Vision Research 206 (2023) 108190

Contents lists available at ScienceDirect

Vision Research

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

Extracting statistical information about shapes in the visual environment

Sabrina Hansmann-Roth [a,*], Andrey Chetverikov [b], Árni Kristjánsson [a]

[a] Icelandic Vision Lab, School of Health Sciences, University of Iceland, Reykjavík, Iceland
[b] Donders Institute for Brain, Cognition, and Behavior, Radboud University, Nijmegen, the Netherlands

ARTICLE INFO

Keywords:
Visual ensembles
Feature distribution learning
Summary statistics
Visual search
Shape perception

ABSTRACT

It is well known that observers can use so-called summary statistics of visual ensembles to simplify perceptual processing. The assumption has been that instead of representing feature distributions in detail the visual system extracts the mean and variance of visual ensembles. But recent evidence from implicit testing using a method called feature distribution learning showed that far more detail of the distributions is retained than the summary statistic literature indicates. Observers also encode higher-order statistics such as the kurtosis of feature distributions of orientation and color. But this sort of learning has not been shown for more intricate aspects of visual information. Here we tested the learning of distractor ensembles for shape, using the feature distribution learning method. Using a linearized circular shape space, we found that learning of detailed distributions of shape does not occur for this shape space while observers were able to learn the mean and range of the distributions. Previous demonstrations of feature distribution learning involved simpler feature dimensions than the more complex shape space tested here, and our findings may therefore reveal important boundary conditions of feature distribution learning.

1. Introduction

The visual information describing the world around us is rich and highly complex. But detailed processing of these billions of bits of information exceeds the capacities of the visual system. Importantly, however, although visual information is noisy, it is not random. It is regularly structured in space and time. We are, for example, often confronted with sets of similar objects in our visual field, like individual leaves on a tree sharing properties such as shape, texture and color. This redundancy can be efficiently compressed using summary statistics of the items such as average and variance, which are more stable and less prone to noise than single representations. This strategy of summarizing and extracting compressed, statistical information (e.g. average or variance) of groups of features has long been studied within the framework of ensemble perception (e.g. Alvarez, 2011; Haberman & Whitney, 2012; Whitney & Leib, 2018). The extraction of mean and variance for groups of objects has been shown for features like orientation (e.g. Miller & Sheldon, 1969; Parkes, Lund, Angelucci, Solomon & Morgan, 2001), hue (e.g. Maule, Witzel & Franklin, 2014; Webster, Kay & Webster, 2014), speed and direction of motion (Watamaniuk & Duchon, 1992; Watamaniuk & McKee, 1998) and size (Ariely, 2001; Chong & Treisman, 2003).

Only a few studies have, however, addressed whether other summary statistics such as variance, skewness or kurtosis of visual ensembles are also encoded. Atchley and Anderson (1995) presented participants with moving dot clouds. The velocity of each dot was drawn from a predefined distribution. Participants were able to detect dot clouds with different mean velocity or different variance but were unable to find dot clouds that had velocity distributions differing in skewness or kurtosis. Dakin and Watt (1997) found similar results for oriented lines. These results led to the conclusion that the visual system is not sensitive to these higher order ensemble statistics (Dakin, 2015).

However, information about feature distributions in natural images is utilized to make perceptual decisions (e.g., orientation, Girshick, Landy, & Simoncelli, 2011) and Chetverikov et al. (2016, 2017a, 2017b, 2021, see also Tanrikulu et al., 2020, 2021) recently introduced a new approach for studying internal representations of feature distributions, based on intertrial priming in visual search (Kristjánsson & Ásgeirsson, 2019). They used “priming of pop out” (Maljkovic & Nakayama, 1994) which shows how switching target and distractor features leads to an increase in response time that is even larger than for new target and distractor features, a so-called role-reversal effect (Kristjánsson & Driver, 2008). Chetverikov et al. (2016) used such role-reversals to reveal observers’ internal models of distractor distributions. Their

* Corresponding author at: School of Health Sciences, University of Iceland, Reykjavik, Iceland.
E-mail address: Sabrina@hi.is (S. Hansmann-Roth).

https://doi.org/10.1016/j.visres.2023.108190
Received 11 August 2022; Received in revised form 1 February 2023; Accepted 1 February 2023
Available online 11 February 2023

participants viewed search displays containing 36 items drawn from a predefined distribution and searched for the oddly oriented item. After a few learning trials with this predefined distractor distribution, encodings of the internal representation of the distribution were measured on a so-called test trial where a target drawn from the previously learned distractor distribution (or close to it in feature space) was presented. Chetverikov et al. found that search times were slower when the target feature was drawn from within the preceding distractor distribution (role-reversal), compared to when it was drawn from outside it. Crucially, response times varied as a function of the distance between the previous distractor mean and the current target (so-called CT-PD distance), so that the response time distribution mimicked the shape of the learned distractor distribution. Similar encoding characteristics were also found in the visual periphery (Tanrikulu et al., 2020). This showed that distributions of feature probabilities are encoded and affect search behavior. Importantly, Hansmann-Roth et al. (2021) then showed that the same observers who showed detailed learning of feature distributions were unable to explicitly judge the same properties when asked to compare two ensembles with different distribution characteristics as is typically done in ensemble encoding studies. And Pascucci et al., (2022) have then shown that active search is not necessary for such learning, opening up further possibilities for how this learning could be used for representing the visual environment. Until now, however, this implicit method for studying representations of feature distributions has only been used to study low-level features such as color and orientation.

Simple summary statistics of multiple features of one or more groups can be encoded simultaneously. Observers can simultaneously extract the average size of two groups of disks that differ in color (Chong & Treisman, 2005) and can even extract statistical information for up to four groups. However, encoding capacity for summary statistics of multiple groups is not unlimited. Simultaneous comparison of multiple averages of different groups resulted in worse performance than when groups were presented in succession (Attarha & Moore, 2015; Attarha et al. 2014). While these studies mainly focused on the ability to extract the same statistics from multiple independent groups, other studies have assessed the ability to extract different feature characteristics simultaneously. For example, Emmanouil & Treisman (2008) examined observers' ability to extract the average speed and size of moving disks. Although, performance was affected when observers judged multiple features rather than only a single feature, some statistical information of multiple features could be encoded and extracted simultaneously. Whether learning of the full feature distributions of different visual features can simultaneously occur, has also been recently assessed: Hansmann-Roth et al. (2019) examined the simultaneous encoding of multiple low-level features of objects using the implicit FDL method. Observers searched for objects consisting of a task relevant and a task irrelevant feature. While the task-relevant feature was encoded in detail, the secondary feature was not encoded and even interfered with the encoding of the relevant feature.

