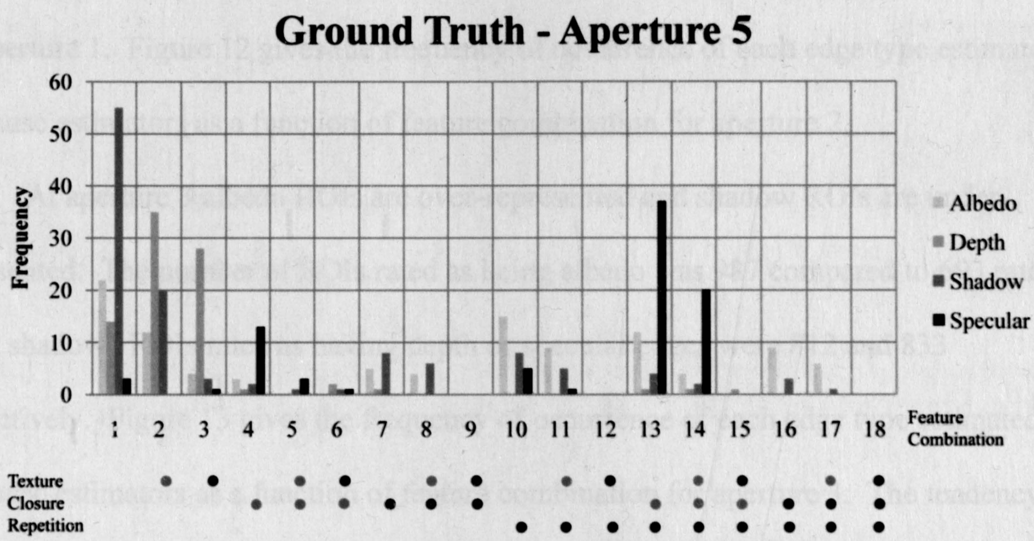

Figure 10. Ground truth aperture 5. Shown is the frequency of each ground truth cause plotted as a function of feature combination for all feature experts at aperture size 5.

Analysis 2

Analysis 2 uses the same descriptive methods to examine the relationship between combinations of closure, texture and repeating elements and the four edge types. However, the observer generated cause estimation data was used in place of the ground truth data. Using the cause estimation data rather than ground truth shows whether or not human subjects are in fact using the three image features to interpret the cause of the edge (rather than whether or not they should). To simplify interpretation, the five cause estimators for each aperture were combined into a single graph. Combining the data from all participants used in the cause estimation simplifies interpretation and represents the average cause estimator.

Both apertures 1 and 2 show nearly identical trends to apertures 1 and 2 of analysis 1. Very few ROIs were rated as having any of the three features present at either aperture. Again, this is due to the small amount of visual information present in the ROI. Figure 11 gives the frequency of occurrence for each edge type as a function of feature combination for aperture 1. Figure 12 gives the frequency of occurrence of each edge type estimated by the cause estimators as a function of feature combination for aperture 2.

At aperture 3 albedo ROIs are over-represented and shadow ROIs are under-represented. The number of ROIs rated as being albedo was 987 compared to 693 rated as being shadow. ROIs rated as having depth or specular edges were 812 and 833 respectively. Figure 13 gives the frequency of occurrence of each edge type estimated by the cause estimators as a function of feature combination for aperture 3. The tendency for participants to rate ROIs as having an albedo edge is most apparent for ROIs rated as having no closure, no texture or one texture and no repeating elements (feature combinations 1 and 2).

ROIs rated as having feature combination 1 are most likely rated as having an albedo edge for the cause estimation data but a shadow edge for the ground truth data. Similarly, albedo edges are as likely as depth edges for feature combination 2 as observed by the cause estimators. This was not the case for feature combination 2 in the ground truth data. Both of these results is likely due to the over estimation of albedo edges by the cause estimation group. The results for aperture 3 for both the cause estimators and the ground truth data are otherwise similar.

At aperture 4, an over-estimation of shadow edges and underestimation of depth edges was observed. The number of ROIs rated as being shadow was 1029 compared to

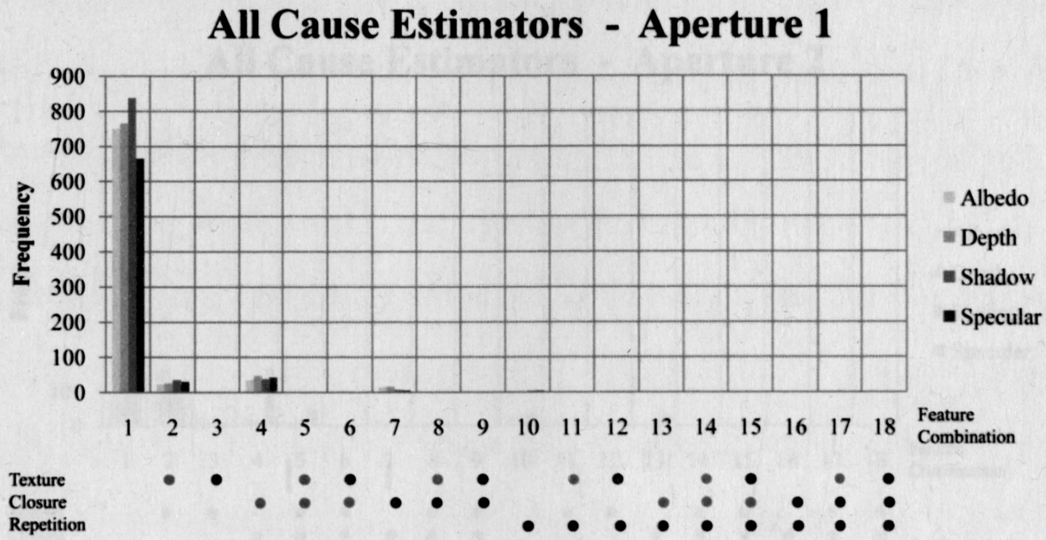

Figure 11. Cause estimators aperture 1. Shown is the frequency of each cause estimation plotted as a function of feature combination for all feature experts at aperture size 1.

