Behavioural Brain Research 356 (2019) 423–434

Contents lists available at ScienceDirect

Behavioural Brain Research

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

Research report

Emerging category representation in the visual forebrain hierarchy of pigeons (*Columba livia*)

Amir Hossein Azizi[a,1,*], Roland Pusch[b,1], Charlotte Koenen[b], Sebastian Klatt[a], Franziska Bröcker[b], Samuel Thiele[b], Janosch Kellermann[c], Onur Güntürkün[b], Sen Cheng[a]

[a] Institute for Neural Computation, Ruhr-University Bochum, 44801 Bochum, Germany
[b] Biopsychology, Institute of Cognitive Neuroscience, Faculty of Psychology, Ruhr-University Bochum, 44801 Bochum, Germany
[c] Chair of Statistics and Econometrics, Faculty of Management and Economics, Ruhr-University Bochum, 44801 Bochum, Germany

ARTICLE INFO

Keywords:
Categorization
Visual forebrain areas
Pigeon

ABSTRACT

Recognizing and categorizing visual stimuli are cognitive functions vital for survival, and an important feature of visual systems in primates as well as in birds. Visual stimuli are processed along the ventral visual pathway. At every stage in the hierarchy, neurons respond selectively to more complex features, transforming the population representation of the stimuli. It is therefore easier to read-out category information in higher visual areas. While explicit category representations have been observed in the primate brain, less is known on equivalent processes in the avian brain. Even though their brain anatomies are radically different, it has been hypothesized that visual object representations are comparable across mammals and birds. In the present study, we investigated category representations in the pigeon visual forebrain using recordings from single cells responding to photographs of real-world objects. Using a linear classifier, we found that the population activity in the visual associative area *mesopallium ventrolaterale* (MVL) distinguishes between animate and inanimate objects, although this distinction is not required by the task. By contrast, a population of cells in the *entopallium*, a region that is lower in the hierarchy of visual areas and that is related to the primate extrastriate cortex, lacked this information. A model that pools responses of simple cells, which function as edge detectors, can account for the animate vs. inanimate categorization in the MVL, but performance in the model is based on different features than in MVL. Therefore, processing in MVL cells is very likely more abstract than simple computations on the output of edge detectors.

1. Introduction

Possibly all vertebrates need to categorize the various stimuli surrounding them. Categorization is an important step beyond mere discrimination and constitutes the ability to respond equivalently to members of the same class, to respond differently to members of a different class, and to transfer this distinction to novel members of these classes [1]. As such, categorization is a key component of cognitive systems since it drastically reduces information load [2].

Category-relevant neuronal representations were mostly studied in primates [e.g., [3–5]]. Using fMRI, Huth et al. [6] found that category representations in the human inferior temporal cortex (ITC) form a semantic map. Intriguingly, Kriegeskorte et al. [7] found similar representations in macaque and human ITC using single-unit recordings and imaging, respectively. These findings suggest a common code of object representation across primate species.

On a behavioral level, categorization has been studied in many non-primate species (for review see [8]). Pigeons were the first non-human animals in which the ability to categorize was demonstrated [9] and are presently the best-studied animal model for categorization at behavioral level. Pigeons successfully learn to categorize vertebrates [10], martial arts poses [11], malignant cancer cells [12] and Cubist painting styles [13]. Pigeons even successfully learn to discriminate English words from non-words [14]. Taken together, pigeons have unexpectedly large resources to learn various perceptual categories.

On an anatomical level, the brain structures involved in the process of categorization differ between birds and mammals. The telencephalon of both groups is composed of two major subdivisions: the pallium and the subpallium. Subpallial components are relatively conserved among birds and mammals. In contrast, there are remarkable differences in the pallial organization: in mammals it includes the six-layered cortex. In contrast, the pallium in birds is organized in a nuclear fashion [15]. A morphological identifiable layered cortex is absent and there is no structure that directly corresponds to the primate ITC. Despite the

* Corresponding author.
E-mail address: amir.azizi@ruhr-uni-bochum.de (A.H. Azizi).
[1] These authors contributed equally to this work.

https://doi.org/10.1016/j.bbr.2018.05.014
Received 28 March 2018; Received in revised form 14 May 2018; Accepted 15 May 2018
Available online 06 June 2018
0166-4328/

differences in gross anatomy, nuclear structures in the pallium of birds share homology with laminar specific neuronal populations in mammals: the avian entopallium may share homologies to layer 4 input neurons of the neocortex, receiving sensory thalamic input. Some neurons in arcopallium share homology with layer 5 output neurons projecting to the brainstem [16]. Albeit sharing their pallial origin, no one-to-one homologies of cortical areas and counterparts in the avian pallium can be derived. However, multiple studies show that mammals and birds command similar cognitive abilities [17,18], suggesting that cognitive functions are to some extent invariant to neuroanatomical differences. Consequently, it is conceivable that birds might represent categories with neuronal population level principles similar to primates. The current study investigates this possibility.

Several structures in the pigeon brain are involved in visual processing. We will focus on the tectofugal system since it controls visual processing of stimuli in the frontal visual field, where stimuli are usually presented during behavioral tasks [19,20]. The tectofugal pathway (Fig. 3) ascends from the retina via the midbrain optic tectum and the thalamic nucleus rotundus to the telencephalic entopallium [21]. The entopallium is anatomically and functionally similar to the primate extrastriate cortex [22–26]. Entopallial neurons modify their firing patterns according to reward-associated task contingencies [27] and are involved in visual working memory [28,29].

Entopallial neurons have reciprocal connections with surrounding associative areas like the nidopallium frontolaterale (NFL); the mesopallium ventrolaterale (MVL), and the nidopallium intermediale pars lateralis (NIL) [30–32]. The NFL, for example, processes motion and color [32] and NFL population responses distinguish between simple stimulus features, such as color and spatial frequency [33].

Apart from these studies, the physiology of visual areas beyond the entopallium is mostly unknown, and there is no evidence to date for high-level categorical object representation in any region in the avian brain. To close this gap, we investigated at which stage of the tectofugal pathway category information is represented in a way that it can be easily read out. In the experiment, pigeons viewed photographs of real-world objects, while we recorded from single neurons in the primary visual entopallium and the higher-order visual associative area MVL. Using multi-variate analyses, we found that the population of recorded MVL, but not of entopallium cells represents the distinction between animate vs. inanimate objects. Our results thus show for the first time that the visual processing hierarchy of the pigeon transforms visual inputs such that category information become easier to read out.

2. Materials and methods

2.1. Subjects

Four adult homing pigeons (*Columba livia*) were obtained from local breeders. The pigeons were housed in individual wire-mesh cages with a 12 h light–dark cycle beginning at 08.00 a.m. They were fed with a mixture of grains and maintained at approximately 80–90% of their free-feeding body weight. They had free access to water. Handling of the pigeons was in accordance with the National Institute of Health Guide for Care for Laboratory Animals, and the experiment was approved by the state authorities of North-Rhine Westphalia, Germany.

2.2. Apparatus and stimuli

All sessions took place in a custom-built operant chamber (35 cm × 35 cm × 35 cm) with three horizontally aligned pecking keys (4 cm × 4 cm, 17 cm above the floor). Stimuli were presented on a central pecking key and a food hopper delivered mixed grains as reward. Stimuli were presented on a LCD flat screen monitor mounted behind the pecking keys.

The stimulus set was the same as used in [7]. Fig. 1 shows all 96 color photographs of isolated real-world objects on a gray background. The pictures are divided into animate objects and inanimate objects. These classes are further subdivided into human and non-human objects, and in natural and artificial objects. Pigeons were not required to perform categorization of these stimuli in the behavioral task.

2.3. Experimental design and statistical analysis

2.3.1. Experimental paradigm

Experiments were conducted daily, five days a week. Each daily session consisted of 960 trials, which were divided into 10 blocks. In each block, the entire stimulus set was presented in a randomly permuted order. The order of presentations was therefore different in each block. A schema of an individual trial is depicted in Fig. 2. Each trial began with the appearance of an initialization image which was visible for up to 2 s and disappeared once pecked. After a short delay of 0.2 s, one of the sample stimuli was presented for a fixed interval of 2 s. During stimulus presentation, the pigeon was allowed to peck the presented stimulus without restriction. To ensure that the pigeons were looking at the stimulus, they had to peck the stimulus at least once to proceed in the trial. After the stimulus presentation and a short delay (0.2 s), the initialization key was presented again. Again, this image was presented for up to 2 s and disappeared once pecked. If the pigeons successfully pecked at all three images within a trial, they were rewarded. Access to food was granted with a probability of 55–70% according to the weight and the performance of the individual pigeon. The feeding period lasted for 2 s. Regardless of the delivery of a reward, a feeding light affixed above the food hopper was always on during this period, serving as a secondary reinforcement. After the reward period, a 6 s inter-trial-interval (ITI) preceded the next trial. When the pigeons failed to peck in any phase as required, the trial was aborted and the ITI followed immediately. All experimental hardware was controlled with custom-written Matlab code (MathWorks, Natick, MA, USA) with the aid of the Biopsychology toolbox [34].

