A Computational Method to Find Salient Features

Na-Yung Yu[1*], Takashi Yamauchi[1], Huei-Fang Yang[2], Yen-Lin Chen[2], Ricardo Osuna-Gutierrez[2]

[1] Department of Psychology, Texas A&M University, TX 77801 USA
[2] Department of Computer Science, Texas A&M University, TX 77801 USA
* Correspondence should be addressed to nayungyu@gmail.com

Abstract

Although it is well known that people selectively attend to salient features in similarity judgment, no clear method of identifying "salient features" has been proposed. In this study, we present a new computational technique to identify salient features. First, we collected behavioral data from human participants, and this data was simulated with machine learning techniques, which determined optimal allocations of weights of candidate features. Results revealed image-specific sets of salient features for similarity perception, and suggested that people exaggerate differences between features while computing similarity.

Introduction

Similarity is the bedrock of human induction and generalization. How do people perceive the similarity between objects? A traditional assumption in similarity research is that people attend to matching and mismatching attributes selectively (Tversky, 1977). Although it is well known that people selectively attend to some salient features while ignoring others in similarity judgment, no clear method of identifying salient features has been proposed. In this brief article, we present a new computational method for identifying "significant features" in similarity judgments. In brief, we first collected behavioral data from human participants and applied a machine learning technique to identify 1) what kind of features people use for similarity judgment and 2) how they process them to perceive the similarity between animal faces.

Behavioral Data

We collected ten original animal faces and scaled the images to grayscale, 307 pixel X 307 pixel images. We paired these ten images whose outlines were sufficiently similar, creating five animal pairs (i.e. bear-fox, cow-pig, hippo-sheep, koala-rat, and lion-horse pair). For each pair, one original picture (i.e. Source) was merged with the other original picture (i.e. Target) in Morphman 4.0 (2003). Eighteen morphed pictures were generated for each animal pair. Altogether, 90 morphed pictures (5 animal pairs × 18 morphed pictures) and 10 original pictures were used in the experiment (Figure 1).

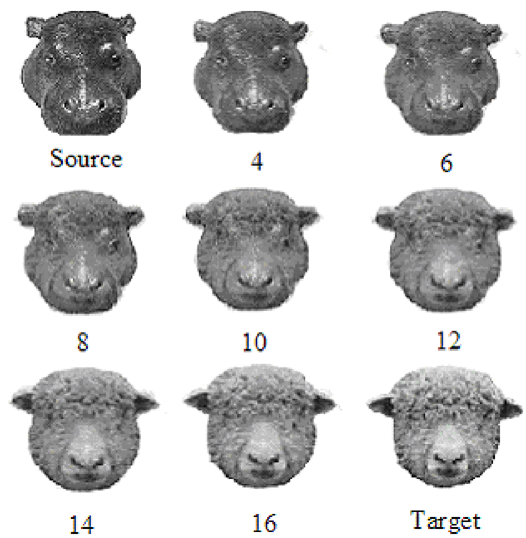

Figure 1: Morphed images for the hippo-sheep animal pair; numbers represent morphing progression

Participants saw two original pictures of each pair on the top and one morphed picture of the pair on the bottom. Their task was to decide which original picture (i.e., source or target) was more similar to the morphed picture (i.e., input image; for a similar task, see Sloutsky and Fisher, 2004). The proportion of participants selecting the source image was recorded (i.e. proportion of selecting hippo for hippo-sheep pair). In the next stage, computational analysis simulated behavioral data, in order to identify the facial features salient in human similarity judgment.

Computational Analysis

Feature Extraction and Simulation

To simulate the behavioral data computationally, we first listed the 34 potential facial features that people are likely to use for their similarity judgments. These candidate features include texture difference, relative color and size. For texture features, we explored Gabor texture (Manjunath & Ma, 1996) and the rate of co-occurring features (Howarth & Ruger, 2004), which have been commonly used in image retrieval field. The averaged gray values of each image were used as a color feature. The ratio of width to height of the face was used to represent the size information. The texture, color and size features were extracted for a whole image as well as its nine sub-regions where each region represents a different part of an animal face.

To identify "salient features" among the candidate features, the weights for each feature were calculated. Salient features should have higher weight values than other features. Because it is computationally prohibitive to evaluate all possible combinations of weights for all features, a modified simulated annealing (SA) algorithm was used to extract the optimal weight combination of feature vectors for each image pair.

We first measured the Euclidean distance between the input image and the source image $d(X_i, S)$, and the input image and the target image $d(X_i, T)$, and then compared the predicted similarity (i.e., $sim(X_i, S)$) to the similarity scores obtained in the behavioral experiment. In equation (1), S denotes the source image, and T denotes the target image in the same animal pair. P denotes the exponential power of the measured distances. When P equals 1, the distance measurement was the Euclidean distance.

$$sim(X_i, S) = \frac{sim(X_i, T)^P}{sim(X_i, S)^P + sim(X_i, T)^P}$$

(1)

The pattern of feature weights that minimized the energy E was deemed an optimal allocation of feature weights, and those features that garnered large weights were extracted as salient features (Russell & Norvig, 2002).

Results and Discussion

The computational simulation on the human behavioral data suggests that people use different features to perceive the similarity between animal faces. Our simulation results also suggest that people are likely to exaggerate similarities and differences. As the power of the Euclidean distance between two animal faces was increased (i.e., parameter P see Equation 1), the SSE between the observed and predicted data decreased considerably. Finally, our procedure combining a machine learning technique with a behavioral experiment provides a new avenue to identify "salient features" in similarity judgments.

References

Howarth P., & Ruger, S. (2004). Evaluation of texture features for content-based image retrieval, *The Proceedings of International Conference on Image and Video Retrieval* (pp. 326-334).

Manjunath B. S. & Ma, W., Y. (1996). Texture features for browsing and retrieval of image data, *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI), 18*, 837-842.

Russell, S., & Norvig, P. (2002). Artificial intelligence: A modern approach (2nd ed.). Prentice Hall.

Sloutsky, V., M., & Fisher, A., V. (2004). Induction and categorization in young children: a similarity-based model, J*ournal of Experimental Psychology: General, 133*, 166-188.

Tversky, A. (1977). Features of similarity, Psychological Review, 84, 327-352.