

# Perceptual Representations of Parametrically-Defined and Natural Objects

## Comparing Vision and Haptics

Nina Gaißert<sup>1</sup>

Max Planck Institute for Biological Cybernetics

Christian Wallraven<sup>2</sup>

Max Planck Institute for Biological Cybernetics

University of Korea

### ABSTRACT

Studies concerning how the brain might represent objects by means of a perceptual space have primarily focused on the visual domain. Here we want to show that the haptic modality can equally well recover the underlying structure of a physical object space, forming a perceptual space that is highly congruent to the visual perceptual space. By varying three shape parameters a physical shape space of shell-like objects was generated. Sighted participants explored pictures of the 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 correlated and resulted in very similar visual and haptic MDS maps. To investigate to which degree these results are transferable to natural objects, we performed the same visual and haptic similarity ratings and multidimensional scaling analyses using a set of natural sea shells. Again, we found very similar perceptual spaces in the haptic and visual domain. Our results suggest that the haptic modality is capable of surprisingly acute processing of complex shape.

**KEYWORDS:** Perceptual space, multidimensional scaling, vision, haptics, computer generated objects, natural objects

**INDEX TERMS:** I.4.7 [Image Processing and Computer Vision]: Feature Measurement-Feature Representation; I.4.10 [Image Processing and Computer Vision]: Image Representation-Multidimensional; I.5.3 [Pattern Recognition]: Clustering-Similarity Measures

### 1 INTRODUCTION

Early on, infants start to actively explore their surroundings. They reach for objects and try to grasp them. Grasping an object makes it possible to turn the object and view it under different angles, but it also provides tactile information about texture, weight, material, size and shape of an object. Humans are therefore able to form a combined visual and haptic representation of objects as soon as active exploration becomes possible. Hence, it is not astonishing that haptic recognition of real, familiar objects is both fast and accurate [9]. Here, we will report that even uncommon tasks such as judging similarities between objects and reconstructing a perceptual similarity space from complex shape are performed astonishingly well by the haptic modality.

Combining computer graphics modeling with 3D printing techniques, we generated a set of complex shell-shaped objects, where every object has several local and global features that are detectably visually as well as haptically. Using the software ShellyLib (© Randolph Schultz, [www.shelly.de](http://www.shelly.de)), it is possible to generate shell-shaped objects and to change the shape of these objects in well-defined steps. Altering three shape parameters we constructed a three-dimensional object space of new but naturalistic objects.

In psychophysical experiments, participants explored these 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 such similarity ratings, the MDS output map can be understood as perceptual space, i.e. topological representations of object properties in the brain [2, 15, 16]. 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 in principle 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. Here, we wanted to know if the visual and the haptic modality extract the same number of shape dimensions, if these dimensions correspond to each other and if so, whether the two modalities weight these dimensions equally or whether we can find shape features that are more salient to one modality than to the other.

To investigate how findings of the computer generated objects can be transferred to real objects, we collected a set of natural sea shells, a set of even more complex objects. While the computer generated objects gave us the opportunity to compare the perceptual spaces of both modalities to the defined physical object space, the natural sea shells make it possible to compare clusters within the perceptual space to family resemblance of the sea shells, i.e. to the biologically-defined taxonomy of the sea shells. Using this analysis we can test if Shepard's hypothesis [17] that objects of the same basic kind (sharing important features like shape or even more complex: being edible) generally form local regions in perceptual space, can hold on a small scale.

In this paper we want to address the question to what extent the haptic modality is able to recover the structure of a physical object space of computer generated objects and how this haptic perceptual space differs from the visual space. Subsequently we compare the weighting of the shape features between the two modalities. Furthermore we compare the findings of the computer generated objects with natural objects.