2.3.2. Surgery

After behavioral training, pigeons were implanted with custom-built microdrives [35–38]. Each microdrive housed seven electrodes made of 25 µm formvar-coated nichrome wires (Science Products GmbH, Hofheim, Germany) and one additional 76 µm heavy-polyimide-coated stainless steel wire serving as reference (Franco Corradi, Milano, Italy), which were connected to microconnectors (Omnetics Connector Corporation, Minneapolis, USA). All pigeons were implanted with two microdrives, one in each brain hemisphere. Pigeons were initially anesthetized with a mixture of ketamine and xylazine (ratio 7:3) with 0.075 ml per 100 g body weight. Anesthesia was maintained with isoflurane during the entire surgery. Feathers overlaying the ears and on the scalp were trimmed and the pigeons were placed in a stereotaxic apparatus. The scalp was then cut and retracted to expose the skull. Six stainless steel screws (Small Parts, Logansports, USA) were subsequently placed into the skull to hold the implant. Two holes were drilled into the skull above the right and left MVL/entopallium. Coordinates were chosen according to the stereotactic atlas of the pigeon brain [39]. Coordinates of the exact electrode placement are given separately for each pigeon in Table 1 and the histological electrode track reconstruction is depicted in Fig. 3. The tips of the electrodes were then lowered into the brain to position them above the MVL. An additional hole was drilled into the skull to insert a 200 µm teflon-coated silver wire (Science Products GmbH, Hofheim, Germany) with its tip melted to a ball serving as ground for electrophysiological recordings. Dental acrylic was used to attach the microdrives to the screws and skull. The incision was sutured and covered with antibiotic ointment. Once fully alert and mobile, the pigeons were returned to the home cage and treated with analgesics (Carprofen, 10 mg/kg) for three consecutive days. After a recovery period with unlimited access to food for at least seven days, recording sessions started.

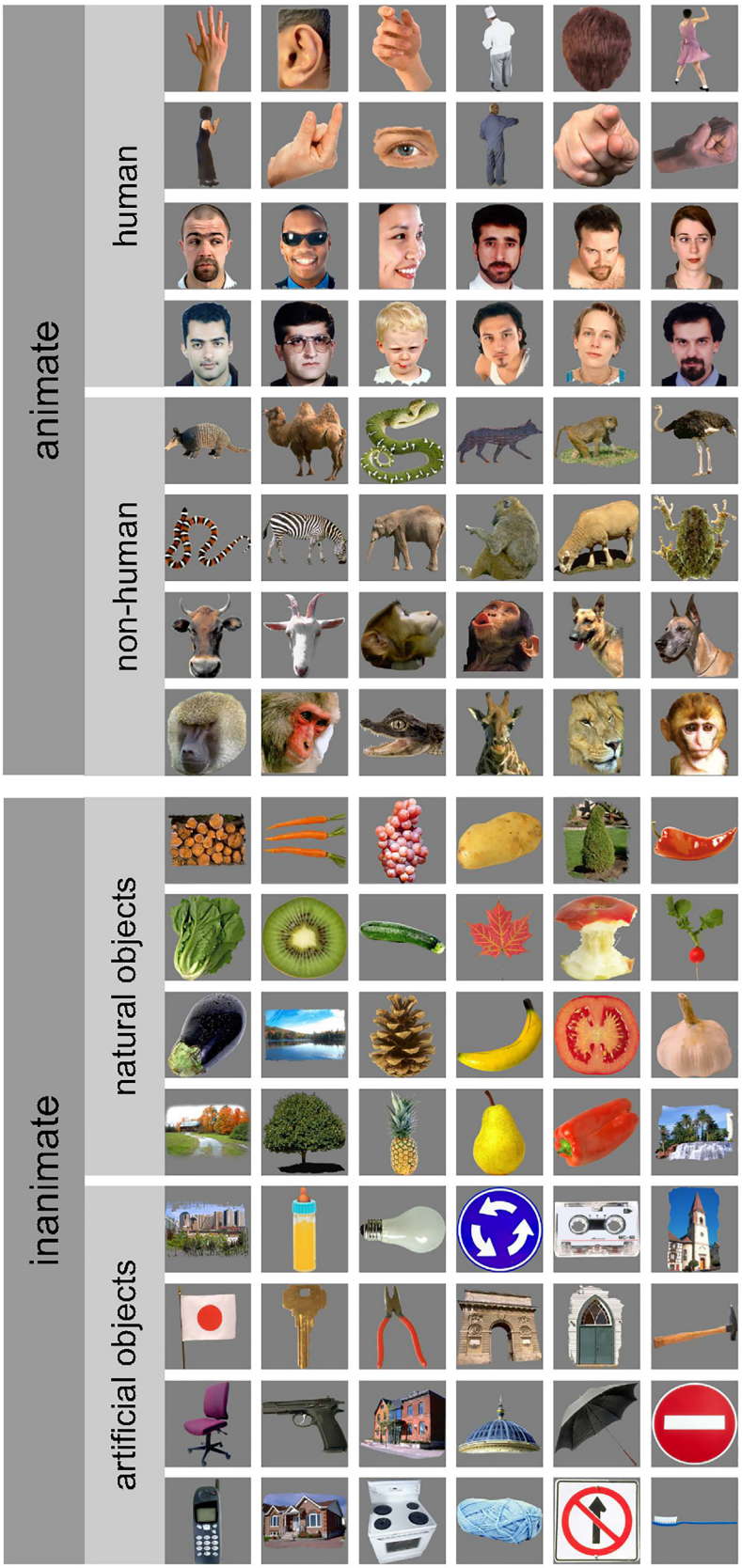

Fig. 1. The 96 visual stimuli used in the experiment. The stimuli are color photographs of isolated real-world objects on a gray background. They belong to different categories. At the coarsest level, they can be classified into animate and inanimate categories. These two classes, in turn, are further divided into different sub-categories as indicated in the figure. Reproduced with permission from [7].

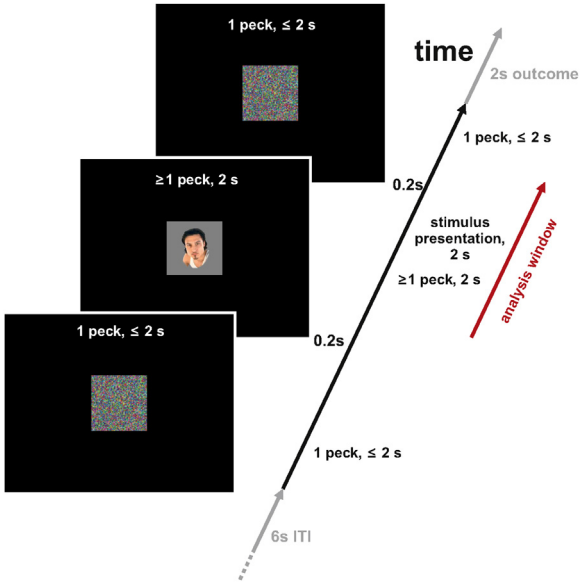

Fig. 2. Sequence of one trial. The initialization key was visible for up to 2 s and disappeared if pecked once. After a short delay (0.2 s) one of the 96 sample stimuli was shown for a fixed interval of 2 s in randomized order and had to be pecked at least once. At the time of stimulus presentation neuronal responses were analyzed. After another delay of 0.2 s the initialization key was shown again and had to be pecked once to finalize the trial. After pecking on all three stimuli, a food-reward was delivered and the inter-trial interval (6 s) followed.

Table 1
Number of cells recorded from MVL or entopallium (ENT) in the two hemispheres.

	Left	Right	Σ
MVL	32	4	36
ENT	6	29	35
Total	38	33	71

2.3.3. Neuronal recording

We recorded from both the right and the left hemisphere in four birds. In every session, we stored the output of eight electrodes, one serving as reference. Signals were fed through a miniature preamplifier (10×), subsequently amplified 1000× and filtered online (300 Hz high-pass filter, 5 kHz low-pass filter; Multi Channel Systems MCS GmbH, Reutlingen, Germany), and digitized using an analog-to-digital converter (sampling rate: 20 kHZ; Micro 1401 mkII, Cambridge Electronic Design, Cambridge, UK). Neuronal activity was recorded with Spike2 Version 7.06a (Cambridge Electronic Design, Cambridge, UK). Spike sorting was performed offline using amplitude thresholds for initial spike detection and a principal component analysis for manual sorting. Sorting quality was checked with custom-written Matlab code (The Mathworks, Natick, MA, USA) and the aid of the MLIB toolbox (written by Maik Stüttgen, available at Matlab central file exchange). A conservative approach to classify neuronal activity as originated from single units was adopted. Single units had to show (1) a clearly distinguishable cluster in principal component space, (2) show a symmetrical and unimodal distribution of peak waveform amplitudes without evidence of false negative classifications, (3) show no sign of overlapping multiple units in the waveform overlay and density plots, (4) have inter spike intervals larger than 1 ms, and (5) a signal-to-noise ratio of at least 2. The signal-to-noise ratio was calculated as the difference between the minimum and maximum of the averaged waveform, divided by the central 95% range in the noise distribution. Additionally, we visually checked for movement related artifacts (from e.g. key pecking and wing flapping) and by inspecting peri-peck time histograms for peaks or troughs of spiking activity near time point 0. For a more detailed description of materials, recording, and spike sorting procedure see [37,38].

2.3.4. Histology

For histological electrode track reconstruction, the pigeons were deeply anesthetized with Equithesin (4.5–5.5 ml/kg body weight) and transcardially perfused with physiological saline followed by 4% formaldehyde. 0.1 ml heparin was injected prior to anesthesia to prevent blood coagulation. Brains were embedded in gelatin and sectioned at 40 µm, Fig. 3. Slices were stained with cresyl violet. The position of the electrode tracks were analyzed under a microscope according to the pigeon brain atlas of [39].

