Visual and Haptic Perceptual Spaces
From Parametrically-Defined to Natural Objects

Nina Gaißert, Kirstin Ulrichs and Christian Wallraven

Max Planck Institute for Biological Cybernetics
Spemannstraße 38
Tübingen, Germany
Nina.Gaissert@tuebingen.mpg.de, Kirstin.Ulrichs@tuebingen.mpg.de, Christian.Wallraven@tuebingen.mpg.de

Abstract

In this study we show that humans form very similar perceptual spaces when they explore parametrically-defined shell-shaped objects visually or haptically. A physical object space was generated by varying three shape parameters. Sighted participants explored pictures of these objects while blindfolded participants haptically explored 3D printouts of the objects. Similarity ratings were performed and analyzed using multidimensional scaling (MDS) techniques. Visual and haptic similarity ratings highly correlate and resulted in very similar visual and haptic MDS maps providing evidence for one shared perceptual space underlying both modalities. To investigate to which degree these results are transferrable to natural objects, we performed the same visual and haptic similarity ratings and multidimensional scaling analyses using a set of natural sea shells.

Introduction

In this paper we want to shed light on how humans represent complex objects. Whereas much research in the past has been devoted to addressing this question in the visual domain, we will compare visual and haptic processing. Our particular interest in this context lies in examining multisensory perceptual spaces of objects, i.e. topological representations of object properties in the brain. Our study follows prior investigations by Cooke (2005; 2006; 2007). Using parametrically-defined objects, these studies have shown that humans can create visual, haptic, and visuo-haptic perceptual spaces of three-dimensional objects that are highly congruent to the underlying physical object space. However, these studies are based on objects that varied in two very intuitive dimensions – shape and texture. With our research we try to extend these findings to more complex and to natural objects. Combining computer graphics modeling with 3D printing techniques, we generated a set of complex objects spanning a three-dimensional object space. These objects have a shell-shaped form for several reasons: the software ShellyLib gave us full control over shape parameters, the objects are naturalistic but not too familiar to participants, and every object has several local and global features that are detectably visually as well as haptically.

To investigate how findings of the computer generated objects can be transferred to real objects, we also collected a set of natural sea shells, a set of potentially even more complex objects (see Figure 4).

In psychophysical experiments, participants explored the objects visually or haptically and rated similarities between pairs of objects. We used multidimensional scaling (MDS) techniques to analyze these similarity ratings. MDS takes distances between pairs of objects as input and returns coordinates of the objects and their relative positions in a multidimensional space. Using human similarity ratings, the MDS output map can be understood as perceptual space (Borg, 1997; Shepard, 1987). This perceptual space provides information about how many dimensions are apparent to the participants, about the weighting of these dimensions, and whether these dimensions correspond to the dimensions of the object space.

Visual and haptic sensory systems are able to extract many of the same object properties, e.g. shape and texture, although they use different types of input information: visual perception has a large spatial extent, while haptic perception is limited to near-body space. Vision is based on fast, parallel processing of two-dimensional retinal input, while touch operates with tactile receptors on 3D objects in a slow, sequential fashion.

In this paper we want to address the question to what degree humans are able to reconstruct a complex stimulus space of parametrically-defined, artificial objects visually or haptically, whether and how the perceptual spaces of the two modalities differ, and furthermore, if these findings are transferrable to real-world objects.

Parametrically-Defined Objects

Stimuli

For the experiments described here we generated a three-dimensional objects space of 21 complex, shell-shaped objects (Figure 1).

The objects were generated using the mathematical model by Fowler, Meinhardt and Prusinkiewicz (1992) and the software ShellyLib. The mathematical model is based on equation 1 and constructs a shell-like shape by shifting an ellipse along a helicon spiral to form the surface of the shell. Three parameters (A, sin β, and $\varepsilon^{\cot\alpha}$) were altered in five defined equidistant steps to construct a three-dimensional object space of 5x5x5=125 objects. To reduce the amount of stimuli to a reasonable amount for haptic experiments we chose 21 objects. These 21 objects span three orthogonal planes arranged in a Y-shaped form (see (Cutzu, 1998) for a similar approach). The center stimulus in the object space is the center stimulus of every plane.

$$r = A * \sin\beta * \varepsilon^{\cot\alpha} \qquad (1)$$

For the visual stimuli, object meshes were imported into the 3D modeling software 3D Studio Max. The object material was set to a white and non-glossy material, resembling the plastic material used by the 3D printer. The camera was positioned at a distance of 50 cm from the object with a field of view of 45 degrees. The lighting was a standard omni-illuminant of 3D Studio Max with an *intensity* multiplier of 1.1. 2D views of the objects were then rendered such that the shape features were clearly visible. The objects were rendered to 1280 x 1024 pixel 2D images on black background.