But how well is statistical information encoded for objects that cannot be described on a single perceptual dimension? While lines can be distinguished by their orientation and colored patches by their hue, other visual properties such as shape or material properties can differ on multiple perceptual dimensions. The perceived gloss of a surface is modulated by the specular reflectance of the material, but also by the albedo and the shape of the surface (Hansmann-Roth & Mamassian, 2017; Nishida and Shinya, 1998; Olkkonen & Brainard, 2011; Pellacini et al. 2000). The roughness of a surface is not only modulated by its physical roughness or bumpiness, but also depends on the specular reflectance and albedo (Ho et al. 2008). Similarly, simple 2D shape representations are also a combination of multiple shape descriptors (compactness, perimeter, area, or the principal orientation axis) that the visual system uses to judge and compare shapes (e.g., Huang, 2020; Morgenstern et al. 2021, Zhang & Lu, 2004).

While size, color and orientation are well studied in ensemble perception, only a handful of studies are available on the ensemble encoding of shape. This lack of studies might be explained by the fact that studying shapes is difficult because they are multidimensional. But, at higher levels of the visual processing hierarchy, the brain might, nevertheless, combine multiple dimensions to achieve properties desirable for real-world behavior, such as object recognition (Di Carlo, Zoccolan, Rust, 2012). Nevertheless, Sweeny et al. (2021) and Elias & Sweeny (2020) have shown that observers can successfully extract the average aspect ratio of groups of ellipses. Khayat et al. (2021) presented novel shapes to observers in RSVP and tested them afterwards on a memory task. Observers implicitly represented the group mean and category boundaries of the ensembles.

A recently proposed shape space (Li et al. 2020) that includes 2D contours in a perceptually uniform space enables new ways of studying the encoding of shape ensembles, including with FDL methods. Li and colleagues (2020) started with 12 prototype shapes and morphed them to ultimately create a 360° circular shape space. Using similarity judgments, they refined their shape space so that eventually similar angular distances corresponded to similar perceptual differences. This shape space therefore allows the application of FDL methodology, just as we have done with other circular shape spaces like orientation and color.

1.1. Current aims and predictions

How precisely can ensembles of differently shaped 2D contours be encoded? We presented observers with differently shaped objects drawn from a Gaussian or a uniform distribution, measuring how rich the representations of such ensembles are. By using the newly proposed perceptually uniform shape space (Li et al., 2020), we can utilize our feature distribution learning (FDL) method to examine the richness of shape representations. Based on our previous results showing encoding of full feature distributions we can hypothesize what the shape of our RT-functions on test trials will look like depending on how much statistical information is encoded (Fig. 2). In previous experiments, the shape of the RT-function on test trials resembled the shape of the underlying feature distribution (Fig. 2d). In contrast, if no information about ensembles on learning trials is encoded, the RT-function should be flat (not affected by mean, range or probabilities; Fig. 2a). If observers encode the summary statistics of the ensemble (mean and range) their RT-functions should show a flat part with higher search times within the range than outside it (Fig. 2b). Moreover, if only mean and range are encoded, we expect no differences between different distractor distributions (Gaussian versus uniform were contrasted) and the first part of the RT-function should be flat since probabilities are not encoded. Alternatively, observers might have prior expectations for the distribution shape of an ensemble, such as a Gaussian prior which is defined by only two parameters. If mean and range are encoded and such priors exist, then the RT-functions on the test trial should resemble the shape of a Gaussian distribution for both distribution shapes (Fig. 2c). And to reiterate, if the shape of the distractor distributions is encoded during the learning trials, then RT-functions on test-trials should resemble the shape of the two distractor distributions tested (Fig. 2d).

We presented observers with ensembles of 2D shapes (Fig. 1) and tested their implicit encoding of the distractor distribution using our recently developed feature distribution learning (FDL) paradigm. Throughout the experiment, observers searched for an oddly-shaped target among a set of distractors of varying shapes. This enables careful examination of the richness of the encoded distractor distributions by comparing their RT-functions to the predictions in Fig. 2.

2. Material and methods

2.1. Procedure/Task

All participants completed two sessions which lasted about 50–60 min each on separate days. These sessions were preceded by a training

Fig. 1. Screenshot of a stimulus on a single learning trial. The odd-one-out target is in the bottom row, fourth column in this example, while all other shapes are distractors drawn from a Gaussian distribution.

session consisting of the same amount of blocks. A single session consisted of 432 blocks and 1944 search trials in total. One block consisted of 3–4 learning trials and one test trial. On each trial, participants were instructed to find the odd-one-out shape (the target) and indicate its position by pressing the "i" key if it was in the upper three rows and "j" if it was in the lower three rows. The stimuli were presented until response and participants were encouraged to respond as quickly and accurately as possible. After each button press the next trial began. If participants made a mistake, the word ERROR appeared on the screen for 1 s. For motivational purposes participants' performance was scored and their score on the last trial was presented in the upper left corner of the screen. The score was computed as Score = 10 + (1-RT) *10 for a correct response and Score = -|Score|-10 for an error, where RT is the response time in seconds. Both training and test sessions were interrupted by three breaks, when the total score was displayed on the screen.

2.2. Stimuli

Our search displays consisted of 36 shapes that appeared on an invisible 6 × 6 grid that subtended 16 × 16° of visual angle on the center of the screen. We used the Validated Circular Space (Li et al. 2020) that contains a set of 360 shapes organized in a circular space similar to the color wheel and orientation space. This space has been validated to ensure that it is perceptually uniform: Similar distances in degrees are visually like similar distances for other shapes on the shape wheel (Li et al. 2020). We used 180 different shapes from this space (every second shape on the wheel). This allowed us to use the same distribution parameters (SD, distance between target and distractors) as in our orientation learning experiments (Chetverikov et al. 2016, 2017b).

