Comparing haptic, visual, and computational similarity-based maps of novel, 3D objects

Theresa Cooke, Christian Wallraven, Heinrich H. Bülthoff

Graduate School of Neural & Behavioural Sciences, Universität Tübingen

MPI FOR BIOLOGICAL CYBERNETICS

At a glance...

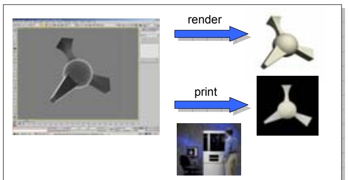
Questions: How do similarity relationships between objects differ when objects are seen and touched? What aspects of perceptual similarity can be captured using machine vision techniques?

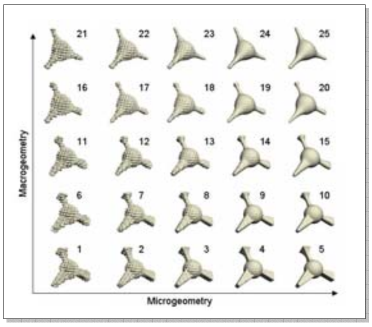
Approach: Gather similarity ratings, use multi-dimensional scaling (MDS) to derive maps of stimuli, and then compare maps based on human vision, human touch, and machine vision

Introduction

Similarity has been proposed as an organizational principle for representing objects in the brain [1-2]. But how does similarity vary as a function of perceptual modality? Here, we investigated this question by parametrically varying two object properties, shape and texture, gathered similarity ratings between pairs of objects, and used these to obtain modality-specific stimulus maps. Comparing the map obtained from visual similarity ratings against the map obtained by haptic ratings revealed differences in the weightings of shape and texture in the two modalities. We then compared these perceptual maps against maps derived from various **computational measures** of similarity to search for features/computations which may explain the perceptual similarities.

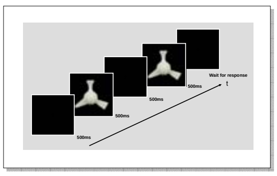
Parametrically-defined stimuli





Our **stimuli** consist of novel, three-dimensional objects whose shape and texture were varied parametrically: a **macrogeometric** smoothing operation gradually averaged out sharp edges in the objects' global shape, while a **microgeometric** smoothing operation gradually reduced the objects' local texture.

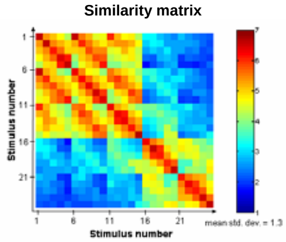
Visual Similarity Ratings: Experimental Design



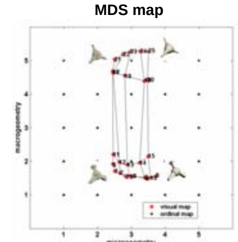
- 10 subjects **rated similarity** on a 7-point scale (1 = low similarity, 7 = high similarity)
- Stimuli were photographs of printed objects
- Bounding box of stimuli subtended 7° x 7° visual angle
- In one block, each stimulus appeared once with itself and once with every other stimulus, i.e., 0.5 x 25 x 24 + 25 = 325 trials/block
- 6 blocks with random trial order
- Afterwards, subjects filled out a questionnaire asking them to describe the objects and how they judged similarity

Visual Similarity Ratings: Results

Similarity matrix



MDS map



- Large box pattern in upper left ↔ effect of shape "groups"
- Fading off-diagonals ↔ effect of parametric shape change
- 5x5 box patterns ↔ effect of parametric texture change
- Similarity data analyzed by **MDS** to obtain **perceptual stimulus map**

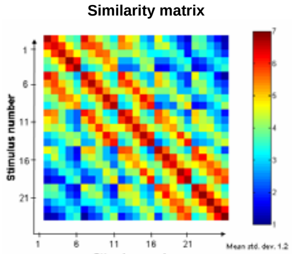
- Subjects **can extract a low-dimensional representation** from a high-dimensional measurement space
- Ordinal relationships between stimuli are preserved
- **Shape** is the dominant perceptual dimension, but subjects still recover order along texture axis
- Clear shape-based groupings (top 2 rows and bottom 3 rows) suggest a link to **category formation**

Haptic Similarity Ratings: Experimental Design

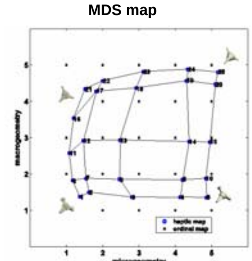
- 10 subjects **rated similarity** on a 7-point scale (1 = low similarity, 7 = high similarity)
- Subjects were given up to 10 s to trace the contour of each object with eyes closed
- Contour-following was chosen because it has been shown to allow for haptic extraction of a wide range of object properties, including local texture and global shape [3]
- 3 blocks of 325 randomized trials over five 2-hour sessions on consecutive days
- Afterwards, subjects filled out a questionnaire asking them to describe the objects and how they judged similarity

Haptic Similarity Ratings: Results

Similarity matrix



MDS map



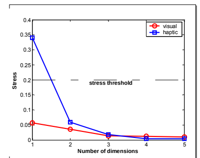
- Absence of large box pattern seen in visual data
- Fading off-diagonals ↔ effect of parametric shape change
- 5x5 box patterns ↔ effect of parametric texture change
- Similarity data analyzed by **MDS** to obtain **perceptual stimulus map**

- Subjects **can extract a low-dimensional representation** from the high-dimensional **haptic** measurement space
- Ordinal relationships between stimuli are preserved
- **Both shape and texture** are important perceptual dimensions
- Same groupings present as in visual data, but less pronounced along the shape axis

Perceptual dimensions in vision and touch

MDS stress plot

- **Visual** ratings: a single dimension suffices to explain similarity data (shape)
- **Haptic** ratings: two dimensions are required



MDS stress plot obtained using Young's S-stress formula 1. Stress should fall below 0.2 for an adequate number of dimensions [4].

