Analyzing Perceptual Representations of Complex, Parametrically-Defined Shapes Using MDS

Nina Gaißert, Christian Wallraven, and Heinrich H. Bülthoff

MPI for Biological Cybernetics,
Spemannstraße 38, 72076 Tübingen, Germany
{nina.gaissert,christian.wallraven,heinrich.buelthoff}@tuebingen.mpg.de
http://www.kyb.mpg.de

Abstract. In this study we show that humans are able to form a perceptual space from a complex, three-dimensional shape space that is highly congruent to the physical object space no matter if the participants explore the objects visually or haptically. The physical object space consists of complex, shell-shaped objects which were generated by varying three shape parameters. In several psychophysical experiments participants explored the objects either visually or haptically and performed similarity ratings. Multidimensional scaling (MDS) analyses showed high congruency of the visual and haptic perceptual space to the physical object space. Additionally, visual and haptic exploration resulted in very similar MDS maps providing evidence for one shared perceptual space underlying both modalities.

Keywords: haptic perception, visual perception, multidimensional scaling, similarity, psychophysics.

1 Introduction

One of the core questions of cognitive neuroscience is: How does the human brain represent and recognize objects? Whereas much research in the past has been devoted to addressing this question in the visual domain, recent studies in the last two decades have started to investigate haptic and visuo-haptic processing [15]. Our particular interest in this context lies in examining multisensory "perceptual spaces" of objects, i.e., topological representations of object properties in the brain. Perceptual spaces can be used as elegant and powerful representational systems which allow for fundamental, perceptual processing such as judging the similarity of two objects by determining their spatial distance [17], or categorizing objects by clustering them according to their proximity relationships.

Our study here follows the methodology laid out in prior studies by Cooke et al. [5,4,6,3]. Using parametrically-defined objects, these studies have shown that humans can create haptic, visual, and visuo-haptic perceptual spaces of three-dimensional (3D) objects that are highly congruent to the physical stimulus space. The physical space was defined by two different parameters or dimensions

M. Ferre (Ed.): EuroHaptics 2008, LNCS 5024, pp. 265–274, 2008.

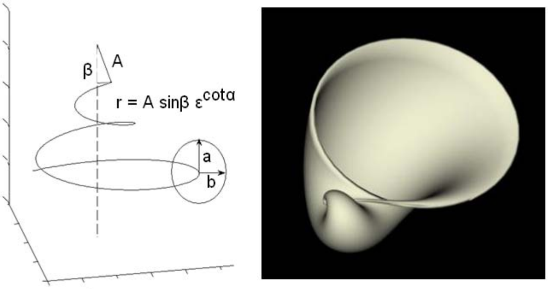

Fig. 1. Stimulus Construction Left: the mathematical, parametric shell model by Fowler, Meinhardt and Prusinkiewicz [9]. Right: center stimulus of the experimental stimulus space.

consisting of global shape and local texture. Given the importance of shape and texture for both visual and haptic modalities and that there were only those two stimulus dimensions considered, the fact that participants were able to form perceptually congruent representations was certainly an interesting finding, but it might not be too surprising. In this study, we are therefore interested in the following two questions: What happens if the stimulus space varies in more than two dimensions? And what happens if those dimensions are not as intuitive as "global shape" and "local texture"? To answer these questions, we created a complex space of shell-shaped objects which varied along three shape dimensions. We used shell-shaped objects for several reasons: they resemble natural objects, they are not too familiar to participants, and we have access to a highly realistic, biologically plausible parametric model for shell shape (and texture) [9]. With the software ShellyLib it is therefore possible to change parameters of these shell-shaped objects in well-defined steps. Combining this software with 3D printing technology gave us full control over the constructed 3D objects.

In psychophysical experiments, participants explored the objects visually or haptically and rated similarity between pairs of stimuli. To analyze these similarity ratings multidimensional scaling (MDS) techniques were used. MDS takes distances between pairs of objects in space as input and returns coordinates of the objects and their relative positions in a multidimensional space. Using human similarity ratings as input the output configuration can be interpreted as a map of objects in a psychological or perceptual space [17,1]. A wide range of multivariate stimuli was studied by psychologists using MDS techniques [11,10,17,3]. MDS provides information about the number of dimensions that are apparent to the participants and whether these dimensions correspond to the manipulated dimensions of the physical object space. Interstimulus distances in the psychological space and relative weights of the dimensions become visible in the output maps.

