PERCEPTUAL BIASES IN THE INTERPRETATION OF NON-RIGID STRUCTURE FROM MOTION

By

RYNE CHOI

A thesis submitted to the

School of Graduate Studies

Rutgers, The State University of New Jersey

In partial fulfillment of the requirements

For the degree of

Master of Science

Graduate Program in Psychology

Written under the direction of

Dr. Manish Singh

And approved by

New Brunswick, New Jersey

May 2020

ABSTRACT OF THE THESIS

Perceptual biases in the interpretation of non-rigid structure from motion

by Ryne Choi

Thesis Director: Dr. Manish Singh

While structure-from-motion (SFM) studies have largely focused on the perception of rigidly moving 3D objects, recent studies have shown that observers are good at perceiving certain kinds of non-rigid transformations (Jain & Zaidi, 2011). Our overarching goal is to understand what types of non-rigid transformations are perceivable in SFM. As a first step towards this goal, this study investigates the role of biological plausibility of shape transformations.

Experiment 1 compared the perception of two non-rigid transformations: part-orientation change (which is more common in animate motion, e.g. part-wise articulation of limbs) and part-length change (which is less common). Stimuli consisted of an ellipsoid with a protruding part that underwent a non-rigid length change as the whole object rotated back and forth. We manipulated the extent of length change and subjects judged whether the part was undergoing a length or orientation change. In this experiment, the image of the part was always contained within the silhouette of the ellipsoid. Results showed a clear range where length change was misperceived as orientation change.

Experiment 2 further investigated the misperception of non-rigid length change. The part was now visible in the silhouette. We observed a misperception of non-rigid length change as a rigidly-attached part with an "illusory" non-orthogonal horizontal angle relative to the ellipsoid. Observers adjusted the perceived angle between the part and the

ellipsoid. We then compared the perceived horizontal angle to model predictions based on a reinterpretation of the length change in terms of a fixed horizontal angle between the part and the main body, with no length change. Even with no free parameters, the model closely tracked observers' data.

In summary, an SFM stimulus in which a part undergoes a length change, tends to be misperceived either as an orientation change (Exp. 1), or as a fixed but "illusory" orientation (Exp. 2). The results suggest that the visual system may be biased towards more biologically plausible interpretations of non-rigid motion.

ACKNOWLEDGEMENTS

I would like to express my sincere appreciation to my advisor, Manish Singh, for his guidance through each step of this process. I would also like to specially thank Jacob Feldman, for his thoughtful comments and feedback. My sincere thanks also go to Eileen Kowler, for her helpful comments and suggestions. I want to also thank my labmates in the Perceptual Science Lab and the Visual Cognition Lab for their helpful discussions. And finally, I would like to thank my family.

TABLE OF CONTENTS

Abstract ii

Acknowledgments iv

List of Figures vi

1. Introduction 1

2. Experiments 7

2.1 Experiment 1: Yes/No Task 8

2.1.1 Method 8

2.1.2 Results and Discussion 11

2.2 Experiment 2: Horizontal Angle Adjustment 12

2.2.1 Method 12

2.2.2 Model Predictions 16

2.2.3 Results and Discussion 18

3. General Discussion 20

3.1 Conclusion 23

Bibliography 25

LIST OF FIGURES

1 Cross section of CAD object 9

2 Stimuli shown to subjects and their underlying 3D shape 10

3 Results for Experiment 1:“Yes/No” Task 13

4 Stimulus and adjustment display in Experiment 2 15

5 Predictions of the re-interpretation model 17

6 Results for Experiment 2: Horizontal Angle Adjustment 19

1. Introduction

Structure from motion (SFM) is the ability to infer the 3D structure of an object from 2D motion cues alone. SFM displays eliminate other cues to depth, such as binocular disparity, shading, or perspective. A single frame of an SFM display carries no information about the 3D structure of the object. But from the motion of the display, subjects are able to perceive the 3D structure of the object. Psychophysical studies have shown that human observers are good at determining the 3D structure of many different types of objects from motion cues alone. In one of the earliest studies of SFM, subjects were able to determine the 3D structure of rotating wireframe objects that were projected onto a translucent screen (Wallach & O'Connell, 1953). From a static view of the object's projection onto the screen, subjects were unable to determine the structure of the object. But, as the object started moving and its 2D projection began to deform, subjects were able to determine its structure (Wallach & O'Connell, 1953). In another study, three points were arranged in space and rotated around a central axis. Subjects were asked to determine which of the three points was located "midway in depth between the other two". Subjects were able to perform accurately on this task, demonstrating an ability to perceive the 3D structure involving the relative 3D locations of the three points from 2D projected motion (Hildreth et al., 1990). Even when the experimenters added noise to the x, y position of points in the picture plane independently for each frame, subjects were able to accurately report the relative depths of the points (Hildreth et al., 1990). In the studies just mentioned, the stimuli consisted of connected line segments or individual dots arranged in space. However, most SFM studies use stimuli that consist of 3D objects that are defined by a random-dot texture sprinkled on its surface. Sperling, Landy, Dosher, and Perkins (1989) presented subjects on each trial with one of 53 different possible 3D surfaces that were defined by their dot motion alone. Subjects were asked to identify the shape. Even with a large set of possible shapes, subjects were able to accurately identify the 3D surfaces (Sperling et al., 1989).

The psychophysics literature demonstrates that subjects are good at perceiving a variety of complex 3D shapes from SFM, but the question remains: how does the visual system determine the 3D structure of an object from motion cues alone? The problem the visual system faces is underdetermined by the stimuli: for any 2D motion profile there is an infinite number of possible 3D interpretations. The visual system must use constraints in order to specify a solution. Ullman (1979) proposed rigidity as a constraint the visual system might use in SFM. He was motivated by environmental plausibility: many of the objects in the environment move rigidly; and stationary objects move rigidly in the optical image because of observer motion. He also proved mathematically that rigidity alone could uniquely determine a 3D interpretation. The rigidity assumption states: “Any set of elements undergoing a 2-D transformation which has a unique interpretation as a rigid body moving in space, should be interpreted as such a body in motion” (Ullman, 1979). He then incorporated the rigidity assumption into his structure-from-motion theorem. The theorem states that if there is a rigid interpretation (i.e., if the 2D motion can be interpreted as a 3D object moving rigidly) given three distinct views of four non-coplanar points, that interpretation is uniquely determined. To determine the structure of the whole object, divide the object into “nuclei” of 4 non-coplanar points, test for rigidity in those groupings, and assign those groupings to that rigid interpretation if it exists. This theorem proved a theoretical minimum requirement for three views of four non-coplanar points. It assumes that if there is a rigid interpretation of the motion, the visual system will take that interpretation. Ullman later developed an incremental rigidity scheme to try to deal with non-rigid motion (Ullman, 1983). That model could only handle small deviations from rigid body motion from frame to frame and did not always converge on a stable solution (Todd, 1995; Ullman, 1983).

