Vision Research 62 (2012) 148–161

Contents lists available at SciVerse ScienceDirect

Vision Research

journal homepage: www.elsevier.com/locate/visres

View-invariance learning in object recognition by pigeons depends on error-driven associative learning processes

Fabian A. Soto [a,*], Jeffrey Y.M. Siow [b], Edward A. Wasserman [b]

[a] Department of Psychological and Brain Sciences, University of California Santa Barbara, Santa Barbara, CA 93106, USA
[b] Department of Psychology, University of Iowa, Iowa City, IA 52242, USA

ARTICLE INFO

Article history:
Received 2 January 2012
Received in revised form 22 March 2012
Available online 17 April 2012

Keywords:
Associative learning
Prediction error
Object recognition
View invariance
Categorization
Pigeon

ABSTRACT

A model hypothesizing that basic mechanisms of associative learning and generalization underlie object categorization in vertebrates can account for a large body of animal and human data. Here, we report two experiments which implicate error-driven associative learning in pigeons' recognition of objects across changes in viewpoint. Experiment 1 found that object recognition across changes in viewpoint depends on how well each view predicts reward. Analyses of generalization performance, spatial position of pecks to images, and learning curves all showed behavioral patterns analogous to those found in prior studies of relative validity in associative learning. In Experiment 2, pigeons were trained to recognize objects from multiple viewpoints, which usually promotes robust performance at novel views of the trained objects. However, when the objects possessed a salient, informative metric property for solving the task, the pigeons did not show view-invariant recognition of the training objects, a result analogous to the overshadowing effect in associative learning.

1. Introduction

Visually recognizing objects in the environment confers a clear advantage for the survival and reproduction of any animal. Among many functions, object recognition allows the animal to detect food, conspecifics, and predators.

An important computational problem posed by object recognition (Rust & Stocker, 2010) is that of invariance: the same object can project very different images to the retina, depending on such factors as viewpoint, position, scale, clutter, and illumination. The present work focuses on understanding how a biological visual system (i.e., the pigeon) learns to recognize objects across variations in viewpoint.

Several experiments have explored whether pigeons show view-invariant object recognition after being trained with only one object view. These experiments have uniformly found that pigeons do not show one-shot view invariance, regardless of the type of object used to generate the experimental stimuli (Cerella, 1977; Friedman, Spetch, & Ferrey, 2005; Lumsden, 1977; Peissig et al., 1999, 2000; Wasserman et al., 1996). However, pigeons do show above-chance performance with novel views of the training object after training with just one view and they exhibit generalization behavior that is closer to true view invariance as the number of training views is increased (Peissig et al., 2002, 1999; Wasserman et al., 1996).

These and other studies suggest that the pigeon's recognition of objects from novel viewpoints depends on similarity-based generalization from the training views (Spetch & Friedman, 2003; Spetch, Friedman, & Reid, 2001), prompting these questions: Which object properties do pigeons use to generalize performance from training images to novel images? How are such properties extracted from images? How are such properties selected during training to guide performance in a particular task?

Regarding the first question, evidence suggests that pigeons extract view-invariant properties from images and rely heavily on them for object recognition (Gibson et al., 2007; Lazareva, Wasserman, & Biederman, 2008). For example, Gibson et al. (2007) trained pigeons (and people) to discriminate four simple volumes shown from a single viewpoint in a four-alternative forced-choice task. Once the subjects attained high performance levels in the task, the researchers used the Bubbles technique (Gosselin & Schyns, 2001) to determine which properties of the images the subjects used for recognition. The results showed that pigeons (and people) relied more heavily on properties that are relatively invariant across changes in viewpoint, such as cotermination and other edge properties, than on properties that vary across changes in viewpoint, such as shading.

However, pigeons show pronounced decrements in recognition performance when they are tested with novel object views that retain view-invariant properties (Peissig et al., 1999, 2000, 2002) and

* Corresponding author. Fax: +1 805 893 4303.
E-mail address: fabian.soto@psych.ucsb.edu (F.A. Soto).

0042-6989/$ - see front matter © 2012 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.visres.2012.04.004

with images that have been manipulated only in view-specific properties, such as shading (e.g., Young et al., 2001). In sum, the evidence suggests that pigeons recognize objects from novel viewpoints through a generalization-based mechanism, and that generalization might be based on the extraction of both view-invariant and view-specific shape properties.

We propose that this evidence is best interpreted within the framework of the recently proposed "Common Elements Model" (Soto & Wasserman, 2010a, 2012) of object categorization learning, which is based on the idea that basic mechanisms of associative learning and generalization underlie pigeons' ability to classify natural objects. Because such basic mechanisms are widespread among vertebrate species, they might also play a role in object categorization by other species, including humans (see Soto & Wasserman, 2010b).