Visual and haptic sensory systems are able to extract many of the same object properties, e.g. global 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; additionally, vision uses a two-dimensional (2D) retinal input and objects are processed holistically, while touch operates with tactile receptors on 3D objects sequentially. Different tactile impressions have to be integrated to form the sensory percept of one object [14]. It has been shown repeatedly that humans can use both visual and haptic input for object recognition [15].

In this paper we investigated if people can identify the dimensions of the complex stimulus space and if these dimensions are stable over modalities used for object exploration. Comparing MDS output maps from visual object exploration with haptic object exploration can tell us if both modalities are able to identify the dimensions of the physical object space and if these dimensions are equally weighted across modalities. If both perceptual maps are highly similar this would provide evidence that one perceptual space underlies both sensory modalities.

2 Methods

2.1 Stimulus Space

Complex shell-shaped objects were used for the experiments described in this paper. The objects were generated using the mathematical model described in [9] (figure 1) 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 stepwise to construct a three-dimensional object space of 5x5x5 = 125 objects.

$$r = A * \sin\beta * \epsilon^{\cot\alpha} \tag{1}$$

Since for the visual and haptic experiments pairwise similarity ratings needed to be performed, the amount of stimuli had to be reduced in order to be able to conduct the experiments in a reasonable time period. We decided to use three orthogonal planes of the object space with seven objects per plane instead of twenty-five objects per plane (figure 2). These seven objects were arranged in a Y-shaped form that is easily detectable in the MDS maps (see also [8] for a similar approach). The center stimulus of the object space is the center stimulus of every plane (figure 1).

For the *visual stimuli*, object meshes were imported into the 3D modeling software 3D Studio Max. The material of the stimuli was set to a white and non-glossy material, resembling the plastic material used by the 3D printer. The camera was positioned in 50 cm distance of the object with a field of view of 45. 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 their features were clearly visible. The objects were rendered to 1280 x 1024 pixel 2D images on a 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.

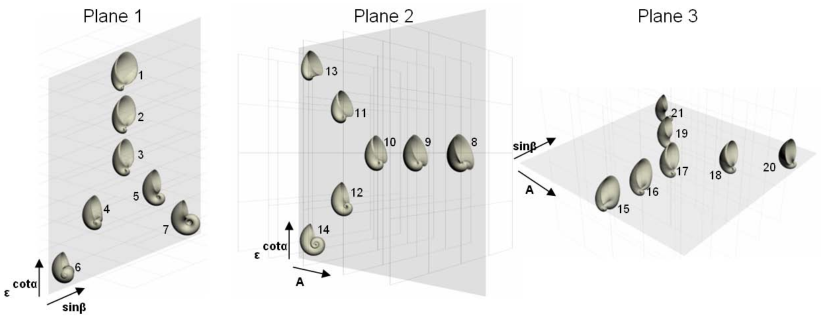

Fig. 2. Stimulus Space. The complex, three-dimensional stimulus space which was constructed from the parametric shell model. The three-dimensional parameter space was split into three two-dimensional planes, which are shown here. Following [8], every plane is defined by seven objects in a Y-shaped form.

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. The resulting 3D objects weighed about 40 g. The maximum dimensions were 5 cm in depth, 10 cm in height and 15 cm in width.

2.2 Visual Similarity Ratings

In the visual similarity rating task, 2D views of the objects were presented to 10 naïve participants with normal or corrected-to-normal vision. The participants were undergraduate students and were paid 8€ per hour. The objects were presented on a Sony Trinitron 21" monitor with a resolution of 1024 x 768 pixels using the *Psychtoolbox* extension for MATLAB [2,16]. 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 in front of a participant. Participants used a chin rest to align the line of sight to the centre of the screen.

The task was to rate the similarity between pairs of objects on a scale from low similarity (1) to high similarity (7). Before the experiment itself started, participants performed some test trials where pairs of objects where shown to make them familiar with the range of objects and to become accustomed to the task. For the experimental trials, 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 moving a bar along a slider.