All individual shape positions on the monitor were jittered by adding random values between +/- 0.5° to the horizontal and vertical coordinates. The distractors were drawn from either a truncated Gaussian or a uniform distribution. These distributions had an SD of 15° and the uniform distribution a range of 60° (values outside this range of the Gaussian distribution were removed, so that both distributions had the same range).

During test trials, distractors were drawn from a Gaussian distribution with an SD of 10°. The mean of the distractor distribution at the beginning of each learning streak and the target shape were randomly selected with the restriction that the target had to be 70-110° away from the distractor mean to ensure that the target differed sufficiently from the distractors (see Fig. 3). This learning streak was followed by a test trial, in which the target was positioned at particular probe points from within and around the previously presented distractor distribution. The probe points ranged from −80° to 80° in steps of 20° (with an added random value between −10° to 10°). We refer to this as the current target to previous distractor distance (CT-PD). The distractor mean was selected randomly, again with the restriction that the minimum distance was 70° and the maximum distance was 110°.

2.3. Observers

11 participants (mean age: 27.1, 7 females) completed the experiment. All (except for one author) were naïve to the purpose of the study and all had normal or corrected-to-normal vision. They all gave written, informed consent. All experiments conformed to the requirements of the local ethics committee and the Declaration of Helsinki.

2.4. Apparatus

All stimuli were displayed on a 24-inch LCD monitor (ASUS, VX248h) with a resolution of 1920 × 1080. All stimuli were displayed using Matlab R2016a and Psychtoolbox-3 (Brainard, 1997) that ran on a Desktop PC with Windows 10.

2.5. Data analysis

Data analyses closely followed the approach introduced to study the

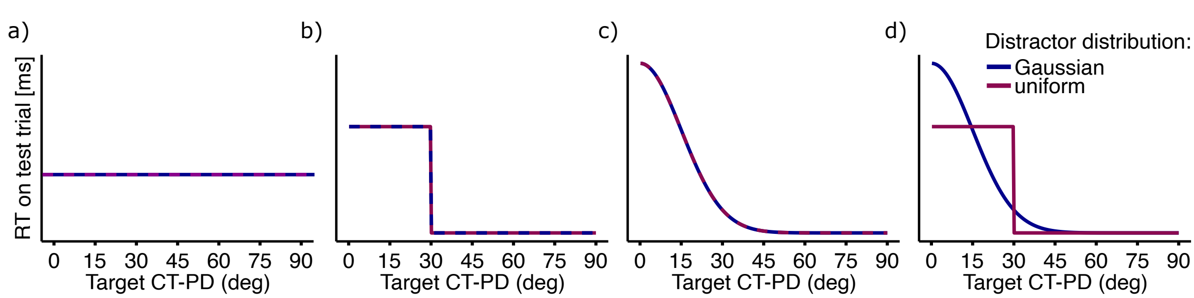

Fig. 2. Overview of the potential search time functions on the test trials depending on how much information about the shape distributions is encoded. Search times should reflect the amount of information encoded about the distractor distribution and are plotted against the distance between the current target and previous distractor distribution (Target CT-PD). a) If observers do not encode any distractor distribution characteristics, the RT function would be flat. B) shows an RT function if information about the mean and range of the distractor features is encoded but not the probabilities of the individual distractors. c) Observers might also encode information about the mean and range of the distractors and assume a Gaussian distribution as a prior. d) If observers encode the full distractor distributions, then RT functions on test trials should differ depending on whether the distractor distribution was Gaussian or uniform.

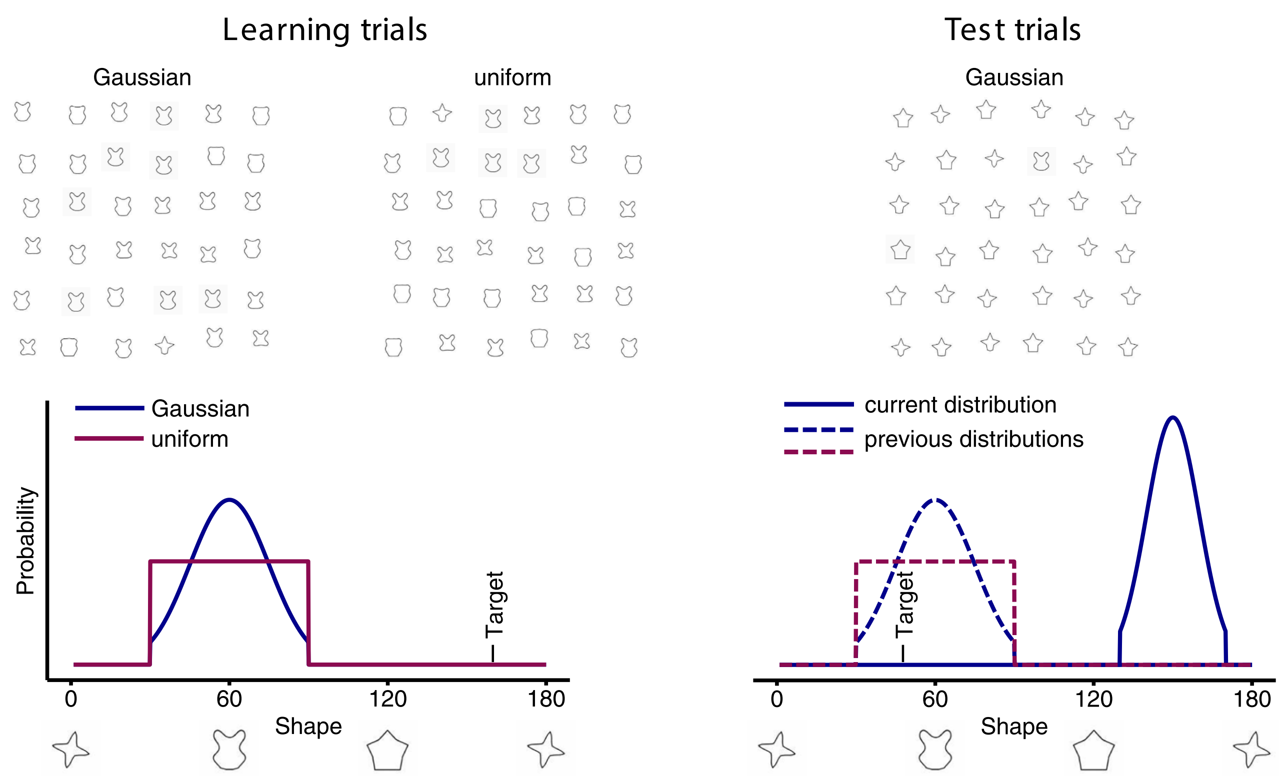

Fig. 3. Overview of the distractor distributions on learning and test trials along with examples of target positions relative to the distractor distributions.