602 rated as being depth. ROIs rated as having albedo or specular edges were 756 and 728 respectively. Figure 14 gives the frequency of occurrence of each edge type estimated by the cause estimators as a function of feature combination for aperture 4. The over-estimation of shadow edges occurs most often when ROIs are rated as having no closure, 1 texture and no repeating elements (feature combination 2).

ROIs rated as feature combination 2 are most likely to be rated as being shadow edges for the cause estimation data. This contrasts with the ground truth data where ROIs rated as feature combination 2 are most likely to be rated as being depth edges. The difference between the cause estimation and ground truth data for feature combination 2 is likely due to the tendency for participants to over-estimate the occurrence of shadow edges and under-estimate depth edges. The data for aperture 4 in both analyses is otherwise similar.

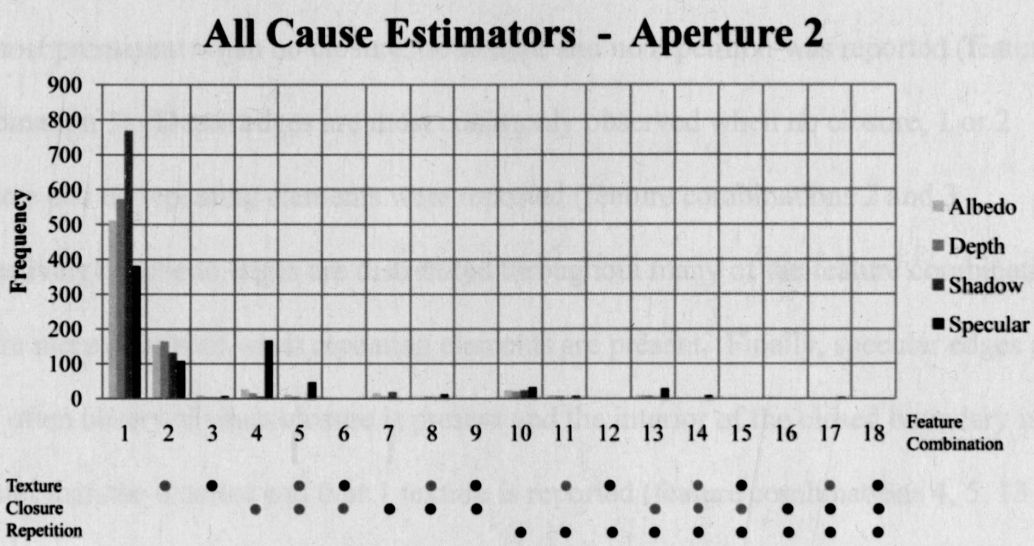

Figure 12. Cause estimators aperture 2. Shown is the frequency of each cause estimation plotted as a function of feature combination for all feature experts at aperture size 2.

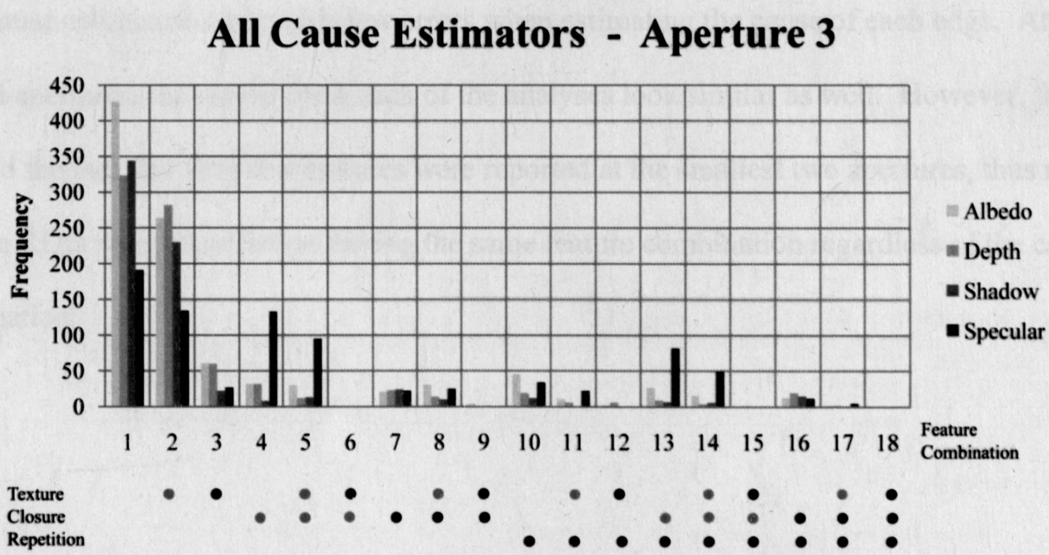

Figure 13. Cause estimators aperture 3. Shown is the frequency of each cause estimation plotted as a function of feature combination for all feature experts at aperture size 3.