For the haptic stimuli the wall thickness of the objects was increased by 6 per cent using the shell modifier of 3D Studio Max. The surface was smoothed using two iterations of the *meshsmooth* modifier. The objects were printed using the EDEN250TM 16 micron layer 3-Dimensional Printing System of Objet, Belgium. The manufacturing process was performed in "high quality mode" with a white acrylic-based photopolymer material, resulting in a hard, white, and opaque plastic model. Each 3D object weighed about 40g. The maximum dimensions were 5 cm in depth, 10 cm in height and 15 cm in width.

Similarity Ratings

The task was to rate the similarity between pairs of objects on a scale from low similarity (1) to high similarity (7). In the visual similarity rating task, 2D pictures of the objects were presented to 10 naïve participants with normal or corrected-to-normal vision. To avoid order effects, 10 other participants (also naïve to the stimulus set) performed the haptic similarity ratings. Participants were blindfolded and explored 3D plastic models of the objects with both hands and no restrictions to the exploratory procedure. Participants were undergraduate students and were paid 8€ per hour.

The visual stimuli were presented on a Sony Trinitron 21" monitor with a resolution of 1024 x 768 pixels using the *Psychtoolbox* extension for MATLAB (Brainard, 1997; Pelli, 1997). The image size was between 9-12 times 9-12 degrees of visual angle resulting in about the same visual impression, as if a 3D object was laying on a table in front of the participant. Participants used a chin rest to align the line of sight to the centre of the screen. Participants had to fixate a fixation cross for 0.5 seconds before the first object appeared on the screen for 3 seconds. Then the screen turned black for 0.5 seconds before the second object was presented for 3 seconds. After seeing both objects, participants had to rate the similarity between these two objects by choosing a number between 1 (fully dissimilar) and 7 (fully similar).

In the haptic similarity rating task, blindfolded participants were seated in front of a table with a sound-absorbing surface. The first object was placed on the table and participants were instructed to start the object exploration. After 8 seconds, participants were instructed to put the object back on the table. The object was replaced by the second object and again participants had 8 seconds to explore the second object. After putting the second object back on the table, the experimenter recorded the rating, which was given verbally.

Before the experiment itself started, participants performed some test trials in which pairs of objects were

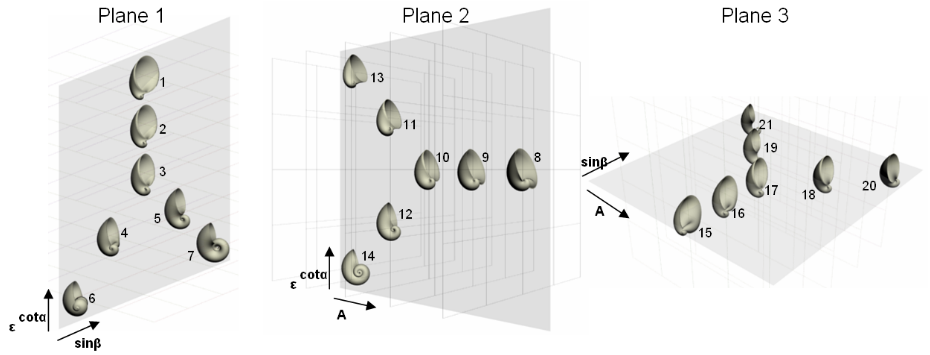

Figure 1 Parametrically-defined Objects Three-dimensional object space, spanned by 21 computer generated objects. The objects are arranged in Y-shaped forms within the three orthogonal planes.

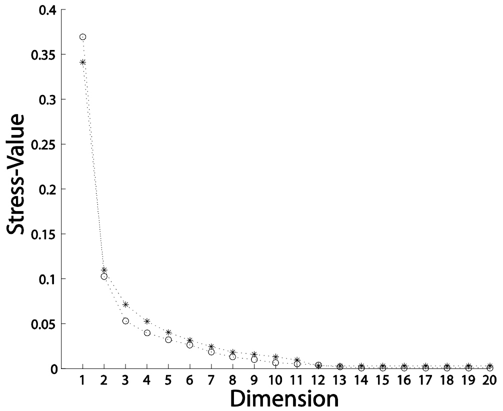

Figure 2 Stress-Values The stress-values for one to twenty dimensions were plotted for the visual modality (stars) and the haptic modality (circles). In both cases the elbow shows that three dimensions were apparent to the participants.

shown to make the participants familiar with the range of objects and to become accustomed to the task. For the ratings, every object was paired once with itself and once with every other object resulting in 231 trials. These trials were shown in randomized order in one block. Participants had to perform three blocks and were allowed to have a break after every block. Due to the length of the experiment, the haptic similarity ratings were split on two consecutive days and both sessions were started with the same test trials.

