Vision Research 109 (2015) 77–86

Contents lists available at ScienceDirect

Vision Research

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

Effects of spatiotemporal consistencies on visual learning dynamics and transfer

Hila Harris, Dov Sagi [*]

Department of Neurobiology, The Weizmann Institute of Science, Rehovot 76100, Israel

ARTICLE INFO

Article history:
Received 19 June 2014
Received in revised form 20 February 2015
Available online 28 February 2015

Keywords:
Perceptual learning
Sensory adaptation
Specificity and generalization
Pre-training
Training structure

ABSTRACT

Spatiotemporal interactions affect visual performance under repeated stimulation conditions, showing both incremental (commonly related to learning) and decremental (possibly sensory adaptation) effects. Here we examined the role of spatiotemporal consistencies on learning dynamics and transfer. The backward-masked texture-discrimination paradigm was used, with stimulus onset asynchrony (SOA) controlling the observers' performance level. Temporal consistencies were examined by modifying the order in which SOA was varied during a training session: gradually reduced SOA (high consistencies) versus randomized SOA (low consistencies). Spatial consistencies were reduced by interleaving standard target trials with oriented 'dummy' trials containing only the background texture (no target, oriented 45° relative to the target's orientation). Our results showed reduced improvement following training with gradual SOA, as compared with random SOA. However, this difference was eliminated by randomizing SOA only at the initial and final segments of training, revealing a contaminating effect of temporal consistencies on threshold estimation rather than on learning. Inserting the 'dummy' trials (reduced spatial consistencies) facilitated both the learning and the subsequent transfer of learning, but only when sufficient pre-training was provided. These results indicate that visual sensitivity depends on a balance between two opposing processes, perceptual learning and sensory adaptation, both of which depend on spatiotemporal consistencies. Reducing spatiotemporal consistencies during training reduces the short-term spatiotemporal interactions that interfere with threshold estimation, learning, and generalization of learning. We consider the results within a theoretical framework, assuming an adaptable low-level network and a readout mechanism, with orientation and location-specific low-level adaptation interfering with the readout learning.

1. Introduction

In perceptual learning, repetition-based training is typically applied, resulting in long lasting improved sensitivity (Fahle & Poggio, 2002; Sagi, 2011). During visual training, sequential discrete visual stimuli are presented to observers. In such training the inter-trials' spatiotemporal dependencies are assumed to accumulate and integrate to facilitate learning. The outcome of this training (i.e., learning) is measured by sampling sensitivity, using a similar structure of repeated performance, so that training and testing the outcome of training are done using the same procedure.

Sensory adaptation experiments have demonstrated that spatiotemporal interactions, enabled by an extensive presentation of trials (many repetitions) or extended exposure to a given stimulus lead to reduced neural response and consequently, to reduced visual performance (Webster, 2011). Similarly, in visual learning, extensive within-day training was shown to hamper performance (Mednick et al., 2002; Censor, Karni, & Sagi, 2006). The within-day deterioration was shown to be location specific and was independent of monetary reward (Mednick et al., 2002). It was also shown (Censor et al., 2006) that the number of trials applied in training affects both the discrimination thresholds and the amount of learning obtained; increasing the number of trials elevated the discrimination threshold, possibly due to sensory adaptation, and, counterintuitively, reduced learning. Although it is assumed that thresholds are objectively captured, in fact, the measurements may be contaminated by short-term spatiotemporal interactions, such as sensory adaptation (Ludwig & Skrandies, 2002; Ofen, Moran, & Sagi, 2007). Here, we attempted to provide estimates of learning gains that are minimally affected by short-term effects. While the mechanisms underlying these deteriorative effects are yet to be determined, the experimental methods employed in the present work show that within-day performance decrements can be alleviated.

---

[*] Corresponding author.
E-mail address: Dov.Sagi@weizmann.ac.il (D. Sagi).

http://dx.doi.org/10.1016/j.visres.2015.02.013

Recently a method for reducing the within-day performance decrements that are attributed to sensory adaptation has been demonstrated (Harris, Gliksberg, & Sagi, 2012). The method is based on earlier findings showing reduced contrast adaptation when sequentially adapting to two gratings differing by 45°, possibly due to inhibitory interactions between orientation-tuned neurons in the primary visual cortex (Greenlee & Magnussen, 1988). This was implemented in the texture-discrimination task (Karni & Sagi, 1991) by interleaving background only (no target) trials with target trials. The observers were required to respond to these trials the same way that they respond to target trials, i.e., to identify the missing target; thus, these additional trials were termed "dummy" trials. The bars presented on those 'dummy' trials were oriented 45° relative to the targets' local orientation. Training with the 'dummy' trials eliminated within-day performance decrements and resulted in generalization of learning across retinal space. Since the effectiveness of the dummy trials was found to be dependent on their local orientation (Harris et al., 2012), similarly to contrast adaptation (Greenlee & Magnussen, 1988), it was suggested that they reduce the effects of sensory adaptation during training. The improved performance with the insertion of dummy trials contradicts an explanation of within-day deterioration in terms of general fatigue or attention lapses, since, in addition to the above-mentioned local orientation dependency, the addition of dummy trials doubles the length of the training sessions, thus predicting greater stimulus-independent fatigue and an increasing number of attention lapses.