At aperture 5, we see similar results to those observed in analysis 1. Shadow edges are most prominent when no closure, no texture and no repetition was reported (feature combination 1). Depth edges are most commonly observed when no closure, 1 or 2 textures and no repeating elements were reported (feature combinations 2 and 3 respectively). Albedo edges are distributed throughout many of the feature combinations, but are more dominant when repeating elements are present. Finally, specular edges are most often observed when closure is present and the interior of the closed boundary is brighter than the exterior and 0 or 1 texture is reported (feature combinations 4, 5, 13 and 14). Figure 15 gives the frequency of occurrence of each edge type estimated by the cause estimators as a function of feature combination for aperture 5.

Recall that in the preliminary study, participants performed quite well at the cause estimation task when sufficient visual information was present (i.e. aperture 5). Essentially the cause estimation data is similar to the ground truth data at large aperture sizes because the cause estimators make very few errors when estimating the cause of each edge. At small apertures, the results from each of the analyses look similar as well. However, this is due to the fact that very few features were reported at the smallest two apertures, thus most of the ROIs were classified as having the same feature combination regardless of the cause estimation.

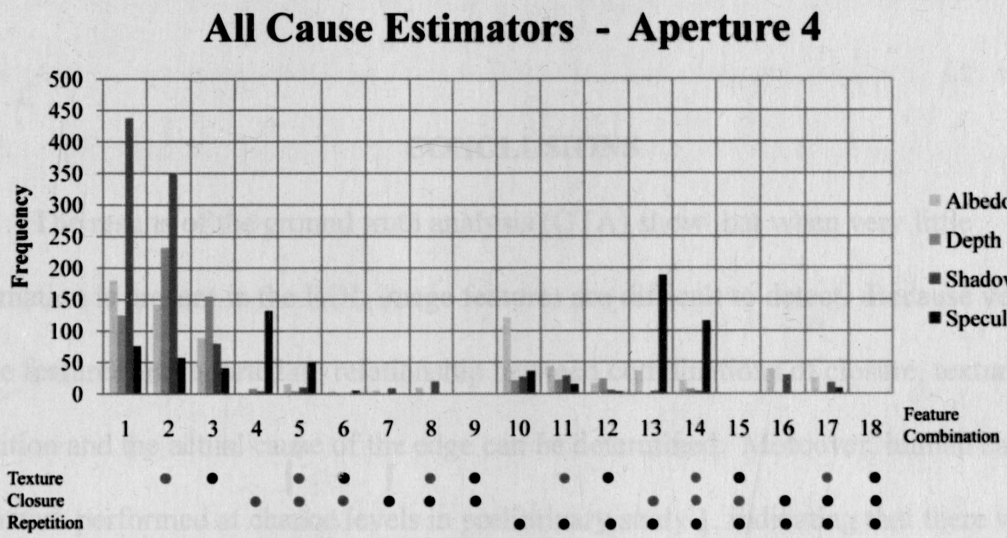

Figure 14. Cause estimators aperture 4. Shown is the frequency of each cause estimation plotted as a function of feature combination for all feature experts at aperture size 4.

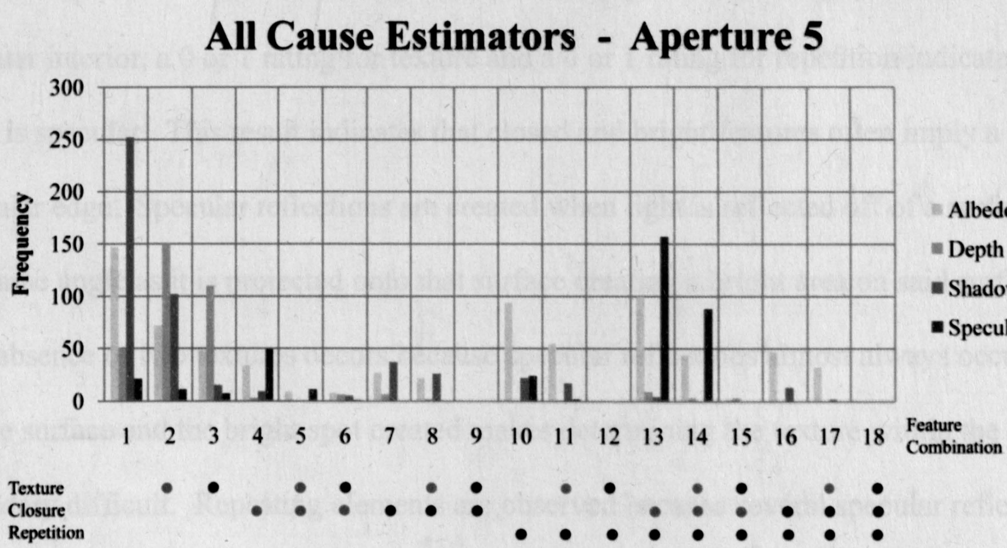

Figure 15. Cause estimators aperture 5. Shown is the frequency of each cause estimation plotted as a function of feature combination for all feature experts at aperture size 5.