Several subsequent studies have investigated the psychological relevance of Ullman’s structure-from-motion theorem. Hildreth et al. (1990) found that subjects’ accuracy on their SFM task increased with increasing numbers of distinct frames, consistent with Ullman's incremental rigidity scheme. Todd and Bressan (1990), however, argued that the increase in frames was confounded with the increase in total stimulus duration. Their study challenged the idea of a three distinct frame minimum requirement for human observers. They showed that while three distinct views are necessary for a Euclidean representation of structure, subjects are much better at tasks involving affine structure which can be derived from just two distinct viewing frame (Todd & Bressan, 1990). In both tasks that required specifying Euclidean or affine structures they found that subject performance did not improve when increasing the number of frames beyond two. Hoffman and Bennett (1986) also explored the minimum conditions needed in order to determine a solution in SFM. Like Ullman they used a rigidity constraint but combined it with a fixed-axis motion constraint (that the axis that the object rotates about, does not change orientation with respect to the observer). They found that under these constraints, when angular velocity is constant (which was the case in the stimuli used in Ullman, 1979), only three orthographic projections of two points are needed.

There have also been differing ideas on what the appropriate input representation is to the SFM system. In Ullman's model and the other studies mentioned above, the representation is location based: the image motion is represented as the (x, y) coordinates of the dots at each time frame t. There are also several models that posit velocity-based inputs: the image motion is represented as velocity fields $\boldsymbol{v}(x, y)$, i.e. a motion vector as a function of image location (Braunstein & Andersen, 1984; Domini & Caudek, 2003; Koenderink & Van Doorn, 1986; Koendreink, 1986). While these approaches differ in terms of input representation, they still assume rigidity.

Rigidity-based models have several limitations. Rigidity was proposed as a constraint that "almost always holds true in the environment" (Ullman, 1979). But, many of the objects we encounter in the environment move non-rigidly, especially biological entities. Rigidity based models are unable to interpret non-rigid motion. However, human observers have demonstrated the ability to perceive non-rigidity based on image motion. Sinha and Poggio (1996) demonstrated a compelling counterexample to the rigidity assumption. In their demonstration they presented subjects with rigid wireframe stick figures that resembled the human form. As the figure was rotated back and forth about its vertical axis subjects reported perceiving the figure as a human walking - hence changing its 3D shape non-rigidly. The rigidity assumption states that if a rigid interpretation is available the visual system will take that interpretation. Their demonstration showed a case where a rigid interpretation was available but the visual system chose a non-rigid interpretation. But, in their demonstration, the stimuli used were human stick figures which could be readily recognized as such (indeed their main aim was to investigate the role of learning in perception). Moreover, stick figures are different than the typical SFM stimuli defined by dots covering the surface of a 3D object. In the literature on biological motion (which is related to SFM but has been historically disconnected from it) subjects have been shown to have vivid non-rigid percepts from dots placed at the joints of humans and animals (e.g. Johansson, 1973). Hence, the class of stimuli used in the biological motion literature is a quite different from SFM. Studies on biological motion demonstrate that a rigidity assumption is not necessary for perceiving 3D shape and motion correctly. Even when no rigid interpretation is available subjects can still accurately perceive the structure of the object. It may be that biological motion stimuli engage a specialized module that is specifically tuned to recognize animal motion.

A recent study by Jain and Zaidi (2011), showed that people can be just as good at determining the structure of 3D objects undergoing certain non-rigid motions in SFM displays. In their study they used 3D tubes/generalized cylinders defined by the dot textures on their surface. They found that subjects were just as sensitive to differences in aspect ratio of cross sections of tubes that were non-rigidly flexing in the image plane about the vertical and depth axes, as they were to tubes that were rotating rigidly about those same axes. This study shows that human observers can determine 3D structure from objects moving non-rigidly in an SFM context. And, it again illustrates that the rigidity assumption is too restrictive as far as human perception is concerned.

Our overarching goal is to understand what types of non-rigid transformations are perceivable in a SFM context. To begin to understand the types of non-rigid transformations that are perceivable in a SFM context we started by categorizing the space of non-rigid transformations into those that are more biologically plausible (such as orientation, or curvature change of a part) and those that are not (such as length, location, or width change). Our focus on biological plausibility comes from the idea that the human visual system evolved in an environment where the motion of other biological entities (to which they must have been frequently exposed to and particularly interested in) were constrained to a certain set of non-rigid shape transformations. The movement of these biological entities, particularly other humans, can be defined mostly in terms of part-orientation and part-curvature changes (i.e., the articulation of limbs, such as the raising of an arm, is a non-rigid part orientation change; the bending of a finger, or curling of a tail is a curvature change). These are also the shape transformations that are particularly well captured by a skeleton-based representation of shape (Blum, 1973; Feldman & Singh, 2006). We believe that the visual system may be biased toward (have strong priors for) these types of non-rigid motions. So, we are interested in whether there are differences in perceiving biologically plausible vs. biologically implausible non-rigid transformations in SFM. This could be viewed as a general prior internalized over the course of evolution.

The two experiments presented in this paper focused on two non-rigid transformations we believe to be categorically different: part orientation change and part length change. Part orientation change is a part-wise rigid transformation, meaning that the individual parts of the object move rigidly, but the transformation of the whole object is non-rigid. Part orientation change, is like the articulation of limbs, and we consider it a biologically plausible transformation. Part length change is not a part-wise rigid transformation because the inter-point distances of points within the part itself are changing. We postulate part-length change to be less biologically plausible than part-orientation change, and thus that the visual system is likely to have a bias/preference for perceiving part orientation change over part-length change. [1].