---

Max Planck Institute for Biological Cybernetics  
Spemannstraße 38, Tübingen, Germany

<sup>1</sup>e-mail: [Nina.Gaissert@tuebingen.mpg.de](mailto:Nina.Gaissert@tuebingen.mpg.de),

<sup>2</sup>e-mail: [Christian.Wallraven@tuebingen.mpg.de](mailto:Christian.Wallraven@tuebingen.mpg.de)

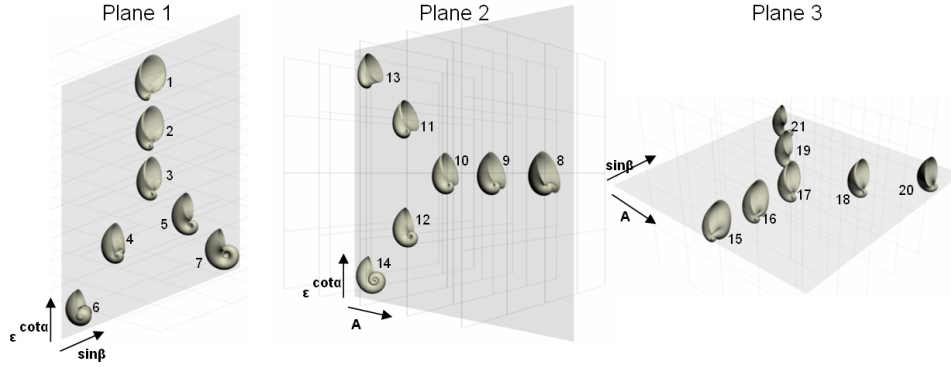


Figure 1. Parametrically-defined Objects: A three-dimensional object space was generated. The space was split in three orthogonal planes and seven objects per plane were selected. These seven objects span a Y-shaped form.

## 2 PARAMETRICALLY-DEFINED OBJECTS

### 2.1 Stimuli

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

The objects were generated using the mathematical model by Fowler, Meinhardt and Prusinkiewicz [8] 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 ( $A$ ,  $\sin \beta$ , and  $\epsilon^{\cot \alpha}$ ) were altered in five defined equidistant steps to construct a three-dimensional object space of  $5 \times 5 \times 5 = 125$  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 [5] for a similar approach). The center stimulus in the object space is the center stimulus of every plane.

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

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 of 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 percent 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.

### 2.2 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. In the haptic task, 10 blindfolded participants (also naïve to the stimulus set), explored plastic models of the objects with both hands and no restriction 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 [3, 14]. 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 would lie on a table at arm’s length distance 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 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 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 into two sessions on consecutive days with both sessions started with the same test trials. After performing the similarity ratings, participants had to answer a questionnaire.

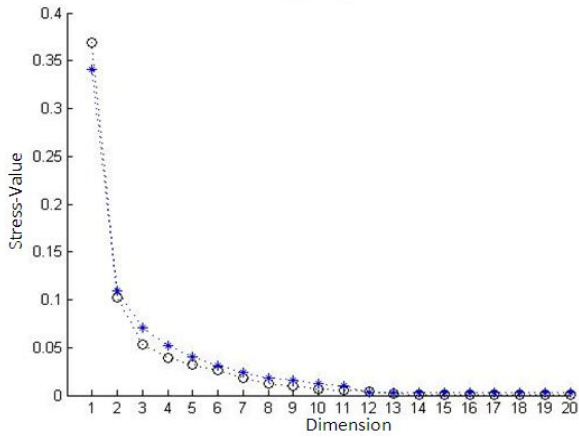


Figure 2. Stress-Values: The stress-values for one to twenty dimensions were plotted for the visual modality (blue stars) and the haptic modality (black circles). In both cases the elbow shows that three dimensions were apparent to the participants.

### 2.3 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 (MDSSCALE) 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 [4]. To determine how many dimensions were necessary to explain the data, the stress-values from one to twenty dimensions were plotted. An "elbow" in the plot indicates how many dimensions are sufficient to explain the data.

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. A goodness-of-fit-

criterion (the sum of squared errors) between the perceptual spaces of both modalities and the physical object space, as well as between the two perceptual spaces was calculated using the *procrustes* function of MATLAB.

Using the coordinates of all stimuli within the three-dimensional perceptual spaces, we correlated the stimulus positions along the three dimensions of the perceptual spaces with the stimulus positions along the three dimensions of the physical space, to determine the saliency of the different shape parameters.

### 2.4 Results

In the visual modality 10 participants performed similarity ratings. Ten other participants performed haptic 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-values. 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 ( $d=0.088$ ).