CONCLUSIONS

The results of the ground truth analysis (GTA) show that when very little information is present in the ROI, image features are difficult to detect. Because very few image features are reported no relationship between combinations of closure, texture and repetition and the actual cause of the edge can be determined. Moreover, human cause estimators performed at chance levels in preliminary study 1, indicating that there was not enough information present in the ROIs to determine the cause of each edge. The same can be said for the cause estimation analysis (CEA). Both analyses provide evidence that when little information is available (i.e. the aperture is small), closure, texture and repeating elements cannot be detected. Moreover, a relationship between closure, texture and repeating elements cannot be established.

The GTA for aperture 3 indicates that ROIs reported as having closure with a brighter interior, a 0 or 1 rating for texture and a 0 or 1 rating for repetition indicate that the edge is specular. This result indicates that closed and bright features often imply a specular edge. Specular reflections are created when light is reflected off of a surface at the same angle as it is projected onto that surface creating a bright area on said surface. The absence of two textures occurs because specular reflections almost always occur on a single surface and the bright spot created makes determining the texture within the closed boundary difficult. Repeating elements are observed because several specular reflections can occur on a single surface in close proximity.

At aperture 3 there are still a majority of ROIs reported as having no closure, no texture and no repeating elements. This is likely due to the limited amount of information present within the ROI. However, a relationship between closed and bright features and specular edges emerges at aperture 3 for both analyses. Closure is more easily detected when other image features are not because closure does not necessarily require a large viewing area to be present. Moreover, specular reflections are generally small, meaning that an entire specular reflection boundary can often be observed when the viewing aperture is small. This means that closed boundaries formed from specular reflections generally make-up only a small portion of a surface. It is not surprising then that this relationship would appear at a smaller aperture.

The CEA for aperture 3 gives similar results to those from the GTA. Again, the majority of the ROIs are reported as having no closure, no texture and no repeating elements. However, a relationship was observed between the closed and bright feature and specular cause estimation. Additionally, the cause estimators from preliminary study 1 are more sensitive to specular edges than any other edge type at aperture 3. Both analyses provide evidence that humans may in fact use closure when determining whether an edge is caused by a specular reflection.

The GTA for aperture 4 further indicates a relationship between closure and specular edges. Albedo edges are often observed when there is no closure or when closed and dark boundaries are present, 0 or 1 texture is reported and a repeating element is observed. Put simply, when a repeating element is present and a closed and bright feature is not present, ROIs are most likely to be albedo. It is helpful to think of both shadow and depth edges to understand why albedo edges are most likely when a repeating element is present. Shadow edges are less likely to have a repeating element given the size of the ROI. Though repetition would be present in say the shadow cast from a tree branch, the area of the image necessary to determine repetition when a shadow edge is present is often much greater. In addition, shadow edges should almost always be rated as having no texture. Depth edges generally lack repeating elements as well. Based on the instructions of the experiment, a repeating element must be a defining factor in the formation of the edge. Depth edges are rarely formed from a repeating element. In addition, those that are formed from repeating elements will likely be rated as having at least one distinct texture. Albedo edges can be formed from many different combinations of features. It is not surprising that albedo edges are more common than depth and shadow edges when a repeating element is present and more common than specular edges when closed and dark boundaries are observed.

Aperture 4 also provides evidence that when closure, texture and repetition are not present, shadow edges are most likely the cause of the edge. Given that most shadows are large, and create an area where less light interacts with the surface being interfered with, closed and bright features are never reported as being shadow edges. In addition, because the area the shadow is cast on is almost exclusively a single texture, the absence of texture as a defining feature of the edge is expected. Finally, it is unlikely that a repeating element is present in a shadow edge when only a small amount of visual information is present. It is not surprising then that shadow edges are 3 times more likely than any other edge when no image features are reported.

Depth edges are most likely found when no closure is present, 1 or 2 textures are observed and no repeating elements are present. It is quite common to observe an object with one texture that occludes another with a different texture (2 texture case). Given the definition of texture, it is also common to see an object with no texture occluding or being occluded by an object that is textured. The size of the viewing aperture likely limits the occurrence of repeating elements being observed when a depth edge is present. Put simply, it appears that the presence of at least one texture is most important in determining a depth edge. Moreover, when two distinct textures are present, depth edges are 5 times more likely than any other edge type for the ground truth analysis (GTA).

The cause estimation analysis (CEA) for aperture 4 provides support for the majority of the results found in the GTA 1 exception. Feature combination 2 (no closure, 1 texture and no repeating element) shows a strong association for shadow edges rather than depth edges. As explained above, this likely occurred because of an over-estimation of shadow edges by the cause estimation group. When the over-estimation of shadow edges is considered, the results for aperture 4 of the CEA are similar to those obtained in the GTA.

Finally, both analyses for aperture 5 appear similar to the results of the GTA for aperture 4. This is the case even though 40 of the 96 ROIs were eliminated because the feature expert group reported seeing an object (i.e. answered "YES" to question 7 in the feature estimation experiment). At aperture 4, only 7 ROIs were eliminated. It appears that at aperture 4, features are able to be detected even when there is not enough information to determine an object in the visual scene. Moreover, these features provide enough information to assist in edge classification.