In the first visual experiment (visual I) every object was compared once with itself and once with every other object of the same plane, resulting in 84 trials. These 84 object pairs were shown randomly in one block. Altogether, the experiments consisted of three blocks, where each block was randomized differently.

In the second visual experiment (visual II) the sequence of stimulus presentation was changed to see if a different sequence would influence the resulting

MDS maps. All objects from each plane were paired again once with itself and once with every object of the same plane, but this time pairs were blocked *by plane*. Again three blocks were run. In both cases it took participants about 45 minutes to perform the three blocks of similarity ratings.

2.3 Haptic Similarity Ratings

In the haptic similarity rating task, 3D objects were presented to a different group of 10 participants. The participants were undergraduate students and were paid 8€ per hour. They were blindfolded before the experiment, which started with a number of practice trials to help the participants to become accustomed to the haptic exploration task and to get familiar with the range of objects. In the experimental trials one object was put on the desk in front of the participants. Then participants were allowed to explore the object for 10 seconds with both hands and without any restrictions to the exploratory procedure. After the exploration the object was put down by the participants and exchanged with the second object. Again participants were allowed to explore the object with both hands for 10 seconds without any restrictions to the exploratory procedure. After putting the second object down, the experimenter recorded the rating, which was given verbally. For the ratings, every object was paired once with itself and once with every object of the same plane, resulting in 84 trials. The order of pairing was performed as described for the first visual experiment, meaning that pairs were shown randomly and *not* blocked by plane. It took participants 45 minutes to perform one set of ratings. The rating was followed by a 15 minutes break before the whole rating was repeated with a newly randomized order. Before the third repetition participants took a second 15 minutes break and then rated the stimuli again with a newly randomized order.

2.4 Analysis of Similarity Data

Subjects' similarity ratings ranging from 1 to 7 were converted to dissimilarities and these pairwise dissimilarities were then averaged for each plane over all subjects and over all trials. This dissimilarity data were analyzed separately for each plane of the three-dimensional object space using 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 [3].

To determine how many dimensions were necessary to explain the data, the stress-value was calculated for every plane using one to five dimensions and then plotted. An "elbow" in the plot indicates how many dimensions are sufficient to explain the data. It is also common practice to use a stress-value below 0.2 as evidence that the number of dimensions is sufficient to explain the data [7]. In order to quantify the stability of the MDS solution, we performed MDS 500 times using matrices created by randomly perturbing the average dissimilarity matrices with the standard error of the mean matrices (perturbation was in the range of +/- 1 SEM). Each of the 500 MDS solutions was fit to the MDS solution from the unperturbed matrix using the *procrustes* function of MATLAB. This Monte-Carlo-like technique can be used to define areas around the location of objects in the MDS map. The size of the area (visualized in figure 4 as contour lines), resembles the standard error of the mean over participants. The outermost contour line encloses 80% of the calculated perturbed matrices.

3 Results

3.1 Dimensionality

For all experiments and all conditions stress values were calculated for an MDS with one to five dimensions (see figure 3). An elbow in the plot shows which amount of dimensions is sufficient to explain the data. For all conditions except for plane 3 in the visual conditions two dimensions explain the data sufficiently. For plane 3 in both visual conditions the stress value for one dimension is already below 0.2, whereas the elbow is at two or rather at three dimensions. Taking just the stress value into account, one dimension would actually be enough to explain the data, whereas the other criterion would point towards two dimensions being sufficient (see next section for an explanation). For the sake of easier comparison, we nevertheless decided to conduct all MDS analyses in a two-dimensional parameter space (most importantly, this concerns the Monte-Carlo-like determination of map stability mentioned in section 2.4).