For measuring the performance threshold, stimulus magnitude needs to be varied, covering a range of magnitudes. In a typical perceptual-learning experiment, training starts with high magnitude stimuli (an easy level of difficulty, for example, target trials that are not followed by a mask in a backward-masking paradigm) and gradually proceeds to stimuli of lower magnitude (more difficult) either until a close-to-chance level performance is reached (Karni & Sagi, 1991), or, in the case of adaptive procedures, until the observer converges to an above-chance performance level (e.g., 75% correct), defined as the threshold (Levitt, 1971). An alternative method distributes all stimulus magnitudes randomly over the course of the training session (Dosher & Lu, 2000; Harris et al., 2012; Watanabe, Nanez, & Sasaki, 2001). Thus, although learning was shown using both gradual and random presentation methods, their effects on learning were never compared.

The gradually decreasing magnitude method has the advantage of allowing the observers to familiarize themselves with the stimuli and the task, thus stabilizing their performance level. Extra trials for familiarization purposes are commonly given in the pre-training phase. It is the earliest familiarization stage, which serves either as the initial practice or as a baseline performance measure. Common research practice shows that this pre-training phase is required for the observers so that they will be able to perform the trained task, without which subsequent training is inefficient. However, there is currently no research available that indicates the amount of pre-training that should be used; thus, it is not clear how these procedural manipulations affect future learning and transfer. It was previously shown that a minimal number of trials is required for long-term learning to occur (Aberg, Tartaglia, & Herzog, 2009), which varies across different tasks (Wright & Sabin, 2007). Hussain, Sekuler, and Bennett (2009) reported that a remarkably small amount of practice (5 trials per condition) was required to induce learning in a texture identification task and they found that further increasing the number of trials on the first day of training reduced learning during the 2nd day while preserving the total amount of learning. Pre-training was also found to be a crucial factor in measuring specificity and the transfer of learning, when administered at both trained and transfer locations. Zhang, Xiao, Klein, Levi, and Yu (2010) demonstrated that a pre-training phase (pre-test) of a peripheral stimulus induced the transfer of foveal learning to the pre-trained peripheral location. Typically in the pre-training phase a series of above-threshold (easy) stimuli with a constant stimulus magnitude is provided, thus introducing both spatial and temporal (spatiotemporal) consistencies.

Here we examined the effect of temporal consistencies and spatial consistencies on visual learning dynamics and transfer. The temporal consistencies were enhanced using a gradually decreasing SOA (experiment 1) instead of a randomly varying SOA (Harris et al., 2012). Next, to test how temporal interactions influence performance estimation (experiment 2), the effect of reduced temporal consistencies at the times that learning was measured (the start and end of training) was examined. Last, the influence of spatiotemporal repetitions in the early pre-training phase was tested by reducing the number of trials provided in the pre-training phase at the trained location (experiment 3). In each of these experiments the role of spatial consistencies was investigated in a separate group of observers. Spatial consistencies were reduced using 'dummy' trials (see above). The 'dummy' trials reduced the spatial consistencies via interrupting the repeated exposure of the target element, previously shown to counteract the within-day performance deterioration (Harris et al., 2012).

2. Methods

2.1. Apparatus

The stimuli were presented on a 19″ Mitsubishi Diamond Pro 930SB color monitor, using a PC with an Intel processor. The monitor refresh rate was 100 Hz. The mean luminance of the stimulus (line textures) was 63–65 cd/m^2 in an otherwise dark environment.

2.2. Stimuli and task

Observers were trained with the standard texture discrimination task (TDT, Karni & Sagi, 1991). In Fig. 1A the target frame (10 ms) and the mask frame (100 ms) are schematically shown. Each target frame was followed by a patterned mask. Observers had to judge the arrangement of a peripheral target (an array of three diagonal bars) and report whether it was horizontal or vertical. The peripheral target is embedded in a background (19 × 19 array of horizontal bars, 0.5° × 0.035°, and spaced 0.72° apart, 0.05° jitter) and its position was centered at 5.3° of a visual angle relative to the center of display. Mask patterns were 19 × 19 arrays of randomly oriented 'V'-shaped patterns. The display size was 14° by 13.5° of the visual angle, viewed from a distance of 100 cm. Fixation was enforced at the center of the display by a forced-choice letter discrimination task between a "T" and an "L". Next, they had to report whether the peripheral bar's array was horizontal or vertical. Responses were provided by pressing a computer mouse click. Auditory feedback was provided in case of an incorrect response for the fixation (T/L) task. Each trial was self-initiated by the observer. Target and mask stimuli were separated by a time interval (stimulus onset asynchrony, SOA) ranging from 10 to 300 ms. Each daily session (1–8) consisted of three consecutive sub-sessions (A, B, C). For each sub-session, SOA was either randomized across trials, with 6 presented trials per SOA (Harris et al., 2012) or decreased gradually. The gradually decreasing SOA training sub-sessions started at the highest SOA (300 ms) and ended at the observer's chance level SOA (the SOA for which the performance level was $\leqslant$60% correct) with 10 trials presented per SOA. The measured psychometric functions were fitted with