2.3.5. Analysis of category information in neuronal populations

For every recorded single unit, we calculated the average number of spikes per stimulus presentation. For each stimulus, we collected the average spike numbers of all recorded neurons into a population vector **x**, even if neurons were recorded in different sessions. The population

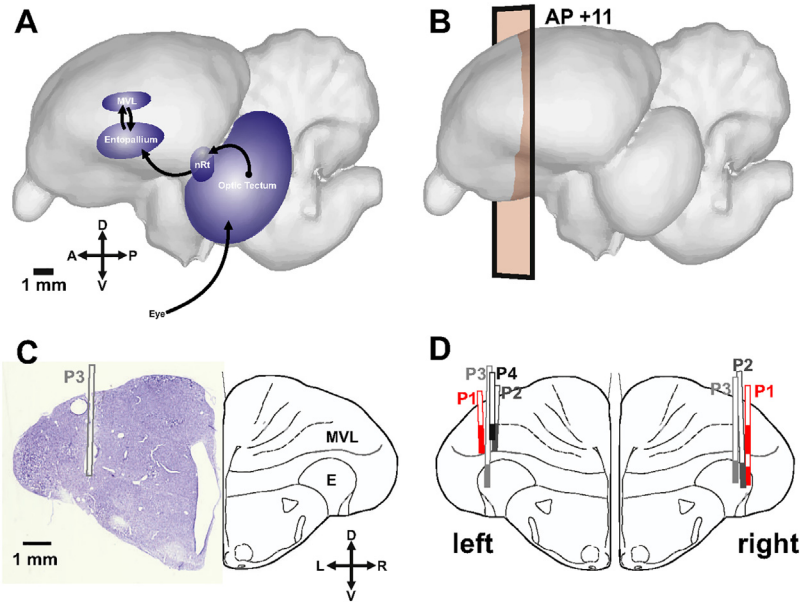

Fig. 3. Visual pathways in the pigeon brain and histological electrode track reconstruction. (A) Sagittal view of the pigeon brain [40]. The structures involved in visual processing are highlighted to give an overview of the visual pathways in the pigeon brain. nRt is nucleus rotundus. (B) Sagittal view of the pigeon brain including the coronal plane of the electrode positions. (C) Left: Coronal section of the plane highlighted in B (AP +11). In the photograph, one hemisphere is shown and the electrode track can be seen reaching into the entopallium. Right: Schematic drawing of the structures in this section of the brain. Drawings are based on the pigeon brain atlas by Karten and Hodos [39]. (D) Electrode track reconstruction for all pigeons used in this study collapsed on the AP +11 coronal section of the pigeon brain. Electrode tracks of different pigeons are indicated by unfilled bars and labelled with the corresponding pigeon number. The filled parts of the bars show the recording sites.

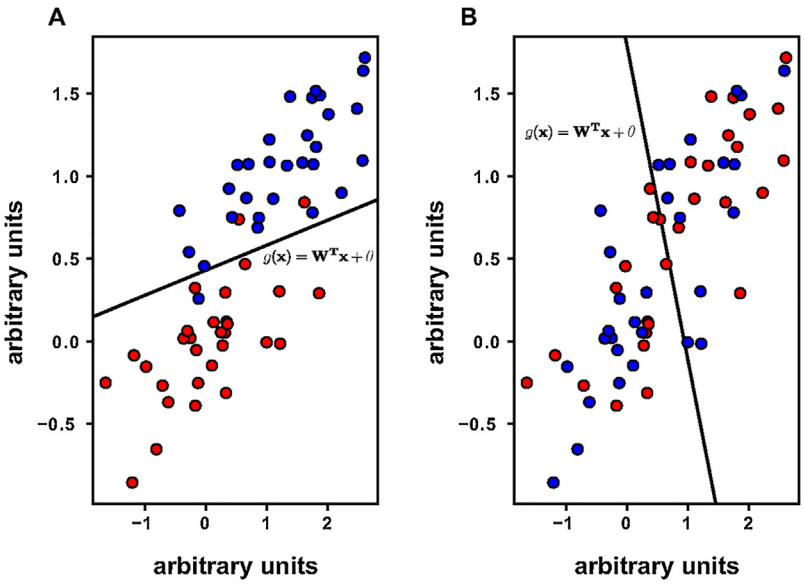

Fig. 4. Illustration of linear discriminant analysis (LDA). LDA is a linear classifier that assigns data points to two different classes based on a linear category boundary (black line). The classifier has to be trained with labelled data (the class label is indicated by the color of the data points). LDA will find a category boundary, regardless of whether the data are well segregated by a linear boundary (A), or not (B). The quality of the classifier is measured by the fraction of correctly classified data points (0.95 for A versus 0.58 for B).

vector is a window into the neural representation of that particular stimulus, which we focus on in this study. Our goal is to study whether the population vector of neurons in the pigeon brain represents categorical information. Specifically, we view each population vector **x** as a point in a high dimensional space, representing one visual stimulus, and ask whether the points can be segregated (classified) based on the category of the stimulus. As a classifier, we adopt linear discriminant analysis (LDA), which identifies a linear boundary between data points from two categories (Fig. 4). Since our stimulus set contained several (sub-)categories, we trained the LDA classifier on each category of interest, in comparison with all other stimuli that were not in that category, i.e., the labels for each population vector was either in-class or out-class.

In LDA, the boundary between categories is calculated as:

$$g(\mathbf{x}) = \mathbf{W}^T\mathbf{x} + \theta, \tag{1}$$

where **W** defines the direction of the linear boundary and θ its offset. A population vector belongs to the training category, if $g(\mathbf{x}) \geq 0$, otherwise it does not. LDA finds the linear function that yields the maximum ratio of between-class scatter to within-class scatter on the training data [41], i.e.,

$$J(\mathbf{W}) = \frac{|\widetilde{m}_1 - \widetilde{m}_2|^2}{\tilde{s}_1^2 + \tilde{s}_1^2} \tag{2}$$

Here, $\widetilde{m}_i$ is the mean for the projected data points of sample *i*, and $\tilde{s}_i^2 = \sum_{y \in Y_i} (y - \widetilde{m}_i)^2$ is the scatter for projected sample *i*. The vector **W** that maximizes *J* in Eq. (2) is

$$\mathbf{W} = \mathbf{S_W}^{-1}(\mathbf{m}_1 - \mathbf{m}_2) \tag{3}$$

Here $\mathbf{S_W} = \sum_{i=1}^{2} \sum_{\mathbf{x} \in D_i} (\mathbf{x} - \mathbf{m}_i)(\mathbf{x} - \mathbf{m}_i)^T$ is the scatter matrix. The solution for the threshold is

$$\theta = \frac{1}{2}(\mathbf{m}_1 + \mathbf{m}_2)\mathbf{S_W}^{-1}(\mathbf{m}_1 - \mathbf{m}_2) \tag{4}$$

The LDA classifier can learn to distinguish two categories at a time. To evaluate the performance of the LDA classifier on multiple categories, we calculated a confusion matrix [4]. We randomly assigned half of the population vectors to the training data, and the remaining population vector to the test data. We labeled the stimuli in the training data as in-class or out-class separately for each category, trained a LDA on this data, and then used that classifier to classify the population vectors in the test data. Test performance was evaluated separately for each category and quantified as the fraction of correct classifications. Since the results of this procedure depends on the (random) assignment of population vectors to training or test data, this entire procedure was repeated 1000 times, and the average classification score was calculated for all categories, which form the confusion matrices.

To determine the statistical significance of the classifier performance, we computed the classification score on permuted data, where the class labels of the images in the test sets were randomly permuted. As in this case no categorical information was present in the test set, the classifier should perform at the level of random guessing. When categories have an equal number of stimuli, the most naïve approach would be guessing with equal probability, so chance performance is 50%. If the number of stimuli in the two categories are different, the optimal strategy would be to guess each category according to the fraction of stimuli in that category, which is also chance performance. If, however, the neuronal activity represents categorical information, the classifier performance on the recorded data with correct class labels should be significantly better than the performance on the permuted data. We used the two-sample Kolmogorov–Smirnov test to compare the classifier performance between the two cases, real and permuted data, separately for each category ($\alpha = 0.01$).

3. Results

3.1. Histology, neuronal recordings, and visual responsiveness

We recorded from a total of 71 single units in 4 pigeons. The single units fired with a mean rate of 4.63 Hz (range: 0.2–27.3 Hz). Comparison of histological track reconstruction with the pigeon brain atlas confirmed the position of the electrodes in the anterior entopallium and the MVL (from 10.5 + AP to 11.25 + AP, Fig. 3). From depth records we reconstructed the recording sites in each pigeon and each hemisphere. According to this analysis 36 MVL single units and 35 entopallial single units were recorded. The number of recorded cells per hemisphere is summarized in Table 1.

Consistent with the visual function of entopallium and MVL, the recorded cells were activated by visual stimulation (Fig. 5). The instantaneous firing rate *r* of a representative cell increased when the visual stimulus was presented to the pigeon at $t = 0$ s (Fig. 5A). To verify the visual responsiveness of all of the recorded cells, we calculated the response index (RI),