3.2 MDS Output Maps

Participants performed similarity ratings on pairs of objects of the three-dimensional stimulus space which was split into three two-dimensional planes to reduce the amount of trials in every experiment. The data of ten participants were averaged and fed into a two-dimensional MDS analysis which yields a map showing the location of the seven objects of one plane in space (figure 4). Each "original"

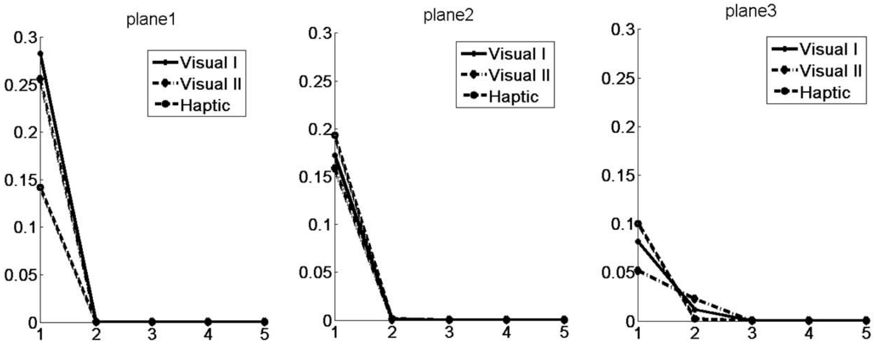

Fig. 3. Stress Values. Stress Values for the first and second visual condition and for the haptic condition plotted for each plane as a function of number of dimensions

object is identified in figure 4 by the number used in figure 2. As can be seen, for all experiments and for all planes, the ordering of stimuli was correctly identified, i.e., every object had the same neighbors in the MDS output space as in the physical object space.

In addition, each object in figure 4 is surrounded by density fields visualizing the stability of the map under perturbation as described in section 2.4. The outermost contour line surrounds 80% of the calculated values[1]. From the size of the density fields, one can immediately see that for all three planes the stability is *higher* in the second visual experiment than in the first visual experiment. We hypothesize that this due to the increased regularity in the sequence of stimulus presentation, which in turn leads to a higher stability of the obtained data as the brain is able to better extract the parameter dimensions from data coming from one plane at a time (even with the participants not being aware of this fact). Additionally, across all experiments stability is lowest for plane 3, which is due to a comparatively larger amount of inter-participant variance.

Using the density fields, we can also approach the question whether the MDS maps from the different conditions are "statistically different" or whether - due to the variance introduced by individual participants - perceptual spaces across conditions are the same. This can be done by overlaying the (properly rotated) maps and determining whether the density fields overlap. If they do overlap for *all* reconstructed object locations, this is a good indication that the perceptual spaces are highly similar. Comparing the MDS output maps in this fashion for the two visual conditions we obtained a complete overlap. We therefore conclude that the topology of the perceptual space is the *same* for all three planes for both visual experiments. This means that although the change in presentation order changed stability it did *not* change the perceptual maps[2].

The haptic experiment was performed with the same sequence of stimuli as the first visual experiment. For plane 1 and plane 2, stability is comparable across both modalities. Interestingly we found that for plane 3, stability is actually *higher* for the haptic condition as for the visual condition. Rating similarity between these shell-shaped objects therefore seems to be more consistent using touch than using vision. Again, using the density fields, we find that for all planes, the maps overlap between the visual and the haptic experiment. For all three planes we can therefore conclude that *highly similar* perceptual maps were formed for the visual and for the haptic condition.

For all three planes, figure 4 also shows the Y-shape form used to construct the stimulus space. The rotation of the Y-shape is a little different for each condition due to the fact that MDS cannot uniquely define the rotation of the map in space. In addition, we distorted the Y-shape such that it provided the best least-squares fit to the reconstructed, average perceptual space. As figure 4 shows, the

---

[1] Note that stars do not lie in the centers of the density maps due to the highly nonlinear nature of the MDS algorithms.

[2] Note that even though it seems that the density fields for plane 3 are more stretched in the horizontal dimension than for the other experiments, the difference is not statistically significant.