[1] Of course, biological entities do grow their limbs over time, but this does not occur on a time scale that is relevant for visual perception

2. Experiments

To compare the perception of non-rigid part-orientation change and part-length change we started by creating a two-part 3D object consisting of a large vertically oriented ellipsoid with a narrow half ellipsoid part protruding from its center (see Figure 1). This object was defined by a random-dot texture on its surface.

The object was animated to rotate back-and-forth about its vertical axis while the protruding part underwent either an orientation change in the plane of the axis of rotation, or a length change. When we developed the non-rigid length change stimuli we noticed a misperception of non-rigid length change as a non-rigid orientation change. This misperception (or perceptual re-interpretation) of a part-length change in terms of a part-orientation change is consistent with the idea that the visual system is biased toward biologically plausible shape transformations. In the following two experiments we used this two-part 3D object undergoing a non-rigid part length change to further investigate this misperception. In the first experiment we used a yes/no task and seven levels of length change to explore the range of length change magnitudes that lead to a misperception of length change as an orientation change. In the second experiment we used the same type of stimuli, but with the part oriented perpendicular to the main body and visible in the silhouette. In the second experiment we again used seven levels of length change but used an adjustment task to explore a different misperception of length change in terms of part orientation. Both experiments explore how the visual system may be biased towards certain interpretations of non-rigid motion.

2.1 Experiment 1: Yes/No Task

2.1.1 Method

Subjects

Five Rutgers University graduate students with normal, or corrected to normal, vision participated in the study. They were paid for their participation.

Stimulus and Design

The basic stimulus used in this experiment was a two-part 3D object. The object was a large ellipsoid with a protruding half ellipsoid part. The object was drawn in the computer assisted design software, OnShape, and exported as a .STL file to the animation software, Blender. In Blender, the object was animated and was rendered with an orthographic projection. The larger ellipsoid body of the object was $4.58°$ visual angle in width and had a height of $7.62°$ (Figure 1). The protruding part made a vertical angle of $30°$ with respect to the main body (the part was oriented $30°$ upward along the vertical plane) but had no horizontal angle (no orientation along the horizontal plane).

The two-part mesh object was covered in a random-dot texture. The dot texture was created by generating a grid of dots and randomly jittering the x, y position of the dots. We used texture mapping to apply the image of the random dot texture to the 3D object.

At the start of each motion sequence, the object was turned so that the protruding part was facing the observer (though it maintained an elevation angle of $30°$ throughout the sequence). The whole object rotated back and forth about its vertical axis $\pm 25°$. The object rotated to the right (observer's right) $25°$, then back to center, then to the left $25°$, then ended its motion when it rotated back to center. The object rotated with a constant speed throughout the motion. Throughout the motion sequence the protruding part never extended past the 2D projection of the larger body. The stimulus thus always appeared on every frame as an elliptical region filled with dots.

As the whole object rotated back and forth, the protruding part underwent a non-

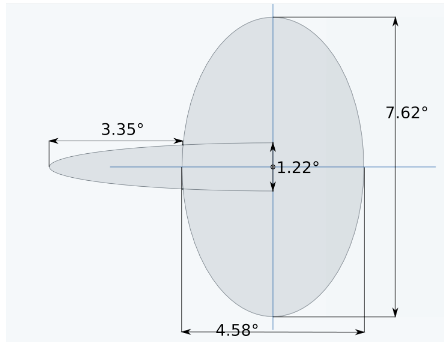

Figure 1: Cross section of two-part object drawn in CAD program OnShape. Dimensions of each part shown in visual angles. Each part was revolved around its axis to create a 3D object and the intersection between the smaller part and body was filleted to create a smooth intersection.

rigid length change. The part length was always longest when the part was facing the observer. As the object rotated away from the observer the part would decrease in length. As the object rotated back to its original position (facing the observer) the part non-rigidly increased in length back to its original length. The degree of length change was manipulated. Figure 2 shows the textured object and its underlying 3D surface when the object was front facing (left) and when the object was rotated $25°$ (right).

There were seven levels of length change used in this experiment. In this experiment we refer to the length changes by length-change ratio. The length-change ratio is the ratio of the part length at its longest (the starting length when the part is front-facing) to its shortest length (length when the whole object is rotated to $\pm 25°$). The starting length of the part was always fixed across all levels of length change. The seven levels of length ratio were: 1.1976, 1.2903, 1.3986, 1.5267, 1.6807, 1.8692, and 2.1053. The experiment was divided into two blocks. In each block each level of length ratio was shown 20 times for a total of 140 trials per block and 280 trials total.

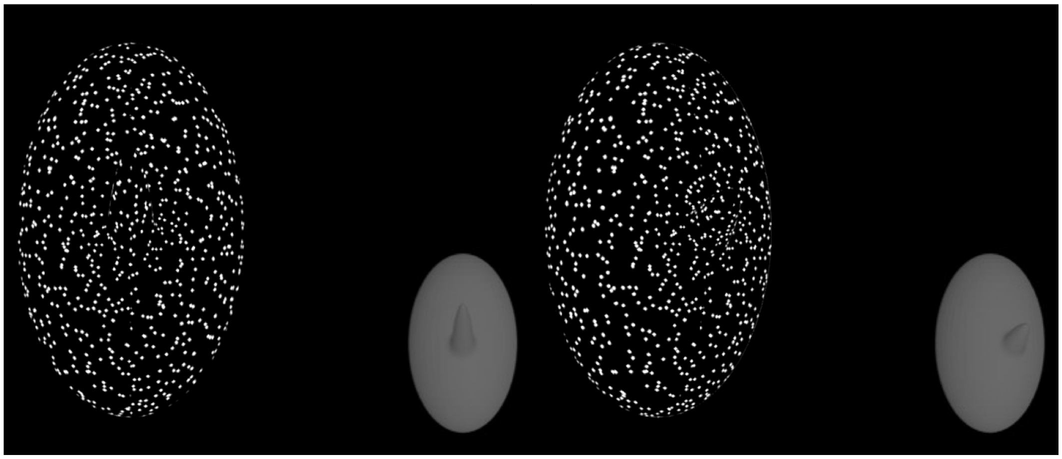

Figure 2: The pair of objects on the left show the first frame of the stimulus when the part is front facing. The textured object is what subjects actually see. The grey object in the bottom corner is the underlying 3D mesh object (subjects never see this). On the right, the object is rotated $25°$ to the right.