The Common Elements Model proposes that each image in a categorization task is represented by a set of "elements," which can be interpreted as coding visual properties in a training image. These properties vary widely in the level to which they are repeated across members of the category. In the case of view-invariant object recognition, properties would show varied levels of view invariance, going from relatively view-invariant properties, which are repeated across many views of the same object, to view-specific properties, which are idiosyncratic to a particular object view. Importantly, the model also suggests a mechanism that selects which properties should control the performance of each available response in a recognition task. This selection process is carried out through associative error-driven learning, which selects those properties that are more informative as to whether the response will lead to a reward.

Consider how the Common Elements Model would explain the effect of training with multiple views of an object on the later recognition of novel views. Training with different views of the same object would lead to a "repetition advantage" effect for view-invariant properties. View-invariant properties are often repeated across different training views and they are frequently paired with the correct responses. This repetition gives them an advantage in controlling performance over view-specific properties, which are not common to many views and therefore do not frequently get paired with the correct response. Even if both types of property are informative as to the correct responses in the task, learning continues only until there is no error in the prediction of reward. At this point, view-invariant properties block view-specific properties from acquiring an association with the correct responses. When novel views of the object are presented in testing, these testing views are likely to share some view-invariant properties with the training views, leading to the successful generalization of performance.

Thus, the hypothesis that pigeons extract visual properties with differing levels of view invariance can explain why they show, on the one hand, extremely view-dependent recognition after training with a single view of an object and, on the other hand, they show high sensitivity to view-invariant object properties. Training with a single view of an object leads to the control of behavior by both view-invariant and view-specific properties, because there is no repetition advantage for the former. Even if view-invariant properties are highly salient, behavior must generalize imperfectly to novel views of an object, which do not share the same view-specific properties as the training view.

The present work presents the results of two experiments testing the predictions of the Common Elements Model of view-invariance learning by pigeons. Specifically, these experiments test the assumption that visual properties are selected to control performance in object recognition tasks through an associative learning mechanism driven by reward prediction error. Error-driven learning should lead to competition between view-invariant and view-specific properties for control of performance in object recognition tasks. The experiments show that results analogous to the relative validity effect (Wagner et al., 1968) and overshadowing (Pavlov, 1927) from the associative learning literature can be found in object recognition experiments. These are two key effects that proved to be central to the development of associative learning theories based on the notion of prediction error. If view-invariance learning in pigeons is driven by reward prediction error, then such experimental designs should reveal competition for behavioral control between view-invariant and view-specific object properties.

2. Experiment 1

A relative validity experiment in Pavlovian conditioning (Wagner et al., 1968; Wasserman, 1974) involves training with two compound stimuli: AX and BX. In the Uncorrelated condition, each compound is reinforced 50% of the time. In the Correlated condition, AX is always reinforced and BX is never reinforced. Even though, in both conditions, X is reinforced 50% of the time—and hence its absolute predictive value is the same—subjects in the Uncorrelated condition respond more to this stimulus than do subjects in the Correlated condition. Thus, conditioning to X depends on the informative value of the *other* stimuli that are presented in compound with it. When A and B are reliable predictors of the

Table 1
Design of Experiment 1, which tested an analog of the relative validity effect in view-invariant object recognition. "Rf" stands for reinforcement and "NRf" stands for nonreinforcement.

Training	Generalization test
Uncorrelated	
Geon 1 – 0°/50% Rf	Training trials +
Geon 1 – 120°/50% Rf	Geon 1 – Rotated around *x*-axis at 30°, 90°, 150°, 210°, 270°, 330°/NRf
Geon 1 – 240°/50% Rf	Geon 1 – Rotated around *y*-axis at 30°, 90°, 150°, 210°, 270°, 330°/NRf
Geon 1 – 60°/50% Rf	
Geon 1 – 180°/50% Rf	
Geon 1 – 300°/50% Rf	
Correlated	
Geon 2 – 0°/Rf	Training trials +
Geon 2 – 120°/Rf	Geon 2 – Rotated around *x*-axis at 30°, 90°, 150°, 210°, 270°, 330°/NRf
Geon 2 – 240°/Rf	Geon 2 – Rotated around *y*-axis at 30°, 90°, 150°, 210°, 270°, 330°/NRf
Geon 2 – 60°/NRf	
Geon 2 – 180°/NRf	
Geon 2 – 300°/NRf	