Individual shape/texture weighting

- **Visual** ratings: consistent dominance of shape over texture
- **Haptic** ratings: broad range of relative weightings → subjects use different strategies

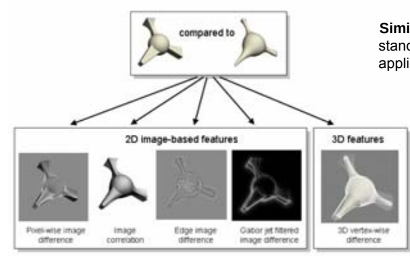
Individual shape/texture tradeoff values are based on a ratio of dimension weights provided by the INDSCAL algorithm [6].

Responses to questionnaire

- Subjects perceive two dimensions of variation as "shape" and "texture"
- **Visual** ratings: Both shape and texture were used to describe objects, but shape was used more often to describe how similarity was judged
- **Haptic** ratings: Texture was used more often to describe objects, but shape and texture were mentioned equally often to describe how similarity was judged

	VISUAL		HAPTIC	
	Object description	**Similarity judgment**	**Object description**	**Similarity judgment**
Mention **"shape"** or shape-related properties	90%	90%	70%	100%
Mention **"texture"** or texture-related properties	80%	60%	100%	100%
Mention other properties (e.g., colour, weight)	40%	0%	30%	10%

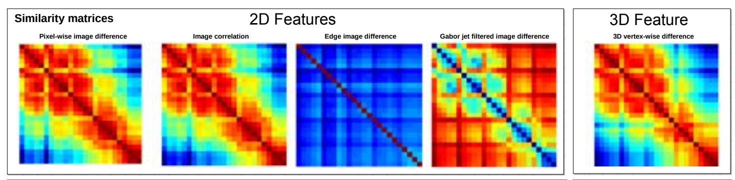
Computational Similarity Measures: Method

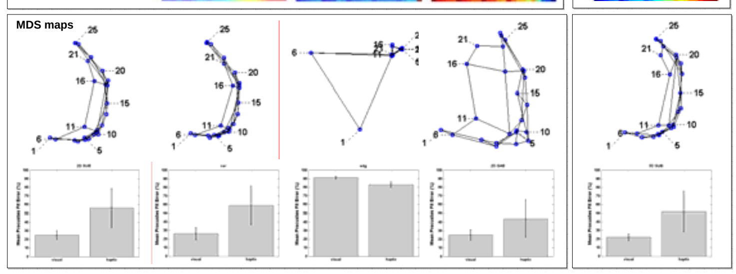


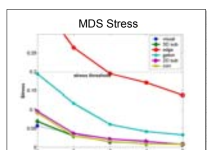
Similarity ratings were generated using the following standard **computer vision techniques** and **MDS** was applied to generate maps of the stimuli:

1) Sum of squared differences (SSD) between pixel values of two images;
2) Correlation between two images;
3) SSD between two images generated by running Canny edge detector on stimulus images;
4) SSD between two images generated by filtering stimuli with Gabor jets;
5) SSD between two object meshes' 3D vertex positions.

Computational Similarity Measures: Results







For all measures except the edge detector, a single dimension suffices to explain the similarity data.

Similarity matrices and maps
- Edge detection correlates poorly with human data
- 2D/3D subtraction, correlation, and Gabor jets yield comparable correlations with human data

Shape vs. texture
- All except edge detector dominated by shape dimension
- Recovery of shape-based groupings (lower 3 and upper 2 rows)

Computational vs. human haptic/visual maps
- Fit evaluated using mean Procrustes fit error across individual maps
- Better correlations with human visual than human haptic data

⇒ A new technique for **perceptually validating** computational features

Conclusions and Outlook

Extraction of low-dimensional variation from high-dimensional measurement spaces
- In both modalities, subjects were able to extract two stimulus variations, which they referred to as changes in "shape" and "texture": non-trivial given the high-dimensionality of measurement spaces → motivates a comparison against computational measures

Visual vs. haptic similarity representations
- Shape dominated visual representations, while both shape and texture were important for haptic representations
- Stimuli clustered in both similarity spaces, suggesting a link between similarity relationships and category structure [5]

Methodological advantages
- Using this approach, object representations/topologies can be compared across modalities
- Using this approach, human and computationally-derived representations can be compared, e.g., for perceptual validation of computational features

Future work
- Validation of a wider range of computational features
- Studies to explore the generalizability of these results to other stimulus sets
- Studies to explore relationship between similarity and categorization for vision vs. touch

References

1. Edelman, S, Representation and recognition in vision, MIT Press, 1999
2. Cutzu, F and Edelman, S, Representation of object similarity in human vision: psychophysics and a computational model, Vision Res, 1998, 38, 2229-2257
3. Lederman, SJ and Klatzky, RL. Extracting object properties through haptic exploration Acta Psychol, 1993, 84, 29-40
4. Clarke KR, Warwick RM, Change in Marine Communities: An Approach to Statistical Analysis and Interpretation, 2nd edition, PRIMER-E, Plymouth, UK, 172pp, 2001
5. Hahn, U, and Ramscar, M (Eds). Similarity and categorization. Oxford University Press, 2001
6. Carroll, JD and Chang, JJ. Analysis of individual differences in multidimensional scaling via an n-way generalization of "Eckart-Young" decomposition. Psychometrika, 35, 283-319, 1970