Multidimensional Scaling

Participants' similarity ratings, ranging from 1 to 7, were converted to dissimilarities which were then averaged for both modalities over all participants and all trials. The correlation between average dissimilarity matrices of both modalities was calculated.

For the multidimensional scaling (MDS) analysis we used the non-metric MDS algorithm (MDSCALE) in MATLAB. Non-metric MDS takes the rank-order of the pairwise proximity values into account and thus fits the human similarity data better than classical, metric MDS (Cooke, 2007). To determine how many dimensions were necessary to explain the data, the stress-value from one to twenty dimensions was plotted. An "elbow" in the plot indicates how many dimensions are sufficient to explain the data (details concerning the dimensionality can be found in (Gaißert, 2009)).

The stress values indicated that three dimensions were apparent to the participants and thus we plotted the visual and haptic perceptual spaces for three dimensions. The goodness-of-fit-criterion or sum of squared errors between the perceptual spaces of both modalities was calculated using the *procrustes* function of MATLAB.

Results

In the visual and in the haptic modality 10 participants performed similarity ratings. These ratings were averaged across participants. The correlation between the visual and the haptic similarity ratings is very high (r=0.929, p=0.00) and shows that humans perceive similarities visually and haptically in a very similar fashion.

Using the average matrices of both modalities, we performed MDS for one to twenty dimensions and plotted the stress-vlaues. The elbows in both plots show that the data can be explained by three dimensions sufficiently (Figure 2). Following this result we plotted the MDS for three dimensions for both modalities (Figure 3). The visual as well as the haptic perceptual spaces show a high congruency to the underlying physical object space (goodness-of-fit-criterion d=0.186, d=0.147 respectively, d=0 would indicate perfect alignment). Moreover, the visual and the haptic perceptual spaces are even more similar as the criterion shows (d=0.088). This finding provides evidence that one perceptual space is formed that is accessible by the visual and the haptic modality.

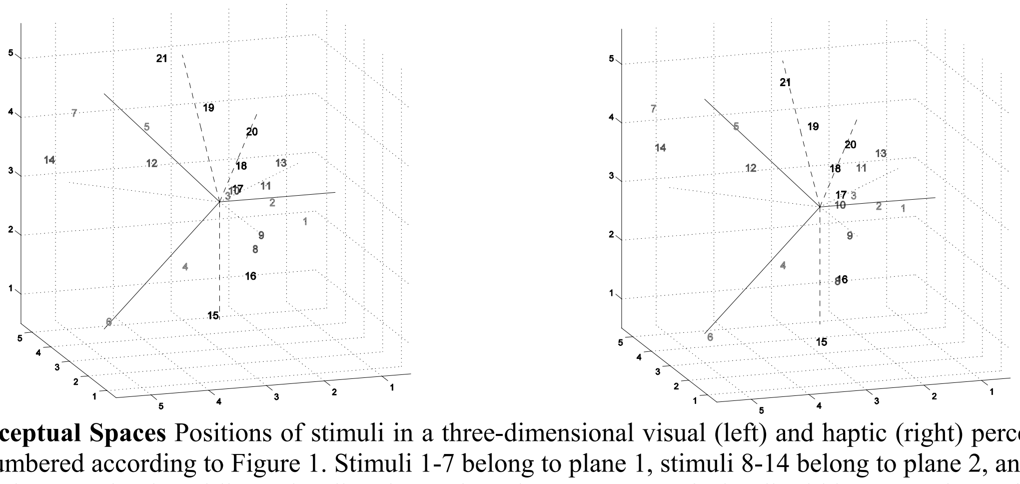

Figure 3 Perceptual Spaces Positions of stimuli in a three-dimensional visual (left) and haptic (right) perceptual space. Stimuli are numbered according to Figure 1. Stimuli 1-7 belong to plane 1, stimuli 8-14 belong to plane 2, and stimuli 15-21 belong to plane 3. The dotted lines visualize the Y-shaped arrangement of stimuli within every plane of the physical object space. The three planes and thus the three Y-shapes are orthogonal to each other.