Procedure

The experiment was run in a dark room. Subject were seated 75 centimeters from an iMac monitor with a refresh rate of 100Hz and screen dimension of 800x1280 pixels. The experiment was programmed and run in Matlab using the Psychtoolbox libraries (Brainard, 1997).

Instructions were presented before the start of each block. In the first block, subjects were instructed to make a judgment about the protruding part and to respond “yes” or “no” to the following question: Did you see an orientation change? (i.e., a change in the orientation of the protruding part)? Subjects were then shown two examples of orientation change. In these examples, a two-part object made of a cylinder body with a protruding cylinder part underwent a non-rigid orientation change. The protruding cylinder part changed its vertical orientation up and down. In the example demos, the whole object did not rotate back and forth, and the object had a shaded (non-textured) surface. The two examples showed the same non-rigid orientation change but from different viewing angles.

Before the start of the second block, subjects were instructed to make a judgment about the protruding part and to respond yes or no to the following question: Did you see a length change? (i.e., a change in the length of the protruding part)? Subjects were then shown two examples of length change. The same two-part cylindrical object was used in these example stimuli. But, in these examples the part had no vertical or horizontal orientation with respect to the body (the part was perfectly perpendicular to the body) and the part non-rigidly increased and decreased in length.

In both blocks, after watching the two example demos, subjects could press any key to move on to the experimental trials. At the beginning of each trial, one of the seven possible length change ratio stimuli described above was presented. The stimulus was then removed from the screen and a question screen appeared asking in the first block: 'Did you see an orientation change? (i.e., a change in the orientation of the protruding part)?' and in the second block: 'Did you see a length change? (i.e., a change in the length of the protruding part)?'. Subjects used the Right Arrow key to respond 'yes' and the Left Arrow key to respond 'no'. No time limit to respond was imposed. Subjects could take a break between blocks. In each block the order of the seven different length ratio stimuli was randomized. It took approximately 40 minutes for subjects to complete the experiment.

2.1.2 Results and Discussion

The graphs in Figure 3 plot proportion of 'yes' responses as a function of length ratio. Each plot represents data from one subject. The solid blue line represents the proportion of 'yes' responses made by subjects in the first block to the question 'Did you see an orientation change? (i.e., a change in the orientation of the part)?'. The dashed black line represents the proportion of 'yes' responses made by subjects in the second block to the question 'Did you see a length change? (i.e., a change in the length of the part)?'. The error bars represent standard error.

The length response for each subject, represented by the dashed black line, is highly consistent across observers. As the degree of length change (length ratio) increases, the proportion of 'yes' responses to the question 'Did you see a length change?' increases monotonically.

The orientation response, represented by the solid blue line, is less consistent across subjects. However, the difference across observes occurs primarily at the highest levels of length change and there seem to be two response patterns. For some subjects, the “illusion” of length change as orientation change seems to break down at the highest levels of length change (hence the curve goes down after reaching a peak). For other subjects the illusion remains and they continue to perceive an orientation change no matter how large the length change gets.

For each subject there is a clear range of length ratio values where the proportion of ‘yes’ responses to the question ‘Did you see an orientation change?’ is significantly higher than the proportion of ‘yes’ responses to the question ‘Did you see a length change?’. The red shaded regions on the plot in Figure 4 represents regions where the proportion of ‘yes’ responses in the first block are higher than the proportion of ‘yes’ responses in the second block and where the standard errors do not overlap. We observe that even at substantial levels of length change subjects reported seeing an orientation change instead.

2.2 Experiment 2: Horizontal Angle Adjustment

2.2.1 Method

Subjects

Five Rutgers University graduate students with normal, or corrected-to-normal, vision participated in the study. They were paid for their participation.

Stimulus and Design

The same type of two-part 3D objects used in the first experiment were used for this second experiment. The same dot texture used in the first experiment was also used for these stimuli. However there were several key differences in the stimuli for the second experiment. First, in this experiment the protruding part was not pointing upward with respect to the main body. The protruding part was perpendicular to the body in both the

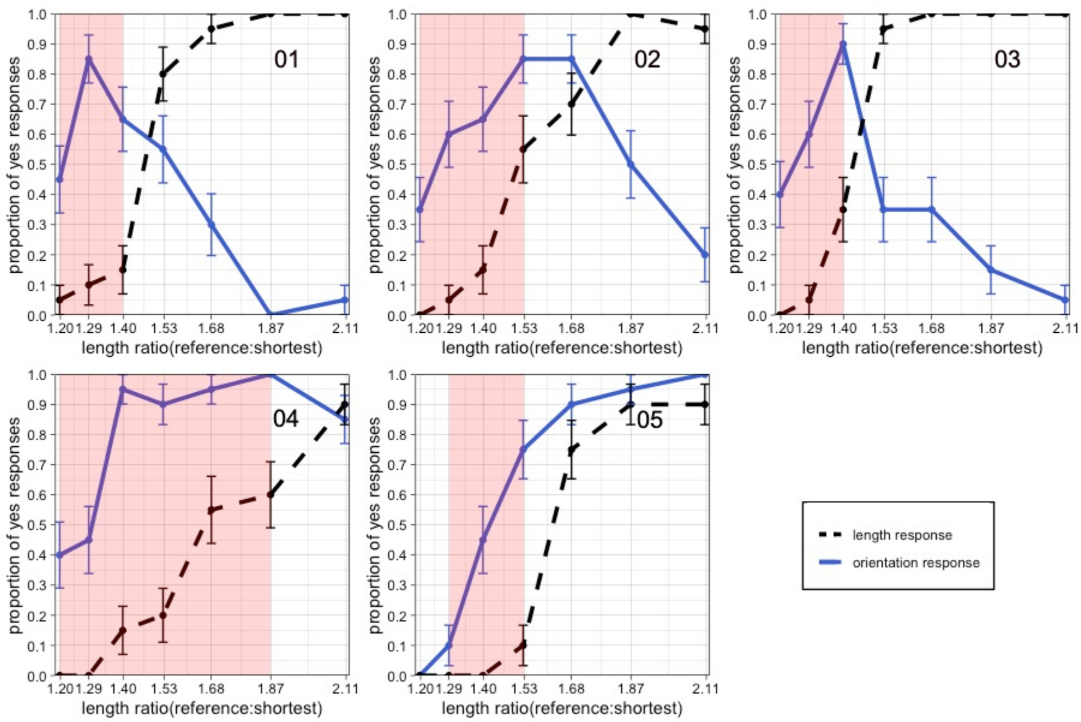

Figure 3: Proportion of "yes" responses to the prompt "Did you see an orientation change?" (plotted as a blue solid line) and to the prompt "Did you see a length change?" (plotted as a dashed black line) as a function of increasing length ratio (magnitude of length change). Each plot shows the data of an individual observer. Error bars represent standard error. Red region represents regions where orientation response was higher than length response and where standard errors do not overlap.