encoding of orientation (Chetverikov et al. 2016, 2017a, 2017c) and color distributions (Chetverikov et al. 2017b; Hansmann-Roth et al., 2019, Hansmann-Roth et al. 2022). Reaction times were log-transformed for the final data analysis. Trials with incorrect responses and the trial immediately following an incorrect response were removed (because of potential post-error slowing). To assess any influence of distribution shape and the effects of repetition within a learning streak we conducted two-way repeated-measures ANOVAs, with Greenhouse–Geisser corrections, where applicable, after testing for sphericity using Mauchly tests. ANOVAs were conducted in R (R Development Core Team, 2012) from the ez package (Lawrence, 2016). We compared the shapes of the RT CT-PD function using segmented regressions (Muggeo, 2008). Confidence intervals are presented on the non-log data, but all statistical tests are done on log-transformed search times. We also fitted pre-specified models to our data that corresponded to actual distribution shapes: a uniform model with a fixed range of 30°, a half-Gaussian model with SD = 15°, a linear model and a "uniform with decrease model", which contains a flat part within the distribution range and a linear decrease outside it. Each model includes a Gaussian-distributed error term (see supplementary material for equations). We fitted the different models to our data and obtained the best-fitting parameters using Maximum Likelihood Estimation and used the Bayesian Information Criterion for comparing models.

3. Results

Our previous research has shown that the visual system can encode intricate features of color and orientation distributions (Chetverikov et al., 2016, 2017; Hansmann-Roth et al., 2019; 2021; Tanrikulu et al., 2020; 2021). Our goal was to examine how much statistical information (mean, range or shape distribution) of ensembles of differently shaped objects is encoded.

We first analyzed observers search performance on learning trials where they searched for an oddly-shaped target. Learning streaks consisted of 3–4 learning trials where the distractor distribution characteristics were constant. Fig. 4 plots the average search times and error rates on these learning streaks separately for both distractor distribution types.

We analyzed average error rates for both distractor distributions and compared the change in error rates over the trials in the learning streaks. A two-factor (distribution type × trial number within learning streak) repeated-measures ANOVA revealed a main effect of trial number during the learning streak, $F(3, 30) = 21.93$, $p < .001$, $\eta^2_G = 0.17$ but no effect of distractor distribution type $F(1, 10) = 2.67$, $p = .133$, $\eta^2_G = 0.02$ nor an interaction, $F(3, 30) = 0.06$, $p = .950$, η^2_G less than 0.01. Overall, the error rates were very small. Observers responded more accurately on later trials within a learning streak but the average error rates did not differ significantly between the two distractor distributions.

Furthermore, we analyzed average search times for both distractor distributions, comparing the change in search time over the trials within the learning streaks. A two-factor (distribution type × trial number in learning streak) repeated-measures ANOVA revealed a main effect of trial number, $F(3, 30) = 91.03$, $p < .001$, $\eta^2_G = 0.17$, a significant effect of distribution type $F(1, 10) = 164.38$, $p < .001$, $\eta^2_G = 0.13$ but no interaction, $F(3, 30) = 0.5$, $p > .05$, $\eta^2_G = 0.00$. Participants responded faster during later trials within a streak and faster for targets drawn from a Gaussian distribution. These search time differences between the uniform and the Gaussian distribution during the learning might be induced by the small difference in standard deviation. Both distributions had the same stimulus range which causes the standard deviation for the uniform distribution to be slightly larger than the standard deviation for the Gaussian distractor distribution (17.3° vs 15°).

In previous FDL studies, search times on test trials followed the shape of the underlying distractor distribution: RT functions on test trials that followed a Gaussian distractor distribution monotonically decreased, while RT functions following uniform distractor distributions consisted of a two-step function: a flat part followed by a linear decrease with a significant breakpoint close to the edge of the underlying physical distribution.

To what extent were observers able to encode certain characteristics

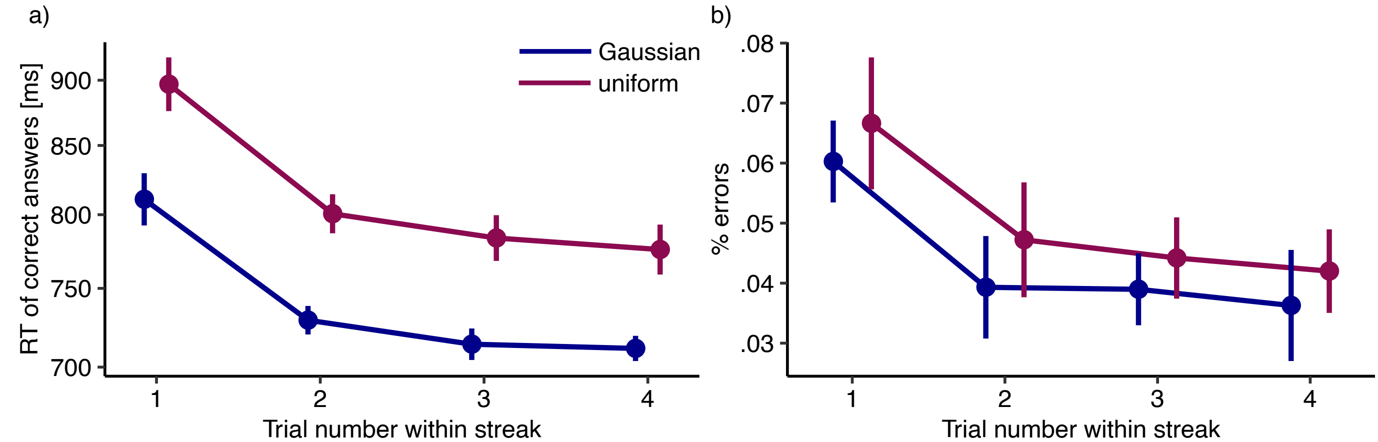

Fig. 4. Search times and accuracy on the learning trials. Mean search times (a) and mean error rates (b) across all observers are plotted as a function of the trial number on the learning trials. Colors correspond to the two different distractor distribution types. Error bars correspond to the 95% CI adjusted for within-subject variability (Morey, 2008).