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 [5, 6].

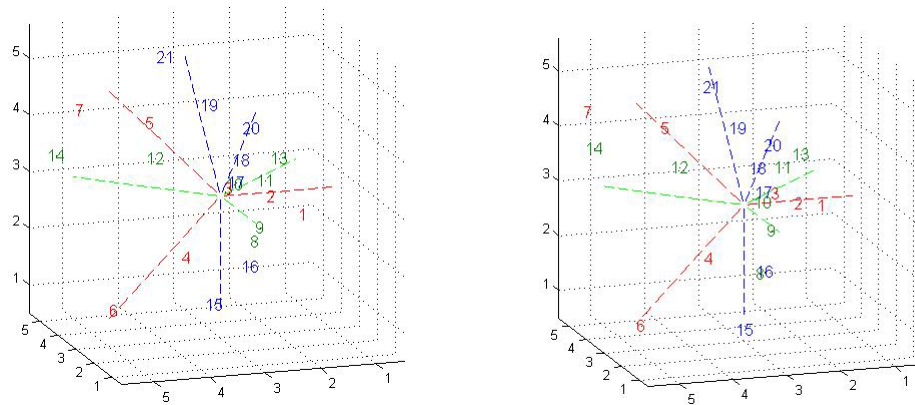


Figure 3. Perceptual Spaces: Positions of stimuli in a three-dimensional visual (left) and haptic (right) perceptual space. Stimuli numbered according to Figure 1. Red stimuli belong to plane 1, green stimuli belong to plane 2, and blue stimuli belong to plane 3. The dotted lines visualize the Y-shaped 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.



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: *Macra stultorum*. 22: *Pecten maximus*. 23: *Acanthocardia tuberculata*. 24: *Glycymeris insubrica*.

To determine the saliency of the shape parameters used for object generation, we correlated the dimensions of the perceptual spaces to the dimensions of the physical object space. In both modalities the third dimension correlates best to the parameter A. Visually  $\sin \beta$  correlates to dimension one and  $\epsilon^{\text{cota}}$  to dimension two while haptically  $\epsilon^{\text{cota}}$  defines dimension one and  $\sin \beta$  is dimension two.

By varying the shape parameters along one dimension and observing the resulting shape changes, as well as evaluating the questionnaires filled in by the participants, it was possible to verbalize the shape parameters (note that this was only possible by understanding the process of object generation, which was not transparent for the participants).  $\sin \beta$  corresponds to the symmetry of the objects, while  $\epsilon^{\text{cota}}$  is the number of convolutions, mostly referred to as “bulkiness” in the questionnaires. Modifying the parameter A changes the distance between aperture and tip of the shell.

### 3 NATURAL SEA SHELLS

#### 3.1 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 described for the computer generated objects.

We chose four bivalve molluscs: *Macra 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.

#### 3.2 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 two different perspectives of the object were presented to show all features 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 signaled 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 signaled 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.

After performing similarity ratings participants had to answer a questionnaire where they were asked to rate how strongly they relied on special object features to perform their similarity ratings. The answers were given on a scale from zero (feature is not important at all) to six (feature is very important) (Figure 8).

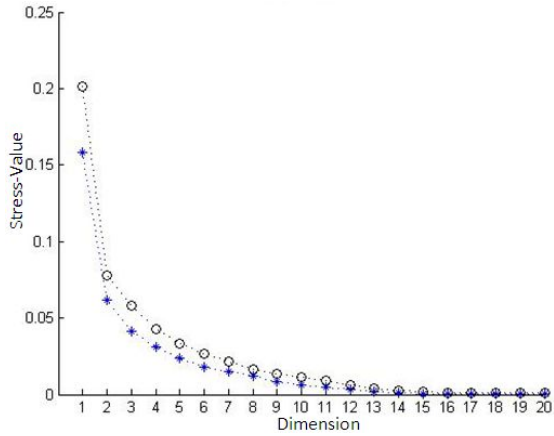


Figure 5 Stress-Values: The stress-values for one to twenty dimensions were plotted for the visual modality (blue stars) and the haptic modality (black circles).

