Features for visual object recognition

Martin Arguin, Marie-Audrey Lavoie, Mélanie Lévesque, Gabriela Milanova, Pénélope Pelland-Goulet, Marc-Antoine Akzam-Ouellette and Youri Tassé

Université de Montréal

cerebrum

CIRCA

Centre de recherche iugm Institut universitaire de gériatrie de Montréal

Visit us at: www.neurocognitionvisionlab.org

INTRODUCTION

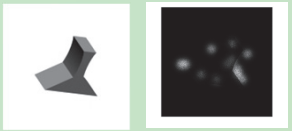
Human visual object recognition largely relies on shape information. However, the nature of the shape features that actually underlie this task remain largely unknown despite a wealth of competing theories aiming to account for the code by which human vision represents shape. Here, we report a series of five object recognition experiments using a spatial sampling paradigm (cf. Bubbles) to calculate classification images (CIs) that demonstrate the efficient features used by human participants.

METHODS

In all experiments, the targets were behind an occluding mask and partially revealed for 100 ms by a collection of 12 randomly located circular gaussian apertures of 0.8° in diameter. Participants pressed a keyboard key to indicate the identity of the target. Response accuracy was maintained at 50% correct by manipulating the degree of degradation of the target image (contrast of overlaid white noise). The experiments differ from one another in terms of the class of stimuli and the exposure of instances from various viewpoints or not.

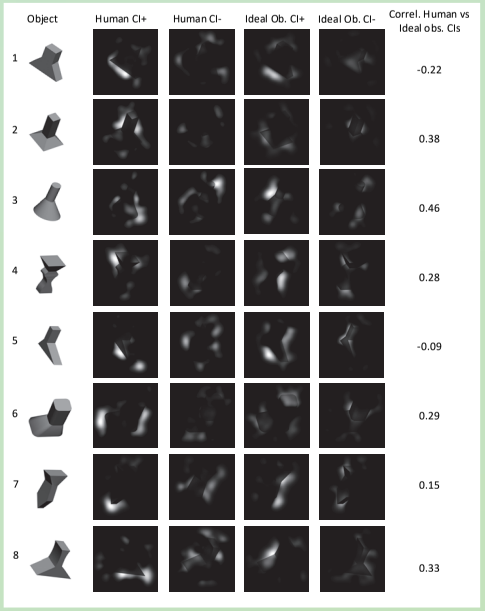
For all experiments, the results from human observers were compared to those of an ideal observer.

The results of each experiment are illustrated in terms of the classification images associated with significantly high (CI+) or low (CI-) processing efficiency for each stimulus for the human and ideal observers and the correlation between them.

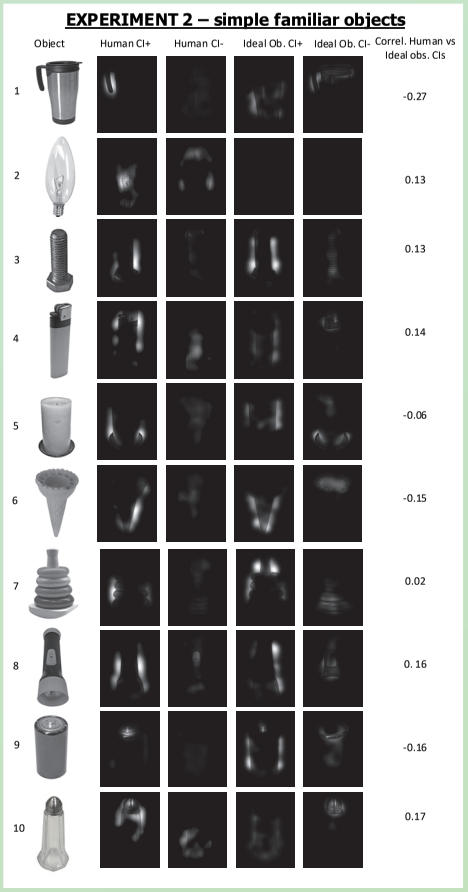


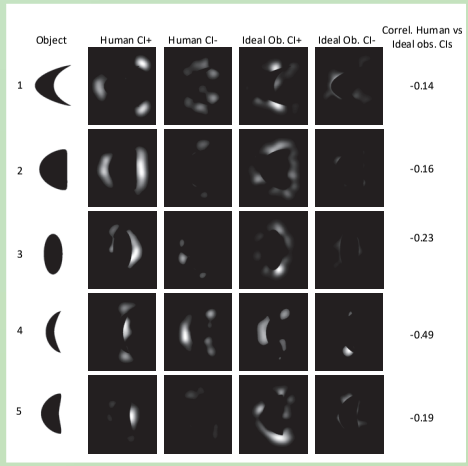
Example of a target object and of its spatial sampling using 12 randomly located gaussian 0.8° apertures.