DISCUSSION

This study brings to light some interesting relationships between image features and the cause of edges in natural visual scenes. Regardless, there are several limitations that need to be addressed. First, this study was designed as a precursor to developing a probabilistic model with image features as predictor variables of edge type. The groups used for data collection were designed with the intention of using the data in the development of such a model. The methods used for this study limited the type of statistical analyses that could be used, thus the interpretation of the results is largely data driven and descriptive. Future studies of this kind should be designed in a manner appropriate for inferential statistical analysis.

Second, the definition used to define texture was open to interpretation by the observer. While guidelines were given as to what should be considered a texture, the goal was to learn how human subjects interpret visual scenes. Thus, some participants did not agree on texture ratings. Regardless, a more concise definition may have helped provide greater separation of edge types in the final analysis.

Another limitation was the stimuli used in the experiment. Several different objects were photographed in a well controlled environment to represent the edge types and image features thought to be most important in this study. Nevertheless, the stimuli used may have underrepresented the features found in the natural world. Moreover, while the ground truth cause for each edge was established for each of the ROIs, we did not establish a ground truth for the image features present in each of the ROIs. Future studies would benefit from using a greater variety of stimuli as well as establishing a ground truth for the image features.

Finally, though it was necessary to use two different groups of participants for our study, it would be interesting (as well as beneficial) to use the same group of participants for both the feature estimation and cause estimation data. This could be accomplished by presenting the data in two different studies separated by a large enough period of time or by altering the presentation of ROIs in a manner that does not alter the information present in each. In a related study, McDermott (2004) rotated image ROIs to alter presentation of images so participants were unaware that they had viewed the same image more than once. This was effective and would have allowed for the use of inferential statistical methods and a more meaningful interpretation of the results.

Even with these limitations, the results of this study seem to indicate that humans may use closure, texture and repetition when segmenting objects. Perhaps more interestingly, humans appear to do this only when enough information is present to determine these features but not enough information is present to determine specific objects in the visual scene. At apertures 1 and 2, very little separation is observed between edge types as a function of feature combination (see figures 6, 7, 11 & 12). This occurs because almost all of the ROIs are rated as having no closure, no texture and no repetition. However, at aperture 4 (figures 9 & 14), the majority of the ROIs are still unrecognizable, yet there is greater separation between edge types as a function of feature combination. At aperture 5 (figures 10 & 15), 40 of the 96 images were omitted because participants recognized an object. This seems to support the idea of scale posed by Marr (1982) in that certain features are more salient as more information becomes available. Humans do seem to use different strategies for identifying the type of edge present in a visual scene depending on the amount of visual information available.

Perhaps the most interesting relationship observed in the current study is that of closure and specular edges. This relationship began to emerge as early as aperture 2 (figures 7 & 12) and was consistent through aperture 5 (figures 10 & 15). In preliminary study 1, specular edges were the most accurately detected as indicated by greater sensitivity, (ability to correctly detect specular edges) greater specificity (ability to correctly identify edges that are not specular) and greater d' values when compared to the other 3 edge types (see Table 1). In preliminary study 2, the MSSE classification algorithm performed best when specular edges were the cause of the comparison ROI. Given the relationship observed in the current study, it appears that humans do use closure (with brighter interior) as an important feature for distinguishing specular edges from albedo, depth and shadow edges.

While there were other relationships between image features and edge types observed in the current study, it would be difficult to conclude that these relationships were as meaningful as the one seen for closure and specular edges. This is due to the large number of more than one edge type found for most of the feature combinations. For example, at aperture 4 for the ground truth data, shadow edges are 3 times more likely than both albedo and depth edges when no closure, no texture and no repeating elements are reported (see figure 9). While this is interesting, there are still a significant number of albedo and depth ROIs related to the same set of image features. Because edge types often share some of the combinations of these image features, future directions should include probabilistic modeling to best represent human performance. In addition, probabilistic models can be easily adapted to accommodate new information obtained from other studies of a similar nature. For example, if a second study was using color as a feature, the results obtained could be added to the model to determine whether color leads to greater separation, less separation or has no effect on edge classification when the information from closure, texture and repetition are also present.

Future directions regarding object recognition should include additional image features (e.g. color, junction type) to compliment the three used in the current study. This may lead to greater separation of edge types based on feature combinations. Human subjects do not perform perfectly when identifying different types of edges when limited information is available. Because human performance is not perfect, the development of a probabilistic model would be a useful and logical next step in developing a theory relating image features to image segmentation. Such a model would provide a realistic representation of how the human visual system segments a visual scene based on varying amounts of information.

REFERENCES

Brady, M. & Padmanabhan, G. (2007, May). Estimation of Contrast Origin in Natural Images.Poster session presented at the annual meeting of the Vision Sciences Society, Sarasota, Florida.

Elder, J.H. & Zucker, S.W. (1998). Evidence for boundary specific grouping. *Vision Research, 38(1),* 143-152.

Gamble, E.B., Geiger, D., Poggio, T. & Weinshall, D. (1989). Integration of vision modules and labeling of surface discontinuities. *IEEE Transaction on Systems, Man, and Cybernetics, 19(6)*, 1576-1581.

Goldstein, E.B. (2002). Sensation and Perception (6th Ed.). Pacific Grove:Wadsworth.

Hubel, D.H. & Wiesel, T.N. (1959). Receptive fields of single neurones in the cat's striate cortex. *The Journal of Physiology, 148*, 574-591.

Kuffler, S.W. (1953). Discharge patterns and functional organization of mammalian retina. *The Journal of Physiology, 16*, 37-68.