(mean, range and/or distribution shape) of the distractor distributions? We used our previously developed analysis method to assess the shape of the RT functions on the test trials and their similarity to the underlying distractor distributions from the learning trials. Fig. 5 plots search times on the test trials as a function of the distance between the target on the test trial and the mean of the previous distractor distribution. The upper panel plots the underlying physical distractor distributions for

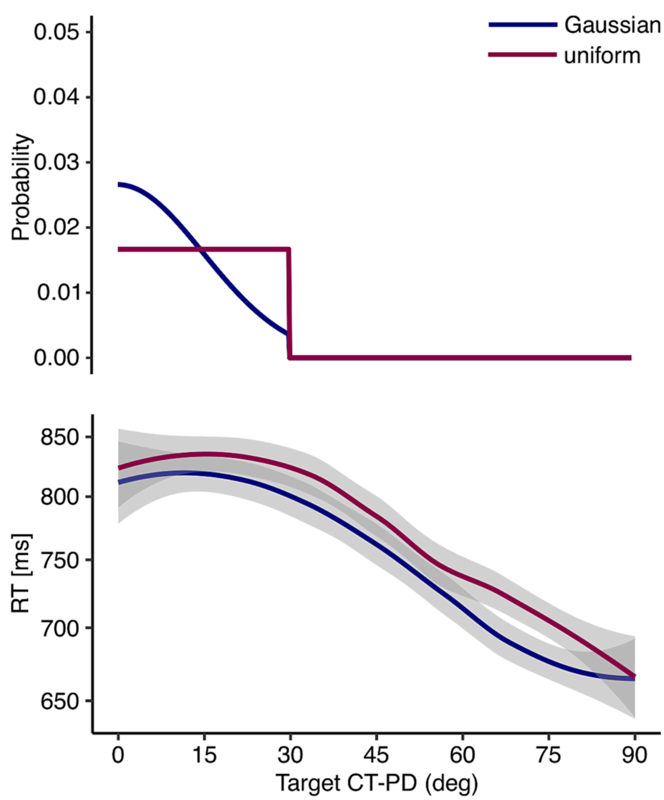

Fig. 5. Mean search times on the test trials across all observers against the distance between the current target and the previous distractor distribution mean (CT-PD). Only trials where observers found the correct target are included. Gray areas show the 95% CI's based on the fitted loess functions. Each fitted curve corresponds to one of the distractor distribution shapes on the learning trials. The upper panel corresponds to the actual distractor distribution shape. We plot only the absolute CT-PD since both distractor distributions are symmetrical.

comparison. To compare the RT functions to our different hypothetical results plotted in Fig. 1, we first conducted a segmented regression and then fitted predefined models to our data that corresponded to different distribution shapes.

The segmented regression systematically searches for changes in the slope of the RT function. Both RT functions were very similar showing an initial flat part followed by a steep decrease. Our segmented regression analyses confirmed these informal observations: Following a uniform distractor distribution, the Davies test confirmed a significant breakpoint at 26° away from the distribution mean (p =.004). The segment before the breakpoint was not significantly different from zero, b = -0.08, CI = [-2.17,2.01] and the segment after the breakpoint showed a significantly negative slope, b = -2.65, CI = [-3.19, −2.11]. Following a Gaussian distribution, we observed the same pattern: a significant breakpoint, confirmed by the Davies test (p =.007), at 27.6° from the distractor distribution mean. The first segment was flat and not significantly different from zero, b = -0.49, CI = [-2.40, 1.42] and the second segment after the breakpoint showed a significantly negative slope, b = -2.73, CI = [-3.27, −218].

In an additional step we also fitted our predefined models to the RT-functions. These predefined models have been used to study distribution learning for color and orientation (Chetverikov et al., 2019). Our predefined models corresponded to different distribution shapes and were compared to a Null model that assumes no encoding of the statistical characteristics of the presented ensemble. We used a half-Gaussian model with a SD = 15, a uniform model with a fixed range of 30 and a linear model and a "uniform with decrease" model. Each model entails a Gaussian-distributed error term. We then extracted the best fitting parameters using Maximum Likelihood Estimation and used a Bayesian Information Criterion for comparison. Fig. 6 shows observers' data and the resulting fits. For both RT functions the "uniform with decrease" model provided the best fit (uniform: BIC = 3008.27; Gaussian: 3246.25), followed by the linear model (uniform: ΔBIC = 12.8; Gaussian: 6.6). Moreover, we fitted these models separately to the data for individual observers. Following a uniform distractor distribution, the best model fit for most observers was provided by the "uniform with decrease" model and the uniform model (N = 7 (uniform model for one observer)), followed by the linear model which provided the best fit for 3 observers. For the remaining observer the Null model provided the best fit, indicating that this observer did not encode any information about the distribution statistics. There was no difference between the best model fit and the Null model fit: ΔBIC = 1.3.

Following a Gaussian distractor distribution, the best model fit was also provided by the "uniform with decrease" and the uniform model (N = 7 (uniform model for one observer, although the difference between