vertical and horizontal directions. Second, the object did not start off with the protruded part facing the subject. Instead, the object was oriented $45°$ degrees to the right, facing away from the subject. Third, the whole object rotated back-and-forth about the vertical axis of the large ellipsoid between $45°$ and $90°$. When the object was rotated $90°$ the part was fully visible in the 2D projected silhouette.

The second experiment also included stimuli where the length of the part would increase non-rigidly. In the first experiment, the length of the part always decreased as the whole object rotated away from its starting position (front facing) and increased as it came back to front. This experiment also had stimuli where the length of the protruding part non-rigidly increased in length as the whole object rotated away from its starting position (at $45°$) till the end of its rotation at ($90°$). Then as the whole object rotated back to its starting position, the length of the protruding part decreased back to its original length.

There were seven levels of the degree of length change i.e., non-rigid length changes of the protruding part. In this experiment we take the length ratio to be the ratio of the length of the part at $90°$ (peak of whole body rotation) to the length of the part at $45°$ (starting point). So in the length increasing conditions the ratios were greater than 1 whereas in the length decreasing conditions the ratios were less than 1. We had 3 levels of increasing length where the part non-rigidly increased to $1.24, 1.48$ or 1.72 times its original length. There was one level of no length-change, where the length ratio is 1. And there were 3 levels of decreasing length where the part non-rigidly decreased to $\frac{1}{1.24}$, $\frac{1}{1.48}$, and $\frac{1}{1.72}$ times its original length. Each of the seven levels were presented ten times each for a total of 70 trials.

To report perceived horizontal angle, subjects were asked to make adjustments to a schematic figure. The figure, as seen on the right side of Figure 4, was a white circle with a protruding white line part. The schematic represents a view of the 3D object from the top: the circle represents the main body while the line represents the protruding part. The diameter of the circle was $2.14°$ while the line was $2.35°$ long. Subjects adjusted the orientation of the white line in the adjustment display to match the perceived horizontal orientation of the part relative to the main body in the SFM display (see below for details).

Procedure

The experiment was run in a dark room. Subjects were seated 75 cm from an iMac monitor with a refresh rate of 100Hz and screen dimension of 800x1280 pixels. The experiment was programmed and run in Matlab using the PsychToolbox libraries (Brainard, 1997).

Instructions were presented at the start of the experiment. Subjects were asked to make a judgment about the horizontal orientation of the protruding part with respect to the main body. They were asked to adjust the schematic figure on the right of the screen to match their percept of the horizontal angle. The arrow keys were used to adjust the angle

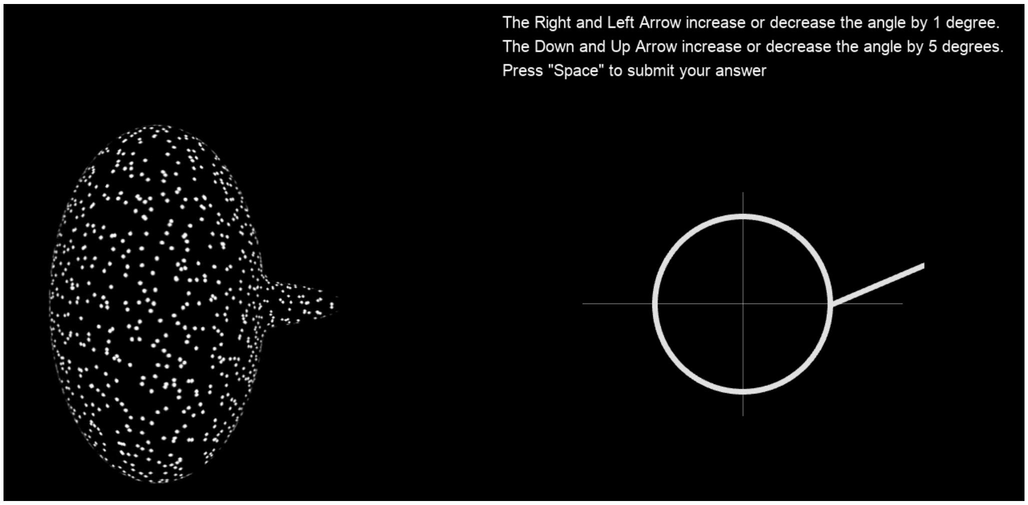

Figure 4: Subjects matched the perceived horizontal orientation of the part with respect to the main body in the SFM display on the left using the adjustment display on the right. The adjustment display is a schematic of a top view of the 3D object. Stimulus remained on display during the whole trial. Adjustment instructions remained above adjustment display.

of the line (with respect to the circle) on the schematic. It was explained to the subjects that the schematic represented a top-down (looking down on) view of the object. Adjusting the line downward meant that the part was perceived to be angled toward the observer and adjusting the line upward meant that the part was perceived to be angled away from the observer. The angle could be adjusted in increments of either $1°$ or $5°$ (using two different sets of keys). The $1°$ increments allowed subjects to make finer adjustments once their setting was in the right ballpark.

At the bottom of the instructions screen were 5 examples of two-part 3D objects with different horizontal angles between protruding part and body. The two-part objects were cylinders with protruding cylinder parts. They were not defined by a random dot texture but were given a wood grain texture with shading. Each figure was oriented 45 degrees away from the observer. From left to right the parts had a horizontal angle of $-40°$, $-20°$, $0°$, $20°$, and $40°$. Above each 3D figure was a schematic showing the "correct answer" to each of the oriented 3D objects. Part of the goal of these examples was to make clear to the subject that the relevant angle they were meant to match was the (horizontal) orientation of the part relative to the main body of the object, and not the orientation of the part in space with respect to the observer. Once subjects understood the task they could press any key to continue on to the practice trials. There were seven practice trials. On each trial one of the seven levels of length ratio were shown in a random order. Once the practice trials were over, the subjects continued on to the real trials.