$$\mathrm{RI} = \frac{|r(s) - r(b)|}{r(s) + r(b)}, \tag{5}$$

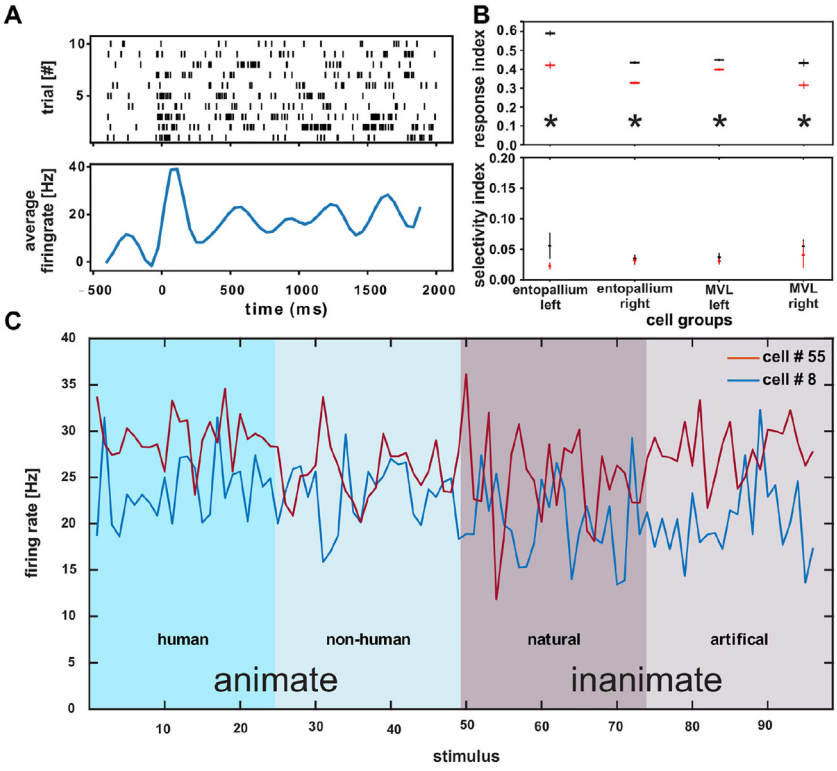

Fig. 5. Individual cells were responsive to visual stimulation, but not selective for categories. (A) Post-stimulus time histogram of a typical cell. The response of this cell to a stimulus was recorded in 10 different trials. Each tick-mark shows one spike of the cell. The spike-density function of the cell, smoothed with a Gaussian kernel ($\sigma = 60$ ms) is shown in the lower panel. (B), top panel The response index (RI) of the cells in each of the recorded areas. The cells in all regions exhibit a significant RI around stimulus onset (black marks) as compared to the RI in the ITI, when there was no visual stimulus (red marks) (p-values: 1.08×10^{-17}, 3.93×10^{-34}, 4.39×10^{-10}, 1.35×10^{-5}, Wilcoxon rank-sum test). Bottom panel The average selectivity index (SI) of the cells in each of the recorded areas (black marks). None of the SI is significantly different from a random shuffle (red marks), where the animate and inanimate labels have been randomly permuted (p-values: 0.52, 0.87, 0.67, 0.39, Wilcoxon rank-sum test). Vertical bars show the s.e.m. (C) The response of two selected cells. The average firing rate does not distinguish between different stimulus categories. Blue regions mark animate stimuli (dark blue for human images and light blue for non-human images) and the purple regions mark inanimate stimuli (dark purple for natural and light purple for artificial objects). (For interpretation of the references to color in this legend, the reader is referred to the web version of the article.)

where $r(b)$ and $r(s)$ are the average firing rates of a given cell in the time-window of 200 ms before and after the stimulus onset across trials. As a control, we calculated RI for two time windows in the ITI during which there was no stimulus presented on the screen. The population of cells in the entopallium or in the MVL in both hemispheres show significantly higher RI around stimulus onset as compared to the ITI control (Fig. 5B, top panel), suggesting that cells in all recorded areas were responsive to visual stimulation.

Even though the activity levels of cells were modulated by visual stimulation, their mean activity levels did not differ systematically between categories. For the coarsest category distinction, i.e., animate versus inanimate stimuli, we recorded 29,065 neural responses to animate and 29,054 responses to inanimate objects (total 58,119). The firing rate of cells were indistinguishable when averaged over all recorded cells (4.38 Hz for animate vs. 4.36 Hz for inanimate objects, $p = 0.52$; paired t-test). Similarly, The average firing rates of cells in entopallium and MVL were similar (4.76 Hz for entopallium and 4.14 Hz for MVL, $p = 0.51$; paired t-test). We next asked whether individual cells are selective for the animate vs. inanimate distinction (e.g., Fig. 5C). To examine this systematically, we calculated the selectivity index (SI) for each cell,

$$\mathrm{SI} = \frac{|\bar{r}_{\mathrm{anim}} - \bar{r}_{\mathrm{inanim}}|}{\bar{r}_{\mathrm{anim}} + \bar{r}_{\mathrm{inanim}}}. \quad (6)$$

$\bar{r}_{\mathrm{anim}}$ and $\bar{r}_{\mathrm{inanim}}$ are the firing rates of a cell in response to 2 s of stimulus presentation, averaged over animate or inanimate stimuli, respectively. As a control, the selectivity index was calculated on the same neural responses with the animate and inanimate labels randomly shuffled. These selectivity indices in all four regions were not significantly different from the random shuffle (Fig. 5B, bottom panel). So neither entopallial nor MVL cells seem to carry category individually.

3.2. Population activity reveals category information

While we found no evidence of category information in individual cells, it might be evident in the population response. We therefore used linear discriminant analysis (LDA) to analyze the population activity. We trained a linear classifier on half of the labeled data, and tested the performance on the other half of the data. When analyzing population responses of entopallial cells, LDA did not classify the stimuli better than chance for any (sub-)category (Fig. 6). The performance on the test data is summarized in the confusion matrix (Fig. 6A). The diagonal cells show the fraction of correct classifications (indicated by the orange box), and the off-diagonal cells the fraction of incorrect classifications. The performance on the animate vs. inanimate classification of 0.62 was not significantly higher than expected by chance. Significance was evaluated by comparing the classification performance to the distribution of the classification scores on 1000 randomly labeled test sets (Fig. 6B). Here, the control classifier scores were calculated using correctly labeled data as the training set and randomly labeled ones as the test set for the LDA. By contrast, the population responses of MVL cells significantly distinguished between animate and inanimate categories based on LDA with a hit rate between 0.68 and 0.70 (Fig. 6). So, category information was more accessible in MVL than in the processing stage below, in the entopallium. This difference cannot be explained by a difference in visual responsiveness of the two areas since cells in the entopallium exhibited similar, if not higher, responsiveness scores than cells in the MVL (Fig. 5B). Another way to show the difference in classification performance between MVL and entopallium is by looking at the receiver operating characteristic (ROC). Each point in this graph corresponds to randomly splitting the data into training and test sets (Fig. 6C, F). The ROC curve for the MVL is further away from the diagonal, which represents random guessing, than the ROC curve for the entopallium (area under curve: 0.68, 0.75 for entopallium and MVL, respectively).

It is conceivable that the stimuli in the animate and inanimate categories have many different low-level features, and that the MVL cells were sensitive to these differences (Fig. 1). However, Kriegeskorte et al. [7], from whom we obtained the stimulus material, confirmed that low-level features in the stimuli such as, e.g., color and image resolution, cannot distinguish on their own between animate and inanimate categories (cf. their supplementary figures S6 and S7). It is thus more likely that the MVL population represents more abstract features of the

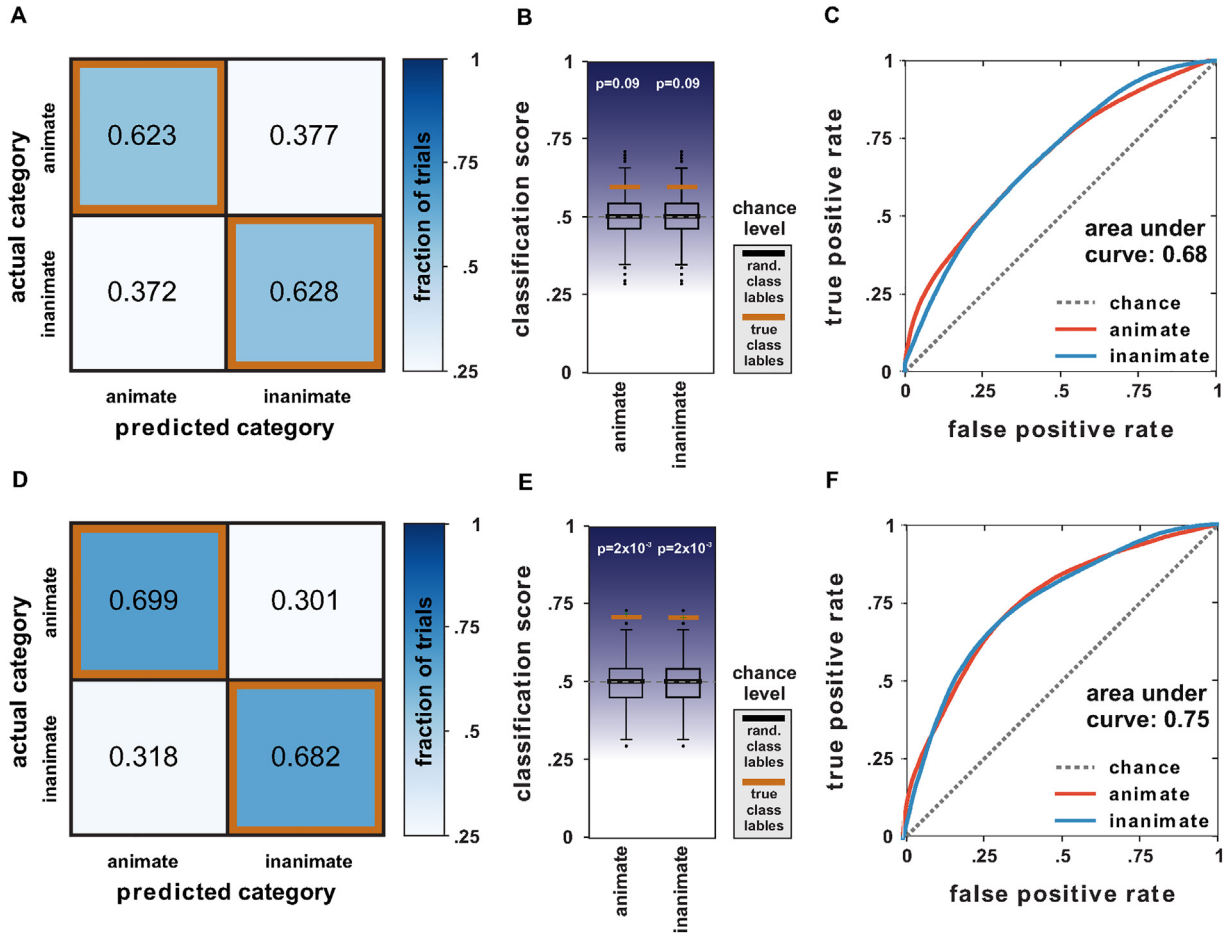

Fig. 6. The population response of MVL, but not entopallial cells revealed category information in a linear discriminant analysis. (A) The confusion matrix of the classifier when distinguishing between animate and inanimate objects, using the population response of recorded neurons in the entopalium. Diagonal cells indicate rates of correct classifications (indicated by the orange box), off-diagonal cells the rates of incorrect classifications. (B) The classification score for the classifier on the entopallial population and correctly labeled test data is marked with an orange bar. The distribution of classification scores on randomly labeled test sets is shown by the box-whisker plot. If the orange bar is clearly outside the range indicated by the whiskers, the performance of the classifier is significantly better than chance. In this case, the classification performance is not significant. (C) The receiver operating characteristic curve for entopallium on the animate and inanimate categorization. Area under curve: 0.68. The diagonal dashed line represents random guessing. (D) The confusion matrix of the classifier when distinguishing between animate and inanimate objects, using the population response of recorded neurons in MVL. Diagonal cells indicate rates of correct classifications (indicated by the orange box), off-diagonal cells the rates of incorrect classifications. (E) The performance of the classifier on the MVL population is significantly better than chance. Plotting convention as in B. (F) The receiver operating characteristic curve for MVL on animate and inanimate categorization. Area under curve: 0.75. (For interpretation of the references to color in this legend, the reader is referred to the web version of the article.)

stimuli that allow LDA to classify the stimuli.