Marr, D. (1982). Vision. New York:W.H. Freeman and Company.

McDermott, J. (2004). Psychophysics with junctions in real images. *Perception, 33,* 1101-1127.

Nawrot, M. (2003). Depth from motion parallax scales with eye movement gain. *Journal of Vision, 3,* 841-851.

Nawrot, M. (2003). Eye movements provide the extra-retinal signal required for the perception of depth from motion parallax. *Vision Research, 43,* 1553-1562.

Nawrot, M. & Joyce, L. (2006). The pursuit theory of motion parallax. *Vision Research, 46,* 4709-4725.

Nawrot, M. & Stroyan, K. (2009). The motion/pursuit law for visual depth perception from motion parallax. *Vision Research, 49*, 1969-1978.

Pinel, John J. (2009). Biopsychology (7^{th} ed.). Boston: Pearson.

Poggio, G.F. & Poggio, T. (1984). The analysis of stereopsis. *Annual Review of Neuroscience, 7, 379-412.*

Poggio, T., Torre, V. & Koch, C. (1985). Computational vision and regularization theory. *Nature, 317*, 314-319.

Richards, W.A., Koenderink, J.J. & Hoffman, D.D. (1987). Inferring three-dimensional shapes from two-dimensional silhouettes. *Journal of the Optical Society of America, 4(7),* 1168-1175.

Wheatstone, C. (1838). Contributions to the physiology of vision: Part I. On some remarkable and hitherto unobserved phenomena of binocular vision. *Philosophical Transactions of the Royal Society of London, 128*, 371–394.

Zhaoping, L. (2003) V1 mechanisms and some figure-ground and border effects. *Journal of Physiology, 97,* 503-515.

APPENDIX A. PARTICIPANT INSTRUCTIONS

Edge Feature Classification

You will be asked to look at an image of a scene as viewed through an aperture in a wall. To enhance the illusion, you will be required to view the screen through a stereoscope. Seat yourself comfortably in front of the screen. Adjust the seat or monitor height for optimum viewing. You will be asked to answer questions about a number of different scenes. Follow the instructions below when examining each scene. There are controls to start the session, to advance to the next scene, to revert to the previous scene, to end the session, to toggle the edge location indicator, and to hide the controls.

Controls:

START button – Starts the session.

NEXT button – Advances to the next scene.

PREVIOUS button – Reverts to the previous scene.

DONE button – You may press this at any time to end a session. All your answers are recorded, and restored if started again at a later time.

SPACEBAR – toggles the appearance of a red dot indicating where the edge of interest is located.

HIDE CONTROLS – Hides the controls.

1. **Closure** – Open or Closed

This indicates whether the most salient contour to which the edge belongs is open or closed. If the edge is part of a contour which wraps around within the aperture to form a closed boundary, the answer is closed. If the edge is part of a contour which continues onward to the border of the aperture, or simply dissipates, the answer is open.

2. **Brighter Inside** – Yes or No

This question is only applicable if the edge belongs to a closed contour. A closed contour will always form an inside region and an outside region. If the inside region appears brighter, then indicate yes, otherwise, indicate no.

3. **Contour Trace** – Trace the contour to which the edge belongs.

This question is only applicable if the edge belongs to a closed contour. Use the mouse to drag out the trace of the closed contour. You will generally begin at the edge of interest and trace along the contour until reaching the edge of interest again. If successful, the message box will indicate the area of the closed region. If not, add strokes to complete the closure, or restart by pressing the clear button.

4. **Texture** – 0, 1, or 2

The contour to which the edge belongs may form a boundary between two regions. These regions may or may not be textured. If both regions bordering the contour are comprised of the same texture, or neither region bordering the contour is textured, indicate 0. If exactly one region is textured, indicate 1. If exactly two regions are textured, indicate 2.

If the contour boundary is open, but does not extend across the aperture, there is only one region formed, and so you must indicate 0, since the contour cannot form a boundary between any textures.

Regarding the extent of the regions:

If the contour boundary is not part of a repeating structure, and is closed then the two regions are simply "the inside" and "the outside" of the boundary. The outside region continues to the aperture boundary.

If the contour boundary is not part of a repeating structure, is open, and extends across the aperture, then the two regions are simply those on opposing sides of the boundary.

If the contour boundary is part of a repeating structure, whether open or closed, the two regions formed are those of the repeating structures, and everything excluding the repeating structures.

Regarding texture definitions:

Once the regions have been identified, each region must be characterized as being textured or not. We are using a specific criteria based on our quadratic modeling to define what is required for a region to be considered textured. A region is textured if it cannot be modeled by a quadratic luminance profile. In more practical terms, a region can be profiled using a quadratic model of luminance if

1) The region varies smoothly in luminance. (no step edges)
2) The region has at most only one change in the direction of luminance. That is, only the following types of changes in luminance are allowed.

   a) a ramp gradually from dark to light

   b) a ramp gradually from light to dark

   c) a peak gradually from dark to light to dark

   d) a valley gradually from light to dark to light

   e) a uniform luminance

3) The changes must occur in a direction orthogonal to the contour at the point where the edge is located.

So, again, if a region does not meet the above criteria, it is considered textured.