EXPERIMENT 1 – two-part novel objects

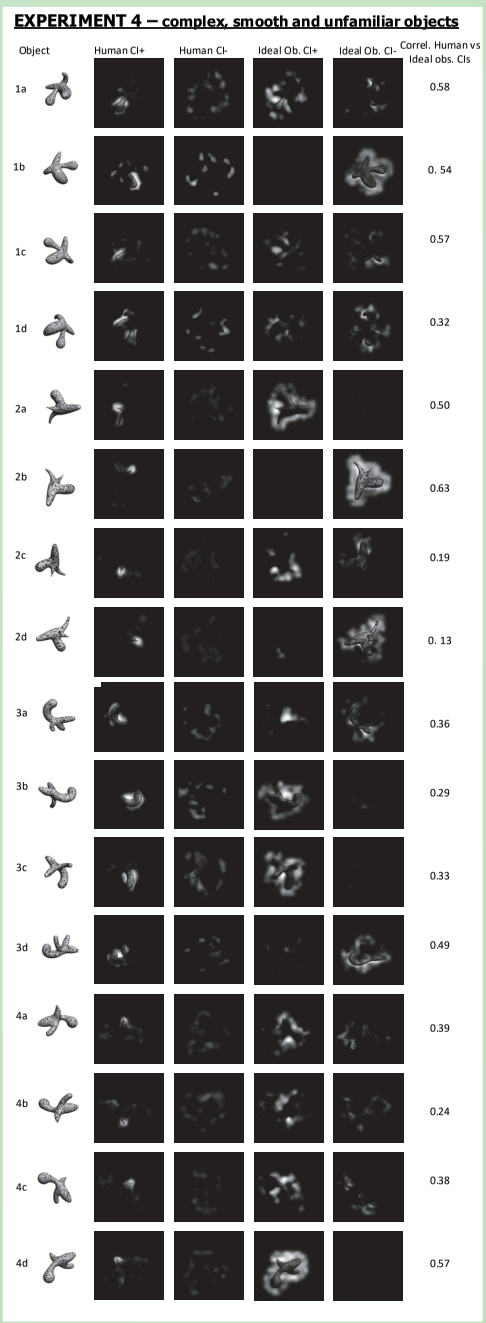


EXPERIMENT 2 – simple familiar objects

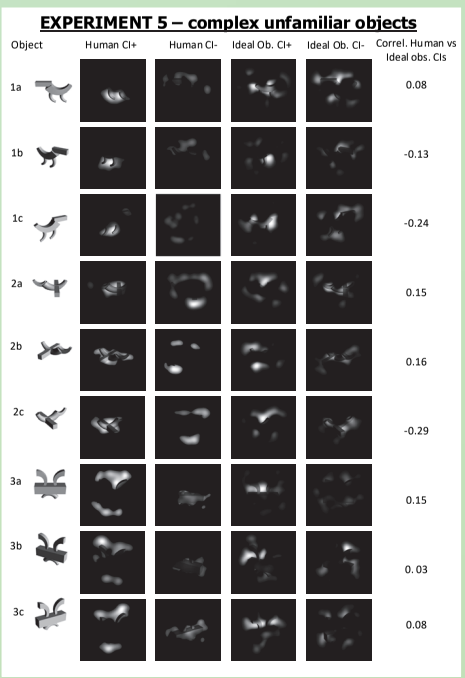




EXPERIMENT 4 – complex, smooth and unfamiliar objects



EXPERIMENT 5 – complex unfamiliar objects



DISCUSSION

Within-group coherence analyses show that the mean CIs in all experiments constitute a proper representation of those of individual participants.

The CIs of human participants generally correlate poorly with those of the ideal observer. Thus, limitations of the human visual system impose constraints on the effective features.

For all objects that were presented along variable viewpoints (Exps. 4 and 5), the regions on the object's surface which constituted the effective features were well matched across viewpoints.

There is no particular type of feature such as those proposed by major shape perception theories (e.g. concavities, convexities, edge intersections, object parts, etc.) that dominate in the mean CIs and feature sizes are quite variable.

Overall, the features uncovered appear most compatible with the 'image fragment' or the 'minimal recognizeable configurations (MIRCs)' approaches proposed by Ullman and collaborators.

Financial support:

NSERC CRSNG

Thanks to Charles E. Connor for the software used in the generation of stimuli for Exp. 4.