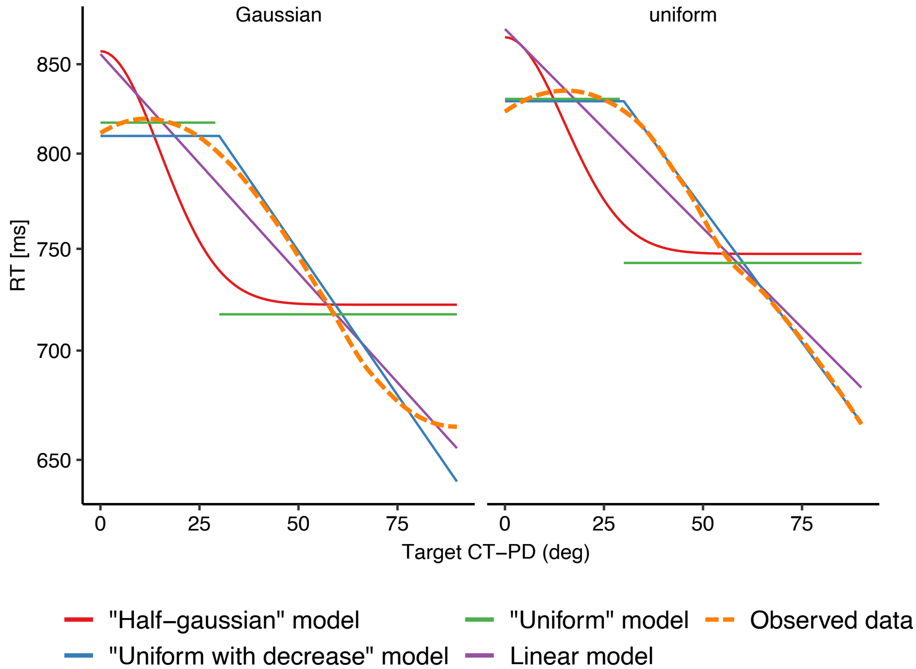

Fig. 6. Mean search times on the test trials across all observers are plotted against the distance between the current target and the previous distractor distribution mean (CT-PD). The observed data is plotted along with the modelling fits using maximum likelihood estimation. The observed data is plotted in orange and the best fit of the different models to the observed data is plotted in blue, green, purple and red. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

the uniform and the "uniform with decrease" model for this observer was small: $\Delta BIC = 1.05$)). For the remaining 4 observers the linear model provided the best fit.

Overall, our results show that observers can encode information about shape ensemble statistics on the learning trials. In contrast with previous results for orientation and color, however, we found no differences between encodings of Gaussian and uniform distractor distributions. This suggests that the distractor distributions of the 2D shapes were not fully encoded but were approximated by their mean and range. Both the segmented regression and the model comparisons show that observers successfully learned the mean and the range of the distractor distributions, but distribution shapes were not encoded.

4. Discussion

We have previously shown how observers can learn surprisingly intricate details of distributions of simple features (color and orientation). Here, we expanded our paradigm to study the encoding of feature distributions of the dimension of shape using a circular and perceptually uniform shape space (Li et al., 2020). This is a more elaborate stimulus set, that includes a higher number of dimensions that can vary.

We applied our feature distribution learning method (Chetverikov et al., 2016, 2019), presenting observers with ensembles of distractors sampled from either a uniform or a Gaussian distribution from the shape space and a single target that observers had to find as quickly and accurately as possible. We examined search times on test trials as a function of the feature distance between the target on the test trial and the previous distractor distribution to assess how much information about the distractor distributions was encoded. Our previous work showed that observers can encode the full probability density functions of feature distributions for orientation and color. Conversely, the current results show that while observers can implicitly encode summary statistics (mean and range) for differently shaped distractors, there were no differences in the shape of the RT-functions on test trials indicating that the full probabilistic representation of the distractors was not encoded.

Our distractors were defined by shape (contour) differences. While color and orientation are low-level visual features, shape as defined is this stimulus set, is a far more complex entity. The objects vary in the shape of the contours, in convexity and concavity, and in the area of the stimulus. Multiple features of the contour are therefore likely to contribute to the percept of the individual objects of the ensemble. As we previously showed (Hansmann-Roth et al., 2019) there are limits to what can be learned in FDL studies. There we tested whether two probability distributions of different features could be learned simultaneously. Observers performed a search task with lines of a particular orientation and color drawn from Gaussian or uniform distributions, either searching for the oddly oriented or oddly colored line. Observers were able to learn task-relevant distributions, but unable to learn the shape of the irrelevant feature distribution, although they were able to learn its summary statistics (mean and range). What is visible from the examples in Fig. 3 is that the target shape might not pop-out to the same degree as e.g., a red colored circle among greenish circles. Extracting statistical information from visual scenes differs depending on whether global or local attention is deployed. Chong and Treisman (2005) combined two different tasks within a trial: First, participants performed either serial search (closed circle among open circles) or a parallel search (open circle among closed circles). Immediately afterwards, participants were presented with two test circles and either performed a member identification task or a mean discrimination task. They were asked to indicate which test circle corresponded to the mean size of circles, and in the member identification task, they indicated which test circle had been presented at a particular location within the display. Their results showed that after parallel search, participants performance on the mean discrimination task was better than if it was preceded by a serial search task that requires more local attention than a parallel search task. In previous research, we also showed the importance of parallel processing for distribution learning. Representing the distribution shape requires large set sizes during the visual search task. Visual search displays with less than 36 items resulted in no encoding of the shape of the feature distribution, demonstrating that sampling only a small number of items from the display is not sufficient (Chetverikov et al. 2017c). Moreover, our previous also showed that encoding the shape of a feature distribution requires a large set of stimuli (distractors) drawn from that distribution. Observers did not encode any information about the shape of the distractor distributions for set size smaller than 36 items. These results furthermore demonstrate that distribution learning needs parallel processing of the distractors (Chetverikov et al. 2017c). If those complex shapes applied here prevent parallel search and the target can only be found through serial search, then this may prevent distribution learning as well.

The circular shape space applied in this study was developed based on an initial morphing of different shapes from a set of prototypes. The authors then used participants similarity ratings and reconstructed a shape space using multidimensional scaling. After multiple iterations they obtained their circular shape space, which is comparable to the color wheel in CIELAB color space. The angular distance between shapes along their circle is a proxy of their overall visual similarity. The successful creation of such a space suggests that the shape space is encoded by the visual system.

Two pairs of shapes with the same angular distance in this shape space also exhibit a similar visual dissimilarity to human observers, even though the shape of the contours, features such as curvature, symmetry, complexity might completely differ. Based on this, we assume that these shapes can be processed holistically and the lack of learning the distractor distribution shape is therefore not caused by the particular shape space we applied.

The shape space has also been tested in a study investigating interference effects in working memory and results in that study were comparable to the results they obtained using the color wheel (Li et al. 2019). It appears that the shape space affects memory in a similar way as color, highlighting that this shape space is like a color wheel and can be used to study ensemble perception, working memory etc., similarly to other visual features.