5. **Repeating** – Yes or No

The contour to which the edge belongs forms a border along one of many repeated structures. The contour may be open or closed. The regions within the repeated structures should "feel" like they can be grouped together by the way they look. They need not be regular in their shape, spacing, or size.

6. **Object** – Yes or No

If one or more regions formed by the contour to which the edge belongs stands out as being part of, or being an entire, nameable object, indicate yes, otherwise no. Geometric shapes do not constitute objects, nor do categories like wedge.

7. **Object Description** - Free Response

Only applicable if one or more regions formed by the contour to which the edge belongs stands out as being part of, or being an entire, nameable object. If this is the case, describe the object recognizable in less than a few words.

APPENDIX B. EXPERIMENT PROTOCOL FOR PRELIMINARY STUDY 1

Image Interpretation Study
Experimental Protocol

Before the participant arrives:

1. Setup the training image on the experimenter's computer.
    a. Open the folder containing the experiment (C:\Lab\Experiments\VSS2007\Experiment)
    b. Open the training image "VSS2007-MilkCoke-annotated.tif"
2. Setup **Filter View** on the experimenter's computer to track the edges that the participant is viewing.
    a. Start Matlab.
    b. Change the current working directory in Matlab to the experiment folder.
    c. Add Filter View to the Matlab path by entering the following, "path(path, 'C:\Lab\Tools\Filter View 1.0')" If it has been typed in before, you may simply type "p" and then the up arrow key until the command shows up.
    d. Start Filter View by entering "filterview".
    e. Load the patch sequence data for the block that the subject is running.
    File->Patch Sequence Data->Open
    Open Block1, Block2, Block3, Block4, or Block5 in the experiment folder, depending on which block is being run.
    When prompted for the location of the fvd files, click OK and select the "fvd" folder in the experiment folder.
    When prompted for the location of the image files, click OK and select the "lum" folder in the experiment folder.
    It takes a few minutes to load all the files, be patient.
3. Setup **Patch Recorder** on the experimenter's computer to be ready to record responses.
    At the Matlab prompt enter "recordResponses".
    When prompted, enter the subject group and the subject number. The group will be the block number being run. Verify the block number and subject number by looking at the sign-in sheet. Patch recorder will be loaded up and ready to record responses.
4. Setup the stimulus display computer.
    a. Load up the experiment on the stimulus display computer by navigating to the experiment folder (\\bradyprogrammer\lab\experiments\FolderName\Experiment) and double-clicking the "ContrastOrigin.exp" file.
    b. Configure the experiment by clicking the editor tab, and double-clicking the "ContrastOrigin.sce" file. After it appears in the editor window, scroll down to the line which says "psdFilename = ...". The filename should be

changed to block1.psd, block2.psd,...etc. as required for this experimental run.
    c. Run the scenario, and bring up the trial.

5. Verify you have the correct scenario running by comparing the edge you see on the stimulus display with the edge you see in filter view for trial 1. They should be identical.

After the participant arrives.

1. Greet participant. Make sure he/she is here for the Image Interpretation study.
2. Ask participant whether they have 20/20.
3. Ask participant if they have any vision problems, problems seeing in depth.
4. Ask participant if they wear glasses or contacts, and if so if their prescription is up to date.
5. If they do have corrected vision but don't have their glasses or contacts, reschedule. Otherwise, tell them they don't qualify for the study.
6. Verify their visual acuity by testing their vision with the eye chart. Fill out an eye-chart scoring sheet. Be sure they can't see the sheet. Make sure the chart is illuminated properly and seat them at eye-level with the chart. Stand to the side and help them in aligning the front of their eyes with the tape on the floor indicating the distance from the chart to read from. Ask them to read the chart starting with the left-most column. Score their performance on the eye-chart sheet. Ask them to do the best they can all the way through the columns. If they don't score at least 2 out of 3 at the 20/20 level, they must be disqualified.
7. If they can proceed, ask them to read and sign the consent form.
8. Fill out the sign-in sheet with the information required.
9. Collect demographic information on the back of the eye chart scoring sheet. M/F, Ethnicity, Age.
10. Show the participant the training image for edge types.
    Explain to them the four types of edges and refer to examples in the ones numbered 1-4. Be very terse, but clear. Use the following wording for each edge type.
    a. "You will be looking at parts of different images and asked to identify the cause of the edge you see inside a red circle. The parts you will see will not be from this training image."
    b. "If the edge is caused by a change in depth like in number 1, say depth."
    c. "If the edge is caused by a change in coloring like in number 2, say albedo."
    d. "If the edge is caused by a shadow, like in number 3, say shadow."
    e. "If the edge is caused by a shiny spot like in number 4, say specular."

    Gage whether they understand and explain as necessary.

    Then quiz them on numbers 5-12. Repeat until they get them all correct.

11. Make sure the experimenter's monitor is tilted away from their view.
12. Seat them in front of the monitor with the stereoscope positioned properly for viewing. Make sure they are at the correct height to be viewing comfortably. Verify

that they can see clearly in stereo by bringing up trial 1, and asking them if the red circle is sharp, and in focus. Be sure that they are only seeing one red circle. Make adjustments as necessary before proceeding. If possible, view the image, and adjust the stereoscope so that you can see the image clearly. Then check if that works for the participant.