These questionnaires were analyzed to better understand which shape features participants used for forming their perceptual spaces.

### 3.3 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. Using MDS, the stress-values for one to twenty dimensions were plotted (Figure 5). Here the elbow is not as pronounced as in the experiments on computer generated shells, thus we decided to visualize the MDS output maps for two (Figure 6) and for three dimensions (Figure 7). The goodness-of-fit-criterion between the perceptual spaces of both modalities was determined for the three-dimensional perceptual spaces.

By correlating the dimensions of the visual and the haptic perceptual spaces we analyzed if there is a difference in the saliency of the first three dimensions. We then tried to find shape features determining these dimensions by looking at the objects and how the features changes along the dimensions. In addition to that we analyzed the questionnaires filled in by the participants.

#### 3.3.1 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 two to 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 at these details in further analyses. Based on the fact that the elbow in the stress-value plot is not as pronounced as for the computer generated shells, we decided to plot the MDS output map for two

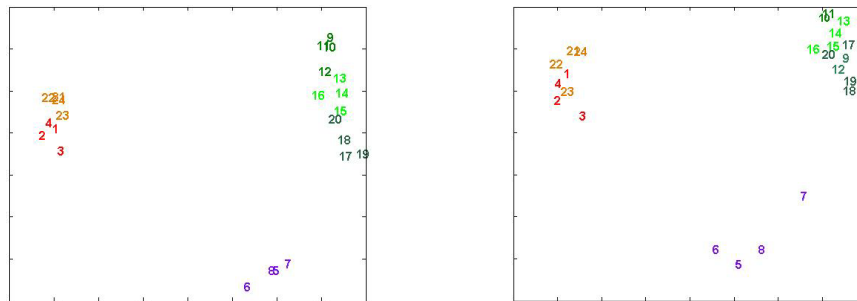


Figure 6 Perceptual Spaces: Positions of stimuli in a two-dimensional visual (left) and haptic (right) perceptual space. Objects numbered according to Figure 4. Shells within one column of Figure 4 are closely related and are marked with the same color.

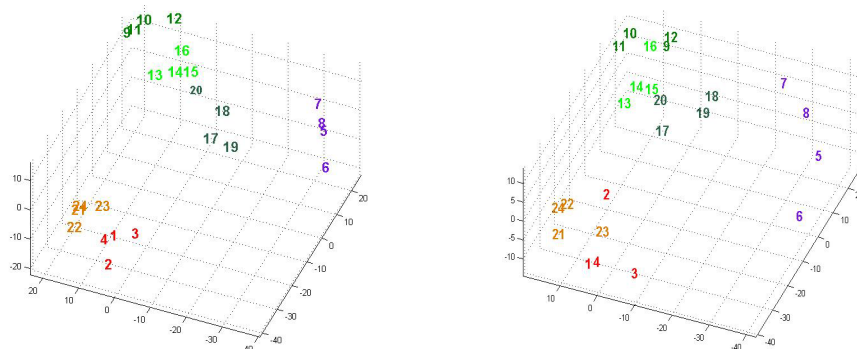


Figure 7 Perceptual Spaces: Positions of stimuli in a three-dimensional visual (left) and haptic (right) perceptual space. Objects numbered according to Figure 4. Shells within one column of Figure 4 are closely related and are marked with the same color.

and three dimensions (Figure 6 and Figure 7 respectively). As can be seen in both figures, 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. To compare the visual and the haptic modality, we calculated the goodness-of-fit-criterion for the two three-dimensional output maps. In line with the high correlations of the similarity ratings, 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  $d=0.072$ ). This result again highlights the astonishing fact that that our haptic modality is able to precisely recover the same perceptual space as the visual modality although humans have so little experience in touching shell-shaped objects.

Taking a closer look at Figure 6 shows that visually, as well as haptically, participants form three clusters of object shapes. The first cluster contains objects 1-4 and 21-24. Although objects 1-4 are gastropods while objects 21-24 are bivalves the proximity within the perceptual space 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 cluster 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. This feature is closely related to the tip of the shell. All shells with a circular aperture have a pronounced tip while the tip is less pronounced for the shells 13-20, having a groove-like aperture. Objects 9-12 do not even have a tip but have a very pronounced groove as aperture.