Our findings are rather surprising taking into account the highly unintuitive shape-parameters we altered to generate the stimulus space and also considering the fact that the participants had never explored the objects haptically before and presumably also had little experience in touching shell-shaped objects at all. Nevertheless, the haptic modality not only reconstructed the underlying stimulus dimensionality, but also was able to faithfully represent the topology of the stimuli in feature space. This result demonstrates the astonishing capabilities of the haptic modality in shape processing which are on par with those of the visual modality (Cutzu, 1998; Edelman, 1999).

Natural Sea Shells

Stimuli

The visual and the haptic perceptual spaces of shell-shaped, parametrically-defined objects are highly congruent. To see how this finding is transferrable to natural objects we gathered a set of 24 natural sea shells (Figure 4) and performed the same tasks visually and haptically as with the computer generated objects.

We chose four bivalve molluscs: *Mactra stultorum*, *Pecten maximus*, *Acanthocardia tuberculata*, and *Glycymeris insubrica*. All other objects are gastropods. Four objects have a conical shell: *Patella barbara*, *Patella longicosta*, *Patella granularis*, and *Patella vulgata*. Four shells have a turban like shell: *Turbo argyrostomus*, *Turbo coronatus*, *Turbo crassus*, and *Turbo setosus*. Four objects are extremely smooth and shiny: *Cypraea eglantine*, *Cypraea histrio*, *Cypraea lynx*, and *Ovula ovum*. Four members of the olive shells were selected: *Oliva irisans*, *Oliva miniacea*, *Agaronia gibbosa*, and *Olivancillaria vesica auricularia*. Four objects have a cone like shell: *Conus figulinus*, *Conus malacanus*, *Conus marmoreus* and *Conus textile*. Every group of four is a group of objects belonging to the same superfamily.

Similarity Ratings

The task was to rate the similarity of pairs of objects on a 7 point scale from low similarity (1) to high similarity (7). Twelve participants with normal or corrected-to-normal vision performed the visual similarity ratings. Twelve other participants were blindfolded and performed the haptic similarity ratings, palpating the objects with both hands. All participants were naïve to the stimuli and were paid 8€ per hour.

The experiment was started by introducing the objects to the participants. Every object was presented to the participants in a randomized order. In the visual modality one object was placed on a black plateau, a black curtain was automatically opened, and the participant was able to explore the object visually for 12 seconds before the curtain closed automatically. During this time the object was rotated by the experimenter to show all sides of the object. This was done as the natural objects were richer in both shape and textural features than the computer generated models. For haptic exploration the object was placed on the same plateau. A beep gave the signal to start the haptic, unrestricted exploration. 15 seconds later a second beep signalized the end of the exploration. Again, more time was given in both conditions to allow observers to sample all potentially informative stimulus properties.

In the experimental trials every object was paired once with itself and once with every other object. The pairs of objects were shown in randomized order. In contrast to the experiments using computer generated objects, here every participant had to rate every pair just once instead of three times because the previous experiments showed that the judgments did not vary over repetitions. The objects were placed on the plateau. In the visual modality the curtain was opened for 6 seconds. The object was rotated by the experimenter, who also recorded the rating of the participant. In the haptic modality, beeps signalized the beginning and the end of the exploration which lasted for 8 seconds. The exploration times were kept similar to the previous experiment to facilitate comparison.

Figure 4 Natural Sea Shells Gastropods: 1: *Patella barbara*, 2: *Patella longicosta*, 3: *Patella granularis*, 4: *Patella vulgata*. 5: *Turbo argyrostomus*, 6: *Turbo coronatus*, 7: *Turbo crassus*, 8: *Turbo setosus*. 9: *Cypraea eglantine*, 10: *Cypraea histrio, 11: Cypraea lynx*, 12: *Ovula ovum*. 13: *Oliva irisans*, 14: *Oliva miniacea*, 15: *Agaronia gibbosa*, 16: *Olivancillaria vesica auricularia*. 17: *Conus figulinus*, 18: *Conus malacanus*, 19: *Conus marmoreus*, 20: *Conus textile*. Bivalves: 21: *Mactra stultorum*, 22: *Pecten maximus*, 23: *Acanthocardia tuberculata*, 24: *Glycymeris insubrica*.