To further analyze the nature of this information, we subdivided the stimuli in the animate category into human and non-human sub-categories, and those in the inanimate category into artificial and natural sub-categories. We then applied the binary LDA to these sub-categories (Fig. 7). Only the human subcategory can be distinguished from the other subcategories when using MVL population activity, and the selectivity of MVL cells to stimuli in the human subcategory are responsible for the animate vs. inanimate distinction (Fig. 7). The average firing rate of the cells in response to stimuli in the human category was not significantly different from the responses to the other stimuli (4.36 Hz for human, 4.40 Hz for the other stimuli, $p = 0.25$; two-sample t-test). Hence, the information of the human category was only apparent in the population response, not at the single cell level.

To directly test whether the classification of animate vs. inanimate objects was solely based on the classification of the human subcategory, we repeated the animate vs. inanimate classification without the stimuli of the human subcategory (Fig. 8, “non-human vs. inanimate”, last column). The classifier could then no longer significantly distinguish between the remaining animate (i.e., non-human) and inanimate stimuli. This shows that the features that are readable by a simple linear classifier are expressed in the responses of the cells to images of human faces and human body-parts. The limited number of stimuli per category do not allow us to further subdivide our stimuli to pinpoint more precisely what category feature the population responds to.

It is conceivable that the difference in the representation of category information between MVL and entopallium arises from the random sampling of recorded neurons, which happened to select more, and/ or more strongly, category-coding neurons in MVL than in the entopallium. This scenario is not highly likely since the number of recorded neurons in MVL and entopallium were well balanced in our study ($n = 36$ and $n = 35$, respectively). Nevertheless, as the number of recorded cells was somewhat small in both regions, we performed a subsampling analysis to examine whether the entopallial population reflected any category information that does not rise to the level of significance in our analysis. If this were the case, we would expect that including more and more cells in the analysis would make the analysis more and more probable to yield a significant classification. Specifically, we randomly subsampled different numbers of neurons from the recorded populations and repeated our classification analyses

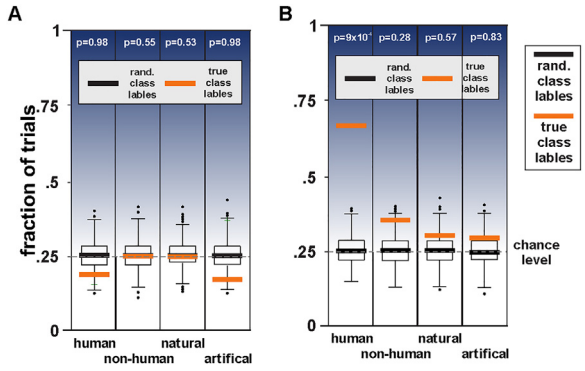

Fig. 7. Classification of subcategories is apparent in MVL, but not in the entopallium. The performance of the classifier when using true-labeled data in four subcategories (red bars) compared to the distribution of classification scores on data with shuffled labels. (A) For entopallial cells, the classifier did not perform better than the chance level for any subcategory. (B) For MVL cells, linear discriminant analysis could significantly distinguish the human subcategory. (For interpretation of the references to color in this legend, the reader is referred to the web version of the article.)

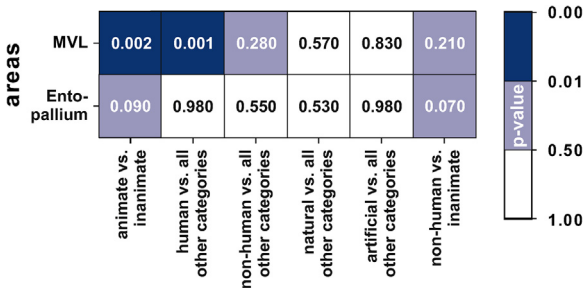

Fig. 8. Summary of the significance of classification based on different categories and for different neural populations. The numbers are the *p*-values of the shuffle test, the background color of each cell indicates three ranges of the *p*-value as shown in the color bar. For MVL cells, the classifier could significantly distinguish between animate and inanimate objects. By contrast, entopallial cells did not perform significantly better than chance on any classification. Stimuli in the human subcategory accounted for the animate–inanimate classification in the MVL population. No other subcategory within the animate and inanimate categories lead to a significant classification. Also, when excluding images of human faces and body parts from the animate category, the classification (non-human vs. inanimate) failed.

(Fig. 9). For each subsample size, the fraction of 300 random sampling that lead to significant classification was calculated. The subsampling analysis revealed that, for the entopallial population, the fraction of significant classification was constant at a low value as the subsample size increased (Fig. 9, blue line), suggesting that the entopallial population indeed did not contain category information that LDA can extract. By contrast, the subsampling analysis on the MVL population revealed that the likelihood of significant classification clearly increased monotonically with subsample size (Fig. 9, orange line). While this result was expected, if at least a subpopulation of MVL neurons encode category information, the subsampling analysis worked surprisingly well for the small total number of MVL cells in our recordings.

3.3. Comparing category coding in MVL and entopallium to a model of simple cells

Little is known about the functional selectivity of cells along the avian tectofugal pathway [42]. In an attempt to better understand the neuronal responses that we have recorded in the MVL and entopallium, we compared our experimental results with a computational model of visual processing in the pigeon inspired by the one proposed by Soto

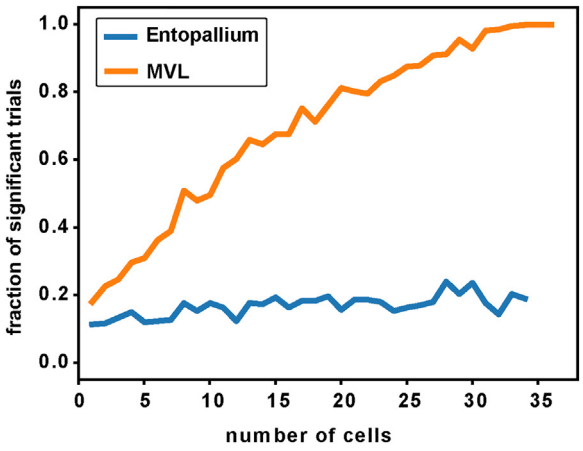

Fig. 9. Subsampling analysis of classification performance in MVL and entopallium. Shown is the fraction of analysis trials that reach significance ($\alpha = 0.01$) as a function of the number of cells included in the subsample. For each subsample, cells are drawn randomly from the indicated population. Classification scores are assessed on the subsample in the same way as on the complete dataset. This analysis is repeated 300 times for each subsample size, which allows for the calculation of the fraction of analysis trials that yield significant category information that LDA can extract. MVL population (orange line). Entopallial population (blue line). (For interpretation of the references to color in this legend, the reader is referred to the web version of the article.)

and Wasserman [43]. In this model, the function of the first stages of processing are comparable to the responses of simple cells in the primary visual cortex of primates. A common model for their response function has been Gabor filters [44,45] (see Fig. 10).

If Gabor filtering is an applicable model for visual processing in the MVL and entopallium, or downstream of them, we would expect that a LDA on the output of a bank of Gabor filters applied to the visual stimuli used in our experiment would produce similar classification results as observed in our experiment. We used a bank of 68 Gabor filters with different orientation and frequency. Each stimulus was convolved with these filters to obtain a representation of the stimulus that accentuates the edges of the image (Fig. 10C). We then calculated the sum of the absolute values of all pixels in the results of these convolutions. This sum represents the response of a single model cell, which might, or might not, bear similarity to MVL or entopallial responses. The vector of 68 responses was treated and analyzed in exactly the same way in which we analyzed the experimentally recorded population responses (Fig. 11).

To our surprise, the performance of the classifier on these model cell responses was similar to that on MVL responses for the animate–inanimate classification (Fig. 11A and B). However, when examining the performance of the model cells on the subcategories, we found that the model cells significantly code for the subcategories “humans” as well as “artificial objects”. In addition, the classification score was much higher for the artificial category than for the other subcategories (Fig. 11C and D). This pattern contrasts with the experimental data on the MVL population (Fig. 8). So, the Gabor filter model accounts for the coarser distinction between animate and inanimate categories in MVL cells, but differs in performance on the more fine-grained subcategorization (human, non-human, natural objects, and artificial objects). In other words, the Gabor filter model captures some aspect, but not the details, of the computation giving rise to MVL responses. Our results therefore indicate that MVL cells are not merely summing the output of edge detectors, and that different or additional functional processing must occur in the visual system of the pigeon leading up to, and/ or in the, MVL.