On each trial the SFM stimulus was continuously displayed on the left side of the screen while subjects made their angle adjustments to the schematic on the right side of the screen. Once they felt their adjustment of the schematic matched their percept of the orientation of the protruding part with respect to the main body of the 3D object, they submitted their response by pressing the space button and moved on to the next trial. On each trial the adjusted angle response was collected. There was no time limit on each trial. 70 trials were presented in random order for each subject. The experiment took approximately 25 minutes.

2.2.2 Model Predictions

We create a simple model based on the assumption that that length change is being reinterpreted as a fixed horizontal angle with no length change. Under these assumptions the model makes predictions of perceived horizontal angle given some magnitude and direction of length change. Figure 5 illustrates our model.

The top row of the figure represents what is "actually" going on in our stimuli. The blue lines represent the part. On the left side of the figure is an schematic illustration of the length decreasing case. The part length goes from L, its original length, to $x \cdot L$ (decreasing so $x < 1$) from $45°$ to $90°$. On the right side of the diagram the part is increasing to x times its length from $45°$ to $90°$. The two schematics at the bottom of the figure show the projected lengths of the part under the new reinterpretation of the part having a fixed horizontal angle with no length change. In both schematics the length of the part (again represented by the blue line) stays fixed from $45°$ to $90°$. On the left, in the decreasing length case, the part is perceived to be angled away from the observer. We use θ to represent

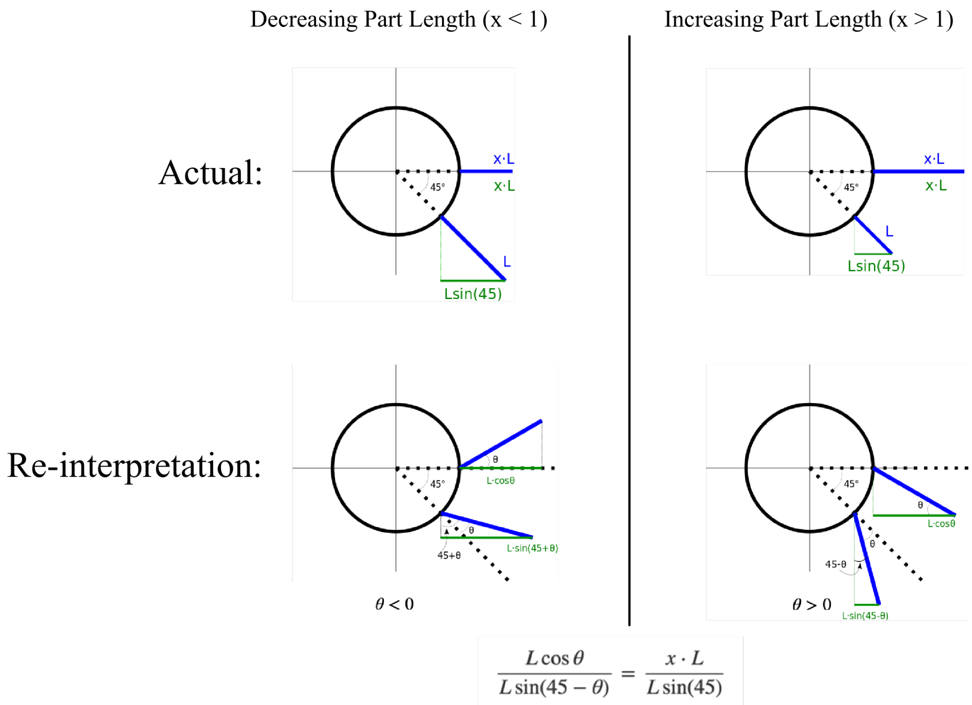

Figure 5: Top row represents the actual length change of the figure as the part goes from $45°$ to $0°$. Bottom row represents the model's reinterpretation of the length change as a part with a fixed horizontal angle and no length change. Column on the left is the decreasing case, the right column is the increasing case. Blue lines represent the part, green lines represent the part's projection on the image plane.

the perceived part angle. The green solid lines represent the part's projected lengths in the image plane. The projected length changes in the image plane for the actual length change (the two schematics in the top row) and the projected length changes in the image plane for the reinterpretation (bottom row) must be equated. The ratios between the lengths of the parts in the image plane from start to end must be the same for both "what's really happening" (top row) and for the "reinterpretation" (bottom row). So we equate the ratios and obtain the equation:

$$\frac{L \cdot cos(\theta)}{L \cdot sin(45 - \theta)} = \frac{x \cdot L}{L \cdot sin(45)} \tag{1}$$

$$\frac{cos(\theta)}{sin(45)cos(\theta) - sin(\theta)cos(45)} = \frac{2}{\sqrt{2}}x \tag{2}$$

$$\frac{cos(\theta)}{\frac{\sqrt{2}}{2}(cos(\theta) - sin(\theta))} = \frac{2}{\sqrt{2}}x \tag{3}$$

$$\frac{cos(\theta)}{cos(\theta) - sin(\theta)} = x \tag{4}$$

$$\theta = tan^{-1}(1 - \frac{1}{x}) \tag{5}$$

In the decreasing case $x = \frac{1}{1.72}, \frac{1}{1.48}$, or $\frac{1}{1.24}$. In the increasing case, $x = 1.24, 1.48$, or 1.72. In the no length change case $x = 1$. Solving for θ, we input the seven levels of length ratio used in our experiment to get our predictions of perceived part angle (see Figure 6).

2.2.3 Results and Discussion

In Figure 6 we plotted mean angle settings as a function of log length ratio. Each subplot shows the data from one observer. On the x-axis the negative values represent the decreasing length cases (the length ratios were $\frac{1}{1.24}, \frac{1}{1.48}$, and $\frac{1}{1.72}$) and the positive values represent the length increasing conditions. For the mean angle responses, the negative values mean the subjects adjusted the angle of the white line upward from horizontal (i.e. the part was seen as angled away from the observer). The positive values mean the subjects adjusted the angle of the white line downward from horizontal (i.e. seen as angled towards the observer). The dashed green line is our model prediction. The error bars are 95% confidence intervals.