While in the visual perceptual space the three clusters are easily divisible into the groups of four stimuli represented as columns in Figure 4 and marked within the same color in Figure 6, these objects are more intermingled in the haptic two-dimensional perceptual space. Plotting the haptic perceptual space using three dimensions (Figure 7) makes it possible to divide the three clusters into six groups of objects as already possible for the two-dimensional visual perceptual space (Note that Figure 7 may lead to the impression that object 2 is more closely related to objects

21-24 than to 1, 4, and 3. This is an effect of the perspective of the perceptual space, however. In fact from a different perspective object 2 is easily recognizable as a member of group 1-4. The same accounts for object 16.). The stress-values for two dimensions of the visual perceptual space are roughly the same as for the three-dimensional haptic space. Thus these two perceptual spaces can explain the same amount of noise contained within similarity ratings. These two results show that visually as well as haptically participants need two dimensions to form three distinct clusters - to distinguish all six groups of shells from each other, however, participants need three dimensions haptically while visually this is already possible with two dimensions.

As a next step we tried to relate the dimensions of the perceptual spaces to object features. Visually as well as haptically the first dimension divides shells with convolutions from shells without convolutions (flat shells). The second dimension, again visually as well as haptically, corresponds to the tip of the shell, which is closely related to the form of the aperture. Shells with a pronounced tip have a round aperture and shells with a less pronounced tip have a groove-like aperture. So far it was not possible to correlate more dimensions to object features.

The saliency of the shape features found for the natural sea shells correlates very well with the findings of the computer generated shells. In both experiments the dimension concerning the convolutions is more important than the dimension concerning aperture or tip - both visually as well as haptically. It seems that the findings of our computer generated objects about shape processing are therefore largely transferrable to natural objects.

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 but rather analyzed *the shape* of objects in both modalities. This is confirmed by the questionnaires participants had to fill in after the experiments (Figure 8). Participants rated shape as significantly more important than size, patterning, color, texture, material, and weight, while they did not show a difference between the different shape features: convolutions, aperture, and tip.

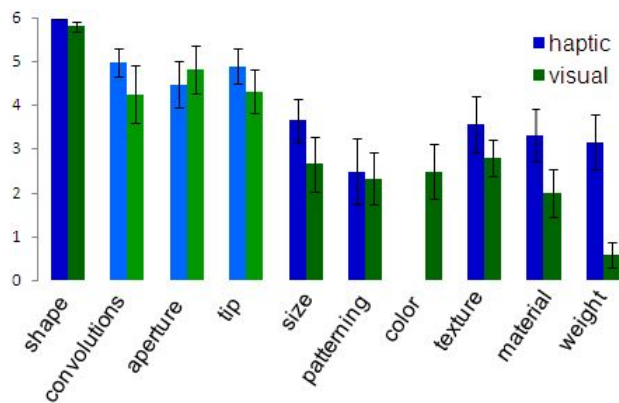


Figure 8. Questionnaires: Participants were asked to rate the importance of the following object features for performing similarity ratings: shape, size, patterning, color, texture, material, and weight (dark colors). Since we expected shape to play a major role in the similarity judgments, we asked for more details concerning shell shape: convolutions, aperture, tip (bright colors). Bars represent mean ratings across participants (0=not important, 6=very important). Error bars represent SEM.



Figure 9. Dendrograms: Left: Using participants' similarity ratings we performed a cluster analysis resulting in the same dendrogram visually as well as haptically. The six groups, visualized in Figure 4 as columns, are recovered almost correctly besides the objects marked by grey squares. Right: phylogenetic tree visualized in a form to allow for easy comparison to the visual and haptic performance.

### 3.4 Cluster Analysis

Using the similarity ratings of the visual and haptic exploration averaged across participants we performed a cluster analysis using the *linkage* algorithm of MATLAB. These clusters were visualized as dendrograms in Figure 9. Additionally we plotted the phylogenetic relation between the sea shells in a form that allows easy comparison to the dendrograms recovered using MATLAB.