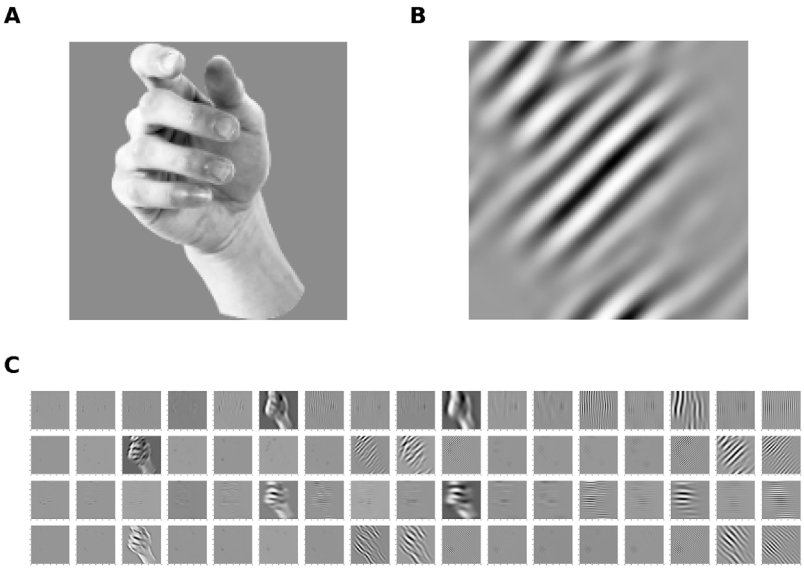

Fig. 10. Illustration of Gabor filters in image processing. (A) An example of an image. (B) The result of the convolution between the image and a Gabor filter. (C) A bank of 68 Gabor filters, each with different orientation and frequency, was used as a model of visual processing by simple cells. Shown are the results of the convolutions between these Gabor filters and the image depicted in (A).

4. Discussion

We found categorical representations of complex visual stimuli in a visual association area of the pigeon brain. Pigeons were exposed to photographs depicting real-world objects while single-unit activity was recorded in two visual forebrain structures of the tectofugal pathway: the entopallium and the mesopallium ventrolaterale (MVL), a visual associative area. Importantly, the birds only had to perceive the stimuli and categorization was not required in the task. In our experiment, response rates of individual neurons did not significantly distinguish between visual categories. In the entopallium, there also was no distinction between categories at a population level, at least none that could be extracted using linear discriminant analysis (LDA). In contrast, the MVL population showed significant categorization along the animate/inanimate border, mainly driven by the category "human". The reliability of correct categorization increased monotonically with the number of MVL neurons such that only tens of neurons are sufficient to build up reliable categorical object representation.

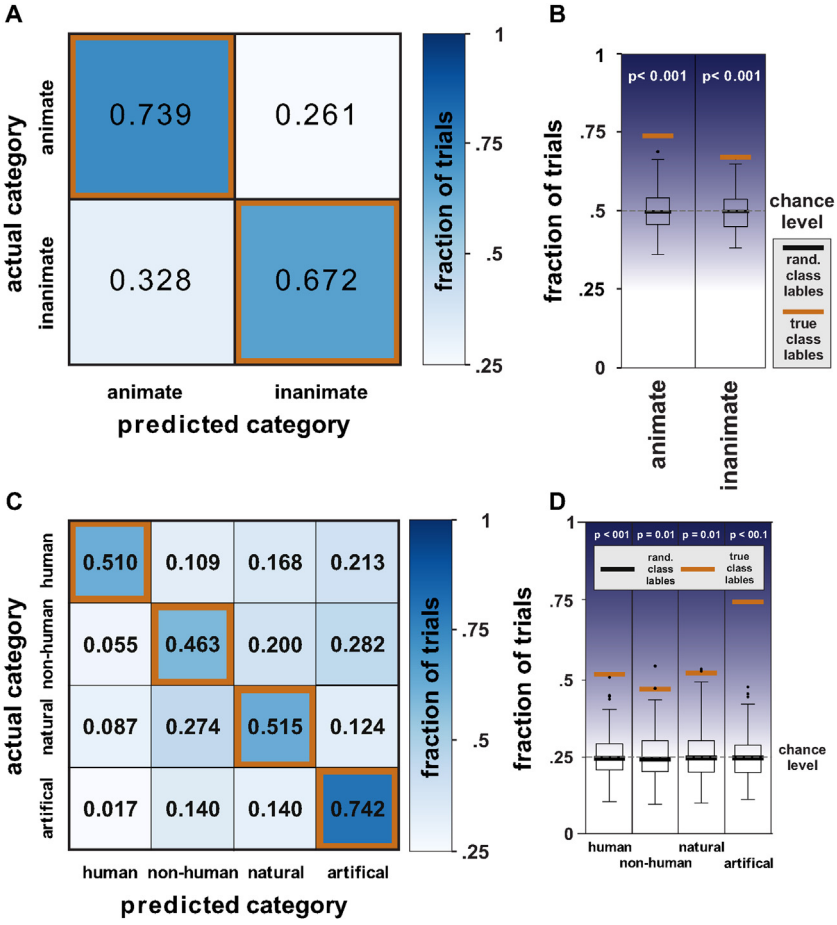

Fig. 11. Classification performance of a model based on Gabor filters. (A) Confusion matrix for the LDA classification between animate and inanimate categories. (B) Comparison to a distribution of classification score on randomly labelled data shows that the classification based on model cell responses is significantly better than chance. (C, D) Classification performance on subcategories is significantly better than chance for the subcategories "humans" and "artificial objects". (For interpretation of the references to color in this legend, the reader is referred to the web version of the article.)

4.1. Implications of the LDA results for neural coding in MVL and entopallium

The fact that LDA cannot classify the population responses in the entopallium does not necessarily mean that there is no category information. Category information is obviously present in the retina and all early stages of visual processing, even if it cannot be extracted from the entopallial cells using a linear classifier. So, what does our finding that LDA reveals category information in MVL, but not in the entopallium, tell us about neural coding in these areas?

The visual system in vertebrate brains is hierarchical. Canonical neural computations repeatedly applied in different stages of the visual hierarchy enable higher areas to perform more sophisticated operations on the input data than lower areas can [44]. In other words, each cell in the hierarchical layers of the visual system pools the activity of many cells in the lower input layer. The visual system does this, supposedly, to recode the same information in a different format that can be read out by downstream neurons. So, it is not so much the information that is contained in the population *per se*, but rather the neural coding of that information that matters. In this light, what we found is a hint that the MVL and entopallium employ different codes for representing visual information and that the code employed by MVL, but not of entopallium, makes it possible to read out categorical information easily, i.e., using LDA. Furthermore, we note that pigeons in our experiment were not required to learn the categories of the stimuli. It is conceivable that learning would change the category coding properties of the entopallium and MVL, and that one would find different results from the reported ones in a learning task.

4.2. "Humans" as a potential special category for laboratory pigeons

Why might MVL cells encode the human category, even though the animals were not required to learn any categories in our study? Since the pigeons in our study had rich experience with humans handling them, at least parts of the "human"-specific coding in MVL recordings might have resulted from experience with humans. Consequently, the absence of population responses to the other categories could be partly due to the absence of respective experiences with these objects. Aust and Huber [46] have suggested that pigeons' representation of humans was shaped by visual experience with live humans and their constituent parts by investigating two groups of pigeons that had different pre-experience with live humans. Pigeons therefore likely recognize the relation between pictures of humans and their 3-D-referents ([47,48] but see [49]). When pigeons were conditioned to discriminate between hundreds of photographs of which some depicted humans, they transferred this knowledge to novel photographs [9], demonstrating that pigeons acquired the category "human". Subsequent studies confirmed that indeed human features had driven the categorization process [50,51]. Furthermore, Dittrich et al. [52] found that pigeons mostly pecked on the head of human figures in a people-present/people-absent task and that removal of the heads impaired performance, while removal of other parts of the human figures did not. These studies taken together indicate to us that pigeons in our study might have formed a representation of the humans who handle them and human body parts in 3-D, and that this representation generalized to the new 2-D visual stimuli presented during our experimental study even though it was not required by the task.

4.3. Comparison of the avian and mammalian visual systems

In birds, the outer tectal layers constitute a two-dimensional map of the visual field, as transmitted from the contralateral retina [53]. The tectum is composed of 15 layers, with the tectal receptive field widths gradually increasing from the superficial to the deep layers [54]. In contrast to this retinotopic organization, the thalamic n. rotundus as the next tectofugal entity is characterized by functional domains in which different visual features such as movement, color, or luminance are processed in parallel [55]. Thus, a precise topographic tectal code is translated into a functionotopic thalamic code at the mesopallial/diencephalic junction. This re-organization appears to be achieved by five different retinotopically organized tectal cell populations that project in parallel to different rotundal and subsequently also to entopallial domains [56,57]. Consequently, the entopallium is also constituted by functional domains that possibly still possess visual topography. The entopallium itself is subdivided into different layer-like components that partly project to different associative telencephalic regions of which the MVL is one [31]. Neurons of the MVL have been found to code for simple object features like motion and color [32].

The mammalian extrageniculocortical pathway (colliculo-pulvino-extrastriate system) is also constituted by parallel and functionally diverse thalamo-cortical projections, although not much is known how the retinotopic collicular is translated to a functionotopic thalamic system [58,59]. The major difference between the pigeon and the mammalian visual system is the functional embedding of the two ascending visual pathways to the forebrain. The pigeons tectofugal pathway (homologous to the mammalian extrageniculocortical pathway) primarily processes detailed object vision [21], while this is the domain of the mammalian geniculocortical pathway. Despite this major difference, our results open the possibility that avian and mammalian visual pathways may process category-relevant object vision in a similar way at a higher processing level. Let us first outline category coding in the primate temporal cortex.