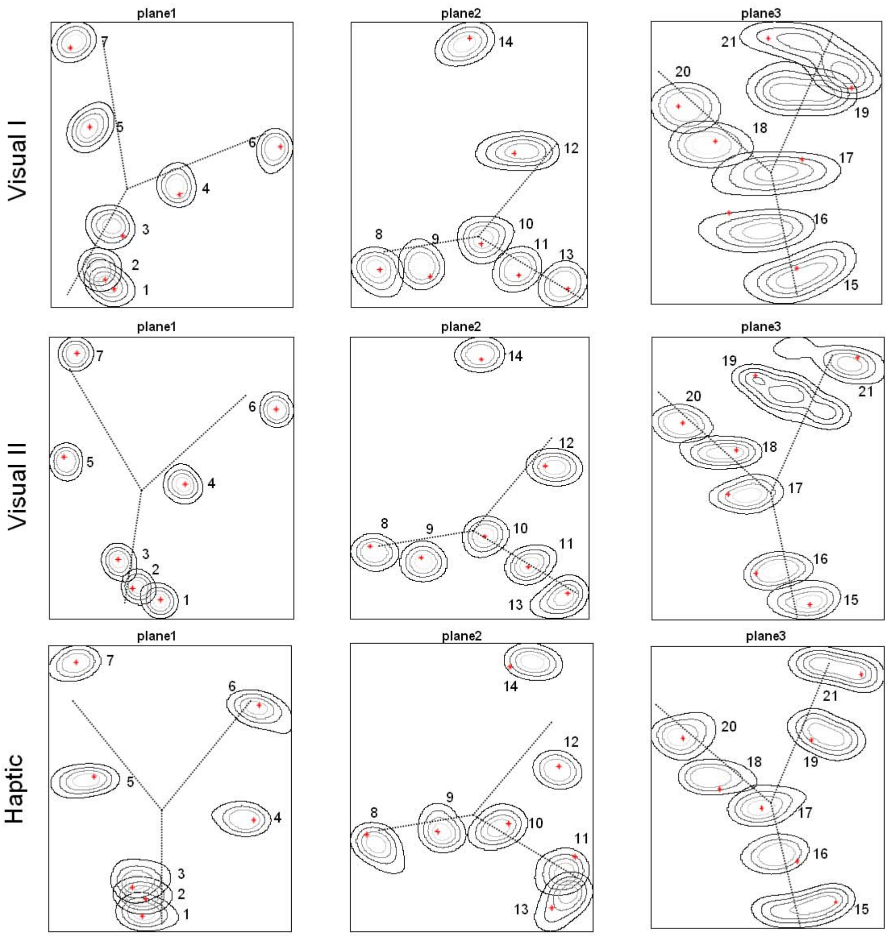

Fig. 4. MDS Output Maps. The figure plots MDS output maps for each of the three planes for the first and second visual condition and the haptic condition. The unperturbed MDS from the average data is marked with stars each of which is numbered to identify the stimuli according to figure 2. The density fields visualize the stability of the map under perturbation. The outermost line encloses 80% of the perturbed MDS maps. Additionally, the figure shows the best fit of the original, physical parameter map (a Y-shape) to the perceptual map.

Y-shape was perceived by the participants but distorted differently for every plane. For plane 1, the Y-shape is not stretched meaning that both dimensions were weighted equally. For plane 2, the Y-shape is stretched horizontally which corresponds to the parameter change of $\epsilon^{\cot\alpha}$. For plane 3, the Y-shape is also stretched horizontally which corresponds to a parameter change of $\sin\beta$. In both planes the second dimension corresponds to a parameter change of A. This leads to the assumption that changes in A (at least with the parameter range used in this experiment) are not as easy to perceive as the other two parameters. This parameter is thus not weighted as strongly as the other dimensions also resulting in a larger amount of inter-participant variance.

A parameter change in $\epsilon^{\cot\alpha}$ determines the number of convolutions the shell contains. The two other parameters determine the shape of the shells in a more complex way, however. After the similarity ratings participants, answered a questionnaire in which they described how they rated similarity and what features of the objects they used. Due to the complexity of the parameter space, however, it was not possible to form a clear correlation between participants' descriptions to the weighting of the different dimensions. This is in contrast to the earlier experiments by Cooke et al. [3] in which the questionnaires clearly contained descriptions relating to the physically manipulated parameters of shape and texture.