Each subject perceived some "illusion" of horizontal part angle when there was none (recall that the part was always perpendicular to the main body in all cases - which corresponds to a horizontal angle of 0 on this scale). Although the veridical percept would lead to horizontal angle settings of $0°$ for each level of length ratio, we can see that subjects consistently reported perceiving a horizontal angle. Subjects also perceived the horizontal angle in the direction that was predicted. In the decreasing part length cases (negative log length ratio), subjects adjusted the angle in the adjustment display upward (negative mean adjusted angle response) meaning they perceived the part to be pointing away from them. In the increasing length conditions subjects adjusted the angle in the adjustment display downward meaning they perceived the part to be pointing toward them. This was consistent for all subjects. Given that the model has no free parameters, and thus involves no ”fitting” to the data, it does surprisingly well in predicting subject’s angle settings. Recall that the model was based on the single assumption of reinterpreting length change as a (fixed) horizontal angle with no length change. Although some deviations are apparent for extreme negative values, the overall fit of the model suggests that this single assumption goes a long way in explaining observers’ perception of part orientation in these displays.

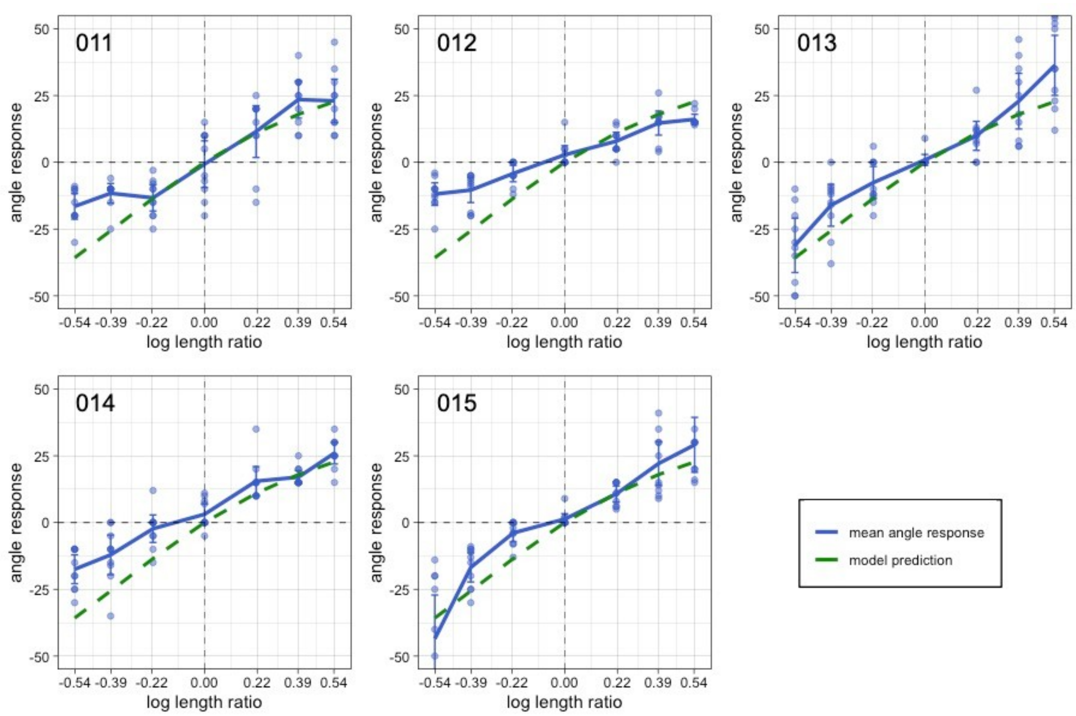

Figure 6: Each plot represents a different subject. Angle response is plotted as a function of log length ratio. Each point represents a response on a trial. Mean of the angle responses are plotted as the blue solid line while the model prediction is plotted as the dashed green line. Vertical reference line separates decreasing length ratio (negative) from increasing. Horizontal reference line separates upward angle adjustments (negative) from downward (positive).

3. General Discussion

The SFM literature has primarily focused on rigid body motion. While there are various models for how SFM might be computed given different input representations, and many different demonstrations of how well subjects can perceive 3D structure in SFM displays, these models and studies have all been done primarily with rigid objects. Most of the models and psychophysical observations either implicitly assumed rigidity or at least stayed within the boundaries of investigating the ability to perceive the structure of rigidly moving objects. While the assumption of rigidity is useful and environmentally plausible most of the time, focusing solely on rigidity leaves out the enormous set of possible non-rigid motions. As we have mentioned, many of the objects in the environment move non-rigidly, especially biological entities. Claiming environmental plausibility and ignoring non-rigidity (or certain types of non-rigidities) is incomplete. We have seen from demonstrations like those by Sinha and Poggio (1996) and more recent studies, like Jain and Zaidi (2011), that the rigidity assumption is too restrictive.

The experiments presented here are motivated by an overarching goal to begin exploring the space of non-rigid transformations in SFM. We started by asking: What types of non-rigid transformations are perceivable in an SFM context? To begin to tackle the space of non-rigid transformations we tried to look at environmental plausibility. Objects in the environment that move non-rigidly are typically biological entities, such as other humans and animals. The motions of these biological entities are usually constrained to certain types of non-rigidities which we refer to as biologically plausible motions. Previous work has shown that biological plausibility influences the perception of apparent motion (Shiffrar & Freyd, 1990) and sensitivity to shape change (Denisova et al., 2016). From this slightly different perspective on environmental plausibility we asked: are there differences in the way we perceive biologically plausible versus biologically implausible (or less plausible) non-rigid transformations in SFM? To investigate the difference we looked at orientation change vs. length change.

The findings from the two experiments presented here would suggest that there are differences in the way that we perceive biologically plausible vs. biologically implausible transformations in SFM. In both of the studies we focused on exploring our initial observation that when our two-part SFM stimulus underwent a non-rigid length change, which we consider to be less biologically plausible, it tends to be reinterpreted in a way that is more biologically plausible - that is as a part-wise rigid orientation change. In both experiments we manipulated the magnitude of length change. In the first experiment we did this to observe the range of length change ratios where the misperception of non-rigid length change as part-wise rigid orientation change persisted. In the second experiment the range of length change ratios not only illustrated a range where an illusion existed but was also produced parametric data that could be compared against the predictions of a simple model that re-interprets part length change in terms of an “illusory” part orientation.