Studies in human subjects show that object classification is achieved in about 150 ms or less [60–62]. Such a fast processing time likely requires feed-forward propagation to higher regions within the ventral visual pathway [63]. Along this stream, receptive fields of neurons increase in size and cells respond to more complex features [64,65], which implies that lower stages of the ventral visual stream must represent a large dynamic range of stimulus features. In higher stages like the ITC, only binary information of presence or absence of a stimulus features need to be represented [66].

Recent studies in primates show that populations of ITC cells rapidly converge towards a population vector that distinguishes between categories of object features [67]. This process is extremely fast such that populations of ITC-cells contain information about a certain object category just 50 ms after arrival of the spike front in temporal cortex [68]. For example, in the study of [4] a linear classifier that was trained with the data from the recorded cortical cells successfully categorizes all stimuli with extremely high accuracy. Similar to our observation in MVL, categorization performance also increased linearly with the number of recorded neurons such that just 100 cells were sufficient to successfully categorize all stimuli [4]. This is strong evidence that the ventral visual pathway supports bottom-up visual object categorization [69].

This finding resembles ours in pigeons and is also reminiscent to a recent study in which a representational dissimilarity analysis of recorded neurons from the forebrain area NFL revealed population coding for some low-level features like color and spatial frequency [33]. If indeed neurons in the various visual associative areas of the avian forebrain processed stimuli as a set of representational elements, stimuli that share these elements could be perceived as perceptually coherent and would define a category. Indeed, pigeons spontaneously group related stimuli, even when such grouping is unrelated to the prevailing reward contingencies [70]. In categorization experiments, learning takes place when category-specific elements acquire control over behavior. This is usually achieved by rewarding the animals for choosing the category-relevant stimuli [71]. Based on an error-driven learning rule in this supervised learning account, the critical elements of the rewarded stimuli would acquire associative value, and thus the ability to faithfully predict an outcome [72].

To summarize, we showed that a small neuronal population in a visual association area of pigeons contains easily accessible information

about the category “human”, whereas a lower visual area did not. Since our experimental approach prevented reward contingencies from influencing the neuronal responses, it is likely that this population code was mainly driven by feed-forward processes within the tectofugal pathway. Thus, despite the important anatomical differences in the organization of avian and mammalian forebrains, the encoding of categorical information in associative neurons in pigeons might arise in a similar fashion as in monkeys and humans [7,73].

Acknowledgements

We thank Nikolaus Kriegeskorte for kindly providing the visual stimuli that were used in this study. This work was supported by grants from the German Research Foundation (DFG) through the SFB 874, projects B2 (S.C.) and B5 (O.G.). We thank Alexis Garland for critically reading a previous version of the manuscript.

References

[1] H. Cohen, C. Lefebvre, Handbook of Categorization in Cognitive Science, Elsevier Science, 2005 ISBN 9780080457413, URL https://books.google.de/books?id=5WDfl14RgKMC.
[2] E. Wasserman, R.E. Kiedinger, R.S. Bhatt, Conceptual behavior in pigeons: categories, subcategories, and pseudocategories, J. Exp. Psychol. Anim. Behav. Process. 14 (3) (1988) 235–246, http://dx.doi.org/10.1037/0097-7403.14.3.235.
[3] D.J. Freedman, M. Riesenhuber, T. Poggio, E.K. Miller, Experience-dependent sharpening of visual shape selectivity in inferior temporal cortex, Cereb. Cortex 16 (11) (2006) 1631–1644, http://dx.doi.org/10.1093/cercor/bhj100.
[4] C.P. Hung, G. Kreiman, T. Poggio, J.J. DiCarlo, Fast readout of object identity from macaque inferior temporal cortex, Science 310 (5749) (2005) 863–866, http://dx.doi.org/10.1126/science.1117593.
[5] R. Kiani, H. Esteky, K. Mirpour, K. Tanaka, Object category structure in response patterns of neuronal population in monkey inferior temporal cortex, J. Neurophysiol. 97 (6) (2007) 4296–4309, http://dx.doi.org/10.1152/jn.00024.2007.
[6] A.G. Huth, S. Nishimoto, A.T. Vu, J.L. Gallant, A continuous semantic space describes the representation of thousands of object and action categories across the human brain, Neuron 76 (6) (2012) 1210–1224.
[7] N. Kriegeskorte, M. Mur, D.A. Ruff, R. Kiani, J. Bodurka, H. Esteky, K. Tanaka, P.A. Bandettini, Matching categorical object representations in inferior temporal cortex of man and monkey, Neuron 60 (6) (2008) 1126–1141, http://dx.doi.org/10.1016/j.neuron.2008.10.043.
[8] O.F. Lazareva, E.A. Wasserman, Categories and concepts in animals, in: J.H. Byrne (Ed.), Learning and Memory: A Comprehensive Reference, Volume One: Learning Theory and Behavior, 2nd ed., Elsevier, Amsterdam, 2017, pp. 111–139 (Chapter 7).
[9] R.J. Herrnstein, D.H. Loveland, Complex visual concept in the pigeon, Science 146 (3643) (1964) 549–551.
[10] D.F. Kendrick, A.A. Wright, R.G. Cook, On the role of memory in concept learning by pigeons, Psychol. Rec. 40 (3) (1990) 359–371.
[11] M.A. Qadri, R.G. Cook, Pigeons and humans use action and pose information to categorize complex human behaviors, Vision Res. 131 (2017) 16–25.
[12] R.M. Levenson, E.A. Krupinski, V.M. Navarro, E.A. Wasserman, Pigeons (*Columba livia*) as trainable observers of pathology and radiology breast cancer images, PLoS ONE 10 (11) (2015) e0141357.
[13] S. Watanabe, J. Sakamoto, M. Wakita, Pigeons’ discrimination of paintings by Monet and Picasso, J. Exp. Anal. Behav. 63 (2) (1995) 165–174.
[14] D. Scarf, K. Boy, A. Uber Reinert, J. Devine, O. Güntürkün, M. Colombo, Orthographic processing in pigeons (*Columba livia*), Proc. Natl. Acad. Sci. U. S. A. 113 (40) (2016) 11272–11276, http://dx.doi.org/10.1073/pnas.1607870113.
[15] J. Harvey, Karten, Vertebrate brains and evolutionary connectomics: on the origins of the mammalian “neocortex”, Philos. Trans. R. Soc. Lond. B: Biol. Sci. 370 (1684) (2015), http://dx.doi.org/10.1098/rstb.2015.0060.
[16] J. Dugas-Ford, J.J. Rowell, C.W. Ragsdale, Cell-type homologies and the origins of the neocortex, Proc. Natl. Acad. Sci. U. S. A. 109 (42) (2012) 16974–16979, http://dx.doi.org/10.1073/pnas.1204773109.
[17] O. Gntrkn, T. Bugnyar, Cognition without cortex, Trends Cogn. Sci. (Regul. Ed.) 20 (4) (2016) 291–303.
[18] O. Güntürkün, F. Strckens, D. Scarf, M. Colombo, Apes, feathered apes, and pigeons: differences and similarities, Curr. Opin. Behav. Sci. 16 (2017) 35–40, http://dx.doi.org/10.1016/j.cobeha.2017.03.003.
[19] M. Remy, O. Güntürkün, Retinal afferents to the tectum opticum and the nucleus opticus principalis thalami in the pigeon, J. Comp. Neurol. 305 (1) (1991) 57–70.
[20] O. Güntürkün, U. Hahmann, Functional subdivisions of the ascending visual pathways in the pigeon, Behav. Brain Res. 98 (2) (1999) 193–201, http://dx.doi.org/10.1016/S0166-4328(98)00084-9.
[21] H. Mouritsen, D. Heyers, O. Güntürkün, The neural basis of long-distance navigation in birds, Annu. Rev. Physiol. 78 (2016) 133–154.
[22] D. Scarf, M. Stuart, M. Johnston, M. Colombo, Visual response properties of neurons in four areas of the avian pallium, J. Comp. Physiol. A Neuroethol. Sens. Neural. Behav. Physiol. 202 (3) (2016) 235–245.
[23] J. Engelage, H.-J. Bischof, Single-cell responses in the ectostriatum of the zebra finch, J. Comp. Physiol. A 179 (6) (1996) 785–795, http://dx.doi.org/10.1007/BF00207357.
[24] Y. Gu, Y. Wang, T. Zhang, S.R. Wang, Stimulus size selectivity and receptive field organization of ectostriatal neurons in the pigeon, J. Comp. Physiol. A Neuroethol. Sens. Neural. Behav. Physiol. 188 (3) (2002) 173–178.
[25] Q. Xiao, D.P. Li, S.R. Wang, Looming-sensitive responses and receptive field organization of telencephalic neurons in the pigeon, Brain Res. Bull. 68 (5) (2006) 322–328.
[26] Q. Xiao, B.J. Frost, Looming responses of telencephalic neurons in the pigeon are modulated by optic flow, Brain Res. 1305 (2009) 40–46.
[27] J. Verhaal, J.A. Kirsch, I. Vlachos, M. Manns, O. Güntürkün, Lateralized reward-related visual discrimination in the avian entopallium, Eur. J. Neurosci. 35 (8) (2012) 1337–1343.
[28] M. Johnston, C. Anderson, M. Colombo, Neural correlates of sample-coding and reward-coding in the delay activity of neurons in the entopallium and nidopallium caudolaterale of pigeons (*Columba livia*), Behav. Brain Res. 317 (2017) 382–392.
[29] M. Johnston, C. Anderson, M. Colombo, Pigeon NCL and NFL neuronal activity represents neural correlates of the sample, Behav. Neurosci. 131 (3) (2017) 213–219.
[30] S.A. Husband, T. Shimizu, Efferent projections of the ectostriatum in the pigeon (*Columba livia*), J. Comp. Neurol. 406 (3) (1999) 329–345.
[31] N.O. Krtzfeldt, J.M. Wild, Definition and novel connections of the entopallium in the pigeon (*Columba livia*), J. Comp. Neurol. 490 (1) (2005) 40–56.
[32] M. Stacho, F. Strockens, Q. Xiao, O. Güntürkün, Functional organization of telencephalic visual association fields in pigeons, Behav. Brain Res. 303 (2016) 93–102.
[33] C. Koenen, R. Pusch, F. Brker, S. Thiele, O. Güntürkün, Categories in the pigeon brain: a reverse engineering approach, J. Exp. Anal. Behav. 105 (1) (2016) 111–122, http://dx.doi.org/10.1002/jeab.179.
[34] J. Rose, T. Otto, L. Dittrich, The biopsychology-toolbox: a free, open-source Matlab-toolbox for the control of behavioral experiments, J. Neurosci. Methods 175 (1) (2008) 104–107, http://dx.doi.org/10.1016/j.jneumeth.2008.08.006.
[35] D.K. Bilkey, G.M. Muir, A low cost, high precision subminiature microdrive for extracellular unit recording in behaving animals, J. Neurosci. Methods 92 (1) (1999) 87–90, http://dx.doi.org/10.1016/S0165-0270(99)00102-8.
[36] D.K. Bilkey, N. Russell, M. Colombo, A lightweight microdrive for single-unit recording in freely moving rats and pigeons, Methods 30 (2) (2003) 152–158, http://dx.doi.org/10.1016/S1046-2023(03)00076-8.
[37] S. Starosta, O. Güntürkün, M.C. Stttgen, Stimulus–response-outcome coding in the pigeon nidopallium caudolaterale, PLOS ONE 8 (2) (2013) 1–11, http://dx.doi.org/10.1371/journal.pone.0057407.
[38] S. Starosta, M.C. Stttgen, O. Güntürkün, Recording single neurons’ action potentials from freely moving pigeons across three stages of learning, J. Vis. Exp. 88 (2014) 51283.
[39] H. Karten, W. Hodos, A Stereotaxic Atlas of the Brain of the Pigeon (Columbia livia), Johns Hopkins Press, Baltimore, 1967.
[40] O. Güntürkün, M. Verhoye, G. De Groof, A. Van der Linden, A 3-dimensional digital atlas of the ascending sensory and the descending motor systems in the pigeon brain, Brain Struct. Funct. 218 (1) (2013) 269–281.
[41] R.O. Duda, P.E. Hart, D.G. Stork, Pattern Classification, 2nd ed., Wiley-Interscience, 2000 ISBN 0471056693.
[42] G.J. Marin, E. Duran, C. Morales, C. Gonzalez-Cabrera, E. Sentis, J. Mpodozis, J.C. Letelier, Attentional capture? Synchronized feedback signals from the isthmi boost retinal signals to higher visual areas, J. Neurosci. 32 (3) (2012) 1110–1122.
[43] F.A. Soto, E.A. Wasserman, Missing the forest for the trees: object-discrimination learning blocks categorization learning, Psychol. Sci. 21 (10) (2010) 1510–1517.
[44] T. Serre, A. Oliva, T. Poggio, A feedforward architecture accounts for rapid categorization, Proc. Natl. Acad. Sci. U. S. A. 104 (15) (2007) 6424–6429, http://dx.doi.org/10.1073/pnas.0700622104.
[45] T. Serre, T. Poggio, E. McDermott, Learning a Dictionary of Shape-Components in Visual Cortex: Comparison With Neurons, Humans and Machines, Massachusetts Institute of Technology, 2006.
[46] U. Aust, L. Huber, Representational insight in pigeons: comparing subjects with and without real-life experience, Anim. Cognit. 13 (2) (2010) 207–218, http://dx.doi.org/10.1007/s10071-009-0258-4.
[47] S.L. Astley, E.A. Wasserman, Visual discrimination of real objects and pictures in pigeons, Anim. Learn. Behav. 25 (2) (1997) 185–192.
[48] M.L. Spetch, A. Friedman, Pigeons see correspondence between objects and their pictures, Psychol. Sci. 17 (11) (2006) 966–972, http://dx.doi.org/10.1111/j.1467-9280.2006.01814.x.
[49] L. Dittrich, R. Adam, E. Unver, O. Güntürkün, Pigeons identify individual humans but show no sign of recognizing them in photographs, Behav. Process. 83 (1) (2010) 82–89, http://dx.doi.org/10.1016/j.beproc.2009.10.006.
[50] U. Aust, L. Huber, The role of item- and category-specific information in the discrimination of people versus nonpeople images by pigeons, Anim. Learn. Behav. 29 (2) (2001) 107–119, http://dx.doi.org/10.3758/BF03192820.
[51] U. Aust, L. Huber, Target-defining features in a “people-present/people-absent” discrimination task by pigeons, Anim. Learn. Behav. 30 (2) (2002) 165–176, http://dx.doi.org/10.3758/BF03192918.
[52] L. Dittrich, J. Rose, J.U. Buschmann, M. Bourdonnais, O. Gunturkun, Peck tracking: a method for localizing critical features within complex pictures for pigeons, Anim. Cogn. 13 (1) (2010) 133–143.
[53] O. Hardy, N. Leresche, D. Jassik-Gerschenfeld, Postsynaptic potentials in neurons of the pigeon's optic tectum in response to afferent stimulation from the retina and other visual structures: an intracellular study, Brain Res. 311 (1) (1984) 65–74,