Our current results may therefore reveal further limits of feature distribution learning. The shape variation that we tested here is multidimensional. The shapes are assembled from different parts, and these parts may have different salience (Biederman, 1987; Hoffman & Richards, 1984; Hoffman & Singh, 1997; Marr & Nishihara, 1978). Different individual parts might therefore correspond to the final percept making any learning of the feature space of shapes far more difficult than the more straightforward one-dimensional features that we have tested previously.

Another related issue is whether the search that was tested here may differ from the search tasks in our previous demonstrations of feature distribution learning. For example, when the odd-one-out item contains a salient convexity versus when the odd-one-out approaches the shape of a pentagon may involve a very different search scenario. Visual similarity is an important determinant of visual search performance (Duncan & Humphreys, 1989; Hout et al. 2016). While orientation space may contain anisotropies in the form of cardinal biases, and color space may contain salient color category boundaries, the net effect may not be as damaging to feature distribution learning as for the shape space that we test here. It is also interesting that the search time differences between uniform and Gaussian distributions that we observe here are similar to those that we have found in previous FDL studies, yet no encoding of the different distractor distribution shapes occurred.

Based on the results one might ask whether the shape space we applied was therefore not encoded? However, based on the predictions described in the Introduction (Fig. 2), our results fit best with the assumption that observers encoded the mean and the range of the distractor distributions but could not encode the probabilities of the individual distractors. Moreover, the decrease in response times and the increase in accuracy during the learning blocks also shows that information about the target and distractor shape was encoded, namely the mean and range.

To sum up, we tested learning of distributions for a newly proposed linearized shape space (Li et al., 2020). While observers were able to learn the mean and range, unlike previous results for orientation and color space, they were unable to learn the distribution shape. Shape space contains more variation than the simpler spaces that were tested previously and the current results may therefore reveal important boundary conditions on possibilities for feature distribution learning.

CRediT authorship contribution statement

Sabrina Hansmann-Roth: Conceptualization, Methodology, Software, Validation, Data curation, Formal analysis, Writing – original draft. **Andrey Chetverikov:** Conceptualization, Methodology, Writing – review & editing, Project administration, Funding acquisition. **Árni Kristjánsson:** Conceptualization, Methodology, Writing – review & editing, Project administration, Funding acquisition.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

All data have been made publicly available via the Open Science Framework and can be accessed at https://osf.io/wema7/

Acknowledgements

We thank Karólína Jóhannsdóttir and Sæunn Jódís Sunnudóttir for help with the data collection. SHR and AK were supported by grant IRF #173947-052 from the Icelandic Research Fund, and by a grant from the Research Fund of the University of Iceland.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.visres.2023.108190.

References

Alvarez, G. A. (2011). Representing multiple objects as an ensemble enhances visual cognition. *Trends in Cognitive Sciences, 15*(3), 122–131.
Ariely, D. (2001). Seeing sets: Representation by statistical properties. *Psychological Science, 12*, 157–162.
Attarha, M., & Moore, C. M. (2015). The capacity limitations of orientation summary statistics. *Attention, Perception, & Psychophysics, 7*(4), 1116–1131.
Attarha, M., Moore, C. M., & Vecera, S. P. (2014). Summary statistics of size: Fixed processing capacity for multiple ensembles but unlimited processing capacity for single ensembles. *Journal of Experimental Psychology: Human Perception & Performance, 40*(4), 1440–1449.
Biederman, I. (1987). Recognition-by-components: A theory of human image interpretation. *Psychological Review, 94*, 115–147.
Brainard, D. (1997). The Psychophysics Toolbox. *Spatial Vision, 10*, 433–436.
Chetverikov, A., Campana, G., & Kristjánsson, Á. (2016). Building ensemble representations: How the shape of preceding distractor distributions affects visual search. *Cognition, 153*, 196–210.
Chetverikov, A., Campana, G., & Kristjánsson, Á. (2017a). Rapid learning of visual ensembles. *Journal of Vision, 17(2):21*, 1–15. https://doi.org/10.1167/17.2.21
Chetverikov, A., Campana, G., & Kristjánsson, Á. (2017b). Representing color ensembles. *Psychological Science, 28*(10), 1510–1517. https://doi.org/10.1177/0956797617713787
Chetverikov, A., Campana, G., & Kristjánsson, Á. (2017c). Set size manipulations reveal the boundary conditions of perceptual ensemble learning. *Vision Research, 140*, 144–156. https://doi.org/10.1016/j.visres.2017.08.003
Chetverikov, A., Hansmann-Roth, S., Tanrikulu, O. D., & Kristjánsson, Á. (2019). Feature distribution learning (FDL): A new method to study visual ensembles with priming of attention shifts. In S. Pollman (Ed.), *Spatial Learning and Attention Guidance*. New York, NY: Neuromethods, Springer Nature.
Chong, S. C., & Treisman, A. (2003). Representation of statistical properties. *Vision Research, 43*(4), 393–404.
Chong, S. C., & Treisman, A. (2005). Statistical processing: Computing the average size in perceptual groups. *Vision Research, 45*(7), 891–900.
Dakin, S. C., & Watt, R. J. (1997). The computation of orientation statistics from visual texture. *Vision Research, 37*(22), 3181–3192.
Dakin, S. C. (2015). Seeing statistical regularities: Texture and pattern perception. In J. Wagemans (Ed.), The Oxford Handbook of Perceptual Organization (pp. 150–167).
Duncan, J., & Humphreys, G. W. (1989). Visual search and stimulus similarity. *Psychological review, 96*(3), 433.
Elias, E., & Sweeny, T. D. (2020). Integration and segmentation conflict during ensemble coding of shape. *Journal of Experimental Psychology: Human Perception and Performance, 46*(6), 593.

Girshick, A. R., Landy, M. S., & Simoncelli, E. P. (2011). Cardinal rules: Visual orientation perception reflects knowledge of environmental statistics. *Nature Neuroscience, 14*(7), 926–932.

Haberman, J., & Whitney, D. (2012). Ensemble perception: Summarizing the scene and broadening the limits of visual processing. In J. Wolfe, & L. Robinson (Eds.), *From perception to consciousness: Searching with Anne Treisman* (pp. 1–21). Oxford, UK: Oxford University Press.