#### 3.4.1 Results

Due to the fact that visual and haptic similarity ratings are highly correlated, we recovered identical dendrograms for the visual and the haptic performance, which are visualized in Figure 9 (left). In addition to this we plotted the phylogenetic tree in a way to allow comparison of the visual and the haptic performance with the taxonomy of the shells to better understand if human do recover the family resemblance of the shells and thus use the biology of the objects for object categorization. Although there is some agreement between the phylogenetic tree and the visual and haptic dendrograms it seems clear, that perceptual similarity ratings were not able to fully recover the structure of the phylogenetic tree. Especially groups 1-4 and 21-24 were put in closely related clusters, whereas biologically group 1-4 contain gastropods which are more closely related to all other objects *except* to objects 21-24 which are bivalves. This finding is also clearly visible in the perceptual spaces where these eight objects are in direct proximity to each other. Objects 5-8 form their own group in both perceptual spaces but the cluster in Figure 9 is assigned to be closer related to objects 9-20, which is taxonomically correct. Although the exact family resemblance of objects 9-20 is not correct, the cluster analysis correctly identified that these twelve objects are closely related and distinct from all other objects.

## 4 SUMMARY AND OUTLOOK

Participants performed similarity ratings on a set of computer generated, shell-shaped objects visually and haptically. The haptic modality recovered the structure of the underlying physical object space astonishingly well, even slightly better than the visual modality. Overall we found a high correlation between visual and

haptic similarity perception and between the visual and the haptic perceptual spaces.

To investigate the extent to which these findings generalize to natural objects, we collected a set of natural sea shells. Again, participants performed similarity ratings which were analyzed using multidimensional scaling techniques. The similarity ratings as well as the perceptual spaces showed a high correlation for the visual and the haptic modality and result in identical dendrograms in our performed cluster analysis.

The good performance of the haptic modality is especially astonishing when compared to haptic performance in recognition of 2D raised-line depictions, where touch alone performs quite poor [11, 13]. We assume that this difference is mostly due to the fact that we used naturalistic and natural 3D shapes in our experiments and during evolution the haptic modality evolved to explicitly perceive 3D shapes. In addition to that, participants were allowed to palpate the objects in a very natural way, with both hands and no restrictions to the exploratory procedure.

When exploring computer generated shells, the topology of the perceptual spaces was very similar, but the weighting of the stimulus features was dependent on the modality as described earlier [10]. The saliency of the first two dimensions was reversed between the two modalities (visually first dimension: symmetry, second dimension: convolutions, haptically first dimension: convolutions, second dimension: symmetry). The third dimension (aperture and tip) was the same for both modalities and thus the fact that the dimension concerning the convolutions was more salient than the dimension concerning the aperture and tip was preserved. In the experiments using natural sea shells we found the same pattern of convolutions being more salient than aperture or tip, visually as well as haptically. Cooke et. al. [4] 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. As we used only shape variations of similar spatial scales, our results show that the visual and the haptic modality both collect similar, redundant information about shape. How this information is integrated, e.g. in an optimal fashion based on noise levels [7], needs to be determined in further experiments.

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 perceived as not very diagnostic for shells (e.g. an algae cover could easily change the color and the reflectance of a sea shell) or animals in general (e.g. the fur of an ermine is brown in summer and white in winter). 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. Another reason might lie in the fact that the shape similarities between the different families in our stimulus set turn out to provide a much better, implicit clustering within similarity space.

The high similarity between the visual and the haptic perceptual spaces of both computer generated and natural shell-shaped objects provide evidence that one perceptual space is formed that is accessible to both modalities. If it is one underlying space or just congruent, but separate, unimodal spaces, needs to be determined in future fMRI experiments which will show if visual and haptic exploration of these objects activates the same or different brain areas.

Shepard [17] proposes that objects of the same basic kind generally form connected local regions in perceptual space. This hypothesis is true for our natural sea-shells. Objects closely related in taxonomy form clusters in the visual and haptic perceptual spaces. In a way, of course, this is a circular finding because when Carl von Linné started his *Systema Naturae* he used visual similarity to group plants and animals and name them [12] - indeed, it is easy to see how shape similarity would group the two groups of bivalves and gastropods together, whereas from a biological standpoint they are put into different groups based on their anatomy. Visual similarities used for taxonomy, however, are more and more replaced by screening for genetic markers or correlations of genetic sequences –it will be interesting to see how far “form follows function” with these new taxonomies.