http://dx.doi.org/10.1016/0006-8993(84)91399-4.
[54] B.J. Frost, E. Diane, DiFranco, Motion characteristics of single units in the pigeon optic tectum, Vis. Res. 16 (11) (1976) 1229–1234, http://dx.doi.org/10.1016/0042-6989(76)90046-8.
[55] Y.C. Wang, S. Jiang, B.J. Frost, Visual processing in pigeon nucleus rotundus: luminance, color, motion, and looming subdivisions, Vis. Neurosci. 10 (1) (1993) 21–30.
[56] B. Hellmann, O. Güntürkün, Structural organization of parallel information processing within the tectofugal visual system of the pigeon, J. Comp. Neurol. 429 (1) (2001) 94–112.
[57] G. Marin, J.C. Letelier, P. Henny, E. Sentis, G. Farfan, F. Fredes, N. Pohl, H. Karten, J. Mpodozis, Spatial organization of the pigeon tectorotundal pathway: an interdigitating topographic arrangement, J. Comp. Neurol. 458 (4) (2003) 361–380.
[58] M. Tohmi, R. Meguro, H. Tsukano, R. Hishida, K. Shibuki, The extrageniculate visual pathway generates distinct response properties in the higher visual areas of mice, Curr. Biol. 24 (6) (2014) 587–597.
[59] H. Nakamura, H. Hioki, T. Furuta, T. Kaneko, Different cortical projections from three subdivisions of the rat lateral posterior thalamic nucleus: a single-neuron tracing study with viral vectors, Eur. J. Neurosci. 41 (10) (2015) 1294–1310.
[60] S. Thorpe, D. Fize, C. Marlot, Speed of processing in the human visual system, Nature 381 (6582) (1996) 520–522.
[61] M. Riesenhuber, T. Poggio, Models of object recognition, Nat. Neurosci. 3 (Suppl) (2000) 1199–1204.
[62] K. Grill-Spector, N. Kanwisher, Visual recognition: as soon as you know it is there, you know what it is, Psychol. Sci. 16 (2) (2005) 152–160, http://dx.doi.org/10.1111/j.0956-7976.2005.00796.x.
[63] T. Serre, M. Kouh, C. Cadieu, U. Knoblich, G. Kreiman, T. Poggio, A Theory of Object Recognition: Computations and Circuits in the Feedforward Path of the Ventral Stream in Primate Visual Cortex, In AI MEMO, 2005.
[64] N.K. Logothetis, D.L. Sheinberg, Visual object recognition, Annu. Rev. Neurosci. 19 (1996) 577–621.
[65] K. Tanaka, Inferotemporal cortex and object vision, Annu. Rev. Neurosci. 19 (1996) 109–139.
[66] J.J. DiCarlo, D.D. Cox, Untangling invariant object recognition, Trends Cogn. Sci. (Regul. Ed.) 11 (8) (2007) 333–341.
[67] J.J. DiCarlo, D. Zoccolan, N.C. Rust, How does the brain solve visual object recognition? Neuron 73 (3) (2012) 415–434.
[68] J.J. DiCarlo, J.H. Maunsell, Form representation in monkey inferotemporal cortex is virtually unaltered by free viewing, Nat. Neurosci. 3 (8) (2000) 814–821.
[69] N. Li, D.D. Cox, D. Zoccolan, J.J. DiCarlo, What response properties do individual neurons need to underlie position and clutter “invariant” object recognition? J. Neurophysiol. 102 (1) (2009) 360–376.
[70] S.L. Astley, E.A. Wasserman, Categorical discrimination and generalization in pigeons: all negative stimuli are not created equal, J. Exp. Psychol. Anim. Behav. Process. 18 (2) (1992) 193–207.
[71] Y. Yamazaki, U. Aust, L. Huber, M. Hausmann, O. Gunturkun, Lateralized cognition: asymmetrical and complementary strategies of pigeons during discrimination of the “human concept”, Cognition 104 (2) (2007) 315–344.
[72] F.A. Soto, E.A. Wasserman, A category-overshadowing effect in pigeons: support for the common elements model of object categorization learning, J. Exp. Psychol. Anim. Behav. Process. 38 (3) (2012) 322.
[73] D.E. Stansbury, T. Naselaris, J.L. Gallant, Natural scene statistics account for the representation of scene categories in human visual cortex, Neuron 79 (5) (2013) 1025–1034, http://dx.doi.org/10.1016/j.neuron.2013.06.034.