In Experiment 1 we observed non-rigid length change being reinterpreted as a non-rigid part orientation change. We were able to show a clear range of length change values where this misperception or bias towards an interpretation of part-orientation change existed for each of our subjects. In Experiment 2, where we made the non-rigid length change fully visible in the silhouette, we observed non-rigid length change reinterpreted as a part with no length change and a fixed horizontal angle. For our second experiment we created a simple geometric model that reinterprets length change in terms of a fixed horizontal angle between the part and the body with no length change. The predictions of the model closely matched the orientation settings of subjects suggesting that their visual systems may be making similar assumptions. In both experiments we see a clear bias against a length change interpretation in favor of an interpretation involving either a changing part orientation (in Exp. 1) or a fixed but “illusory” part orientation (in Exp. 2) between the part and the main body.

The question we are asking in this study is whether there are differences in the way the visual system perceives biologically plausible vs. less biologically plausible non-rigid transformations in SFM. However, biologically plausible and less plausible transformations are each their own categories within the larger space of non-rigid transformations and we have only started by comparing two: orientation and length change. An obvious next step would be to explore other non-rigid transformations within the two subcategories of biologically plausible and less plausible transformations. Examples of other biologically plausible transformations to consider could be curvature change (like the curving of a cat's tail) while a less plausible transformation could be e.g. width change of a part. Such transformations are readily expressed as specific changes to the parameters of a skeleton-based representation of shape (Blum, 1973; Feldman & Singh, 2006). Hence skeleton-based representations of shape may provide a natural framework for modeling various perceptually-relevant non-rigid shape transformations. Further investigation of how subjects perceive orientation change in SFM may also fill in some possible limitations on how we addressed our question. While we investigated a misperception of length change as orientation change the current study never actually presented our subjects with orientation change. To further demonstrate a difference in how subjects perceive biologically plausible and less biologically plausible transformations in SFM, a more direct overt comparison could be done. However, given that length change has been misperceived as orientation change we would expect there would be large differences in performance on any task that directly compared the ability of subjects to determine the structure of objects undergoing orientation vs. length change.

By exploring the observed misperception of non-rigid length change as orientation change in our two experiments we begin to address the question of whether there are preferences or visual biases for perceiving biologically plausible non-rigid transformations. Our results suggest that there are differences and that the visual system may be biased towards interpretations of non-rigidity that are more biologically plausible.

3.1 Conclusion

Structure from motion is the fascinating ability to perceive the 3D structure of objects from 2D image motion. It has been a topic of study for a long time but has somehow largely restricted itself to the domain of rigid motion. However, it is clear that some forms of non-rigid motion are readily perceivable in SFM. Our overarching goal is to understand what types of non-rigid transformations are perceivable in SFM. The current study suggests that the distinction between biologically plausible and biologically implausible shape transformations may be an important factor in addressing this larger question.

Bibliography

Blum, H. (1973). Biological shape and visual science (part i). *Journal of theoretical Biology*, *38*(2), 205–287.

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

Braunstein, M. L., & Andersen, G. J. (1984). Shape and depth perception from parallel projections of three-dimensional motion. *Journal of Experimental Psychology: Human Perception and Performance*, *10*(6), 749–760.

Denisova, K., Feldman, J., Su, X., & Singh, M. (2016). Investigating shape representation using sensitivity to part-and axis-based transformations. *Vision research*, *126*, 347–361.

Domini, F., & Caudek, C. (2003). 3-D structure perceived from dynamic information: A new theory. *Trends in Cognitive Sciences*, *7*(10), 444–449.

Feldman, J., & Singh, M. (2006). Bayesian estimation of the shape skeleton. *Proceedings of the National Academy of Sciences*, *103*(47), 18014–18019.

Hildreth, E. C., Grzywacz, N. M., Adelson, E. H., & Inada, V. K. (1990). The perceptual buildup of three-dimensional structure from motion. *Perception & Psychophysics*, *48*(1), 19–36.

Hoffman, D. D., & Bennett, B. M. (1986). The computation of structure from fixed-axis motion: Rigid structures. *Biol. Cybern*, *54*, 71–83.

Jain, A., & Zaidi, Q. (2011). Discerning nonrigid 3D shapes from motion cues. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(4), 1663–1668.

Johansson, G. (1973). Visual perception of biological motion and a model for its analysis. *Perception & Psychophysics*, *14*(2), 201–211.

Koenderink, J., & Van Doorn, A. (1986). Depth and shape from differential perspective in the presence of bending deformations. *J. Opt. Soc. Am. A*, *3*(2), 242–249.

Koendreink, J. (1986). Optic flow. *Vision Res*, *26*(1), 161–180.

Shiffrar, M., & Freyd, J. J. (1990). Apparent motion of the human body. *Psychological science*, *1*(4), 257–264.

Sinha, P., & Poggio, T. (1996). Role of learning in three-dimensional form perception. *Nature*, *384*(6608), 460–463.

Sperling, G., Landy, M. S., Dosher, B. A., & Perkins, M. E. (1989). Kinetic Depth Effect and Identification of Shape. *Journal of Experimental Psychology: Human Perception and Performance*, *15*(4), 826–840.

Todd, J. T. (1995). *The visual perception of three-dimensional structure from motion*.

Todd, J. T., & Bressan, P. (1990). The perception of 3-dimensional affine structure from minimal apparent motion sequences. *Perception & Psychophysics*, *48*(5), 419–430.

Ullman, S. (1979). The interpretation of structure from motion. *Proceedings of the Royal Society of London. Series B, Containing papers of a Biological character. Royal Society (Great Britain)*, *203*(1153), 405–426.

Ullman, S. (1983). Maximizing rigidity: The incremental recovery of 3-d structure from rigid and nonrigid motion. *Perception*, *13*(3), 255–74.

Wallach, H., & O'Connell, D. N. (1953). The kinetic depth effect. *Journal of Experimental Psychology*, *45*(4), 205–217.