4 Summary and Outlook

From our experimental results, we can conclude that participants were able to identify the number of dimensions used to construct the complex, three-dimensional object space and that participants formed a perceptual space that is topologically highly similar to the physical object space. The non-intuitive parameters used to construct the shell-shaped objects were nevertheless perceived quite "faithfully" by the participants in both the visual and the haptic domain. Despite earlier results about the "restricted" capabilities of haptic perception [13], we actually found that even for this hard task, it was as good as visual perception when no restrictions were made to the exploratory procedures. In addition, we found that a change of the randomization procedure from "completely randomized" to "randomized by plane" allowed participants to rate similarity in a much more consistent fashion. Finally, from the high congruency between the visual and the haptic perceptual maps we conclude - taking also the earlier evidence from [3] into account - that visuo-haptic processing of similarity might be based on *one underlying space* which is accessible to both modalities.

To further investigate the underlying perceptual space, in one of our next experiments participants will learn the objects visually and later identify them haptically and vice versa. In subsequent experiments, we will also allow participants to actively manipulate the objects such that they can analyze and compare them from all viewpoints. In addition, we are planning neuro-imaging experiments using fMRI which would allow us to show if visual and haptic exploration of the shell-shaped objects would activate the same brain areas and thus if the underlying map is processed by similar brain structures for both modalities [12]. Additionally, any differences in activating these brain structures could be correlated to modality.

References

1. Borg, I., Groenen, P.J.F.: Modern Multidimensional Scaling: Theory and Applications. Springer, Heidelberg (2005)
2. Brainard, D.H.: The psychophysics toolbox. Spat Vis. 10(4), 433–436 (1997)
3. Cooke, T., Jäkel, F., Wallraven, C., Bülthoff, H.H.: Multimodal similarity and categorization of novel, three-dimensional objects. Neuropsychologia 45(3), 484–495 (2007)
4. Cooke, T., Kannengiesser, S., Wallraven, C., Bülthoff, H.H.: Object feature validation using visual and haptic similarity ratings. ACM Trans. Appl. Percept. 3(3), 239–261 (2006)
5. Cooke, T., Wallraven, C., Bülthoff, H.H.: A comparison of visual and haptic object representations based on similarity. In: 9th International Conference on Information Visualisation, pp. 33–40. IEEE Computer Science Society, Los Alamitos (2005)
6. Cooke, T., Wallraven, C., Bülthoff, H.H.: Characterizing perceptual differences due to haptic exploratory procedures: An mds approach. In: EuroHaptics 2006 Conference, July 2006, pp. 11–19 (2006)
7. Cox, T.F., Cox, M.A.: Multidimensional Scaling, vol. 2. Chapman and Hall, Boca Raton (2001)
8. Cutzu, F., Edelman, S.: Representation of object similarity in human vision: psychophysics and a computational model. Vision Res. 38(15-16), 2229–2257 (1998)
9. Fowler, D.R., Meinhardt, H., Prusinkiewicz, P.: Modeling seashells. In: SIGGRAPH 1992: Proceedings of the 19th annual conference on Computer graphics and interactive techniques, New York, NY, USA, pp. 379–387. ACM Press, New York (1992)
10. Garbin, C.P., Bernstein, I.H.: Visual and haptic perception of three-dimensional solid forms. Percept Psychophys 36(2), 104–110 (1984)
11. Hollins, M., Faldowski, R., Rao, S., Young, F.: Perceptual dimensions of tactile surface texture: a multidimensional scaling analysis. Percept Psychophys 54(6), 697–705 (1993)
12. James, T.W., Humphrey, G.K., Gati, J.S., Servos, P., Menon, R.S., Goodale, M.A.: Haptic study of three-dimensional objects activates extrastriate visual areas. Neuropsychologia 40(10), 1706–1714 (2002)
13. Klatzky, R.L., Lederman, S.J.: Touch in Experimental Psychology. John Wiley & Sons, New York (2003)
14. Loomis, J.M., Klatzky, R.L., Lederman, S.J.: Similarity of tactual and visual picture recognition with limited field of view. Perception 20, 167–177 (1991)
15. Newell, F.N., Bülthoff, H.H., Ernst, M.O.: Cross-modal perception of actively explored objects. In: Eurohaptics 2003 Conference Proceedings, pp. 291–299 (2003)
16. Pelli, D.G.: The videotoolbox software for visual psychophysics: transforming numbers into movies. Spat Vis. 10(4), 437–442 (1997)
17. Shepard, R.N.: Toward a universal law of generalization for psychological science. Science 237(4820), 1317–1323 (1987)