Hansmann-Roth, S., Chetverikov, A., & Kristjánsson, A. (2019). Representing color and orientation ensembles: Can observers learn multiple feature distributions? *Journal of Vision, 19*(9), 2. https://doi.org/10.1167/19.9.2

Hansmann-Roth, S., Kristjánsson, Á., Whitney, D., & Chetverikov, A. (2021). Dissociating implicit and explicit ensemble representations reveals the limits of visual perception and the richness of behavior. *Scientific Reports, 11*, 3899.

Hansmann-Roth, S., & Mamassian, P. (2017). A glossy simultaneous contrast: Conjoint measurements of gloss and lightness. *I-Perception, 8*(1), 1–16.

Hansmann-Roth, S., Þorsteinsdóttir, S., Geng, J. J., & Kristjánsson, Á. (2022). Temporal integration of feature probability distributions. *Psychological Research*, 1–15.

Hout, M. C., Godwin, H. J., Fitzsimmons, G., Robbins, A., Menneer, T., & Goldinger, S. D. (2016). Using multidimensional scaling to quantify similarity in visual search and beyond. *Attention, Perception, & Psychophysics, 78*(1), 3–20.

Ho, Y.-X., Landy, M. S., & Maloney, L. T. (2008). Conjoint measurement of gloss and surface texture. *Psychological Science, 19*(2), 196–204.

Hoffman, D. D., & Richards, W. A. (1984). Parts of recognition. *Cognition, 18*(1–3), 65–96.

Hoffman, D. D., & Singh, M. (1997). Salience of visual parts. *Cognition, 63*(1), 29–78.

Huang, L. (2020). Space of preattentive shape features. *Journal of Vision, 20*(4), 1–20.

Khayat, N., Fusi, S., & Hochstein, S. (2021). Perceiving ensemble statistics of novel image sets. *Attention, Perception, & Psychophysics, 83*(3), 1312–1328.

Kristjánsson, Á., & Ásgeirsson, Á. G. (2019). Attentional priming: Recent insights and current controversies. *Current Opinion in Psychology, 29*, 71–75.

Kristjánsson, Á., & Driver, J. (2008). Priming in visual search: Separating the effects of target repetition, distractor repetition and role-reversal. *Vision Research, 48*(10), 1217–1232.

Li, A. Y., Liang, J., Lee, A. C. H., & Barense, M. D. (2019). Visual interference can help and hinder memory: Measuring memory fidelity using a novel circular shape space. bioRxiv. *Advance online publication.*. https://doi.org/10.1101/535922

Li, A. Y., Liang, J. C., Lee, A. C., & Barense, M. D. (2020). The validated circular shape space: Quantifying the visual similarity of shape. *Journal of Experimental Psychology: General, 149*(5), 949.

Marr, D., & Nishihara, H. K. (1978). Representation and recognition of the spatial organization of three-dimensional shapes. *Proceedings of the Royal Society of London. Series B. Biological Sciences, 200*(1140), 269–294.

Maule, J., Witzel, C., & Franklin, A. (2014). Getting the gist of multiple hues: Metric and categorical effects on ensemble perception of hue. *JOSA A, 31*(4), A93–A102.

Maljkovic, V., & Nakayama, K. (1994). Priming of pop-out: I Role of features. *Memory & Cognition, 22*(6), 657–672.

Miller, A. L., & Sheldon, R. (1969). Magnitude estimation of average length and average inclination. *Journal of Experimental Psychology, 81*(1), 16.

Morey, R. D. (2008). Confidence intervals from normalized data: A correction to Cousineau (2005). *Tutorial in Quantitative Methods for Psychology, 4*(2), 61–64.

Morgenstern, Y., Hartmann, F., Schmidt, F., Tiedemann, H., Prokott, E., Maiello, G., & Fleming, R. (2021). An image-computable model of human visual shape similarity. *PLoS Computational Biology, 17*(6), e1008981.

Muggeo, V. M. R. (2008). segmented: An R package to fit regression models with broken-line relationships. *R News, 8*(1), 20–25.

Nishida, S. Y., & Shinya, M. (1998). Use of image-based information in judgments of surface-reflectance properties. *JOSA A, 15*(12), 2951–2965.

Olkkonen, M., & Brainard, D. H. (2011). Joint effects of illumination geometry and object shape in the perception of surface reflectance. *i-Perception, 2*(9), 1014–1034.

Parkes, L., Lund, J., Angelucci, A., Solomon, J. A., & Morgan, M. (2001). Compulsory averaging of crowded orientation signals in human vision. *Nature Neuroscience, 4*, 739–744.

Pascucci, D., Ceylan, G., & Kristjánsson, Á. (2022). Feature distribution learning by passive exposure. *Cognition, 227*, Article 105211.

Pellacini, F., Ferwerda, J. A., & Greenberg, D. P. (2000). Toward a psychophysicallybased light reflection model for image synthesis. In Proceedings of the 27th Annual Conference on Computer Graphics and Interactive Techniques (pp. 55–64).

R Development Core Team (2012). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from http://www.Rproject.org/.

Sweeny, T. D., Bates, A., & Elias, E. (2021). Ensemble perception includes information from multiple spatial scales. *Attention, Perception, & Psychophysics, 83*(3), 982–997.

Tanrıkulu, Ö. D., Chetverikov, A., & Kristjánsson, Á. (2020). Encoding perceptual ensembles during visual search in peripheral vision. *Journal of Vision, 20*, 1–18.

Tanrıkulu, Ö. D., Chetverikov, A., & Kristjansson, A. (2021). Testing temporal integration of feature probability distributions using role-reversal effects in visual search. *Vision Research, 188*, 211–226.

Watamaniuk, S. N., & Duchon, A. (1992). The human visual system averages speed information. *Vision Research, 32*(5), 931–941.

Watamaniuk, S. N., & McKee, S. P. (1998). Simultaneous encoding of direction at a local and global scale. *Perception & Psychophysics, 60*(2), 191–200.

Webster, J., Kay, P., & Webster, M. A. (2014). Perceiving the average hue of color arrays. *JOSA A, 31*(4), A283–A292.

Whitney, D., & Leib, A. Y. (2018). Ensemble perception. *Annual Review of Psychology, 69*, 105–129.

Zhang, D., & Lu, G. (2004). Review of shape representation and description techniques. *Pattern Recognition, 37*(1), 1–19.