13. Explain to them what they will be doing.
    "You will be presented with a series of images. When looking through the stereoscope, you will see a wall with a hole in it. Inside the hole, there should be an image. Your task is to tell me the cause of the edge that you see within the red circle. If you see more than one edge, or no edge at all inside the red circle, ask us to clarify what you should be seeing. Also, if any of the images appearing in the hole look like an object, a part of an object, or some material, tell us what it looks like to you. Between trials, there will be an image containing a coke bottle, a milk bottle, and an apple. This is to remind you of how those familiar objects would appear in the images you see. If at any time you forget the meaning of the four possible responses depth, albedo, shadow, and specular, ask us, and we will remind you. When you complete each trial, press space, and wait for me say ok before continuing, then press space to continue."
14. As each trial is presented, press **NEXT** in the patch recorder to advance to that trial. Before recording every response, be sure that the trial number shown in the patch recorder is the same as the trial number shown at the bottom of the stimulus presentation screen. Record a subject's response by highlighting the region number appearing in the list box at the right in the patch recorder window, and then pressing **Cause** button. Assign one cause only, the participant's response, and press **OK**.
15. Whenever the participant tells you they recognize the image as an object, part of an object, or material, record what it reminds them of in the notes edit box of patch recorder.
16. If the participant has difficulty seeing an edge or sees more than one edge in the ROI. Bring up that patch in filter view, look at the actual cause of the edge listed, and identify the edge that they should be attending to. Describe its orientation and location so they can find it.
17. Make a record if anything goes out of the ordinary in the notes edit box.
18. When finished with the last trial, press **Finish**.
19. Ask them if they want to come back and be paid to do similar sessions later. If so, record contact information for them.
20. Thank the participant.

APPENDIX C. INFORMED CONSENT

CONSENT TO PARTICIPATE IN RESEARCH:

Object Recognition

Mark Brady, Ph.D.
Assistant Professor
Department of Psychology
North Dakota State University
Fargo, ND 58105
231-5752

Research Study
You are invited to participate in a research study of perception being conducted by Dr. Mark Brady, Assistant Professor of Psychology, North Dakota State University.

Method of Participant Selection
You have been selected because you are 18 to 35 years of age, are in general good health, and have 20/20 or corrected to 20/20 vision.

Purpose of Study
Image contrast arises from various sources, including changes in depth, surface pigment, surface curvature, shadow edges, and light reflection edges. Observers are adept at determining the source of contrast when global image information is available. However, it is unknown if contrast origin can be determined by local information alone. The purpose of this study is to determine if contrast origin can be determined from local information alone.

Determining the source of an edge in a 3d scene is a simple task because there is so much global image information available. The purpose of this study is to determine if small parts of the scene can be interpreted correctly in 2d by an observer. Without information from the scene, it is unknown if the observer is able to discern edges due to lighting, color or object contour.

Explanation of Procedures
You will be shown a series of image fragments. For each fragment, you will be asked to identify the source of image contrast. For training purposes, examples of each of the possible contrast sources will provided prior to the start of the experiment. The same task will then be repeated using entire images. The image fragments of interest will be outlined in the whole images. Percent correct will be measured in each case and the results compared. The experiment will take approximately 60 minutes.

Potential Risks and Discomforts

There is no deception whatsoever used in these experiments. As in any experiment, participants may experience fatigue. If the experiment takes longer than anticipated, you will be permitted to rest every 30 minutes.

Potential Benefits
While it is unlikely that you will directly benefit from the results of basic research such as this, it is hoped that through your participation you may gain insight into the subject matter and methodology of experimental psychology. It is also hoped that through continued research on the mechanisms of object recognition we may better understand the how the brain accomplishes this feat.

Alternatives to Participation
If you have been recruited from the NDSU Department of Psychology subject pool, and are participating in fulfillment of a course requirement you should know that the Department of Psychology maintains a policy that the course requirement may be fulfilled in alternate ways (i.e., writing papers, doing presentations or reading extra material). The available alternatives are at the discretion of the instructor, whom you should contact for further information in the event you decide not to participate in this experiment.

Compensation for Participation
Student participants recruited from the NDSU Psychology Department subject pool will receive extra credit in their courses for their participation at the rate of 4 credits per hour.

Assurance of Confidentiality
Any information that is obtained from you during this study and that can be identified with you will remain strictly confidential, and will not be disclosed without your expressed permission. Any publication resulting from this study will identify you by alpha-numeric code only. Data and records created by this project are the property of the University and the investigator. You may have access to information collected on or about you by making a written request to the investigator, but not to information collected on or about others participating in the project.

Voluntary Participation and Withdrawal from the Study
Your participation in this experiment is voluntary. Your decision whether or not to participate will not directly affect your course grade, your present or future relationship with North Dakota State University, the Department of Psychology, or the experimenter. If you decide to participate, you are free to withdraw your consent and discontinue your participation at any time.

Offer to Answer Questions
Please feel free to ask questions now or at any time during the study. If you have questions about this study, you can contact Dr. Mark Brady (231-5752). If you have questions about the rights of research subjects or research-related injury, contact the NDSU IRB Office, (701) 231-8908.

Consent Statement

You are voluntarily making a decision whether or not to participate. Your signature indicates that you have decided to participate, having read the information provided above. You will be given a copy of this consent form to keep. If there are things you don't understand about the study, please ask the researchers before you sign the form.

____

Name of Participant (print) Date

____

Signature of Participant Date

____

Researcher obtaining consent (print) Date

____

Signature of Researcher obtaining consent Date