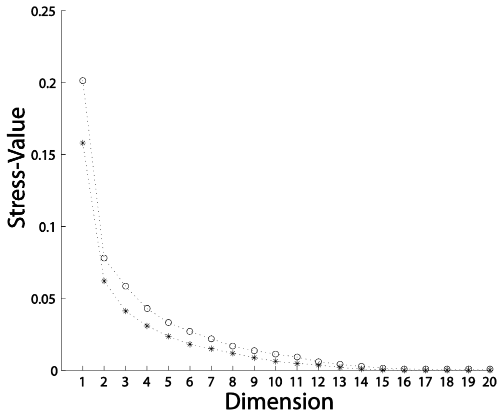

Figure 5 Stress-Values The stress-values for one to twenty dimensions were plotted for the visual modality (stars) and the haptic modality (circles). In both cases the elbow shows that three dimensions were apparent to the participants.

Multidimensional Scaling

We analyzed the similarity ratings performed on the natural sea shells as described previously for the computer generated objects. The similarity ratings were converted to dissimilarities and averaged across participants. The correlations between both modalities were calculated. For the MDS analysis, the stress-values were plotted and the MDS output map was visualized for three dimensions. The goodness-of-fit-criterion between the perceptual spaces of both modalities was also determined.

Results

Two groups of twelve participants rated the similarity between pairs of objects in a visual and haptic condition, respectively. These ratings were averaged across participants and correlated. The correlation is even higher for the natural stimuli compared to the computer generated objects (r=0.968, p=0.00), again showing how similar visual and haptic shape perception was in these experiments.

The stress-values indicate that visually and haptically participants mostly relied on three dimensions (Figure 5). Other dimensions are apparent to the participants as well but play a minor role in judging similarities. We will have a closer look into these details in further analyses. Based on this result we plotted the MDS for three dimensions (Figure 6) and calculated the goodness-of-fit-criterion for the two output maps. In line with the high correlations, the two perceptual spaces were found to be even more similar for the natural objects than for the computer generated objects (goodness-of-fit of only d=0.072).

As can be seen in Figure 6, the perceptual spaces of both visual and haptic exploration not only are highly congruent, but also exhibit a very consistent clustering of the different stimulus groups. Objects 1-4 and 21-24 result in two distinct groups but are direct neighbors in the perceptual spaces although objects 1-4 are gastropods while objects 21-24 are bivalves. The proximity within the perceptual spaces can be explained by the fact that all of these shells are not convoluted while all other shells have a distinct convolution. Objects 5-8 form their own cluster in the perceptual spaces while objects 9-20 form a large group within the visual and the haptic perceptual spaces. The feature most likely to explain this pattern is the form of the aperture. Objects 5-8 have a circular aperture while the aperture of objects 9-20 resembles a groove. Further analysis will compare the shape features to the dimensions of the perceptual spaces in more detail as well as compare the clusters within the perceptual spaces to genetic relations between the sea shells.

Object Feature Validation

MDS provides information about how many dimensions are apparent to participants, but more importantly provides a weighting of these dimensions. Exploring the computer generated shells, participants perceived three dimensions. By correlating the dimensions of the perceptual spaces of the visual and the haptic modality to the dimensions of the physical object space, it will be possible to determine the saliency of the three shape parameters. Cooke (2007) found that shape dominated texture when objects were explored visually while texture dominated shape when objects were explored haptically for objects that varied along the two dimensions shape and texture and concluded that the two modalities complement each other. Further analysis of our data will show if visual and haptic object exploration has a complementary character for our stimulus set as well.

Analyzing the MDS maps of the real shells experiments will also show if visual and haptic object exploration weighted the dimensions equally or not. However, beyond the already highlighted clustering, it will be challenging to correlate the dimensions of the perceptual spaces to clearly defined object features as there is no physical object space with which we can correlate the perceptual spaces. A further, more detailed analysis taking into account also the questionnaires of the participants will show which shape features were salient.

Summary and Outlook

Participants performed similarity ratings on a set of computer generated, shell-shaped objects visually and haptically. In both modalities participants were able to reconstruct the complex structure of the three-dimensional object space although the shape-parameters we altered were very unintuitive. Furthermore the visual and the haptic perceptual spaces were very similar to each other providing evidence that one perceptual space is formed that is accessible to both modalities.

With the set of natural sea shells we were able to extend these findings to natural objects. Again, the results show that humans perceive similarity in the same fashion in both modalities.