In addition, Shepard proposes that basic level categories form connected regions within the perceptual space [17]. This means that whenever two objects A and B are taken out of one basic category and morphs are generated between these objects A and B, every newly generated object has to belong to the same basic category. With our computer generated objects and the here presented perceptual spaces we can test if this hypothesis can hold as well.

Touch is an expert system in identifying everyday objects [9]. Our research showed that also difficult tasks as rating similarities, forming a perceptual space and identifying object parameters of the shell-shaped objects are performed astonishingly well by the haptic modality. These results open the way for novel application scenarios of the haptic modality such as advanced human-machine interfaces or even teaching facts to children [1].

## 5 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. Furthermore we want to thank Kirstin Ulrichs for conducting the visual and haptic similarity ratings using natural objects.

## REFERENCES

- [1] Bara, F., et al., *The visuo-haptic and haptic exploration of letters increases the kindergarten-children's understanding of the alphabetic principle*. Cognitive Development, 2004. **19**(3): p. 433-449.
- [2] Borg, I. and P. Groenen, *Modern Multidimensional Scaling*. 2005, New York: Springer.
- [3] Brainard, D.H., *The Psychophysics Toolbox*. Spat Vis, 1997. **10**(4): p. 433-6.
- [4] Cooke, T., et al., *Multimodal similarity and categorization of novel, three-dimensional objects*. Neuropsychologia, 2007. **45**(3): p. 484-95.
- [5] Cutzu, F. and S. Edelman, *Representation of object similarity in human vision: psychophysics and a computational model*. Vision Res, 1998. **38**(15-16): p. 2229-57.
- [6] Edelman, S., H.H. Bühlhoff, and I. Bühlhoff, *Effects of parametric manipulation of inter-stimulus similarity on 3D object categorization*. Spatial Vision, 1999. **12**(1): p. 107-123.
- [7] Ernst, M.O. and M.S. Banks, *Humans integrate visual and haptic information in a statistically optimal fashion*. Nature, 2002. **415**(6870): p. 429-433.
- [8] Fowler, D.R., H. Meinhardt, and P. Prusinkiewicz, *Modeling seashells*. ACM transactions on computer graphics, 1992. **26**(2): p. 379-387.
- [9] Klatzky, R.L., S.J. Lederman, and V.A. Metzger, *Identifying objects by touch: an "expert system"*. Percept Psychophys, 1985. **37**(4): p. 299-302.
- [10] Lakatos, S. and L.E. Marks, *Haptic form perception: Relative salience of local and global features*. Perception & Psychophysics, 1999. **61**(5): p. 895-908.
- [11] Lederman, S.J., et al., *Visual mediation and the haptic recognition of two-dimensional pictures of common objects*. Percept Psychophys, 1990. **47**(1): p. 54-64.
- [12] Linnaeus, C., *Systema Naturae per Regna Tria Naturae, Secundum Classes, Ordines, Genera, Species cum Carateribus, Differentiis, Synonymis, Locis*. Editio Decima, Reformata. Vol. 1 Regnum Animale. 1758, Stockholm: Laurentii Salvii.
- [13] Loomis, J.M., R.L. Klatzky, and S.J. Lederman, *Similarity of Tactual and Visual Picture-Recognition with Limited Field of View*. Perception, 1991. **20**(2): p. 167-177.
- [14] Pelli, D.G., *The VideoToolbox software for visual psychophysics: transforming numbers into movies*. Spat Vis, 1997. **10**(4): p. 437-42.
- [15] Shepard, R.N., *The analysis of proximities: Multidimensional scaling with an unknown distance function. II*. Psychometrika, 1962. **27**: p. 219-246.
- [16] Shepard, R.N., *Toward a universal law of generalization for psychological science*. Science, 1987. **237**(4820): p. 1317-23.
- [17] Shepard, R.N., *Perceptual-cognitive universals as reflections of the world*. Behav Brain Sci, 2001. **24**(4): p. 581-601; discussion 652-71.