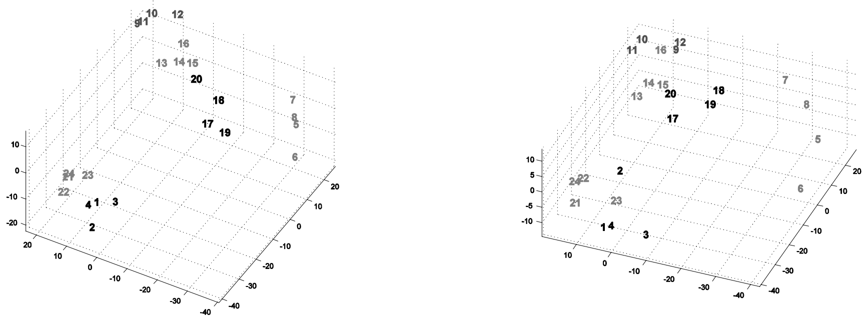

Figure 6 Perceptual Spaces Positions of stimuli in a three-dimensional visual (left) and haptic (right) perceptual space. Objects are numbered according to Figure 4. Shells within one column of Figure 4 are closely related and are marked with the same shade.

The results show highly congruent perceptual spaces in both modalities. Whether it is one underlying space or just congruent, but separate spaces, needs to be determined in future experiments. Participants will learn objects visually and haptically and afterwards recognize them using the same or the other modality. One underlying space should show no effect of modality change while congruent spaces should show an effect. Furthermore, fMRI experiments could show if visual and haptic exploration of these objects activates the same or different brain areas.

Further analysis of the presented results is necessary as well. The perceptual spaces of the natural objects clearly showed that participants did not focus on color in the visual domain and not on material properties in the haptic modalities (also shown by questionnaires). Why did participants clearly focus on shape to judge similarities? One reason might be that color is not a very diagnostic feature for shells in general (e.g. an algae cover could easily change the color and the reflectance of a sea shell). Similarly, relying on fine-grained texture of the object may also be unhelpful because water and sand can smoothen the surface of the sea shells. But both features may become more salient if more objects of only one superfamily are explored. This will be investigated in upcoming studies.

We suggest that humans rely on shape when judging similarities of objects because shape is determined by evolution and by physical parameters. Therefore we are very interested in seeing if cluster analyses of the similarity ratings can predict family resemblance of the objects and is correlated to genetic variation. Upcoming analyses will show if this hypothesis can hold.

Acknowledgements

All sea shells are items of loan from the natural history museum Stuttgart, Germany (Am Löwentor, Staatliches Museum für Naturkunde Stuttgart). We want to thank Hans-Jörg Niederhöfer for providing the sea shells and helping to select an adequate set of stimuli.

References

Borg, I. and Groenen P. 1997. *Modern Multidimensional Scaling*. Springer.

Brainard, D.H. 1997, The Psychophysics Toolbox. *Spat Vis*. 10(4): 433-6.

Cooke, T., Wallraven, C., and Bülthoff, H.H. 2005. A Comparison of Visual and Haptic Object Representations Based on Similarity. *9th International Conference on Information Visualisation* (IV'05): 33-40.

Cooke, T., Kannengiesser, S., Wallraven, C., Heinrich, H.H. 2006, Object feature validation using visual and haptic similarity ratings. *ACM Transactions on Applied Perception* 3(3): 239-261.

Cooke, T., Jäkel, F. Wallraven, C., Heinrich, H.H. 2007, Multimodal similarity and categorization of novel, three-dimensional objects. *Neuropsychologia* 45(3): 484-95.

Cutzu, F. and Edelman, S. 1998. Representation of object similarity in human vision: psychophysics and a computational model. *Vision Res*, 38(15-16): 2229-57.

Edelman, S., Bülthoff, H.H., and Bülthoff, I. 1999. Effects of parametric manipulation of inter-stimulus similarity on 3D object categorization. *Spatial Vision* 12(1): 107-123

Fowler, D.R., Meinhardt, H., and Prusinkiewicz, P. 1992, Modeling seashells. *ACM transactions on computer graphics* 26(2): 379-387.

Gaißert, N., Wallraven, C., and Bülthoff, H.H. 2009. Visual and Haptic Perceptual Spaces Show High Similarity in Humans. Forthcoming.

Pelli, D.G. 1997. The VideoToolbox software for visual psychophysics: transforming numbers into movies. *Spat Vis* 10(4): 437-42.

Shepard, R.N. 1987. Toward a universal law of generalization for psychological science. *Science* 237(4820): 1317-23.